RADAR

Research Archive and Digital Asset Repository

OXFORD
BROOKES
UNIVERSITY

This version is available: https://radar.brookes.ac.uk/radar/items/9138ffd6-b883-4429-9927-8984713aa6ad/1/

WWW.BROOKES.AC.UK/GO/RADAR

Investigating hydrological contributions to volcano monitoring signals. A time-lapse gravity example.

B. Hemmings[1], J. Gottsmann[1], F. Whitaker[1], A. Coco[2]

[1]*School of Earth Sciences, University of Bristol, UK*
[2]*Oxford Brookes University, UK*

20 July 2016

SUMMARY

Geophysical techniques are widely used to monitor volcanic unrest. A number of studies have also demonstrated that hydrological processes can produce or trigger geophysical signals. Hydrologically-induced gravity signals have previously been recorded by specifically designed gravity surveys as well as, inadvertently, by volcano monitoring studies. Water table corrections of microgravity surveys are commonplace. However, the fluctuations of the water table beneath survey locations are often poorly known, and such a correction fails to account of changes in water-mass storage in the unsaturated zone. Here, we combine 2D axis-symmetrical numerical fluid-flow models with an axis-symmetric, distributed-mass, gravity calculation to model gravity changes in response to fluctuating hydrological recharge. Flow simulations are based on tropical volcanic settings where high surface permeabilities promote thick unsaturated zones. Our study highlights that mass storage (saturation) changes within the unsaturated zone beneath a survey point can generate recordable gravity changes. We show that for a tropical climate, recharge variations can generate gravity variations of over 150 μGal; although, we demonstrate that for the scenarios investigated here, the probability of recording such large signals, is low. Our modelling results indicate that microgravity survey corrections based on water table elevation may result in errors of up to 100 μGal. The effect of inter-annual recharge fluctuations dominate over seasonal cycles which makes prediction and correction of the hydrological contribution more difficult. Spatial hydrogeological heterogeneity can also impact on the accuracy of relative gravity surveys, and can even result in the introduction of additional survey errors. The loading fluctuations associated with saturation variations in the unsaturated zone may also have implications for other geophysical monitoring techniques, such as geodetic monitoring of ground deformation.

Key words: Hydrogeophysics – Volcano monitoring – Time variable gravity – Permeability and porosity

1 INTRODUCTION

Volcanic hazard monitoring utilises a number of geophysical techniques. These monitoring tools are sensitive to physical changes within the sub-surface such as density, electrical properties and elastic deformation. In volcanic settings geophysical signals are regularly interpreted to reflect changes within the magmatic system and to track the movement of magma (e.g. Chouet 2003; Dzurisin 2003). As technology, data collection and processing strategies have improved, volcanic monitoring techniques are able to resolve increasingly small signals (e.g. McNutt 2005; Battaglia et al. 2008; Parker et al. 2014; Muller et al. 2015). This ability to resolve finer signals increases the challenge to decipher underlying physical sources that may or may not be related to the magmatic system and volcanic unrest. One non-magmatic source of geophysical signals are temporal and spatial hydrological variations. These hydrologically-induced signals may reflect a perturbation of the volcanic hydrothermal system and therefore can indicate changes in volcanic behaviour (e.g. Bianco et al. 2004; Todesco 2009). However, non-volcanic hydrological fluctuations, driven by weather and climate, also generate recordable geophysical signals (e.g. GPS, Argus et al.2014, Fu et al.2015; seismicity, Jiménez & Garca-Fernández2000; gravity, Jacob et al.2010). This can present a risk of misinterpretation where the purpose of the geophysical campaigns is to monitor pre-eruptive or eruptive activity.

This study focusses on geophysical signals resulting from climatically-driven subsurface water mass variations. Mass variations are typically recorded by gravimetric monitoring campaigns. Therefore, we target this investigation on exploring gravimetric changes associated with variations in sub-surface water storage due to variations in hydrological recharge and hydrogeological properties. We assess the potential for contamination of volcanigenic gravimetric signals with signals resulting from seasonal and inter-annual recharge variations that are typical of tropical climates.

1.1 Gravimetry used in volcanic monitoring

Gravity measurements are often an important component of the geophysical monitoring suite at active volcanoes. Static gravity data can be used to image and identify structural and compositional heterogeneities in the subsurface (Hautmann et al. 2013). Re-occupying a survey network over periods of months to decades can provide a record of mass or density changes over space and time. Such dynamic, or time-lapse, gravity monitoring has been used to track the complex evolution of volcanic unrest (Gottsmann et al. 2006a,b) and observe potential eruption precursors (Rymer 1994; Carbone et al. 2003a). Table 1 provides examples of the amplitude and timescales of observed gravity changes at volcanoes around the world.

The nature of postulated magma-related gravity sources varies. Microgravity signals associated with eruptive phases at Mt Etna (Italy) have been attributed to basic magma injection both within and below the edifice (as summarised by Budetta et al.2004). In more silicic systems, a more complex relationship between magma injection and recorded gravity signals has been suggested. Eggers (1983) proposed that magma injection at Pacaya Volcano, Guatemala generates a gravity decrease as high density, degassed magma is displaced by vesiculated, low density magma. Subsequent degassing and crystallisation of intruded magma would result in a density increase and a positive residual gravity anomaly, as is suggested by Jousset et al. (2000) to explain the residual gravity signal observed at Merapi in 1993 to 1994.

Gravity surveys are typically combined with geodetic data to account for gravity changes related to ground deformation. A number of studies have used these datasets to distinguish magmatic signals from hydrothermal sources of deformation, often in large caldera forming systems (Berrino et al. 1984; Battaglia 1999; Battaglia et al. 2008). Studies that reveal a residual gravity anomaly after correction for elevation changes associated with ground deformation often implicate a magmatic source for the signal. For example, Battaglia (1999) suggest that intrusion of silicate magma at Long Valley caldera can explain both uplift and a residual gravity change of up to 64 μGal between 1982 and 1998. A magmatic source for deformation was initially favoured to explain deformation and gravity relationships observed at Campi Flegrei in the early 1980s (Berrino et al. 1984). Later studies, however, have suggested that the deformation between 1983 and 1984 was due to migration of hydrothermal fluids (Battaglia et al. 2006).

The simple assumption that residual gravity

Table 1. Examples of gravity signal amplitudes and timescale. Note, attempts to account for water table fluctuation may already be incorporated in the quoted Δg.

Location	Time period	Δg (μGal)	Timescale (years)	Reference
Campi Flegrei	1982-1984	150	2	Berrino (1994)
Mount Etna	2000-2001	80	1	Carbone et al. (2003a)
Long Valley Caldera	1982-1998	64	16	Battaglia (1999)
Merapi	1993-1994	60	1	Jousset et al. (2000)
Montserrat	2007-2008	74	1.5	Hautmann et al. (2010)
Sakurajima	1975-1985	120	10	Yokoyama (1989)
Unzen	1999-2004	79, 70	1, 5	Saibi et al. (2010)

anomalies - those that remain after correction for deformation - are due to magmatic processes can be questioned, as potential non-magmatic mass change sources have been identified. A number of authors have proposed that induced gravity changes may be due to dynamic behaviour in the hydrothermal system itself (Gottsmann et al. 2005; Todesco & Berrino 2005; Todesco et al. 2010). To explain short period (less than 1 hour) gravity variations of up to 20 μGal at Nisyros caldera, Gottsmann et al. (2005) suggest movement and accumulation of steam pockets. With addition of other geophysical data records, Gottsmann et al. (2007) propose that these gravity cycles are related to deformation associated with pressure changes within the hydrothermal system on Nisyros. Using TOUGH2 fluid-flow models, Todesco & Berrino (2005) demonstrate that fluid density changes associated with longer period perturbation of the hydrothermal system beneath Solfatara, Campi Flegrei may be responsible for the observed gravity residual fluctuations in the years after the early 1980s bradyseism.

The importance of understanding the distribution of fluids in the subsurface and the effect that it has on gravity measurements is now well acknowledged. Variations in water table elevation are an important correction during the reduction of gravity data (Jachens & Roberts 1985; Battaglia et al. 2003). However, direct measurements of water table and aquifer storage potential (porosity), concurrent with gravity surveys, are rarely available. Assumptions and workarounds used to mitigate this data deficiency include spatially and temporally interpolating the water table elevation, where some measurements do exist (Battaglia 1999), and assuming that the changes in water level are seasonal, secular and consistent (Carbone et al. 2003a; Budetta et al. 2004). Some studies simply demonstrate that the observed gravity signals are too large to

be solely due to water table elevation variations, estimated from rainfall records and recharge assumptions (Jousset et al. 2000; Hautmann et al. 2009). Budetta et al. (1999) recorded seasonal gravity variations of 20 μGal (peak to peak) on Mt Etna. Although this hydrological gravity signal obscured volcanically induced gravity signals at distal stations, where the volcanic signals are smaller, they were still able to distinguish magma-induced gravity changes in more proximal localities. Their study demonstrates the importance of understanding hydrological component of recorded gravity changes, particularly where the expected signals of volcanic unrest are small.

1.2 Known hydrological influence

Currently, compensation for hydrological influences on gravity measurements at active volcanoes consists of accounting for variations in the elevation of the water table. This is normally achieved by representing the water table aquifer as an infinite slab. Under the infinite slab assumption, the gravity correction due to a change in water table elevation of δz, is given by:

$$\Delta g_{wt} = 2\pi G \rho_w \phi \delta z \,, \qquad (1)$$

where Δg_{wt} is the water table gravity correction, G is the universal gravitational constant (6.67×10^{-11} m^{3}kg^{-1}s^{-2}), ρ_w is the density of water and ϕ is the effective porosity of the aquifer. However, even where accurate Δg_{wt} estimates are possible, through concurrent water table and gravity measurements, this hydrological correction fails to account for water-mass storage above the water table, in perched aquifers and within the unsaturated (vadose) zone. An increasing number of studies are using gravity measurements to assess and track water storage changes in the sub-

surface (e.g. Jacob et al. 2010; Naujoks et al. 2010; Creutzfeldt et al. 2012; Pfeffer et al. 2013; Hector et al. 2015). In a survey designed to assess water storage changes in karst environments in the south of France, Jacob et al. (2010) highlight the importance of mass storage in the vadose zone on gravity measurements. They suggest that the potential for mass storage in the epikarst, above the water table, is significant and in some sites is responsible for the entire seasonal gravity signal (Jacob et al. 2009). High precision, superconducting gravimeter experiments in Moxa, Germany have revealed seasonal, hydrologically induced, gravity signals $\sim$ 3 to 4 μGal (Naujoks et al. 2010). However, these signals are often masked in such sensitive gravimetric data by local hydrological effects, such as water content variations in the soil and disaggregated bedrock immediately surrounding the gravimeter (Krause et al. 2009).

