CUMULATIVE DISSERTATION

Advances in hydrogravimetry

Terrestrial gravimeters as field tools for hydrological applications

by Marvin Reich

A thesis presented for the degree of
Doctor of Natural Sciences (Dr. rer. nat.)

Supervisor:

Andreas Güntner   University of Potsdam and GFZ, Germany

Referees:

Jacques Hinderer   University of Strasbourg, France
Peter Bauer-Gottwein   Technical University of Denmark, Denmark

Summary

The interest of the hydrological community in the gravimetric method has steadily increased within the last decade. This is reflected by numerous studies from many different groups with a broad range of approaches and foci. Many of those are traditionally rather hydrology-oriented groups who recognized gravimetry as a potential added value for their hydrological investigations. While this resulted in a variety of interesting and useful findings, contributing to extend the respective knowledge and confirming the methodological potential, on the other hand, many interesting and unresolved questions emerged. This thesis manifests efforts, analyses and solutions carried out in this regard. Addressing and evaluating many of those unresolved questions, the research contributes to advancing hydrogravimetry, the combination of gravimetric and hydrological methods, in showing how gravimeters are a highly useful tool for applied hydrological field research.

In the first part of the thesis, traditional setups of stationary terrestrial superconducting gravimeters are addressed. They are commonly installed within a dedicated building, the impermeable structure of which shields the underlying soil from natural exchange of water masses (infiltration, evapotranspiration, groundwater recharge). As gravimeters are most sensitive to mass changes directly beneath the meter, this could impede their suitability for local hydrological process investigations, especially for near-surface water storage changes (WSC). By studying temporal local hydrological dynamics at a dedicated site equipped with traditional hydrological measurement devices, both below and next to the building, the impact of these absent natural dynamics on the gravity observations were quantified. A comprehensive analysis with both a data-based and model-based approach led to the development of an alternative method for dealing with this limitation. Based on determinable parameters, this approach can be transferred to a broad range of measurement sites where gravimeters are deployed in similar structures. Furthermore, the extensive considerations on this topic enabled a more profound understanding of this so called *umbrella effect*.

The second part of the thesis is a pilot study about the field deployment of a superconducting gravimeter. A newly developed field enclosure for this gravimeter was tested in an outdoor installation adjacent to the building used to investigate the umbrella effect. Analyzing and comparing the gravity observations from both indoor and outdoor gravimeters showed performance with respect to noise and stable environmental conditions was equivalent while the sensitivity to near-surface WSC was highly increased for the field deployed instrument. Furthermore it was demonstrated that the latter setup showed gravity changes independent of the depth where mass changes occurred, given their sufficiently wide horizontal extent. As a consequence, the field setup suits monitoring of WSC for both short and longer time periods much better. Based on a coupled data-modeling approach, its gravity time

series was successfully used to infer and quantify local water budget components (evapotranspiration, lateral subsurface discharge) on the daily to annual time scale.

The third part of the thesis applies data from a gravimeter field deployment for applied hydrological process investigations. To this end, again at the same site, a sprinkling experiment was conducted in a 15 x 15 m area around the gravimeter. A simple hydro-gravimetric model was developed for calculating the gravity response resulting from water redistribution in the subsurface. It was found that, from a theoretical point of view, different subsurface water distribution processes (macro pore flow, preferential flow, wetting front advancement, bypass flow and perched water table rise) lead to a characteristic shape of their resulting gravity response curve. Although by using this approach it was possible to identify a dominating subsurface water distribution process for this site, some clear limitations stood out. Despite the advantage for field installations that gravimetry is a non-invasive and integral method, the problem of non-uniqueness could only be overcome by additional measurements (soil moisture, electric resistivity tomography) within a joint evaluation. Furthermore, the simple hydrological model was efficient for theoretical considerations but lacked the capability to resolve some heterogeneous spatial structures of water distribution up to a needed scale. Nevertheless, this unique setup for plot to small scale hydrological process research underlines the high potential of gravimetery and the benefit of a field deployment.

The fourth and last part is dedicated to the evaluation of potential uncertainties arising from the processing of gravity observations. The gravimeter senses all mass variations in an integral way, with the gravitational attraction being directly proportional to the magnitude of the change and inversely proportional to the square of the distance of the change. Consequently, all gravity effects (for example, tides, atmosphere, non-tidal ocean loading, polar motion, global hydrology and local hydrology) are included in an aggregated manner. To isolate the signal components of interest for a particular investigation, all non-desired effects have to be removed from the observations. This process is called *reduction*. The large-scale effects (tides, atmosphere, non-tidal ocean loading and global hydrology) cannot be measured directly and global model data is used to describe and quantify each effect. Within the reduction process, model errors and uncertainties propagate into the residual, the result of the reduction. The focus of this part of the thesis is quantifying the resulting, propagated uncertainty for each individual correction. Different superconducting gravimeter installations were evaluated with respect to their topography, distance to the ocean and the climate regime. Furthermore, different time periods of aggregated gravity observation data were assessed, ranging from 1 hour up to 12 months. It was found that uncertainties were highest for a frequency of 6 months and smallest for hourly frequencies. Distance to the ocean influences the uncertainty of the non-tidal ocean loading component, while geographical latitude affects uncertainties of the global hydrological component. It is important to highlight that the resulting correction-induced uncertainties in the residual have the potential to mask the signal of interest, depending on the signal magnitude and its frequency. These findings can be used to assess the value of gravity data across a range of applications and geographic settings.

In an overarching synthesis all results and findings are discussed with a general focus on their added value for bringing hydrogravimetric field research to a new level. The conceptual and applied methodological benefits for hydrological studies

are highlighted. Within an outlook for future setups and study designs, it was once again shown what enormous potential is offered by gravimeters as hydrological field tools.

Zusammenfassung
(in allgemeinverständlicher Sprache)

Gravimetrie ist eine geophysikalische Methode, bei der Massen und deren Veränderungen beobachtet und gemessen werden. Die Messgeräte der Gravimetrie heißen Gravimeter. Wenn man diese Methode in der Erforschnug von Wasser-relevanten Fragestellungen, Prozessen und Zustände einsetzt (Hydrologie), spricht man auch von *Hydrogravimetrie*. Die vorliegende Disseration beschäftigt sich damit wie diese hydrogravimetrische Methode für angewandte Forschung im Feld benutzt wird und weiterentwickelt werden kann.

Zuerst wird thematisiert, wie konventionelle Aufbauten mit Gravimetern aussehen und was daran aus der hydrologischen Perspektive problematisch ist. Das Gebäude in dem sich das Gravimeter befindet, stellt eine große versiegelte Fläche dar, die es verhindert, dass in der direktem Umgebung natürliche Prozesse ablaufen. Das ist so problematisch, weil das Gravimeter besonders empfindlich auf Massänderungen in nächster räumlicher Nähe reagiert. Als Lösung wird mit Hilfe einer neuen Methode aufgezeigt, wie man unter Benutzung von traditionellen hydrologischen Messinstrumenten um das Gebäude herum diese verhinderten natürlichen Prozesse beschreiben kann.

Darauf folgend wird anhand eines erfolgreich getesteten Aufbaus eines Gravimeters außerhalb von einem Gebäude, also direkt im Gelände, demonstriert, was solch eine Außeninstallation für einen großen Vorteil für die hydrologische Feldforschung mit sich bringt. Darüberhinaus wird gezeigt, dass dieser alternative Aufbau keinerlei Nachteile hinsichtlich Genauigkeit, Qualität, Rauschen oder Beherschbarkeit von Umwelteinflüssen mit sich bringt, sondern vor allem die Empfindlichkeit für Messungen von Wassermassenänderungen in Oberflächennähe stark verbessert.

Anhand eines Beregnungsexperiments auf der Fläche um dieses im Gelände installierten Gravimeters, werden die Vorzüge der gravimetrischen Methode für die hydrologische Prozessforschung aufgezeigt. Verschiedene mögliche Ausbreitungen des verregneten Wassers im Untergrund können mittels dieser Methode charakterisiert und identifiziert werden.

Im letzen Teil wird das Problem von Unsicherheiten besprochen, die aus der notwendigen Datenbearbeitung resultieren. Um die gravimetrischen Beobachtungen auf die Anteile zu reduzieren, die innerhalb einer Studie betrachtet werden sollen, müssen alle Komponenten die das Gravimeter misst, die aber die hydrologische Interpretation stören, beseitigt werden. Dabei handelt es sich vor allem um globale Komponenten wie Gezeiten, Luftdruckschwankungen, Gezeiten-unabhägige Meeresströmungen und globale Hydrologie. Es wird untersucht, welche Unsicherheiten bei deren Korrektur auftreten, wenn verschiedene Zeitintervalle von zu beobachtenden hydrologischen Signalen vorherrschen.

Alle gewonnenen Resulte und Erfahrungen werden in einer gesamtheitlichen Be-

trachtung dahingegehend diskutiert, wie die hydrogravimetrische Methode aufgrund dieser neuen Erkenntnisse verbessert und vorangebracht werden konnte.

Contents

List of Publications

M. Reich, M. Mikolaj, T. Blume, and A. Güntner. 2019. Reducing gravity data for the influence of water storage variations beneath observatory buildings. *Geophysics*, 84:EN15–EN31. doi: 10.1190/geo2018-0301.1

Author's contribution: Concept of study, programming, data processing and analysis, setup and evaluation of hydrological model, preparation and writing of the manuscript, corresponding author for publishing procedure.

A. Güntner, M. Reich, M. Mikolaj, B. Creutzfeldt, S. Schroeder, and H. Wziontek. 2017. Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure. *Hydrology and Earth System Sciences*, 21(6):3167–3182. doi: 10.5194/hess-21-3167-2017

Author's contribution: Concept of study (jointly with Güntner and Mikolaj), instrumental field setup and fieldwork, partially data preparation and processing, analysis of results (jointly with Güntner and Mikolaj), preparing and writing parts of the manuscript, involved in publishing procedure.

M. Reich, M. Mikolaj, T. Blume, and A. Güntner. 2021b. Field-Scale Subsurface Flow Processes Inferred From Continuous Gravity Monitoring During a Sprinkling Experiment. *Water Resources Research*, 57(10):e2021WR030044. doi: https://doi.org/10.1029/2021WR030044

Author's contribution: Concept of study (jointly with Güntner, Blume and Mikolaj), design and installation of field setup (jointly with Güntner and Mikolaj), field measurements (jointly with Güntner and Mikolaj), development and programming of hydro-gravitational model, data processing and analysis, preparation and writing of the manuscript, corresponding author for publishing procedure.

M. Mikolaj, M. Reich, and A. Güntner. 2019b. Resolving Geophysical Signals by Terrestrial Gravimetry: A Time Domain Assessment of the Correction-Induced Uncertainty. *Journal of Geophysical Research: Solid Earth*, 124:2153–2165. doi: 10.1029/2018jb016682

Author's contribution: Concept of study (jointly with Mikolaj), partially data preparation and processing, analysis of results (jointly with Güntner and Mikolaj), preparing and writing parts of the manuscript, involved in publishing procedure.

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Gravimetry: concept and method

Gravitation is one of the four fundamental forces. It is a natural phenomena which describes the attraction of masses and is present everywhere in the universe. On Earth, all objects are attracted by and drawn towards its core, which constitutes its biggest and most condensed mass.

After the famous but unproven experiments by Galileo Galilei [Viviani, 1717], dropping objects from the tower of Pisa, it was Sir Isaac Newton who combined theoretical work by Galilei and Kepler with his own mathematical findings into the first modern theory of gravitation [Newton, 1687]. Newton states with the law of gravitation, that two point masses are attracted by each other due to the gravitational force:

$$F = G \times \frac{m_1 \times m_2}{R^2} \tag{1.1}$$

with G as the gravitational constant ($6.673 10^{-11} \frac{N \times m^2}{kg^2}$), m_1 and m_2 two masses with a distance R between them.

When combining Eq.1.1 with Newtons law of motion:

$$F = m \times a \tag{1.2}$$

we can describe the gravitational attraction at a certain distance R resulting from mass m to be:

$$a = G \times \frac{m}{R^2} \tag{1.3}$$

When using the mass of the earth as m and its radius as R, we obtain the gravity on the shell of a perfect sphere due to the earths core in its center. This concept implies perfectly concentrated mass in points with no other matter in between or around it. It was only a matter of time before these equations were used to fit the needs for applied geodetic science in smaller regions of the earth.

For local to regional scales, it is often convenient to grid, or discretize, the area of interest. In the case of terrestrial gravity, the interest is generally to validate or describe the gravitational attraction of all surrounding masses at a single point. At this single point, usually a gravimeter is situated. A gravimeter is an instrument used to measure gravity, equal to the sum of the vertical component of all gravitational

attractions at the sensor location (gravimeters are only able to measure the vertical component of gravity, which is almost always several orders of magnitude larger than the horizontal component). An overview of different gravimeters is given in chapter 1.1.2. To calculate the gravity effect of mass (or mass change) within a constrained volume, e.g. one grid cell, Forsberg [1984] developed an analytical solution in relation to Eq.1.3:

$$\Delta g = \Delta \rho \times G \times \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} \frac{(z - z_g)}{((x - x_g)^2 + (y - y_g)^2 + (z - z_g)^2)^{3/2}} \, dx \, dy \, dz \quad (1.4)$$

$\Delta \rho$ is the density change of the volume element and the gravimeter is located at (x_g, y_g, z_g). The total gravity effect of all mass changes at (x_g, y_g, z_g) is the sum of all individual grid-cells' gravity effects.

The SI unit for gravity is m/s^2. Nevertheless, for convenience and taking into account the magnitude of the phenomena of interest, instrumental precision and historic reasons, the main units found in most studies are: mGal (10^{-5} m/s^2), µGal (10^{-8} m/s^2) and nm/s^2 (10^{-9} m/s^2).

Gravity is measured with absolute and relative instruments. Absolute gravimeters measure the absolute acceleration at its position and time of measurement. In turn, relative gravimeters can only detect changes in the gravitational acceleration between two locations, or over time, depending on the type of meter. The focus of this present work is on relative gravimetry, using superconducting gravimeters. Furthermore, as satellite measurement techniques also exist, this work addresses gravity and gravimetry in its terrestrial context, if not stated otherwise. 'Terrestrial' hereby refers to in-situ measurements with gravimeters directly deployed on the surface of the earth.

1.1.1   Influences on the magnitude of gravity

The gravitational attraction is influenced by different phenomena which magnitude mainly depends on the location of the point of interest (with respect to a reference ellipsoid). They are all based on the concept of Newtonian attraction (see Eq. 1.1) and the centrifugal force due to Earth's rotation: (a) the latitude effect (due to the latitude coordinate of the location of the gravimeter), (b) terrain effects (missing mass due to valleys next to the gravimeter or excess mass due to a hill), (c) Bouguer plate effect (layer of certain thickness and density between gravity sensor and reference ellipsoid) and (d) free-air effect (due to elevation above or below reference ellipsoid). A good and detailed overview is given by Lowrie [2007]. In the context of time-variable gravity measurements, they can be neglected as they are assumed to not change in the observed temporal and spatial scales (they are relevant and important for investigations about e.g. geologic structures).

More important for relative gravimetry are influences with respect to the instrumental behavior (drift), disturbances of the point of interest (earthquakes, microseismic, surrounding noise) and other natural phenomena: tides, atmosphere, hydrology, polar motion, non-tidal ocean loading, land subsidence or uplift. These are discussed in detail in chapter 1.2.5.

1.1.2   Terrestrial gravimeter types

A brief overview of both absolute and relative gravimeter types in its historical context is given. A more comprehensive description can be found in Van Camp et al. [2017] and Crossley et al. [2013].

The first gravimeters were based on a pendulum, a concept developed already by Galilei [Chapin, 1998]. Knowing the pendulums length and weight, it was possible to calculate the gravitational attraction via measurements of its oscillation time. In the 1930's springs where introduced to measure the change in gravity, detectable via the change of the spring's extension. Much advancement has been made since the invention of the Zero-Length-Spring by LaCoste [1934] but the concept of springs for relative gravimetry is still used until today in modern gravimeters like the CG-6, gPhone or Burris gravimeter. These gravimeters can reach an accuracy of up to 10 nm/s^2 (although usually a bit lower) [Fores et al., 2019], depending mainly on measurement setup, campaign design and frequent instrument calibration. In 1990 a new gravimeter type was developed by a consortium of institutes (NIST, USA; NOAA, USA; IFAG, Germany): the FG-5 [Niebauer et al., 1995]. This absolute gravimeter uses lasers to measure the time a falling reference object (corner-cube) needs to cover a predefined vertical distance. The instrument precision is much higher (10 nm/s^2) and as a result, FG-5 measurements and comparisons are the accepted reference standard until present times. Another absolute gravimeter, developed in the late 1990's, is the A10 [Liard and Gagnon, 2002]. Despite its lower precision (100 nm/s^2), it is still used in many campaigns as it is designed to target outdoor and rather mobile measurements, as opposed to the FG-5. In 1968, a new concept for relative gravimeters was introduced by Prothero and Goodkind [1968]. They build a gravimeter based on superconductors creating a magnetic field which hosts a small, hovering test mass. Changes in gravitational attraction, in contrast to the concept of springs, are measured via the necessary voltage to keep the test mass centered. These SG's (Superconducting Gravimeters) have a much larger and complex setup than other relative gravimeters but reach an accuracy of up to 1 nm/s^2 [Hinderer et al., 2015b]. They are the standard for relative gravimeter types until today. The latest instrumental development was the introduction of quantum gravimeters (e.g. the AQG by Ixblue). These are also absolute gravimeters where the falling test mass is not a solid body but a cloud of cold atoms. The first device available commercially in 2018 reaches a precision of up to 10 nm/s^2 as stated by the manufacturer [Ménoret et al., 2018]. Nevertheless this is still under ongoing discussion and testing by the community.

Besides instrument precision and investment costs, gravimeters can be differentiated by (a) its type (absolute or relative) and (b) its application range (stationary or field surveys). Relative gravimeters have the disadvantage of instrumental drift. This means every device has an individual, usually linear deviation of its measured quantity with time. Absolute gravimeter do not show this behaviour and are therefore used to regularly estimate this drift component of the relative devices. An advantage of relative gravimeters (except Burris) is their capability to perform continuous measurements. With respect to the area of application, short term field campaigns can be conducted with the CG-6, the Burris gravimeter and the A10 [Francis, 2021; Dykowski et al., 2014]. Outdoor measurements are also possible with a variant of the SG, the iGrav [Güntner et al., 2017], as well as the gPhone [Fores et al., 2019], but at a rather permanent location. Measurements with classical SG's

and FG-5 are possible only within observatory buildings. The AQG is a hybrid instrument, enabling both deployment in observatory buildings (A-series) as well as field surveys (B-series).

1.2 Terrestrial gravimetry in hydrological investigations

Hydrogravimetry can generally be defined as the combination of gravimetry and hydrology. This can be approached from the geodetic perspective, considering hydrological signals as noise to be eliminated from gravity observations. In most occasions, it is however seen from the hydrological point of view: making use of the gravimetric method for hydrological investigations.

1.2.1 Storage changes in hydrology

Hydrology is the understanding and description of all water-related states, fluxes, and processes over a range of spatio-temporal scales. Without making the claim to be complete, this comprises the water cycle, water budgets, water resources, redistribution and storage of water, water management and floods and droughts. Despite differences in content, foci of investigation, target quantities, and spatial extents, all of the above depend on water storage changes (WSC) over a certain period of time. Hydrological fluxes cause these changes in storage. For example, an extraction of water out of the soil due to evapotranspiration leads to a decrease in soil water storage and an increase in the atmospheric compartment, or percolating water from the soil compartment to the groundwater results in a decrease in the first and an increase in the latter storage compartment. The water cycle describes this context of water fluxes between different storage compartments within the hydrosphere. Water budgets quantify the gain or loss of water within a spatially limited domain in a certain period of time. A domain could be as small as a soil sample in the lab, an agricultural unit, an entire catchment or even as large as a whole continent. A general equation for the water balance is:

$$P + Q_{in} = ET + Q_{out} + \Delta S \qquad (1.5)$$

where P is precipitation, Q_{in} all other inflow into the domain, ET evapotranspiration, Q_{out} all outflow out of the domain and ΔS total storage changes of water in soil, groundwater or biomass. The change in storage is therefore defined as the difference of input and output of a domain. Quantifying this storage and monitoring its development is a key for understanding water reservoirs and managing water resources.

For gravity, these WSC are highly relevant because they represent mass changes. Consequentially, their increase or decrease leads to a respective change in gravity (see Eq. 1.4). This is also one of the keys of the gravimetric method for hydrological applications: water storage changes can be measured directly without knowing the underlying flux. Enhancements of gravimeters within the last 2 decades from a technical and instrumental point of view (e.g. device handling, outdoor usage, precision), sparked the interest of this geophysical measurement method within the hydrological research community. As time-variable quantities are mostly the basis

of interest in hydrological research, relative gravimetry was favored in many approaches over the absolute variant. Suitability of devices for field applications was another aspect which influenced the choice of instrumentation.

1.2.2   The integrative nature of the gravimetric method

The big difference of gravimetry over traditional hydrological measurement methods is its integrative nature. This can be both an advantage and disadvantage at the same time. Instead of information about single points or transects, as for example acquired by soil moisture sensors, it allows to capture the dynamics of water storage as a whole (e.g. groundwater, vadose zone, soil moisture, superficial runoff, evapotranspiration or complete catchments) [Creutzfeldt et al., 2010a] within the region of sensitivity of the gravimeter. In this sense it is partly comparable to discharge measurements where recorded water volumes over time are related to the volume of water stored in the catchment. However, gravimetry extends this application with its ability to simultaneously provide information about other storage compartments at different spatio-temporal scales. At the same time, this last point also demonstrates the difficulty of this integrative nature: it can be cumbersome to impossible to disentangle the observational data.

1.2.3   From hydrological storage variations to gravity effect

Hydrological storage changes occur in all the different storage compartments. For simplicity this could be seen as a change in the height of a horizontal layer of water in different scenarios: unconfined aquifer storage at the water table, soil moisture change close to the land surface, humidity change in the atmosphere or tree water storage changes. All increases or decreases of these water-storage-layers, which are the WSC, are jointly sensed by the gravimeter. If a certain magnitude of mass change in one of these compartments can be resolved by the gravimeter, depends on the instrumental precision and varies among their types (compare chapter 1.1.2.

With individual and specifically adapted post-processing, a superconducting gravimeter can reach a precision up to 0.1 nm/s^2 [Van Camp et al., 2016b] which enables to look at water storage changes as small as 0.2 mm (height change of the water layer). This estimate is calculated using Bouguer's approximation of an infinite slab:

$$\Delta g = 2\pi \times \rho \times G \times \Delta h \tag{1.6}$$

with ρ being the water density, G the gravitational constant and Δh the change in water storage height (thickness of the slab). Applying Eq. 1.6 presumes idealized conditions as in an infinite extent of the storage change within a completely homogeneous storage medium which is perfectly flat. While this is not the case for any site, using Bouguer's approach is sufficient for many applications (e.g. deep aquifers) or is at least a good first attempt to judge expected gravity changes related to water storage changes of the signal of interest.

If there is need for a more precise conversion or the domain of interest is rather small (e.g. plot-scale, field-scale) or near the gravimeter, a more physically-based and adoptable approach is necessary. This can be achieved by subdividing the domain into a 3D-grid with multiple grid cells. Applying equation 1.4 enables the

conversion of grid-cell WSC into grid-cell based gravity change. The total change in gravity for the entire domain is obtained by adding the calculated gravity changes of all grid cells. For example, one can imagine a certain spatial extent on the surface of the earth where water storage changes are observed e.g. by groundwater wells and soil moisture sensors. The grid-based approach takes into account spatially distributed heterogeneous WSC and facilitates the calculation of the resulting total change in gravity.

The aforementioned *region of sensitivity*, where the gravimeter senses all mass changes, is a cone shape region above and below the gravimeter (with the peaks of the cones connecting at the gravity sensor) [Kennedy et al., 2014; Van Camp et al., 2017]. Consequentially, mass changes in the horizontal plane cannot be sensed. The region of sensitivity is also often called footprint (of the gravimeter).

1.2.4   Influence of topography

Topography not only influences hydrological processes but also impacts the composition and consequently the interpretation of the resulting gravity signal from water mass changes. There are two idealized situations which help to understand this: a gravimeter either (a) on a hilltop or (b) at the bottom of a steep valley. For (a), a possible mass increase (e.g. precipitation leading to an increase in soil water storage) would take place only below the gravity sensor, resulting in an increase of observed gravity. For (b) and when neglecting the soil below the gravimeter, the mass increase in the hill slopes, above the gravity sensor, would result in a negative change in gravity. In mixed situations, it is complicated to impossible to spatially attribute the origin of gravity changes, observed by the instrument. It is therefore crucial to have a good understanding of site-surrounding topography [Creutzfeldt et al., 2008]. Furthermore, topography highly influences the spatial extent up to which the gravimeter is sensible to mass changes (besides the influence of the precision of the gravimeter itself). This footprint can be up to several hundreds of meters for high elevated locations on hilltops and as small as only several ten-meters in steep valleys [Naujoks et al., 2010].

1.2.5   Extracting hydrological signals of interest from gravity data

Different hydrological processes have different spatio-temporal dynamics and magnitudes. All of them are included in an aggregated way in the gravimeter data, due to its integrative nature. Other non-hydrological signals are also included in the observations. A good overview of magnitudes and frequencies of these phenomena is given by Van Camp et al. [2017]. Consequently, it is important to design a gravimetric study with a clear idea of the expected target quantity of interest. This is true in general and not only for hydrological research. With respect to water related study subjects, the prevailing phenomena can be classified into two major components: (i) the large-scale effects (hereafter referred to as $g_{hyd\,global}$), and (ii) local effects (hereafter referred to as $g_{hyd\,local}$). The local-scale in this context describes the area directly around the location of a gravimeter (up to a radius of around 10 km, depending on topography), whereas the large-scale (or global-scale) covers the remaining domain of the earth. These large-scale effects are therefore water mass

changes related to the global hydrological cycle and comprise variations with semi-annual, annual and inter-annual frequencies as well as long-term trends. The gravity signal from local hydrology on the other hand is dominated by processes leading to WSC present in the direct vicinity (up to several km) of the instrument. These are WSC in groundwater (depletion or recharge), soil moisture storage increase (resulting from precipitation) or decraese (due to evapotranspiration) and within biomass (e.g. transpiration, water extraction by plants). They encompass frequencies from minutes to daily, semi-annual, annual, inter-annual and long-term. Eq. 1.7 includes all effects composing the observable gravity signal.

$$g_{obs} = g_{tides} + g_{atmosphere} + g_{hyd\,global} + g_{hyd\,local} + g_{drift} +$$
$$g_{polar\,motion} + g_{non\,tidal\,ocean\,loading} + g_{other} \qquad (1.7)$$

g_{tides} is the effect due to solid earth tides and ocean loading, $g_{atmosphere}$ describes the effect of atmospheric pressure variations, $g_{hyd\,global}$ is the global hydrological effect, $g_{hyd\,local}$ relates to local hydrological mass changes, g_{drift} is the drift of the instrument itself, $g_{polar\,motion}$ relates to motion of the earth's rotational axis, $g_{non\,tidal\,ocean\,loading}$ is the effect of all non-tide related water mass changes in the oceans (e.g. ocean currents) and g_{other} are all other possible phenomena which influence the measurement location (e.g. earthquakes, geothermal activity, volcanoes, subsidence, uplift, micro-seismic noise, human activity related noise).

Generally speaking, to isolate the signal of interest from the observed gravity time series, all other effects have to be subtracted. Consequently, all non-desired effects have to be known when processing the gravity data. Two different approaches exist: (i) direct measurements of the effect and (ii) modelling of the effect. The approach applied usually depends on the non-desired effect itself but furthermore on the study setup, data availability and temporal data resolution. Sometimes it is also possible to combine both approaches in individual ways. This is for instance the case for the atmospheric component where post-processing could be carried out with centralized provided, worldwide available model data (e.g. ATMACS, Klügel and Wziontek [2009]), the locally recorded air pressure time series in combination with an admittance factor [Warburton and Goodkind, 1977], or even the combination of both (thus including a local pressure residual).

An example approach for gravity data processing in order to isolate the local hydrological component is given in the following. From the observed gravity data (usually device-internally filtered for micro-seismic frequencies), tides, polar motion and drift is subtracted. Tides can be calculated based on location-specific estimated tidal parameters, polar motion coordinates are provided by IERS [Petit and Luzum, 2010], and the drift component is know from regular absolute gravity measurements at a site. Next, the remaining gravity time series is scanned for steps, earthquakes and possible effects of maintenance times at the instrument, after which a further reduction for the atmospheric (e.g. ATMACS, Klügel and Wziontek [2009]), the global hydrological (e.g. ERA5, Hersbach et al. [2020]) and the non-tidal ocean loading (e.g. OMCT6, Dobslaw et al. [2017]) component is carried out. The remaining time series is called the *gravity residual* and comprises all effects of local hydrological processes (and other uncorrected processes, if present) relevant at the site.

If the signal of interest is specific to a certain part of the local hydrological gravity

component only, the residual can be further reduced, using Eq. 1.8, to focus only on the desired target quantity.

$$g_{hyd\,local} = g_{gw} \; + \; g_{vz} \; + \; g_{sm} \; + \; g_{biomass} \tag{1.8}$$

Individual components represent a change in gravity due to WSC in the groundwater compartment (g_{gw}), the vadose zone (g_{vz}), the soil water storage (g_{sm}) or within above ground vegetation (e.g. plant transpiration, water storage in trees) ($g_{biomass}$).

Usually the quantification of local hydrological components is based on in-situ measurements. It is also possible to model local components impossible to measure, if reasonable. Coming from the example back to more general remarks, it is important that applied filters and temporal data resolution is coherent and reasonable between all component time series. An extended overview of gravity data processing can be found in Hinderer et al. [2015b]; Crossley et al. [2013]; Van Camp et al. [2017].

1.2.6   Current state of applications of terrestrial gravimetry within hydrological research

One of the first mentions of a hydrological influence on gravity measurements goes back to Bonatz [1967]; Lambert and Beaumont [1977]. These early studies looked at the impacts from individual water compartments, soil moisture [Bonatz, 1967] and groundwater [Lambert and Beaumont, 1977], on gravity and compared their findings with traditional hydrological measurements (e.g. piezometers). Since the initial publications on this topic, the hydrogravimetric method has been an increasingly utilized tool as well as a methodological study objective for numerous hydrological-oriented investigations at different spatio-temporal scales, environments and research foci.

Using a **hydrological model** to describe and quantify processes and states is a traditional approach in hydrology. With the rise of the gravimetric method, this aspect has been revisited with a new perspective in numerous studies. Hasan et al. [2006] showed that a linear transfer function is capable of modeling the gravity response to medium to high rainfall events. They included other water compartments such as interception, snow cover and soil moisture in a simple GIS-based model. The need to include a rather complete picture of all water compartments was also shared by Krause et al. [2009] who looked at the influence of modeled and measured soil moisture. Naujoks et al. [2010] evaluated a quite complex 3D forward gravity model with local hydrological measurements. They underlined the importance of local and independent measurements for qualitatively high model calibrations as well as the ability to carry out a reduction of the observed gravity for the local hydrological component. A diverse perspective, also with respect to independent local observations, was given by numerous studies of Creutzfeldt. The hydrological influence on gravity observations was analysed with the use of lysimeter data in combination with a simple 1D hydrological model [Creutzfeldt et al., 2010c]. At the study site they found the gravity residuals to be dominantly affected by local water storage changes. The model calibration was highly improved by using the lysimeter data whereas short term variations of gravity observations could be well explained by snow cover and soil moisture measurements. In a different approach where the local water storage change model was constrained by gravity data only,

the understanding of maximum water storage capacity of the hydrological system was highly improved in comparing to an analysis on the basis of only soil moisture and groundwater data [Creutzfeldt et al., 2010b]. This study furthermore showed the potential how gravity data can be useful for upscaling point measurements to the field scale. Another valuable study with respect to hydrological model and parameter calibration was conducted by Christiansen et al. [2011a]: a forced infiltration experiment, combined with a vadose zone model which was coupled with a hydrogeophysical inversion based on ground penetrating radar, facilitated the constraint of traditional Van Genuchten soil-water retention parameters.

Calibrating or estimating **hydrological parameters** with gravity data has also been the focus of many other studies. Gehman et al. [2009]; Pool and Eychaner [1995] used relative gravity surveys to reveal aquifer storage variations for inferring the specific yield. Pool and Eychaner [1995] showed that this method was sufficiently accurate when significant variations of water levels within the aquifers occurred. Wilson et al. [2012] confirmed this by stating that the uncertainty in estimating the specific yield was always high when only small groundwater level variations took place. Gehman et al. [2009] compared their results with the traditional approach of aquifer pumping tests, finding similar parameter values. This approach was also carried out by Pendiuk et al. [2020] but with gravity data not from relative field surveys but a permanently installed superconducting gravimeter. Forward modeling of gravity was constrained by pumping tests data and lead to the inference of specific yield. With another study, Creutzfeldt et al. [2013] showed that gravity data from a superconducting gravimeter can be successfully used to define storage-discharge relationships. In a nested catchment approach, with catchments increasing in size, they found this method to be suitable for large catchments ($> 100\ km^2$) but not applicable for areas smaller than $1\ km^2$.

Besides hydrological models and parameters, **hydrological processes** have also been intensively addressed by various gravity studies. Average daily gravity residuals stacked from a 9 year long time series of a superconducting gravimeter enabled Van Camp et al. [2016b] to infer evapotranspiration amounts and rates. Hydrological modeling of the unsaturated zone at an artificial recharge site showed gravity data being more sensitive to water movements than sparsely distributed groundwater level measurements [Kennedy et al., 2016]. With their approach is was furthermore possible to determine infiltration rate and the depth of the wetting front. Combining a controlled experiment in an indoor basin which simulated the dynamics of a pumping test along with water-storage-change measurements and simultaneous CG-5 measurements showed that at most gravity stations the modeled values were in good agreement with hydrological observations [Christiansen et al., 2011b]. Subsurface water storage changes were also the subject of a study by Creutzfeldt et al. [2015] where a gravity data based hydrological model was calibrated together with data from tree-ring growth, allowing to extend the capability to describe water storage changes 63 years into the past. Another study focusing on water storage changes in a mountain catchment during the drought of 2018 showed the advantage of gravity observations in combination with modeled catchment water balance as opposed to using remote sensing data (MERRA2) [Quentin et al., 2020]. The drought of 2003 and its long lasting effects on the evolution of different hydrological compartments was the focus of a study by Creutzfeldt et al. [2012]. Depth-integrative gravity residuals, indicating a several years lasting depletion of the subsurface water stor-

age, turned out to be a tremendous advantage over superficial point measurements (e.g. soil moisture). With flash floods, another hydrological extreme was addressed in a study combining superconducting gravity data and groundwater levels in an environment dominated by karst cave systems [Watlet et al., 2020]. Observing these events in all datasets allowed to identify a large number of unknown and unmapped voids in the karst system, suggesting additional temporal water storage.

