

The application of self-validation to wireless sensor networks

Michael A Collett^{†§}, Maurice G Cox[†], Mihaela Duta[‡],
Trevor J Esward[†], Peter M Harris[†], and Manus P Henry[‡]

[†] National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK

[‡] Department of Mechanical Engineering, Oxford University

Abstract. Self-validation is a valuable tool for extending the operating range of sensing systems and making them more robust. Wireless sensor networks suffer many limitations meaning that their efficacy could be greatly improved by self-validation techniques. We present two independently developed data analysis techniques and demonstrate that they can be applied to a wireless sensor network. Using an acoustic ranging application we demonstrate an improvement of more than ten-fold in the uncertainty of a single measurement where multiple sensor readings are appropriately combined. We also demonstrate that of the two methods observed for determining the largest consistent subset one is more rigorous in dealing with correlation, and the other more suited to time-series data.

Keywords: wireless sensor networks, SEVA, self-validation, data fusion, data exclusion, largest consistent subset, uncertainty evaluation, information hierarchy

Submitted to: *Measurement Science and Technology*

1. Introduction

Wireless sensor networks are set to become a significant enabling technology in many areas of measurement, security and environmental monitoring. In the next decade industry will begin to benefit from autonomous sensor networks that can determine their position and adapt to their environment. This development raises a number of metrological issues relating to reliability, system confidence, calibration and validation. There is a need for

- (i) metrologically sound methods for quantifying and improving the robustness and reliability of sensor networks and related systems;
- (ii) investigation of the specific metrological and communication problems introduced by the wireless medium that affect confidence and efficacy of measurement, e.g., electromagnetic compatibility issues;

§ To whom correspondence should be addressed (michael.collett@npl.co.uk)

- (iii) development of numerical indications to represent the uncertainty in high-level outputs and robustness of WSN measurement systems to enable well-informed decisions to be taken for control and design purposes.

1.1. The enabling technology of the future

Wireless sensor networks combine sensing, computation and communication into a single device that may be very small. By means of advanced networking protocols and radio link redundancy, the connectivity properties of such devices and networks can make them more robust and efficient. Multi-hop routing can contribute to this flexibility as well as allow remote data aggregation, filtering and processing. The capabilities of any single device may be limited, but a network of many devices offers radical new technological possibilities. “The power of wireless sensor networks lies in the ability to deploy large numbers of tiny nodes that assemble and configure themselves” [1]. Sensor networks will be smart and able to change functions in response to changing conditions. Networks will be autonomous, self-configuring, self-calibrating and self-reorganising for optimum performance.

1.2. Advantages of distributed sensing

One of the most distinctive aspects of WSNs compared with traditional single sensors is distributed sensing. In a distributed sensing system, the measurement tasks are fulfilled through collaborative measurement by several sensors. Quite often, several sensors, either homogeneous or heterogeneous, detect a common phenomenon, and data from various sensors are aggregated to produce more accurate information. Distributed sensing offers several advantages:

- improved uncertainty: the weighted mean of the measurement is obtained through data fusion of multiple independently obtained measured values, and the uncertainty associated with the aggregated output is smaller than that associated with a single measured value;
- increased confidence: measurement confidence is increased as the measured value from one sensor is confirmed by those from other sensors;
- higher robustness and reliability: since an abundance of data leads to redundancy the system can still provide information even if part of the network fails.

1.3. Challenges introduced by distributed sensing

As well as the advantages listed above, there are a number of new challenges introduced by these novel intelligent distributed systems:

- restricted access: traditional instruments can be brought to a laboratory and calibrated in controlled conditions using a well-understood reference instrument or artefact. WSN nodes are commonly too numerous and too remote to make such an approach viable;

- limited access to raw data: owing to communication and data storage limitations, the user may only have access to processed or aggregated data;
- resource constraint: such systems are subject to limited processing, power supply, communications bandwidth and sensor quality. This can result in loss of stability of measurement, making some form of regular calibration or validation necessary;
- harshness of environment: many remote sensing applications will be deployed in harsh, uncontrollable, dynamic environments, which can lead to sensors being damaged, destroyed or forced to operate outside their intended ranges;
- abundance of data: as well as an advantage, it becomes ever more important to fuse data into meaningful information with quantifiable uncertainty.

1.4. The aim of the current work

This paper seeks to extend the concept of self-validation to networks of sensors and hence to overcome some of the challenges listed above. Self-validation has already been applied successfully to single instruments [2, 3, 4, 5]. The question arises as to whether the concept can be transferred to networks of instruments and to wireless sensor networks in particular. We show that for a demonstrator network that has been given the task of acoustic location of a target, self-validation can be built into the distributed measurement task. Moreover, combined with simple calibration methods, the network is capable of more accurate location measurement than would be possible with a single sensor, and can report when the output of specific sensors is unreliable. The work builds on methods for combining redundant measured values that have been reported in [6, 7, 8].

2. Fault detection in sensor networks

To improve measurement confidence and ensure reliability of WSNs, it is especially important to integrate suitable fault detection schemes into the network. There can be various types of fault, such as noise, sensor drift, complete loss of a sensor and saturation of a sensor. The causes of these faults can include low batteries, interference, ageing/fouling of the sensor or environmental conditions lying outside specified bounds. The faults can be temporary or permanent depending on their source.

One of the methodologies used in WSNs to tackle this problem is through dense deployment of sensors; if the sensor spacing is smaller than the spatial variation in the physical quantity being measured, then neighbouring sensor readings should be comparable^{||}. Although individual sensors may fail, the network can exploit such redundancy and collaborate to detect the fault. Thus, improving measurement confidence in WSNs largely depends on the ability to detect faults reliably and react accordingly by correcting, discarding or augmenting the uncertainty associated with

^{||} Similar techniques can be used in the time domain. Sampling much faster than the variation of the physical quantity allows application of filtering for outliers.

discrepant data points. A variety of classes of fault detection technique has been proposed for WSNs:

Data-based fault detection: These generic methods perform bulk statistical analysis on data sets to verify and reject outliers. They work if there is a large set of data constituting measured values of nominally the same quantity, and the technique is generally performed at a central aggregating base-station.

Model-based fault detection: A fault detection model is generated based upon a variety of factors such as physical theory, and machine learning techniques like genetic algorithms, neural networks and decision trees. The model is applied to identify any inconsistency in data from several (potentially heterogeneous) sensors, so as to detect and report faults. This technique carries out fault detection at the network level. It can be applied to a wider range of situations, but the model is application-specific, so it is necessary to build a model for each application, which can be challenging and time-consuming to do reliably. Models can also be applied to time-series data from a single sensor, for example if a measured temperature varies faster than is physically possible, then readings can be marked as discrepant and treated accordingly.

Smart “SEVA” sensors: In contrast to model-based fault detection, smart sensors and SEVA (SELF-VALIDating) sensors house the intelligence not solely at the network or base-station level but also close to the measurement sensor itself. The sensor provides extra functions beyond those necessary for generating a correct representation of the sensed quantity and act as more than simple transducers. A sensor may provide information about signal condition, context (such as other related physical quantities) and even measurement uncertainty. Local data filtering or smoothing can also be performed. Knowledge of sensor-specific fault modes is the basis for building a successful SEVA sensor. For example, a pressure sensor may give faulty data when its working temperature is outside a normal range. This simple fault can be detected by monitoring the temperature in the pressure sensor so as to gain information about its reliability. Temperature information may be used to help correct clock drift [9]. In networks such information may of course come from adjacent sensors rather than the specific sensor of interest. Several authors have described self-validating sensors of varying complexity from thermocouples to flow meters and their associated software [10]-[18].

3. The two data-combining techniques

There are many situations in which it is necessary to aggregate a large number of measured values of a single measurand. These may be repeated indications from a single sensor or multiple sensors. Two techniques have been developed independently at NPL and Oxford University for the validation of such data sets. The remainder of this paper describes these methods and demonstrates that they can be used effectively for an

acoustic localisation application of wireless sensor networks.

