Early warning meets optimization: reimagining proactive policy responses to agricultural risk

Djavan De Clercq
Reuben College
University of Oxford

Supervisor: Prof. Adam Mahdi

Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Hilary Term 2026

Contents

Abstract

Disasters like droughts inflict about $123 billion in agricultural losses globally each year. But targeted, preemptive aid programs—the focus of this thesis—can mitigate their impacts.

One approach to tackling these hazards is leveraging *early warning systems*, which combine satellite data, climate models, and algorithms to predict disasters before they escalate. Properly used, these systems can trigger *anticipatory action*, a new type of humanitarian aid that allocates aid resources *before* a hazard occurs: for instance, by distributing cash or drought-resistant seeds ahead of expected weather extremes. This proactive approach is distinct from traditional humanitarian aid efforts, which typically allocate aid *after* disasters occur.

Major global humanitarian organizations are accelerating efforts to develop early warning systems and anticipatory action initiatives. However, important challenges need to be addressed as these systems scale. First, early warning forecasts are often inaccurate, biased, or generic (e.g., "a drought is expected") rather than fully *impact-based* (e.g., "an expected drought will cause crop losses of 30% in this region"). And anticipatory action itself often employs *blanket approaches* (e.g., providing the same amount of aid to everyone) rather than targeting resources to the most vulnerable. In many contexts, early warning systems are only loosely connected to anticipatory action frameworks, rather than functioning as a seamless, end-to-end approach.

A comprehensive review of the literature clarified the key challenges outlined above and informed the overarching research question that guides the thesis: *How can early warning systems and anticipatory action be designed to maximize the impact of aid delivered to vulnerable agricultural communities?* This question is addressed through four research papers: Paper I highlights how early warning systems and anticipatory action can be integrated; Paper II and Paper III advance scalable, impact-focused forecasts that predict actual crop losses, moving beyond generic alerts; and Paper IV integrates impact-based forecasts with mathematical optimization to target the most vulnerable, and evaluates the practical effectiveness of the framework using a simulation-based approach.

This DPhil research aims to make anticipatory action more timely, equitable, and targeted. The research uses India as a case study, due to heightened climate vulnerability in its rice-growing areas. The proposed approaches are broadly applicable to other agricultural settings confronting climate risks. [Thesis word count: 47,464]

1 Introduction

1.1 What is this thesis about?

This doctoral thesis introduces an *approach to deliver targeted preemptive aid* in the context of agricultural disaster relief by *addressing gaps in early warning systems and anticipatory action.* Disasters inflict about $123 billion in agricultural losses each year [1], but targeted aid programs can mitigate these impacts while ensuring that limited humanitarian budgets are spent effectively.

Anticipatory action is a new type of humanitarian aid that disburses resources, such as cash, drought-resistant seeds, or fertilizer, *before* a hazard occurs, when some given disaster threshold is met [2]. For instance, if a flood forecast for a region passes a certain level, a specified payout is automatically released. The payout is typically disbursed via a blanket approach, in that the same level of funds are provided to all recipients in an at-risk area. This proactive strategy is distinct from traditional humanitarian efforts, which typically respond *after* disasters occur.

For anticipatory action to work, the forecasts that trigger the release of resources need to be reliable; otherwise, “false alarms” can occur. This is the job of *early warning systems*, which use satellite data, climate models, and algorithms to produce forecasts that warn of impending hazards like droughts or floods [3].

Anticipatory action and early warning systems are exciting new fields in the humanitarian sector. But two important challenges need to be addressed to ensure that preemptive aid is disbursed in a targeted, transparent, and equitable manner (Figure 1.1).

The first challenge is that early warning systems have problems that constrain their utility for anticipatory action schemes. Some early warning systems are inaccurate, complicating the targeted release of funds [4]; some are biased, in that they perform better in some regions than others, impacting the fairness of aid allocation [5]; and some are not impact-based, meaning that they produce generic flood or drought warnings, rather than quantifying the real impact of these hazards, complicating efforts to give aid to those who need it the most [6, 7]. Lastly, early warning systems often operate in siloes from anticipatory action. This lack of harmonization means that warnings do not consistently inform decisions on early action.

The second challenge is that anticipatory action itself is *imprecise*: aid is often distributed uniformly across at-risk regions, as if the population were monolithic, overlooking differences in vulnerability. For instance, when facing drought, farmers without irrigation may need more support than those with it. Incorporating vulnerability assessments allows for more targeted aid, ensuring resources reach those most in need and preventing waste of limited humanitarian budgets, which is crucial for donors, humanitarian organizations, and governments [8, 9].

This thesis directly tackles these two challenges by answering a straightforward research question: *How can early warning systems and anticipatory action be designed to maximize the impact of aid delivered to vulnerable agricultural communities?*

The thesis answers this question by making the following contributions.

- First, I address gaps in agricultural early warning systems by introducing an end-to-end architecture that fuses publicly available geospatial datasets with machine learning to generate impact-based crop forecasts. This framework doesn't just warn of "floods" or "droughts" in general terms; it pinpoints *how* specific areas will be affected, providing actionable intelligence rather than broad-stroke alerts.
- Second, I introduce a computational framework that integrates impact-based forecasts into optimization-based aid allocation schemes, enabling timely, equitable, and cost-effective anticipatory action under constrained humanitarian budgets. The practical effectiveness of this approach is demonstrated using a simulation-based approach that compares it to traditional modes of anticipatory action.

This interdisciplinary research is at the intersection of humanitarian relief, geocomputation, machine learning, meteorology, and operations research. India's rice-producing regions serve as a case study, given their vulnerability to extreme weather and the global importance of rice as a staple crop.

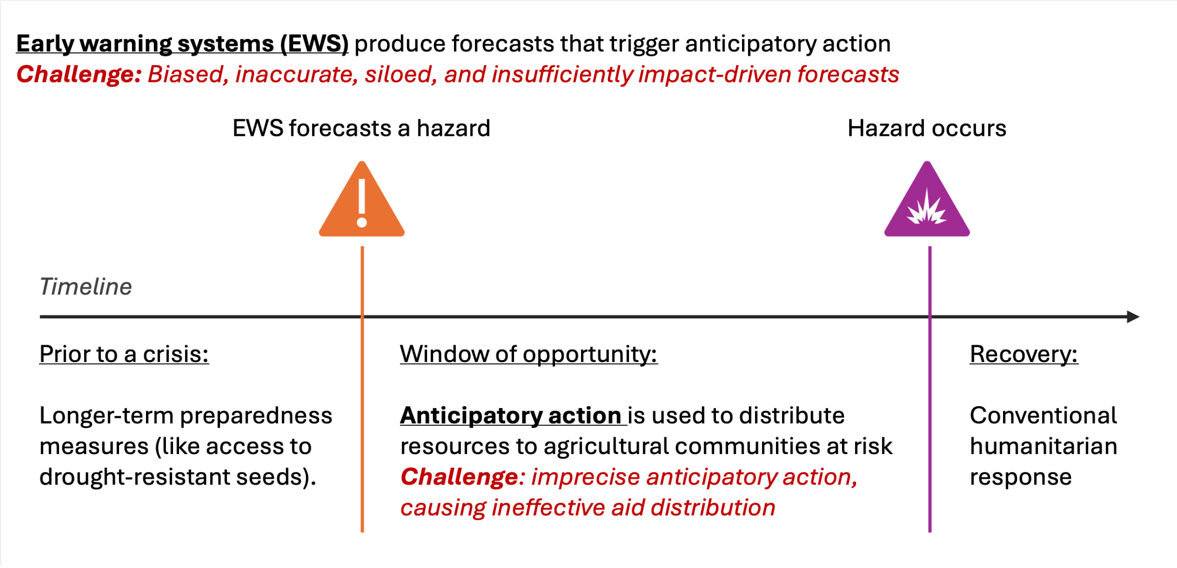


Figure 1.1: Early warning systems and anticipatory action face challenges that limit their ability to allocate aid precisely and effectively.

1.2 Context, gaps, and research question

1.2.1 Early warning systems and anticipatory action in brief

Investments in early warning systems can mitigate disasters in agriculture. Disasters like floods, droughts, and storms are increasingly devastating agriculture worldwide, climbing from about 100 events per year in the 1970s to nearly 400 per year in the past two decades, data from the United Nations show [10]. From 1991 to 2021, they have caused an estimated $3.8 trillion worth of crop and livestock losses [1], an average of roughly $123 billion per year in that time period. To put this into perspective, this annual loss represents about 3% of global agricultural GDP, which stood at 3.7 trillion in 2021 [11]. Lower- and lower-middle-income countries are hardest hit, losing 10% to 15% of their agricultural GDP [1]. In India, climate shocks affected 23.2 million hectares of crops between 2019 and 2023, with 87% of districts vulnerable to drought [12].

A promising approach for addressing these disasters is to use *early warning systems* to anticipate risks like droughts and floods before they escalate. There is strong evidence that investments in early warning systems are very cost-effective: for every $1 billion spent annually on early warning systems and disaster response in developing countries, an estimated $4 to $36 billion could be saved in avoided losses [13]. Recognizing this potential, the United Nations launched a $3.1 billion “Early Warnings for All” initiative in 2022 to ensure that every person on Earth is protected by some form of early warning systems by 2027 [14]. While progress has accelerated since its launch, with over 119 countries reporting multi-hazard early warning systems by late 2025, recent assessments indicate that significant gaps in coverage and implementation persist, particularly in vulnerable and low-income countries [15].

Notable agriculture-focused early warning systems include global systems such as the United Nations' *Global Information and Early Warning System* (GIEWS) [16], and regional systems such as the United Nations' *Somalia Water and Land Information Management* (SWALIM) system. Alerts from such systems have helped farmers mitigate risks by taking proactive measures, such as reallocating fungicide stocks to contain wheat rust outbreaks in Ethiopia and harvesting crops early to prevent weather damage in Zimbabwe [17, 18].

Anticipatory action can translate early warnings into concrete humanitarian impact. Early warnings alone rarely suffice to prevent humanitarian or economic crises [19]. While a drought forecast may provide critical lead time for decision-makers to plan their response, acting on it effectively requires more than just awareness. It demands a structured response.

This is the premise of *anticipatory action*, an emerging field that seeks to turn early warnings into concrete, preemptive measures. Instead of waiting for a crisis to unfold and reacting after the fact, policymakers and aid organizations use forecasts to issue cash transfers before droughts take hold, distribute drought-resistant seeds, or advise farmers to harvest early [20]. This approach, championed by institutions such as the United Nations and the Red Cross, shifts disaster management from reaction to preemption, ensuring that interventions are deployed where they can do the most good.

Anticipatory action has been used in diverse settings, from protecting livestock and boosting milk production in Kenya and Sudan [21, 22], to distributing drought-resistant seeds and irrigation equipment in the Philippines and Madagascar [23, 24], to mitigating extreme cold impacts on herders in Mongolia [25]. In Afghanistan, cash and livestock support improved household food security [23], while in flood-prone areas, such as Bangladesh, simple interventions like waterproof storage drums helped families protect critical supplies [23].

Anticipatory action is still in its infancy, with critical gaps that must be addressed for meaningful impact as it scales up. In 2022, governments and international organizations allocated about $390.6 billion for *international development financing*, covering everything from economic growth initiatives to disaster response

[26]. Within this, \$277.1 billion went towards *official development assistance* (ODA), which primarily funds long-term development projects, while \$39.2 billion went to *humanitarian aid*, defined as emergency relief for crises like natural disasters and conflicts.

Focusing on disaster response, total *crisis financing*, defined as funds used to prevent, prepare for, and respond to disasters, accounted for about \$76 billion [26]. However, much of this funding is reactive, disbursed after disasters strike.

In contrast, only \$0.9 billion was allocated to *pre-arranged financing* (PAF), which ensures funds are available in advance, but only released when a disaster occurs. For instance, an automatic payout triggered by a flood. An even smaller portion, just \$148 million in 2023, was allocated to *anticipatory action* [27], which goes a step further by distributing aid ahead of disasters based on forecasts, rather than waiting for them to happen. In 2023, anticipatory action funds reached approximately 10.9 million people across 47 countries [27].

Because anticipatory action is still in its early stages, there is a unique opportunity to shape its development, especially as the United Nations and other global bodies call for more proactive humanitarian responses. In 2014, the UN General Assembly urged "a shift towards an anticipatory approach to humanitarian crises," and the Sendai Framework for Disaster Risk Reduction 2015–2030, guided by the UN Office for Disaster Risk Reduction (UNDRR), encourages states to invest in anticipatory methods [28].

1.2.2 Gaps in early warning and anticipatory action

Realizing the promise of anticipatory action will require revamping early warning systems by improving forecast accuracy, timeliness, and data quality, as well as designing targeted approaches to allocating aid that genuinely reach those most in need. My review of the literature revealed two overarching challenges that must be addressed to maximize the impact of anticipatory action [29].

First, many early warning systems are not actionable, which limits their ability to reliably inform anticipatory action. This is due to several factors such as inaccuracy, bias, and a lack of impact-based predictions.

Early warnings are often *inaccurate* and *biased* [3]. One study found in much of the tropics, weather forecasts have relatively poor skill in forecasting extreme temperature and rainfall events [30], especially over longer lead times [31]. Another study found that, between 2009 and 2020, the widely used Famine Early Warning Systems Network (FEWS NET) for 25 African countries achieved around 84% accuracy, and unanticipated weather shocks further hampered its projections [32]. Furthermore, in Ethiopia, warnings proved more reliable in western regions than in the more food-insecure northeast, likely due to insufficient information. Similar issues appear in Europe: an analysis of nearly 20,000 production forecasts showed that forecasts systematically underestimated production for wheat, rapeseed, and sugar beet [5]. Such regional or crop-specific discrepancies in accuracy risk triggering unfair or misplaced humanitarian responses, while repeated false alarms can create a “cry-wolf effect,” eroding trust and impeding timely action during genuine crises [33, 34].

Many early warning systems are *not impact-based*, meaning that they predict the weather without showing who will be hit hardest [35]. This can leave governments and aid agencies unsure where to act. As a United Nations manual highlights, effective warnings must move from stating what the weather will be (e.g., “drought is coming”) to what it will do (e.g., “30% of crops will fail”) [36]. If a drought alert treats a commercial farm with deep wells the same as a smallholder reliant on rainfall, even though the former can withstand weeks of dryness while the latter faces total crop failure, anticipatory action may be misallocated. Impact-based forecasting fixes this by replacing generic hazard maps with vulnerability maps, which helps aid organizations understand where aid is needed most [37]. For agricultural early warning systems specifically, an important gap is crop yield forecasting, which converts weather data into expected impacts on crops [16, 38], helping to anticipate food production risks or losses of livelihood.

Second, anticipatory action frameworks’ capacity to protect the most vulnerable is undermined by two design flaws: weak targeting of vulnerable groups and poor integration with early warning systems.

Effective targeting of vulnerable populations remains difficult in practice. In most anticipatory action schemes, blanket aid distribution remains common, meaning that the same level of support is given to everyone despite differences in vulnerability. For example, in a drought-hit district with 10,000 farmers and a $10 million budget, each farmer might receive $1,000, even though a small-scale farmer reliant on rainfall faces a far higher risk of losing the entire crop than an irrigated farmer using drought-resistant seeds. A more structured plan could allocate $2,000 to those at greatest risk, $1,000 to farmers with moderate risk, and $500 to those better prepared, all while staying within the same $10 million budget. The need to target anticipatory aid more effectively has been articulated by the United Nations: “simply acting early does not guarantee that all benefit equally from anticipatory actions,” so “humanitarian actors need to actively strive to be inclusive in all their efforts” [39]. Without it, already finite resources risk being misdirected, arriving too soon, too late, or going to the wrong people, undermining both trust and impact [40, 41]. Designing well-targeted aid allocation strategies is critical in a world where humanitarian budgets are already constrained [9]. For instance, funding gaps in humanitarian aid reached $32 billion in 2023 [42].

Moreover, due to its relative novelty, *anticipatory action programs remain disconnected from early warning systems*. Forecasts for disasters like Cyclone Nargis (2008), the Somalia famine (2011), and the 2021 floods in Germany were all in place, yet catastrophic losses occurred due to breakdowns in communication, coordination, or institutional readiness [43]. Simply possessing a forecast is not enough; rather than asking, "What forecasts do we have?" a more fitting question is, "What decisions need to be made, and what forecast information is required to support them?" [19]. As the Centre for Disaster Protection emphasizes, anticipatory action must be fully integrated into early warning systems rather than functioning as a separate add-on [44]. Without this shift, early warnings will remain passive alerts rather than the catalysts for decisive action.

1.2.3 Research questions and case study region

My review of the literature clarified the key challenges outlined above and informed the research questions that this thesis addresses [29]. The overarching research question is:

> *RQ: How can early warning systems and anticipatory action be designed to maximize the impact of aid delivered to vulnerable agricultural communities?*

This broad question is further divided into two subsidiary research questions, emphasizing the thesis' focus on computational and data-driven approaches to designing early warning systems and decision-making structures for anticipatory action:

> *SQ1: How can early warning systems deliver scalable, impact-based forecasts that anticipate crop production shocks?*
> *SQ2: How can early warning forecasts power computationally optimized aid allocation decisions to maximize the effectiveness of anticipatory action?*

India serves as the empirical case study through which these research questions are operationalized. While the analytical framework developed in this thesis is intended to be broadly applicable, India offers a particularly instructive context for anticipatory action, given its scale, climatic volatility, and the centrality of agriculture to rural livelihoods.

Over half of India's land area is devoted to agriculture, and approximately 70% of rural households rely on farming as their primary source of income [45, 46]. Much of this agriculture is rainfed, leaving production highly exposed to climate shocks that are projected to intensify in coming decades [47].

Despite productivity gains since the 1950s, driven by expanded use of fertilizers, irrigation, and improved seeds, extreme weather still causes sharp fluctuations in the yields of key crops such as rice, wheat, and maize [48]. One study of 58 years of yield data (1961–2017) across seven major crops (e.g., rice, pulses, wheat) found rainfall and extreme temperatures to be major drivers of yield variability [49]. More recently, record-breaking heat in March 2022 scorched wheat fields, cutting yields by up to 15% and prompting India to ban wheat exports for domestic food security [50, 51].

Climate models suggest that India will face more frequent extreme precipitation events and hotter, drier summers, creating further risks for its food systems [52, 53, 48]. Such extremes can slash household incomes by 25–60%, trigger conflicts over water, and force families to withdraw children from school or migrate in search of relief [54, 55, 56]. Although early warning systems have been highlighted as crucial for building resilience [57], current systems in India still struggle with limited impact-based forecasting, poor communication of alerts, and a lack of localized response plans [58, 59]. Strengthening these systems is vital for paving the way toward more robust anticipatory action, helping ensure timely, targeted interventions that protect vulnerable farmers before disasters escalate.

1.3 Thesis structure and contributions

This integrated thesis is organized around four research papers that I led as first author, in collaboration with international experts in agricultural early warning systems, anticipatory action, and disaster risk management. Each paper addresses a distinct aspect of the research question and contributes to the broader goal of integrating early warning systems and anticipatory action to maximize resilience in agricultural communities:

Core Paper I (Chapter 2): “Towards optimal anticipatory action: maximizing the effectiveness of agricultural early warning systems with operations research”
(Published in the International Journal of Disaster Risk Reduction).

What it does: This “background” paper provides both the foundational context for this thesis and the central research question by introducing early warning systems and anticipatory action, and discussing their role in mitigating climate shocks in farming communities. The paper makes a case not only for *integration* between early warnings and anticipatory action, but also for major advancements in the computational underpinnings of forecasting and decision-making to pave the way for proactive, data-driven resilience in agricultural systems.

Why it matters: This study pinpoints the broader disconnect between having a hazard forecast (e.g., for droughts or floods) and using that information to effectively protect agricultural livelihoods. It demonstrates how to move beyond "blanket" aid distributions, where each farmer gets the same payout regardless of risk, and adopt a more nuanced approach that matches higher payouts to those facing greater vulnerabilities (e.g., rain-fed smallholders), while minimizing waste in regions better equipped to withstand shocks. For farmers, that translates into more timely, targeted relief when forecasts indicate trouble ahead; for aid agencies, it offers a transparent, data-driven way to explain who gets how much aid, and why; for funding organizations, it provides greater confidence that every dollar spent tangibly boosts climate resilience. Additionally, the paper clarifies opportunities for developers of early warning systems and anticipatory action protocols to refocus future research and development efforts, ensuring that improved forecasting actually leads to more targeted interventions.

Core Paper II (Chapter 3): "Modern computational approaches for rice yield prediction: a systematic review of statistical and machine learning-based methods" (*Published in Computers & Electronics in Agriculture—April 2025 issue*).

What it does: This paper systematically surveys 156 studies on statistical and machine learning-based approaches for crop yield forecasting. It not only highlights which techniques and datasets are typically used at different scales (field-level vs. region-level predictions) but also exposes critical stumbling blocks like limited ground-truth data, inconsistent approaches for model validation, and difficulty generalizing across diverse regions and cultivars. By clarifying these data and algorithmic gaps, the review offers a practical foundation for effective design of yield-focused forecasting systems that can integrate into agricultural early warning systems.

Why it matters: By documenting the chief stumbling blocks in rice yield forecasting, from inadequate ground-truth data to limitations of certain algorithms, this review points the way toward more powerful yield forecasting approaches that work as part of real-world, operational solutions. Better forecasting systems mean agricultural producers receive earlier, more accurate warnings about potential crop losses, and governments can mobilize resources in a more targeted and timely manner. The review also guides research and development toward scalable approaches, rather than one-off proofs of concept. Ultimately, bridging these gaps can elevate yield forecasting into a core component of agricultural early warning systems.

Core Paper III (Chapter 4): "Feasibility of Machine Learning-Based Rice Yield Prediction in India at the District Level Using Climate Reanalysis and Satellite Data" *(Published in Agricultural Systems—October 2024 issue).*

What it does: Building on insights from the systematic review, this study implements a fully programmatic, end-to-end workflow to forecast district-level rice yields in India using climate reanalysis data (e.g., ERA5) and satellite imagery (e.g., NASA MODIS). By automating data ingestion, preprocessing, and validation within a single pipeline, it demonstrates that subnational yield predictions can be generated at scale without relying on commercial or proprietary tools. Machine learning models are trained on these openly available datasets, and an interactive visualization tool reveals where forecasts tend to over- or underestimate yields to highlight any potential regional patterns of bias. The paper provides a reference architecture for integrating large-scale geospatial data and machine learning models into production-grade early warning systems.

Why it matters: By harnessing freely available climate and satellite data, this paper complements more generic hazard alerts (e.g., "drought on the way") with location-focused yield forecasts (e.g., "rice shortfall likely in District X"). Yield prediction is often a missing piece in early warning systems, and filling that gap can guide aid agencies toward areas at the highest risk of crop losses. In addition, because all tools used are open source, the work offers a practical starting point for research and development teams looking to refine and scale similar yield-forecasting approaches in real-world anticipatory action settings.

Core Paper IV (Chapter 5): "Planning ahead: optimizing early warning-driven anticipatory action cash transfers via an open-source web application" *(Submitted to the International Journal of Disaster Risk Reduction; currently under revision to address reviewer comments prior to resubmission).*

What it does: This paper applies mathematical optimization (e.g., linear programming) to determine not just *how much* cash assistance to allocate, but also *where* and *when* to release it in phased tranches. A central feature is a user-friendly, web-based tool that lets practitioners or policymakers easily compare different allocation scenarios without needing to write code. Built on open-source optimization and visualization libraries, this system demonstrates how inputs from early warning models can be translated into a transparent, spatially explicit plan of who gets aid, in what amount, and at which stage. Additionally, a simulation-based approach is used to test our optimization strategy against a simpler, population-based "blanket" method, allowing us to pinpoint when and where the optimized plan consistently improves outcomes.

Why it matters: Accurate hazard forecasts alone won't prevent wasteful or poorly timed disbursements. Even with the right early warning, if distributions are ad hoc or "one-size-fits-all," limited budgets can still be squandered. By systematically balancing vulnerability scores, budget constraints, and equity criteria, the proposed tool helps ensure resources flow to the farmers or districts most at risk, and at moments when it can do the most good. In so doing, it supports more credible, impact-oriented anticipatory action, giving aid agencies a practical way to justify decisions, and creating a clearer path from forecast to meaningful, targeted interventions.

Discussion and Conclusions (Chapter 6): Integrates the thesis findings across forecasting and optimization, assesses methodological and institutional constraints, and reflects on the implications for institutionalizing anticipatory, decision-oriented early warning systems.

Together, these papers offer a structured foundation for better integrating early warning systems with anticipatory action in agriculture, while also strengthening the algorithms that underpin both forecasting and structured decision-making. The methodologies and tools developed are designed for practical implementation, and the overall framework is adaptable beyond India's rice sector, potentially supporting more data-driven, risk-informed disaster preparedness initiatives in diverse agricultural contexts worldwide.

1.4 Engagement with academic experts

This thesis benefited from the expert review and insights of leading researchers across multiple disciplines. They co-authored papers and provided feedback on key methodological components. Jonas Jagermeyer (NASA Goddard Institute for Space Studies and Columbia University) and Marijn van der Velde (Joint Research Centre of the European Commission) reviewed sections on agricultural monitoring using satellite and climate data. Jim Hall (Professor of Climate and Environmental Risk, University of Oxford), Marc van Homberg (Scientific Lead of the 510 Netherlands Red Cross and Professor, University of Twente) and Marleen de Ruiter (Assistant Professor of Disaster Risk Management, Vrije Universiteit Amsterdam) contributed expertise in disaster risk monitoring and climate risk. Sagar Surendra Deshmukh (Assistant Professor of Agricultural and Rural Technology, IIT Guwahati) contributed expertise on agricultural challenges in India. The optimization components of this work benefited from feedback

Table 1.1: Summary of research papers that form the integrated thesis

RQ: How can early warning systems and anticipatory action be integrated to maximize resilience in vulnerable agricultural communities?

	Paper I	**Paper II**	**Paper III**	**Paper IV**
Title	*Towards optimal anticipatory action: maximizing the effectiveness of agricultural early warning systems with operations research*	*Modern computational approaches for rice yield prediction: a systematic review of statistical and machine learning-based methods*	*Feasibility of Machine Learning-Based Rice Yield Prediction in India at the District Level Using Climate Reanalysis and Satellite Data*	*Planning ahead: optimizing early warning-driven anticipatory action cash transfers via an open-source web application*
Research questions addressed	**RQ:** How can early warning systems and anticipatory action be designed to maximize the impact of aid delivered to vulnerable agricultural communities?	**SQ1:** How can early warning systems deliver scalable, impact-based forecasts that anticipate crop production shocks?	**SQ1:** How can early warning systems deliver scalable, impact-based forecasts that anticipate crop production shocks?	**SQ2:** How can enhanced early warning forecasts power computationally optimized aid allocation decisions to maximize the effectiveness of anticipatory action?
Methods	Literature review	Systematic review Comparative analysis of forecasting techniques	Machine learning model development Climate and satellite data analysis Model interpretability and bias detection	Mathematical optimization (linear programming) Simulation of cash transfer strategies Web-based decision-support tool
Relevance	Bridges hazard forecasts and structured decision-making for more effective anticipatory action	Reveals key data and algorithmic gaps in yield forecasting for early warning systems	Implements district-level yield predictions using open data and ML for large-scale early warning adoption	Uses data-driven optimization to allocate resources efficiently based on improved forecasts
Publicat. status	**Published:** *International Journal of Disaster Risk Reduction*, March 2025	**Published:** *Computers & Electronics in Agriculture*, April 2025	**Published:** *Agricultural Systems*, October 2024	**Published:** *International Journal of Disaster Risk Reduction*, September 2026

from Burcu Balçık (Professor of Humanitarian Relief Chains, Özyeğin University) and Lily Xu (Postdoctoral Researcher, University of Oxford; Incoming Assistant Professor of Industrial Engineering and Operations Research, Columbia University).

1.5 Insights from professional work

Throughout my part-time DPhil, I have also served as a Director of Data Science at McKinsey & Company, where I was part of ACRE, a specialized group focused on leveraging artificial intelligence to address agricultural challenges. This position offered a unique vantage point on problems that cut across global agricultural systems. In addition, I gained access to cutting-edge datasets spanning agriculture, climate, and earth observation, as well as hands-on experience deploying predictive models and optimization algorithms at scale across food systems.

My work at McKinsey included predicting disease pressure on rice in China, building digital twins of sugarcane farming operations in India, estimating field-level weekly blueberry production and quality in Peru, evaluating food security monitoring systems in East Africa, forecasting sugar beet supply in the United Kingdom, estimating wheat production in France, monitoring rice production in the Philippines, modeling micronutrients in the world's soils, and forecasting citrus production in Brazil (Figure 1.2). It also included structuring investments into the Saudi Arabian food production systems and optimizing the flow of crops from farm to processing mills in various contexts. Visiting agricultural production sites across Peru, India, the United Kingdom, Saudi Arabia, and elsewhere, I gained first-hand insight into the priorities of farmers, government officials, investors, and humanitarian practitioners, all of whom shape the complex ecosystem in which data-driven agricultural innovations must operate.

A common theme across these projects was the disconnect between forecasting and decision-making. Even in highly commercialized agricultural settings, stakeholders often had access to good forecasts, such as next week's crop yield and quality, but subsequent decisions, like where to export given prevailing market prices and logistics costs, were still made using simple heuristics or gut-feel rules. In cases where mathematical optimization was introduced, significant increases in value were observed. However, optimization remains underutilized in agriculture compared to industries like aerospace, e-commerce, semiconductors, or automobile manufacturing, where it is more deeply embedded in decision-making processes.

These experiences reinforced two central pillars of this thesis. First, they underscored the importance of good forecasting techniques as the backbone of effective agricultural

Figure 1.2: Examples of agricultural production sites visited during my professional career, such as blueberry farms in Peru (left) and rice paddy fields in India (right).

monitoring systems, particularly in anticipating shocks such as droughts or disease outbreaks. Second, they highlighted the need for structured decision frameworks to translate these forecasts into timely, targeted interventions (in both private and public sector contexts), aligning closely with the thesis' focus on anticipatory action.

1.6 Concluding thoughts

This thesis demonstrates how coupling impact-oriented forecasting with computationally optimized anticipatory action can transform limited humanitarian resources into more precise and timely interventions for those most at risk. By linking forecasts directly to decisions, it offers scalable methods that strengthen early warning systems and improve the targeting of aid.

No single framework can deliver flawless early warnings or perfect anticipatory action. Challenges remain, from engineering global-scale monitoring systems to coordinating stakeholders. Yet by addressing key methodological gaps and providing practical tools, this thesis lays the groundwork for more decisive, data-driven responses to agricultural crises, complementing broader efforts in disaster risk reduction and climate resilience.

2 Towards optimal anticipatory action: maximizing the effectiveness of agricultural early warning systems with operations research

A version of this chapter is published online in the International Journal of Disaster Risk Reduction (March 2025 issue).

Abstract

Agricultural communities worldwide face escalating climate-induced hazards, including droughts, floods, and extreme temperatures. Early warning systems have emerged as a key defense, providing forecasts that trigger *anticipatory action*, such as cash transfers or distributing drought-tolerant seeds, before a crisis materializes. However, these interventions often rely on static thresholds and blanket aid allocation strategies, limiting their overall effectiveness. This paper shows how operations research and data-driven optimization can refine resource allocation by accounting for varying vulnerabilities, budgets, and logistics. We first review agricultural risks and the evolving role of early warning systems and anticipatory action. We then introduce key optimization methods (e.g., linear programming, stochastic models) and illustrate their potential benefits with regard to more targeted aid funding. We conclude with research directions that set the stage for subsequent chapters, which delve deeper into designing impact-based algorithms for agricultural early warning systems, and integrating these methods into decision frameworks that enable equitable, targeted anticipatory action.

2.1 Introduction

In agricultural regions around the world, communities increasingly face climate-related crises such as droughts, floods, and erratic temperature shifts [60, 61, 62]. Anticipatory action, which is a type of preemptive intervention informed by early warnings, such as distributing cash transfers to farmers ahead of a predicted drought, has been used to reduce the impacts of these crises by enabling preparedness before disasters strike. From drought in East Africa to Mongolia's severe winter "dzud," this proactive approach to disaster management has gained popularity across international organizations and governments, offering lower cost and greater efficiency over the reactive approach of addressing damages and losses after they have already occurred.

Yet translating early warnings into effective anticipatory action remains challenging. A game developed by the Red Cross Red Crescent Climate Centre simulates this complexity: participants, acting as either farmers or donors, allocate limited resources across flood and drought scenarios, balancing immediate risks with resource constraints to safeguard lives and livelihoods [63]. Many different real-world scenarios involve similar decisions: should funds go toward cash subsidies, early harvest advisories, or micro-irrigation tools? Effective action depends on the ability to allocate scarce resources across regions with widely varying needs.

The reality of anticipatory action often involves difficult trade-offs under time pressure and uncertainty. Resources are finite, and decisions about where to direct aid, how much to allocate, and when to act are complex [20]. Currently, many response plans depend on pre-defined thresholds or simplified heuristics to guide decisions, which can overlook nuanced needs across different areas or population groups. For example, one study of anticipatory action concluded that there remain significant deficiencies in "understanding the optimal balance between in-kind/service delivery and cash assistance" and "determining the frequency and amount of distributions" [64]. As anticipatory action scales up and becomes more widespread, minimizing subjective judgment in resource allocation can help create systems that are more just, adaptive to evolving conditions, and scalable.

Despite its increasing popularity, anticipatory action remains constrained by limited funding and the complexity of coordinating diverse stakeholders. In 2023, only \$148 million, representing less than 1% of the \$70 billion in global disaster financing, was allocated to anticipatory action, supporting just 10.9 million people across 47 countries [65, 66]. Even with efforts from 125 organizations, NGOs, and local partners, responders face persistent difficulties in aligning resources with needs, addressing contingent

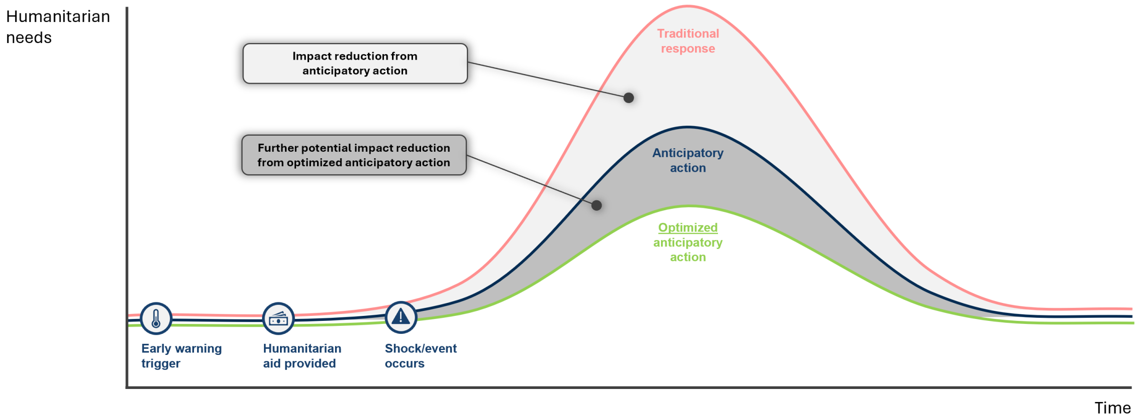


Figure 2.1: Comparative humanitarian needs over time under different response approaches. The red curve represents traditional reactive responses, the blue curve represents anticipatory action based on early warnings, and the green curve illustrates the potential for further impact reduction through optimized anticipatory action.

liabilities, and adapting interventions to regional disparities. Achieving greater impact will require more precise and equitable methods to allocate scarce resources under these constraints.

In this article, we discuss how mathematical optimization can improve the effectiveness of agricultural anticipatory action, as illustrated in 2.1. We begin by examining the current state of early warning systems, followed by an exploration of how optimization methods have been applied in other fields and the potential of these methods to improve anticipatory action. We then outline directions for future research: enhancing early warning systems, integrating them with optimization frameworks, and applying optimization to improve the effectiveness of anticipatory action in the humanitarian sector.

While the goal of this paper is not to present optimization as a panacea, we aim to highlight a useful tool that anticipatory action practitioners should consider adding to their toolkit. The practical application of optimization is already gaining momentum in the humanitarian sector. The World Food Programme's "Optimus" system demonstrates how optimization methods can save millions in donor funding by identifying the most cost-effective strategies for food distribution [67]. Through an online decision-support tool, Optimus integrates data on population needs, transport routes, and local market conditions to optimize food basket design, sourcing, and delivery networks, allowing WFP to reach more people at lower costs. Used in over 44 operations, Optimus has saved more than $50 million to date by improving efficiencies and enabling quick, evidence-based decisions, such as sourcing locally to cut lead times in Madagascar and diversifying food baskets in Ethiopia to respond to supply chain disruptions [68]. This success underscores optimization's potential to enhance anticipatory action, including in agricultural contexts such as preemptively disbursing aid to farmers affected by hazards like drought or flood, ensuring timely support and minimizing losses.

2.2 Failure to plan for weather shocks has a cost

Extreme weather events such as floods and windstorms, alongside cascading phenomena like agricultural droughts or locust outbreaks, inflict diverse material and non-material losses on farming communities. These range from direct economic costs, such as destroyed crops and lost income, to indirect damages like food price inflation and long-term degradation of soil quality. Non-economic effects, including mental health crises and reduced social cohesion, further compound the challenges faced by vulnerable populations. While not all these consequences can be mitigated through anticipatory action, understanding the breadth and complexity of these impacts is essential for designing effective interventions. Coordinated, equitable resource allocation requires recognizing how these losses differ significantly depending on regional conditions, hazard types, and the socio-economic fabric of the affected areas, as these factors shape both the need for preemptive measures and the feasibility of response strategies.

2.2.1 Effects on agricultural yields and economies

In 2020, billions of locusts descended on parts of East Africa, decimating cropland with the worst swarms in decades [69]. Locust swarms, which can contain up to 80 million locust adults per square kilometer, consume the same amount of food in a day as 35,000 people [70]. The livelihoods of nineteen million farmers and herders across Ethiopia, Kenya, and Somalia were severely affected as a result [71]. The colossal swarms that hit these countries were two years in the making, as major storms inundated a remote area of Saudi Arabia in 2018, leading to an 8,000-fold surge in the desert locust population. Strong winds subsequently drove the insects into East Africa in mid-2019, where wet weather further bolstered their numbers.

Disruptions to agriculture caused by extreme weather events are not limited to East Africa, as countries around the world have seen major shocks to agricultural supply due to heavy rains and extreme temperatures. Vegetable prices in China's Shandong province doubled in 2021 as unusually heavy rains drenched northern swathes of China in September, causing vegetables to lie 'dead in the ground' [72]. British sugar beet farmers lost 45 million pounds in 2020, as sugar beet yields decreased by 61% due to exceptional weather conditions, including the driest May since 1868 and the wettest February since 1914 [73]. The windstorms in the United States in 2020, which tore across the U.S. Corn Belt at speeds upwards of 100 miles per hour, left millions of acres of corn and soybean fields destroyed, grain bins pummeled, and buildings leveled, plunging many farmers in the region into economic uncertainty [74]. In Tamil Nadu, India, there were 67 extreme weather days in 2024, nearly double the 27 days reported in 2023. This led to the loss of 1,039 hectares of crops, significantly impacting farmers' livelihoods [75].

These events illustrate some of the challenges faced by global agricultural producers. There will be 10 billion people on Earth by 2050, or around 2 billion more people than there are today [76]. This will test the global agricultural system and the farmers who supply that system. Research has warned that agricultural production worldwide will face challenges in meeting global demand for food and fiber [77, 78], with food demand estimated to increase by more than 70% by 2050 [79]. Such challenges include climate-related pressures and extreme weather, biotic threats such as pests and diseases, decreasing marginal productivity gains, soil degradation, water shortages, nutrient deficiencies, urban population growth, rising incomes, and changing dietary preferences [80, 81, 82].

Extreme weather events, like those mentioned above, often dominate newsfeeds, underscoring the broader variability and unpredictability that research has shown to impact global agriculture. One study showed that climate variation accounts for roughly a third (32–39%) of observed variability in crop yields [83]. Another study found that growing season climate factors explain 20–49% of the variance of global yield anomalies, depending on the crop type [84]. Moreover, recent research has shown that climatic hazards such as drought, extreme precipitation, and extreme temperature events exceeding historical ranges are expected to become more common globally [85]. Increasingly severe drought conditions could reduce economic output in the agriculture sector by 10% in the absence of climate mitigation efforts [86]. Other regions where research has demonstrated an increasing vulnerability to drought risk include South America, East Africa, China's Pearl River Basin, northern India, and Australia [87, 88, 89, 90].

Climate-induced destruction of crops has significant economic costs. For example, the United States Federal Crop Insurance Program, which is a federally subsidized program that pays farmers when their crop yields or revenues decrease, has paid out $143.5 billion in federal crop insurance payments to farmers from 1995 through 2020, with two-thirds of that amount paid out for crop damage from drought, excess moisture, hail, heat, excess wind, and cold weather exacerbated by climatic change [91]. Moreover, analyses of climate model simulations provide high confidence that anthropogenic climate forcing has increased U.S. crop insurance losses, with temperature trends alone in the U.S. contributing $27 billion of the national-level crop insurance losses over the 1991–2017 period [92]. Global insurer Munich Re has estimated that the worldwide economic cost of natural disasters from extreme weather increased from $25 billion per year during the 1980s to $175 billion in 2016 [93]. Globally, three-quarters of harvested areas (454 million hectares) between 1983 and 2009 experienced drought-induced yield losses, translating to crop production losses of $166 billion. Globally averaged, a single drought event decreases agricultural gross domestic production by 0.8% [94].

2.2.2 The human cost of climate extremes

Farming is hard work: debt, loneliness, long days, and stress [95]. Consider for example the economic lives of the two billion people who live in about 475 million small farm households on farming plots of land smaller than two hectares [96]. Many are poor, food insecure, and have constrained access to markets and services. They operate their farms as business owners, providing an estimated 80% of the food produced in Asia and sub-Saharan Africa [97], and raise capital from multiple sources to invest in productive assets, such as a bicycle, spade, or bag of seeds. They take decisions which might lead to financial gain or financial hardship.

And to produce food, smallholder farmers must grapple with numerous decisions: what to plant; when to plough, seed, and harvest; which fertilizers and crop protection products to use and when; how much of the harvest to retain for household consumption and how much to sell on the market to raise funds; and how much of the harvest to keep in storage. Smallholder farmers make such decisions in economic environments where markets function sub-optimally, and they are exposed to risks such as weather shocks, geopolitical risk, and price fluctuations. Ultimately, the decisions that smallholders make under such uncertainty have tangible implications for their livelihoods, such as their ability to invest in themselves and their children [98].

Given the challenging conditions in which farmers operate, it may come as no surprise that researchers have identified farming as a particularly stressful occupation associated with a number of occupational and mental health risks [99, 100, 101, 102], and that farmers have been found to have worse mental health issues than the general population [103]. Studies on mental health in agricultural communities around the world in the last 40 years have revealed numerous mental health risk factors, such as financial hardship [104], bad weather [105], government policy [106], lack of control over market conditions [107], paperwork [108], financial struggles [109], pesticide exposure [110], limited social mobility [111], anti-immigrant discourse [112], increased workload at peak times [113], poor housing conditions [114], exposure to loud machinery [115], isolation and lack of social support [116, 117], drought [118], water scarcity [119], concerns about children [120], rural-urban schisms [121], substance abuse [122], time pressure [123], farming bureaucracy [124], marital problems [125], crowded living conditions [126], and occupational injury [117]. Such mental health struggles have also been found to be linked to suicide risk in farming populations [127, 128, 129, 130, 131, 132, 133, 134], and research has shown that farmers face significant barriers to help-seeking behaviour [135, 136, 137, 138].

Compounding the litany of day-to-day stressors that farmers face, they must also contend with increasingly erratic weather patterns induced by climate change and the effect on their livelihoods. One study noted that in developing nations, farmers' constant exposure to weather extremes is a significant factor in the low productivity of agriculture, the slow pace of economic development, and persistent poverty [139]. More specifically, studies have shown that extreme climate events have led to: declines in up to 50% in household income from tea sales for Chinese tea farmers [140]; increased anxiety due to perceived impacts on profitability of farming operations for farmers in the United States [141]; heightened worries about the weather, undermined ideas of self-identity, and greater perceived risk of depression and suicide for farmers in Australia [142]; decreased crop yields and reduced income for smallholder farmers in six Central American regions [143]; greater mental stress due to crop loss, income loss, ineffective disaster management, destruction of infrastructure and homesteads, and limited capacity to deal with shocks [144]; crop and income losses for Malagasy farmers [145]; low crop yields and high crop failure in South Africa [146]; recurring crop failures from typhoons, flooding, and dry spells, hampering the smallholders' farming economy in the Philippines [147]; and increased financial losses between 2000 and 2020 for Dutch farmers [148]. In India, climate shocks have affected 23.2 million hectares of crops between 2019 and 2023, with 87% of districts vulnerable to drought. This has forced many rural households to sell livestock at distress prices and shift to casual labor, exacerbating their economic instability [12].

2.3 Agricultural early warning systems and anticipatory action

Early warning systems and anticipatory action are fundamentally interconnected, with the former acting as the foundation for the latter by providing advance notice of potential hazards that enables timely interventions. Early warning systems monitor and forecast risks such as droughts, floods, or locust outbreaks, offering the information needed to activate targeted measures before disasters occur. For example, a drought forecast can trigger the preemptive distribution of drought-tolerant seeds or the scheduling of cash transfers to protect vulnerable households. Effective anticipatory action relies on this connection, as it transforms early warnings into actionable steps that reduce the impact of hazards. In this section, we explore how existing agricultural early warning systems function and the types of anticipatory actions they facilitate, highlighting their critical role in mitigating weather-induced shocks to agricultural systems.

2.3.1 Early warning systems are a cost-effective means to mitigate climate risk

Given the large economic and human cost of climate-induced pressure on agricultural production systems, understanding, responding to, and mitigating the impact of extreme climate events is important for a wide range of stakeholders in the global food system, including farmers in both advanced and less advanced economies [149]. According to the Global Commission on Adaptation, there are several methods that can help agricultural communities become more resilient to the effects of climate change. These include making new infrastructure more resilient, improving dryland agriculture crop production, making water resource management more robust, and crop diversification. An example of the latter includes intercropping in Pakistan and Bangladesh, where farmers grow more than one crop in the same field to minimize risks associated with yield and income loss [150]. In Northern China and Nepal, farmers have adopted more resilient crop varieties such as cotton and coarse cereals [151, 152].

To increase farmers' resilience to climate shocks, a range of strategies, such as those mentioned above, must be deployed. However, evidence shows that investing in early warning systems (EWSs), which provide advance notice of potential hazards or disasters, may offer a high return on investment.

The United Nations Office for Disaster Risk Reduction (UNDRR) defines an early warning system as "an integrated system of hazard monitoring, forecasting and prediction, disaster risk assessment, communication and preparedness activities systems and processes that enables individuals, communities, governments, businesses and others to take timely action to reduce disaster risks in advance of hazardous events" [153]. Effective early warning systems should be multi-hazard, designed to detect a variety of hazards that may occur alone, simultaneously, or in a cascading manner.

These systems are built upon four critical pillars: (1) *Disaster risk knowledge*, which involves gathering and assessing disaster risk information; (2) *Detection, monitoring, and forecasting* of hazards; (3) *Warning dissemination and communication*, ensuring timely and actionable warnings reach at-risk populations; and (4) *Preparedness and response capabilities*, empowering individuals and organizations to take necessary actions before a disaster occurs [154]. These pillars form the foundation of the Early Warnings for All (EW4All) initiative, which aims to ensure global multi-hazard early warning coverage by 2027. At COP27 in November 2022, the United Nations announced an Executive Action Plan which aims to ensure that all people on Earth are protected by early warning systems in five years, through targeted investments of $3.1 billion between 2023 and 2027 [14].

Early warning systems that focus on agriculture specifically come in many different forms depending on their scope and focus. Some systems are designed to warn of specific hazards, such as drought, floods, or locust outbreaks, while others address multiple hazards simultaneously, reflecting the interconnected risks faced by farming communities. The “agricultural” aspect of these systems may refer to their target audience, such as smallholder farmers or large commercial operations, or to the specific agricultural impacts they monitor, like crop losses or disruptions caused by geopolitical or environmental factors. For instance, systems like the U.S. Drought Early Warning System (DEWS) focus on meteorological droughts, while multi-hazard systems, such as the Famine Early Warning Systems Network (FEWS NET), combine drought, rainfall, and other factors to assess agricultural impacts. The following examples illustrate the diversity of these systems, highlighting different approaches to addressing agricultural risks globally:

- The Food and Agriculture Organization’s (FAO) Global Information and Early Warning System (GIEWS) was one of the first global sources of information on food production and food security. GIEWS uses geospatial data to monitor the world’s food supply and demand conditions and identify weather-related issues that may impact food security in member countries [16].
- The European Commission’s (EC) Joint Research Centre (JRC) maintains the MARS Crop Yield Forecasting System (MCYFS), which predicts end-of-season crop production levels in Europe as the growing season progresses [5].
- In addition to MCYFS, the JRC also maintains the ASAP (Anomaly hot Spots of Agricultural Production) early warning system which aims to provide timely information about possible crop production anomalies globally, primarily based

on drought conditions [155].

- The GEOGLAM Crop Monitor is a tool that provides information on the current state and potential future impact of crops in major producing and trading countries, as well as the agro-climatic factors that may affect production. It is part of the G20 Agricultural Market Information System (AMIS) and focuses specifically on wheat, maize, rice, and soybean. The information provided by the GEOGLAM Crop Monitor is based on multiple sources and is intended to be a consensus assessment of global crop conditions [156].
- The Somalia Water and Land Information Management (SWALIM) project, established by FAO, developed the Combined Drought Index (CDI) to monitor drought conditions in Somalia. The CDI integrates hydro-meteorological data to assess drought severity and assist in preemptive response measures across vulnerable regions [157].

Early warning systems can monitor both slow-onset events, such as droughts, and rapid-onset crises, such as floods or storms. The lead time associated with each type of hazard can result in different warning message strategies or anticipatory actions being taken.

Over the years, evidence has emerged that early warning systems are a cost-effective means of building climate resilience. One study estimates that investments of around $1 billion per year into early warning systems and disaster response systems in developing countries could deliver benefits ranging from $4 to $36 billion [13]. Another study shows that preparedness investments have saved an average of $2 in humanitarian cost for every $1 spent, while other research shows that early warning systems have a median benefit-to-cost ratio of 5:1 [158]. In Ethiopia, a drought early warning system that reduces livelihood losses and dependence on assistance has a benefit-cost ratio of 3:1 to 6:1. The benefit-cost ratio of improving national hydrometeorological services in developing countries is also in the range of 4:1 to 36:1. In line with one of the goals of the United Nations' Sendai Framework for Disaster Risk Reduction, increasing investment in early warning systems would contribute to greater availability and access to disaster risk information, which is necessary for attaining the sustainable development goals [159].

2.3.2 Early warnings systems in practice: messaging and anticipatory action

The cost-effectiveness of agricultural early warning systems is achieved through two main avenues: direct farmer communication and support for anticipatory actions by organizations. Timely, actionable information helps farmers reduce losses by adjusting harvests, planting schedules, or protecting livestock ahead of adverse weather. Simultaneously, these systems enable governments, NGOs, and other agencies to allocate resources and launch interventions proactively. By triggering early actions like cash transfers, such systems can prove more effective than delayed, post-disaster aid, which often arrives too late [160, 161].

Dissemination of early warning messages

Specific examples of the cost-effectiveness of early warning systems have been documented in farming communities in both lower-income and higher-income nations. Take for example wheat rust, a disease which spreads via fungal spores dispersed by wind and has exacerbated severe crop losses in Ethiopia. One study developed an early warning system which helped policymakers allocate limited stocks of fungicide to combat the spread of wheat rust during the 2017 and 2018 wheat seasons in Ethiopia. Wheat rust alerts and agronomic advisories were sent via short messages to 10,000 development agents and 275,000 smallholder farmers who rely on wheat for economic livelihoods and subsistence [17].

A different program called the Scaling up Early Warning and Anticipatory Action for Agriculture and Food Security Project (EWAA) provided early warning messages to farmers in Zimbabwe to help them take action to protect their crops, livestock, and assets. For example, one farmer said that she was able to quickly harvest her sorghum before it was damaged by rain after receiving an SMS alert about the forecasted weather. Another farmer reported that they decided to delay applying fertilizer when they were notified through SMS that there would be high temperatures and no rain in the coming days. Thus, the SMS alert system appears to have had some impact on the short-term farming practices of the targeted group. In total, 13,790 recipients received early warning messages through SMS and other formats during the 2021/22 agricultural season [18].

The United Nations Development Programme (UNDP) and the Government of Sri Lanka implemented an early warning system that provides agronomic advisories via WhatsApp and Facebook to Sri Lankan field officers, who in turn circulate alerts to farming communities during cultivation planning meetings. Regional radio programmes and YouTube channels also disseminate the messages [162].

Agricultural communities in Ethiopia and Nicaragua have benefitted from community-based actions informed by early warnings and forecasts, both before and during the agricultural season. While extreme events such as drought depressed yields for all crops, farmers with access to early action advice were more able to mitigate impacts, organized in accessing relief and recovered more effectively [163].

In some regions, there is also evidence that households are willing to pay for the alerts from early warning systems. In Nepal's flood-prone Lower Karnali River Basin, early warning systems (EWS) have demonstrated significant cost-effectiveness in reducing flood impacts. A study found that households saved an average of NPR 117,027 (USD 1,083) in movable property, livestock, vehicles, and health costs during floods, with a benefit-cost ratio of 24 to 73 across different scenarios. Additionally, 98% of respondents expressed willingness to pay an annual fee of NPR 79 (USD 0.70) to fund community-managed EWS, sufficient to cover system maintenance and operating costs [164].

Anticipatory action for rapid-onset and slow-onset agricultural disasters

Anticipatory action is defined as "a set of interventions that are carried out when a hazard poses imminent danger based on a forecast, early warning or pre-disaster risk analysis" and is taken "before an anticipated disaster to mitigate its impact on people, assets and infrastructure that are likely to be affected" [20].

Public sector organizations such as the World Food Programme (WFP), the International Federation of Red Cross and Red Crescent Societies (IFRC), the Food and Agriculture Organization (FAO), and government agencies typically lead efforts in implementing anticipatory action. These actions can range from short-term interventions, like cash transfers and asset protection, to longer-term measures, such as distributing drought-resistant seeds or vaccinating livestock. Anticipatory action operates in the critical window between preparedness and response in the broader disaster risk management cycle. While preparedness builds general readiness, anticipatory action is activated by specific forecasts, allowing targeted measures just before a disaster strikes. This proactive approach reduces the need for expensive post-disaster responses and strengthens overall disaster response by ensuring communities are better equipped to face the imminent threat [64].

Anticipatory action is initiated when an early warning or forecast crosses a predefined threshold (for instance, *"release funds to region R if expected drought conditions sur-*

pass threshold T"). These thresholds are part of an Early Action Protocol (EAP), which guides the steps taken to protect communities before a disaster strikes. The EAP outlines the specific actions to be taken, the roles and responsibilities of each organization involved, and the financial mechanism that releases the necessary funds once the trigger is activated. For example, the Drought Early Action Protocol (EAP) used by the Zimbabwe Red Cross is based on forecast and observational data predicting drought hazards. It ensures that early actions, such as cash transfers or the distribution of drought-tolerant seeds, are deployed promptly to mitigate the impact of droughts [165].

Anticipatory action can be implemented in response to both rapid-onset and slow-onset hazards. For rapid-onset disasters, such as floods or cyclones, anticipatory actions are typically triggered just days or even hours before the event. In contrast, slow-onset hazards, like droughts or extreme winter conditions, offer longer lead times, sometimes allowing for interventions months ahead. For slow-onset disasters, phased triggering mechanisms are often used to align anticipatory actions with forecast lead times. These mechanisms typically include a soft trigger at the seasonal scale, providing early indications of potential risks, followed by a hard trigger closer to the event based on short-term forecasts [166]. For example, subseasonal-to-seasonal (S2S) forecasts can provide initial warnings months in advance, allowing for preparatory actions like distributing drought-resistant seeds, while short-term forecasts refine these actions as the hazard approaches. This phased approach enhances the efficiency and timing of interventions, reducing the overall impact of slow-onset hazards. The ability to act early in either rapid-onset or slow-onset scenarios is critical in protecting lives and livelihoods [64].

The Food and Agriculture Organization has compiled a series of case studies where early warning systems helped trigger anticipatory actions across various regions. For instance, in Kenya, anticipatory actions protected livestock, resulting in cows producing nearly an additional liter of milk per day, 80% of which was consumed by households, primarily benefiting children under 5 [21]. In Sudan, early warnings triggered interventions such as livestock vaccination and feed distribution, preventing severe losses. Each household saved on average one animal, avoided significant losses, and benefited from increased milk production, improving both nutrition and livelihoods during the drought [22]. In Mongolia, early feed distribution ahead of a winter "dzud" (cold wave) reduced livestock mortality by the equivalent of four cattle per household [25]. Similarly, in the Philippines, stress-tolerant seeds and micro-irrigation equipment distributed ahead of El Niño–induced drought allowed families to maintain food production and avoid negative coping strategies [23]. In Madagascar, the distribution of vegetable seeds and irrigation equipment increased production sixfold in some cases, significantly reducing crop losses

[24]. In Afghanistan, cash and livestock interventions during the La Niña–induced drought significantly increased household food consumption, boosting milk production [23]. Finally, waterproof drums distributed ahead of floods helped households save critical goods, with over fifty percent of beneficiaries noting that the drums were used to save food items with an average market value of about $9 per household [23]. Additional examples of impact can be found in the World Food Programme's *Anticipatory Action Year in Focus 2023* report [167].

2.4 Towards "optimal" anticipatory action

Anticipatory action is increasingly being recognized as a proactive way to mitigate the most severe impacts of crises by acting early on forecast information. Yet, effectively reaching the most vulnerable in a context of finite humanitarian resources remains a pressing challenge. Research has highlighted challenges related to the risk of blanket approaches, where aid is distributed broadly rather than strategically targeting those most in need [168], and in determining the right balance between in-kind and cash assistance, as well as in fine-tuning the timing and scale of distributions [64]. Put simply, "anticipatory action models tested to date face challenges in effectively identifying where and when to act, whom to pay, and how to quickly disseminate effective early warning messages and deliver cash to widespread geographies before floods hit" [169].

Targeting challenges have been a persistent issue across both anticipatory action pilot programs and standard humanitarian response efforts worldwide. In Bangladesh, aid agencies struggled to coordinate targeting strategies in time [170], while in Ethiopia, ensuring sufficiently granular targeting in drought-affected regions proved difficult [171]. In South Sudan, aid workers faced painful decisions about prioritization amid widespread need [172], and in Bangladesh, blanket approaches to assistance often left the most vulnerable at a disadvantage, particularly among displaced populations [173].

A senior UN OCHA employee highlighted the extreme end of these challenges in the context of flooding in South Sudan, where limited funding had forced particularly difficult trade-offs in anticipatory action [172]:

> *We only have one pot of money to respond to all shocks—climate, conflict, and so forth. The funding comes in and we are trying to allocate based on level of needs. There is no easy way to do this, even though we created a methodology to help prioritise. At the cluster and inter-cluster level, we have had to reallocate resources from one area to another. We have been clear from the beginning that we are taking from the very, very vulnerable to give*

> *to the very, very, very vulnerable.*

Efficient allocation of limited humanitarian budgets is essential for equity and impact. Targeting aid based on risk ensures it reaches those most in need while maximizing coverage within budget constraints [174]. This is especially crucial as evidence shows large-scale aid, including World Bank funding, is often misallocated, failing to prioritize the most vulnerable even amid budget cuts [175]. At the same time, inclusive approaches are necessary to ensure marginalized groups are not left behind [169].

Case studies in humanitarian response highlight the effectiveness of better targeting: in Togo, refined aid allocation strategies during COVID-19 could reduce exclusion errors by up to 21% [176], while in Nigeria, optimizing resource distribution could prevent approximately 84,000 deaths and 15.7 million malaria cases over five years [177]. In some contexts, severe funding shortages have made transitioning from blanket aid distribution to targeted assistance not just beneficial, but unavoidable, such as in Syria, where humanitarian agencies were forced to adopt stricter prioritization measures due to resource constraints [178].

As we will show in this section, optimization is an important technique that can help address gaps in aid targeting by allocating limited resources, such as cash, water, and seeds, in a way that aligns with the specific needs of affected communities. This approach makes interventions more targeted and resource-efficient, enhancing timeliness and impact.

Mathematical optimization models offer a structured approach to these resource allocation challenges, simulating entire systems under constraints like resource availability, transportation logistics, and regional needs [179]. Such models support the simultaneous optimization of factors such as cash disbursement amounts, timing for maximum impact, prioritization of resources like food or fodder in specific regions, and dynamic adjustments as conditions evolve. Moreover, embedding early warning system outputs into impact-oriented optimization models ensures that early warnings directly inform decisions that minimize humanitarian costs or maximize recipient benefits, as opposed to just producing alerts. This approach transforms early warning system outputs from general predictions into valuable inputs to downstream decision models.

2.4.1 Making better decisions with operations research

The need for structured approaches to complex decision problems is the essence of *operations research*, a field which emerged in the 1940s. Operations research, a discipline at the intersection of management science, optimization, and decision science, emerged as the scientific study of making good decisions in complex systems [180, 181]. The term "operations" originally referred to military operations, as the field developed during World War II when British scientists like Patrick Blackett and his team pioneered methods to optimize military logistics, improving tactics by applying scientific principles to operational problems [181]. Around the same time, in the Soviet Union, Leonid Kantorovich developed mathematical methods for optimizing resource allocation in the Soviet Union under various constraints in his 1939 work, *Mathematical Methods of Organizing and Planning of Production* [182]. Kantorovich and Dutch economist Tjalling Koopmans were awarded the Nobel Prize in Economics in 1975 for their contributions to optimal resource allocation [183].

Kantorovich's work laid the foundation for linear programming, arguably the most important method in operations research. Linear programming is a mathematical technique used to determine the optimal outcome in a system with limited resources by maximizing or minimizing an objective (such as cost or output) within a set of linear constraints (such as labour or materials). In this context, the term "programming" refers not to computer programming but rather the formulation of a detailed plan or schedule of activities. The method was rediscovered and given its current name in 1947 by George Dantzig, who introduced the simplex algorithm, an efficient way to solve these linear optimization problems [184]. Dantzig's contributions to linear programming paved the way for its application across a wide range of fields, including transportation, manufacturing, and resource management.

These multiple strands of wartime operational research, resource allocation in economies, and production planning converged into what we now recognize as the field of operations research, with wide-ranging applications in industry, logistics, and public policy.

Linear programming paved the way for a wide range of algorithms collectively known as **mathematical programming**. These include **integer programming**, which addresses decisions involving discrete variables, such as whether to undertake specific investments or allocate resources to distinct regions [185]. **Stochastic programming** allows decision-makers to account for uncertainty in variables by optimizing over multiple scenarios [186], while **network optimization** focuses on optimizing flows within networks, such as transportation or communication systems [187]. **Queuing theory** models systems with waiting lines, optimizing service efficiency in settings like healthcare or telecommunications, and **dynamic programming** tackles problems involving sequential decisions over time, commonly used in inventory and resource management [188]. Furthermore, **robust optimization** provides solutions that remain effective under uncertain or variable conditions, offering greater resilience in planning

[189, 190, 191]. **Nonlinear programming** extends these techniques to problems where the relationships between variables are not linear, allowing decision-makers to optimize in more complex systems involving nonlinear constraints or objectives [192]. Additionally, **multi-objective optimization** techniques, such as the ant colony optimization and NSGA-III algorithms, provide frameworks for finding optimal solutions in scenarios with several competing objectives, balancing trade-offs between them [193].

In addition to these mathematical programming approaches, other decision-making frameworks, such as **data envelopment analysis** [194], enable efficiency evaluation across multiple decision-making units [195]. Meanwhile, **multi-criteria decision analysis** (MCDA) tools, such as TOPSIS and VIKOR, help decision-makers evaluate multiple options systematically [196]. Data envelopment analysis (DEA) and multi-criteria decision analysis (MCDA) differ from mathematical programming techniques in that they focus on evaluating performance across multiple criteria or units, rather than directly optimizing an objective function within a set of constraints.

Collectively, these extensions of linear programming and related optimization methods have significantly broadened the set of tools available for aiding decision-makers in real-world planning scenarios. The right optimization approach depends on the specific nature of the problem: whether it involves uncertainty, multiple objectives, discrete decisions, or dynamic systems, as well as factors like the scale of the problem, available data, and the need for robustness in the face of variability [197]. The appeal of mathematical programming lies in its structured approach and transparency. The solutions it provides are not only optimal according to a set of criteria but also interpretable, making mathematical programming particularly relevant in fields requiring both precision and accountability.

As mathematical programming has evolved, it has been applied in fields as diverse as planning the pre-positioning of personnel in fire departments [198], water resource management [199], micro-grid optimization [200], national energy system design [201], liquid natural gas supply chain planning [202], construction planning [203], and production scheduling in mining operations [204].

While mathematical programming is a powerful tool for solving complex decision problems, it is not without limitations. First, misspecified objectives or constraints can lead to models that omit critical imperatives or misrepresent them, resulting in flawed recommendations [205]. Second, overly simplistic assumptions, such as fixed decision boundaries or inadequate feedback loops, may fail to capture the dynamic and interconnected nature of real-world systems [206]. Third, these models can struggle to accommodate the divergent goals and priorities of multiple stakeholders, potentially leading to solutions that favor narrow interests at the expense of broader equity [207]. Fourth, the quality of optimization outputs is heavily dependent on data, meaning issues like mislabeling, mismeasurement, or incomplete datasets can introduce bias and imprecision that undermine the model's validity [208]. Fifth, reliance on approximation or relaxation techniques, while computationally efficient, may yield results that lack formal guarantees, reducing their reliability in critical applications [209]. Finally, treating optimization outputs as definitive without incorporating the necessary contextual information can lead to ethically questionable or impractical decisions [210]. Addressing these challenges requires not only technical rigor in model development but also a deep understanding of the decision-making context to ensure that mathematical programming delivers actionable and equitable solutions.

2.4.2 Operations research in humanitarian aid and agriculture

In addition to the broader applications described above, mathematical programming has also been applied in disaster planning, though typically in a reactive, post-disaster context, as well as in agricultural settings. Examples are given below.