More studies in **karst environments** have been conducted. Fores et al. [2016] confirmed, based on hydro-gravitational 1D modelling, that the karst storage capacity in their study area was generally low, allowing for a fast transfer of water. Time-lapse gravimetry with CG-5 gravimeters in 3 independent karst systems was used to infer karst system properties such as its transmissive function [Champollion et al., 2018]. A joint-interpretation of data from various geophysical methods (seismic, electric restistivity tomography, gravity) in combination with tracers and hydrological modeling in an easily accessible karst tunnel allowed for a spatial differentiation of hydrological processes and a quantitative interpretation of heterogeneity in the karst system.

Furthermore the subject of hydrological masses related to **volcanoes** has been investigated at a few research sites. Gravity residuals and GPS measurements were related to topographical variations due to growth and collapse of volcanic structures [Jousset et al., 2000]. Additionally, they found a correlation in terms of an admittance factor between the gravity residual drift and volcanic activity. Kazama et al. [2015] showed that after correcting for hydrologically-induced local disturbances, the gravity time series could serve to monitor changes in magma masses. These findings as well as the urgent need for a detailed knowledge of the local hydrological gravity component was confirmed by Carbone et al. [2019], using a multi-gravimeter approach. At the same site, Chauhan et al. [2020] combined gravity data in a joint analysis with data of ground deformation during a larger eruption of 2018, concluding that interpretations of the existing data with respect to volcanic processes are more complex than thought by previous studies. The difficulty in interpreting gravity data for volcanic signals without having a good understanding of local hydrological signals was also confirmed by Hemmings et al. [2016].

In general, this topic of **local hydrology** around gravimetric measurements is of high importance: either because a precise quantification is desired for actively investigating hydrological processes and dynamics or because these local phenomena mask the actual quantity of interest and should therefore be removed. Creutzfeldt et al. [2010a] showed in a comprehensive study that the largest part of the gravity signal can be explained by local water storage changes. Their study covered most hydrologically relevant compartments (snow, soil, saprolite, groundwater) and included an uncertainty and error assessment in comparing the gravimetric approach to traditional hydrological methods of investigating water storage changes. Seasonal water storage changes at local to regional scales were modeled by Van Camp et al. [2006] with a focus on short-period rainfall events. They concluded that only with detailed geological and geophysical investigations in combination with environmental monitoring allows to sufficiently quantify local mass induced gravity changes. Applied experiments in the direct vicinity of a superconducting gravimeter with a special focus on implications with respect to local hydrology and its gravitational quantification was addressed by Kroner and Jahr [2006]. Local water storage variations with respect to its correction, a local but non-hydrological part, also in combination

with atmosphere and other local phenomena relevant for processing superconducting gravity data, was covered in Hector et al. [2014]. Mikolaj et al. [2015] introduced an approach of how to deal with hydrologically-induced gravity variations when there is no comprehensive local monitoring network of hydrological variables. For a detailed local model, including numerous and high precision (spatial and temporal) observations, the computational resources can be a limiting factor. Leirião et al. [2009] introduced a now widely spread and used approach for (spatially) nested gravity forward modelling. They showed how computations could be optimized by using different equations (prism, MacMillan, point-mass) for calculating the gravity effect depending on the distance of the mass change to the gravimeter. This was evaluated by aquifer pumping tests.

A focus on **local effects** and its impact on gravity observations can also be found in the following studies. A rain admittance factor was discussed by Meurers et al. [2007], with a focus on how its application prior to tidal analysis highly improved results. A strong emphasis on the importance of detailed knowledge about the local topography and its possible implications for the quality of interpreting results of the local hydrological gravity component was given by Creutzfeldt et al. [2008]. Their study also firstly addressed the impact of the shielding observatory building around a gravimeter, called the *umbrella effect*. In addition to topography, this shielding or masking effect was also treated by Deville et al. [2013]. Concluding that the Bouguer approach was not always suitable to transfer water height in mm to units of gravity, they proposed a daily-precipitation based modelling approach, separating the gravity signal in a shallower and a deeper component while applying a 2-bucket reservoir model. The need for a reliable quantification of the shielding effect a building has on gravity measurements conducted inside it was also expressed by other studies [Hector et al., 2014; Kazama et al., 2012; Fores et al., 2016]. While Kazama et al. [2012] used mathematical constraints in combination with observed saturation states outside of the building in order to quantify mass changes below it, Fores et al. [2016] compared the impact of the masking effect for different gravimeters at different locations within the same building. They furthermore concluded that the *umbrella effect* reduces the sensitivity to local precipitation events by up to 80%.

In the last decade but mostly recently, more studies were carried out combining **gravimetric surveys** with a hydrological point of view. Using absolute-gravity meters, Ferguson et al. [2008] mapped and monitored waterflooding in gas caps of hydrological reservoirs. The importance of hydrological models for improving traditional survey processing methods (such as network adjustment) was shown by Kennedy and Ferré [2015], thus enabling to reveal areas of preferential water infiltration as well as areas of higher groundwater storage. Improving the standard deviations of relative gravity field survey data by an enhanced ambient temperature correction, lead to an improvement of interpreting gravity observations with respect to hydrological signals [Fores et al., 2017]. A comprehensive gravity survey with a focus on hydrology was carried out by Pfeffer et al. [2013]. They used **hybrid gravimetry**, the combination of stationary installations and field surveys by gravimeters, in looking at spatial water storage changes which could be attributed mainly to the vadose zone. The same method was used in a study where a superconducting gravimeter was combined with spatially distributed CG-5 measurements, as well with a focus on spatial heterogeneity of water storage [Quentin et al., 2021]. These water storage changes jointly with the topographic wetness index was com-

pared to a physically-based hydrological model. A good overview of the method of hybrid gravimetry was provided by Hinderer et al. [2016], including an exhaustive list of applications in the fields of hydrology, ice-melting, volcanology and geothermics.

1.3 Open questions and challenges of terrestrial gravimetry in hydrological studies

Many obstacles for an enhanced application of hydrogravimetry relate to the fact that an instrument designed and mostly applied in one field (Geodesy) is now widely used in another (Hydrology). Research and application in geodesy had and has a different focus than in hydrology. Certainly overlaps exist and as demonstrated by the numerous studies in hydrogravimetric topics discussed above, many methods and applications could already be transformed, modified or adapted in order to work out in the context of hydrology. The advantage of this method for hydrological research is enormous and so is its potential. Nevertheless, remaining open questions and challenges need further investigation in order to find solutions acceptable from both the geodetic and hydrological point of view.

Much fieldwork and gravimeter installations by hydrologists aim at investigating phenomena on a different temporal scale or focus than those of geodetic ones. **Pillars for field campaigns** targeted at hydrological extreme events or other temporally limited events might be needed only for some few campaigns in a certain area but still have to provide the standard stability for gravimetric surveys. Similarly problematic are short term and semi-permanent occupations of a gravity (reference) station with respect to obtaining a **precise data-based tide model** which might be necessary to reach a certain reduction precision depending on the hydrological dynamic of interest. **Networks for relative field surveys** are preferably large enough to cover several spatial scales but have to be small enough for studies to adapt to timing and duration of hydrological events. The need for **improved knowledge on spatially distributed local hydrology** for different study areas and objectives was mentioned by Hemmings et al. [2016]; Chauhan et al. [2020]; Kennedy and Ferré [2015]. A similar need but independent of spatial properties, was stated by other studies [Carbone et al., 2019; Hemmings et al., 2016], concluding that a more detailed estimation of the local hydrological gravity component would have improved the study's interpretations. This is directly linked to the fact that buildings or shelters are needed to guarantee favorable conditions for some gravimeters but its sealed surfaces and **masking effect** possibly hinder to investigate some hydrological questions as regard to content [Creutzfeldt et al., 2008; Deville et al., 2013; Hector et al., 2014; Kazama et al., 2012; Förste et al., 2016].

Furthermore, resolving certain small signals might also be problematic with respect to instrumental precision of some gravimeters or the size of **uncertainty due to large-scale, global signal corrections** [Pool and Eychaner, 1995; Wilson et al., 2012; Creutzfeldt et al., 2013]. With respect to these necessary (large-scale) corrections, availability of near-real time (NRT) gravity correction data for NRT provisioning of hydrological events is still challenging. This is especially true for events on short notice (e.g. potential flooding) where fast results and interpretations might be expected.

Last but not least, the problem of **equifinality** that comes with the gravimet-

ric method must be considered when attributing water mass changes to observed changes in gravity. Equifinality stands for the principle that the description or calculation of a certain end state of a system (e.g. the integrative gravity signal) is possible by multiple solutions of conditions (e.g. spatially distributed WSC). Often it is still necessary to complement gravity observations with additional geophysical measurements or supportive monitoring data sets [Van Camp et al., 2006]. The idea of taking advantage of the integrative nature of gravity measurements aims at **reducing the need for massive, invasive hydrological sensor installations**. The goal is to have the smallest possible number of additional hydrological installations, which on the same time still provide enough information about non-desired local components (in a sufficient spatio-temporal scale).

1.4   Research questions and general aim

The general aim of this thesis is to further advance the gravimetric method for hydrological studies, provide solutions for field deployments of gravimeters, and develop methods for hydrological-focused processing of gravity data. With all hydrogravimetric applications carried out up to now in numerous studies, the focus of this work is about using gravimeters as applied hydrological field measurement tools.

To this end, based on the state of the art applications (Chapter 1.2.6) and remaining open questions (Chapter 1.3), existing concepts of hydrogravimetry were examined, adapted, transferred, applied and evaluated from a field hydrologist's point of view. This was carried out within this thesis in the 3 main parts (4 Chapters) where the following research questions were addressed and answered.

The first part encompasses the topics "Hydrology at conventional gravimetric observatory buildings" (Chapter 2) and "The novelty of outdoor measurements with superconducting gravimeters" (Chapter 3):

- How can the impact of the hydrological signal in the direct vicinity of a gravimeter in a shielding building be described?

- Is it possible to attribute water storage changes below these buildings without hydrological sensor installations?

- Would an outdoor installation, avoiding this masking effect, add value and improve the usefulness of gravimetric data for hydrological signals components?

- What is necessary to operate superconducting gravimeters not in an observatory building but directly in the field?

- What are the advantages and disadvantages of such a setup and how does this influence the hydrological context of investigations?

The second part addresses how the gravimetric method successfully contributes to investigate "Hydrological subsurface flow processes at a field site" (Chapter 4):

- What is the added value of gravity data in the scope of sprinkling experiments?

- Can the integral nature of the gravimetric method at the plot scale be an advantage with respect to model evaluation?

- Are subsurface flow processes identifiable by looking at the gravimetric response of water mass redistribution?

- Is it feasible to use a gravimeter as a standalone method for this type of process investigation and lead to invasive (hydrological) sensor installations becoming redundant?

In the third and last part, the focus is on difficulties and problems due to "Correction-induced uncertainties when processing terrestrial gravimetry data" (Chapter 5):

- How big are uncertainties of (modelled) global gravimetric effects, using the default corrections within gravity observation data processing?

- How do these correction-induced uncertainties depend on the time period of the signal of interest?

- Could the uncertainty of these necessary global-effect corrections impede investigations of hydrological processes and dynamics small in magnitude or duration?

At a higher abstraction level, all findings and individual conclusions will enable a generic synthesis, illustrating how answering these questions contributed to advancing hydrogravimetry.

Chapter 2

Hydrology at conventional gravimetric observatory buildings

This chapter has been published as:

Reducing gravity data for the influence of water storage variations beneath observatory buildings

Marvin Reich, Michal Mikolaj, Theresa Blume, Andreas Güntner

Abstract

Ground-based gravimetry is increasingly used to study mass distributions and mass transport below the Earth surface. The gravity effect of local water storage variations can be large and should be accounted for in the interpretation of this data. However, the effect of hydrological mass changes in the immediate vicinity of the gravimeter is not considered in standard routines for separating unwanted signal components. This applies in particular to the effect of the buildings in which gravimeters are installed. The building shields the underlying soil from precipitation and evapotranspiration and thus directly affects the water storage dynamics in the near-field of the gravimeter. A combined approach of in-situ soil moisture observations and hydrological modeling was used to quantify the altered water storage variations below observatory buildings. Subsequently, the errors caused by different estimation approaches for this umbrella effect in hydro-gravitational computations were assessed. Depending on the site characteristics, the errors range from 4.1 to 105.3 nm/s^2 for the intra-annual amplitude when natural soil moisture data is considered for modeling the umbrella effect, and from 4.1 to 12.8 nm/s^2 when assuming no gravity change within 5 m below the building. These results were condensed to general recommendations, leading to a new simple and broadly applicable method to reduce observed gravity data for building effects, given basic information about gravimeter location, building dimensions, climatic regime and soil type of the observation site. This new reduction approach shows errors of the intra-annual amplitude from 1.9 to 7.8 nm/s^2.

2.1   Introduction

In the last decade, gravimeters have been increasingly used in Earth sciences, especially in the fields of geodynamics, hydrology, geothermal energy, volcanology, and seismology. This is mainly owed to instrumental improvements in terms of accuracy, resolution or flexibility of operation such as for superconducting gravimeters [Warburton et al., 2010] and to the refinement of correction methods for other gravity signal components [Klügel and Wziontek, 2009; Boy and Hinderer, 2006; Mikolaj et al., 2015]. Standard corrections now include Earth and ocean tides, atmosphere, polar motion and, in case of relative gravimeters, instrumental drift (see Crossley et al. [2013] for an overview).

In addition, for many applications of gravimetry, a reduction for the effect due to local hydrology is necessary. Ferguson et al. [2008] measured ice thickness and water levels to account for possible mass variations due to hydrology. For investigations on volcanoes, Greco et al. [2012] tried to avoid effects of changing groundwater tables by conducting all surveys in the same season of the year. For a study on volcanoes, Jousset et al. [2000] considered hydrological processes to occur at much smaller time scales than volcanic activity and consequentially did not apply any direct hydrological reduction on their gravity measurements. In a geothermal study, Hinderer et al. [2015a] discussed the necessary reduction of local hydrology that, however, were not possible due to missing groundwater and soil moisture data. For isolating the volcanic activity, Kazama et al. [2015] reduced their observed absolute gravity signal for variations of soil moisture and groundwater table.

Generally, water storage changes (WSC) are traditionally derived from solving the water balance equation where other components (precipitation, evapotranspiration, runoff) are quantified based on local hydrological measurements and models [Hector et al., 2014]. WSC need to be converted into a corresponding gravity response in order to use them as a reduction for the measured gravity signal. Typically, this transformation is done either by i) direct forward modeling of gravity, using hydrological measurements of different storage compartments in combination with a gravity component grid [Creutzfeldt et al., 2010a] or by ii) feeding observations into hydrological models in order to use the model output in combination with the gravity component grid [Hasan et al., 2006; Naujoks et al., 2010; Creutzfeldt et al., 2010c]. The gravity component grid describes the gravity effect of a unit mass change as a function of its position in space relative to the sensor. It is usually the result of a nested grid approach with higher discretization in the close vicinity of the gravity sensor where the impact of mass changes on the gravity measurements is higher (e.g. Leirião et al. [2009]). Furthermore, the conversion of WSC into a gravity response requires knowledge about the three-dimensional distribution of WSC in the research area. To this end, various approaches have been made to upscale hydrological point measurements (e.g. soil moisture) to the gravimeter footprint [Creutzfeldt et al., 2010c; Naujoks et al., 2010; Hector et al., 2013] or to simplify model structures to lower dimensional concepts [Harnisch and Harnisch, 2002; Kroner and Jahr, 2006; Creutzfeldt et al., 2010b].

The methodology of gravity forward calculations with different mesh setups and element sizes and the way how hydrological measurements are incorporated has improved considerably over the last years [Hasan et al., 2006; Creutzfeldt et al., 2008; Leirião et al., 2009; Naujoks et al., 2010; Kennedy et al., 2016]. The hydrological

influence of the near-field (i.e., a zone with a radius of some 10s to 100s meter around the instrument) on the gravity signal has been discussed with a focus on the local topography by Meurers et al. [2007], Creutzfeldt et al. [2008], Deville et al. [2013]. Similarly, the influence of the building in which the gravimeter is deployed on the local hydrology and thus on the gravity measurements has been discussed in some studies (e.g., Meurers et al. [2007], Hector et al. [2014]). The building isolates the soil or rock beneath it from direct infiltration of rain or snowmelt. This effect was first introduced as the "umbrella effect" by Creutzfeldt et al. [2008]. For the gravity calculations, the authors of this and further studies for the Wettzell site (e.g. Creutzfeldt et al. [2010b]) set mass changes beneath the building either completely equal to zero or they allowed WSC to occur in the whole domain. Deville et al. [2013] proposed a method to correct for building effects based on daily precipitation and evapotranspiration data. An approach by Kazama et al. [2012] to consider the geometry of the gravimeter building, was limited to flat areas and fitted soil parameters. Up to now, there is no reported site where the umbrella effect has been quantified based on direct hydrological observations. This lack of measurements hinders hydrological modeling for the determination of site-specific local effects [Hector et al., 2014].

This study investigates the time-variable effect of gravimeter buildings on hydrological mass changes and thus on gravity observations. The focus is on the area beneath the observatory building where soil moisture dynamics are altered, resulting in no or reduced WSC. The aim is to develop a comprehensive and flexible approach to reduce gravity signals for the building effects on local hydrology, applicable to gravimeter sites worldwide.

To this end, we followed both an observation-based and a model-based approach. First, we assessed the water fluxes below a gravimeter building with soil moisture observations for the example of the Geodetic Observatory Wettzell. Second, for the model-based approach, we set up a physically-based 2D hydrological model for the direct surroundings of gravimeter buildings in general, thus moving beyond the Wettzell test site. The influence of site-specific setups, such as building dimensions, position of the gravimeter within the building, soil type, anisotropy of soil hydraulic properties, and climate regimes was estimated based on scenario simulations. We determined the errors of three approaches for the assessment of WSC below the building when forward modeling gravity: i) setting all mass variations below the building equal to zero (e.g. Creutzfeldt et al. [2008], Hector et al. [2013]), ii) setting all variations below the building to values modeled for the area next to the building (e.g. Creutzfeldt et al. [2010b]), and iii) combining the two approaches (e.g., Creutzfeldt et al. [2010c], Deville et al. [2013], Fores et al. [2016], Mikolaj et al. [2016]). We finally introduce a new reduction method for the local hydrological gravity effect beneath the gravimeter buildings based on soil moisture measurements outside of the buildings.

2.2   Methods & Data

2.2.1   Data at test site

The test site with soil moisture observations beneath a gravimeter building is located in the Bavarian Forest in southeastern Germany at the Geodetic Observatory

Wettzell, which is operated by the German Federal Agency for Cartography and Geodesy (BKG) [Schlüter et al., 2007]. This mid-mountainous region (altitude of the study site 606 m) is characterized by a temperate climate with mean annual precipitation of 863 mm and a potential evapotranspiration of 433 mm.

Soil moisture was measured with 15 min resolution with Time Domain Reflectometry (TDR). For the first time, soil moisture sensors were deployed directly in the space beneath an observatory gravimeter building. They were installed in 4 vertical profiles, 1 outside and 3 beneath the building, up to a depth of 2 m (Figure 2.1). Hourly groundwater levels were available at a nearby well (Figure 2.1). Precipitation was measured with a heated tipping bucket rain gauge with a temporal resolution of 30 minutes. A lysimeter (depth 1.5 m, surface area 1m^2) with undisturbed soil and grass cover is used to estimate actual evapotranspiration at 15 min intervals [von Unold and Fank, 2007]. Gravity was measured with a dual-sphere superconducting gravimeter (SG30), located in the observatory building. The time series was decimated to one hour and corrected for all known processes besides local hydrology, i.e., Earth and ocean tides, polar motion, atmosphere and global hydrology.

An extensive description of all available datasets and specifications of the observatory building can be found in Appendix A.
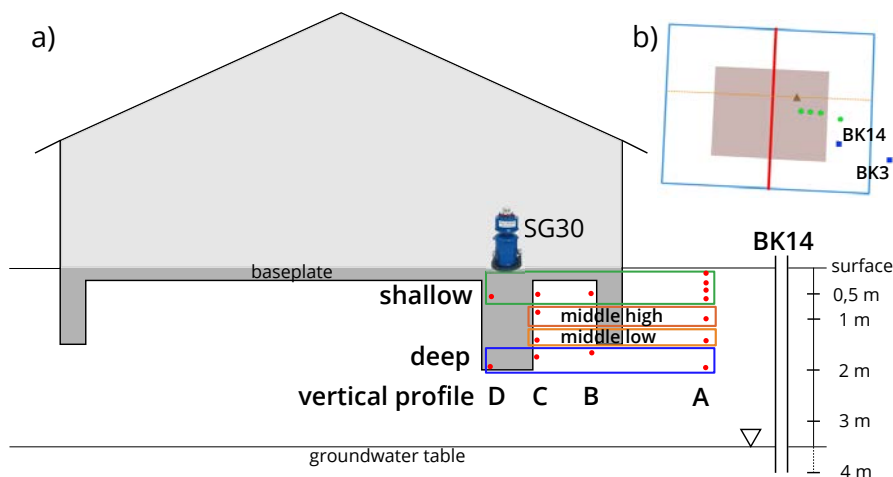
Figure 2.1: a) Cross section through the gravimeter building at test site along the orange dashed line (visible in part b) with location of the groundwater observation well BK14 and the soil moisture sensors (red dots). Profile D is located in front of the gravimeter pillar. All horizontal and vertical dimensions are true to scale, corresponding to the vertical extent shown on the right side. b) Gravimeter building from above (filled square), including model domains (red line: 2D, blue box: 3D, cross section: orange dashed line) and sensors (triangle: gravimeter, green dots: soil moisture sensors, blue square: groundwater wells).

2.2.2   Observation-based approach

To investigate flow processes beneath the gravimeter building for the example of the test site in Wettzell, a cross-correlation analysis of the observed soil moisture time series was carried out. Correlations between the soil moisture dynamics outside and below the building were considered as an indicator for lateral water flow from the outside to below the building, and lag times as an indicator for flow velocities.

Soil moisture time series of all sensors were corrected for outliers and filtered to remove noise. The input for each cross-correlation consisted of pairs of time series from sensors of the same depth. One sensor was always located in profile A (next to building) and the other one in profile B, C or D (beneath building, see Figure 2.1 for sensor notation). With 4 depth levels (8 time series pairs) and 6 yearly data intervals, a total of 48 cross-correlations were calculated. After averaging over the six yearly values, a pair of correlation coefficient and lag time was obtained for each sensor location.

2.2.3   Model-based approach

For the model-based approach of this study on investigating 2D and 3D flow processes beneath gravimeter buildings for different environmental conditions, hydrological models were set up using the finite-element software HYDRUS [Šimůnek et al., 2013]. Model scenarios included a model for the actual conditions at the Wettzell study site, as well as various model scenarios for varying soil characteristics, climate conditions and building dimensions.

The 2D model domain for all scenarios had a horizontal extent of 20 m and a vertical extent of 5 m. The geometry of the building was based on the observatory building at the test site (Figure 2.1). This includes a 1.1 m by 1.1 m square pillar for the gravimeter, reaching 1.2 m below the baseplate of the building. A flat terrain was assumed as the uttermost universal topographic condition. The mesh of the model had a varying discretization for different vertical layers (horizontally, vertically): 0 – 0.1 m (0.25 m, 0.025 m), 0.1 – 0.5 m (0.5 m, 0.05 m), 0.5 – 2.0 m (0.5 m, 0.05 m) and 2.0 – 5.0 m (1.0 m, 0.1 m). All vertical boundary conditions were set to no flow. The upper boundary was no flow for the area of the gravimeter building and atmospheric for the remaining top nodes (2.2). The lower boundary was set to free drainage. Soil scenarios and scenarios for anisotropy of hydrological conductivity were parameterized with 3 common soil textures (sandy loam, silty loam and clay loam). Scenarios for different climate conditions and building dimension were set up with the parameterization of a sandy loam soil (Table 2.2). Model forcing was based on time series measured at the test site Wettzell and consisted of observed precipitation, evapotranspiration and groundwater levels for the period 01/2009 to 01/2012, with a temporal resolution of 1 hour. For the climate scenarios, a simple approach to generate synthetic forcing data was used by scaling the meteorological input of the measured time series with the aridity index [Barrow, 1992], i.e., the relationship of average annual precipitation to potential evapotranspiration. For each climate scenario, the same evapotranspiration time series was used but precipitation was scaled according to the aridity index to mimic more dry or wet climate conditions.

For the scenario of the test site Wettzell, the model was modified in the following two ways: soil layering and parameters from local observations at the site were used (Table 2.1), and the lower boundary was set to variable head condition defined by the measured groundwater table. For all scenarios, the initial conditions were set to field capacity with the intention to reach stable conditions after the spin-up period of 2 years. The following third year was used for further analysis. The HYDRUS model output on the irregular 2D mesh was converted to a regularly spaced 3D grid in order to facilitate gravity forward calculations with the gravity model. This new grid has a discretization of 0.1 m in all directions and a spatial extent of 21 * 18 * 5

m. For this purpose, the 2D mesh was first interpolated to the new spatial resolution using inverse distance weighting. Then the values of the 2D profile were distributed in a radially symmetrical way to the 3D space according to their depth and distance to the building edges. The resulting 3D model domain is shown in Figure 2.2.
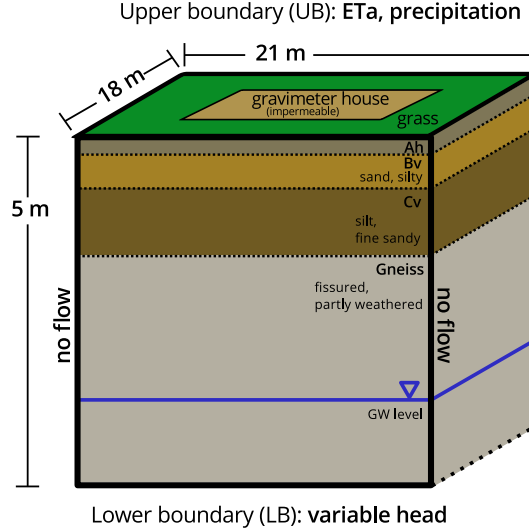
Figure 2.2: Exemplary model setup (after distributing 2D model output into 3D space) with lateral and vertical extents, layers and their soil type and boundary conditions. In the center of the top layer the area of the gravimeter building is shown. This setup is used for the Wettzell scenario. All other scenarios have a free drainage condition at the lower boundary and no soil layering.

Table 2.1: Van Genuchten parameters at the test site Wettzell for different soil layers, based on the analysis of undisturbed soil samples [Creutzfeldt et al., 2010c]. Θr represents the residual water content, Θs the saturated water content of the soil. α is related to the inverse of the pore-entry-pressure, n is a measure for the pore-size distribution and K_{sat} the hydraulic conductivity.

Layer [depth range in m]	Θr [-]	Θs [-]	α [1/m]	n [-]	K_{sat} [m/h]
Ah $[0 - 0.1]$	0.01	0.47	1.8	1.14	0.0329
Bv $[0.1 - 0.5]$	0.01	0.36	63.3	1.06	0.0151
Cv $[0.5 - 2.0]$	0.01	0.37	20.3	1.08	0.0656
weathered gneiss $[2.0 - 12.8]$	0.01	0.04	13	1.3	0.0342

The hydrological response simulated with the hydrological model was then converted to gravity by forward modeling: The corresponding gravity effect of each 3D grid cell was multiplied by its soil moisture content and the sum over all cells of the complete model domain was calculated for each time step of the modeling period. The space of the gravimeter pillar was assumed to have no mass variations (Θ g = 0). For each hydrological model scenario, gravity effects were calculated for different positions of the gravimeter within the building (Figure 2.3).

Table 2.2: Hydrological model setups and parameters of individual scenarios.

Scenario	Soil setup				Anisotropy	Lower Boundary Conditon		Building dimension					Climate (scaling)				
	Sandy loam	Silty loam	Clay Loam	Wettzell soil layers	Anisotropy (x direction)	free drainage	variable head	small: 4m x 4m	medium: 8m x 8m	Wettzell: 8m x 11 m	big: 14m x 14m	huge: 17m x 17m	AI = P/ET = 0,35	AI = 0,575	AI = 1,0	AI = 1,5 (Wettzell)	AI = 2,0
Soil Texture	x					x				x						x	
		x				x				x						x	
			x			x				x						x	
Anisotropy	x				x	x				x						x	
		x			x	x				x						x	
			x		x	x				x						x	
Building Dimension	x					x		x								x	
	x					x			x							x	
	x					x				x						x	
	x					x					x					x	
	x					x						x				x	
Climate	x					x				x			x				
	x					x				x				x			
	x					x				x					x		
	x					x				x						x	
	x					x				x							x
Wettzell				x			x			x						x	

Quantifying the umbrella effect: error assessment

The hydrological model results were used to estimate mass variations in the space below the gravimeter building. The space of influence of the umbrella effect is the zone were water storage variations are altered compared to the natural conditions next to the building. The aim of the error assessment is to quantify the gravity contribution due to WSC in the zone below the building and to estimate errors that arise from simplified reduction approaches commonly used in previous studies: i) setting WSC in the whole zone below the building equal to zero, ii) neglecting the presence of the building, and iii) combining i) and ii), i.e. setting WSC below the building up to a certain depth to zero and assuming that in deeper zones they are identical to outside of the building. The full error assessment for all three approaches was achieved by vertically moving a virtual boundary line starting from the building baseplate down to the deepest layer (5 m). The boundary line splits the space below the building into two zones. Two variants were used to fill these zones with soil moisture information: 1) zero WSC above the line, and simulated WSC from beneath the building below the line, or 2) simulated WSC from beneath the building above the line, and WSC from outside the building below the line (Figure 2.4). For the lowermost position of the boundary line, variant 1) results in approach (i) above, i.e., the whole space is set to 0. For the uppermost position of the boundary line, variant 2) results in approach ii) above, i.e., discarding any umbrella effect of the building. The same routine was applied to all hydrological
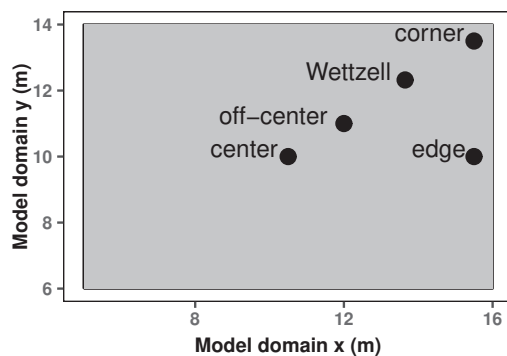
Figure 2.3: Different scenario for the positions of the gravimeter inside the observatory building (gray rectangle). "Wettzell" is the actual position of the gravimeter at the test site.

model scenarios.
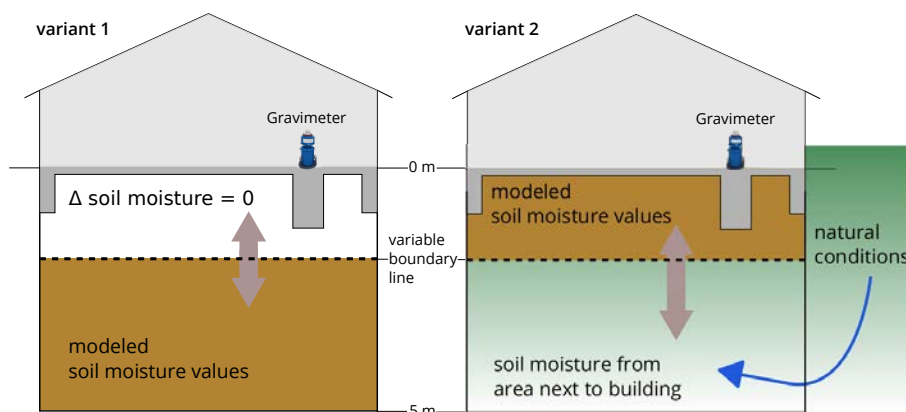
Figure 2.4: Schematic for the error calculation of the umbrella effect for approaches 1 and 2. The virtual boundary line between the two zones is vertically moved downward for individual gravity computations. The dark gray building baseplate, walls and the gravimeter pillar were excluded from the calculation as no mass changes occur here.

Quantifying the umbrella effect: reduction method

As an alternative to the existing simplified reduction approaches, a new reduction procedure was developed here. It uses reduction ratios to determine time series of mass variations below the gravimeter building, solely based on simulated or observed soil moisture data next to the building that are not influenced by the umbrella effect.

The following routine was applied to determine these ratios: In a first step, the gravity response of the simulated soil moisture variations below the building was calculated (real case). In a second step, the gravity response for the soil moisture variations simulated next to the building but assumed to occur below the building was calculated (hypothetical case). In both calculations, the same soil volume was used, including building foundation, gravimeter pillar and position of the gravity sensor. As a third step, the dynamic ratio was determined as the ratio between the obtained gravity response for the real and the hypothetical case at each time

step. Finally, a linear regression was set up, relating these dynamic ratios at each time step to the corresponding mean soil moisture simulated for the zone outside of the building. With the resulting regression parameters, the ratio could thus be expressed as a function of the moisture state of the zone that surrounds the gravimeter building. These steps were carried out for all hydrological model scenarios given in Table 2.2, different gravimeter locations (Figure 2.3), and different depths of the model domain ranging from 0.5 to 5.0 m.