In many volcanic settings high permeabilities, combined with high relief, provide a potential for a deep water table and a relatively thick vadose zone, on the order of several tens to hundreds of meters (e.g. Yucca Mountain, Nevada, Klavetter & Peters1986; Hawaii, Ingebritsen & Scholl1993; Oregon Cascades, Hurwitz et al.2003; Masaya Volcano, Nicaragua, Pearson et al.2012). Capillary pressures in unsaturated material can promote water retention, especially in more fine grained material. Therefore, a thick vadose zone can have significant mass storage potential. Such hydrogeology, coupled with temporally and spatially variable recharge, has the potential to generate recordable and significant gravity signals independent of any change in water table elevation. For example, a change in liquid saturation fraction of 0.2 in the vadose zone with porosity of 0.25 results in a density change ($\Delta\rho$) of $\sim$50 kg/m^3. Using a Bouguer slab approximation (Equation 2) and an estimated vadose zone thickness of h = 50 m, this change in water saturation would produce a gravity signal of over 100 μGal.

$$\Delta g = 2\pi \; G \; \Delta\rho \; h \,. \qquad (2)$$

Even where saturation changes throughout the subsurface are integrated to provide effective water table changes, the Bouguer approximation has been shown to be inaccurate, especially where surface topography is significant (e.g. Creutzfeldt et al. 2008; Leirião et al. 2009; Creutzfeldt et al. 2012). By deriving time constants relating rainfall to gravity change, Cross-ley et al. (1998) developed an empirical method to correct for saturation and water table changes. Estimation of the time constants that relate rainfall to gravity changes requires regularly sampled (at least daily) and concurrent, gravity and rainfall timeseries for a number of years (e.g. Mouyen et al. 2013). These time-constants are not necessarily uniform across a gravity survey network; the collection of many years of continuous gravity data across a 4D time-lapse microgravity network, with many tens of benchmark survey locations is, unfortunately, not feasible. Here, we couple hydrological simulations with gravity calculations to model potential gravity signals associated with temporal and spatial variations in groundwater recharge that are typical of a number of volcanic environments. We detail a method for converting a finite element distribution of saturation changes into modelled gravity signals. This method can provide a means to identify and correct hydrological contributions to recorded gravity signals.

2   METHODS

2.1   Hydrological model

Hydrological simulations are performed using the TOUGH2 code (Pruess 2004) for an isothermal system at 25 °C. Two model geometries are presented here. Type *flat* have a flat ground surface at $z = 400$ m and base at $z = -50$ m (Fig. 1a). Type *topo* incorporates topography, reflecting a cross-section through Centre Hills, Montserrat, as an example of a typical volcanic island. Local topographic variations of 100 m are superimposed on a slope of 700 m across the model domain (Fig. 1b). For both model types, the domain is 2D axisymmetric about $x = 0$ m. The distal lateral boundary at $x = 3600$ m is open to flow, as summarised in Fig. 1. The distal boundary condition is designed to mimic a coastal boundary with the water table at sea level ($z = 0$ m). The basal boundary is closed and the top surface is open to flow with a gas filled atmosphere (at atmospheric pressure). Recharge is modelled by generating water in the surface cells of the domain.

The domain is discretised into 51,457 cells for *flat* models and 41,609 for *topo* models. In the axis-symmetric models cell volumes generally increase with radial distance from the axis at $x = 0$ (Figs. 2a and 3a). We refine the column widths below the modelled survey benchmark locations
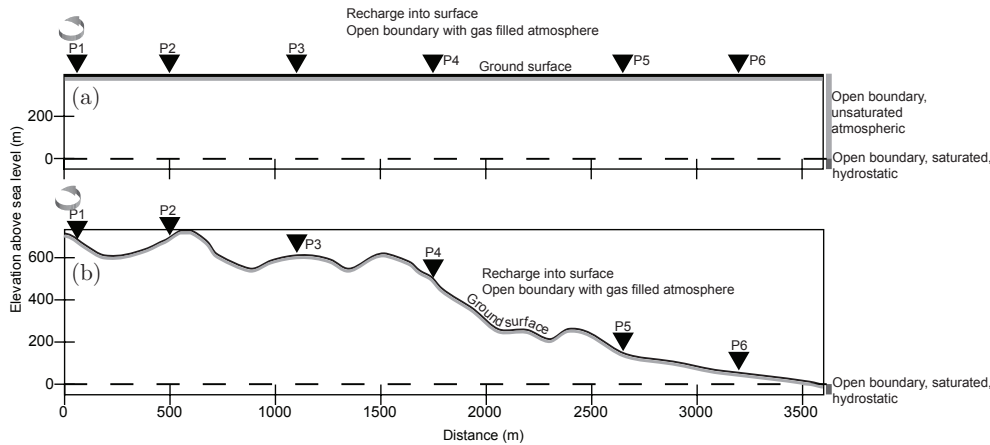
Figure 1. 2D model domain for gravity simulations for (a) *flat* and (b) *topo* models. All models are axis-symmetric about $x = 0$ m. Recharge is simulated by water generation in the cells comprising the ground surface boundary, which is open to a gas filled atmosphere. The right-hand (distal) boundary is open to flow with hydrostatic pressure below $z = 0$ m and atmospheric pressure above. Black inverted triangles represent the six synthetic gravity survey locations (P1-P6) for which gravity time-series are derived.

from 10 m down to 1.25 m (beneath black inverted triangles in Figs. 1 - 3). Cell thicknesses range from 2 to 10 m. Absolute permeability is described as an exponential function with depth, d, after Saar & Manga (2004):

$$k = k_0 e^{-\lambda d}, \qquad (3)$$

where k is permeability in m^2, at depth d in metres, relative to the ground surface. k_0 is the surface permeability (5×10^{-13} m^2) and $\lambda = 0.004$. Key hydrological properties are described in Table 2. In some simulations this basic permeability distribution is modified to include a low permeability core in the region $0 \le x \le 1400$ m and $-50 \le z \le 250$ m (Figs. 2b and 3b), reflecting the conceptual model Type 1, described in Hemmings et al. (2015). This conceptual model cites the presence of a low permeability core of intrusive and altered volcanic material to explain high discharge springs at high elevations around the extinct volcanic complex of Centre Hills, Montserrat. The hydrological properties of this region are defined as "low-k core" in Table 2

2.1.1 Recharge

Recharge, reflecting the proportion of precipitation that enters the groundwater system, is specified as water generation in the upper-most cells (below the atmosphere). Volcanic settings are

commonly associated with high relief topography. There are often strong relationships between elevation, precipitation, and associated recharge. Spatial recharge variation can also be strongly controlled by evapotranspiration which reflect natural and anthropogenic changes in land-use and vegetation distribution. Recharge models can be used to estimate spatially distributed recharge rates, accounting for variable precipitation, as well as temperature and land-use controlled evapotranspiration (e.g. Hughes et al. 2008; Hemmings et al. 2015). Here, we use recharge model results for Montserrat, presented in Hemmings et al. (2015), to define an elevation dependant recharge rate::

$$r_z = 3.15 \times 10^{-3} z + 0.186, \qquad (4)$$

where r_z is recharge in mm/day at elevation z in metres above sea level.

Each model scenario is subject to an initial simulation phase which is run to steady-state. In this first simulation phase the initially saturated models drain under constant recharge to establish a steady-state water table elevation and saturation condition. The models are then subject to 100 years of six-monthly and annually varied recharge. The varied recharge input is pseudo-synthetic, derived from observations from Montserrat, where annual rainfall varies spatially and temporally from 1000 to 3000 mm/yr, and
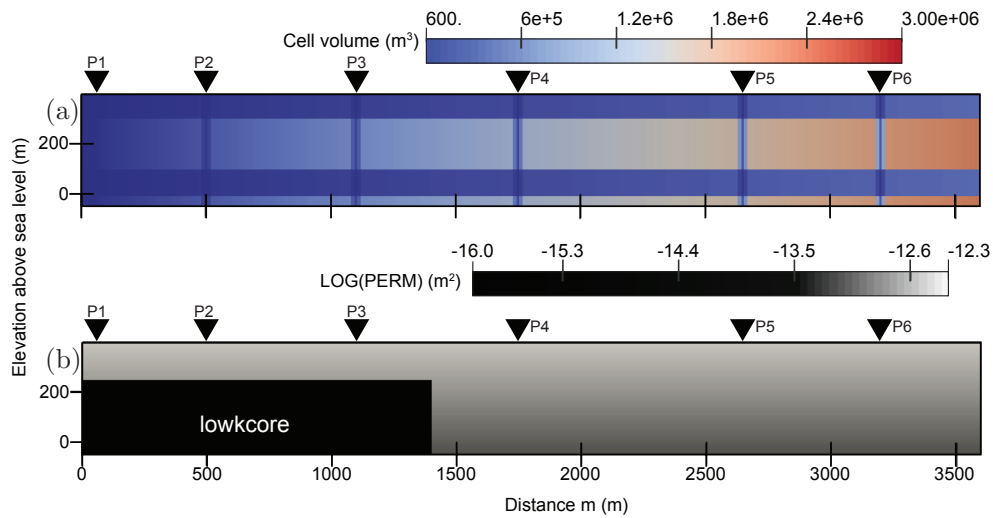
Figure 2. *flat*-type models showing cell volume (a) and permeability distribution (b). a) demonstrates the increasing cell volume with distance from the axial centre (at $x = 0$ m). Cell volume is less below the survey locations (black inverted triangles) as the cell width is reduced to 1.25 m. Cell thickness is reduced from 10 m to 2 m in the upper 100 m of the domain and at $0 < z < 100$ m. Permeability in b) reduces exponentially with depth (d), according to Equation 3. The plot here includes a low permeability region (low-k core) at $0 < x < 1400$ m and $-50 < z < 250$ m.

recharge from <50 to >2000 mm/yr. A series of 100 annual factors are generated from random gaussian distribution, based on the modelled annual variation in the recharge for Montserrat (Hemmings et al. 2015), with standard deviation of 0.5 about a mean of 1. The recharge rate in each cell, as defined by Equation 4, is multiplied by the same series of annual factors, producing a 100 year varied recharge input for each surface cell. We also explore the effect of seasonality of recharge on subsurface saturation by superimposing a regular seasonal variation, whereby 15% of the annual recharge is assigned evenly across the first six months of each year and 85% evenly across the second six months. This is achieved by adjusting the annual recharge rate to a six-monthly step function; the total volume of recharge for each year is preserved.

Eight contrasting model scenarios are simulated, exploring the effects of topography and spatially varied recharge on the gravity signal at the surface. The differences between these model scenarios are summarised in Table 3. For model *flat*1 and *flat*2, the recharge rate is consistent across all surface cells. The initial steady-state recharge rate is 1.44 mm/day (525 mm/yr, 1.66×10^{-5} kg/s/m^2). The 100 year variable recharge time-series is displayed in Fig. 4.

