Microgravity surveying before, during and after distant large earthquakes

Boddice, Daniel; Metje, Nicole; Tuckwell, George

DOI:
10.1016/j.jappgeo.2022.104542

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Boddice, D, Metje, N & Tuckwell, G 2022, 'Microgravity surveying before, during and after distant large earthquakes', *Journal of Applied Geophysics*, vol. 197, 104542. https://doi.org/10.1016/j.jappgeo.2022.104542

Journal of Applied Geophysics 197 (2022) 104542

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

Microgravity surveying before, during and after distant large earthquakes

Daniel Boddice [a,*], Nicole Metje [a], George Tuckwell [a,b]

[a] School of Engineering, College of Engineering and Physical Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[b] RSK, 18 Frogmore Road, Hemel Hempstead, Hertfordshire HP3 9RT, UK

ARTICLE INFO

Keywords:
Microgravity
Noise
Earthquake

ABSTRACT

Microgravity surveying is a widely used geophysical method, especially for detection of deeper features which are difficult to detect with other techniques. However, the accuracy of the readings is strongly affected by the amount of microseism noise, which is typically rejected by using long integration times for each measurement cycle. Large seismic events such as earthquakes create large ground acceleration waves which have been known to affect gravity readings. While previous data which have been collected serendipitously, tend to only record the final readings, unique long-duration datasets during and after three earthquakes including the raw data (sampled at 6–10 Hz) from field gravimeters (Scintrex CG-5 and CG-6) in static locations recorded during surveys are presented. The aim of the current study is to characterise both long and short-term effects of the generated seismic accelerations on measurement accuracy and repeatability, and identify changes to commercial practice to mitigate these effects.

During earthquakes, the nature of the microseism noise was fundamentally altered by each of the different associated seismic waves. Minor effects were found for body P- and S-waves, but much larger effects were found for surface waves, especially the Rayleigh waves which gave errors many times those which would occur in normal conditions. These accelerations persisted in the data for several hours or even days after the earthquake affecting instrument performance. The main finding is that the optimum course of action is to identify the earthquake early by analysing the data in the frequency domain through FFT, switch to 60 s cycles and return to the base station until the strongest waves have passed. Surveying on subsequent days was affected by lower frequency free earth oscillations requiring removal of the unwanted signals using linear trends between at least hourly base stations. Using these techniques will both facilitate data collection as well as improve data confidence in these challenging conditions. Instruments operating in a gradiometer configuration were shown to be comparatively unaffected by the increased noise for typical commercial integration times paving the way for the next generation of instruments to operate successfully, even with this challenging environmental noise.

1. Introduction

Microgravity surveying (detection of signals from subsurface features of a few microGals) is widely used commercially for mineral exploration, environmental studies and the detection of subsurface voids, mine workings, sinkholes and other subsurface hazards (Hinze, 1990; Nabighian et al., 2005; Tuckwell et al., 2008). It is, especially useful for deeper void targets where ground conditions can make the use of other active geophysical techniques difficult as the depth of investigation is limited by signal attenuation, for example when using Ground Penetrating Radar (GPR) in conductive ground.

Microgravity survey involves taking a series of discrete points over the survey site using a precision gravimeter instrument to collect a measurement of the local gravitational field (e.g. Debeglia and Dupont, 2002; Tuckwell et al., 2008; Kaufmann, 2014; Martínez-Moreno et al., 2016). After applying corrections to remove the influence of terrain and other unwanted signals such as the tides, the resulting map is proportional to the underlying mass distribution of the ground. Dense materials cause the instrument to record higher acceleration (a positive gravity anomaly), whereas low-density features (e.g. air-filled voids) create a relative gravity low (or negative gravity anomaly) (Reynolds, 2011; Long and Kaufmann, 2013). However, one of the limiting factors impacting the resolution of the technique is that it is strongly affected by the amount of microseismic noise which is recorded by the instruments mass and spring system, as vibrations in the ground generate accelerations which are indistinguishable from those caused by gravity due to

* Corresponding author.
E-mail address: d.boddice@bham.ac.uk (D. Boddice).

https://doi.org/10.1016/j.jappgeo.2022.104542
Received 16 April 2021; Received in revised form 8 October 2021; Accepted 11 January 2022
Available online 15 January 2022

the equivalence principle. These accelerations affect both the accuracy of an individual measurement cycle (a group of measurements over a time period which is used to average out ground acceleration noise),as well as the repeatability of multiple measurement cycles (Boddice et al., 2018), therefore affecting the reliability of the final gravity map.

Recent use of gravity measurements, with respect to earthquakes, has been focused on future earthquake prediction by identifying which ancient faults are under stress and most likely to be active (Levandowski et al., 2017), and tracing changes in mass distribution (Barnes, 1966; Yiqing et al., 2011), the progression of fault movements using satellite measurements (De Viron et al., 2008; Grocholski, 2018), and developing early warning systems using the instantaneous gravity signal (Zhu and Zhan, 2012; Montagner et al., 2016). Much less attention has been given to the effects of the resulting seismic accelerations on the accuracy of field measurements in typical microgravity surveys, particularly with commercially viable measurement times. Whilst large magnitude earthquakes in the UK are rare, significantly large earthquakes can be detected seismically around the globe using seismometers, and have the potential to affect the usual microseismic noise which can be detected by the gravimeter, both in terms of its amplitude and frequency. These additional accelerations can affect the accuracy of the measurements and even render the technique futile for a span of time, both during the earthquake and during subsequent bursts of seismic accelerations caused by the earth ringing (e.g. Rymer, 1989; Seigel, 1995). From a commercial perspective, even when the triggering earthquake takes place far from the site in question, the seismic waves present an additional source of acceleration noise on surveys which can affect the instrument performance and the optimum integration times. Experiments to quantify these effects are hard to set up due to the unpredictable nature of earthquakes, and suitable datasets can only be acquired serendipitously, making systematic studies difficult if not impossible.

This paper takes unique, long-term microgravity data recorded using commercial field gravimeters in static locations that by chance covered time intervals before, during and after distant earthquakes and examines the subsequent effect on data quality, as well as considering if any potential changes to commercial surveying practice are required when surveying in the immediate aftermath of a distant large earthquake. The objectives of the current study were:

- To identify how the noise caused by accelerations due to different seismic modes affects the gravimeter performance both during and after a distant triggering earthquake and the duration of these effects.
- To identify the best course of action for a field operator when an earthquake occurs during a survey including determining the optimum cycle times to be used in the period following an earthquake, the necessary intervals between base station reoccupations and suitable drift removal approaches
- Examine the effects of earthquakes on instruments in a gradiometer configuration with a view towards characterising the limitations on performance for future instruments.

2. Gravimeters and noise response

Most field gravimeters in use today, such as the Scintrex CG-5 and CG-6, work using a mass mounted on a quartz spring in a near vacuum to increase sensitivity, the force on which is proportional to the strength of the gravitational field at the measurement location, which is used to measure differences in acceleration. However, due to the equivalence principle, measurements of gravity are strongly affected by ground vibrations which are indistinguishable from gravitational acceleration, and cause the mass-spring system to oscillate, affecting the measurement. Typically, the dominant source of these vibrations comes from microseismic accelerations caused by the interaction between waves with the fixed ocean floor and other waves (Ardhuin et al., 2011; Ardhuin and Herbers, 2013; Ardhuin et al., 2015), which typically have a peak frequency of between 0.1 and 0.3 Hz with a bandwidth of approximately 0.05–0.5 Hz. However, the instruments are also affected by a broadband range of frequencies from other seismic modes and anthropogenic vibrational sources (Goncharenko et al., 2018). In surveying practice, the higher frequency acceleration noises (>0.1 Hz) are removed by using a long measurement window (30–60 s) for each measurement cycle to average the effects out by measuring several cycles of the noise, whereas longer period diurnal noise sources as well as the drift of the instrument (caused by the slow relaxation of the quartz spring) are removed by making repeated measurements at a single point (base station) over the course of the survey day and removing the underlying trends (Gabalda et al., 2003; Tuckwell et al., 2008). One common estimation of the error for a single cycle of microgravity measurement is the root mean square (RMS) error, which can be estimated using Eq. (1) (Scintrex Ltd., 2006).

$$RMS\ Error = \frac{SD}{\sqrt{N_{readings}}} \tag{1}$$

Under normal survey conditions, the typical RMS error for a 30 s measurement will be less than 10 μGals. The presence of different frequencies of ground acceleration noise on the gravimeter can be assessed using the instrument's raw data which records at 6 Hz (CG-5) and 10 Hz (CG-6) during each measurement cycle. The visibility of accelerations in the instrument's raw data is affected by the sampling rate of the sensor, which limits it to those with frequencies less than the Nyquist frequency (i.e. half the sampling frequency) (Leis, 2011). Accelerations with a frequency greater than the Nyquist frequency are misidentified through “aliasing” and appear as lower frequencies. The sampling rates of the CG5 and CG6 instruments used in the current study cause accelerations with frequencies over 3 and 5 Hz respectively to appear as lower frequencies or as a broadband range of different frequency noise signals within the visible range. An example of how this behaviour may affect the recorded accelerations is shown in Fig. 1. For a simulated hypothetical sinusoidal acceleration signal with a frequency of 7.02 Hz. Lower frequencies are generally more of an issue for gravity measurements because they result in less complete cycles making the integration procedure less effective at removing them.

3. Effects of earthquakes on gravity readings

Earthquakes introduce a whole range of different frequency accelerations (Kulhánek, 1990; Greco et al., 2008; Niebauer et al., 2011; Ghobadi-Far et al., 2019) which can influence both the accuracy of individual readings as well as the repeatability of measurement cycles. The effect of large earthquakes on microgravity measurements is strongly dependent on the magnitude and depth of the earthquake and distance of the recording instrument from the epicentre. There are four main types of signal produced with the potential to affect microgravity measurements:

1. Instant static (permanent) (Barnes, 1966; Montagner et al., 2016) and transient changes (Harms et al., 2015) in gravity due to the redistribution of mass in the ground which is instantaneous and therefore detectable in advance of the arrival of seismic waves. These gravity signals are typically only fractions of a μGal but can be measured using high accuracy superconducting gravimeters if sufficiently close to the triggering earthquake, yet far enough to generate a distinct difference in arrival time with the seismic waves (Montagner et al., 2016). For distant earthquakes, the static signal is unnoticeable during surveys using current field instruments as the resulting signal size falls below the practical resolution of the instrument, although these effects have been noticed for local earthquakes using repeated surveys (Tiwari and Mishra, 1997). For the purposes of the current work looking at distant earthquakes, static signals are therefore unlikely to be relevant.

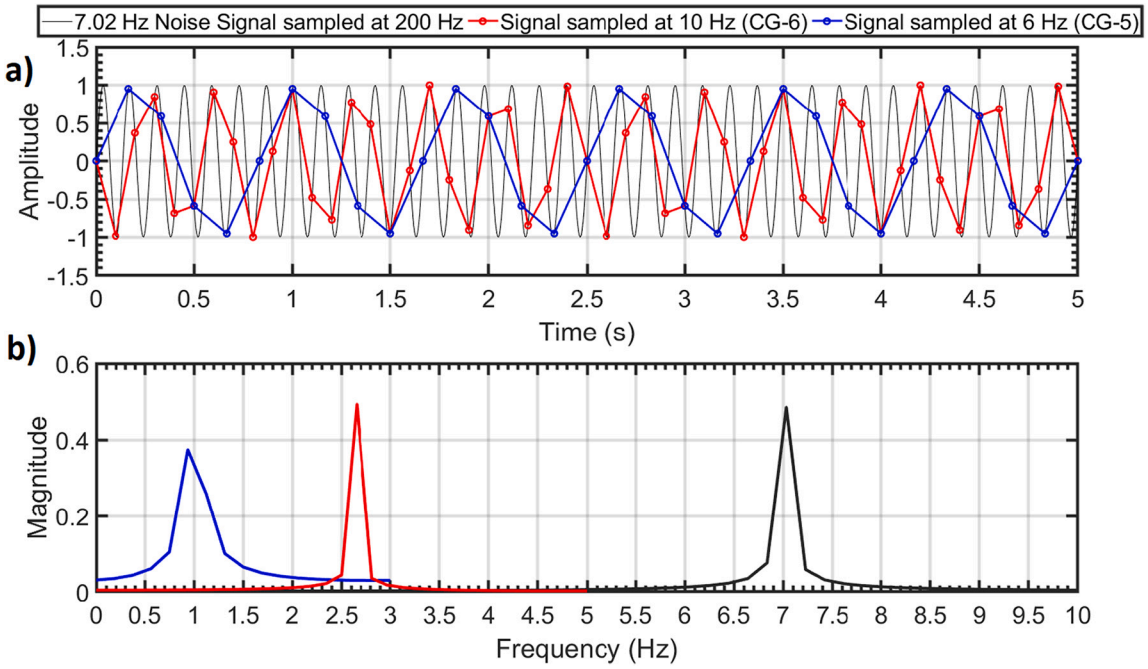

Fig. 1. The effects of aliasing on accelerations above the Nyquist frequency when recorded at the sampling rate of the instruments in this study (6 and 10 Hz) a) the simulated time domain data and b) the FFT of the simulated data. Notice how the 7.02 Hz noise appears as lower frequency noise.