NPL mathematicians have developed a methodology for identifying the Largest Consistent Subset (LCS) of measured values for two cases:

- the values to be combined can be regarded as independently obtained, or
- there is some logical correlation between the measured values; e.g., they were obtained using the same sensor, or more than one sensor has a calibration that is traceable to a common source.

The uncorrelated and correlated cases are considered in references [7] and [6] respectively. The methodology described in section 3.1 for determining a largest consistent subset of measured values having associated correlation was used to study the performance both of the network as a whole and of individual sensors under resource-constrained circumstances. The Sensor Validation Research Group at Oxford University, UK, has considered the issue from the point of view of online sensing [19]. The method examines the consistency of each pair of measured values and determines the largest consistent subset as the largest collection of mutually consistent measurements: the ‘Maximum Clique’. The methodology is described in Section 3.2.

3.1. Determining the largest consistent subset: chi-squared method

Following the approach set out in reference [6], consider a set of data consisting of N measured values x_i , provided independently, together with the associated standard uncertainty $u(x_i)$, for $i = 1, \dots, N$. The weighted mean y of these measured values is formed, with weights inversely proportional to the squares of the $u(x_i)$. If this weighted mean is consistent with the measured values according to a chi-squared test, it can be accepted as an estimate of the measured quantity. Otherwise, a largest consistent subset (LCS) of the measured values is determined. This is a subset from the complete set that corresponds to as many measurement results as possible and that is consistent with the weighted mean of the subset. Define the ‘sum of squares function’ $F_J(\eta)$ by

$$F_J(\eta) = \sum_{j \in J} \left(\frac{x_j - \eta}{u(x_j)} \right)^2, \quad (1)$$

where J is a subset of the indices $\{1, \dots, N\}$. Then, the consistency test for the N measurement results involves comparing the *observed chi-squared value* $F_{\{1, \dots, N\}}(y)$ with a percentage point of the chi-squared distribution with $\nu = N - 1$ degrees of freedom [20].

The LCS may be formed using a fully enumerative approach [21]. If the N measured values taken together fail the consistency test, examine all possible subsets of $N - 1$ values (obtained by omitting one value at a time), and conduct the consistency test on each subset. For example, for the subset $J = \{1, \dots, N - 1\}$ obtained by omitting the N th value, the test involves comparing the observed chi-squared value $F_J(y_J)$, where y_J is the weighted mean of the values in the subset J , with a percentage point of the chi-squared

distribution with $\nu = N - 2$ degrees of freedom. If none of these subsets passes the test, consider all subsets of $N - 2$ measured values, and so on.

Let R denote the largest integer for which there is a subset of R measured values that passes the consistency test. If this subset is the only one that passes the test, it constitutes the required LCS. Otherwise, select that with the smallest observed chi-squared value, i.e., satisfying the consistency test by the largest margin.

The computational time for determining the LCS by full enumeration may be excessive in some instances, especially when N is large, with several values that lie outside the LCS. The paper [7] describes an efficient approach, based on the properties of the sum of squares function (1), for determining the LCS having smallest observed chi-squared value, and applies it to measurement results from several interlaboratory comparisons. For a given η , let $J = J_r(\eta)$ denote the subset of the measured values of size r for which the sum of squares function (1) has smallest value taken over all subsets of size r . The approach exploits the result that $J_{r+1}(\eta)$ contains $J_r(\eta)$.

When the measured values are not provided independently, the generalized weighted mean is used in place of the weighted mean, and the sum of squares function (1) takes the form

$$F_J(\eta) = \mathbf{e}_J^T(\eta) \mathbf{U}_J^{-1} \mathbf{e}_J(\eta), \quad (2)$$

where $\mathbf{e}_J(\eta)$ contains the deviations $x_j - \eta$, $j \in J$, and \mathbf{U}_J is the uncertainty (covariance) matrix associated with the deviations, comprising the variances associated with the measured values x_j , $j \in J$, in its diagonal elements and the covariances associated with pairs of the measured values elsewhere. For such measured values, the above result regarding properties of the subsets $J_r(\eta)$ no longer generally holds, and the appendix to reference [7] describes an example to illustrate this. Consequently, an algorithm not requiring the property that $J_{r+1}(\eta) \supset J_r(\eta)$, e.g., a fully enumerative approach, is required to determine the LCS.

3.2. Determining a consistent subset: maximum clique with uncertainty augmentation

Consider again the set of N measured values x_i and an associated standard uncertainty $u(x_i)$ for $i = 1..N$, which have been obtained independently and nominally relate to the same quantity. It is also assumed at this stage that these uncertainties have been generated by each of the SEVA sensors themselves, with limited local processing, so that both x_i and $u(x_i)$ are potentially suspect.

Each pair of measured values is checked for consistency, by calculating the Moffat distance $d_M(i, j)$ [22] defined here as

$$d_M(i, j) = \frac{|x_i - x_j|}{\sqrt{u(x_i)^2 + u(x_j)^2}}, \quad (3)$$

where i and j indicate any pair of indices from $1 \dots N$. For consistency at a stated level of confidence it is required that,

$$d_M(i, j) < k, \quad (4)$$

where k is a coverage factor corresponding to that level of confidence. For example, under a Gaussian assumption, $k = 1.96$ for a 95 % level of confidence. Note that this is equivalent to the chi-squared test applied to the pair of measured values. Once all the pair-wise consistency relations are evaluated, the largest consistent subset can be determined, either by exhaustive search, or through a linear approximation using overlapping intervals, as described in reference [5]. However, when dealing with a time-series of measurement results from a set of sensors, there is a high probability that one or more sensors will be found to be inconsistent at some time step, due to statistical variability, so there can be a constantly changing set of sensors within a consistent subset (CS). If the weighted mean y is based only on the CS, then as membership of the CS changes randomly over time, jumps could be introduced into the time series for y , which in many measurement and control contexts is undesirable.

Accordingly, additional steps are taken to ensure the y values vary smoothly with time. The most important of these is the notion of uncertainty augmentation when a measured value, say x_{out} , falls outside the CS. This occurs because at least one of its Moffat distances in relation to the members of the CS exceeds the selected threshold. It is simple to demonstrate that if the standard uncertainty $u(x_{out})$ associated with x_{out} , is multiplied (augmented) by a function of its largest Moffat distance to any of the CS measurement results, then it becomes consistent with all of them and can thus be included in the subset itself. Of course a consequence of augmenting the uncertainty in this way is that the corresponding weight in the calculation of y is reduced. Where random effects cause relatively small variations in a Moffat distance, which occasionally just exceeds k , then this approach yields only slight variations in the weighting of the associated measurement, so there is relatively little variation in y (compared with excluding the measurement entirely). However, if the measured values from one sensor are gradually drifting away from those of the majority of sensors, the corresponding weighting is successively reduced. In the case of more than one consistent subset of the same size, the intersection of the two sets is used as the consistent subset, and uncertainty augmentation is applied to the residual members of both sets to make them consistent. Again this strategy avoids jumps between different consistent subsets in the time series for y .

4. The test application: wireless sensor networks, position measurement and ranging

A common application is the use of a network of sensors to locate a target, for the purposes of tracking the target or identifying the positions of the sensors. In such cases, each sensor typically provides data that, when processed, enables the position of the target to be estimated, and a weighted mean estimate of the position is calculated using a data fusion algorithm. We show how the two data fusion algorithms described in section 3 can be applied to an acoustic ranging task in a manner that makes individual sensors and also the network of sensors self-validating, leading to improved accuracy of

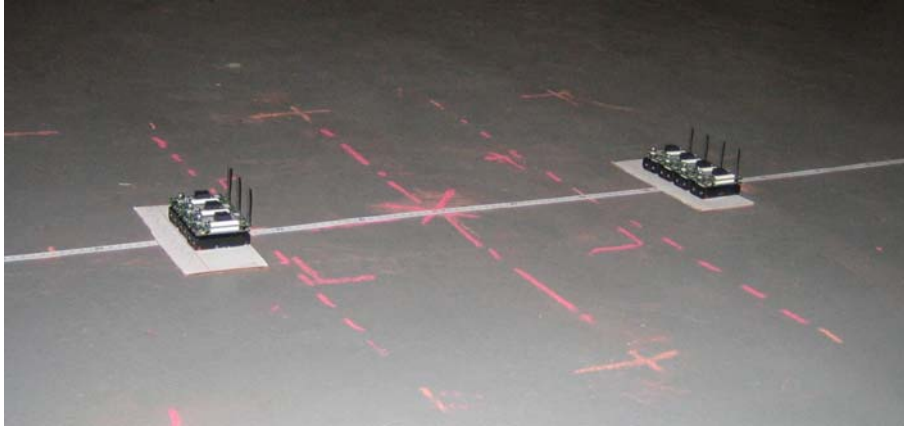


Figure 1. Layout for ranging calibrations in the hemi-anechoic chamber

measurement.