To include information about the size of the observatory building, the above mentioned first three steps were carried out for 5 different building dimensions (see Table 2.2). Before calculating the dynamic ratio, time series were normalized by the equivalent time series of each scenario for the building dimension of the scenario for the test site Wettzell that serves as a reference here. An exponential function described the relation between the mean of the dynamic ratio for each of these scenarios and the corresponding building dimension. This relationship enabled a further adjustment of the gravity response due to mass variations from below the gravimeter building as a function of individual building sizes.

2.3   Results & Discussion

2.3.1   Observed soil moisture: correlation and time lag

Soil moisture at the Wettzell test site in profile A (next to gravimeter building) generally was most dynamic, slightly decreasing with depth (Figure 2.5). Profile B (on building edge) also had high dynamics of shallow and deep soil moisture. Towards the center of the building (profiles C and D), there were no dynamics at shallow depths, but an increase in dynamics with depth. The deep sensors of profiles B, C and D had similar dynamics as the deep sensor next to the building (A). The shallow sensors below the building showed very low correlation to the outside sensors (Figure 2.6). This indicates that no water from outside of the building reached these locations. Deep sensors below the building, however, correlated well with the deep sensor next to the building, reaching correlation values of up to 0.89. Hence, at some depth between shallow and deep sensors, water starts to move laterally under the building. At profile C, this resulted in increasing correlation coefficients with depth. For the time lags of maximum correlation, negative time lags signify a lead in the signal of profile A, located outside of the building area. Time series of deep sensors beneath the building had the least delay with average values of -4 to -2 hours. In contrast, water needed more time to reach the shallow sensors B and sensor C mid high as flow path lengths bypassing the subsurface walls were longer. Similar characteristics could be observed at the temporal scale of individual rainfall events (not shown).
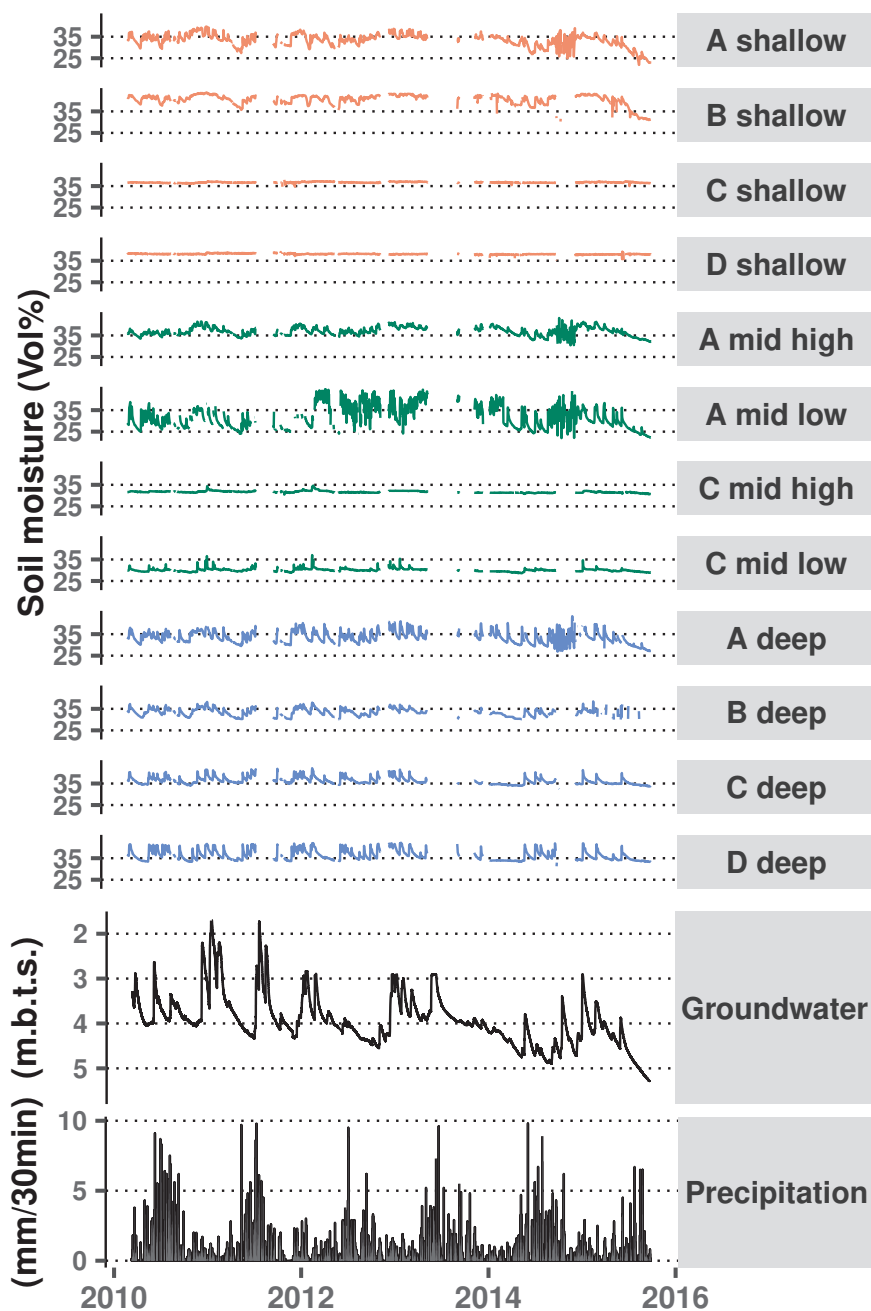
Figure 2.5: Time series of selected soil moisture sensors (for locations see 2.1), groundwater level (in meter below terrain surface) and precipitation at the test site Wettzell for the complete time period. Soil moisture is shown in red (shallow sensors), green (mid-sensors) and blue (deep sensors).
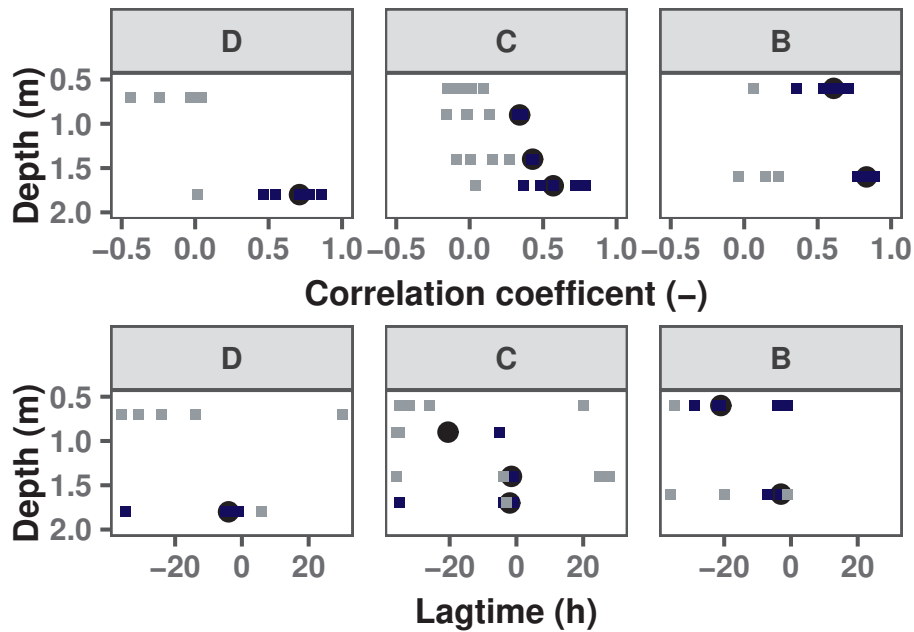
Figure 2.6: Results of the cross-correlation analysis of yearly intervals of soil moisture time series under and next to the building at the site Wettzell. The gray (correlation <0.3) and blue (>0.3) squares show the individual years, black dots the average over all years. Top row shows correlation coefficients, lower row shows lag times relative to the time series in the the corresponding depth of profile A.

2.3.2   Hydrological modeling

For all model scenarios, temporal mean, standard deviation and range of the soil moisture values for the second year of the simulation period (corresponding to the year 2011 for the test site Wettzell) were calculated at every grid point (Figure 2.7). These three parameters notably differ when comparing areas beneath the building and next to it throughout all the scenarios. Differences in soil moisture states and dynamics diminish with increasing depth, similar to what was observed in the measured soil moisture. In terms of the mean soil moisture, the absolute values of these differences vary between scenarios. For the soil texture and anisotropy scenarios, this is due to the soil parametrization (van Genuchten parameters), in the climate scenarios due to varying water availability. In all scenarios, the shape of the zone of influence of the umbrella effect is a cone pointing downwards. This indicates that the space where mean soil moisture below the gravimeter building differs from the surroundings becomes smaller with increasing depth below the building. In the anisotropy scenarios, this cone is narrower and does not reach as deep. This is a direct consequence of the assumed anisotropy which favors lateral relative to vertical water movement. Thus, water from the areas next to the building reaches the zone below the building at shallower depth as in other scenarios. When comparing the soil moisture dynamics below and next to the building in terms of the parameters standard deviation and range, the zone with no or very small soil moisture variations extends deep below the building for most scenarios (Figure 2.7). The patterns tend to resemble an upward pointing cone. This signifies that dampening of soil moisture variations due to water flow into the drier zone below the building also affects the subsurface outside of its footprint.

The root mean square error (RMSE) between the simulated soil moisture time series of the scenario "Wettzell" and the observed soil moisture at the test site was 1 to 3 VWC (in Vol%). Seasonal dynamics as well as dynamics and amplitudes associated with major wet and dry spells were reasonably represented by the model. The gravity response inferred from these modeled water storage changes compares to the observed gravity signal of SG30 at the Wettzell test site for the example of year 2011 resulted with a RMSE of 7.6 nm/s^2 for hourly time series (see Appendix B for details).
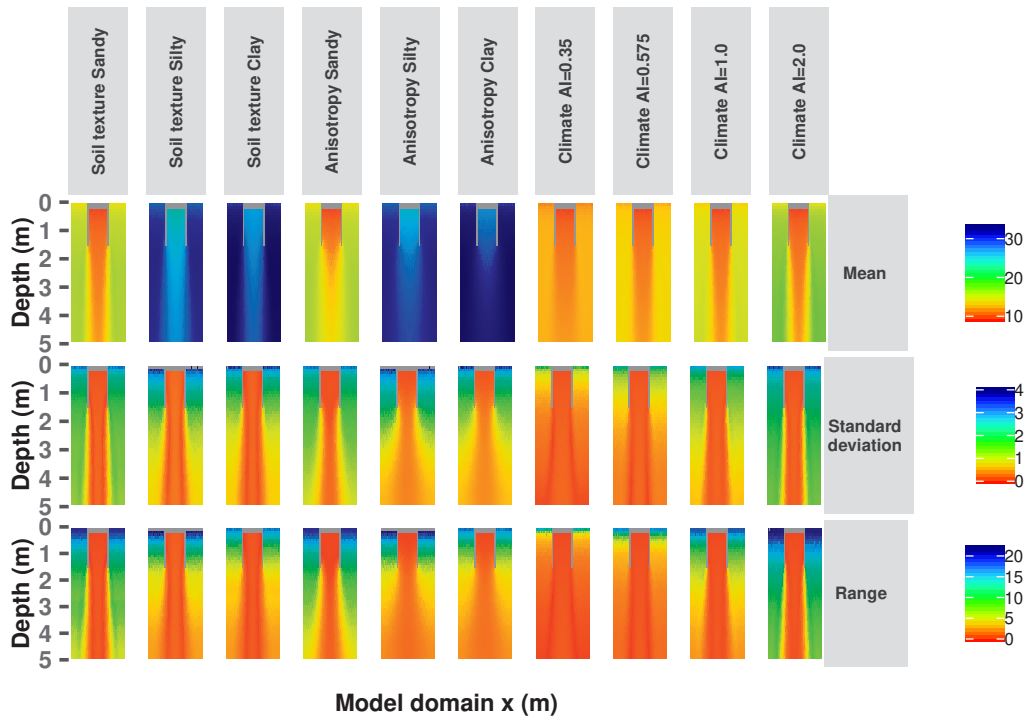
Figure 2.7: Interpolated 2D HYDRUS model output for all scenarios, showing parameters temporal mean, standard deviation and range for the second year of the simulation period in VWC in Vol%.

2.3.3   Quantifying the umbrella effect: error assessment

Errors made by the simplified reduction approaches (see Figure 2.4) are given for the maximum gravity signals, i.e., the annual amplitude of soil moisture as the difference between the maximum and minimum soil moisture values in the second year of the simulation period (Figure 2.8). For variant 1 (Figure 2.8, left), the error that is committed by setting soil moisture variations to 0 below the gravimeter building increases the deeper one assumes this zone of zero change to extend below the building. For the maximum depth of 5 meters considered here, the error is 4.1 to 12.8 nm/s^2, depending on the scenario. For variant 2 (Figure 2.8, right), the error that is committed by setting soil moisture variations below the gravimeter building to the variations that occur in the undisturbed surroundings of the building, increases the closer this zone of replacement comes to the baseplate of the building. In the case that the presence of the building is completely neglected (except of the gravimeter pillar) and outside soil moisture variations are taken for the entire space

below the building (here 5 m), the errors reach 11.2 to 105.3 nm/s^2, depending on the scenario. Overall, the approach relying on soil moisture data from areas next to the building (variant 2) leads to much bigger errors when reducing the gravity time series for local hydrology than assuming no storage variations below the building (variant 1). Thus, variant 1, while not the perfect solution, should be preferred over variant 2, when choosing one of these simplified reduction approaches. The results of the favored new approach developed in this study for reducing gravity data for the umbrella effect are presented in the following.
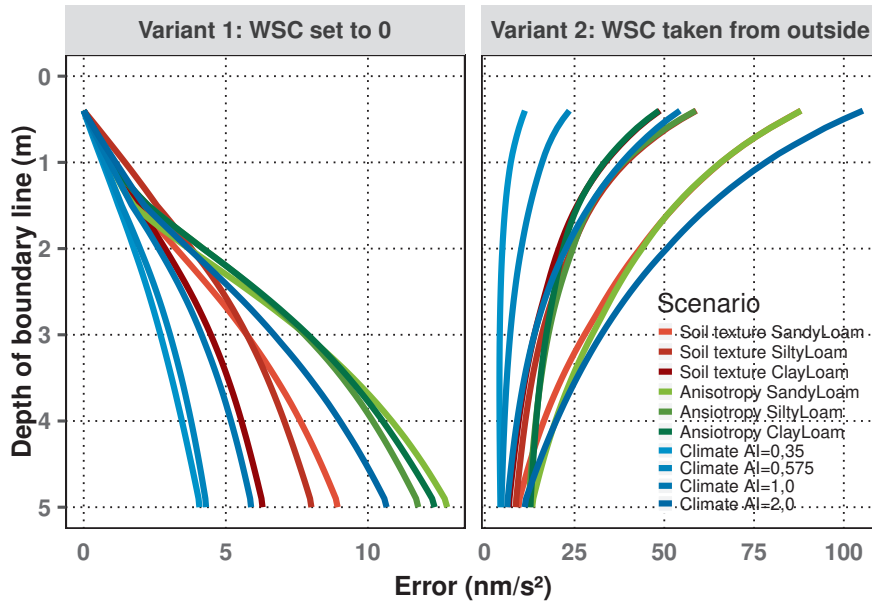
Figure 2.8: Errors of simplified reduction approaches for water storage variations below the gravimeter building, following the two reduction setups shown in Figure 2.4.

2.3.4   Quantifying the umbrella effect: reduction method

For all hydrological model scenarios (Table 2.2) and different positions of the gravime-ter (Figure 2.3), the corresponding gravity responses of WSC below the building were calculated for (i) the real case with the umbrella effect, and (ii) the hypothetical case where the WSC dynamics outside of the building were applied to below the building. In the following, results are discussed for the example of the scenario "soil type sandy loam", but they are similarly valid for all other scenario combinations.

   Gravity signals for the hypothetical case are much higher and with much more dynamics than the gravity responses of the real case (Figure 2.9). A delay in the response can be identified below the building due to the travel time of water moving underneath the building. Absolute values of the gravity signal highly depend on the position of the gravimeter. With respect to the umbrella effect, this means that the more of the volume close to and below the gravity sensor is covered by the building foundation, the bigger the influence of the shielding of the building on the gravity signal.
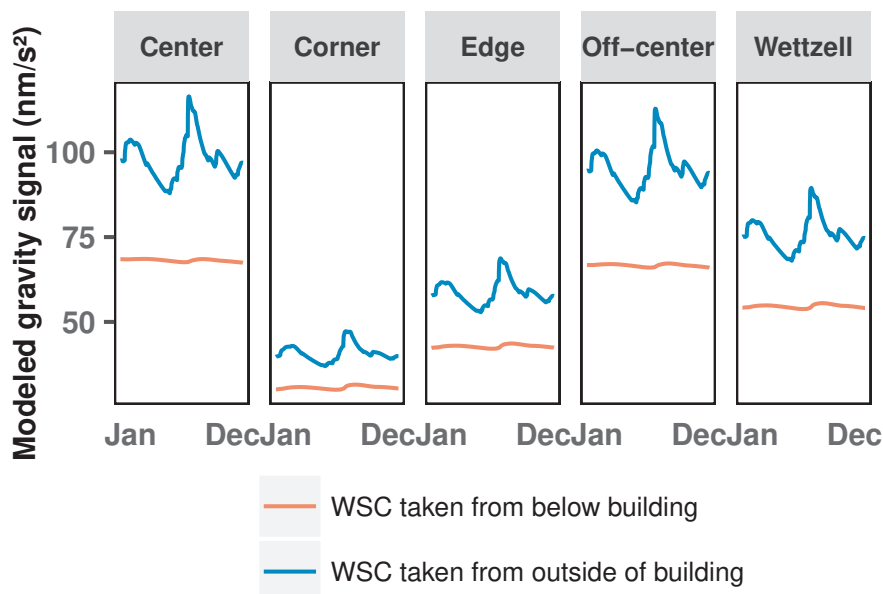
Figure 2.9: Modeled gravity response below the building for the hydrological scenario "Soil type sandy loam" throughout the second year of the simulation period. For each of the possible gravimeter positions, the gravity response of WSC in the space below the building is shown, using either the real WSC (red) or assuming the hypothetical WSC taken from outside of the building (blue).
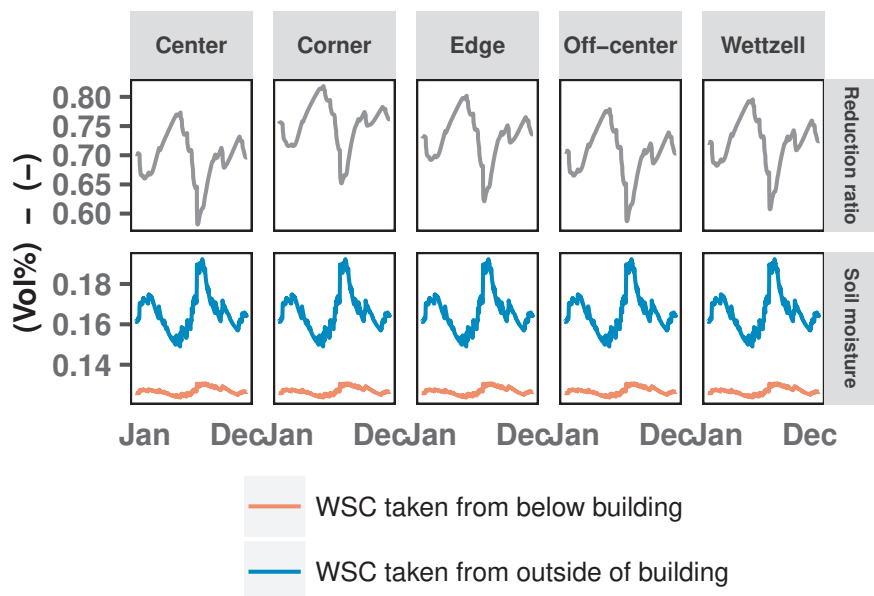
Figure 2.10: Upper panels: Dynamic reduction ratio, resulting from the ratio of the real to the hypothetical gravity response below the building. Lower panels: Mean simulated soil moisture based on WSC of the real case(red) and the hypothetical case (blue). All time series are for the second year of the simulation period, showing the hydrological scenario "Soil type sandy loam" and different gravimeter positions.

This is reflected in the dynamic reduction ratios, relating the real to the hy-

pothetical gravity response below the building for every time step of the model. The ratios range from 0.58 to 0.81 depending on the gravimeter position, shown for the example of the hydrological model "Soil type sandy loam" (Figure 2.10),. The marked variations of the ratio in time highlight the importance of reducing the gravity observations for the umbrella effect in a time-varying way. This time dependency has already been stated in earlier studies by Hector et al. [2014] and Fores et al. [2016]. In the lower part of Figure 2.10, mean soil moisture content of the corresponding model is shown for every time step. As expected, larger and more frequent dynamics are found in the hypothetical scenarios (in blue). An anti-correlation between soil moisture and reduction ratios can be identified, following a linear relationship. For the model "soil type sandy loam", a linear regression results in an R^2 from 0.91 to 0.93, depending on the gravimeter position, with the parameters intercept I and slope S ranging from 1.39 to 1.40 and -4.21 to -3.86, respectively. Thus, the dynamic reduction ratio can be determined as a function of the depth-averaged water content Θ of the soil surrounding the gravimeter building. The two regression parameters vary between the different scenarios (Figure 2.11), with scenario "Climate AI = 0.35" being most different from the others. Within each scenario, the gravimeter position 'corner' results in regression parameters that tend to differ from the ones for other gravimeter positions. For each hydrological scenario hydScen and gravimeter position Gpos, the reduction ratio $RR_{hydScen,Gpos}$ at each timestep t can thus be calculated as:

$$RR_{hydScen,Gpos}(t) = I_{hydScen,Gpos} + S_{hydScen,Gpos} \times \Theta(t) \qquad (2.1)$$

with Θ(t) in VWC / 100, I the intercept and S the slope of the regression.

Similar to the gravimeter position, the footprint of the building has a direct impact on the magnitude of the umbrella effect. This effect has again been estimated by the above modelling approach for all hydrological scenarios and gravimeter positions, in this case all normalized by the respective time series of the scenario for the test site Wettzell building size. The resulting reduction ratios $RR_{footprint}$ for the 5 building sizes (see Table 2.2) decrease with an increasing footprint of the building (Figure 2.12), ranging from 1.35 (small building) to 0.96 (huge building). This relationship is in agreement with findings by Deville et al. [2013]. For a particular building size, the range of the reduction ratio between dry and wet soil moisture conditions is expressed by the error bars in Figure 2.12. This enables the estimation of possible errors for $RR_{footprint}$ with respect to the actual moisture state at the observation site. $RR_{footprint}$ can be described by an exponential function of the building footprint A (R^2 of 0.99) as:

$$RR_{footprint} = a + b \times e^{1/A} \qquad (2.2)$$

with building footprint A in m^2, a = -5.57 and b = 6.5.

All calculations above were done for a vertical model extent of 5 m. For smaller vertical extents, regression parameters slightly vary among the different depth scenarios but all following a linear trend (intercept increasing with larger extents, slope decreasing). Details are described in Appendix C.
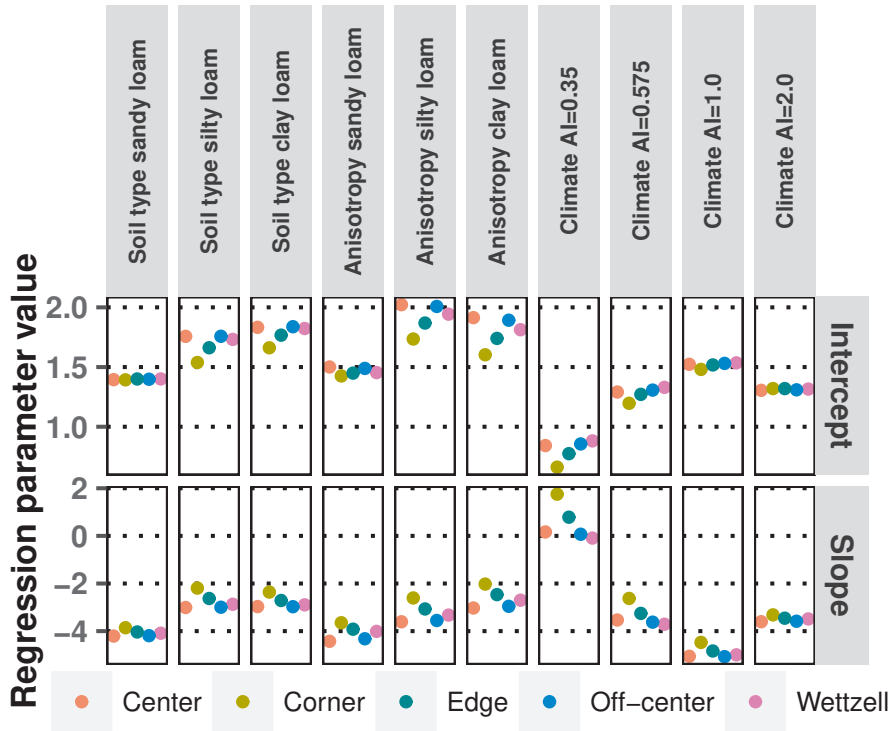
Figure 2.11: Regression parameters intercept and slope, used in Eq.2.1 to calculate the dynamic reduction ratio for each hydrological scenario at all gravimeter locations (see Figure 2.1).
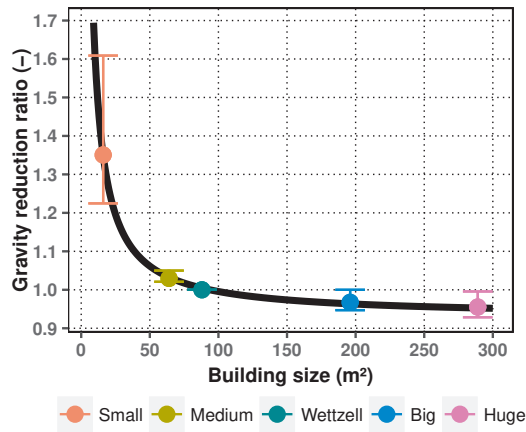
Figure 2.12: Building size in relation to its mean gravity correction factor, normalized by the Wettzell building size scenario, shown for 5 different hydrological model building dimension scenarios. The black line represents the exponential function which was fitted to the data points. For each scenario, the range of its dynamic reduction ratios is shown in colored error bars. Upper end of error bar corresponds to the driest soil moisture state, the lower end to wettest state.

Multiplicative combinations of both reduction ratios $RR_{hydScen,Gpos}$ and $RR_{footprint}$ for the hydrological model "Soil type sandy loam" result in values in the range between 0.56 (huge building with gravimeter at center) and 1.09 (small building with

gravimeter in corner). These values describe in a multiplicative way how much smaller (or larger in the case of values > 1) the gravity response of the space below the gravimeter building is relative to the gravity response if one assumes water mass variations below the building to be the same as outside of the building. The reduction ratios found here are in good agreement with values found in earlier studies: Kazama et al. [2012] mention a reduction ratio of 0.67 due to their 400 m² building, Fores et al. [2016] estimated a ratio of about 0.8 for a 56 m² building and a depth of 5 m and Deville et al. [2013] report a ratio of 0.9 for a 10 m² building. In summary, following the approach explained above (and visualized in a flow chart in Figure 2.13) by choosing the hydrological scenario that is most appropriate for the site of interest, the position of the gravimeter and the size of the observatory building leads to the two reduction ratios that can then be used to calculate the gravity response corresponding to mass variations below the building according to the following equation:

$$g\ response_{belowbuilding}(t) = g\ response_{outsidebuilding}(t) \times RR_{hydScen,Gpos}(t) \times RR_{footprint}$$
(2.3)

This time series can then be used to reduce the observed gravity signal for local near-field hydrology that causes water mass changes below the building:

$$g\ observations_{reduced}(t) = g\ observations_{measured}(t) - g\ response_{belowbuilding}(t)$$ (2.4)

The possibility to adjust parameters to the specific conditions of the research site gives this reduction approach the flexibility to be applicable to almost any gravimeter site worldwide, except for underground sites and sites on steep slopes.

2.4 Conclusions

When exploiting terrestrial gravity measurements for geophysical applications, reductions to isolate the signal of interest from the integrative measured gravity signal are necessary. Water storage variations in the very near-field around the gravimeter have a strong effect, but are usually influenced by the observatory building. This study presents a new approach of calculating the gravity response that is caused by water mass variations below the observatory building based on measured or simulated data outside of the building, which can then be used to reduce the observed gravity signal.

The gravimeter building protects the subsurface below it from hydro-meteorological effects. Data from soil moisture sensors installed beneath the gravimeter building at the test site Wettzell documented these reduced hydrological dynamics. Hardly any dynamics are observed closer to the base plate of the building while the soil moisture time series below and next to the building are more strongly correlated at greater depths.

Reduce gravity signal for local hydrology beneath gravimeter building

Existence of local hydrological model for space below the observatory building?

— no → ① Forward model gravity effect with data from outside of the gravimeter building (model or measured) but with same volume as below the building

② Choose correction parameters, according to
1) hydrological site conditions
2) gravimeter location within building, taking into account the vertical extent of the model domain, and calculate reduction ratio, using mean soil moisture:
$RR_{hydScen,Gpos}(t) = I + S * \Theta(t)$

③ Calculate reduction ratio for building size (A), taking into account the vertical extent of the model domain, acording to:
$RR_{footprint} = a + b * e^{1/A}$

④ Multiply forward modeled gravity time series (1) with ratios obtained from (2) and (3):
$g_{belowBuilding}(t) =$
$g_{modeled}(t) * RR_{hydScen,Gpos}(t) * RR_{footprint}$

yes ↓

Use model soil moisture output to forward model gravity residuals below the building

Use this reduced gravity time series to account for mass variations occuring below the gravimeter building:
$g_{reduced}(t) = g_{obs}(t) - g_{belowBuilding}(t)$

Figure 2.13: Flow chart showing how to calculate a gravity response time series for the space below gravimeter buildings as a consequence of the umbrella effect. These time series can then be used to reduce observed gravity data for the near-field local hydrology below observatory buildings.

Based on hydrological models with different soil textures, lateral anisotropies, climate regimes, and building dimensions, more pronounced water storage changes (WSC) (more water migrating below the building) were observed in models with lateral anisotropies, for climate regimes with high water availability, and for smaller building sizes. Generally, the extent of the space of modified soil moisture dynamics relative to open-air conditions reaches much deeper than may be expected, partly exceeding the total model depth of 5 m used here. If the vertical extent is shallower, e.g. due to the presence of bedrock, the distortion effect of the building compared to open-air conditions slightly increases, but the differences are small and conditions are generally well represented by the reduction factors obtained for the vertical model domain used here.

The error of simplified reduction approaches for the umbrella effect used in previous studies was found to range between 4.1 to 105.3 nm/s^2 in terms of the intra-annual range for the example of the test site Wettzell and a maximum soil depth of 5 m. The approach that assumes identical variations of soil moisture below and next to the building leads to largest errors. Lower but still considerable errors propagate into the reduction when setting the variations of soil moisture below building to zero up to a certain depth and filling the remaining soil layers with values observed or simulated next to the building. Smallest errors result when setting the whole volume below the building to zero storage variations and thus to no gravity variations. Nonetheless, errors from 4.1 to 12.8 nm/s^2 in the intra-annual range for this approach are larger than the accuracy of superconducting gravimeters, and are thus relevant for the interpretation of the signal of interest.

As an alternative to these simplified reduction approaches, the new method presented in this study uses soil moisture data from the area next to the observatory building with natural open-air conditions and the umbrella effect is considered via reduction ratios. These reduction ratios of the gravity response below the building are derived as a function of site-specific conditions and the size of the observatory building. As differences in WSC between open-air conditions and below the building vary in time, depending on the actual soil moisture state, the reduction ratios for the site-specific conditions are dynamic in time. With the application of the reduction ratios to the estimated gravity signal for open-air conditions, mass variations corresponding to the space beneath the observatory building are obtained, which can then be used to reduce the observed gravity signal for the umbrella effect. This new method led to errors of the intra-annual amplitude from 1.9 to 7.8 nm/s^2, ranging between scenarios. The RMSE, calculated on the basis of daily aggregated gravity time series, ranged from 0.4 to 1.8 nm/s^2. Besides limitations and uncertainties of the specific hydrological model used to estimate these ratios (e.g. tempo-spatial resolution, soil parameter transfer functions, boundary conditions, etc.), factors such as water drainage of the roof of the building, estimation of soil parameters and other site-specific setups have to be considered as uncertainties of the method.

Main advantages of the presented approach are its flexibility to adjust to individual site setups and conditions, and its applicability even after the construction of the gravimeter building. The latter means that there is no need to install soil moisture sensors below the building already during construction. A model or observation-based estimate of WSC in the surroundings of the building, however, is a necessary input to the approach. This permits its application even for short study periods. In-depth hydrological understanding or actual hydrological modeling to adapt the approach to the specific site conditions is not necessary because different scenarios are presented here. While these do not cover all possible conditions, they are presumably broad enough so that reduction ratios for the actual conditions of a monitoring site can at least be interpolated from the values given here. We recommend this quantification of near-field hydrological signal components and, consequently, the signal reduction for building effects in future gravity studies.

Chapter 3

The novelty of outdoor measurements with superconducting gravimeters

This chapter has been published as:

Andreas Güntner, Marvin Reich, Michal Mikolaj, Benjamin Creutzfeldt, Stephan Schröder, Hartmut Wziontek

Abstract

In spite of the fundamental role of the landscape water balance for the Earth's water and energy cycles, monitoring the water balance and its components beyond the point scale is notoriously difficult due to the multitude of flow and storage processes and their spatial heterogeneity. Here, we present the first field deployment of an iGrav superconducting gravimeter (SG) in a minimized enclosure for long-term integrative monitoring of water storage changes. Results of the field SG on a grassland site under wet-temperate climate conditions were compared to data provided by a nearby SG located in the controlled environment of an observatory building. The field system proves to provide gravity time series that are similarly precise as those of the observatory SG. At the same time, the field SG is more sensitive to hydrological variations than the observatory SG. We demonstrate that the gravity variations observed by the field setup are almost independent of the depth below the terrain surface where water storage changes occur (contrary to SGs in buildings), and thus the field SG system directly observes the total water storage change, i.e., the water balance, in its surroundings in an integrative way. We provide a framework to single out the water balance components actual evapotranspiration and lateral subsurface discharge from the gravity time series on annual to daily time scales. With about 99% and 85% of the gravity signal due to local water storage changes originating within a radius of 4000 and 200 meter around the instrument, respectively, this setup paves the road towards gravimetry as a continuous hydrological field monitoring technique at the landscape scale.