2. Body seismic waves in the ground in the form of Primary (P)-waves and Secondary (*S*)-waves (Kulhánek, 1990). P-waves are compressional and travel faster than S-waves, although their amplitude is smaller and the frequency higher. In contrast S-waves, formed by the transverse movement of material produce slightly larger acceleration signals which travel slower. These waves are typically recorded using seismometer instruments which are specifically optimised for their frequencies (0.1–10 Hz) and designed not to saturate their response. In contrast, gravimeters are usually designed to filter this noise, either through sensor design or applied filters, and consequently, only specifically designed gravimeters can detect these waves easily (Niebauer et al., 2011).
3. Surface seismic waves which travel only through the crustal surface of the Earth in the form of Love (side to side motion) and Rayleigh (rolling motion) waves which shake the ground (Kulhánek, 1990). Whilst Love waves are unlikely to be detected as gravimeters typically measure the vertical gravity component, Rayleigh waves should affect the instrument. Indeed, gravimeters have been shown to be more sensitive to these waves than even conventional seismometers due to their low frequency response, with the instruments capable of detecting several additional Rayleigh wave arrivals compared to the seismometer (Niebauer et al., 2011). These Rayleigh waves produce the largest ground movements and therefore will have the biggest effect on the measurement accuracy in the period of time immediately following an earthquake.
4. Free earth oscillations in the aftermath of the waves from the main earthquake caused by the interference of P, vertically polarised S and Rayleigh waves (spherical oscillations) and Love and horizontally polarised S-waves (toroidal oscillations), which create a set of standing waves at discrete frequencies in the elastic earth. Since gravimeters only record vertical changes in acceleration due to their design, and toroidal waves do not induce radial deformation or change in volume (or density), only spherical modes can be observed (Arora et al., 2008; Ghobadi-Far et al., 2019). Although these oscillations are typically small in amplitude, excitement caused by an earthquake of vertical motion has been detected using static gravimeters (Arora et al., 2008), who observed waves with periods between 3.4 min ($_0S_{42}$ mode) and 54 min ($_0S_2$ mode), with the largest amplitudes being on oscillations with periods of around 20 min ($_0S_0$ mode). The oscillations can persist for several days after a strong earthquake, and the comparatively long periods in comparison to the instrument measurement window length makes their effects hard to recognise or account for using integration, but potentially causes variance between subsequent measurement cycles. Free earth oscillations are likely to have the most persistent effect on instrument accuracy during a gravity survey as they continue to propagate for a long time after the earthquake, introducing instability on gravity readings.

Together these effects result in a potentially increased error on the measurements in addition to the error due to the presence of usual microseismic and ground vibration noise. Variations in the recorded seismic signal of up to an hour in duration have been observed for localised seismic activity during continuous field recording using field gravimeters in active seismic zones (e.g. Bonvalot et al., 1998; Tikku et al., 2006). The authors typically found elevated seismic noise at higher frequencies than the microseismic background which increased the RMS error by up to 50–350 μGals for a 30 s cycle time. However, much less attention has been given to the effects of distant teleseismic events. Due to the unpredictable nature of earthquakes, few studies have been carried out to specifically quantify the effects of teleseismic activity, although several authors have recorded earthquakes during measurements for other purposes (e.g. Bonvalot et al., 1998; Debeglia and Dupont, 2002; Tikku et al., 2006; Greco et al., 2008) and justified the higher noise associated with them or removed the data from the study. Whilst almost all of the authors recorded an increase in the RMS error during the earthquake and the immediate aftermath, the duration of the increased microseismic noise and the effect on in the aftermath varied greatly between different earthquakes, with some studies recording no noticeable effect and others showing heightened RMS noise levels and an increased spread of gravity readings for several hours.

However, almost all of these studies focused on the final output RMS of the instrument, but did not identify the frequencies of the earthquake related noise from the raw data and how that interacts with the specific design of the different gravimeter sensors used, which had different measurement times and sampling rates, often had limited dynamic

range and are designed to filter out seismic sources of noise. Niebauer et al. (2011), using a specially designed gravimeter with a flat frequency response below 1 Hz, showed that P-, S- and Rayleigh waves were all visible on the measured data. However the particular design of the gravimeter used for the study is atypical and interaction in a field gravimeter may be different due to design considerations which may affect the way different amplitudes and frequencies of noise are manifested. Greco et al. (2008) showed that the period of oscillation relative to some large earthquake components corresponded closely with the main resonance frequency of the field gravimeters being used, effectively amplifying the noise in these frequency bands. It has also been shown that the effects of seismic noise on the accuracy of gravity readings are non-linear and strongly dependent on both the amplitude and frequencies of the noise (Greco et al., 2014), based on an analysis of localised seismic disturbances. However, their analysis was limited to the range of 3–25 Hz as it was directed at localised seismic activity, whereas distant teleseismic events tend to produce much lower frequency sources of noise which may affect the instruments differently. Furthermore, it is anticipated that the frequencies and amplitudes of the noise will change over time due to attenuation and interference between different wave types.

By analysing the raw data and deploying instruments used most commonly for field surveys, the effects of earthquake seismic signals can be more accurately characterised with a full frequency range of noise sources between the reciprocal of the measurement time and Nyquist frequency. Using this information, the effects of using different measurement strategies can be assessed to devise an optimum strategy by assessing the accuracy of using different cycle times. Additionally, as all of the above mentioned studies stopped 18–22 h after the earthquake, by continuing to collect data on subsequent days after the earthquakes, the evolution of the noise over time and the effects of long period noises sources such as free earth oscillations were also studied.

4. Earthquakes analysed and processing methods

Two different earthquakes in Chile in 2015 were recorded using Scintrex CG-5 gravimeters which were used to record continuous data in static locations in the UK; one from the basement of a building in the University of Birmingham on data that was collected to investigate long term instrument stability (36 h of data) and the other earthquake in 2015 was recorded from a commercial site in Thurrock, Kent, where one instrument was left recording continuously at the base station to collect information on background noise for research (data collected for 6–8 h a day for the day before, and 1, 2 and 5 days after the earthquake). An additional 2019 earthquake from Indonesia was captured during a gravity gradient survey using two Scintrex CG-6 instruments, also at the University of Birmingham. The earthquakes are summarised in Table 1. It should be noted that for all the measured earthquakes, the measurement site lies in the shadow zone (104–140° for P waves and above 104° for S-waves), meaning that any effect of these body waves is likely to be due to reflected waves as opposed to direct arrivals.

The CG-5 instruments were set to record raw data, which is sampled at 6 Hz from the internal analogue to digital converter and converted from digitiser counts to μGal using a scaling factor. In all cases the raw data was pre-processed before further analysis to remove known sources of noise (temperature of the spring, tilt of the instrument (Scintrex Ltd., 2006), celestial and ocean tidal effects using a tidal model (Hartmann and Wenzel, 1995) and linear drift of the instrument spring using a linear fit to the data. Further details on processing raw data from the Scintrex CG-5 to remove instrumental effects are provided by Sugihara (2004).

The data used for the gradiometer measurements in Section 6 of the paper consists of shorter readings (three minutes). Three Scintrex CG-6 instruments were used with the trident gravity gradient stand to create gravity gradiometers with separations of 0.5 m at two different heights and 1 m between the top and bottom sensors (Fig. 2). The sampling rate

Fig. 2. An example of the trident stand setup used to collect gradiometer data during the study. The three instruments were controlled using the data collection tablet interface and fired simultaneously.

Table 1
Summary of the earthquakes recorded for this study.

Earthquake number	Measurement location	Date	Earthquake location	Earthquake magnitude (Mw)	Distance from triggering epicentre	Duration of data recorded
1	University of Birmingham (52.453276°N, 1.927410°W)	20th June 2015 02:10:07 (UTC)	73 km WNW of Talcahuano, Chile[a] (36.360°S, 73.812°W, 11.0 km depth)	6.4	12,070 km (angular distance = 108.5092°)	36 h
2	Thurrock, Kent (51.479232°N 0.275479°E)	16th September 2015 22:54:33 (UTC)	48 km West of Illapel in Chile[b] (31.573°S, 71.674°W, 22.4 km depth)	8.3 (aftershocks 6–7)	11,590 km (angular distance = 104.1955°)	6–8 h per day (day before and 1,2 and 5 days after the earthquake)
3	University of Birmingham (52.453276°N, 1.927410°W)	2nd August 2019 12:03:27 (UTC)	106 km WSW of Tugu Hilir, Indonesia (7.282°S, 104.791°E, 49.0 km depth)[c]	6.9	11,780 km (angular distance = 105.9257°)	2 × 3 min readings (1 before and 1 during earthquake)

[a] https://earthquake.usgs.gov/earthquakes/eventpage/us10002ke8#executive.
[b] https://earthquake.usgs.gov/earthquakes/eventpage/us20003k7a#executive.
[c] https://earthquake.usgs.gov/earthquakes/eventpage/us60004zhq/executive.

of the CG-6 was 10 Hz. Data were processed by aligning the measurement times of the two sensors, applying tilt and temperature corrections and removing the drift from the instrument using linear fits between base station readings. Gravity gradients were calculated using Eq. (2) with the heights (h) in metres and gravity of the two sensors (g) in μGal to give a measured gravity signal in Eotvos (1 E = 10 μGal/m = 10^{-9} s^{-2}).

$$g_z = \left(\frac{(g_{bottom} - g_{top})}{(h_{top} - h_{bottom})}\right) * 10 \quad (2)$$

5. Surveying during and immediately after an earthquake

To identify the effects of an earthquake on the performance of field gravimeters, data were examined from an instrument measuring before, during and after a teleseismic event (earthquake 1) using the maximum measurement cycle time (256 s) allowed by the instrument. The arrival times of the different wave types have also been calculated. Fig. 3 shows approximately 2 h of data taken before, during and for several hours after earthquake 1. A Fast Fourier transform (FFT) of each of these intervals has also been provided to view the different frequency components.

In the time interval before the earthquake, the typical microseismic noise is visible as a broadband noise, with the highest peak at approximately 0.2 Hz. The nature of this noise is in line with the findings of other authors for microseisms from ocean wave sources (Ardhuin et al., 2011; Ardhuin and Herbers, 2013; Ardhuin et al., 2015) and the instrument is specifically designed to reject this noise through the filtering and integration process. Once the waves from the earthquake arrive, whilst the typical microseism noise is still apparent, all of the different waves associated with the earthquake create additional low frequency noise below 0.1 Hz on the data, with the largest peak being caused by the Rayleigh wave arrivals. Th can also be seen in the increased standard deviations on these readings with slight increases being recorded after the P and S-waves but a very large increase being recorded following the arrivals of the Rayleigh waves. In the aftermath of the earthquake and after the initial Rayleigh wave arrivals have decayed, the nature of the microseismic noise has been altered, with a 0.05 Hz acceleration signal still visible (as shown by the peak in the FFT data), in addition to the usual broadband microseismic noise peak at 0.2 Hz. This means that there is an additional source of low frequency noise which affects the measurements, even after the obvious effects visible in the time domain are no longer apparent.