Models *topo*1, *topo*2 use the same spatially uniform recharge rate, neglecting the elevation control on recharge. In models *topo*3 and *topo*4, the recharge rate in each surface cell varies with elevation, according to Equation 4. The same spatial recharge variation is forced on the 'flat' models *flat*3, *flat*4. In the models with spatially varied recharge, the phase of the 100 year recharge time-series is consistent across the surface and only the amplitudes of the recharge rate differs between recharge cells.

2.1.2 Simulation time-steps

The time steps during the flow simulations are automatically adjusted, depending on the convergence rate of the Newton-Raphson iterations used to solve the nonlinear, coupled flow equations. The during our simulations the time-step ranges from 1 second to an enforced maximum of 25 days. Time step lengths are modified to honour the time of generation rate changes. The mean and modal time-step length for each simulation range from 4.7 to 18.9 days, and 1.5 to 23.5 days, respectively. The mean time-step length of all simulations is 8.9 days. To prevent aliasing of saturation changes when formulating the gravity
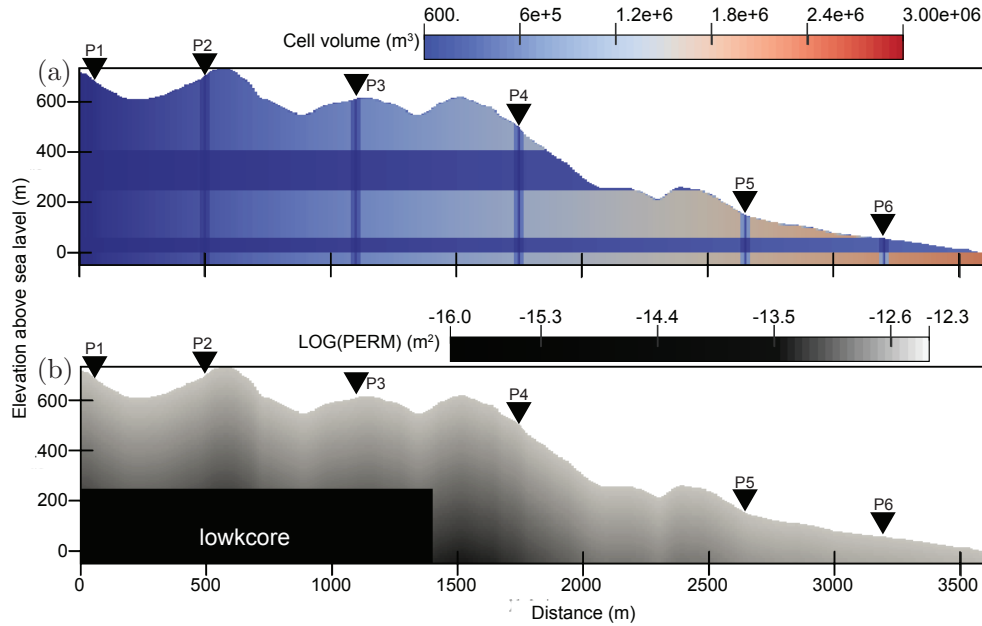
Figure 3. *topo*-type models showing cell volume (a) and permeability distribution (b). a) demonstrates the increasing cell volume with distance from the axial centre (at $x = 0$). Cell volume is less below the survey locations (black inverted triangles) as the cell width is reduced to 1.25 m. Cell thickness is reduced from 10 m to 2 m in the regions where $250 < z < 400$ m and $0 < z < 100$ m. Permeability in b) reduces exponentially with depth (d), according to Equation 3. The plot here includes a low permeability region (low-k core) at $0 < x < 1400$ m and $-50 < z < 250$ m.

signal, the modelled saturation data is outputted at least once every 3 months.

2.2 Gravity formulation

Leirião et al. (2009) present a detailed method for converting three-dimensional finite-difference hydrological model results into simulated gravity changes. Their method, also adopted by Piccolroaz et al. (2015), calculates the contribution of each element to the simulated gravity change at any given point by using either in the a prismatic volume-mass approach (after, Nagy 1966), the MacMillan equation (Creutzfeldt et al. 2008) or a point-mass solution. Both the prism and MacMillan equations incorporate information about the dimension or shape of the elements. Which method is used for each element is related to the shape of the element and the distance to the simulated gravity survey point. The relatively computationally intensive prism equation is used for the most proximal elements, the point-mass equation for the most distal, and the

MacMillan equation is used for intermediate elements.

Similar to the approaches outlined by Leirião et al. (2009), we develop a method for calculating the gravity signal at a given surface location (P) that results from the distribution of mass throughout the entire model domain. Our approach is an extension of the point-mass method. Apart from computational efficiency, the advantage of the point-mass method for our purposes is geometrical flexibility; it allows the calculation of simulated gravity changes from linear and axisymmetric two-dimensional hydrological models.

The contribution of each element (i), with density ρ_i, centred at cylindrical coordinates (α_i, z_i), to the gravity measurement recorded at $P(\alpha_p, z_p)$ is calculated by considering the element as a ring centred on $\alpha = 0$ (Fig. 5). The vertical component of gravitational acceleration (in m/s^2), recorded at point P, associated with a ring of radius α_i, radial thickness $d\alpha$ and elevation thickness dz, is defined by integrating around the ring for angles $0 \leq \theta < 2\pi$ radians:

Table 2. Permeability, porosity and van Genuchten (capillary pressure and relative permeability) parameters (α, m,S_{lr} and S_{gr}) used in simulations. α is related to air-entry pressure, and m to pore size distribution. S_{lr} and S_{gr} are residual liquid and gas saturations, respectively. The models implement the modified van Genuchten curves of Luckner et al. (1989) and incorporate the method of Webb (2000) which uses a logarithmic extension of the capillary pressure curve for saturations below $S_{lr} + \varepsilon$.

Material	Permeability (k_0) (m^2)	Porosity	α (MPa^{-1})	m	S_{lr}	S_{gr}	ε
main	5×10^{-13}	0.34	200	0.85	0.1	0	0.001618
low-k core	1×10^{-16}	0.1	200	0.85	0.1	0	0.001618

Table 3. Summary of differences between model simulations. *flat*-type models have uniform ground surface elevation at 400 m (see Figs 1a and 2) and are denoted by horizontal line in topography column. The top surface of *topo*-type models is based on a topographic profile through a typical volcanic arc island (see Figs 1b and 3) as denoted by the profile line in the topography column. Horizontal lines and profile lines in the Recharge column reflect spatially homogeneous and elevation dependant recharge, respectively. A checkmark in "low-k core" column indicates that the simulation includes a low permeability unit in the region $0 \leq x \leq 1400$ m and $-50 \leq z \leq 250$ m.

Simulation	Topography	Recharge	low-k core
*flat*1	———	———	✗
*flat*2	———	———	✓
*topo*1	∿∿	———	✗
*topo*2	∿∿	———	✓
*flat*3	———	∿∿	✗
*flat*4	———	∿∿	✓
*topo*3	∿∿	∿∿	✗
*topo*4	∿∿	∿∿	✓

Figure 5. Geometrical illustration of the gravity due to an elemental ring, centred at $\alpha = 0$ and elevation z_i, with radius α_i, elevation thickness dz and radial thickness $d\alpha$, recorded at point P. $S(\theta)$ is the distance from P to any point on the ring and is defined in Equations 6.

$$\Delta g(P, \Delta t) = G \sum_{\alpha_i} \sum_{z_i} \alpha_i \, \Delta\rho_i \, d\alpha_i \, dz_i \, (z_p - z_i)$$
$$\cdot \int_0^{2\pi} \frac{1}{(\alpha_i^2 - 2\alpha_i\alpha_p cos\theta + \alpha_p^2 + (z_p - z_i)^2)^{\frac{3}{2}}} d\theta \tag{7}$$

For the purpose of modelling gravity changes due to water-mass storage, the change in density of an element is defined by:

$$\Delta\rho_i = \Delta S_{li}\rho_w\phi_i \,, \tag{8}$$

where ΔS_{li} is the change in liquid saturation, ρ_w is the density of water and ϕ_i is the porosity of the cell i.

Equations 7 and 8 are used to compute the

$$g_i(P) = G\rho_i\alpha_i d\alpha dz \int_0^{2\pi} \frac{(z_p - z_i)}{s(\theta)^3} \, d\theta \,, \tag{5}$$

where $s(\theta)$ is a function defining the distance between benchmark survey location and each elemental component on the ring:

$$s(\theta) = (\alpha_i^2 - 2\alpha_i\alpha_p cos\theta + \alpha_p^2 + (z_p - z_i)^2)^{\frac{1}{2}} \,. \tag{6}$$

The total gravity change recorded at point P due to a change in mass or density within the model is calculated by summing the contribution of all the elemental rings:
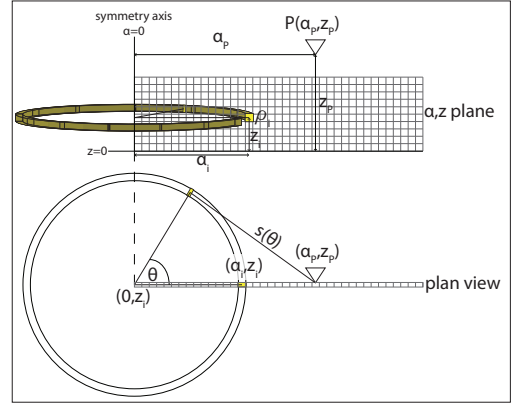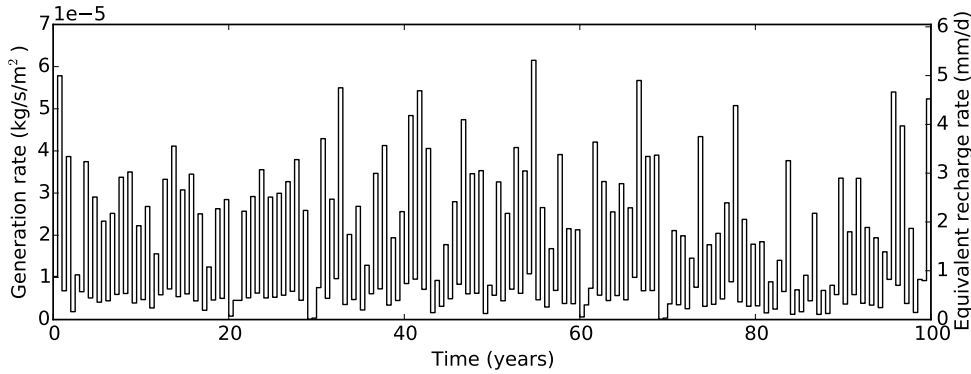
Figure 4. Synthetic generation rate time-series for cell at 400 m elevation.

gravity signal created by temporal variations in saturation at 6 survey locations on the surface of the model domain (black inverted triangles in Fig. 1). For comparison, we also compute the gravity correction (Δg_{wt}, Equation 1) associated with the modelled change in the water table elevation below each survey location. Modelled gravity signals are presented in μGal (1 μGal $= 1 \times 10^{-8}$ m/s^2).

