

Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations

Sarah Hable,¹ Karin Sigloch,² Guilhem Barruol,³ Simon C. Stähler⁴ and Céline Hadziioannou⁵

¹Department of Earth and Environmental Sciences, Ludwig-Maximilians-University, 80333 Munich, Germany. E-mail: shable@geophysik.uni-muenchen.de

²Department of Earth Sciences, University of Oxford, Oxford OX1 3AN, UK

³Laboratoire GéoSciences Réunion, Université de La Réunion, Institut de Physique du Globe de Paris, Sorbonne Paris Cité, UMR CNRS 7154, Université Paris Diderot, 97744 Saint-Denis, France

⁴Department of Earth Sciences, Eidgenössische Technische Hochschule, 8092 Zurich, Switzerland

⁵Institute of Geophysics, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, 20146 Hamburg, Germany

Accepted 2018 June 14. Received 2018 January 9; in original form 2018 June 11

SUMMARY

Many applications in seismology rely on the accurate absolute timing of seismograms. However, both seismological land stations and ocean bottom seismometers (OBSs) can be affected by clock errors, which cause the absolute timing of seismograms to deviate from a highly accurate reference time signal, usually provided by GPS satellites. Timing problems can occur in land stations when synchronization with a GPS signal is temporarily or permanently lost. This can give rise to complicated, time-dependent clock drifts relative to GPS time, due to varying environmental conditions. Seismometers at the ocean bottom cannot receive GPS satellite signals, but operate in more stable ambient conditions than land stations. The standard protocol is to synchronize an OBS with a GPS signal immediately before deployment and after recovery. The measured timing deviation, called ‘skew’, is assumed to have accumulated linearly over the deployment interval, an assumption that is plausible but usually not verifiable. In recent years, cross-correlations of ambient microseismic noise have been put to use for correcting timing errors, but have been limited to interstation distances of at most a few tens of kilometres without reducing the temporal resolution. We apply noise cross-correlations to the evaluation of clock errors in four broad-band land stations and 53 wideband and broad-band OBSs, which were installed on and around the island of La Réunion in the western Indian Ocean during the RHUM-RUM (Réunion Hotspot and Upper Mantle-Réunions Unterer Mantel) experiment. We correlate all three seismic components, plus a hydrophone channel in OBS stations. Daily cross-correlation functions are derived for intermediate distances (~ 20 km) for land-to-land station pairs; stable, 10 d stacks are obtained for very large interstation distances up to >300 km for land-to-OBS and OBS-to-OBS configurations. Averaging over multiple station pairs, and up to 16 component pairs per station, improves the accuracy of the method by a factor of four compared to the single-channel approaches of prior studies. The timing accuracy of our method is estimated to be ~ 20 ms standard deviation or one sample at a sampling rate of 50 Hz. In land stations, nonlinear clock drifts and clock jumps of up to 6 min are detected and successfully corrected. For 52 out of 53 OBSs, we successfully obtain drift functions over time, which validate the common assumption of linear clock drift. Skew values that were available for 29 of these OBSs are consistent with our independent estimates within their observational error bars. For 23 OBSs that lacked skew measurements, linear OBS clock drifts range between 0.2 and 8.8 ms d⁻¹. In addition to linear drift, three OBSs are affected by clock jumps of ~ 1 s, probably indicating a missing sample problem that would otherwise

have gone undetected. Thus we demonstrate the routine feasibility of high-accuracy clock corrections in land and OBSs over a wide range of interstation distances.

Key words: Time-series analysis; Seismic instruments; Seismic interferometry; Seismic noise.

1 INTRODUCTION

A crucial requirement of many seismological processing methods is the correct absolute timing of seismograms. For the internal clocks of seismological land stations, accurate timing can usually be ensured by frequent synchronization with GPS satellites acting as a highly accurate, external reference clock. This synchronization can fail when reception of the GPS signal is lost, in which case the internal clock usually starts to drift notably relative to reference time (e.g. Sens-Schönfelder 2008). The estimation and correction of such clock errors is a long-standing problem in seismology.

An analogous timing problem occurs in ocean bottom seismometers (OBSs), because GPS satellite signals do not reach the seafloor. Only two GPS synchronizations can be attempted: immediately before deployment and after recovery. If both GPS connections are successful, a timing deviation between the internal clock of the data logger that records the seismometer and hydrophone components and GPS clock is obtained, the so-called skew value. A first-order correction of internal clock timing can then be attempted by assuming that the skew accumulated gradually and linearly over the deployment interval (e.g. Gouédard *et al.* 2014; Hannemann *et al.* 2014). This is plausible because the rate of quartz oscillators in station clocks is expected to depend mainly on ambient temperature, which tends to remain very stable in deep water, although ultimately these assumptions remain unverified. If one or both GPS synchronizations failed, then not even this simple linear clock correction is possible.

Recent years have seen the rise of an independent observational method for estimating and correcting station clock errors, using cross-correlation functions (CCFs) of ambient seismic noise. It is based on the principle that noise CCFs feature a causal and an acausal part, which should occur at time lags of equal magnitude but opposite sign. Violation of this expectation can indicate the presence of clock errors, an ‘unphysical’ biasing process in the sense that it is unrelated to seismic wave propagation (Stehly *et al.* 2007). Since Lobkis & Weaver (2001) developed the basic principle of noise cross-correlations in ultrasound experiments, CCFs have found wide application in seismology, including structural studies of the crust (e.g. Shapiro *et al.* 2005; Sabra *et al.* 2005b), global tomography (e.g. Haned *et al.* 2016) and the monitoring of seismic velocity changes along faults (e.g. Brenguier *et al.* 2008b; Wegler *et al.* 2009) and on volcanoes (e.g. Sens-Schönfelder & Wegler 2006; Brenguier *et al.* 2008a; Sens-Schönfelder *et al.* 2014). A smaller number of studies investigated the use of noise cross-correlations for the correction of seismometer clocks. Stehly *et al.* (2007) and Sens-Schönfelder (2008) used CCFs to detect clock errors in land stations, while Sabra *et al.* (2005a), Gouédard *et al.* (2014), Hannemann *et al.* (2014) and Le *et al.* (2018) measured clock drifts of OBSs. In most cases, CCFs were calculated between stations with relatively small interstation distances, ranging from several metres (Sabra *et al.* 2005a), to several kilometres (Sens-Schönfelder 2008; Gouédard *et al.* 2014) and up to tens of kilometres (Hannemann *et al.* 2014; Le *et al.* 2018). Obtaining CCFs for larger interstation distances usually came at expense of temporal resolution (Stehly *et al.* 2007: distance 200 km, monthly CCFs).

Such long time averaging intervals may obscure the true time dependence and magnitude of clock drifts (Xie *et al.* 2018). Prior work used only a single-instrument channel: either the vertical seismic component of land stations (Sens-Schönfelder 2008) and OBSs (Hannemann *et al.* 2014) or the hydrophone component of OBSs (Sabra *et al.* 2005a; Gouédard *et al.* 2014; Hannemann *et al.* 2014; Le *et al.* 2018). We are aware of only one study (Stehly *et al.* 2007) that uses several component pairs in order to discriminate between Rayleigh and Love waves.

In this study, we demonstrate the great potential of the noise cross-correlation technique for even larger interstation distances (>300 km) and finer time resolution (CCFs from daily or 10 d intervals). We also demonstrate a several-fold increase in accuracy of clock error estimates when correlating all three seismogram components in the case of land stations, plus a hydrophone channel (H) in the case of OBSs. We successfully retrieve CCFs between OBSs and land stations; prior work on this is very limited (e.g. Corela *et al.* 2017; Tian & Ritzwoller 2017).

Section 2 describes our data, recorded by seismometers that were operating for several months to 3 yr on and around the island of La Réunion in the western Indian Ocean, as part of the RHUM-RUM experiment (Réunion Hotspot and Upper Mantle-Réunions Unterer Mantel, 2011–2015). Section 3 describes the various processing steps for obtaining multichannel CCFs and estimating the clock errors they imply. Section 4 presents results, subdivided by the three types of station pairs encountered: land-to-land clock errors; land-to-OBS; and OBS-to-OBS. The island stations on La Réunion were spaced by tens of kilometres, versus 200 km or more between two OBSs. Section 5 is a discussion, followed by conclusions in Section 6.

We used the open-source toolboxes Python (Rossum 1995), ObsPy (Beyreuther *et al.* 2010; Megies *et al.* 2011; Krischer *et al.* 2015) and obspyDMT (Hosseini & Sigloch 2017) for data downloading, processing and plotting all figures except Figs 1 and 11, which were generated with the open-source mapping toolbox GMT (Wessel *et al.* 2013).

2 DATA DESCRIPTION

The RHUM-RUM experiment was deployed between 2011 and 2015 with the primary objective of imaging crust and mantle beneath the volcanic hotspot of La Réunion. The island, 70 km long and 50 km wide, is located in the western Indian Ocean, 800 km east of Madagascar and 200 km southwest of Mauritius (Fig. 1a). Its hotspot volcanism, time-progressive hotspot track and associated large igneous province (Deccan Traps) have long marked it as a strong candidate for hosting a deep mantle plume underneath (Barruol & Sigloch 2013). In order to investigate this hypothesis, 48 wideband OBSs from the German DEPAS pool and 9 broad-band OBSs from the French INSU pool were deployed around the island for roughly 13 months (October 2012–November 2013), at water depths of 2200–5400 m. An extensive performance report of the OBSs is given by Stähler *et al.* (2016). In addition, 10 broad-band land stations were installed on La Réunion, operating for roughly 3

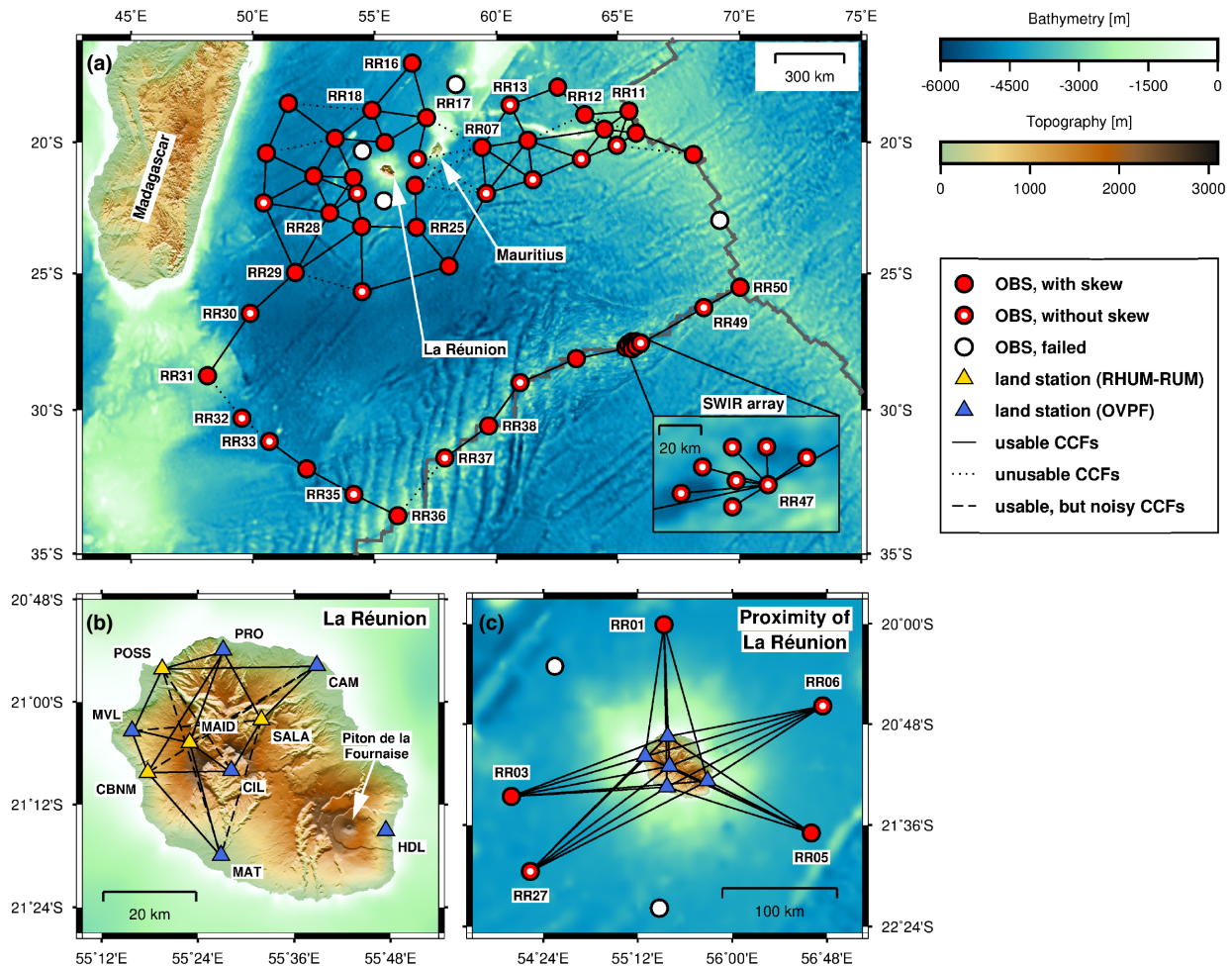


Figure 1. Maps of seismological broad-band stations used in this study, installed on and around La Réunion island. Circles represent ocean bottom seismometers (OBSs) deployed by the RHUM-RUM experiment: coloured red if a skew measurement was available; red with a white dot if no skew had been obtained; white if the OBS failed completely. Yellow triangles denote RHUM-RUM land stations; blue triangles denote permanent land stations operated by the OVPF volcano observatory. Stations explicitly mentioned in the text are labelled by their station names. Black lines connect pairs of neighbouring stations for which cross-correlations were attempted: OBS-to-OBS correlations in panel (a); land-to-land correlations in panel (b); and land-to-OBS correlations in panel (c). Solid black lines denote usable correlations; dotted lines denote unusable correlations and dashed lines usable but noisy correlations. Grey lines in (a) indicate plate boundaries. Figure generated with the GMT toolbox (Wessel *et al.* 2013).

yr (mid-2012–mid-2015). Additional RHUM-RUM island stations on Mauritius, Rodrigues, the Seychelles, Madagascar and the Îles Éparses were not part of our investigation. In total, the RHUM-RUM array covered $2000 \times 2000 \text{ km}^2$ of the Indian Ocean and ocean islands.

This study arose from the initial discovery of timing errors of up to several minutes in 4 out of 10 stations on La Réunion (yellow triangles in Fig. 1b). All four stations featured the same type of data logger (RefTek RT 130), in which a software failure was causing loss of GPS synchronization. Upon detection of these timing problems, the GPS receiver units of the four malfunctioning stations were repaired while the data loggers and seismic sensors were kept running in the field, with their internal clocks adrift relative to GPS reference time. After re-installation, the four repaired GPS units provided proper clock synchronizations until the entire array was dismantled approximately 3 months later.

The OBS deployment also encountered an unexpectedly high rate of GPS synchronization failures. Successful synchronization both before and after deployment, and thus a skew value, could be obtained for only 29 of the 57 OBSs (red circles in Fig. 1a). For these

stations, the timing of the seismograms was corrected based on the assumption of linear clock drift (Stähler *et al.* 2016). For another 24 OBSs, GPS synchronization upon recovery failed because the OBS clocks had shut down prematurely (red circles with white dot in Fig. 1a). The likely reasons were sudden, sharp voltage drops of the (poor quality) lithium batteries of the OBS clocks (Stähler *et al.* 2016). High-quality seismograms but unknown clock errors for these 24 OBSs prompted the extension of our terrestrial clock study to the oceanic realm. The remaining four OBSs had failed completely (white circles in Fig. 1a).

Our clock error study is divided into three parts: first, we evaluate the clock errors of the four land stations (CBNM, MAID, POSS and SALA) whose GPS units temporarily failed, setting their internal clocks adrift. We calculate daily CCFs between these erroneous stations and five reference land stations (CAM, CIL, MAT, MVL and PRO) whose clocks were properly synchronized to GPS throughout the RHUM-RUM deployment. These five permanent stations are operated by the volcano observatory of La Réunion, the Observatoire Volcanologique du Piton de la Fournaise (OVPF). Inter-station distances on the island were 19 km on an average (Fig. 1b).