The majority of localisation methods rely on the multi-lateration of multiple ranging estimates made between a number of sensors [23]. It has been shown that position estimation is highly unstable with respect to errors in ranging [24, 25]. So, for such lateration-based location methods, accurate ranging techniques are required. The chosen method of ranging is one proposed by the Institute for Software Integrated Systems at Vanderbilt University [26], which uses the time of flight of an acoustic signal to give an estimate of the distance between a sounder and a listener. A ‘sounder’ node transmits a radio message and an acoustic pulse simultaneously. A ‘listener’ node estimates the time delay of arrival of the acoustic pulse with respect to the radio message and uses it to estimate the distance to the sounding node.

The sensor platform employed was the MICAz mote developed at the University of California at Berkeley to act as a research tool for low-power wireless sensor networks [27]. Each mote contains a number of on-board sensors as well as a sounding device. The specification for the system is given in reference [28].

The acoustic ranging application utilizes a broad range of functionality of the system and encompasses issues such as synchronisation, radio communications, high sample rate sensing and actuation, on-board processing and data analysis. The main reason for choosing acoustic ranging as the application of interest was that it can be used to expose a large number of the general issues involved in measurement using WSN systems, and will hence be subject to many of the inherent problems and issues described in section 1.3. Some further investigations into these issues using the NPL testbed are described in reference [29].

In the context of acoustic ranging measurement made with networks, there are a number of failure modes which can manifest themselves in erroneous data and where self-validation would prove beneficial. These could be at the point of raw data gathering if, for example, there were too much background noise to determine the presence of the acoustic pulse. Another fundamental point of potential failure is the onboard algorithm [26] for judging the start of the acoustic pulse based on a stored buffer of

microphone data. The task of judging the start of a pulse when there is a noisy signal is very challenging, even in non resource-constrained situations. One characteristic of the implementation of this algorithm is to give strong discretization in the output ranging estimates, which must be accounted for during any subsequent analysis.

4.1. Calibrating the sensor ranging

Preliminary measurements were carried out by comparing the ranging estimates of pairs of nodes to a reference ruler (as shown in figure 1). The measured values are shown in figure 2. It can be seen that all the ranging estimates have an appreciable offset from the actual separation (shown by the broken line). Some pairings produce intermittently high estimates relative to the rest of the nodes. Based upon these measured values, it was found that one of the sounders gave the best linearity in estimates from the other node listeners. Hence, the sensor network to be used for the measurement stage comprised seven devices, with one (called device 1) acting as a sounder and the other six (called devices 2–7) acting as listeners. To measure accurately the offset between actual separation and reported range detailed calibrations were carried out. Example data from a calibration run between nodes 1 and 2 is shown in figure 3, which shows the range indications returned by node 2 at separations of 55, 60...100 cm, along with an applied straight-line fit. Similar calibration processes and resulting linear fits were used to correct for any offsets from ranging pairs, as well as evaluate the uncertainty associated with their outputs.

4.2. Measurement stage

Consideration is given to how correlation associated with pairs of estimates provided by a single sensor arises through a calibration of the sensor, and to the use of a method to determine a largest consistent subset to test the consistency of measurement data provided by a single sensor and by multiple sensors.

It is assumed that irrespective of (a) the distance between the sounder and a listener, and (b) the choice of listener, the indications provided by the listener have the same associated uncertainty. It is further assumed that pairs of indications, provided by the same listener or by different listeners, have zero associated correlation. The problem and its mathematical treatment are described at length elsewhere [6] and will not be repeated here. This paper concentrates on presenting the key results that are relevant to the question of self-validation in sensor networks.

In the measurement stage, the sounder and listeners are located in several different configurations. For each configuration, the listeners are located uniformly around a circle with the sounder at the centre of the circle (figures 4 and 5). The configurations differ in terms of the ordering of the listeners around the circle, i.e., the mapping from locations A–F on the circle to listeners 2–7, but they are such that the circles all have the same nominal radius of 67 cm (measured from the centre of the sounder to the centre of each listener as in the calibration stage).

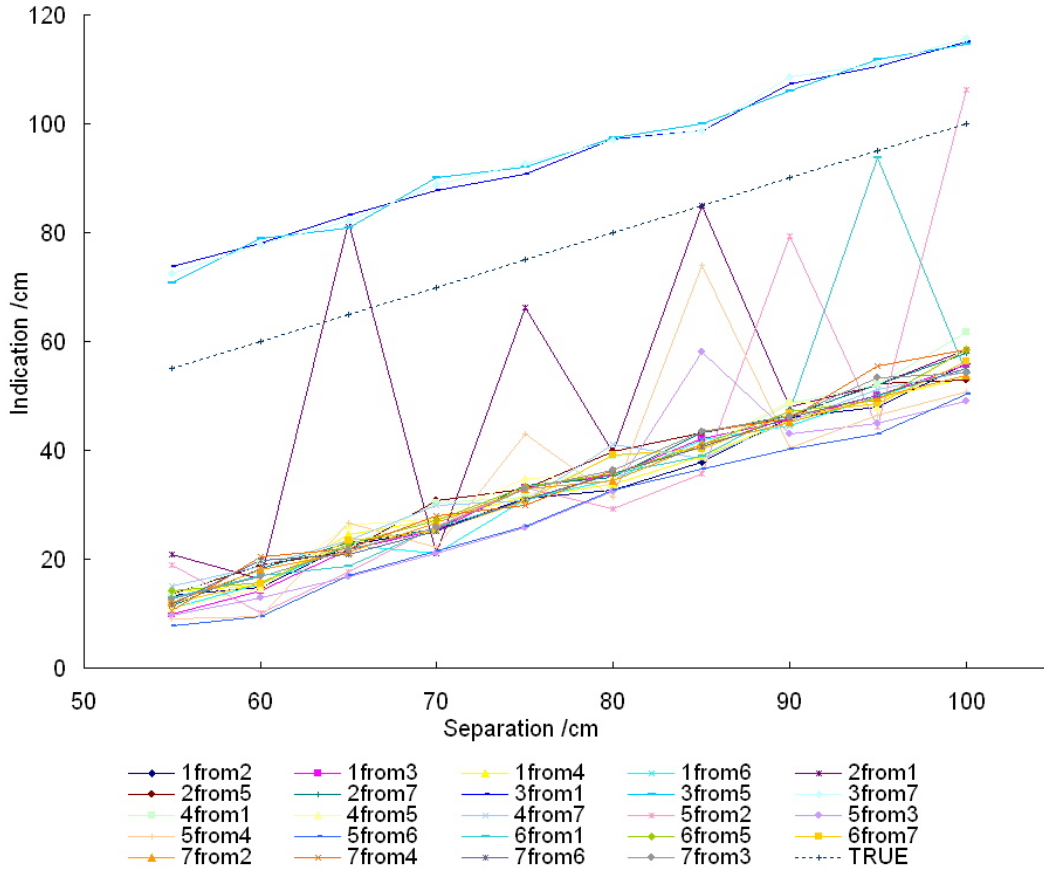


Figure 2. A set of ranging estimates at separations of 55 cm to 100 cm. Legend uses the convention “A from B” where A is the sounder and B the listener.

For the i th configuration ($i = 1, \dots, 10$) and the j th listener ($j = 2, \dots, 7$), repeated indications y_{ijk} ($k = 1, \dots, m_{ij}$) are recorded. Each stage of this example application results in a number of data sets which can be aggregated and checked for consistency: repeat indications of a listener in a given configuration, all indications from an individual listener in all configurations and all indications from all listeners in all configurations. This is one example of the information hierarchy [30] which characterizes the path of data and levels of aggregation available in WSN systems.