Applications in humanitarian response to disasters Mixed-integer linear programming (MILP) has been used to optimize vehicle routes and minimize aid distribution times in a case study in Chile [211]. Similarly, in Korea, MILP was applied to maximize the number of evacuees transported via helicopters based on aircraft capacity and urgency levels [212]. In another case, a stochastic programming model was developed using mobile phone location data to enhance disaster preparedness and post-disaster response [213]. A different study used MILP to design an efficient relief network for Iran, optimizing relief operations in the wake of disasters [214]. Other applications of MILP include the development of a two-echelon pre-disaster relief network for rare disasters in Turkey [215], and a bi-objective MILP model aimed at minimizing the total completion time of rescue operations during natural hazards in Iran [216]. Additionally, mixed-integer non-linear programming (MINLP) was applied to real-time location data to optimize rescue plans for trapped individuals [217], while MILP was used in Brazil to allocate shelters and distribute relief supplies during emergencies [218]. Linear programming was used to reduce procurement and logistics costs for 12 million mosquito nets in Ivory Coast [219]. Moreover, robust optimization has been used to address uncertainty in food procurement prices during humanitarian emergencies, demonstrating improved outcomes compared to traditional approaches in scenarios like the Syrian food aid operation [220].

Applications in commercial supply chain planning for agriculture Optimization models are widely applied in agricultural supply chains to address uncertainties and improve decision-making. For example, a recent review outlined the progress in optimization for fresh fruit supply chains under uncertainty [221]. Mathematical programming models have also been used to plan the harvest of fresh tomatoes in multi-farmer supply chains under uncertainty [222]. In the soybean industry, a two-stage stochastic linear programming model with fixed recourse was proposed for tactical planning under take-or-pay contracts [223]. In Mexico, a MILP model optimized the bioethanol supply chain by determining the optimal number and location of storage centers, biorefineries, and mixing plants, as well as the flow of biomass and bioethanol [224]. MILP was also used to optimize production scheduling in a grain processing facility in Uruguay, minimizing costs while considering practical constraints such as shelf life and storage capacity [225]. In Chile, a mathematical model was developed to coordinate harvest planning between farmers and wholesalers, optimizing decisions under uncertainty in crop yield, prices, and demand [226]. In another study, a mixed-integer programming model incorporated vehicle routing problem concepts to optimize the transportation of harvested products from fields to a central depot [227]. Beyond purely commercial objectives in supply chains, mathematical programming has also used to address environmental sustainability in agriculture. For instance, a bilevel linear programming model was developed to formulate subsidy policies aimed at minimizing the environmental impact of agricultural activities [228]. Another study employed MILP to optimize a closed-loop supply chain for perishable agricultural products, balancing cost minimization, CO_2 emissions reduction, and responsiveness to demand [229]. In Brazil, a multi-product MILP model optimized fertilizer logistics, minimizing costs and CO_2 emissions while strategically placing fertilizer mixing factories [230].

Applications in long-term agricultural resource allocation Optimization models have also been used to address longer-term structural issues in agricultural resource allocation. For example, a case study in China employed an optimization model for agricultural water and land resource allocation under uncertainty [231]. In Bangladesh, a mixed-integer non-linear programming model was formulated to optimize environmental sustainability for flood control, drainage, and irrigation projects in deltaic regions [232]. A multi-level multi-objective stochastic programming model was developed to formulate sustainable water-allocation schemes for arid agricultural regions [233]. Another multistage stochastic programming model optimized farmland irrigation management under uncertainty, considering crop prices, precipitation, and irrigation water availability to maximize farmers' profits [234]. In Rajasthan, India, linear programming was used to identify the most efficient crop combinations based on resource availability, labor, water, and land, with particular attention to the benefits of crop rotation [235].

Applications in responding to acute risks Several studies have focused on optimizing resources based on short-term uncertainties in agriculture. For instance, a multi-objective linear programming input-output method was applied to allocate water across various economic sectors during drought conditions [236]. In another study, a hydroclimatic monitoring system was integrated with linear programming to optimize agricultural net benefit under scenarios of rainfall delay or reduction [237]. In Haryana, India, linear programming was used to determine optimal land and water resource allocations, maximizing net returns from an irrigated area [238]. Similarly, a model for managing drought risk through optimized water distribution under different drought scenarios was developed and tested in Heilongjiang, China [239].

2.4.3 Opportunities for optimizing anticipatory action in agriculture

Similarly to the examples provided above related to humanitarian aid, anticipatory action in agricultural contexts involves complex decisions that must balance multiple, often competing factors under constraints [240]. Mathematical programming methods such as linear programming and stochastic linear programming, are designed to address these types of problems.

For instance, consider a drought scenario where cash transfers and drought-tolerant seeds need to be distributed across multiple regions. The objective is to maximize the impact of these interventions while minimizing costs and ensuring timely deployment. A mathematical program, such as a mixed-integer linear program, could model this by setting an objective function that maximizes the number of beneficiaries reached or minimizes the cost of aid distribution. Constraints can include limited budgets, transportation capacity, and time windows for intervention (e.g., ensuring seeds are distributed before the planting season begins or cash transfers are made before household resources are exhausted). A stochastic linear program can further enhance this model by incorporating uncertainty, such as varying rainfall forecasts or fluctuating market prices for seeds, allowing decision-makers to optimize across multiple scenarios [241]. The model can dynamically adjust as new forecast data arrives, ensuring that both resources and timing are optimized across different phases of the interven-

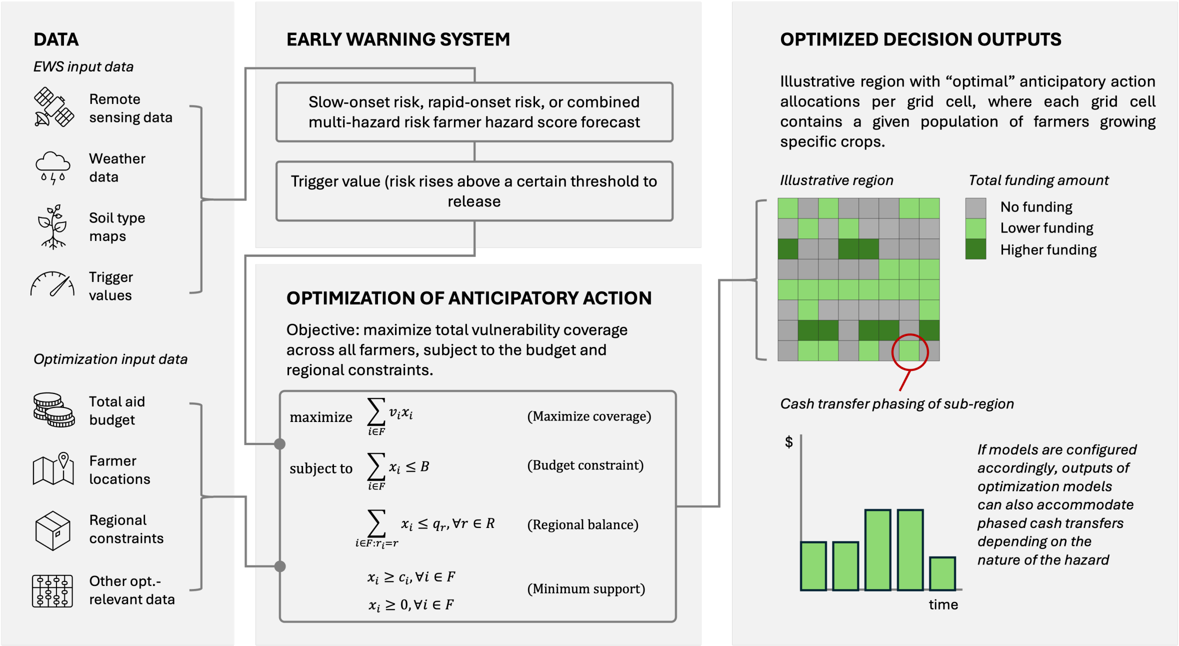


Figure 2.2: Schematic framework illustrating one possible configuration linking early warning systems to optimized anticipatory action. In this example, the early warning system utilizes data inputs such as remote sensing, weather data, and soil type maps to trigger risk-based thresholds, forecasting hazard scores for farmers. Optimization models allocate resources (e.g., cash transfers) by maximizing vulnerability coverage under constraints such as total budget, farmer locations, and regional balance. The outputs are illustrative decision maps indicating funding levels across grid cells, with the potential for phased cash transfer strategies based on hazard dynamics. Other configurations linking early warning systems to optimization models could be envisaged, depending on the specific use case and decision-making context.

tion, from initial preparation to final deployment. By explicitly framing the problem in terms of objectives, constraints, and uncertainty, mathematical programming offers a structured, data-driven approach to making complex anticipatory action decisions more efficient and impactful (see example A and Figure 2.2).

Example: A mathematical program for cash transfer payouts to farmers in drought-affected areas

Consider a government planning to allocate cash transfers to farmers who are expected to be impacted by drought, based on predictions from an early warning system. The goal is to distribute funds optimally, given a limited budget and the varying levels of vulnerability among farmers.

Let us define the following sets and parameters:

- **Sets:**
  - $F = \{1, 2, \ldots, N\}$: the set of farmers.
  - R: the set of all regions.
- **Parameters:**
  - x_i: the amount of cash transfer to be allocated to farmer $i \in F$.
  - v_i: the drought vulnerability score of farmer i based on EWS outputs, calculated by overlaying a probabilistic drought severity index, such as the Standardized Precipitation Index (SPI) or the Standardized Precipitation-Evapotranspiration Index (SPEI), with a spatial grid of farmer locations to capture localized risk levels and exposure.
  - c_i: the minimum cash required to adequately support farmer i.
  - B: the total budget available for cash transfers.
  - r_i: the region where farmer i is located.
  - q_r: the maximum allowable total cash transfer for region $r \in R$, which accounts for regional balance constraints.

The objective is to maximize the total vulnerability coverage across all farmers, subject to the budget and regional constraints. This can be formulated as the following linear program:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i \in F} v_i x_i \\
\text{subject to} \quad & \sum_{i \in F} x_i \leq B \quad \text{(Budget constraint)} \\
& \sum_{i \in F : r_i = r} x_i \leq q_r, \quad \forall r \in R \quad \text{(Regional balance constraint)} \\
& x_i \geq c_i, \quad \forall i \in F \quad \text{(Minimum support per farmer)} \\
& x_i \geq 0, \quad \forall i \in F
\end{aligned}$$

In this formulation, the decision variables x_i represent the cash transfers allocated to each farmer i. The objective function $\sum_{i \in F} v_i x_i$ seeks to allocate funds to maximize the total vulnerability-weighted support, ensuring that farmers with higher vulnerability receive more resources. The constraints ensure that the total budget is not exceeded, that each region receives a fair distribution of aid via the regional balance constraint (q_r), and that each farmer receives at least the minimum cash required (c_i). Additionally, a realistic constraint for logistics or delivery capacity can be added, such as limiting the maximum amount of cash that can be distributed based on logistical constraints in each region or incorporating path constraints. In this example, a hypothetical machine learning model (such as XGBoost) can be used to predict drought severity for each region, generating a vulnerability score for each farmer based on their location and relevant features (e.g., historical climate data, soil conditions). There are an exponential number of possible allocations of cash transfers, rendering it impossible to enumerate and evaluate the benefit of every possible allocation. However, this problem can be efficiently solved using mathematical programming techniques, making it a powerful tool for integrating early warning system data into equitable and targeted decision-making processes. Illustrative extensions of the model are provided in the supplementary information.

While the core benefits of mathematical programming revolve around resource allocation, these techniques can be applied across multiple stages of the process, depending on the decision-maker's priorities. Examples are provided in Table A.1. Optimization can address decisions at the level of protocol design, where triggers for interventions like cash transfers are set based on forecast data, or in real-time adjustments to the timing and scale of aid delivery as new information becomes available. It can also be used to optimize the mix of interventions, balancing immediate cash relief with longer-term resource needs, such as seed distribution or infrastructure like water storage. Each stage, from pre-disaster planning to post-disaster recovery, offers distinct opportunities for optimization, with models tailored to maximize impact while minimizing costs, logistical complexities, and uncertainty.

Table 2.1: Optimization opportunities tied to anticipatory action enabled by early warning systems (adapted from [242]). This table highlights how mathematical programming can support anticipatory action decisions based on early warning systems. For specificity, the illustrative examples focus on the perspective of rice farmers in South Asia and the challenges they face.

Stage	**Intervention Type**	**Illustrative examples of mathematical programming applications**
Pre-planning and scenario readiness	Pre-positioning aid for anticipatory action	**Objective:** Optimize the allocation and pre-positioning of resources, such as wind barriers or drought-resistant seeds, to ensure rapid and effective responses to forecasted hazards. For instance, strategically place wind barriers in regions with a high probability of damaging winds to protect rice seedlings, or stockpile drought-resistant seeds in areas projected to experience rainfall deficits. This approach improves the speed and efficiency of anticipatory action. **Constraints:** Storage and transport capacity, budget availability, and lead time for deployment. **Required Data:** Historical hazard profiles, real-time seasonal forecasts, resource vulnerability indices, and logistics network capabilities.
	Trigger threshold optimization	**Objective:** Identify and dynamically adjust EWS trigger levels to balance the risks of over-action (exhausting resources prematurely) and under-action (allowing preventable losses). For example, setting a low drought severity index (DSI) threshold, such as 40, may lead to unnecessary cash transfers that deplete resources too early, while a high threshold, such as 80, could delay critical interventions, resulting in significant household losses. Optimizing these thresholds ensures timely and effective actions tailored to evolving hazard conditions. **Constraints:** Budget limits, lead time for resource mobilization, and regional equity considerations. **Required Data:** Historical drought impact data, real-time hazard forecasts, household vulnerability indices, cost-benefit analyses of early interventions, and adaptive decision rules for seasonal updates.

Table continued on next page

Table 2.1 continued from previous page.

Stage	**Intervention Type**	**Illustrative examples of mathematical programming applications**
Rapid-onset hazard response	Early harvest and crop protection	**Objective:** Minimize rice crop losses by timing interventions based on short-term EWS outputs. For instance, in the event of forecasted hailstorms within 5–10 days [243], recommend early harvesting if crops are mature enough, or deploy protective measures such as temporary coverings to shield vulnerable rice seedlings. **Constraints:** Farmer labor availability, access to protective materials, transport logistics, and market readiness. **Required Data:** 5–10 day hailstorm forecasts, crop maturity data, availability of protective materials, and local road network conditions.
	Emergency input deployment	**Objective:** Rapidly support farmers in replanting after severe weather events, such as high winds or hailstorms, that destroy rice seedlings. For instance, provide immediate cash disbursements to enable farmers to purchase replacement seedlings from local markets or ensure rapid delivery of seedlings to affected areas for replanting early in the season. **Constraints:** Availability of replacement seedlings, timely cash transfer mechanisms, transportation logistics, and market access. **Required Data:** Real-time EWS alerts, local seedling supply data, farmer vulnerability indices, and transport network conditions.
Slow-onset hazard mitigation	Phased risk-based transfers	**Objective:** Gradually deploy cash transfers to stabilize households as conditions deteriorate. For example, allocate larger payments during peak drought periods while reserving resources for later needs [64]. **Constraints:** Fixed total budget, regional equity, and timing constraints. **Required Data:** Seasonal drought forecasts, household income data, and effectiveness of past transfers.
	Proactive water management for paddy farming	**Objective:** Optimize the distribution and replenishment of water tanks to sustain rice paddies during delayed monsoons or prolonged dry spells. For instance, strategically allocate water tanks to high-vulnerability areas with limited groundwater availability, enabling farmers to irrigate crops during critical growth stages [244]. **Constraints:** Tank availability, replenishment schedules, transport capacity, and equitable distribution across regions. **Required Data:** Seasonal drought forecasts, village-level crop vulnerability indices, water tank storage capacities, transport network information, and real-time monitoring of water use.

Beyond optimizing outcomes, mathematical programming can also bring additional transparency to decision-making, which can be opaque in large-scale interventions. One study highlights that “the design of any anticipatory action projects should consider the practicalities of explaining the transfer amount to local authorities, community leaders, people not targeted by the interventions, and other stakeholders who are part of the approval process for the transfer value” [20]. Targeting vulnerable households for anticipatory action before disasters is fraught with challenges, particularly when aid decisions risk creating tensions among communities. In Madagascar, a cyclone season affecting nearly a million people exposed the difficulty of selecting recipients when resources were insufficient to support all those impacted, necessitating the use of risk pooling. Similarly, in Bangladesh, local NGOs emphasized the need for sensitive targeting to avoid conflict, highlighting that anticipatory action often contrasts with traditional aid operations, requiring careful planning and community engagement to ensure fairness and trust [41].

Mathematical programming forces decision-makers to clearly articulate their goals and constraints, such as maximizing the number of beneficiaries within a fixed budget or minimizing logistical costs. This clarity not only helps explain why certain decisions are made but also makes the decision-making process more defensible, as the assumptions embedded in the model are made explicit. Optimization models allow for the transparent inclusion of potentially conflicting stakeholder objectives [245], as well as explicit considerations of equity [246], enhancing clarity in balancing priorities like minimizing costs or ensuring fairness. Increased trust between stakeholders due to greater transparency may play a role in improving the overall effectiveness of subsequent humanitarian operations [247].

One example of how transparency can be applied is in dealing with the uncertainty of trigger values, the critical thresholds used to initiate anticipatory actions. Stakeholders may disagree on the appropriate course of action when trigger values are too low (e.g., initiating cash transfers prematurely when a drought is forecast but not severe) or too high (e.g., delaying action until the drought has already caused significant harm). As one study notes, “the negative consequences of acting in vain are significant and must be factored in” [40]. Optimization can help address these challenges by openly articulating the trade-offs between competing priorities, such as maximizing the prevention of harmful events while minimizing unnecessary expenses.

For instance, in a flood scenario, optimization could show how reallocating aid based on updated rainfall forecasts can reduce wasted resources and ensure the hardest-hit regions receive priority. Similarly, in a drought, optimization models could adjust cash transfer amounts dynamically, ensuring the most vulnerable households are supported while staying within budget constraints. By making assumptions, constraints, and objectives explicit, optimization not only provides a framework to recalibrate decisions as conditions evolve but also fosters trust among stakeholders through open and transparent decision-making processes. This approach ensures that even when forecasts are uncertain or imperfect, interventions remain both defensible and impactful.

2.5 The path forward

To enhance the effectiveness of agricultural early warning systems and anticipatory action, we must focus on improving existing systems and integrating them with optimization frameworks. This section outlines key areas for advancement, aiming to build more resilient and adaptive agricultural systems in the face of increasing climate risks.

2.5.1 Improve accuracy, lead time, and adoption of early warning systems

Early warning systems offer valuable input to anticipatory action but require improvements in areas such as data quality and modeling accuracy to more reliably guide optimization-based decisions that enhance both the timing and precision of interventions. Research to improve early warning systems might involve efforts to:

Address gaps in data and methods. One study examined 8 major global and regional agricultural monitoring systems, and found that gaps in data included crop calendars, cropland maps, satellite imagery, mobile phone data, and reliable meteorological data. Gaps in methods include the need for better crop yield and crop production forecasts, the operationalization of research methods at scale [16], the integration of near-real-time information to allow crop yield forecasting models to deal with increasingly unprecedented weather extremes [248], and methods that allow users to interpret why certain forecasts are being generated [249, 250]. Closing these gaps could yield early warning systems with greater accuracy in predicting agricultural production shocks and extended lead times for more proactive responses.

Address model performance issues and regional biases. Another study conducted a statistical analysis on outputs of the Famine Early Warning Systems Network (FEWS NET) for 25 African countries from 2009 to 2020. FEWS NET projections are a widely used resource for monitoring, planning, interventions, and resource allocation. The analysis showed that the overall system accuracy was 84%, but drops sharply with ascending food insecurity, biasing toward over-projection [4]. Moreover, unanticipated weather shocks also hindered projection accuracy [32]. Another evaluation of FEWS NET showed that the system provided more accurate warnings in the western parts of Ethiopia than the generally food insecure northeastern regions, likely due to insufficient information. Early warning systems have also seen performance issues in Europe. One study examined 19,980 crop yield forecasts published by member states of the European Union across seven crops, and found that forecasts systematically underestimated yields for wheat, rapeseed, and sugar beet crops [5]. Regional discrepancies in accuracy may trigger unfair anticipatory actions for populations in these regions; false pessimism may trigger expensive humanitarian response, potentially reducing funding for future crises, while false optimism may fail to trigger a response when it is most needed [33]. False alarms in early warning systems may also lead to the ’cry-wolf effect,’ where individuals become less responsive to future warnings due to reduced trust in the system, potentially hindering timely adaptation decisions and risking greater losses during actual crises [34].

Address user adoption issues. Depending on the region, agricultural stakeholders may not interact with the outputs of early warning systems in the intended manner. In one region of Ghana, for instance, one study showed that just 60% of farmers paid no or little attention to early warning drought information [251]. Furthermore, studies in Ghana have also shown gender-based discrepancies in farmers’ willingness to pay for climate-related warnings, with women less likely to pay for such information [252]. Previous research has highlighted the importance of setting application design standards in early warning systems to enhance usability by anticipating the complexity of possible user demands [253]. Other research has also explored the role of gamification to encourage citizens to share the outputs of early warning systems with each other to foster wider adoption [254].

Develop objective-driven early warning systems for targeted decision-making. Traditional agricultural early warning systems like ASAP and GIEWS serve as general portals for users to access weather and yield projections, providing broad overviews rather than specific actionable insights. However, evolving toward an impact-driven approach could strengthen the role of EWS in guiding anticipatory action [255, 256]. A promising example is the Netherlands Red Cross 510 initiative, which employs impact-based forecasting (IBF) to predict drought consequences on specific outcomes, such as crop yield anomalies, directly linking these forecasts to pre-defined, region-specific early actions, such as resource allocation to mitigate food insecurity or prevent livestock losses. Such targeted, impact-oriented early warning systems could pave the way for more integrated planning and optimized decisions [165].

2.5.2 Integrate early warning systems with accessible optimization frameworks

While related fields have made progress in connecting forecasting systems with optimization frameworks [237], significant advancements are still required to fully integrate early warning systems with accessible, effective optimization tools tailored specifically for anticipatory action. By strengthening these connections, early warning systems can more directly support data-driven decision-making, enabling timely, targeted interventions that maximize humanitarian impact. The following initiatives are proposed to advance this integration:

Develop modular optimization toolkits building on open-source packages. Create open-source toolkits with adaptable optimization templates for common scenarios such as resource allocation, logistics, and prioritization. Leveraging Python-based packages like PuLP [257] or Pyomo [258], these templates can provide flexible, high-level modeling tools without requiring extensive customization or mathematical expertise. An adjacent example in another field is the Route Optimization Tool (RoOT), a simple, open-source toolkit designed to reduce the need for advanced expertise in optimizing health product distribution. RoOT uses a modular interface that enables government logisticians to optimize routes for distributing health supplies, even in emergency settings like cyclone responses, by selecting options aligned with available vehicles and road conditions [259]. Each template in these toolkits should specify sections for incorporating early warning data and input from multiple organizations, ensuring that relevant anticipatory information and inter-agency considerations are embedded in each planning scenario. In addition, these toolkits should specify sections for incorporating early warning data and input from multiple organizations, and they should account for trade-offs in multi-hazard scenarios where responses to one hazard might conflict with another.

Ensure that the right stakeholders are involved in designing optimization–driven systems. To maximize the adoption and utility of anticipatory action systems, leveraging best practices in human-centered design (HCD) is crucial. HCD emphasizes iterative engagement with end-users throughout the design lifecycle to ensure systems are tailored to local needs. For example, a study on emergency preparedness in rural communities showed that involving users in the design phase enhanced both the usability and adoption of a mobile application by integrating their feedback into system features [260]. Similarly, research on humanitarian IT adoption highlighted the importance of co-designing tools with stakeholders and fostering trust, as this encourages practical and effective integration of new technologies into decision-making processes [261]. Finally, insights from digital decision support systems for population health demonstrate that engaging users early during needs assessment and consistently throughout prototyping and evaluation phases ensures that systems align with both cognitive and contextual requirements, thereby fostering greater adoption and impact [262]. Involving the right stakeholders is also critical to ensure that optimization only occurs for relevant anticipatory actions that are appropriate for the local context.

Build capacity in optimization for anticipatory action. Conducting targeted training sessions in optimization for anticipatory action practitioners can foster effective and confident application of these techniques in humanitarian contexts. Evidence from diverse fields underscores the value of using engaging, context-rich examples to enhance understanding and application of optimization methods. For example, role-playing games inspired by *Game of Thrones* have been used successfully to introduce linear programming concepts, allowing learners to solve fictional optimization problems like blending poisons and scheduling troops for battle [263]. In another case, optimization principles were taught using scheduling challenges from the 18-team Jupiler Pro League, Belgium's highest professional soccer division, which provided a real-world framework to make complex constraints and objectives tangible [264]. Finally, a study on wheelchair rugby demonstrates how predictive analytics can directly inform optimization, as player performance predictions are integrated to optimize lineups, a useful parallel to the integration of early warning systems with optimization in anticipatory action [265]. In the field of humanitarian aid specifically, optimization education could be embedded within existing educational initiatives like the Red Cross Red Crescent Climate Centre's simulation game, where participants, acting as farmers or donors, allocate limited resources to manage flood or drought risks [63].

Establish a standardized evidence base for optimization assumptions. Developing a centralized, standardized evidence base for optimization model parameters would enable practitioners to incorporate data-driven assumptions into their models. This resource would document the empirical impacts of specific interventions, allowing users to ground optimization assumptions in real-world outcomes and maintain transparency by tracing these assumptions back to empirical studies. For example, when building a model to allocate cash assistance in response to a forecasted drought, stakeholders might disagree on the potential impact of cash transfers on food security and asset protection. With access to an evidence base, practitioners could reference studies such as one that found cash transfers helped protect household assets and reduced reliance on high-interest loans [266], another that observed reduced debt accumulation and increased food security following cash interventions [267], or a study demonstrating the effectiveness of anticipatory actions in protecting livestock [268]. By consulting a curated database of impact studies, practitioners can set more reliable parameters, thereby enhancing the accuracy and credibility of optimization models for anticipatory action.

2.5.3 Concluding remarks

While optimization holds promise for enhancing anticipatory action, its application should be approached with caution. As one optimization practitioner advises, seeking to optimize processes that should not exist in the first place may risk encouraging interventions that lack meaningful impact [269]. Rather than viewing optimization as a comprehensive solution, it is better seen as a structured approach to improve decision-making in contexts that may benefit from a rigorous and flexible approach to planning. While optimization has proven useful in fields like supply chain logistics and emergency response, its value in anticipatory action will depend on carefully aligning methods with humanitarian objectives and ensuring that these tools are leveraged in a manner that remains adaptable to the complex needs of crisis management.

Supplementary information

This section provides additional extensions to the core cash transfer optimization model introduced in the main text. These extensions demonstrate how the model can be refined to address equity, uncertainty, multi-period planning, and logistical constraints. While not exhaustive, the goal is to illustrate how mathematical programming models can be used to flexibly and transparently incorporate new constraints and structures, adapting the model to align with practical policy objectives and operational considerations.

Farmer equity constraints

To promote fairness in allocating resources to farmers with similar vulnerability scores, constraints can be imposed to limit the disparity in cash transfers among farmers with similar vulnerability levels. Let δ be a threshold indicating “similar” vulnerability, and let ϵ be the maximum allowable difference in cash transfers for such farmers. A conceptual constraint is:

$$|x_i - x_j| \;\le\; \epsilon \quad \text{for all } i, j \text{ where } |v_i - v_j| \le \delta.$$

Since absolute values are not linear, practitioners typically linearize this via

$$x_i - x_j \;\le\; \epsilon, \qquad x_j - x_i \;\le\; \epsilon \quad \text{for all } i, j \text{ with } |v_i - v_j| \le \delta.$$

This ensures that farmers with similar vulnerability receive transfers that do not differ by more than ϵ.

Risk aversion in budget allocation

When drought forecasts are uncertain, the model can include a term that discourages large allocations to regions with high forecast uncertainty. Let σ_r denote the uncertainty in the forecast for region r, and λ be a risk-aversion parameter. The objective function can be modified to

$$\max \sum_{i \in F} \bigl(v_i - \lambda\, \sigma_{r_i}\bigr)\, x_i,$$

where σ_{r_i} is the uncertainty corresponding to the region of farmer i. A larger λ penalizes allocations to uncertain regions, prioritizing more reliable areas while still balancing overall vulnerability coverage.

Dynamic budget allocation across multiple periods

In some cases, planners may wish to allocate transfers over multiple time steps (e.g., months or seasons) as forecasts evolve. Let T be the set of time periods, and B_t the total budget available in period t. Define x_i^t as the transfer to farmer i in period t. Then

$$\sum_{i \in F} x_i^t \;\le\; B_t, \quad \forall t \in T,$$

$$x_i \;=\; \sum_{t \in T} x_i^t, \quad \forall i \in F.$$

The total amount x_i allocated to farmer i is the sum of allocations over all periods. Forecast information can be updated each period, allowing the model to be rerun or adapted as conditions change.

Regional prioritization with vulnerability targets

To ensure that highly vulnerable regions receive a minimum level of support, one may add lower-bound constraints on regional allocations proportional to the region's aggregate vulnerability. Define

$$\sum_{i:\, r_i = r} x_i \;\geq\; \alpha \sum_{i:\, r_i = r} v_i, \quad \forall r \in R,$$

where α is a fraction indicating what portion of the total regional vulnerability must be covered. This enforces a baseline of resource allocation in alignment with regional vulnerability.

Logistics and delivery constraints

Realistic planning must account for logistical limitations, such as staffing, payment infrastructure, or local capacity, that cap the feasible amount of transfers in each region. Let L_r represent this logistical capacity for region r. The original regional constraint

$$\sum_{i:\, r_i = r} x_i \;\leq\; q_r$$

can be tightened by incorporating L_r:

$$\sum_{i:\, r_i = r} x_i \;\leq\; \min\bigl(q_r,\; L_r\bigr), \quad \forall r \in R.$$

This ensures allocations to each region do not exceed either the equity-driven limit q_r or the local delivery capacity L_r.

Summary

Each extension is an example of addressing real-world trade-offs in resource allocation decisions. Fairness constraints balance equity and efficiency by promoting uniform coverage among farmers with similar vulnerability scores, though this may limit aid to the most vulnerable. Risk-aversion terms prioritize reliability in uncertain regions but may reduce broader coverage. Dynamic budget allocation allows flexible, multi-period distributions but adds computational complexity and requires frequent updates. Regional minimums protect highly vulnerable areas but may divert resources from others in need. Delivery constraints ensure practical implementation by accounting for logistical limits, though this may restrict allocations based solely on vulnerability.

Together, additional constraints like these can increase the practicality of optimization models and their alignment with policy and operational needs. However, larger models may require additional expertise to formulate, more data, and greater computational resources to solve.

3 Modern computational approaches for rice yield prediction: a systematic review of statistical and machine learning-based methods

---

A version of this chapter was published in Computers & Electronics in Agriculture (April 2025 issue).

Abstract

This systematic review examines the performance of various methods for rice yield forecasting, an essential component of *impact-based* early warning in agriculture. Synthesizing 156 studies from 2015 to 2025, it evaluates both conventional (e.g., linear regression) and advanced machine learning models (e.g., random forests, neural networks), which draw on climate data, remote sensing indices (NDVI, MODIS), and UAV imagery. Key challenges identified include limited data availability, overfitting, and difficulty in generalizing models across regions and rice cultivars. Several gaps remain in standardizing evaluation metrics, incorporating ground-truth data, and adapting models for real-world agricultural decision-making. Future research should focus on integrating multiple data sources, enhancing model scalability, and developing frameworks for better uncertainty quantification. Addressing these gaps would boost scalability, enhance trust in predictive accuracy, and position yield forecasting as a pivotal component of early warning systems. The insights here provide the methodological underpinning for the machine learning models developed in Chapter 4, ultimately guiding more precise, impact-driven interventions that strengthen resilience in vulnerable farming communities.

3.1 Introduction

Rice is the most consumed grain worldwide, serving as a staple for half of the global population. In 2022, it provided 17.2% of global caloric supply and 25% of calories in Asia, home to over half of the world's population [270, 271]. Demand for rice is expected to increase by 25% by 2030 [272, 273].

Accurate estimations of how much rice a country will produce in advance of the actual harvest can provide a number of benefits. Good forecasts help farmers improve management, traders and insurers adjust pricing, suppliers manage stocks, and logistics companies plan routes. National authorities can balance food imports and exports, ensuring food security, while international aid organizations can swiftly mobilize resources in anticipation of potential shortfalls [149, 64]. The benefits of advance prediction of rice yields depends on the time horizon (a week ahead, a month ahead, etc.) and the spatial granularity (country-level, district-level, field-level, etc.) of the predictions.

Despite these potential benefits, there have been limited reviews focused on the effectiveness of machine learning-based approaches for rice yield prediction. Schauberger et al. (2020) conducted a review of local to regional yield forecasting approaches across 44 crops in 71 countries [149]. Van Klompenburg et al. (2020) reviewed 80 studies of machine learning and deep learning applied to crop yield prediction more broadly [274]. Wu et al. conducted a topical review on rice yield estimation based specifically on space-borne synthetic aperture radar (SAR) data [275]. Katsantonis et al. (2017) reviewed approaches for forecasting rice blast, a type of rice disease which causes multi-million dollar losses every year [276]. Darra et al. (2023) reviewed the extent to which yield prediction can be fully digitized, across a number of crops and approaches [277]. Torre et al. (2021) reviewed remote sensing-based estimation of rice yields using various models [278]. Other reviews have focused on rice area mapping [279, 280], mechanistic approaches to rice production forecasting [281], deep learning approaches for rice leaf disease detection [282]; rice type classification [283]; rice nutrient deficiency disorders [284]; smart farming applications such as smart irrigation [285]; and phenotype prediction [286, 287].

This review makes several contributions that build on the previous literature by providing a comprehensive synthesis of 156 studies, offering four key contributions. First, it employs a rigorous PRISMA-guided methodology, ensuring transparency and reproducibility. Second, it systematically evaluates data sources, methods, performance metrics, and challenges across studies, extending its scope beyond the temporal and geographic boundaries of prior reviews. Third, the review covers diverse data types, such as remote sensing, meteorological, and field data, across the studies reviewed, offering a broader perspective compared to prior reviews that often focus on a single category of data. Finally, it identifies critical gaps, including data limitations, overfitting, and challenges in generalizing models across regions and cultivars, while proposing actionable directions for future research, such as uncertainty quantification and decision-support integration.

3.2 Methods

3.2.1 Overview

A systematic review is a structured method of reviewing literature that ensures a thorough, unbiased, and transparent evaluation of research on a specific question [288]. In the context of machine learning for rice yield prediction, this involved systematically identifying, selecting, and analyzing relevant studies to provide a comprehensive and objective overview of the topic.

This research followed a step-by-step process based on guidelines for conducting systematic literature reviews proposed across several studies to ensure that the review is thorough and reproducible [289, 290, 291, 292, 293, 294, 295]. The process began with clearly defining the study's scope and research questions, ensuring that the review addressed significant gaps in the literature. Next, we implemented a comprehensive search strategy to identify and select relevant studies based on predefined inclusion and exclusion criteria (following PRISMA guidelines, as shown in Figure 3.1). Following this, we systematically extracted data from included studies using a structured template. The extracted data was then synthesized and categorized for further analysis. The final steps involved detailed analysis, such as comparing data, methods, and results across studies to draw meaningful conclusions and recommendations.

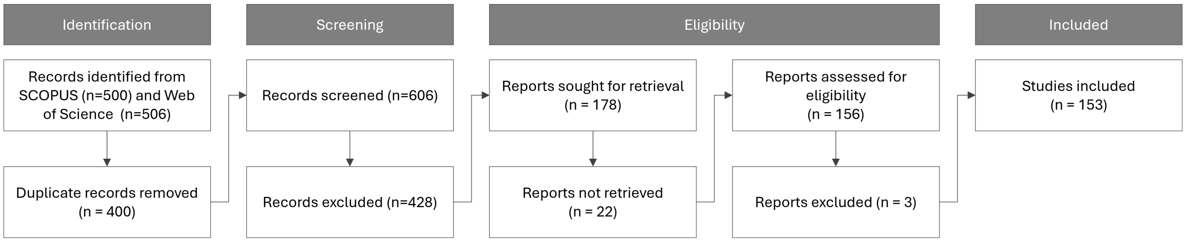


Figure 3.1: PRISMA flow diagram illustrating the study selection process. A total of 1,006 records were identified through SCOPUS and Web of Science. After removing 400 duplicates, 606 records were screened, leading to the exclusion of 428 records. Of the 178 reports sought for retrieval, 22 were not retrieved, and 156 reports were assessed for eligibility. Following the exclusion of 3 reports, 153 studies were included in the final review.

3.2.2 Detailed breakdown of methodological approach

1. Search strategy and pre-screening A systematic search was conducted in the Web of Science and SCOPUS databases using specific search queries. For both databases, the query focused on titles containing the terms "(rice OR paddy) AND (yield OR production) AND (prediction OR forecasting OR estimation OR modeling OR simulation)," and was limited to publications between January 1, 2015, and January 1, 2025.

The decision to limit the systematic review to studies published between 2015 and 2025 reflects the rapid advancements in machine learning (ML) during this period, marked by milestones like the open-sourcing of TensorFlow (2015) and PyTorch (2016) [296]. These years saw the emergence of more sophisticated algorithms, improved computational infrastructure, and access to larger, diverse datasets, significantly enhancing ML's applicability in agriculture. This timeframe also aligns with practical scope management, avoiding an unmanageable number of articles while focusing on recent advancements most relevant to current research and practice, as suggested in [297]. Furthermore, previous systematic reviews have similarly restricted their scopes, with periods like 2016–2020 for ML in rice phenotype prediction [286], and 2008–2019, 2016–2021, and 2018–2023 for crop yield prediction broadly [274, 298, 299]. Thus, a 10-year timeframe was chosen to maintain relevance and rigor.

The initial searches yielded 1006 results, which were reduced to 606 after removing duplicates. As shown in Figure 3.1, further filtering excluded an additional 428 articles, leaving 178 relevant studies. Of these, full texts for 156 were accessible through the authors' current journal subscriptions. Three records were unavailable as full texts in English. The additional filtering of the 428 articles was made on the basis of the criteria outlined in Table 3.1.

Table 3.1: Systematic review inclusion criteria. These criteria outline the specific requirements for studies to be included in the review to ensure a consistent focus on preharvest rice yield prediction using machine learning-based models and relevant datasets.

Category	**Inclusion Criteria**
Forecast Requirements	Presents at least one "true" forecast for rice yield during the growing season, made prior to harvest, and does not rely on post-forecast information such as weather or crop growth data.
Methodology	Uses empirical methods (e.g., machine learning or deep learning, including techniques such as Neural Networks or XGBoost) or statistical models (e.g., linear regression, stepwise regression) provided they involve prediction. Metrics like MAE, RMSE, or R^2 are mentioned for model validation. Reference yields must be based on observed data from field studies, experimental research, or administrative sources. Simulated data for reference yields is not accepted. Studies utilizing a wide variety of data such as UAVs, satellite data, or remote sensing data are included.
Focus	Explicitly focuses on predicting rice yield (absolute or relative), not proxy metrics like biomass or Net Primary Productivity (NPP). Avoids crop simulation models, field experiments, or exploratory relationships without prediction or forecasting. Avoids ambiguous studies lacking clear prediction or forecasting. Greenhouse or pot experiments are excluded.
Publication Requirements	Must be published in English and peer-reviewed. Review articles are included only if they introduce a novel method or provide a quantitative comparison of multiple methods.

2. Data extraction and analysis After applying the specified filters, 156 relevant studies were identified and systematically analyzed. Data were extracted on various variables, including study information (lead author, publication year, journal, title), experimental setup (study region, rice type, forecast horizon, spatial resolution), input data (types and sources including exogenous variables such as satellite or weather data), methodology (model type used), evaluation methods (e.g., cross-validation, validation sets), evaluation metrics, and qualitative information such as study limitations and suggestions for future research. The choice of these data extraction fields builds on previous systematic reviews of (not rice-specific) crop yield forecasting [274, 149].

Insights were consolidated into a master table and further processed or visualized using Python packages like NumPy, Pandas, and Altair [300]. A thematic analysis was then conducted, categorizing the studies into key themes such as modeling approaches, data sources, and performance evaluation methods. This analysis allowed for the identification of common patterns, emerging trends, and gaps in the existing research, providing an overview of the current state of rice yield forecasting methodologies.

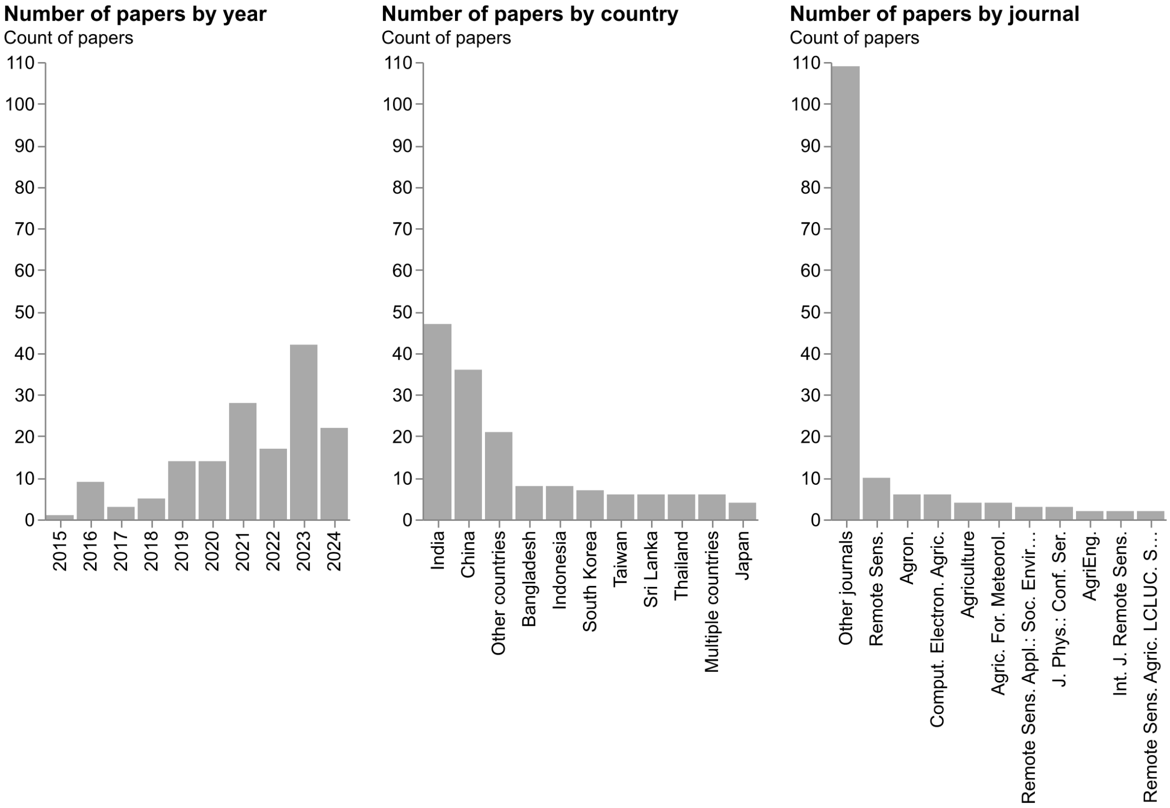


Figure 3.2: Overview of articles included in this review. The panels show the distribution of papers by year (left), by country (center), and by journal (right). The charts highlight the concentration of research in specific countries and journals, with a notable rise in publications in recent years. Note that the 2024 cutoff is the month of September.

3.3 Results

3.3.1 Description of the data

Figure 3.2 shows a clear upward trend in the number of research papers on rice yield prediction, with a noticeable increase in publications starting in 2019 and reaching a peak in 2023 (the cutoff of 2024 data is the month of September). The data also highlights that the majority of these studies are contributed by researchers from India and China, indicating strong regional engagement in this field. Additionally, a significant concentration of these publications appears in the journal *Remote Sensing*, suggesting that this journal is a primary outlet for rice yield prediction research. The full list of papers can be found in Table A.1.

3.3.2 Datasets and methods

The charts in Figure 3.3 present a quantitative overview of the variables and datasets most commonly employed in recent rice yield prediction studies. In the domain of remote sensing variables, the Normalized Difference Vegetation Index (NDVI) is the most frequently used, appearing in nearly 50 studies (Figure 3.3, top left). Other commonly used remote sensing variables include the Enhanced Vegetation Index (EVI), RGB imagery, and Synthetic Aperture Radar (SAR) data, though these are utilized less frequently than NDVI. Regarding the datasets from which these variables are derived, the MODIS (Moderate Resolution Imaging Spectroradiometer) dataset is the most frequently used for remote sensing data, appearing in over 60 studies (Figure 3.3, bottom left). UAV (Unmanned Aerial Vehicle) data and Sentinel-2 are also notable sources but are used less frequently.

Climatic and meteorological variables are also heavily utilized, with precipitation being the most commonly included variable, appearing in approximately 50 studies (Figure 3.3, top center). Temperature and relative humidity follow as the next most frequent variables, highlighting their importance in understanding the environmental conditions that influence rice growth. Other variables such as soil moisture, wind speed, and evapotranspiration are included in fewer studies. For climatic and meteorological datasets, local meteorological stations are the most common source, utilized in nearly 80 studies (Figure 3.3, bottom center). The CRU (Climatic Research Unit) dataset, as well as datasets from the ECMWF (European Centre for Medium-Range Weather Forecasts) and NASA's POWER project, are also used, though to a lesser extent.

Measurement data variables show a more focused application, with biomass and Leaf Area Index (LAI) being the most frequently used, though they appear in fewer than 10 studies each (Figure 3.3, top right). Other variables such as plant height and yield data are less commonly used. Measurement data is predominantly collected through field surveys, with experimental data and soil data being utilized in fewer studies (Figure 3.3, bottom right).

Figure 3.4 presents the frequency of modeling methods used in the studies papers evaluated. Neural Networks (NN) are the most frequently employed, appearing in nearly 70 studies, followed by Random Forest, which is used in about 50 studies. Linear Regression and Support Vector Machines (SVM) are also commonly used, with occurrences in approximately 40 and 30 studies, respectively. Long Short-Term Memory (LSTM) models and Convolutional Neural Networks (CNNs) are featured in over 30 studies, indicating their application in handling sequential and spatial data. Other methods such

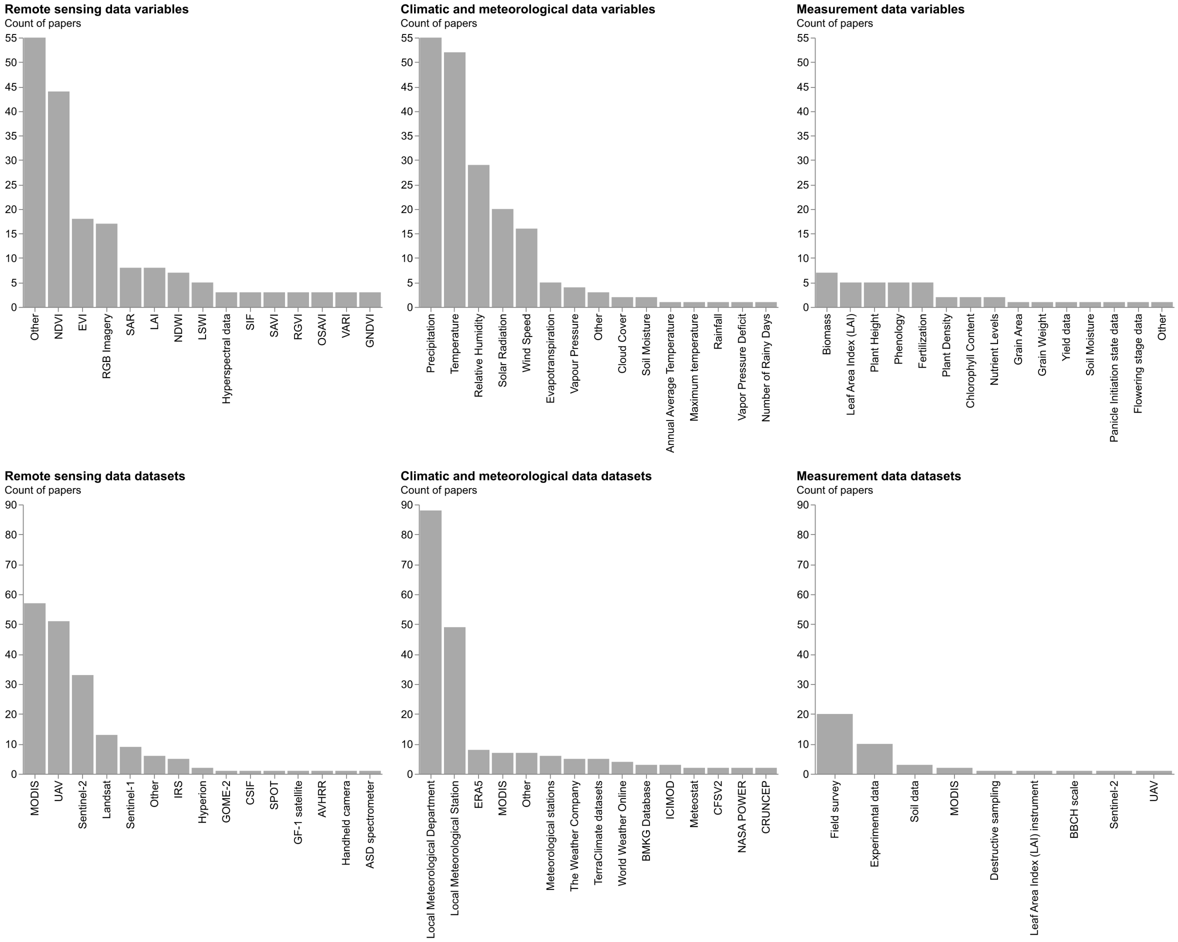


Figure 3.3: Overview of the most frequently used variables in the articles included in this review. The panels show the top 15 subcategories for each of the three main data categories: Remote sensing data (left), Climatic and meteorological data (center), and Measurement data (right). In each panel, subcategories outside the top 15 are grouped into an "Other" category, providing a concise view of the most commonly studied variables across different domains.

as Multiple Linear Regression (MLR), Polynomial Regression, and XGBoost appear less frequently, with each being used in fewer than 30 studies. Less common methods include Lasso Regression, Ridge Regression, and Decision Trees, which are used in fewer than 20 studies. A variety of other methods, including Gradient Boosting Machines (GBM), Bayesian Regression, and LightGBM, appear sporadically across the literature, each featuring in fewer than 15 studies.

Figure 3.5 presents the distribution of R^2, MAPE (Mean Absolute Percentage Error), and MAE (Mean Absolute Error) metrics from rice yield prediction studies, organized by country. While these metrics are influenced by the specific machine learning methods and study designs employed, such as variations in sample sizes, feature sets, validation

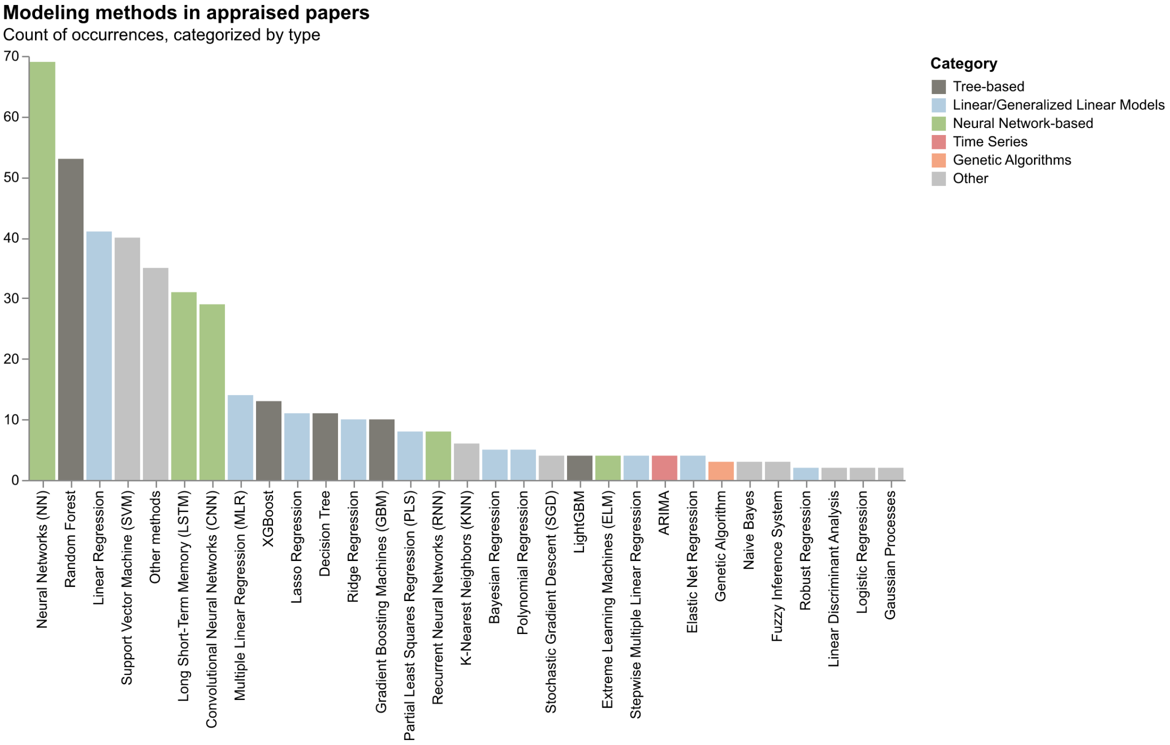


Figure 3.4: Distribution of the top 30 modeling methods employed in the research papers appraised in this review. The chart groups methods into categories such as tree-based, linear/generalized linear models, neural network-based, time series, genetic algorithms, and others. Neural Networks (NN) and Random Forest are the most frequently used methods. “Other methods” represents a collection of diverse approaches utilized less frequently across the reviewed studies.

methods, and forecast horizons, they provide insight into the range of performance achieved across diverse contexts. Direct cross-country comparisons may not always be appropriate due to these methodological differences, but the reported statistics remain valuable for identifying general trends and benchmarking model effectiveness.

Other systematic reviews of crop yield forecasting research have noted similar challenges in comparing performance metrics. For instance, one study noted that “The forecasting performance depends on various factors, including crop, region, method, lead time to harvest and input diversity” and “the majority of studies (52%) does not provide an out-of-sample assessment of their performance” [149].

The R^2 values, reflecting the proportion of variance explained by the models [301], are consistently high in training datasets (upper left panel of Figure 3.5) across most countries, with values often exceeding 0.65 for countries like China, Japan, and Spain. This suggests strong model fits for training data in these regions. Higher ranges in R^2 values are observed for countries like Bangladesh and Thailand. In testing datasets (bottom left panel of Figure 3.5), R^2 values exhibit greater variability, with countries such as Bangladesh, China, India, and the Philippines showing a broader range of performance, reflecting potential challenges in model generalization. By contrast, countries like Japan, South Korea, and Pakistan showed narrower R^2 bounds.

The MAPE values, representing the percentage error in predictions, reveal considerable variability across countries in the testing datasets (bottom middle panel of Figure 3.5). Countries like China and Spain exhibit consistently low MAPE values, often below 6%, indicating strong relative prediction accuracy. However, countries such as India and Indonesia show higher MAPE ranges. The variability in MAPE highlights the potential influence of factors such as data quality and model parameterization, particularly in regions with greater agricultural heterogeneity or data sparsity.

The MAE values, which measure the absolute error in yield predictions, also vary substantially in the testing datasets (bottom right panel of Figure 3.5). Research papers on countries such as Bangladesh, Myanmar, and Spain demonstrate low MAE values. In contrast, countries like China, India, and Sri Lanka exhibit higher MAE values, with some exceeding 0.8 t/ha, indicating substantial absolute prediction errors.

Figure 3.6 shows the most common challenges across studies, with limited generalization across crops and regions being the most frequently cited issue, followed by concerns about model sustainability and overfitting. Data-related problems such as the absence of agronomic data and poor data quality also feature prominently, emphasizing the ongoing difficulty of accessing reliable inputs for modeling. Computational processing requirements and continuous data integration are key engineering challenges, highlighting the scalability barriers of deploying models in real-world settings. Additionally, the chart reveals significant operational challenges, including the difficulty of integrating model outputs into decision-making frameworks and effectively communicating uncertainty to non-experts, underscoring the need for better alignment between model outputs and stakeholder needs.

3.4 Discussion

3.4.1 Strengths and limitations of the current review

In this systematic review, we aimed to compile, categorize, and evaluate the AI-based approaches used in rice yield prediction studies. Despite the comprehensive coverage of recent research on the topic, several limitations must be acknowledged. First, the inclusion of studies was restricted to those published in English, which may have led to

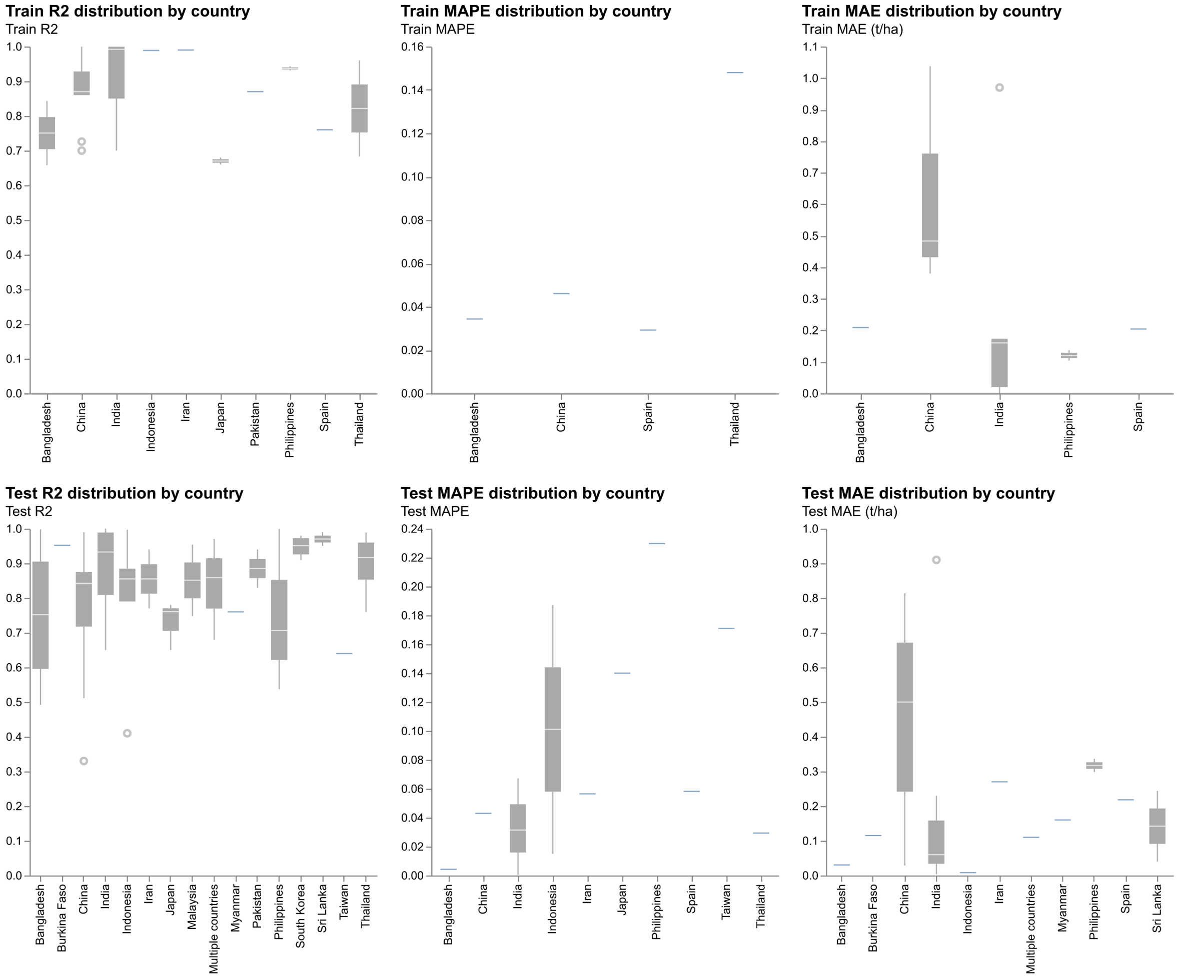


Figure 3.5: Distribution of R2, MAPE, and MAE values by country across the reviewed research papers (for papers where these metrics were unambiguously reported). The box plots illustrate the variability of these metrics within different countries, showcasing the differences in model performance. Each subplot represents a different metric, with R2, MAPE, and MAE on the left, middle, and right, respectively. The top row provides the training data metrics, while the bottom provides them for testing data. Direct cross-study comparisons may not always be appropriate due to methodological differences across papers.

an under-representation of research conducted in non-English-speaking regions. This is particularly relevant for regions like Southeast Asia or Latin America, where significant rice cultivation occurs, and valuable local insights might be missed.

Additionally, the focus of this review on peer-reviewed literature may have excluded valuable contributions from non-academic sources, such as government reports, industry-led initiatives, or unpublished data from NGOs. These non-peer-reviewed sources often offer practical insights into the operational implementation and scalability of machine learning models in real-world settings, particularly in resource-constrained environ-

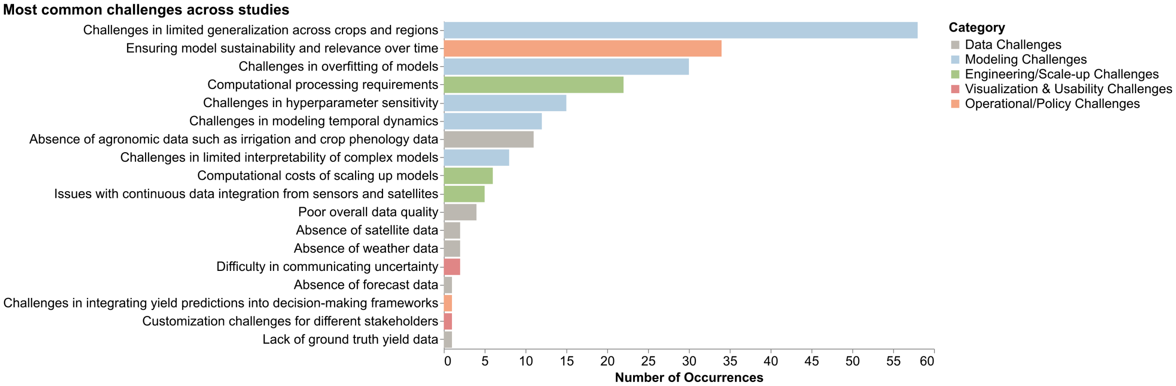


Figure 3.6: Distribution of the most common challenges identified across studies, categorized by challenge type, including data challenges, modeling challenges, engineering/scale-up challenges, visualization/usability challenges, and operational/policy challenges.

ments [302]. For example, some startups such as CropIn have developed commercial AI-based models that address challenges such as data integration and real-time forecasting, issues that are less frequently addressed in academic literature [303].The omission of these perspectives might narrow the scope of understanding, especially when considering the deployment and adoption of these models in regions where food security is a pressing issue. Expanding future reviews to include gray literature and field-level case studies could provide a more comprehensive and actionable understanding of the barriers and opportunities in deploying crop yield prediction models at scale.

3.4.2 Limitations in the articles reviewed

Comparing studies on crop yield predictions poses significant challenges due to variations in study design, geographic focus, and data quality.

Study design Machine learning models that integrate remote sensing and meteorological data to predict yields at a national or regional scale may not be directly comparable with studies focusing on yield predictions at the individual field level. Various studies, such as Islam et al. (2023) and De Clercq et al. (2024), used machine learning techniques and earth observation data to predict rice yields at the district level, measuring performance with metrics like RMSE, MAE, and MAPE [304, 250]. These models, while offering large-scale insights, rely on coarser spatial resolution data that can obscure local variabilities critical for field-level predictions. In contrast, Basir et al. (2021) applied neural networks to predict yields based on transplanting parameters in Bangladesh, relying on field-collected data, which offers high resolution but lacks the scalability needed for broader regional applications [305]. The differences in data types, spatial resolution, and modeling approaches between these studies make direct comparisons of their outcomes problematic.

For instance, while district-level predictions provide useful information for policymakers managing food security, they may overlook the nuances of individual field conditions such as microclimate effects or local soil variability. On the other hand, field-level predictions using high-resolution UAV-based imagery, as done by Mia et al. (2023) in Japan, can capture specific crop health and growth patterns but may struggle to generalize findings across larger, diverse regions [306]. Similarly, Ghose et al. (2021) used a model based on atmospheric oscillations to predict regional yields in Bangladesh, where atmospheric and climatic factors play a dominant role [307]. The use of large-scale atmospheric drivers versus on-the-ground parameters further complicates the ability to generalize results across different scales. These differences in methodology underscore the importance of tailoring yield models to the scale and purpose of the study. Without careful consideration of these factors, direct performance comparisons may lead to misleading conclusions about model efficacy and the applicability of findings across scales.

Geographic focus Geographic variability also introduces challenges when comparing rice yield prediction models across different regions. Models tailored to specific environmental conditions may not perform well when applied to areas with different climates, soils, or agricultural practices. For instance, Fu et al. (2023) used multispectral and SAR data to estimate rice yields in Thailand based on rice phenology stages, an approach suited to the monsoon-driven climate of the region but potentially less applicable to arid areas like Iran, where Asmar et al. (2024) integrated physical and deep learning models to forecast yields under water-scarce conditions [308, 309]. Even within the same region, variability can affect model performance. Mokhtar et al. (2024) showed that adding soil data improved model accuracy in parts of China, highlighting how local environmental factors like soil quality can significantly impact prediction outcomes [310]. Another study by Brinkhoff et al. (2024) on field-level rice yield forecasting notes that their method "relies on rice-specific phenology models and sowing date, which may not be available in other rice-growing areas, potentially limiting scalability" [311]. These differences underline the need for region-specific model calibration, as cross-regional comparisons without accounting for local conditions can lead to misleading conclusions about model performance.

Despite the limitations on the geographic scalability of some studies, other research has explicitly focused on national/regional scalability. One study on rice yield prediction at the district level in India concludes that, despite the focus on Kharif-season rice in India, the approach "could be replicated and scaled to other regions" due to the use of globally available input features like ERA5 climate data and MODIS satellite imagery [250]. Another study applying convolutional neural networks to paddy rice in Iran noted potential for the model to be "adapted and applied to different regions and crop years" [309]; similar claims on the regional scalability of field-level rice yield prediction in China were made in Yu et. al. (2023) [312]. Nevertheless, there is limited evidence that globally generalizable models have been developed that would be useful for decision-makers in different countries at arbitrary levels of granularity.