Second, we calculate CCFs between the five OBSs located closest to La Réunion (RR01, RR03, RR05, RR06 and RR27) and five GPS-synchronized land stations operated by OVPF (CIL, HDL, MAT, MVL and PRO), in order to estimate the OBS clock drifts and to compare these estimates to measured skew values, which are available for RR01, RR03 and RR05 (Fig. 1c). Despite significant larger interstation distances of ~ 140 km, we retrieve reliable results with a temporal resolution of 10 d for the CCFs. This prompted us to take the third step of expanding the clock error study to the entire OBS network. We compute CCFs between neighbouring OBSs with interstation distances of up to 370 km (Fig. 1a). At this scale, verified skew-corrected OBSs serve as reference stations for OBSs that lack skew measurements.

3 METHODOLOGY

3.1 Pre-processing

Before the CCF calculation can be performed, continuous broadband waveform data are pre-processed in order to remove contaminating signals that could overprint ambient noise, such as earthquakes. Our procedure largely follows Bensen *et al.* (2007). The pre-processing steps are applied to day-long time-series for each station and component (land stations: E, N and Z; OBSs: 1, 2, Z and H). Time-series with missing samples must either be rejected or filled with some *ad hoc* number. Filling of large gaps could however distort the resulting CCF. To overcome this problem and to reduce the effect of the numerous small gaps in our data, we divide the 24 h long time-series into 1 hr windows with an overlap of 50 per cent, which yields 47 hr long windows per day (records of 00:00 a.m.–00:30 a.m. and 11:30 p.m.–00:00 a.m. are used only once). Gaps of less than 500 samples (corresponding to 5 and 10 s for sampling rates of 100 and 50 Hz, respectively) are filled with interpolated values, whereas windows featuring longer gaps are rejected. The amount of rejected windows is less than 1 per cent for the land stations. The majority of the OBSs contain no gaps, except for three stations, where we exclude 13 per cent (RR39), 24 per cent (RR53) and 33 per cent (RR33) of the hour-long windows. Even then a sufficient number of windows is obtained each day to calculate daily CCFs (see Section 3.2). The division into hour-long windows is redundant if the number of gaps is negligible. Next, the instrument response is removed from each hour-long window, a standard processing step (e.g. Bensen *et al.* 2007), even though strictly speaking it may not be necessary since we only rely on the CCF stability over time, not on its physical correctness. The instrument correction is followed by the removal of each window's mean value and trend. Then a zero-phase bandpass filter from 0.01–10 Hz is applied. This wide frequency band gives us more flexibility in testing various narrower frequency ranges without repeating the entire, computationally expensive pre-processing before settling on a specific frequency band (see below: 0.05–0.5 Hz) for further processing.

A subsequent optional downsampling is performed as follows: the RHUM-RUM OBSs sampled at 50, 62.5 or 100 Hz, depending on instrument type (Stähler *et al.* 2016). For consistency and usability, all OBS data are downsampled to 50 Hz; for all other land stations, their original sampling rate of 100 Hz is retained. Although more severe downsampling would render the CCF computations less time consuming, it is avoided in order to estimate clock errors to the highest possible temporal resolution.

To reduce the effect of highly energetic signals such as earthquakes, each hour-long seismogram is amplitude clipped at twice its standard deviation of that hour-long time window. These clipped time-series are then spectrally whitened between 0.05 and 0.5 Hz (2–20 s period). This period range is chosen because it contains the primary (~ 14 s) and secondary microseisms (~ 7 s), which are present in the La Réunion region throughout the year (Davy *et al.* 2015). Finally, 1-bit normalization is applied to the seismograms with the same purpose as the amplitude clipping. Larose *et al.* (2004) demonstrated in acoustic laboratory experiments that this technique improves the signal-to-noise ratio (SNR). It is therefore, a widely used method in ambient noise pre-processing (e.g. Shapiro & Campillo 2004; Shapiro *et al.* 2005; Hobbiger *et al.* 2012).

3.2 CCF calculation

CCFs are generally calculated using

$$\text{CCF}_{k_1 k_2}(s_1, s_2, t) = \int_{\tau_1}^{\tau_2} d_{k_1}(s_1, \tau) \cdot d_{k_2}(s_2, \tau - t) d\tau, \quad (1)$$

where $d_{k_1}(s_1, \tau)$ represents the seismogram of station s_1 and component k_1 , while $d_{k_2}(s_2, \tau - t)$ denotes the time-reversed seismogram of station s_2 and component k_2 . The times τ_1 and τ_2 indicate the start and end times of the CCF, respectively, and its lapse time is given by t . CCFs consist of a causal and an acausal part, which should be symmetric for homogeneously distributed noise sources. The causal part represents the response (the so-called Green's function) of station s_2 to a delta pulse at location of station s_1 , while the acausal part corresponds to the impulse response of s_1 with a source at position s_2 .

In nature, noise sources are often distributed unevenly, which leads to asymmetric CCFs (Stehly *et al.* 2007). Temporal CCF changes can be caused by several mechanisms, which are summarized by Stehly *et al.* (2007). Velocity changes in the subsurface would affect a CCF's causal and acausal parts in analogous ways, by dilating or compressing the CCF's waveforms on either side of $t = 0$, corresponding to slower or faster wave propagation, respectively. Such changes in wave propagation are expected in the vicinity of faults after large earthquakes (e.g. Brenguier *et al.* 2008b) or close to volcanoes, reflecting magma movements (e.g. Brenguier *et al.* 2008a; Sens-Schönfelder *et al.* 2014). However, there are no large earthquake faults on La Réunion and the island's active volcano, the Piton de la Fournaise, is too far away to have an effect on our CCFs (see Fig. 1b).

Temporal changes in the locations of noise source should be observed as shape changes of a CCF's causal and acausal parts independently of each other. This effect can be neglected because our tropical and subtropical stations record microseismic noise sources (2–20 s) that show little spatial seasonality (Davy *et al.* 2015). This is confirmed by the marked stability over time in our CCFs (Fig. 2).

In contrast to the above, instrument clock drifts manifest themselves by time-shifting the entire CCF, so that its causal part moves closer to $t = 0$ and its acausal part moves further away, or vice versa. We evaluate such shifts as described in Section 3.3. CCFs would be affected in the same way by the occurrence of a phase change in the instrument response.

We compute CCFs of our pre-processed, hour-long traces according to eq. (1). Daily CCFs are obtained by averaging all hourly CCFs in a day (up to 47). For the three OBSs (RR33, RR39 and RR53) affected by numerous gaps, we can use the sufficient number of ~ 30 – 40 windows per day for the calculation. The start time τ_1 and end time τ_2 of the hourly CCFs depend on interstation distance

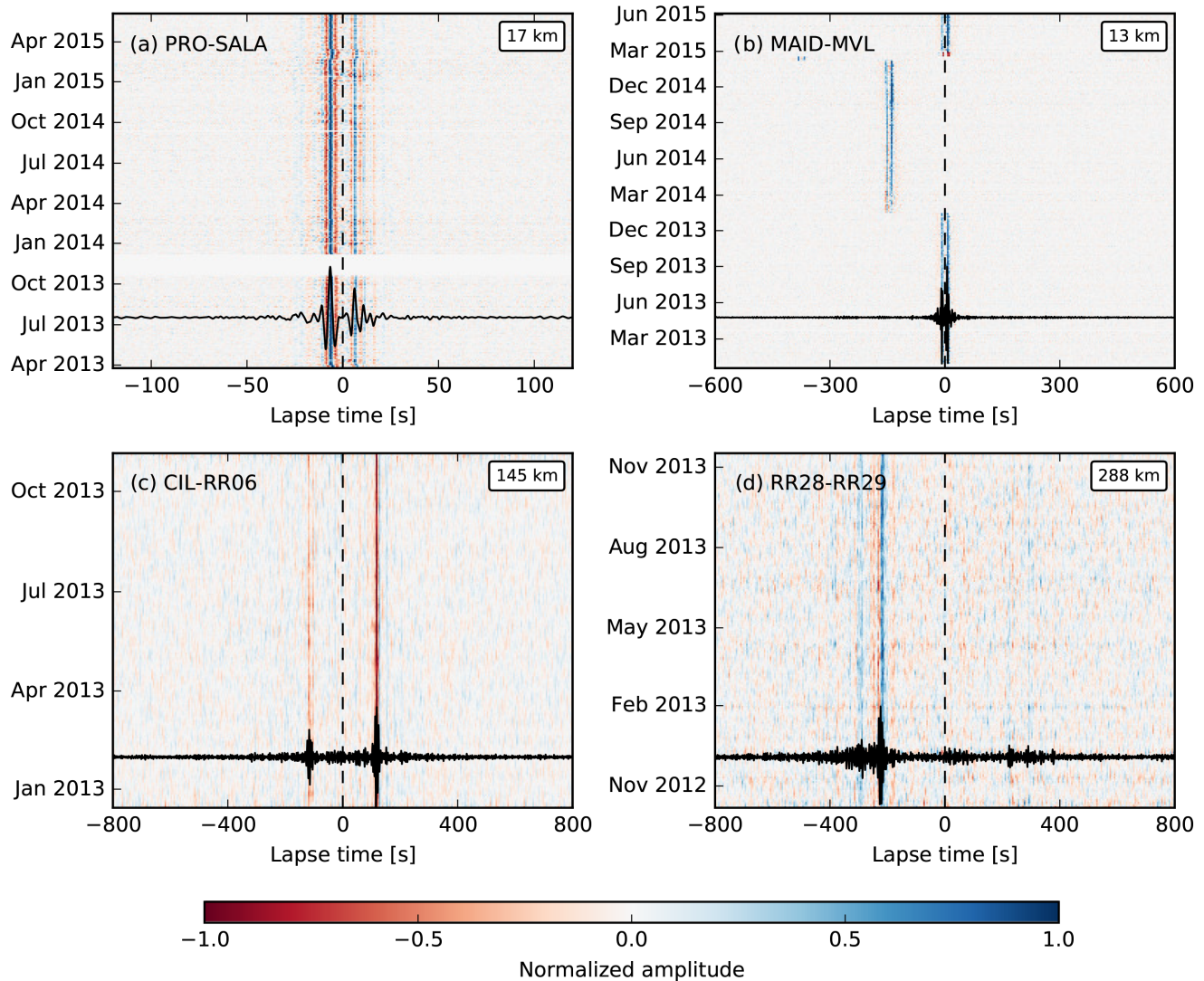


Figure 2. Examples of cross-correlation functions (CCFs) for the different types of station configurations and interstation distances. Colour shading represents the normalized amplitudes of the time-series. (a) Station pair PRO-SALA (land-to-land correlation, distance 17 km, ZN-component, daily stacks). (b) Station pair MAID-MVL (land-to-land, distance 13 km, ZE-component, daily stacks). (c) Station pair CIL-RR06 (land-to-OBS, 145 km, EZ-component, 10 d stacks). (d) Station pair RR28-RR29 (OBS-to-OBS, 288 km, HH-component, 10 d stacks). In each panel, the black time-series represents the RCF, that is, the stack over all individual CCFs. The dashed vertical line indicates zero lapse time; interstation distance is given in the top right of each panel.

and on the expected clock error: we use lapse times from -120 to $+120$ s for correlations of two land stations (Fig. 2a). An exception was station MAID, which showed severe clock jumps of several minutes, necessitating start and end times of ± 600 s (Fig. 2b). For the larger interstation distances of land-to-OBS and OBS-to-OBS correlations, lapse times run from -800 to $+800$ s (Figs 2c and d).

We characterize the quality of the daily CCFs by defining their SNR as the ratio of the maximum absolute value of a signal window to the standard deviation of a noise window:

$$\text{SNR} = \frac{\max(|\text{CCF}_{\text{signal}}(t)|)}{\text{std}(\text{CCF}_{\text{noise}}(t))}. \quad (2)$$

The time windows are based on interstation distance. The signal window is chosen in the early lapse times (land-to-land correlations: -25 to $+25$ s, land-to-OBS: -150 to $+150$ s, OBS-to-OBS: -400 to $+400$ s) expected to contain high energetic wave trains; the noise window is defined in the late lapse times (land-to-land: ± 80 to ± 120 s, land-to-OBS and OBS-to-OBS: ± 650 to ± 800 s), where

noise is dominant. For land-to-land station pairs, we reject daily CCFs if their SNR is below 7. For land-to-OBS and OBS-to-OBS correlations, this SNR threshold is chosen as 1, since we expect a generally lower SNR for these large-distance correlations. The mean SNR (SNRs averaged over the data period) ranges from 3.5–20, but we chose this much lower rejection threshold of 1 in order to reject only obviously useless windows, not those that are noisy but may still contain usable information. To increase the SNR of land-to-OBS and OBS-to-OBS correlations, we stack up to 10 daily CCFs (of $\text{SNR} \geq 1$) from a 10 d time window into 10 d stacks, where the 10 d stacking window is moved in increments of 1 d over the entire data period.

Next we calculate a reference correlation function (RCF) by stacking all daily CCFs that passed the SNR thresholding. For the RCF calculation for station MAID we exceptionally consider only the time period prior to the first extreme clock jump in January 2014 (about 1 yr of data). RCFs are displayed as black curves in Fig. 2.

Comparing the different station pairs of Fig. 2 reveals that clear wave arrivals in the CCFs can be retrieved continuously over the entire operation period. The SNR of daily CCFs decreases with increasing interstation distance (compare PRO-SALA and RR28-RR29). The examples of CIL-RR06 (land-to-OBS) and RR28-RR29 (OBS-to-OBS) demonstrate that stable, 10 d stacked CCFs can be obtained for interstation distances of several hundred kilometres. The high-amplitude wave trains of PRO-SALA and MAID-MVL most likely correspond to Rayleigh and Love waves, whereas the wave packages of CIL-RR06 and RR28-RR29 arrive significantly later than expected for surface wave velocities of roughly 3 km s^{-1} . They may be associated with Scholte waves (Scholte 1947) because their velocity is in the range of $0.8\text{--}1.5 \text{ km s}^{-1}$ (Flores-Mendez *et al.* 2012; Le *et al.* 2018). Scholte waves of similar velocities were used by Le *et al.* (2018) to examine OBS clock errors. For estimating clock errors, it is not necessary to know the exact nature of the slow wave packages as long as we are able to retrieve clear and stable wave trains throughout the study period, which is clearly the case in Fig. 2.

Clock drifts are often so tiny that they cannot be detected by eye in the CCF plots, for example, Fig. 2(c). Minor CCF drifts are visible for PRO-SALA (e.g. in March 2015). By contrast, MAID-MVL is dominated by several large clock jumps: +2.5 min in January 2014, an additional 3.5 min jump in early February 2015 and a negative jump back to the initial state in late February 2015, after reinstallation of the repaired GPS unit.

3.3 Clock error measurement

For every station pair and component pair, the clock error of each day is measured by performing a cross-correlation between the RCF and the daily CCFs (for land-to-land station pairs) or the 10 d stacks (for land-to-OBS and OBS-to-OBS pairs). The clock error is taken to be the time-shift that maximizes the Pearson correlation coefficient (CC), which is the maximum amplitude of the cross-correlation between RCF and CCF. Each daily or 10 d CCF is associated with a CC value; the average over all such values is CC_{av} , the average CC achieved over the duration of the deployment, for a given station pair and component pair. If the CC of an individual CCF drops below a certain threshold, chosen as 85 per cent of CC_{av} , then this CCF is rejected in order to ensure a consistently high quality of the daily (or 10 daily) clock error estimates.

Per station pair, this yields up to 9 daily clock error measurements ϵ_i from 3×3 component pairs for land-to-land correlations (EE, EN, EZ, NE, NN, NZ, ZE, ZN and ZZ) and also for land-to-OBS correlations (E1, E2, EZ, N1, N2, NZ, Z1, Z2 and ZZ). OBS-to-OBS station pairs yield up to 16 clock error estimates ϵ_i from 4×4 component pairs (11, 12, 1H, 1Z, 21, 22, 2H, 2Z, H1, H2, HH, HZ, Z1, Z2, ZH and ZZ).