4.3. Measurement results

Below are results from the ranging application described in section 4.2, along with the determination of the LCS and implementation of the methods described in sections 3.1 and 3.2. These data sets were chosen to emphasize differences in the two validation techniques, as well as demonstrate data validation at many levels of the information hierarchy.

Figure 6 shows the aggregation of repeat indications from a single sensor (listener 4) in a single configuration (position 5), with validation carried out using the chi-squared method (6(a)) and the maximum clique method (6(b)). This data set was found by

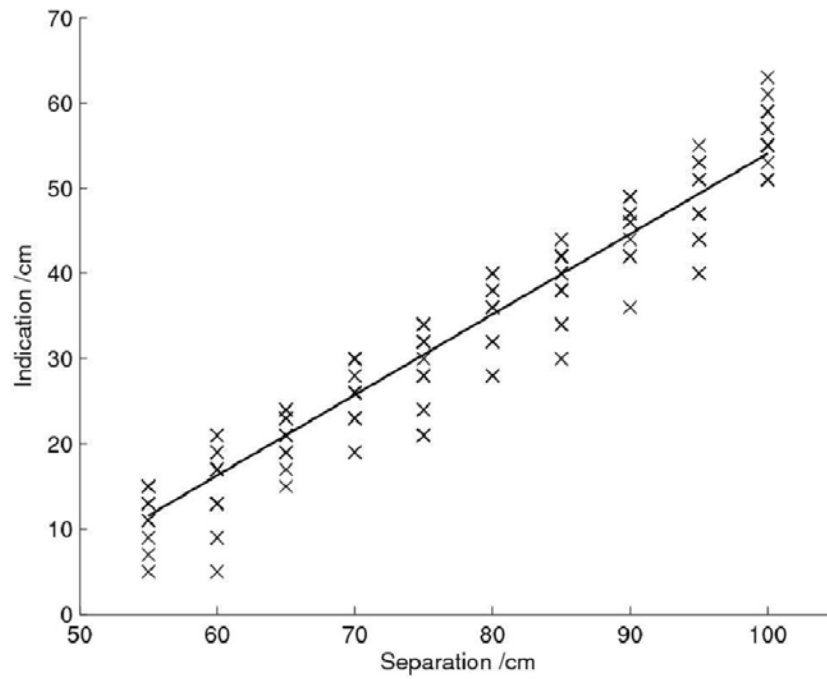


Figure 3. All the calibration data for node 2 listening to node 1, together with the applied straight-line fit (s.d. not shown). Note this includes some numerically identical repeat indications.

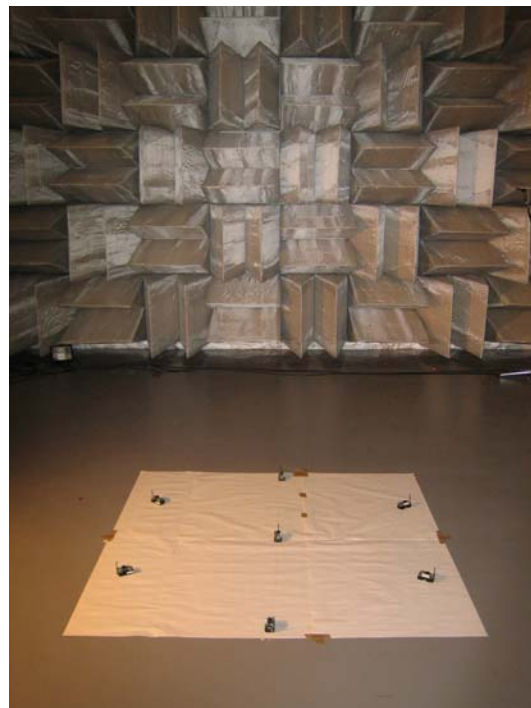


Figure 4. Layout for gathering largest consistent subset data in the hemi-anechoic chamber

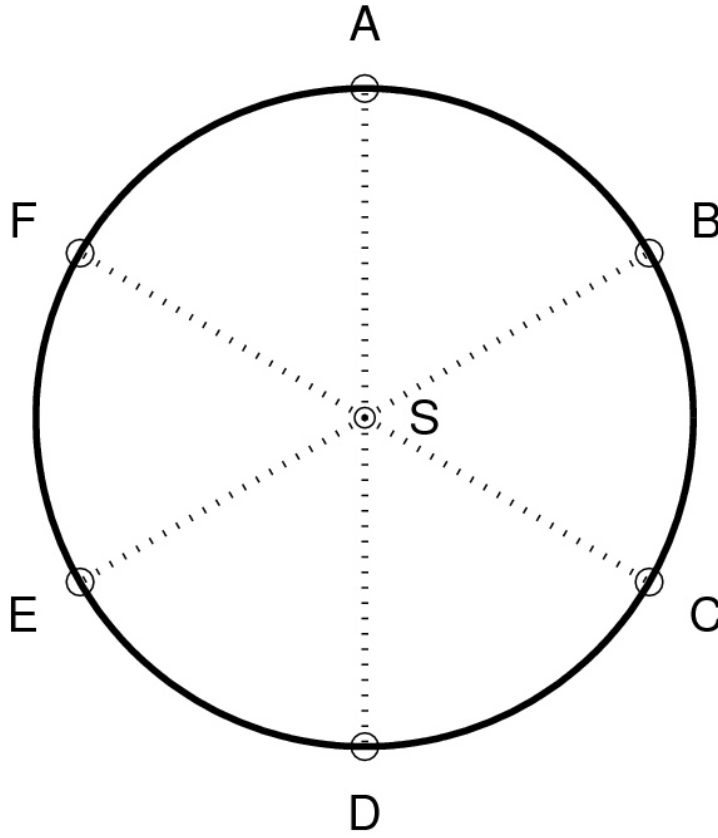
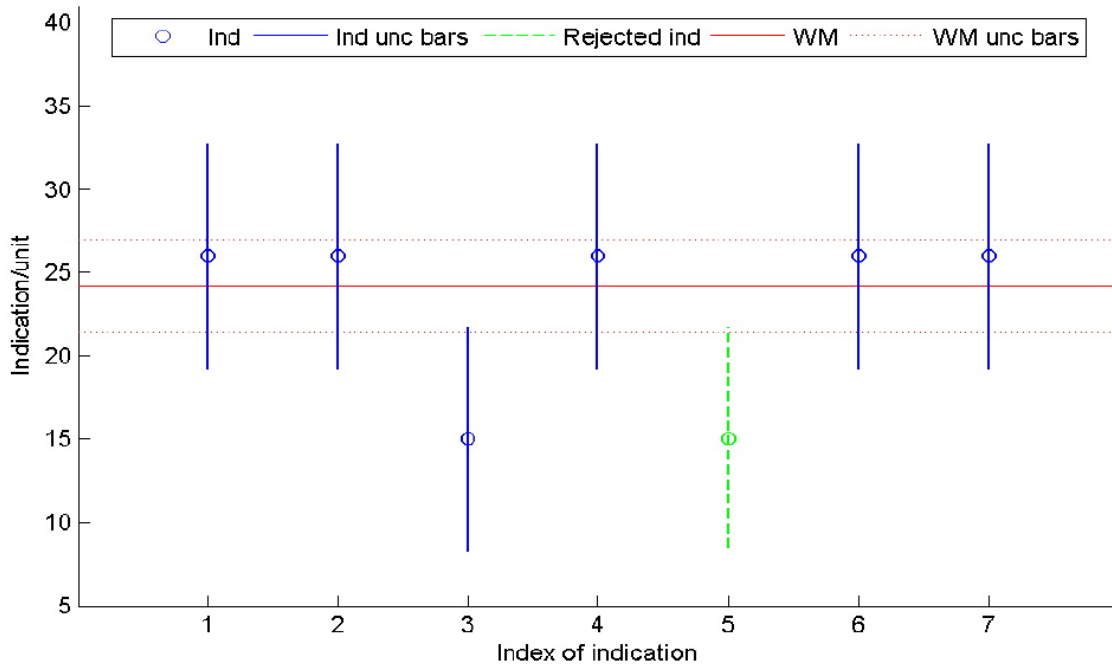