3   RESULTS

Each model scenario simulation (Table 3) results in a 100-year Δg time-series at every survey location (P1-P6). All model scenarios produce a variation in Δg of over 150 μGal during the 100 years of varied recharge for at least one survey location. The simulations produce a maximum Δg between 55-60 years and 65-70 years and a minimum Δg between 85 and 95 years, with respect to a value of $\Delta g = 0$ at time zero. The minimum and maximum of each simulated time-series is displayed in Fig. 6 and in tabular form in Table **??**.

The gravity signals in simulations $flat1$ and $flat2$, without spatially variable recharge and with flat topography, are similar for each simulated survey location. (e.g. $flat2$, Fig. 7). Consequently, the minimum and maximum gravity changes in the 100 years of simulation are relatively consistent for survey locations P1-P6 (Fig. 6). The model scenarios that incorporate topography and/or the spatially varied recharge input demonstrate more gravity signal variability between different survey locations (e.g. $topo4$, Fig. 8). Signal amplitudes are greater for the stations closer to the axis.

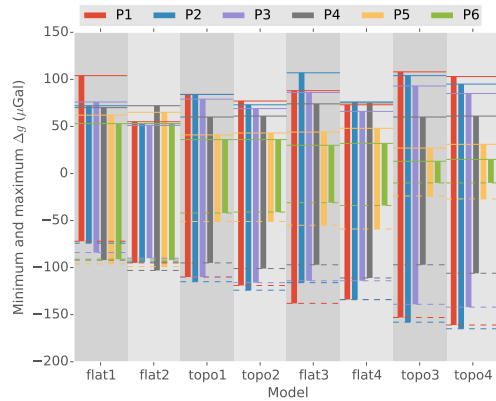Compared to the spatially uniform recharge

Figure 6. Minimum and maximum of simulated gravity signals at the 6 survey locations (P1-P6) for each model.

scenarios ($flat1$, $flat2$, $topo1$ and $topo2$), recharge is higher towards the axial boundary of the spatially varied scenarios ($flat3$, $flat4$, $topo3$ and $topo4$). As a result, the magnitude of the temporal variation is also higher. Similarly, recharge, and therefore the magnitude of storage variations, is lower in surface cells towards the distal boundary in these models (where elevation is lower). This contributes to higher Δg amplitudes for the more axial stations in models $flat3$, $flat4$, $topo3$, and $topo4$. The presence or absence of a core with low permeability and porosity appears to have a very limited effect on the simulated gravity signals in these models, although it does raise the water table by up to 135 m in the interior of the model. Modelled Δg time-series for all scenarios are presented in Appendix **??** (Figs. **??** to **??**).
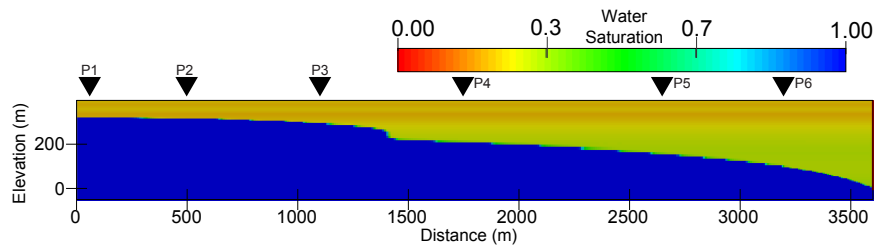
Liquid saturation at 70 yrs.
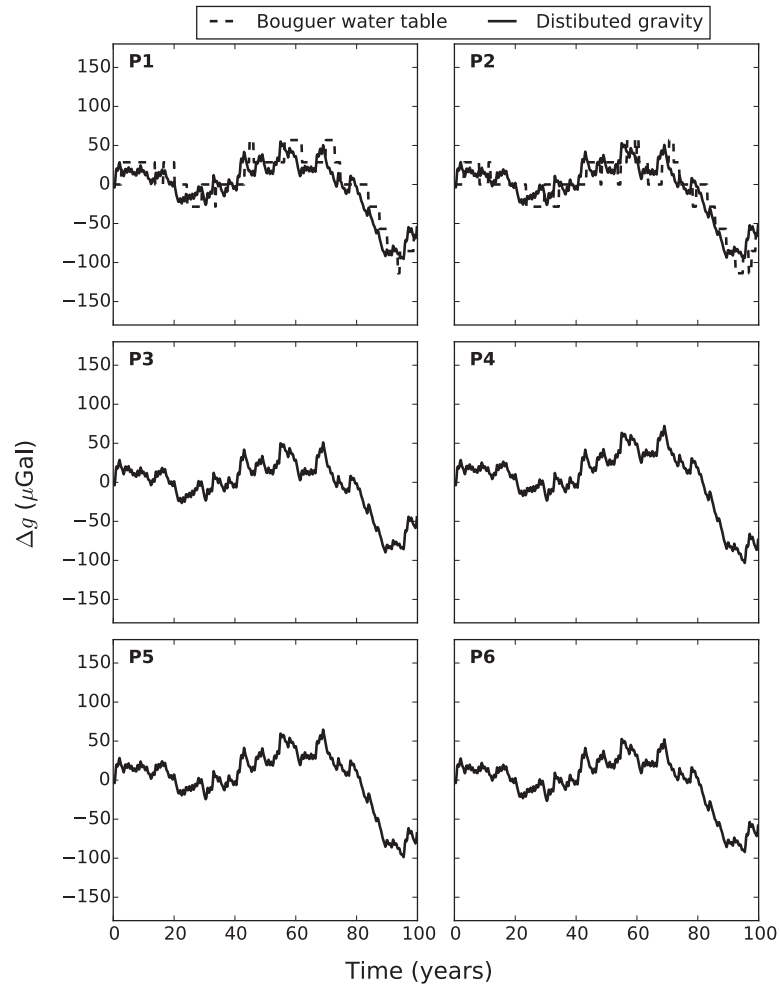
Modelled gravity changes as measured at model survey locations P1-P6.

Figure 7. Results from simulation *flat2*. Step in water table at $x \sim 1400$ m in a) is due to low permeability unit. Solid line in b) is the gravity signal calculated based on the distributed changes in saturation (Equations 7 and 8). The dashed line for P1 and P2 represents the commonly applied correction, based on water table elevation (Equation 1). Poor resolution in the model below P3 to P6 at the elevation of the water table causes large step artefacts in the water table correction; therefore, it is not displayed.
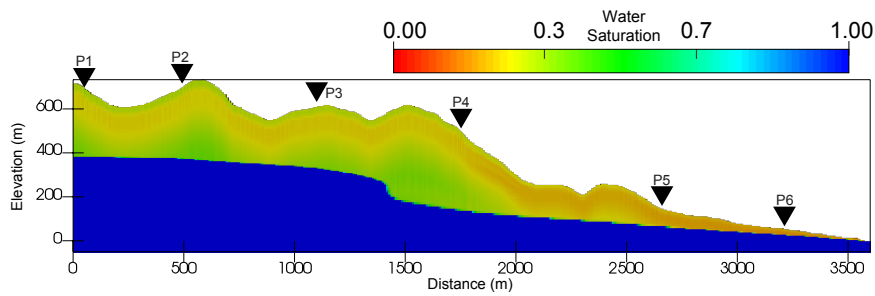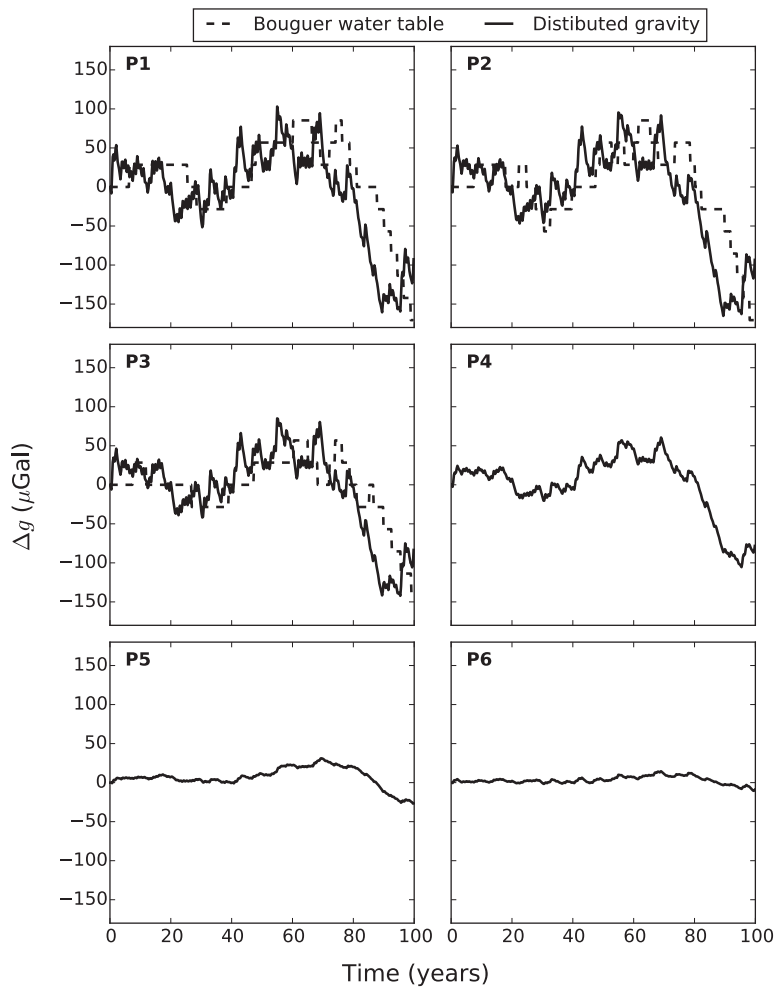
Liquid saturation at 70 yrs.

Modelled gravity changes as measured at model survey locations P1-P6.

Figure 8. Results from simulation *topo*4. Step in water table at $x \sim 1400$ m in a) is due to low permeability unit. Solid line in b) is the gravity signal calculated based on the distributed changes in saturation (Equations 7 and 8). The dashed line for P1 to P3 represents the commonly applied correction, based on water table elevation (Equation 1). Poor resolution in the model below P4 to P6 at the elevation of the water table causes large step artefacts in the water table correction; therefore, it is not displayed.