The relatively large scatter that affects clock error estimates derived from a single component can be decreased substantially by forming a weighted average over estimates from all 9 or 16 components, following Hobiger *et al.* (2012):

$$\epsilon(t) = \frac{\sum_{i=1}^N CC_i^2(t) \cdot \epsilon_i(t)}{\sum_{i=1}^N CC_i^2(t)}, \quad (3)$$

where ϵ denotes our best estimate of the clock error of one station pair, CC_i is the CC of component pair i and N is the number of usable component pairs. The correlation coefficients CC_i^2 serve as

weights for the individual clock errors ϵ_i . Similarly, a weighted average $CC(t)$ is calculated for each station pair by averaging over its component pairs (Hobiger *et al.* 2012):

$$CC(t) = \frac{\sum_{i=1}^N CC_i^3(t)}{\sum_{i=1}^N CC_i^2(t)}. \quad (4)$$

Fig. 3 visualizes the substantial benefit of averaging over component pairs, on the example of station pair CIL-SALA (time period December 2013–February 2015). Fig. 3(a) shows $\epsilon(t)$, our best estimate of the CIL-SALA clock error obtained as the weighted average (eq. 3) over the error estimated from nine components, which are individually shown in Figs 3(b)–(j). The orange curve, repeated in all 10 panels, is the best spline fitting function of polynomial degree 3 to $\epsilon(t)$ in Fig. 3(a). Standard deviations σ to this curve are calculated for the averaged clock error $\epsilon(t)$ as well as for individual component pairs, and are shown in the bottom right of each panel (Figs 3a–j). It is evident from these values and also from visual comparison that Fig. 3(a) shows a much reduced scatter compared to the other nine panels. In general, the σ of a weight-averaged $\epsilon(t)$ is typically two to three times smaller than the σ of its individual constituent components. This observation can be expected from statistics, where it is known as standard error: the standard deviation is reduced by the square root of the number of measurements (Hughes & Hase 2010):

$$\sigma = \frac{\sigma_i}{\sqrt{n}} \quad (5)$$

In our case, σ_i represents the standard deviation of an individual channel pair and n indicates the number of channel pairs used for averaging. Thus, we would expect that the standard deviation is reduced by a factor of three when nine component pairs are used instead of one.

For the average and each component in Figs 3(a)–(j), we also show the time-averaged correlation coefficient CC_{av} , that is, CC values averaged over the period of December 2013–February 2015. High CC values are seen to generally coincide with tightly clustered curves, and are thus an indication for a high confidence clock error measurement. The exception is component pair NE with a relatively low CC_{av} of 0.73 (e.g. compared to ZZ with $CC_{av} = 0.90$), and yet its standard deviation is significantly lower (NE: $\sigma = 55.8 \text{ ms}$ and ZZ: $\sigma = 78.8 \text{ ms}$).

The clock error estimate of a station can be further improved by averaging over all station pairs including this station. This step is analogous to the average over component pairs in eqs (3) and (4), but with N now indicating the number of station pairs. Averaging over multiple station pairs can only be performed when it is known that only the station common to all station pairs can be affected by timing problems, whereas the timing of all other stations is assured (e.g. due to GPS synchronization over the entire recording period). Else more than one clock error would be present, with no clear path separating them out. Fig. 4 demonstrates the benefit on the example of station SALA, where the averaged error estimate in Fig. 4(a) is much less scattered, and has a lower σ of 18.4 ms, than individual estimates from station pairs CAM-SALA ($\sigma = 36.6 \text{ ms}$), CIL-SALA ($\sigma = 29.8 \text{ ms}$) or PRO-SALA ($\sigma = 44.4 \text{ ms}$) in Figs 4(b)–(d). Jointly considering Fig. 3(j); $\sigma = 78.8 \text{ ms}$ for component pair ZZ) and Fig. 4(a; $\sigma = 18.4 \text{ ms}$), our example shows that compared to the conventional method of using only the ZZ-component of one station pair, clock error accuracy can be improved by a factor of 4 by averaging over all available components and station pairs.

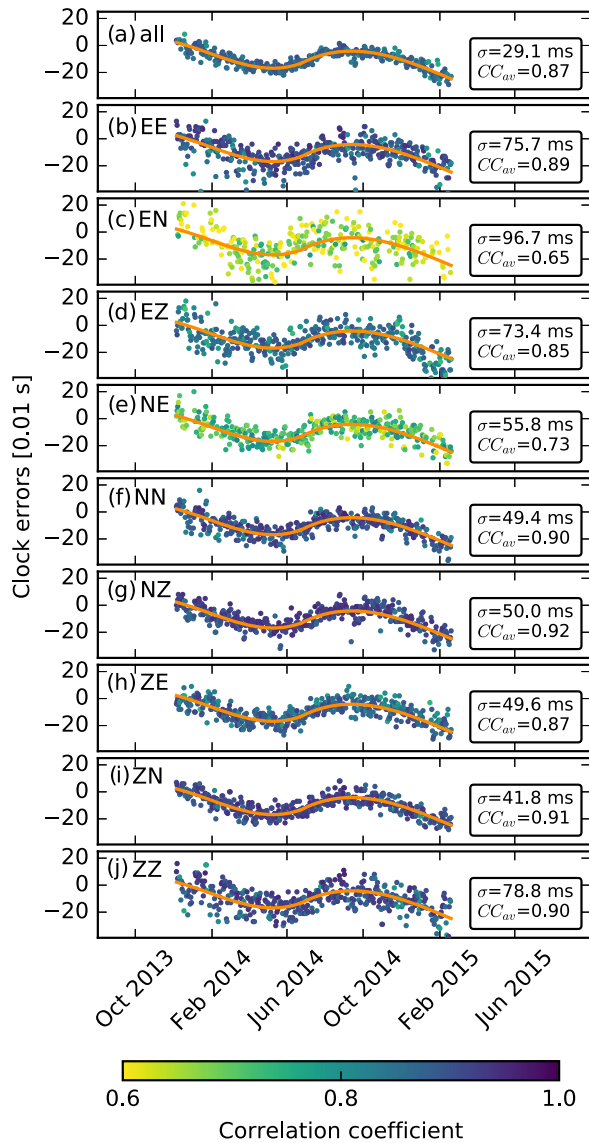


Figure 3. Clock drift of island station SALA relative to island station CIL from December 2013 to February 2015, as estimated from multicomponent cross-correlations. (a) shows $\epsilon(t)$, the clock drift estimates over time obtained from averaging over all nine individual component pairs. Each dot represents the result from one daily CCF. (b–j) show the same clock drift as estimated by the nine individual component pairs (east-with-east EE, east-with-north EN, etc.). Clock drift estimates are given in units of 0.01 s, corresponding to one sample. The clock errors of (a) are fitted by a polynomial function given by the orange curve, which is also plotted in all subsequent panels. Boxes on the right of each panel state measurement uncertainty σ in milliseconds, defined as the standard deviation of the ensemble of daily measurements from the orange fitting curve. σ thus quantifies the accuracy of our clock error estimates $\epsilon(t)$: the σ and point scatter in panel (a) are strongly reduced compared to those of any individual component in panels (b–j). CC_{av} is the average cross-correlation coefficient of the daily CCFs (average taken over the entire deployment period).

4 RESULTS

4.1 Correlations of land stations to land stations

The four problematic RHUM-RUM island stations SALA, CBNM, POSS and MAID, which had suffered from temporary outages of their GPS units, were correlated with five stations (CAM, CIL, MAT,

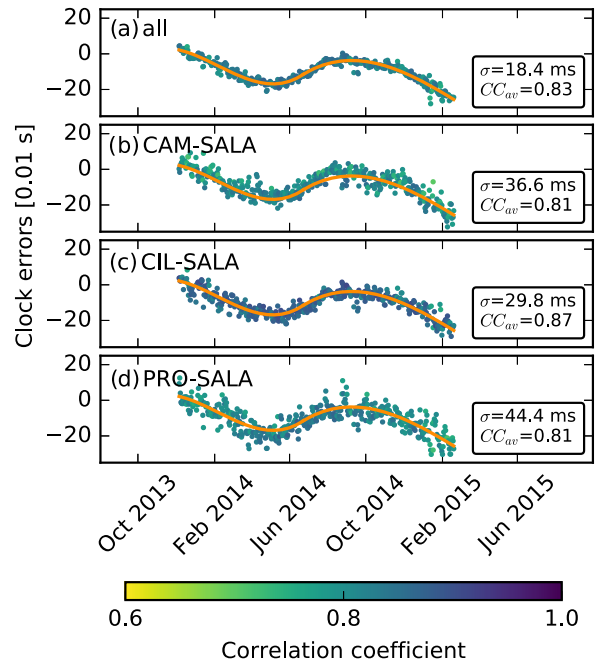


Figure 4. Clock drift of island station SALA relative to the three island stations CAM, CIL and PRO used for correlation, for the same time period as in Fig. 3. (a) shows daily $\epsilon(t)$ estimates after averaging over all three station pairs, together with the best-fitting polynomial curve (in orange). Panels (b–d) show daily $\epsilon(t)$ estimates for the three individual station pairs (CAM-SALA, CIL-SALA and PRO-SALA), each already averaged over up to nine component pairs, and superimposed by the same polynomial curve as in (a). Boxes on the right contain uncertainty estimates σ and averaged CCF correlation coefficients, as defined in caption of Fig. 3. Note that panel (c), CIL-SALA, is identical to panel (a) in Fig. 3.

MVL and PRO), which are permanently operated by the OVPF volcano observatory and served us as reference stations. This yielded 20 station pairs, with interstation distances ranging between 10 and 43 km. Measurements over the nine component pairs were averaged as described in Section 3.3. For averaging over station pairs, only the three or four best station pairs were taken into account, as indicated by the solid black lines connecting station locations in Fig. 1(b). The OVPF stations are designed to record data only when GPS signal is available (Valérie Ferrazzini, personal communication, 2015 September 15), and should therefore never be afflicted by clock errors. We confirmed this by correlating the reference stations with each other, and not observing any clock errors. (Note that this also serves as an independent check on the validity of our method.) We conclude that the clock error measurements presented below originated solely from the RHUM-RUM island stations.

Fig. 5 shows our best estimates of the relative clock errors of SALA, CBNM, POSS and MAID over time. Each dot represents the clock error of one day, so that a rising trend indicates an increasing cumulative error. Clock errors are given in units of samples (1 sample \equiv 0.01 s), except for station MAID, whose large errors are plotted in seconds. Positive clock errors are caused by clocks that run fast, causing the waveforms to appear delayed. Gaps in the data reflect periods when the recording stopped until it was re-started during station servicing visits, which happened every few months. Since the clock errors of these four stations were caused by the loss of GPS synchronization, the time period immediately after re-installation of a repaired GPS unit can be regarded as reference

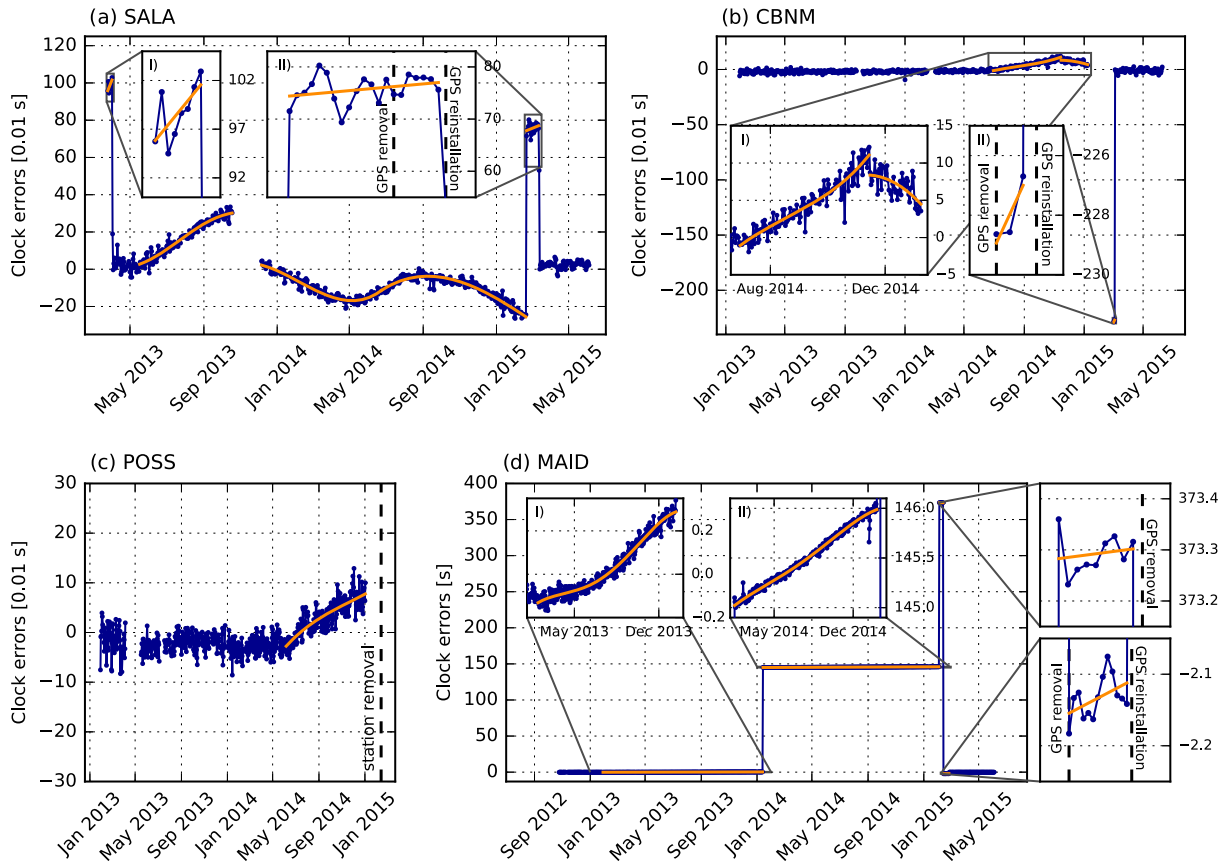


Figure 5. Best estimates of clock drift for all four investigated island stations, obtained by averaging over all available component and station pairs. Blue dots represent daily clock error estimates for (a) SALA, (b) CBNM and (c) POSS, given in units of 0.01 s, corresponding to one sample. Panel (d) shows results for station MAID, in units of seconds. The dates of GPS unit removals and re-installations are represented by vertical dashed lines. Time intervals during which the clocks were unsynchronized and thus adrift are fitted by polynomial functions (orange curves). Note that part of panel (a), drift of station SALA from December 2013 to February 2015, is identical to panel (a) in Fig. 4.

period during which the station clock is known to have worked properly. The clock error estimates for such a reference period need not be centred around time zero, because the clock error measurement is a relative measurement: 1 or 10 d CCFs are correlated with an RCF obtained by averaging CCFs over much or all of the operating period. This means that in general RCFs can be contaminated by clock errors as well.

For station SALA in Fig. 5(a), comparison of the reference period (after GPS re-installation in mid-March 2015, see zoom inset II on the right) with earlier times indicates that the station had no GPS signal when it was installed in late March 2013. After 8 d (zoom inset I on the left), the station probably latched onto a satellite signal and synchronized, which caused a clock jump of around 1 s, followed by a period of no significant drift during April 2013. The GPS connection appears to have been lost again in May 2013, as evidenced by the onset of a gradually accumulating, positive clock error. Subsequent clock errors manifest as gradual, nonlinear clock drifts (fitted by orange polynomial splines), as well as clock jumps of 0.3–1 s.

In stations CBNM (Fig. 5b) and POSS (Fig. 5c), the clock worked normally during most of the operation period, until it started to drift in late June 2014 in CBNM (left inset I and orange fitted curve in Fig. 5b). After a recording stop of CBNM in January 2015, the station’s GPS sensor was removed and re-installed 3 d later,

which manifests as a jump of 2.3 s (right inset II), back to relative clock error values around zero. Similarly, the clock of station POSS started to drift in early June 2014 (orange line in Fig. 5c). Station POSS was removed from the field in February 2015, and its repaired GPS sensor was re-installed at station MAID, whose GPS sensor could not be repaired. Station MAID (Fig. 5d) exhibits striking clock jumps of up to 6 min, which had already been identified in the CCF plot of Fig. 2(b; see Section 3.2). Zooming into time periods between jumps reveals nonlinear drift behaviour (orange curves), as for the other stations.