Temporal coverage of datasets Differences in data length and availability complicate the comparison of rice yield prediction models, as varying datasets can produce misleading impressions of model performance. For example, Sathya et al. (2023) used an MLR-LSTM model trained on seven years of data from the Tamil Nadu Delta, reporting an R^2 of 0.93, while Ekanayake et al. (2021) applied Random Forest and other techniques to 11 years of data in Sri Lanka, finding R^2 values ranging from 0.61 to 0.99 across districts [313, 314]. To take a hypothetical example: although the Tamil Nadu model appeared to perform better, this could be an artifact of the smaller dataset, where reduced variability might artificially inflate R^2 values. In contrast, the longer dataset in the Sri Lankan study introduces greater variability, which may lower R^2 without necessarily reflecting poorer predictive capability. These differences illustrate that metrics like R^2 are context-sensitive, and high scores in shorter datasets do not automatically signify model superiority. Consequently, comparisons across studies should account for dataset variability and length to avoid drawing flawed conclusions about model effectiveness.

Challenges in inconsistent performance metrics and cross validation methods Comparing rice yield prediction models across studies presents significant challenges due to the inconsistency in evaluation metrics, dataset lengths, and reporting methods. Metrics like RMSE, MAPE, and R^2, though frequently employed, are not universally comparable. For instance, MAPE is particularly prone to distortions in low-value (i.e., low-yield) conditions [315], making it difficult to compare directly with RMSE, which emphasizes overall error magnitude without being sensitive to specific yield ranges. Moreover, R^2, while commonly used, often fails to capture key aspects of model performance such as bias and error distribution [316]. Additionally, variations in how studies perform train-test splits, whether temporal or random, further complicate comparisons. Some studies emphasize training set performance, while others highlight test set outcomes, leading to discrepancies in how model generalizability and robustness are interpreted. These inconsistencies in metric selection, data handling, and reporting approaches contribute to divergent conclusions about model performance across regions and study conditions.

For instance, regarding performance metrics, in the study by Zhou et al. (2023), rice yield prediction in Hubei Province was evaluated using CNN-LSTM, CNN, and ConvLSTM models, and evaluation metrics included RMSE, MAE, and the correlation coefficient, with data split into training (2000–2013), validation (2014–2016), and testing (2017–2019) sets [317]. The study highlighted the superior performance of the CNN-LSTM model, reporting RMSE, MAE, and R values for the test set. In Tanaka et al. (2023), the authors reported using relative root mean square error (rRMSE) and R^2 as evaluation metrics for their CNN-based model for rice yield estimation, explaining 68% of the yield variation with an rRMSE of 0.22 [318]. Pankaj et al. (2024) evaluated multiple regression techniques for rice yield prediction using R^2, RMSE, and MAE, with models split into a 70% training and 30% test set, and averaged over ten trials for robustness [319]. Similarly, Marong et al. (2024) applied RMSE and R^2 to assess Random Forest models in Malaysia, achieving an R^2 of 0.954, but without addressing low-yield distortions that might arise with MAPE [320]. These examples illustrate how variations in metric choices and dataset handling can obscure cross-study comparisons, underscoring the need for standardized evaluation frameworks to avoid skewed interpretations of model effectiveness across diverse agricultural conditions and data environments.

Further complicating the comparability of model performance across studies is the choice of cross-validation methodology. As mentioned above, some studies adopt a train-test split methodology where earlier years are used for training and subsequent years are reserved for testing, ensuring a clear temporal separation between datasets [317]. This approach contrasts with studies that employ leave-one-year-out (LOYO) cross-validation, where each year is systematically left out as a test set while all other years are used for training. Both methods address temporal dependencies but reflect different strategies: one provides a fixed temporal partition, while the other tests a model's generalizability across all years. [312, 311]. Each approach may have merits depending on the specific goals of the study and the characteristics of the available data.

3.4.3 Implications for future research and development

In the following section, we explore potential areas of future research that could enhance the real-world impact of rice yield prediction models, including improving data availability and integration (e.g., expanding access to ground truth data, building centralized repositories, leveraging hyperspectral imaging and SAR); advancing model development and scalability (e.g., enhancing generalization across regions and cultivars, integrating process-based and machine learning approaches, exploring frontier deep learning models, standardizing reporting guidelines); and enabling user-centric operationalization and deployment (e.g., adopting modern data engineering practices, embedding models into decision-making systems, and improving the visualization of model outputs).

A. Data availability and integration

Expand access to ground truth data The availability and quality of ground truth data on yield, phenological stages, and agronomic practices represent important bottlenecks in the performance of rice yield forecasting models. Yu et al. (2023) highlighted that the lack of multi-year field-level yield data across large regions remains a major obstacle for calibrating and evaluating robust data-driven models [312]. Future studies could focus on establishing collaborative data-sharing platforms, integrating crowd-sourced and sensor-based data collection, and leveraging remote sensing technologies to fill gaps in ground truth data and improve the accuracy and scalability of yield predictions. In addition, collaboration with private sector sources of ground truth data, from companies such as OneSoil, may enable more rapid experimentation [321].

Build a centralized repository of global rice-related datasets Future efforts could also focus on building a centralized, API-accessible repository that consolidates up-to-date, relevant data sources specifically for rice yield prediction. This would enable seamless ingestion of critical data into models, including crop masks from sources like MAPSPAM or CROPGRIDS, and weather data from programs like ERA5 [322, 323, 324]. By ensuring researchers have easy access to these data updates, the repository might enable faster experimentation in the field.

The relatively low incidence of research papers in this review which utilized advanced data inputs such as hyperspectral imagery underscores the need for such as centralized repository to include more advanced and experimental data types such as UAV data or hyperspectral data. Hyperspectral imaging (HSI) captures data across hundreds of narrow spectral bands, offering greater detail than conventional imaging techniques that use the visible spectrum (RGB) or multispectral data.

For instance, Burglewski et al. (2023) demonstrated that silage maize yield modeling approaches leveraging hyperspectral data that include the spectral red edge may outperform approaches based on Landsat imagery [325].

Clarify data and algorithm selection criteria for yield prediction in the context of an evolving data landscape Determining the optimal combination of data sources and algorithms for rice yield prediction remains a significant challenge, as it depends on factors such as spatial granularity (e.g., field-level versus regional), project budget, time horizon, accuracy requirements, and intended end use. For instance, while NDVI derived from NASA MODIS may be sufficient for coarse regional predictions, more granular imagery from Sentinel or Planet Labs might be necessary for field-level applications requiring higher precision [326]. Similarly, hyperspectral sensors have become attractive options due to the ability to obtain high spectral resolution data via both satellite and UAV [327]; but they can come with increased costs and complexity, which may not always justify their use [328]. The choice of algorithms also requires careful consideration, with deep learning models excelling in capturing complex patterns in large datasets but often demanding significant computational resources, while traditional machine learning or hybrid models may provide more practical solutions under constrained budgets or data limitations [329]. Future research could therefore focus on developing decision trees or frameworks that guide the selection of data sources and algorithms for different yield prediction scenarios, enabling institutions to more effectively trade off cost, accuracy, and scalability in operational systems. This could build on existing research that has developed decision frameworks for choosing appropriate methods to solve general data science problems [330]. Such frameworks would help determine, for example, when advanced models are justified over simpler alternatives and how best to deploy available data to meet specific objectives.

Looking ahead, the increasing availability of advanced satellite platforms, such as Sentinel-2 with its higher revisit frequency and spatial resolution, and new constellations like PlanetScope, which offers daily global coverage at high resolution, presents significant opportunities for rice yield modeling [331]. Emerging hyperspectral satellites, such as those in the EnMAP and PRISMA missions, are expected to provide a level of spectral granularity that enables more precise differentiation of crop conditions and stress factors [332]. Moreover, synthetic Aperture Radar (SAR) has also introduced new opportunities to improve crop yield forecasting applications. SAR data, with its ability to maintain consistent acquisition schedules regardless of weather conditions or the time of day, addresses some of the limitations associated with multispectral sensors, such as cloud cover interference, background noise, aerosol effects, and signal saturation in areas with dense biomass [333, 334]. The potential of SAR becomes even more significant when combined with optical sensors such as Landsat or Sentinel-2 through optical-SAR data fusion. This approach leverages the complementary strengths of both data types, with SAR providing all-weather capabilities and sensitivity to surface and subsurface moisture, while optical sensors offer detailed spectral information critical for assessing vegetation health and stress [335, 336]. Research evidence shows that combining SAR and optical data has provided strong performance in crop classification [337], crop monitoring [338], and yield prediction [312]. These advances could dramatically improve the ability to monitor dynamic crop growth processes and incorporate real-time updates into yield prediction models, making it essential for future frameworks to be adaptable to increasingly granular and high-frequency datasets.

B. Model development and scalability

Enhance model scalability to other regions and cultivars Limited generalization of models across cultivars and regions remains a major challenge, as rice cultivars respond differently to environmental factors like soil type and water availability. Nguyen-Thanh Son et al. (2022) highlighted issues with applying machine learning models for rice yield prediction, such as cloud cover and mixed-pixel problems in fragmented fields. While their SVM model performed well in Taiwan, limitations like small sample sizes and environmental variability may hinder broader generalization [339]. Similarly, Zheng et al. (2022) found that while UAV multispectral imagery improved yield prediction in specific conditions, model accuracy declined when applied to different years, cultivars, and sensor data, emphasizing the need for more adaptable models across diverse environments [340]. Tanaka et al. (2023) demonstrated that although their deep learning model successfully predicted rice yields across diverse environments using ground-based RGB images, accuracy was lower in conditions not represented in training data, underscoring the need for models adaptable to a broader range of conditions [318]. Aziz et al. (2023) also suggested that future models should account for irrigated and non-irrigated regimes more explicitly [341]. Parreno et al. (2024) highlighted challenges in model over-fitting and sensitivity to machine learning model hyperparamers as important factors impeding the scalability of their models [342]. Future research should focus on developing models that incorporate a wider variety of environmental conditions, including diverse cultivars and irrigation regimes, while improving hyperparameter tuning methods to reduce overfitting and enhance model scalability across regions.

Explore synergies between process-based and machine learning-based models Machine learning-based models, which are typically based on limited historical observations, may face challenges in accurately predicting future yield-climate relationships [343]. For example, boosted regression trees (BRTs), a popular ML technique, excel at capturing non-linear relationships and spatial variability, making them more accurate than traditional linear regression in many cases. However, as highlighted by Sidhu et al. (2023), BRTs can conflate correlated variables, such as time and temperature trends, reducing their reliability in attributing yield changes to specific climate factors [344].

To address this challenge, future research in rice yield forecasting could focus on investigating how to combine process-based models with machine learning techniques. This hybrid approach holds promise for improving model accuracy and interpretability by leveraging the strengths of both methods. Process-based models (PBMs) like DSSAT and APSIM capture crop physiology and biophysical relationships, while machine learning (ML) models excel at handling large datasets and uncovering nonlinear patterns. As seen in Zhang et al. (2023), PBMs provide valuable insights into plant growth mechanisms but struggle with parameterization and scalability [345]. On the other hand, ML models can suffer from interpretability issues and require large datasets for training. Integrating these two approaches, as suggested by Maestrini et al. (2022), can address these challenges by using PBMs to guide feature engineering or as input for ML models, ensuring biologically grounded predictions while mitigating overfitting [346]. Early research on the topic has shown that such combined approaches can reduce model error, as in the case of corn modeling in the US corn belt [347].

Explore the effectiveness of frontier deep learning models Transformer architectures are emerging as a promising research avenue in agriculture [348], particularly due to their ability to model long-range dependencies in data sequences [349]. In crop yield prediction, one study demonstrated the potential of combining a transformer with a temporal convolutional network (TCN) to integrate climatic, soil, and satellite data, thereby enhancing temporal feature extraction and capturing complex dependencies [350]. Another study proposed the SSA-LSTM-transformer (SLTF) model, optimized using a sparrow search algorithm, to estimate winter wheat yield with remotely sensed variables such as leaf area index (LAI), fraction of photosynthetically active radiation (FPAR), and vegetation temperature condition index (VTCI), achieving improved accuracy and generalizability [351]. Transformers have also shown promise in crop row detection, where a model employing polynomial parameter learning efficiently approximated crop row curves, delivering robust results in complex environments without traditional feature extraction or post-processing [352]. These emerging studies highlight potential to boost accuracy in yield prediction problems.

Enhance reproducibility of experiments with standardized reporting guidelines Improving the transparency and reproducibility of models requires a shift towards standardized reporting metrics. Studies in rice yield prediction use metrics such as RMSE, MAPE, and R^2 inconsistently, making cross-study comparisons difficult. This is a challenge which has been identified in the crop yield production more broadly, beyond just rice [149]. Future research should emphasize standardized metric reporting and the use of shared benchmarks for training and validation, allowing for better model comparability. Moreover, the development of open-source platforms that centralize rice yield prediction data and models, such as repositories hosted on GitHub or specialized agricultural data platforms, will promote consistency in model evaluation.

C. User-centric operationalization and deployment

Investigate model performance in production using state-of-the-art data engineering practices To ensure more reliable deployment of rice yield models in production, future research could incorporate modern data engineering practices. Such practices might include: automated pipelines to ingest and process agricultural data like satellite imagery, weather, and soil metrics; scheduled updates of datasets to maintain the timeliness of predictions; using MLOps tools such as MLflow to track model experiments and performance across diverse environments [353]; containerizing models with tools like Docker to ensure portability and scalability [354]; leveraging cloud platforms like AWS or GCP for flexible computing infrastructure; utilizing workflow orchestration tools like Apache Airflow or AWS Step Functions for efficient task management; implementing CI/CD pipelines to automate model retraining and mitigate performance drift due to seasonal variations; and adopting Git-based workflows to enable collaboration and version control across large research teams working on challenging yield prediction problems at scale [355, 356].

Enhance the visualization and communication of rice yield prediction model outputs Effective visualization and communication of yield prediction outputs may facilitate broader adoption of rice yield prediction models. Visualization tools such as Tableau, Streamlit, Dash, and PowerBI have proven effective in building interactive web applications and dashboards in fields like bioinformatics, bacterial testing, and financial auditing [357, 358, 359, 360, 361, 362]. Applying these tools to rice yield prediction could allow users to visualize spatial yield variability, track temporal trends, and assess prediction uncertainties in an intuitive manner, as demonstrated in a recent applied rice yield prediction study [250].

Incorporate rice yield prediction into decision-making systems Beyond predictive accuracy, future research could focus on integrating rice yield models with optimization techniques to help various stakeholders in government and the private sector make better decisions. For instance, tools such as linear programming, stochastic models, and mixed-integer linear programming (MILP) have been widely used in other fields to optimize resource allocation and decision-making. In humanitarian contexts, [211] used MILP to optimize vehicle routes and minimize aid distribution times in Chile, and [212] used MILP to maximize helicopter evacuation efficiency based on aircraft capacity and victim urgency levels. Similarly, [213] developed a stochastic model to enhance disaster preparedness and response using mobile phone location data. Other studies have focused on optimizing relief networks, pre-disaster preparedness, and rescue operations [214, 215, 216, 217, 218]. Integrating these techniques into rice yield prediction models could enable public sector decision-makers to adjust policies and resource allocation strategies based on forecasted yields. For instance, in the case of lower-than-expected yields, optimization models could suggest reallocating subsidies, adjusting irrigation plans, or prioritizing regions for agricultural support programs. Conversely, in scenarios with higher yield predictions, these models could assist in planning for adequate storage, distribution networks, and food security measures, ensuring that surplus production benefits the wider population. This integration would create a more dynamic, responsive framework for managing agricultural resources like rice.

3.5 Conclusion

This review synthesizes machine learning-based approaches for pre-harvest rice yield prediction, evaluating a wide range of methods like neural networks, random forests, and support vector machines across diverse regions. It highlights key trends in performance, identifies common data sources, and addresses challenges such as data scarcity and model generalization. By comparing studies systematically, the review reveals both the potential and limitations of current models, offering targeted recommendations for improving data integration, model adaptability, and practical application. These findings are important for advancing the accuracy and utility of yield forecasts in policy and agricultural management.

4 Feasibility of machine learning-based rice yield prediction in India at the district level using climate reanalysis and satellite data

---

Published in Agricultural Systems (October 2024 issue).

Abstract

This study builds on the systematic review of rice yield forecasting methods (Chapter 3) to examine whether machine learning models, trained on openly accessible climate reanalysis (ERA5) and satellite (MODIS) data, can reliably predict Kharif-season rice yields for 247 districts in India several months prior to harvest. Nineteen algorithms, ranging from Orthogonal Matching Pursuit to CatBoost, were tested on a 20-year dataset. Top models achieved out-of-sample R^2, MAE, and MAPE values as high as 0.82, 0.29, and 0.16, respectively, comparable to or exceeding benchmarks from prior literature. Further interpretability was provided via SHAP analysis, which highlighted temperature, soil water content, and leaf area index among key drivers of yield. To facilitate practical use, an interactive dashboard visualizes regional predictions and uncertainties, revealing, for instance, higher accuracy in states like Uttarakhand than in Uttar Pradesh. These findings offer evidence that scalable, district-level yield forecasts can become a robust component of India's agricultural early warning systems, augmenting crop monitoring with proactive, impact-based alerts. As a bridge to the next chapter (Chapter 5), we also note that this forecasting framework can be integrated into optimization-based decision tools, enabling policymakers and humanitarian agencies to allocate resources such as cash transfers or fertilizer more precisely in anticipatory action contexts.

4.1 Introduction

4.1.1 The societal implications of accurate crop yield forecasting in India

Yield forecasting is the science of predicting agricultural productivity as measured by crop yield – the ratio of the total mass of the harvested product (such as rice) to the area used to cultivate the crop – before the harvest takes place, typically a few months in advance [363].

Pre-harvest prediction of crop yields is important in helping a wide range of stakeholders make better decisions around agricultural planning. For farmers, accurate crop yield forecasts can facilitate decision-making around what to grow and when to grow it [274]. In addition, near real-time monitoring of crop growth can inform the use of preventive measures such as irrigation and fertilization to boost agricultural productivity where needed [364]. For governments, yield prediction is relevant to the formulation of policies related to national food security, such as pricing policies for domestic markets, and policy decisions on the import and export of different crops [365].

Accurate crop yield forecasting may also enable better design of insurance products that mitigate climate risks and stabilize farmer incomes [366]. Weather-based crop insurance, for instance, uses a weather index such as total precipitation to determine payments to farmers, meaning that insurance companies do not need to visit farmers to assess damages and arbitrate claims. Rather, if the weather reaches a certain threshold, rapid automatic payments can be distributed to farmers, who avoid the need to sell assets to survive due to adverse climate events [367].

The need for accurate information on crop yields is particularly important in countries like India, where the agricultural sector provides livelihoods for hundreds of millions of farmers, with 70% of rural households depending on agriculture for their main source of income. One of India's major staple crops is rice, which contributes to 30% of calories consumed in India and is a key export commodity for the country [368]. India cultivates rice on about 45 million hectares of land, with a total production of 178 million metric tonnes in 2020 [369]. In addition, the distribution of monsoon rainfall, which is a major source of water for rice cultivation, has become erratic in recent years due to climate variability [370, 371]. In such contexts, crop yield predictions may be able to supplement agricultural early warning systems (AEWSs) that give advanced notice of potential risks to crop productivity, enabling preemptive action in affected areas. Previous research has shown that current agricultural monitoring systems lack robust crop yield and crop production forecasts, as well as the operationalisation of such methods at scale [372].

4.1.2 Overview of approaches and variables used to model crop yields

Crop yield prediction is a challenging problem in precision agriculture, as final yields depend on a variety of factors such as weather, climate, soil, seed type, and agronomic practices such as irrigation and fertilizer use [373].

This complexity is evident from the variety of variables included and methods applied in the growing body of literature on crop yield forecasting. For example, recent examples in literature involving deep learning approaches include corn and soybean yield forecasting in the US based on convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [374], soybean yield forecasting in Argentina based on deep transfer learning [375], and vineyard grape yield estimations based on CNNs [376]. Recent examples based on machine learning approaches include sugarcane yield prediction using random forests [377], prediction of wheat, barley, and canola yields in Western Australia using random forest [378], yield forecasting of spring maize in Pakistan based on LASSO regression and support vector machine [379], and Jojoba yield prediction in Israel based on gradient boosted regression trees [380]. Other examples of the machine learning approaches that have been applied to yield prediction have been summarized in a systematic literature review, which also analyzed the variables most frequently included in crop yield prediction studies. Across 50 studies between 2008 and 2019, features used as predictors of yield have included temperature, soil type, rainfall, humidity, pH-value, NDVI, wind speed, and more [274].

For studies specific to rice, the staple crop of over half the world's population [381], a number of approaches have been applied to yield forecasting in recent years. Recent examples include rice yield prediction for 81 counties in southern China based on recurrent neural networks [382]; application of the ecological distance algorithm to model rice yields [383]; field and county-level rice yield prediction based on synthetic aperture radar (SAR), optical and meteorological data [384]; random forest yield prediction based on high-resolution imagery collected from unmanned aerial vehicles (UAVs) [385]; simulation of yields using the Cropping System Model-CERES-RICE [386]; pixel-scale rice yield prediction in South Korea based on a combination of deep learning and crop models [387]; rice yield estimation at 500m spatial resolution based on gradient boosted regression and vegetation indices derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) [388]; and rice paddy yield prediction using sentinel-based optical and SAR data in India based on random forest [389]. To the authors' knowledge, there has been less research on rice prediction at the district-level in India.

4.1.3 Research contributions

This study marks an advancement in the field of rice yield prediction in India by building a proof-of-concept approach capable of predicting rice yields at the district level for 247 rice-producing districts across India.

First, a novel combination of data sources is used to predict Indian rice yields. These include data from ERA5, a climate re-analysis product developed by the European Centre for Medium Range Weather Forecasts (ECMWF), which combines observations with modelled data to provide hourly data on atmospheric, land-surface, and sea-state parameters globally [390]. Vegetation data was derived from the MODIS sensors onboard NASA's TERRA and AQUA satellites, and a cropland mask (CROPGRIDS) was used to filter earth observation data according to where rice is grown at a pixel level. Collectively, these data enhance the model's ability to capture the intricate effects of climate variables and vegetation health indices on crop yields.

Secondly, the study creates a spatially matched dataset for rice crop yields in India, tackling a common problem in yield prediction research: a lack of training datasets where yield data is accurately aligned with specific geographic locations. Typically, discrepancies between the names and boundaries in geographic datasets (like shapefiles) and official agriculture statistics can hinder the effective use of data in research. Fuzzy matching algorithms were used to align Indian shapefiles (which may have variations in district names) with official rice yield data from the Indian government to produce a dataset where yield information is precisely matched to its geographic location. This matching is important for researchers aiming to spatially aggregate earth observation data in a manner consistent with available yield data. Making this matched dataset open source specifically aids rice-related research in India, enabling more accurate earth observation studies by providing a reliable foundation for correlating satellite data with actual agricultural outputs.

Thirdly, the study provides a new benchmark for district-level rice yield prediction in India based on predictions made exploring the predictive effectiveness of 19 machine learning models such as LightGBM, Bayesian Ridge Regression, and others. This methodological approach facilitates the identification of the most effective models for rice yield prediction at the district level. Previous literature on crop yield prediction in India has largely focused on using a narrower range of algorithms such as random forest or support vector machine [274]. A detailed evaluation of model performance is provided, coupled with the interpretability provided by SHAP values.

Fourth, beyond model development, an interactive dashboard tool was developed, not only to visualize the yield predictions across each of India's districts, but also to allow for detailed diagnosis of model performance. It serves as a practical tool for model evaluation, offering insights into regional performance variations and facilitating the identification of areas where predictions may be improved.

Overall, this study offers evidence that scalable crop yield prediction models have the potential to be integrated into agricultural early warning systems in India, which, as noted in previous research, currently lack such forecasting capabilities and the means to operationalize these methods at a scale that contributes towards more resilient agricultural practices and food security planning [16].

4.2 Methods

4.2.1 Study region and brief overview

In this study, climate and remote sensing data were used as predictors to model rice yields for the kharif season (wet summer monsoon season) from 2001 to 2020 at the district level in India (India consists of 36 states and 684 districts). In India, more than half of the annual rice crop is grown during kharif [391], a season which is characterized by high temperature, high humidity, and medium to high rainfall [392]. Kharif season rice is typically sowed between the start of June to the end of August and harvested between the end of September to early January, depending on the region. During the 2019-2020 season, harvesting of Kharif rice was completed in February 2020.

The study involved key steps: ingesting 20 years of climate, vegetation, and rice yield data via API, pre-processing and aggregating data to match rice yields, training and evaluating models like Bayesian ridge regression and LightGBM, creating an interactive dashboard to visualize results and errors, and discussing implications for agricultural early warning systems. The methods outlined in this research are fully reproducible. All data, python code, and dashboards can be found on GitHub: `https://github.com/DPhilCode/india_rice_early_warning`.

4.2.2 Data collection

As shown in Table A.1, a range of datasets were used in the rice yield modelling, including climate reanalysis data, remote sensing data, district-wise historical rice yield data, and cropland masks. A visual overview of the geospatial data used in this research are also provided in Figure 4.1.

Climate reanalysis data. Daily climate reanalysis data on temperature, potential evaporation, surface pressure, leaf area index, total precipitation, and soil water content was obtained from ERA5 data from the European Center for Medium-Range Weather Forecasts (ECMWF), which provides global estimates of surface and atmospheric parameters since 1950 at a resolution of approximately 30*30 km [393]. Climate reanalysis data, which are often freely available, provide temporally and spatially homogenous data [394], which makes them suitable for applications such as crop yield prediction in contexts where in-situ weather station measurements are inadequate or incomplete. In addition, weather stations vary in their accuracy and generally record a limited number of variables, such as rainfall, temperature, pressure, and wind speed; variables that are more technically demanding to measure, such as humidity and solar radiation, may be lacking [395].

The climate variables used in this study were selected due to their influence on rice yields. An extensive body of research has shown that rice growth is affected by factors such as soil water content [396], temperature [397], potential evaporation (as a proxy for transpirational demand) [398], surface pressure [399], and precipitation [400]. These variables can impact rice development across phenological stages. Past case studies in India have shown the approximate number of days taken for each stage: the sowing to tillering phase (P1) can range from 30 to 60 days, and the rate of tillering tends to increase under higher temperatures; the tillering to panicle initiation phase (P2) can range between 42 to 49 days; the panicle initiation to flowering phase (P3) can range from 12 to 28 days; the flowering to milk phase (P4) can range from 7 to 20 days; and the mil to physiological maturity phase (P5) can range from 17 to 31 days [401]. One study showed reported a linear relationship between the days taken from sowing to flowering and average air temperature [402].

Remote sensing data. Normalized Difference Vegetation Index (NDVI) is a dimensionless index that describes the difference between visible and near-infrared reflectance of vegetation cover, and can be used to estimate the density of green on an area of land [403]. To determine the density of green on a patch of land, the wavelengths of visible and near-infrared sunlight reflected by the plants are observed. NDVI values range

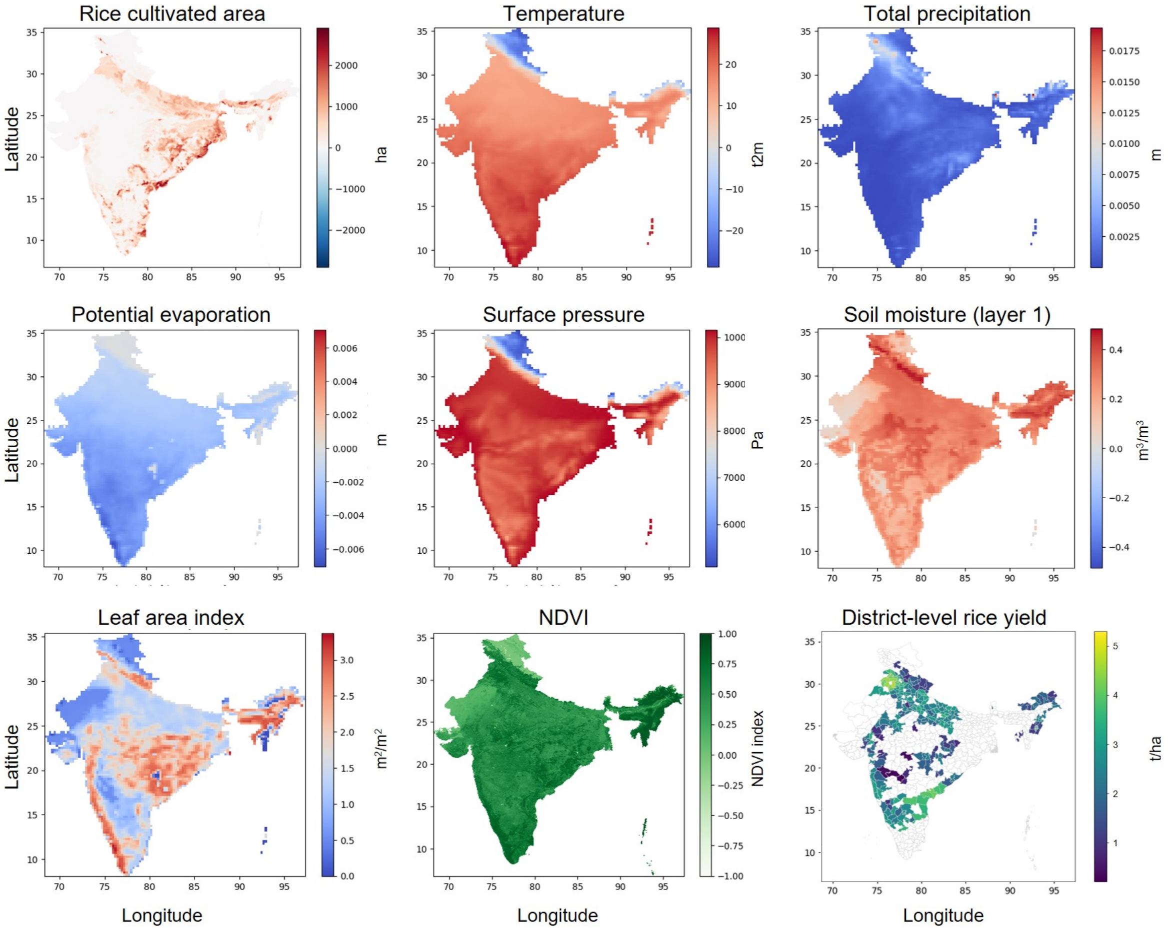


Figure 4.1: Overview of the geospatial data used in this research. The panels show, respectively: cultivated rice area in India; snapshots of temperature, total precipitation, potential evaporation, surface pressure, soil moisture, and leaf area index from the ECMFW (average values of January 2022); NDVI data from NASA's MODIS (average values of January 2022); and district-level rice yield data from India's Ministry of Agriculture and Farmers Welfare (in 2020).

from -1 to +1; higher values of NDVI imply healthy and dense vegetation, whereas lower NDVI values indicate sparser vegetation. NDVI data was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard NASA's Terra and Aqua satellites, due to their wide coverage and temporal resolution. The MOD13A3 v061 product was used to provide monthly NDVI data at a per pixel basis at 1km resolution. Monthly averages of vegetation indices can help to mitigate the impact of short-term missing values. By averaging over a month, the product reduces the influence of occasional cloud cover or other issues that might affect individual daily observations.

There are several examples in academic literature of using NDVI to investigate the progress of crops, such as for wheat in Argentina [404], cereals in Europe [405], and rice in Vietnam [406]. NDVI data was masked using CROPGRIDS, a global, geo-referenced dataset providing information on areas for 173 crops circa the year 2020, at a resolution of 0.05° ($\sim$ 5.55 km at the equator) [407].

Yield data. District-level rice production and yield data from 1995 to 2021 for 367 districts were obtained from the APY dataset of the Directorate of Economics and Statistics in India's Ministry of Agriculture and Farmers Welfare [408]. In this dataset, the year denotes the year in which the crop was harvested. For kharif season rice, the sowing is in the previous calendar year [409]. The raw yield data were aligned with a shapefile representing Indian districts. This alignment was achieved using the FuzzyWuzzy library in Python, which employs the Levenshtein distance to address the challenge of slight spelling discrepancies in the district names between the yield data and the shapefile [410].

Data ingestion. Data used in this analysis were programmatically downloaded via API and automated Python scripts. ERA5 data was ingested via the CDS API, while NDVI data was ingested via USGS' AppEEARS API.

Table 4.1: Earth observation datasets used in the machine learning models. A range of agronomically-relevant datasets were used as predictors of the target variable (district-level rice yield in India); rice areas masks were used to filter NDVI data by rice-growing area.

Data type	**Parameter**	**Description**	**Unit**	**Source**
Climate reanalysis	Potential evaporation (pev)	A measure of the extent to which near-surface atmospheric conditions are conducive to the process of evaporation.	m	ECMFW (ERA5)

Table 4.1 continued from previous page.

Data type	**Parameter**	**Description**	**Unit**	**Source**
	2m-temperature	The temperature of air at 2m above the surface of land, sea or inland waters. 2m temperature is calculated by interpolating between the lowest model level and the Earth's surface, taking account of the atmospheric conditions.	K	ECMFW (ERA5)
	Total precipitation	The accumulated liquid and frozen water, comprising rain and snow, that falls to the Earth's surface. It is the sum of large-scale precipitation and convective precipitation.	m	ECMFW (ERA5)
	Leaf area index, low vegetation (LAI)	The surface area of one side of all the leaves found over an area of land for vegetation classified as 'low'. 'Low vegetation' consists of crops and mixed farming, irrigated crops, short grass, and more.	m^{2}-m^{-2}	ECMFW (ERA5)
	Surface pressure	The pressure (force per unit area) of the atmosphere on the surface of land, sea and inland water. It is a measure of the weight of all the air in a column vertically above the area of the Earth's surface represented at a fixed point.	Pa	ECMFW (ERA5)
	Volumetric soil water (SWVL1)	The volume of water in soil layer 1 (0 - 7cm, the surface is at 0cm).	m^{3}-m^{-3}	ECMFW (ERA5)
Remote Sensing	Normalized Difference Vegetation Index (NDVI)	A dimensionless index that describes the difference between visible and near-infrared reflectance of vegetation cover, and can be used to estimate the density of green on an area of land. This study leveraged the MOD13A3 v061 product, which provides 16-day and monthly vegetation index values at a per pixel basis at 1 kilometer (km) spatial resolution.	(-)	NASA EOSDIS (AQUA MODIS)
Crop mask	Rice crop mask	A comprehensive global, geo-referenced dataset providing information on areas for 173 crops circa the year 2020, at a resolution of 0.05° (5.55 km at the equator).	ha	CROP

4.2.3 Data pre-processing and feature engineering

Climate data from ERA5 in NetCDF format over a bounded area comprising India was clipped to the Indian country boundary. NDVI data from NASA's AQUA MODIS satellite in NetCDF format were clipped to the Indian country boundary and masked with a rice cropland layer.

Next, the climate variables and NDVI were aggregated to the district level based on zonal statistics. The vector geometry data for India's ADM2 (district-level) boundaries which raster pixels were aggregated to were obtained from the Database of Global Administrative Areas (GADM) [411]. District-level yield data from APY was then merged to the climate and remote sensing data aggregated at the district level to produce a spatially consistent geodataframe. Yield outliers beyond three standard deviations were removed as they were assumed not achievable at the district level in India [388].

Feature engineering was conducted to produce monthly averages for the climate and NDVI parameters for every month between May and November, corresponding to the full sowing and growing period for kharif rice [388]. This process was repeated for all variables to produce a set of 52 features used as input for the modelling. The months selected for climate and NDVI feature aggregation were chosen to reflect the full range of rice growth stages, including the grain filing, vegetative, and reproductive stages [412].

The scope of this study was limited to weather and satellite data because these data sources are available at large scales and can be accessed programmatically via APIs. The aim of this study is to demonstrate whether weather and satellite data alone can provide strong predictive performance for rice yields. This focus ensures that the model remains practical and scalable. Nevertheless, the district identifier was used as a feature in the model to account for district-specific characteristics. This approach is consistent with yield prediction focused on other crops beyond rice [274]. An assumption here is that by including the district, the model indirectly incorporates region-specific management practices such as fertilizer application, tillage practices, and other crop management techniques that vary across districts.

4.2.4 Model development and interpretation

This study developed and tested the performance of multiple rice yield prediction models based on a variety of machine learning models. These included LightGBM, an efficient and distributed gradient boosting framework that uses tree-based learning, Bayesian ridge regression, which has been recognized for its ability to deal with hierarchical data structures, gradient boosting regression, random forest, Huber regression, decision tree regression, elastic net regression, AdaBoost, orthogonal matching pursuit, and extremely randomized trees.

The models above were trained on district-level data for 2001 to 2018 (4,606 observations), and validated on out-of-sample test data for 2019 and 2020 (502 observations). The data was split in a manner that reflects how yield prediction models may be used in practice, avoiding random splits in favor of chronological splits to help ensure the model's robustness to future, unseen data. This out-of-sample approach to testing regression models with temporal dependency has been shown to be more robust than cross-validation approaches tailored to time series problems [413]. For model training, default hyperparameters were used for consistency across models and to enable fair comparison. A large-scale study has shown that hyperparameter tuning often results in marginal performance gains compared to default parameters for regression models such as random forest or XGBoost [414]. The full code behind the model development is provided in the GitHub link in section 2.1.

The top-performing models were evaluated based on three out-of-sample performance measures including R2, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). Also reported were Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Logarithmic Error (RMSLE). In addition, model results were evaluated based on prediction error plots, residual plots, and spatio-temporal plots of prediction error to evaluate potential model bias (for example, better model performance for certain rice-growing regions of India). Lastly, Shapley Additive exPlanations (SHAP) were used to explore the impact of features on model output [415]. SHAP values are a model-independent methodology used for quantifying the significance of features in predictive modeling. The SHAP of feature j for observation $\phi_j(\mathbf{x})$ is defined as

$$\phi_j(\mathbf{x}) = \sum_{S \subseteq \{1,\ldots,M\} \setminus \{j\}} \frac{|S|! \cdot (M - |S| - 1)!}{M!} \left(f(\mathbf{x}_{S \cup \{j\}}) - f(\mathbf{x}_S) \right) \tag{4.1}$$

where j is the feature evaluated, M the total number of features, S a subset of the full feature set {1,. . . ,M} that does not include the feature j, $\mathbf{x}_S$ a subset of features in S, and f the model's prediction function [416].

4.2.5 Computation

Data ingestion, pre-processing, and modelling was conducted in a conda-based python environment with a diverse set of python libraries. Data processing and geospatial operations were carried out using python libraries including numpy, xarray, pandas, rasterio, rasterstat, and geopandas. Modelling and visualization was conducted using python libraries including scikit-learn, pycaret, matplotlib, and seaborn.

4.2.6 Interactive visualization for model evaluation and decision-making

Visual dashboards can serve as a helpful tool for making the outputs of analytical models more comprehensible to stakeholders by converting predictive analytics into easily interpretable visual formats. This transformation is particularly beneficial for users from various backgrounds, including farmers, policymakers, and researchers, as it facilitates their engagement with and comprehension of the data. For example, platforms like Streamlit and PowerBI have been used to build visual web applications and dashboards across diverse fields, such as bioinformatics, bacterial testing, financial auditing, twitter sentiment analysis, credit card fraud detection, drug target prioritization, and pharmaceutical sales forecasting [357, 358, 359, 360, 361, 362].

In this study, we utilize the dashboard to not only visualize the model's predictions, such as forecasted yields in each Indian district, but also to offer insights into the model's diagnostic aspects, like its varying predictive accuracy across different regions. We developed two distinct visual dashboards: the first provides a clear, spatially detailed overview of the model's forecasts, and the second aids in identifying and understanding any potential errors in the model's performance.

4.3 Results and discussion

This study assessed the feasibility of predicting pre-harvest rice yields in India using machine learning models, with a focus on satellite and climate reanalysis data. The results, which are detailed in the subsequent sections and compared to other global studies of rice yield prediction, show that the models achieve strong performance across several metrics in out-of-sample tests. Such results affirm the potential of these models for rice yield forecasting and provide a benchmark for predictive precision in the field.

4.3.1 Overview of out-of-sample model performance

Table 4.2 summarizes the out-of-sample (validation set) performance across the models tested: R2, MAE, and MAPE values of up to 0.82, 0.29, and 0.16 respectively were achieved. Compared to out-of-sample results reported in previous literature on rice yield prediction in different parts of the world, the models perform well. For instance, one study which developed rice yield prediction models for China based on support vector machine regression, neural networks, and random forest, achieved R2 values ranging from 0.24 to 0.31 and MAE values ranging from 0.58 to 0.66 t/ha [417]. Another study estimating rice yields in Vietnam's Mekong Delta reported out-of-sample MAE values ranging from 0.46 to 0.55 t/ha for Winter and Summer rice models [418]. A study on county-level rice yield prediction in China's Jiangsu province reported out-of-sample R2 values of 0.39 to 0.59 on an independent holdout set [384]. A study on pixel-scale rice yield prediction in South Korea reported test-set R2 values of 0.80 [387]. One image-driven yield prediction study reported test-set R2 values of 0.65 [419]. A study using multi-temporal UAV-based multispectral vegetation indices reported test set R2 values of up to 0.80 [420].

Table 4.2: Model performance on out-of-sample test data shows that the top three (based on R2 and MAPE) models include Random Forest, CatBoost, and Light Gradient Boosting. Experiment 1 ("all features") shows results for model runs including climate/satellite-derived features and the "Year" and "District" features. Experiment 2 shows results for model runs trained exclusively using climate and satellite data observations ("EO features only").

Model	**Experiment 1 – all features**				**Experiment 2 – EO features only**			
	MAE	**RMSE**	**R2**	**MAPE**	**MAE**	**RMSE**	**R2**	**MAPE**
Random Forest Regressor	0.31	0.41	0.80	0.16	0.45	0.56	0.63	0.25
CatBoost Regressor	0.29	0.39	0.82	0.18	0.43	0.53	0.67	0.24
Light Gradient Boosting Machine	0.31	0.41	0.80	0.19	0.44	0.56	0.64	0.25
Extreme Gradient Boosting	0.33	0.43	0.78	0.19	0.44	0.58	0.61	0.23
Orthogonal Matching Pursuit	0.33	0.46	0.76	0.20	0.76	0.96	-0.08	0.49
Decision Tree Regressor	0.41	0.56	0.63	0.20	0.58	0.81	0.24	0.33
Bayesian Ridge	0.33	0.46	0.76	0.21	7.51	11.81	-161.70	4.15
Gradient Boosting Regressor	0.32	0.41	0.80	0.21	0.50	0.62	0.55	0.28
Ridge Regression	0.34	0.47	0.75	0.21	0.61	0.78	0.30	0.37
Huber Regressor	0.33	0.46	0.75	0.21	0.75	0.95	-0.06	0.49
K Neighbors Regressor	0.39	0.51	0.70	0.21	0.55	0.70	0.44	0.31
Linear Regression	0.36	0.48	0.73	0.21	0.61	0.78	0.29	0.36
AdaBoost Regressor	0.45	0.55	0.65	0.27	0.63	0.75	0.34	0.40
Passive Aggressive Regressor	0.48	0.62	0.56	0.31	0.70	0.91	0.04	0.51
Elastic Net	0.66	0.81	0.24	0.41	0.74	0.94	-0.04	0.49
Lasso Regression	0.80	0.99	-0.13	0.53	0.74	0.94	-0.03	0.49
Lasso Least Angle Regression	0.80	0.99	-0.13	0.53	0.76	0.96	-0.07	0.48
Dummy Regressor	0.80	0.99	-0.13	0.53	0.80	0.99	-0.13	0.53
Least Angle Regression	1.90	2.35	-5.42	0.99	2.33	2.98	-9.34	1.21

The results also perform well compared to studies which only reported in-sample performance metrics. One study on rice yield modelling in Bangladesh reported in-sample R2 values ranging from 0.44 to 0.91; out-of-sample performance was not reported [421]. Another study on rice yield in China report in-sample R2 values of 0.77, lower than the out-of-sample R2 performance achieved in this study of 0.82 [422]. For rice yield prediction in the Philippines, one study reported an in-sample RMSE of 0.46 t/ha [423]. Another study using drones reported in-sample R2 values of 0.60 to 0.81 for rice yield prediction in Japan based on NDVI [424].

While these are not direct comparisons (i.e., other studies were done for rice in different regions and had different experimental setups), we have provided them to give the reader a range of the current state of the literature. This helps contextualize our findings within the broader scope of rice yield prediction studies, highlighting typical performance metrics achieved across various models.

4.3.2 Errors, residuals, and SHAP value analysis

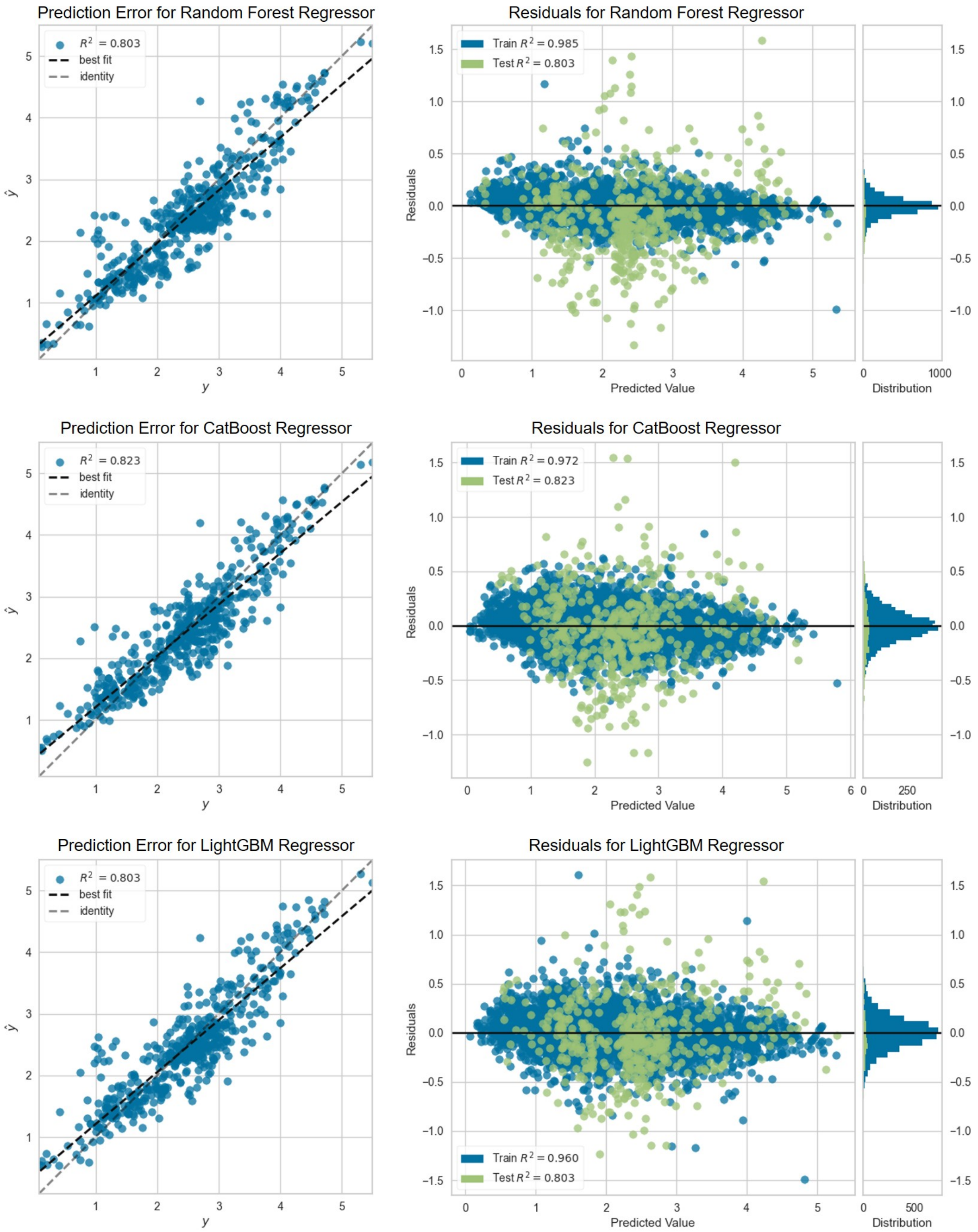
As shown in Figure 4.2, observed and simulated yields show a high level of agreement for some of the top-performing models including the random forest, CatBoost, and LightGBM regressors. In addition, the residual plot residual plots show that the majority of both training set and test set observations are randomly dispersed along the horizontal axis, indicating a reasonable low level of bias and homoscedasticity. The distribution of residuals here is roughly centred around zero but with some skewness, indicating the potential presence of outliers.

In addition to the errors and residuals, the SHAP summary plots in Figure 4.3 concisely display the magnitude, prevalence and direction of a variable's effect on final rice yield. The plots reveal that important variables include temperature, soil water volume ("SWVL1"), NDVI, and LAI in selected months. The importance ranking of these variables corroborates previous findings that factors such as soil water content, temperature, and NDVI are important factors in estimating rice growth [396, 397].

A closer analysis of the specific impacts of features on rice yield is shown in the right-hand side panels of Figure 4.3. For instance, increases in temperature in August, which coincides with the sowing to panicle initiation phases of rice growth, are associated with higher yields, corroborating previous findings in India that above average yields may be associated with higher maximum temperatures [401]. The full set of SHAP plots for all features is available in the analysis output on GitHub.

4.3.3 Interactive visualization tool

In addition to the SHAP plots, which provide insight into how variables drive yield outcomes, two visual dashboards were developed to (a) provide an easy-to-understand,



Prediction Error for Random Forest Regressor

Residuals for Random Forest Regressor

Prediction Error for CatBoost Regressor

Residuals for CatBoost Regressor

Prediction Error for LightGBM Regressor

Residuals for LightGBM Regressor

Figure 4.2: Comparative evaluation of selected models, including Random Forest, CatBoost, and LightGBM regressors using prediction error and residuals analysis. Two years of observations (502 observations in total) where used for the out-of-sample validation data, on which the Random Forest, CatBoost, and LightGBM models have test R^2 values of 0.80, 0.82, and 0.80 respectively. Residuals are mostly centered around zero, but CatBoost shows a skewness in error distribution. The histogram of residuals indicates Random Forest and CatBoost have a tighter error distribution compared to LightGBM's broader range.

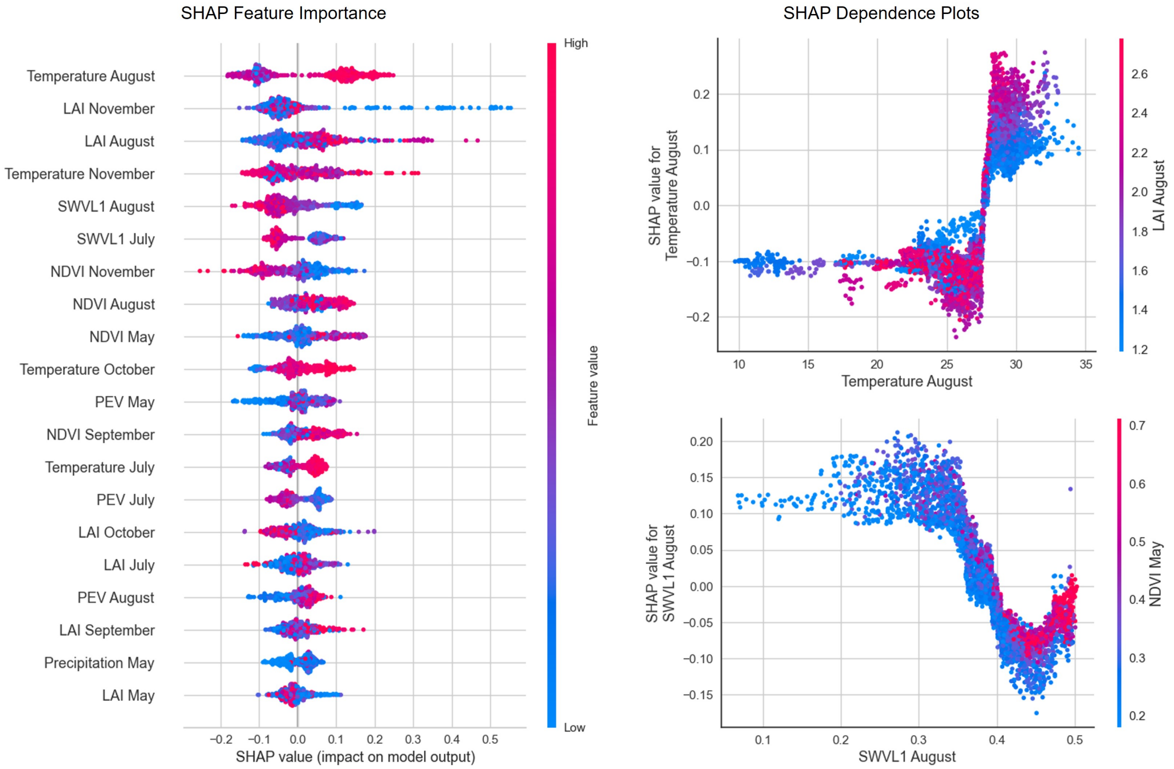


Figure 4.3: Interpretation of SHAP values for selected features in the rice yield prediction random forest model. The SHAP feature importance plot (left panel) exhibits the impact of various features on the model's output. Higher SHAP values indicate a greater influence on the predicted yield. The coloring on the feature importance plot represents the value of the feature for each data point. Blue points indicate low feature values, while pink points represent high feature values. This color gradient allows us to visualize not only the impact (magnitude of the SHAP value) each feature has on the model output but also the distribution of the feature's values. For instance, when examining 'Temperature August', we can see a mix of pink and blue points across a range of SHAP values, indicating a diverse range of temperatures in August within the dataset and how these varying temperatures correlate with the rice yield prediction. The top right panel presents a SHAP dependence plot for temperature in August, illustrating a correlation between higher temperatures and increased SHAP values for rice yield. The intensity of the color indicates the interaction effect, with a notable interaction with LAI in August, as higher LAI values (depicted in red) intensify the impact of temperature on yield. The bottom right panel depicts a SHAP dependence plot for soil water volume (SWVL1) in August, showing the relationship between SWVL1 values and SHAP values. This plot reveals that certain values of SWVL1 are associated with lower or higher SHAP values, indicating its varying influence on yield predictions, with the color intensity representing the interaction with NDVI in May.

spatially explicit summary of model predictions, and (b) to help to identify potential biases in model performance. These are shown in Figure 4.4 and Figure 4.5.

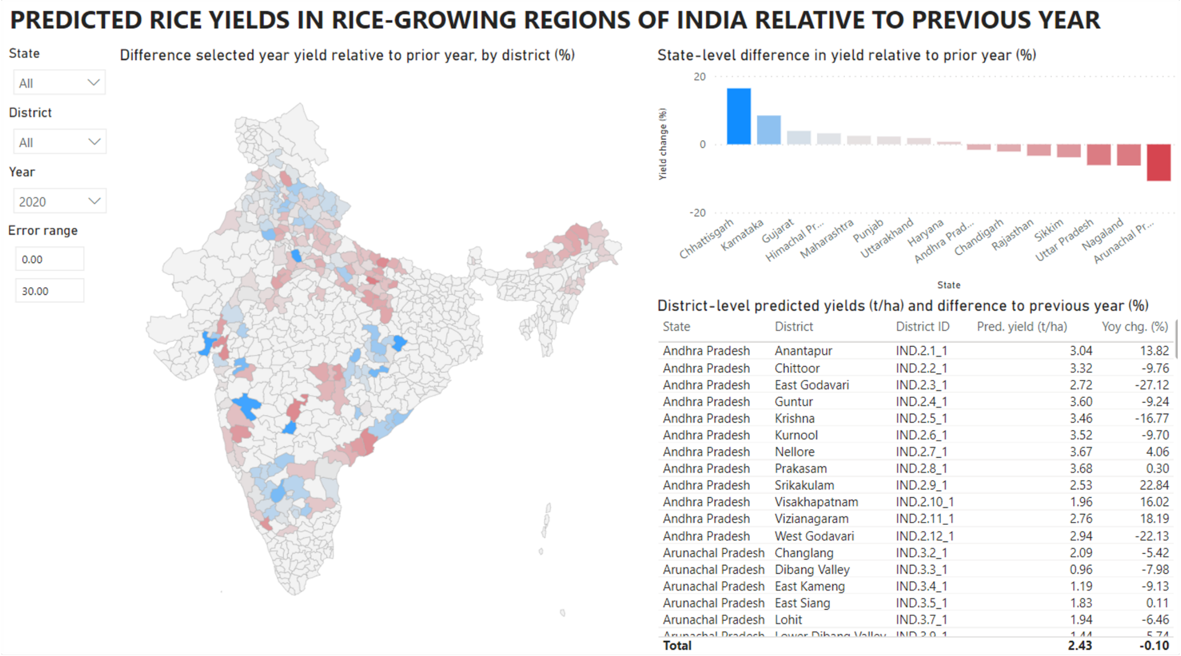
For example, in Figure 4.4 for the year 2020 (one of the test set years) the dashboard shows that districts in the state of Chhattisgarh with expected yield increases relative to the previous year included Jashpur, Korba, and Koriya, where yields were expected to increase by 52%, 22%, and 22%, respectively. In another example, districts in the state of Gujarat such as Kheda and Sabar Kantha were expected to see yields decrease by 27% and 21% respectively according to the model.

In addition to a visual representation of the predictions, the dashboard also provides a spatial view of prediction errors in order to more easily identify areas where model predictions may be inaccurate (Figure 4.5). For instance, the dashboard shows that on average, the absolute percentage error for predicted versus actual yields ranged from an average of 7.1% in districts in Uttarakhand to an average of 14.7% in Uttar Pradesh, implying that the model may perform better in some regions than others.

4.3.4 Limitations and future research directions

Below, we present a non-exhaustive list of potential future research directions to build on the results of this research, including: refining model accuracy by exploring additional variables; communicating the outputs of yield prediction models in early warning systems by leveraging large language models; modeling how policymakers can help to disseminate the yield predictions-based early warning tools; and combining yield prediction modeling results with optimization approaches to support to anticipatory aid allocation efforts.

Modeling enhancements to increase predictive power. Increased predictive power of rice yield model may potentially be achieved by incorporating a wider array of agronomically relevant climatological variables. Prior investigations have highlighted the influence of thermal extremes, precipitation extremes, daytime humidity variations, and solar radiation on rice yields [401]. In addition, agricultural yield models – particularly those that are grounded in time series analysis – may stand to gain from incorporating additional auto-regressive elements, rolling averages, or cumulative indices, such as total sunshine hours or cumulative rainfall since rice sowing [425]. Beyond remotely sensed variables, yield models may also benefit from the inclusion of variables related to socio-economic changes in farmer populations, cultural and agronomic management practices that affect rice cultivation, plant varieties, and market or policy-related shifts that provide incentives for farmers to cultivate their crop in different ways [426, 427, 428, 429]. In addition, model experimentation could be conducted with different combinations of these features in order to determine how to minimize the difference in prediction accuracy across different regions of India.



PREDICTED RICE YIELDS IN RICE-GROWING REGIONS OF INDIA RELATIVE TO PREVIOUS YEAR

State: All

District: All

Year: 2020

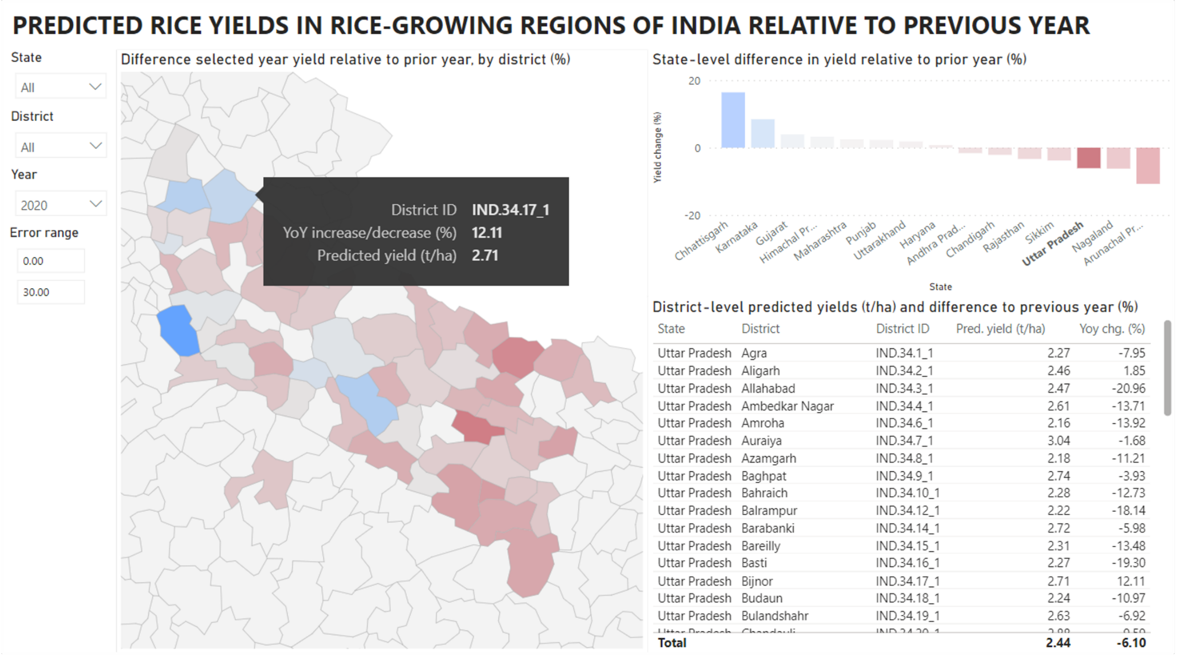
Error range: 0.00 – 30.00

Difference selected year yield relative to prior year, by district (%)

State-level difference in yield relative to prior year (%)

District-level predicted yields (t/ha) and difference to previous year (%)

State	District	District ID	Pred. yield (t/ha)	Yoy chg. (%)
Andhra Pradesh	Anantapur	IND.2.1_1	3.04	13.82
Andhra Pradesh	Chittoor	IND.2.2_1	3.32	-9.76
Andhra Pradesh	East Godavari	IND.2.3_1	2.72	-27.12
Andhra Pradesh	Guntur	IND.2.4_1	3.60	-9.24
Andhra Pradesh	Krishna	IND.2.5_1	3.46	-16.77
Andhra Pradesh	Kurnool	IND.2.6_1	3.52	-9.70
Andhra Pradesh	Nellore	IND.2.7_1	3.67	4.06
Andhra Pradesh	Prakasam	IND.2.8_1	3.68	0.30
Andhra Pradesh	Srikakulam	IND.2.9_1	2.53	22.84
Andhra Pradesh	Visakhapatnam	IND.2.10_1	1.96	16.02
Andhra Pradesh	Vizianagaram	IND.2.11_1	2.76	18.19
Andhra Pradesh	West Godavari	IND.2.12_1	2.94	-22.13
Arunachal Pradesh	Changlang	IND.3.2_1	2.09	-5.42
Arunachal Pradesh	Dibang Valley	IND.3.3_1	0.96	-7.98
Arunachal Pradesh	East Kameng	IND.3.4_1	1.19	-9.13
Arunachal Pradesh	East Siang	IND.3.5_1	1.83	0.11
Arunachal Pradesh	Lohit	IND.3.7_1	1.94	-6.46
Total			**2.43**	**-0.10**



PREDICTED RICE YIELDS IN RICE-GROWING REGIONS OF INDIA RELATIVE TO PREVIOUS YEAR

State: All

District: All

Year: 2020

Error range: 0.00 – 30.00

Difference selected year yield relative to prior year, by district (%)

District ID IND.34.17_1

YoY increase/decrease (%) 12.11

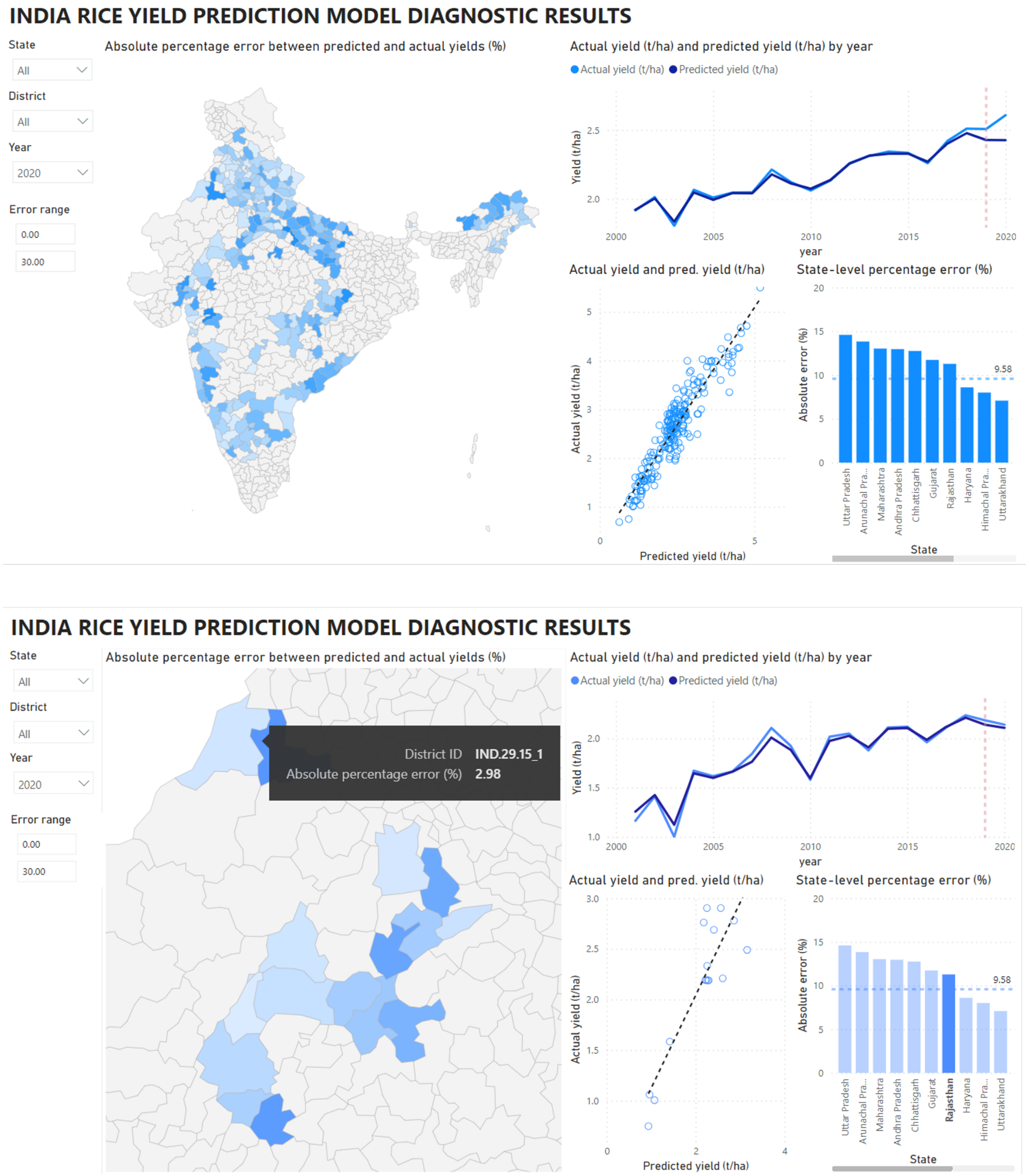
Predicted yield (t/ha) 2.71

State-level difference in yield relative to prior year (%)

District-level predicted yields (t/ha) and difference to previous year (%)

State	District	District ID	Pred. yield (t/ha)	Yoy chg. (%)
Uttar Pradesh	Agra	IND.34.1_1	2.27	-7.95
Uttar Pradesh	Aligarh	IND.34.2_1	2.46	1.85
Uttar Pradesh	Allahabad	IND.34.3_1	2.47	-20.96
Uttar Pradesh	Ambedkar Nagar	IND.34.4_1	2.61	-13.71
Uttar Pradesh	Amroha	IND.34.6_1	2.16	-13.92
Uttar Pradesh	Auraiya	IND.34.7_1	3.04	-1.68
Uttar Pradesh	Azamgarh	IND.34.8_1	2.18	-11.21
Uttar Pradesh	Baghpat	IND.34.9_1	2.74	-3.93
Uttar Pradesh	Bahraich	IND.34.10_1	2.28	-12.73
Uttar Pradesh	Balrampur	IND.34.12_1	2.22	-18.14
Uttar Pradesh	Barabanki	IND.34.14_1	2.72	-5.98
Uttar Pradesh	Bareilly	IND.34.15_1	2.31	-13.48
Uttar Pradesh	Basti	IND.34.16_1	2.27	-19.30
Uttar Pradesh	Bijnor	IND.34.17_1	2.71	12.11
Uttar Pradesh	Budaun	IND.34.18_1	2.24	-10.97
Uttar Pradesh	Bulandshahr	IND.34.19_1	2.63	-6.92
Total			**2.44**	**-6.10**

Figure 4.4: Interactive dashboard of yield prediction model outputs. The top panel shows a map of India with a colour coding applied to different districts to indicate the predicted yield values. Shades of blue indicate an increase in yield and shades of red denote a decrease in yield compared to the prior year's yield. Accompanying the map is a bar chart that provides a state-level summary and a table that enumerates the district-level predicted yields and the percentage change from the previous year across all states and districts. The bottom panel provides a similar comparative yield prediction, but focuses on the state of Uttar Pradesh. The dashboard is downloadable on GitHub.