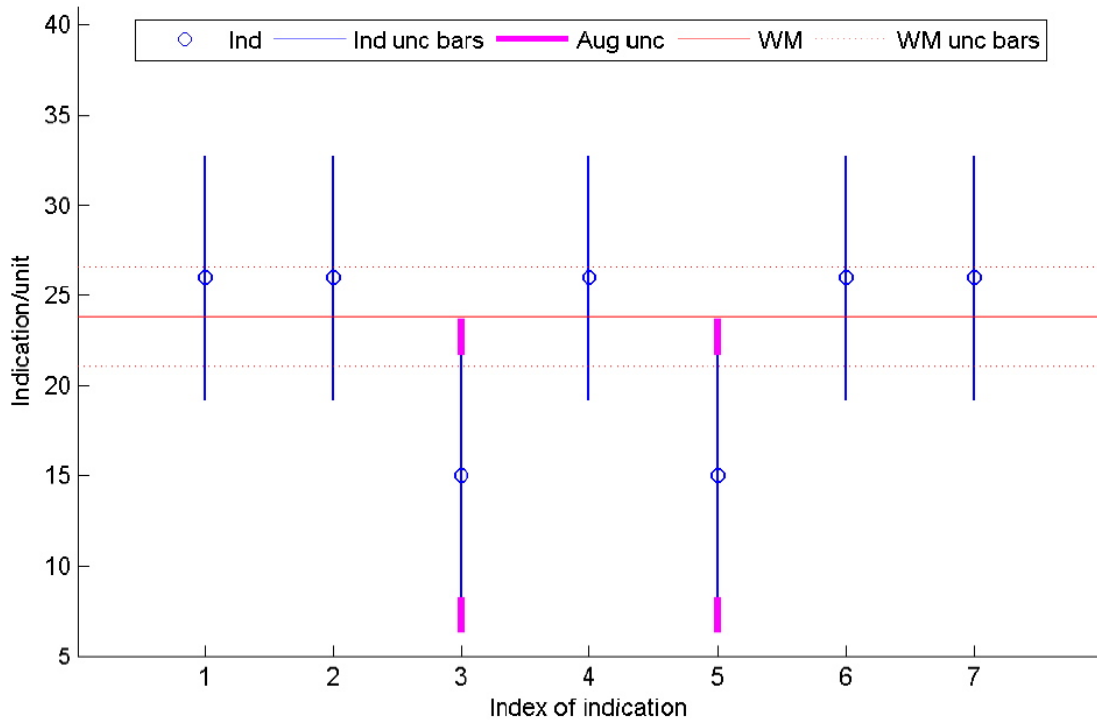
Figure 5. Configuration of the sounder and listeners in the measurement stage. The sounder (S) is at the centre of a circle of nominal radius 67 cm. The listeners are located uniformly around the circle (A–F).

both validation techniques to have inconsistencies: indications 3 and 5 are statistical outliers, failing the chi-squared and maximum clique tests. This arises from the strong discretisation in outputs from the ranging algorithm (section 4, [26]). The two analysis methods differ in their handling of these outliers. As seen in 6(a), the chi-squared method would exclude one of these data points, but has no way to choose between them, meaning the LCS cannot be uniquely defined. However, 6(b) shows that the method of uncertainty augmentation would allow both points to remain, albeit with reduced weighting.

Figure 7 shows the aggregation of repeat indications from listener 7 in configuration 7, with validation carried out using the two methods. This data set was found by both validation techniques to have inconsistencies: indication 3 is a statistical outlier, failing both the chi-squared and maximum clique tests. The two analysis methods differ in their handling of this outlier: either rejecting it, or reducing its weight in the final aggregated result. This causes the two methods to produce differing numerical outputs: the chi-squared method gives 24.14 with an associated standard uncertainty of 1.31 cm and the maximum clique method gives 23.58 with an associated standard uncertainty of 1.28 cm.

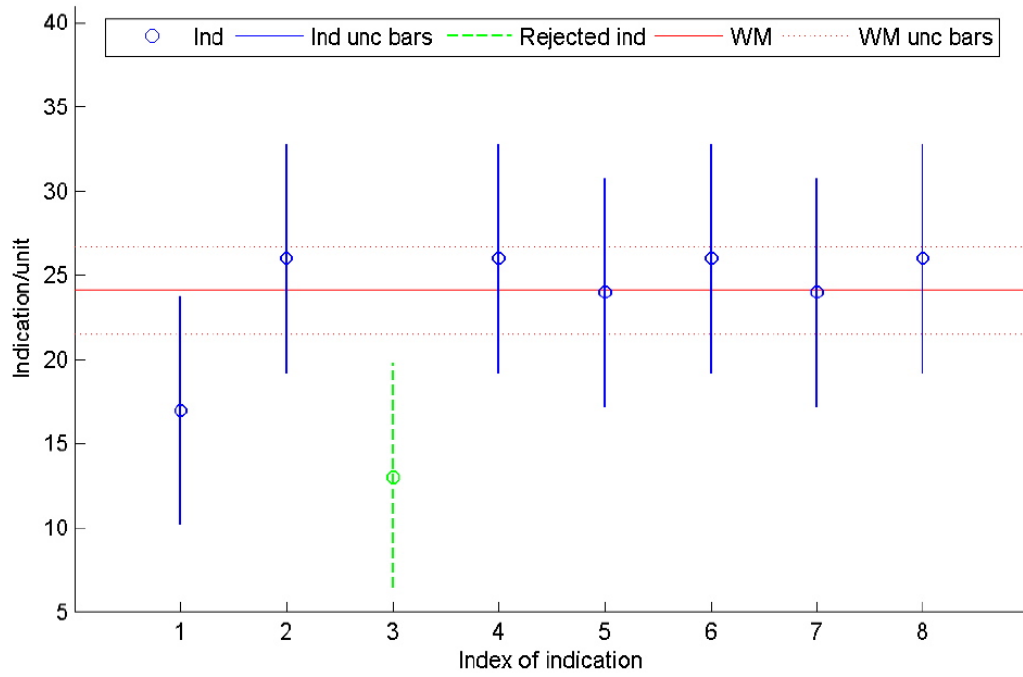


(a) LCS based on chi-squared test. Weighted mean: 24.17 ± 1.42 cm

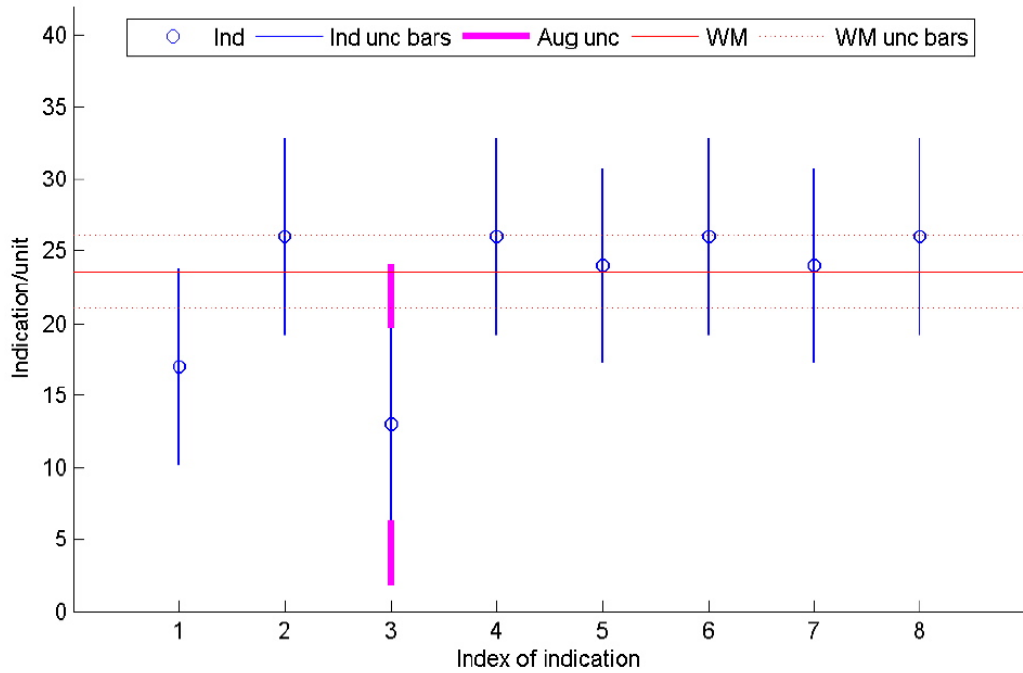


(b) LCS based on maximum clique method with uncertainty augmentation. Weighted mean: 23.85 with an associated standard uncertainty of 1.39 cm

Figure 6. Weighted mean (WM) with the associated 95 % uncertainty band for repeated indications of listener 4 in configuration 5. Results are shown for both the chi-squared and maximum clique methods.