5.1. Standard deviation (SD) and RMS as a measurement of error

Standard deviation or standard error is often used as an on-site quality control measure, with large values over a certain threshold being rejected and the point re-attempted. However, as the standard deviations for the arrivals of the early seismic waves fall well within the typical range of values which the instrument operator could expect in normal surveying conditions, especially on a day with a moderate breeze, it is unlikely that the SD will allow the earthquake to be detected at an early stage after the event. Furthermore, as the frequencies of these earthquake related noise sources are lower than typical microseismic vibrations, they may have a disproportionate effect on the measurement accuracy as less complete cycles can be measured within the measurement window.

5.2. Allan deviation as a measure of error

One problem with using the standard deviation or RMS error as a measure of measurement accuracy is that the method makes the assumption that the noise is white noise (i.e. that the spectral power is the same across all frequencies), whereas in reality the noise is defined by noise bands peaking at specific frequencies as shown in Fig. 3. The optimum time for a measurement cycle using a gravimeter is a balance between two competing effects on the instrument; the integration of the high frequency microseismic noise sources by using a long enough integration time to average their effects, and the effects of lower frequency noise sources with periods longer than the integration time which reduce the accuracy as the measurement time is increased (Boddice et al., 2018). Taking these different noise sources into account, a better estimate for the accuracy of the measurements and determination of the optimum measurement time for a reading can be assessed using the Allan deviation (Allan, 1966) which shows the resulting noise for different measurement cycle times. The Allan deviation is determined by calculating sets of mean values for a range of different window sizes (equivalent to different averaging time periods) and then calculating the variation in the means for each window size using the standard deviation to determine the stability of the readings for a range of different integration times. Plotting the variance against the time for each window length, an initial decrease can be observed as the high frequency noise sources are averaged followed by an increase as the effects of lower frequency noise sources dominate the improvement from further averaging can be observed.

The Allan deviation of the time intervals containing different waves before, during and after the earthquake is shown in Fig. 4 and the results for the use of different integration times shown in Table 2. The results show that all of the different waves created a short term increase in the measurement errors for short averaging windows which can be divided into three distinct groups, shown by the similarity of the different lines in Fig. 4: 1) readings without the presence of major earthquake waves from before and after the earthquake (blue and red lines), 2) readings with body (P- and S-) waves which had a small effect on the data (yellow and purple lines) and 3) much larger effects from the surface Love and Rayleigh waves (green and cyan lines).

The first of these groups, where no earthquake waves were present, are characterised by low short-term errors with the Allan deviation reaching a flat state within 20 s after which the improvements are minimal. The interval following the final Rayleigh waves shows very little difference to the data taken before the earthquake, with comparable size errors for all measurement cycle times of 20–90 s, although there is some variation at very short measurement time windows. This suggests that the additional low frequency noise shown in Fig. 3 during the time following the earthquake is not having a significant effect on the measurements when taken using commercially viable measurement cycle times. Larger differences are apparent for longer cycle times, suggesting that although this noise has little effect on the performance of the instrument in a commercial context, the noise may provide a limitation on the maximum accuracy which can be obtained if no limitations exist on integration time. The optimum performance capped at 1.48 μGals (90 s integration time) before the errors began to rise again, as opposed to measurements taken before the earthquake where the errors continued to fall for all measurement times up to the longest measurement cycle times possible by the recording instrument.

5.3. Effects of body waves on measurement precision

For the second of these groups containing the body P- and S-waves, the Allan deviation is characterised by having a higher short term measurement error and taking longer to reach the same accuracy as the periods without earthquake noise, although the measurement continues to improve with longer measurement times instead of reaching a steady state. For 30 s readings, these waves introduced a considerable increase in the measured errors, taking the measurement errors from 1.7 μGals to 2.34 and 3.04 μGals respectively (increases of 37.6% and 78.8% respectively compared to the errors seen before the earthquake). For longer readings of 60 s, the P-waves actually showed a slight improvement to the data and the decrease in accuracy for the S-waves was more modest (1.49 to 1.95 μGals; an increase of 30.8%). The longer 60 s

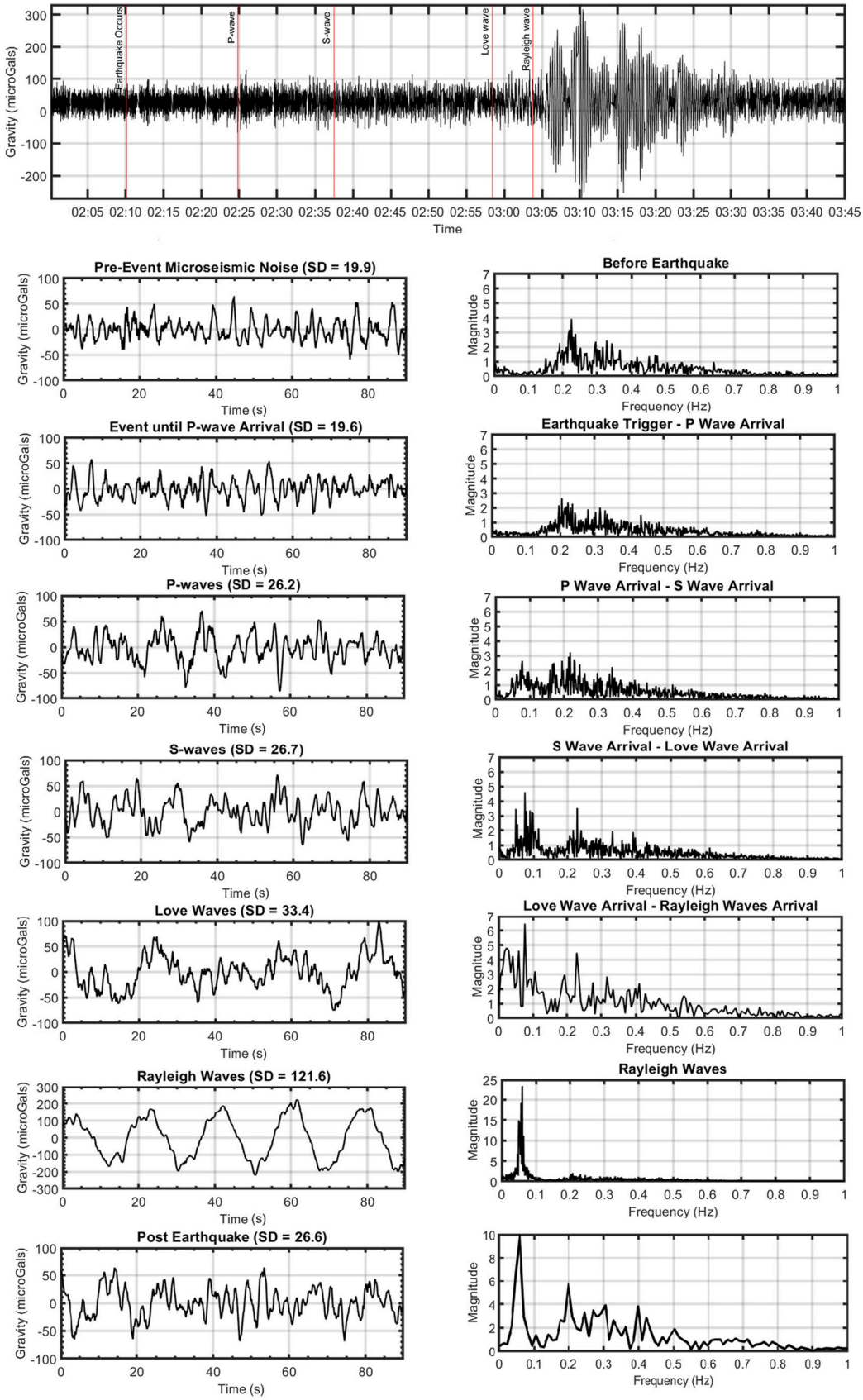

Fig. 3. The effect of different wave arrivals on frequencies of recorded noise on the gravimeter taken from data of earthquake 1 (Mw =6.4). The left hand column shows data from the different wave arrivals in the time domain and the right hand column shows the corresponding frequency domain data.

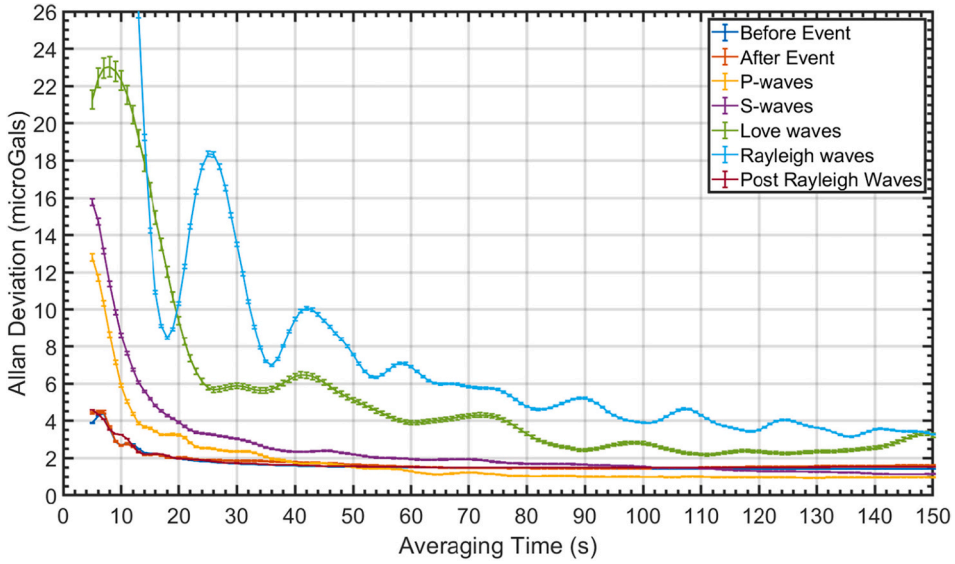

Fig. 4. Overlapping Allan deviation of earthquake 1 data taken at different stages of the earthquake.

Table 2
Table showing the Allan Deviation for different averaging times for different stages of earthquake 1. N.B. the length of time for the Love wave period was too short to obtain Allan deviations for longer than 150 s.

Measurement cycle time (s)	Allan deviation (errors) (μGal)						
	Before the earthquake	After the earthquake	P-waves	S-waves	Love waves	Rayleigh waves	Post Rayleigh waves
10	2.69	2.67	5.89	8.62	22.30	50.84	3.26
20	1.97	2.04	3.25	3.93	9.40	10.33	1.98
30	1.70	1.86	2.34	3.04	5.88	13.51	1.73
60	1.49	1.53	1.28	1.95	3.90	6.90	1.51
90	1.44	1.44	1.00	1.64	2.43	5.20	1.48
120	1.43	1.53	0.98	1.30	2.35	3.50	1.50
180	1.41	1.71	1.06	1.06	N/A	2.82	1.57
250	1.39	1.63	0.99	1.24	N/A	2.17	1.65

readings may therefore be more appropriate to use during periods where these waves are present, as they give a performance which is roughly equivalent to that which would be obtained when using 30 s readings, which is the minimum amount of time typically used on a survey and a much lower drop in performance.

5.4. Effects of surface waves on measurement precision

The final group of waves (surface waves) are characterised by having a vastly higher short term error, as well as much less stable values across the whole measurement range, never reaching the same accuracy levels as for the other periods, as well as not showing a consistent decrease in the measurement errors with increasing measurement times unlike with the other groups. This is likely to be the result of averaging incomplete cycles of low frequency high amplitude noise. The resulting performance is 2.5 times worse for the Love waves and 7 times worse for the Rayleigh waves when using 30 s cycles, with large errors still present for 60 s cycles (1.5 and 3.5 times worse for Love and Rayleigh waves respectively). It is therefore likely that it is impossible to measure accurately within these periods and that the resulting measurements will be unlikely to be useful, especially during the Rayleigh wave period.