3.1 Introduction

Water storage is the fundamental state variable of the global water cycle. It is a key state that governs processes of land-atmosphere water and energy exchange, runoff generation, groundwater recharge, as well as matter and solute transport in the Earth's biogeochemical cycles. Quantifying water storage is the basis for water resources assessment and management. Water storage dynamics reflect the net effect of all water fluxes acting in the landscape, balancing precipitation, evapotanspiration and runoff. It has been suggested for a long time that direct measurements of total water storage variations are needed for closing the water budget at spatial scales of practical relevance such as the forest stand, landscape or catchment scale, and for understanding the relationships between storage and water fluxes [Beven, 2002; Davies and Beven, 2015].

The major obstacles for integrative monitoring of water storage variations at the field or landscape scale are, first, that total water storage is a complex state of the hydrological system, composed of various individual storage compartments that would need to be monitored individually. This includes interception storage, soil moisture, vadose zone, groundwater, surface water bodies, snow and ice, with varying contributions depending on the environmental and climatic conditions [Güntner et al., 2007]. Secondly, considerable heterogeneity even at small spatial scales makes it challenging to infer representative storage dynamics at larger scales from traditional point-scale measurements. While progress has been made during the last years with satellite-based and geophysical methods at larger scales [Ochsner et al., 2013; Bogena et al., 2015], these techniques measure the soil moisture component with limited integration depth only. Total water storage variations are available from satellite gravimetry at regional to continental scales [Tapley et al., 2004], however, with low spatial and temporal resolution. Terrestrial gravimetry, in turn, i.e., measuring with gravimeters on the ground (see Crossley et al. [2013] and Niebauer [2015]) for an overview), is an emerging technology for non-invasive monitoring of water storage variations at the landscape scale of some hundreds to thousands of meters in an integrative way over all storage compartments [Bogena et al., 2015].

Terrestrial gravimetry is the measurement of the acceleration of gravity at the Earth's surface, varying in space and time according to Newton's law of mass attraction and due to the Earth's rotation. Gravity changes are determined by measuring the impact of the resulting force changes on a test mass. In absolute gravity measurements, the magnitude of the gravity vector is deduced by observing the trajectory of a free moving object along the vertical. Relative gravimeters measure gravity differences between stations or over time. Stationary relative gravity measurements are carried out continuously and recorded with discrete sampling rate, resulting in time series of minute or hourly resolution, for instance. In contrast, current technology in absolute gravimetry is restricted to periodically repeated observations (time-lapse measurements). For the continuous monitoring gravity variations due to water mass changes in the surroundings of the instrument, which is about seven orders of magnitude smaller than the attraction by the Earth mass, most stable and sensitive relative gravimeters are required. Even though todays spring type gravimeters are well advanced, superconducting gravimeters (SGs) show highest sensitivity and long term stability [Neumeyer, 2010; Hinderer et al., 2015a]. In SGs, the conventional spring-mass system is replaced by a superconducting sphere that floats in the mag-

netic field generated by superconducting coils. Both spring gravimeters and SGs are relative gravimeters.

Time-lapse gravity measurements have been applied in hydrology, for instance, for studying karst systems (e.g., Jacob et al. [2009]), analysis of water flow and storage processes at the hillslope and small catchment scale [Hector et al., 2013; Pfeffer et al., 2013; Piccolroaz et al., 2015], or groundwater model calibration [Christiansen et al., 2011a]. Despite recent improvements of processing strategies towards hydrological applications [Kennedy et al., 2016], the use of time-lapse gravimetry is limited by instrument accuracy and low temporal resolution. High-precision and time continuous monitoring of gravity variations with SGs has been shown to be sensitive enough to resolve water storage variations at seasonal and event time scales occurring within a radius of few 100 meters around the instrument [Creutzfeldt et al., 2008]. To this end, all non-hydrological gravity effects, in particular tidal variations, atmospheric changes and polar motion have to be carefully removed [Hinderer et al., 2015a], as these signals are up to two orders of magnitude larger than the signal of interest. The same applies to seasonal large scale hydrological effects, both in terms of their mass attraction effect as well as in terms of continental loading variations, i.e., deformation of the Earth surface [Boy and Hinderer, 2006]. Continuous measurements with SGs have been used in hydrology to study local water storage variations [Creutzfeldt et al., 2010a; Hector et al., 2014; Fores et al., 2016], validation and calibration of hydrological models [Naujoks et al., 2010], or for unraveling the long-term effects of hydrological extremes [Creutzfeldt et al., 2012]. Recently, Van Camp et al. [2016b] presented the possibility of monitoring evapotranspiration with SGs, albeit limited in their study by the need to stack the gravity time series over several years in order to isolate the signal, an unusual and for most other cases impractical deployment of the SG in a cave about 50 m below the terrain surface, and by the difficulty of correcting for mass changes due to lateral subsurface flow processes.

In spite of these and other studies, the potential of superconducting gravimetry for hydrological applications has not yet been fully explored. Main reasons are that, first, the SG monitoring sites have rarely been selected or optimized for the purpose of monitoring water storage, but for other geodetic or geophysical interests. Secondly, given the highly sensitive SG technology and the instrument size, SGs usually are permanently installed in buildings or in underground observatories under temperature controlled and low-noise conditions. Thus, a field deployment of a SG as a hydrological sensor has been beyond of what was feasible from a technological point of view. And third, the hydrological interpretation of the SG time series is in most cases hindered by disturbance of natural local hydrology due to the observatory building itself and the distance of the hydrological variations of interest to the instrument. While the last aspect known as umbrella effect [Creutzfeldt et al., 2010c; Deville et al., 2013] is particularly relevant because of the important effect of mass changes in the near-field (several meters) around the gravimeter, it is usually ignored in view of missing moisture measurements and unknown flow dynamics below or above the observation room. In the present study, we minimize the umbrella effect by deploying the SG in a compact field enclosure. The latest generation of superconducting gravimeters (iGrav by GWR Instruments, Inc.; Warburton et al. [2010]) being smaller and lighter than previous observatory SGs facilitates this development. First attempts of using superconducting gravimeters as hydrological

sensors placed into small shelters with climate control in the field have been reported by Wilson et al. [2012] and Hector et al. [2014]. An even smaller enclosure, similar to the one presented in this study, was used in a prototype version by Kennedy et al. [2014] in an arid environment for assessing artificial groundwater recharge. Here, for the first time, we test the performance and assess the hydrological value of a long-term SG field installation for the example of a grassland site under wet-temperate climate conditions.

3.2   Study site and hydrological data

The Geodetic Observatory Wettzell [Schlüter et al., 2007], operated by the Federal Agency for Cartography and Geodesy (BKG), is located on a mountain ridge in the Bavarian Forest of south-eastern Germany (Figure 3.1). The crystalline basement of metamorphic rocks (Gneiss) in Wettzell is covered from bottom to top by weathering zones of fractured gneiss, saprolite, periglacial weathering layers, and soil, with Cambisols making up the predominant soil type. The climate of the study area is temperate with mean annual precipitation of 995 mm and mean annual temperature of 7°C. Land cover in the surroundings of the observatory is dominated by a mosaic of grassland and forest, while grassland, gravel and sealed surfaces of roads and buildings alternate on the grounds of the observatory. For a detailed description of the environmental and hydro-meteorological conditions of the study area, its hydrological dynamics including water storage variations and the hydro-meteorological monitoring systems including weather stations, clusters of soil moisture probes, groundwater observation wells, a lysimeter, and snow monitoring, see Creutzfeldt et al. [2010a] and Creutzfeldt et al. [2012].

Table 3.1: Parameters used at the Wettzell observatory to correct precipitation (P) for under-catch by applying the equation $Pcor = P + b * P^\epsilon$. Liquid precipitation between March and November is treated as summer rain; liquid precipitation during the other months as winter rain.

Precipitation type	ϵ	b
Liquid/summer	0.38	0.28
Liquid/winter	0.46	0.24
Mixed	0.55	0.305
Snow	0.82	0.33

At the Wettzell station, BKG operates two superconducting gravimeters of the observatory type (SG029 and SG030) in dedicated gravimeter buildings. At a distance of 41 m from SG030 (Figure 3.1), the new field setup with an iGrav has been installed in February 2015. The closest groundwater observation well (called BK3) is at a distance of 19 m from the iGrav location as part of a network of 9 observation wells on the observatory area with continuous hourly water level monitoring. During the study period, the groundwater level at BK3 is 6.2 m below the terrain surface in average, with peak-to-peak amplitude of 3.5 m.

Time series of precipitation, air temperature, humidity, wind speed and net radiation are available from meteorological stations on the grounds of the observatory
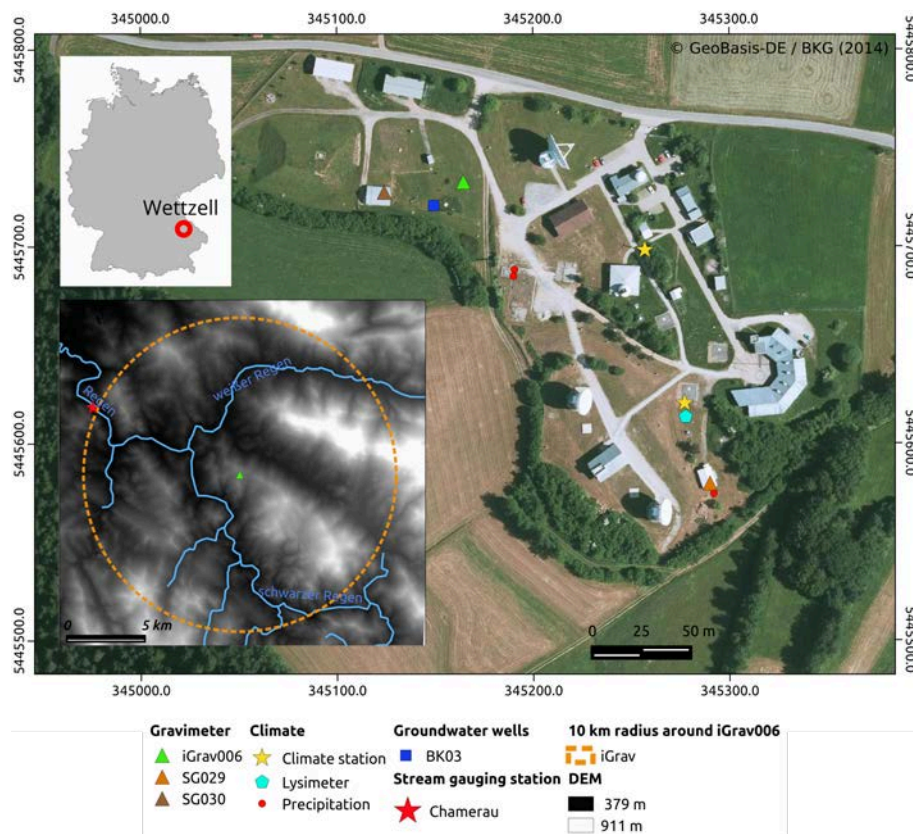
Figure 3.1: The Geodetic Observatory Wettzell, including the position of hydro-logical and gravimetric monitoring systems used in this study. Insert figure shows the region surrounding the observatory with Digital Elevation Model (DEM) with minimum (black) and maximum (white) elevation of 379 and 911 m, respectively. Coordinates for UTM, Zone 33N, in meters.

with hourly temporal resolution. Precipitation data obtained with a Hellmann-type gauge have been corrected for systematic under-catch errors due to wind and evap-oration effects by applying the approach of Richter [1995] as recommended by the German Weather Service (DWD). The correction relies on two parameters (Table 3.1), one site-specific parameter accounting for wind exposition and an empirical exponent. Both parameters depend on the precipitation type. Precipitation is as-sumed to be snow at air temperatures below 0 °C, mixed precipitation between 0 and 4 °C, and liquid precipitation above 4 °C. The correction method is designed for daily precipitation. To correct the hourly values, each daily correction volume is uniformly distributed over all hours with precipitation of the same day.

Since 2007, a weighing UMS lysimeter with a 1.5 m deep undisturbed soil mono-lith and a surface area of 1 m² with grass cover similar to the surrounding grassland sites is operated at the Geodetic Observatory [von Unold and Fank, 2007; Creutzfeldt et al., 2010c]. For this study, the recorded time series of both the monolith weight and the drainage tank weight were filtered for noisy data. The correction method applied follows Peters et al. [2014] as part of their adaptive window and adaptive threshold (AVAT) filter. It consists of a moving average smoothing routine where the window length is adjusted dynamically for each timestep. For the recorded lysimeter time series with a temporal resolution of 1 minute, we used values of 1 to 7 for the

possible order of the fitted polynom, and a maximum window width of 31 minutes. The filtered lysimeter data were aggregated to one hour resolution, from which time series of precipitation and actual evapotranspiration were extracted by considering the overall lysimeter system weight (sum of both monolith and drainage tank) at each point in time, defining precipitation and evapotranspiration as an increase or decrease in weight, respectively. Furthermore, independent from the lysimeter, grass reference evapotranspiration was calculated from the available meteorological data with the Penman-Monteith approach following the FAO-56 standard [Allen et al., 1998].

Streamflow time series were available at the gauging station Chamerau of river Regen, the main river that drains the mountainous area of the Bavarian forest surrounding the observatory. Wettzell is located in a headwater catchment contributing to the river Regen, at a distance of about 10 km to the gauging station. The total catchment area at the Chamerau station is 1356 km², mean streamflow is 26.2 m³/s (period 1931-2013) [Bayerisches Landesamt für Umwelt, last access: 27 June 2017].

3.3   Gravimeter system and performance

3.3.1   Configuration of the gravimeter (iGrav) field monitoring system

The monitoring system is a two enclosure system, comprising the SG itself in its dome-shaped field enclosure (iGFE – iGrav field enclosure) and an external box for peripheral hardware (iBox) (Figure 3.2). The instrument with serial number #006 (iGrav006) was deployed here. A SG records time series of gravity variations as voltage changes in an electronic feedback loop. This feedback loop keeps the levitated superconducting sphere in a constant position by adjusting an ultra-stable and homogenous magnetic field to compensate external force changes. The magnetic field is generated by Niobium coils. The whole sensor system is temperature stabilized by liquid Helium at about 4.7 Kelvin. The system is actively cooled by cryocooler which is re-liquefying evaporating Helium gas and enables a closed system with only 16 litres of liquid Helium.

The main function of the iGFE field enclosure is to protect the iGrav from environmental effects, in particular humidity in form of precipitation and dew, and wind load, and to provide a stable and isolated casing for efficient temperature control in its interior. The enclosure with an outer diameter of 0.9 m is made of double aluminium walls with isolation foam in between. The iGrav (baseplate and body diameter of 0.55 and 0.36 m, respectively) and the field enclosure were mounted separately on a concrete pillar in such a way that there is no transfer of enclosure vibrations and deformations, due to wind stress for instance, to the instrument. Similarly, to minimize noise transfer to the sensor unit, the cryocooler was attached to the field enclosure (via the red platform in Figure 3.2a). The pillar itself has a cylindrical shape with a total height of 2 m, thereof 0.8 m above the ground. The diameter of the pillar is 1 m, leaving only little space around the enclosure. For maintenance, the instrument can be accessed from top via a removable cap (Figure 3.2a). The iGFE also houses controllers with temperature sensors ranging from the bottom to the top, as well as a PC, heating and cooling grills. The cooling grills are connected to a water chiller inside the iBox. The iBox also contains the

compressor to drive the cryocooler, a gas bottle for re-liquefying Helium in the iGrav or recharging the compressor, the power supply including an UPS backup system, controller, temperature sensors and a PC for the remote control of the monitoring system (Figure 3.2b). The entire system requires AC line power with a demand of about 1.5 kW. The temperature inside the iBox is passively regulated with fans. The umbrella effect of the iBox on the gravity observations is negligible, as the footprint is 1 m² and the length of connection hoses for cooling water and Helium gas allows a distance of up to 15 m from the gravimeter.

Figure 3.2: Photographs of the gravimeter monitoring system with (a) the iGrav inside the iGFE field enclosure, (b) peripheral hardware inside the iBox, and (c) an overall view of iBox (foreground), dome-shaped iGFE, and the yellow connection hoses. Photos taken on a grassland site of the Geodetic Observatory in Wettzell, Germany.

3.3.2 iGrav data processing

The voltage changes measured by the gravimeter have to be transformed into a gravity signal by calibration. Usually, the respective scale factor of the SG is determined by co-located measurements of a dominant gravity signal at daily time scales, i.e., the tidal variations with either an absolute or a well calibrated relative gravimeter [Meurers, 2012; Van Camp et al., 2016c]. Although two well-calibrated SG are operated nearby, it was decided for the present study to determine the scale factor and phase delay by regression and cross-correlation against a tidal model resulting from the harmonic analysis of a nine years gravity record by SG029 at the same site. This pragmatic approach effectively minimized any tidal residuals in the gravity time series after reducing the tidal signal, in particular at diurnal to semi-diurnal frequencies which may interfere with hydrological mass variations at these frequencies, especially evapotranspiration. To validate this approach, the same procedure was applied to the nearby SG030. The scale factor computed in this way differed by only 0.8 per mill from the value obtained from calibration using absolute gravity measurements. The calibration factor for iGrav006 used here was 914.416 ±0.005 nm/s²/V with a phase delay of 11.7 s.

As the hydrological signal rarely exceeds 10% of the total measured gravity signal, other gravity effects caused by Earth and ocean tides, Earth rotation and atmospheric variations have to be carefully removed. The local tide model mentioned above was used and atmospheric effects were corrected for with Atmacs (Atmospheric attraction computation service, Klügel and Wziontek [2009]), supplemented by in-situ observations of atmospheric pressure variations to enhance the temporal resolution. The polar motion effect was computed based on the Earth orientation

parameters provided by IERS (International Earth Rotation and Reference Systems Service, www.iers.org). For further analysis, the 1 second gravity time series was filtered using a low-pass filter and decimated to 1 minute temporal resolution. Nine steps found in the gravity residuals were corrected manually via visual inspection. Five of the steps were caused by maintenance work, such as cold-head exchange. Two steps were of unknown source, and the remaining two steps were caused by power surge and iGrav software upgrade. Two steps (in June 2015 and May 2016) out of the nine occurred during rain events so that the instrumental error could not be separated unequivocally from a hydrological mass effect. Hence, a small uncertainty with respect to the level of the residual time series after these events remains.

The instrumental drift of a SG is usually obtained by repeated co-located absolute gravity observations over longer time spans, assuming identical gravity variations at both sensor locations. At the iGrav site, measurements with an absolute gravimeter could not be carried out under the field conditions. Drift determination for iGrav006 based on the drift corrected signal of SG030 was not possible either, as hydrological near-field effects turned out to be too different at both locations. Therefore, the drift of iGrav006 was estimated based on two epochs for which the same total water storage was estimated from independent observations. We assume here that total water storage is the sum of (i) soil moisture storage observed by the lysimeter and (ii) groundwater storage derived from the groundwater level observed at BK3. Based on these data, the same total water storage was found for the days 19 May 2015 and 12 April 2016. This resulted in an iGrav006 drift rate of $+94$ nm/s^2/yr. This drift rate is higher than a first long-term drift of 45 nm/s^2/yr derived for iGrav002 from precise absolute gravity measurements over a four year period by Fores et al. [2016]. The difference is not surprising as it is known from observatory SGs that drift rates vary among the individual instruments [Crossley et al., 2013]. However, an uncertainty of the drift estimation in our study originates from neglecting possible storage changes in the vadose zone in between the lower boundary of the lysimeter and the groundwater level, as well as from uncertainty of the elimination of instrumental steps. The drift was removed from the gravity time series before further analysis. Furthermore, all iGrav006 measurements recorded before 01 May 2015 were discarded because of vibration effects due to inadequate initial mounting of the cryocooler, and several steps related to system maintenance during the initial phase of the field deployment. Figure 3.3 shows the final time series of the gravity residuals of iGrav006 in comparison to the observatory SG030. These time series represent mainly hydrological mass effects as all other gravity effects have been removed as described above.

3.3.3 System performance and data noise

One of the main technical challenges arising from the compact iGFE design under outdoor conditions is the efficient temperature control within the enclosure during all seasons of the wet temperate climate. After some minor modifications, the system was able to stabilize the temperature during most of the time as discussed in the following.

The electronics board (PCB) is mounted below a sealed cover around the neck of the gravimeter and is flooded with Helium gas to avoid humidity. Stable temperature

Figure 3.3: Time series of gravity residuals of iGrav006 in the field enclosure and of SG030 in the observatory building, and daily precipitation rates (from top and right axis) at the Observatory Wettzell.

inside this casing is actively achieved by a heater for which a constant set point (here 32 °C) above the air temperature inside the iGFE was defined. Together with the general heating system of the iGFE, this heater showed sufficient performance to keep the PCB temperature constant during the winter season with minimum outside air temperature of -13°C. However, an unwanted temperature increase inside the iGFE was observed in particular during warm summer days with high insolation. Under these conditions, the performance of the water-based cooling system with regard to the redistribution of cooled air within the iGFE was not sufficient. The resulting increase of the temperature on the PCB was found to have direct effects on the recorded gravity data. A PCB temperature increase caused an apparent decrease of gravity. Significant diurnal temperature variations of several degrees inside the iGFE exceeded the set point of the PCB temperature control and thus translated into PCB temperature patterns and related diurnal variations in the gravity time series. To determine a regression parameter between PCB temperature and gravity, the PCB temperature was artificially increased via the control software. This experiment showed a non-linear response, i.e., the regression parameters varied between -4.4 and -3.2 nm/s^2 per °C temperature increase for different temperature levels. Thus, a direct correction for the spurious diurnal temperature effects was possible with low accuracy only. In July 2016, the PCB temperature issue could be solved by installing extra fans inside the field enclosure and by increasing the set point of the PCB temperature regulation to 34 °C. The fans ensure a better circulation of the cooled air inside the iGFE and thus avoid PCB temperatures exceeding the set point even during hot summer days (Figure 3.4). In turn, no disturbing effect on gravimeter noise due to the fans could be observed. It should also be noted that the temperature variations inside the field enclosure causes tilt effects due to thermal expansion of the mount of the gravimeter and heat transfer from the thermal levellers. This resulted in notable variations in the control values of the active tilt feedback system which was able to compensate these effects.

To characterize the performance of iGrav006 in the field enclosure, its noise level was compared with those of the nearby dual sphere gravimeters SG029 and SG030, both located in a controlled environment of buildings at the observatory. As a common quality indicator for the sensor, the power spectral density (PSD) of the gravity time series is considered at periods from 3 hours to 1 minute in the (sub-)seismic frequency range [Banka and Crossley, 1999; Rosat and Hinderer,
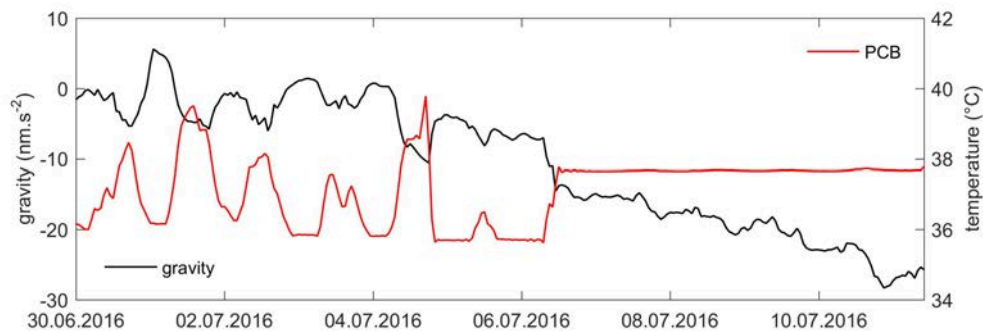
Figure 3.4: Comparison of gravity residuals and PCB temperature (electronics board) before and after improvement of temperature control inside the field enclosure on July 7, 2016. Note that there is an offset of about 4 °C between the PCB set point temperature (34°C) and the actually recorded temperature.

2011]. While the noise is the combination of instrument and site noise, at frequencies larger than about 1 mHz the instrument noise tends to be higher than seismic or environmental noise. Similar to the procedure described in Rosat and Hinderer [2011], tidal and atmospheric effects were removed by the described models before the PSD was estimated by averaging 12 segments overlapped by 75% from a period of 6 quiet days with low seismic activity (Feb 24th to 29th 2016). The results for all five sensors are shown in Figure 3.5, together with the New Low Noise Model (NLNM) of Peterson, 1993 as a reference for the lowest background noise levels of the Global Seismographic Network. The noise level of iGrav006 is very similar to those of the observatory gravity for most parts of the spectrum, with slightly higher values at periods longer than half an hour. This demonstrates the high performance of the iGrav sensor and the quality of the field enclosure system with reasonable reduction of cryocooler vibrations and no visible additional noise from environmental effects such as wind load. The slightly higher PSD values at hourly scales may indicate very small diurnal temperature effects or a higher sensitivity of the field system to hydrological variations. The peaks at periods shorter than a few minutes indicate the resonance frequency (parasitic modes) of the respective sensor, which were not excited for all spheres during the analysis period.

The range of environmental conditions under which the iGrav could be successfully operated in the field enclosure at the Wettzell site was -13.3 to 28.6 °C for air temperature (after fan installation inside iGFE), maximum 8.2 m/s wind speed, and 18.4 to 100% for relative humidity of the air.

3.4 Hydrological value

3.4.1 Sensitivity to water storage variations

It is well known that the gravitational force reduces by the square of the distance from the source. Furthermore, a gravimeter is sensitive only in the direction of the vertical and insensitive to any horizontal components of the gravity vector. Both aspects are important if the sensitivity of a gravimeter to water storage variations is considered. Following Bonatz [1967], a first approximation can be given by neglecting topography and assuming a homogenous layer where water storage changes
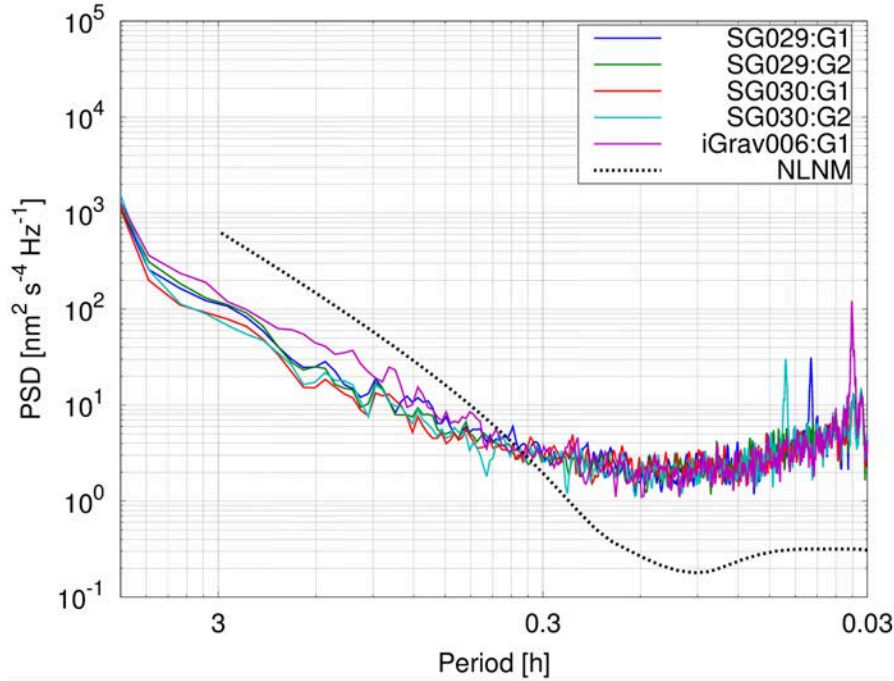
Figure 3.5: Noise level of the dual sphere observatory gravimeters (SG029 and 030) with sensors G1 and G2, and of gravimeter inside the field enclosure (iGrav006), expressed as the power spectral density (PSD) for the time period 24 February 2016 to 29 February 2016. The New Low Noise Model (NLNM) (Peterson, 1993) has been added as a reference for minimum seismic noise.

occur. The mass attraction effect can then be described by a homogenous cylinder with the sensor on its symmetry axis following Eq.3.1 [Heiskanen and Moritz, 1967]:

$$g_c = 2\pi G\rho \left[d + \sqrt{r^2 + h^2} - \sqrt{r^2 + (h+d)^2} \right] \tag{3.1}$$

where d is the thickness and r the radius of the cylindrical layer. h is the distance of the sensor along the symmetry axis (i.e., the height of the sensor above a soil layer where water storage changes occur), while ρ is the density of the layer (as a function of its water content), and G the universal gravimetric constant. Increasing the radius to infinity results in the well-known Bouguer plate following Eq.3.1:

$$g_c = 2\pi G\rho d \tag{3.2}$$

for which the mass attraction effect only depends on the thickness of the layer and on its density. Accordingly, if the radius of the region is chosen sufficiently large, the gravity effect does not depend on the distance of the sensor to the layer. The solid lines in Figure 3.6 illustrate that this is the case for a radius of about 100 m to 200 m for sensor heights of up to 5 m above the cylinder, as the resulting gravity effect converges asymptotically to the effect of the Bouguer plate (dotted purple line). These results changed significantly if the concrete monument (gravimeter pillar) on which the gravimeter is installed were considered. Assuming no water storage changes within the pillar volume, the total gravity effect reduces considerably for sensor heights below 1 m and the effect of an infinite plate is never reached. However, for sensor heights above 2 m both curves come very close at the same radius, since

the effect of the monument decreases rapidly with distance.
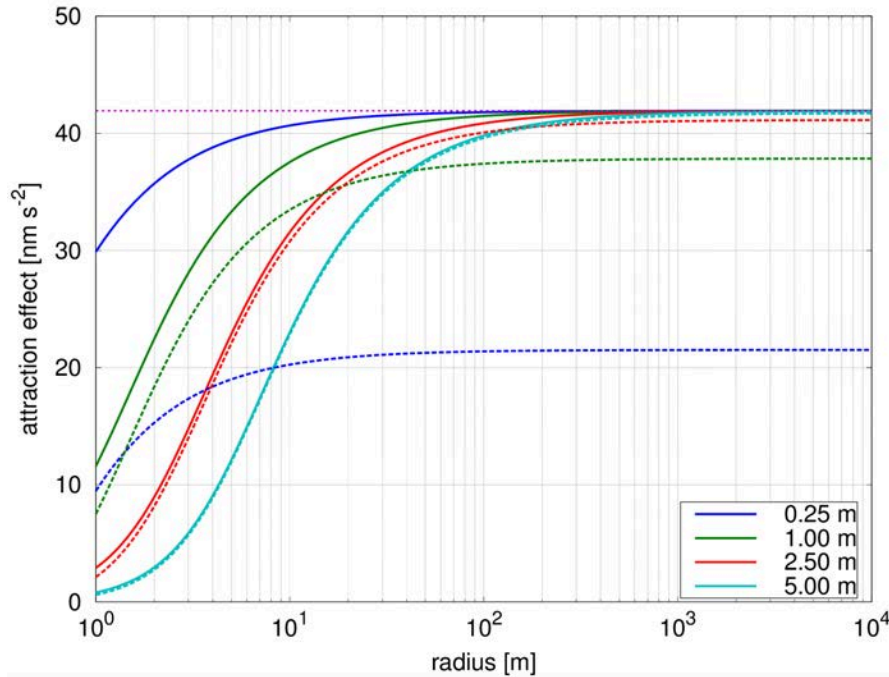
Figure 3.6: Gravity effect of a homogenous cylinder with a thickness of 10 cm, a density of 1000 kg/m³ and varying radii on a sensor placed at different heights (0.25 m to 5 m) above the cylinder (solid lines). Dashed lines show the same gravity effects but reduced for a cylinder with a radius of 0.5 m and a depth of 1.2 m, corresponding to the dimensions of the pillar used for the installation of iGrav006 at Wettzell. Dashed purple line indicates the gravity effect of the cylinder with infinite radius.

Thus, amplifying the gravity signal that is recorded by a gravimeter due to water storage variations in its surroundings can basically be achieved in two ways: (i) by reducing the sealed area of pillar and housing around the gravimeter (i.e., minimizing the umbrella effect), and (ii) by positioning the gravimeter sensor in a suitable position within the local topography. While (i) is the main motivation for the compact design of the field enclosure system described in this study, (ii) has also been considered with the iGrav deployment at the Wettzell site. Both issues are discussed in the following in comparison to the nearby observatory SG030. For the calculation of the gravity effect on the gravimeter sensor, a gravity model with a prism approximation was used [Nagy, 1966]. The location of each prism with respect to the gravimeter sensor is defined by a high-resolution local digital elevation model. The size of individual prisms is smaller the closer they are to the sensor. Given the location and the water mass change in the prism, the gravity effect of each prism on the sensor can be integrated analytically based on Newton's law of mass attraction, and finally summed up for all prisms to get the overall gravity effect of water storage changes in the surroundings of the gravimeter.

The area sealed by foundations and the roof of the observatory building of SG030 equals 88 m² while the iGrav pillar covers about 0.8 m². Soil moisture sensors installed beneath the SG030 building show that soil moisture variations in the first two meters below the building are absent or markedly smaller than for outside

sensors under natural conditions in the soil surrounding the building [Reich et al., 2019]. As an example, a water storage change of 10 mm in the first two meters below the SG building and below the iGrav pillar would cause a gravity effect of 2.79 nm/s^2 and 0.15 nm/s^2 for SG030 and iGrav006, respectively. In other words, a natural storage change of 10 mm will result in a gravity signal that is about 18 times smaller for SG030 than for iGrav006 due to the umbrella effect of housing or pillar. For the further analysis in this study, we set the depth of the umbrella space, i.e. the depth below housing or pillar in which no soil moisture variations take place, to 2 m.