The purpose of computing the fitting polynomials is to correct for clock errors according to these best-fitting curves, instead of using the scattered daily estimates. This follows the rationale that true clock drifts are probably rather smooth because the physical oscillators are unlikely to run fast on one day and slow on the next. Polynomial degrees are chosen according to the length of the drift interval: linear if the interval is of only a few days (e.g. SALA, insets I and II of Fig. 5a), whereas for longer periods a polynomial degree up to 4 is used (e.g. MAID, insets I and II of Fig. 5d). The nonlinear clock drifts for longer periods are in the range of -2.1 to $+2.4$ ms d⁻¹ for SALA and -1.3 to $+1.5$ ms d⁻¹ for CBNM. POSS shows a moderate clock drift with 0.5 ms d⁻¹ on an average, while the drifts for MAID range between 0.4 and 2.9 ms d⁻¹.

4.1.1 Implementation of clock corrections on the archived RHUM-RUM data and meta-data

The ultimate purpose of estimating clock errors is to correct seismograms archived at the data centre such that end users can rely on downloading correctly timed data and meta-data. For implementing this timing correction, we use the polynomials fitted to the estimated clock error curves, rather than individual daily values of the clock error curves themselves. The RHUM-RUM time-series are archived as miniSEED at the RESIF data centre (<http://dx.doi.org/10.15778/RESIF.YV2011>) in data records of 10–15 s length. Hence the correction procedure (Wayne Crawford, personal communication, 2017 December 17) looks up the value of a fitting curve in absolute time increments of 10–15 s and modifies the header of each data record accordingly, more specifically header fields 8 (start time of record); 16 (time correction); and 12 (bit 1: time correction applied). The time-corrected RHUM-RUM seismograms will be made available through the RESIF data centre in 2018. Time-corrected seismograms at the data centre have a data quality flag of ‘Q’ whereas the uncorrected seismograms have a data quality flag of ‘D’.

The user experience of this clock correction might differ depending on a user’s handling software, and by the length of the time-series requested. For example, ObsPy (Beyreuther *et al.* 2010; Megies *et al.* 2011; Krischer *et al.* 2015) works with a single sampling rate per time-series. Hence if user requests 2 d of data, divided into two consecutive chunks, then the last sample of the first chunk will be wrong by the equivalent of one day’s drift, whereas the first sample of the second time-series will be essentially correct, and spacing of these two samples will differ from the (constant) sampling interval within each chunk.

In order to verify our clock error corrections, the entire cross-correlation procedure was repeated on the time-corrected seismograms, divided into 24 hr chunks. Fig. 6 shows the results: flat clock error values around zero and over each station’s entire operation period indicate that all clock errors have been corrected successfully, even the large clock jumps of station MAID. The scatter of the residual clock errors around zero can be used to assess the accuracy of our method, as discussed in Section 5.1.

This same approach to re-writing meta-data at the data centre and verification of the corrected seismograms has been applied to clock error estimates obtained for OBS stations, as described in Sections 4.2 and 4.3.

4.2 Correlations of land stations to the nearest OBSs

Clock drifts of the five OBSs deployed closest to the island of La Réunion (RR01, RR03, RR05, RR06 and RR27) are evaluated by computing their CCFs with five reference land stations (CIL, HDL, MAT, MVL and PRO). These are essentially the same OVPF reference stations as used for land-to-land correlations, except that HDL is used instead of CAM, because CAM was not installed until July 2013. Correlation of five OBSs with five reference land stations yielded 25 station pairs, with interstation distances ranging between 99 and 174 km. Clock error averaging over the nine component pairs and five station pairs is performed according to eq. (3). We chose not to include the hydrophone component in land-to-OBS correlations because those already benefitted from the large number of station pairs to average over (see Section 5.1).

Immediately before deployment, each OBS clock was synchronized with the GPS clock, so that the clock error is known to have

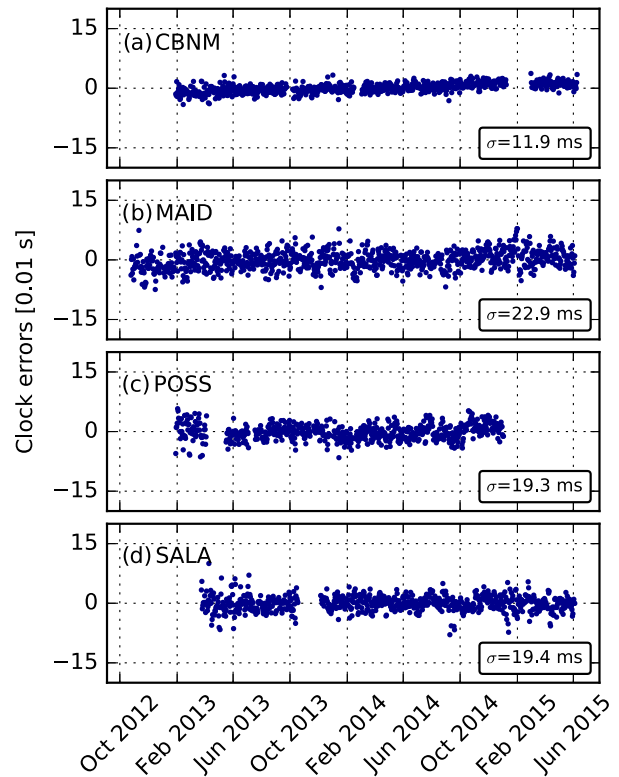


Figure 6. Residual clock errors of land stations (a) CBNM, (b) MAID, (c) POSS and (d) SALA after timing correction, that is, subtraction of the orange curve segments from the daily estimates in Fig. 5. Estimation uncertainty σ , here defined as standard deviation from zero, is given in the bottom right of each panel. They quantify the accuracy of our method for clock error estimation.

been zero at the very beginning of the operation period. Immediately after OBS recovery, a second GPS synchronization attempt was made; if successful, the difference of OBS clock and GPS clock yielded the ‘skew’ or cumulative OBS clock error. Skew is defined as OBS time minus GPS time, so that a positive skew value indicates an OBS clock that runs fast compared to the GPS clock. The time difference (i.e. the OBS time error) is thought to be due to the temperature-dependent oscillator of the internal OBS clock. Due to near-constant temperatures at the bottom of the deep ocean (4 °C), a constant clock drift rate during the operation period is usually anticipated (e.g. Gouédard *et al.* 2014; Hannemann *et al.* 2014; Stähler *et al.* 2016), which suggests that seismograms should be corrected by linear interpolation of the skew for the total period of OBS operation (Lingering doubts remained whether relatively large clock drifts might occur during an OBS free fall to the seafloor upon deployment, and/or rapid rise during recovery, discussed in Section 5.5).

Prior to this study, the timing of RHUM-RUM stations RR01, RR03 and RR05, for which a skew had been successfully obtained, had been corrected by a linear skew interpolation of daily time-series (Stähler *et al.* 2016). No correction was possible for RR06 and RR27, because the OBS clock had stopped working before recovery and thus the second GPS synchronization had failed.

Fig. 7 shows daily clock error estimates for the five OBSs. To the extent that OBS clocks actually drift linearly, we would expect to observe no clock errors for RR01, RR03 and RR05 in Figs 7(a)–(c). This is indeed the case to good approximation, with flat curves scattering around zero, and in particular no or only weak differing

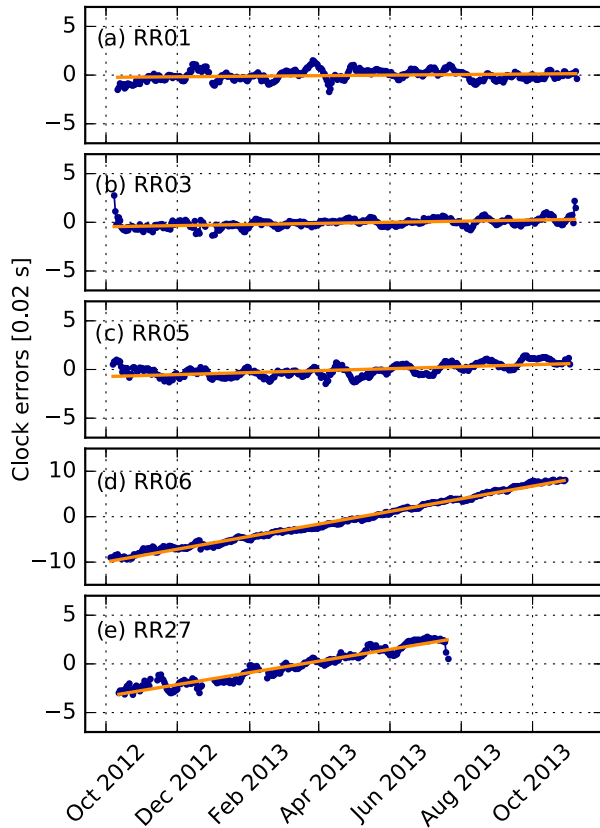


Figure 7. Daily clock drift estimates from the first iteration for ocean bottom stations, obtained through correlation with nearby island stations (land-to-OBS). (a) Station RR01, (b) RR03, (c) RR05, (d) RR06 and (e) RR27. Drift values are given in units of 0.02 s, corresponding to one sample. Seismograms of RR01, RR03 and RR05 had been linearly skew corrected prior to correlation. The observation that their daily estimates scatter around zero with no discernible trend over deployment time implies that their clock drift rate was indeed almost constant and accurately quantified by the GPS skew measurement. No such independent skew measurements were available for RR06 and RR27. Their daily estimates show clear, almost linear trends over deployment time, again confirming that clock drift rates were almost constant. Drift rates are estimated as the slopes of the best-fitting orange lines, and listed in Table 1 for all OBSs.

trends near the very beginning and end of the recording periods. This provides independent confirmation that linear skew interpolation is an appropriate clock error correction for OBSs. The tiny jumps at the beginning and end are very likely processing artefacts, rather than a true clock drift, due to the fact that the first and last stacks contain only one CCF, whereas the second and second-to-last stacks contain two CCFs.

Processing the non-corrected seismograms of RR06 and RR27 with our method reveals clear clock drifts that are well approximated by a linear fit, with slopes of 0.91 ms d^{-1} for RR06 and 0.40 ms d^{-1} for RR27 (Figs 7d and e). After applying these linear timing corrections and re-running the processing (calculation of daily CCFs, computation of 10 d stacks and RCF, clock error estimation), a small clock drift remained measurable. This may be due to the usage of 10 d CCF stacks (instead of daily), which flattens the clock drift slightly. In addition, the correlation with an RCF affected by clock errors may play a role (Sens-Schönfelder 2008; Gouédard *et al.* 2014). A second iteration of linear clock drift correction is necessary and sufficient to remove the clock error completely. Thus the

full clock drift is the sum of drifts from the first and second iterations (RR06: $0.91 + 0.15 \text{ ms d}^{-1} = 1.06 \text{ ms d}^{-1}$ and RR27: $0.40 + 0.11 \text{ ms d}^{-1} = 0.51 \text{ ms d}^{-1}$). The necessity of several iterations will be explained in more detail in Section 4.3.

These are very encouraging results, indicating that our 10 d CCF stacking period still offers sufficiently high timing resolution to successfully estimate and remove OBS clock drifts. The 10 d stacking interval is necessary to obtain sufficiently high SNRs for stable CCFs when OBS interstation distances are up to several hundred kilometres. Our study is the first to present successful OBS cross-correlations and clock estimates over such large distances with the comparatively short stacking length of 10 d.

4.3 Correlations of OBSs to OBSs

The encouraging outcome from land-to-OBS correlations in Section 4.2 led us to expand our study to the entire OBS network. This permits to verify measured skews, the general applicability of linear skew interpolation, and the estimation of clock drifts for OBSs that lack skew measurements.

We calculate 10 d stacked CCFs of 89 OBS station pairs, using neighbouring stations with distances ranging from 16 to 374 km, averaging $\sim 209 \text{ km}$ (Fig. 1a). We include the hydrophone component, which is particularly useful because the seismic channels of several OBSs failed, whereas the hydrophones worked very reliably (Stähler *et al.* 2016). Up to 16 component pairs are averaged according to eq. (3), as in previous sections.

By comparing the clock error curves over time of individual component pairs, we detected that the H-component of several OBSs shows a behaviour that differs from the seismometer components. In most cases, the deviations between the curves were limited to the first few months of the operation period. We find that this behaviour can be attributed to a change in hydrophone noise levels, probably related to a protracted settling period on the seafloor, which amounts to an unexpected malfunctioning of the hydrophone model used. Details are still under investigation, see Appendix A. To avoid a degradation of our clock error estimates, we do not admit CCFs that included a hydrophone channel prior to its settling into proper operation, which is characterized by a higher dynamic range of the instrument response (Fig. A1c). Eq. (5) predicts a reduction in standard deviation by a factor of $\sqrt{12} \approx 3.5$ from averaging over 12 component pairs, compared to the hypothetically possible $\sqrt{16} = 4$ in the case of 16 component pairs. This relatively small loss in estimation accuracy led us to exclude the questionable hydrophone components.

We proceed in two stages: (1) For the 29 OBSs where skew values had been obtained, we verify these skews and the validity of their linear interpolation over the recording period for the purpose of clock correction. (2) For the 24 OBSs without skew, we estimate and correct clock drifts, using as reference stations the previously verified OBSs with skews.

For a given OBS with a skew value, Stage 1 estimates the full time dependence of its clock errors by attempting to average over several correlating OBS pairs, as in previous sections (eq. 3 again, where N is now the number of neighbouring stations). However, the requirement that all correlating OBSs need to be in the ‘with skew’ group, means that only a few correlating pairs may be available (e.g. only RR16-RR17 and RR16-RR18 for station RR16). In some cases, only an indirect skew verification can be carried out, for example, for RR31, which can only be correlated with OBSs that lack a skew measurement. RR29-RR30 and RR30-RR31 show the

same clock drift, indicating that the source of the clock drift is the station lacking a skew measurement (RR30). The timing of the station RR31 can therefore be inferred as accurate, especially since the timing of RR29 can be verified directly with several other station pairs. We find that 26 out of 29 skew-corrected stations show no additional clock drift as expected, which verifies both the measured skew values and the validity of their linear interpolation.

The three exceptions are station RR12, which shows a small linear clock drift of -0.65 ms d^{-1} , and stations RR07 and RR11, each of which experienced one clock jump of roughly 1 s during their respective operation periods (Fig. 8a). By comparing raw data with skew-corrected data, we find that the negative clock drift of RR12 was caused by a sign error: seismograms had mistakenly been corrected and archived using a skew value of $+0.11 \text{ s}$ instead of -0.11 s (the latter reported correctly by Stähler *et al.* 2016). This mistake produced an induced skew of -0.22 s , which corresponds to a clock drift of -0.57 ms d^{-1} . This induced clock drift is in good agreement with our measured drift of -0.65 ms d^{-1} . The presumable origin of the clock jumps in RR07 and RR11 is discussed in Section 5.4. Due to the usage of 10 d CCF stacks, the date of the jumps can only be estimated within a 2 d time window. Clock error corrections for RR07, RR11 and RR12 were conducted as for OBSs without skew, see discussion of Stage 2 below.

For the estimation and correction of clock errors of 24 OBSs without skews in Stage 2, the 26 OBSs with successfully verified skews from Stage 1 are regarded as reference stations. As before, we attempt to average eq. (3) over several station pairs in order to estimate an accurate clock drift time-series for a ‘skewless’ OBS (e.g. the pairs RR29–RR30 and RR30–RR31 for ‘skewless’ station RR30). However, the requirement that one station in each station pair needs to be a reference station tends to cut down on the number of successful station pairs due to larger interstation distances. We exclude from the averaging process station pairs with low CC s ($CC_{av} < 0.35\text{--}0.4$, dotted black lines in Fig. 1), because this tends to flatten the clock drift curve. This flattening behaviour can particularly be observed in the first iteration (Fig. 9). The station pairs can be included in the averaging process in later iterations if the flattening behaviour has disappeared. We suspect that this flattening effect stems from the alteration of the RCF by clock errors, with the effect that some parts of the RCF may be destructively stacked rather than constructively, such that the RCF does not resemble individual CCFs. This explanation is supported by the observation that the flattening effect is mainly visible in the first iteration. Moreover, the effect is more severe the lower the CC and the higher the drift rate.

Special consideration was required for the SWIR array, a subarray of 8 OBSs on the southwest Indian Ridge that were spaced at much closer distances (16–42 km) than the other OBSs (see inset in Fig. 1a). For all 8 OBSs, the skew measurement had failed. Hence, we initially estimated and corrected the clock drift of only one SWIR array station (RR47) by correlating with reference stations outside the subarray. RR47 was then used as a reference station to correlate with the remaining seven SWIR array OBSs. Note that errors made on the clock drift of RR47 (e.g. drift underestimation) would propagate to the clock drift estimation of the other SWIR OBSs.