(a) LCS based on chi-squared test. Weighted mean: 24.14 ± 1.31 cm



(b) LCS based on maximum clique method with uncertainty augmentation. Weighted mean: 23.58 with an associated standard uncertainty of 1.28 cm

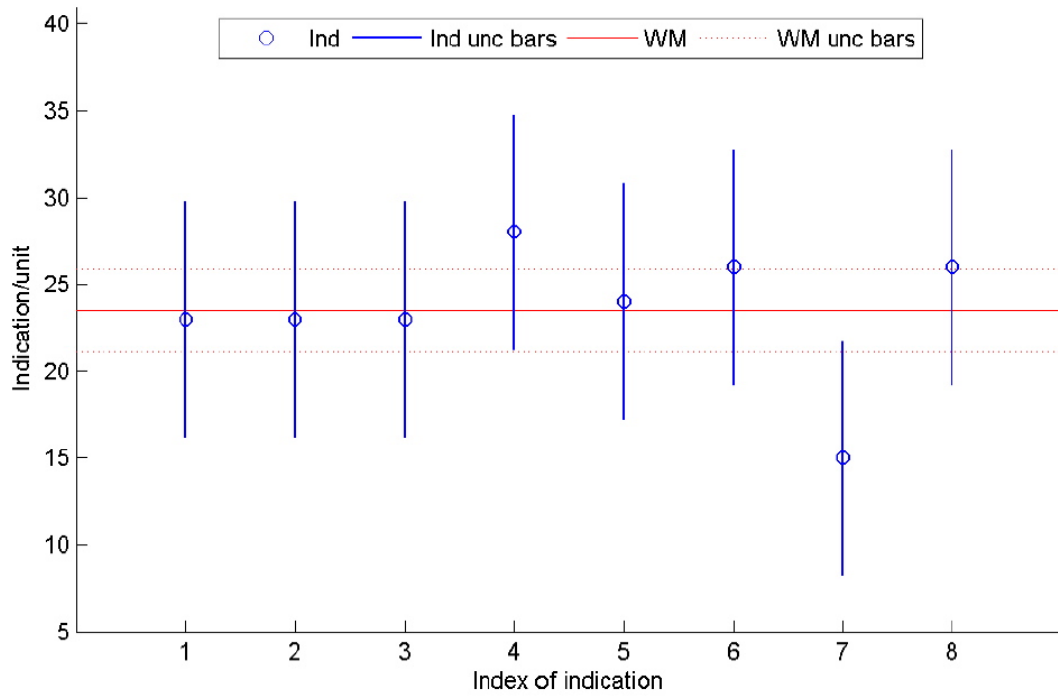
Figure 7. Weighted mean (WM) with the associated 95 % uncertainty band for repeated indications of listener 7 in configuration 7. Results shown for both chi-squared and maximum clique methods.

Figure 8 shows the aggregation of repeat indications from listener 6 in configuration 8, again with validation carried out using the chi-squared method 8(a) and the maximum clique method 8(b). This data set is notable because it passes the chi-squared test but fails the maximum clique test. Augmentation of the uncertainty of indication 7 brings it into the maximum clique. There are another 10 configuration/listener combinations which fail the maximum clique test while passing the chi-squared test. This indicates that the NPL LCS method allows for slightly larger variations before performing any exclusion or correction of data values.

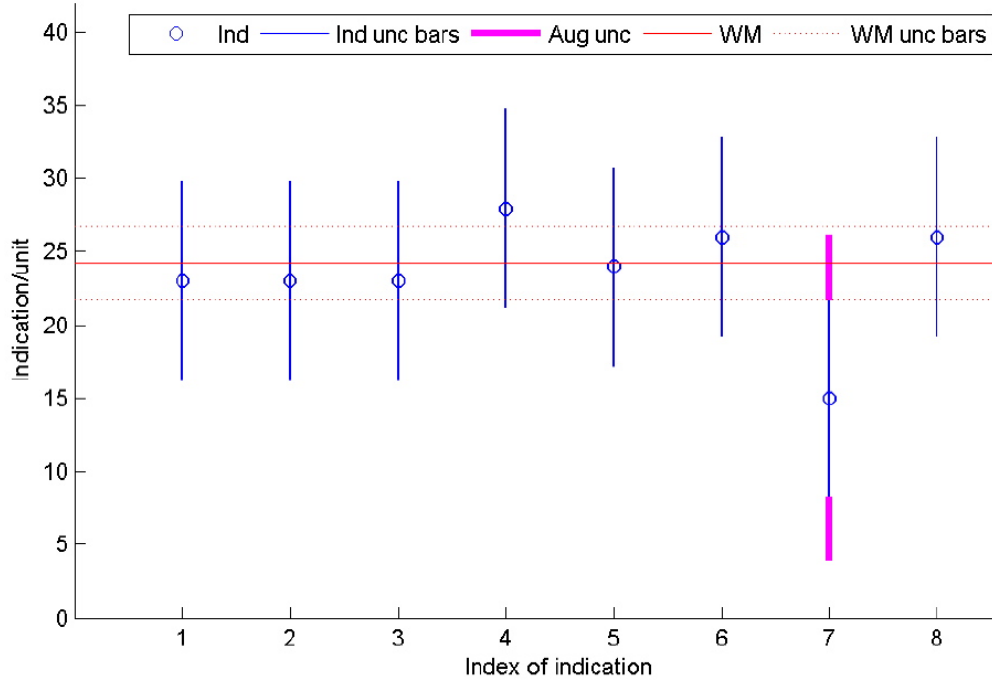
Figure 9 shows the aggregation of the combined indications of listener 2 over all the configurations around the circle. This is the next level up the information hierarchy from that shown in previous figures as it includes some extra stages of aggregation and processing. Firstly, the value at each configuration results from the aggregation of repeat raw readings, which have themselves been combined using the two aggregation methods and secondly, the calibration factor calculated in figure 3 and its associated uncertainty have been factored into the readings, introducing correlation between all readings from each sensor. Despite these differences it is still valid to apply the aggregation methods to these “higher-level” data sets. This data set was found to pass both tests for consistency. This is likely to be due to the fact that the outlier rejection and data validation had already been carried out on the lower-level data, meaning that any inconsistencies had already been removed. This may indicate that for this application it is not necessary to carry out validation at all levels. The determination of when and where validation would be beneficial would depend on the application. Factors such as information propagation errors or combination of multiple sensors may necessitate validation above the raw data level. For comparison, the nominal value of 67 cm for the separation between the sounder and the listener is shown as a broken horizontal line.