The topographic effect reflects the spatial distribution of hydrological mass changes outside of the building or pillar relative to the position of the gravimeter sensor. While the building with SG030 is located in a topographically low position, the iGrav was intentionally placed on an adjacent upslope location. For SG030, water storage changes partly take place at topographic positions above the gravimeter sensor with a gravity effect that is opposite in sign to the same changes occurring topographically below the gravimeter sensor. In total, the gravity effects of near-surface soil moisture variations in the landscape cancel out to some extent for SG030. In contrary, for iGrav006 all near-field mass changes are located below the gravity sensor so that no cancelling effect occurs. Furthermore, following the theoretical considerations above (Figure 3.6), the instrument was placed on a pillar of 0.8 m height, leading to an effective height of the gravity sensor of 1.05 m above the terrain surface. This further amplifies the sensitivity of the instrument to near-surface soil moisture variations (Figure 3.7). While this sensitivity increases markedly within the first meter, it levels off at even higher sensor positions. Similarly, the increase of sensitivity with sensor height is less pronounced if the water storage variations occur in larger soil depths [Hector et al., 2014]. In this study, the chosen sensor height is a compromise between signal sensitivity, on the one hand, and system operability with regard to ease of access to the field enclosure and the gravimeter, and stability of the concrete pillar, on the other hand.
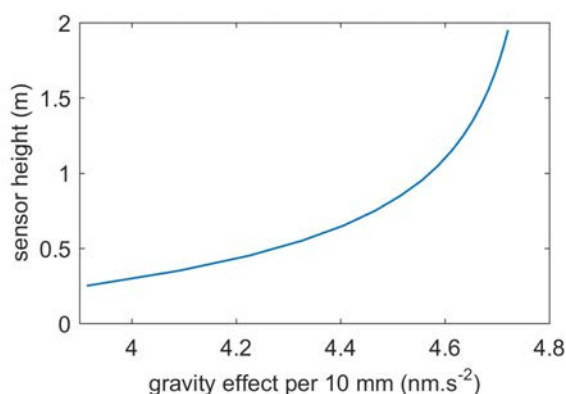
Figure 3.7: Gravity effect of a water storage change of 10 mm in the uppermost soil layer (1 m thickness, uniformly distributed) as a function of the height of the gravity sensor above the terrain surface, for the iGrav location at Wettzell.

Using the prism-based gravity model, the gravity effects on SG030 and iGrav006 for water storage variations in different depths below the terrain surface are shown in Figure 3.8 for different integration radii, i.e., the distance from the gravimeter
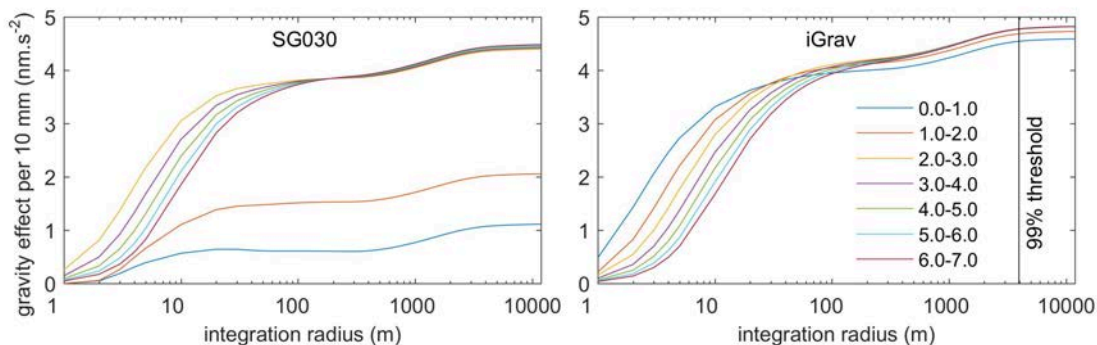
Figure 3.8: Gravity effect of a 10 mm water storage change in different depths below the terrain surface (0 – 7 m depth) at Wettzell, considering the real topography and the umbrella effect of SG030 gravimeter building and the iGrav pillar, respectively (no storage change within 2 m underneath building and pillar).

that is considered for calculating the gravity effect. All soil layers are assumed to be parallel to the surface, follow the topography given by the elevation model for the entire domain, and the water is uniformly distributed inside each layer. Both the umbrella effect and the topographic effect as explained above are considered for these calculations. As a result, soil moisture variations in the near-surface layers (0-2 m) have a considerably smaller gravity effect for SG030 than for the iGrav due to the umbrella effect. For soil moisture variations at larger depths, the iGrav also exhibits a slightly larger gravity effect than SG030 due to the topographic influence, but the difference between the two gravimeters is much smaller than for near-surface mass effects. Similar to Creutzfeldt et al. [2008], Figure 3.8 also shows that most of the local gravity effect originates from a distance of up to about 200 m around the instrument, which is consistent with the simple cylinder model described above. A second-order increase of gravity is due to storage changes at a distance of about 1-2 km from the sensors. These areas have a higher gravity effect due to their lower topographic positions in valleys around the gravimeters that are located on top of a topographic ridge at the landscape scale. Increasing the integration radius beyond 4 km has only a minor impact on either iGrav or SG030 because at this distance the gravity effect reaches 99% of the total effect computed for radius of 12 km. Here, only the mass attraction effect but not the surface deformation caused by large-scale hydrological loading of the Earth crust were considered. The overall higher sensitivity of the iGrav in the field enclosure to water storage changes is also expressed in markedly higher gravity amplitudes of iGrav006 in the residual gravity time series, both for event and for seasonal time scales (Figure 3.3).

The most interesting result of the sensitivity analysis shows up when comparing the gravity effects at the maximum integration radius (or in approximation at any radius at the landscape scale of larger than about 200 m that integrates over most of the gravimeter signal) (Figure 3.8). In this case, for iGrav006, the gravity effect of each layer is almost identical regardless of its depth. For example, the effect of a 10 mm water storage change is 4.6 and 4.8 nm/s^2 for the uppermost layer and the deepest (groundwater) layer, respectively. For SG030, in contrast, the effect is 1.1 and 4.4 nm/s^2, respectively. This means that the iGrav006 in its field enclosure setup is rather insensitive to the depth below the terrain surface where the water storage

change occurs, as the footprint of the monument is rather small and the sensors position sufficiently high. In turn, this means that once the water has infiltrated into the soil and increased the water storage, the vertical redistribution of water by hydrological flow processes does not influence the observed gravity signal, unless the water exits the domain again by evapotranspiration or by lateral flow.

We confirm and illustrate this feature by a virtual experiment using a hydrological model, based on HYDRUS-1D [Šimůnek et al., 2016]. The vertical extent of the model domain of 10 m is discretized into 1 cm intervals. A highly conductive sandy-loamy soil with a saturated hydraulic conductivity of 5.5 *10-4 m/s and a porosity of 37 Vol% was chosen for the entire profile. The boundary conditions were set to "atmospheric" for the upper boundary and "no flow" for the lower one. The model was driven with an artificial precipitation input over a period of 15 days and total sum of 361 mm of rain (24 mm/d). In model run 1, a constant evaporation rate of 12 mm/d was set for the following 15 days. In model run 2 with the same precipitation for the first 15 days, zero evaporation was set for the following 15 days. No groundwater variations were considered in this experiment. The simulated profile soil moisture variations were then converted into gravity effects for the locations of both gravimeters SG030 and iGrav006 using the prism-based approach mentioned above. To this end, the simulated 1D soil moisture variations were transferred to the entire domain and the real topography and building or pillar dimensions of Wettzell including an umbrella effect of 2 m in depth as described above were considered. The simulated profile soil moisture changes and related gravity effects are shown in Figure 3.9. The continuous wetting front advancement to larger depths during the entire experiment is obvious, as well as the drying topsoil layers due to evaporation in model run 1 after day 15.

The storage increase by precipitation and the subsequent decrease by evaporation cause a close to linear gravity increase/decrease for iGrav006 in the field enclosure. The advancement of the wetting front to larger depths and the redistribution of water within the soil profile does not change the gravity signal for iGrav006. This can be clearly seen for model run 2 after day 15 where in the absence of evaporation or precipitation the total water storage in the system remains constant, as does gravity in the case of iGrav006. In contrast, the redistribution of water within the soil profile causes a further increase of gravity for SG030 even without net mass change, because the wetting front advancement moves water from top soil layers with lower gravity sensitivity for SG030 due to the umbrella effect of the observatory building to deeper layers with higher sensitivity. As a consequence, the water mass loss due to evaporation after day 15 in model run 1 is not visible for SG030 as it is masked by the water redistribution in the profile that even causes an increase of gravity during the first days after evaporation kicked in. The complex interplay of (i) the hydrological processes of water redistribution within the profile with (ii) the varying sensitivity to hydrological mass changes in depth due to the umbrella effect causes a non-linear gravity response of SG030 from which it is difficult to disentangle the underlying water storage variations. Similar behaviour was observed during rain events using iGrav002 installed inside a building in the Larzac plateau, France [Fores et al., 2016]. In contrast, the iGrav006 setup in the field enclosure allows for monitoring the variations of total water storage within its sensitivity domain without the need to know the vertical distribution of hydrological mass changes. It is thus an unprecedented means of assessing the landscape water balance in an

integrative way as the net effect of all water inflows and outflows.
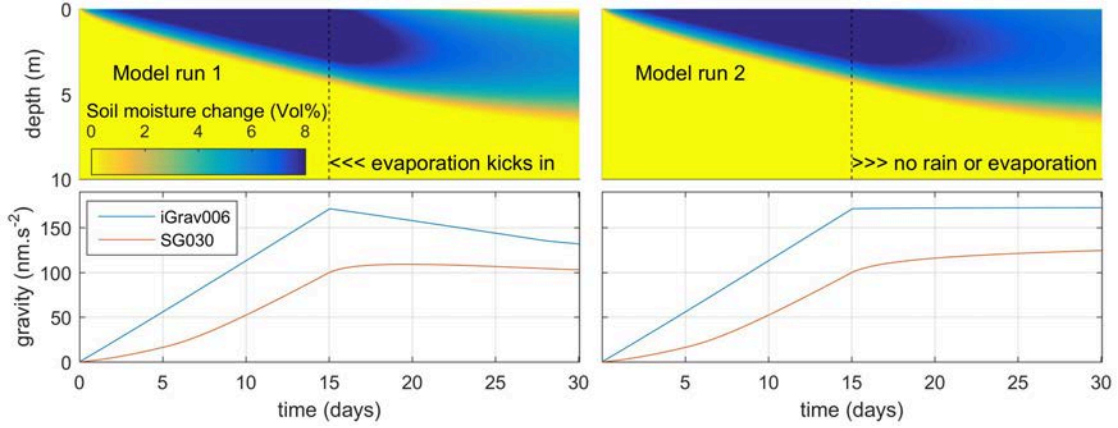
Figure 3.9: Simulated profile soil moisture changes (upper plots) and gravity response for SG030 and iGrav006 at Wettzell (lower plots) for two model experiments, both with artificial rainfall during the first 15 days, and evaporation during the second 15 days (model run 1 only).

3.4.2   Resolving water balance components – annual scale

To demonstrate the value of the gravimeter field deployment for the direct analysis of the landscape water balance and its components, we set up the water balance equation in a way that the left-hand side of Eq.3.3 is water storage change (dS/dt) as given by the change of the iGrav006 gravity residuals (dg/dt). A constant mean sensitivity factor s = 0.478 nm/s^2/mm derived from the above sensitivity analysis (compare Figure 3.8) is used to convert a gravity change into an equivalent water storage change. It should be noted that the gravity residuals dg/dt represent the gravimeter signal that was reduced for non-hydrological mass effects from Earth and ocean tides, Earth rotation and atmospheric variations. Thus, dg/dt still comprises gravity effects () from mass attraction and surface loading effects of non-local (i.e., continental to global scale) water storage variations which have to be removed for the present application. For this purpose, the mGlobe software package [Mikolaj et al., 2016] is used, considering simulated water storage variations on the global scale by four land surface models of the GLDAS system [Rodell et al., 2004].

The right-hand side of the water balance in Eq.3.3 is composed of precipitation (P) minus actual evapotranspiration (E) minus runoff (R) from the area contributing to gravity variations seen by the iGrav006. As introduced in Sect. 2, as a first guess of the vertical fluxes at the land surface, P was taken from local gauge measurements with under-catch correction. For E, the potential evapotranspiration without water limitation in the form of the Penman-Monteith grass reference evapotranspiration (E_{ref}) was taken because it can be quantified from meteorological observables alone. In view of the predominance of grassland and partly forest with high infiltration capacity in the surroundings of the gravimeter location, and based on own field observations during rainfall events, surface runoff is considered to be negligible at the site so that R encompasses subsurface runoff only. Given the specific topographic situation of the observatory on a ridge of the hilly mountain range, negligible lateral subsurface inflow is expected for the site because there is

hardly any upslope contributing area. Thus, R in the water balance equation can be assumed to be dominated by landscape-scale subsurface runoff leaving the headwater area. This runoff can finally be expected to enter nearby rivers that drain the mountain range. Streamflow time series measured at the gauge Chamerau (Sect. 2), converted to specific runoff in mm water equivalent, are thus used as the basis for quantifying the runoff component in the water balance of Eq.3.3:

$$dS/dt = s \times dg/dt - v \times dg_{glob}/dt = u \times P - a \times E_{ref} - c \times R \qquad (3.3)$$

with dS/dt water storage change over time dt, dg/dt change of iGrav006 gravity residuals over time dt, gravity change due to large-scale hydrological variations (by mass attraction and loading), s sensitivity (scale) factor of gravimeter P, E_{ref}, R precipitation, reference evapotranspiration, runoff and v,u,a,c optimization parameters (see plain text for detailed explanations).

Although the problem in Eq.3.3 is linear, the parameters are not linearly independent. The optimization problem was therefore solved by introducing additional parameter constraints and applying a non-linear optimization approach using the interior-point algorithm (Matlab R2015b). To evaluate the statistical match of the daily water storage time series of the left and the right side of Eq.3.3 we apply the Kling-Gupta efficiency (KGE) [Gupta et al., 2009] and use KGE as the performance criterion to be maximized during optimization. The optimization was performed by adjusting the parameters a, c, u and v in Eq.3.3 without explicitly enforcing the closure of the water budget over the analysis period. Setting the closure of the water budget as a constraint would imply that the gravity residuals were not affected by errors, which apparently cannot be assumed so that imperfect instrumental (drift and steps) and gravity (e.g. atmosphere) corrections would directly propagate into the estimated parameter. Table 3.2 shows the a-priori defined parameter ranges. In case of evapotranspiration, the factor a converts E_{ref} into actual evapotranspiration E and, hence, a can vary between 0 and 1. The precipitation factor u accounts for inaccuracies in the under-catch correction. The lower bound is set to 0.9 which approximates precipitation without under-catch correction, the upper bound is set to 1.1 to account for a possible underestimation of the correction. The runoff factor c can be interpreted as a correction for conceptual mismatches (e.g., river runoff at the large catchment scale may not be fully equivalent to the subsurface runoff component considered at the gravimeter scale) and for conversion errors (e.g., inaccurate catchment area). The lower and upper bounds of c were initially set to allow for a maximum change of specific runoff by 10% in both directions. These bounds were further kept for the analysis as they were never reached during the optimization process. In addition, an uncertainty factor v for the contribution of the large-scale hydrological gravity effect on the left-hand side of the equation is included. Given large difference in estimates of this gravity effect when different global hydrological models are used [Mikolaj et al., 2016], varying v during optimization allows for accounting for this uncertainty. The parameter bounds of v were derived from the minimum and maximum multiplicative factors that were needed to convert the large-scale hydrological gravity effect of each of the four different hydrological models used here to the mean effect of the four models (dg_{glob}/dt).

The optimization (period 19 June 2015 – 29 June 2016) resulted in a very high Kling-Gupta efficiency of 0.98, mainly due to the good fit of the dominant seasonal storage variations and no considerable bias (Figure 3.10). An exception with a larger

Table 3.2: Parameters of the water balance equation adjusted during optimization (see text for detailed explanations).

Parameter scope	Name	Parameter range	Optimized value	Independent value from lysimeter
Evapotranspiration	a	$0.00 - 1.00$	0.69	0.68
Precipitation	u	$0.90 - 1.10$	1	1.02
Runoff	c	$0.90 - 1.10$	1.08	-
Global hydrological gravity effect	v	$0.49 - 1.28$	1.28	-

Figure 3.10: Time series of water storage at Wettzell (as a deviation from the initial storage value at the beginning of the study period, arbitrarily set to 0), with the measured iGrav006 gravity-based storage time series (blue line) (left side of Eq.3.3) and the optimized storage time series (red line) after optimization following the right side of the water balance equation (Eq.3.3).

bias is towards the end of the study period where major steps in the gravity time series had to be removed manually prior to optimization (see Chapter 3.2). The gravimeter-based time series tends to miss some of the higher frequency dynamics that exist in the P-E-R time series. This is partially caused by setting a constant gravimeter sensitivity factor s. The mean factor s equals 0.478 nm/s^2/mm, while s equals 0.441 nm/s^2/mm for layer between 0 and 1 cm. Thus, a sudden water storage increase due to precipitation can be underestimated by 8% as long as the precipitated water is concentrated on or close to the surface. Furthermore, the v,u,a,c parameters were optimized for the whole time series, leading to parameter values that primarily reflect the dominant seasonal variations and may underestimate shorter-term storage changes. Nevertheless, the very good overall optimization results can be demonstrated with basic statistics. The correlation coefficient between the optimized storage time series on the left and right side of Eq.3.3 is 0.95, the mean difference is 0.02 mm, and the standard deviation is 13 mm. The optimized parameter values are listed in Table 3.2. The runoff factor c = 1.08 means that runoff from the gravimeter footprint is slightly increased relative to measured streamflow within the optimization procedure. Reasons may include unaccounted groundwater discharge in the valley bottom of the river gauging station or higher than average gradients for subsurface flow processes in the iGrav headwater region. The precipitation factor u = 1.00 indicates that the under-catch correction by Richter (1995) is reasonable for this site so that no further adjustment of precipitation volumes by the

optimization approach is required. The evapotranspiration factor a = 0.69 means that the actual evapotranspiration of the landscape around the Wettzell observatory is about 69% of the (potential) grass reference evapotranspiration. This in turn shows that the hydrological system was at least partly water limited for evapotranspiration during the study period, in spite of its location in a mountainous humid temperate climate regime. One contributing factor is that the period included an exceptional drought in summer 2015 that hit in particular Southern Germany and the area of the Czech Republic [Laaha et al., 2017] where Wettzell is part of. To assess the validity of the latter two factors, we compared them to values derived in a completely independent way from the lysimeter time series at the Wettzell observatory. The corresponding lysimeter-based factor u* was computed as the ratio between lysimeter precipitation (determined from its mass increase during rainfall events) and gauge precipitation corrected for under-catch. The factor a* is the ratio between the lysimeter-based actual evapotranspiration and the grass reference evapotranspiration. The lysimeter-based factors u* and a* for the study period are 1.02 and 0.68, respectively, and thus very close to the gravity-based optimization results (Table 3.2). These results show that the superconducting gravimeter in the field deployment is very well suited for quantifying the annual constituents of the water balance equation. For the 1-year period (mid June 2015 – mid June 2016) in Wettzell, the annual values derived from the gravity-constraint approach for P, E, R and for gravity-based dS/dt were 829, 412, 394 and -8 mm, respectively. Thus, in total, the mismatch of P – E – R versus dS/dt amounted to 31 mm. This value, corresponding to about 4% of annual precipitation, can be considered as the error in closing the water balance at the annual scale by the gravity approach.

3.4.3   Resolving water balance components – daily scale

Quantifying actual evapotranspiration is of particular interest due to the lack of other direct observation techniques at the stand or landscape scales, with the exception of the eddy covariance method [Baldocchi et al., 1988]. The questions arises whether the gravity-based water balance approach presented above can be used to quantify E over shorter periods of time, ideally on a daily basis to assess the landscape E response to changing conditions in terms of meteorological drivers, water availability and the physiological status of plants. The main obstacles are instrumental issues such as noise, the spurious temperature effects on the gravity time series mentioned above, and a deficient correction of non-hydrological effects, e.g. atmospheric and Earth tides or ocean loading, that partly exhibit a similar daily period as E. Van Camp et al. [2016b] recently needed to stack the gravity time series over several periods without rainfall to isolate a mean daily value for E. For the optimization presented here, we thus solve the water balance within moving windows of several days in length instead of using day to day differences that are particular sensitive to the noise components mentioned above. Different from the optimization described above, only the evapotranspiration factor a is adjusted (here in a time-variable way, i.e., for each moving window) while the other three factors are taken as constant values as derived in the optimization before (Table 3.2). Feeding the moving window with an a-priori estimate of E_{ref} acts like a physical low-pass filter that minimizes the noise of physical or instrumental origin. The longer the window, the higher the reduction of noise. On the other hand, however, a longer window decreases the

accuracy at the daily level. To assess the quality of the window-based optimization, we test the following question: Does this method of scaling E_{ref} within in a moving window approach result in actual evapotranspiration at the daily scale? To answer this question, we run a moving window optimization for the time-varying factor a using the E_{ref} time series as input and the observed time series of lysimeter E as the target. Then, daily E is estimated by multiplying a with E_{ref} for the central day of the considered window. The difference between estimated and actual lysimeter evapotranspiration is the error of the method. The root mean square error (RMSE) for windows of 9, 11 and 13 days in length is 0.16, 0.17 and 0.18 mm/d, respectively. A further increase of the window length gradually degrades the accuracy. This shows that the method itself is indeed capable of resolving daily evapotranspiration rates with sub-mm accuracy when setting the window to a reasonable length.

The results of the final time-varying optimization of a in Eq.3.3 with fixed factors for precipitation, streamflow and gravity-based storage change are shown in Figure 3.11 for a moving window length of 11 days. The estimated actual evapotranspiration fits the lysimeter observations well, both with regard to the magnitude of daily E rates as well as their temporal variations. In particular during the summer season (July-August), estimated E usually is considerably smaller than E_{ref}, as also indicated by comparatively small values of a. This demonstrates the water-limited state of the hydrological system during this period. E (both from gravity and from lysimeter) tends to be close to E_{ref} during the autumn season with overall much smaller daily E rates.

An exception of the overall good performance is the period between August 22 and August 30 where estimated daily E rates are always equal to E_{ref} and systematically overestimate observed lysimeter E (Figure 3.11). Within this time period, the RMSE of estimated versus lysimeter evapotranspiration is 0.92 m/d, while it is 0.42 mm/d outside of this interval. This discrepancy is related to a strong decrease of storage as given by the gravimeter in this period, with an average of 4 mm/d. Unrealistically high runoff rates of more than 1 mm/d would be required to match these storage change rates with the observed E of the lysimeter which is in the order of 2.5 mm/d. Hence, the apparent storage decrease is probably related to insufficient correction of gravity effects in the iGrav time series that are not related to local water storage. However, the strong storage decrease is neither reduced by the use of different global model for the large-scale hydrological effects, nor by the inclusion of non-tidal ocean loading effects or by a change of gravimeter drift rate. We furthermore checked possible atmospheric effects, motivated by the fact that the atmosphere contains the evaporated mass which affects the gravity with opposite sign due to the change of its position relative to gravimeter sensor (from the soil below to the atmosphere above the instrument). The applied 3D Atmacs correction was replaced by (i) a 3D correction using the ERA Interim model [Mikolaj et al., 2016] and (ii) a simple approach using in-situ observed air pressure solely instead of a full 3D field. None of these changes, however, affected the E estimations in this dry period in a considerable way so that the reason for the discrepancies remains unresolved. Large-scale atmospheric mass increase by evaporated water during the strong evapotranspiration period that could not be included in atmospheric models nor in the local air pressure observations remain an unproven hypothesis. It should, however, also be noted that the observation data taken here from the lysimeter are very different in spatial scale than the E effect seen by the gravimeter. Differences

in E between the lysimeter and the landscape scale may thus also contribute to the differences in the time series. This is not necessarily a limited performance of the gravimetry-based approach, but an expression of its better suitability for quantifying landscape-scale hydrological dynamics.
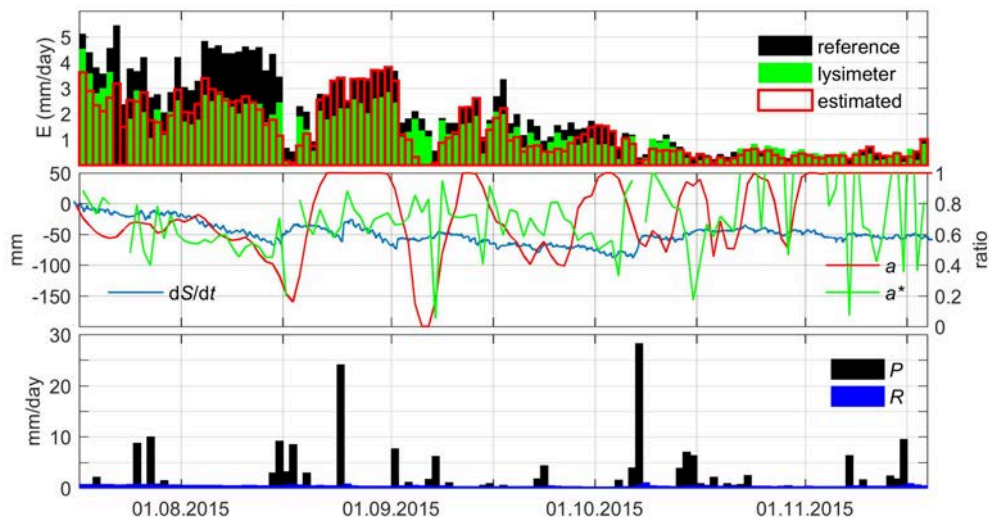
Figure 3.11: Upper panel: Comparison of (Penman-Montheith grass) reference evapotranspiration, observed lysimeter-based actual evapotranspiration (zero values of lysimeter E indicate missing data), and estimated actual evapotranspiration (from gravity-constraint optimization). Middle panel: Gravity-based storage anomaly relative to 10.07.2016, the optimized factor a of the water balance equation (ratio of gravity-based E to E_{ref}) and the factor a* based on the lysimeter time series (ratio of lysimeter-based E to E_{ref}). Lower panel: Input time series of precipitation and runoff.

3.5 Conclusions

Observing mass budgets in environmental systems is a fundamental challenge and rarely possible in a comprehensive way due to the multitude and complexity of pools and fluxes involved. This also applies to the hydrological cycle when it comes to monitoring the water balance at spatial scales in between the point and the river basin scale. With the field observation technique presented here, based on continuous gravimetry with a superconducting gravimeter (SG), an unprecedented means of directly monitoring the water balance at the $100 - 1000$ meter scale becomes available.

By deploying a SG in a small field enclosure, we demonstrate for the first time that a continuous and stable outdoor operation of a SG is feasible for a long time (here more than one year) under humid environmental conditions with marked daily and seasonal temperature variations. At the same time, the quality of the gravity time series does not degrade in comparison to the standard SG deployment under controlled conditions in observatory buildings. The field enclosure design proves to shield the instrument sufficiently well from temperature variations, wind pressure or other environmental effects that may cause vibrations, instrument tilt or other spurious effects. Thus, the tiny hydrology-induced gravity signal of interest is only

marginally obscured by instrument noise. We show that the deployment of the SG in a field enclosure conveys other advantages relative to the existing SGs in observatory buildings. First, being spatially closer to the signal of interest, the field SG is more sensitive to local water storage variations and it is not affected by unnatural and usually unknown storage variations below a building. Secondly, we demonstrate that the gravity residual time series of the field SG are a direct expression of the total water mass change in the surroundings of the instrument, almost independent of the depth below the terrain surface where the storage changes occur. Thus, with the field SG, we present the first continuous and integrative monitoring technique of the landscape water budget. It should be noted that this conclusion applies if the storage changes occur within the full integration radius of the SG (few hundreds of meters) with low horizontal heterogeneity. In contrast, spatially constrained water storage changes, such as in the case of an artificial sprinkling experiment in the vicinity of the gravimeter, result in a sensitivity of the SG to the depth of the storage change, and can be used for identifying the infiltration process. Also, we point out that the gravimeter setup with a small enclosure on a tall pillar as presented here may not be most appropriate for groundwater (saturated zone) applications, such as assessing groundwater recharge. In this case, moisture variations in the unsaturated zone are unwanted signal that needs to be removed in order to identify recharge. Thus, the increased near-surface sensitivity of the present setup is a disadvantage for such applications.

With the gravity monitoring system presented here, we show that the annual water balance can be closed within 4% of annual precipitation. This error results from imperfect reduction of mass signals other than the local hydrological ones in the gravity time series, and of instrumental effects such as drift and steps. We provide a framework to quantify the individual components of the water balance from the gravity observations at annual to daily time scales. Notably, expanding the potential as pointed out by Van Camp et al. [2016b], we demonstrate the value of the field SG as a technique for assessing actual evpotranspiration. The accuracy of the approach when evaluated against daily ET rates from lysimeter time series was in the order of 0.5 to 1 mm/d, but the different spatial footprints of both methods limit their direct comparability. In turn, if reasonable data of actual ET were available from other observation techniques such as the eddy covariance method, a collocated field SG will offer a unique means of estimating subsurface runoff via the gravity-based water balance approach presented here, even for a thick unsaturated zone or deep groundwater tables.

From a practical perspective, and compared to other hydrological field monitoring techniques, the widespread and flexible deployment of the field SG system proposed here may currently be hampered by the need of a solid gravimeter monument, the weight and complexity of the monitoring system, the power requirements, and by its costs. Nevertheless, this study lays out the potential of high-precision gravimetry in the field as a non-invasive observation method that fills gaps in the spectrum of existing hydrological and hydro-geophysical methods with respect to the target observables and the spatial scale to be captured. Ongoing technological development towards smaller gravimeters including alternative techniques for high-precision gravity measurements such as quantum gravimeters give prospect for a much broader future application of the hydro-gravimetric principles developed here.

Chapter 4

Hydrological subsurface flow processes at a field site

This chapter has been published as:

FIELD-SCALE SUBSURFACE FLOW PROCESSES INFERRED FROM CONTINUOUS GRAVITY MONITORING DURING A SPRINKLING EXPERIMENT

Marvin Reich, Michal Mikolaj, Theresa Blume, Andreas Güntner

Abstract

Field scale subsurface flow processes are difficult to observe and monitor. We investigated the value of gravity time series to identify subsurface flow processes by carrying out a sprinkling experiment in the direct vicinity of a superconducting gravimeter. We demonstrate how different water mass distributions in the subsoil affect the gravity signal and show the benefit of using the shape of the gravity response curve to identify different subsurface flow processes. For this purpose, a simple hydro-gravimetric model was set up to test different scenarios in an optimization approach, including the processes macropore flow, preferential flow, wetting front advancement, bypass flow and perched water table rise. Besides the gravity observations, electrical resistivity and soil moisture data were used for evaluation. For the study site, the process combination of preferential flow and wetting front advancement led to the best correspondence to the observations in a multi-criteria assessment. We argue that the approach of combining field-scale sprinkling experiments in combination with gravity monitoring can be transferred to other sites for process identification, and discuss related uncertainties including limitations of the simple model used here. The study stresses the value of advancing terrestrial gravimetry as an integrative and non-invasive monitoring technique for assessing hydrological states and dynamics.

4.1   Introduction

Subsurface flow processes govern water re-distribution within the soil, i.e., percolation to larger depths as well as lateral fluxes. Rainfall characteristics, vegetation water use, surface properties and the soil physical characteristics define how water propagates through the subsurface. Common processes include wetting front advancement, preferential flow paths or macropore infiltration, bypass flow or the development of a perched water table [Beven and Germann, 2013]. These subsurface flow processes together with infiltration, evapotranspiration and surface runoff are key components of the hydrological cycle. Their understanding is essential for assessing water storage dynamics at different scales which in turn impacts agriculture, water supply, natural hazards and engineering, amongst others.

Whereas infiltration can be measured manually at small spatial footprints by standard hydrological field methods (e.g. ring infiltrometer, amoozemeter, hood infiltrometer), process dynamics in the subsurface are more difficult to assess. At the plot or hillslope scale, sprinkling experiments have been used to look at characterizing flow processes [McGlynn et al., 2002; Wienhöfer et al., 2009; Angermann et al., 2017], control of boundary conditions [McDonnell et al., 2007; Wienhöfer and Zehe, 2014], subsurface structures [Jackisch et al., 2017], small-scale infiltration patterns [van Schaik et al., 2010], large scale conductivity and event water fractions [Jackson et al., 2016].

While sprinkling experiments capture more of the horizontal and vertical spatial heterogeneity of most soils and sites than the point-scale methods, monitoring of sprinkling experiments usually involves invasive methods such as profiles of soil moisture sensors, trenches for subsurface runoff collection, or excavations after the experiments to identify flow patterns in the case of dye tracing. We therefore suggest gravimetry as a promising alternative allowing for non-invasive and integrative monitoring of the mass variations caused by water storage changes (WSC). With precise corrections of the measured gravity signal (for Earth and ocean tides, atmosphere, polar motion and instrumental drift; see Crossley et al. [2013] for an overview), the gravity residuals predominantly reflect near-field hydrological mass variations, including infiltrating and percolating water.

Only few studies have so far exploited gravity measurements with a focus on infiltration and subsurface processes. Pool and Eychaner [1995] concluded that temporal gravity surveys can help to estimate aquifer storage changes and specific yield. Christiansen et al. [2011b] evaluated the usefulness of a relative spring gravimeter (Scintrex CG-5) for monitoring WSC in small-scale field campaigns using a closed laboratory basin with controlled water table heights as an approximation. Using drip irrigation on a 10 m by 10 m field site, Christiansen et al. [2011a] combined relative gravity data, using again a Scintrex CG-5, with ground penetrating radar measurements in a coupled 1D inversion approach to calibrate van Genuchten parameters. Kennedy et al. [2016] inferred the infiltration rate and wetting front depth by measuring relative gravity at various locations near an artificial recharge facility. Their results indicated that the accuracy of the gravity measurements is the key factor for determining parameters that are relevant for the infiltration and percolation process (hydraulic conductivity, specific yield, Brooks-Corey exponent). In an earlier study, Kennedy et al. [2014] stated that a multi-gravimeter approach with gravimeters placed at different distances from the recharge site is favourable to a

single-gravimeter approach in terms of identifiability of the velocity of the wetting front as well as reduction of effects which are not related to the mass change of interest. Furthermore, such an approach enables the possibility to localize mass changes, depending on the magnitude of the increase compared to the sensitivity of the instrument and correction-induced uncertainties [Mikolaj et al., 2019b]. However, the multi-gravimeter setup is significantly more time and cost intensive. Apart from these controlled infiltration setups using artificial recharge basins, even less studies used gravimetry to study more natural infiltration conditions. Pool [2005] looked at channel recharge, using hydrological data in combination with repeated gravity surveys and climate indicators. The main reason for the low number of studies on this subject, besides the limited accuracy of field gravimeters in the past, is the difficulty to localize mass variations in space due to the ambiguities originating from the integrative nature of gravimetry. Güntner et al. [2017] showed that gravity measurements of field gravimeters are close to insensitive to the depth of the water storage change if this change is spatially uniformly distributed across the measurement footprint (simliar to a Bouguer slab). This is not the case for a very local, spatially limited, water mass increase from a sprinkling experiment. Here, the gravimeter is sensitive to the actual location of the infiltrated water, as a function of its distance to the gravimeter both in vertical and horizontal direction.