INDIA RICE YIELD PREDICTION MODEL DIAGNOSTIC RESULTS

State: All
District: All
Year: 2020
Error range: 0.00 – 30.00

INDIA RICE YIELD PREDICTION MODEL DIAGNOSTIC RESULTS

State: All
District: All
Year: 2020
Error range: 0.00 – 30.00

District ID: IND.29.15_1
Absolute percentage error (%): 2.98

Figure 4.5: Interactive dashboard showing spatial view of yield prediction model error. The top panel provide model diagnostic information. The visual shows a map with the average percentage error by region; a scatter plot comparing the predicted yield and the actual yield; and a line graph showing the actual yield and predicted yield each year. The lower panel shows a similar view, zoomed in to districts within the state of Rajasthan. The dashboard is downloadable on GitHub.

However, the augmentation of the feature set should also be approached with caution, as an expanded variable set can inadvertently lead to over-fitting, introduce feature redundancy, and complicate model interpretability. In scenarios where model clarity and understandability are important, a more conservative approach with a reduced set of features might be advisable to streamline the interpretive process, allowing for clearer insights and more straightforward decision-making [430]. Approaches which balance model interpretability with overall accuracy may be more helpful to decision makers.

Leveraging large language models to communicate yield prediction results. Additional research is needed on how access to climate and yield anomaly information via user-tailored interfaces can help mitigate meteorological shocks for agricultural communities. Studies have shown that inadequate access to climate information in South Asia has been observed as a factor driving perceived losses in farming communities [431].

Large language models (LLMs) may have potential to enhance the utility and accessibility of crop yield prediction models. By integrating LLMs with rice crop yield models, the data these yield models generate can be transformed into concise reports disseminated through channels such as text message-based farmer advisories or agricultural extension services. LLMs are able to process both structured and unstructured data, such as summarizing tabular data on yield predictions for various regions and translating these summaries into different local languages, thus increasing their accessibility and ease of understanding [432, 433].

Moreover, LLMs may have the potential to offer personalized agronomic advice based on location-specific yield predictions. Their proven capability in handling expert-level tasks across various fields, including agronomy, suggests they could be effectively deployed in agricultural settings [434]. For instance, existing digital solutions like KissanGPT, a chatbot designed to assist farmers with queries such as optimal fertilizer application, indicate the practical applications of LLMs in agriculture [435]. LLMs could help make rice yield prediction models more accessible to farmers through web platforms or mobile messaging applications, like WhatsApp, allowing for interaction in their native languages. Examples of mobile-based agricultural early warning systems have included wheat rust disease alerts in Ethiopia and weather alerts in Zimbabwe [436].

Yield prediction systems that leverage LLM technology to communicate outputs may also help bolster agricultural extension services. Bangladesh, for instance, has boosted support services for farmers, including enhanced information access and extension services, via the government's 'Info Sarkar' project, which aims to link government offices nationwide and has established over 4,500 internet-equipped Union Digital Centres (UDCs) to aid rural communities. Studies indicate these centers are being effectively utilized by educated youths, suggesting a potential for these individuals to lead in disseminating climate adaptation strategies to farmers [431].

Such agricultural extension services could benefit from additional digital aid such as LLMs that are able to distill agronomic science in a manner that is understandable and actionable for farmers, especially in contexts where agronomic advice should be disseminated in a context-specific manner. Advisories could be targeted based on whether farmers socio-economic status, the size of their farms, the level of internet connectivity, household income, amongst other factors. Past research has shown that differences in farmers backgrounds can significantly affect adoption of agricultural technologies, suggesting that tailored advice may be an important prerequisite for helping farmers adapt their agronomic practices in the face of climate-related risk [437]. The outputs of LLMs could be further bolstered with specialized recommendations by analytical decision support tools such as Nutrient Expert, a nutrient management tool that helps farmers boost yields [438].

Modelling the dissemination of yield prediction outputs. Building yield prediction systems that are tailored to local needs may help to boost resilience in agricultural communities. However, good design alone may be insufficient for widespread adoption. There is also a need to anticipate how external factors (such as mobile network coverage or regional agroclimatic variation) might affect the diffusion of this technology in a society across time and space, and how policymakers can enhance and sustain adoption.

The adoption of these systems could potentially be modeled using principles of the Diffusion of Innovation theory [439]. This theory provides a framework to understand how innovations are taken up in a population, highlighting the role of early adopters and the subsequent spread through social and communication networks. In the context of yield prediction systems, this might involve analyzing how the perceived advantages and compatibility with existing practices influence the rate of adoption among farmers or agricultural extension workers.

Furthermore, computational models of diffusion, such as the Bass diffusion model, can offer insights into the expected rate of technology uptake. These models provide a means to simulate and anticipate the adoption curve, taking into account various societal and technological factors [440]. By understanding the likely trajectory of technology diffusion, policies can be tailored to support and accelerate adoption, ensuring that the benefits of yield prediction systems are maximized. Applications of Bass-like models are numerous in the literature, but have yet to be applied to agricultural early warning technologies. Studies have applied the methodology to forecasting the diffusion of innovations including novel foods such as edible insects in the Netherlands, groundwater pumps in Pakistan, preterm birth screening technology, fuel cell vehicles in China, and solar water heaters in Brazil [441, 442, 443, 444, 445, 446].

Leveraging yield prediction models for anticipatory aid allocation. Another potential area of research could involve investigating how to effectively integrate yield prediction algorithms into decision-making tools for the public sector to enhance the strategic allocation of humanitarian aid in response to weather-related challenges. For example, algorithms that produce anticipated rice yields in specific areas could be combined with optimization algorithms can assist government decisions on the anticipatory allocation of resources such as financial support, fertilizers, fungicides, and water resources among affected farming communities.

While examples of such optimization applied specifically to anticipatory aid allocation in agricultural settings are limited, there is a growing body of research on applications of operations research applied to humanitarian aid [447]. Examples of using optimization approaches (such as linear mixed integer programming, stochastic programming, or multi-objective optimization) in relevant contexts include: aid disbursement following the 2010 Haiti earthquake [448]; optimal aid disbursement in response to internal displacement in northwest Syria [449]; post-disaster distribution of essential humanitarian aid (medicine, food, and water) from temporary warehouses to points of demand in Peru [450]; optimal location planning of warehouse locations to store relief items in Thailand [451]; logistics distribution of essential relief items during COVID-19 lockdowns in Bangladesh [452]; optimization of United Nations Humanitarian Response Depot distribution plans [453]; humanitarian relief logistics in both pre- and post-disaster situations in presence of uncertainty [454]; and the World Food Programme's Optimus tool, which leverages linear programming to optimize food aid operations [455].

4.4 Conclusion

This study advances the state of the art in district-level rice yield prediction in India through an integrated approach, combining ERA5 climate reanalysis, MODIS satellite vegetation indices, and a novel, spatially matched yield dataset. By evaluating 19 machine learning models, the research establishes benchmarks for accuracy, achieving out-of-sample R2, MAE, and MAPE values of up to 0.82, 0.29, and 0.16, respectively.

The development of an interactive dashboard tool offers a means for visualizing yield predictions and assessing model performance across regions. This approach not only demonstrates the feasibility of using machine learning for rice yield forecasting at the district level in India, but also provides a benchmark for predictive accuracy against which further algorithmic innovations can be easily compared. The results offer evidence that machine learning-based rice yield prediction may have the potential to augment Indian agricultural early warning systems with robust crop yield prediction capabilities. While the focus of this study was Kharif-season rice in India, the findings suggest that the approach could be replicated and scaled to other regions, warranting further investigation.

5 Planning ahead: optimizing early warning-driven anticipatory action cash transfers via an open-source web application

---

Chapter currently under review in *International Journal of Disaster Risk Reduction*.

Abstract

This chapter unites the earlier developments in early warning and yield forecasting with targeted anticipatory action, showing how mathematical programming can harness those forecasts to strategically allocate cash transfers in advance of disasters. By merging vulnerability data, whether drawn from simulated hazard maps or operational early warning systems, our open-source decision-support tool distributes pre-disaster funds over multiple time steps, ensuring that the most at-risk farming communities receive timely and proportionate support. Using West Bengal in India as a case study, we present an interactive web application geared toward humanitarian aid stakeholders, from policymakers to NGO field officers, that enables parameter tuning (e.g., budgets, hazard intensity) and displays optimized payout maps without requiring coding expertise. Our results also indicate that this approach scales efficiently to thousands of locations, setting the stage for wider deployment. The approach is tested in real-world applications via simulations across a range of scenarios and compared to a blanket aid allocation approach. By linking impact-based early warnings to a straightforward, rigorously tested allocation framework, this work advances the overarching thesis objective: designing anticipatory measures that are not only prompt but also equitable, data-driven, and highly effective in safeguarding vulnerable agricultural livelihoods.

5.1 Introduction

5.1.1 The case for optimizing anticipatory action

In agricultural regions worldwide, communities are increasingly vulnerable to climate-related crises, including droughts, floods, and erratic temperature shifts [60, 61, 62, 84, 456, 457]. Anticipatory action, a proactive approach to disaster management informed by early warnings, aims to reduce the impacts of these crises by enabling preparedness before disasters strike [20]. For example, distributing cash transfers to farmers ahead of a predicted drought or subsidizing seeds for flood-resistant crops has proven to be more cost-effective than reacting after the fact [458]. This approach has gained traction among governments and international aid organizations, reducing response costs while safeguarding livelihoods in regions such as East Africa and Mongolia [459, 64]. In Bangladesh, where the 2020 monsoon floods were among the most severe in decades, anticipatory cash transfers of BDT 4,500 (about \$53) helped nearly 3,800 vulnerable households evacuate, protect assets, and avoid high-interest debt, demonstrating the potential of forecast-based funding to mitigate disaster impacts [266]. Anticipatory action frameworks now operate in 47 countries, with 107 programs safeguarding 10.9 million people using \$147.8 million in pre-agreed funding. In 2023 alone, 98 activations reached 12.8 million people with \$198 million. Successful examples include Kenya's livestock protection (boosting milk yields) and Mongolia's winter feed distribution (reducing livestock deaths) [64].

Anticipatory action holds promise but can fall short when implemented with overly rigid triggers and blanket interventions. Traditional approaches often deploy support, such as cash assistance, after rainfall drops below a fixed threshold for droughts or when flood predictions reach a certain probability. Although these triggers help mobilize aid, they do not always account for the diversity of needs within affected regions. Two neighboring villages might both experience drought conditions that activate cash transfers, for instance, yet one village may have reliable groundwater resources while the other lacks them entirely. Distributing the same level or type of assistance in both cases could lead to inefficiencies or unintended inequities. For example, one study highlighted the gaps in determining "the optimal balance between in-kind service delivery and cash assistance" and "the frequency and amount of distributions" [64]. As highlighted in recent studies, decisions around how much aid to provide, what form it should take (e.g., cash versus in-kind), and how often it should be distributed must adapt to local realities. To achieve fairer, more impactful outcomes as anticipatory action scales globally, organizations can minimize subjective judgment by integrating nuanced, data-driven criteria, ensuring support is tailored to varying community contexts rather than triggered by a single indicator alone [31, 460].

Optimized anticipatory action harnesses tools from operations research and mathematical optimization to allocate resources strategically, moving beyond fixed triggers and uniform responses to evaluate complex trade-offs under uncertainty. For instance, policymakers must decide whether to prioritize cash subsidies in one district while distributing water pumps in another, or how to allocate limited funds to achieve maximum impact under tight budget constraints. These challenges, mirrored in simulation games like those developed by the Red Cross Red Crescent Climate Centre [63], reflect real-world pressures faced by governments and humanitarian organizations, such as investing in cash transfers, distributing seeds, or funding micro-irrigation systems with incomplete data and limited time. In India, for example, a recent monsoon failure left farmlands "reduced to rubble," yet only 10% of affected farmers reportedly received any compensation, with some still awaiting support more than a year later [461].

While methods in mathematical programming, such as linear and stochastic programming, have been extensively studied in humanitarian logistics, such as in optimizing food distribution networks and supply chain operations [211, 214, 215], their application to anticipatory action remains largely unexplored. By explicitly framing decisions in terms of objectives, constraints, and uncertainties, optimization offers a structured framework to make anticipatory action more scalable, equitable, and impactful, addressing the inherent trade-offs and resource constraints that limit its effectiveness.

5.1.2 Allocating anticipatory funds under early warning

When organizations involved in anticipatory action (ranging from national Red Cross and Red Crescent societies to Caritas affiliates, and from local governments working in partnership with UN agencies or international NGOs) receive an impact-based forecast that a flood, cyclone, or drought is imminent, they unlock a ring-fenced pot of "anticipatory" money [64]. Their task is to turn the rapid flow of early-warning intelligence such as ensemble weather forecasts, satellite rainfall estimates such as CHIRPS [462], river-basin flood models like GloFAS [463], and static layers on poverty or coping capacity [464] into a concrete plan that specifies who should receive what support, where, and when. In settings where fairness across subregions matters, the affected area may also be partitioned into analyst-defined groups, such as administrative zones, spatial bands, or livelihood categories, so that allocations can be compared across those groupings as well as across individual pixels. Together, these data streams highlight geographic hotspots where lives and livelihoods are most at risk days to weeks before landfall.

The organization must then decide how to spend a limited, pre-agreed budget: for example, whether to send cash, fertilizer, or water pumps; how much to send to each village; and how to do so fairly across districts while staying within logistical and stock limits. The overarching objective is simple in principle: shield the greatest number of vulnerable people before the hazard hits. But this objective must be balanced against real-world constraints on money, time, and delivery capacity. The methodology section that follows shows how this decision problem can be formalised and, in the context of India, solved with optimisation tools to produce a transparent, auditable allocation plan.

5.1.3 Contributions of this research

In this decision-making context, the present study centers on India, a country where heightened climate variability and an extensive agricultural sector underscore the urgency of effective anticipatory action strategies [465]. West Bengal, in particular, exemplifies the challenges and opportunities for pre-disaster interventions. With a history of cyclones, floods, and crop losses, regions such as Purba Medinipur illustrate how optimizing anticipatory aid allocations could mitigate adverse impacts on vulnerable farming communities [466]. By examining this specific setting, the study advances understanding of how mathematical optimization can refine and scale anticipatory action efforts where constraints on time and resources are pronounced.

This paper makes six main contributions, organized into advances for research and practical tools for policymakers and practitioners:

New research contributions

- **Pixel-level, time-phased MILP for anticipatory aid allocation.** We extend existing optimisation-based approaches by formulating a spatially disaggregated, multi-phase mixed-integer linear programming (MILP) model that balances equity, vulnerability, and logistical constraints. This advances anticipatory action research by explicitly integrating dynamic spatial-temporal trade-offs into resource allocation decisions.
- **Pixel-level hazard vulnerability simulation pipeline.** We develop a hazard simulation module that generates detailed, pixel-level vulnerability maps from synthetic disaster scenarios. Structured to accept real-time early warning inputs, this pipeline bridges the gap between scenario-based vulnerability assessment and real-time operational readiness, allowing rigorous validation of anticipatory opti-

misation methods.

- **Scalability testing across spatial and institutional dimensions.** We systematically test the model's computational performance across varying spatial resolutions, solver settings, and constraint configurations. This evaluation contributes novel insights into model tractability, clarifying computational trade-offs crucial for scaling anticipatory optimisation methods to broader geographic and policy contexts.

New practical tools for policymakers and practitioners:

- **Open-source dashboard for policy testing.** We provide an interactive, open-source web dashboard enabling policymakers and humanitarian practitioners to dynamically visualise and adjust resource allocation scenarios. This practical tool lowers technical barriers to implementing optimisation-driven anticipatory action, facilitating broader stakeholder engagement in evidence-based decision-making.
- **Empirical case study in coastal India.** We illustrate the framework's practical applicability through an empirical study in Purba Medinipur, West Bengal, a climate-vulnerable agricultural region. This case study demonstrates the operational feasibility of optimized anticipatory aid allocations, validating its relevance and potential effectiveness in real-world settings.
- **Comparative evaluation against a population-based baseline.** We benchmark the proposed optimisation framework against a conventional per-capita allocation rule across multiple synthetic hazard scenarios. The comparative evaluation clearly highlights the practical benefits of constraint-aware optimization, particularly improved vulnerability targeting and expanded coverage, underscoring its value for policy implementation.

5.2 Methods

This section outlines the methods developed to link early warning system outputs with optimization-based anticipatory aid allocations. First, the rationale for selecting India, particularly Purba Medinipur district in West Bengal, as a case study is explained, highlighting the region's vulnerability to climate-induced agricultural risks. Next, the simulation approach for generating farmer vulnerability data is described, detailing how spatially explicit hazard scenarios are constructed. Following this, a mixed-integer linear programming model designed for multi-item, multi-phase anticipatory aid allocation is introduced, focusing on maximizing humanitarian impact while respecting practical constraints like budget limitations and equity requirements. The computational framework and implementation strategy using Python, Pyomo, and Streamlit are then discussed, emphasizing reproducibility and ease of use. Lastly, methods for scalability testing and comparative simulations against simpler allocation strategies are presented, allowing for comprehensive evaluation of the optimization approach.

While the methods presented here are illustrated through the Indian context, they are designed to be adaptable and transferable to other regions facing similar anticipatory action needs.

5.2.1 Study region

Purba Medinipur district in West Bengal, India

India was chosen as a case study region due to the escalating vulnerability of its agricultural sector to climate-induced challenges, such as more frequent heat waves, droughts, extreme precipitation events, and intense cyclonic activities [465]. Agriculture, which accounts for 15% of India's GDP and employs 40% of its population—rising to 70% among rural households—is particularly exposed to these risks. Between 2015 and 2021, India lost nearly 34 million hectares of crops to excess rainfall and an additional 35 million hectares to drought, with economic losses across sectors, including agriculture, amounting to $159 billion in 2021 alone [467]. Farmers have reported facing challenges such as erratic rainfall, reduced rainy days, and shifting seasonal cycles, which disrupt rice yields and compound existing socio-economic and ecological stressors [468].

India's agricultural sector, particularly its vulnerable potato belt, has increasingly recognized the need for anticipatory action to safeguard farmers from climate shocks [469]. The state of West Bengal for instance, including Purba Medinipur district (the focus of this study), is highly vulnerable to natural disasters, such as floods, cyclones, hailstorms, droughts, and erosion, which pose significant risks to the livelihoods of farming communities [470]. For instance, in that state in 2023, Cyclone Michaung devastated rice and potato crops due to inadequate early warning systems, leaving already heavily indebted farmers reeling from poor access to seeds, limited export opportunities, and inadequate access to cold storage [469]. Farmers in regions like Jalpaiguri struggled to store their harvests due to inequitable access to cold storage bonds, forcing them to sell at losses or take on additional loans in an area where interest rates can range from 21-27% [471].

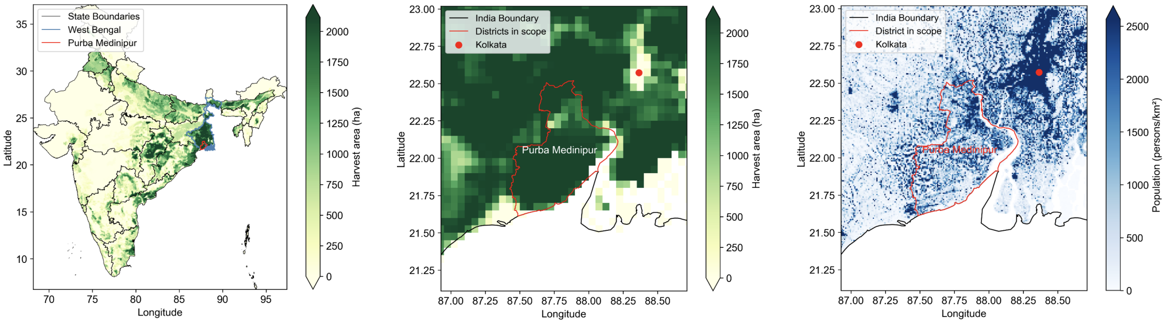


Figure 5.1: Overview of harvest area and population in the study area. The left panel shows harvested rice area (hectares) across India. The middle and right panels respectively show harvested rice area and population density in the Purba Medinipur district in West Bengal. The red dot denotes Kolkata, highlighting its proximity to key agricultural zones. Raw data obtained from MapSPAM [472] and WorldPop [473].

Purba Medinipur is situated along the Bay of Bengal in the state of West Bengal, India, and is recognized as a significant rice-producing district [474]. This district, spanning 4,736km^2 between 21°36'N–22°57'N latitude and 86°33'E–88°12'E longitude, features a coastal landscape that is representative of the broader challenges faced by farming communities in India's coastal regions. Approximately 35% of India's population resides in coastal zones within 100km of the seashore [475], highlighting the broader national importance of developing effective agricultural and climate-resilience strategies in such areas.

The climate of Purba Medinipur is tropical, with pronounced wet and dry seasons marked by high humidity and variable annual temperatures. The monsoon season (June–October) delivers most of the district's rainfall, supporting rice cultivation. However, this coastal environment is also exposed to tropical cyclones, including Aila (2009), Phailin (2013), Titli (2018), Fani (2019), Amphan (2020), and Yaas (2021). These recurrent cyclonic events significantly affect agricultural activities, infrastructure, and livelihoods.

Agriculture, aquaculture, fishing, and tourism are among the district's main economic activities [477]. Rice (paddy) cultivation is particularly prominent during the Aman season, which runs from July to November; seedbeds are established in July, followed by transplanting in August and harvesting in November. Several factors favor paddy cultivation in Purba Medinipur, including reliable monsoon rains, accessible irrigation facilities, suitable soil types, and a long-standing preference for rice among local farmers [478]. Soil composition varies across the district, with sandy soils covering about 46,270 ha and clay soils covering about 31,169 ha. Certain areas such as Deshpran and Egra-II

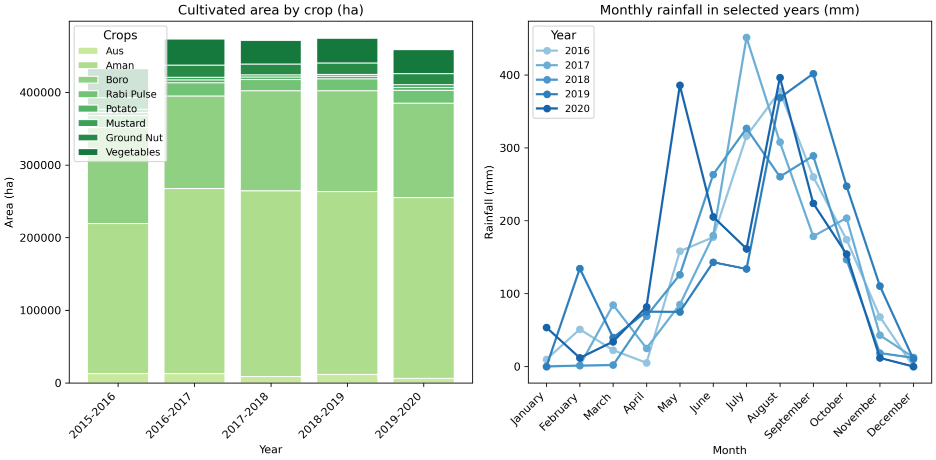


Figure 5.2: Cultivated area (in hectares) for different crops across years (left) and monthly rainfall (in mm) for selected years (right) in Purba Medinipur. The left plot visualizes the distribution of cultivated land among crops, showing year-on-year trends, while the right plot highlights rainfall variability over months and across five years (2016–2020) [476].

are predominantly clayey, whereas Contai and Ramnagar have mixtures of clay loam and sandy loam [479].

From a socio-economic perspective, the district had a population of 5.1 million as of the 2011 Census. Of its total land area, 430,140ha are cultivated [480]. Despite strong agricultural potential, farming communities here are regularly exposed to climate variability, marine hazards, and natural calamities. These stressors disrupt not only farming practices but also broader aspects of socioeconomic stability [481]. In response, farmers frequently shift or diversify their livelihoods seasonally, although the extent of success varies, influencing food security, income stability, and social well-being.

Challenges in traditional relief and emerging anticipatory action strategies in India

West Bengal has faced mixed success with traditional aid delivery following climate-induced disasters, highlighting systemic challenges and the potential for anticipatory action to improve humanitarian outcomes. Cyclone Yaas in May 2021 affected over 10 million people across the state, damaged more than 300,000 homes, and inundated 221,000 hectares of cropland and 71,560 hectares of horticultural land. The total economic losses were estimated at Rs. 20,000 crore (approximately $2.7 billion). Farmers in coastal regions like Purba Medinipur and South 24 Parganas suffered substantial livelihood losses due to saltwater intrusion into farmlands, fishponds, and aquaculture systems, with recovery times for soil and water systems stretching to three years. Relief efforts, such as the *Duare Tran* scheme, provided cash assistance ranging from Rs. 1,000 to Rs. 20,000 ($13 to $270) depending on the severity of losses. However, numerous challenges undermined their effectiveness, including delays in aid disbursement (some farmers waited more than two years to claim crop insurance compensation after heavy flooding), inadequate general compensation, and bureaucratic barriers such as the need for repeated visits to government offices to receive funding [466].

Compounding these challenges, rigid eligibility criteria excluded farmers in "less-affected" districts like Malda, where Cyclone Yaas destroyed jute and paddy crops but failed to meet thresholds for prioritization under the *Duare Tran* relief programme. While priority was given to districts like South 24 Parganas, which suffered catastrophic damage from Cyclone Yaas, Malda's farmers, who lost crops to both flooding and hailstorms, were deemed less critical. As one farmer lamented, "The harvest was hardly due by a month. But the heavy rain during Yaas submerged my entire paddy field... As the produce isn't of good quality, I can't take it to government mandis, so I will sell it to middlemen who usually give Rs. 1,000 for a quintal," considerably below the Rs. 1,800 offered in government procurement in the previous year. With no immediate relief forthcoming, many in Malda resorted to annual state schemes like *Krishak Bandhu*, which provide Rs. 4,000–Rs. 10,000 (54˘135) per year [466].

In Bihar, similar shortcomings have plagued drought mitigation initiatives like the diesel subsidy scheme introduced in 2008. Intended to support farmers in irrigating their paddy fields during droughts, the program faced issues of low awareness, high transaction costs, and delays in fund disbursement. A 2014 survey of 240 farmers in Nalanda District revealed that many marginal farmers were unaware of the scheme or its eligibility criteria, with 85% citing difficulties in providing required documentation, such as land revenue receipts. Even when farmers did apply, the process involved multiple bureaucratic steps, ranging from village-level approvals to district-level fund transfers, resulting in delays of weeks or months. Farmers often received subsidies after the cropping season, rendering the financial support ineffective in preventing crop losses. Furthermore, the program disproportionately benefited wealthier farmers with access to diesel pumps, while smaller landholders reliant on hired irrigation were largely excluded. These inefficiencies highlight the broader challenge of designing equitable and timely aid systems in states with limited administrative capacity [482].

Amidst these challenges, India is beginning to experiment with anticipatory action as a complementary strategy to traditional disaster relief. Pilot initiatives, such as Caritas India's 2024 landslide response in Wayanad, demonstrate potential: preemptive cash transfers of Rs. 9,500 to 525 households enabled families to secure food, repair homes, and avoid high-interest loans [483]. While still nascent, these efforts align with broader regional momentum: eight South Asian countries, including India, are now developing national anticipatory action frameworks to institutionalize protocols for early warnings and pre-agreed financing [484]. Neighboring Bangladesh provides a proven model, where pre-disaster cash transfers during floods reduced hunger by 36% and prevented distress sales of livestock [485]. For India, scaling such approaches could address challenges like those faced in West Bengal.

5.2.2 Simulation of farmer vulnerability data

In this section, we present a method to emulate the outputs of an early warning system for climate shocks. This approach aligns with recent literature reviews on humanitarian logistics, which emphasize the value of simulation techniques for stress-testing scenarios, analyzing hypothetical situations, and evaluating the impact of parameterized decisions across disaster life cycle stages [486].

We assume that, within a specified bounding box, certain "centers" of flooding or other hazards emerge. Around these centers, vulnerability decays gradually (or steeply) as one moves away in space. This modeling approach allows us to simulate various flood shapes—circular, elliptical, linear, or even wavy swaths—thus providing flexible stress-test scenarios for downstream optimization. Ultimately, these vulnerability scores serve as simulated inputs to an optimization model that allocates anticipatory aid actions optimally.

Suppose we have n_{centers} random hazard locations

$$\left\{(c_{x_i}, c_{y_i})\right\}_{i=1}^{n_{\text{centers}}},$$

drawn uniformly from the bounding box

$$[x_{\min},\, x_{\max}] \times [y_{\min},\, y_{\max}].$$

For each pixel $\mathbf{p} = (p_x, p_y)$, we define a distance to each center (c_{x_i}, c_{y_i}) as follows. If *circular* distances are desired, we use:

$$d_i(\mathbf{p}) \;=\; \sqrt{(p_x - c_{x_i})^2 + (p_y - c_{y_i})^2}\,. \tag{5.1}$$

If *elliptical* shapes are desired, we stretch the y-axis by a factor $\alpha > 0$ (here called `ellipse_ratio`), giving:

$$d_i(\mathbf{p}) \;=\; \sqrt{(p_x - c_{x_i})^2 + \bigl(\alpha\,(p_y - c_{y_i})\bigr)^2}\,. \tag{5.2}$$

Finally, for *vertical line* swaths, the distance depends only on the horizontal gap:

$$d_i(\mathbf{p}) \;=\; \bigl|\,p_x - c_{x_i}\bigr|. \tag{5.3}$$

We then take the minimum distance over all centers:

$$d(\mathbf{p}) \;=\; \min_{1 \le i \le n_{\text{centers}}} d_i(\mathbf{p}).$$

Given a user-specified *spread* parameter $\sigma > 0$, the *vulnerability* at pixel $\mathbf{p}$ is modeled via exponential decay:

$$\text{Vulnerability}(\mathbf{p}) \;=\; \max\Bigl\{0,\; 100\,\exp\bigl(-\tfrac{d(\mathbf{p})}{\sigma}\bigr)\Bigr\}. \tag{5.4}$$

Hence, pixels very close to a flood center exhibit vulnerability near 100, while more distant pixels drop rapidly (or gradually) toward zero, depending on σ.

In the implementation, σ is defined on normalized spatial coordinates rather than physical distance units such as kilometers. This keeps scenario design intuitive across different map extents. In the reported scenario experiments, spread values are drawn from approximately 0.03 to 0.14, which spans relatively compact hotspots as well as broader hazard swaths. Accordingly, σ should be interpreted here as a scenario-shape parameter unless a hazard-specific calibration is available.

Multiple shapes: To allow more complex scenarios with overlapping hazards of different shapes, we may define a set of hazards $\mathcal{H}$, each with potentially distinct parameters (shape, number of centers, or spread). Let Vulnerability$^h(\mathbf{p})$ be computed via (5.4) for each $h \in \mathcal{H}$. We then combine them using:

$$\text{Vulnerability}(\mathbf{p}) \;=\; \begin{cases} \max\limits_{h \in \mathcal{H}} \text{Vulnerability}^h(\mathbf{p}), & (\text{“max” rule}), \\ \min\Bigl\{100,\; \sum\limits_{h \in \mathcal{H}} \text{Vulnerability}^h(\mathbf{p})\Bigr\}, & (\text{“sum” rule}). \end{cases} \tag{5.5}$$

If multiple hazards overlap, the resulting pixel vulnerability is determined by the *strongest* local hazard. This approach creates a patchwork of hazards where a pixel is dominated by whichever shock is most severe (“max" rule). In contrast, overlapping hazards *add* together, meaning that a pixel affected by two hazards of vulnerability 60 each might reach the (capped) value of 100 (“sum” rule). This method highlights multi-hazard “hotspots,” where floods, storms, or other events coincide, leading to highly vulnerable regions. In practice, the choice of combining rule depends on how the hazards reinforce or compete. Such a framework allows one to place, for instance, multiple circular floods and a wavy swath on the same map, merging their pixelwise vulnerability via either the maximum or the sum rule.

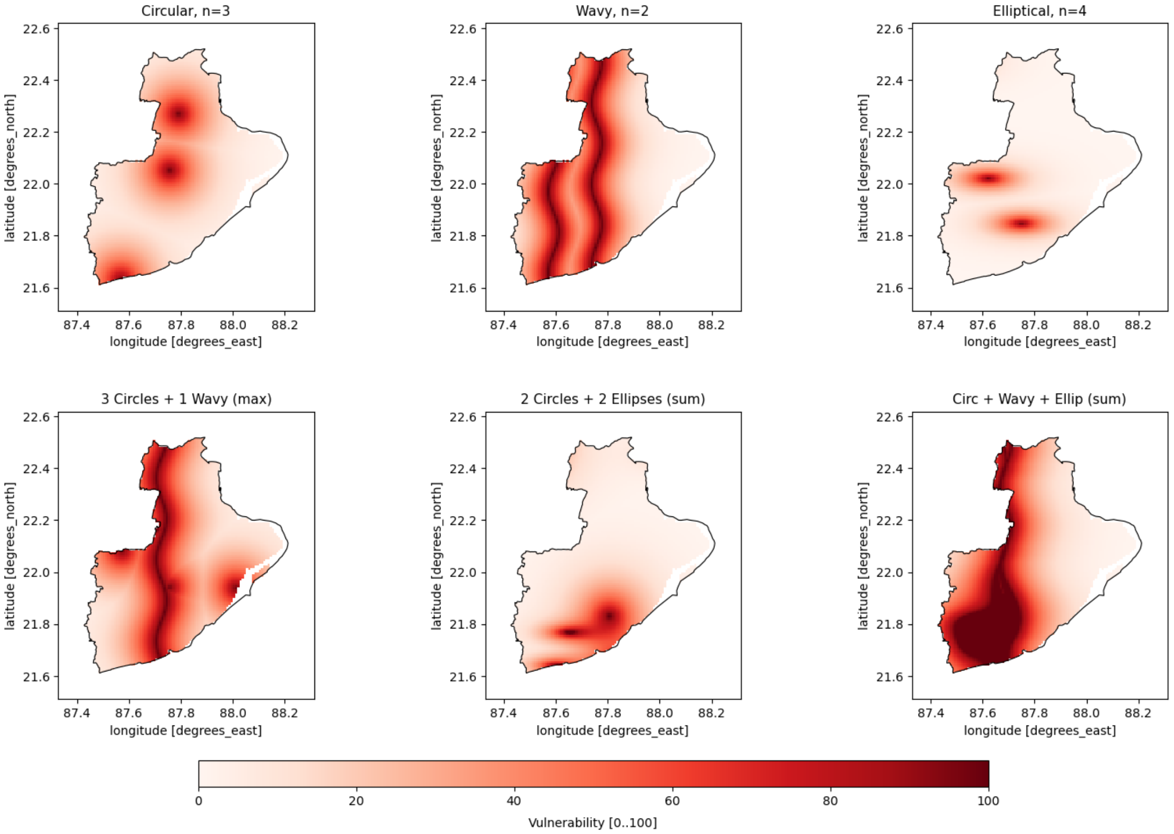


Figure 5.3: Examples of simulated vulnerability maps illustrating different hazard configurations in Purba Medinipur. The top row shows single-shape hazards: (1) circular with 3 centers, (2) wavy with 2 centers, and (3) elliptical with 4 centers. The bottom row demonstrates multi-shape configurations: (4) three circular hazards combined with one wavy hazard using the max rule, (5) two circular and two elliptical hazards combined using the sum rule, and (6) a combination of circular, wavy, and elliptical hazards using the sum rule. The shared colorbar below represents vulnerability scores ranging from 0 to 100.

The resulting vulnerability maps (see figure 5.3) effectively simulate a hypothetical early warning system’s output, identifying “hotspot” regions around one or multiple hazard centers. These pixel-level vulnerability scores are then fed into an optimization model. This approach makes it straightforward to create a variety of plausible and distinct vulnerability scenarios for that optimization procedure.

5.2.3 Mixed-integer program for multi-item, multi-phase allocation

We formulate a mixed-integer linear program (MILP) for allocating multiple aid items across micro-regions and planning phases. The model is designed to prioritize locations that combine larger populations with higher vulnerability, while respecting budget limits, equity requirements, item capacities, and timing restrictions. It also distinguishes between objective-side welfare weights and reporting-side coverage conversions, allowing the allocation problem and the reporting metrics to be defined transparently.

Notation. We first define the main sets, parameters, and decision variables, and then present the objective function and constraints.

Sets and Indices

- $I = \{1, 2, \ldots, N\}$: Set of micro-regions (or "pixels"), indexed by i.
- G: Set of groups used in the equity constraints, indexed by g and h. In the empirical application, these groups are analyst-defined spatial partitions of the study area, and each pixel belongs to exactly one group.
- $T = \{1, 2, \ldots, M\}$: Set of discrete time steps (phases), indexed by t.
- K_{all}: Set of all aid types, indexed by k; these may include *cash*, *seedlings*, *protective kits*, and other items.
- $K_{\text{int}} \subseteq K_{\text{all}}$: Subset of aid items allocated in integer units.
- $K_{\text{real}} = K_{\text{all}} \setminus K_{\text{int}}$: Subset of aid items allocated fractionally.
- $K_{\text{cap}} \subseteq K_{\text{all}}$: Subset of aid items with finite total availability.
- $K_{\text{early}} \subseteq K_{\text{all}}$: Optional subset of aid items that may only be allocated in the first phase.
- $H = \{i \in I : v_i \geq \tau\}$: Set of high-vulnerability pixels, defined by vulnerability threshold τ.

In practice, the groups in G are not meant to represent a fixed universal category. Rather, they are a flexible way to encode the kinds of subpopulations or areas that a policymaker may wish to monitor for fairness when distributing aid. Depending on the application, a group could correspond to an administrative zone, a rural-versus-urban partition, a livelihood category, or another operationally relevant classification. In our illustrative case study, groups are constructed as simple spatial partitions of the district so that the equity mechanism can be demonstrated transparently. The purpose is therefore not to claim that these exact group boundaries are fixed categories, but to show how the model can incorporate fairness checks across meaningful subregions when such group definitions are available in practice.

Parameters

- p_i: Population of pixel i. This is the number of people living in micro-region i.
- $v_i \in [0,1]$: Vulnerability score of pixel i. Larger values indicate greater modeled exposure and need.
- B_t: Budget available at time step t. This determines the maximum expenditure in phase t under the at-most budget rule, or the exact expenditure under the exact-budget rule.
- c_k: Cost per unit of aid item k. This converts physical allocations into spending and enters the budget, equity, and high-vulnerability constraints.
- b_k: Benefit weight for aid item k in the objective function. This parameter captures the relative contribution of one additional unit of item k to the objective.
- q_k: Coverage-per-unit conversion for aid item k. This parameter maps one unit of item k into people-equivalent coverage for reporting and for the coverage cap.
- minAlloc_k, maxAlloc_k: Baseline lower and upper bounds on item k per pixel and per phase.
- ω_t: Phase weight in the objective function. Larger values of ω_t place greater value on earlier action. A convenient specification is

  $$\omega_t = 1 + \lambda \frac{M-t}{M-1}, \qquad t = 1, \ldots, M, \quad M > 1,$$

  where $\lambda \geq 0$ governs the strength of the preference for earlier action.

- d_g: Group denominator used in the equity metric. Depending on the equity mode,

$$d_g = \begin{cases} 1, & \text{for absolute-spend equity,} \\ \sum_{i \in g} p_i, & \text{for per-capita-spend equity,} \\ \sum_{i \in g} p_i v_i, & \text{for per-need-spend equity.} \end{cases}$$

- ε: Equity-tolerance parameter. This bounds pairwise differences in the normalized group allocation metric.
- $\overline{U}_k$: Global capacity of aid item k for each $k \in K_{\text{cap}}$. This limits the total quantity of any stock-constrained item that can be allocated across all pixels and phases.
- τ: High-vulnerability threshold. Pixels with $v_i \geq \tau$ belong to the protected set H.
- $\gamma \in [0, 1]$: Protected high-vulnerability budget share. This is the minimum share of each phase budget reserved for the high-vulnerability set H.
- $\bar{c} \geq 0$: Coverage-cap ratio. This controls the maximum amount of people-equivalent coverage that can be attributed to any pixel.

Because $x_{i,t,k}$ may represent different physical units across aid types, the model uses separate parameters to keep aggregation interpretable. The cost parameter c_k converts allocated units into spending, the benefit weight b_k is a dimensionless priority weight used in the objective, and q_k converts allocated units into people-equivalent coverage for reporting. The objective value is therefore best interpreted as a relative welfare score, not as a direct headcount or monetary outcome.

Table 5.1 reports the aid-type assumptions used in the case-study experiments and as defaults in the dashboard. In the comparison experiments, we use a mild timing preference of $\lambda = 0.1$, which yields phase weights of 1.10, 1.05, and 1.00 over the three planning phases. When $\lambda = 0$, the model is time-neutral across phases.

Decision variables

Allocation variables. For each pixel $i \in I$, time $t \in T$, and item type $k \in K_{\text{all}}$, define

$$x_{i,t,k} = \begin{cases} x^R_{i,t,k} \in \mathbb{R}_{\geq 0}, & \text{if } k \in K_{\text{real}}, \\ x^I_{i,t,k} \in \mathbb{Z}_{\geq 0}, & \text{if } k \in K_{\text{int}}. \end{cases} \tag{5.6}$$

The variable $x_{i,t,k}$ is the quantity of item k allocated to pixel i in time step t.

Table 5.1: Aid-type assumptions used in the case-study experiments and dashboard defaults. Costs are expressed in normalized budget units.

Aid type	c_k	b_k	q_k	maxAlloc$_k$	$\overline{U}_k$
cash	1.0	1.00	1.00	2,000	–
seedlings	0.5	0.90	0.60	500	5,000
protective	2.0	0.80	0.50	300	2,000

Group-level variables. To express the equity constraint compactly, define

$$y_{g,t} \;=\; \sum_{k \in K_{\text{all}}} \sum_{i \in g} c_k \, x_{i,t,k}, \quad \forall g \in G,\ \forall t \in T. \tag{5.7}$$

The variable $y_{g,t}$ represents the cost-weighted amount allocated to group g in period t.

Objective function

We seek to maximize

$$\max \sum_{t \in T} \sum_{k \in K_{\text{all}}} \sum_{i \in I} \big(\omega_t \, b_k \, v_i \, p_i\big) \, x_{i,t,k}. \tag{5.8}$$

The objective assigns greater value to allocations directed toward populous and highly vulnerable locations. The phase weight ω_t allows earlier periods to receive greater emphasis when anticipatory action is preferred. The coverage conversion q_k does not enter the objective and is used only for coverage accounting and the coverage cap.

Constraints

(1) Budget constraints per time step.

$$\sum_{k \in K_{\text{all}}} \sum_{i \in I} c_k \, x_{i,t,k} \;\le\; B_t, \quad \forall t \in T. \tag{5.9}$$

This is the at-most budget formulation. Under exact spending, (5.9) is replaced by equality.

(2) Minimum and maximum item constraints.

$$\underline{x}_{k,t} \;\le\; x_{i,t,k} \;\le\; \overline{x}_{k,t}, \quad \forall i,\ \forall t,\ \forall k, \tag{5.10}$$

where the effective time-dependent bounds are

$$(\underline{x}_{k,t}, \overline{x}_{k,t}) \;=\; \begin{cases} (0,0), & \text{if } k \in K_{\text{early}} \text{ and } t > 1, \\ (\text{minAlloc}_k, \text{maxAlloc}_k), & \text{otherwise.} \end{cases} \tag{5.11}$$

(3) Coverage-cap constraint.

$$\sum_{k \in K_{\text{all}}} \sum_{t \in T} q_k \, x_{i,t,k} \;\leq\; \bar{c}\, p_i, \quad \forall i \in I. \tag{5.12}$$

This cap limits reported people-equivalent coverage in each pixel.

(4) Group sum and equity constraints.

$$y_{g,t} = \sum_{k \in K_{\text{all}}} \sum_{i \in g} c_k \, x_{i,t,k}, \qquad \forall g,\ t, \tag{5.13}$$

$$-\varepsilon \;\leq\; \frac{y_{g,t}}{d_g} - \frac{y_{h,t}}{d_h} \;\leq\; \varepsilon, \qquad \forall g, h,\ t. \tag{5.14}$$

This formulation enforces equity on normalized, cost-weighted group allocations.

(5) Availability constraints. For stock-constrained items, total allocation cannot exceed available supply. In the empirical application, this applies to items such as seedlings and protective kits:

$$\sum_{i \in I} \sum_{t \in T} x_{i,t,\text{seedlings}} \;\leq\; \overline{U}_{\text{seedlings}}, \tag{5.15}$$

$$\sum_{i \in I} \sum_{t \in T} x_{i,t,\text{protective}} \;\leq\; \overline{U}_{\text{protective}}. \tag{5.16}$$

More generally, the same structure applies to any $k \in K_{\text{cap}}$.

(6) High-vulnerability budget-floor constraint.

$$\sum_{i \in H} \sum_{k \in K_{\text{all}}} c_k \, x_{i,t,k} \;\geq\; \gamma B_t, \quad \forall t \in T. \tag{5.17}$$

This optional rule reserves a minimum share γ of each phase budget for the high-vulnerability set H.

Full mixed-integer program

Combining all elements, the at-most-budget version of the model is

$$\begin{aligned}
\text{Maximize} \quad & \sum_{t\in T}\sum_{k\in K_{\text{all}}}\sum_{i\in I}\big(\omega_t\, b_k\, v_i\, p_i\big)\, x_{i,t,k}, && && (5.18)\\
\text{subject to} \quad & \sum_{k\in K_{\text{all}}}\sum_{i\in I} c_k\, x_{i,t,k} \;\le\; B_t, && \forall t\in T, && (5.19)\\
& \underline{x}_{k,t} \;\le\; x_{i,t,k} \;\le\; \overline{x}_{k,t}, && \forall i\in I,\ \forall t\in T,\ \forall k\in K_{\text{all}}, && (5.20)\\
& \sum_{k\in K_{\text{all}}}\sum_{t\in T} q_k\, x_{i,t,k} \;\le\; \bar{c}\, p_i, && \forall i\in I, && (5.21)\\
& y_{g,t} = \sum_{k\in K_{\text{all}}}\sum_{i\in g} c_k\, x_{i,t,k}, && \forall g\in G,\ \forall t\in T, && (5.22)\\
& -\,\varepsilon \;\le\; \frac{y_{g,t}}{d_g} - \frac{y_{h,t}}{d_h} \;\le\; \varepsilon, && \forall g,h\in G,\ \forall t\in T, && (5.23)\\
& \sum_{i\in I}\sum_{t\in T} x_{i,t,k} \;\le\; \overline{U}_k, && \forall k\in K_{\text{cap}}, && (5.24)\\
& \sum_{i\in H}\sum_{k\in K_{\text{all}}} c_k\, x_{i,t,k} \;\ge\; \gamma B_t, && \forall t\in T \quad \text{if } \gamma>0, && (5.25)\\
& x_{i,t,k}\in\mathbb{Z}_{\ge 0}, && \forall i\in I,\ \forall t\in T,\ \forall k\in K_{\text{int}}, && (5.26)\\
& x_{i,t,k}\in\mathbb{R}_{\ge 0}, && \forall i\in I,\ \forall t\in T,\ \forall k\in K_{\text{real}}. && (5.27)
\end{aligned}$$

where the effective time-dependent bounds are given by

$$(\underline{x}_{k,t}, \overline{x}_{k,t}) \;=\; \begin{cases} (0,0), & \text{if } k\in K_{\text{early}} \text{ and } t>1,\\ (\text{minAlloc}_k, \text{maxAlloc}_k), & \text{otherwise.} \end{cases} \tag{5.28}$$

Under exact spending, constraint (5.19) is replaced by

$$\sum_{k\in K_{\text{all}}}\sum_{i\in I} c_k\, x_{i,t,k} = B_t, \qquad \forall t\in T. \tag{5.29}$$

The reporting-side coverage delivered to pixel i is computed ex post as

$$\min\left\{\sum_{k\in K_{\text{all}}}\sum_{t\in T} q_k\, x_{i,t,k},\ \bar{c}\, p_i\right\}. \tag{5.30}$$

This formulation accommodates continuous and integer allocations, explicit budget policy, normalized equity across groups, stock limits, protected spending for high-vulnerability locations, and time-sensitive anticipatory action.

5.2.4 Computation

Many different tools can be used to formulate mixed-integer linear programs, including LINGO, AMPL, and Excel Solver [487]. In this study, Python was selected to ensure a fully reproducible optimization pipeline that other researchers can easily adopt. Specifically, the mixed-integer programming model was implemented with the Python-based Pyomo framework, an open-source library that provides a flexible environment for defining and executing mathematical optimization problems. Pyomo's modular approach allows sets, variables, and constraints to be expressed cleanly, facilitating both readability and extensibility [258].

The current implementation uses a single shared codebase for the manuscript experiments, the reproducibility notebooks, and the web application. In the local environment used for this study, the core packages were Pyomo 6.8.2, HiGHS 1.9.0 via `highspy`, and Streamlit 1.41.1. Open-source solver backends exposed by the code include HiGHS [488], CBC [489], and GLPK [490], although the reported runtime comparison focuses on HiGHS and CBC. In the current codebase, no explicit solver time limit, thread override, or custom MIP gap is imposed in the reported experiments; runs follow the solvers' default termination logic, and infeasible instances are reported as such. All computations were conducted on a Macbook Pro M4 Max (16 CPU cores, 40 GPU cores) with 128GB of unified LPDDR5 RAM.

5.2.5 Web application development and deployment

To enhance user interaction with the optimization model, an interactive web-based application was developed using the Streamlit framework [491]. Streamlit is a Python-based platform designed for rapidly building data-driven applications and has been adopted in various fields, including smart farming for crop recommendation systems [492], molecular structure visualization [493], remote monitoring of building temperature through operations dashboards [494], and medical diagnostics such as early detection of lung cancer and pneumonia [495]. Its ability to integrate user-defined parameters and display real-time computational outputs makes it well-suited for applications focused on anticipatory action, where testing different scenarios can aid decision-makers.

The web application described in this work exposes the main modeling choices through a plain-language interface. Users can configure hazard shapes, time-phased budgets, budget mode, grouping strategy, equity mode, equity tolerance, high-vulnerability thresholds and budget floors, solver selection, and the item-level assumptions for costs, benefit weights, coverage conversions, capacities, and timing restrictions. These inputs are mapped directly to the Pyomo model, and the resulting allocations $\{x_{i,t,k}\}$ are summarized through tables, charts, and spatial maps. The application also provides side-by-side comparison with a blanket baseline, a budget-mode sensitivity view comparing at-most and exact spending, perturbation-based scenario stress tests, and runtime benchmarks. This immediate feedback loop allows decision-makers to explore how different policy assumptions affect both allocations and reported outcomes.

To support broad accessibility, the application is packaged for browser-based deployment on Render [496] and similar cloud platforms [496, 497]. This removes the need for local setup when a hosted version is provided, while still allowing the full stack to be run locally from the open-source repository. The repository includes the Pyomo model, the Streamlit interface, and the reproducibility notebooks used for the comparison and runtime analyses. This setup lowers technical barriers for non-technical users, including policymakers and field practitioners, who can interact with the model without requiring specialized expertise in programming or optimization [498].

The entire codebase, including the Pyomo model and Streamlit application, is open-source and can be accessed via GitHub. This open distribution fosters transparency, reproducibility, and collaboration among researchers and practitioners seeking to adapt or extend the methodology for their specific contexts.

5.2.6 Scalability tests with increasing spatial granularity

Even if a mathematical model for humanitarian aid allocation is conceptually simple, its practical implementation can become computationally demanding once high spatio-temporal granularity is introduced across many regions and multiple planning horizons. Decision-makers in humanitarian settings often operate with limited computing resources [499] and may need to evaluate multiple scenarios within short time windows. Similar concerns arise in large-scale energy system models, where relatively small changes in formulation or solver settings can substantially affect runtime [500].

To examine how the proposed allocation model scales with problem size, we vary $|I|$, the number of pixel-level units, from 1,000 to 30,000. These experiments represent increasingly fine spatial resolutions and larger intervention footprints. For each problem size, we record build time, solve time, and total runtime under alternative solver backends.

The scalability experiments use synthetic datasets with the same structure as the operational model inputs. Each synthetic pixel is assigned a population, a vulnerability score, a group label, and geographic coordinates. Unless otherwise stated, the experiments use a three-phase budget schedule (1,000,000, 500,000, 300,000), exact spending in each phase, integer restrictions for discrete aid items, and the same general class of aid items as in Section 5.2.3. We compare two open-source solvers, CBC and HiGHS, under three representative policy configurations:

1. **Configuration A (no equity, no high-vulnerability floor).** This setting omits the equity rule and imposes no minimum budget share for high-vulnerability pixels. The model is therefore governed only by the phase budgets, item bounds, integrality conditions, and any relevant item capacities.

2. **Configuration B (per-capita equity, no high-vulnerability floor).** This setting introduces group equity while keeping $\gamma = 0$. Equity is enforced on per-capita, cost-weighted group spending with tolerance $\varepsilon = 5.0$. Let

   $$d_g = \sum_{i \in g} p_i$$

   denote the population of group g. Then, for each time step t,

   $$-\varepsilon \;\le\; \frac{y_{g,t}}{d_g} - \frac{y_{h,t}}{d_h} \;\le\; \varepsilon, \quad \forall g, h,\ t, \tag{5.31}$$

   where $y_{g,t} = \sum_{i \in g} \sum_{k \in K_{\text{all}}} c_k x_{i,t,k}$ is the cost-weighted allocation to group g at time t.

3. **Configuration C (tighter per-capita equity plus a high-vulnerability floor).** This setting strengthens the equity rule to $\varepsilon = 1.0$ and introduces a protected spending floor for high-vulnerability pixels. Let

   $$H = \{\, i \in I : v_i \ge \tau \,\}$$

   denote the set of pixels whose vulnerability exceeds threshold τ. In the reported experiments, $\tau = 0.80$. The model then imposes

   $$\sum_{i \in H} \sum_{k \in K_{\text{all}}} c_k \, x_{i,t,k} \;\ge\; \gamma B_t, \quad \forall t \in T, \tag{5.32}$$

   with $\gamma = 0.20$. Thus, at least 20% of each phase budget must be directed to the high-vulnerability set.

Each configuration is evaluated for varying $|I|$ across both solvers. Together, these experiments show how computational burden evolves as spatial resolution increases and as the policy design becomes more restrictive.

5.2.7 Simulation-based comparison with a blanket approach

To evaluate the practical value of the optimization model, we compare it with a simple blanket allocation rule that distributes support in proportion to population alone. The comparison addresses a straightforward policy question: how much changes when aid is targeted using vulnerability, item-specific costs, group equity, and phase-by-phase planning, rather than being spread uniformly on a per-capita basis? By repeating this comparison across multiple synthetic hazard scenarios, we quantify the gains from structured optimization in terms of coverage, targeting, and spending patterns. The procedure is summarized in Algorithm 1.

Let I denote the set of pixels and $T = \{1, \ldots, T\}$ the set of decision phases. Each pixel $i \in I$ has population p_i and normalized vulnerability $v_i \in [0, 1]$. For presentation in maps and figures, vulnerability may be shown on a 0–100 scale, but all calculations use the normalized representation. Each aid item $k \in K$ has a unit cost c_k, a benefit weight b_k, and a coverage conversion q_k. These definitions are consistent with the model in Section 5.2.3.

1. Generating hazard scenarios. Building on the hazard simulation framework described in Section 5.2.2, we generate multiple disaster profiles by combining randomly selected hazard components from the circular, elliptical, vertical-line, and wavy-line families. For each scenario s, between one and three hazard components are drawn, each with its own shape, number of centers, and spread parameter. In the reported experiments, spread values are drawn from approximately 0.03 to 0.14 on normalized coordinates, and the resulting surfaces are combined using either a pixelwise maximum rule or a capped sum rule. Formally,

$$v_i^{(s)} = \begin{cases} \max_{\ell \in \mathcal{H}_s} \big(h_\ell^{(s)}(i) \big), & \text{if combine} = \text{max}, \\ \min\Big\{ 1, \; \sum_{\ell \in \mathcal{H}_s} h_\ell^{(s)}(i) \Big\}, & \text{if combine} = \text{sum}, \end{cases} \tag{5.33}$$

where $h_\ell^{(s)}(i) \in [0, 1]$ is the contribution of hazard component ℓ to pixel i in scenario s, and $\mathcal{H}_s$ is the set of hazard components active in that scenario. Population values

$\{p_i\}$ are derived from the clipped LandScan raster. Each scenario therefore represents a distinct spatial risk pattern defined on the same population grid.

2. Blanket allocation versus optimization.

(a) Blanket allocation. The blanket baseline is deliberately simple. In the reported experiments, it allocates the baseline cash item only and does so strictly in proportion to population, without using vulnerability information. Let k^0 denote the baseline item. Then, for each phase t,

$$\widehat{x}_{i,t,k^0} \;=\; \frac{B_t}{c_{k^0}} \cdot \frac{p_i}{\sum_{j \in I} p_j}, \tag{5.34}$$

so that the corresponding spending assigned to pixel i in phase t is

$$c_{k^0} \widehat{x}_{i,t,k^0} = B_t \frac{p_i}{\sum_{j \in I} p_j}.$$

This baseline preserves the intuition of uniform per-capita spending. If a per-cell bound becomes binding, the proportional allocation is clipped accordingly, but it remains untargeted with respect to vulnerability.

(b) Optimization-based approach. We next solve the multi-item, multi-phase MILP introduced in Section 5.2.3, which chooses allocations $\{x_{i,t,k}\}$ subject to budget rules, group equity, item capacities, integrality restrictions, and any high-vulnerability spending floor. In the comparison experiments, the model is solved under exact spending in each phase, per-capita group equity, a high-vulnerability threshold $\tau = 0.70$, a protected budget share $\gamma = 0.20$, and a mild preference for earlier action through the phase weights ω_t. The solution is denoted by $\{x^{\star}_{i,t,k}\}$.

3. Coverage and metrics. We compare the blanket baseline and the optimized allocation using three metrics.

(a) High-vulnerability coverage ratio C_{HV}. Let

$$H = \{\, i \in I \mid v_i \geq \tau \,\}$$

be the set of high-vulnerability pixels. For any allocation plan, the people-equivalent coverage delivered to pixel i is

$$c_i \;=\; \min\Big\{ \bar{c}\, p_i,\; \sum_{t \in T} \sum_{k \in K} q_k\, x_{i,t,k} \Big\}, \tag{5.35}$$

where $\bar{c}$ is the coverage-cap ratio. The corresponding high-vulnerability coverage rate is

$$C_{HV} \;=\; \frac{\sum_{i\in H} c_i}{\sum_{i\in H} p_i} \quad \in [0,1]. \tag{5.36}$$

This metric measures the share of the population living in the highest-risk pixels that receives effective, capped coverage.

(b) Targeting metric $\boldsymbol{\rho}$. Targeting is measured using total allocated spend. Define

$$s_i \;=\; \sum_{t\in T}\sum_{k\in K} c_k\, x_{i,t,k}, \tag{5.37}$$

the total spending assigned to pixel i. The targeting metric is then

$$\rho \;=\; \operatorname{corr}\big(\{s_i\},\, \{v_i\}\big), \tag{5.38}$$

where higher values indicate stronger alignment between spending and vulnerability.

(c) High-vulnerability spend share $\boldsymbol{\phi}$. The fraction of total spending directed to the high-vulnerability set is

$$\phi \;=\; \frac{\displaystyle\sum_{i\in H}\sum_{t\in T}\sum_{k\in K} c_k\, x_{i,t,k}}{\displaystyle\sum_{i\in I}\sum_{t\in T}\sum_{k\in K} c_k\, x_{i,t,k}}. \tag{5.39}$$

This metric complements C_{HV} by quantifying the share of the total budget directed to the highest-risk locations.

4. Multiple scenarios and summary plots. The above steps are repeated for S hazard scenarios spanning different combinations of shapes, severities, and overlap patterns. For each scenario, we compute (C_{HV}, ρ, ϕ) for both the blanket baseline and the optimized allocation. The resulting values are summarized in bar charts and tables to compare the two approaches across scenarios.

For a given vulnerability surface, the optimization itself is deterministic. Forecast uncertainty is therefore represented here through repeated synthetic hazard scenarios rather than through a stochastic or robust objective. In the application, this same logic is extended through perturbation-based stress tests that re-solve the model under modest changes to the vulnerability field.

Algorithm 1 Population-proportional blanket baseline versus MILP allocation

```
Require: vulnerability $v_i \in [0, 1]$, population $p_i$ for each pixel $i \in I$; budgets $B_t$ for
    $t = 1, \ldots, T$; item data $c_k$, $b_k$, and $q_k$ for $k \in K$; threshold $\tau$; coverage cap $\bar{c}$;
    additional policy settings for the MILP.

BLANKET BASELINE
Choose baseline item $k^0$
$PopTot \leftarrow \sum_{j \in I} p_j$
for $t = 1$ to $T$ do
    for $i \in I$ do
        $\widehat{x}_{i,t,k^0} \leftarrow \frac{B_t}{c_{k^0}} \frac{p_i}{PopTot}$
    end for
end for

MILP ALLOCATION
Solve the model in Section 5.2.3
Denote the optimal solution by $x^\star_{i,t,k}$

METRICS (FOR EITHER ALLOCATION $z \in \{\widehat{x}, x^\star\}$)
$H \leftarrow \{\, i \in I \mid v_i \geq \tau \,\}$
for $i \in I$ do
    $c_i(z) \leftarrow \min\Big(\bar{c} p_i, \sum_{t \in T} \sum_{k \in K} q_k \, z_{i,t,k}\Big)$
    $s_i(z) \leftarrow \sum_{t \in T} \sum_{k \in K} c_k \, z_{i,t,k}$
end for
$C_{HV}(z) = \frac{\sum_{i \in H} c_i(z)}{\sum_{i \in H} p_i}$
$\rho(z) = \mathrm{corr}\big(\{s_i(z)\}, \{v_i\}\big)$
$\phi(z) = \frac{\sum_{i \in H} s_i(z)}{\sum_{i \in I} s_i(z)}$
Compare $(C_{HV}, \rho, \phi)$ for blanket versus MILP across scenarios
```

5.3 Results and discussion

This section provides a structured overview of the interactive decision-support application presented in Figure 5.4. We first describe its core modules, highlighting how vulnerability assessments and optimization modeling combine to identify targeted anticipatory aid allocations. We then benchmark the computational scalability of the tool across varying problem sizes and configurations, and compare its performance against conventional blanket allocation strategies. Finally, we outline practical extensions and suggest future research directions to enhance anticipatory action planning and decision-making.

5.3.1 Overview of web application features and possible extensions

Vulnerability module

The *vulnerability module* provides a structured approach for generating spatially explicit vulnerability maps, which serve as inputs for the optimization model. This module synthesizes hypothetical hazard impact scenarios by simulating localized vulnerability hotspots based on spatial decay functions and configurable hazard distributions. Users can further adapt or refine this module to reflect region-specific risk profiles, incorporating different spatial structures, exposure assumptions, or vulnerability aggregation methods.