All investigated OBSs show clock errors that clearly follow a linear trend over time (Fig. 8b). For most stations, the correction process required to completely remove these trends consists of several iterations of the kind described in Section 4.2. The required number of iterations typically increases if the clock drift is large, and/or if the CCFs have low SNRs (linked to large interstation

distances and associated low CC_{av} ; Gouédard *et al.* 2014). For instance, four iterations are required for RR47 ($CC_{av} = 0.44$, total drift: 8.03 ms d^{-1}), whereas one iteration is sufficient for RR12 ($CC_{av} = 0.60$, total drift: -0.65 ms d^{-1}). Fig. 10 visualizes this iterative process of clock drift reduction on the example of station RR13. In each iteration, we correct the absolute timing of the seismograms until the slope fitted to the clock drift measurements has dropped to less than 0.1 ms d^{-1} . This value is considered to be indistinguishable from zero clock drift, from comparisons to clock drift estimates performed on skew-corrected OBSs. Total clock drift is obtained as the sum of drift estimates from the individual iterations, as in Section 4.2. On the example of RR13 in Fig. 10, total clock drift is the sum of the drifts from the first, second and third iterations ($4.09 + 1.03 + 0.24 \text{ ms d}^{-1}$). Since the drift estimate of the fourth iteration (0.06 ms d^{-1}) has dropped below the threshold value, this iteration is regarded as a control iteration, which signals that the correction procedure has converged. To generate the graphical representation of drifts in Fig. 8, total clock drifts are determined by comparing the CCFs of the first iteration with the RCF of the last included iteration (for RR13: third iteration). In contrast to individual iterations, clock drift starts at approximately zero units because the RCF of the last iteration is almost free from contaminating effects, given that clock errors have nearly been removed prior to stacking individual CCFs into the RCF. The clock drifts in Fig. 8 can be fitted by linear drift rates that essentially correspond to the sum of drifts from individual iterations.

We successfully estimate total absolute clock drifts for 23 out of 24 OBSs that lacked a skew measurement. These absolute drift rates range from 0.21 to 8.77 ms d^{-1} . For one OBS (RR32), no reliable clock error estimates could be obtained because this station’s seismometer failed, its hydrophone was affected by high noise, and correlating adjacent stations RR31 and RR33 also happened to be noisy (Stähler *et al.* 2016).

For RR35, we detect a clock jump of 0.94 s in addition to a linear clock drift (Fig. 8b). This resembles the jumps at ‘with skew’ stations RR07 and RR11, see discussion in Section 5.4.

Table 1 summarizes our estimates of total clock errors (both drifts and jumps) for all RHUM-RUM OBSs (including results from land-to-OBS correlations), together with parameters such as the number of required iterations and the average CC . The high CC for the SWIR OBSs can be attributed to their small distances to reference station RR47 ($\sim 25 \text{ km}$). The timing of the OBS records affected by clock errors was corrected according to calculated linear drifts (see OBSs marked blue in Table 1) analogous to the procedure of the land stations described in Section 4.1.

5 DISCUSSION

5.1 Method accuracy

Using cross-correlations of ambient noise, we successfully correct timing errors of broad-band land seismometers as well as OBSs, including hydrophone channels. Per station pair, we average up to 16 component pairs and multiple station pairs, with interstation distances ranging from 10 to 374 km. The accuracy of our method can be evaluated by the standard deviation σ of the clock errors calculated from the time corrected seismograms. Thus, we assess the method accuracy after clock correction by calculating the standard deviation of the residual clock errors. The evaluation of this residual scattering around the baseline of zero is the same accuracy criterion as calculating the standard deviation of the clock errors

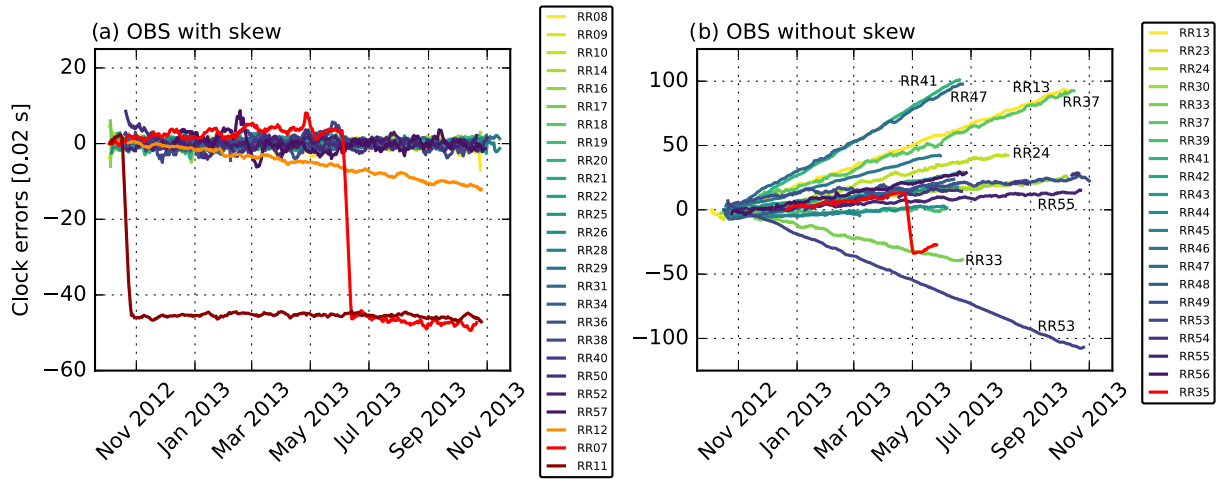


Figure 8. Daily clock drift estimates for ocean bottom stations, obtained through correlation with neighbouring OBSs (OBS-to-OBS). (a) Clock errors of all OBSs that could be linearly time-corrected prior to correlation, thanks to existing skew measurements. Drift values are given in units of 0.02 s, corresponding to one sample. The vast majority of stations show no trend away from zero, providing direct observational evidence for the validity of linear skew correction. Three aberrant stations displaying clock drift or jumps are discussed in the text. (b) Clock error estimates for OBSs that could not be time corrected *a priori*, for lack of skew measurements. Almost all stations that clearly deviate from zero do so in a near-linear fashion, an observational confirmation of near-constant rates of clock drift. RR35 shows a clock jump, discussed in the text.

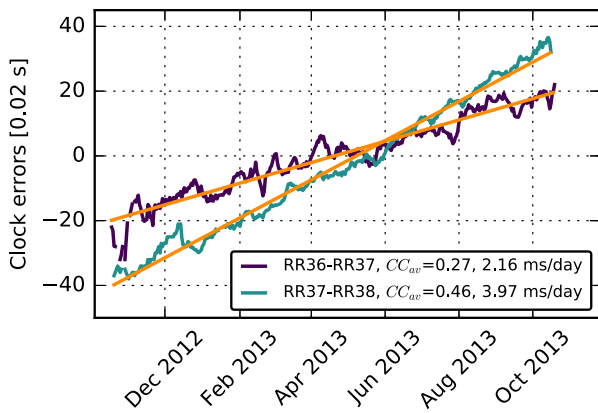


Figure 9. Practical complications in estimating drift rates of OBS clocks, on the example of station RR37. Plot shows estimates from the first iteration for two correlation pairs RR36-RR37 and RR37-RR38 (OBS-to-OBS correlations), and their best linear fits (orange lines) with slopes of 2.16 and 3.97 ms d^{-1} . These slopes are clearly different and yet should be identical because both RR36 and RR38 are skew corrected (thus not contributing any drift). The explanation is that clock drift tends to be underestimated for low $CC_{av} < 0.35-0.4$, such as here. The solution is to execute several iterations of CCF/RCF computation, daily drift estimates and linear timing correction, as discussed in Section 4.3. Table 1 gives the required number of iterations for each OBS (3 for RR37).

to the fitting function (land stations: polynomial function, OBSs: linear function), as performed in Section 3.3 (see Figs 3 and 4). These residual clock error variations are caused by the imperfect convergence of the CCFs to a perfectly stable function, where large interstation distances, short time windows for CCF stacking and the number of usable station and component pairs (N_{sc}) determine the limits of achievable accuracy in clock error estimates (Stehly *et al.* 2007; Sens-Schönfelder 2008).

The σ estimates for clock errors of the land stations range from 11.9 to 22.9 ms (see Fig. 6), while individual component pairs show σ values of 30–114 ms. This enhancement stems from our averaging process, where we took 27 or 36 station and component pairs into

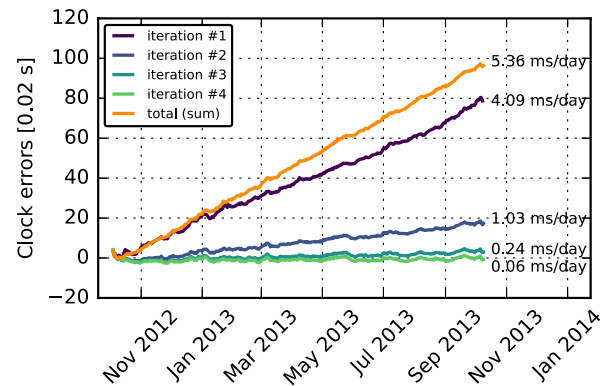


Figure 10. Iterative procedure for estimating OBS clock drift rate on the example of RR13. Coloured lines are estimates of clock error over deployment time. Three iterations of linear estimation and correction are necessary. Drift rate estimates decrease in every iteration (#1: 4.09 ms d^{-1} ; #2: 1.03 ms d^{-1} ; #3: 0.24 ms d^{-1}) until the rate of the fourth iteration (0.06 ms d^{-1}) is indistinguishable from residual scatter, signalling convergence. Our best estimate of the actual (total) clock drift rate is 5.36 ms d^{-1} , the sum of partial rates from iterations 1–3. For better visual reference, the clock drifts of individual iterations are aligned to approximately share the same starting point as the summed drift curve.

account which should reduce the standard deviation statistically by $\sqrt{27}$ or $\sqrt{36}$ (see eq. 5).

The σ estimates for the OBSs immediately adjacent to La Réunion (which were correlated with land stations) are even smaller, ranging between 8.4 ms (RR06) and 12.8 ms (RR05), see Table 1. This can most likely be attributed to the use of 10 d CCF stacks in the land-to-OBS correlations, as opposed to daily CCFs in the land-to-land correlations. Additionally, we were able to average over a higher number of station and component pairs ($N_{sc} = 45$), which clearly improves the accuracy compared to individual component pairs (σ values of 22–62 ms).

The σ estimates for clock errors of more distant OBSs (from OBS-to-OBS correlations and 10 d stacks) range between 10.2 and 43.9 ms, with a mean of 20.7 ms (Table 1). These σ values are found

Table 1. Summary of clock drift estimates for all 57 RHUM-RUM OBS stations. Clock drifts are stated in units of ms d^{-1} , that is, as best-fitting slope to the temporal succession of individual clock error estimates, and also as cumulative drift in ms over a hypothetical, 365 d long deployment. σ is the estimated uncertainty on individual clock error estimates, which is not time-dependent and quantifies our method's accuracy. For example, the clock of RR06 ran fast, accumulating an error of 388.0 ± 8.4 ms over the course of 1 yr. Light blue shading indicates stations affected by statistically significant clock drift. Non-shaded rows denote OBSs where drift rates are not distinguishable from zero within their σ bounds, they coincide with stations that were linearly skew corrected prior to processing. Three exceptions, discussed in the text, are skew-corrected stations that were additionally affected by apparent clock jumps (RR07 and RR11) or by a clock drift due to a sign error of the applied skew correction (RR12). In the case of clock jumps (RR07, RR11 and RR35), the affected stations are indicated by dark blue shading, and two clock drift rate estimates are given: for before and after the jump. Jump occurrences can be constrained to within the 2 d intervals given in brackets below the jump magnitudes. CC_{av} is the average cross-correlation coefficient of a station's CCFs from the first iteration compared to the RCF of the last included iteration. D_{av} is the average distance to neighbouring stations; N_{it} is the number of CCF iterations required to converge on the stated clock drift estimates. N_{sc} denotes the number of station and component pairs used for averaging, for example, RR25 was correlated with three stations, but for one of these stations only the H-component was usable resulting in a N_{sc} of 36 (16+16+4 component pairs) for RR25. 'Skew correction' states the magnitude of an *a priori* linear timing correction, if applied. In notes, OBSs correlated with land stations are marked 'land-to-OBS', whereas the default is OBS-to-OBS. 'SWIR' denotes stations in a densely spaced subarray.

OBS	clock drift rate [ms/day]	drift per year [ms]	uncertainty σ [ms]	clock jump [s] (date of jump)	CC_{av}	D_{av} [km]	N_{it}	N_{sc}	skew correction [ms/day]	notes
RR01	0.019	6.9	10.7	-	0.60	125	1	45	1.688	land-to-OBS
RR02	-	-	-	-	-	-	-	-	-	OBS failed
RR03	0.038	13.9	9.0	-	0.50	144	1	45	2.032	land-to-OBS
RR04	-	-	-	-	-	-	-	-	-	OBS failed
RR05	0.067	24.5	12.8	-	0.50	137	1	45	2.341	land-to-OBS
RR06	1.063	388.0	8.4	-	0.55	144	2	45	-	land-to-OBS
RR07	0.261/-0.440	95.3/-160.6	23.0	-0.999 (2013-06-07/08)	0.48	199	2	16	1.357	-
RR08	-0.071	-25.9	25.2	-	0.31	335	1	16	3.585	-
RR09	0.019	6.9	18.3	-	0.49	236	1	32	5.588	-
RR10	-0.022	-8.0	20.1	-	0.50	200	1	32	0.999	-
RR11	1.896/-0.015	692.0/-5.5	10.2	-0.990 (2012-10-21/22)	0.59	117	2	32	1.554	-
RR12	-0.652	-238.0	11.7	-	0.60	137	1	24	correct: -0.284 applied: +0.284	-
RR13	5.365	1958.2	16.3	-	0.42	197	3	32	-	-
RR14	0.098	35.8	19.5	-	0.53	168	1	12	6.041	-
RR15	-	-	-	-	-	-	-	-	-	OBS failed
RR16	-0.020	-7.3	18.1	-	0.62	257	1	32	4.111	-
RR17	0.035	12.8	12.3	-	0.64	219	1	64	4.714	-
RR18	-0.020	-7.3	11.1	-	0.47	244	1	80	0.914	-
RR19	0.085	31.0	18.4	-	0.49	208	1	80	4.324	-
RR20	0.063	23.0	20.0	-	0.46	244	1	32	1.046	-
RR21	0.021	7.7	18.0	-	0.56	230	1	32	0.689	-
RR22	-0.075	-27.4	14.9	-	0.57	190	1	76	2.287	-
RR23	1.522	555.5	24.8	-	0.50	264	2	16	-	-
RR24	3.494	1275.3	36.5	-	0.47	272	2	5	-	-
RR25	0.025	9.1	20.2	-	0.45	206	1	36	1.088	-
RR26	-0.074	-27.0	13.0	-	0.55	213	1	76	1.605	-
RR27	0.516	188.3	9.6	-	0.49	156	2	45	-	land-to-OBS
RR28	0.012	4.4	14.2	-	0.56	185	1	76	7.789	-
RR29	0.075	27.4	16.8	-	0.48	312	1	32	8.467	-
RR30	1.464	534.4	16.8	-	0.45	279	2	28	-	-
RR31	-0.192	-70.1	18.7	-	0.42	306	1	12	-2.085	-
RR32	-	-	-	-	-	-	-	-	-	no reliable result
RR33	-3.298	-1203.8	20.4	-	0.37	180	4	16	-	-
RR34	-0.098	-35.8	20.4	-	0.37	180	1	16	-3.233	-
RR35	1.878/6.367	685.5/2324.0	22.5	-0.942 (2013-04-26/27)	0.45	197	3	8	-	-
RR36	0.200	73.0	33.6	-	0.44	187	1	4	7.669	-
RR37	5.108	1864.4	36.3	-	0.48	213	3	4	-	-
RR38	-0.050	-18.3	36.3	-	0.48	213	1	4	-0.150	-
RR39	0.299	109.1	22.4	-	0.44	230	1	8	-	-
RR40	-0.056	-20.4	17.4	-	0.44	247	1	16	0.475	-
RR41	8.773	3202.1	16.3	-	0.78	42	2	16	-	SWIR
RR42	2.928	1068.7	12.9	-	0.71	32	2	4	-	SWIR
RR43	1.404	512.5	12.7	-	0.81	25	1	16	-	SWIR
RR44	0.723	263.9	21.1	-	0.84	18	1	16	-	SWIR
RR45	0.209	76.3	18.8	-	0.78	16	1	16	-	SWIR
RR46	3.902	1424.2	14.7	-	0.85	20	2	16	-	SWIR
RR47	8.026	2929.5	17.4	-	0.44	247	4	16	-	SWIR
RR48	2.572	938.8	19.2	-	0.81	23	1	16	-	SWIR
RR49	0.913	333.2	36.8	-	0.61	171	1	4	-	-
RR50	-0.083	-30.3	36.8	-	0.60	171	1	4	4.339	-
RR51	-	-	-	-	-	-	-	-	-	OBS failed
RR52	0.051	18.6	19.7	-	0.36	263	1	16	2.419	-
RR53	-6.135	-2239.3	11.0	-	0.60	93	2	32	-	-
RR54	1.746	637.3	25.0	-	0.49	203	2	8	-	-
RR55	0.827	301.8	17.8	-	0.53	169	2	12	-	-
RR56	2.594	946.8	22.3	-	0.44	287	2	20	-	-
RR57	-0.127	-46.4	43.9	-	0.46	211	1	4	3.419	-