In this study, we assess whether gravimetry is a suitable non-invasive technique for distinguishing different infiltration and subsurface flow processes. To this end, we monitor the gravity signal during sprinkling experiments with a superconducting gravimeter in a field enclosure in the centre of the experimental plot. First, using virtual experiments with a hydro-gravimetric model, we explore the theoretical subsurface mass re-distribution of idealized, single infiltration processes and their resulting gravity responses. Then, applying the model in an inverse mode and in combination with an optimization approach, we identify the dominant subsurface flow processes for a real-world sprinkling experiment and evaluate the results with a multi-criteria approach of independent observations from electrical resistivity transects and soil moisture measurements.

4.2 Methods

4.2.1 The hydro-gravimetric model

As water infiltrates into the soil during rainfall or sprinkling, it is added to the subsurface and thus results in a local mass change that can be observed by a gravimeter. To translate water storage changes in the subsurface into gravity signals and vice versa, we developed a simple grid-based model [Reich, 2021], using the software R [R Core Team, 2020; Wickham et al., 2020; Wickham, 2016, 2007], that converts water mass changes within each cell of a 3D grid into their gravity signals as sensed at the location of the gravity sensor (figure 4.1). After defining the grid extent and resolution to fit the planned experiment, the static gravity effect grid with cell-specific conversion factors is calculated using the prism approach [Forsberg, 1984]. Grid cells without mass variations in the area of the concrete base of the gravimeter are set to zero mass change. In the next step, a sequence of water storage grids is generated based on the temporal dynamics of the idealized subsurface flow processes during the sprinkling experiment, defined by a) the lateral extent of the infiltration

area, b) the spatial pattern of artificial rainfall intensity taken as infiltration rate for the grid cells at the surface and determined using a grid of collecting beakers, c) the spatial extent of infiltration (defined by the parameter called *process depth* or *clustering factor*, explained below), d) the maximal allowed saturation deficit of each cell (*max* $\Delta\Theta$, i.e. the available pore space that can be filled with water until saturation is reached), and e) the sprinkling water volume input per time step. For each time step, every water storage change value of a grid cell is multiplied with the corresponding cell-specific conversion factor. The sum of all these multiplied cell values results in the total gravity change at the location of the sensor for the respective time step.

Besides applying the hydro-gravimetric model in the forward modelling mode as explained above, the model can also be applied in an inversion mode to identify the water distribution and subsurface flow processes that fit the observed gravity time series (see section 4.2.6). In an extended version of the hydro-gravimetric model, lateral flow is included as a complementary process of subsurface water redistribution. For this, a supplementary factor is needed (lateral redistribution factor), defining the percentage of water which is passed vertically or horizontally to neighbouring cells, once the saturation deficit of the cell is zero. The routine of lateral water distribution includes all neighbouring cells with a water content smaller than their maximal allowed saturation deficit. The amount of water for orthogonal and diagonal cell neighbours is scaled with 0.6 and 0.4, respectively, according to the multiple flow direction algorithm by Quinn et al. [1991].
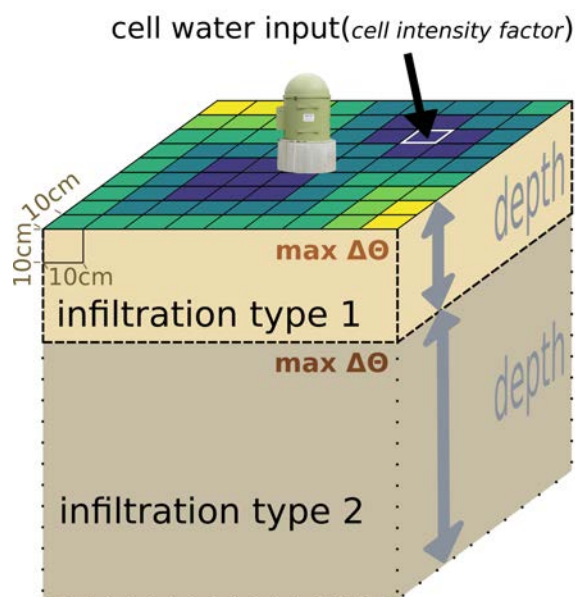
Figure 4.1: Structure of the model domain (with a 15 x 15 m area and 5 m depth) and necessary input parameters. "max $\Delta\Theta$" refers to the saturation deficit of each grid cell. The cell water input depends on the total amount of water for the sprinkling area per timestep, multiplied with the cell intensity factor that represents the spatial distribution of water within the sprinkling area. The depth of each infiltration process can be varied manually or within the optimization algorithm. The small green dome is the gravimeter enclosure on a concrete pillar. The figure shows, as an example, the combination of a shallow and a deep infiltration type, for a heterogeneous sprinkling water distribution on the surface.

4.2.2   Virtual Experiments of individual infiltration- and sub-surface flow processes

In order to characterize the gravity responses of different subsurface flow types, virtual experiments were run with the hydro-gravimetric model. Process types considered here are: wetting front advancement, bypass flow, perched water table, preferential flow and macropore infiltration. These process representations are strongly simplified in the model, resulting in a certain limitation of interpretation. Wetting front advancement (WFA) refers to a classical view of subsurface water redistribution as for example described by Green and Ampt [1911]. The water content of a grid cell is increased until its saturation deficit reaches zero. When this is the case, the wetting front advances towards greater depths, increasing the water content of the next lower grid cell. This way water propagates downwards sequentially (figure 4.2). WFA starts either at the surface (figure 4.1, infiltration type 1) or directly below another process (figure 4.1, infiltration type 2). To simulate bypass flow [Beven and Germann, 2013], the water is directly routed to a defined soil depth (process depth), from where it advances downwards following the wetting front advancement concept. For the bypassed soil zone no change in water storage is assumed. For the perched water table scenario [Todd and Mays, 2004], water is also routed directly to a certain soil depth (process depth) which is assumed to be impermeable, and ponds on this layer until saturation is reached, after which it advances upwards into the upper-next grid cell layer. In the case of macropore infiltration [Beven and Germann, 2013], soil pores are filled with water and the water moves laterally from the pore into the matrix. For the virtual experiments, macropores and exfiltration into the matrix are combined into one single, idealized process and assumed to be uniformly distributed across the entire horizontal modelling domain (example of figure 4.1, infiltration type 1). For the macropore scenario, the water content of all cells within the macropore layers are increased simultaneously. Preferential flow [Beven and Germann, 2013] is represented in the model in a similar way as macropores, with the difference that water can only infiltrate into certain columns of the model domain and not over its entire horizontal extent. The clustering factor defines how many laterally adjacent grid cells are combined together where each column extends over the entire depth of the model domain. This results in a chessboard-like pattern with an equal number of vertical structures of a) preferential flow path cells and b) no-flow cells. In the preferential flow path cells, the water content increases simultaneously until the saturation deficit reaches zero. As for the gravity response of the different subsurface process scenarios, the most important factor is the distance between the mass change (water content increase) and the gravity sensor. For the macropore and preferential flow scenario this distance is fixed, whereas the distance increases for WFA and bypass flow with the downward advancement of the wetting front and decreases for the perched water table scenario with its rise during the sprinkling experiments. Virtual experiments were conducted for each process type, a first set with vertical flow only and a second one including lateral flow. In order to receive a clear and unique response for each process type, either a lot of water or a long sprinkling time is necessary. Figure 4.2 shows exemplary water distributions and their progression over time for these virtual experiments.
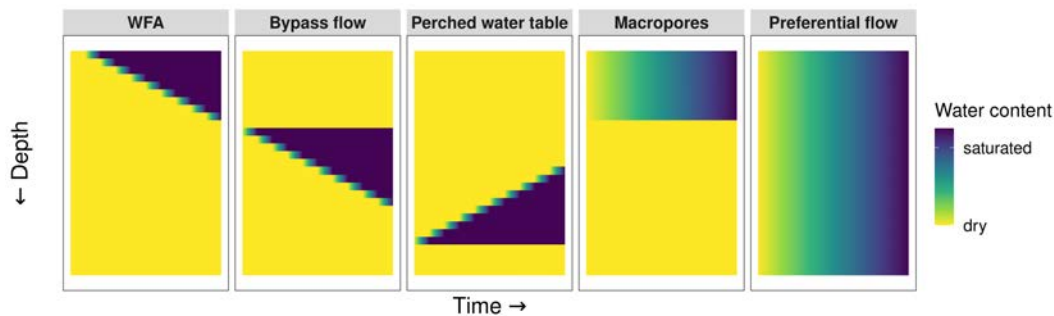
Figure 4.2: Schematic representation of the virtual experiments, showing water content with depth and over time for each of the subsurface flow processes (no lateral flow considered). Wetting front advancement is abbreviated with 'WFA'. For preferential flow, one preferential flow column is shown as an example while there is no wetting in a neighbouring column.

4.2.3   Field Site and Real-World Sprinkling Experiments

The sprinkling experiment was carried out at the Geodetic Observatory Wettzell (operated by the German Federal Agency for Cartography and Geodesy) in south-eastern Germany. The surroundings of the observatory in this mid-mountainous region of the Bavarian Forest are dominated by grassland and forest, with some interlaced agricultural fields. The mean annual temperature is 7°C, annual precipitation amounts to 863 mm and potential evapotranspiration to 433 mm (climate station Höllenstein-Kraftwerk $1947 - 2005$). The soil at the study site consists of loamy-sandy brown Cambisol, underlain by a saprolite layer of weathered gneiss, fractured gneiss and gneiss as the underlying basement Creutzfeldt et al. [2010a]. At the area of the sprinkling experiment, groundwater levels of the unconfined and fractured aquifer are between -5.7 and -7.4 m. The sprinkling site itself, with the gravimeter located in its center, is a flat terrain grassland area (see figure 4.3).

On this sprinkling site, 8 sprinkling units of the type miniWobbler [Jomaa et al., 2010] were arranged in a square (15 m x 15 m) around the gravimeter. A fire hydrant at a distance of about 150 m from the sprinkling area was used as water supply. Water volume over time was measured, using a conventional water flow meter, independently in two subsystems consisting of 4 sprinkler units each (see figure 4.4). The sprinkling experiment was conducted on August 26 of 2016 during nighttime with a duration of 6 hours and a total sprinkled water volume of 14.5 m^3. With 21%VWC as an average over all sensors, the antecedent soil moisture content prior to the experiment was comparatively low with respect to the yearly range of values for this site.

4.2.4   Monitoring

Sprinkling Water Distribution

In order to know the real distribution of the sprinkled water, 24 beaker glasses (250 ml, gray triangles in figure 4.4) were distributed on the plot, based on the theoretical distribution of water due to the setup of the sprinkling units. Frequently recorded weights of all 24 beaker glasses during the experiment (by quickly removing beakers, weighing them and placing them back in their original location) were then

Figure 4.3: Setup of the sprinkling experiment at Wettzell Observatory (Bavaria, Germany), with an iGrav gravimeter in its field enclosure (green dome, center), sprinkling units on tripods and electrode profiles for electrical resistivity monitoring (orange).

interpolated via inverse distant weighting (IDW). As statet by Yao et al. [2013]; Chen et al. [2017], IDW is the method best suitable for this task. Errors of interpolating precipitation or soil moisture data with IDW have been quantified, albeit at larger spatial scales, with an RMSE in the range of 3 - 13% [Yao et al., 2013; İmamoğlu and Sertel, 2016; Chen et al., 2017]. We then use the interpolated values, normalized by their total amount, to derive the spatial distribution of irrigation intensity over the sprinkling area at the resolution of the grid used for the hydro-gravimetric model. This approach further reduces the error introduced by the individual measurements as we do not use the absolute values but only their spatial pattern based on the relative differences.

Gravity

The iGrav superconducting gravimeter (#006, GWR) was installed on a 2 m deep concrete pillar, 0.80 m thereof above the ground, with a diameter of 1 m. Taking into account the sensor height within the gravimeter (0.25 m), the absolute gravity sensor height above the ground is 1.05 m. This newest generation of superconducting gravimeters [Warburton et al., 2010] can be operated in the field with a small field enclosure, ensuring a controlled environment of the sensor. All details about the gravimeter setup, instrumental parameters as well as observations over one year of operation in Wettzell can be found in Güntner et al. [2017]. To eliminate all known sources of gravity variations not related to the sprinkling experiment, the observed gravity was corrected for Earth and ocean tides, polar motion and atmospheric effects. The latter effect was corrected using the Atmacs model [Klügel and Wziontek, 2009], while polar motion was modeled using IERS Earth rotation parameters (International Earth Rotation and Reference Systems Service, `www.iers.org`). The tidal signal was removed using a local tidal model derived from a harmonic analysis of long-term gravity measurements by another superconducting gravimeter (SG029) located at a distance of about 200 m from the field gravimeter. The modeled tidal and atmospheric effect was also used to calibrate the gravimeter, see Güntner et al. [2017] for details. The instrumental drift was removed assuming a linear trend as given in Güntner et al. [2017]. Thanks to the short duration of the sprinkling experiment, possible errors due to the instrumental drift can be neglected. The same holds true for the large-scale continental water storage variation effect. For the same
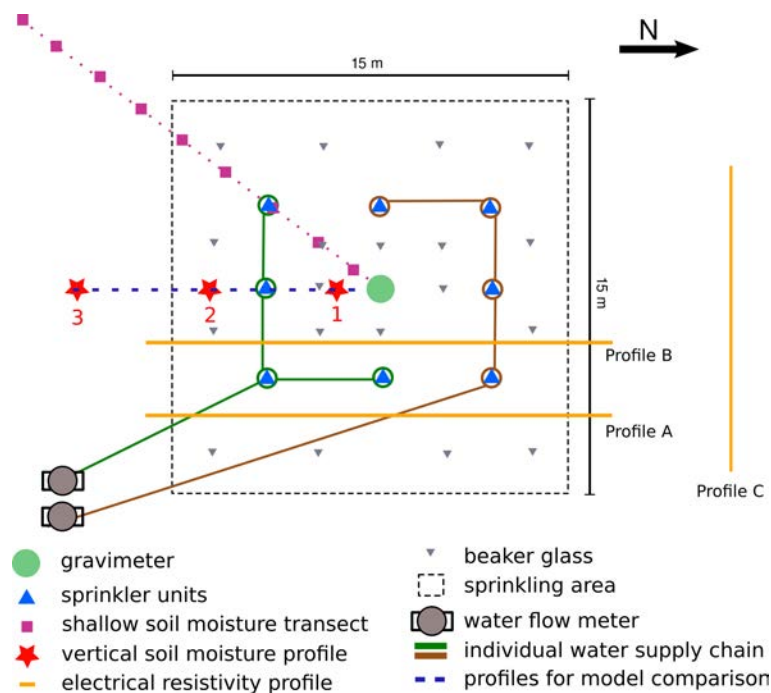
Figure 4.4: Setup of the sprinkling experiment including sprinkling unit layout and monitoring equipment (except superficial TDT soil moisture sensors). See text for a detailed explanation of each measurement setup and properties.

reason, a reduction of the gravity time series for attraction and loading effects by large-scale continental water storage variations was not carried out.

After eliminating all the effects mentioned above, the obtained residual gravity time series, hereafter called "gravity signal", represents the mass changes due to local hydrological dynamics.

Soil Moisture

In south-western direction from the gravimeter, a transect of 10 soil moisture sensors (measurement principle: Frequency Domain Reflectometry (FDR), Campbell Scientific, 1-3% accuracy) at a depth of 5 to 10 cm below the surface and a distance of 1 m between the sensors was set up (purple squares in figure 4.4). South of the gravimeter, 3 vertical soil moisture profiles were installed (red stars in figure 4.4). Each profile contains 3 FDR sensors at the depths of 10, 20 and 30 (one sensor at 40) cm below the surface. All FDR sensors measure soil moisture at 1 minute intervals. Superficial soil moisture was measured at 30 locations (measurement principle: Time Domain Transmissometry (TDT), TOMST, 1% accuracy). They are randomly distributed on the sprinkling area and its near surrounding, recording data every minute within an integration radius of 15 cm.

Electrical Resistivity Tomography

Electrical resistivity tomography (ERT) was used as a second non-invasive method to estimate the spatio-temporal distribution of water in the subsurface of the sprinkling area from measured 2D resistivity profiles. For this purpose, two parallel ERT profiles (A and B) were set up on the sprinkling area and third one as a reference

profile (C) outside of the sprinkling area (orange lines in figure 4.4). Profiles A and B each consist of 75 electrodes, laid out with a spacing of 0.25 m, resulting in a total profile length of 18.75 m. Profile C has 50 electrodes with the same spacing and thus a total length of 12.5 m. The measurements were carried out using a GeoTom (GEOLOG, 0.5% system accuracy as stated by manufacturer), recording data with a Wenner array for each profile at hourly intervals. This temporal resolution had to be chosen in order to meet the required measurement time for all 2D profiles points and due to the fact that no simultaneous measurements are possible when using only one GeoTom device. For final results, apparent resistivity values were inverted using the software BERT [Rücker and Günther, 2011] in time-lapse mode.

4.2.5   Uncertainties inherent to Sprinkling Experiments

Uncertainties inherent to the sprinkling experiment and their impacts on the observed gravity response are discussed in the following. We examined four sources of uncertainty: Quantification of the total sprinkling water volume, wind causing non-uniform input, uncertainty in the spatial distribution of sprinkled water and uncertainty in determining the outer limits of the sprinkling area. To assess their potential impact on the gravity signal the model was run with a range of different scenarios describing these uncertainties. (a) The error of measuring the total water volume used during an entire sprinkling experiment originates in imprecise flow meters and/or losses in hoses and connectors. To estimate the impact of this error, different possible error values are considered, based on device-specific experiences ($\pm$ 5, 10 and 15% of the registered volumes). (b) To estimate the influence of wind, a constant wind direction with a constant wind velocity is assumed, causing a spatial displacement of sprinkling water. The resulting modified spatial distribution of sprinkled water and the previously used undisturbed one are compared in order to asses its impact. (c) The spatial distribution of the sprinkling units can lead to heterogeneous water input intensities. We assess the impact of two variations: (c1) a theoretical distribution of intensities provided by the manufacturer of the sprinkling units, based on number and location of the units and (c2) a distribution based on measured and interpolated intensities from a real sprinkling experiment. (d) Due to the issues mentioned above for (b) and (c), the exact boundaries of the area that receives input of sprinkling water are not well defined. This delineation of the experimental area, however, is important for calculating the gravity signal. We discuss some suggestions of how to deal with this issue. For each uncertainty case and each subsurface flow process type, the assumed spatio-temporal distribution of sprinkling water was taken as an input for the hydro-gravimetric model and the corresponding gravity response was calculated.

4.2.6   Model Optimization for Identifying Dominant Subsurface Flow Processes

To identify the best-matching spatio-temporal distribution of water in the subsurface of the experimental site, the hydro-gravimetric model (see section 2.1) was run in the optimization mode with the following input: total sprinkled water volume, water input intensity distribution on the sprinkling area and observed gravity signal. For the inversion, the parameters of the subsurface process grid are optimized by

fitting the modeled gravity response to the observed gravity signal. As objective function the modified Nash-Sutcliffe-Efficiency (mNSE) [Legates and McCabe Jr., 1999] was used, ranging from 1 (perfect fit), via 0 (model as good as mean values of observations) towards -Inf (mean values of observations better than model). The optimization routine was used from the R-package "ppso" [Francke, 2016]. Different types of infiltration dynamics were tested for their ability to match the observed gravity signal: (i) individually without considering lateral flow processes, (ii) individually including lateral flow and (iii) in combinations, with a maximum of two infiltration processes per model run. All used parameter ranges are listed in table 4.1.

Table 4.1: Parameter ranges for the optimization algorithm of the hydro-gravimetric model. Allowed parameter ranges are given in columns "value max" and "value min". The parameter value used by the algorithm in the first iteration step is given in column "initial value". For a description of the parameters see text in section 4.2.1 and 4.2.2.

Scenario	Parameter	value max	value min	initial value	unit
Single process	max $\Delta\Theta$ of cells	0.25	0.05	0.15	% VWC
	process depth	4.5	0.1	-1	m
	clustering factor	1	7	3	-
Single process, Including lateral flow	max $\Delta\Theta$ of cells	0.25	0.05	0.15	% VWC
	process depth	4.5	0.1	-1	m
	clustering factor	1	7	3	-
	lateral redistribution factor	0	1	0.4	-
Two processes combined (process 1 + process 2)	max $\Delta\Theta$ of cells for process 1	0.25	0.05	0.15	% VWC
	process 1 depth	1.5	0.1	-0.5	m
	clustering factor	1	7	3	-
	max $\Delta\Theta$ of cells for process 2	0.25	0.05	0.15	% VWC
	process 2 depth	4.9	1.5	-1.5	m
	lateral redistribution factor	0	1	0.4	-

4.3   Results and Discussion

4.3.1   Theoretical Characteristic Gravity Response of Individual Subsurface Flow Processes

Forward modeled gravity responses of the 5 different scenarios of subsurface flow processes reveal characteristic shapes of the response curves (figure 4.5). Three types of shapes can be identified: linear, concave and convex. The shape is linked to the spatial-temporal dynamics of subsurface water storage, namely the distance of the mass change with respect to the gravity sensor. In the case of macropore and preferential flow scenario, this distance is fixed so the increasing soil moisture content leads to a linear response in the gravity signal. The modeled gravity increase of the macropore scenario as compared to the preferential flow scenario is larger because the water mass is located closer to the sensor. For wetting front advancement and bypass flow, water infiltrates deeper into the soil over time and the distance of the mass change to the gravity sensor increases. As gravity components are inversely proportional to the squared distance between mass change and sensor, the relative gravity signal decreases over time as the mass moves deeper into the soil and away from the sensor. In the curve shown in figure 4.5, this results in a convex shape. The

opposite is true for water redistribution following a perched water table scenario: decreasing distance over time between mass change and sensor leads to a concave shape of the response curve. The absolute values of the total gravity signal of the individual processes differ, however, subsurface flow process identification is mainly based on the typical shape of the response curve.
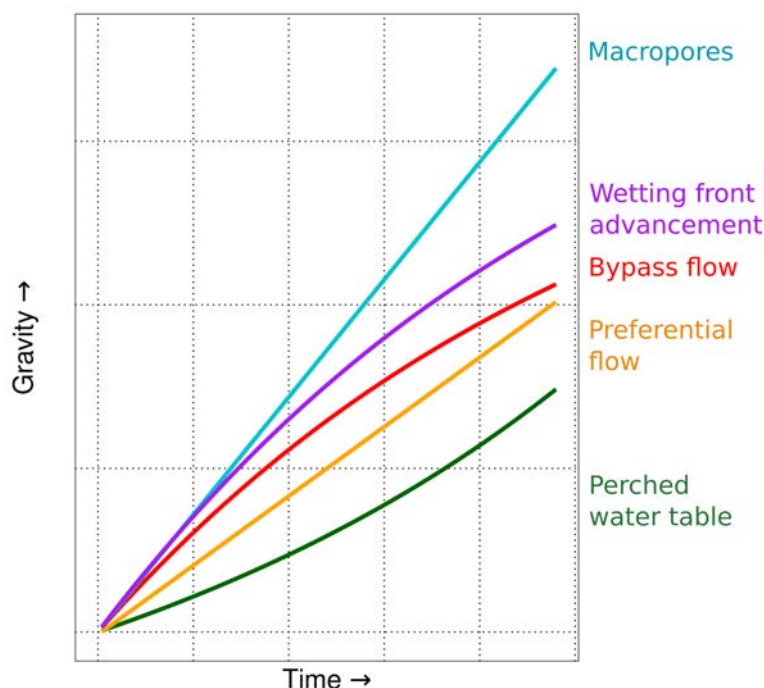
Figure 4.5: Example gravity response over time for single, isolated subsurface flow processes. For the process descriptions, see section 4.2.2.

4.3.2   Uncertainties inherent to Sprinkling Experiments and their Impact on the Gravity Response

To investigate the influence of external factors likely to impact field experiments, the input of the hydro-gravimetric model was modified according to the possible variations linked to the three possible error sources, laid out in section 4.2.5: (i) The measurement error of the total sprinkling water volume influences the resulting gravity response in a linear way: the modeled gravity values will differ from the observed ones in absolute values but will not influence the shape of the response curve. An error will consequently not influence the identification of the dominating infiltration dynamics and subsurface flow processes (figure 4.6a). (ii) As for the influence of wind, blowing steadily in one single direction, the resulting mass redistribution of water in the soil according to individual infiltration dynamics leads to a linear influence on the modeled gravity response (figure 4.6b). Again, absolute modeled gravity values are influenced but not the shape of the response curve. (iii) When assessing the influence of the geometry of the distribution of sprinkling units on the sprinkling area, we found slightly different results. Firstly, considering a theoretical superposition of water concentrations based on sprinkling unit properties and their geometry (exemplary shown for 4 units set up in a square, figure 4.7,

plot entitled "Ideal") helps to obtain a rough idea about the spatial distribution of intensities but will most likely differ from reality. Not completely horizontally installed sprinkling units, pressure differences within the hose system, temporal instabilities in the water supply or even imperfections of individual nozzles impact the real distribution of sprinkled water. We therefore measured cumulative water amounts at representative points of the sprinkling area. Secondly, compared to this theoretical water distribution, the actual distribution during the experiment results in higher amounts of water masses at some locations of the sprinkled area (figure 4.7, plot entitled "Measured"). We find that this locally increased input of water leads to the process-representative shape of the response curve being reached after less time and/or lower water input amount. Consequentially, a heterogeneous water input distribution does not hinder but even favors the identification of dominant infiltration dynamics (if high input locations are closer to the gravimeter and assuming that the infiltration process observed at the high input locations would also be dominant over the entire sprinkling area).

With respect to the limits of the sprinkling area, we set the size of the model domain depending on the maximum sprinkling distances around the sprinkler units. However, in other locations, when high (subsurface) lateral flow is expected, the extent of the sprinkling area could and should be enlarged accordingly, in order to include all resulting mass transfers. Concerning evapotranspiration it should be noted that it is difficult to determine its amount throughout the course of the experiment with a precision needed in order to correct model water input. We therefore recommend to conduct experiments after sunset or even during night times when evapotranspiration will be as low as 1.7 to 14% of daytime rates [Iritz and Lindroth, 1994; Malek, 1992]. Generally speaking, all these factors and error sources have a more direct impact if we would want to obtain an absolute quantification of water masses in the subsurface but are not influencing or hindering the qualitative identification of dominant subsurface flow processes.
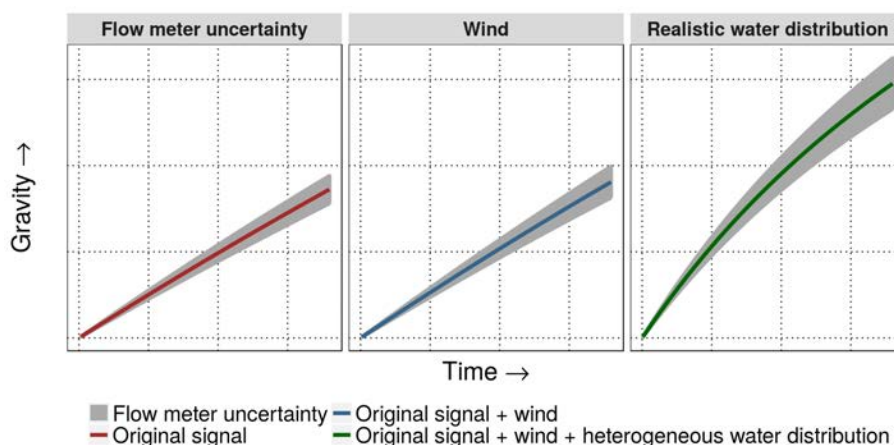
Figure 4.6: Impact of uncertainties due to the experimental setup on the gravity response. Uncertainty due to inaccurate estimation of total sprinkling water (10% scenario shown, left), wind (middle) and a heterogeneous water distribution due to sprinkling unit setup (right). This is exemplarily shown for the process wetting front advancement but affects equally all isolated process dynamics. For all three plots, the x-axis shows equal time periods.
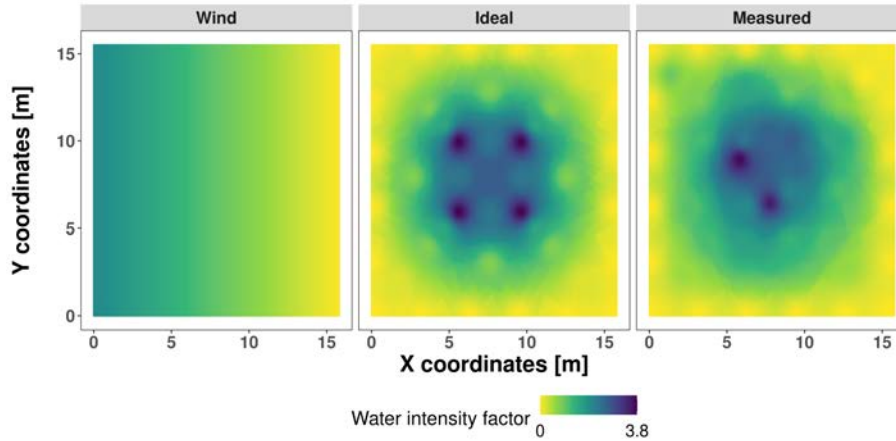
Figure 4.7: Sprinkling water intensity distribution grids for wind (left), a theoretical distribution based on the sprinkler unit setup (middle) and the real distribution based on interpolated intensity measurements (right). Provided values represent the water intensity factor which describes the relation of measured or interpolated water input of each grid cell to average input based on the total sprinkling volume. The gravimeter is always located in the center of the grid and absolute water volumes are identical for all three plots.

4.3.3   Real-World Experiment

Model Evaluation by Response Curve Shapes and Performance Measures

Constrained by the observed gravity signal, the hydro-gravimetric model (see section 4.2.1) was run in inversion mode with a one minute resolution for the 6 hours duration of the sprinkling experiment. The modeled scenarios covered the 5 subsurface flow dynamics (see section 4.2.2) and were categorized in 3 process groups: (i) single processes excluding lateral flow, (ii) single processes including lateral flow and (iii) a combination of 2 infiltration processes (3 with macropore flow in the upper soil layers and wetting front advancement (WFA), bypass flow and perched water table in the lower ones and 1 with preferential flow in the upper and wetting front advancement in the lower soil layers, all including lateral flow). For simplicity we decided to combine not more than 2 processes per model run. For all scenarios, table 4.2 summarizes the fitted values of the parameters maximal saturation deficit, process depth, clustering factor and vertical to lateral water redistribution factor (where applicable). For the parameter ranges see table 4.1.

Within all three scenario groups, the process bypass flow results in best performance measures (see table 4.2). The optimized parameter values *maximal saturation deficit* (max $\Delta\Theta$) and *process depth* are similar for all process groups: 0.06, 0.06, 0.09% VWC and and -1.17, -1.21, -1.82 m, respectively. The simulated gravity time series confirm that every subsurface flow process results in a characteristic shape of the gravity response with time: convex for wetting front advancement and bypass flow, linear for macropores and quasi-linear for preferential flow and concave for perched water table (see figure 4.8). The superposition of the processes macropores and perched water table results in a more complex shape of the gravity response curve, so does preferential flow in combination with bypass or WFA. This latter is the only scenario that results in saturation deficit parameter ranges similar to in-situ

observations (max $\Delta\Theta > 0.12\%$VWC). These in-situ observations were evaluated by comparing their Θ_{max} and Θ_{min} values during the sprinkling experiment. Comparing process combinations which include bypass flow or preferential flow, we find similarly high performance measures but differences in the parameter values. This points to the equifinality of optimizing a model [Beven, 2006] with gravity data: a similar gravity signal may result from water mass changes of different magnitude that occur at different distance to the gravimeter. This fact complicates identifying "the best" result, i.e., the dominant flow process. Therefore, additional data from independent sources need to be considered, as discussed below, or the quantity of optimized parameters reduced [Schumacher et al., 2018].