5.5. Effects on measurement repeatability

To test the actual measured spread of the mean values (measurement repeatability) and to take into account other instrumental effects, in addition to the increased noise which would be expected, the different periods were isolated and 30 and 60 s moving mean window functions over the sections of the data with the presence of the different types of earthquake waves were determined. The probability density functions have been calculated using a normal kernel smoothing function in Matlab (ksdensity), and are shown in Fig. 5. The standard deviations of the moving means and the mean RMS error calculated for each period are shown in Table 3. F tests between the distribution of data collected before the earthquake and each of the periods using the Brown Forsyth test (chosen because it is more robust than other methods to datasets which violate the assumption of normality) were also conducted to test for significant differences in the measured variances during each period, as well as between the 30s and 60s results and p values are also shown in Table 3. These tests all rejected the null hypothesis (i.e. that there was no difference in the variances between the datasets) with a confidence of more than 99% suggesting that any visible differences between the distributions are statistically valid.

For both the periods before and the period after the earthquake, both 30s and 60s window lengths showed comparable performances, suggesting that in normal surveying conditions, there is little practical benefit on the measurements, although using longer measurement times results in lower measurement standard deviations and RMS errors. The differences in the standard deviations in these periods ranged from 0.2 to 0.35. Since it could be expected that 95% of the data would be within 1.96 standard deviations, and the resolution of the CG-5 instrument is limited to 1 μGal, the difference in the errors would likely to be unnoticeable (0.39–0.68 μGals) and not significantly impact performance. This is also shown by the similarity in the probability density functions for 30 s and 60 s moving windows in these periods.

The results show very small effects on the data for the initial body P-

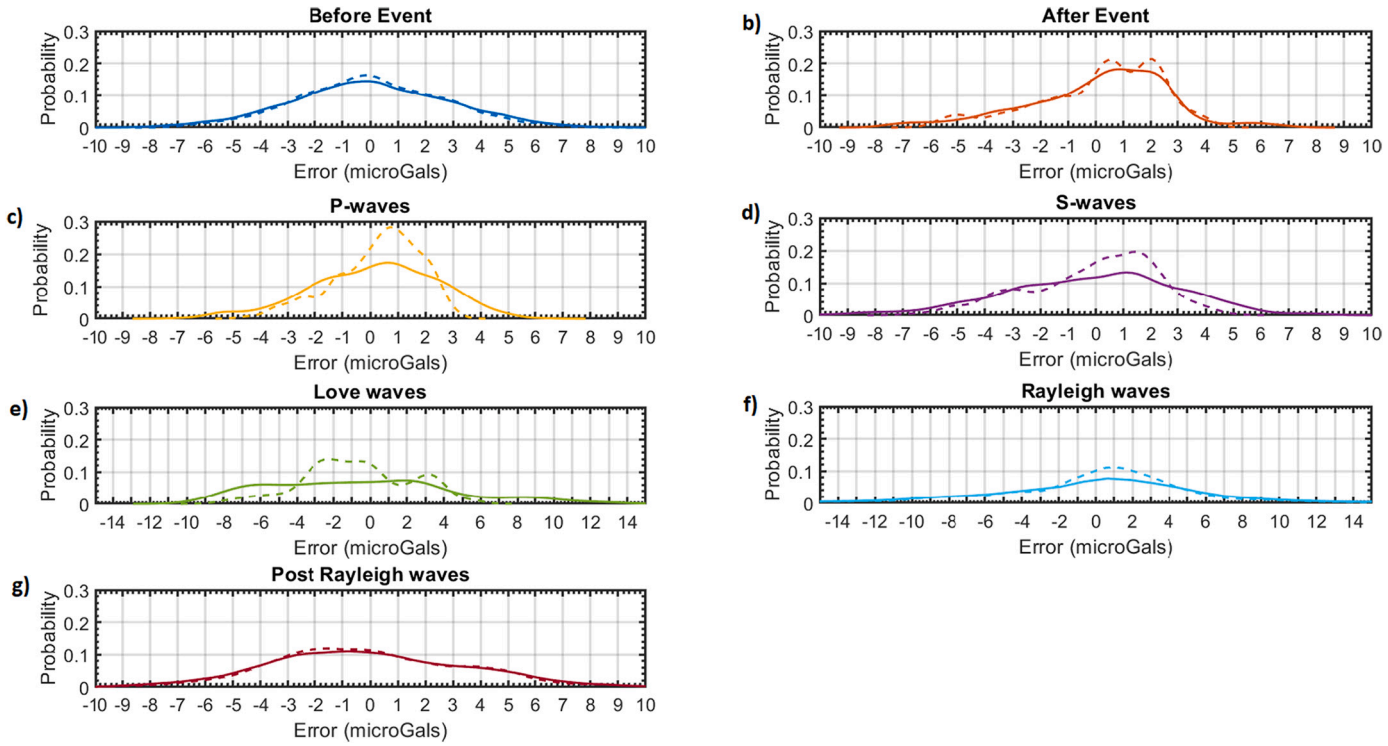

Fig. 5. Probability density functions for moving mean windows for the different periods of the earthquake (30s windows as solid lines and 60s windows as dashed lines) a) before the earthquake, b) after the earthquake, c) P-waves, d) S-waves, e) Love waves, f) Rayleigh waves and g) after the Rayleigh waves.

Table 3
Statistics from the moving means for 30 s and 60 s windows for data collected at different times during and after earthquake 1.

	SDs of means		Interquartile range		Mean RMS errors		P values for f-test with before earthquake		P values for f-test between 30 and 60 s windows
	30s	60s	30s	60s	30s	60s	30s	60s	–
Before the earthquake	2.90	2.64	3.89	3.58	1.69	1.20	–	–	0
Earthquake until P-wave arrival	2.59	2.24	3.23	3.09	1.56	1.11	7.62e-39	4.18e-55	1.27e-23
P-waves	2.30	1.57	3.23	2.10	1.98	1.41	1.38e-64	4.52e-255	2.91e-128
S-waves	3.15	2.30	4.29	3.18	1.90	1.35	1.56e-23	2.52e-40	9.72e-146
Love waves	5.05	2.91	7.26	4.20	2.26	1.63	0	2.47e-09	4.99e-99
Rayleigh waves	11.15	5.88	7.94	5.48	4.27	3.11	0	0	0
After the earthquake waves	3.82	3.62	5.06	4.77	1.51	1.07	0	0	0

and S-waves, with the P-waves even offering a slight improvement on the initial measurement conditions shown by the smaller spread of data during this period. Nevertheless, a more significant advantage from the longer cycle times is apparent by the difference in the probability density functions which has a higher number of readings falling near the centre of the distribution for the 60s windows. Furthermore, the standard deviation of the measurements during these periods remains comparable to the period before the earthquake and in the case of the P-waves, actually shows an improved performance. These results are in line with the results from the Allan deviations which also showed improved performance at longer time windows during the P-waves.

Both the Love and Rayleigh waves have had a more significant effect on the measured data which is shown by the large increases to the SD (30s: 5.05 for Love and 11.15 for Rayleigh from 2.90 before the earthquake 60s: 2.91 for Love and 5.88 for Rayleigh from 2.64) and the increased spread of data in Fig. 5e and f. During the Love waves, 60s readings showed only a slightly worse performance which was almost comparable to readings taken without earthquake waves (within 10%; increase in SD of only 0.27), whereas the errors for 30s showed an increase of more than 2 μGals in SD. Calculating the 95% error window indicates the noise would create an additional spread of 4 μGals which is unlikely to give stable enough readings for detecting small targets, as these errors are comparable in size to the targets of interest in a high accuracy civil engineering survey. During the Rayleigh waves, neither measurement gives acceptable errors with the SD which are 284% higher for 30s cycles and 122% for 60s readings. Although the 60s shows an improvement visible in the probability density function in Fig. 5f, the 95% error bracket gives errors of 21.8 and 11.5 μGals for the 30s and 60s cycles respectively, which are too high to obtain good stable results for individual measurements as the number of repeat readings to obtain an acceptable level of accuracy would be commercially unviable.

The period after the earthquake showed a decrease in accuracy compared to the period before the earthquake (30s: 3.82 compared to 2.90 before and 60s: 3.62 compared to 2.64). This is likely to be the result of the continuing low frequency noise which was identified in the data in Fig. 3. Taking into account that 95% of the readings should fall within 1.96 SD, the difference in values of 0.92 for 30s readings and 0.98 for 60s readings would result in an additional 1.8–1.9 μGals in error for a single cycle during this period. The higher errors on individual cycles may result in an increased number of repeated readings during this period or less accurate survey outcomes. Interestingly, the increased errors are not identified by either the RMS error which recorded lower values than the period before the earthquake or the Allan deviation which showed no large differences at these averaging times between this

period after and the period prior to the earthquake. The persistence of these effects following earthquake 2 is examined in Section 6.

6. Surveying during the days following an earthquake

As shown in Section 5, the effects of earthquakes persist beyond the end of the easily identifiable waves manifesting in the data and persist over a longer time period of several days. The majority of existing studies which contain earthquakes in the data are limited to a few hours after the earthquake, but it is of interest to investigate how the noise evolves and the resulting noise effects over a longer time frame. A static Scintrex CG-5 gravimeter had been deployed during the previous day of a commercial microgravity survey as part of an investigation into background noise on the day before earthquake 2 in 2015. After the earthquake was reported as taking place overnight, the gravimeter was again deployed over the following survey days (17th 18th and 21st September 2015) to collect comparable datasets over a number of days after the earthquake. Data were collected using 60s cycle times, comparable to those used in a typical survey. A wind shield was used throughout the survey to protect the instrument from these additional sources of vibration.

FFTs of all of the datasets are displayed in Fig. 6, showing the evolution of the noise spectra over the measurement period. The day after the earthquake, a small peak can be observed at a frequency of 0.05 Hz (similar to that identified after earthquake 1), before disappearing on subsequent days. This is likely to be the result of the low frequency seismic waves, and the smaller amplitude is a result of a greater amount of time elapsing since the triggering earthquake which has allowed the energy in the waves to decay. These waves seem to rapidly decay and are no longer visible 2 days after the earthquake. In normal surveying conditions, the peak of the seismic noise usually falls in the range of 0.2–0.3 Hz as can be seen on the day before the earthquake. A more persistent effect can be seen in the main peak of the microseism noise, which is larger in the two days after the earthquake when the energy in the microseism has shifted to higher frequencies, peaking between 0.3 and 0.4 Hz before returning to normal by the fifth day after the earthquake. The reasons for this shift are unknown although it may be related to a localised seismic source of noise which may be unrelated to the highlighted triggering earthquake, as the frequencies seem to be too high to be related to normal seismic modes associated with an earthquake.