4 DISCUSSION

The flow simulations with 100 years of variable recharge presented above display appreciable and recordable changes in gravity, on the order of 10s to 100s of μGal. The results indicate that topography, even in the absence of any associated spatial recharge variation, can produce significant spatial variations in gravity changes (compare P1 and P6 in models *topo*1 and *topo*2, Figs. **??** and **??**). This is related to differences in vadose zone thickness, induced by high topographic relief, and relatively high and homogeneous permeabilities. The vadose zone thickness at P6 in the topographic models is just 20 m, compared to between 300 and 400 m at P1 and P2 (with and without the lower permeability core, respectively). With the relatively thin vadose zone beneath P6, recharge reaches the water table relatively quickly; from here it is transported laterally rather than stored, and the recharged mass has less effect on the gravity recorded at a fixed location on the surface. The mass storage potential beneath the survey location is relatively low and therefore the amplitude of the gravity signal is low, compared to P1. The tendency for vadose zone thickness to correlate with elevation means that gravity changes associated with saturation variations are likely to be stronger towards volcano summits, potentially mimicking the spatial patterns commonly associated with magmatic signals. Spatial variations in recharge and storage also have a strong control on spatial differences in modelled gravity. The amplitude of the gravity signal is largely dependent on the magnitude of the temporal recharge variation. Compare, for example, the results from the flat model with spatially varied recharge (*flat*3 and *flat*4, Figs. **??** and **??**) to those with spatially uniform recharge (*flat*1 and *flat*2, Figs. **??** and 7).

The modelled gravity changes are dominated by inter-annual recharge variations. The seasonal gravity fluctuations are generally low amplitude (< 10 μGal). Small seasonal gravity variations are comparable with early estimates from Mount Etna (Sanderson 1982), however, more recents studies suggest that the seasonal hydrological component of gravity measurements from Etna may be on the order of 20 μGal (Budetta et al. 1999; Carbone et al. 2003b). None of the model scenarios presented here produce seasonal changes of that magnitude. However, it may be possible to generate such signals under alternative hydrological regimes that are not investigated in this model suite.

4.1 Comparison to common water table correction

Correction of gravity signals for hydrology, using Equation 1, most commonly involves some estimate of water table elevation changes (δz). More rarely, direct measurements of δz are used. In our simulation we calculate this correction, under the idealised scenario of knowing δz beneath each survey location, through time. We can compare this water table correction with the modelled distributed gravity signal that also incorporates vadose zone mass changes.

Unfortunately, evaluation of δz in the models, and therefore the estimation of the associated gravity correction, is limited by the vertical resolution of the models cells at the level of the water table. Where cell thickness at the water table is coarse (10 m), the apparent water table elevation in the model can jump as cells cross a threshold between saturated and unsaturated conditions. The associated gravity correction (Equation 1) is sensitive to this artefact and direct comparison with the distributed saturation calculation is not appropriate. The result of this calculation is only displayed where the modelled water table coincides with finer cell thickness (2 m). In these cases it is possible to compare the axis-symmetric distributed gravity calculation with the water table correction (Equation 1). For the flat topography models (*flat*-type) the distributed gravity calculation matches the water table correction well (e.g P1 and P2 in Figs. 7b). However, for models with topography (*topo*-type) the water table exhibits a phase-lag of ~ 5 yrs compared to the distributed gravity calculation (e.g. P1 - P3 in Fig. 8b). With this phase-lag the water table correction miss-corrects by up to 100 μGal.

4.2 Relative gravity changes

In order to correct for instrument drift and large spatial scale gravity variations, it is common to record gravity changes relative to a reference site or base station. In order to mimic this technique we calculate relative gravity signals using the most distal station (P6) as the reference site. The reference time-series is subtracted from the signal from the other stations to provide time-series of relative gravity changes. The minimum
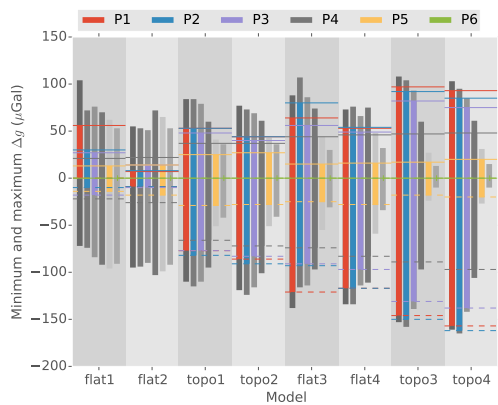
Figure 9. Minimum and maximum of simulated gravity (grey-scale) and relative gravity (colour) signals at the 6 survey locations (P1 - P6) for each model.

and maximum of each relative gravity time-series is displayed in Fig. 9 and in tabular form in Table **??**. The relative gravity time-series for *flat*2 and *topo*4 are presented in Figs 10 and 11; relative gravity time-series plots for the remaining simulations are provided in Appendix **??** (Figs. **??** - **??**).

Calculating relative gravity, with P6 as the reference station reduces the signal amplitude produced by the recharge variation for all stations in all the simulations. The effect is most significant for the flat models, with spatially uniform recharge (*flat*1 and *flat*2, e.g. Fig. 10). For *topo*3 and *topo*4 the Δg signal is low amplitude at P6 and the relative gravity reduction has limited effect on the gravity change signal for each station (e.g. Fig. 11). For these scenarios, with topography and spatially varied recharge, relative gravity changes still reach 100 - 150 μGal during the 100 year simulation.

The suite of scenarios investigated here demonstrates that when there are spatial variations in recharge and storage potential, especially between survey sites and the reference location, calculating relative gravity does not effectively remove the signal associated with temporally varied recharge. Importantly, in the models presented here, the phase of the input recharge signal is consistent across the surface. As a result, the modelled gravity changes over time at each station are also in-phase. Where spatial variations in recharge and storage potential are more complex, and the temporal variations across the survey sites are not necessarily in-phase, the calcula-

tion of relative gravity changes could introduce an erroneous signal. Spatially complex recharge and storage potential is not uncommon in high relief volcanic regions. Additionally, reference stations are often necessarily distal, or, as can be the case on small volcanic islands, they may be located on a different island with a very different hydrogeological regimes.

4.3 Implications for gravity surveys

The simulations presented here demonstrate that temporal and spatial hydrological recharge variations can generate recordable gravity changes. However, in order to assess the implications of recharge-induced signals on gravity monitoring surveys, it is important to consider the timescales of these signals relative to the timescales of the surveys.

Continuous gravity records of over long time periods are rare, but some do exist with lengths on the order of years (e.g. Budetta et al. 2004; Jousset et al. 2000). More common continuous gravity surveys in volcanic environments are at short timescales of hours to days (e.g. Gottsmann et al. 2007). The timescales of the flow investigations here do not provide insights on such short timescales. The most widely used gravity survey technique is discrete reoccupation of a survey network at intervals on the order of weeks, months and years (e.g. Carbone et al. 2003a; Hautmann et al. 2010). Such studies attempt to interpret the changes in the gravity signal that occur on the timescales of the reoccupations. Table 1 provides examples of the gravity change amplitudes recorded at volcanoes experiencing unrest (60 - 150 μGal). Also provided are the timescales that the changes are recorded over by volcano monitoring campaigns (1 - 16 years). Gottsmann et al. (2005) highlight the issue of the Nyquist frequency for dynamic gravity surveys; only signals with a period greater than two times the occupation interval can be unambiguously resolved. The propagated uncertainty of measurements from individual stations within a typical dynamic gravity volcano monitoring network can be up to 10 μGal (Battaglia et al. 2008; Hautmann et al. 2010).

To illuminate the amplitude of our modelled gravity changes, over typical field campaign intervals, we calculate changes over 1, 2, 5 and 10 year windows, rolling over every data point in the modelled gravity time series. Performing this operation, over these four campaign intervals (Δt)
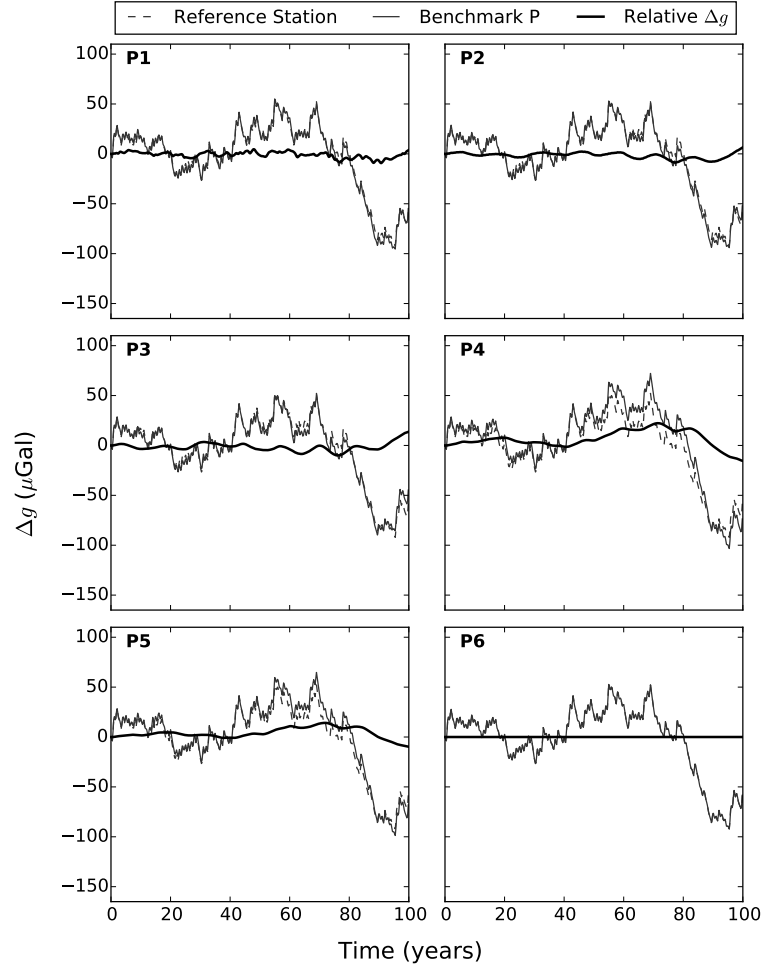
Figure 10. *flat*2 relative gravity change time-series (thick black lines). Relative gravity change at Pn ($n = 1, 2, ..., 6$) are calculated with P6 as the reference station by $\Delta g_{rel}^{Pn} = \Delta g^{Pn} - \Delta g^{P6}$. Δg^{Pn} is shown as the thin solid line on each plot; Δg^{P6}, the reference signal, is shown by the thin dashed line. As P6 is the reference station $\Delta g_{rel}^{Pn} = 0$ for that station.

on the gravity change times series ($f(t)=\Delta g(t)$) results in four discrete-time-type derivative time-series ($\frac{\Delta f(t)}{\Delta t}$) which can be described by:

$$\frac{\Delta f(t)}{\Delta t} = f(t + \Delta t/2) - f(t - \Delta t/2). \quad (9)$$

Using the discrete-time derivatives we can assess the statistical probabilities associated with the magnitude of gravity changes recorded over campaign intervals of 1, 2, 5 and 10 years. We compute distribution and exceedance probabilities for $|\Delta g|$ over these fixed time intervals.