Table 4.2: Optimized parameter values for all modeled scenarios. "cf" is the clustering factor and "lrf" is the lateral redistribution factor. For a description of the parameters see text in section 4.2.1 and 4.2.2. All parameters are unitless, except "depth of process" which is provided in meter below the surface.

scenario	subsurface flow process	mNSE	max $\Delta\Theta$	depth of process	cf	lrf
Single process	wetting front advancement	0.53	0.05	-	-	-
	bypass	0.98	0.06	1.17	-	-
	perched water table	0.81	0.25	2.45	-	-
	preferential flow	0.84	0.06	-	1	-
	macropores	0.85	0.05	4.42	-	-
Single process, including lateral flow	wetting front advancement	0.53	0.05	-	-	0.81
	bypass	0.98	0.06	1.21	-	0.61
	perched water table	0.81	0.25	2.36	-	0.79
	preferential flow	0.84	0.06	-	2	0.98
	macropores	0.85	0.05	4.27	-	0.12
Two processes combined, including lateral flow	macropores	0.47	0.05	1.50	-	0.26
	wetting front advancement		0.05	-	-	
	macropores	0.97	0.05	0.10	-	0.40
	bypass		0.09	1.82	-	
	macropores	0.77	0.05	1.50	-	1.0
	perched water table		0.05	1.73	-	
	preferential flow	0.91	0.12	0.83	7	0.54
	wetting front advancement		0.17	-	-	

Model Evaluation by other Monitoring Data

Depth-specific maximal saturation deficits (table 4.2), estimated by the optimization routine of the hydro-gravimetric model, vary in the range of 5 to 25% VWC. Except for the combined scenario with preferential flow and for the perched water table scenario, the optimized values are at the lower boundary of this range. For the uppermost 10 cm of profiles 1 and 2 (first and second closest to gravimeter, see figure 4.4) observed soil moisture increases by 14% VWC and 15% VWC (figure 4.9), respectively. After this increase, the soil moisture content stayed constant for the remaining duration of the experiment, indicating a saturation of those layers. At a depth of 20 cm, at profile 2, observed soil moisture increased by 12% VWC. At the lowest sensor locations (30 and 40 cm depth), the increase was 8% to 9% VWC at both profiles before it leveled out during the experiment. With 7.8 m distance to the gravimeter, profile 3 is located outside of the sprinkling area. A soil moisture increase at profile 3 was observed at a depth of 10 cm only. The soil moisture response started about 70 min later than in the uppermost layers of profiles 1 and
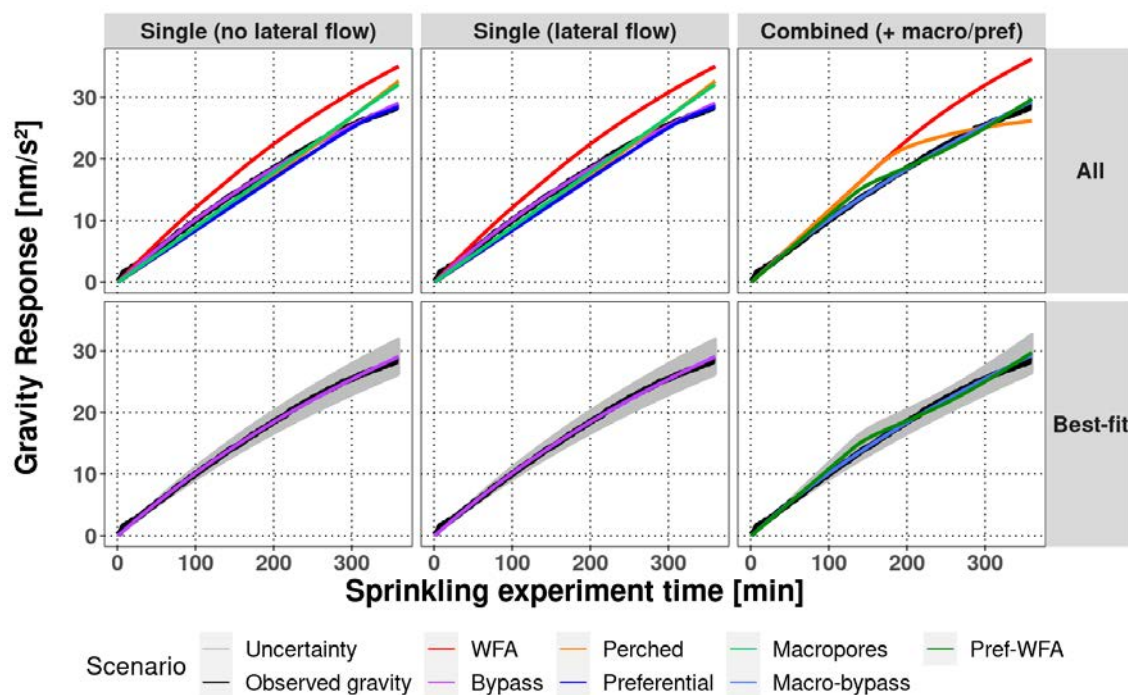
Figure 4.8: Observed (black) and modeled gravity response for the duration of the sprinkling experiment and for all scenarios. The optimized gravity responses for all flow processes of the different process groups (upper row) and the best-fit scenarios of each group (bottom row). Note that for the combined processes there are two best-fit scenarios. For the best-fit results, the uncertainty interval (grey area) represents the gravity response based on the uncertainty of sprinkling water volumes of $\pm\,10\%$.

2 and gradually increased by 11% VWC until the end of the sprinkling period. It did not reach a plateau value, indicating that a saturation/steady state was not reached. As there is no direct input of sprinkling water at the location of profile 3, these observations indicate water input by near-surface lateral flow after surface saturation is reached within the sprinkling area. Comparing the observed values of soil moisture increase with the modeled soil moisture increase which is constrained by the optimized maximal allowed saturation deficits of the different model scenarios, it can be stated that: (a) max $\Delta\Theta$ of nearly all modeled scenarios strongly underestimate the observed soil moisture changes and (b) only the combined scenario of preferential flow and wetting front advancement results in values which are comparable to the measured soil moisture changes. It has to be noted that the model is incapable of estimating spatially varying saturation deficits within the layer of one subsurface flow process. The observed soil moisture, in contrast, indicates saturated conditions within the sprinkling area and no saturation in the area receiving lateral flow. These point observations however might not be representative for the entire area monitored by the gravimeter, especially with the high soil heterogeneity and the resulting complexity for small-scale model representation [Vereecken et al., 2016] and no observations below depths of 40 cm.

For an additional evaluation of the model results, we compare 2D segments of the modeled 3D changes in soil moisture distribution with segments of ERT profile B and with an interpolation of measured soil moisture to a depth of 30 cm based on all soil moisture sensors at the experimental site, i.e., the near-surface sensors,
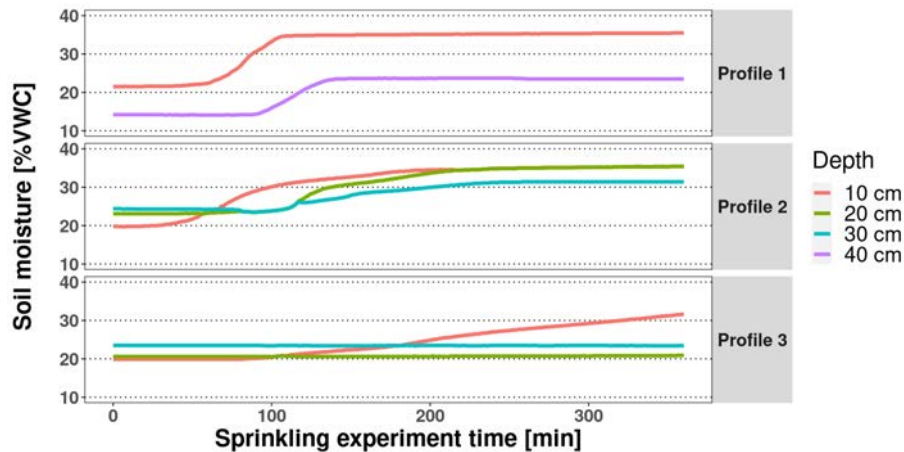
Figure 4.9: Measured soil moisture dynamics at three vertical profiles for the duration of the sprinkling experiment (with sprinkling starting at t=0 minutes and ending at t=360 minutes). The distance of the profiles from the center of the sprinkling grid increases from Profile 1 to 3. For the exact location of the profiles see figure 4.4.

the vertical sensor profiles and the diagonal transect of near-surface sensors (figure 4.4). For ERT, we inverted the electrical resistivity data up to a depth of 5 m in a time-lapse mode and derived the differences between individual time steps. The differences between individual time steps and pre-experiment conditions indicate a decrease in resistivity and thus a likely increase in soil moisture. The more negative the difference, the higher the increase in soil moisture at a particular location. Inverted resistivity values (see figure 4.10) show a general increase in soil moisture throughout the 2D profile but predominantly close to the surface. There are two particular distinct areas of much higher increase. One marked patch with an estimated vertical extent of 2.5 m is situated along the profile between meter 10 and 15. A less pronounced and spatially much smaller patch can be identified at the profile position of 5 m. ERT Profile A shows similar results (data not shown, see figure 4.4 for its location). Figure 4.10 furthermore includes modeled soil moisture differences of two scenarios with high performance measures: combined bypass flow and combined preferential flow. The model-based patterns do not match well with the ERT data. The combined macropore-bypass flow scenario (left part of figure 4.10) shows almost no soil moisture increase in the upper layer of the macropore process and a small increase at lower depths (bypass flow process) which is laterally more spread out than the structure from ERT. For the combination of preferential flow and wetting front advancement (right part of figure 4.10) the model results in a soil moisture increase at shallower depths (preferential flow process) which is limited to the preferential vertical flow paths, and an underlying increase with more horizontal spread but limited vertical extent caused by the wetting front advancement process. For this scenario max $\Delta\Theta$ values match the maximum soil moisture changes from sensor observations but do not match the spatial structure inferred from ERT measurements. It should be noted, however, that the assumed simplified chess-board like distribution of preferential flow paths in the model is artificial and random, and thus cannot represent the real-world ERT patterns at the right locations. Nevertheless, the larger size of the patterns in the ERT data compared to the model patterns may

indicate that even a larger size of the preferential flow pathways might be adequate for the model. On the other hand, depending on soil properties and the used array type, ERT data might also be limited in their capability to adequately represent small-scale vertical heterogeneities [Samouëlian et al., 2005].

For interpolated soil moisture measurements and model results from the same two scenarios (shown in figure 4.11), we can clearly identify that the increase of soil moisture in the combined model scenario macropore-bypass flow is too small in both amount and depth in the top 30cm of the model domain. For scenario preferential flow and WFA, a reasonable but spatially patchy increase in soil moisture within the first 30 cm of the model domain can be observed. This patchiness is caused by the model setup of preferential flow columns. It comes closer to the ERT patterns than the other scenarios, but the comparatively small size of the modelled patches indicates that an even higher value of the clustering factor could result in even higher correspondence of model results and soil moisture and ERT observations. On the other hand, one has to note that the ERT setup itself may not allow for a more detailed differentiation of individual preferential flow paths due to its horizontal resolution of 0.25 m. As Tso et al. [2019] stated, there is also a considerably high uncertainty when using ERT data to interpret soil moisture changes.
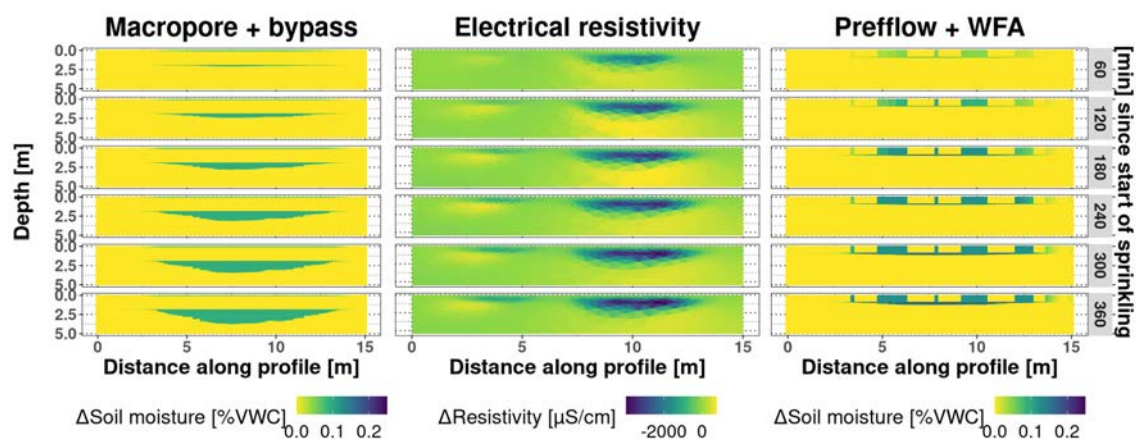
Figure 4.10: Hourly sequence of inverted electrical resistivity out of ERT measurements (middle) and modeled soil moisture increase for the best-fit scenarios "macropores + bypass flow" (left) and "preferential flow + wetting front advancement" (right) during the sprinkling experiment for the transect of ERT profile B (see figure 4). The values indicate the change compared to the state prior to the start of sprinkling. Negative changes in ERT data indicate an increase in soil moisture. The gravimeter is located on the x-axis at position 7.5 m.

For an extended evaluation of the above results, a synthetic spatial observation-based soil moisture distribution was set up, called virtual reality (VR) in the following. The spatial dimensions of VR are identical to the modeled scenarios. The magnitude of soil moisture changes in VR is inferred from the soil moisture observations, whereas the spatial structure of the water distribution is based on the ERT observations. For both, the observed states at the end of the sprinkling experiment were used. Figure 4.12 illustrates VR in a view from above and by an exemplary 2D-profile. The overall soil moisture changes considered for VR correspond to the total amount of sprinkled water used in the real-world experiment. The combination of
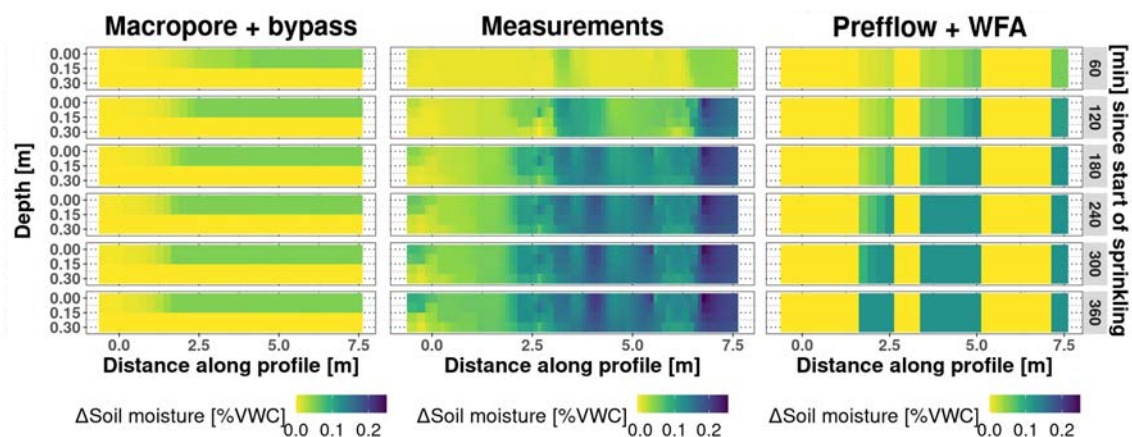
Figure 4.11: 2D profiles showing the changes of soil moisture measurements (middle) and modeled soil moisture for the best-fit scenarios macropores + bypass flow (left) and preferential flow + wetting front advancement (right). Note the different y-axis scale compared to Figure 10. The values indicate the change with respect to the state prior to the start of sprinkling. The gravimeter is located on the x-axis at position 7.5 m.

the real water mass increase in VR together with the observed spatial distribution in terms of distribution with depth and distance to the gravimeter result in a forward calculation of its gravity effect to match the observed gravity change at the end of sprinkling experiment with an error of less than 1%. While VR does not allow for reconstructing the temporal evolution of the gravity increase during the sprinkling experiment, it supports the assumption that the additional observation data provide a good representation of the actual changes in soil moisture distribution resulting from the artificial rainfall. Furthermore, in view of the various scenarios that have been tested for infiltration and subsurface flow processes leading to the observed soil moisture distribution and gravity response, the VR results give additional support for the validity of the dominant processes selected with by the hydro-gravimetric approach. At the same time it illustrates the limitations and ambiguity of this approach when no additional observations are available.

Combined Evaluation of Dominant Subsurface Flow Pattern

The evaluation of the model results and the comparison with the measured datasets of gravity, soil moisture and electrical resistivity indicate that only a combined evaluation of a) model performance measures, b) the shapes of the gravity response curves, c) the amount and d) spatial distribution of water mass increase leads to a comprehensive assessment of the hydro-gravimetric process inference. For our modelling results, all four criteria are best met by the scenario preferential flow + wetting front advancement: The shape of the gravity response of this scenario matches the observed gravity data during the sprinkling experiment (see figure 4.8), the scenario results in one of the highest overall mNSE (see table 4.2) and also shows the by far highest, although not perfect, accordance with measured soil moisture changes over time in different layers (compare figures 4.9, 4.11 and table 4.2). Concerning the ERT data and the ability of the model to reproduce these spatio-temporal variations of soil moisture, the results are less satisfying (see figure 4.10). The reasons for
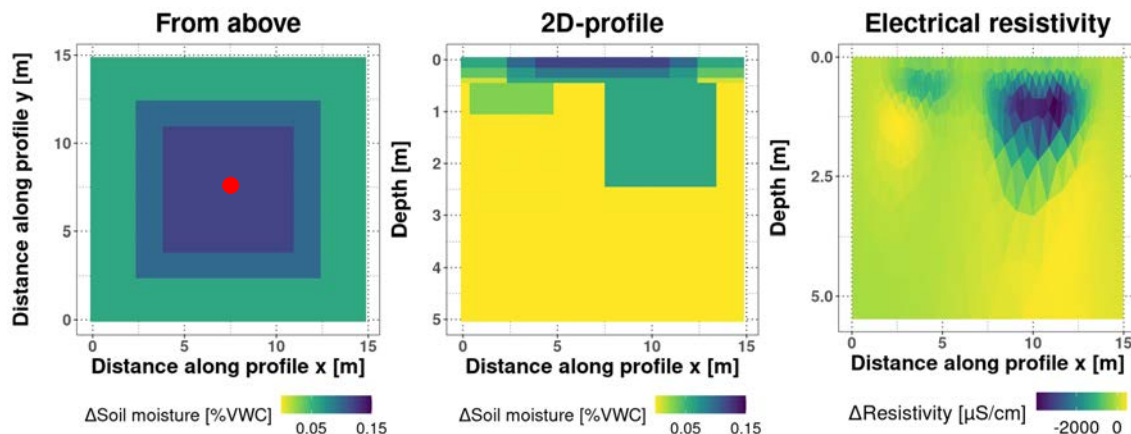
Figure 4.12: Virtual reality (VR) changes in water distribution based on ERT measurements and observed soil moisture changes. As stated in the text, the gravity response of the setup shown in the 3 plots matches the observed gravity signal by the gravimeter. The left plot shows a view from above onto the sprinkling area whereas the middle one shows a vertical 2D profile through the gravity sensor location (red dot). For a comparison of spatial patterns, the right plot shows the measured changes in electrical resistivity between the end and the start of the sprinkling.

this are manifold: ERT data are only available at three 2D profiles but soil heterogeneity at the site is quite high. Also, ERT itself is limited in adequately resolving small-scale moisture patterns. Furthermore, the model is also limited in resolving 3-dimensional complex soil structures and the corresponding water storage dynamics. Nevertheless, in line with the observations, the optimum gravity-based scenario shows near-surface wetting in addition to a deeper soil moisture increase which is not the case for any other scenario. The result is also in accordance with previous studies at the same site where Müller [2010] described highly heterogeneous soils throughout the area with active lateral flow and the presence of macropores and preferential flow structures with spatially variable densities and sizes. The spatial heterogeneity of the soil and a complex system of interaction of lateral flow processes were also described by Heim [2010].

4.3.4   Advantages and Disadvantages of Gravimetry in Combination with Sprinkling Experiments

We summarize the pros and cons of combining sprinkling experiments with gravimetric measurements as a tool for assessing dominant hydrological processes in table 4.3. As illustrated in this study, the integrative nature of gravimetry provides well-suited and unique complementary information to traditional hydrological field methods (e.g. soil moisture sensors, lysimeters) as well as other non-invasive geophysical techniques (e.g. ERT). On the other hand, these additional observations are necessary to reduce the equivocality when interpreting the integrative gravity signal. While some prerequisites for the high-quality application of gravimetry in field experimental setups such as a stable based plate or pillar for the instrument, are inherent to this technique and will remain, others such as high instrument precision in combination with minimized instrument size for ease of field application

may be alleviated in future by ongoing technological advances [Ménoret et al., 2018; Middlemiss et al., 2022].

Table 4.3: Advantages and disadvantages of combining gravimetry with sprinkling experiments.

Advantages
non-invasive method
field-scale monitoring footprint (200-300 m radius)
integrative information on water storage changes
separation of storage compartments when combined with other integrative methods (e.g. CRNS)
multi-gravimeter approach could provide spatially distributed storage changes
lower precision of certain gravimeter types can be compensated with larger sprinkling water amount
water storage (mass) budgeting at larger scales than with traditional field methods
Disadvantages
additional measurements needed for signal separation
water storage changes in gravimeter footprint outside sprinkling area need to be minimized
large sprinkling volumes needed for good signal-to-noise ratio
base plate / pillar construction required within sprinkling area
high investment costs
expertise in gravity data processing needed

4.4   Summary and Conclusions

A sprinkling experiment was conducted around a superconducting gravimeter (iGrav#006). In addition to the continuous monitoring of gravity, soil moisture sensors were installed in three vertical profiles and one near surface horizontal transect across the sprinkling area. Furthermore soil resistivity was monitored hourly at 3 ERT profiles. The goal of the experiment was to investigate the usefulness and feasibility of terrestrial gravimeters as a non-invasive, integrative field method for subsurface process identification and quantification of spatio-temporal water mass re-distributions. Five hydrological infiltration and distribution process scenarios were defined and included in the analysis: macropores, preferential flow, wetting front advancement, bypass flow and a perched water table. We found that a preliminary identification of dominant water distribution processes in the subsurface can be directly inferred from the gravity observations. The shape of the gravity response curve, serves as a proxy: i) a linear shape indicates a water content increase at rather constant distances to the gravimeter as is the case for e.g. near-surface macropore filling or preferential flow paths, ii) a convex shape indicates mass increases at distances increasing over time, e.g. wetting front advancement or bypass flow and iii) a concave shape indicates mass increases at distances decreasing with time, e.g. perched water table scenarios where water ponds on top of a non-conductive soil layer. To better resolve spatio-temporal water distributions, we developed a simplified hydro-gravimetric model which converts grid-based soil moisture changes to its corresponding gravity value. This model was run inversely within an optimization algorithm, using the modified Nash-Sutcliffe-Efficiency as its objective function to compare modeled gravity responses to observed data from the sprinkling experiment. It was applied for different setups and combinations of all 5 process scenarios with varying constrains of parameter sets. It was then evaluated using model performance measures, gravity data, soil moisture and ERT time series. We found that a combination of scenarios preferential flow and wetting front advancement led to the overall

best results within this multi-criteria assessment. It is furthermore in accordance with subsurface properties and dynamics found in previous studies at the same site. Amplitudes and values of in-situ measured soil moisture were reasonably well reproduced by the model. However, the spatial structure of water redistribution did not match the observations from ERT measurements. Nevertheless, the corresponding modeled gravity response curve matched gravity observations whereas others could clearly be discarded.

Although the integrative nature of the gravimetric method and the resulting equifinality still causes some complications in the interpretation, this can be compensated by additional measurements. The developed pseudo-physical model was intentionally simple. This was helpful for rapid process understanding based on gravity data but also showed some limitations and disadvantages. Its incapability to resolve complex subsurface soil structures and storage dynamics was reflected in the spatial water redistribution patterns of the inverse model results. The model setup of a virtual reality, based on observational data from soil moisture and spatial structures from ERT, confirmed these limitations but at the same time underlined the feasibility of our general approach. In order to further develop this gravimetric approach as a stand-alone solution for sprinkling experiments or similar setups with the intention to describe the subsurface flow processes non-invasively and without the need of additional measurements, it is necessary to conduct similar studies but with a more complex numerical, physically-based hydrological model (e.g. Hydrus). This is planed for follow-up studies. It would help to overcome the discussed spatial limitations of the simple hydro-gravimetric model, allow for a standardized process definition and thus comparison with other studies and enable a more contained water budgeting (incl. ET and root water uptake) while keeping the coupling to gravity observation data as the basis for model optimization.

We also discussed uncertainties and possible errors for setting up real-world sprinkling experiments in combination with gravimetry. We found that an inaccurate estimation of sprinkled water volume does not hinder the identification of the dominant process because the shape of gravity response curve is not directly affected. The curve also remains unaffected by the displacement of irrigation water by steady wind from one direction. Furthermore we showed the necessity to observe and quantify spatial sprinkling water distributions during experiments, as highly heterogeneous distributions of applied water have a big influence on the shape of the gravity response curve. It is also recommended to conduct experiments on cloudy days and/or during nighttime to minimize the effect of evapotranpiration, unless evapotranspiration can be quantified.

In general, our findings showed that using a gravimeter in combination with a sprinkling experiment is highly beneficial to assess processes in the subsurface. Its non-invasive character makes it perfectly suitable for accompanying these types of field experiments. The discussed setup and analysis can be transferred to sites with similar conditions. At sites with very different conditions, adaptions of the setup might be necessary. Necessary infrastructure for any site is a stable pillar for gravimeter installation, electricity and water supply. Depending on available water volume, on-site electricity and required accuracy, it is possible to also use other types of gravimeters (e.g. gPhone, CG-6) for the same type of experiment. With promising results, discussed uncertainties and limitations of the presented approach and a clear perspective for further development of similar experiments in the future,

the study contributes towards paving the road for gravimeters to become a tool for field hydrologists.

Chapter 5

Correction-induced uncertainties when processing terrestrial gravimetry data

Michal Mikolaj, Marvin Reich, Andreas Güntner

Abstract

Terrestrial gravimetry is increasingly used to monitor mass transport processes in geophysics boosted by the ongoing technological development of instruments. Resolving a particular phenomenon of interest, however, requires a set of gravity corrections of which the uncertainties have not been addressed up to now. In this study, we quantify the time domain uncertainty of tide, global atmospheric, large-scale hydrological, and non-tidal ocean loading corrections. The uncertainty is assessed by comparing the majority of available global models for a suite of sites worldwide. The average uncertainty expressed as root-mean-square error equals $5.1\,\mathrm{nm/s^2}$, discounting local hydrology or air pressure. The correction-induced uncertainty of gravity changes over various time periods of interest ranges from $0.6\,\mathrm{nm/s^2}$ for hours up to a maximum of $6.7\,\mathrm{nm/s^2}$ for six months. The corrections are shown to be significant and should be applied for most geophysical applications of terrestrial gravimetry. From a statistical point of view however, resolving subtle gravity effects in the order of few $\mathrm{nm/s^2}$ is challenged by the uncertainty of the corrections.

5.1   Introduction

The high precision of state-of-the-art gravimeters has contributed to their widespread use in geophysics for measuring spatial and temporal variations of gravity of the Earth. The common denominator of studies exploiting terrestrial gravity measurements is the relation of gravity variations to mass transport. Besides the more traditional use of gravimeters in geodesy and geodynamics, e.g., to map changes of the Earth's shape [Plag and Pearlman, 2009], in tectonics [Van Camp et al., 2011], or to assess land uplift due to glacial isostatic adjustment [Ophaug et al., 2016; van Dam et al., 2017], terrestrial gravimetry is an emerging monitoring technique in other fields of research. These include monitoring of geothermal fields [Kao et al., 2014], CO_2 storage reservoirs [Sugihara et al., 2017], volcanoes [Carbone et al., 2017; Poland and Carbone, 2016], evapotranspiration rates [Güntner et al., 2017; Van Camp et al., 2016b], pre- and co-seismic changes [Imanishi et al., 2004; Vallée et al., 2017], or the definition of the kilogram [Stock, 2013]. Many more applications of absolute and relative gravity measurements have been reviewed by Crossley et al. [2013]; Van Camp et al. [2017]. Given the ongoing technological advancements in terrestrial gravimetry towards smaller, more precise, and more affordable instruments, including the development of new technologies such as atom quantum instruments [Gillot et al., 2014; Ménoret et al., 2018] and micro-electromechanical devices [Middlemiss et al., 2016], fields of research and application for terrestrial gravimetry can be expected to grow further in the future.

All applications of terrestrial gravimetry share the common challenge of resolving the signal of interest by applying a set of time-variable gravity corrections. Taken into account should be all other effects that are part of the integrative gravity measurements due to both mass attraction and Earth's surface deformation. The standard procedure for reducing the observed signal to gravity residuals of interest comprises Earth and ocean tides, variations of the Earth orientation parameters and atmospheric loading [Hinderer et al., 2015b]. Especially if small-scale phenomena are of interest, further corrections for gravity effects of continental (large-scale) hydrology and non-tidal ocean loading should be considered. While there are many models available for these processes, there is no recommended procedure which model should be applied to reduce the gravity observations. This has led to inconsistent approaches among studies so far. For reducing tides, Tanaka et al. [2013], for instance, used the Baytap program [Tamura et al., 1991], other studies preferred ETERNA [Wenzel, 1996], e.g., in Meurers et al. [2016], or VAV program [Venedikov et al., 2005], e.g., in Arnoso et al. [2011]. Much of the same variety can be found even in studies focusing on identical or similar phenomena, like local water storage changes, where for the correction of global atmospheric gravity effects, some studies [Hector et al., 2015] chose the EOST atmospheric loading [Boy and Hinderer, 2006], and others [Creutzfeldt et al., 2013] preferred Atmacs [Klügel and Wziontek, 2009]. Similar examples can by found for continental hydrology and non-tidal ocean loading [Hector et al., 2014; Mikolaj et al., 2015, e.g]. In view of differences between the various models, it often remains unclear to what extent the result depends on the chosen correction model and whether the signal of interest can be reliably resolved.

The signal of interest can be confidently resolved only if the instrumental precision (or reproducibility), uncorrected ambient and environmental noise, and the uncertainty of applied corrections allow such conclusion from a statistical point of
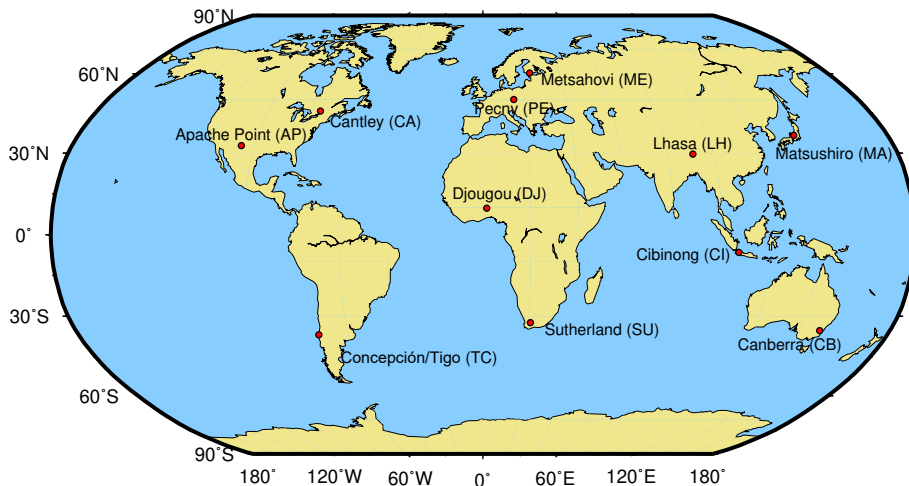
Figure 5.1: Location of study sites (map created using P. et al. [2013]).

view. The estimation of the instrumental precision requires separation of the noise sources. This can be done by comparing the gravity records of multiple gravimeters operated at the same location. Recently presented results in Rosat and Hinderer [2018] obtained after comparing six SGs suggested that the instrumental self-noise can reach $0.003\,\mathrm{nm/s^2}$ in amplitude (or $1\,(\mathrm{nm/s^2})^2/\mathrm{Hz}$ power spectral density) at periods between approximately 20 second up to 1 hour. The noise amplitude increases for daily periods approximately by a factor 10 (100 for power spectral density). In the time domain, the SG precision was in Hinderer et al. [2015a] estimated to be 1 $\mathrm{nm/s^2}$ based on a review of several studies comparing co-located and dual-shere SG measurements. The long-term precision of SG records is directly dependent on the reproducibility of FG5 absolute gravimeters used to correct the instrumental drift of SGs. The long-term reproducibility of FG5 absolute gravimeters was estimated in Van Camp et al. [2005] to be at the level of $16\,\mathrm{nm/s^2}$. The aforementioned values do not take systematic effects or the uncertainty of gravity corrections into account.

The aim of this study is to quantify the uncertainty of selected gravity corrections. The uncertainty is assessed comparing commonly applied correction models. The acquired differences are in accordance with JCGM [2008] used to compute the root-mean-square error characterizing the dispersion of correction models. The results reveal to which extent the correction procedure contributes to the overall uncertainty of gravity residuals, and which are the main components affecting the resulting uncertainty. For the correction procedure, we follow an average user's perspective who selects one or another correction model without prior knowledge of the true and actually unknown value. We consider only those corrections for which a variety of basically indistinguishable models is available, i.e., where a regular user cannot infer a preference for a distinct model from literature. This limits the range of examined corrections to global atmospheric, tidal, and non-tidal ocean mass, and large-scale hydrological effects. Local hydrological effects contributing significantly to the variation of gravity are beyond the scope of this study. Contributions such as polar motion and length of day are not considered because of the agreement on correction procedures [Hinderer et al., 2015a; Van Camp et al., 2017]. Details on the uncertainty of some gravity processing procedures such as calibration of relative gravimeters can be found in Meurers [2012]; Van Camp et al. [2016c], while

the uncertainty of gap fillings and discontinuity corrections are to a certain extend discussed in Hinderer et al. [2002].

In order to cover a broad range of conditions, the correction uncertainty is determined for 11 sites worldwide, located in a wide range of latitudes, altitude and in all 5 major climate zones except of polar according to Kottek et al. [2006] (Figure 5.1). The sites were selected among the stations of the International Geodynamics and Earth Tide Service IGETS [Voigt et al., 2016]. Several studies provided information on the combined instrumental and environmental noise level at these sites [Rosat and Hinderer, 2011; Van Camp et al., 2010; Zürn and Widmer, 1995, e.g.,]. These studies, however, did not quantify the uncertainty of the corrections that were applied to reduce the environmental noise. A different approach to assess the uncertainty of long-term trends in terrestrial gravity measurements was applied in Van Camp et al. [2016a] by considering the impact of water storage estimates of one model or observations (i.e., GRACE) on the trend estimate. Here, we present the statistical uncertainty of a set of gravity corrections commonly applied to terrestrial gravity observations by comparing correction models that meet criteria as discussed in the following section. Furthermore, the uncertainty is quantified in the time domain and for temporal changes in the range between one hour and one year, and for the long-term trend.