To understand how these changes to the microseismic noise have affected the accuracy of the gravimeter, the overlapping Allan deviation has been calculated for each of the survey days (Fig. 7) with the results for a selection of integration times shown in Table 4. The optimum performance for the gravimeter is unsurprisingly on the day before the earthquake, with a significant degradation in performance the day after the earthquake. Over integration times consistent with those usually used for commercial microgravity surveys (i.e. 60s or less), only the day immediately following the earthquake (17th September 2015) showed a large drop in instrument performance. Subsequent days recorded similar errors to the day before the earthquake suggesting that the instrument performance was only majorly affected on the first day after the earthquake. This is also apparent from the similar shapes of the curves in Fig. 7. Since the 18th of March also showed a shift in the microseism peak to higher frequencies but showed no obvious degradation of performance, it is likely that the poor performance of the instrument on the day following the earthquake is due to the lower frequency noise (0.05 Hz) which was only apparent on that day. These accelerations are consistent with the noise caused by earthquake 1 in the previous section, and the frequency correlates well with those caused by Rayleigh waves as seen by other authors (e.g. Niebauer et al., 2011). The higher frequency microseism noise seen on the days following the earthquake is more easily accounted for by the integration process due to the larger

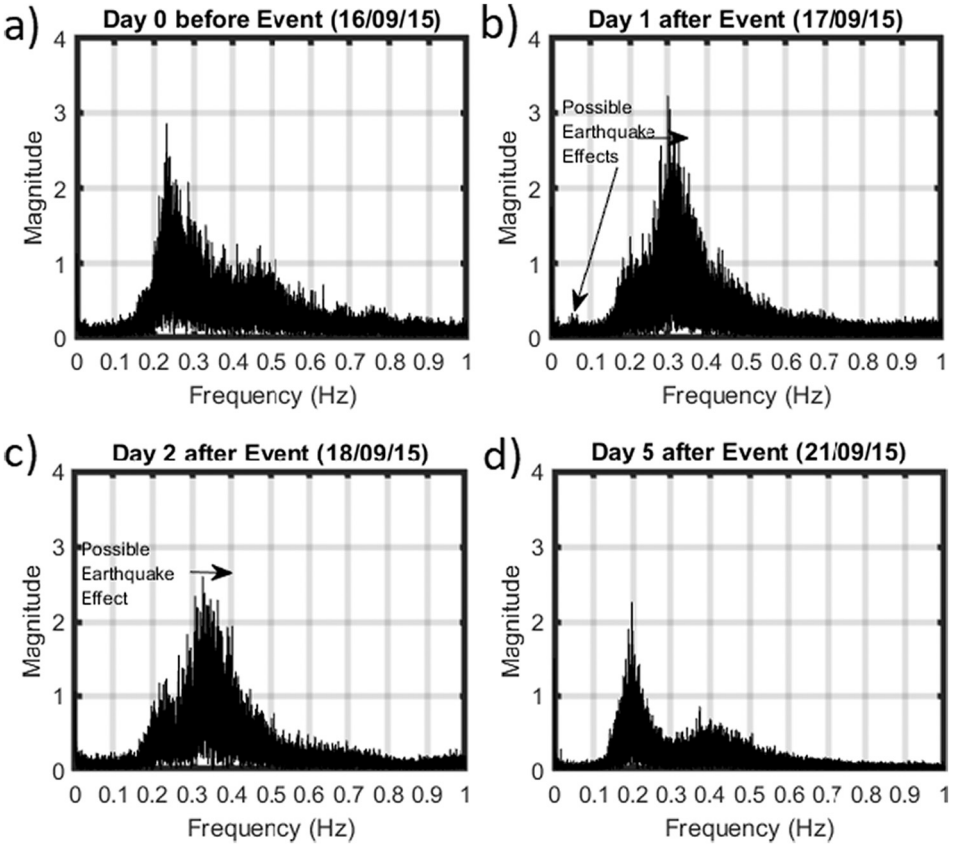

Fig. 6. FFT of the data from the 2015 earthquake 2 survey a) immediately before and b), c) and d) for several days after the earthquake.

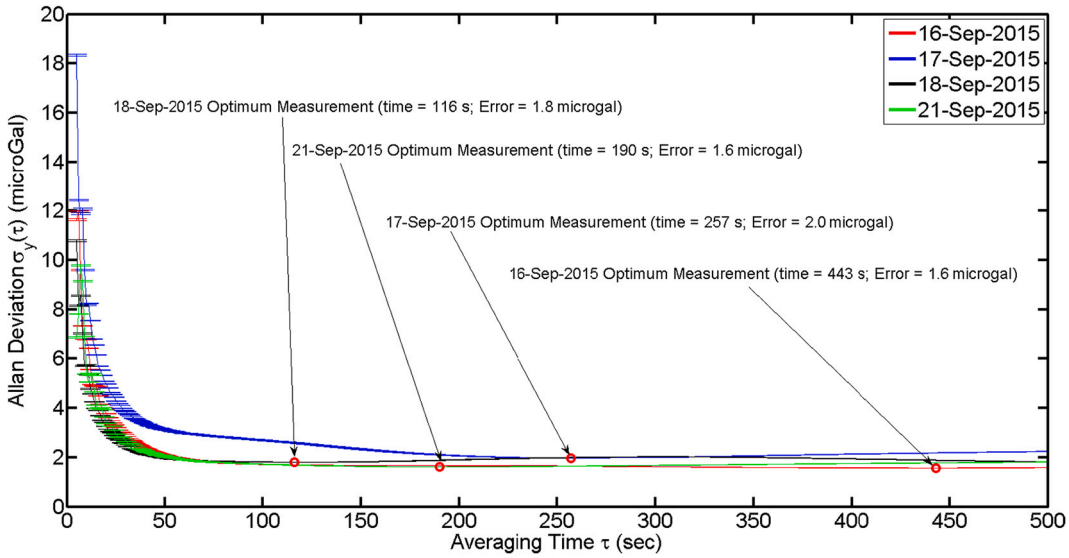

Fig. 7. Overlapping Allan deviation of earthquake 2 data taken before, during and after the Earthquake with the optimum measurement cycle times indicated for each day.

Table 4
Table showing the Allan deviation for different averaging times for earthquake 2 data.

Measurement cycle time (s)	Allan deviation errors (μGal)			
	16th September 2015	17th September 2015	18th September 2015	21st September 2015
10	6.78	8.19	5.74	5.37
20	3.79	4.97	3.27	3.38
30	2.85	3.85	2.50	2.66
60	1.92	2.98	1.90	1.91
90	1.75	2.76	1.83	1.75
120	1.69	2.56	1.81	1.68
180	1.67	2.15	1.87	1.61
250	1.65	1.98	1.99	1.64

number of complete cycles to be averaged and hence has a more minimal effect.

The overall effect on the accuracy has been calculated using the probability density distributions (Fig. 8) as with earthquake 1. The statistics relating to these distributions are also displayed in Table 5. The results showed a similar distribution for the day before and the days 2 and 5 after the earthquake with only slight differences in the instrument performance. Using Brown-Forsyth f-tests to assess the equality of the variances between day 0 (before the earthquake) and each of the other days accepted the null hypothesis at the 95% confidence level (i.e. that the distributions were the same) for day 3. The other two days (day 1 and day 5) rejected the null hypothesis, showing a significant difference between these days. As shown by the probability density distributions in Fig. 8 and the difference in the standard deviations in Table 5, the difference between the period before the earthquake and for the 5th day after the earthquake is small with less than 1 μGal in SD (around 21%). This variation is similar to those which would be seen due to natural variations in the microseismic noise between survey days due to changes in environmental conditions, and between instruments due to differences in the sensor springs (see for example difference between instruments in Boddice et al. (2018) or Jousset et al. (1995)). It is therefore unlikely that these effects would be noticeable, especially if multiple instruments are being used on the survey.

In contrast, the day immediately following the earthquake shows a marked decrease in sensor accuracy, with a broader distribution of values shown by the SD which is 39% higher (4.49 to 3.23) than the day prior to the earthquake. By using the 95% confidence intervals (calculated by multiplying these SDs by 1.96) as a benchmark for the largest errors to be expected on a measurement, the expected maximum errors are 2.47 μGals higher for the day immediately after the earthquake in comparison to the day preceding the event for the 60s cycle times used. This increase is also reflected in the higher RMS values seen in Table 5. These larger errors may be significant for high accuracy microgravity projects and result in either lower accuracy measurements, or the need to take an increased number of repeats to achieve stable values which makes surveying slower.

7. Long period variations

In addition to the high frequency noise which affects the accuracy of individual measurement cycles, longer period variations can be induced by the free earth oscillations caused by the earthquake. Typically, these oscillations will manifest as a longer period noise between measurement cycles as the periods typically range from 3 to 50 min. These frequencies are difficult to separate from existing low frequency sources of noise, especially the instrument noise, which tends to dominate at periods longer than the maximum window length (256 s) (Boddice et al., 2018). Another concern is the effects of drift on the instrument as the internal quartz spring relaxes. As the rate at which the spring relaxes is reasonably constant, drift on a survey over a long timeframe can be considered linear and is removed in the instrument software by applying a linear trend removal function (Scintrex Ltd., 2006). However, this correction only approximates the drift in survey conditions, and daily variation in the drift rate and additional non-linearities, such as vibration induced short term drifts, are removed using repeated base station readings (Seigel, 1995) and either fitting linear trends between each point (Tuckwell et al., 2008), or using all of the points to fit a higher order polynomial (Gabalda et al., 2003). These two methods are compared whilst looking at the time gaps between base station readings to see if there are any significant changes to survey practice needed in the aftermath of a large earthquake to account for very low frequency effects from free earth oscillations.

The dataset taken on site during earthquake 2 has been desampled to use only small selections of readings as base station readings. Different intervals of time have been taken between base station readings used to remove the drift from the whole datasets taken on each survey day, and the resulting errors have been calculated using the standard deviation of the corrected measurement cycles for both linear fitting and polynomial

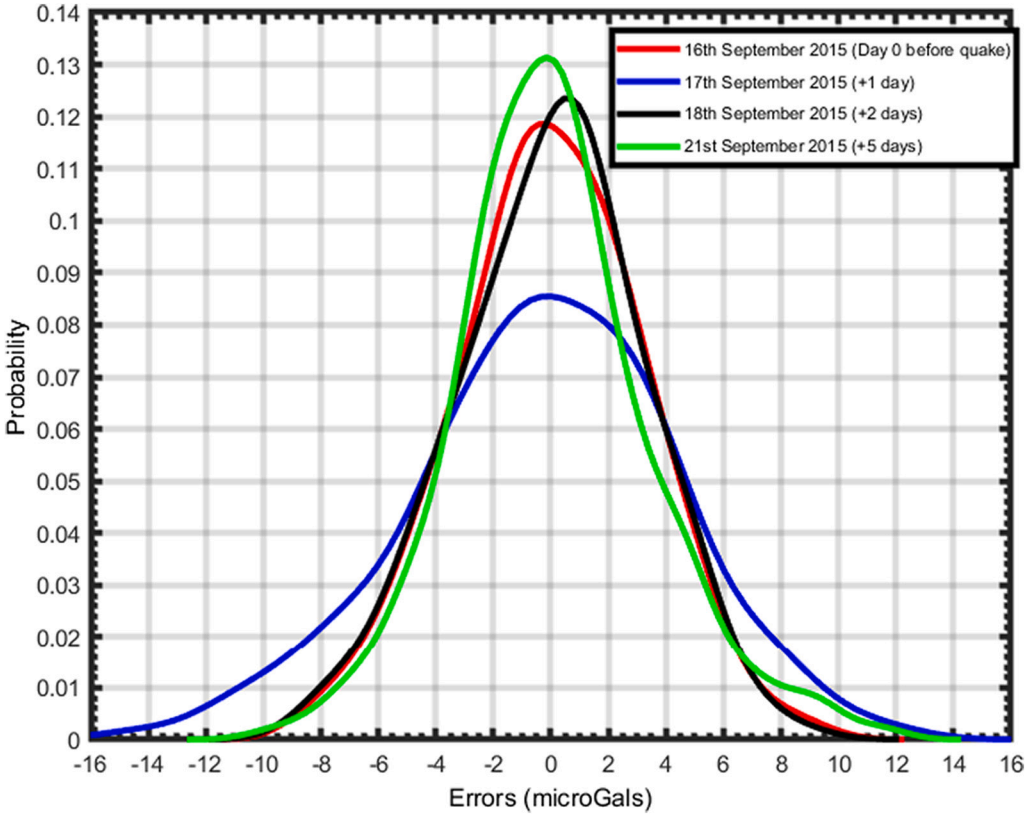

Fig. 8. The probability density functions for the measurement cycle means on different days before and after earthquake 2 in 2015.

Table 5
SD, interquartile range, maximum and minimum values and overall range for the measurements taken on each day before and after the earthquake.

	SD of means	IQR	Min	Max	Range	Mean RMS	*p*-Value of f-test
16th September	3.23	4.33	−8.69	9.08	17.77	3.31	–
17th September	4.49	5.49	−15.01	10.05	25.06	4.44	5.52e-10
18th September	3.23	3.99	−8.97	9.60	18.57	3.31	0.96
21st September	3.91	4.66	−12.10	12.87	24.97	2.73	2.60e-4

fitting methods. The results are shown in Table 6. Using the linear fit between base station readings approach produces superior results to the polynomial fit approach in nearly all cases, even when 3rd order polynomial fits are used. To ensure the highest accuracy, it is recommended that the linear fit approach is used for all surveys, regardless of whether a large teleseismic event has occurred.