Fig. 12 shows the derivative time-series and exceedance probability plots for model *topo*4. If two gravity readings, one year apart ($\Delta t = 1$ yr), were made at site P1 during our 100 year simulation, the probability that $|\Delta g|$ would exceed 25 μGal is 17 %. If the interval between campaigns is 10 years, the 25 μGal exceedance probability is 60 %. This analysis for model *topo*4 suggests that $|\Delta g|$ recorded at typical campaign intervals can exceed 50 μGal for 1 year, 70 μGal for 2 years, 100 μGal for 5 years and 160 μGal for 10 years. However, the majority of the changes recorded at station P1 over these windows would be 6 - 20, 10 - 30, 10 -
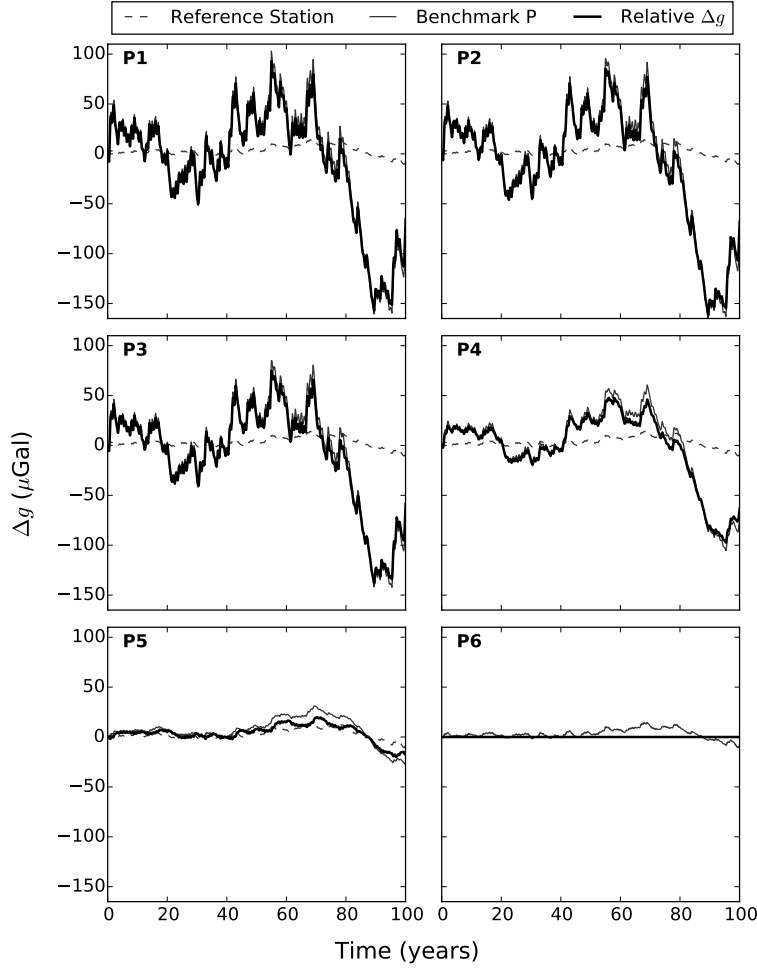
Figure 11. *topo*4 relative gravity change time-series (thick black lines). Relative gravity change at Pn ($n = 1, 2, ..., 6$) are calculated with P6 as the reference station by $\Delta g_{rel}^{Pn} = \Delta g^{Pn} - \Delta g^{P6}$. Δg^{Pn} is shown as the thin solid line on each plot; Δg^{P6}, the reference signal, is shown by the thin dashed line. As P6 is the reference station $\Delta g_{rel}^{Pn} = 0$ for that station.

45 and 15 - 60 μGal, respectively (Fig. 13). Table 4 illustrates that hydrologically induced gravity changes modelled at P1 of model *topo*4 are unlikely to produce the amplitude of gravity changes on the timescales provided as examples in Table 1. However, where the time intervals between surveys are large and the gravity changes are low there is increased potential of recording hydrologically induced gravity changes and misinterpreting them as volcanic signals. The equivalent figures for models *flat*1 to *flat*4 and models *topo*1 to *topo*3 are provided in Appendix **??**, Figs. **??** to **??**

Table 4. Probability of exceeding (Prob.Ex.) the Table 1 gravity change amplitudes under the hydrological scenario simulated at P1 in *topo*4

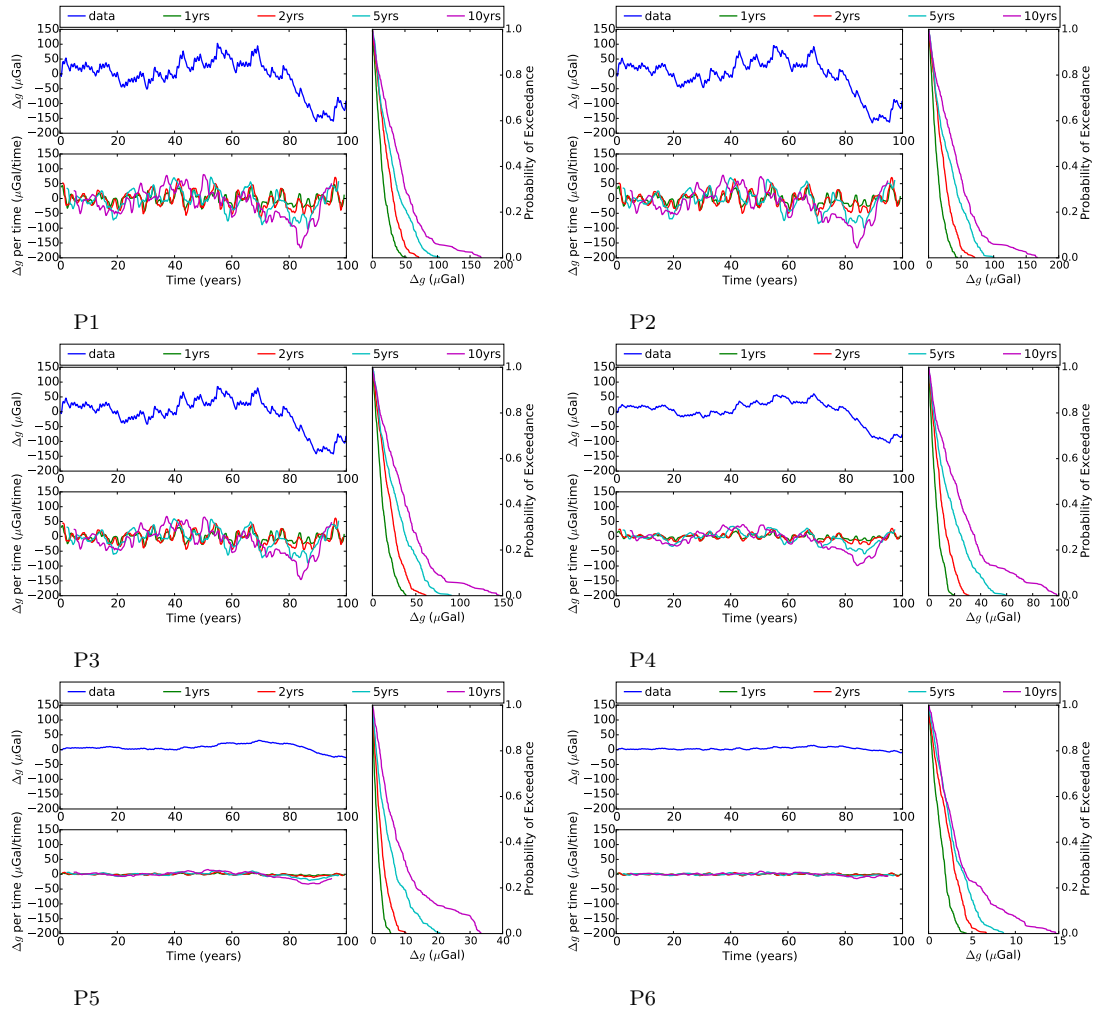| Location	Δg (μGal)	Timescale (years)	Prob.Ex. (%)
Campi Flegrei	150	2	0
Mount Etna	80	1	0
Long Valley Caldera	64	16	33
Merapi	60	1	0
Montserrat	74	1.5	0.5
Sakurajima	120	10	5
Unzen	79, 70	1, 5	0, 8.1

Figure 12. *topo*4 gravity signals over 1, 2, 5 and 10 year time-series (lower-left subplots), together with exceedance probability (right-hand subplots) for gravity changes over these campaign intervals.

The geometry, hydrology and recharge variations in the simulations presented here are based on data and observations from the sub-tropical volcanic island of Montserrat. The results clearly have implications for gravity surveys of a number of volcanic systems with similar recharge regimes to Montserrat. Our results suggest that time-lapse gravity surveys are still a relevant and valuable volcano monitoring tool. However, we demonstrate the importance of understanding the hydrological dynamics when interpreting gravity signals, especially where large temporal and spatial recharge variations exist. The magnitude of inter-annual recharge variations have a critical control on the amplitude of hydrologically induced Δg signals. An additional suite of models has shown that hydraulic properties also exert a control on the amplitude of water-mass storage variations resulting from fluctuating recharge (Hemmings 2014). For example, lower permeability slows the transfer of recharging groundwater, producing low frequency but large amplitude Δg signals. Conversely, porosity has the opposite effect; lower porosity reduces storage capacity but also promotes rapid saturation variations, and therefore more rapid water migration through the vadose zone, producing lower amplitude but higher frequency Δg signals.
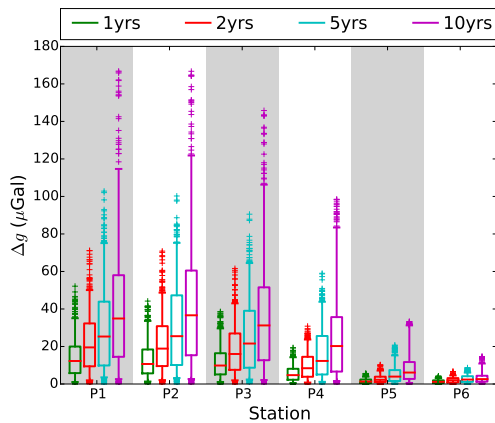
Figure 13. Boxplot of gravity changes over 1, 2, 5 and 10 year campaign intervals for model *topo*4. The boxes cover the second and third quartile, with the median gravity change is marked by a red line. Whiskers covers the range between the 5th and 95th percentile. Outliers, marked by crosses are outside this range.

4.4 Dynamic gravity survey recommendations

The insights gained from the numerical simulations presented here allow us to provide some recommendations for performing and interpreting dynamic gravity surveys in areas where spatial and temporal hydrological variations are likely to be present.

(i) Collect data on water table elevation and rainfall variability.