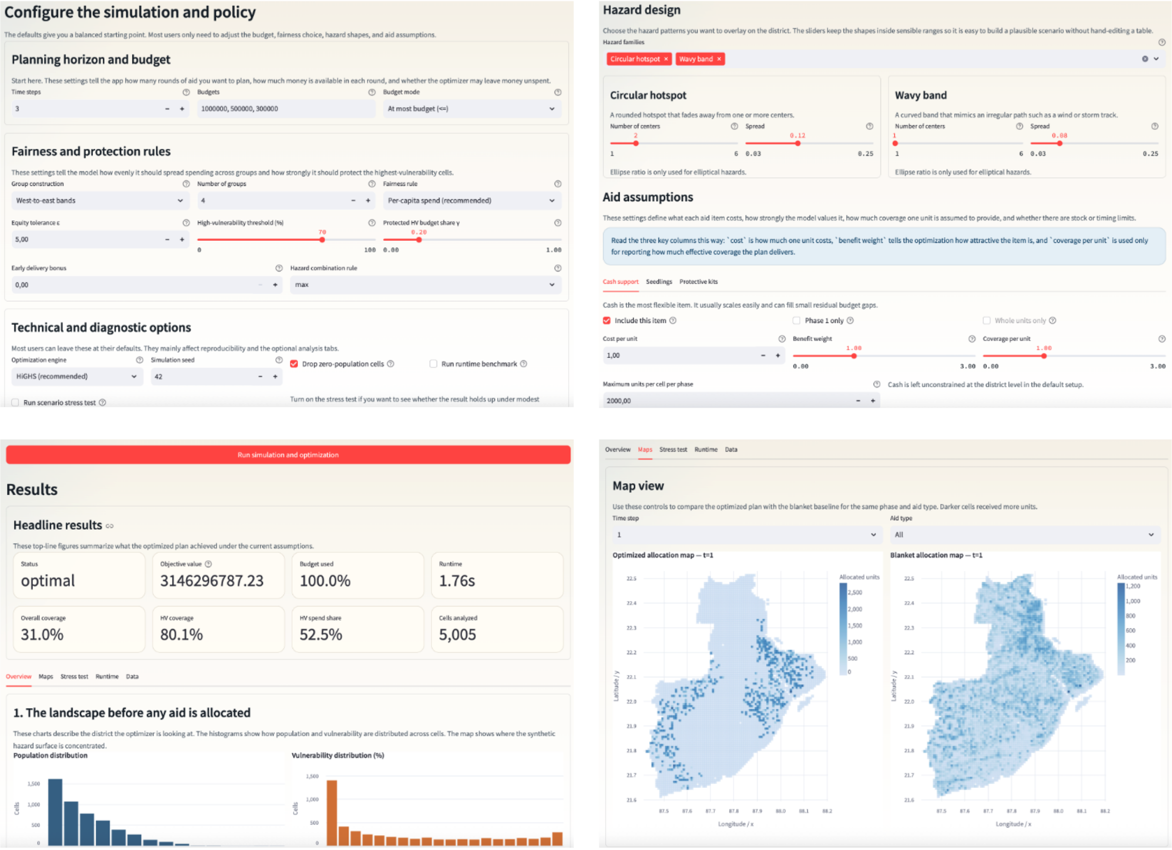
Alternatively, users can replace this module with externally derived vulnerability scores from an integrated early warning system, allowing for real-time data ingestion from meteorological models, satellite-based hazard detection, or government risk assessments. For India specifically, vulnerability scores might be derived from drought monitoring systems based on indices such as the Normalized Vegetation Supply Water Index (NVSWI) [501]; machine learning-based flood vulnerability maps [502], or remote-sensing based crop yield forecasting models which can anticipate potential crop yield decreases [250, 503]. Integrating nowcasts (i.e., predicting events with a lead time of three hours or less) or forecasts directly into the tool would allow policymakers to simulate optimal aid allocation based on predicted agricultural shortfalls rather than relying solely on historical or hypothetical vulnerability patterns. This flexibility ensures that the framework can be used both as a standalone exploratory tool for scenario analysis and as a decision-support system when coupled with operational forecasting pipelines.

The effectiveness of vulnerability modules that integrate existing early warning systems depends on the availability, accuracy, and geographic coverage of these systems. For example, in the case of flooding, a research report indicates that although nearly two-thirds of India's population faces exposure to extreme flood events, only about one-third currently benefits from flood early warning systems [504]. Constraints such as these may limit the extent to which optimized anticipatory action schemes can be designed.

Optimization module

The *optimization module* converts simulated vulnerability data into actionable aid allocation plans by solving the MILP problem under specified budget, equity, and capacity constraints. Its central objective is to route limited resources toward the most vulnerable populations, while respecting fairness criteria across regions, explicit spending rules, and aid-type-specific operational limits.

A distinguishing feature of this module is its spatially explicit allocation output, wherein each pixel or administrative unit receives an assigned level of support in each time period. This enables policymakers to visualize how resource distribution evolves over time and across the study area, facilitating comparisons of alternative scenarios in the Streamlit-based dashboard. By adjusting model parameters such as budget mode, group-level equity rules, high-vulnerability protection thresholds, aid-type assumptions, or prioritization weights, users can iteratively refine allocation schemes to align with evolving policy objectives or logistical considerations.



Configure the simulation and policy

The defaults give you a balanced starting point. Most users only need to adjust the budget, fairness choice, hazard shapes, and aid assumptions.

Planning horizon and budget

Start here. These settings tell the app how many rounds of aid you want to plan, how much money is available in each round, and whether the optimizer may leave money unspent.

Time steps: 3
Budgets: 1000000, 500000, 300000
Budget mode: At most budget (<=)

Fairness and protection rules

These settings tell the model how evenly it should spread spending across groups and how strongly it should protect the highest-vulnerability cells.

Group construction: West-to-east bands
Number of groups: 4
Fairness rule: Per-capita spend (recommended)
Equity tolerance ε: 5,00
High-vulnerability threshold (%): 70 (0–100)
Protected HV budget share γ: 0.20 (0.00–1.00)
Early delivery bonus: 0,00
Hazard combination rule: max

Technical and diagnostic options

Most users can leave these at their defaults. They mainly affect reproducibility and the optional analysis tabs.

Optimization engine: HiGHS (recommended)
Simulation seed: 42
☑ Drop zero-population cells
☐ Run runtime benchmark
☐ Run scenario stress test

Turn on the stress test if you want to see whether the result holds up under modest

Hazard design

Choose the hazard patterns you want to overlay on the district. The sliders keep the shapes inside sensible ranges so it is easy to build a plausible scenario without hand-editing a table.

Hazard families: Circular hotspot, Wavy band

Circular hotspot

A rounded hotspot that fades away from one or more centers.

Number of centers: 2 (1–6)
Spread: 0.12 (0.03–0.25)

Ellipse ratio is only used for elliptical hazards.

Wavy band

A curved band that mimics an irregular path such as a wind or storm track.

Number of centers: 1 (1–6)
Spread: 0.08 (0.03–0.25)

Ellipse ratio is only used for elliptical hazards.

Aid assumptions

These settings define what each aid item costs, how strongly the model values it, how much coverage one unit is assumed to provide, and whether there are stock or timing limits.

Read the three key columns this way: `cost` is how much one unit costs, `benefit weight` tells the optimization how attractive the item is, and `coverage per unit` is used only for reporting how much effective coverage the plan delivers.

Cash is the most flexible item. It usually scales easily and can fill small residual budget gaps.

☑ Include this item
☐ Phase 1 only
☐ Whole units only

Cost per unit: 1,00
Benefit weight: 1.00 (0.00–3.00)
Coverage per unit: 1.00 (0.00–3.00)
Maximum units per cell per phase: 2000,00

Cash is left unconstrained at the district level in the default setup.

Results

Headline results

These top-line figures summarize what the optimized plan achieved under the current assumptions.

Status optimal	Objective value 3146296787.23	Budget used 100.0%	Runtime 1.76s
Overall coverage 31.0%	HV coverage 80.1%	HV spend share 52.5%	Cells analyzed 5,005

1. The landscape before any aid is allocated

These charts describe the district the optimizer is looking at. The histograms show how population and vulnerability are distributed across cells. The map shows where the synthetic hazard surface is concentrated.

Map view

Use these controls to compare the optimized plan with the blanket baseline for the same phase and aid type. Darker cells received more units.

Time step: 1
Aid type: All

Figure 5.4: User interface of the anticipatory aid allocation dashboard. The upper-left panel shows the main policy configuration menu, where the user selects the planning horizon, phase budgets, budget mode, fairness settings, high-vulnerability protection rules, and technical options such as solver choice and diagnostic runs. The upper-right panel shows the hazard-design and aid-assumption controls, where the user specifies hazard shapes and item-level parameters such as cost, benefit weight, coverage per unit, stock limits, and timing restrictions. The lower-left panel shows the results overview, including headline metrics and explanatory charts that summarize the simulated landscape before allocation. The lower-right panel shows the map view, where optimized and blanket allocations can be compared side by side for the same phase and aid type. Together, these panels allow the user to define assumptions, run the optimization, and interpret the resulting allocation pattern in both summary and spatial form.

Beyond its core functionalities, the web application is designed to be adaptable, allowing users to customize it for different anticipatory action needs. Governments, NGOs, and researchers can modify constraints, refine objectives, or explore alternative optimization approaches to better align with policy priorities and operational challenges. Specifically, users could integrate:

- **Additional constraints:** The model could integrate *spatial contiguity* requirements to ensure that allocated resources form coherent, contiguous distribution zones across panchayat or district lines, potentially reducing logistical inefficiency or administrative infeasibility that arises from fragmented aid delivery across non-contiguous regions. Drawing from spatial optimization frameworks used in land-use planning, the model could enforce contiguity constraints similar to those applied in wildfire mitigation and habitat conservation, where maintaining uninterrupted resource patches enhances management effectiveness [505].

  Furthermore, a *multi-hazard interaction constraint* could be added to boost allocations in regions where overlapping risks such as cyclones and floods compound vulnerability, as has occurred in West Bengal [466]. Another constraint could mandate that a specified minimum share of the total budget be directed to predefined high-risk or high-vulnerability zones (e.g., coastal areas or regions with limited irrigation in West Bengal) [506].

- **Enhanced objective functions:** Moving beyond a single criterion (e.g., vulnerability-weighted impact), future models could adopt *multi-objective frameworks* balancing equity, speed of impact, and cost-efficiency [507]. In contexts like West Bengal, where recurrent cyclones demand both short-term relief and longer-term resilience, a *resilience-based scoring* approach could reward interventions that mitigate future risks or facilitate faster recovery. The objective function might also integrate indicators of agricultural productivity, market connectivity, or healthcare availability to capture a more holistic measure of community well-being [508].

- **Alternative mathematical programming approaches:** While mathematical programming suits many anticipatory transfer scenarios, *stochastic and robust optimization* techniques could be beneficial when climate forecasts are uncertain or budgets fluctuate [509]. Moreover, rather than developing fully stochastic formulations, one study suggests that leveraging high-performance computing to simulate thousands of scenarios using the same optimization model can yield valuable insights into the distribution of near-optimal solutions, providing a practical alternative when the uncertainty is difficult to quantify explicitly [510]. For large-scale

> deployments spanning multiple districts or states, *metaheuristics* like genetic algorithms can handle complex constraints and speed up solutions in practice [511]. Additionally, in contexts where aid packages include not only cash but also in-kind transfers (e.g., seeds, seedlings, or fertilizer), employing *mixed integer programming* can effectively model discrete decisions associated with selecting different types of interventions [230]. For instance, farmers in Bihar have benefited from electric pumps that help in transplanting paddy saplings when there is deficient rainfall [512].

By modifying the application in these ways, users can align the optimization framework with their own anticipatory action contexts, whether adjusting constraints for different geographic regions, redefining objective functions to prioritize specific vulnerabilities, or integrating alternative optimization techniques to better reflect operational realities.

5.3.2 Computational scalability across an increased surface area

To investigate how the multi-item, multi-phase MILP scales with increasing spatial resolution, we ran tests varying the number of pixels from 1,000 up to 30,000, applying two different solvers (`cbc` and `appsi_highs`). Figure 5.5 illustrates computation times under three representative model setups: one without equity or high-vulnerability requirements, one with per-capita, cost-weighted equity and tolerance $\varepsilon = 5.0$, and one with stricter per-capita equity $\varepsilon = 1.0$ plus a mandatory 20% protected budget share for pixels with $v \geq 0.80$.

Overall, runtime increases with problem size and with the addition of stricter policy constraints, as expected. Solver choice also matters materially, with the two open-source backends exhibiting different scaling patterns across configurations. Even so, the experiments show that the proposed formulation remains tractable over large synthetic instances, supporting its use for high-resolution scenario analysis. GLPK was not included in the main comparison because it was materially slower in preliminary testing, especially under the more constrained model variants, which is consistent with prior comparative studies [513, 514].

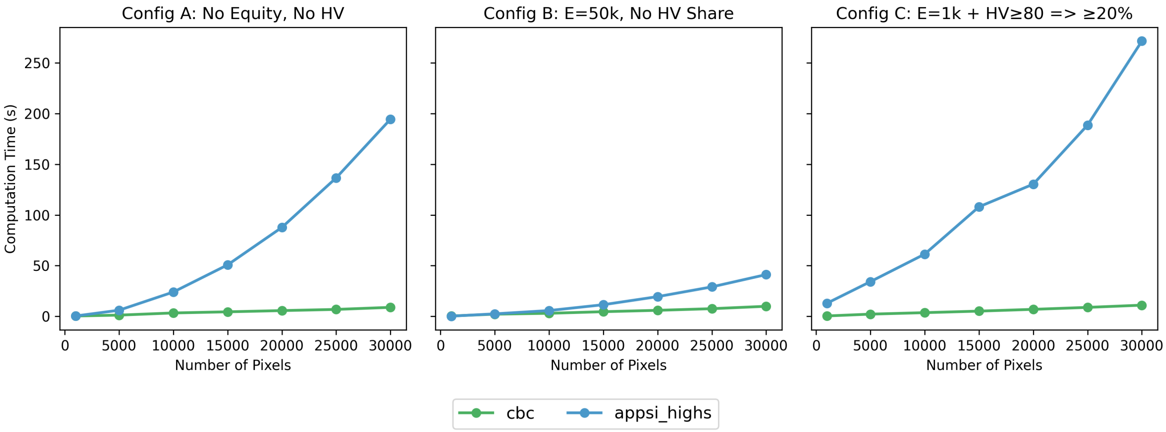


Figure 5.5: Computation times for two solvers (`cbc` in green, `appsi_highs` in blue) at different problem sizes (1k–30k pixels) and three constraint setups. The left-most panel imposes no equity or high-vulnerability constraints; the middle panel uses per-capita, cost-weighted equity with $\varepsilon = 5.0$; and the right-most panel sets $\varepsilon = 1.0$ and requires a 20% minimum budget share for high-vulnerability pixels with $v \geq 0.80$.

5.3.3 Comparison of a blanket allocation approach against an optimized approach

To assess how a multi-phase optimization model compares with a simpler population-based policy, we conducted ten synthetic hazard scenarios in which (i) a *blanket* approach allocated funds strictly proportional to population, and (ii) an *optimized* approach leveraged vulnerability scores, equity constraints, and item-specific capacities. Figure 5.6 illustrates the outcomes for three key metrics: (1) fraction of at-risk individuals in high-vulnerability (HV) areas receiving sufficient coverage, (2) correlation between vulnerability and pixel-level spend, and (3) fraction of the total spend directed to those HV areas.

Across the ten scenarios, the blanket baseline delivered a nearly constant high-vulnerability coverage rate of about 31%, whereas the optimized approach averaged about 66% and ranged from roughly 36% to 98%. The targeting metric shows the same pattern. Under the blanket rule, the mean correlation between vulnerability and pixel-level spend was slightly negative at about -0.03, reflecting the fact that population alone does not systematically line up with the simulated hazard field. Under the optimized rule, the mean targeting correlation rose to about 0.34, and it was positive in every scenario. These results show that once vulnerability is brought explicitly into the allocation problem, spending becomes materially more aligned with modeled need.

The spending-share metric points in the same direction. On average, the blanket base-

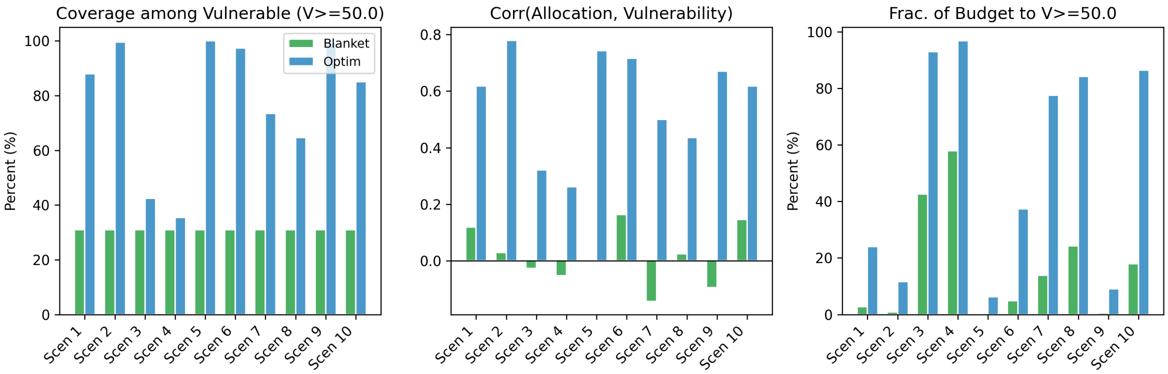


Figure 5.6: Comparative performance of blanket and optimization-based allocations across ten synthetic hazard scenarios. Shown are (a) the fraction of at-risk individuals in high-vulnerability areas who receive sufficient coverage, (b) the correlation between vulnerability scores and pixel-level spend, and (c) the share of the total spend directed to high-vulnerability pixels. In each metric, the optimization-based approach outperforms the population-proportional blanket strategy.

line directed about 40% of total spend to the high-vulnerability set, while the optimized model directed about 70%. In several scenarios, the optimized share exceeded 90%. This indicates that the model is not merely spreading resources more thinly across all exposed areas. Rather, it is systematically concentrating spending where vulnerability is greatest, while still respecting the budget, fairness, and capacity structure of the optimization problem.

Figure 5.7 offers a spatial viewpoint of these differences in Purba Medinipur. The left panel displays a simulated vulnerability map, with darker regions indicating higher risk levels. A purely population-based blanket approach allocates many resources to populous but less-vulnerable areas, spreading limited funds too thinly across the district. In contrast, the optimized model clusters a greater share of investments in pockets of high hazard exposure, reflecting the region's actual vulnerability patterns. Consequently, more severely threatened pixels achieve coverage levels closer to their population capacity, thereby enhancing resilience for those communities.

Overall, these findings underscore the limitations of uniform per-capita aid in complex disaster contexts, where vulnerability can vary dramatically across neighboring locations. By leveraging techniques from multi-phase mixed-integer optimization, decision-makers can satisfy simultaneous constraints such as integer item allocations, budget ceilings, and equity considerations, while still achieving a stronger alignment between resource distribution and ground-level risk. In practice, this approach can expedite the recovery of hardest-hit populations and reduce the need for subsequent emergency interventions.

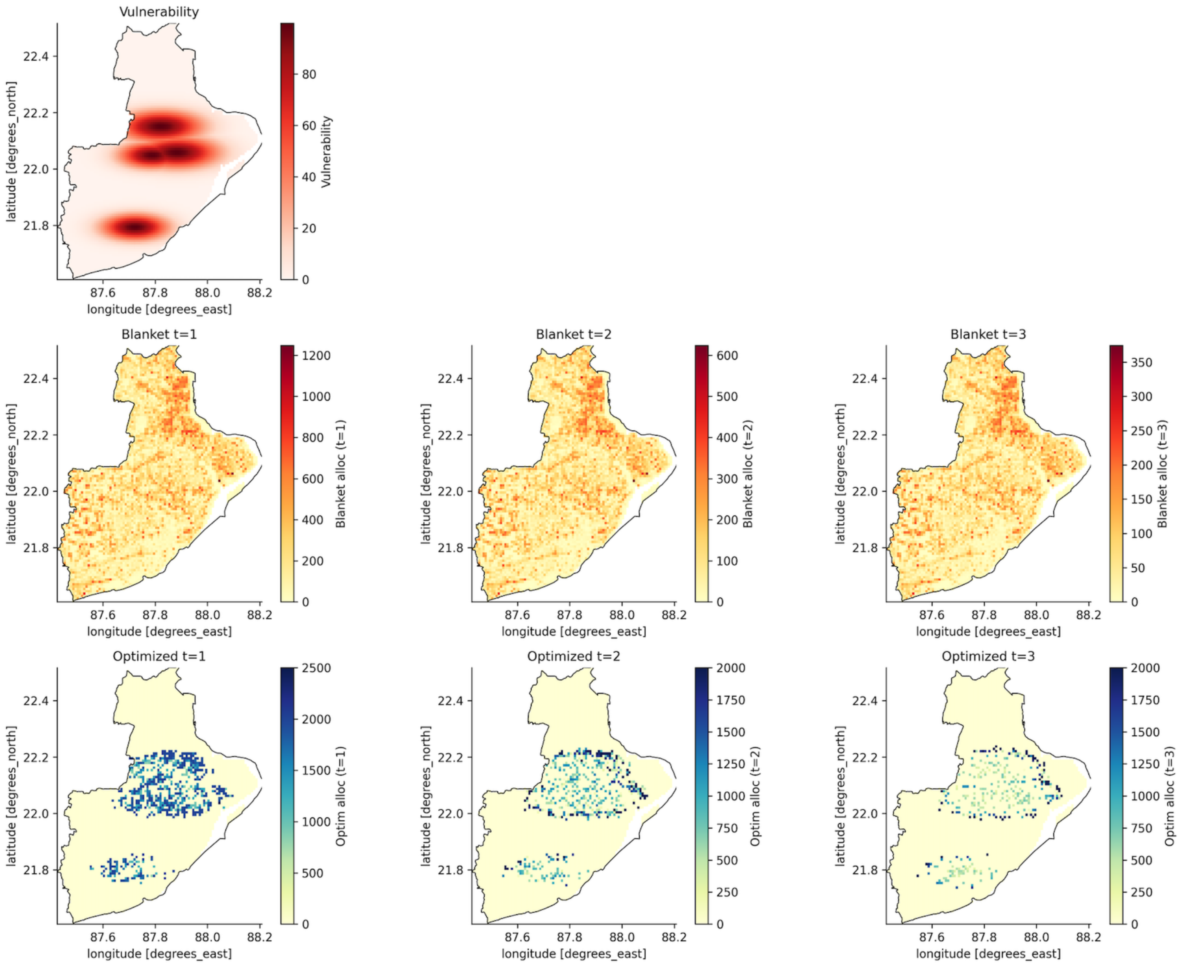


Figure 5.7: Spatial illustration of blanket vs. optimized allocations in Purba Medinipur. The upper panel shows a simulated vulnerability map in a selected scenario; darker shades indicate greater hazard intensity. The middle row provides a spatial view of blanket allocation distributed solely by population density at each time step. The bottom row shows the optimization-based allocations that favor highly vulnerable pixels.

5.4 Discussion: areas for future research in data-driven anticipatory action

To maximize the effectiveness and impact of targeted, data-driven anticipatory action, several critical research gaps must be addressed. Future efforts should focus on advancing the integration of early warning systems with optimization-based decision support tools, bridging practical barriers related to implementation scale and data availability, and strengthening interdisciplinary collaboration among climate scientists, data engineers, policymakers, and local stakeholders. Specifically, future research should prioritize:

Making early warning systems more people-centric for improved anticipatory action. A major challenge in anticipatory action is both the accurate prediction of hazards and their meaningful translation into people-centric vulnerability assessments. As one study notes, "there is a conceptual gap between 'impact-based' forecasts and actual hazard impact," with current systems relying on meteorological thresholds while failing to incorporate real-time vulnerability data such as soil moisture depletion, groundwater levels, and socio-economic stressors [256]. Integrating multi-layered institutional data—such as farmer debt burdens and fluctuating market access—into forecasting models could help translate early warnings into more actionable vulnerability assessments. Additionally, one study highlight the inaccuracy of drought early warnings in South India, reporting a high false alarm rate (0.81 for the Kharif season, 0.30 for the Rabi season), demonstrating the need for greater accuracy to ensure that anticipatory action interventions are not misguided [515].

Connecting optimized allocation outputs with digital cash distribution mechanisms. Mobile-based direct cash transfers have proven effective in delivering fast financial relief to vulnerable populations. Countries like Kenya (M-Pesa) and Brazil (Bolsa Família) have demonstrated how digital payment networks can ensure that funds reach intended recipients without intermediaries or bureaucratic delays [516, 517]. India's Direct Benefit Transfer (DBT) system, launched in 2013, was designed to reduce inefficiencies in welfare distribution, addressing long-standing issues where government funds often failed to reach intended beneficiaries due to leakages, corruption, and administrative inefficiencies. Before DBT, studies found that 58% of subsidized food grains never reached below-poverty-line households, and the government spent INR 3.65 for every INR 1 actually received by beneficiaries. To fix these inefficiencies, DBT was implemented to digitally transfer welfare payments directly into recipients' bank accounts, using a biometric verification system (Aadhaar), financial inclusion initiatives (Jan-Dhan accounts), and widespread mobile connectivity. This system has scaled rapidly, disbursing over INR 5.5 trillion (as of 2020–21) across hundreds of welfare schemes, covering subsidies for food, fuel, education, and employment guarantees. However, regional disparities in banking infrastructure and mobile penetration persist, limiting DBT's reach in rural and marginalized communities [518].

Disbursement of cash in anticipatory action efforts could potentially leverage existing DBT initiatives, with optimization models helping to allocate funds to recipients. However, challenges remain in ensuring inclusion, as some smallholder farmers may lack formal land titles or mobile banking access, potentially complicating disbursement. Additionally, blockchain technology has been proposed as a solution to enhance the transparency and speed of anticipatory cash transfers, with one study claiming that "decentralized finance applications can expedite and reduce the costs of funds transfer, exemplified in blockchain-enabled cash transfers" [459]. Further research is needed to establish seamless integration between anticipatory action planning frameworks—such as those proposed in this study—and operational disbursement systems, ensuring efficiency, transparency, and equitable access to preemptive funds.

Involve various stakeholders in the formulation of optimized anticipatory action algorithms. To ensure that optimized anticipatory action initiatives can be adapted to new contexts where they have not yet been deployed, future research should prioritize the inclusion of policymakers and stakeholders in developing a diverse repository of scenarios. By "diverse repository," we refer to a structured collection of predefined optimization setups that future users can select from, facilitating rapid deployment of anticipatory action strategies tailored to different needs. Including this into the web application could lead to an interactive dropdown menu where users could choose from a wide range of pre-configured optimization models based on their specific operational constraints and objectives. These setups would include variations such as: (i) cash-only allocations versus mixed interventions that incorporate resource distribution (e.g., water pumps, fertilizer); (ii) inclusion or exclusion of equity constraints to balance funding across regions or vulnerable groups; and (iii) multi-phase versus single-phase allocations to account for staggered funding releases. By building this repository with direct input from policymakers, humanitarian actors, and technical experts, future users—whether governments, NGOs, or local responders—would have access to a comprehensive suite of optimization frameworks that are both adaptable and practically relevant.

The experience of Bangladesh's Central Emergency Response Fund (CERF)-funded anticipatory action framework illustrates the importance of stakeholder involvement in designing flexible and scalable intervention models. For example, a significant proportion of Bangladesh's CERF allocation was directed toward cash transfers, yet an impact study found that the 2020 transfer values were too small, necessitating a review and adjustment of the amounts in subsequent phases. This underscores the need for an optimization framework where users can test different allocation schemes, adjusting parameters dynamically based on empirical evidence. Moreover, planned anticipatory action activities should be clearly distinguished from those better suited for post-shock response, as interventions such as water treatment activities or dignity kit distribution may not fit within the anticipatory action framework [519]. Local stakeholders can also play a critical role in defining equity constraints within anticipatory action frameworks, as evidenced by Start Network's experience in Bangladesh, where "the increased use of community-owned triggers helped enhance understanding and acceptance of technical, scientific meteorological knowledge" [41]. A similar logic applies to the design of optimization algorithms for anticipatory aid allocation. Just as participatory approaches to early warning system triggers can build trust and improve decision-making, involving local actors in defining optimization parameters—such as equity constraints, eligibility thresholds, and allocation priorities—ensures that allocation models, including cash-based delivery components, align with on-the-ground realities. This participatory approach may enhance both the legitimacy and effectiveness of optimization-driven interventions.

Incorporating lessons from a broad range of local contexts into a repository of optimization algorithms may help ensure that anticipatory action strategies are tailored to the specific coping mechanisms and economic realities of different regions. For instance, while flood-affected communities in Bangladesh might prioritize direct cash transfers for immediate asset protection, drought-prone regions in Namibia may benefit more livestock feed distribution [520].

Technical governance of open-source optimization frameworks for anticipatory action. From a technical standpoint, scaling solutions beyond pilot environments requires disciplined software engineering practices and thoughtful governance structures. Maintaining a single open-source repository with clear pull request protocols and defined merge rules ensures code integrity and traceability, especially as more contributors—ranging from data scientists to NGO partners—begin to collaborate. Establishing a centralized project ownership or steering committee can help manage the direction of new features, maintain quality standards, and coordinate updates [521, 522]. Rigorous version control is particularly important for auditing model changes over time.

Engineering solutions for operational scale and cost efficiency. Moving from pilot projects to broader coverage demands a structured approach to cloud infrastructure, data management, scheduling, and user interfaces. In India, for example, real-time alerts from an early warning system could be fed into a container-based workflow that ingests hazard data, runs the optimization model, and provides recommended allocations. Meanwhile, secure APIs with role-based authentication and encryption help guard sensitive beneficiary information, allowing only authorized agencies to access optimized allocation plans [523]. Simple dashboards or mobile apps could then present these allocation results in real time to local decision-makers, ensuring transparency and timely disbursements. By integrating data ingestion, compute orchestration, security protocols, and user-facing tools into a cohesive system, large-scale anticipatory action initiatives can remain both cost-effective and adaptive, even as hazard conditions evolve.

Searching for the right policy. Another complementary approach is to view anticipatory action through the lens of *sequential decision-making.* Rather than finding a single best allocation for one point in time, the goal is to identify the best *policy*, a rule for making allocation decisions that performs well across many possible futures. In this view, our MILP represents one such policy: it maps the current state S_t (budgets, capacities, vulnerability maps, and constraints) into an allocation $X_t^\pi(S_t) = \{x_{i,t,k}\}$. Powell's framework [524] formalizes this by optimizing directly over policies:

$$\max_{\pi \in \Pi} \mathbb{E}\Big[\sum_{t=1}^{T} C\big(S_t,\, X_t^\pi(S_t)\big)\Big] \quad \text{s.t.} \quad S_{t+1} = S^M\big(S_t,\, X_t^\pi(S_t),\, W_{t+1}\big), \tag{5.40}$$

where $C(\cdot)$ is the contribution function, here the time-weighted vulnerability-population benefit in Eq. (5.8), S^M describes how the system evolves, and W_{t+1} represents new information such as changing forecasts or updated vulnerability data. To operationalize this idea, we can define a family of candidate policies $\Pi_\Theta = \{\pi_{m,\theta}\}$, each corresponding to either a different parameterization or an alternative objective formulation, and perform a meta-search:

$$(m^\star, \theta^\star) \in \arg \max_{m \in \mathcal{M},\, \theta \in \Theta_m} \mathbb{E}\Big[\sum_{t=1}^{T} C^{(m)}\big(S_t,\, X_t^{\pi_{m,\theta}}(S_t)\big)\Big]. \tag{5.41}$$

Here, m indexes the choice of objective function and θ its parameters. For instance, our baseline objective $C^{(\mathrm{VB})}$ maximizes time-weighted benefit in vulnerable populations, while an alternative could emphasize *equitable resilience*:

$$C_t^{(\mathrm{EQ})} = \sum_{i,k} \big(\omega_t b_k v_i p_i\big) x_{i,t,k} - \xi \sum_{g<h} \left| \frac{y_{g,t}}{d_g} - \frac{y_{h,t}}{d_h} \right|, \tag{5.42}$$

where the second term penalizes disparities in normalized group spending, controlled by ξ. Another variant could reward *temporal stability*:

$$C_t^{(\mathrm{ST})} = \sum_{i,k} \big(\omega_t b_k v_i p_i\big) x_{i,t,k} - \eta \sum_{i,k} \big| x_{i,t,k} - x_{i,t-1,k} \big|, \tag{5.43}$$

discouraging abrupt reallocations that may be operationally infeasible. By evaluating such policy variants $\pi_{m,\theta}$ on historical or stress-tested hazard sequences and comparing common KPIs (coverage C_{HV}, targeting ρ, spend share ϕ, equity dispersion), policymakers can identify not just the best model configuration but the best *decision rule.* This embodies Powell's principle of "optimizing over policies" rather than single decisions, reframing anticipatory optimization as a continuously improving, evidence-driven policy search process [525].

Building on this view, it is useful to contrast our proposed policy-based framework with how anticipatory action is typically operationalized in practice. Most existing systems rely on *trigger rules*—fixed thresholds on one or two indicators (e.g., forecasted rainfall or flood probability) that, once crossed, activate a pre-defined response such as cash disbursement [64]. In Powell's terminology, these represent narrow, single-condition policies: $\pi_{\text{trigger}}(S_t)$ = "act if indicator exceeds threshold". In contrast, our approach formulates a more holistic *allocation policy*, implemented through a mixed-integer optimization model that determines *how much to allocate, where, and in what form*, given budgets, vulnerabilities, and equity requirements. The trigger still governs *when* anticipatory action begins, but the optimization policy defines *what happens once it is activated*, optimizing resources under multiple simultaneous constraints. Looking forward, Powell's sequential decision framework offers a natural extension, by treating this optimization model as one among many competing policies that can be tested, compared, and improved across successive hazard activations [**?**]. In this way, anticipatory action can evolve from static trigger-based activation toward a broader system of *learned decision rules*, continuously refined through data, simulation, and real-world feedback.

5.5 Conclusion

This chapter developed a spatially explicit, time-phased mixed-integer programming framework for optimizing anticipatory aid allocation in disaster-prone regions, demonstrating how structured, equitable, and timely planning can enhance the effectiveness of humanitarian interventions. By integrating simulated vulnerability data and explicit equity constraints, the model moves beyond traditional trigger-based systems, enabling targeted allocations that are responsive to both local vulnerability and operational constraints. The case study in Purba Medinipur, India, showed that optimized multi-phase allocations redirected up to 80% of resources toward the most vulnerable pixels, doubled high-risk coverage relative to a uniform per-capita rule (from 31% to over 65%), and achieved strong equity across administrative groups. The interactive web-based decision-support tool further enhances the usability of this model, enabling policymakers and practitioners to dynamically explore various allocation scenarios. Computational testing confirmed that the proposed optimization model is scalable, efficiently handling large-scale, high-resolution problems across diverse solver configurations.

Looking forward, several avenues can enhance this research: incorporating real-time early warning data will bolster the accuracy and responsiveness of vulnerability assessments; integrating optimized aid-allocation plans with existing digital financial infrastructure, such as India's Direct Benefit Transfer system where appropriate, could streamline disbursements; and expanding stakeholder involvement in defining scenario constraints and objectives would ensure contextual relevance and acceptance. Additionally, refining the technical governance of open-source optimization frameworks and advancing the engineering of scalable cloud-based deployment solutions will be essential for widespread adoption and operational impact. Looking ahead, in terms of formulating and deploying adaptive optimization models, the next frontier is moving from single optimization runs toward *learning across policies*. Future work could evaluate alternative policy formulations within Powell's sequential decision framework, allowing governments and humanitarian organizations to test, compare, and refine decision rules over repeated activations. Combined with real-time early warning data and digital disbursement systems such as India's Direct Benefit Transfer network, such adaptive policy learning could transform anticipatory action from a threshold-based activation mechanism into a continuously improving, evidence-driven system for proactive risk management.

6 Discussion and conclusion

This chapter brings together the main findings of the thesis and reflects on their implications. It first summarises the key contributions across the forecasting, optimisation, and integration chapters. It then assesses the strengths and limitations of the approach, including its empirical scope, modelling assumptions, and institutional realism. The chapter next examines the practical conditions required to translate a working prototype into an operational government capability, focusing on governance, finance, and implementation capacity. It concludes by outlining priorities for future research aimed at embedding anticipatory, optimisation-guided decision-making within routine agricultural disaster response.

The structure of this discussion section is informed by the framework proposed by Docherty and Smith (1999) [526]. While not followed rigidly, their approach provides a useful organising principle, which is adapted here to suit the context of this thesis.

6.1 Key contributions and integrative synthesis

Data science in development is rarely a solitary or purely technical enterprise. In public-sector settings, value is created at the boundary between data, institutions, and domain expertise rather than inside the model alone. That is especially true here, where crop science, climate analytics, remote sensing, optimization, humanitarian operations, and public administration must be brought together before an early warning can become an action.

The contribution of this thesis therefore lies not in any single model viewed in isolation, but in the construction of a socio-technical architecture that links prediction to decision-making. Framed this way, the thesis sits most clearly within data science for public decision-making and development, rather than claiming equal disciplinary contributions to agronomy, meteorology, and humanitarian practice all at once.

This positioning also clarifies the thesis's main contributions. Its central methodological contribution is the end-to-end coupling of impact-based yield forecasting and optimization. Its central theoretical contribution is the reframing of early warning systems as decision architectures rather than information pipelines. Its central policy contribution is to make the forecast-to-action link explicit, auditable, and open to revision. The argument advanced here is that predictive information creates greater public value when it is embedded into a structured decision-making process that accounts for real-world constraints.

The opening perspective paper establishes the conceptual foundation for this argument, arguing that while global investment in early warning capacity has expanded, its real impact remains constrained by the absence of decision-making mechanisms that connect forecasts to resource allocation. It positions early warning systems not as information systems but as potential decision architectures capable of guiding timely, evidence-based interventions when linked with formal optimization [527].

Building on this foundation, the three subsequent papers develop the core layers of the proposed architecture. The first conducts a systematic review of agricultural early warning systems, synthesizing evidence on how advances in earth observation, machine learning, and digital infrastructure have expanded forecasting capacity, while systematically identifying the structural gap between predictive outputs and actionable decision-making [528]. Although motivated by the context of rice production and food security in India, the review derives generalizable insights about the institutional, technical, and operational barriers that prevent forecast information from translating into policy action. The second paper operationalizes this agenda by constructing spatially explicit yield prediction models for rice in India using modern machine-learning methods, demonstrating how satellite and climatic data can detect early-season production risks at scales relevant for government intervention [250]. The third paper completes the architecture by designing optimization models that translate forecast signals into anticipatory policy strategies, allocating resources such as funds or other interventions to minimize expected welfare losses. Together, these studies articulate and implement a systematic integration of forecasting and optimization, transforming early warning systems from predictive tools into structured decision frameworks capable of guiding timely and cost-effective government responses.

Beyond its applied contributions, the thesis advances a conceptual reframing of early warning systems as decision-analytic structures rather than information pipelines. It proposes that the value of forecasting should be assessed not solely by predictive accuracy, but by its capacity to improve downstream policy decisions under uncertainty. This perspective builds on work in decision science and environmental information theory that emphasizes value-of-information analysis as the appropriate metric for informational worth [529]. It also aligns with emerging research in agricultural data science, which locates data value in its ability to generate decision-relevant insight through prediction and optimization rather than through data accumulation alone [530]. By explicitly linking anticipatory information to welfare-based optimization, the thesis establishes a framework for evaluating early warning systems in terms of decision value rather than signal precision.

6.2 Strengths of this study

This thesis offers three principal strengths: conceptual clarity, methodological rigor, and policy relevance.

First, it advances a conceptual contribution by reframing early warning systems as decision-analytic architectures rather than information pipelines. While substantial progress has been made in improving forecast accuracy through earth observation and machine learning, the thesis formalizes the link between predictive models and welfare-based policy optimization. In doing so, it clarifies that the value of forecasting lies not solely in signal precision, but in its capacity to improve downstream decision quality under uncertainty. This perspective aligns with recent work on the value of information and data-to-decision pathways [529], and contributes a structured framework for evaluating early warning systems in terms of decision impact rather than predictive performance alone.

Second, the thesis demonstrates methodological rigor through the operational coupling of machine-learning yield models and optimization-based allocation strategies. Rather than treating forecasting and planning as separate analytical domains, the work shows how probabilistic yield predictions can be embedded directly into anticipatory resource allocation models. This approach reflects broader disciplinary convergence, as global machine learning-based yield forecasting increasingly achieves policy-relevant lead times [531], and operations research applications in humanitarian logistics demonstrate how uncertainty can be formally incorporated into anticipatory decision models [532]. By integrating these strands within a coherent architecture, the thesis moves from theoretical possibility to implementable design.

Third, the thesis prioritizes transparency and reproducibility. All analytical components, from model development to optimization routines, are implemented using open-source tools, documented methods, and traceable data pipelines. This directly addresses persistent reproducibility gaps in geospatial and remote-sensing research. A review of 200 remote-sensing-related papers published between 2014 and 2022 found that only 37 provided access to source data and just 16 shared replicable code or materials [533]. Adherence to the FAIR principles, ensuring that workflows and data are Findable, Accessible, Interoperable, and Reusable, enhances both scientific credibility and policy uptake. As Wilkinson et al. (2025) argue, transparent and well-documented workflows maximize their value as research assets and facilitate adoption by the wider community [534].

6.3 Limitations of this study

This thesis has three principal limitations: empirical scope, institutional realism, and structural simplification.

First, empirical validation is conducted in a single context, namely rice production in India. This setting was chosen deliberately because it offers unusually rich agricultural, climatic, and administrative data, as well as a policy environment where food security decisions are closely linked to seasonal forecasts. However, testing the framework in one national context limits external validation. Applying the architecture in other regions would require adaptation to differences in data availability, institutional capacity, and agro-ecological structure. In lower-data settings, implementation may necessitate coarser spatial inputs, simplified models, or greater reliance on expert judgment, potentially affecting predictive precision and policy granularity. As systematic reviews of machine-learning-based yield prediction note, models often exhibit limited transferability across crops, regions, and environmental conditions without careful recalibration [274].

Second, the optimization models abstract from behavioural and political dynamics that shape real-world decision-making. While simulated hazard and production-deficit scenarios enable transparent evaluation of anticipatory allocation under uncertainty, actual government responses operate within institutional environments that are often more complex than represented here. Budget cycles, administrative capacity, coordination across agencies, and political constraints influence how early warnings translate into action. As research in humanitarian logistics has shown, limited decision-maker involvement in model design is associated with weak implementation and misalignment between analytical objectives and practitioner priorities [535]. Optimization models developed in controlled settings may therefore require substantial adaptation before becoming operationally useful. Their effectiveness depends not only on technical soundness, but on early engagement with the institutional actors who must ultimately adopt and apply them.

Third, the framework simplifies the structure of the policy problem in order to remain analytically tractable. The welfare function is reduced to a limited set of objectives, and parameters such as risk exposure, response effectiveness, and resource availability are treated as fixed within each simulation. In practice, governments operate in environments characterized by interacting hazards, multiple actors, and evolving institutional and socio-economic conditions. Recent reviews of adaptive pathways and deep uncertainty tools emphasize that increasing structural complexity challenges existing modelling approaches and requires explicit attention to interacting uncertainties and cross-sector dynamics [536, 537]. By holding key parameters constant, the model isolates the value of anticipatory optimization, but does not capture dynamic feedbacks, evolving objectives, or shifting governance constraints. Extending the framework to incorporate adaptive priorities and time-varying conditions might enhance realism and strengthen its relevance for long-term policy design.

6.4 Institutional and operational challenges

This discussion explores what it may take to move from a working prototype to an institutional capability that repeatedly turns early warnings into timely, equitable actions. We proceed from higher-level constraints (law, governance, finance) to lower-level delivery (policies, engineering, human adoption), and conclude with near-term researchable priorities. The aim is not to prescribe a single path but to surface practical open questions that, if answered, could enable responsible and sustained scale-up of anticipatory action systems which integrate early warning with quantitative decision-making.

6.4.1 System-level constraints around governance and finance

From humanitarian pilots to government ownership As the IFRC observes, "government ownership of anticipatory action is essential for it to reach scale and sustainability," yet most systems are still "externally driven, project-based approaches rather than routine, government-led components of disaster response" [538]. Humanitarian pilots proved that acting early saves lives, but the report stresses that long-term impact depends on embedding this logic in domestic law, policy and finance. The challenge, it notes, is not to transfer a "tool" (the anticipatory action framework) but to “durably strengthen the underlying capacities, legal and policy frameworks, ownership and confidence of government institutions to use the tool” so that early warnings routinely trigger preventive action when risk levels demand it [538]. In short, the IFRC suggests that the centre of gravity must shift: from humanitarian agencies demonstrating what is possible, to governments making it standard practice within their disaster management systems.

Several countries have begun to close this gap by embedding anticipatory action in national law. Fiji’s 2024 *Disaster Risk Management Bill* now allows state resources to be used “in anticipation of a potential emergency,” formally recognizing the legitimacy of pre-impact spending [539]. Bangladesh represents a slower, incremental path: over two decades, its disaster management law has evolved from reactive relief to proactive risk governance. As Zaman et al. (2022) describe, reforms culminating in the *Disaster Management Act* (2012), *Standing Orders on Disasters* (2019), and *National Plan for Disaster Management* (2021–2025) established the legal basis for acting early, with mandates, financing rules, and inter-ministerial coordination clearly defined [540].

Who has the legal authority to act early? Legal recognition of anticipatory action is only the first step; it must be matched by clear institutional authority. Anticipatory action often stalls because no single body controls the triggers, budgets and technical oversight [541, 542]. The IFRC report makes clear that turning anticipatory action from a project into a policy means navigating complex laws, fiscal rules, and institutional rivalries. Disaster management agencies may have the mandate to coordinate, but not the authority to spend; finance ministries may control the funds, but not the triggers; and line ministries often act only after formal declarations. As a result, even where governments support the principle of acting early, their systems are still built for reacting late [538].

To bridge this gap, national disaster management agencies are increasingly taking on the role of coordinating anticipatory action schemes, chairing cross-ministerial working groups that include meteorological services, finance ministries and humanitarian partners [538]. Sierra Leone’s *Disaster Risk Financing Strategy 2024–2029* explicitly assigns this role to its National Disaster Management Agency, which is directed to “facilitate the use of anticipatory finance, including forecast-based triggers for the early release of funds ahead of the impact of disasters” [543]. Mozambique’s Law No. 10/2020 on Disaster Risk Management and Reduction expands the country’s legal framework beyond rapid-onset emergencies. It assigns the National Institute of Meteorology and the National Institute for Disaster Management responsibility for establishing an alert system for slow-onset hazards such as droughts, enabling “early warning and early action.” In doing so, it strengthens earlier legislation that had focused mainly on sudden disasters [544].

Yet even with such mandates in place, the tension between national consistency and local flexibility remains unresolved. Kenya’s drought management system illustrates the trade-off: county committees can interpret indicators using local knowledge, improving contextual relevance but sometimes blurring accountability and delaying payouts [545]. Finding the right balance may require a hybrid model: one in which a national agency maintains the official trigger list and budget framework, while subnational authorities can propose context-specific adjustments under defined rules and review cycles. Determining how frequently to revise triggers, and how to validate local adaptations without undermining speed, remains an open governance question. If funders increasingly link pre-arranged finance to documented governance structures, clarity over who maintains and updates these rules may become a precondition for scaling anticipatory action.

This fragmentation of authority is not only a coordination challenge; it also creates a design risk for integrated early warning and decision systems. When mandates are split across agencies, it becomes unclear who defines the objective function (e.g., vulnerability coverage versus equity), who curates the trigger list that enters the model, and who authorizes the budget vector on which the optimization depends.

Who has the operational capability to act early? Moreover, as the complexity and coordination demands of anticipatory systems grow, these responsible agencies must also build and retain the technical expertise needed to manage them. The effectiveness of trigger models and decision policies, such as those discussed in Chapter 5, may ultimately depend less on their mathematical design than on institutional capability: how well staff can interpret forecasts, calibrate thresholds, and adjust resource-allocation frameworks over time. Delivering humanitarian aid is already a notoriously difficult endeavour, as seen in crises such as the 2010 Haitian earthquake, where sprawling networks of relief agencies and NGOs, from the UN and USAID to the WHO, struggled within layers of bureaucracy, fragmented coordination, and host-country dissonance that often blunted or even undermined their own interventions [546]. Another challenge is brain drain: governments often invest in training staff on anticipatory methods, only to lose that capacity when personnel move to other ministries or roles [547].

How can laws and policies guarantee money moves early? Some governments struggle to act before disasters strike as financial strain and rigid funding rules keep disaster funds locked until after crises hit [548]. One potential solution is to ring-fence national budgets for pre-arranged, anticipatory spending to avoid costly delays [538]. Fiji is emerging as a rare success story. Working with the World Food Program and other UN partners, the country has introduced a system that provides anticipatory financial assistance to vulnerable households before predicted disasters. The government draws on its own contingency funds to issue early cash transfers, which are later reimbursed by the WFP, ensuring rapid and predictable support when threats arise [549]. In practical terms, this kind of pre-approved liquidity is what enables anticipatory allocation models to function. In Chapter 5, the "phase budgets" in the optimization framework represent exactly these ring-fenced early-action funds. Without committed, upfront financing, even the most carefully designed allocation model cannot be implemented in time.

How do governments maintain trust while acting on forecasts? Credibility is important for well-functioning anticipatory action systems. When forecasts prove wrong or data is mishandled, public trust can collapse, a dynamic often called the "cry-wolf effect," where people stop responding to warnings after too many false alarms [550]. Safeguards such as phased alerts, independent verification and clear "stop" mechanisms help prevent that loss of confidence. Reliable and transparent trigger models are also crucial: acting on forecasts that don't materialize wastes money and weakens support for early action. Few programmes track false alarms, but learning from them is essential to strengthen forecast-to-action systems [551]. Trust also depends on visibility: publishing trigger criteria, disbursement logic, activation logs, and spending summaries helps show that early action is evidence-based, not arbitrary. At regional level, organizations such as IGAD and ASEAN are beginning to harmonize trigger methodologies and share lessons between governments [552]. An open question is how to institutionalize this transparency without overwhelming overstretched agencies.

Within the architecture developed in this thesis, credibility determines whether the forecasting and optimization components are actually used. Even technically sound predictions and welfare-improving allocations can lose influence if repeated forecast errors or opaque model outputs undermine institutional confidence. Over time, this may make finance ministries more hesitant to release funds or decision-makers more cautious in acting early. Anticipatory systems therefore require not only accurate models and efficient allocations, but also transparent communication and systematic learning from false alarms to sustain trust in model-based decision-making.

Who pays? A further implementation question concerns not only whether early action is worth financing, but who should pay for which part of the system. There are at least two distinct layers of finance here. The first is the finance required to maintain the system itself: data acquisition, model updating, dashboard maintenance, validation, staffing, coordination, and periodic review. The second is the finance required for activation, that is, the actual disbursement of cash, seed, input support, procurement, or other anticipatory measures once risk thresholds are reached. Much of the anticipatory-action literature shows that pilots often concentrate on the second layer while underestimating the first. Yet without stable institutional funding for the predictive and decision-support apparatus itself, a system remains an intermittent project rather than an operational public capability [553, 554].

For that reason, a durable financing model will likely need to be hybrid. Over time, the core costs of maintaining the warning-and-decision system should be internalised within domestic public budgets if governments are genuinely to own the architecture rather than merely host externally funded pilots. At the same time, the funding of specific activations may require a more flexible blend of sources, including state contingency funds, disaster-management allocations, pre-arranged humanitarian finance, and, in some contexts, insurance-linked or other risk-financing instruments. This matters especially in lower-capacity settings, where even if the analytical logic of anticipatory action is accepted, fiscal rigidity may prevent timely release of funds. The key point, therefore, is that financing cannot be treated as an afterthought to modelling; it is part of the design problem from the outset [553].

6.4.2 Ethical and legal questions in targeted anticipatory action

A central ethical issue concerns the question of who should receive aid once a forecast indicates heightened risk. A narrow answer would be to prioritise those with the highest predicted production losses. But that is unlikely to be sufficient in all cases. In practice, exposure to agricultural shocks is mediated by vulnerability, coping capacity, tenure status, indebtedness, gendered access to assets, and digital inclusion. A system that targets only predicted yield loss may therefore miss households that are highly vulnerable but poorly represented in the data, including tenant cultivators, land-poor households, women with weaker formal asset claims, or those with limited digital access. Research on targeted cash transfers shows that selective inclusion can produce tensions and perceptions of unfairness even when programmes generate real benefits, while work on digital inequality in India shows that access to and effective use of digital systems are already stratified by caste, education, and occupation [555, 556].

For that reason, the ethics of implementation should be framed not only in terms of efficiency, but also in terms of fairness, transparency, and procedural protection. Any operational version of the system would need explicit allocation criteria, periodic bias and exclusion audits, and some form of human review for borderline or contested cases. It would also need a grievance mechanism through which those excluded could challenge or query the basis of decisions. Public-administration research on algorithmic decision-making suggests that legitimacy does not arise automatically from technical performance; it depends on whether decision processes are understandable, accountable, and regarded as procedurally fair by affected publics. Where automated or semi-automated systems are introduced without such safeguards, they risk undermining trust even when they improve aggregate performance [557, 558].

A related issue is legal. If a forecast crosses a specified probability threshold, does that create an entitlement to compensation, or does it merely authorise discretionary action by the state? This thesis does not need to settle that question, but it should raise it explicitly. Comparative work on automated administrative decision-making, especially in European public-law settings, suggests that once algorithmic systems influence the allocation of public benefits, issues of reason-giving, due process, hearing rights, and contestability become unavoidable [559]. Questions may arise wherever predictive systems are allowed to shape eligibility, prioritisation, or benefit delivery. This is therefore an area where future work would need collaboration not only with implementers, but also with legal scholars and public-administration specialists.

6.4.3 Uneven implementation capacity across Indian states

A potential concern is that implementation capacity in India is too uneven for a single deployment model to be realistic. The architecture developed in this thesis may be analytically general, but its operationalisation will depend on the strength of state institutions, the flexibility of budgetary systems, the quality of beneficiary registries, the reach of agricultural extension, the reliability of payment channels, and the ability of district administrations to coordinate action under uncertainty. Evidence from Indian state performance more broadly shows that programme outcomes are positively associated with initial administrative capacity, development, and accountability, while state-level fiscal conditions also vary significantly across the federation [560, 561]. This means that the same forecast-to-decision pipeline may function as a practical policy instrument in one state and remain aspirational in another.

Implementation capacity also needs to be understood organisationally, not just territorially. A functioning system would require coordination across meteorological and remote-sensing actors, agricultural departments, disaster-management authorities, district administrations, payment systems, and front-line knowledge networks. Research from Karnataka shows that access to agrarian knowledge systems already varies by caste, gender, class, geography, and distance from administrative centres, and that information is often insufficiently localised or untimely [562]. This suggests that even a technically accurate forecasting system will underperform if it is not linked to institutions that can translate risk information into locally intelligible advice and action. Put differently, implementation capacity resides not only in state capitals, but in the everyday ecology of extension, local administration, and trusted intermediaries.

6.4.4 Policy design: from single triggers to portfolios

Chapter 5 introduced an optimization framework for allocating anticipatory resources once an early-action protocol has been activated. In that formulation the policy was fixed in advance. Once a forecast crossed a trigger, a predefined set of actions, mainly phased transfers subject to budget and equity constraints, became available, and the model optimized where, how much, and when resources should be allocated. This addresses an important allocation problem, but only after the intervention logic has already been chosen. In practice, governments may face an earlier design problem. They may need to decide which trigger rules, intervention bundles, and sequencing rules should exist in the first place. The challenge, therefore, is not only how to optimize allocation within a single protocol, but how to design and compare alternative response rules before crises occur.

In the decision-analytic sense used in operations research, particularly in the work of Warren Powell, a policy is a rule that maps the current state of a system to an action [524]. Given the available information, such as a forecast, a vulnerability profile, a budget, and the time remaining before impact, the policy specifies what should be done next. Many anticipatory-action systems still operate with a narrow policy rule: a single forecast threshold that triggers a predefined response. Such rules are simple to communicate and implement, but they compress heterogeneous risk conditions into a single decision point. Emerging applications suggest more flexible approaches are possible. For example, drought early-action protocols in Mozambique use multiple trigger levels linked to escalating responses. This allows interventions to adapt to different levels of forecast risk rather than relying on a single binary activation rule [563].

Optimizing across policies Once policy is framed as a rule rather than a threshold, the analytical task changes beyond simply optimizing resource allocation within a fixed policy. It is an ex ante policy design problem: governments must decide in advance which intervention packages should exist, and which forecast conditions should trigger each of them. Candidate policy rules can be stress-tested using hindcasts, simulated shocks, and recent near misses, with the aim of retaining a small portfolio of rules that performs robustly across plausible futures. In the thesis case study, for example, district-level rice yield forecasts could be linked not to one predetermined response, but to a small menu of policy packages. A moderate projected shortfall in a district with functioning markets might trigger temporary cash support or input vouchers. A larger projected deficit might activate a broader package combining cash transfers with subsidized seed or fertilizer for the next planting season. In districts facing severe projected losses and weak market access, the response could expand further to include procurement, transport support, or emergency seed distribution. The distinction is important. Chapter 5 asks how limited funds should be allocated once a policy has already been chosen. The problem here is prior and broader: it asks which policy rules should be available in the first place, and how forecasts should be mapped to different response packages under uncertainty.

Recent literature helps clarify this shift. Work on contextual optimization formulates the problem as learning decision rules that map observed information, such as forecasts or local conditions, to actions, rather than selecting one action as universally optimal [564]. Multistage disaster-planning models similarly show why this matters by comparing adaptive policies, which can change as new information arrives, with static plans fixed in advance and rolling-horizon approaches that update only incrementally; their central insight is that flexibility can materially improve performance when uncertainty resolves over time [565]. Adaptation research makes a related point from a different angle, arguing that decision-makers should map the feasible solution space across scenarios, timescales, and measures, rather than assume that one warning signal should correspond to one response [566]. Evidence from forecast-based drought interventions in East Africa further shows why a portfolio perspective is needed: the value of action depends strongly on the hazard, the geography, and the forecast horizon, and in many cases interventions generate the greatest benefit closer to impact, when forecast skill is higher, even if earlier action can still provide some value [166, 567].

More importantly, these response rules should not be treated as fixed protocol choices, but as hypotheses to be tested and revised over time. The question after each season is not only whether a trigger was activated, but whether the underlying rule produced timely and proportionate action, or whether a different rule would have performed better under the realized conditions. This could involve adjusting thresholds, changing lead times, revising the sequence in which instruments are introduced, or redesigning which measures are bundled together under different forecast states. In this sense, the simulation framework developed in Chapter 5 is valuable not because it can optimize every activation in real time, but because it provides a disciplined testbed for comparing candidate policy rules before the next season begins. Governments could use such a framework to maintain a small policy playbook and periodically revalidate it against hindcasts, simulated shocks, false alarms, missed events, and operational lessons from previous activations.

This interpretation is consistent with recent work on anticipatory action and warning systems. OCHA's synthesis of anticipatory action pilots argues that historical forecasts are essential for estimating how often triggers correctly and incorrectly recommend action, that trigger performance should be evaluated after each risk season, and that phased approaches can link signals over time to different activities [31]. Work on drought triggers in Mozambique likewise shows that trigger design is fundamentally a problem of balancing hit rates, false alarms, lead times, and decision scales, and recommends refining triggers and investing in monitoring, evaluation, and learning to improve future operations [563, 165]. Related research on impact-based forecasting highlights the need for explicit evaluation methods and dynamic exposure and vulnerability data [568], while iterative risk-governance research emphasizes repeated co-production and refinement of fit-for-purpose solutions as conditions, risks, and institutional arrangements evolve [569]. The objective, then, is not to identify one perfect trigger once and for all, but to establish a disciplined process for testing, comparing, and revising a small portfolio of response rules over time.

6.4.5 Linking farmer communication pathways

A further question concerns how the decision-support framework developed in this thesis could be connected to farmers themselves. The architecture proposed here is primarily designed for public decision-makers, such as agricultural departments, disaster management authorities, and humanitarian actors, who must interpret forecasts and coordinate anticipatory responses. This institutional orientation is intentional. Early warning and anticipatory action systems operate through the decisions of governments and implementing organisations, which determine when to trigger interventions, how resources are allocated, and which policy instruments are deployed. In that sense, the framework developed in this thesis should be understood first and foremost as a decision-support tool for institutions, rather than as a direct farmer-facing technology.

Nevertheless, the functioning of anticipatory action systems will in many cases require some form of communication with farmers themselves. Certain types of interventions (such as early harvesting advice, input distribution, temporary cash transfers, or targeted advisories) may depend on farmers receiving timely information or notifications. The analytical outputs produced by the system may therefore need to be translated into farmer-facing communication channels when an intervention is activated. Research on agricultural mobile advisory services suggests that the effectiveness of such communication depends not only on the accuracy of the underlying information, but also on trust, perceived usefulness, connectivity, and the presence of social support networks such as advisers, peers, or family members [570, 571]. For this reason, a realistic implementation architecture may involve a two-layer structure, in which the logic underpinning allocation mechanisms remains institution-facing while farmers receive simplified notifications through channels such as SMS alerts, voice calls, extension visits, or messaging platforms.

That in turn implies a substantial qualitative research agenda. Future work should include co-design with farmers, field-based usability testing, experiments in communicating uncertainty, and attention to who within the household actually receives, understands, and acts upon notifications. Recent participatory design work in Maharashtra shows that trust and perceived ease of use are decisive for adoption, and that multilingual and audio-visual features can materially improve usability [571]. Survey evidence from South India similarly suggests that familiarity, interest, and perceived economic benefits matter at least as much as classical usability variables [572]. Evidence from existing mobile advisory systems in India also warns that women, elderly farmers, and smallholders may engage less unless message design explicitly accounts for unequal access to phones, literacy, and control over devices [573].

6.5 Avenues for further research in the near-term

In the near term, an important research priority is to strengthen the link between forecasting and decision-making *under uncertainty*. This includes moving from static, one-off optimization to repeated decision cycles in which policies are evaluated and refined over time. Embedding learning into the architecture, so that models update not only forecasts but also allocation rules based on past activations, would make anticipatory systems more adaptive and credible. Closely related is the need to better quantify how forecast uncertainty affects allocation outcomes, including robustness to miscalibration and the trade-offs between speed, caution, and welfare gains. While extensions to new crops and regions are important for external validation, advancing the way uncertainty is handled within the forecast-to-decision pipeline is likely to generate the largest conceptual and practical returns.

A second priority is scaling the architecture across institutional contexts while preserving transparency and usability. This means testing the framework beyond rice and India, and integrating it more closely with real budgeting systems, contingency funds, and operational planning tools. Research could explore portfolio-based policy design, where governments compare multiple anticipatory instruments such as cash transfers, input subsidies, and insurance triggers within a shared allocation framework. Other extensions include multi-hazard settings, cross-sector coordination, and alignment with climate adaptation pathways. The goal is not to expand the model in every direction, but to focus on extensions that make anticipatory optimization durable, transferable, and embedded within routine public decision-making.

6.6 Concluding remarks

As global hazards grow more frequent and complex, the gap between knowing a crisis is coming and acting in time remains one of the most persistent challenges in disaster policy. This thesis has shown that progress in forecasting must be matched by equally deliberate advances in decision design if early warnings are to lead to early action rather than delayed response. Around the world, governments and international partners are inching toward this integration. For example, recent national dialogues on anticipatory action in Nepal and new legislative frameworks in the Philippines signal growing recognition that forecast-informed decision-making cannot remain a niche pilot but must be embedded in public systems [574]. At the same time, broad initiatives such as the IFRC's global work on anticipatory action and scaling efforts across regions attest both to the promise and the unfinished business of this agenda [575].

This thesis contributes not only a set of models, but a framework for embedding early warning within real-world decision-making systems. Effective anticipatory action requires reconciling technical uncertainty, legal authority, budgetary constraints, and public trust. The real test for the next phase of research and practice will be whether these components can be brought together into systems that reliably convert prediction into timely, equitable, and funded action.