The probability density distributions for readings corrected using the linear fit approach on each of the survey days for a range of different base station measurement time spacings are shown in Fig. 9. For the two days with no earthquake effects visible in the previous analysis (day 0 and day 5), 2 h separations between base station readings do not give notably worse results than shorter separations. For example, the differences between 2 and 1 h measurement gaps on these days are just 0.15 and 0.17 μGals respectively which is unlikely to have a significant effect on instrument performance. Furthermore, no mean shifts are visible on either of these days, other than when only three base stations at the start, middle and end of the survey day are used. This shows that the drift removal is sound, even for comparatively long gaps between base station readings during normal surveying conditions. Notable improvements in instrument performance are only visible when the interval between base station readings is brought down to 15 min. However, these improvements are generally small (SD reduction is less than 1 μGal for all but the second day after the earthquake), whilst the survey time will be drastically increased as 4 times as many base station readings will be needed, making the strategy unlikely to be commercially viable for most applications.

In the days immediately following the earthquake, using longer gaps between measurement stations leads to much larger errors. On the first day after the earthquake for example, using a 2 h measurement separation, causes a 2 μGal shift in the mean value of the data in comparison to the use of base station measurement frequencies of 1 h or less. On the second day after the earthquake, the shift in the data is even worse, with shifts of 5 μGals in the mean value for all base station gaps greater than 1 h. Furthermore, the spread of the data is also much greater with the difference in SD as high as 2.5 μGals on day 2 after the earthquake. For base stations taken with gaps of at least an hour, the performance of the drift correction in the days after the triggering earthquake is only marginally worse than during the day before the event, with no notable mean shifts and increases in the data spread in the order of 1 μGal. These results show that the drift correction is likely to leave large errors in the data if base stations are not taken with sufficient frequency after a large earthquake.

One interesting finding is that the worst drift corrections occurred on the second day after the earthquake, rather than the day immediately following the earthquake. One possible explanation for this is the attenuation and consequent reduction of the effective frequency of the seismic waves which was hard to detect using the previous FFT analysis due to the comparatively long period of noise. The consequences of this

Table 6
Standard deviations of the measurements over different base station time gaps using different linear and polynomial fitting methods.

		Start, middle and end of day	2 h	1.5 h	1 h	30 min	15 min
Linear fits between stations	Day 0 (16/09/15)	3.40	3.38	3.04	3.23	2.96	2.43
	Day 1 (17/09/15)	4.71	4.68	4.90	3.91	3.98	3.50
	Day 2 (18/09/15)	4.49	5.92	4.59	4.22	3.16	2.68
	Day 5 (21/09/15)	3.86	3.59	3.36	3.42	3.11	2.51
Polynomial -1st order fit	Day 0 (16/09/15)	3.40	3.11	3.20	3.13	3.05	3.11
	Day 1 (17/09/15)	4.80	4.78	4.81	4.78	4.81	4.78
	Day 2 (18/09/15)	8.98	9.42	8.77	8.65	8.49	8.46
	Day 5 (21/09/15)	3.99	3.95	4.03	3.99	3.97	3.98
Polynomial -2nd order fit	Day 0 (16/09/15)	3.40	3.08	3.11	3.01	2.96	3.01
	Day 1 (17/09/15)	4.62	4.60	4.72	4.58	4.64	4.63
	Day 2 (18/09/15)	4.63	4.68	4.31	4.38	3.58	3.47
	Day 5 (21/09/15)	3.65	3.73	3.53	3.51	3.44	3.47
Polynomial -3rd order fit	Day 0 (16/09/15)	N/A[a]	3.40	3.12	3.28	3.06	3.02
	Day 1 (17/09/15)	N/A[a]	4.61	4.89	4.55	4.64	4.62
	Day 2 (18/09/15)	N/A[a]	N/A[a]	3.66	4.07	3.63	3.41
	Day 5 (21/09/15)	N/A[a]	3.40	3.68	3.65	3.35	3.37

[a] Insufficient data available.

low frequency noise are a larger spread of values when using time gaps of more than 1 h between base station readings, even in comparison to the previous day (SD 4.26 compared to 3.91 for day 1 after the earthquake). In light of these findings and the previous observations, whilst it is usually sufficient to take base stations every two hours, it is recommended that for all surveying after a large teleseismic event base stations are repeated at least hourly, and a linear rather than a polynomial fitting method between base stations is used fit to achieve a satisfactory level of accuracy. Based on the available data, the duration of this survey strategy of hourly base station measurements is hard to precisely determine, but is for between 2 and 5 days after the earthquake.

8. Effect of earthquakes on gradient or difference measurements

It has been identified that for teleseismic events, unlike for more localised seismic activity, there is a strong correlation between gravity instruments, even those separated by large distances (Greco et al., 2008). With this in mind, it may also be possible to reduce the impact of earthquake noise by using an additional instrument to record and subtract the noise by using a gradient measurement or by recording the noise on a separate static instrument (variometer configuration). To investigate this seismic noise cancellation technique, data taken during a gravity gradient survey before and during earthquake 3 was used and is shown in Fig. 10. However, similar results could be expected for a variometer configuration for two adjacent instruments if the signal generated from accelerations with frequencies above the DC gravity value was used to correct the measuring instrument. Once again, it is apparent from comparison of the data taken before and during the earthquake that there is a significant shift in the nature of the microseismic noise, specifically manifesting as a low frequency noise with a frequency between 0 and 0.1 Hz. Whilst this frequency is a similar range as previously seen, the noise appears to cover a broader range of frequencies in comparison to the equivalent noise seen during earthquake 1, further highlighting differences in instrument responses to different earthquakes as previously discussed.

The correlation coefficient between different pairs of these sensors was calculated and is shown in Table 7. Crucially, the noise shows good correlation between the different sensors, with the correlations even improving for the measurements taken during the earthquake, allowing subtractions of the noise signal to be made in order to improve the accuracy of the readings. To test the ability of the gradient configuration to suppress the effects of noise and improve instrument performance, the Allan deviations both for the individual gravity sensors and the gravity gradients taken between different sensors were calculated and are displayed in Fig. 11 and the errors for some key values shown in Table 8.

Analysis of the gravity data from all three sensors shows a similar performance to each other for both measurements taken before (Fig. 11a) and during earthquake 3 (Fig. 11b). However, the performance of the instrument is greatly affected during the earthquake, with the best recorded instrument performance being almost 3 times worse than during the measurements preceding the earthquake. For commercially viable integration times (< 60s) the performance is 4–7 times worse, rendering a useful survey result impossible. This is similar to the results seen in the previous dataset during earthquake 1.

In contrast, the Allan deviations for the gradient readings for the same integration times show a very similar performance both before and during the earthquake, showing the benefits of noise cancelling. At longer integration times above 100 s, a more significant effect of the earthquake noise can be noticed on the gradient readings which is reflected in the difference between the data taken before (Fig. 11c) and during the earthquake (Fig. 11d). Whilst data taken during the period before the earthquake shows a continuous improvement with increasing measurement times throughout the range of integration times, the period during the earthquake shows a decrease in accuracy after the optimum averaging time has been reached, suggesting a low frequency noise source on the measurements. As all the instruments are measuring simultaneously, the real world vibrational effects on the sensors are near identical and these different responses are the result of their interaction with the sensor technology, probably due to slight differences in the mechanical properties of the instrument quartz sensor springs. A new generation of gravity gradient instruments which are based on quantum technology (e.g. Freier et al., 2016; Hinton et al., 2017) rely on using free falling atoms as test masses which are expected to overcome these effects. The use of a gradient or difference measurement shows great potential for taking measurements with the same accuracy for periods with and without teleseismic activity for the range of integration times typically used.

9. Discussion

The results show that the seismic waves associated with earthquakes have a considerable impact on data quality for several days after the earthquake. In an ideal scenario, it is advantageous to spot the

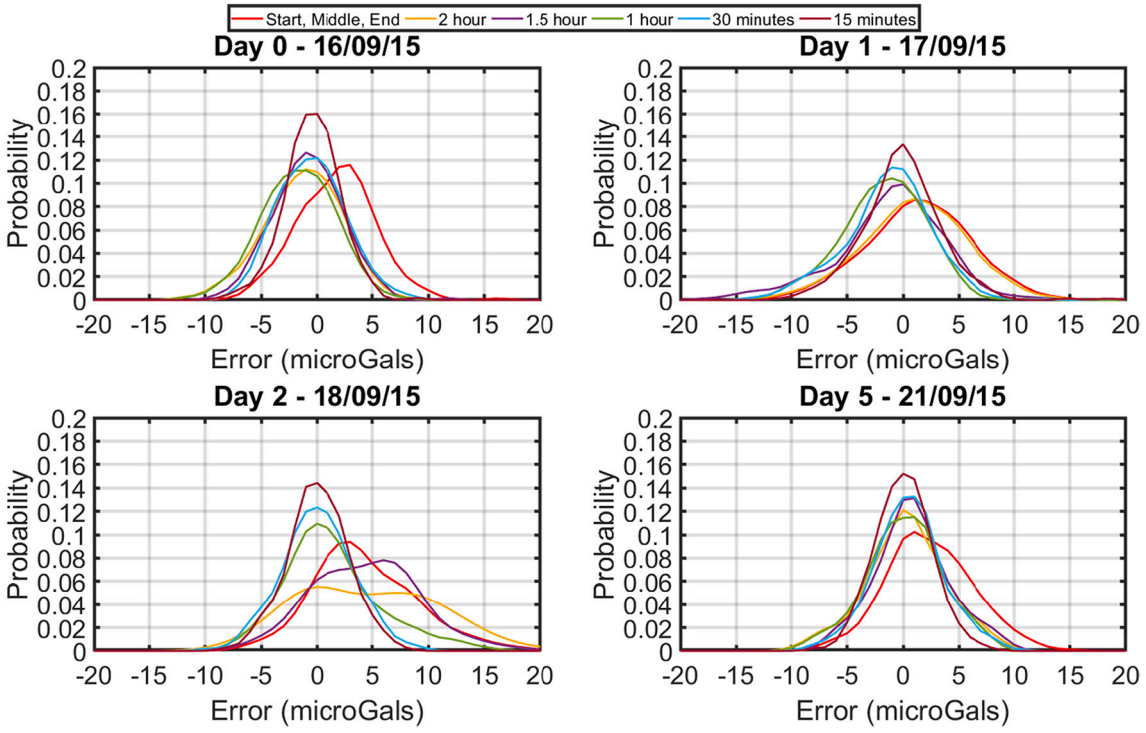

Fig. 9. The probability density distribution of the measurement cycle means on different days before and after earthquake 2 in 2015 when using linear fits between base stations to remove the drift.

earthquake as early as possible as the effect on accuracy is small enough to continue getting reasonable results before the surface waves (i.e. Love and Rayleigh waves) arrive, especially if longer readings (60s instead of 30s) are used during the body waves as shown here. Early detection allows the instrument operator to take the instrument immediately back to the base station to compensate for instrument drift, allowing the readings taken since the previous base station measurements to still be usable in the final dataset. In practice, this may be difficult for all but the most experienced operators as the other waves are difficult to spot in the time domain as their magnitude is comparatively small as shown in Fig. 3. The main diagnostic feature of an earthquake which any surveyor will recognise is a sudden rise in the instrument reading cycle standard deviation, although this is probably only noticeable once the Rayleigh waves arrive due to the small changes in SD for the other wave types. In this scenario, the presence of an earthquake can be confirmed by the operator by checking the data graphically for recent readings using the data recall function on the instrument to identify a low frequency noise with a period of approximately 20 s. If such noise is present, the recommended best course of action is to check an online resource for ongoing earthquakes. If an earthquake is found, increase the measurement cycle time to 60s (if using 30s readings) and move the instrument back to the base station to take continuous readings until the standard deviation returns to a lower level and the low frequency effects are no longer visible, which may take 1–2 h. Similar to the other waves, noise in the period after the initial Rayleigh waves, such as that shown in Fig. 3, will also be hard to identify by visual inspection in the time domain using the instruments display due to their small amplitude and may only be recognisable during post processing. However, for normal commercial integration times, it was not seen to have a significant effect on the accuracy of individual measurement cycles.