(ii) Incorporate continuous gravimeter data, where possible, to test reference station stability and response to rainfall fluctuations.

(iii) Perform frequent reoccupations and attempt to illuminate relationships between Δg and seasonal/inter-annual rainfall variation.

(iv) If possible, perform numerical fluid-flow to Δg simulations, such as those presented here, to assess the potential magnitude of hydrologically induced Δg signals, and ideally correct for distributed water-mass storage changes above and below the water table. A priori information on the distribution of hydrogeological heterogeneity can and should be incorporated into the flow models.

4.5 Broader relevance for geophysical monitoring

The combined fluid-flow and gravity simulation results presented here also have implications for other geophysical volcano monitoring techniques that are sensitive to hydrological flow and fluctuations. Saturation has a strong control on electrical properties of the subsurface. A common assumption is that electrical resistivity is proportional to S_l^{-n}, following Archie's law (Archie 1942). Although a number of studies have demonstrated that the relationship between electrical resistivity and saturation is more complicated, order of magnitude scale resistivity changes are regularly observed over the saturation ranges produced by our models (e.g. Frohlich & Parke 1989; Knight 1991; Khalil & Monterio Santos 2009).

Mass storage variations, such as those presented in these models, also produce geomechanical loading effects that can generate geodetic deformation signals. Hydraulic (un)loading effects related to seasonal, and longer term variations in rainfall, snowfall, snowpack thickness, and groundwater abstraction have been implicated in fluctuating GPS signals recorded in California (Argus et al. 2014; Amos et al. 2014) and the Cascades (Fu et al. 2015). Equivalent water table changes on the order of 0.5 m, similar to our simulations, have been linked to vertical ground surface oscillations on the order of 10 mm in the Californian mountains (Argus et al. 2014). Whether related to stress field changes associated with seasonal loading or lubrication of faults by groundwater, seasonal seismicity has also been reported in the literature (e.g Jiménez & Garca-Fernández 2000). Regular seasonal geophysical signals are relatively easy to identify and can be filtered out of monitoring data. However, the results we present here suggest that inter-annual recharge variations can generate higher amplitude saturation fluctuations. The geophysical signals associated with these inter-annual fluctuations are not necessarily predictably cyclical. Such signals are therefore less easily identified in monitoring data and can be difficult to remove using standard signal processing techniques alone.

When interpreting geophysical monitoring signals, care should be taken to ensure that non-cyclical hydrological fluctuations are not misinterpreted as geological or volcanic signals. This study indicates that the magnitude of a hydrological contribution to a geophysical monitoring signal will be site-specific; dependent on local cli-

mate and weather variations, hydrogeology, and the amplitude of the geophysical signal of interest. Incorporating numerical flow modelling into geophysical monitoring campaigns may be useful for assessing the likelihood of hydrological contamination of a particular geophysical monitoring record.

5   CONCLUSIONS

The flow simulation scenarios investigated here demonstrate that variations in groundwater recharge can generate changes in subsurface mass distributions that are large enough to produce recordable gravity signals. Inter-annual recharge variations dominate modelled gravity changes, with signals on the order of 10s - 100s of μGal. Gravity changes result from mass (saturation) changes within the vadose zone as well as water table elevation changes. Spatial variations in recharge can generate significant spatial differences in recorded gravity changes, which are sufficient to produce artefacts in relative gravity measurements.

For the 100-year varied recharge simulations explored here, the probability of recording exceptionally large gravity changes between typical monitoring campaigns is relatively low. However, we have demonstrated that recordable gravity signals, on the order of 10s of μGal can be generated by simple and not atypical variations in recharge. This demonstrates the importance of understanding spatial and temporal groundwater behaviour when interpreting gravity signals in monitoring locations where the hydrology is dynamic.

Numerical fluid-flow models present a valuable tool for investigating and estimating the effect of hydrological processes on gravity measurements. As such, the simulations and work flow presented here can provide a method for assessing the potential for hydrologically generated gravity signals and correcting for them, where necessary. The approach presented here can also be adapted to explore potential hydrological contributions to other geophysical monitoring signals.

REFERENCES

Amos, C. B., Audet, P., Hammond, W. C., Bürgmann, R., Johanson, I. a., & Blewitt, G., 2014. Uplift and seismicity driven by groundwater depletion in central California, *Nature*, **509**(7501), 483–486.

Archie, G., 1942. The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics, *Trans. AIME*, **146**(01), 54–62.

Argus, D. F., Fu, Y., & Landerer, F. W., 2014. Seasonal variation in total water storage in California inferred from GPS observations of vertical land motion, *Geophys. Res. Lett.*, **41**(6), 1971–1980.

Battaglia, M., 1999. Magma Intrusion Beneath Long Valley Caldera Confirmed by Temporal Changes in Gravity, *Science (80-.)*., **285**(5436), 2119–2122.

Battaglia, M., Segall, P., & Roberts, C., 2003. The mechanics of unrest at Long Valley caldera, California. 2. Constraining the nature of the source using geodetic and micro-gravity data, *J. Volcanol. Geotherm. Res.*, **127**(3-4), 219–245.

Battaglia, M., Troise, C., Obrizzo, F., Pingue, F., & De Natale, G., 2006. Evidence for fluid migration as the source of deformation at Campi Flegrei caldera (Italy), *Geophys. Res. Lett.*, **33**(1), L01307.

Battaglia, M., Gottsmann, J., Carbone, D., & Fernández, J., 2008. 4D volcano gravimetry, *Geophysics*, **73**(6), WA3–WA18.

Berrino, G., 1994. Gravity changes induced by height-mass variations at the Campi Flegrei caldera, *J. Volcanol. Geotherm. Res.*, **61**(3-4), 293–309.

Berrino, G., Corrado, G., Luongo, G., & Toro, B., 1984. Ground deformation and gravity changes accompanying the 1982 Pozzuoli uplift, *Bull. Volcanol.*, **47**(2), 187–200.

Bianco, F., Del Pezzo, E., Saccorotti, G., & Ventura, G., 2004. The role of hydrothermal fluids in triggering the JulyAugust 2000 seismic swarm at Campi Flegrei, Italy: evidence from seismological and mesostructural data, *J. Volcanol. Geotherm. Res.*, **133**(1-4), 229–246.

Budetta, G., Carbone, D., & Greco, F., 1999. Subsurface mass redistributions at Mount Etna (Italy) during the 1995-1996 explosive activity detected by microgravity studies, *Geophys. J. Int.*, **138**(1), 77–88.

Budetta, G., Carbone, D., Greco, F., & Rymer, H., 2004. Microgravity studies at Mount Etna (Italy), in Mt. Etna Volcano Lab., pp. 221–240, eds Bonaccorso, A., Calvari, S., Coltelli, M., Del Negro, C., & Falsaperla, S., American Geophysical Union, Washington, D. C.

Carbone, D., Budetta, G., & Greco, F., 2003a. Bulk processes prior to the 2001 Mount Etna eruption, highlighted through microgravity studies, *J. Geophys. Res.*, **108**(B12), 2556.

Carbone, D., Budetta, G., & Greco, F., 2003b. Possible mechanisms of magma redistribution under Mt Etna during the 1994-1999 period detected through microgravity measurements, *Geophys. J. Int.*, **153**(1), 187–200.

Chouet, B., 2003. Volcano Seismology, *Pure Appl. Geophys.*, **160**(3), 739–788.

Creutzfeldt, B., Guntner, A., Klugel, T., & Wziontek, H., 2008. Simulating the influence of water storage changes on the superconducting gravimeter of the Geodetic Observatory Wettzell, Germany, *Geophysics*, **73**(6), WA95.

Creutzfeldt, B., Ferre, T., Troch, P., Merz, B., Wziontek, H., & G??ntner, A., 2012. Total water storage dynamics in response to climate variability and extremes: Inference from long-term terrestrial gravity measurement, *J. Geophys. Res. Atmos.*, **117**(8), 1–17.

Crossley, D., Xu, S., & Van Dam, T., 1998. Comprehensive analysis of 2 years of SG data from Table Mountain, Colorado, *Proc. 13th Int. Symp. Earth Tides, Brussels*, pp. 659–668.

Dzurisin, D., 2003. A comprehensive approach to monitoring volcano deformation as a window on the eruption cycle, *Rev. Geophys.*, **41**(1).

Eggers, A. A., 1983. Temporal gravity and elevation changes at Pacaya volcano, Guatemala, *J. Volcanol. Geotherm. Res.*, **19**(3-4), 223–237.

Frohlich, R. K. & Parke, C. D., 1989. The Electrical Resistivity of the Vadose Zone - Field Survey, *Ground Water*, **27**(4), 524–530.

Fu, Y., Argus, D. F., & Landerer, F. W., 2015. GPS as an independent measurement to estimate terrestrial water storage variations in Washington and Oregon, *J. Geophys. Res. Solid Earth*, **120**(1), 552–566.

Gottsmann, J., Rymer, H., & Wooller, L. K., 2005. On the interpretation of gravity variations in the presence of active hydrothermal systems: Insights from the Nisyros Caldera, Greece, *Geophys. Res. Lett.*, **32**(23), L23310.

Gottsmann, J., Camacho, A. G., Tiampo, K. F., & Fernández, J., 2006a. Spatiotemporal variations in vertical gravity gradients at the Campi Flegrei caldera (Italy): a case for source multiplicity during unrest?, *Geophys. J. Int.*, **167**(3), 1089–1096.

Gottsmann, J., Wooller, L., Martí, J., Fernández, J., Camacho, a. G., Gonzalez, P. J., Garcia, A., & Rymer, H., 2006b. New evidence for the reawakening of Teide volcano, *Geophys. Res. Lett.*, **33**(20), L20311.

Gottsmann, J., Carniel, R., Coppo, N., Wooller, L., Hautmann, S., & Rymer, H., 2007. Oscillations in hydrothermal systems as a source of periodic unrest at caldera volcanoes: Multiparameter insights from Nisyros, Greece, *Geophys. Res. Lett.*, **34**(7), L07307.

Hautmann, S., Gottsmann, J., Sparks, R. S. J., Costa, A., Melnik, O., & Voight, B., 2009. Modelling ground deformation caused by oscillating overpressure in a dyke conduit at Soufrière Hills Volcano, Montserrat, *Tectonophysics*, **471**(1-2), 87–95.

Hautmann, S., Gottsmann, J., Camacho, A. G., Fournier, N., Sacks, I. S., & Sparks, R. S. J., 2010. Mass variations in response to magmatic stress changes at Soufrière Hills Volcano, Montserrat (W.I.): Insights from 4-D gravity data, *Earth Planet. Sci. Lett.*, **290**(1-2), 83–89.

Hautmann, S., Camacho, A. G., Gottsmann, J., Odbert, H. M., & Syers, R. T., 2013. The shallow structure beneath Montserrat (West Indies) from new Bouguer gravity data, *Geophys. Res. Lett.*, **40**(19), 5113–5118.