A Appendices

A.1 Studies included in the systematic review

Table A.1: Summary of selected research articles on rice yield prediction. A total of 156 articles were included in the review, covering various methodologies and data sources.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Analysis and forecasting of Australian rice yield using phenology-based aggregation of satellite and weather data	James Brinkhoff	2024	Agricultural and Forest Meteorology	Australia	[311]
The effect of dataset construction and data pre-processing on the eXtreme Gradient Boosting algorithm applied to head rice yield prediction in Australia	A. Clarke	2024	Computers and Electronics in Agriculture	Australia	[576]
Integrating Climate and Satellite Data for Multi-Temporal Pre-Harvest Prediction of Head Rice Yield in Australia	Clarke A.	2024	Remote Sensing	Australia	[576]
Rice yield responses in Bangladesh to large-scale atmospheric oscillation using multifactorial model	Bonosri Ghose	2021	Theoretical and Applied Climatology	Bangladesh	[307]
Artificial neural network model in predicting yield of mechanically transplanted rice from transplanting parameters in Bangladesh	Md Samiul Basir	2021	Journal of Agriculture and Food Research	Bangladesh	[305]
Forecasting the Impact of Climate Change on Rice Crop Yields under RCP4.5 and RCP8.5 Scenarios in Central Luzon, Philippines, Using Machine Learning Algorithms.	Rizza G. Baltazar	2024	Int. J. Agric. Nat. Resour.	Philippines	[577]
Ensemble of Machine Learning Algorithms for Rice Grain Yield Prediction Using UAV-Based Remote Sensing	Tapash Kumar Sarkar	2024	Journal of Biosystems Engineering	South Korea	[578]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
A novel machine learning approach for rice yield estimation	Surabhi Lingwal	2024	Journal of Experimental and Theoretical Artificial Intelligence	India	[579]
Securing China's rice harvest: unveiling dominant factors in production using multi-source data and hybrid machine learning models	Ali Mokhtar	2024	Scientific Reports	China	[310]
Remote sensing-based paddy yield estimation using physical and FCNN deep learning models in Gilan province, Iran	Ehsan Asmar	2024	Remote Sensing Applications: Society and Environment	Iran	[309]
An Innovative Method for Paddy Yield Prediction Based on DCNN-ELM Approach	Dr. Mohd. Asif Gandhi	2024	Proceedings of the 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT-2024)	India	[580]
A rotated rice spike detection model and a crop yield estimation application based on UAV images	Yunting Liang	2024	Computers and Electronics in Agriculture	China	[581]
Data-Driven Rice Yield Predictions and Prescriptive Analytics for Sustainable Agriculture in Malaysia	Muhammad Marong	2024	International Journal of Advanced Computer Science and Applications	Malaysia	[320]
New approach for forecasting rice and corn production in the Philippines through machine learning models	Samuel John Estenor Parreño	2024	Multidisciplinary Science Journal	Philippines	[342]
Rice Yield Prediction Based on Deep Learning	Xuning Chang	2024	Artificial Intelligence Technologies and Applications	China	[582]
Artificial Bee Colony Algorithm-based Feature Selection and Hybrid ML Framework for Efficient Rice Yield Prediction	Manasa Chitradurga Manjunath	2024	International Journal of Electrical and Computer Engineering Systems	India	[583]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Hybrid Deep Neural Networks with Multi-Tasking for Rice Yield Prediction Using Remote Sensing Data	Che-Hao Chang	2024	Agriculture	Taiwan	[584]
Phenological analysis and yield estimation of rice based on multi-spectral and SAR data in Maha Sarakham, Thailand	Tingyan Fu	2024	Journal of Spatial Science	Thailand	[308]
Rice Yield Estimation Using Multi-Temporal Remote Sensing Data and Machine Learning: A Case Study of Jiangsu, China	Liu	2024	Agriculture	China	[585]
Combining satellite data and artificial intelligence with a crop growth model to enhance rice yield estimation and crop management practices	Nguyen-Thanh Son	2024	Applied Geomatics	Taiwan	[586]
Comparative analysis of SMLR, ANN, Elastic net and LASSO based models for rice crop yield prediction in Uttarakhand	Parul Setiya	2024	MAUSAM	India	[587]
Application of UAV-Borne Visible-Infared Pushbroom Imaging Hyperspectral for Rice Yield Estimation Using Feature Selection Regression Methods	Shen Y.	2024	Sustainability	China	[588]
Enhancing direct-seeded rice yield prediction using UAV-derived features acquired during the reproductive phase	Guodong Yang	2024	Precision Agriculture	China	[589]
Paddy yield prediction based on 2D images of rice panicles using regression techniques	Pankaj	2024	The Visual Computer	China, India	[319]
Multi-Objective Optimization with Artificial Neural Network Based Robust Paddy Yield Prediction Model	S. Muthukumaran	2023	Intelligent Automation and Soft Computing	India	[590]
Multimodal Deep Learning for Rice Yield Prediction Using UAV-Based Multispectral Imagery and Weather Data	Mia	2023	Remote Sensing	Japan	[306]
Rice Yield Estimations Based on Transformed Surface Reflectance from Orbital Hyperspectral Remote Sensing	Miphokasap P.	2023	Chemical Engineering Transactions	Thailand	[591]
Rice Yield Prediction Using Sentinel-1 Radar Vegetation Indices and XGBoost	Chunling Sun	2023	IEEE SAR in Big Data Era (BIGSARDATA)	China	[592]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Research on a Hyperspectral Rice Yield Estimation Model Based on Random Forest	Xiaopan Wang	2023	2023 3rd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)	China	[593]
Smart Digital-Twin hub Concept for Rice yield prediction and monitoring from multivariate time series data	Yahjeb Bouha Khatraty	2023	2023 24th International Conference on Control Systems and Computer Science (CSCS)	Mauritania	[594]
Paddy Yield Forecasting using Regression Techniques	Chandrakumar T	2023	IEEE Delhi Section Flagship Conference (DELCON)	India	[595]
Rice yield prediction using Bayesian analysis on rainfed lands in the Sumbing-Sindoro Toposequence, Indonesia	Abdul Aziz	2023	Scientific Horizons	Indonesia	[341]
IoT-Enabled Smart Solution for Rice Disease Detection, Yield Prediction, and Remediation	Wanninayake K.M.I.S	2023	5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)	Sri Lanka	[596]
Responsible AI for Sustainable Agriculture: Forecasting Rice Yield to Combat Global Food Insecurity	Rahul Bhattacharyya	2023	14th International Conference on Computing Communication and Networking Technologies (ICCCNT)	India	[597]
Ensemble Deep Learning Algorithm for Forecasting of Rice Crop Yield based on Soil Nutrition Levels	M. Chandraprabha	2023	EAI Endorsed Transactions on Scalable Information Systems	India	[598]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Deep Learning Enables Instant and Versatile Estimation of Rice Yield Using Ground-Based RGB Images	Yu Tanaka	2023	Plant Phenomics	Côte d'Ivoire, Senegal, Japan, Kenya, Madagascar, Nigeria, Tanzania	[318]
Multi-Agent System Based on Stacking Technique for Rice Yield Prediction	Hayam R. Seireg	2023	3rd IEEE International Conference on Electronic Engineering	Egypt	[599]
Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularized Extreme Learning Machine	Mr. S. Thirumal	2023	Proceedings of the 7th International Conference on Intelligent Computing and Control Systems (ICICCS-2023)	India	[600]
Rice Yield Prediction in Sumatra Indonesia Using Machine Learning and Climate Data	Aldian Nurcahyo	2023	2023 3rd International Conference on Intelligent Cybernetics Technology and Applications (ICICyTA)	Indonesia	[601]
2023 10th International Conference on ICT for Smart Society (ICISS)	Erio Yoshino	2023	2023 10th International Conference on ICT for Smart Society (ICISS)	Indonesia	[602]
Automated Hyperparameter Tuned Stacked Autoencoder based Rice Crop Yield Prediction Model	Mr. S. Thirumal	2023	Proceedings of the 7th International Conference on Trends in Electronics and Informatics (ICOEI 2023)	India	[603]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Prediction of Paddy Yield based on IoT Data using GRU Model in Lowland Coastal Regions	Arepalli Peda Gopi	2023	Proceedings of the 5th International Conference on Smart Systems and Inventive Technology (ICSSIT 2023)	India	[604]
Comparison of Machine Learning Algorithms for the prediction of Rice Crop Yield in Karnataka	Dr. Senthil. S	2023	Fourth International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)	India	[605]
Weather-Based Statistical and Neural Network Tools for Forecasting Rice Yields in Major Growing Districts of Karnataka	Thimmegowda M.N.	2023	Agronomy	India	[606]
Rice Yield Prediction in Hubei Province Based on Deep Learning and the Effect of Spatial Heterogeneity	Zhou S.	2023	Remote Sensing	China	[317]
SSE-Based Evolutionary Algorithm for Hyper-parameter Optimization of LightGBM on Paddy Rice Yield Prediction Problem	Ayana Takai	2023	2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)	Japan	[607]
Improved information dissemination services for the agricultural sector in Thailand: development and evaluation of a machine learning based rice crop yield prediction system	Sumanya Ngandee	2023	Information Development	Thailand	[608]
Rice Yield Modeling Using Machine Learning Algorithms Based on Environmental and Agronomic Data of Pampanga River Basin, Philippines	V. David	2023	Universal Journal of Agricultural Research	Philippines	[609]
Comparative Analysis of Statistical and Machine Learning Techniques for Rice Yield Forecasting for Chhattisgarh, India	Satpathi A.	2023	Sustainability	India	[610]
Rapid Rice Yield Estimation Using Integrated Remote Sensing and Meteorological Data and Machine Learning	Islam	2023	Remote Sensing	Nepal	[304]
A tree based eXtreme Gradient Boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh	Mst Noorunnahar	2023	PLOS ONE	Bangladesh	[611]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Ensemble Feature Selection Framework for Paddy Yield Prediction in Cauvery Basin using Machine Learning Classifiers	P Sathya	2023	Cogent Engineering	India	[313]
Rice Yield Estimation Based on Machine Learning Approaches using MODIS 250 m Data	Woranan Mongkolnithithada	2023	Engineering Access	Thailand	[612]
Teaching and Learning based Optimization with Deep Learning Model for Rice Crop Yield Prediction	S. Thirumal	2023	SSRG International Journal of Electrical and Electronics Engineering	India	[613]
Rice Yield Production Forecasting using Deep Learning Models	Vincent Junior Halim	2023	2023 International Conference on Networking, Electrical Engineering, Computer Science, and Technology (IConNECT)	Indonesia	[614]
A robust rice yield estimation framework developed by grading modeling and normalized weight decision-making strategy using UAV imaging technology	Rui Yang	2023	Computers and Electronics in Agriculture	China	[589]
Prediction of Rice Yield Based on Multi-Source Data and Hybrid LSSVM Algorithms in China	Long Zhao	2023	International Journal of Plant Production	China	[615]
Temporal convolutional network based rice crop yield prediction using multispectral satellite data	Alkha Mohan	2023	Infrared Physics and Technology	India	[616]
Paddy Yield Prediction in Tamilnadu Delta Region Using MLR-LSTM Model	Sathya P	2023	Applied Artificial Intelligence	India	[617]
ICT for Intelligent Systems	Sakshi Gandotra	2023	ICT for Intelligent Systems, Smart Innovation, Systems and Technologies	India	[618]
Improved prediction of rice yield at field and county levels by synergistic use of SAR, optical and meteorological data	Weiguo Yu	2023	Agricultural and Forest Meteorology	China	[312]
Predicting rice yield based on weather variables using multiple linear, neural networks, and penalized regression models	Parul Setiya	2023	Theoretical and Applied Climatology	India	[619]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Rice yield prediction model using normalized vegetation and water indices from Sentinel-2A satellite imagery datasets	Aung Myint Htun	2023	Asia-Pacific Journal of Regional Science	Myanmar	[620]
Rice variety classification and yield prediction using semantic segmentation of agro-morphological characteristics	Bharati Patel	2023	Multimedia Tools and Applications	India	[621]
Comparison of yield prediction models and estimation of the relative importance of main agronomic traits affecting rice yield formation in saline-sodic paddy fields	Baishun Liu	2023	European Journal of Agronomy	China	[622]
Improved Yield Prediction of Ratoon Rice Using Unmanned Aerial Vehicle-Based Multi-Temporal Feature Method	Zhou Longfei	2023	Rice Science	China	[623]
Impacts of meteorological variables and machine learning algorithms on rice yield prediction in Korea	Subin Ha	2023	Int. J. Biometeorol.	South Korea	[624]
Research on Rice Yield Prediction Model Based on Deep Learning	Xiao Han	2022	Computational Intelligence and Neuroscience	China	[625]
Assessing the Potentials of Multi-temporal Sentinel-1 SAR Data for Paddy Yield Forecasting Using Artificial Neural Network	Pavan Kumar Sharma	2022	Journal of the Indian Society of Remote Sensing	India	[626]
Comparative Evaluation of Neural Networks in Crop Yield Prediction of Paddy and Sugarcane Crop	K. Krupavathi	2022	The Digital Agricultural Revolution	India	[627]
Field-scale rice yield prediction from Sentinel-2 monthly image composites using machine learning algorithms	Nguyen-Thanh Son	2022	Ecological Informatics	Taiwan	[628]
Remote Sensing Time Series Analysis for Early Rice Yield Forecasting Using Random Forest Algorithm	Nguyen-Thanh Son	2022	Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries	Taiwan	[339]
Transferability of Models for Predicting Rice Grain Yield from Unmanned Aerial Vehicle (UAV) Multispectral Imagery across Years, Cultivars and Sensors	Zheng H.	2022	Drones	China	[340]
Random Forest for rice yield mapping and prediction using Sentinel-2 data with Google Earth Engine	K. Choudhary	2022	Advances in Space Research	China	[629]
Rice Yield Estimation Based on Continuous Wavelet Transform With Multiple Growth Periods	Gu C	2022	Frontiers in Plant Science	China	[630]

Table A.1 continued from previous page.

Title	Author	Date	Journal	Country	Cit.
Estimating Yield-Related Traits Using UAV-Derived Multispectral Images to Improve Rice Grain Yield Prediction	Bascon M.V.	2022	Agriculture	Japan	[631]
Rice yield estimation using a CNN-based image-driven data assimilation framework	Jingye Han	2022	Field Crops Research	China	[632]
Rice Yield Prediction and Model Interpretation Based on Satellite and Climatic Indicators Using a Transformer Method	Liu	2022	Remote Sensing	India	[633]
Prediction models on biomass and yield of rice affected by metal (oxide) nanoparticles using nano-specific descriptors	Jing Li	2022	NanoImpact	China	[634]
Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries	Dharmesh Verma	2022	Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries	India	[635]
2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)	Astrid Noviana Paradhita	2022	2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)	Indonesia	[636]
Rice Yield Estimation Using Deep Learning	Niyati Mishra	2022	CCIS	India	[637]
A Time-Series Based Yield Forecasting Model Using Stacked Lstm To Predict The Yield Of Paddy In Cauvery Delta Zone In Tamilnadu	M. Geetha	2022	First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)	India	[638]
Deep Learning Based Yield Prediction Model To Predict The Yield of Paddy In Cauvery Delta Region	M. Geetha	2022	2022 International Conference on Computer Communication and Informatics (ICCCI)	India	[639]
Estimation and Forecasting of Rice Yield Using Phenology-Based Algorithm and Linear Regression Model on Sentinel-II Satellite Data	Abid Nazir	2021	Agriculture	Pakistan	[640]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches	Juan Cao	2021	Agricultural and Forest Meteorology	China	[641]
Two-Stage Spatiotemporal Time Series Modelling Approach for Rice Yield Prediction and Advanced Agroecosystem Management	Santosha Rathod	2021	Agronomy	India	[642]
Rice Yield Forecasting in West Bengal Using Hybrid Model	Aishika Banik	2021	Lecture Notes in Networks and Systems	India	[643]
Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea	Seungtaek Jeong	2021	Science of the Total Environment	South Korea, North Korea	[644]
Remote Sensing Based Yield Estimation of Rice (Oryza Sativa L.) Using Gradient Boosted Regression in India	Arumugam P.	2021	Remote Sensing	India	[645]
Prediction of Rice Yield in East China Based on Climate and Agronomic Traits Data Using Artificial Neural Networks and Partial Least Squares Regression	Guo Y.	2021	Agronomy	China	[646]
Use of regression techniques for rice yield estimation in the North-Western province of Sri Lanka	E.M.P. Ekanayake	2021	Ceylon Journal of Science	Sri Lanka	[647]
A Classifier Ensemble Approach for Prediction of Rice Yield based on Climatic Variability for Coastal Odisha Region of India	Subhadra Mishra	2021	Informatica	India	[648]
A Novel Borda Count based Feature Ranking and Feature Fusion Strategy to Attain Effective Climatic Features for Rice Yield Prediction	Subhadra Mishra	2021	Informatica	India	[649]
Rice Yield Simulation and Planting Suitability Environment Pattern Recognition at a Fine Scale	Li	2021	ISPRS International Journal of Geo-Information	China	[650]
Prediction of rice yield based on LSTM long short term memory network	Haoming Mo	2021	Journal of Physics: Conference Series	China	[651]
Rice-Yield Prediction with Multi-Temporal Sentinel-2 Data and 3D CNN: A Case Study in Nepal	Ruben Fernandez-Beltran	2021	Remote Sensing	Nepal	[652]
Assessment of Regression Models for Predicting Rice Yield and Protein Content Using Unmanned Aerial Vehicle-Based Multispectral Imagery	Kang Y.	2021	Remote Sensing	South Korea	[653]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
A Deep Learning Approach for Yield Estimation and Phenotype Analysis in Rice Crops	Dr. P. Devaki	2021	International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)	India	[654]
Exploration of Machine Learning Approaches for Paddy Yield Prediction in Eastern Part of Tamilnadu	Vinson Joshua	2021	Agronomy	India	[655]
Within-Field Rice Yield Estimation Based on Sentinel-2 Satellite Data	Franch B.	2021	Remote Sensing	Spain	[656]
Neural Network Autoregression And Classical Time Series Approaches For Rice Yield Forecasting	Bhardwaj N.	2021	The Journal of Animal and Plant Sciences	India	[657]
Estimation of Normal Rice Yield Considering Heading Stage Based on Observation Data and Satellite Imagery	Yuki Sofue	2021	2021 IEEE International Geoscience and Remote Sensing Symposium	Indonesia	[658]
Artificial neural network model in predicting yield of mechanically transplanted rice from transplanting parameters in Bangladesh	Md Samiul Basir	2021	Journal of Agriculture and Food Research	Bangladesh	[305]
Comparation of rice yield estimation model combining spectral index screening method and statistical regression algorithm	Wang Yaomin	2021	Transactions of the Chinese Society of Agricultural Engineering	China	[659]
Development of remote sensing-based yield prediction models at the maturity stage of boro rice using parametric and nonparametric approaches	Md. Monirul Islam	2021	Remote Sensing Applications: Society and Environment	Bangladesh	[660]
Grain Yield Estimation in Rice Breeding Using Phenological Data and Vegetation Indices Derived from UAV Images	Ge	2021	Agronomy	China	[661]
Modeling the Relationship between Rice Yield and Climate Variables Using Statistical and Machine Learning Techniques	Lasini Wickramasinghe	2021	Journal of Mathematics	Sri Lanka	[662]
Machine Learning Modelling of the Relationship between Weather and Paddy Yield in Sri Lanka	Piyal Ekanayake	2021	Journal of Mathematics	Sri Lanka	[314]
Genomic Prediction of Arsenic Tolerance and Grain Yield in Rice: Contribution of Trait-Specific Markers and Multi-Environment Models	Nourollah Ahmadi	2021	Rice Science	France, Scotland	[663]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Integrated phenology and climate in rice yields prediction using machine learning methods	Yahui Guo	2021	Ecological Indicators	China	[646]
Rice yield responses in Bangladesh to large-scale atmospheric oscillation using multifactorial model	Bonosri Ghose	2021	Theoretical and Applied Climatology	Bangladesh	[307]
Prediction of Rice Production in India Using Artificial Neural Network with Genetic Algorithm	Surjeet Kumar	2020	2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)	India	[664]
RiceSAP: An Efficient Satellite-Based AquaCrop Platform for Rice Crop Monitoring and Yield Prediction on a Farm- to Regional-Scale	Watcharee Veerakachen	2020	Agronomy	Thailand	[665]
Integrated Application of Remote Sensing and GIS in Crop Information System—A Case Study on Aman Rice Production Forecasting Using MODIS-NDVI in Bangladesh	B. M. Refat Faisal	2020	AgriEngineering	Bangladesh	[666]
Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan	Nguyen-Thanh Son	2020	International Journal of Remote Sensing	Taiwan	[667]
Yield prediction model of rice and wheat crops based on ecological distance algorithm	Li Tian	2020	Environmental Technology and Innovation	China	[668]
Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer – a case study of small farmlands in the South of China	Liang Wan	2020	Agricultural and Forest Meteorology	China	[669]
Prediction of Rice Yield via Stacked LSTM	Xiangyan Meng	2020	International Journal of Agricultural and Environmental Information Systems	China	[670]
Application of phenology-based algorithm and linear regression model for estimating rice cultivated areas and yield using remote sensing data in Bansloi River Basin, Eastern India	Gopal Chandra Paul	2020	Remote Sensing Applications: Society and Environment	India	[671]
Rice yield response forecasting tool (YIELDCAST) for supporting climate change adaptation decision in Sahel	Seydou Traore	2020	Agricultural Water Management	Burkina Faso	[672]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Incorporation of parental phenotypic data into multi-omic models improves prediction of yield-related traits in hybrid rice	Yang Xu	2020	Plant Biotechnology Journal	China, USA, Mexico	[673]
Improving the UAV-based yield estimation of paddy rice by using the solar radiation of geostationary satellite Himawari-8	Akira Hama	2020	Hydrological Research Letters	Japan	[674]
Artificial Neural Network to Estimate the Paddy Yield Prediction Using Climatic Data	Vinushi Amaratunga	2020	Mathematical Problems in Engineering	Sri Lanka	[675]
Artificial neural network to estimate the paddy yield prediction using remote sensing, weather and non weather variable in Ampara district, Sri Lanka	W. M. R. K Wanninayaka	2020	2020 5th International Conference on Information Technology Research (ICITR)	Sri Lanka	[676]
An end-to-end model for rice yield prediction using deep learning fusion	Zheng Chu	2020	Computers and Electronics in Agriculture	China	[677]
Boro Rice Yield Estimation Model Using Modis Ndvi Data for Bangladesh	Md. Samiul Alam	2019	IEEE International Geoscience and Remote Sensing Symposium	Bangladesh	[678]
Machine learning approach for Kharif rice yield prediction integrating multi-temporal vegetation indices and weather and non-weather variables	Aditi Chandra	2019	International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences	India	[679]
Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India)	Avinash Kumar Ranjan	2019	Spatial Information Research	India	[680]
Yield Estimation of Paddy Rice Based on Satellite Imagery: Comparison of Global and Local Regression Models	Yi-Shiang Shiu	2019	Remote Sensing	Taiwan	[681]
Adaptive Neuro Fuzzy Inference System (ANFIS) approach for modeling paddy production data in Central Java	Tarno	2019	Journal of Physics: Conference Series	Indonesia	[682]
Remotely Sensed Boro Rice Production Forecasting Using MODIS-NDVI: A Bangladesh Perspective	B.M. Refat Faisal	2019	AgriEngineering	Bangladesh	[683]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Machine Learning Methodologies for Paddy Yield Estimation in India: a Case Study	Ranjini B Guruprasad	2019	2019 IEEE International Geoscience and Remote Sensing Symposium	India	[684]
Rice yield estimation at pixel scale using relative vegetation indices from unmanned aerial systems	Feilong Wang	2019	8th International Conference on Agro-Geoinformatics	China	[685]
Rapid yield prediction in paddy fields based on 2D image modelling of rice panicles	Sanqin Zhao	2019	Computers and Electronics in Agriculture	China	[686]
Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel	Lei Zhang	2019	Computers and Electronics in Agriculture	Burkina Faso	[687]
Combining UAV-based vegetation indices, canopy height and canopy coverage to improve rice yield prediction under different nitrogen levels	Liang Wan	2019	ASABE Annual International Meeting	China	[688]
Modelling and predicting wetland rice production using support vector regression	Muhammad Alkaff	2019	TELKOMNIKA	Indonesia	[689]
Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images	Qi Yang	2019	Field Crops Research	China	[690]
Plot-scale rice grain yield estimation using UAV-based remotely sensed images via CNN with time-invariant deep features decomposition	Qi Yang	2019	2019 IEEE International Geoscience and Remote Sensing Symposium	China	[691]
A Technique of Fuzzy C-Mean in Multiple Linear Regression Model toward Paddy Yield	Nur Syazwan Wahab	2018	Journal of Physics: Conference Series	Malaysia	[692]
Regional-scale rice-yield estimation using stacked auto-encoder with climatic and MODIS data: a case study of South Korea	Jong-Won Ma	2018	International Journal of Remote Sensing	South Korea	[693]
Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India	Bappa Das	2018	International Journal of Biometeorology	India	[694]
Integrated model for predicting rice yield with climate change	Jin-Ki Park	2018	International Agrophysics	South Korea	[695]
Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images	Reza M. N.	2018	Biosystems Engineering	South Korea	[696]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Automatic Rice Yield Estimation Using Image Processing Technique	Md Nasim Reza	2017	Intelligent Environments	South Korea	[697]
Energy flow modeling and predicting the yield of Iranian paddy cultivars using artificial neural networks	Alireza Taheri-Rad	2017	Energy	Iran	[698]
A robust and novel regression based fuzzy time series algorithm for prediction of rice yield	Anshul Garg	2017	2017 International Conference on Intelligent Communication and Computational Techniques (ICCT)	India	[699]
Rice crop yield prediction in India using support vector machines	Niketa Gandhi	2016	2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)	India	[700]
Rice crop yield prediction using artificial neural networks	Niketa Gandhi	2016	IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2016)	India	[701]
Metabolomic prediction of yield in hybrid rice	Shizhong Xu	2016	The Plant Journal	United States of America, China	[702]
Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques	Niketa Gandhi	2016	2016 IEEE International Conference on Advances in Computer Applications (ICACA)	India	[703]
An efficient rice yield predictive model in lower northern Thailand	Jaratsri Rungrattanaubol	2016	AIP Conference Proceedings	Thailand	[704]

Table A.1 continued from previous page.

Title	**Author**	**Date**	**Journal**	**Country**	**Cit.**
Rice Yield Estimation Using Below Cloud Remote Sensing Images Acquired by Unmanned Airborne Vehicle System	Teoh C	2016	International Journal on Advanced Science, Engineering and Information Technology	Malaysia	[705]
Assessment of Multimodel Ensemble Seasonal Hindcasts for Satellite-Based Rice Yield Prediction	Jong Ahn Chun	2016	Journal of Agricultural Meteorology	Republic of Korea	[706]
Weed management through herbicide application in direct-seeded rice and yield modeling by artificial neural network	Dibakar Ghosh	2016	Spanish Journal of Agricultural Research	India	[707]
Development of a remote sensing-based rice yield forecasting model	Mostafa K. Mosleh	2016	Spanish Journal of Agricultural Research	Bangladesh	[708]
Rice yield estimation using Landsat ETM + Data	Altaf Ali Siyal	2015	Journal of Applied Remote Sensing	Pakistan	[709]

A.2 Author contributions by chapter (CRediT)

Several of the papers included in this thesis were the result of collaborative efforts with leading researchers in the field. To provide full transparency and formal attribution, detailed author contribution tables are presented below for each chapter, structured according to the CRediT taxonomy.

CRediT is an internationally recognised taxonomy developed to increase transparency in scholarly publishing. Instead of treating authorship as a simple list of names, it distinguishes between specific roles such as *Conceptualization*, *Methodology*, *Software*, *Validation*, *Formal analysis*, *Investigation*, *Data curation*, *Writing – original draft*, *Writing – review and editing*, *Visualization*, *Supervision*, *Project administration*, and *Funding acquisition* [710].

By reporting contributions in this structured way, readers and examiners can see more precisely who was responsible for which aspects of the research. This is especially important in multi-author projects where collaboration spans disciplines and institutions.

This doctoral research was also conducted *in the age of large language models*. Halfway through the degree, tools such as ChatGPT and specialized academic assistants (e.g. DeepResearch, Elicit) became widely available. As this was my second attempt at obtaining a doctoral degree (my first PhD was completed at Tsinghua University, without assistance from such language models), working with such tools presented both a fascinating new challenge and an opportunity. They were used primarily to accelerate literature review, helping locate obscure research papers (including those in other languages) and anticipatory action case study reports from around the world. LLMs were also employed for brainstorming and as a sparring partner in refining arguments, formulating models, and testing framing options. Their use was supplementary and interpretive; all final methodological, empirical, and theoretical contributions remain my own.

Chapter 2: Towards optimal anticipatory action

This chapter was co-authored by Djavan De Clercq, Lily Xu, Marleen C. de Ruiter, Marc van den Homberg, Marijn van der Velde, Jim W. Hall, Jonas Jaegermeyr, and Adam Mahdi. Djavan led the *Conceptualization*, *Formal Analysis*, *Investigation*, *Data Curation*, *Visualization*, and was primarily responsible for the *Writing – Original Draft*. Lily Xu, Marleen C. de Ruiter, Marc van den Homberg, Marijn van der Velde, Jim W. Hall, and Jonas Jaegermeyer contributed to *Conceptualization* and *Writing – Review*

& Editing. Adam Mahdi contributed through *Supervision* and *Writing – Review & Editing*. Large language models were employed to support the identification of emerging case studies on anticipatory action, many of which are not yet documented in peer-reviewed journals.

Chapter 3: Modern computational approaches for rice yield prediction

This chapter was co-authored by Djavan De Clercq and Adam Mahdi. Djavan led the *Conceptualization*, *Methodology*, *Software*, *Formal Analysis*, *Investigation*, *Data Curation*, and *Visualization*, and was primarily responsible for the *Writing – Original Draft*. Adam contributed through *Supervision* and *Writing – Review & Editing*, and provided input during *Conceptualization* and *Methodology*. Project administration was shared between both authors. Large language models were used to support literature search and brainstorming ideas.

Chapter 4: Feasibility of machine learning-based rice yield prediction in India

This chapter was co-authored by Djavan De Clercq and Adam Mahdi. Djavan led the *Conceptualization*, *Methodology*, *Software*, *Formal Analysis*, *Investigation*, *Data Curation*, and *Visualization*, and was primarily responsible for the *Writing – Original Draft*. Adam contributed through *Supervision* and *Writing – Review & Editing*, and provided input during *Conceptualization* and *Methodology*. Project administration was shared between both authors. Large language models were used to support python programming.

Chapter 5: Optimizing early warning-driven anticipatory action cash transfers

This chapter was co-authored by Djavan De Clercq, Burcu Balcik, Marleen C. de Ruiter, Sagar Surendra Deshmukh, Marc van den Homberg, and Adam Mahdi. Djavan led the *Conceptualization*, *Methodology*, *Formal Analysis*, *Investigation*, *Data Curation*, *Visualization*, and was primarily responsible for the *Writing – Original Draft*. Burcu Balcik, Marleen C. de Ruiter, Marc van den Homberg, and Sagar Surendra Deshmukh contributed to *Methodology* and *Writing – Review & Editing*. Adam Mahdi contributed through *Supervision* and *Writing – Review & Editing*. Large language models were employed to assist in reviewing grey literature and brainstorming model framing options.

References

---

[1] Food and Agriculture Organization of the United Nations. *The Impact of Disasters on Agriculture and Food Security 2023*. FAO, Edinburgh, Scotland, UK, 2023.

[2] Food and Agriculture Organization of the United Nations. Anticipatory action: Changing the way we manage disasters. `https://www.fao.org/emergencies/our-focus/anticipatory-action/en`, 2024. Accessed: 13 Feb 2026.

[3] Md. Mujahidul Islam, Mehedi Hasan, Md. Saju Mia, Abdullah Al Masud, and Abu Reza Md Towfiqul Islam. Early Warning Systems in Climate Risk Management: Roles and Implementations in Eradicating Barriers and Overcoming Challenges. *Natural Hazards Research*, January 2025.

[4] David Backer and Trey Billing. Validating famine early warning systems network projections of food security in Africa, 2009–2020. *Global Food Security*, 29:100510, June 2021.

[5] Marijn van der Velde and Lorenzo Nisini. Performance of the MARS-crop yield forecasting system for the European Union: Assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015. *Agricultural Systems*, 168:203–212, January 2019.

[6] Husain Najafi, Pallav Kumar Shrestha, Oldrich Rakovec, Heiko Apel, Sergiy Vorogushyn, Rohini Kumar, Stephan Thober, Bruno Merz, and Luis Samaniego. High-resolution impact-based early warning system for riverine flooding. *Nature Communications*, 15(3726):1–12, May 2024.

[7] Sanjay Srivastava, SungEun Kim, Lam Hoang, Jin Rui Yap, Jaehee Hwang, and Maria Bernadet Dewi. *Manual for Operationalizing Impact-based Forecasting and Warning Services (IBFWS)*. Viet Nam, October 2024. Developed with support from WMO (Climate Risk Early Warning Systems-Canada Fund).

[8] Gabrielle Smith, Lili Mohiddin, and Laura Phelps. *Targeting in Urban Displacement Contexts: Guidance Note for Humanitarian Practitioners*. London, UK, June 2017. Developed under the Stronger Cities Initiative with support from the International Rescue

Committee (IRC), Norwegian Refugee Council (NRC), and World Vision International (WVI).

[9] Paolo Verme and Chiara Gigliarano. Optimal targeting under budget constraints in a humanitarian context. *World Development*, 119:224–233, July 2019.

[10] United Nations Office for Disaster Risk Reduction. Global Assessment Report on Disaster Risk Reduction 2022: Our World at Risk: Transforming Governance for a Resilient Future, 2022.

[11] Food and Agriculture Organization of the United Nations. Gross domestic product and agriculture value added 2012–2021: Global and regional trends, 2023.

[12] Sravanthi Choragudi and Anirban Kundu. Safeguarding agricultural households against climate shocks. *Ideas for India*, September 2024. [Online; accessed 19 Sep. 2024].

[13] Stéphane Hallegatte. A cost effective solution to reduce disaster losses in developing countries: Hydro-meteorological services, early warning, and evacuation. *World Bank Policy Research Working Paper*, (6058), 2012.

[14] World Meteorological Organization. Early Warnings for All: The UN global early warning initiative for the implementation of climate adaptation, executive action plan 2023–2027. `https://library.wmo.int/idurl/4/58209`, 2022. 56 pages.

[15] World Meteorological Organization. Early warning systems reach new heights, but critical gaps jeopardize global progress, November 2025. Press release.

[16] Steffen Fritz, Linda See, Juan Carlos Laso Bayas, François Waldner, Damien Jacques, Inbal Becker-Reshef, Alyssa Whitcraft, Bettina Baruth, Rogerio Bonifacio, Jim Crutchfield, Felix Rembold, Oscar Rojas, Anne Schucknecht, Marijn van der Velde, James Verdin, Bingfang Wu, Nana Yan, Liangzhi You, Sven Gilliams, Sander Mücher, Robert Tetrault, Inian Moorthy, and Ian McCallum. A comparison of global agricultural monitoring systems and current gaps. *Agricultural Systems*, 168:258–272, January 2019.

[17] Clare Allen-Sader, William Thurston, Marcel Meyer, Elias Nure, Netsanet Bacha, Yoseph Alemayehu, Richard O. J. H. Stutt, Daniel Safka, Andrew P. Craig, Eshetu Derso, Laura E. Burgin, Sarah C. Millington, Matthew C. Hort, David P. Hodson, and Christopher A. Gilligan. An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11):115004, October 2019.

[18] Food and Agriculture Organization of the United Nations. After-action review of the early warning messaging activity under the scaling up early warning and anticipatory action for agriculture and food security project (EWAA) in Zimbabwe, August 2022. Published: 5 Aug 2022.

[19] Erin Lentz and Daniel G. Maxwell. Anticipating, Mitigating, and Responding to Crises: How Do Information Problems Constrain Action?, February 2022. [Online; accessed 13 Feb. 2025].

[20] Asia-Pacific Technical Working Group on Anticipatory Action and Asia-Pacific Regional Cash Working Group. Anticipatory action and cash transfers for rapid-onset hazards: Practitioners' note for field testing. Technical report, Food and Agriculture Organization of the United Nations (FAO) and International Federation of Red Cross and Red Crescent Societies (IFRC), Bangkok, 2022. Co-authored by Vlad Sorin Cozma and Catherine Jones.

[21] Food and Agriculture Organization of the United Nations. *Horn of Africa: Impact of early warning early action*. FAO, Rome, 2018.

[22] Food and Agriculture Organization of the United Nations. The Sudan: Impact of early warning early action. *FAO*, 2019.

[23] Food and Agriculture Organization of the United Nations. The impact of disasters on agriculture and food security 2023: Avoiding and reducing losses through investment in resilience, 2023.

[24] Food and Agriculture Organization of the United Nations. Madagascar: Impact of early warning early action. *FAO*, 2019.

[25] Food and Agriculture Organization of the United Nations. Mongolia: Impact of early warning early action. Rome, FAO, 2018. Licence: CC BY-NC-SA 3.0 IGO.

[26] Michèle Plichta and Lydia Poole. The State of Pre-Arranged Financing for Disasters 2024. Technical report, Centre for Disaster Protection, London, 2024. A comprehensive analysis of pre-arranged financing trends, instruments, and coverage for disaster risk management.

[27] Anticipation Hub. Anticipatory action in 2023: A global overview. Technical report, Anticipation Hub, Bonn, Germany, April 2023.

[28] United Nations Office for Disaster Risk Reduction and United Nations Office for the Coordination of Humanitarian Affairs. Briefing Note: Anticipatory Action – An Innovative Tool at the Intersection of Disaster Risk Reduction and Humanitarian Response. United Nations Office for Disaster Risk Reduction, Geneva, 2024. Available under the Creative Commons Attribution-Non-Commercial 3.0 IGO licence (CC BY-NC IGO).

[29] Djavan De Clercq, Lily Xu, Marleen C. de Ruiter, Marc van den Homberg, Marijn van der Velde, Jim W. Hall, Jonas Jaegermeyr, and Adam Mahdi. Towards optimal anticipatory action: Maximizing the effectiveness of agricultural early warning systems

with operations research. *International Journal of Disaster Risk Reduction*, 119:105249, March 2025.

[30] Erin Coughlan de Perez, Laura Harrison, Kristoffer Berse, Evan Easton-Calabria, Joalane Marunye, Makoala Marake, Sonia Binte Murshed, Shampa, and Erlich-Honest Zauisomue. Adapting to climate change through anticipatory action: The potential use of weather-based early warnings. *Weather and Climate Extremes*, 38:100508, December 2022.

[31] Juan Chaves-Gonzalez, Leonardo Milano, Dirk-Jan Omtzigt, Daniel Pfister, Josee Poirier, Ashley Pople, Julia Wittig, and Zinta Zommers. Anticipatory action: Lessons for the future. *Frontiers in Climate*, 4:932336, December 2022.

[32] P. Krishna Krishnamurthy, Richard J. Choularton, and Peter Kareiva. Dealing with uncertainty in famine predictions: How complex events affect food security early warning skill in the Greater Horn of Africa. *Global Food Security*, 26:100374, September 2020.

[33] Richard J. Choularton and P. Krishna Krishnamurthy. How accurate is food security early warning? Evaluation of FEWS NET accuracy in Ethiopia. *Food Security*, 11(2):333–344, April 2019.

[34] Katharina Hembach-Stunden, Tobias Vorlaufer, and Stefanie Engel. False and missed alarms in seasonal forecasts affect individual adaptation choices. *Q Open*, 4(1):qoad031, January 2024.

[35] World Meteorological Organization. WMO and the Early Warnings for All Initiative, August 2024. [Online; accessed 20 Feb. 2025].

[36] World Meteorological Organization (WMO). *Manual for Operationalizing Impact-based Forecasting and Warning Services (IBFWS)*, 2021. Developed to support the Viet Nam Meteorological and Hydrological Administration (VNMHA).

[37] Red Cross Red Crescent Climate Centre. *Impact-based Forecasting for Early Action: The Future of Forecasts*, 2021. Developed in collaboration with the UK Met Office and other partners.

[38] F. Sabo, M. Meroni, H. Kerdiles, P. Vojnovic, M. Piles, J. Munoz-Mari, A. Mateo Sanchis, and F. Rembold. Machine learning workflow for predicting national-level crop yield across 77 countries. Technical Note JRC137518, European Commission, Joint Research Centre, Ispra, Italy, 2024. Technical note on Large Scale Yield Forecasting v1.0.

[39] Catherine Jones, Nora Guerten, Marya Hillesland, and Elizabeth Koechlein. Applying an inclusive and equitable approach to anticipatory action. Technical report, FAO, Rome, 2020.

[40] Ana Lopez, Erin Coughlan de Perez, Juan Bazo, Pablo Suarez, Bart van den Hurk, and Marteen van Aalst. Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Weather and Climate Extremes*, 27:100167, March 2020.

[41] Josh Ayers, M. A. Oenone Chadburn, and Kazi Amdadul Hoque. Behaviourally-informed early warning and anticipatory action: Unleashing the potential of locally-led pre-emptive response and the forecast-based financing it requires. *Mavs Open Press*, August 2024.

[42] Niklas Rieger, Mike Pearson, Suzanna Nelson-Pollard, and Elise Belcher. Falling Short? Humanitarian Funding and Reform, October 2024. A comprehensive analysis of humanitarian funding trends and reform efforts.

[43] Erin Coughlan de Perez, Kristoffer B. Berse, Lianne Angelico C. Depante, Evan Easton-Calabria, Elton Pierre R. Evidente, Theodore Ezike, Dorothy Heinrich, Christopher Jack, Alfredo Mahar Francisco A. Lagmay, Selma Lendelvo, Joalane Marunye, Daniel G. Maxwell, Sonia Binte Murshed, Christopher Garimoi Orach, Mecthilde Pinto, Leah B. Poole, Komal Rathod, Shampa, and Carolyn Van Sant. Learning from the past in moving to the future: Invest in communication and response to weather early warnings to reduce death and damage. *Climate Risk Management*, 38:100461, January 2022.

[44] Zoë Scott. The future of Anticipatory Action: Four challenges to reaching scale and sustainability, August 2023. [Online; accessed 20 Feb. 2025].

[45] Food and Agriculture Organization of the United Nations. India at a glance, September 2024. [Online; accessed 15 Sep. 2024].

[46] Aditya R. Khanal and Ashok K. Mishra. Enhancing food security: Food crop portfolio choice in response to climatic risk in India. *Global Food Security*, 12:22–30, March 2017.

[47] Pratap S. Birthal and Jaweriah Hazrana. Crop diversification and resilience of agriculture to climatic shocks: Evidence from India. *Agricultural Systems*, 173:345–354, July 2019.

[48] Vimal Mishra, Kaustubh Thirumalai, Deepti Singh, and Saran Aadhar. Future exacerbation of hot and dry summer monsoon extremes in India. *npj Climate and Atmospheric Science*, 3(10):1–9, March 2020.

[49] Raju Guntukula. Assessing the impact of climate change on Indian agriculture: Evidence from major crop yields. *Journal of Public Affairs*, 20(1):e2040, February 2020.

[50] Rajendra Jadhav, Mayank Bhardwaj, and Nigam Prusty. India bans wheat exports as heat wave hurts crop, domestic prices soar. *Reuters*, May 2022.

[51] Pratik Parija. Scorching Heat in India to Deal New Blow to Global Wheat Supply. *Bloomberg*, May 2022.

[52] Haider Ali, Parth Modi, and Vimal Mishra. Increased flood risk in Indian sub-continent under the warming climate. *Weather and Climate Extremes*, 25:100212, September 2019.

[53] Sourav Mukherjee, Saran Aadhar, Daithi Stone, and Vimal Mishra. Increase in extreme precipitation events under anthropogenic warming in India. *Weather and Climate Extremes*, 20:45–53, June 2018.

[54] Sushil Pandey, Humnath Bhandari, Ramesh Sharan, Sudhir K. Taunk, Dibakar Naik, Preeda Prapertchob, and Shijun Ding. Economic costs of drought and rice farmers risk coping mechanisms: A cross country comparative analysis. *Agricultural and Applied Economics Association*, 2004.

[55] Sebastian Jülich. Drought triggered temporary migration in an East Indian village. *International Migration*, 49(s1):e189–e199, June 2011.

[56] Parmeshwar Udmale, Yutaka Ichikawa, Sujata Manandhar, Hiroshi Ishidaira, and Anthony S. Kiem. Farmers' perception of drought impacts, local adaptation and administrative mitigation measures in Maharashtra State, India. *International Journal of Disaster Risk Reduction*, 10:250–269, December 2014.

[57] Chandra Sekhar Bahinipati, Vijay Kumar, and P. K. Viswanathan. An evidence-based systematic review on farmers' adaptation strategies in India. *Food Security*, 13(2):399–418, April 2021.

[58] Tanvi Deshpande. Explained: Why India's early warning systems for floods and cyclones fall short. *IndiaSpend*, July 2022.

[59] Disha Shetty. India's investments in early warning systems have unintended benefits for women. *The Fuller Project*, February 2022. Accessed: February 25, 2022.

[60] Henrique M. D. Goulart, Karin van der Wiel, Christian Folberth, Juraj Balkovic, and Bart van den Hurk. Storylines of weather-induced crop failure events under climate change. *Earth System Dynamics*, 12(4):1503–1527, December 2021.

[61] Raed Hamed, Anne F. Van Loon, Jeroen Aerts, and Dim Coumou. Impacts of compound hot–dry extremes on US soybean yields. *Earth System Dynamics*, 12(4):1371–1391, November 2021.

[62] Alessia Matanó, Wouter R. Berghuijs, Maurizio Mazzoleni, Marleen C. de Ruiter, Philip J. Ward, and Anne F. Van Loon. Compound and consecutive drought-flood events at a global scale. *Environmental Research Letters*, 19(6):064048, May 2024.

[63] Red Cross Red Crescent Climate Centre. Dissolving disasters: Disaster risk reduction game kit, 2024. [Online; accessed 9 Oct. 2024].

[64] Asia-Pacific Technical Working Group on Anticipatory Action and Asia-Pacific Regional Cash Working Group. Anticipatory action and cash transfers for slow-onset hazards: Practitioners' note for field testing. Technical report, Food and Agriculture Organization of the United Nations (FAO) and International Federation of Red Cross and Red Crescent Societies (IFRC), Bangkok, 2024. Co-authored by Catherine Jones, Vlad Sorin Cozma, and Mulugeta L. Handino.

[65] Martin Plichta, Sam Cook, and Carlene Rhodes. The state of pre-arranged financing for disasters 2023, August 2023.

[66] European Humanitarian Forum. Addressing the global humanitarian funding gap, March 2024. Accessed: 2024-11-20.

[67] Koen Peters, Luk N. Van Wassenhove, Wim Van den Heuvel, and Ludo De Boeck. The nutritious supply chain: optimizing humanitarian food assistance. *INFORMS Journal on Optimization*, 3(2):200–226, 2021.

[68] Paul Ngosa Mboshya. The use of optimus in Madagascar and Ethiopia emergencies, 2024. Accessed: 2024-10-28.

[69] Sharmila Devi. Locust swarms in East Africa could be "a catastrophe". *The Lancet*, 395(10224):547, 2020.

[70] Abdi Latif Dahir. "Like an umbrella had covered the sky": Locust swarms despoil Kenya. *The New York Times*, February 2020.

[71] Rachel Nuwer. As locusts swarmed East Africa, this tech helped squash them. *The New York Times*, April 2021.

[72] Dominique Patton. Vegetables pricier than pork worry Chinese consumers as costs swell. *Reuters*, October 2021.

[73] Judith Evans. "It's the end of the road": UK sugar beet farmers face rising pressures. *Financial Times*, April 2021.

[74] Will Wright. A bitter wind at a shaky time, and Iowa is left reeling. *The New York Times*, October 2020.

[75] The Hindu Bureau. Tamil Nadu sees surge in extreme weather events in 2024: report. The Hindu, November 2024. Published by the Centre for Science and Environment and Down to Earth.

[76] Tim Searchinger, Richard Waite, Craig Hanson, Janet Ranganathan, and Emily Matthews. Creating a sustainable food future, July 2019.

[77] Deepak K. Ray, Nathaniel D. Mueller, Paul C. West, and Jonathan A. Foley. Yield trends are insufficient to double global crop production by 2050. *PLoS One*, 8(6):e66428, June 2013.

[78] Hugo Valin, Ronald D. Sands, Dominique van der Mensbrugghe, Gerald C. Nelson, Helal Ahammad, Elodie Blanc, Benjamin Bodirsky, Shinichiro Fujimori, Tomoko Hasegawa, Petr Havlik, Edwina Heyhoe, Page Kyle, Daniel Mason-D'Croz, Sergey Paltsev, Susanne Rolinski, Andrzej Tabeau, Hans van Meijl, Martin von Lampe, and Dirk Willenbockel. The future of food demand: Understanding differences in global economic models. *Agricultural Economics*, 45(1):51–67, January 2014.

[79] Rajendra P. Sishodia, Ram L. Ray, and Sudhir K. Singh. Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19):3136, September 2020.

[80] Julia Bailey-Serres, Jane E. Parker, Elizabeth A. Ainsworth, Giles E. D. Oldroyd, and Julian I. Schroeder. Genetic strategies for improving crop yields. *Nature*, 575(7781):109–118, November 2019.

[81] Food and Agriculture Organization of the United Nations. The future of food and agriculture – Alternative pathways to 2050, 2018. [Online; accessed 12 Sep. 2024].

[82] National Academies of Sciences, Engineering, and Medicine. *Science breakthroughs to advance food and agricultural research by 2030*. National Academies Press, July 2018.

[83] Deepak K. Ray, James S. Gerber, Graham K. MacDonald, and Paul C. West. Climate variation explains a third of global crop yield variability. *Nature Communications*, 6:5989, January 2015.

[84] Elisabeth Vogel, Markus G. Donat, Lisa V. Alexander, Malte Meinshausen, Deepak K. Ray, David Karoly, Nicolai Meinshausen, and Katja Frieler. The effects of climate extremes on global agricultural yields. *Environmental Research Letters*, 14(5):054010, May 2019.

[85] Yusuke Satoh, Kei Yoshimura, Yadu Pokhrel, Hyungjun Kim, Hideo Shiogama, Tokuta Yokohata, Naota Hanasaki, Yoshihide Wada, Peter Burek, Edward Byers, Hannes Muller Schmied, Dieter Gerten, Sebastian Ostberg, Simon N. Gosling, Julien Eric Stanslas Boulange, and Taikan Oki. The timing of unprecedented hydrological drought under climate change. *Nature Communications*, 13(1):3287, June 2022.

[86] Gustavo Naumann, Carmelo Cammalleri, Lorenzo Mentaschi, and Luc Feyen. Increased economic drought impacts in Europe with anthropogenic warming. *Nature Climate Change*, 11(6):485–491, June 2021.

[87] Vivek Gupta and Manoj Kumar Jain. Investigation of multi-model spatiotemporal mesoscale drought projections over India under climate change scenario. *Journal of Hydrology*, 567:489–509, December 2018.

[88] Gebremedhin Gebremeskel Haile, Qiuhong Tang, Seyed-Mohammad Hosseini-Moghari, Xingcai Liu, T. G. Gebremicael, Guoyong Leng, Asfaw Kebede, Ximeng Xu, and Xiaobo Yun. Projected impacts of climate change on drought patterns over East Africa. *Earth's Future*, 8(7):e2020EF001502, July 2020.

[89] Zhaoli Wang, Ruida Zhong, Chengguang Lai, Zhaoyang Zeng, Yanqing Lian, and Xiaoyan Bai. Climate change enhances the severity and variability of drought in the Pearl River Basin in South China in the 21st century. *Agricultural and Forest Meteorology*, 249:149–162, February 2018.

[90] Lei Xu, Nengcheng Chen, and Xiang Zhang. Global drought trends under 1.5 and 2 °C warming. *International Journal of Climatology*, 39(4):2375–2385, March 2019.

[91] Anne Schechinger. Crop losses from climate crisis cost billions of dollars in insurance payouts, September 2024. [Online; accessed 12 Sep. 2024].

[92] Noah S. Diffenbaugh, Frances V. Davenport, and Marshall Burke. Historical warming has increased U.S. crop insurance losses. *Environmental Research Letters*, 16(8):084025, July 2021.

[93] Edel Doherty, Sinead Mellett, Daniel Norton, Thomas K. J. McDermott, Denis O' Hora, and Mary Ryan. A discrete choice experiment exploring farmer preferences for insurance against extreme weather events. *Journal of Environmental Management*, 290:112607, July 2021.

[94] Wonsik Kim, Toshichika Iizumi, and Motoki Nishimori. Global patterns of crop production losses associated with droughts from 1983 to 2009. *Journal of Applied Meteorology and Climatology*, 58(6):1233–1244, June 2019.

[95] Pat Finnigan. Mental health: A priority for our farmers. *Report of the Standing Committee on Agriculture and Agri-food*, 8(99):19–22, 2019.

[96] Shenggen Fan and Christopher Rue. The role of smallholder farms in a changing world. In *The Role of Smallholder Farms in Food and Nutrition Security*, pages 13–28. Springer, Cham, Switzerland, May 2020.

[97] Sarah K. Lowder, Jakob Skoet, and Saumya Singh. What do we really know about the number and distribution of farms and family farms in the world? Background paper for The State of Food and Agriculture 2014. *AgEcon Search*, April 2014.

[98] George Rapsomanikis. The economic lives of smallholder farmers. *Food and Agriculture Organization of the United Nations*, 2015.

[99] Christine E. Fraser, Kathryn B. Smith, Frances Judd, John S. Humphreys, Lyn J. Fragar, and Annette Henderson. Farming and mental health problems and mental illness. *International Journal of Social Psychiatry*, 51(4):340–349, December 2005.

[100] Anthony Gregoire. The mental health of farmers. *Occupational Medicine*, 52(8):471–476, December 2002.

[101] Linda Price and Nick Evans. From stress to distress: Conceptualizing the British family farming patriarchal way of life. *Journal of Rural Studies*, 25(1):1–11, January 2009.

[102] Philippe Roy, Gilles Tremblay, John L. Oliffe, Jalila Jbilou, and Steve Robertson. Male farmers with mental health disorders: A scoping review. *Australian Journal of Rural Health*, 21(1):3–7, February 2013.

[103] Sahar Daghagh Yazd, Sarah Ann Wheeler, and Alec Zuo. Key risk factors affecting farmers' mental health: A systematic review. *International Journal of Environmental Research and Public Health*, 16(23):4849, December 2019.

[104] Brit Logstein. Farm-related concerns and mental health status among Norwegian farmers. *Journal of Agromedicine*, 21(4):316–326, 2016.

[105] Chipo Plaxedes Mubaya, Jemimah Njuki, Eness Paidamoyo Mutsvangwa, Francis Temba Mugabe, and Durton Nanja. Climate variability and change or multiple stressors? Farmer perceptions regarding threats to livelihoods in Zimbabwe and Zambia. *Journal of Environmental Management*, 102:9–17, July 2012.

[106] Huat Bin, Felicity Lamm, and Rupert Tipples. The impact of stressors on the psychological wellbeing of New Zealand farmers and the development of an explanatory conceptual model. *Figshare*, January 2008.

[107] James L. Walker, Lilly Schubert Walker, and Patricia M. MacLennan. An informal look at farm stress. *Psychological Reports*, 59(2):427–430, October 1986.

[108] Gary Raine. Causes and effects of stress on farmers: A qualitative study. *Health Education Journal*, 58(3):259–270, September 1999.

[109] Sue Simkin, Keith Hawton, Joan Fagg, and Anders Malmberg. Stress in farmers: A survey of farmers in England and Wales. *Occupational and Environmental Medicine*, 55(11):729–734, November 1998.

[110] Parichat Ong-Artborirak, Waraporn Boonchieng, Yuwayong Juntarawijit, and Chudchawal Juntarawijit. Potential effects on mental health status associated with occupational exposure to pesticides among Thai farmers. *International Journal of Environmental Research and Public Health*, 19(15):9654, August 2022.

[111] William A. Vega, George J. Warheit, and Rodolfo Palacio. Psychiatric symptomatology among Mexican American farmworkers. *Social Science & Medicine*, 20(1):39–45, 1985.

[112] Scott C. Carvajal, Clara Kibor, Deborah Jean McClelland, Maia Ingram, Jill Guernsey de Zapien, Emma Torres, Floribella Redondo, Kathryn Rodriguez, Raquel Rubio-Goldsmith, Joel Meister, and Cecilia Rosales. Stress and sociocultural factors related to health status among US–Mexico border farmworkers. *Journal of Immigrant and Minority Health*, 16(6):1176–1182, October 2014.

[113] Hilda M. Firth, Sheila M. Williams, G. Peter Herbison, and Robert O. McGee. Stress in New Zealand farmers. *Stress and Health*, 23(1):51–58, February 2007.

[114] Dana C. Mora, Sara A. Quandt, Haiying Chen, and Thomas A. Arcury. Associations of poor housing with mental health among North Carolina Latino migrant farmworkers. *Journal of Agromedicine*, 21(4):327, 2016.

[115] Suda Hanklang, Orawan Kaewboonchoo, Ikuharu Morioka, and Suwan-ampai Plernpit. Gender differences in depression symptoms among rice farmers in Thailand. *Asia-Pacific Journal of Public Health*, 28(1):83–93, December 2015.

[116] Joseph D. Hovey and Cristina G. Magaña. Psychosocial predictors of anxiety among immigrant Mexican migrant farmworkers: Implications for prevention and treatment. *Cultural Diversity and Ethnic Minority Psychology*, 8(3):274–289, August 2002.

[117] Athena K. Ramos, Dejun Su, Lina Lander, and Roy Rivera. Stress factors contributing to depression among Latino migrant farmworkers in Nebraska. *Journal of Immigrant and Minority Health*, 17(6):1627–1634, December 2015.

[118] Kate M. Gunn, Lisa J. Kettler, Gemma L. A. Skaczkowski, and Deborah A. Turnbull. Farmers' stress and coping in a time of drought. *Rural and Remote Health*, 12:2071, 2012.

[119] Sarah Ann Wheeler, Alec Zuo, and Adam Loch. Water torture: Unravelling the psychological distress of irrigators in Australia. *Journal of Rural Studies*, 62:183–194, August 2018.

[120] Florence Becot, Shoshanah Inwood, and Hannah Budge. "The source of all my joy and all my stress": Children and childcare as underappreciated sources of stress that affect farm women. *Journal of Agromedicine*, 30(1):114–131, January 2025.

[121] Delwar Hossain, Rob Eley, Jeff Coutts, and Don Gorman. Mental health of farmers in Southern Queensland: Issues and support. *Australian Journal of Rural Health*, 16(6):343–348, December 2008.

[122] Ann E. Hiott, Joseph G. Grzywacz, Stephen W. Davis, Sara A. Quandt, and Thomas A. Arcury. Migrant farmworker stress: Mental health implications. *Journal of Rural Health*, 24(1):32–39, 2008.

[123] Heather Schlesser, Jayne Krull, Katie Wantoch, Alana Voss, Stef Morrill, and Joy Kirk-Patrick. Wisconsin farmers share their stressors and mental healthcare needs through focus groups. *Journal of Agromedicine*, 29(4):626–635, October 2024.

[124] Ian J. Deary, Joyce Willock, and Murray McGregor. Stress in farming. *Stress Medicine*, 13(2):131–136, April 1997.

[125] Dalva Hedlund and Alan Berkowitz. The incidence of social-psychological stress in farm families. *International Journal of Sociology of the Family*, pages 233–243, 1979.

[126] Joseph G. Grzywacz, Sara A. Quandt, Haiying Chen, Scott Isom, Lisa Kiang, Quirina Vallejos, and Thomas A. Arcury. Depressive symptoms among Latino farmworkers across the agricultural season: Structural and situational influences. *Cultural Diversity and Ethnic Minority Psychology*, 16(3):335, July 2010.

[127] Urska Arnautovska, Samara McPhedran, and Diego De Leo. A regional approach to understanding farmer suicide rates in Queensland. *Social Psychiatry and Psychiatric Epidemiology*, 49(4):593–599, April 2014.

[128] Annette L. Beautrais. Farm suicides in New Zealand, 2007–2015: A review of coroners' records. *Australian & New Zealand Journal of Psychiatry*, 52(1):78–86, April 2017.

[129] Prakash B. Behere, Himanshu Mansharamani, Aniruddh P. Behere, and Richa Yadav. Suicide and self-harms in rural setting. In *Mental Health and Illness in the Rural World*, pages 151–167. Springer, Singapore, April 2020.

[130] Manik Changoji Bhise and Prakash Balkrushna Behere. Risk factors for farmers' suicides in central rural India: Matched case–control psychological autopsy study. *Indian Journal of Psychological Medicine*, 38(6):560, November 2016.

[131] Pennan Chinnasamy, Minna J. Hsu, and Govindasamy Agoramoorthy. Groundwater storage trends and their link to farmer suicides in Maharashtra State, India. *Frontiers in Public Health*, 7, August 2019.

[132] Fiona Judd, Henry Jackson, Caitlin Fraser, Greg Murray, Garry Robins, and Angela Komiti. Understanding suicide in Australian farmers. *Social Psychiatry and Psychiatric Epidemiology*, 41(1):1–10, January 2006.

[133] Sarantis Michalopoulos. One French farmer commits suicide every two days, survey says. *Euractiv*, October 2018.

[134] Wendy Ringgenberg, Corinne Peek-Asa, Kelley Donham, and Marizen Ramirez. Trends and characteristics of occupational suicide and homicide in farmers and agriculture workers, 1992–2010. *Journal of Rural Health*, 34(3):246–253, June 2018.

[135] Bronwyn Brew, Kerry Inder, Joanne Allen, Matthew Thomas, and Brian Kelly. The health and wellbeing of Australian farmers: A longitudinal cohort study. *BMC Public Health*, 16(1):1–11, December 2016.

[136] John David Polain, Helen Louise Berry, and John Oliver Hoskin. Rapid change, climate adversity and the next 'big dry': Older farmers' mental health. *Australian Journal of Rural Health*, 19(5):239–243, October 2011.

[137] Alice K. Staniford, Maureen F. Dollard, and Bernard Guerin. Stress and help-seeking for drought-stricken citrus growers in the Riverland of South Australia. *Australian Journal of Rural Health*, 17(3):147–154, June 2009.

[138] Dongxin Wang, Jing Ma, Lihong Tan, Yan Chen, Xiaosong Li, Xuefei Tian, Xuhui Zhou, and Xuejun Liu. Epidemiology of severe mental illness in Hunan province in central China during 2014–2015: A multistage cross-sectional study. *PLoS One*, 12(11):e0188312, November 2017.

[139] Michael Carter, Alain de Janvry, Elisabeth Sadoulet, and Alexander Sarris. Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling. *Revue d'économie du développement*, 23(1):5–57, June 2015.

[140] Selena Ahmed, John Richard Stepp, Colin Orians, Timothy Griffin, Corene Matyas, Albert Robbat, Sean Cash, Dayuan Xue, Chunlin Long, Uchenna Unachukwu, Sarabeth Buckley, David Small, and Edward Kennelly. Effects of extreme climate events on tea (Camellia sinensis) functional quality validate indigenous farmer knowledge and sensory preferences in tropical China. *PLoS One*, 9(10):e109126, October 2014.

[141] Meredith Howard, Selena Ahmed, Paul Lachapelle, and Mark B. Schure. Farmer and rancher perceptions of climate change and their relationships with mental health. *Journal of Rural Mental Health*, 44(2):87, 2020.

[142] Neville R. Ellis and Glenn A. Albrecht. Climate change threats to family farmers' sense of place and mental wellbeing: A case study from the Western Australian Wheatbelt. *Social Science & Medicine*, 175:161–168, February 2017.

[143] Celia A. Harvey, Milagro Saborio-Rodríguez, M. Ruth Martinez-Rodríguez, Barbara Viguera, Adina Chain-Guadarrama, Raffaele Vignola, and Francisco Alpizar. Climate

change impacts and adaptation among smallholder farmers in Central America. *Agriculture & Food Security*, 7(1):1–20, December 2018.

[144] Byomkesh Talukder, Gary W. van Loon, Keith W. Hipel, Sosten Chiotha, and James Orbinski. Health impacts of climate change on smallholder farmers. *One Health*, 13:100258, December 2021.

[145] Celia A. Harvey, Zo Lalaina Rakotobe, Nalini S. Rao, Radhika Dave, Hery Razafimahatratra, Rivo Hasinandrianina Rabarijohn, Haingo Rajaofara, and James L. MacKinnon. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. *Philosophical Transactions of the Royal Society B*, 369(1639), April 2014.

[146] Nomcebo R. Ubisi, Paramu L. Mafongoya, Unathi Kolanisi, and Obert Jiri. Smallholder farmers' perceived effects of climate change on crop production and household livelihoods in rural Limpopo province, South Africa. *Change and Adaptation in Socio-Ecological Systems*, 3(1):27–38, 2017.

[147] Jacquelyn F. Escarcha, Jonatan A. Lassa, Eric P. Palacpac, and Kerstin K. Zander. Livelihoods transformation and climate change adaptation: The case of smallholder water buffalo farmers in the Philippines. *Environmental Development*, 33:100468, March 2020.

[148] Anoek J. van Tilburg and Paul F. Hudson. Extreme weather events and farmer adaptation in Zeeland, the Netherlands: A European climate change case study from the Rhine delta. *Science of the Total Environment*, 844:157212, October 2022.

[149] Bernhard Schauberger, Jonas Jägermeyr, and Christoph Gornott. A systematic review of local to regional yield forecasting approaches and frequently used data resources. *European Journal of Agronomy*, 120:126153, October 2020.

[150] Muhammad Ashraf, Jayant K. Routray, and Muhammad Saeed. Determinants of farmers' choice of coping and adaptation measures to the drought hazard in northwest Balochistan, Pakistan. *Natural Hazards*, 73(3):1451–1473, September 2014.

[151] Gopal Datt Bhatta and Pramod Kumar Aggarwal. Coping with weather adversity and adaptation to climatic variability: A cross-country study of smallholder farmers in South Asia. *Climate and Development*, March 2016.

[152] Yongdeng Lei, Chengcheng Liu, Linbo Zhang, and Shanghua Luo. How smallholder farmers adapt to agricultural drought in a changing climate: A case study in southern China. *Land Use Policy*, 55:300–308, September 2016.

[153] Javier Mulero Chaves and Tomaso De Cola. Public warning applications: Requirements and examples. In *Wireless Public Safety Networks 3*, pages 1–18. Elsevier, Waltham, MA, USA, January 2017.

[154] United Nations Office for Disaster Risk Reduction. Early warnings for all (EW4All), 2023. Accessed: 2024-10-10.

[155] Felix Rembold, Michele Meroni, Ferdinando Urbano, Gabor Csak, Hervé Kerdiles, Ana Perez-Hoyos, Guido Lemoine, Olivier Leo, and Thierry Negre. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. *Agricultural Systems*, 168:247–257, January 2019.

[156] Inbal Becker-Reshef, Brian Barker, Michael Humber, Estefania Puricelli, Antonio Sanchez, Ritvik Sahajpal, Katie McGaughey, Christopher Justice, Bettina Baruth, Bingfang Wu, Adam Prakash, Abdolreza Abbassian, and Ian Jarvis. The GEOGLAM crop monitor for AMIS: Assessing crop conditions in the context of global markets. *Global Food Security*, 23:173–181, December 2019.

[157] Food and Agriculture Organization of the United Nations. FAO SWALIM: Somalia Water and Land Information Management – Who we are, September 2024. [Online; accessed 20 Sep. 2024].

[158] Kyla Wethli. Benefit-Cost Analysis for Risk Management: Summary of Selected Examples, 2014. World Bank background paper.

[159] Ademola Braimoh, Bernard Manyena, Grace Obuya, and Francis Muraya. *Assessment of food security early warning systems for East and Southern Africa*. World Bank, 2018.

[160] Gabriela Guimarães Nobre, Frank Davenport, Konstantinos Bischiniotis, Ted Veldkamp, Brenden Jongman, Christopher C. Funk, Gregory Husak, Philip J. Ward, and Jeroen C. J. H. Aerts. Financing agricultural drought risk through ex-ante cash transfers. *Science of the Total Environment*, 653:523–535, February 2019.

[161] Mariette Vreugdenhil, Isabella Pfeil, Luca Brocca, Stefania Camici, Markus Enenkel, and Wolfgang Wagner. Satellite soil moisture for yield prediction in water limited regions. *Copernicus Meetings*, March 2021.

[162] United Nations Development Programme. Early warning systems, September 2024.

[163] Richard Ewbank, Carlos Perez, Hilary Cornish, Mulugeta Worku, and Solomon Woldetsadik. Building resilience to El Niño-related drought: Experiences in early warning and early action from Nicaragua and Ethiopia. *Disasters*, 43(S3):S345–S367, April 2019.

[164] Rajesh Kumar Rai, Marc J. C. van den Homberg, Gopal Prasad Ghimire, and Colin McQuistan. Cost-benefit analysis of flood early warning system in the Karnali River Basin of Nepal. *International Journal of Disaster Risk Reduction*, 47:101534, August 2020.

[165] 510 Data and Digital Initiative of the Netherlands Red Cross and Zimbabwe Red Cross Society and Danish Red Cross and Finnish Red Cross. Forecasting drought in Zimbabwe: Advances, challenges, and future opportunities. Technical report, 510 Data and Digital Initiative, Netherlands Red Cross, The Hague, Netherlands, 2024.

[166] Jonathan Lala, Donghoon Lee, Juan Bazo, and Paul Block. Evaluating prospects for subseasonal-to-seasonal forecast-based anticipatory action from a global perspective. *Weather and Climate Extremes*, 38:100510, December 2022.

[167] World Food Programme. Scaling-up anticipatory actions for food security: Anticipatory action year in focus 2023. Technical report, World Food Programme, Rome, Italy, April 2023.