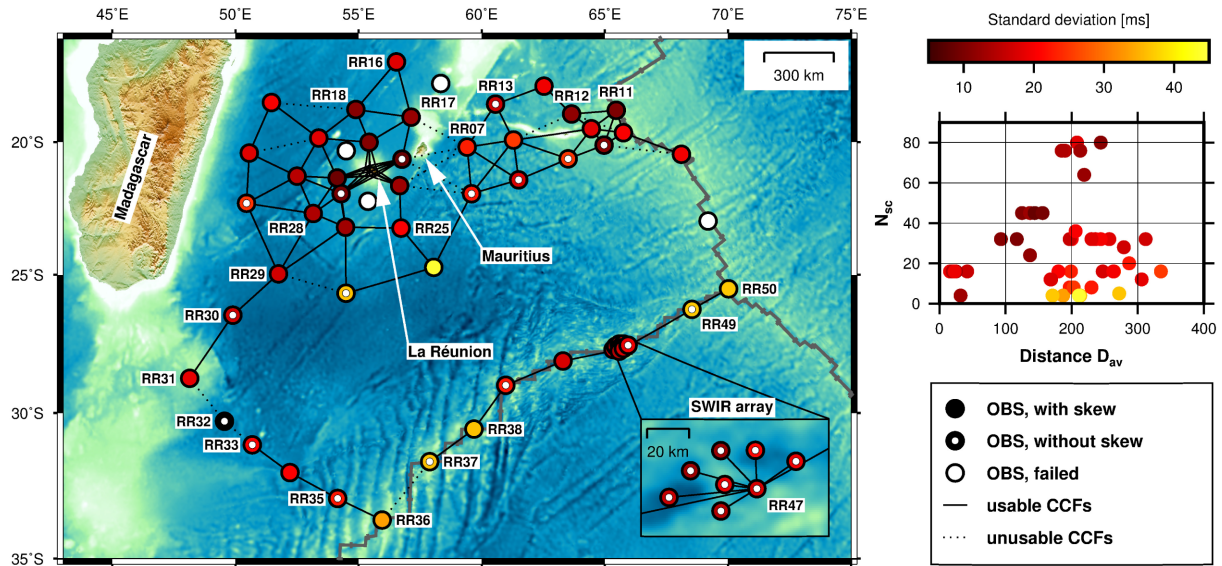


Figure 11. Graphical summary of the uncertainties (standard deviation σ) affecting our OBS clock error estimates. The σ values tend to be low (and estimation accuracy high) where interstation distances D_{av} are small and/or in the centre of the network (corresponding to a high N_{sc}), where many neighbouring stations are available for correlation. Right panel shows σ values (same colouring as in map) as a function of N_{sc} and D_{av} , confirming this relationship. (Solid and dotted black lines denote successful and unsuccessful correlation pairs, respectively.) RR32 is shaded black because no reliable estimates could be obtained. Figure generated with the GMT toolbox (Wessel *et al.* 2013).

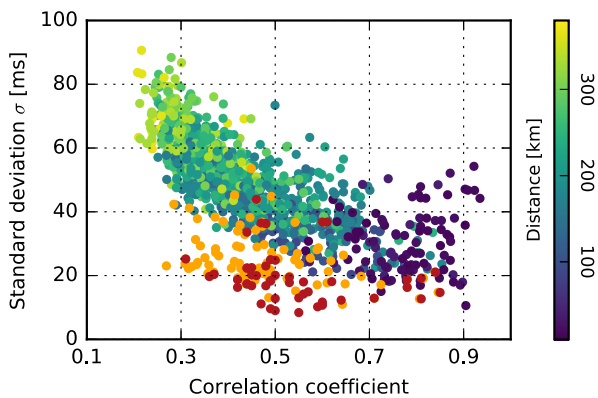


Figure 12. Beneficial effect of averaging on the accuracy of clock drift estimates, on the example of OBS-to-OBS correlations. Standard deviations σ are shown as a function of the correlation coefficient for 1076 component pairs, where colouring indicates distance. Orange dots denote σ values for the 87 station pairs (after averaging over component pairs). Red dots represent σ for the 52 OBSs (after averaging over component and station pairs).

to depend strongly on interstation distance and on the number of station and component pairs (N_{sc} of 4–80) available for averaging, see Fig. 11. As for land-to-land and land-to-OBS correlations, the σ estimates are significantly reduced compared to those of the individual component pairs (11–91 ms) due to the performed averaging. For instance, the σ value of RR45 ($\sigma = 18.8$ ms) is quite similar to the value of RR19 ($\sigma = 18.4$ ms), even though the interstation distance is much shorter for RR45 (16 km) than for RR19 (208 km). However, RR19 benefitted from the higher number of station and component pairs ($N_{sc} = 80$) available for averaging, compared to $N_{sc} = 16$ for RR45. Fig. 12 visualizes the considerable benefit of averaging over multiple component and station pairs for the OBSs: σ estimates are plotted as a function of CC for the 1076 component pairs, for the 87 station pairs (averaged over the component

pairs) and for the 52 stations (averaged over component and station pairs). The σ values decrease with increasing CC , and higher CC s are usually correlated with smaller interstation distances. Averaging over component pairs tremendously reduces the standard deviation (orange dots in Fig. 12, *cf.* Fig. 3); this effect is further enhanced by averaging over multiple station pairs (red dots in Fig. 12, *cf.* Fig. 4). Additional relationships between distance, SNR, CC and σ value of the 1076 component pairs are plotted in Appendix B. Note that σ values and interstation distances are quite similar for the land network and the SWIR OBS array due to a trade-off between N_{sc} and stacking duration. At the SWIR OBSs, the lower N_{sc} (that should lead to a higher σ value compared to the land stations) cancels the effect of using 10 d stacks (that should lower the σ value compared to 1 d stacks).

Compared to the stations’ sampling rates of 50, 62.5 and 100 Hz, that is, sampling intervals of 20, 16 and 10 ms, the σ uncertainties on clock drift estimates are on the order of just one sample. The clock drift estimates themselves are typically of a few milliseconds per day (Table 1), which means that within recording intervals of well over a month, clock drift rates can be reliably estimated. Even for shorter recording durations a clock drift estimate is possible (e.g. Fig. 5a, inset II: clock errors range between 70 and 80 samples). Such approximate clock drift corrections are much preferable to none at all.

5.2 Comparison with prior work

A comparison with the parameters and achieved accuracies of prior work on clock error estimation is given in Table 2. Stehly *et al.* (2007) reported a standard deviation σ of 47–110 ms for an interstation distance between land stations of roughly 200 km, using CCFs stacked over 1 month. Le *et al.* (2018) give a σ value of 200 ms for OBS interstation distances of 60–270 km and using 11 d stacks of the HH-component, as compared to our σ uncertainties of 11.1–43.9 ms on drift estimates derived from OBSs spaced by

Table 2. Accuracy of clock error estimates: comparison of our method to prior studies. The most pertinent parameters are station type (land versus OBS), typical interstation spacing and length of the time windows for CCF stacking. Uncertainty σ is the estimated, typical accuracy of clock error estimates in each study. The exact definition of σ differed somewhat across studies: for Stehly *et al.* (2007) and Sens-Schönfelder (2008) it is very similar to ours. Gouédard *et al.* (2014) and Le *et al.* (2018) used a bootstrapping approach that defines σ as the standard deviation of their 200 bootstrap realizations. Hannemann *et al.* (2014) define σ as deviation of CCF results from measured OBS skews, that is, they take the applicability of strictly linear OBS drift for granted.

Study	Type	Distance (km)	CCF stacking window	Accuracy σ (ms)
This Study	land stations	10–43	1 d	11.9–22.9
	OBSs	16–374	10 d	8.4–43.9
Stehly <i>et al.</i> (2007)	land stations	~200	1 month	47–110
Le <i>et al.</i> (2018)	OBSs	60–270	11 d	~200
Sens-Schönfelder (2008)	land stations	0.2–5	1 d	8–63
Gouédard <i>et al.</i> (2014)	OBSs	~4.5	1 d	~5
	OBSs	10–15	1 month	n/a
Hannemann <i>et al.</i> (2014)	OBSs	<75	20 d	87–129

200–335 km (Table 1). For this we stacked CCFs over only 10 days, but with averaging over as many station and component pairs as available. Sens-Schönfelder (2008) found standard deviations of 8–63 ms for land seismometer spaced by 0.2–5 km and using day-long CCFs, as compared to 11.9–22.9 ms for our land stations, which were spaced by 10–43 km and also used day-long CCFs. Hence we achieved the same or better accuracy as prior work on clock drift estimates, but for larger interstation distances and/or at higher temporal resolution. This demonstrates the significant benefit of averaging clock drift estimates over multiple component and station pairs.

In Fig. 3, we have defined the error σ of our clock drift measurements as the standard deviation of individual daily clock drift estimates from a fitted polynomial spline of degree 3. Our study shows that for any single-component CCF, most of this error (i.e. point scatter in any of panels Figs 3b–j) is random noise, because it averages out to a several times smaller scatter and σ in Fig. 3(a), which shows the average over all nine component pairs. Although the predominantly random nature of measurement noise on single-component estimates could be suspected by prior workers (because they observed increasing σ for decreasing SNR and/or shorter stacking windows), the suppression of this noise by our averaging technique provides direct confirmation that it is indeed mainly random, as opposed to systematic noise or true clock drifting on the timescale of the stacking window. The small clock error wiggles that can be observed in some OBSs (especially in Fig. 7) could possibly be linked to tides, as the duration of one wiggle is roughly 1 month. This observation could merit a future study.

For OBSs, the *a priori* expectation of linear clock drift due to constant temperature at the ocean bottom is clearly confirmed by our results. This is in agreement with the findings of Gouédard *et al.* (2014), Hannemann *et al.* (2014) and Le *et al.* (2018), who obtained OBS clock drifts that are to first-order linear. A temperature dependence is not evident for the investigated land stations. For example, station MAID shows positive clock drift rates over more than 2 yr (February 2013–February 2015, divided by a jump in January 2014) without any seasonal variations (Fig. 5d).

The use of several iterations in estimating clock error measurements was first proposed by Sens-Schönfelder (2008), in order to iteratively suppress the alteration of the RCF by clock errors. We find that this procedure is especially necessary when dealing with CCF stacking windows longer than 1 d, that is, for all OBSs. By contrast, the timing of our land stations can be fully corrected in one iteration (Fig. 6). Besides CCF stacking length, the number of required iterations is primarily determined by the magnitude of the timing errors and by the SNR of the CCFs (Gouédard *et al.* 2014).

We find that poor SNRs (low *CC* values) can lead to an underestimation of clock drift. This can be avoided by excluding station pairs with average *CC* below 0.35–0.4, in agreement with Sens-Schönfelder (2008), who adopted an acceptance threshold of *CC* > 0.4. The OBS (RR32) for which we could not obtain a reliable result exhibits a *CC_{av}* of 0.23, which is clearly below our *CC* threshold.

5.3 Consistency with skew measurements and lab experiments

For OBSs, both the magnitude and sign of estimated clock drifts are in good agreement with measured skew values, where available (Stähler *et al.* 2016). The majority of the RHUM-RUM OBSs were equipped with clocks that ran too fast (Table 1), yielding delayed waveforms.

Stähler *et al.* (2016) made an attempt to measure OBS clock drifts in the laboratory, on the clocks of 7 ‘skewless’ RHUM-RUM recorders (the others were unavailable, having been re-deployed elsewhere). We find that their results are inconsistent with our drift estimates, which is very likely due to the experimental shortcoming of running the recorders at ambient lab temperatures rather than at temperatures characteristic of the ocean bottom (see Appendix C for details). Stähler *et al.* (2016) had similarly noted that the drift rates found in their experiments were much lower than those implied by their skew measurements, and that for the one clock (RR11) that both had a skew and ran in the lab, the two values obtained were inconsistent.

5.4 Apparent clock jumps—true clock failures or missing data samples?

In addition to linear clock drifts, our CCFs detect apparent clock jumps at three OBSs (RR07 and RR11 with skew; RR35 without skew, Fig. 8). If these jumps had been produced by actual jumps of the physical station clock, we would expect the clock error to return to zero upon re-synchronization with GPS after OBS recovery (GPS synchronization was successful for RR07 and RR11). The reason is that the timing of the last sample would have been corrected properly according to the skew, regardless of any clock jumps that happened earlier, as explained by Fig. 13(a). However, such a return to zero drift is not indicated by the CCFs. Hence there must be another cause for the apparent clock jumps. We ruled out the possibility of artificially induced jumps due to an incorrect skew correction. Instead, we suspect that the apparent clock jumps are caused by missing samples, that is, that in each case the data logger failed to write a short chunk of about 50 samples to disk. If this failure

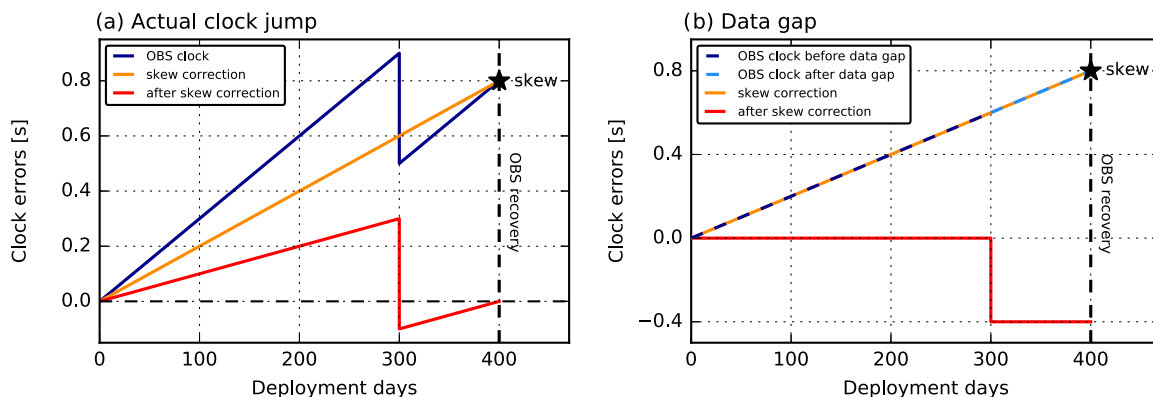


Figure 13. A jump in a clock error curve of an OBS can be due either to a true, physical jump of the data logger clock, or to a batch of missing samples on disk. Our method can distinguish between these two cases if a skew measurement is available (as for stations RR07 and RR11, *cf.* Fig. 8(a)). (a) Schematic visualization of the effect of an actual clock jump. Blue curve depicts a scenario where the recorder clock is affected by a constant drift rate, plus a clock jump 300 d into the deployment. Due to the jump, the measured skew value (star) underestimates the true clock drift rate (slopes of orange versus blue curves). Red curve shows the residual clock error after linear skew correction, which is the curve that our method would (approximately) estimate. The timing of the last sample is corrected properly, but clock drift is underestimated before and after the jump. (b) Clock error curves for the case of missing samples, where 300 d into the deployment, a batch of samples corresponding to 0.4 s recording duration was not written to disk. The physical clock is always correct, and so are the measured skew value (star) and hence the estimated drift rates before and after 300 d (dashed dark blue and light blue curves). However, a constant erroneous timing offset of -0.4 s is assigned to all samples after the data gap so that the linear skew correction (red curve) removes all clock drift, but the constant offset remains for all times after the data gap. The estimated clock drift curves of RR07 and RR11 in Fig. 8(a) resemble the red curve in case (b) rather than in case (a).

went unnoted by the logger, a wrong timing (too early) would have been assigned to all samples after the gap. Indeed, all apparent jumps have this sign towards early times (Table 1). Although this problem was rare (three occurrences across the entire OBS network and deployment period), the fact that it occurred in the exact same manner at three different stations (once per station) also points to a technical weakness of the data logger model used in the German DEPAS OBSs.