Future instrument software for the collection of data may seek to identify the noise from earthquakes using an FFT calculated at the end of the reading set on the measured data to search for noise peaks below 0.1 Hz, either through visual inspection or an automated peak detection algorithm. This would have the advantage of detecting the earthquake earlier (i.e. when the P-waves arrive) which allows the user to return to the base station and record the drift before the readings become impossible, resulting in more useable and reliable collected data. However, calculation of the FFT and the subsequent peak detection is challenging due to the comparatively short reading lengths which gives poor resolution in the FFT at the lower frequency range, in addition to the instrument's low sample rate and high noise floor at low frequencies on the analogue-to-digital converter. Preliminary testing on the data here showed that a period of at least 90s is required for a FFT to show a clearly identifiable spike below 0.1 Hz, which may be possible by performing the process across the data from all of the repeat readings taken on each measurement point (usually 3 x 30s cycles). An example of applying an automatic peak detection technique using a simple magnitude threshold (magnitude $>$0.07 shown as red crosses) on noise below 0.15 Hz is shown in Fig. 12, which demonstrates the potential of the method to identify earthquake waves at an early stage to allow the survey to be paused and the methodology to be adapted in the manner suggested. The method identifies the seismic waves from the earthquake almost immediately after the first P-waves hit, making it faster than could be expected using the time domain readings.

The main limitation of the current work is the limited amount of data availability, due to the difficulty in capturing earthquake events using a controlled reference instrument over a long period of time, especially as the raw data option is often not enabled for most typical commercial surveys. This has limited the sample to three earthquakes in the current study. As the epicentre of each earthquake is a different distance from the measurement location and also generated at a different triggering depth and magnitude, it is difficult to assess how these earthquake specific factors affect the resulting ground accelerations and performance of the gravimeter. To address this, a much larger dataset of earthquakes at a wide range of distances, depths and magnitudes would need to be collected using continuously monitoring instruments, which is challenging due to unpredictability of earthquakes and the resulting amount of time such an endeavour would require to capture a large enough dataset to systematically study these effects.

Another limitation of the work is that the data was collected using only the most popular commercial gravimeter currently in use. Other

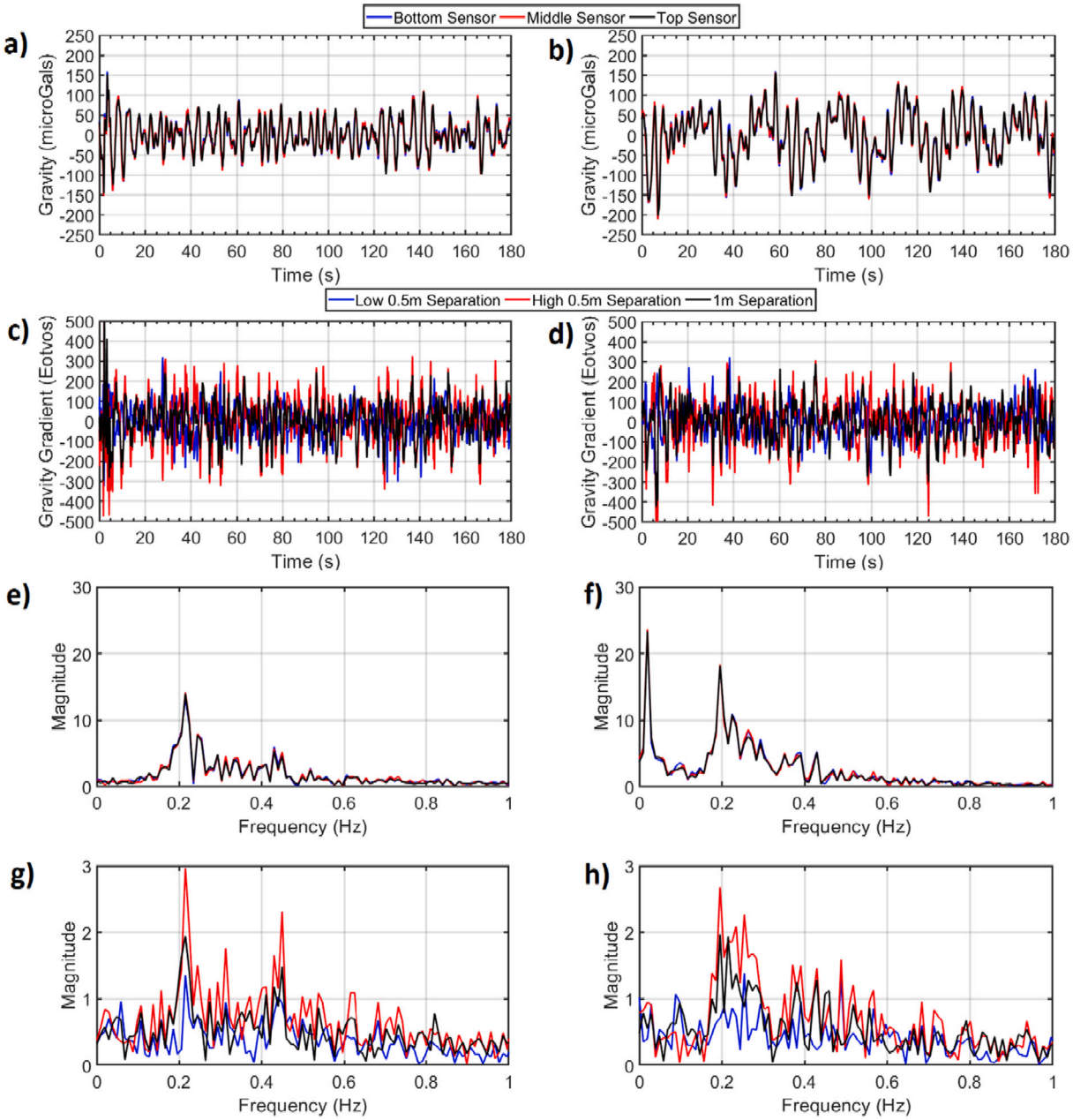

Fig. 10. Gravity data taken from a) before and b) during earthquake 3 in 2019 using three Scintrex CG-6 instruments and the trident stand, Gravity gradient data taken before (c) and during earthquake 3 in 2019 and the FFTs for e) gravity before and f) during earthquake 3 and gravity gradient before (g) and after (h) the earthquake.

Table 7
Pairwise correlations between data from different instrument pairs before and during the event.

Instrument pair	Before earthquake	During earthquake
Bottom and middle	0.986	0.990
Bottom and top	0.966	0.989
Middle and top	0.969	0.984

field gravimeters may have springs with different resonance frequencies which would affect their response to earthquake induced accelerations. In addition, some older gravimeters such as the Lacoste Romberg G and D meters rely on zero length springs (LaCoste & Romberg, 2004) which require manual balancing to collect readings and may experience different challenges, especially if not using electrostatic nulling. Collecting data with a wider range of instruments simultaneously would allow comparisons to be made and assess the wider impact of earthquake induced accelerations on different instrument types.

10. Conclusions

Acquiring suitable data to study the effects of earthquakes is difficult due to difficulty in predicting their occurrence. This paper analysed unique datasets of gravimeter measurements taken before, after and during earthquake events. It has been shown that teleseismic earthquakes change the nature of the usual microseismic noise, affecting the accuracy of gravity measurements taken using field gravimeter instruments. Whilst these changes have previously been noted both anecdotally and in individual study data, studies have not usually looked at long term effects in the days after the earthquake or the frequency

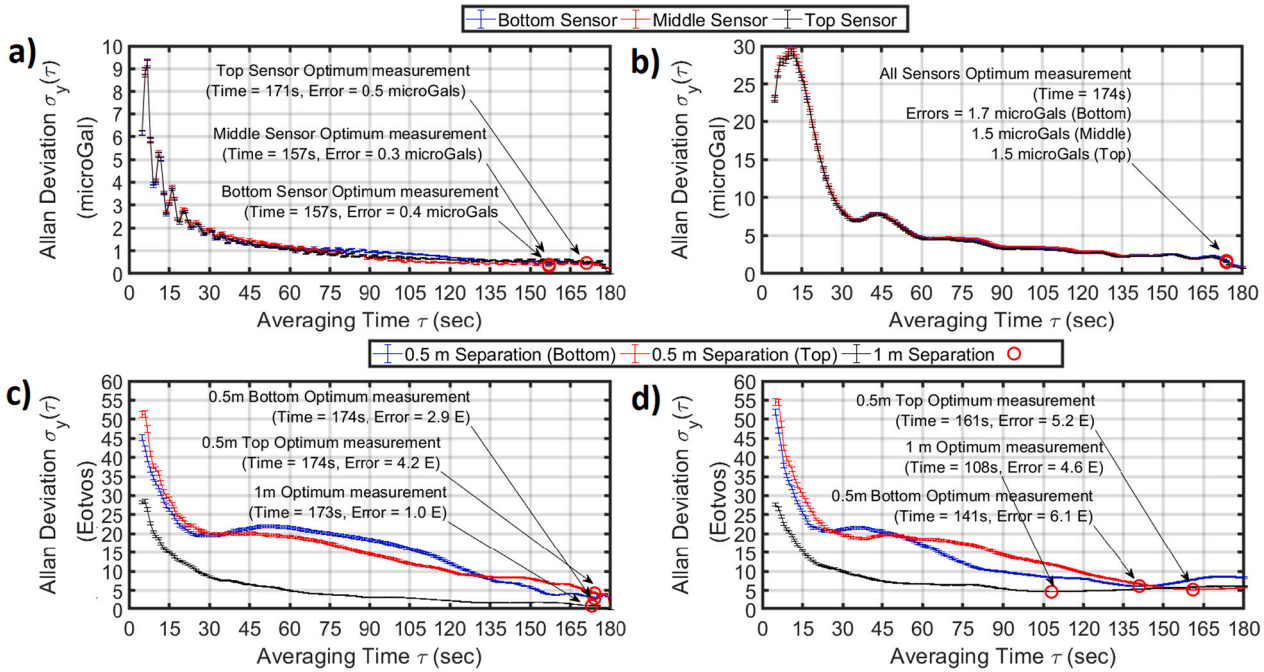

Fig. 11. Allan deviations of gravimeters and gradiometers for measurements taken before and during earthquake 3. a) gravimeters before the earthquake, b) gravimeters during the earthquake, c) gradient measurements before the earthquake and d) gradient measurements after earthquake 3.

Table 8
Table showing the Allan deviation for different averaging times for gravimeters and gradiometers before and during earthquake 3.