Figure 10 shows the estimates \hat{x}_j , $j = 2, \dots, 7$, with the associated 95 % coverage intervals, for the distance from the sounder to the listeners for all six listeners, combined over all ten configurations. The weighted mean \hat{x} is shown as a solid horizontal line. The nominal value is again indicated by a dashed horizontal line. The estimate of the radius of the circle using the chi-square test¶ is $\hat{x} = 67.13$ cm with associated standard uncertainty $u(\hat{x}) = 0.22$ cm, while using the maximum clique technique gives $\hat{x} = 67.23$ cm with associated standard uncertainty $u(\hat{x}) = 0.23$ cm. This result also involves uncertainty augmentation on the results from listeners 3 and 5. This indicates that, despite the previous processing to ensure the data from each sensor are self-consistent, there may be inconsistencies between the sensors themselves. These are relatively small, however, and are not rejected by the chi-squared method. This indicates that there is benefit in applying these techniques to higher-level data, as well as the raw data level. These two methods provide a comparable final indication of the measured quantity. Despite the close agreement in this result, there remain important differences between the two methods, which will determine their applicability

¶ This is the result as implemented by Oxford. The NPL implementation, as quoted in [6] is $\hat{x} = 67.07$ cm with associated standard uncertainty $u(\hat{x}) = 0.23$ cm

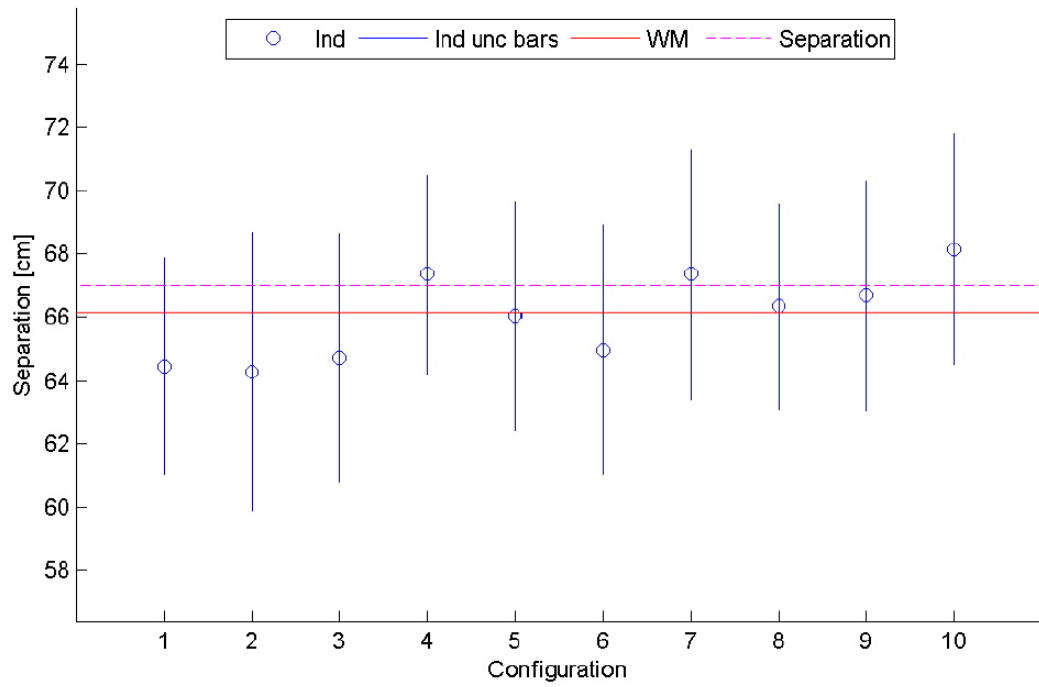


(a) LCS based on chi-squared test. Weighted mean: 23.50 with an associated standard uncertainty of 1.23

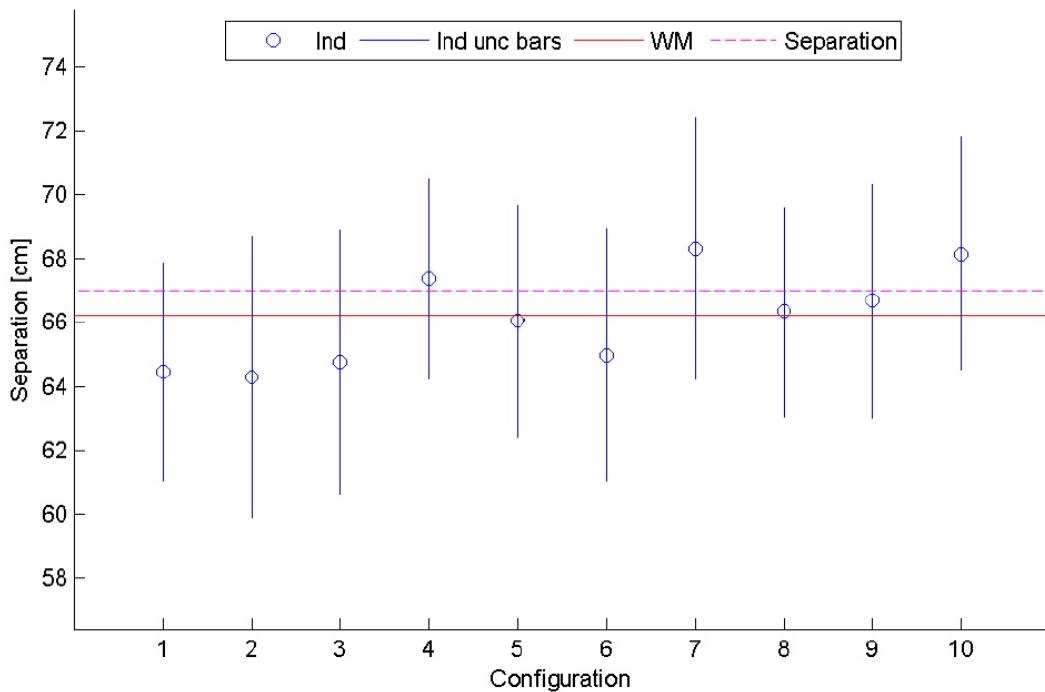


(b) LCS based on maximum clique method with uncertainty augmentation. Weighted mean: 24.22 with an associated standard uncertainty of 1.28

Figure 8. Weighted mean (WM) with the associated 95 % uncertainty band for repeated indications of listener 6 in configuration 8.



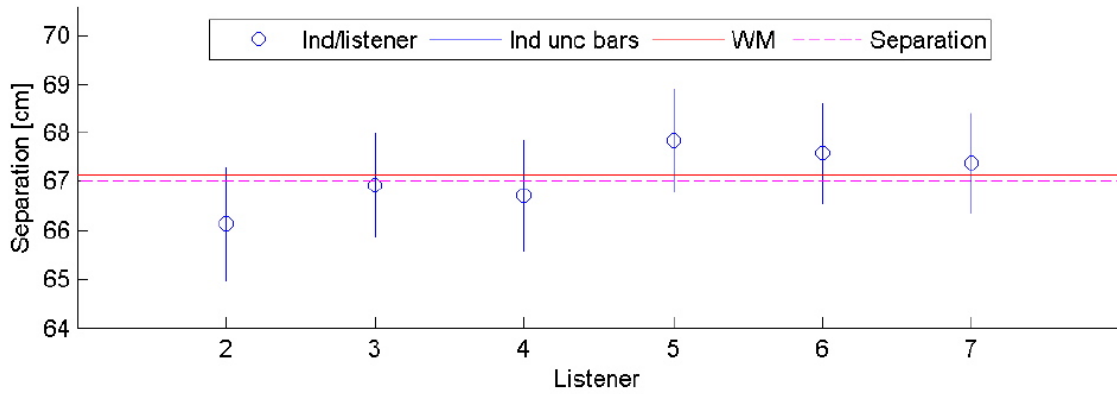
(a) LCS based on chi-squared test. Weighted mean: 66.14 with an associated standard uncertainty of 0.59



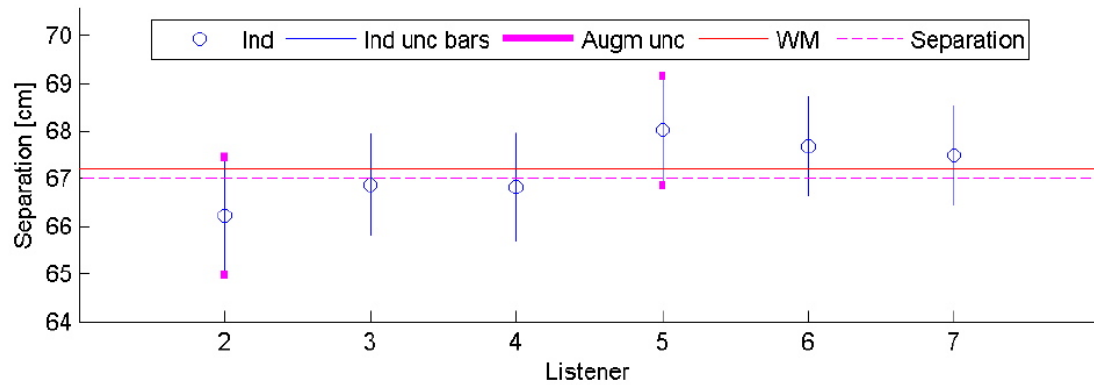
(b) LCS based on maximum clique method with uncertainty augmentation. Weighted mean: 66.22 with an associated standard uncertainty of 0.59

Figure 9. Weighted mean (WM) with the associated 95 % uncertainty band for each configuration of listener 2. The calibration factor measured in figure 3 has also been included at this stage. Results shown for both chi-squared and maximum clique methods.

and validity to wireless sensor networks and their applications.