Hector, B., Séguis, L., Hinderer, J., Cohard, J. M., Wubda, M., Descloitres, M., Benarrosh, N., & Boy, J. P., 2015. Water storage changes as a marker for base flow generation processes in a tropical humid basement catchment (Benin): Insights from hybrid gravimetry, *Water Resour. Res.*, **51**(10), 8331–8361.

Hemmings, B., 2014. *Field and numerical studies of the hydrogeology of an island arc volcano: Insights from Montserrat, BWI*, Phd thesis, University of Bristol.

Hemmings, B., Whitaker, F., Gottsmann, J., & Hughes, A., 2015. Hydrogeology of Montserrat review and new insights, *J. Hydrol. Reg. Stud.*, **3**, 1–30.

Hughes, A. G., Mansour, M. M., & Robins, N. S., 2008. Evaluation of distributed recharge in an upland semi-arid karst system: the West Bank Mountain Aquifer, Middle East, *Hydrogeol. J.*, **16**(5), 845–854.

Hurwitz, S., Kipp, K., Ingebritsen, S., & Reid, M. E., 2003. Groundwater flow, heat transport, and water table position within volcanic edifices: Implications for volcanic processes in the Cascade Range, *J. Geophys. Res.*, **108**(B12), 2557.

Ingebritsen, S. & Scholl, M., 1993. The hydrogeology of Kilauea volcano, *Geothermics*, **22**(4), 255–270.

Jachens, R. C. & Roberts, C. W., 1985. Temporal and areal gravity investigations at Long Val-

ley Caldera, California, *J. Geophys. Res.*, **90**(B13), 11210.

Jacob, T., Chery, J., Bayer, R., Le Moigne, N., Boy, J.-P., Vernant, P., & Boudin, F., 2009. Time-lapse surface to depth gravity measurements on a karst system reveal the dominant role of the epikarst as a water storage entity, *Geophys. J. Int.*, **177**(2), 347–360.

Jacob, T., Bayer, R., Chery, J., & Le Moigne, N., 2010. Time-lapse microgravity surveys reveal water storage heterogeneity of a karst aquifer, *J. Geophys. Res.*, **115**(B6), B06402.

Jiménez, M.-J. & Garca-Fernández, M., 2000. Occurrence of shallow earthquakes following periods of intense rainfall in Tenerife, Canary Islands, *J. Volcanol. Geotherm. Res.*, **103**(1-4), 463–468.

Jousset, P., Dwipa, S., Beauducel, F., Duquesnoy, T., & Diament, M., 2000. Temporal gravity at Merapi during the 19931995 crisis: an insight into the dynamical behaviour of volcanoes, *J. Volcanol. Geotherm. Res.*, **100**(1-4), 289–320.

Khalil, M. A. & Monterio Santos, F. A., 2009. Influence of degree of saturation in the electric resistivity-hydraulic conductivity relationship, *Surv. Geophys.*, **30**(6), 601–615.

Klavetter, E. & Peters, R., 1986. Estimation of hydrologic properties of an unsaturated, fractured rock mass, Tech. rep., Sandia National Laboratories, Albuquerque, New Mexico.

Knight, R., 1991. Hysteresis in the electrical resistivity of partially saturated sandstones, *Geophysics*, **56**(12), 2139.

Krause, P., Naujoks, M., Fink, M., & Kroner, C., 2009. The impact of soil moisture changes on gravity residuals obtained with a superconducting gravimeter, *J. Hydrol.*, **373**(1-2), 151–163.

Leirião, S., He, X., Christiansen, L., Andersen, O. B., & Bauer-Gottwein, P., 2009. Calculation of the temporal gravity variation from spatially variable water storage change in soils and aquifers, *J. Hydrol.*, **365**(3-4), 302–309.

Luckner, L., Van Genuchten, M. T., & Nielsen, D. R., 1989. A consistent set of parametric models for the two-phase flow of immiscible fluids in the subsurface, *Water Resour. Res.*, **25**(10), 2187–2193.

McNutt, S. R., 2005. Volcanic Seismology, *Annu. Rev. Earth Planet. Sci.*, **33**(1), 461–491.

Mouyen, M., Masson, F., Hwang, C., Cheng, C. C., Moigne, N. L., Lee, C. W., Kao, R., & Hsieh, W. C., 2013. Erosion effects assessed by repeated gravity measurements in southern Taiwan, *Geophys. J. Int.*, **192**(1), 113–136.

Muller, C., del Potro, R., Biggs, J., Gottsmann, J., Ebmeier, S. K., Guillaume, S., Cattin, P.-H., & Van der Laat, R., 2015. Integrated velocity field from ground and satellite geodetic techniques: application to Arenal volcano, *Geophys. J. Int.*, **200**(2), 861–877.

Nagy, D., 1966. The Gravitational Attraction of a Right Rectangular Prism, *GEOPHYSICS*, **31**(2), 362–371.

Naujoks, M., Kroner, C., Weise, a., Jahr, T., Krause, P., & Eisner, S., 2010. Evaluating local hydrological modelling by temporal gravity observations and a gravimetric three-dimensional model, *Geophys. J. Int.*, **182**, 233–249.

Parker, A. L., Biggs, J., & Lu, Z., 2014. Investigating long-term subsidence at Medicine Lake Volcano, CA, using multitemporal InSAR, *Geophys. J. Int.*, **199**(2), 844–859.

Pearson, S. C. P., Kiyosugi, K., Lehto, H. L., Saballos, J. a., Connor, C. B., & Sanford, W. E., 2012. Integrated geophysical and hydrothermal models of flank degassing and fluid flow at Masaya volcano, Nicaragua, *Geochemistry Geophys. Geosystems*, **13**(5), Q05011.

Pfeffer, J., Champollion, C., Favreau, G., Cappelaere, B., Hinderer, J., Boucher, M., Nazoumou, Y., Oï, M., Mouyen, M., Henri, C., Le Moigne, N., Deroussi, S., Demarty, J., Boulain, N., Benarrosh, N., & Robert, O., 2013. Evaluating surface and subsurface water storage variations at small time and space scales from relative gravity measurements in semiarid Niger, *Water Resour. Res.*, **49**(6), 3276–3291.

Piccolroaz, S., Majone, B., Palmieri, F., Cassiani, G., & Bellin, A., 2015. On the use of spatially distributed, time-lapse microgravity surveys to inform hydrological modeling, *Water Resour. Res.*, **51**(9), 7270–7288.

Pruess, K., 2004. The TOUGH Codes–A Family of Simulation Tools for Multiphase Flow and Transport Processes in Permeable Media, *Vadose Zo. J.*, **3**(3), 738–746.

Rymer, H., 1994. Microgravity change as a precursor to volcanic activity, *J. Volcanol. Geotherm. Res.*, **61**(3-4), 311–328.

Saar, M. O. & Manga, M., 2004. Depth dependence of permeability in the Oregon Cascades inferred from hydrogeologic, thermal, seismic, and magmatic modeling constraints, *J. Geophys. Res.*, **109**(B4), B04204.

Saibi, H., Gottsmann, J., & Ehara, S., 2010. Posteruptive gravity changes from 1999 to 2004 at Unzen volcano (Japan): A window into shallow aquifer and hydrothermal dynamics, *J. Volcanol. Geotherm. Res.*, **191**(1-2), 137–147.

Sanderson, T. J. O., 1982. Direct gravimetric detection of magma movements at Mount Etna, *Nature*, **297**(5866), 487–490.

Todesco, M., 2009. Signals from the Campi Flegrei hydrothermal system: Role of a magmatic source of fluids, *J. Geophys. Res.*, **114**(B5), B05201.

Todesco, M. & Berrino, G., 2005. Modeling hydrothermal fluid circulation and gravity signals at the Phlegraean Fields caldera, *Earth Planet. Sci. Lett.*, **240**(2), 328–338.

Todesco, M., Rinaldi, A. P., & Bonafede, M., 2010.

Modeling of unrest signals in heterogeneous hydrothermal systems, *J. Geophys. Res.*, **115**(B9), B09213.

Webb, S. W., 2000. A simple extension of two-phase characteristic curves to include the dry region, *Water Resour. Res.*, **36**(6), 1425–1430.

Yokoyama, I., 1989. Microgravity and height changes caused by volcanic activity: Four Japanese examples, *Bull. Volcanol.*, **51**(5), 333–345.

APPENDIX A: ADDITIONAL TABLES AND FIGURES FOR GRAVITY MODELS

APPENDIX B: ADDITIONAL RELATIVE GRAVITY TABLES AND FIGURES

APPENDIX C: ADDITIONAL CAMPAIGN INTERVALS FIGURES

Table A1. Minimum and maximum gravity changes simulated at P1 - P6 for the 8 model scenarios simulated here.

Simulation	P1 Δg μGal		P2 Δg μGal		P3 Δg μGal		P4 Δg μGal		P5 Δg μGal		P6 Δg μGal	
	min.	max.	min.	max.	min.	max.	min.	max.	min.	max.	min.	max.
*flat*1	-72	104	-74	72	-84	76	-92	70	-96	62	-91	53
*flat*2	-95	55	-94	53	-90	51	-103	72	-99	65	-92	53
*topo*1	-110	84	-115	84	-110	79	-95	60	-51	41	-42	36
*topo*2	-119	77	-124	73	-116	69	-101	61	-51	43	-41	36
*flat*3	-138	88	-116	107	-114	86	-97	74	-55	44	-31	30
*flat*4	-134	73	-134	76	-114	66	-111	75	-59	48	-34	32
*topo*3	-153	108	-158	104	-139	93	-97	60	-24	27	-10	13
*topo*4	-161	103	-165	95	-142	85	-106	61	-27	31	-10	15

Table B1. Minimum and maximum relative gravity changes simulated at P1 - P6 for the 8 model scenarios simulated here. With P6 as the reference station.

Simulation	P1 Δg μGal		P2 Δg μGal		P3 Δg μGal		P4 Δg μGal		P5 Δg μGal		P6 Δg μGal	
	min.	max.	min.	max.	min.	max.	min.	max.	min.	max.	min.	max.
*flat*1	0	56	-10	30	-18	27	-22	21	-14	-13	0	0
*flat*2	-9	7	-9	8	-10	14	-26	22	-18	14	0	0
*topo*1	-77	53	-82	53	-77	48	-66	37	-29	25	0	0
*topo*2	-86	44	-91	44	-83	40	-72	37	-28	27	0	0
*flat*3	-121	64	-93	80	-91	56	-74	44	-25	15	0	0
*flat*4	-117	53	-117	54	-97	49	-83	46	-28	16	0	0
*topo*3	-146	97	-150	92	-131	82	-89	47	-18	17	0	0
*topo*4	-157	93	-162	85	-138	75	-97	48	-20	20	0	0