[168] Emilie Gettliffe. UN OCHA Anticipatory Action: Lessons from the 2020 Somalia Pilot. Working paper, Centre for Disaster Protection, October 2021. Funded by UK aid through the UK government.

[169] Jaime Martin Farelo and Francesca Giubilo. Anticipatory Action: Shaping the Future of Humanitarian Response, June 2024. Includes contributions from multiple organizations such as Diakonie Katastrophenhilfe, International Rescue Committee, and Welthungerhilfe.

[170] Emilie Gettliffe. Anticipatory action for monsoon flooding in Bangladesh, December 2020. Blog post, Centre for Disaster Protection.

[171] Aayushi Kachalia and Devin Olmack. Listening to those in crisis: understanding perspectives on anticipatory action in Ethiopia, November 2023.

[172] Evan Easton-Calabria. Possibilities and limitations of anticipatory action in complex crises: acting in advance of flooding in South Sudan. *Disasters*, 49(1):e12654, September 2024.

[173] ACAPS. Vulnerabilities in the Rohingya Refugee Camps. Technical report, ACAPS, December 2019.

[174] Paolo Verme and Chiara Gigliarano. Optimal targeting under budget constraints in a humanitarian context. *World Development*, 119:224–233, 2019. Open access under CC BY-NC-ND.

[175] Ariel BenYishay, Matthew DiLorenzo, and Carrie Dolan. The economic efficiency of aid targeting. *World Development*, 160:106062, December 2022.

[176] Emily Aiken, Suzanne Bellue, Dean Karlan, Christopher Udry, and Joshua E. Blumenstock. Machine learning and phone data can improve targeting of humanitarian aid. *Nature*, 603:864–870, March 2022.

[177] Nick Scott, S. Azfar Hussain, Rowan Martin-Hughes, Freya J. I. Fowkes, Cliff C. Kerr, Ruth Pearson, David J. Kedziora, Madhura Killedar, Robyn M. Stuart, and David P. Wilson. Maximizing the impact of malaria funding through allocative efficiency: using the right interventions in the right locations. *Malaria Journal*, 16(1):1–14, December 2017.

[178] United Nations High Commissioner for Refugees. UNHCR Regional Cash Assistance Update: Syria & Iraq Situations, January to June 2017. Technical report, UNHCR, June 2017.

[179] Ibrahim M. Hezam, Moddassir K. Nayeem, and Gyu M. Lee. A systematic literature review on mathematical models of humanitarian logistics. *Symmetry*, 13(1):11, December 2020.

[180] Warren B. Powell. *A modern approach to teaching introduction to optimization*. Now Publishers, Hanover, MA, 2023.

[181] Saul I. Gass and Arjang A. Assad. History of operations research. In *Transforming research into action*, pages 1–14. INFORMS, 2011.

[182] Leonid V. Kantorovich. The mathematical method of production planning and organization. *Management Science*, 6(4):363–422, 1939.

[183] Kenneth J. Arrow. George Dantzig in the development of economic analysis. *Discrete Optimization*, 5(2):159–167, May 2008.

[184] George B. Dantzig. Linear programming. *Operations Research*, February 2002.

[185] Laurence A. Wolsey. *Integer programming*. John Wiley Sons, 2020.

[186] George B. Dantzig. Linear programming under uncertainty. *Management Science*, April 1955.

[187] Ravindra K. Ahuja, Thomas L. Magnanti, James B. Orlin, and M. R. Reddy. Applications of network optimization. In *Handbooks in Operations Research and Management Science*, volume 7, pages 1–83. Elsevier, Waltham, MA, USA, January 1995.

[188] Richard E. Bellman and Stuart E. Dreyfus. *Applied dynamic programming*, volume 2050. Princeton University Press, 2015.

[189] Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471–483, June 2014.

[190] Bram L. Gorissen, İhsan Yanıkoğlu, and Dick Den Hertog. A practical guide to robust optimization. *Omega*, 53:124–137, June 2015.

[191] Hans-Georg Beyer and Bernhard Sendhoff. Robust optimization — A comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196(33):3190–3218, July 2007.

[192] Mordecai Avriel. Nonlinear programming. In *Mathematical Programming for Operations Researchers and Computer Scientists*, pages 271–367. CRC Press, Boca Raton, FL, USA, November 2020.

[193] Haiping Ma, Yajing Zhang, Shengyi Sun, Ting Liu, and Yu Shan. A comprehensive survey on NSGA-II for multi-objective optimization and applications. *Artificial Intelligence Review*, 56(12):15217–15270, December 2023.

[194] Abraham Charnes, William W. Cooper, and Shanling Li. Using data envelopment analysis to evaluate efficiency in the economic performance of Chinese cities. *Socio-Economic Planning Sciences*, 23(6):325–344, 1989.

[195] Ana Santos Camanho, Maria Conceicao Silva, Fabio Sartori Piran, and Daniel Pacheco Lacerda. A literature review of economic efficiency assessments using data envelopment analysis. *European Journal of Operational Research*, 315(1):1–18, May 2024.

[196] Bartłomiej Kizielewicz, Andrii Shekhovtsov, and Wojciech Sałabun. pymcdm — The universal library for solving multi-criteria decision-making problems. *SoftwareX*, 22:101368, May 2023.

[197] Joaquim R. R. A. Martins and Andrew Ning. *Engineering design optimization*. Cambridge University Press, Cambridge, UK, January 2022.

[198] Gerald G. Brown, Robert A. Koyak, Javier Salmerón, and Zachary Scholz. Optimizing prepositioning of equipment and personnel for Los Angeles County Fire Department to fight wildland fires. *INFORMS Journal on Applied Analytics*, August 2021.

[199] Thomas W. Archibald and Sarah E. Marshall. Review of mathematical programming applications in water resource management under uncertainty. *Environmental Modeling and Assessment*, 23(6):753–777, December 2018.

[200] Adriana C. Luna, Nelson L. Diaz, Moisès Graells, Juan C. Vasquez, and Josep M. Guerrero. Mixed-integer linear programming-based energy management system for hybrid PV-wind-battery microgrids: Modeling, design, and experimental verification. *IEEE Transactions on Power Electronics*, 32(4):2769–2783, June 2016.

[201] Florian Peterssen, Marlon Schlemminger, Clemens Lohr, Raphael Niepelt, Richard Hanke-Rauschenbach, and Rolf Brendel. Impact of forecasting on energy system optimization. *Advances in Applied Energy*, 15:100181, September 2024.

[202] Arun Kumar Sangaiah, Erfan Babaee Tirkolaee, Alireza Goli, and Saeed Dehnavi-Arani. Robust optimization and mixed-integer linear programming model for LNG supply chain planning problem. *Soft Computing*, 24(11):7885–7905, June 2020.

[203] Yasser M. R. Aboelmagd. Linear programming applications in construction sites. *Alexandria Engineering Journal*, 57(4):4177–4187, December 2018.

[204] Firouz Khodayari and Yashar Pourrahimian. Mathematical programming applications in block-caving scheduling: A review of models and algorithms. *International Journal of Mining and Mineral Engineering*, August 2015.

[205] Benjamin Laufer, Thomas Gilbert, and Helen Nissenbaum. Optimization's neglected normative commitments. In *ACM Conference Proceedings*, pages 50–63. Association for Computing Machinery, New York, NY, USA, June 2023.

[206] Chao Ning and Fengqi You. Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers and Chemical Engineering*, 125:434–448, June 2019.

[207] Xinying Violet Chen and J. N. Hooker. A guide to formulating fairness in an optimization model. *Annals of Operations Research*, 326(1):581–619, July 2023.

[208] Harwin De Vries and Luk N. Van Wassenhove. Do optimization models for humanitarian operations need a paradigm shift? *Production and Operations Management*, 29(1):55–61, January 2020.

[209] E. Downey Brill. The use of optimization models in public-sector planning. *Management Science*, May 1979.

[210] Vincent Bebien, Odile Bellenguez, Gilles Coppin, Anna Ma-Wyatt, and Rachel Stephens. Ethical decision-making in human-automation collaboration: A case study of the nurse rostering problem. *AI Ethics*, pages 1–13, April 2024.

[211] Antonia Ilabaca, Germán Paredes-Belmar, and Pamela P. Alvarez. Optimization of humanitarian aid distribution in case of an earthquake and tsunami in the city of Iquique, Chile. *Sustainability*, 14(2):819, 2022.

[212] Byung Hoon Oh, Kwangyeon Kim, Han-Lim Choi, and Inseok Hwang. Integer linear program approach for evacuation of disaster victims with different urgency levels. *IFAC-PapersOnLine*, 50(1):15018–15023, July 2017.

[213] Jing Wang, Jianping Cai, Xiaohang Yue, and Nallan C. Suresh. Pre-positioning and real-time disaster response operations: Optimization with mobile phone location data. *Transportation Research Part E: Logistics and Transportation Review*, 150:102344, June 2021.

[214] Aliakbar Hasani and Hadi Mokhtari. Redesign strategies of a comprehensive robust relief network for disaster management. *Socio-Economic Planning Sciences*, 64:92–102, December 2018.

[215] Gökalp Erbeyoğlu and Ümit Bilge. A robust disaster preparedness model for effective and fair disaster response. *European Journal of Operational Research*, 280(2):479–494, January 2020.

[216] Erfan Babaee Tirkolaee, Nadi Serhan Aydın, Mehdi Ranjbar-Bourani, and Gerhard-Wilhelm Weber. A robust bi-objective mathematical model for disaster rescue units allocation and scheduling with learning effect. *Computers and Industrial Engineering*, 149:106790, November 2020.

[217] Ning Xu, Qiong Zhang, Haoran Zhang, Minsung Hong, Rajendra Akerkar, and Yongtu Liang. Global optimization for multi-stage construction of rescue units in disaster response. *Sustainable Cities and Society*, 51:101768, November 2019.

[218] Juliano Silva Souza, Flávio Araújo Lim-Apo, Leonardo Varella, Antônio Sérgio Coelho, and João Carlos Souza. Multi-period optimization model for planning people allocation in shelters and distributing aid with special constraints. *Socio-Economic Planning Sciences*, 79:101087, February 2022.

[219] Irineu de Brito, Silvia Uneddu, Emma Maspero, and Paulo Gonçalves. Optimizing long-lasting insecticidal nets campaign in Ivory Coast. *Logistics*, 4(3):18, August 2020.

[220] Danique de Moor, Joris Wagenaar, Robert Poos, Dick Den Hertog, and Hein Fleuren. A robust approach to food aid supply chains. *European Journal of Operational Research*, 318(1):269–285, October 2024.

[221] Tri-Dung Nguyen, Tri Nguyen-Quang, Uday Venkatadri, Claver Diallo, and Michelle Adams. Mathematical programming models for fresh fruit supply chain optimization: A review of the literature and emerging trends. *AgriEngineering*, 3(3):519–541, July 2021.

[222] M. M. E. Alemany, Ana Esteso, Ángel Ortiz, and Mariana del Pino. Centralized and distributed optimization models for the multi-farmer crop planning problem under uncertainty: Application to a fresh tomato Argentinean supply chain case study. *Computers and Industrial Engineering*, 153:107048, March 2021.

[223] Silvia Araújo dos Reis, José Eugenio Leal, and Antônio Márcio Tavares Thomé. A two-stage stochastic linear programming model for tactical planning in the soybean supply chain. *Logistics*, 7(3):49, August 2023.

[224] Edgar León-Olivares, Hertwin Minor-Popocatl, Omar Aguilar-Mejía, and Diana Sánchez-Partida. Optimization of the supply chain in the production of ethanol from

agricultural biomass using mixed-integer linear programming (MILP): A case study. *Mathematical Problems in Engineering*, 2020(1):6029507, January 2020.

[225] Gabriel Bayá, Eduardo Canale, Sergio Nesmachnow, Franco Robledo, and Pablo Sartor. Production optimization in a grain facility through mixed-integer linear programming. *Applied Sciences*, 12(16):8212, August 2022.

[226] Víctor M. Albornoz and Patricio I. Vera. Coordinating harvest planning and scheduling in an agricultural supply chain through stochastic bilevel programming. *International Transactions in Operational Research*, 30(4):1819–1842, July 2023.

[227] Harol Mauricio Gámez-Albán, Ruben Guisson, and Annelies De Meyer. Optimizing the organization of the first mile in agri-food supply chains with a heterogeneous fleet using a mixed-integer linear model. *Intelligent Systems with Applications*, 23:200426, September 2024.

[228] Konstantinos Ziliaskopoulos and Konstantinos Papalamprou. A bilevel linear programming model for developing a subsidy policy to minimize the environmental impact of the agricultural sector. *Sustainability*, 14(13):7651, June 2022.

[229] Younis Jabarzadeh, Hossein Reyhani Yamchi, Vikas Kumar, and Nader Ghaffarinasab. A multi-objective mixed-integer linear model for sustainable fruit closed-loop supply chain network. *Management of Environmental Quality: An International Journal*, 31(5):1351–1373, June 2020.

[230] Fernando Pauli de Bastiani, Thiago Guilherme Péra, and José Vicente Caixeta-Filho. Fertilizer logistics in Brazil: Application of a mixed-integer programming mathematical model for optimal mixer locations. *Logistics*, 8(1):4, January 2024.

[231] Mo Li, Qiang Fu, Vijay P. Singh, Dong Liu, Tianxiao Li, and Yan Zhou. Managing agricultural water and land resources with tradeoff between economic, environmental, and social considerations: A multi-objective non-linear optimization model under uncertainty. *Agricultural Systems*, 178:102685, February 2020.

[232] Fahad Khan Khadim and Ramesh S. V. Teegavarapu. Optimal interventions for flood control, drainage and irrigation project improvements. *Hydrological Sciences Journal*, February 2020.

[233] Fan Zhang, Shanshan Guo, Xiao Liu, Youzhi Wang, Bernard A. Engel, and Ping Guo. Towards sustainable water management in an arid agricultural region: A multi-level multi-objective stochastic approach. *Agricultural Systems*, 182:102848, June 2020.

[234] Qi Li and Guiping Hu. Multistage stochastic programming modeling for farmland irrigation management under uncertainty. *PLoS One*, 15(6):e0233723, June 2020.

[235] Mahak Bhatia and Anil Rana. A mathematical approach to optimize crop allocation — A linear programming model. *International Journal of Design Nature and Ecodynamics*, 15(2):245–252, 2020.

[236] Jaume Freire-González, Christopher A. Decker, and Jim W. Hall. A linear programming approach to water allocation during a drought. *Water*, 10(4):363, March 2018.

[237] Igor Boninsenha, Everardo Chartuni Mantovani, Marcos Heil Costa, and Aziz Galvão da Silva Júnior. A linear programming model for operational optimization of agricultural activity considering a hydroclimatic forecast — Case studies for Western Bahia, Brazil. *Water*, 14(22):3625, November 2022.

[238] Ajay Singh and Sudhindra Nath Panda. Development and application of an optimization model for the maximization of net agricultural return. *Agricultural Water Management*, 115:267–275, December 2012.

[239] Xiao Liu, Ping Guo, Qian Tan, Fan Zhang, Yan Huang, and Youzhi Wang. Drought disaster risk management based on optimal allocation of water resources. *Natural Hazards*, 108(1):285–308, August 2021.

[240] Kenneth R. Baker. *Optimization modeling with spreadsheets, 3rd edition.* Wiley, Hoboken, NJ, USA, July 2015.

[241] Alan J. King and Stein W. Wallace. *Modeling with stochastic programming.* Springer, New York, NY, USA, 2012.

[242] Anticipation Hub. Anticipatory action in 2023: A global overview. Technical report, Anticipation Hub, Bonn, Germany, 2023. Published in partnership with the Anticipatory Action Task Force (AATF).

[243] Nabansu Chattopadhyay, S. Sunitha Devi, Gracy John, and V. R. Choudhari. Occurrence of hail storms and strategies to minimize its effect on crops. *Mausam*, 68(1):75–92, January 2017.

[244] Siba Prasad Mishra and Bhumika Bauri. Carving drought impact over Purulia district, West Bengal, India. *Advances in Research*, 25(3):186–205, April 2024.

[245] Fabiola Regis-Hernández, Jaime Mora-Vargas, and Angel Ruíz. A multi-criteria vertical coordination framework for a reliable aid distribution. *Journal of Industrial Engineering and Management*, 10(4):789–815, October 2017.

[246] Wapee Manopiniwes and Takashi Irohara. Stochastic optimisation model for integrated decisions on relief supply chains: Preparedness for disaster response. *International Journal of Production Research*, February 2017.

[247] Muhammad Khan, Hee Yong Lee, and Jung Han Bae. The role of transparency in humanitarian logistics. *Sustainability*, 11(7):2078, April 2019.

[248] Marijn van der Velde, Bettina Baruth, Andras Bussay, Andreja Ceglar, Sergio Garcia-Condado, Sokratis Karetsos, Romain Lecerf, Ricardo Lopez-Lozano, Alessandro Maiorano, Luigi Nisini, Luisa Seguini, and Maja van den Berg. In-season performance of European Union wheat forecasts during extreme impacts. *Scientific Reports*, 8(15420):1–10, October 2018.

[249] Weston Anderson, Shraddhanand Shukla, Jim Verdin, Andrew Hoell, Christina Justice, Brian Barker, Kimberly Slinski, Nathan Lenssen, Jiale Lou, Benjamin I. Cook, and Amy McNally. Preseason maize and wheat yield forecasts for early warning of crop failure. *Nature Communications*, 15(7262):1–11, August 2024.

[250] Djavan De Clercq and Adam Mahdi. Feasibility of machine learning-based rice yield prediction in India at the district level using climate reanalysis and remote sensing data. *Agricultural Systems*, 220:104099, October 2024.

[251] Peter Dok Tindan, Divine Odame Appiah, and Alexander Yao Segbefia. Attentiveness to early warning drought information: Implications for policy support and climate risk reduction in Ghana. *International Journal of Disaster Risk Science*, 13(1):25–37, February 2022.

[252] Philip Antwi-Agyei, Kofi Amanor, Jonathan N. Hogarh, and Andrew J. Dougill. Predictors of access to and willingness to pay for climate information services in north-eastern Ghana: A gendered perspective. *Environmental Development*, 37:100580, March 2021.

[253] Waleed Alsabhan and Basil Dudin. Real-time flood forecasting and warning: A comprehensive approach toward HCI-centric mobile app development. *Multimodal Technologies and Interaction*, 7(5):44, April 2023.

[254] Marco Romano, Paloma Díaz, and Ignacio Aedo. A gamified platform for civic engagement in early warning. In *ACM Conference Proceedings*, pages 1–8. Association for Computing Machinery, New York, NY, USA, September 2018.

[255] Teboho E. Masupha, Mokhele E. Moeletsi, and Mitsuru Tsubo. Prospects of an agricultural drought early warning system in South Africa. *International Journal of Disaster Risk Reduction*, 66:102615, December 2021.

[256] Biswanath Dash. Gap between impact-based and impact forecast and warning: Implications for people-centric early warning system (EWS) in India. *Natural Hazards Research*, 4(1):110–117, March 2024.

[257] Iain Dunning, Stuart Mitchell, and Michael O'Sullivan. PuLP: A linear programming toolkit for Python, September 2011. Department of Engineering Science, The University of Auckland.

[258] Michael L. Bynum, Gabriel A. Hackebeil, William E. Hart, Carl D. Laird, Bethany L. Nicholson, John D. Siirola, Jean-Paul Watson, and David L. Woodruff. *Pyomo — Optimization modeling in Python*. Springer, Cham, Switzerland, 2021.

[259] Zelda B. Zabinsky, Mariam Zameer, Larissa P. G. Petroianu, Mamiza M. Muteia, and Aida L. Coelho. Route optimization tool (RoOT) for distribution of vaccines and health products. *Gates Open Research*, 5(34):34, November 2021.

[260] Aryamala Prasad, Mila Gascó-Hernández, J. Ramon Gil-García, and Qianli Yuan. Human-centric design in applications for emergency preparedness and response in rural communities: The case of the E!App. In *Proceedings of the 24th Annual International Conference on Digital Government Research*, pages 380–388, Gdańsk, Poland, July 2023. ACM.

[261] Gaurav Kabra, A. Ramesh, Pervaiz Akhtar, and Manoj Kumar Dash. Understanding behavioural intention to use information technology: Insights from humanitarian practitioners. *Telematics and Informatics*, 34(7):1250–1261, November 2017.

[262] Holland M. Vasquez, Emilie Pianarosa, Renee Sirbu, Lori M. Diemert, Heather Cunningham, Vinyas Harish, Birsen Donmez, and Laura C. Rosella. Human factors methods in the design of digital decision support systems for population health: A scoping review. *BMC Public Health*, 24:2458, September 2024.

[263] Jeroen Beliën, Liesje De Boeck, Dries Goossens, Nils-Hassan Quttineh, and Jens Vermeir. The game of thrones quest for optimality: A role playing game for teaching the art of linear programming. *INFORMS Transactions on Education*, October 2024.

[264] Dries Goossens and Jeroen Beliën. Teaching integer programming by scheduling the Belgian soccer league. *INFORMS Transactions on Education*, May 2022.

[265] Timothy C. Y. Chan, Craig Fernandes, Albert Loa, and Nathan Sandholtz. Moneyball for murderball: Using analytics to construct lineups in wheelchair rugby. *INFORMS Transactions on Education*, February 2023.

[266] Clemens Gros, Andrea Pronti, Khairul Sheikh, Ahmadul Hassan, and Mohammad Shahjahan. Effects of anticipatory humanitarian cash assistance to households forecasted to experience extreme flooding: Evidence from Bangladesh. *Hydrology Research*, 54(11):1315–1328, November 2023.

[267] Clemens Gros, Meghan Bailey, Saroja Schwager, Ahmadul Hassan, Raymond Zingg, Muhammad Mamtaz Uddin, Mohammad Shahjahan, Hasibul Islam, Stefanie Lux,

Catalina Jaime, and Erin Coughlan de Perez. Household-level effects of providing forecast-based cash in anticipation of extreme weather events: Quasi-experimental evidence from humanitarian interventions in the 2017 floods in Bangladesh. *International Journal of Disaster Risk Reduction*, 41:101275, December 2019.

[268] Clemens Gros, Evan Easton-Calabria, Meghan Bailey, Kadirbyek Dagys, Erin Coughlan de Perez, Munguntuya Sharavnyambuu, and Andrew Kruczkiewicz. The effectiveness of forecast-based humanitarian assistance in anticipation of extreme winters: A case study of vulnerable herders in Mongolia. *Disasters*, 46(1):95–118, January 2022.

[269] Milan Zeleny. On the squandering of resources and profits via linear programming. *Interfaces*, October 1981.

[270] FAO. FAOSTAT: Food balance sheets, 2024. Accessed 8 December 2024. Licence: CC-BY-4.0.

[271] Duangthip Sirikanchanarak, Tanaporn Tungtrakul, and Songsak Sriboonchitta. The future of global rice consumption: Evidence from dynamic panel data approach. In *Predictive Econometrics and Big Data*, pages 629–642. Springer, Cham, Switzerland, December 2017.

[272] Naomi K. Fukagawa and Lewis H. Ziska. Rice: Importance for global nutrition. *Journal of Nutritional Science and Vitaminology*, 65(Supplement):S2–S3, 2019.

[273] Emily Elert. Rice by the numbers: A good grain. *Nature*, 514(7524):S50–S51, October 2014.

[274] Thomas van Klompenburg, Ayalew Kassahun, and Cagatay Catal. Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177:105709, October 2020.

[275] Xuexiao Wu, Prosper Washaya, Long Liu, Kun Li, Yun Shao, Lingyuan Meng, and Bailin Hu. Rice yield estimation based on spaceborne SAR: A review from 1988 to 2018. *IEEE Access*, 8:157462–157469, August 2020.

[276] Dimitrios Katsantonis, Kalliopi Kadoglidou, Christos Dramalis, and Pau Puigdollers. Rice blast forecasting models and their practical value: A review. *Phytopathologia Mediterranea*, 56(2):187–216, September 2017.

[277] Nicoleta Darra, Evangelos Anastasiou, Olga Kriezi, Erato Lazarou, Dionissios Kalivas, and Spyros Fountas. Can yield prediction be fully digitilized? A systematic review. *Agronomy*, 13(9):2441, September 2023.

[278] Daniel Marc G. dela Torre, Jay Gao, and Cate Macinnis-Ng. Remote sensing-based estimation of rice yields using various models: A critical review. *Geo-spatial Information Science*, October 2021.

[279] Rongkun Zhao, Yuechen Li, and Mingguo Ma. Mapping paddy rice with satellite remote sensing: A review. *Sustainability*, 13(2):503, January 2021.

[280] Mostafa K. Mosleh, Quazi K. Hassan, and Ehsan H. Chowdhury. Application of remote sensors in mapping rice area and forecasting its production: A review. *Sensors*, 15(1):769–791, January 2015.

[281] Purba Goswami and Gopal Dutta. Evaluation of DSSAT model (CERES rice) on rice production: A review. *International Journal of Chemical Studies*, 8(5):404–409, 2020.

[282] Chinna Gopi Simhadri, Hari Kishan Kondaveeti, Valli Kumari Vatsavayi, Alakananda Mitra, and Preethi Ananthachari. Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques. *Information Processing in Agriculture*, May 2024.

[283] Kylin Bocalan Felizardo, Angelo Mari Cuevas Paredes, and Edwin Romeroso Arboleda. Advancements in artificial intelligence (AI) for enhanced insights and automation in rice agriculture: A systematic review. *International Journal of Science and Research Archive*, 11(1):444–463, 2024.

[284] Mayuri Sharma, Chandan Jyoti Kumar, and Dhruba K. Bhattacharyya. Machine/deep learning techniques for disease and nutrient deficiency disorder diagnosis in rice crops: A systematic review. *Biosystems Engineering*, 244:77–92, August 2024.

[285] Rayner Alfred, Joe Henry Obit, Christie Pei-Yee Chin, Haviluddin Haviluddin, and Yuto Lim. Towards paddy rice smart farming: A review on big data, machine learning, and rice production tasks. *IEEE Access*, 9:50358–50380, March 2021.

[286] Nicholas Dominic, Tjeng Wawan Cenggoro, and Bens Pardamean. Systematic literature review on statistics and machine learning predictive models for rice phenotypes. *Procedia Computer Science*, 227:1054–1061, January 2023.

[287] Jérôme Bartholomé, Parthiban Thathapalli Prakash, and Joshua N. Cobb. Genomic prediction: Progress and perspectives for rice improvement. In *Genomic Prediction of Complex Traits: Methods and Protocols*, pages 569–617. Springer US, 2022.

[288] Andy P. Siddaway, Alex M. Wood, and Larry V. Hedges. How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70:747–770, January 2019.

[289] Wondimagegn Mengist, Teshome Soromessa, and Gudina Legese. Method for conducting systematic literature review and meta-analysis for environmental science research. *MethodsX*, 7:100777, January 2020.

[290] Gehad Mohamed Tawfik, Kadek Agus Surya Dila, Muawia Yousif Fadlelmola Mohamed, Dao Ngoc Hien Tam, Nguyen Dang Kien, Ali Mahmoud Ahmed, and Nguyen Tien Huy. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health*, 47(1):1–9, December 2019.

[291] Philipp C. Sauer and Stefan Seuring. How to conduct systematic literature reviews in management research: A guide in 6 steps and 14 decisions. *Review of Managerial Science*, 17(5):1899–1933, July 2023.

[292] Edoardo Aromataris and Alan Pearson. The systematic review: An overview. *AJN The American Journal of Nursing*, 114(3):53, March 2014.

[293] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, 2021.

[294] Alex Pollock and Eivind Berge. How to do a systematic review. *International Journal of Stroke*, 13(2):138–156, November 2017.

[295] Theresa L. Charrois. Systematic reviews: What do you need to know to get started? *Canadian Journal of Hospital Pharmacy*, 68(2):144, March 2015.

[296] Jeffrey Dean. A golden decade of deep learning: Computing systems and applications. *Daedalus*, 151(2):58–74, May 2022.

[297] Timothy Meline. Selecting studies for systemic review: Inclusion and exclusion criteria. *ASHA Wire*, March 2006.

[298] D. M. P. W. Dissanayake, R. M. K. T. Rathnayake, and L. L. Gihan Chathuranga. Crop yield forecasting using machine learning techniques: A systematic literature review. *KDU Journal of Multidisciplinary Studies*, 5(1), July 2023.

[299] Md. Abu Jabed and Masrah Azrifah Azmi Murad. Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon*, 10(24):e40836, December 2024.

[300] Jacob VanderPlas, Brian E. Granger, Jeffrey Heer, Dominik Moritz, Kanit Wongsuphasawat, Arvind Satyanarayan, Eitan Lees, Ilia Timofeev, Ben Welsh, and Scott Sievert. Altair: Interactive statistical visualizations for Python. *Journal of Open Source Software*, 3(32):1057, December 2018.

[301] Norman R. Draper and Harry Smith. *Applied Regression Analysis*. John Wiley and Sons, Chichester, England, UK, April 1998.

[302] Arsenio Paez. Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10(3):233–240, August 2017.

[303] Priya Nair Rajeev and Simy Joy. CropIn: Addressing farmer poverty through "connected" farming. In *Socio-Tech Innovation: Harnessing Technology for Social Good*, pages 289–304. Palgrave Macmillan, Cham, Switzerland, April 2020.

[304] Md Didarul Islam, Liping Di, Faisal Mueen Qamer, Sravan Shrestha, Liying Guo, Li Lin, Timothy J. Mayer, and Aparna R. Phalke. Rapid rice yield estimation using integrated remote sensing and meteorological data and machine learning. *Remote Sensing*, 15(9):2374, April 2023.

[305] Md Samiul Basir, Milon Chowdhury, Md Nafiul Islam, and Muhammad Ashik-E-Rabbani. Artificial neural network model in predicting yield of mechanically transplanted rice from transplanting parameters in Bangladesh. *Journal of Agriculture and Food Research*, 5:100186, September 2021.

[306] Md. Suruj Mia, Ryoya Tanabe, Luthfan Nur Habibi, Naoyuki Hashimoto, Koki Homma, Masayasu Maki, Tsutomu Matsui, and Takashi S. T. Tanaka. Multimodal deep learning for rice yield prediction using UAV-based multispectral imagery and weather data. *Remote Sensing*, 15(10):2511, May 2023.

[307] Bonosri Ghose, Abu Reza Md Towfiqul Islam, Roquia Salam, Shamsuddin Shahid, Mohammad Kamruzzaman, Samiran Das, Ahmed Elbeltagi, Mohammed Abdus Salam, and Javed Mallick. Rice yield responses in Bangladesh to large-scale atmospheric oscillation using multifactorial model. *Theoretical and Applied Climatology*, 146(1–2):29–44, July 2021.

[308] Tingyan Fu, Shufang Tian, and Qian Zhan. Phenological analysis and yield estimation of rice based on multi-spectral and SAR data in Maha Sarakham, Thailand. *Journal of Spatial Science*, 69(1):149–165, March 2023.

[309] Ehsan Asmar, Mohammad H. Vahidnia, Mojtaba Rezaei, and Ebrahim Amiri. Remote sensing-based paddy yield estimation using physical and FCNN deep learning models in Gilan province, Iran. *Remote Sensing Applications: Society and Environment*, 34:101199, April 2024.

[310] Ali Mokhtar, Hongming He, Mohsen Nabil, Saber Kouadri, Ali Salem, and Ahmed Elbeltagi. Securing China's rice harvest: unveiling dominant factors in production using multi-source data and hybrid machine learning models. *Scientific Reports*, 14(1), June 2024.

[311] James Brinkhoff, Allister Clarke, Brian W. Dunn, and Mark Groat. Analysis and forecasting of Australian rice yield using phenology-based aggregation of satellite and weather data. *Agricultural and Forest Meteorology*, 353:110055, June 2024.

[312] Weiguo Yu, Gaoxiang Yang, Dong Li, Hengbiao Zheng, Xia Yao, Yan Zhu, Weixing Cao, Lin Qiu, and Tao Cheng. Improved prediction of rice yield at field and county levels by synergistic use of SAR, optical and meteorological data. *Agricultural and Forest Meteorology*, 342:109729, November 2023.

[313] P. Sathya and P. Gnanasekaran. Ensemble feature selection framework for paddy yield prediction in Cauvery Basin using machine learning classifiers. *Cogent Engineering*, 10(2), August 2023.

[314] Piyal Ekanayake, Windhya Rankothge, Rukmal Weliwatta, and Jeevani W. Jayasinghe. Machine learning modelling of the relationship between weather and paddy yield in Sri Lanka. *Journal of Mathematics*, 2021:1–14, May 2021.

[315] Samsul Amar, Andi Sudiarso, and Muhammad Kusumawan Herliansyah. The accuracy measurement of stock price numerical prediction. *Journal of Physics: Conference Series*, 1569(3):032027, July 2020.

[316] Malcolm McPhee, Jonathan Richetti, Barry Croke, and Brad Walmsley. Model evaluation: The misuse of statistical techniques when evaluating observations versus predictions. *SESMO*, 6:18758, September 2024.

[317] Shitong Zhou, Lei Xu, and Nengcheng Chen. Rice yield prediction in Hubei Province based on deep learning and the effect of spatial heterogeneity. *Remote Sensing*, 15(5):1361, February 2023.

[318] Yu Tanaka, Tomoya Watanabe, Keisuke Katsura, Yasuhiro Tsujimoto, Toshiyuki Takai, Takashi Sonam Tashi Tanaka, Kensuke Kawamura, Hiroki Saito, Koki Homma, Salifou Goube Mairoua, Kokou Ahouanton, Ali Ibrahim, Kalimuthu Senthilkumar, Vimal Kumar Semwal, Eduardo Jose Graterol Matute, Edgar Corredor, Raafat El-Namaky, Norvie Manigbas, Eduardo Jimmy P. Quilang, Yu Iwahashi, Kota Nakajima, Eisuke Takeuchi, and Kazuki Saito. Deep learning enables instant and versatile estimation of rice yield using ground-based RGB images. *Plant Phenomics*, 5, January 2023.

[319] Pankaj, Brajesh Kumar, P. K. Bharti, Vibhor Kumar Vishnoi, Krishan Kumar, Shashank Mohan, and Krishan Pal Singh. Paddy yield prediction based on 2D images of rice panicles using regression techniques. *The Visual Computer*, 40(6):4457–4471, October 2023.

[320] Muhammad Marong, Nor Azura Husin, Maslina Zolkepli, and Lilly Suriani Affendey. Data-driven rice yield predictions and prescriptive analytics for sustainable agriculture

in Malaysia. *International Journal of Advanced Computer Science and Applications*, 15(3), 2024.

[321] Varvara Aleksandrovna Vedeneeva and Alexander Vladimirovich Koshelev. Analysis of the seasonal dynamics of the NDVI index of potato agrocenosis in the conditions of the dry steppe zone of the Volgograd region. *Research on Crops*, 24(2):366–372, June 2023.

[322] Jonas Bunsen, Vlad Coroamă, and Matthias Finkbeiner. Input–output global hybrid analysis of agricultural primary production (IO-GHAAPP) database. *Sustainability*, 15(12):9351, June 2023.

[323] Fiona H. M. Tang, Thu Ha Nguyen, Giulia Conchedda, Leon Casse, Francesco N. Tubiello, and Federico Maggi. CROPGRIDS: a global geo-referenced dataset of 173 crops. *Scientific Data*, 11(413):1–14, April 2024.

[324] Meric Yilmaz. Accuracy assessment of temperature trends from ERA5 and ERA5-Land. *Science of the Total Environment*, 856:159182, January 2023.

[325] Nathan M. Burglewski, Quirine M. Ketterings, Sunoj Shajahan, and Jan van Aardt. A comparison of traditional and machine learning corn yield models using hyperspectral UAS and Landsat imagery. In *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXIX*, volume 12519, pages 80–88. SPIE, June 2023.

[326] Preeti Jacob, Pranathi Kommuru, R. Ruchitha, Harshitha Kuracha, and Thamineni Taruni. Comparative study of MODIS, Landsat-8, Sentinel-2B, and LISS-4 images for precision farming using NDVI approach. *E3S Web of Conferences*, 405:01004, 2023.

[327] Yang Liu, Haikuan Feng, Jibo Yue, Xiuliang Jin, Yiguang Fan, Riqiang Chen, Mingbo Bian, Yanpeng Ma, Xiaoyu Song, and Guijun Yang. Improved potato AGB estimates based on UAV RGB and hyperspectral images. *Computers and Electronics in Agriculture*, 214:108260, November 2023.

[328] David J. Mulla. Satellite remote sensing for precision agriculture. In *Sensing Approaches for Precision Agriculture*, pages 19–57. Springer, Cham, Switzerland, November 2021.

[329] Jane Southworth, Audrey C. Smith, Mohammad Safaei, Mashoukur Rahaman, Ali Alruzuq, Bewuket B. Tefera, Carly S. Muir, and Hannah V. Herrero. Machine learning versus deep learning in land system science: A decision-making framework for effective land classification. *Frontiers in Remote Sensing*, 5:1374862, May 2024.

[330] Guanwei Huang and Mavin Buss. Decision-making in machine learning. In *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications*, pages 256–260. IEEE, 2020.

[331] Khilola Amankulova, Nizom Farmonov, Parvina Akramova, Ikrom Tursunov, and László Mucsi. Comparison of PlanetScope, Sentinel-2, and Landsat 8 data in soybean yield estimation within-field variability with random forest regression. *Heliyon*, 9(6):e17432, June 2023.

[332] Massimo Musacchio, Malvina Silvestri, Vito Romaniello, Marco Casu, Maria Fabrizia Buongiorno, and Maria Teresa Melis. Comparison of ASI-PRISMA data, DLR-EnMAP data, and field spectrometer measurements on "Sale 'e Porcus", a salty pond (Sardinia, Italy). *Remote Sensing*, 16(6):1092, March 2024.

[333] Susan C. Steele-Dunne, Heather McNairn, Alejandro Monsivais-Huertero, Jasmeet Judge, Pang-Wei Liu, and Kostas Papathanassiou. Radar remote sensing of agricultural canopies: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(5):2249–2273, 2017.

[334] Mahya G. Z. Hashemi, Ehsan Jalilvand, Hamed Alemohammad, Pang-Ning Tan, and Narendra N. Das. Review of synthetic aperture radar with deep learning in agricultural applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218:20–49, December 2024.

[335] Anesmar Olino de Albuquerque, Osmar Luiz Ferreira de Carvalho, Cristiano Rosa e Silva, Pablo Pozzobon de Bem, Roberto Arnaldo Trancoso Gomes, Díbio Leandro Borges, Renato Fontes Guimarães, Concepta Margaret McManus Pimentel, and Osmar Abílio de Carvalho Júnior. Instance segmentation of center pivot irrigation systems using multi-temporal Sentinel-1 SAR images. *Remote Sensing Applications: Society and Environment*, 23:100537, August 2021.

[336] Bahar Asadi and Ali Shamsoddini. Crop mapping through a hybrid machine learning and deep learning method. *Remote Sensing Applications: Society and Environment*, 33:101090, January 2024.

[337] Stella Ofori-Ampofo, Charlotte Pelletier, and Stefan Lang. Crop type mapping from optical and radar time series using attention-based deep learning. *Remote Sensing*, 13(22):4668, November 2021.

[338] Felix Lobert, Johannes Löw, Marcel Schwieder, Alexander Gocht, Michael Schlund, Patrick Hostert, and Stefan Erasmi. A deep learning approach for deriving winter wheat phenology from optical and SAR time series at field level. *Remote Sensing of Environment*, 298:113800, December 2023.

[339] Nguyen-Thanh Son, Chi-Farn Chen, and Cheng-Cru Chen. Remote sensing time series analysis for early rice yield forecasting using random forest algorithm. In *Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries*, pages 353–366. Springer International Publishing, 2022.

[340] Hengbiao Zheng, Wenhan Ji, Wenhui Wang, Jingshan Lu, Dong Li, Caili Guo, Xia Yao, Yongchao Tian, Weixing Cao, Yan Zhu, and Tao Cheng. Transferability of models for predicting rice grain yield from unmanned aerial vehicle multispectral imagery across years, cultivars and sensors. *Drones*, 6(12):423, December 2022.

[341] Abdul Aziz, Komariah Ariyanto, Dwi Priyo Ariyanto, and Sumani Ariyanto. Rice yield prediction using Bayesian analysis on rainfed lands in the Sumbing-Sindoro toposequence, Indonesia. *Scientific Horizons*, 26(7), July 2023.

[342] Samuel John Estenor Parreño and Maria Cristine Joy Anter. New approach for forecasting rice and corn production in the Philippines through machine learning models. *Multidisciplinary Science Journal*, 6(9):2024168, April 2024.

[343] Xinyao Feng, Haoliang Tian, Jiahui Cong, and Chuang Zhao. A method review of the climate change impact on crop yield. *Frontiers in Forests and Global Change*, 6:1198186, July 2023.

[344] Balsher Singh Sidhu, Zia Mehrabi, Navin Ramankutty, and Milind Kandlikar. How can machine learning help in understanding the impact of climate change on crop yields? *Environmental Research Letters*, 18(2):024008, January 2023.

[345] Ningyi Zhang, Xiaohan Zhou, Mengzhen Kang, Bao-Gang Hu, Ep Heuvelink, and Leo F. M. Marcelis. Machine learning versus crop growth models: an ally, not a rival. *AoB Plants*, 15(2):plac061, February 2023.

[346] Bernardo Maestrini, Gordan Mimić, Pepijn A. J. van Oort, Keiji Jindo, Sanja Brdar, Ioannis N. Athanasiadis, and Frits K. van Evert. Mixing process-based and data-driven approaches in yield prediction. *European Journal of Agronomy*, 139:126569, September 2022.

[347] Mohsen Shahhosseini, Guiping Hu, Isaiah Huber, and Sotirios V. Archontoulis. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Scientific Reports*, 11(1606):1–15, January 2021.

[348] Alvin Inderka, Florian Huber, and Volker Steinhage. On convolutional vision transformers for yield prediction. *arXiv*, February 2024.

[349] Saneesh Cleatus Thundiyil, Seyed S. Shalamzari, Joseph Picone, and Scott McKenzie. Transformers for modeling long-term dependencies in time series data: A review. In *2023 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5. IEEE, 2023.

[350] Benjamin Kwapong Osibo, Tinghuai Ma, Bright Bediako-Kyeremeh, Lorenzo Mamelona, and Kristina Darbinian. TCNT: A temporal convolutional network-transformer

framework for advanced crop yield prediction. *Journal of Applied Remote Sensing*, 18(4):044513, November 2024.

[351] Fengwei Guo, Pengxin Wang, Kevin Tansey, Yue Zhang, Mingqi Li, Junming Liu, and Shuyu Zhang. A novel transformer-based neural network under model interpretability for improving wheat yield estimation using remotely sensed multi-variables. *Computers and Electronics in Agriculture*, 223:109111, August 2024.

[352] Zhiming Guo, Longzhe Quan, Deng Sun, Zhaoxia Lou, Yuhang Geng, Tianbao Chen, Yi Xue, Jinbing He, Pengbiao Hou, Chuan Wang, and Jiakang Wang. Efficient crop row detection using transformer-based parameter prediction. *Biosystems Engineering*, 246:13–25, October 2024.

[353] Marius Schlegel and Kai-Uwe Sattler. MLflow2PROV: extracting provenance from machine learning experiments. In *ACM Conference Proceedings*, pages 1–4. Association for Computing Machinery, New York, NY, USA, June 2023.

[354] Xavier Vasques. Machine learning in production. In *Machine Learning Theory and Applications: Hands-on Use Cases with Python on Classical and Quantum Machines*, pages 375–464. Wiley, 2024.

[355] Ask Berstad Kolltveit and Jingyue Li. Operationalizing machine learning models: a systematic literature review. In *ACM Conference Proceedings*, pages 1–8. Association for Computing Machinery, New York, NY, USA, May 2022.

[356] Beyza Eken, Samodha Pallewatta, Nguyen Khoi Tran, Ayse Tosun, and Muhammad Ali Babar. A multivocal review of MLOps practices, challenges and open issues. *arXiv*, June 2024.

[357] Vipul Jain, H. Kavitha, and S. Mohana Kumar. Credit card fraud detection web application using streamlit and machine learning. In *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, pages 1–5, 2022.

[358] Chanin Nantasenamat, David Mendez, A. Patrícia Bento, Evangelos Karatzas, and Gerald Maggiora. Chapter 27 - Building bioinformatics web applications with Streamlit. In Kunal Roy, editor, *Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development*, pages 679–699. Academic Press, 2023.

[359] Shilpa Patil and V. Lokesha. Live Twitter sentiment analysis using streamlit framework. *SSRN Scholarly Paper*, May 2022.

[360] Mariem Belghith, Salma Bchir, Kaouthar Ben Khalifa, and Hela Zoghlami. A new rolling forecasting framework using Microsoft Power BI for data visualization: a case study in a pharmaceutical industry. *Annales Pharmaceutiques Françaises*, November 2023.

[361] Kathy Abusager, Michael Baldwin, and Vincent Hsu. Using Power BI to inform Clostridioides difficile ordering practices at an acute care hospital in Central Florida. *American Journal of Infection Control*, 48(8):S57–S58, August 2020.

[362] Erin Burrell Nickell, Jason Schwebke, and Paul Goldwater. An introductory audit data analytics case study: using Microsoft Power BI and Benford's law to detect accounting irregularities. *Journal of Accounting Education*, 64:100855, September 2023.

[363] Marie Weiss, Frédéric Jacob, and Grégory Duveiller. Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, 236:111402, January 2020.

[364] Vasit Sagan, Maitiniyazi Maimaitijiang, Sourav Bhadra, Matthew Maimaitiyiming, Davis R. Brown, Paheding Sidike, and Felix B. Fritschi. Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 174:265–281, April 2021.

[365] Ainong Li, Shunlin Liang, Angsheng Wang, and Jun Qin. Estimating crop yield from multi-temporal satellite data using multivariate regression and neural network techniques. *Photogrammetric Engineering and Remote Sensing*, 73(10):1149–1157, October 2007.

[366] Ruixue Wang, Roderick M. Rejesus, and Serkan Aglasan. Warming temperatures, yield risk and crop insurance participation. *European Review of Agricultural Economics*, 48(5):1109–1131, December 2021.

[367] Yvonne Wong Jing Wen, Raja Rajeswari Ponnusamy, and Ho Ming Kang. Application of weather index-based insurance for paddy yield: the case of Malaysia. *International Journal of Advanced and Applied Sciences*, 6:51–59, 2019.

[368] Hari Sankar Nayak, Joao Vasco Silva, Chiter Mal Parihar, Suresh K. Kakraliya, Timothy J. Krupnik, Deepak Bijarniya, Mangi Lal Jat, Parbodh C. Sharma, Hanuman Sahay Jat, Harminder Singh Sidhu, and Tek B. Sapkota. Rice yield gaps and nitrogen-use efficiency in the northwestern Indo-Gangetic Plains of India: evidence-based insights from heterogeneous farmers' practices. *Field Crops Research*, 275:108328, 2022.

[369] FAO. India at a glance, 2018.

[370] Muhammad Ishfaq, Ghulam Akbar, Anwar Ali, Shahid Hussain, Abdul Sattar, Muhammad Zaman, Shah Fahad, Muhammad Aamir Khan, and Muhammad Adnan. Alternate wetting and drying: a water-saving and ecofriendly rice production system. *Agricultural Water Management*, 241:106363, 2020.

[371] Abdus Sattar and R. C. Srivastava. Modelling climate smart rice-wheat production system in the Middle Gangetic Plains of India. *Theoretical and Applied Climatology*, 144(1-2):77–91, 2021.

[372] Steffen Fritz, Linda See, Juan-Carlos Bayas, François Waldner, David Jacques, Inbal Becker-Reshef, Andrew Whitcraft, Bettina Baruth, Rafael Bonifacio, Jonathan Crutchfield, and Felix Rembold. A comparison of global agricultural monitoring systems and current gaps. *Agricultural Systems*, 168:258–272, 2019.

[373] Xiangying Xu, Xuan Zhang, Fulu Tao, Liangliang Zhang, Yuchuan Luo, Jing Zhang, Jichong Han, and Jun Xie. Design of an integrated climatic assessment indicator (ICAI) for wheat production: a case study in Jiangsu Province, China. *Ecological Indicators*, 101:943–953, 2019.

[374] Saeed Khaki, Lizhi Wang, and Sotirios V. Archontoulis. A CNN-RNN framework for crop yield prediction. *Frontiers in Plant Science*, 10:1750, 2020.

[375] Anna X. Wang, Christopher Tran, Nikhil Desai, and David B. Lobell. Deep transfer learning for crop yield prediction with remote sensing data. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, pages 1–5. ACM, 2018.

[376] Hamid Kamangir, Brent S. Sams, Nick Dokoozlian, Luis Sanchez, and J. Mason Earles. Large-scale spatio-temporal yield estimation via deep learning using satellite and management data fusion in vineyards. *Computers and Electronics in Agriculture*, 216:108439, January 2024.

[377] Phusanisa Charoen-Ung and Pradit Mittrapiyanuruk. Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning. In Herwig Unger, Sunantha Sodsee, and Phayung Meesad, editors, *Recent Advances in Information and Communication Technology 2018*, pages 33–42. Springer International Publishing, Cham, 2019.

[378] Patrick Filippi, Elizabeth J. Jones, Brett M. Whelan, Thomas F. A. Bishop, and Jeffrey A. Baldock. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture*, 20(5):1015–1029, 2019.

[379] Ishfaq Ahmad, Aamir Ahmad, Muhammad Shahid, Abdul Wajid, and Shah Fahad. Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan. *Journal of the Indian Society of Remote Sensing*, 46(10):1669–1678, 2018.

[380] Anat Goldstein, Lior Fink, Alon Meitin, Michal Bohadana, Ofer Lutenberg, and Gilad Ravid. Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge. *Precision Agriculture*, 19(3):421–444, 2018.

[381] Asheesh Chaurasiya, Andrew J. McDonald, Timothy J. Krupnik, Harminder Singh Sidhu, Yadvinder Singh, Md. Iqbal Hossain, and Stephen E. Justice. Layering smart management practices to sustainably maintain rice yields and improve water use efficiency in eastern India. *Field Crops Research*, 275:108341, 2022.

[382] Zheng Chu and Jiong Yu. An end-to-end model for rice yield prediction using deep learning fusion. *Computers and Electronics in Agriculture*, 174:105471, 2020.

[383] Li Tian, Chun Wang, Hailiang Li, and Haitian Sun. Yield prediction model of rice and wheat crops based on ecological distance algorithm. *Environmental Technology & Innovation*, 20:101132, 2020.

[384] Weiguo Yu, Gaoxiang Yang, Dong Li, Hengbiao Zheng, Xia Yao, Yan Zhu, Weixing Cao, Lin Qiu, and Tao Cheng. Improved prediction of rice yield at field and county levels by synergistic use of SAR, optical and meteorological data. *Agricultural and Forest Meteorology*, 342:109729, 2023.

[385] Liang Wan, Haiyan Cen, Jiangpeng Zhu, Jiafei Zhang, Yueming Zhu, Da-Wen Sun, Xiaoyue Du, Li Zhai, Haiyong Weng, Yijian Li, Xiaoran Li, Yidan Bao, Jianyao Shou, and Yong He. Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer: a case study of small farmlands in the south of China. *Agricultural and Forest Meteorology*, 291:108096, 2020.

[386] Prakash K. Jha, Pramod Kumar Thakur, Govinda D. Bhatta, Pramod K. Aggarwal, and David B. Lobell. Using daily data from seasonal forecasts in dynamic crop models for yield prediction: a case study for rice in Nepal's Terai. *Agricultural and Forest Meteorology*, 265:349–358, 2019.

[387] Seungtaek Jeong, Jonghan Ko, and Jong-Min Yeom. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. *Science of the Total Environment*, 802:149726, 2022.

[388] Ponraj Arumugam, Abel Chemura, Bernhard Schauberger, and Christoph Gornott. Remote sensing based yield estimation of rice (Oryza sativa L.) using gradient boosted regression in India. *Remote Sensing*, 13(12):2379, 2021.

[389] A. Kumar Ranjan and B. R. Parida. Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). *Spatial Information Research*, 27(4), 2019.

[390] Diego Gómez, Abel Chemura, Bernhard Schauberger, and Christoph Gornott. Regional estimation of garlic yield using crop, satellite and climate data in Mexico. *Computers and Electronics in Agriculture*, 181:105943, 2021.

[391] Maximilian Auffhammer, V. Ramanathan, and Jeffrey R. Vincent. Climate change, the monsoon, and rice yield in India. *Climatic Change*, 111(2):411–424, 2012.

[392] Sheetal Sharma, P. Panneerselvam, Yadvinder Singh, Mangi Lal Jat, Ramesh K. Jat, Andrew J. McDonald, and Timothy J. Krupnik. Field-specific nutrient management using rice crop manager decision support tool in Odisha, India. *Field Crops Research*, 241:107578, 2019.

[393] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Caroline Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gabriele Biavati, Jean-Raymond Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leopold Haimberger, Sean Healy, Robin Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Corentin Lupu, Gábor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

[394] Aleš Urban, Hana Hanzlíková, and Jan Kyselý. Evaluation of the ERA5 reanalysis-based universal thermal climate index on mortality data in Europe. *Environmental Research*, 198:111227, 2021.

[395] Josh M. Colston, Aftab M. S. Ahmed, Shafique B. Soofi, Erik Svensen, Rashidul Haque, Sourav Shrestha, Mahfuzar Rahman Islam, Muhammad Iqbal Nisar, Humayun Kabir, and Stephen P. Luby. Evaluating meteorological data from weather stations, and from satellites and global models for a multi-site epidemiological study. *Environmental Research*, 165:91–109, 2018.

[396] Pramiti Kumar Chakraborty, Pabitra Bandyopadhyay, Kalyan Kumar Jana, and Krishnendu Ray. Assessing congenial soil temperature and its impact on root growth and grain yield of summer rice under varying water stress condition in the lower Gangetic Plain of India. *Journal of the Saudi Society of Agricultural Sciences*, 2021.

[397] Yanyan Jia, Weixing Cheng, Qian Wang, Lijun Yang, and Xiaojun Zhang. Effects of low water temperature during reproductive growth on photosynthetic production and nitrogen accumulation in rice. *Field Crops Research*, 242:107587, 2019.

[398] Tsuneo Kuwagata, Mari Murai-Hatano, Maya Matsunami, Shingo Terui, Atsushi J. Nagano, Atsushi Maruyama, and Sachinobu Ishida. Hydrometeorology for plant omics: Potential evaporation as a key index for transcriptome in rice. *Environmental and Experimental Botany*, 196:104724, 2022.

[399] Yongkang Tang, Shuangsheng Guo, Wenping Dong, Lifeng Qin, Weidang Ai, and Shan Lin. Effects of long-term low atmospheric pressure on gas exchange and growth of lettuce. *Advances in Space Research*, 46(6):751–760, 2010.

[400] Ratneswar Poddar, Biswajit U. Choudhury, Debasish Chakraborty, A. K. Singh, R. Singh, and Mangi Lal Jat. Effect of irrigation regime and varietal selection on the yield, water productivity, energy indices and economics of rice production in the lower Gangetic Plains of eastern India. *Agricultural Water Management*, 251:107327, 2021.

[401] Santanu Kumar Bal, B. K. Bhattacharya, V. M. Sandeep, D. K. Kundu, V. U. M. Rao, and B. Venkateswarlu. Critical weather limits for paddy rice under diverse ecosystems of India. *Frontiers in Plant Science*, 14:1226064, 2023.

[402] José Roberto Alvarado. Influence of air temperature on rice population, length of period from sowing to flowering, and spikelet sterility. In J. E. Hill and B. Hardy, editors, *Proceedings of the Second Temperate Rice Conference*, pages 63–68, Sacramento, California, USA, 1999. International Rice Research Institute.

[403] Leah H. Schinasi, Tarik Benmarhnia, and Anneclaire J. De Roos. Modification of the association between high ambient temperature and health by urban microclimate indicators: a systematic review and meta-analysis. *Environmental Research*, 161:168–180, 2018.

[404] Mariano F. Lopresti, Carlos M. Di Bella, and Américo J. Degioanni. Relationship between MODIS-NDVI data and wheat yield: a case study in northern Buenos Aires Province, Argentina. *Information Processing in Agriculture*, 2(2):73–84, 2015.

[405] Ewa Panek and Dariusz Gozdowski. Analysis of relationship between cereal yield and NDVI for selected regions of Central Europe based on MODIS satellite data. *Remote Sensing Applications: Society and Environment*, 17:100286, 2020.

[406] Nguyen Thanh Son, Chi-Farn Chen, Cheng-Ru Chen, Vo Quang Minh, and Nguyen Huu Trung. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agricultural and Forest Meteorology*, 197:52–64, 2014.

[407] Fiona H. M. Tang, Thu Ha Nguyen, Giulia Conchedda, Leon Casse, Francesco N. Tubiello, and Federico Maggi. CROPGRIDS: a global geo-referenced dataset of 173 crops circa 2020. *Earth System Science Data*, 16:1333–1353, 2024.

[408] Ministry of Agriculture and Farmers Welfare. Crop production statistics information system. Online, 2021.

[409] Geetika Sonkar, R. K. Mall, Tirthankar Banerjee, Nidhi Singh, T. V. Lakshmi Kumar, and Ramesh Chand. Vulnerability of Indian wheat against rising temperature and aerosols. *Environmental Pollution*, 254:112946, 2019.

[410] Chaitali Diwan, Srinath Srinivasa, Gandharv Suri, Saksham Agarwal, and Prasad Ram. AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Computers and Education: Artificial Intelligence*, 4:100110, 2023.

[411] Global Administrative Areas. GADM database of global administrative areas, version 2.0. Online, 2012. Accessed on February 14, 2024.

[412] Longfei Zhou, Ran Meng, Xing Yu, Yigui Liao, Zehua Huang, Zhengang Lü, Binyuan Xu, Guodong Yang, Shaobing Peng, and Le Xu. Improved yield prediction of ratoon rice using unmanned aerial vehicle-based multi-temporal feature method. *Rice Science*, 30(3):247–256, 2023.

[413] Vitor Cerqueira, Luis Torgo, and Igor Mozetič. Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, 109(11):1997–2028, November 2020.

[414] Moshe Sipper. High per parameter: a large-scale study of hyperparameter tuning for machine learning algorithms. *Algorithms*, 15(9):315, September 2022.

[415] Scott M. Lundberg and Su-In Lee. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS '17, pages 4768–4777, Red Hook, NY, USA, 2017. Curran Associates Inc.

[416] Ian Lenaers and Lieven De Moor. Exploring XAI techniques for enhancing model transparency and interpretability in real estate rent prediction: A comparative study. *Finance Research Letters*, 58:104306, December 2023.

[417] Yahui Guo, Yongshuo Fu, Fanghua Hao, Xuan Zhang, Wenxiang Wu, Xiuliang Jin, Christopher Robin Bryant, and J. Senthilnath. Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, 120:106935, 2021.

[418] Kersten Clauss, Marco Ottinger, Patrick Leinenkugel, and Claudia Kuenzer. Estimating rice production in the Mekong Delta, Vietnam, utilizing time series of Sentinel-1 SAR data. *International Journal of Applied Earth Observation and Geoinformation*, 73:574–585, 2018.

[419] Jingye Han, Liangsheng Shi, Qi Yang, Zhuowei Chen, Jin Yu, and Yuanyuan Zha. Rice yield estimation using a CNN-based image-driven data assimilation framework. *Field Crops Research*, 288:108693, 2022.

[420] Xi Su, Jun Zhou, Xiangyu Lu, Jianxun Shen, Huizhe Chen, Mengyuan Chen, Yong He, and Fei Liu. Grain yield prediction using multi-temporal UAV-based multispectral vegetation indices and endmember abundance in rice. *Field Crops Research*, 299:108992, 2023.

[421] Md. Monirul Islam, Shusuke Matsushita, Ryozo Noguchi, and Tofael Ahamed. Development of remote sensing-based yield prediction models at the maturity stage of Boro rice using parametric and nonparametric approaches. *Remote Sensing Applications: Society and Environment*, 22:100494, 2021.

[422] Xiang Zhou, Hengbiao Zheng, Xiaoqing Xu, Jingyi He, Xiukai Ge, Xia Yao, Tao Cheng, Yan Zhu, Weixing Cao, and Yongchao Tian. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130:246–255, 2017.

[423] Tri D. Setiyono, E. D. Quicho, Leonardo Gatti, Manuel Campos-Taberner, Lorenzo Busetto, Filippo Collivignarelli, and Roberto Confalonieri. Spatial rice yield estimation based on MODIS and Sentinel-1 SAR data and ORYZA crop growth model. *Remote Sensing*, 10(2):293, 2018.

[424] Senlin Guan, Katsura Fukami, Tetsuro Matsunaka, Masahiro Okada, and Masayuki Nakagawa. Assessing correlation of high-resolution NDVI with fertilizer application level and yield of rice and wheat crops using small UAVs. *Remote Sensing*, 11(2):112, 2019.

[425] Vitor Cerqueira, Nuno Moniz, and Carlos Soares. VEST: automatic feature engineering for forecasting. *Machine Learning*, 2021.

[426] Jiaping Liang, Xinyu Liu, Zhiyuan Li, Zhen Wang, Yuxin Zhang, and Jun Zhao. Analysis and prediction of the impact of socio-economic and meteorological factors on rapeseed yield based on machine learning. *Agronomy*, 13(7):1867, 2023.

[427] Roel Jongeneel and Ana Rosa Gonzalez-Martinez. Estimating crop yield supply responses to be used for market outlook models: application to major developed and developing countries. *NJAS - Wageningen Journal of Life Sciences*, 92, 2020.

[428] Vololonirina Raharimanana, Tomoaki Yamaguchi, Yasuhiro Tsujimoto, Aung Zaw Oo, Tomohiro Nishigaki, Hobimiarantsoa Rakotonindrina, and Keisuke Katsura. A machine learning approach is effective to elucidate yield-limiting factors of irrigated lowland rice under heterogeneous growing conditions and management practices. *Field Crops Research*, 304:109170, December 2023.

[429] Surabhi Lingwal, Komal Kumar Bhatia, and Manjeet Singh. A novel machine learning approach for rice yield estimation. *Journal of Experimental & Theoretical Artificial Intelligence*, April 2024.

[430] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning.* Springer New York, 2009.

[431] Zeenatul Islam, Mohammad Alauddin, and Md. Abdur Rashid Sarker. Determinants and implications of crop production loss: an empirical exploration using ordered probit analysis. *Land Use Policy*, 67:527–536, 2017.

[432] Vipul Mann, Abhishek Sharma, Pawan Kumar, and Ranjan Bhattacharya. SUSIE: pharmaceutical CMC ontology-based information extraction for drug development using machine learning. *Computers & Chemical Engineering*, 179:108446, 2023.

[433] Han Zhang, Arik Pfisterer, Zihan Xie, Joris Kossen, Xuezhi Kan, Jonas Mueller, Frank Hutter, and Yuyang Wang. Towards foundation models for learning on tabular data. *arXiv*, 2023.

[434] Bruno Silva, Pedro Costa, Pedro Martins, and Tiago Santos. GPT-4 as an agronomist assistant? Answering agriculture exams using large language models. *arXiv*, 2023.

[435] Asaf Tzachor, Medha Devare, Catherine Richards, Pieter Pypers, Aniruddha Ghosh, Jawoo Koo, Simran Johal, and Brian King. Large language models and agricultural extension services. *Nature Food*, 4(11):941–948, 2023.

[436] Clare Allen-Sader, William Thurston, Maria Meyer, Michael Barbetti, Lucy Burgin, and James Brown. An early warning system to predict and mitigate wheat rust diseases in Ethiopia. *Environmental Research Letters*, 14(11):115004, 2019.

[437] Parshuram Samal, Sushil Pandey, S. R. Das, and Satyapriya Barik. Rice ecosystems and factors affecting varietal adoption in rainfed coastal Orissa: a multivariate probit analysis. *Agricultural Economics Research Review*, 24(1):131–140, 2011.

[438] Jagadish Timsina, Sudarshan Dutta, Krishna Prasad Devkota, Somsubhra Chakraborty, Ram Krishna Neupane, Sudarshan Bishta, Lal Prasad Amgain, Vinod K. Singh, Saiful Islam, and Kaushik Majumdar. Improved nutrient management in cereals using Nutrient Expert and machine learning tools: productivity, profitability and nutrient use efficiency. *Agricultural Systems*, 192:103181, August 2021.

[439] Everett M. Rogers. *Diffusion of innovations*. Simon and Schuster, 2010.

[440] Frank M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5):215–227, 1969.

[441] Andrijana Horvat, Vincenzo Fogliano, and Pieternel A. Luning. Modifying the Bass diffusion model to study adoption of radical new foods – the case of edible insects in the Netherlands. *PLOS ONE*, 15(6):e0234538, 2020.

[442] A. Siddiqi, J. Tran, and J. L. Jr. Wescoat. Modeling growth and diffusion of groundwater pumping at multiple sub-provincial scales. In *AGU Fall Meeting Abstracts*, volume 2018, pages GC51I–0906, 2018.

[443] Sabine E. Grimm, John W. Stevens, and Simon Dixon. Estimating future health technology diffusion using expert beliefs calibrated to an established diffusion model. *Value in Health*, 21(8):944–950, 2018.

[444] Yuyang Xian, Hongtao Zhang, Shuang Liu, and Yanjun Wang. Research on the market diffusion of fuel cell vehicles in China based on the generalized Bass model. *IEEE Transactions on Industry Applications*, 58(2):2950–2960, 2022.

[445] Hendrigo Batista da Silva, Wadaed Uturbey, and Bruno M. Lopes. Market diffusion of household PV systems: insights using the Bass model and solar water heaters market data. *Energy for Sustainable Development*, 55:210–220, 2020.

[446] Tiago P. Abud, Wadaed Uturbey, and Bruno M. Lopes. A modified Bass model to calculate PVDG hosting capacity in LV networks. *Electric Power Systems Research*, 209:107966, 2022.

[447] Walter J. Gutjahr and Pamela C. Nolz. Multicriteria optimization in humanitarian aid. *European Journal of Operational Research*, 252(2):351–366, 2016.

[448] Begoña Vitoriano, Javier Montero, and Da Ruan. A multi-criteria optimization model for humanitarian aid distribution. *Journal of Global Optimization*, 51(2):189–208, 2011.

[449] Israa Ismail. A possibilistic mathematical programming model to control the flow of relief commodities in humanitarian supply chains. *Computers & Industrial Engineering*, 157:107305, 2021.

[450] Alvaro Alonso Acero Condor, Cesar Manuel Ramirez Castañeda, and José Antonio Taquía Gutiérrez. Optimization of humanitarian aid resource distribution time through mixed integer linear programming. In *Proceedings of the 2023 9th International Conference on Industrial and Business Engineering*, pages 191–197, New York, NY, USA, 2023. Association for Computing Machinery.