(a) LCS based on chi-squared test. Weighted mean: 67.13 with an associated standard uncertainty of 0.22



(b) LCS based on maximum clique method with uncertainty augmentation. Weighted mean: 67.23 ± 0.23

Figure 10. Weighted sum of the sounder-listener separation for the six listeners. Results shown for both chi-squared and maximum clique methods.

5. Conclusion

The work reported here shows that while self-validation techniques can be applied to wireless sensor networks, and can be implemented at many levels of the information hierarchy [30], careful consideration is required to determine the stages at which validation techniques will provide the most benefit. All implementations of self validation require a methodology for identifying and handling discrepant measured values. In the case of the NPL method the largest consistent subset methodology developed by Cox [7] identifies and excludes such values. The Oxford method, however, avoids the exclusion of individual discrepant values by augmenting the uncertainty associated with the outlying data point. In addition, the Oxford method assumes that all measured values are independently obtained (uncorrelated) whereas the NPL approach

enables correlation to be handled robustly. The practical advantage of the Oxford approach, which was designed for applications that involve time-series measurements, is that one is not continually re-classifying indications as either valid or values that must be excluded, providing stability and smoothness in the output whilst potentially sacrificing some metrological consistency. The choice of procedure, such as the two demonstrated in this paper, may in the end be governed as much by engineering and practical concerns as by metrological ones.

Acknowledgment

The National Measurement System Directorate of the UK's Department of Innovation, Universities and Skills supported this work as part of its Measurements for Emerging Technologies programme. All work by the Oxford research group described in this paper is funded by Invensys. Vanderbilt University developed the majority of the software components which were used in the acoustic ranging example application.

References

- [1] Hill J L System architecture for wireless sensor networks *Ph D thesis University of California, Berkeley, U.S.A.* (2003)
- [2] Henry M P A SEVA Sensor - The Coriolis Mass Flow Meter *Fault detection, supervision and safety for technical processes* Pergamon (1994) 429–434
- [3] Henry M P Self-validating digital Coriolis mass flow meter *Computing and Control Engineering Journal* **11**(5) (2000) 219–227
- [4] Henry M P, Clarke D W, Archer N, Bowles J, Leahy M J, Liu M J, Vignos R P and Zhou J A self-validating digital mass-flow meter: an overview *Control Engineering Practice* **2000** **8**(5) (2000) 487–506
- [5] Henry M P On-line compensation in a digital Coriolis Mass Flow meter *Flow Measurement and Instrumentation* **12**(2) (2001) 147–161
- [6] Collett M A, Cox M G, Esward T J, Harris P M and Sousa J A Aggregating measurement data influenced by common effects *Metrologia* **44** (2007) 308–318
- [7] Cox M G 2007 The evaluation of key comparison data: determining the largest consistent subset *Metrologia* **44** (2007), 187–200
- [8] Duta M and Henry M. The fusion of redundant SEVA measurements. *IEEE Transactions on Control Systems Technology* **13** (2) (2005) 173–184
- [9] Ganeriwal S, Ganesan D, Shim H, Tsiatsis V and Srivastava M Estimating Clock Uncertainty for Efficient Duty-Cycling in Sensor Networks *Sensys 2005, Proceedings of the 3rd ACM Conference on Embedded Networked Sensor Systems, November 2005*
- [10] Boltryk P J, Harris C J and White N M Intelligent sensors - a generic software approach *Journal of Physics: Conference Series* Vol 15 (2005) 155–160
- [11] Frolik J and Kandasamy P A confidence-based approach to the self-validation, fusion and reconstruction of quasi-redundant sensor data *IEEE Transactions on Instrumentation and Measurement* Vol 50, no. 6 (2001) 1761–1769
- [12] Gaura E and Newman R Intelligent diagnostic for sensor arrays *IEEE Mediterranean Conference 2003* T7140 Viewed online on 17th October 2006 <http://med.ee.nd.edu/MED11/pdf/papers/t7-140.pdf>
- [13] Leahy M J, Henry M P and Clarke D W Sensor validation in biomedical applications

- Modelling and Control in Biomedical Systems: 3rd symposium on sensor validation in biomedical application* (1997) 221–226
- [14] Henry M P, Archer N, Atia M R A, Bowles J, Clarke D W, Fraher P M A, Page I, Randall G and Yang J C-Y Y Programmable hardware architectures for sensor validation *Control Engineering Practice* Vol 46, no 10 (1996) 1339–1354
 - [15] Mahmoud W H, Abdelrahman M, and Haggard R L Field programmable gate arrays and implementation of automated sensor self-validation system for cupola furnaces *Computers and Industrial Engineering* Vol 46 (2004) 553–569
 - [16] Menz B Vortex flowmeter with enhanced accuracy and reliability by means of sensor fusion and self-validation *Measurement* Vol 22 (1998) 123–128
 - [17] Ray A and Luck R An introduction to sensor signal validation in redundant measurement systems *IEEE Control Systems Magazine* February 1991 44–49
 - [18] Reed R P Possibilities and limitations of self-validation of thermoelectric thermometry *American Institute of Physics Conference Proceedings, Temperature: Its Measurement and Control in Science and Industry* Vol 684 part 1 (2003) 507–512
 - [19] Duta M, D, Henry M, P The Fusion of Redundant SEVA Measurements *IEEE Transactions on Control Systems Technology, Vol. 13 (2005), pp173-184*
 - [20] Cox M G The evaluation of key comparison data *Metrologia* **39** pp 589-95 2002
 - [21] Iyer H K, Wang C M and Vecchia D F Consistency tests for key comparison data *Metrologia* **41** pp 223-30 2004
 - [22] Moffat R J Contributions to the theory of single sample uncertainty analysis *ASME Journal of Fluid Engineering* Vol 104 (1982) 250–260
 - [23] Allen M, Gaura E, Newman R and Mount S Experimental Localization with MICA2 Motes *Technical Proceedings of the 2006 NSTI Nanotechnology Conference and Trade Show, Volume 3 Vol. 3, p 435-440*
 - [24] Forbes A Multistation Coordinate Metrology *Multilateration Presentation to SIDAnet November 2002* NPL public report
 - [25] Moore D, Leonard J, Rus D and Seller S Robust Distributed Network Localization with Noisy Ranging measurements *SenSys '04, November 3-5, 2004, Baltimore, Maryland, USA*
 - [26] Sallai J, Balogh G, Maroti M and Ledeczi A Acoustic Ranging in Resource Constrained Sensor Networks *Technical Report, ISIS, Vanderbilt University 2003*
 - [27] Hill J L and Culler D Mica: A Wireless Platform for Deeply Embedded Networks. *IEEE Micro*, Vol 22(6), (2002) 12-24.
 - [28] <http://tinyurl.com/239saa>
 - [29] Collett M A, Esward T J, Adamson D B Measurements using the wireless sensor network demonstrator testbed NPL public report
 - [30] Chen Y P, Liestman A L and Liu J Energy-Efficient Data Aggregation Hierarchy for Wireless Sensor Networks *Quality of Service in Heterogeneous Wired/Wireless Networks, 2005*