If the jumps are caused by data gaps, we would expect no clock drift before and after the jump for OBSs subjected to a linear skew correction (RR07 and RR11), as visualized in Fig. 13(b). Indeed we observe almost no clock drift after the jump in RR11. Clock drift in RR11 before the jump is rather high (see Table 1) but this can be explained by the short duration of this period (12 d), which allows no representative drift estimation. For RR07, the estimated clock drift rate before the jump tends to be slightly higher than the skew-derived rate, and after the jump it tends to be slightly lower (see Table 1), unlike the expectation for a physical clock jump (Fig. 13a). The small positive drift before the jump yields a total clock error as high as that produced by the small negative drift after the jump, but of opposite sign. This indicates that the linear skew approximation was applied properly but can be refined with our method. One could argue that this refinement is just a matter of measuring clock drift rates in different windows (before and after the jump). This is true, but the measured drifts appear sufficiently evident to reflect a real refinement of the skew correction. Observational evidence for RR07 is therefore compatible with our suggested explanation of missing samples.

For RR35, the OBS lacking a skew value, we would expect identical clock drift rates before and after the jump, irrespective of its cause. This is not strictly the case in Fig. 8(b), but again the different clock drift rates (see Table 1) may well be due to the short remaining recording duration (25 d) after the jump, which prevents a proper estimate of a long-term trend. *A priori* it is unlikely that the physical clock changed its drift rate in the presumed absence of sudden, significant seafloor temperature variations (only) around the time of the jump.

An undetected data gap manifests as an apparent clock jump and can only be distinguished from a true, physical clock jump when a measured skew value is available (Fig. 13a versus b). This is the case for RR07 and RR11, and their drift curves in Fig. 8(a) clearly support a data gap. In absence of a skew measurement, we cannot totally reject the possibility that a true clock jump occurred at RR35 (red curve in Fig. 8b) but the nearly identical jump magnitude of 1 s at all three OBSs reinforces our inference that the jump has the same technical cause in all three OBSs: a batch of 50 samples (probably corresponding to one data buffer worth of recordings) must not have been written to disk in each instance, given the sampling rate of 50 Hz.

The good news is that clock timing corrections using our CCF method succeed regardless of the nature and cause of these apparent clock jumps, whereas the problem would go undetected and uncorrected in the standard practice of simple, linear skew interpolation. Missing samples (and/or true clock jumps) may or may not be common across the wide variety of OBS recording systems that are in use. The method demonstrated here can shed light on this question both for past and future OBS systems.

5.5 Rapid clock drift immediately after deployment?

We have not considered the possibility of very rapid clock drift at the very beginning of the OBS recording period (amounting to a jump within our temporal resolution). Hypothetically, this could occur due to the very rapid temperature drop during sinking to the ocean bottom. Such a time offset can be detected by comparing the causal and the acausal parts of the CCFs because waves would arrive earlier in the causal part than in the acausal part, or vice versa (Gouédard *et al.* 2014). The time offset is then given by the half-shift between both parts. Applied to daily or 10 d stacked CCFs, this method would have been another possibility to estimate clock errors over time. A detailed method description is given by Gouédard *et al.* (2014). A big disadvantage of this technique is its requirement of highly symmetric CCFs, while for our presented method the stability of

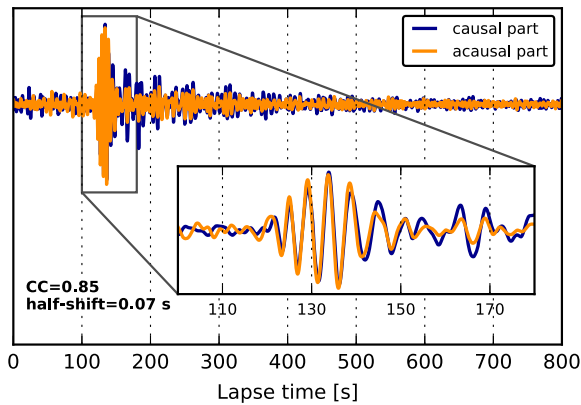


Figure 14. Attempt to detect a hypothesized clock jump in the first hours or days of the deployment through visual comparison of the causal and acausal (time reversed) parts of the RCF of RR49-RR50 (OBS-to-OBS correlation, HH-component). Correlation of the two parts yields a half-shift of 0.07 s with a (high) correlation coefficient of 0.85, indicating no large time offset and hence no indication of abnormal clock behaviour, which might have resulted from the rapid temperature drop during OBS sinking to the seafloor.

(symmetric or asymmetric) CCFs over time is sufficient (Fig. 2). To extract a potential shift that could have occurred at the beginning of the deployment, we correlate the causal and acausal parts of RCFs derived from time-corrected data for all component pairs and station pairs (see example in Fig. 14). However, high variability in the results even between the different channel pairs of one station pair limits the significance of this test. We observe no large shifts (>0.5 s), while the method is not suited to reliably detect real tiny clock jumps, because highly symmetric CCFs of high SNR would be required.

The following considerations show that OBSs are unlikely to accumulate a practically significant clock drift during the first hours or days of (non-stationary) operation. Sinking from surface to ocean bottom occurred within roughly 4 hr, entailing a temperature change from ~ 22 °C (at surface) to 4 °C (at ocean bottom). Presumably, this is accompanied by a proportional change in clock drift rate, since clock frequency is primarily controlled by temperature. In this short thermal equilibration span of hours or even a day, undetected clock drift that is significant (compared to total operation over 1 yr) could only be incurred if drift rates at ambient and transient conditions were one or more orders of magnitude larger than at the seafloor. Lab experiments under ambient conditions showed the opposite to be the case: clock drift were observed to be less severe under ambient conditions (see Appendix C and Table C1).

6 CONCLUSIONS

The cross-correlation of ambient noise is shown to be a routinely applicable method to detect and correct even small clock errors in seismograms and hydrograms of passive experiments with deployment times of several months or more. It allows for a time-continuous monitoring, in contrast to the comparison of earthquake arrival times (Anchietta *et al.* 2011) or burst events of persistent localized microseismic sources (Xia *et al.* 2015; Xie *et al.* 2018). Our estimates of clock drifts as a function of absolute time were successful and quantifiable for all stations in the RHUM-RUM array that had recorded usable data, that is, 4 broad-band island seismometers and 52 broad-band OBSs in deep water.

We show that all three seismometer components and any hydrophone channels can and should be used for the clock

error measurements. Cross-correlating seismograms with hydrophones yield stable and fully usable CCFs despite their different physical nature. The same is true for correlations between land stations and OBSs, which sample vastly different crustal conditions.

For four land stations on La Réunion, interstation distances ranged between 10 and 43 km, for which CCF stacking over short 1 d windows was sufficient. The clocks were affected by complex clock drifts and jumps. During linear drift episodes, typical drift rates ranged between -2.1 and $+2.9$ ms d^{-1} . The uncertainties σ on these estimates were 11.9–22.9 ms and did not vary with time of year.

Near-shore OBSs were correlated with land stations, over distances of 99–174 km, requiring 10 d CCF stacks, with σ between 8.4 and 12.8 ms. OBS-to-OBS distances ranged between 16 km and 374 km, also requiring 10 d time averaging and yielding σ of 10.2–43.9 ms. For all OBSs, we confirm the *a priori* expectation of overwhelmingly linear clock drift, at absolute rates of 0.2–8.8 ms d^{-1} . For the 29 OBSs where a skew value had been obtained upon recovery, the implied linear drifts are consistent with our direct drift estimates.

Typical measurement uncertainties σ for the clock error at any time are thus 20 ms. This corresponds to 1–2 samples (at 50–100 Hz sampling rate); or to 0.3 per cent of the Rayleigh wave traveltime for station distances of typically 20 km on land; or to 0.03 per cent of Rayleigh wave traveltime for typical OBS-to-OBS distances of 200 km (assuming a velocity of 3 kms^{-1}).

Errors on clock drift estimates seem to be caused by imperfect convergence of the CCFs to a function that is perfectly stationary over the recording period. The method reaches its limits when interstation distances increase and/or CCF stacking time windows are shorter than desirable (due to a trade-off with the temporal resolution required to capture actual clock drifting pattern; Stehly *et al.* 2007; Sens-Schönfelder 2008). Within these limiting systematics, the accuracy of clock error estimates can be improved several fold (typically 3–4 fold) by averaging estimates from all seismogram and hydrophone components and from multiple station pairs, where available.

This method is a powerful tool to estimate and correct clock drifts in OBSs and land stations that are beyond the reach of a GPS signal, and to verify the common linear skew correction applied to OBS data. The method can diagnose actual, physical clock drift (e.g. caused by changes in water temperature, which affect clock oscillator frequency), as well as apparent clock jumps, which seem to be due to a data logger's failure to write a batch of samples to disk.

Our method's robust and accurate performance on very heterogeneous seismogram and hydrophone data suggests additional new applications. The timing quality of existing temporary and permanent networks could be assessed and corrected, for ongoing and historical recordings. This would render many problematic data sets first usable, or better usable, for the purpose of structural imaging of mantle and crust, and is particularly pertinent to the expensive but challenging seismograms and hydrograms from OBS deployments. The (re)use of such data will provide valuable contributions to future imaging studies. A different and new imaging application will be the generation of long-range noise correlation functions for the structural imaging of crust and uppermost mantle, most importantly in ocean basins. The routine use of interstation distances exceeding 200 km could previously not be envisaged.

ACKNOWLEDGEMENTS

We are grateful to Valérie Ferrazzini for providing seismological data from the monitoring network of Observatoire Volcanologique du Piton de la Fournaise on La Réunion island. The temporary RHUM-RUM island seismometers were installed and maintained by GB, Céline Davy, Eric Delcher, Fabrice Fontaine and John-Robert Scholz. We thank Henning Kirk at AWI Bremerhaven for discussions on hydrophone responses of the DEPAS OBSs. We are grateful to Wayne Crawford for applying the corrections determined in this study to the archived RHUM-RUM data at RESIF data centre. We thank the participants and crew of cruises *R/V Marion Dufresne* MD192 and *R/V Meteor* M101.

This study arose as part of the RHUM-RUM project (www.rhum-rum.net), funded by Deutsche Forschungsgemeinschaft in Germany (grant nos. SI1538/2-1 and SI1538/4-1) and by Agence Nationale de la Recherche in France (project ANR-11-BS56-0013). KS received additional funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement no. PCIG14-GA-2013-631104 'RHUM-RUM'. CH acknowledges support from the Emmy Noether program (grant no. HA7019/1-1) of the Deutsche Forschungsgemeinschaft.

Additional RHUM-RUM project funding in France came from the CNRS (Centre National de la Recherche Scientifique, France); from TAAF (Terres Australes et Antarctiques Françaises, France); from IPEV (Institut Polaire Français Paul Emile Victor, France); and in Germany from Alfred Wegener Institut Bremerhaven. For the provision of OBS instruments, we acknowledge the DEPAS OBS pool in Bremerhaven (Deutsche Geräte-Pool für Amphibische Seismologie); GEOMAR Helmholtz-Zentrum für Ozeanforschung in Kiel; and INSU-IPGP (Institut National des Sciences de l'Univers - Institut de Physique du Globe de Paris). We thank the universities of Bonn, Münster and La Réunion for providing the terrestrial seismometers.

The RHUM-RUM data are freely available. They are hosted by the RESIF data centre in France (<http://seismology.resif.fr>), where they have been assigned FDSN network code YV. Please support the viability of future data collection efforts by citing RHUM-RUM data using their DataCite identifier <http://dx.doi.org/10.15778/RESIF.YV2011>.

We thank Katrin Hannemann and one anonymous reviewer for their detailed, constructive reviews.

REFERENCES

Anchieta, M.C., Wolfe, C.J., Pavlis, G.L., Vernon, F.L., Eakins, J.A., Solomon, S.C., Laske, G. & Collins, J.A., 2011. Seismicity around the Hawaiian Islands recorded by the PLUME seismometer networks: insight into faulting near Maui, Molokai, and Oahu, *Bull. seism. Soc. Am.*, **101**(4), 1742–1758.

Barruol, G., & Sigloch, K., 2013. Investigating La Réunion hot spot from crust to core, *EOS, Trans. Am. geophys. Un.*, **94**(23), 205–207.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169**(3), 1239–1260.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. ObsPy: a Python toolbox for seismology, *Seismol. Res. Lett.*, **81**(3), 530–533.

Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercissian, A., 2008a. Towards forecasting volcanic eruptions using seismic noise, *Nat. Geosci.*, **1**, 126–130.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N.M., Nadeau, R.M. & Larose, E., 2008b. Postseismic relaxation along the San Andreas Fault at Parkfield from continuous seismological observations, *Science*, **321**(5895), 1478–1481.

Corela, C., Silveira, G., Matias, L., Schimmel, M. & Geissler, W.H., 2017. Ambient seismic noise tomography of SW Iberia integrating seafloor- and land-based data, *Tectonophysics*, **700-701**, 131–149.

Davy, C., Stutzmann, E., Barruol, G., Fontaine, F.R. & Schimmel, M., 2015. Sources of secondary microseisms in the Indian Ocean, *Geophys. J. Int.*, **202**(2), 1180–1189.

Flores-Mendez, E., Carbajal-Romero, M., Flores-Guzmán, N., Sánchez-Martínez, R. & Rodríguez-Castellanos, A., 2012. Rayleigh's, Stoneley's, and Scholte's interface waves in elastic models using a boundary element method, *J. appl. Math.*, **2012**.

Gouédard, P., Seher, T., McGuire, J.J., Collins, J.A. & van der Hilst, R.D., 2014. Correction of ocean-bottom seismometer instrumental clock errors using ambient seismic noise, *Bull. seism. Soc. Am.*, **104**(3), 1276–1288.

Haned, A., Stutzmann, E., Schimmel, M., Kiselev, S., Davaille, A. & Yelles-Chaouche, A., 2016. Global tomography using seismic hum, *Geophys. J. Int.*, **204**(2), 1222–1236.

Hannemann, K., Krüger, F. & Dahm, T., 2014. Measuring of clock drift rates and static time offsets of ocean bottom stations by means of ambient noise, *Geophys. J. Int.*, **196**(2), 1034–1042.

Hobiger, M., Wegler, U., Shiomi, K. & Nakahara, H., 2012. Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan, *J. geophys. Res.*, **117**(B9), B09313.

Hosseini, K. & Sigloch, K., 2017. ObspyDMT: a Python toolbox for retrieving and processing of large seismological data sets, *Solid Earth*, **8**, 1047–1070.

Hughes, I.G. & Hase, T.P.A., 2010. *Measurements and Their Uncertainties: A Practical Guide to Modern Error Analysis*, Oxford University Press.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C. & Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discovery*, **8**(1).

Larose, E., Derode, A., Campillo, M. & Fink, M., 2004. Imaging from one-bit correlations of wideband diffuse wave fields, *J. Appl. Phys.*, **95**(12), 8393–8399.

Le, B.M., Yang, T., Chen, Y.J. & Yao, H., 2018. Correction of OBS clock errors using Scholte waves retrieved from cross-correlating hydrophone recordings, *Geophys. J. Int.*, **212**(2), 891–899.

Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.*, **110**(6), 3011–3017.