Measurement cycle time (s)	Allan deviation errors			
	Gravimeter before the earthquake (μGal)	Gravimeter during the earthquake (μGal)	Gradiometer before the earthquake (Eotvos)	Gradiometer during the earthquake (Eotvos)
10	4.03	29.2	19.33	19.92
20	2.68	18.22	12.56	11.17
30	1.82	8.30	8.52	8.91
60	1.09	4.69	4.93	6.04
90	0.70	3.60	3.21	4.57
120	0.59	2.93	2.28	3.92

response of the noise. The main findings from the work presented in the paper were:

- The main noise effects are caused by additional low frequency accelerations from the different wave types associated with the earthquake (peaking below 0.1 Hz), caused by the arrival of different body and surface waves which was visible in the data both during the earthquake and for several hours afterwards. Of these, Rayleigh waves had the largest effect with significant effects also being noticed for Love waves. Much smaller effects were noticeable for P- and S-waves, especially for longer 60s integration times.
- The use of the RMS error as used by other authors does not give an accurate picture of the accuracy of measurements taken during or in the aftermath of an earthquake as it assumes white noise. The use of the Allan deviation should provide a better estimation of the measurement quality for a given reading, as this measures reading consistency and is recommended that the method is used for future assessments of instrument performance.
- In normal survey conditions prior to and after the earthquake, no significant differences were found between 30s and 60s integration times. In periods where the earthquake waves were present, the use of the longer 60s integration times was found to give more consistent results which were comparable or better than the period prior to the earthquake for P- and S-waves and only marginally worse for Love waves (10% increase in measurement spread). During the Rayleigh waves, although 60s cycles gave slightly better results, the errors were still too large for the readings to be useful being in excess of 10 μGals.
- The main challenge of dealing with a teleseismic event is the difficulty in recognising early waves prior to the arrivals of Rayleigh waves in either the time-domain signal or in the measurement standard deviation (and RMS error). It is recommended that an FFT is added to the measurement software to diagnose these noise effects early enough to take corrective action. The best approach to dealing with an earthquake during a survey, if it can be detected early enough, is to increase the measurement cycle times to at least 60s, and return to the base station to obtain a drift reading before the Rayleigh waves render useful readings impossible. Surveying should only resume once the Rayleigh waves have decayed.
- For very large earthquakes, such as earthquake 2 in this study, results from the day following the event showed a larger spread reflected by a 39% higher standard deviation compared to the day before the earthquake. Applying a 95% confidence threshold indicated that the errors would be up to 2.47 μGals higher the day after the earthquake itself which may compromise very high accuracy surveys and should either be accounted for with more repeat measurements or accounted for when interpreting the results. Subsequent days showed no large differences in performance in comparison to the period before the earthquake. Further investigation is required to determine the relationship between the magnitude, depth and distance of the triggering earthquake and the noise response of the gravimeter.
- In addition to the effects on individual measurement cycles, the study showed an increase in long period drifts over a survey day which both increased the spread of the data and caused shifts in the mean value of the data. A system of linear fits between base station readings was found to be superior to polynomial fitting both with and without earthquake noise on the data. However, whilst gaps of up to two hours could safely be used in normal conditions, gaps of this length created substantial errors in the period after the earthquake. It is recommended that the maximum gap between base

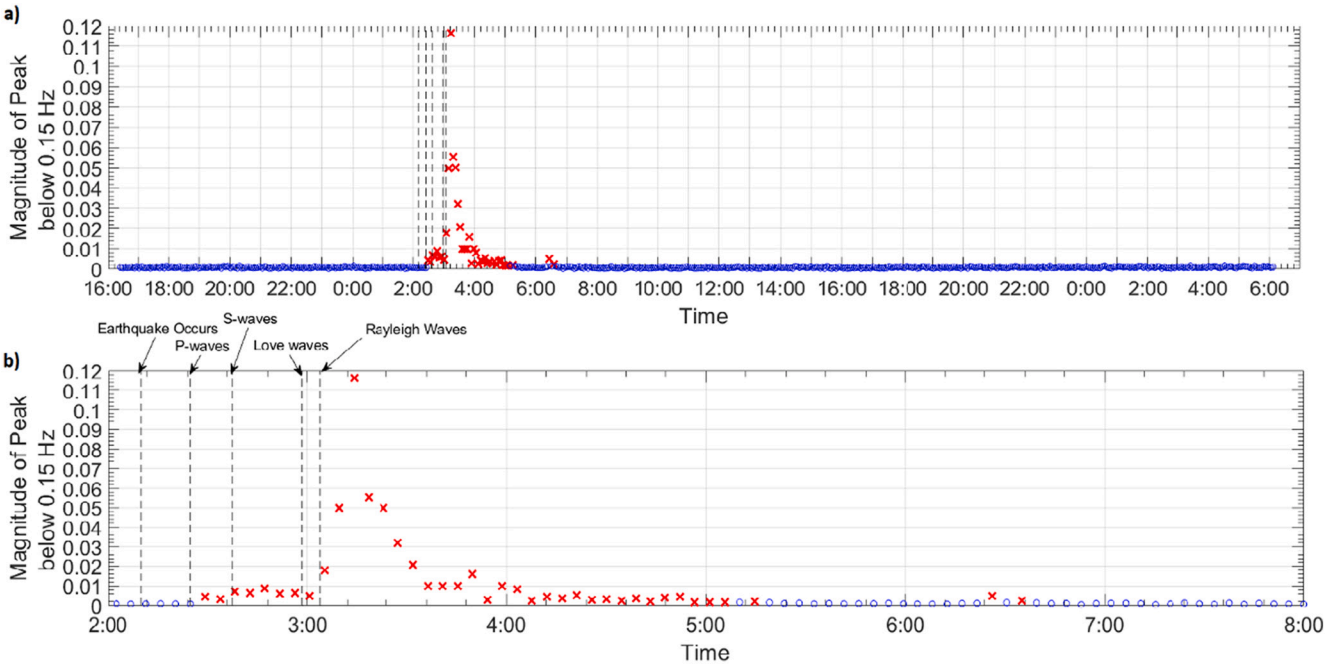

Fig. 12. The magnitude of the FFT for frequencies below 0.15 Hz for the periods before, after and during earthquake 1 in 2015. Magnitudes above 0.07 are flagged as potential removals and are shown as red crosses. The arrival times of the major earthquake wave components are also shown. See text for details. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stations is 1 h which limits the increase in the spread of data, expressed as the standard deviation to below 1 μGal.

- The use of gradiometer configurations drastically reduced the effects of the increased seismic noise on instrument performance, with instrument performance during a major earthquake being similar to the period before the earthquake for typical integration times. Significant degradation of performance was only noticeable for very long integration times and as a consequence of differences in the mechanical properties of the sensor springs between different instruments. This finding bodes well for the next generation of gravity gradient sensors based on quantum technology gravity gradient sensors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the financial support provided by Innovate UK and the Engineering and Physical Sciences Research Council (EPSRC) through the "Study of Industrial Gravity Measurements and Applications (SIGMA+) project" EP/M508378/1, REVEAL project EP/R000220/1 and ISCF Gravity Pioneer Project 104613.

References

Allan, D.W., 1966. Statistics of atomic frequency standards. Proc. IEEE 54 (2), 221–230.
Ardhuin, F., Herbers, T.H.C., 2013. Noise generation in the solid Earth, oceans and atmosphere, from nonlinear interacting surface gravity waves in finite depth. J. Fluid Mech. 716, 316–348.
Ardhuin, F., et al., 2011. Ocean wave sources of seismic noise. J. Geophys. Res. 116, C09004.
Ardhuin, F., et al., 2015. How ocean waves rock the Earth: two mechanisms explain microseisms with periods 3 to 300 s. Geophys. Res. Lett. 42, 765–772.
Arora, B.R., et al., 2008. First observations of free oscillations of the earth from Indian superconducting gravimeter in Himalaya. Curr. Sci. 95 (11), 1611–1617.
Barnes, D.F., 1966. Gravity changes during the Alaska earthquake. J. Geophys. Res. 71 (2), 451–456.
Boddice, D., et al., 2018. A novel approach to reduce environmental noise in microgravity measurements using a Scintrex CG5. J. Appl. Geophys. 152, 221–235.
Bonvalot, S., et al., 1998. Continuous gravity recording with Scintrex CG-3M meters:a promising tool for monitoring active zones. Geophys. J. Int. 135 (2), 470–494.
De Viron, O., et al., 2008. Retrieving earthquake signature in grace gravity solutions. Geophys. J. Int. 174 (1), 14–20.
Debeglia, N., Dupont, F., 2002. Some critical factors for engineering and environmental microgravity investigations. J. Appl. Geophys. 50 (4), 435–454.
Freier, C., et al., 2016. Mobile quantum gravity sensor with unprecedented stability. J. Phys. Conf. Ser. 723, 012050.
Gabalda, G., et al., 2003. CG3TOOL: an interactive computer program to process Scintrex CG-3/3M gravity data for high-resolution applications. Comput. Geosci. 29 (2), 155–171.
Ghobadi-Far, K., et al., 2019. Gravitational changes of the Earth's free oscillation from earthquakes: theory and feasibility study using GRACE inter-satellite tracking. J. Geophys. Res. 124 (7), 7483–7503.
Goncharenko, Y., et al., 2018. Using broadband seismic networks to optimize microgravity survey strategy in the United Kingdom. Near Surf. Geophys. 16 (4), 477–489.
Greco, F., et al., 2008. Seismic-induced Accelerations Detected by Two Parallel Gravity Meters in Continuous Recording with a High Sampling Rate at Etna Volcano, 51(1), p. 17.
Greco, F., et al., 2014. Characterization of the response of spring-based relative gravimeters during paroxysmal eruptions at Etna volcano. Earth Planets Space 66 (1), 44.
Grocholski, B., 2018. Gravity tracking of a great earthquake. Science 361 (6402), 565.
Harms, J., et al., 2015. Transient gravity perturbations induced by earthquake rupture. Geophys. J. Int. 201 (3), 1416–1425.
Hartmann, T., Wenzel, H.-G., 1995. The HW95 tidal potential catalogue. Geophys. Res. Lett. 22 (24), 3553–3556.
Hinton, A., et al., 2017. A portable magneto-optical trap with prospects for atom interferometry in civil engineering. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 375 (2099), 20160238.
Hinze, W.J., 1990. The role of gravity and magnetic methods in engineering and environmental studies. In: Geotechnical and Environmental Geophysics: Volume I, Review and Tutorial, 4. Society of Exploration Geophysicists, pp. 75–126.
Jousset, P., et al., 1995. Performance of two Scintrex CG3M instruments at the fourth International Comparison of Absolute Gravimeters. Metrologia 32 (3), 231.
Kaufmann, G., 2014. Geophysical mapping of solution and collapse sinkholes. J. Appl. Geophys. 111, 271–288.
Kulhánek, O., 1990. Anatomy of Seismograms Amsterdam. Elsevier.
Leis, J.W., 2011. Digital Signal Processing Using MATLAB for Students and Researchers. Wiley.
LaCoste, Romberg, 2004. Instruction Manual for Model G & D Gravity Meters. LaCoste & Romberg, Austin, Texas.

Levandowski, W., et al., 2017. Gravitational body forces focus North American intraplate earthquakes. Nat. Commun. 8, 14314.
Long, L.T., Kaufmann, R.D., 2013. Acquisition and Analysis of Terrestrial Gravity Data. Cambridge University Press, Cambridge.
Martínez-Moreno, F.J., et al., 2016. Collapse susceptibility map in abandoned mining areas by microgravity survey: a case study in Candado hill (Málaga, southern Spain). J. Appl. Geophys. 130, 101–109.
Montagner, J.-P., et al., 2016. Prompt gravity signal induced by the 2011 Tohoku-Oki earthquake. Nat. Commun. 7, 13349.
Nabighian, M.N., et al., 2005. Historical development of the gravity method in exploration. Geophysics 70 (6), 63ND–89ND.
Niebauer, T.M., et al., 2011. Monitoring earthquakes with gravity meters. Geodesy Geodyn. 2 (3), 71–75.
Reynolds, J.M., 2011. An Introduction to Applied and Environmental Geophysics. Wiley.
Rymer, H., 1989. A contribution to precision microgravity data analysis using Lacoste and Romberg gravity meters. Geophys. J. Int. 97 (2), 311–322.
Scintrex Ltd, 2006. Operating Manual for the CG5 Gravity Meter. Available from. http://www.scintrex.com (Document Part No. 867700).
Seigel, H.O., 1995. A Guide to High Precision Land Gravimeter Surveys.
Sugihara, M., 2004. Gravity monitoring with a CG5 Scintrex autogravimeter. ASEG Extend. Abstr. 2004 (1), 1–4.
Tikku, A.A., et al., 2006. Temporal fluctuations of microseismic noise in Yellowstone's Upper Geyser Basin from a continuous gravity observation. Geophys. Res. Lett. 33, L11306.
Tiwari, V.M., Mishra, D.C., 1997. Microgravity changes associated with continuing seismic activities in Koyna area, India. Curr. Sci. 73 (4), 376–381.
Tuckwell, G., et al., 2008. The use of microgravity to detect small distributed voids and low-density ground. Q. J. Eng. Geol. Hydrogeol. 41, 371–380.
Yiqing, Z., et al., 2011. Temporal variation of gravity field before and after Wenchuan Ms8. 0 earthquake. Geodesy Geodyn. 2 (2), 33–38.
Zhu, Y., Zhan, F.B., 2012. Medium-term earthquake forecast using gravity monitoring data: evidence from the Yutian and Wenchuan earthquakes in China. Int. J. Geophys. 2012, 6.