[451] Chawis Boonmee and Chompoonoot Kasemset. The multi-objective fuzzy mathematical programming model for humanitarian relief logistics. *Industrial Engineering & Management Systems*, 19(1):197–210, 2020.

[452] Ziaul Haq Adnan, Mohammad Rafiqul Islam, and Md. Mahfuzur Rahman. Applying linear programming for logistics distribution of essential relief items during COVID-19 lockdown: evidence from Bangladesh. *International Journal of Logistics Economics and Globalisation*, 9(3):191–204, 2022.

[453] İbrahim Miraç Eligüzel, Eren Özceylan, and Gerhard-Wilhelm Weber. Location-allocation analysis of humanitarian distribution plans: a case of United Nations Humanitarian Response Depots. *Annals of Operations Research*, 324(1):825–854, 2023.

[454] Peiman Ghasemi, Fariba Goodarzian, and Ajith Abraham. A new humanitarian relief logistic network for multi-objective optimization under stochastic programming. *Applied Intelligence*, 52(12):13729–13762, 2022.

[455] Koen Peters, Luk N. Van Wassenhove, Wim Van den Heuvel, and Ludo De Boeck. UN World Food Programme: toward zero hunger with analytics. *INFORMS Journal on Applied Analytics*, 52(1):8–26, 2022.

[456] Alice Devot, Lucie Royer, Blandine Arvis, Delphine Deryng, Elisa Caron Giauffret, Laetitia Giraud, Vincent Ayral, and Josselin Rouillard. Research for AGRI Committee – The impact of extreme climate events on agriculture production in the EU. Technical Report PE 733.115, European Parliament, Policy Department for Structural and Cohesion Policies, April 2023.

[457] M. Zenda. A systematic literature review on the impact of climate change on the livelihoods of smallholder farmers in South Africa. *Heliyon*, 10(18):e38162, September 2024.

[458] Amiya K. Chakravarty. Humanitarian response to disasters with funding uncertainty: alleviating deprivation with bridge finance. *Production and Operations Management*, 30(9):3284–3296, September 2021.

[459] Sandra Uwantege Hart, Catherine Jones, and Beatriz Carvalho. Anticipate, automate, accelerate: a framework for blockchain in anticipatory action. *International Journal of Disaster Risk Reduction*, 112:104644, October 2024.

[460] Luca Sergio Italo Parodi, Markus Enenkel, Niccolò Lombardi, and Joshua Ngaina. Anticipatory action for drought in the Sahel: an innovation for drought risk management or a buzzword? *Frontiers in Climate*, 6:1347519, March 2024.

[461] Kundan Pandey. Deficient monsoon rainfall hits paddy farmers in parts of north India. *Mongabay-India*, September 2024.

[462] Laura Harrison, Martin Landsfeld, Greg Husak, Frank Davenport, Shraddhanand Shukla, William Turner, Pete Peterson, and Chris Funk. Advancing early warning capabilities with CHIRPS-compatible NCEP GEFS precipitation forecasts. *Scientific Data*, 9(375):1–13, June 2022.

[463] Feyera A. Hirpa, Peter Salamon, Hylke E. Beck, Valerio Lorini, Lorenzo Alfieri, Ervin Zsoter, and Simon J. Dadson. Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. *Journal of Hydrology*, 566:595–606, November 2018.

[464] Sean Fox, Felix Agyemang, Laurence Hawker, and Jeffrey Neal. Integrating social vulnerability into high-resolution global flood risk mapping. *Nature Communications*, 15:3155, April 2024.

[465] Pritha Datta, Bhagirath Behera, and Dil Bahadur Rahut. Climate change and Indian agriculture: a systematic review of farmers' perception, adaptation, and transformation. *Environmental Challenges*, 8:100543, August 2022.

[466] Rituparna Palit. "Never seen anything like this": how Cyclone Yaas impacted farmers' livelihoods in Bengal. *The Caravan*, July 2021.

[467] World Economic Forum. Income insurance: an innovative tool for climate resilience in India. *World Economic Forum*, August 2024. Accessed: 2024-11-25.

[468] Ranjay K. Singh, Anshuman Singh, Satyendra Kumar, Parvender Sheoran, D. K. Sharma, Lindsay C. Stringer, Claire H. Quinn, Arvind Kumar, and Dheeraj Singh. Perceived climate variability and compounding stressors: implications for risks to livelihoods of smallholder Indian farmers. *Environmental Management*, 66(5):826–844, November 2020.

[469] Mousumi Bhattacharjee and Suresh Chandra Babu. Early warning systems and climate resiliency: recent evidence and digital solutions for India's potato belt. *ANH Academy*, August 2024. Accessed: 2024-11-29.

[470] Srashta Chowdhury, Rwittambhara Kashyap, Pangkhi Medhi, Mohamed Meharaj Bano, Biswajit Paul, and Madhulika Sahoo. Eastern belt report: colleges and universities offering risk, crisis, disaster, and development-related programmes in the Eastern Belt of India, October 2024. Edited by Alexander Skinner, Nibedita S. Ray-Bennett, and Hideyuki Shiroshita.

[471] Moumita Alam. Bengal's potato farmers caught in a web of loans, losses, and political neglect. *Frontline*, October 2023. Accessed: 2024-11-29.

[472] International Food Policy Research Institute (IFPRI). Global spatially-disaggregated crop production statistics data for 2020 version 1.0.0. Harvard Dataverse, V1, 2024. Accessed: 2025-01-27.

[473] Andrew J. Tatem. WorldPop, open data for spatial demography. *Scientific Data*, 4(170004):1–4, January 2017.

[474] Nitiprasad Namdeorao Jambhulkar, Biswajit Mondal, Jaiprakash Bisen, and Asit Kumar Pradhan. Growth and instability of rice production: a district level analysis in West Bengal, India. *ORYZA – An International Journal of Rice*, 60(3):487–494, September 2023.

[475] Ansuman Das and Pranaya Kumar Swain. Navigating the sea level rise: exploring the interplay of climate change, sea level rise, and coastal communities in India. *Environmental Monitoring and Assessment*, 196(11):1–16, November 2024.

[476] District Administration, Purba Medinipur. Agriculture data and reports. https://purbamedinipur.gov.in/agriculture/, 2025. Accessed: 2025-01-27.

[477] S. K. Ariful Hossain, Ismail Mondal, Sandeep Thakur, and Ayad M. Fadhil Al-Quraishi. Coastal vulnerability assessment of India's Purba Medinipur–Balasore coastal stretch: a comparative study using empirical models. *International Journal of Disaster Risk Reduction*, 77:103065, July 2022.

[478] Manishree Mondal, Ramu Guchhait, and Sk. Mustak. Analysis of ecological vulnerability behind the land conversion from agriculture to aquaculture of Purba Medinipur district in West Bengal, India. In *Advanced Remote Sensing for Urban and Landscape Ecology*, pages 251–281. Springer, Singapore, June 2023.

[479] Souvik Chakraborty, Chandi Rajak, and Subhasish Das. The contamination of groundwater and possible remedial measures in the coastal zone of Purba Medinipur, India. *Journal of Coastal Conservation*, 27(4):1–23, August 2023.

[480] Manotosh Das. Fishery in paddy fields in Purba Medinipur district, West Bengal, India. *Research Journal of Animal, Veterinary and Fishery Sciences*, 10(1):25–29, May 2022.

[481] Jasmin Parvin and Ashis Kumar Paul. Changing livelihood security index along the coastal belt of Purba Medinipur district, an assessment using spatial information systems. In *Crisis on the Coast and Hinterland*, pages 417–429. Springer, Cham, Switzerland, November 2023.

[482] Avinash Kishore, P. K. Joshi, and Divya Pandey. Drought, distress, and a conditional cash transfer programme to mitigate the impact of drought in Bihar, India. *Water International*, 40(3):417–431, June 2015.

[483] Caritas India. Caritas India's lifeline Wayanad: delivering timely relief and restoring hope, September 2024. Published September 27, 2024.

[484] Anticipation Hub. A pathway towards national frameworks for anticipatory action: highlights from the 1st South Asia Dialogue Platform, 2024. Accessed: 2024-11-29.

[485] Ashley Pople, Ruth Hill, Stefan Dercon, and Ben Brunckhorst. Anticipatory cash transfers in climate disaster response. Technical report, Centre for Disaster Protection, April 2021. Working Paper 6.

[486] Camila Laura Pareja Yale, Márcia Lorena da Silva Frazão, Marco Aurélio de Mesquita, and Hugo Tsugunobu Yoshizaki. Simulation applications in humanitarian logistics: a systematic literature review. In *Proceedings of the 2020 Winter Simulation Conference*, pages 2589–2598. IEEE, 2020.

[487] Alanoud Alotaibi and Farrukh Nadeem. A review of applications of linear programming to optimize agricultural solutions. *International Journal of Information Engineering and Electronic Business*, 13(2):11, April 2021.

[488] Ivet Galabova. *Presolve, crash and software engineering for HiGHS*. PhD thesis, The University of Edinburgh, Edinburgh, Scotland, UK, January 2023.

[489] John Forrest and Robin Lougee-Heimer. CBC user guide. *INFORMS Tutorials in Operations Research*, 2005(1):257–277, 2005.

[490] Frank Gurski and Jochen Rethmann. Distributed solving of mixed-integer programs with GLPK and Thrift. In *Operations Research Proceedings 2016*, pages 599–605. Springer, Cham, Switzerland, July 2017.

[491] Mohammad Khorasani, Mohamed Abdou, and Javier Hernández Fernández. *Web application development with Streamlit*. Apress, New York, NY, USA, 2022.

[492] Yaganteeswarudu Akkem, Saroj Kumar Biswas, and Aruna Varanasi. Streamlit-based enhancing crop recommendation systems with advanced explainable artificial intelligence for smart farming. *Neural Computing and Applications*, 36(32):20011–20025, November 2024.

[493] J. M. Nápoles-Duarte, Avratanu Biswas, Mitchell I. Parker, J. P. Palomares-Baez, M. A. Chávez-Rojo, and L. M. Rodríguez-Valdez. Stmol: a component for building interactive molecular visualizations within streamlit web-applications. *Frontiers in Molecular Biosciences*, 9:990846, September 2022.

[494] Mohammad Khorasani, Mohamed Abdou, and Javier Hernández Fernández. Streamlit use cases. In *Web Application Development with Streamlit*, pages 309–361. Apress, Berkeley, CA, USA, August 2022.

[495] A. S. Adith Sreeram, Jithendra Sai Pappuri, and Saladi Saritha. An efficient early detection of lung cancer and pneumonia with Streamlit. In *Intelligent Computing for Sustainable Development*, pages 152–164. Springer, Cham, Switzerland, May 2024.

[496] Render. Render – cloud hosting for developers. https://render.com/, 2024. Accessed: 2024-01-31.

[497] Radhika V. Kulkarni, Arjun Thakur, Shravani Nalbalwar, Shubham Shah, and Sankalp Chordia. Exploring scalable and efficient deployment of machine learning models: a comparative analysis of Amazon SageMaker and Heroku. In *2023 International Conference on Information Technology (ICIT)*, pages 746–751. IEEE, 2023.

[498] C. S. Srushti, P. M. Prathibhavani, A. Veena, and K. R. Venugopal. Prediction and analysis of protein 3D structures using protein language model and Streamlit. In *2024 International Conference on Information Management and Technology (ICIMTech)*, pages 636–641. IEEE, 2024.

[499] Noorullah Kuchai, Paul Shepherd, Juliana Calabria-Holley, Alexander Copping, Aude Matard, and David Coley. The potential for computational IT tools in disaster relief and shelter design. *International Journal of Humanitarian Action*, 5(1):1–20, December 2020.

[500] Manuel Bröchin, Bryn Pickering, Tim Tröndle, and Stefan Pfenninger. Harder, better, faster, stronger: understanding and improving the tractability of large energy system models. *Energy, Sustainability and Society*, 14(1):1–21, December 2024.

[501] Peddinti Sandeep, G. Peda Obi Reddy, Ramalingam Jegankumar, and K. Chandra Arun Kumar. Monitoring of agricultural drought in semi-arid ecosystem of Peninsular India through indices derived from time-series CHIRPS and MODIS datasets. *Ecological Indicators*, 121:107033, February 2021.

[502] Prakhar Deroliya, Mousumi Ghosh, Mohit P. Mohanty, Subimal Ghosh, K. H. V. Durga Rao, and Subhankar Karmakar. A novel flood risk mapping approach with machine learning considering geomorphic and socio-economic vulnerability dimensions. *Science of the Total Environment*, 851:158002, December 2022.

[503] Seungtaek Jeong, Jonghan Ko, and Jong-Min Yeom. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. *Science of the Total Environment*, 802:149726, January 2022.

[504] E. T. Bureau. India must invest more in flood warning systems: report. *Economic Times*, July 2023.

[505] Alan T. Murray, Richard L. Church, B. Amelia Pludow, and Peter Stine. Advancing contiguous environmental land allocation analysis, planning and modeling. *Journal of Land Use Science*, January 2022.

[506] Subhra Halder, Subhasish Das, and Snehamanju Basu. Land suitability analysis using a multicriteria decision-making approach for optimal irrigation in a medium irrigation scheme in West Bengal, India. *Irrigation and Drainage*, 72(4):1083–1094, October 2023.

[507] Roberto Da Silva Gervasio Pontes, Diego Nunes Brandão, Fábio Luiz Usberti, and Laura Silva De Assis. Multi-objective models for crop rotation planning problems. *Agricultural Systems*, 219:104050, August 2024.

[508] Assefa A. Berhanu, Zewdu B. Ayele, Dessalegn C. Dagnew, Tadele Melese, Abeje B. Fenta, and Koyachew E. Kassie. Smallholder farmers' vulnerability to climate change

and variability: Evidence from three agroecologies in the Upper Blue Nile, Ethiopia. *Heliyon*, 10(7):e28277, April 2024.

[509] Peter Kall and János Mayer. *Stochastic linear programming: models, theory, and computation.* Springer US, 2011.

[510] Özlem Ergun. Optimizing global food aid operations. BrightTALK Webinar, December 2024. Duration: 72 minutes.

[511] Francisca Santana Robles, Eva Selene Hernández-Gress, Neil Hernández-Gress, and Rafael Granillo Macias. Metaheuristics in the humanitarian supply chain. *Algorithms*, 14(12):364, December 2021.

[512] Mohd. Imran Khan. Bihar: Fear of drought-like situation haunts farmers due to deficient rainfall, July 2023. [Accessed 5 Feb. 2025].

[513] Jared L. Gearhart and Steven M. Adair. Comparison of open-source linear programming solvers. Technical report, Sandia National Laboratories, 2013.

[514] Ricardo Moreno-Chuquen and Sergio Cantillo-Luna. Assessment of a multiperiod optimal power flow for power system operation. *International Review of Electrical Engineering*, 15(6):484–492, December 2020.

[515] Punnoli Dhanya and Vellingiri Geethalakshmi. Reviewing the status of droughts, early warning systems and climate services in South India: experiences learned. *Climate*, 11(3):60, March 2023.

[516] Sara Matos Coelho Bernardo. Cash transfers and post-disaster agriculture: The 'post' is the 'pre', February 2022. [Online; accessed 4 Feb. 2025].

[517] Adrian Cernev. Innovations in conditional cash transfer programmes in the digital age: The payment of Bolsa Familia program via mobile phone, December 2015. [Online; accessed 4 Feb. 2025].

[518] Vandita Dar, Madhvi Sethi, and Saina Baby. Direct cash transfers in emerging economies: the case of India. *Business Perspectives and Research*, 11(2):287–308, May 2022.

[519] Zoë Scott. How can anticipatory action reach scale and sustainability? Learning from CERF in Bangladesh, August 2023. [Online; accessed 5 Feb. 2025].

[520] Evan Easton-Calabria, Shampa Shampa, Mecthilde Pinto, Selma Lendelvo, Sonia Binte Murshed, Leah Poole, Aarti Singh, Irene Amuron, Meghan Bailey, Sayanti Sengupta, and Erin Coughlan de Perez. Targeting aid before disasters occur: perspectives on anticipatory action and social protection in Bangladesh and Namibia, February 2023. [Online; accessed 5 Feb. 2025].

[521] Shurui Zhou, Bogdan Vasilescu, and Christian Kästner. What the fork: a study of inefficient and efficient forking practices in social coding. In *Proceedings of the 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2019)*, pages 350–361, New York, NY, USA, August 2019. Association for Computing Machinery.

[522] Adam Alami, Raúl Pardo, Marisa Leavitt Cohn, and Andrzej Wąsowski. Pull request governance in open source communities. *IEEE Transactions on Software Engineering*, 48(12):4838–4856, November 2021.

[523] Pavan Vovveti. The role of API security in modern enterprise platforms. *International Journal for Research in Applied Science and Engineering Technology*, 12(9):1384–1390, September 2024.

[524] Warren B. Powell. A unified framework for stochastic optimization. *European Journal of Operational Research*, 275(3):795–821, 2019.

[525] Warren B. Powell. From reinforcement learning to optimal control: a unified framework for sequential decisions. In *Handbook of Reinforcement Learning and Control*, pages 29–74. Springer, 2021.

[526] Michael Docherty and Richard Smith. The case for structuring the discussion of scientific papers: much the same as that for structuring abstracts. *BMJ*, 318(7193):1224–1225, 1999.

[527] Djavan De Clercq, Lily Xu, Marleen C. de Ruiter, Marc van den Homberg, Marijn van der Velde, Jim W. Hall, Jonas Jaegermeyr, and Adam Mahdi. Towards optimal anticipatory action: maximizing the effectiveness of agricultural early warning systems with operations research. *International Journal of Disaster Risk Reduction*, 119:105249, 2025.

[528] Djavan De Clercq and Adam Mahdi. Modern computational approaches for rice yield prediction: A systematic review of statistical and machine learning-based methods. *Computers and Electronics in Agriculture*, 231:109852, 2025.

[529] Pierre D. Glynn, Charles R. Rhodes, Scott J. Chiavacci, Jennifer F. Helgeson, Carl D. Shapiro, and Crista L. Straub. Value of information and decision pathways: concepts and case studies. *Frontiers in Environmental Science*, 10:805214, 2022.

[530] Havva Uyar, Ioannis Karvelas, Stamatia Rizou, and Spyros Fountas. Data value creation in agriculture: A review. *Computers and Electronics in Agriculture*, 227:109602, 2024.

[531] Weston Anderson, Shraddhanand Shukla, James Verdin, Andrew Hoell, Christopher Justice, Brian Barker, Kimberly Slinski, Nathan Lenssen, Jiale Lou, Benjamin I. Cook,

and Christopher Funk. Preseason maize and wheat yield forecasts for early warning of crop failure. *Nature Communications*, 15(1):7262, 2024.

[532] Majdi Argoubi and Khaled Mili. Stochastic multi-objective optimization for flood control in multi-reservoir systems: An adaptive progressive hedging approach with scenario clustering. *Frontiers in Water*, 7:1606096, 2025.

[533] Cassandra Howe and Jason A. Tullis. Context for reproducibility and replicability in geospatial unmanned aircraft systems. *Remote Sensing*, 14(17):4304, 2022.

[534] Sean R. Wilkinson, Meznah Aloqalaa, Khalid Belhajjame, Michael R. Crusoe, Bruno de Paula Kinoshita, Luiz Gadelha, Daniel Garijo, Ove Johan Ragnar Gustafsson, Nick Juty, Sehrish Kanwal, Stian Soiland-Reyes, and Carole Goble. Applying the FAIR principles to computational workflows. *Scientific Data*, 12(1):328, 2025.

[535] Oscar Rodríguez-Espíndola, Hossein Ahmadi, Diego Gastélum-Chavira, Omar Ahumada-Valenzuela, Soumyadeb Chowdhury, Prasanta Kumar Dey, and Pavel Albores. Humanitarian logistics optimization models: An investigation of decision-maker involvement and directions to promote implementation. *Socio-Economic Planning Sciences*, 89:101669, 2023.

[536] Patrick H. Curran, Charlotte Brown, Andrew P. Robinson, Anita Wreford, and Tom M. Logan. Evaluating deep uncertainty tools for complex decision problems. *Environmental Science & Policy*, 173:104225, 2025.

[537] Marjolijn Haasnoot, Valeria Di Fant, Jan Kwakkel, and Judy Lawrence. Lessons from a decade of adaptive pathways studies for climate adaptation. *Global Environmental Change*, 88:102907, 2024.

[538] International Federation of Red Cross and Red Crescent Societies. Strengthening national disaster risk management systems through the integration of anticipatory action. Technical report, IFRC, Geneva, Switzerland, 2025. Accessed October 2025.

[539] Government of the Republic of Fiji. National disaster risk management bill 2024 (Bill No. 6 of 2024). Parliamentary Paper, Office of the Attorney-General, 2024.

[540] Towrin Zaman, Khandker Tarin Tahsin, Savio Rousseau Rozario, Adiba Binte Kamal, Mizan R. Khan, Saleemul Huq, and Md. Bodrud-Doza. An overview of disaster risk reduction and anticipatory action in Bangladesh. *Frontiers in Climate*, 4:944736, 2022.

[541] Emily Wilkinson, Tobias Pforr, and Lena Weingärtner. Integrating anticipatory action in disaster risk management. Technical report, Overseas Development Institute, 2020.

[542] Emily Wilkinson, Lena Weingärtner, and Richard Choularton. Forecasting hazards, averting disasters: implementing forecast-based early action at scale. Technical report, Overseas Development Institute, 2018.

[543] Government of Sierra Leone. Disaster risk financing strategy and implementation plan 2024–2029. Technical report, Ministry of Finance, Government of Sierra Leone, Freetown, Sierra Leone, 2024.

[544] Assembleia da República de Moçambique. Lei n.º 10/2020, de 24 de Agosto: Lei de gestão e redução do risco de desastres. Boletim da República, I Série, Número 162, 2020.

[545] Emmah Mwangi, Oliver Taylor, and Martin C. Todd. Mainstreaming forecast based action into national disaster risk management systems: experience from drought risk management in Kenya. *Climate and Development*, 14(8):741–756, 2022.

[546] Krish W. Ramadurai and Sujata K. Bhatia. The humanitarian relief paradigm. In *Reimagining Innovation in Humanitarian Medicine*, pages 1–25. Springer, 2018.

[547] Maliha Muzammil, Karlijn Muiderman, and Joost Vervoort. Approaches through which anticipation informs climate governance in South Asia. Working paper no. 387, CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Wageningen, The Netherlands, 2021.

[548] Michael E. Martinet. The disaster after the disaster: the quicksand of post-disaster recovery funding. *Journal of Business Continuity and Emergency Planning*, 12(4):331–341, 2019.

[549] Emma Youds. The perfect conditions to get ahead of disasters? Introducing groundbreaking work on Fiji's anticipatory action framework, 2024. Accessed October 2025.

[550] Jared LeClerc and Susan Joslyn. The cry wolf effect and weather-related decision making. *Risk Analysis*, 35(3):385–395, 2015.

[551] Swedish International Development Cooperation Agency. Humanitarian brief: anticipatory action. Technical report, Sida Unit for Humanitarian Assistance, September 2024. Accessed October 2025.

[552] Lorenzo Alfieri, Andrea Libertino, Lorenzo Campo, Francesco Dottori, Simone Gabellani, Tommaso Ghizzoni, Alessio Masoero, Luca Rossi, Roberto Rudari, Nicola Testa, Elisa Trasforini, Asnake Amdihun, James Ouma, Yves Tramblay, Hao Wu, and Marco Massabò. Impact-based flood forecasting in the Greater Horn of Africa. *Natural Hazards and Earth System Sciences*, 24(1):199–224, 2024.

[553] Arielle Tozier de la Poterie, Eduardo Castro Jr, Hafizur Rahaman, Dorothy Heinrich, Yolanda Clatworthy, and Luis Mundorega. Anticipatory action to manage climate risks: Lessons from the Red Cross Red Crescent in Southern Africa, Bangladesh, and beyond. *Climate Risk Management*, 39:100476, 2023.

[554] Emmanuel Attoh and Giriraj Amarnath. A framework for addressing the interconnectedness of early warning to action and finance to strengthen multiscale institutional responses to climate shocks and disasters. *Climate Risk Management*, 47:100689, 2025.

[555] Anne Della Guardia, Milli Lake, and Pascale Schnitzer. Selective inclusion in cash transfer programs: Unintended consequences for social cohesion. *World Development*, 157:105922, 2022.

[556] Nidhi Tewathia, Anant Kamath, and P. Vigneswara Ilavarasan. Social inequalities, fundamental inequities, and recurring of the digital divide: Insights from India. *Technology in Society*, 61:101251, 2020.

[557] Pascal D. König and Georg Wenzelburger. The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society*, 67:101688, 2021.

[558] Ulrik B. U. Roehl. Automated decision-making and good administration: Views from inside the government machinery. *Government Information Quarterly*, 40(4):101864, 2023.

[559] Filip Geburczyk. Automated administrative decision-making under the influence of the GDPR: Early reflections and upcoming challenges. *Computer Law & Security Review*, 41:105538, 2021.

[560] Praduymn K. Tripathi and Theocharis N. Grigoriadis. State capacity and the soft budget constraint: Fiscal federalism, Indian style. *Asia and the Global Economy*, 3(2):100058, 2023.

[561] Farzana Afridi, Amrita Dhillon, and Arka Roy Chaudhuri. Measuring performance: Ranking state success over two decades in India. *Journal of Asian Economics*, 83:101550, 2022.

[562] Harpreet Kaur, Arjun Srinivas, and Amir Bazaz. Understanding access to agrarian knowledge systems: Perspectives from rural Karnataka. *Climate Services*, 21:100205, 2021.

[563] Gabriela Guimarães Nobre, Massimiliano Pasqui, Sara Quaresima, Silvia Pieretto, and Rogério Manuel Lemos Pereira Bonifácio. Forecasting, thresholds, and triggers: Towards developing a forecast-based financing system for droughts in Mozambique. *Climate Services*, 30:100344, 2023.

[564] Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 320(2):271–289, 2025.

[565] Murwan Siddig and Yongjia Song. Multistage stochastic programming with a random number of stages: Applications in hurricane disaster relief logistics planning. *European Journal of Operational Research*, 321(3):925–941, 2025.

[566] Frances E. Dunn, Marjolijn Haasnoot, Haomiao Du, Star Karabil, Philip S. J. Minderhoud, Vincent Schippers, Murray Scown, Annisa Triyanti, Trang Vu, and Hans Middelkoop. Mapping the solution space for local adaptation under global change: A test of concept for the Vietnamese Mekong delta. *Global Environmental Change*, 95:103071, 2025.

[567] Tim Busker, Hans de Moel, Bart van den Hurk, and Jeroen C. J. H. Aerts. Impact-based seasonal rainfall forecasting to trigger early action for droughts. *Science of the Total Environment*, 898:165506, 2023.

[568] Sally H. Potter, Thomas Kox, Brian Mills, Andrea Taylor, Joanne Robbins, Carolina Cerrudo, Faye Wyatt, Sara Harrison, Brian Golding, Will Lang, and Florian Pappenberger. Research gaps and challenges for impact-based forecasts and warnings: Results of international workshops for high impact weather in 2022. *International Journal of Disaster Risk Reduction*, 118:105234, 2025.

[569] Janne Parviainen, Stefan Hochrainer-Stigler, Lydia Cumiskey, Sukaina Bharwani, Pia-Johanna Schweizer, Benjamin Hofbauer, and Dug Cubie. The Risk-Tandem framework: An iterative framework for combining risk governance and knowledge co-production toward integrated disaster risk management and climate change adaptation. *International Journal of Disaster Risk Reduction*, 116:105070, 2025.

[570] Ursula Kenny and Aine Regan. Co-designing a smartphone app for and with farmers: Empathising with end-users' values and needs. *Journal of Rural Studies*, 82:148–160, 2021.

[571] Soham Adla, Aiswarya Aravindakshan, Ashray Tyagi, Ramesh Guntha, Mario Alberto Ponce-Pacheco, Anukool Nagi, Prashant Pastore, and Saket Pande. Participatory development of mobile agricultural advisory driven by behavioural determinants of adoption. *Journal of Environmental Management*, 374:124140, 2025.

[572] Marija Cerjak, Marco Medici, Ivica Faletar, Jada Venkata Sundeep, and Maurizio Canavari. Adoption of mobile-based agricultural extension services: evidence from south india. *Journal of rural studies*, 120:103851, 2025.

[573] Monica K Kansiime, Abdillahi Alawy, Catherine Allen, Manish Subharwal, Arun Jadhav, and Martin Parr. Effectiveness of mobile agri-advisory service extension model: Evidence from direct2farm program in india. *World development perspectives*, 13:25–33, 2019.

[574] Henry L. Stimson Center. The evolving landscape of anticipatory action in Nepal, 2026. Accessed: 2026-02-13.

[575] Red Cross Red Crescent Climate Centre. Anticipatory action closes the gap between "we knew" and "we acted", December 2025. Accessed: 2026-02-13.

[576] A. Clarke, D. Yates, C. Blanchard, M. Z. Islam, R. Ford, S. Rehman, and R. Walsh. The effect of dataset construction and data pre-processing on the extreme gradient boosting algorithm applied to head rice yield prediction in Australia. *Computers and Electronics in Agriculture*, 219:108716, April 2024.

[577] Rizza G. Baltazar. Forecasting the impact of climate change on rice crop yields under RCP4.5 and RCP8.5 scenarios in Central Luzon, Philippines, using machine learning algorithms. *International Journal of Agriculture and Natural Resources*, 51(1):10–26, April 2024.

[578] Tapash Kumar Sarkar, Dilip Kumar Roy, Ye Seong Kang, Sae Rom Jun, Jun Woo Park, and Chan Seok Ryu. Ensemble of machine learning algorithms for rice grain yield prediction using UAV-based remote sensing. *Journal of Biosystems Engineering*, 49(1):1–19, December 2023.

[579] Surabhi Lingwal, Komal Kumar Bhatia, and Manjeet Singh. A novel machine learning approach for rice yield estimation. *Journal of Experimental Theoretical Artificial Intelligence*, pages 1–20, April 2022.

[580] Mohd. Asif Gandhi, Lakhwinder Singh, V. Prabakaran, G. Bhupal Raj, Sowmiya, and Harshal Patil. An innovative method for paddy yield prediction based on DCNN-ELM approach. In *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*. IEEE, January 2024.

[581] Yunting Liang, Huifen Li, Han Wu, Yinwei Zhao, Zhijie Liu, Dong Liu, Zongru Liu, Gangao Fan, Zhaoyang Pan, Zhuo Shen, Jia Lv, and Jiyu Li. A rotated rice spike detection model and a crop yield estimation application based on UAV images. *Computers and Electronics in Agriculture*, 224:109188, September 2024.

[582] Xuning Chang. Rice yield prediction based on deep learning. In *Artificial Intelligence Technologies and Applications*. IOS Press, February 2024.

[583] Manasa Chitradurga Manjunath and Blessed Prince Pallayan. Artificial bee colony algorithm-based feature selection and hybrid ML framework for efficient rice yield prediction. *IJECES*, 15(3):235–246, March 2024.

[584] Che-Hao Chang, Jason Lin, Jia-Wei Chang, Yu-Shun Huang, Ming-Hsin Lai, and Yen-Jen Chang. Hybrid deep neural networks with multi-tasking for rice yield prediction using remote sensing data. *Agriculture*, 14(4):513, March 2024.

[585] Zhangxin Liu, Haoran Ju, Qiyun Ma, Chengming Sun, Yuping Lv, Kaihua Liu, Tianao Wu, and Minghan Cheng. Rice yield estimation using multi-temporal remote sensing data and machine learning: A case study of Jiangsu, China. *Agriculture*, 14(4):638, April 2024.

[586] Nguyen-Thanh Son, Chi-Farn Chen, Youg-Sin Cheng, Cheng-Ru Chen, Chien-Hui Syu, Yi-Ting Zhang, Shu-Ling Chen, and Shih-Hsiang Chen. Combining satellite data and artificial intelligence with a crop growth model to enhance rice yield estimation and crop management practices. *Applied Geomatics*, July 2024.

[587] Parul Setiya, Ajeet Singh Nain, and Anurag Satpathi. Comparative analysis of SMLR, ANN, elastic net and LASSO based models for rice crop yield prediction in Uttarakhand. *MAUSAM*, 75(1):191–196, December 2023.

[588] Yiyang Shen, Ziyi Yan, Yongjie Yang, Wei Tang, Jinqiu Sun, and Yanchao Zhang. Application of UAV-borne visible-infrared pushbroom imaging hyperspectral data for rice yield estimation using feature selection regression methods. *Sustainability*, 16(2):632, January 2024.

[589] Guodong Yang, Yaxing Li, Shen Yuan, Changzai Zhou, Hongshun Xiang, Zhenqing Zhao, Qiaorong Wei, Qingshan Chen, Shaobing Peng, and Le Xu. Enhancing direct-seeded rice yield prediction using UAV-derived features acquired during the reproductive phase. *Precision Agriculture*, 25(2):1014–1037, December 2023.

[590] S. Muthukumaran, P. Geetha, and E. Ramaraj. Multi-objective optimization with artificial neural network based robust paddy yield prediction model. *Intelligent Automation and Soft Computing*, 35(1):215–230, 2023.

[591] Poonsak Miphokasap. Rice yield estimations based on transformed surface reflectance from orbital hyperspectral remote sensing. *Chemical Engineering Transactions*, 106:829–834, December 2023.

[592] Chunling Sun, Hong Zhang, Lu Xu, Ji Ge, Jingling Jiang, Mingyang Song, and Chao Wang. Rice yield prediction using Sentinel-1 radar vegetation indices and XGBoost. In *2023 SAR in Big Data Era (BIGSARDATA)*, pages 20–22. IEEE, 2023.

[593] Xiaopan Wang, Xingmei Xi, Songjiang Duan, Zezhong Ma, Xia Long, and Rui Ji. Research on a hyperspectral rice yield estimation model based on random forest. In *2023 3rd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*. IEEE, 2023.

[594] Yahjeb Bouha Khatraty, Nédra Mellouli, Mamadou Tourad Diallo, and Mohamedade Farouk Nanne. Smart digital-twin hub concept for rice yield prediction and

monitoring from multivariate time series data. In *2023 24th International Conference on Control Systems and Computer Science (CSCS)*. IEEE, 2023.

[595] T. Chandrakumar, M. M. Avanthica Sri, K. Mirdula, and K. Monika. Paddy yield forecasting using regression techniques. In *2023 2nd IEEE Delhi Section Flagship Conference (DELCON)*. IEEE, 2023.

[596] K. M. I. S. Wanninayake, L. G. S. W. Bambaranda, T. I. Wickramaarachchi, U. C. S. L. Pathirana, Samitha Vidhanaarachchi, A. A. E. Nanayakkara, K. R. D. Gunapala, S. R. Sarathchandra, Aruna Ishara Gamage, and D. I. De Silva. IoT-enabled smart solution for rice disease detection, yield prediction, and remediation. In *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*. IEEE, 2023.

[597] Rahul Bhattacharyya, Nishit Raj, Siddharta Mukherjee, Prasasti Choudhury, and Rajat Gaur. Responsible AI for sustainable agriculture: forecasting rice yield to combat global food insecurity. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, July 2023.

[598] M. Chandraprabha and Rajesh Kumar Dhanraj. Ensemble deep learning algorithm for forecasting rice crop yield based on soil nutrition levels. *ICST Transactions on Scalable Information Systems*, page e7, May 2023.

[599] Hayam R. Seireg, Yasser M. Omar, and Ahmed Elmahalawy. Multi-agent system based on stacking technique for rice yield prediction. In *2023 3rd International Conference on Electronic Engineering (ICEEM)*. IEEE, October 2023.

[600] S. Thirumal and R. Latha. Automated rice crop yield prediction using sine cosine algorithm with weighted regularized extreme learning machine. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, May 2023.

[601] Aldian Nurcahyo, Haryono Soeparno, Lili Ayu Wulandari, and Widodo Budiharto. Rice yield prediction in Sumatra Indonesia using machine learning and climate data. In *2023 3rd International Conference on Intelligent Cybernetics Technology Applications (ICICyTA)*. IEEE, December 2023.

[602] Erio Yoshino, Felix Indra Kurniadi, and Budi Juarto. Forecasting rice production in Indonesia using regression techniques: a comparative analysis of support vector machine, linear regression, and XGBoost regression. In *2023 10th International Conference on ICT for Smart Society (ICISS)*. IEEE, September 2023.

[603] S. Thirumal and R. Latha. Automated hyperparameter tuned stacked autoencoder based rice crop yield prediction model. In *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*. IEEE, April 2023.

[604] Arepalli Peda Gopi, Varikallu Swathi, Guntupalli Sri Harshitha, Bachina Swetha, and Narra Alekhya. Prediction of paddy yield based on IoT data using GRU model in lowland coastal regions. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, January 2023.

[605] S. Senthil and K. Raja. Comparison of machine learning algorithms for the prediction of rice crop yield in Karnataka. In *2023 Fourth International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*. IEEE, December 2023.

[606] Mathadadoddi Nanjundegowda Thimmegowda, Melekote Hanumanthaiah Manjunatha, Lingaraj Huggi, Huchahanumegowdanapalya Sanjeevaiah Shivaramu, Dadireddihalli Venkatappa Soumya, Lingegowda Nagesha, and Hejjaji Sreekanthamurthy Padmashri. Weather-based statistical and neural network tools for forecasting rice yields in major growing districts of Karnataka. *Agronomy*, 13(3):704, February 2023.

[607] Ayana Takai, Hiroya Makino, and Eisuke Kita. SSE-based evolutionary algorithm for hyper-parameter optimization of LightGBM on paddy rice yield prediction problem. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, October 2023.

[608] Sumanya Ngandee and Attaphongse Taparugssanagorn. Improved information dissemination services for the agricultural sector in Thailand: development and evaluation of a machine learning based rice crop yield prediction system. *Information Development*, November 2023.

[609] Maria Cristina V. David. Rice yield modeling using machine learning algorithms based on environmental and agronomic data of Pampanga River Basin, Philippines. *Universal Journal of Agricultural Research*, 11(5):836–848, October 2023.

[610] Anurag Satpathi, Parul Setiya, Bappa Das, Ajeet Singh Nain, Prakash Kumar Jha, Surendra Singh, and Shikha Singh. Comparative analysis of statistical and machine learning techniques for rice yield forecasting for Chhattisgarh, India. *Sustainability*, 15(3):2786, February 2023.

[611] Mst Noorunnahar, Arman Hossain Chowdhury, and Farhana Arefeen Mila. A tree based extreme gradient boosting (XGBoost) machine learning model to forecast the annual rice production in Bangladesh. *PLOS ONE*, 18(3):e0283452, March 2023.

[612] Woranan Mongkolnithithada, Jurawan Nontapun, and Siwa Kaewplang. Rice yield estimation based on machine learning approaches using MODIS 250 m data. *Engineering Access*, 9(1):75–79, April 2023.

[613] S. Thirumal and R. Latha. Teaching and learning based optimization with deep learning model for rice crop yield prediction. *International Journal of Electrical and Electronics Engineering*, 10(4):105–114, April 2023.

[614] Vincent Junior Halim, Antonio Ryanlie, Kevin Djong, Karli Eka Setiawan, and Muhammad Fikri Hasani. Rice yield production forecasting using deep learning models. In *2023 International Conference on Networking, Electrical Engineering, Computer Science, and Technology (IConNECT)*. IEEE, August 2023.

[615] Long Zhao, Shunhao Qing, Fei Wang, Hui Wang, Hao Ma, Yi Shi, and Ningbo Cui. Prediction of rice yield based on multi-source data and hybrid LSSVM algorithms in China. *International Journal of Plant Production*, 17(4):693–713, August 2023.

[616] Alkha Mohan, M. Venkatesan, P. Prabhavathy, and A. Jayakrishnan. Temporal convolutional network based rice crop yield prediction using multispectral satellite data. *Infrared Physics and Technology*, 135:104960, December 2023.

[617] P. Sathya and P. Gnanasekaran. Paddy yield prediction in Tamilnadu delta region using MLR-LSTM model. *Applied Artificial Intelligence*, 37(1), February 2023.

[618] Sakshi Gandotra, Rita Chhikara, and Anuradha Dhull. Wheat, rice and corn yield prediction for Jammu district using machine learning techniques. In *ICT for Intelligent Systems*, pages 499–512. Springer Nature Singapore, 2023.

[619] Parul Setiya, Anurag Satpathi, and Ajeet Singh Nain. Predicting rice yield based on weather variables using multiple linear, neural networks, and penalized regression models. *Theoretical and Applied Climatology*, 154(1–2):365–375, July 2023.

[620] Aung Myint Htun, Md. Shamsuzzoha, and Tofael Ahamed. Rice yield prediction model using normalized vegetation and water indices from Sentinel-2A satellite imagery datasets. *Asia-Pacific Journal of Regional Science*, 7(2):491–519, May 2023.

[621] Bharati Patel and Aakanksha Sharaff. Rice variety classification and yield prediction using semantic segmentation of agro-morphological characteristics. *Multimedia Tools and Applications*, 82(29):45567–45584, May 2023.

[622] Baishun Liu, Ying Liu, Guangzhi Huang, Xiaotong Jiang, Yanping Liang, Can Yang, and Lihua Huang. Comparison of yield prediction models and estimation of the relative importance of main agronomic traits affecting rice yield formation in saline-sodic paddy fields. *European Journal of Agronomy*, 148:126870, August 2023.

[623] Longfei Zhou, Ran Meng, Xing Yu, Yigui Liao, Zehua Huang, Zhengang Lü, Binyuan Xu, Guodong Yang, Shaobing Peng, and Le Xu. Improved yield prediction of ratoon rice using unmanned aerial vehicle-based multi-temporal feature method. *Rice Science*, 30(3):247–256, May 2023.

[624] Subin Ha, Yong-Tak Kim, Eun-Soon Im, Jina Hur, Sera Jo, Yong-Seok Kim, and Kyo-Moon Shim. Impacts of meteorological variables and machine learning algorithms on rice yield prediction in Korea. *International Journal of Biometeorology*, 67(11):1825–1838, November 2023.

[625] Xiao Han, Fangbiao Liu, Xiaoliang He, and Fenglou Ling. Research on rice yield prediction model based on deep learning. *Computational Intelligence and Neuroscience*, 2022:1–9, April 2022.

[626] Pavan Kumar Sharma, Pratyush Kumar, Hari Shanker Srivastava, and Thota Sivasankar. Assessing the potentials of multi-temporal Sentinel-1 SAR data for paddy yield forecasting using artificial neural network. *Journal of the Indian Society of Remote Sensing*, 50(5):895–907, February 2022.

[627] K. Krupavath, M. Raghu Babu, and A. Mani. Comparative evaluation of neural networks in crop yield prediction of paddy and sugarcane crop. In *The Digital Agricultural Revolution*, pages 25–55. Wiley, May 2022.

[628] Nguyen-Thanh Son, Chi-Farn Chen, Youg-Sin Cheng, Piero Toscano, Cheng-Ru Chen, Shu-Ling Chen, Kuo-Hsin Tseng, Chien-Hui Syu, Horng-Yuh Guo, and Yi-Ting Zhang. Field-scale rice yield prediction from Sentinel-2 monthly image composites using machine learning algorithms. *Ecological Informatics*, 69:101618, July 2022.

[629] Kamal Choudhary, Wenzhong Shi, Yang Dong, and Robert Paringer. Random forest for rice yield mapping and prediction using Sentinel-2 data with Google Earth Engine. *Advances in Space Research*, 70(8):2443–2457, October 2022.

[630] Chen Gu, Shu Ji, Xiaobo Xi, Zhenghua Zhang, Qingqing Hong, Zhongyang Huo, Wenxi Li, Wei Mao, Haitao Zhao, Ruihong Zhang, Bin Li, and Changwei Tan. Rice yield estimation based on continuous wavelet transform with multiple growth periods. *Frontiers in Plant Science*, 13, July 2022.

[631] Maria Victoria Bascon, Tomohiro Nakata, Satoshi Shibata, Itsuki Takata, Nanami Kobayashi, Yusuke Kato, Shun Inoue, Kazuyuki Doi, Jun Murase, and Shunsaku Nishiuchi. Estimating yield-related traits using UAV-derived multispectral images to improve rice grain yield prediction. *Agriculture*, 12(8):1141, August 2022.

[632] Jingye Han, Liangsheng Shi, Qi Yang, Zhuowei Chen, Jin Yu, and Yuanyuan Zha. Rice yield estimation using a CNN-based image-driven data assimilation framework. *Field Crops Research*, 288:108693, November 2022.

[633] Yuanyuan Liu, Shaoqiang Wang, Jinghua Chen, Bin Chen, Xiaobo Wang, Dongze Hao, and Leigang Sun. Rice yield prediction and model interpretation based on satellite and

climatic indicators using a transformer method. *Remote Sensing*, 14(19):5045, October 2022.

[634] Jing Li, Le Yue, Qing Zhao, Xuesong Cao, Weihao Tang, Feiran Chen, Chuanxi Wang, and Zhenyu Wang. Prediction models on biomass and yield of rice affected by metal (oxide) nanoparticles using nano-specific descriptors. *NanoImpact*, 28:100429, October 2022.

[635] Dharmesh Verma, A. N. Singh, Manik H. Kalubarme, G. P. Saroha, Ritesh Sharma, and Brajendra. Varietal discrimination, acreage estimation and yield prediction of basmati aromatic rice in north-western India using satellite data. In *Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries*, pages 163–186. Springer International Publishing, 2022.

[636] Astrid Noviana Paradhita, Myrtana Pusparisti, Yeni Fajariyanti, Agustin Amborowati, Agreianti Puspitasari, Yohanes Wien Tineka, and Haryanto. Rice field production prediction using genetic algorithm (evidence from Sukoharjo district, Indonesia). In *2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS)*. IEEE, November 2022.

[637] Niyati Mishra, Sushruta Mishra, and Hrudaya Kumar Tripathy. Rice yield estimation using deep learning. In *Innovations in Intelligent Computing and Communication*, pages 379–388. Springer International Publishing, 2022.

[638] M. Geetha, R. C. Suganthe, S. K. Nivetha, R. Anju, R. Anuradha, and J. Haripriya. A time-series based yield forecasting model using stacked LSTM to predict the yield of paddy in Cauvery delta zone in Tamilnadu. In *2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*. IEEE, February 2022.

[639] M. Geetha, R. C. Suganthe, R. S. Latha, R. Anju, K. Sastimalar, and P. Shobana. Deep learning based yield prediction model to predict the yield of paddy in Cauvery delta region. In *2022 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, January 2022.

[640] Abid Nazir, Saleem Ullah, Zulfiqar Ahmad Saqib, Azhar Abbas, Asad Ali, Muhammad Shahid Iqbal, Khalid Hussain, Muhammad Shakir, Munawar Shah, and Muhammad Usman Butt. Estimation and forecasting of rice yield using phenology-based algorithm and linear regression model on Sentinel-II satellite data. *Agriculture*, 11(10):1026, October 2021.

[641] Juan Cao, Zhao Zhang, Fulu Tao, Liangliang Zhang, Yuchuan Luo, Jing Zhang, Jichong Han, and Jun Xie. Integrating multi-source data for rice yield prediction across

China using machine learning and deep learning approaches. *Agricultural and Forest Meteorology*, 297:108275, February 2021.

[642] Santosha Rathod, Amit Saha, Rahul Patil, Gabrijel Ondrasek, Channappa Gireesh, Madhyavenkatapura Siddaiah Anantha, Dhumannatarao Venkata Krishna Nageswara Rao, Nirmala Bandumula, Ponnuvel Senguttuvel, Arun Kumar Swarnaraj, Shaik N. Meera, Amtul Waris, Ponnuraj Jeyakumar, Brajendra Parmar, Pitchiahpillai Muthuraman, and Raman Meenakshi Sundaram. Two-stage spatiotemporal time series modelling approach for rice yield prediction and advanced agroecosystem management. *Agronomy*, 11(12):2502, December 2021.

[643] Aishika Banik, G. Raju, and Samiksha Shukla. Rice yield forecasting in West Bengal using hybrid model. In *Data Science and Security*, pages 222–231. Springer Singapore, 2021.

[644] Seungtaek Jeong, Jonghan Ko, and Jong-Min Yeom. Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea. *Science of the Total Environment*, 802:149726, January 2022.

[645] Ponraj Arumugam, Abel Chemura, Bernhard Schauberger, and Christoph Gornott. Remote sensing based yield estimation of rice (Oryza sativa L.) using gradient boosted regression in India. *Remote Sensing*, 13(12):2379, June 2021.

[646] Yahui Guo, Yongshuo Fu, Fanghua Hao, Xuan Zhang, Wenxiang Wu, Xiuliang Jin, Christopher Robin Bryant, and J. Senthilnath. Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, 120:106935, January 2021.

[647] E. M. P. Ekanayake, L. C. D. Wickramasinghe, and R. T. Weliwatta. Use of regression techniques for rice yield estimation in the North-Western province of Sri Lanka. *Ceylon Journal of Science*, 50(4):439, December 2021.

[648] Subhadra Mishra, Debahuti Mishra, Pradeep Kumar Mallick, Gour Hari Santra, and Sachin Kumar. A classifier ensemble approach for prediction of rice yield based on climatic variability for coastal Odisha region of India. *Informatica*, 45(3), September 2021.

[649] Subhadra Mishra, Debahuti Mishra, Pradeep Kumar Mallick, Gour Hari Santra, and Sachin Kumar. A novel Borda count based feature ranking and feature fusion strategy to attain effective climatic features for rice yield prediction. *Informatica*, 45(1), March 2021.

[650] Daichao Li, Jianqin Liang, Xingfeng Wang, Sheng Wu, Xiaowei Xie, and Jiaqi Lu. Rice yield simulation and planting suitability environment pattern recognition at a fine scale. *ISPRS International Journal of Geo-Information*, 10(9):612, September 2021.

[651] Haoming Mo, Ying Zhang, Yifei Liu, and Yanzi Zheng. Prediction of rice yield based on LSTM long short term memory network. *Journal of Physics: Conference Series*, 1952(4):042033, June 2021.

[652] Ruben Fernandez-Beltran, Tina Baidar, Jian Kang, and Filiberto Pla. Rice-yield prediction with multi-temporal Sentinel-2 data and 3D CNN: A case study in Nepal. *Remote Sensing*, 13(7):1391, April 2021.

[653] Yeseong Kang, Jinwoo Nam, Younggwang Kim, Seongtae Lee, Deokgyeong Seong, Sihyeong Jang, and Chanseok Ryu. Assessment of regression models for predicting rice yield and protein content using unmanned aerial vehicle-based multispectral imagery. *Remote Sensing*, 13(8):1508, April 2021.

[654] P. Devaki, Prekshaa Arunachalam, K. S. Arun Sankar, and M. Venu Prasanna. A deep learning approach for yield estimation and phenotype analysis in rice crops. In *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*. IEEE, 2021.

[655] Vinson Joshua, Selwin Mich Priyadharson, and Raju Kannadasan. Exploration of machine learning approaches for paddy yield prediction in Eastern part of Tamilnadu. *Agronomy*, 11(10):2068, October 2021.

[656] Belen Franch, Alberto San Bautista, David Fita, Constanza Rubio, Daniel Tarrazo-Serrano, Antonio Sanchez, Sergii Skakun, Eric Vermote, Inbal Becker-Reshef, and Antonio Uris. Within-field rice yield estimation based on Sentinel-2 satellite data. *Remote Sensing*, 13(20):4095, October 2021.

[657] Nitin Bhardwaj, Muhammed Jaslam P. K. Jaslam, Jitender Kumar Bhatia, Bhupender Parashar, and Salinder. Neural network autoregression and classical time series approaches for rice yield forecasting. *Journal of Animal and Plant Sciences*, 31(4):1126–1134, December 2020.

[658] Yuki Sofue, Chiharu Hongo, Naohiro Manago, Gunardi Sigit, Koki Homma, and Baba Barus. Estimation of normal rice yield considering heading stage based on observation data and satellite imagery. In *2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, July 2021.

[659] Yaomin Wang, Haorui Chen, Junying Chen, Huiyun Wang, Zheng Xing, and Zhitao Zhang. Comparison of rice yield estimation models combining spectral index screening method and statistical regression algorithm. *Transactions of the Chinese Society of Agricultural Engineering*, 37(21):208–216, November 2021.

[660] Md. Monirul Islam, Shusuke Matsushita, Ryozo Noguchi, and Tofael Ahamed. Development of remote sensing-based yield prediction models at the maturity stage of boro rice

using parametric and nonparametric approaches. *Remote Sensing Applications: Society and Environment*, 22:100494, April 2021.

[661] Haixiao Ge, Fei Ma, Zhenwang Li, and Changwen Du. Grain yield estimation in rice breeding using phenological data and vegetation indices derived from UAV images. *Agronomy*, 11(12):2439, November 2021.

[662] Lasini Wickramasinghe, Rukmal Weliwatta, Piyal Ekanayake, and Jeevani Jayasinghe. Modeling the relationship between rice yield and climate variables using statistical and machine learning techniques. *Journal of Mathematics*, 2021:1–9, February 2021.

[663] Nourollah Ahmadi, Tuong-Vi Cao, Julien Frouin, Gareth J. Norton, and Adam H. Price. Genomic prediction of arsenic tolerance and grain yield in rice: contribution of trait-specific markers and multi-environment models. *Rice Science*, 28(3):268–278, May 2021.

[664] Surjeet Kumar, Manas Kumar Sanyal, and Asit Naskar. Prediction of rice production in India using artificial neural network with genetic algorithm. In *2020 International Conference on Computer Science, Engineering and Applications*, pages 13–14. IEEE, 2020.

[665] Watcharee Veerakachen and Mongkol Raksapatcharawong. RiceSAP: An efficient satellite-based AquaCrop platform for rice crop monitoring and yield prediction on a farm- to regional-scale. *Agronomy*, 10(6):858, June 2020.

[666] B. M. Refat Faisal, Hafizur Rahman, Nur Hossain Sharifee, Nasrin Sultana, Mohammad Imrul Islam, S. M. Ahsan Habib, and Tofayel Ahammad. Integrated application of remote sensing and GIS in crop information system—A case study on Aman rice production forecasting using MODIS-NDVI in Bangladesh. *AgriEngineering*, 2(2):264–279, May 2020.

[667] Nguyen-Thanh Son, Chi-Farn Chen, Cheng-Ru Chen, Horng-Yuh Guo, Youg-Sing Cheng, Shu-Ling Chen, Huan-Sheng Lin, and Shih-Hsiang Chen. Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan. *International Journal of Remote Sensing*, 41(20):7868–7888, August 2020.

[668] Li Tian, Chun Wang, Hailiang Li, and Haitian Sun. Yield prediction model of rice and wheat crops based on ecological distance algorithm. *Environmental Technology and Innovation*, 20:101132, November 2020.

[669] Liang Wan, Haiyan Cen, Jiangpeng Zhu, Jiafei Zhang, Yueming Zhu, Dawei Sun, Xiaoyue Du, Li Zhai, Haiyong Weng, Yijian Li, Xiaoran Li, Yidan Bao, Jianyao Shou, and Yong He. Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer. *Agricultural and Forest Meteorology*, 291:108096, September 2020.

[670] Xiangyan Meng, Muyan Liu, and Qiufeng Wu. Prediction of rice yield via stacked LSTM. *International Journal of Agricultural and Environmental Information Systems*, 11(1):86–95, January 2020.

[671] Gopal Chandra Paul, Sunil Saha, and Tusar Kanti Hembram. Application of phenology-based algorithm and linear regression model for estimating rice cultivated areas and yield using remote sensing data in Bansloi River Basin, Eastern India. *Remote Sensing Applications: Society and Environment*, 19:100367, August 2020.

[672] Seydou Traore, Lei Zhang, Aytac Guven, and Guy Fipps. Rice yield response forecasting tool (YIELDCAST) for supporting climate change adaptation decision in Sahel. *Agricultural Water Management*, 239:106242, September 2020.

[673] Yang Xu, Yue Zhao, Xin Wang, Ying Ma, Pengcheng Li, Zefeng Yang, Xuecai Zhang, Chenwu Xu, and Shizhong Xu. Incorporation of parental phenotypic data into multi-omic models improves prediction of yield-related traits in hybrid rice. *Plant Biotechnology Journal*, 19(2):261–272, September 2020.

[674] Akira Hama, Kei Tanaka, Atsushi Mochizuki, Yasuo Tsuruoka, and Akihiko Kondoh. Improving the UAV-based yield estimation of paddy rice by using the solar radiation of geostationary satellite Himawari-8. *Hydrological Research Letters*, 14(1):56–61, 2020.

[675] Vinushi Amaratunga, Lasini Wickramasinghe, Anushka Perera, Jeevani Jayasinghe, and Upaka Rathnayake. Artificial neural network to estimate the paddy yield prediction using climatic data. *Mathematical Problems in Engineering*, 2020:1–11, July 2020.

[676] W. M. R. K Wanninayaka, R. M. K. T Rathnayaka, and E. P. N Udayakumara. Artificial neural network to estimate the paddy yield prediction using remote sensing, weather and non-weather variables in Ampara district, Sri Lanka. In *2020 5th International Conference on Information Technology Research (ICITR)*, pages 1–6. IEEE, 2020.

[677] Zheng Chu and Jiong Yu. An end-to-end model for rice yield prediction using deep learning fusion. *Computers and Electronics in Agriculture*, 174:105471, July 2020.

[678] Md. Samiul Alam, Kazi Kalpoma, Md. Sanaul Karim, Abdullah Al Sefat, and Jun-ichi Kudoh. Boro rice yield estimation model using MODIS NDVI data for Bangladesh. In *IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, July 2019.

[679] A. Chandra, P. Mitra, S. K. Dubey, and S. S. Ray. Machine learning approach for Kharif rice yield prediction integrating multi-temporal vegetation indices and weather variables. *ISPRS Archives*, XLII-3/W6:187–194, July 2019.

[680] Avinash Kumar Ranjan and Bikash Ranjan Parida. Paddy acreage mapping and yield prediction using Sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). *Spatial Information Research*, 27(4):399–410, January 2019.

[681] Yi-Shiang Shiu and Yung-Chung Chuang. Yield estimation of paddy rice based on satellite imagery: comparison of global and local regression models. *Remote Sensing*, 11(2):111, January 2019.

[682] Tarno, A. Rusgiyono, and Sugito. Adaptive neuro fuzzy inference system (ANFIS) approach for modeling paddy production data in Central Java. *Journal of Physics: Conference Series*, 1217(1):012083, May 2019.

[683] B. M. Refat Faisal, Hafizur Rahman, Nur Hossain Sharifee, Nasrin Sultana, Mohammad Imrul Islam, and Tofayel Ahammad. Remotely sensed boro rice production forecasting using MODIS-NDVI: A Bangladesh perspective. *AgriEngineering*, 1(3):356–375, July 2019.

[684] Ranjini B. Guruprasad, Kumar Saurav, and Sukanya Randhawa. Machine learning methodologies for paddy yield estimation in India: A case study. In *IGARSS 2019 - IEEE International Geoscience and Remote Sensing Symposium*, pages 7254–7257. IEEE, 2019.

[685] Feilong Wang, Fumin Wang, Yao Zhang, Jinghui Hu, Jingfeng Huang, Lili Xie, and Jingkai Xie. Rice yield estimation at pixel scale using relative vegetation indices from unmanned aerial systems. In *2019 International Conference on Agro-Geoinformatics*, pages 16–19. IEEE, 2019.

[686] Sanqin Zhao, Haonan Zheng, Mingmei Chi, Xicun Chai, and Yutao Liu. Rapid yield prediction in paddy fields based on 2D image modelling of rice panicles. *Computers and Electronics in Agriculture*, 162:759–766, July 2019.

[687] Lei Zhang, Seydou Traore, Jiankun Ge, Yanbin Li, Shunsheng Wang, Ge Zhu, Yuanlai Cui, and Guy Fipps. Using boosted tree regression and artificial neural networks to forecast upland rice yield under climate change in Sahel. *Computers and Electronics in Agriculture*, 166:105031, November 2019.

[688] Liang Wan, Haiyan Cen, Jiangpeng Zhu, Yijian Li, Yueming Zhu, Dawei Sun, Haiyong Weng, and Yong He. Combining UAV-based vegetation indices, canopy height and canopy coverage to improve rice yield prediction under different nitrogen levels. In *ASABE Annual International Meeting*. American Society of Agricultural and Biological Engineers, 2019.

[689] Muhammad Alkaff, Husnul Khatimi, Wenny Puspita, and Yuslena Sari. Modelling and predicting wetland rice production using support vector regression. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 17(2):819, April 2019.

[690] Qi Yang, Liangsheng Shi, Jinye Han, Yuanyuan Zha, and Penghui Zhu. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Research*, 235:142–153, April 2019.

[691] Qi Yang, Liangsheng Shi, and Lin Lin. Plot-scale rice grain yield estimation using UAV-based remotely sensed images via CNN with time-invariant deep features decomposition. In *IGARSS 2019 - IEEE International Geoscience and Remote Sensing Symposium*, pages 7180–7183. IEEE, 2019.

[692] Nur Syazwan Wahab, Mohd Saifullah Rusiman, Mahathir Mohamad, Nur Amira Azmi, Norziha Che Him, M. Ghazali Kamardan, and Maselan Ali. A technique of fuzzy C-mean in multiple linear regression model toward paddy yield. *Journal of Physics: Conference Series*, 995:012010, April 2018.

[693] Jong-Won Ma, Cong-Hieu Nguyen, Kyungdo Lee, and Joon Heo. Regional-scale rice-yield estimation using stacked auto-encoder with climatic and MODIS data: a case study of South Korea. *International Journal of Remote Sensing*, 40(1):51–71, November 2018.

[694] Bappa Das, Bhakti Nair, Viswanatha K. Reddy, and Paramesh Venkatesh. Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India. *International Journal of Biometeorology*, 62(10):1809–1822, July 2018.

[695] Jin-Ki Park, Amrita Das, and Jong-Hwa Park. Integrated model for predicting rice yield with climate change. *International Agrophysics*, 32(2):203–215, April 2018.

[696] Md Nasim Reza, In Seop Na, Sun Wook Baek, and Kyeong-Hwan Lee. Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images. *Biosystems Engineering*, 177:109–121, January 2019.

[697] Md Nasim Reza, In Seop Na, Sun Wook Baek, and Kyeong-Hwan Lee. Automatic rice yield estimation using image processing technique. In *Intelligent Environments 2017*, pages 59–68. IOS Press, 2017.

[698] Alireza Taheri-Rad, Mehdi Khojastehpour, Abbas Rohani, Surur Khoramdel, and Amin Nikkhah. Energy flow modeling and predicting the yield of Iranian paddy cultivars using artificial neural networks. *Energy*, 135:405–412, September 2017.

[699] Anshul Garg and Bindu Garg. A robust and novel regression based fuzzy time series algorithm for prediction of rice yield. In *2017 International Conference on Intelligent Communication and Computational Techniques*, pages 48–54. IEEE, 2017.

[700] Niketa Gandhi, Leisa J. Armstrong, Owaiz Petkar, and Amiya Kumar Tripathy. Rice crop yield prediction in India using support vector machines. In *2016 International Joint Conference on Computer Science and Software Engineering*, pages 1–5. IEEE, 2016.

[701] Niketa Gandhi, Owaiz Petkar, and Leisa J. Armstrong. Rice crop yield prediction using artificial neural networks. In *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development*, pages 15–16. IEEE, 2016.

[702] Shizhong Xu, Yang Xu, Liang Gong, and Qifa Zhang. Metabolomic prediction of yield in hybrid rice. *The Plant Journal*, 88(2):219–227, August 2016.

[703] Niketa Gandhi and Leisa J. Armstrong. Rice crop yield forecasting of tropical wet and dry climatic zone of India using data mining techniques. In *2016 IEEE International Conference on Advances in Computer Applications*, page 24. IEEE, 2016.

[704] Jaratsri Rungrattanaubol and Anamai Na-udom. An efficient rice yield predictive model in lower northern Thailand. In *AIP Conference Proceedings*. AIP, 2016.

[705] C. C. Teoh, N. Mohd Nadzim, M. J. Mohd Shahmihaizan, I. Mohd Khairil Izani, K. Faizal, and H. B. Mohd Shukry. Rice yield estimation using below cloud remote sensing images acquired by unmanned airborne vehicle system. *International Journal on Advanced Science, Engineering and Information Technology*, 6(4):516–519, 2016.

[706] Jong Ahn Chun, Sung Kim, Woo-Seop Lee, Sang Myeong Oh, and Hyojin Lee. Assessment of multimodel ensemble seasonal hindcasts for satellite-based rice yield prediction. *Journal of Agricultural Meteorology*, 72(3–4):107–115, 2016.

[707] Dibakar Ghosh, Udai P. Singh, Krishnendu Ray, and Anupam Das. Weed management through herbicide application in direct-seeded rice and yield modeling by artificial neural network. *Spanish Journal of Agricultural Research*, 14(2):e1003, June 2016.

[708] Mostafa K. Mosleh, Quazi K. Hassan, and Ehsan H. Chowdhury. Development of a remote sensing-based rice yield forecasting model. *Spanish Journal of Agricultural Research*, 14(3):e0907, August 2016.

[709] Altaf Ali Siyal, Jan Dempewolf, and Inbal Becker-Reshef. Rice yield estimation using Landsat ETM+ data. *Journal of Applied Remote Sensing*, 9(1):095986, November 2015.

[710] Liz Allen, Alison O'Connell, and Veronique Kiermer. How can we ensure visibility and diversity in research contributions? How the contributor role taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*, 32(1):71–74, January 2019.

Acknowledgements

I am deeply grateful to my supervisor, Dr. Adam Mahdi, for his guidance, intellectual generosity, and steady encouragement. His clarity of thought, optimism, and rigorous feedback strengthened the direction, coherence, and analytical depth of this work. It has been a true privilege to learn under his mentorship.

I also thank the Oxford Internet Institute for its commitment to part-time doctoral study for professionals. The opportunity to pursue this research alongside practice was invaluable, and it fundamentally shaped the perspective and applied orientation of this thesis.

This thesis benefited from the expertise and feedback of collaborators across multiple disciplines. I am grateful to Jonas Jägermeyer and Marijn van der Velde for their guidance on agricultural monitoring using satellite and climate data. I thank Jim Hall, Marc van Homberg, and Marleen de Ruiter for their insights on disaster risk and climate resilience. I am also grateful to Sagar Surendra Deshmukh for sharing perspectives on agricultural systems in India, and to Burcu Balçık and Lily Xu for their feedback on the optimization components of this work.

The Dieter Schwarz Foundation generously supported parts of this research through the DSF–OII programme at the Oxford Internet Institute, for which I am sincerely grateful.

Finally, I thank my family for their unwavering encouragement. Their patience and belief in me were a constant source of strength throughout this journey.