Megies, T., Beyreuther, M., Barsch, R., Krischer, L. & Wassermann, J., 2011. ObsPy – what can it do for data centers and observatories? *Ann. Geophys.*, **54**(1), 47–58.

Rossum, G., 1995. Python reference manual, Tech. Rep., CWI (Centre for Mathematics and Computer Science).

Sabra, K.G., Roux, P., Thode, A.M., D'Spain, G.L., Hodgkiss, W.S. & Kuperman, W.A., 2005a. Using ocean ambient noise for array self-localization and self-synchronization, *IEEE J. Ocean. Eng.*, **30**(2), 338–347.

Sabra, K.G., Gerstoft, P., Roux, P. & Kuperman, W.A., 2005b. Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.*, **32**(14), L14311.

Scholte, J.G., 1947. The range of existence of Rayleigh and Stoneley waves, *Geophys. J. Int.*, **5**(5), 120–126.

Sens-Schönfelder, C., 2008. Synchronizing seismic networks with ambient noise, *Geophys. J. Int.*, **174**(3), 966–970.

Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.*, **33**(21), L21302.

Sens-Schönfelder, C., Pomponi, E. & Peltier, A., 2014. Dynamics of Piton de la Fournaise volcano observed by passive image interferometry with multiple references, *J. Volc. Geotherm. Res.*, **276**, 32–45.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, **31**(7), L07614.

- Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, **307**(5715), 1615–1618.
- Stähler, S.C. *et al.*, 2016. Performance report of the RHUM-RUM ocean bottom seismometer network around La Réunion, western Indian Ocean, *Adv. Geosci.*, **41**, 43–63.
- Stehly, L., Campillo, M. & Shapiro, N.M., 2007. Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts, *Geophys. J. Int.*, **171**(1), 223–230.
- Tian, Y. & Ritzwoller, M.H., 2017. Improving ambient noise cross-correlations in the noisy ocean bottom environment of the Juan de Fuca plate, *Geophys. J. Int.*, **210**(3), 1787–1805.
- Wegler, U., Nakahara, H., Sens-Schönfelder, C., Korn, M. & Shiomi, K., 2009. Sudden drop of seismic velocity after the 2004 M_w 6.6 mid-Niigata earthquake, Japan, observed with Passive Image Interferometry, *J. geophys. Res.*, **114**(B6), B06305.
- Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic Mapping Tools: improved version released, *EOS, Trans. Am. geophys. Un.*, **94**(45), 409–410.
- Xia, Y., Ni, S., Zeng, X., Xie, J., Wang, B. & Yuan, S., 2015. Synchronizing intercontinental seismic networks using the 26 s persistent localized microseismic source, *Bull. seism. Soc. Am.*, **105**(4), 2101–2108.
- Xie, J., Ni, S., Chu, R. & Xia, Y., 2018. Assessing the short-term clock drift of early broadband stations with burst events of the 26 s persistent and localized microseism, *Geophys. J. Int.*, **212**(1), 324–332.
- Zhan, Z., Tsai, V.C. & Clayton, R.W., 2013. Spurious velocity changes caused by temporal variations in ambient noise frequency content, *Geophys. J. Int.*, **194**(3), 1574–1581.

APPENDIX A: HYDROPHONE RESPONSE CHANGE

For 12 ocean bottom stations, our analyses uncovered poor performance of hydrophones during their first few months of deployment, followed by settling into high-quality operation for the remainder of the experiment. The 12 affected stations are all German DEPAS instruments, specifically RR01, RR13, RR17, RR18, RR26, RR27, RR35, RR41, RR42, RR54, RR55 and RR57 (the French INSU stations used differential pressure gauges instead of hydrophones).

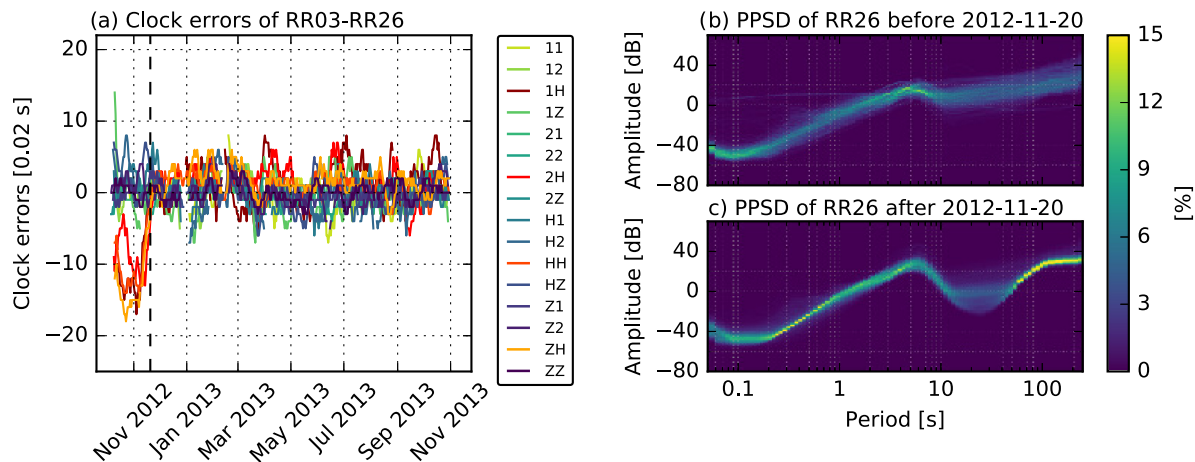


Figure A1. Diagnosis of hydrophone malfunctioning at some stations during the first few months of deployment, on the example of station RR26. (a) Apparent clock drift curves for the 16 individual component pairs of station pair RR03-RR26 (OBS-to-OBS correlation). Up until about 2012 November 20 (dashed vertical line), four clock error curves (in reddish line colours) are seen to deviate from all others. These are the four component pairs involving the hydrophone channel of station RR26 (1H, 2H, HH and ZH). (b) Probabilistic power spectral density plot of the H-channel at RR26 for the time period before 2012 November 20. (c) Same as (b) but for the time period after 2012 November 20. Evidently the hydrophone response changed rapidly and markedly around this date, several weeks into the deployment, for reasons that are unclear. By changing the shapes of the hydrograms and their CCFs, this behaviour translates into apparent clock drifts over time.

Investigation showed that the instrument response of these 12 hydrophones changed markedly over time, and in each case transitioning rapidly from an initial state to a final state, several weeks or months into the deployment. Changes in instrument response causes changes to the recorded waveforms and hence CCFs. The transition event would have caused an apparent time-shift in the CCFs, which is indistinguishable from a clock error in a single CCF involving a hydrophone, but is unmasked by comparison to CCFs that involve only seismometer channels.

Fig. A1(a) shows an example for station pair RR03-RR26. The timing of both stations has already been corrected according to their measured skews, so that we do not expect large clock errors in either station. Nevertheless, clock errors of up to -0.3 s can be observed solely in component pairs including the H-component of RR26 (1H, 2H, HH and ZH), which indicates that the hydrophone channel is responsible for the clock deviations. Probabilistic power spectral densities (PPSD) calculated on the H-channel indeed reveal a marked spectral change around 2012 November 20 (Figs A1b and c), which coincides with the date around which the suspicious CCFs settled into concordant behaviour. More detailed investigation of the hydrophone PPSDs showed that the change in the instrument response (in the period range between ~ 7 and ~ 100 s) occurred gradually over a few days, starting on 2012 November 15 and reaching their final state on 2012 November 20. It is not surprising that the CCFs are sensitive to this noise level change, given that they are computed in the period sub-band of 2–20 s. For example, Zhan *et al.* (2013) showed that a sudden change in frequency content can cause a temporal dilation of CCFs.

We made similar findings for the other 11 OBSs, where the problematic operation interval ranged from 1 to 3 months after deployment. The cause for the initially different instrument responses remains unclear, as does the reason or occasion for their settling into normal operation. Seasonal variations of the noise intensity due to storms (Davy *et al.* 2015) seem to play no role because the issue is limited to 12 out of 52 stations, does not affect the seismometer components, and the dates of the noise transitions differ across affected stations. Moreover, seasonal variations would mainly affect secondary microseisms (~ 7 s), whereas the hydrophone noise

changes also occur at significantly longer periods. From the temporal evolution of the noise, we speculate that certain internal hydrophone components took excessively long to settle into normal working mode. The hydrophones consist of a piezo element that accumulates charge on a capacitor that might need a few hours or even a day to discharge completely, a requirement for proper functioning. Still this process should not take several months. Several hydrophones are of an older but otherwise identical model (HTI-01-PCA instead of HTI-04-PCA in RR03, RR04, RR06, RR11, RR13, RR18, RR20, RR23, RR24, RR35, RR41 and RR53). The older model is slightly over-represented among the faulty ones, but not by much. Hannemann *et al.* (2014) detected also hydrophone malfunctioning at three DEPAS OBSs in the beginning of their operation period, which could be associated with small waveform amplitudes that tend to recover over time. Similarly, we found in some of our erroneous stations that the hydrophone malfunctioning seems to coincide with an amplitude reduction. However, a general correlation with reduced amplitude and stepwise recovery as described by Hannemann *et al.* (2014) cannot be clearly confirmed. Hence further investigation would be needed to clarify the causes of these initially faulty instrument responses in a subset of the German hydrophones.

APPENDIX B: RELATIONSHIPS BETWEEN SNR, CC, DISTANCE AND STANDARD DEVIATION σ

For OBS-to-OBS correlations, we are dealing with a large data set comprising 87 station pairs with interstation distances ranging from 16 to 374 km. For each station pair, we are able to use up to 16 component pairs, yielding a total of 1076 station and component pairs, an unprecedented large data set for a clock error study. Fig. B1 explores the relationships between the four most significant parameters characterizing each correlation pair: interstation distance, SNR, CC and standard deviation σ . SNR and CC values are determined as averages over the individual values of the 10 d stacks. The σ estimates, our measure of estimation accuracy, are calculated from the scatter of the residual clock errors around zero.

Fig. B1(a) shows that high SNRs are associated with high CCs. The relationship can be approximately quantified by fitting a logarithmic curve (orange curve):

$$CC = 0.32 + 0.15 \cdot \ln(\text{SNR} - 3.48). \tag{6}$$

The colouring indicates that high SNRs and high CCs tend to be achieved for small interstation distances.

Fig. B1(b) shows CC values as a function of interstation distance. For each station pair, the CC values of up to 16 component pairs are plotted in a column-like way. Again, high CCs are seen to coincide with small interstation distances. The σ values are lower (and accuracy is higher) for shorter interstation distances. For a given distance, σ estimates tend to be lower for higher CC values. The relationship between σ and CC is presented in Fig. 12. Fig. B1(c) gives the standard deviation σ as a function of distance with CC values represented by the colouring. This plot supports our findings from Fig. B1(b). Both Figs B1(b) and (c) indicate a quasi-linear relationship between CC and interstation distance, as well as between standard deviation σ and interstation distance.

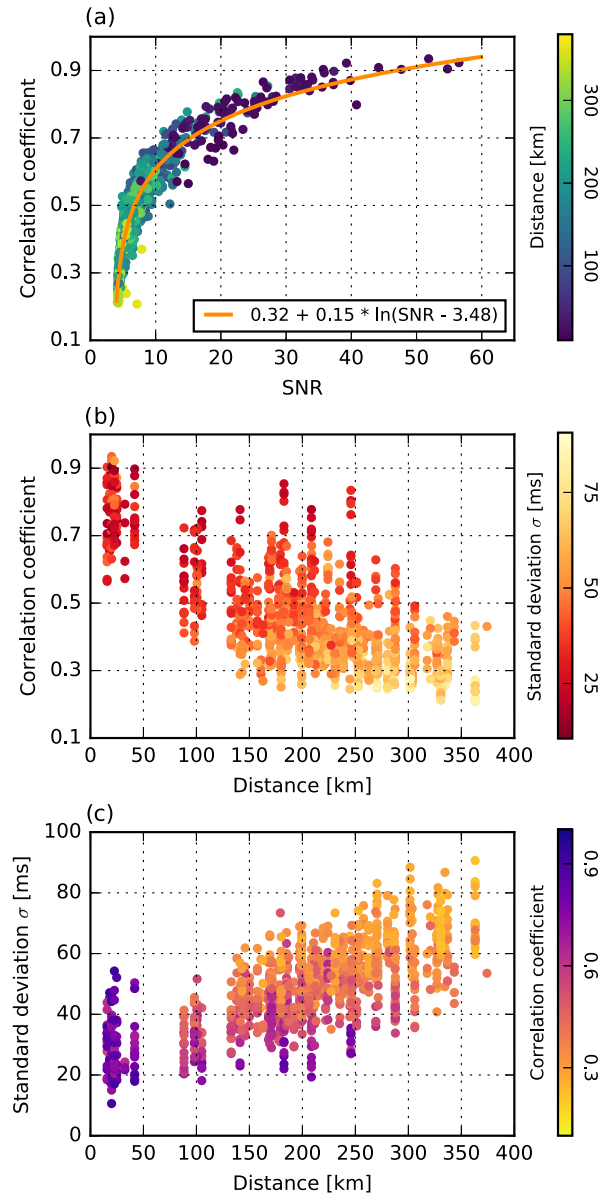


Figure B1. Relationships between SNR, CC, distance and standard deviation σ of the 1076 component pairs from OBS-to-OBS correlations. For details see Appendix B. The relationship between σ value and correlation coefficient is shown in Fig. 12.

APPENDIX C: COMPARISON WITH LABORATORY EXPERIMENTS

For OBSs without skew measurements, Stähler *et al.* (2016) attempted to obtain clock drift estimates by re-running their data loggers in the laboratory. Short of specialized equipment, these tests were run under ambient pressure and temperature conditions. They ran two tests of 7 and 33 d duration, on the seven afflicted stations that had not already been re-deployed at the time (RR06, RR11, RR41, RR43, RR44, RR45 and RR55). The measured skew values were linearly extrapolated to derive a hypothetical skew value for a 365 d long experiment. For direct comparison, we also calculated hypothetical cumulative clock errors after 365 d of deployment from our estimated drift rates (Table 1) for these seven OBSs. The values are given in Table C1, together with the results of the two lab runs.

Table C1. Comparison of our CCF-derived clock drift estimates of Table 1 with direct laboratory measurements of clock skews, as reported by Stähler *et al.* (2016). Seven of the RHUM-RUM OBS data loggers were put through two test runs of 7 and 33 d duration, under ambient temperature and pressure conditions at the facilities of the DEPAS OBS instrument pool at Bremerhaven. For both methods, the estimated/measured clock drift rates are linearly extrapolated to a hypothetical deployment duration of 365 d. The direct skew measurements are inconsistent with our results, presumably because the test runs were not conducted at the low temperatures that define deep sea conditions.

OBS	This study	Lab run (7 d)	Lab run (33 d)
RR06	0.39 ± 0.008 s	0.15 s	0.13 s
RR11	0.57 ± 0.010 s	-0.15 s	-0.21 s
RR41	3.20 ± 0.016 s	0.30 s	0.23 s
RR43	0.51 ± 0.013 s	0.00 s	0.033 s
RR44	0.26 ± 0.021 s	-0.50 s	0.55 s
RR45	0.08 ± 0.019 s	0.045 s	-0.05 s
RR55	0.30 ± 0.018 s	0.0015 s	-0.03 s

Stähler *et al.* (2016) doubted that their laboratory results yielded reasonable approximations of clock drift during the actual ocean

bottom deployment. First, their skew values for station RR44 differed by > 1 s between the two runs. Second, the lab-predicted skew for RR11 differed by ~ 0.8 s from the measured skew. RR11 was the only station among the seven for which a skew had been obtained upon recovery. Our own, noise-based drift estimate for RR11 is consistent with the skew measurement obtained aboard the ship. Third, the skews of both lab runs are generally small in magnitude, compared to skews measured on the ship (for stations other than the seven in question). Consistent with this, our noise-based drift estimates are larger in magnitude than those measured in the lab (except for RR44).

Hence we concur with Stähler *et al.* (2016) that reliable OBS clock drift estimates cannot be gained from laboratory experiments without simulating at least realistic temperature conditions on the ocean bottom, of relatively constant 4°C . Our earlier findings of almost linear OBS clock drifts throughout the deployment indicate that temperature transients upon deployment (sudden decrease from ~ 22 to 4°C) and recovery (sudden increase to $\sim 22^\circ\text{C}$) are too rapid to accumulate a notable clock error.