Table 5.1: Study sites

Site						
Name	ID	λ (^{o})	ϕ (^{o})	H(m)	Climate[a]	Coast[b]
Apache Point	AP	-105.82	32.78	2788	Cfb	706
Cantley	CA	-75.8	45.59	269	Dfb	108
Canberra	CB	149.0	-35.32	763	Cfb	116
Cibinong	CI	106.85	-6.49	138	Af	43
Djougou	DJ	1.61	9.74	483	Aw	385
Lhasa	LH	91.04	29.65	3600	BSk	670
Matsushiro	MA	138.20	36.54	451	Cfa	63
Metsahovi	ME	24.10	60.22	56	Dfb	11
Pecny	PE	14.79	49.91	535	Cfb	410
Sutherland	SU	20.81	-32.38	1791	BSk	219
Concepción	TC	-73.03	-36.84	156	Csb	11

[a] derived after Rubel et al. [2017]
[b] distance (km) to coastline computed using `naturalearthdata.com`

5.2   Data Sets

To simulate realistic scenarios of the gravity correction procedure that apply to any user for any location worldwide, the input models of the geophysical processes to be reduced were selected according to the following criteria: (1) cited and widely recognized model, (2) global spatial coverage, (3) sufficient temporal coverage and resolution, (4) public availability. The latter criterion includes data sets and software that is available upon request. Criterion (3) limits the choice of models to those starting at least in 2008 and providing data continuously until the present. The

starting date was constrained by the availability of actual gravity measurements at the selected sites in the IGETS data base. It should be noted, that the actual gravity time series were needed only as input for tidal analysis programs. The data set and the source code used in this study is provided in Mikolaj et al. [2019a].

5.2.1 Tides

Three different software packages were used to estimate the tidal parameters, namely Baytap(08), ETERNA3.40 and VAV06. All program parameters were set following the recommendation in the respective user manual. This included the tidal wave grouping with the exception of long-periodic waves (up to frequency of 0.7215 cycles per day). The latter were not estimated because of the unknown parameters such as instrumental drift and steps that affect the actual low-frequency variations. The hourly gravity time series of superconducting gravimeters were obtained after downloading from the IGETS database (`isdc.gfz-potsdam.de/igets-data-base/`, access date 5 March 2018, [Boy et al., 2017; Förste et al., 2016; Wziontek et al., 2017a]). Hourly time series were used for all sites except of Djougou and Cibinong, where these were not available. At these sites, one minute data was downloaded, calibrated using parameters in the input file header, filtered, and re-sampled to hourly values. Table S1 in the Supporting information shows the time interval entering the tidal analysis at each site. In situ pressure recordings were also used in the tidal analysis assuming linear relation between gravity and pressure. To remove the unknown low-frequency effects, e.g., instrumental drift, the input time series were high-pass filtered prior to the tidal analysis using a finite impulse response filter proposed in the ETERNA3.40 user manual with a cutoff frequency of 0.583 cycles per day. Time series prepared in such manner were then used as input in all three programs using identical tidal potential development and other settings for each site. The three sets of estimated tidal parameter were then used to predict the tidal signal at each site for the common time period starting 1 January 2008 to 31 December 2015 using TSoft [Van Camp and Vauterin, 2005]. Thus, the programs were used only to derive the tidal correction models. The uncertainty was assessed by comparing all available correction models without considering the actual gravity measurements.

5.2.2 Large-scale hydrology

The large-scale hydrological effects from continental water storage variation were computed with seven global hydrological models, namely four land surface models (CLM, MOS, NOAHv21, VIC) of the GLDAS model [Rodell et al., 2004], ERA Interim [Dee et al., 2011], MERRA-2 [Gelaro et al., 2017], and NCEP-DOE Reanalysis 2 model [Kanamitsu et al., 2002]. The simulated storage variations were converted to gravity effects using the mGlobe toolbox [Mikolaj et al., 2015] (download at `github.com/emenems/mGlobe`, access date 28 February 2018). The large-scale hydrological effects were computed for points with spherical distance greater than 0.1 degree from the study site. Mass conservation was enforced by applying a uniform layer over the ocean determined by total model mass deficit or surplus as compared to the long-term average. Spurious steps in MERRA-2 and a peak in NCEP-DOE Reanalysis 2 that occurred when enforcing mass conservation were corrected. To suppress the impact of unreliable data of global hydrological models in Antarctica

and Greenland, all mass variations in these regions were set to zero. Due to the minimal effect of the temporal resolution on the result (standard deviation below $0.1\,\mathrm{nm\,s^{-2}}$), the large-scale hydrological effects were computed for all models with 24 hour time resolution and resampled to hourly resolution afterwards.

5.2.3 Non-tidal ocean loading

The non-tidal ocean loading effects were acquired from EOST Loading Service [Boy et al., 2009] utilizing ECCO-JPL [Fukumori, 2002], ECCO2 [Menemenlis et al., 2008] and TUGOm [Loren and Florent, 2003] models (download at `loading.u-strasbg.fr`, access date 28 February 2018). In addition, non-tidal ocean mass variations of the OMCTv06 model [Dobslaw et al., 2017] were converted to their gravity effect using mGlobe. Identical minimal integration radius of 0.1 degree from the study site and maximal available model temporal resolution was used in all cases. Linear interpolation was used to resample all time series to hourly resolution. Further details on the effect of the interpolation method on the uncertainty can be found in Mikolaj et al. [2015].

5.2.4 Global atmosphere

Empirical or physical approaches can be followed to correct the atmospheric effect on gravity. The empirical approaches called also single admittance utilize in situ air pressure observations converted to a gravity effect by assuming linear relation between these variables. This approach can be further refined by introducing a dependence of the admittance factor on the frequency of the pressure variations [Crossley et al., 1995]. As shown in Hinderer et al. [2014], the estimated frequency dependence can be modelled using mass variation output of a physical atmospheric model. The advantage of global atmospheric models is that the computed gravity corrections account also for atmospheric processes occurring far beyond the regional zone. Merriam [1992] showed that single admittance can account for only around 90% of the total atmospheric effect. Hence, the second approach based on modelling of global atmospheric mass variation is analyzed in this study. The disadvantage of global models is the lower temporal resolution compared to in situ pressure, potentially leading to insufficient modelling of processes at shorter periods. This issue is addressed in a separate analysis combining in situ pressure with global models as recommended in the Atmacs model [Klügel and Wziontek, 2009, `atmacs.bkg.bund.de/data/`, last access date 2 July 2018] and Mikolaj et al. [2015].

The main analysis outlined in the Methods section uses following corrections: The EOST service provides global atmospheric effects utilizing ECMWF Operational and ERA Interim models. The latter model was also utilized for an independent computation using mGlobe as, unlike in case of hydrological and oceanic effects, the computation approaches of EOST (2D) and mGlobe (3D) differ. Maximal model temporal resolution and the inverted barometer assumption [Egbert and Erofeeva, 2002] was adopted to separate oceanic and atmospheric loading. Linear interpolation was used to re-sample global corrections to hourly samples that then entered the main uncertainty analysis.

The Atmacs global model was used in addition to the three main models to assess the effect of model spatial resolution on the uncertainty of the correction. The

Atmacs model [Klügel and Wziontek, 2009] did not enter the main global analysis because criterion (3) was not met (starting in February 2010 or 2012, see Atmacs global model at `atmacs.bkg.bund.de/data/`, access date 2 July 2018). The discontinuities of the Atmacs model caused by the change of the input weather forcing data were removed using mean difference to both ECMWF Operational and Interim models.

The separate aforementioned analysis combining in situ pressure with global model was carried out only at nine of the eleven study sites where global model data overlapped with available in situ pressure observations (see Table S1 in the Supporting information). A single admittance factor equal to $-3.0\,\mathrm{nm/s}^2$ per hPa was applied to account for the differences between in situ and model pressure data, called pressure residuals. The choice of the admittance factor affects the uncertainty. A more objective estimation of the admittance factor would require local digital elevation as suggested in Mikolaj et al. [2015]. The pressure residuals multiplied by the admittance factor were added to the total correction that was computed using each global atmospheric model.

5.3   Methods

The uncertainty of the correction models was addressed both in terms of the uncertainty at any point in time (non-differential mode) and the uncertainty of gravity differences (differential mode). The latter refers to many time domain applications of terrestrial gravimetry where the relative gravity change over a certain time is of major interest to, e.g., quantify the change of mass in given interval. These time intervals are hereafter denoted as periods of interest. Furthermore, the uncertainty of the long-term trend was estimated by looking at the slope of a degree one polynomial fitted to the deviation between models. The uncertainty of trend estimation in absence of the hydrological correction is discussed in Van Camp et al. [2010, 2016a]. In the differential mode, each period of interest was separated from the previous one (with shorter period) by an anti-aliasing zero-phase low-pass filter (see Supporting Information Figure S1). Due to the need for an efficient number of filter coefficients and the corresponding edge effect, the filtered time series were shortened to four years in the differential mode. The differences were then computed for a set of periods of interest ranging from one hour up to one year. The maximal analyzed period of interest in the analysis combining in situ pressure and global atmospheric model was set to two weeks. This limitation reflected the gaps in pressure observations, preventing efficient application of anti-aliasing filters for long periods. To ensure that the two uncertainty modes can be directly compared, an identical time span was used to compute the uncertainty indicators in the non-differential mode. This did not not apply to the assessment of the long-term trend where the full 8 year-long time series were used.

In the following, the differences between correction models affecting their uncertainty are denoted as deviations. The deviations were computed at each site in a loop, selecting one reference (i.e., one particular combination of tides, non-tidal ocean, hydrology and atmosphere models taken as the virtual true correction) and one alternative (another combination of the four component models). The alternative was then subtracted from the reference, resulting in a time series of deviations. This loop was repeated for all unique combinations. Combinations leading to zero
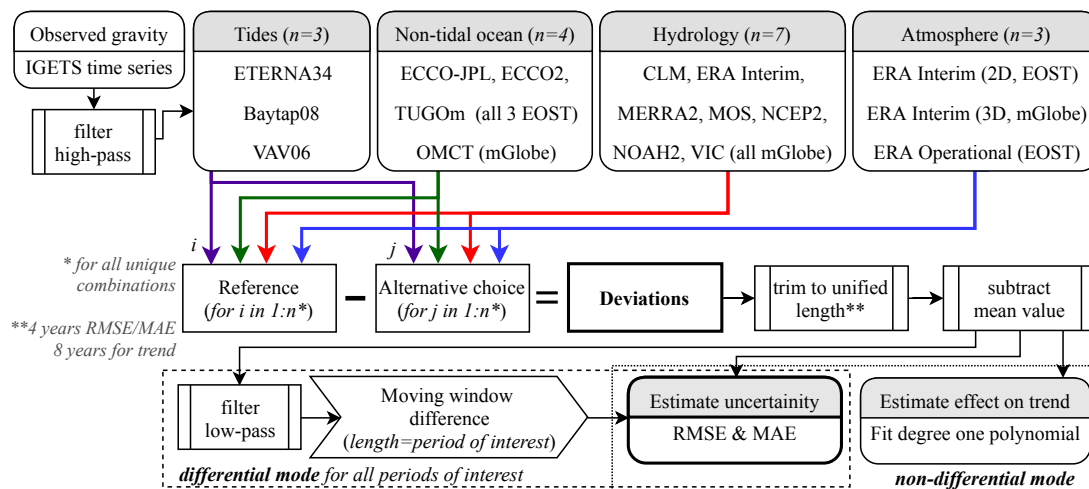
Figure 5.2: Uncertainty analysis work-flow diagram for a single site.

deviations due to the identical reference and alternative models were not treated as unique. Nonetheless, combinations that result in zeros for individual components (e.g. because the same hydrological model was taken) but non-zero deviations for the full correction were not excluded. This approach allows for selecting combinations where only one component (correction) affects the deviations. Thus, the simplified rule of uncorrelated uncertainty propagation can be applied to estimate the arbitrary combination of components. Combinations leading to deviations with opposite sign were treated as identical, i.e., used only once in computing uncertainty indicators. These criteria led to 13695 combinations at each site considering all four components. As a measure of uncertainty, the root-mean-square error (RMSE) of the deviations was computed. Provided here are the averaged RMSE values over all combinations. Similarly, the mean absolute error (MAE) was computed as an additional uncertainty indicator (results presented in the Supporting Information). Due to the subtraction of mean value of deviations prior to uncertainty analysis, the RMSE is in essence identical to standard deviation. Only negligible differences were observed between the averaged RMSE and standard deviation caused by the application of filters after mean subtraction in the differential model. Figure 5.2 summarizes the complete work-flow of the global uncertainty analysis applied at each study site.

To assess the significance of each correction in relation to its amplitude, ratios between the non-differential RMSE and the peak-to-peak amplitude of the correction were computed. Figure S1 shows the mean amplitudes (half the peak-to-peak range) of each effect in the frequency domain. The time domain amplitude as well as the non-differential RMSE were computed as a mean over all available models for the respective correction. Any correlation in this study is expressed in terms of the Pearson correlation coefficient.

5.4   Results

The uncertainties of the correction models in the non-differential mode are shown in terms of RMSE and for the long-term trend in Figure 5.3. Results for MAE are shown in Figure S2. The uncertainties of the tidal corrections are the smallest
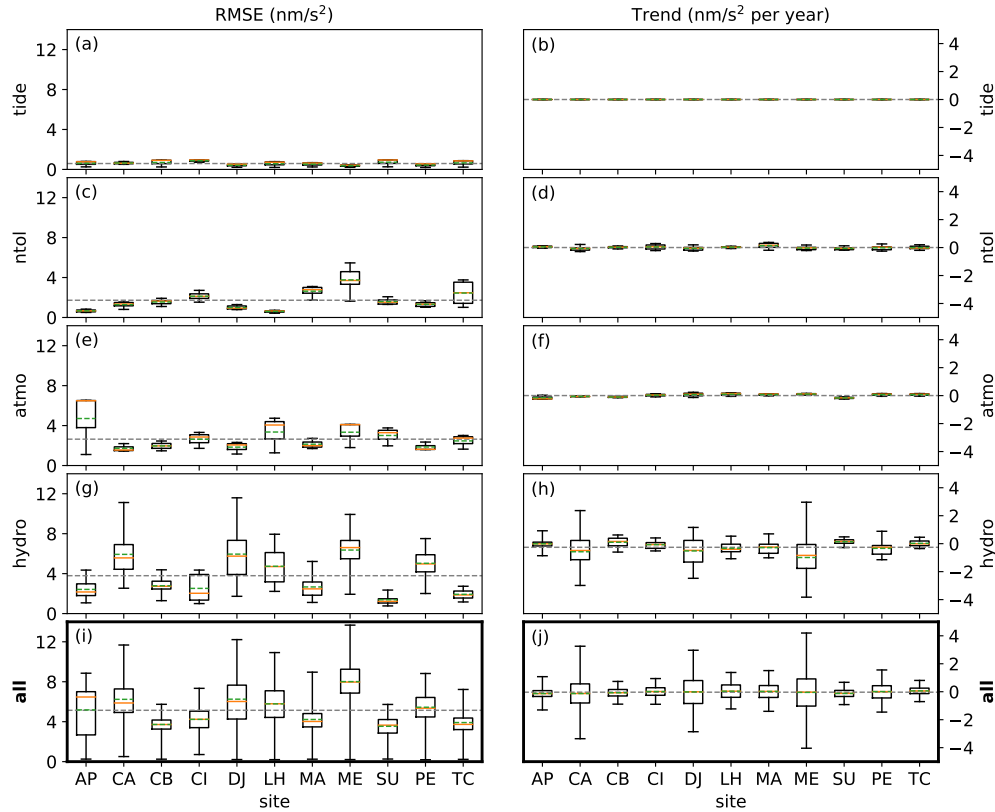
Figure 5.3: Correction uncertainties in non-differential mode expressed as RMSE (a,c,e,g,i) and for the long-term trend (b,d,f,h,j) for each study site (columns), for each component (rows), and for the combined uncertainty from all components (lowest row denoted as 'all'). Boxes mark the 1st and 3rd quartile of all possible combinations, whiskers show the range, median in orange, mean in green. The gray line is the mean over all sites.

among all components. The mean RMSE value over all sites lies in case of tides below $0.6 \, \mathrm{nm/s^2}$. The mean RMSE of non-tidal ocean loading effect equals $1.7 \, \mathrm{nm/s^2}$, $2.6 \, \mathrm{nm/s^2}$ in case of atmosphere, and $3.8 \, \mathrm{nm/s^2}$ for large-scale hydrology.

Above-average RMSE values of the non-tidal ocean loading effect are anti-correlated with the site distance to the ocean ($r = -0.8$, p-value $= 0.003$). The distance of the above-average RMSE sites Cibinong, Matsushiro, Metsahovi and Concepción to the coastline is below 63 km and only 11 km for the latter two sites with the highest uncertainty. Accordingly, the uncertainty of the non-tidal ocean loading correction is smallest (below $1 \, \mathrm{nm/s^2}$) for sites far from the coastline such as Lhasa and Apache Point. These two sites show the highest uncertainty related to correction of atmospheric effects. This uncertainty is mainly driven by one outlying model, namely ECMWF Operational (Figure 5.3(e)). With the exception of Metsahovi, all sites showing this feature (Apache Point, Lhasa, and Sutherland) are located in high altitude (Table 5.1). As shown in Mikolaj et al. [2015], the low resolution of ERA Interim leads to insufficient modelling of atmospheric masses close to the actual topography of the Earth. This effect is minimized combining atmospheric model parameters with in situ pressure observations leading to decreased RMSE with the exception of Metsahovi as shown in Figure 5.5(a). Thus, the increased uncertainty at Metsahovi is related to discrepant modelling of atmospheric circulation patterns

Table 5.2: Ratio of RMSE and peak-to-peak amplitude of the corrections (rounded to 1 decimal place)

site ID	tide	ntol[a]	atmo[b]	hydro[c]
AP	0.2‰	14.7%	5.2%	17.2%
CA	0.2‰	17.6%	0.9%	14.4%
CB	0.2‰	14.0%	1.2%	17.3%
CI	0.3‰	12.9%	8.2%	17.5%
DJ	0.1‰	15.6%	5.3%	18.0%
LH	0.2‰	17.1%	3.9%	22.5%
MA	0.2‰	17.2%	1.3%	26.8%
ME	0.2‰	16.2%	1.2%	18.2%
PE	0.2‰	15.6%	1.0%	13.6%
SU	0.2‰	13.3%	3.0%	20.4%
TC	0.2‰	20.6%	2.0%	11.6%
mean[d]	0.2‰	15.9%	3.0%	18.0%

[a]non-tidal ocean loading, [b]atmosphere
[c]continental hydrology, [d]average over all sites

in the particular region rather than to the topographic effect. Overall, the inclusion of in situ pressure decreases the mean RMSE (computed for all 11 sites) from $2.6\,\mathrm{nm/s^2}$ to $1.4\,\mathrm{nm/s^2}$ (computed for 9 sites).

The highest contribution to the total uncertainty comes from large-scale hydrology. The mean site-dependent RMSE ranges from $1.4\,\mathrm{nm/s^2}$ (Sutherland) to $6.4\,\mathrm{nm/s^2}$ (Metsahovi) and reaches a maximum of $11.6\,\mathrm{nm/s^2}$ (Djougou) for individual combinations. Similarly to the non-tidal ocean loading effect, the uncertainty correlates with the mean peak-to-peak amplitude of the correction ($r = 0.92$, p-value $< 10^{-5}$). In addition, the RMSE for large-scale hydrology correlates with latitude ($r = 0.71$, p-value $= 0.015$), indicating higher uncertainties in the gravity residuals after hydrology correction in higher latitudes. Due to the major contribution of large-scale hydrology, significant correlation with latitude also translates to total uncertainty ($r = 0.78$, p-value $= 0.005$). At four out of eleven sites, the maximum RMSE exceeds $10\,\mathrm{nm/s^2}$ with an absolute maximum of $13.7\,\mathrm{nm/s^2}$. The overall mean RMSE over all sites and corrections equals $5.1\,\mathrm{nm/s^2}$.

At none of the sites and for none of the corrections, the ratio between the RMSE and the peak-to-peak amplitude exceeds 30%. Minimal ratios close to zero were found for tides. On average, the ratio between the RMSE and peak-to-peak amplitude for atmosphere equals 3.0%, 15.9% for non-tidal ocean loading and 18.0% for large-scale hydrology. Hence, the peak-to-peak amplitude of the modelled continental hydrological effect is on average 5 times higher than its uncertainty underlining the significance of the correction despite the increased RMSE. A list of all ratios can be found in Table 5.2.

The non-differential results for long-term trends (Figure 5.3(b,d,f,h,j)) imply that the deviations between models do not have systematic effects on the trend of corrected gravity. The effect on the trend is on average close to zero for all sites and corrections. The trend for tides, non-tidal oceanic and atmospheric loading is insignificant also when looking at individual model combinations. Only the large-scale hydrology seems to affect the long-term trend for some combinations, with a
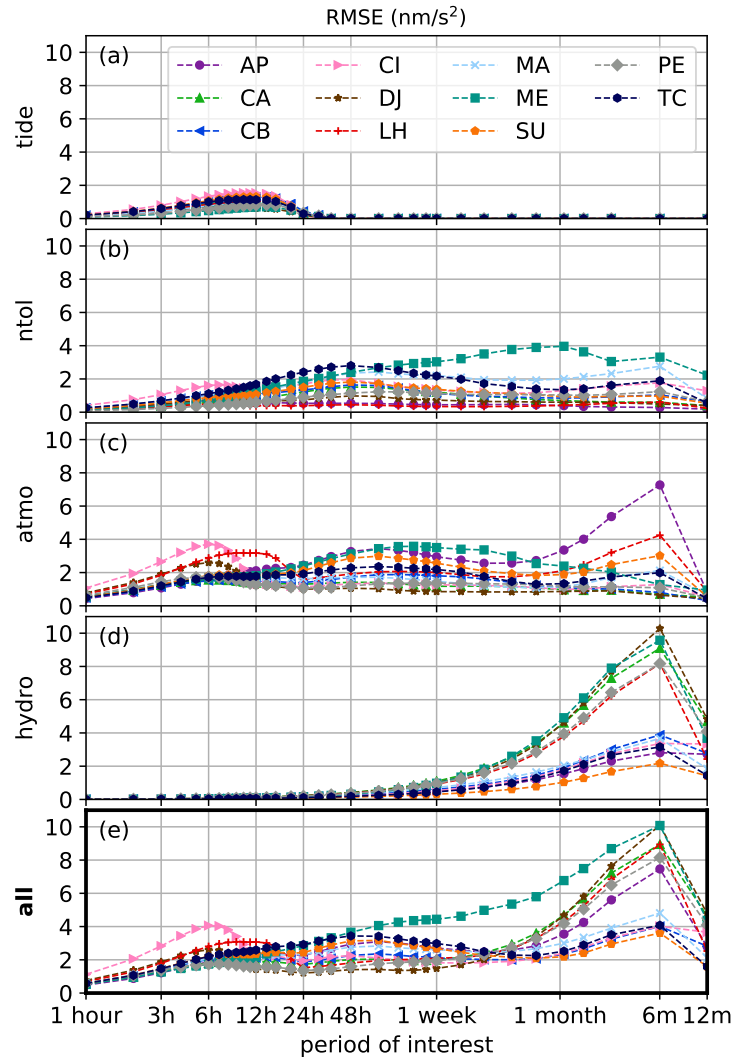
Figure 5.4: Correction uncertainties in differential mode, for each study site and each component (a-d), and for the combined uncertainty from all components (e). X-axis in logarithmic scale.

maximum effect within $\pm 4\,\text{nm/s}^2$ per year and the zero line within the 1st and 3rd quartile (Figure 5.3(h,j)).

Uncertainties of correction models in the differential mode are shown in terms of RMSE in Figure 5.4 and for MAE in Figure S3. For the sake of clarity, these figures present only the mean value over all combinations of the respective correction components at the given periods of interest. Like in the non-differential mode, tide models play only a minor role for total uncertainty. This applies especially to the uncertainty assessed by comparing the individual tidal models as presented in this study. Actual environmental signals such as evapotranspiration may still interfere with tides, leading to systematic effects that are not reflected in the uncertainty values estimated here. The RMSE reaches a maximum of $1.5\,\text{nm/s}^2$ at 12 hours in Cibinong. For the same period, the minimum RMSE lies at $0.6\,\text{nm/s}^2$ in Djougou. The mean over all worldwide sites equals $1.1\,\text{nm/s}^2$ (Figure 5.6 and S4). Tidal waves with a period larger than 33 hours were not considered here because of the unknown instrumental effects interfering with the long-periodic tidal waves

(RMSE equals 0 in Figure 5.4(a)). The non-tidal ocean loading correction affects the uncertainties of gravity changes across all periods of interest. The RMSE does not exceed $2\,\text{nm/s}^2$ for most sites and periods of interest. This value is exceeded only at three sites close to the coastline (Matsushiro, Metsahovi, Concepción). The mean RMSE reaches a maximum of $1.6\,\text{nm/s}^2$ at 48 hour period. The uncertainty of the global atmospheric correction on gravity in the differential mode shows several local maxima. These maxima highlight the different dominant atmospheric processes at the different sites. At Cibining, Djougou, Sutherland and Lhasa, the atmospheric mass transport is driven mainly by solar atmospheric tides S2 and S1. Discrepancies in modelling approaches of these phenomena by different atmospheric models leads to increased uncertainty. In turn, clear maximum RMSE of 3.7, 2.6 and $1.8\,\text{nm/s}^2$ is found at 6 hours for the former three sites and $3.2\,\text{nm/s}^2$ at 12 hours in Lhasa. The maximum uncertainty lies at the half of the dominant period because deviations of gravity changes in time are assessed in the differential mode. Therefore, a deviation with a clear 24 hour period will lead to high uncertainty for 12 hour differences, computed in the worst case from the difference at the maximum and minimum of a periodic oscillation. An identical phenomenon is observed for large-scale hydrological corrections with a dominant 12 month period. For the remaining sites, local maximum can be found between 6 hours to 5 days ranging from $1.7\,\text{nm/s}^2$ (Pecny) to $3.6\,\text{nm/s}^2$ (Metsahovi) and a mean maximum value over all sites of $2.1\,\text{nm/s}^2$ around 72 hours (Figure 5.6)). The combination of a global atmospheric model with in situ pressure data reduces the uncertainty at all site and for all analyzed periods of interest with the exception of Metsahovi at periods of around one week (see Figure 5.5(b)). On average, the RMSE in the differential mode is reduced after inclusion of in situ pressure by a factor of 2.4. The uncertainty related to the correction of the large-scale hydrological effect shows a distinct maximum at 6 months and exceeds $10\,\text{nm/s}^2$ at one site (Djougou). The mean maximum RMSE for large-scale hydrology over all sites equals $5.9\,\text{nm/s}^2$. The total uncertainty in the differential model over all sites and corrections (Figure 5.6) is dominated by the large-scale hydrology with maximum of $6.7\,\text{nm/s}^2$ at the 6 month period. The total RMSE exceeds $2\,\text{nm/s}^2$ at all periods longer then 5 hours. The minimum uncertainty of $0.6\,\text{nm/s}^2$ is found for hourly gravity variations owing especially to the atmospheric correction. This value is reduced to $0.2\,\text{nm/s}^2$ if the correction is computed by combining the in situ pressure data with the global model.

5.5   Discussion

The comparatively high uncertainty of the atmospheric correction is mainly caused by the deviation between ERA Operational and Interim in high-altitude regions. The prioritization of high-resolution models one may argue for, however, does not necessarily lead to more trustful results. The uncertainty in terms of RMSE is not reduced when replacing the correction based on ERA Interim model by the Atmacs correction which makes use of a high-resolution weather model similar to ERA Operational (see Figure S5 in Supporting Information). This example underlines the inherent difficulty to assess which correction model might performance best at a certain study site. Unless all corrections are taken into consideration, criteria such as reduction of residuals used in similar studies [Xu, 2017] cannot by reasonably employed in case of terrestrial gravity measurements. This can be illustrated for un-
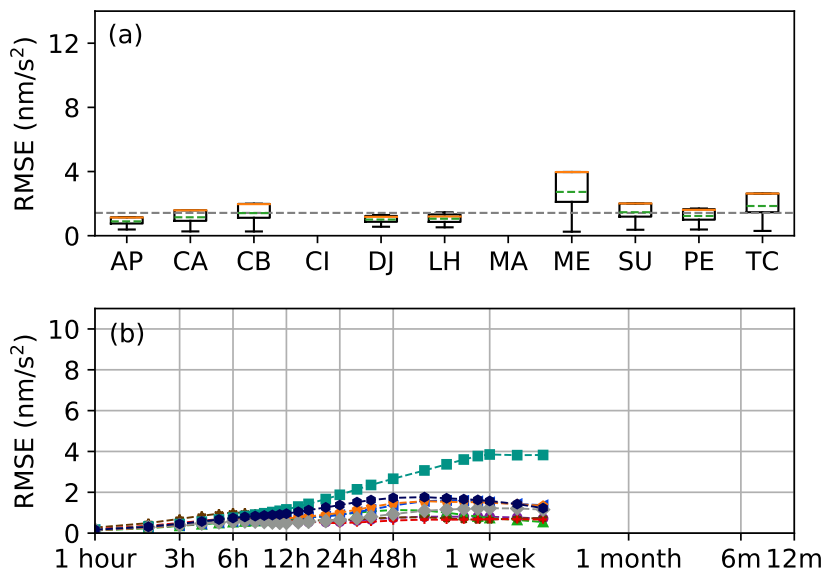
Figure 5.5: Non-differential (a) and differential mode (b) RMSE of the atmospheric correction computed combining in situ pressure and global model at selected study site. See Figure 5.3 and 5.4 for legend and other details describing the (a) and (b) plots respectively.

derground gravity sites where the local hydrology interferes with large-scale signals and any correction of large-scale effects leads to increased variability of residuals regardless of the precision of the used model [Longuevergne et al., 2009]. Different criteria for model performance, such as choice of model closest to overall mean, are also problematic, especially because of systematic errors. Any model-specific deficits, e.g., missing groundwater compartment in hydrological models, the tidal wave grouping, non-consideration of long-periodic tides, choice of pressure admittance, or quality of gravity series pre-processing (gap filling, discontinuity correction etc), are not reflected in the uncertainty indicators. Furthermore, the estimated uncertainties reflect the dispersion among the models used here within the studied period. Any new development of correction models may lead to new estimates. The dispersion of the atmospheric corrections is significantly reduced when combining global model with in situ pressure. Such an approach tends to diminish the model discrepancies towards a given reference (in situ pressure) without knowledge of the true total atmospheric gravity effect. Thus, the results do not describe the absolute precision of gravity residuals, but rather the precision of corrections assessed using state-of-the-art global models available to users worldwide.

Uncertainty assessed in this manner is limited by the rather small number of study sites. The low number of IGETS sites needed for the tidal analysis, especially in latitude range ±20 degrees, manifests itself in clustering of the RMSE due to large-scale hydrology correction into the two groups prominent in Figure 5.4(d). Nevertheless, the study sites cover an extended latitude and altitude range, distance to coastline, all continents and climate zones with the exception of polar region. Hence, the extreme values are well represented by the characteristic sites of the respective region. The use of regional models that assimilate local observations might be advisable in certain regions, e.g., at Metsahovi with distinct non-tidal and
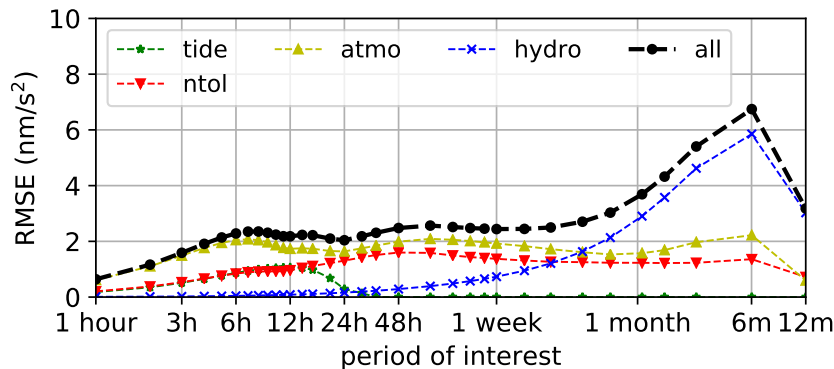
Figure 5.6: Correction uncertainties in differential mode, computed as mean over all study sites for each component (combined contribution denoted as 'all') using global atmospheric correction without in situ pressure. X-axis in logarithmic scale.

atmospheric effects in the Baltic Sea. These models should, however, be combined with large-scale models to account for global mass transport processes, especially when using high-precision superconducting gravimeters capable of detecting such phenomena.

Depending on the field of research, the uncertainties in gravity units $(\mathrm{nm/s^2})$ obtained here can be converted to units of interest, such as kg, mm of water, or mm of land uplift, using gravity modelling, or they can be directly compared to the expected signal amplitudes of interest, as listed in the review article of Van Camp et al. [2017]. Such comparisons show that the uncertainty related to gravity corrections does not hinder the resolution of substantial effects of interest in most of the research fields including hydrology, volcanology, reservoir monitoring, or local subsidence, all exceeding tens of $\mathrm{nm/s^2}$. However, subtle effects in the order of few $\mathrm{nm/s^2}$, i.e., on the lower part of the expected signature level, will be difficult to be resolved. This is because of the uncertainty of the gravity corrections adding to the instrumental precision (see Introduction). Therefore, a meaningful use of gravimeters in studies focusing on subtle changes requires employment of advanced processing techniques such as stacking, filtering or use of additional constrains suppressing random errors, e.g., as employed in Güntner et al. [2017]; Van Camp et al. [2016b].

5.6   Conclusions

The uncertainty of terrestrial gravity correction models for tides, atmospheric, large-scale hydrological and non-tidal ocean loading effects was assessed using a large set of available global models. The presented values can be used to estimate the correction-induced uncertainty in cases where users apply one or all of the mentioned corrections in order to reduce the environmental effect on observed gravity. The obtained root-mean-square error at 11 sites worldwide equals $5.1\,\mathrm{nm/s^2}$. This value gives the average uncertainty that affects gravity residuals without taking local corrections of hydrology or in situ pressure into account. The estimate does not include other random and systematic errors such as instrumental effects or near-field noise. The uncertainty varies significantly with location and is correlated to geographical latitude owing especially to the large-scale hydrological effects. Depending on the

study site, the uncertainty can range from $3.5\,\mathrm{nm/s}^2$ up to a maximum of $8.0\,\mathrm{nm/s}^2$. A high negative correlation has also been found between the site distance to the ocean and the uncertainty due to non-tidal ocean loading correction. The total correction uncertainty is mainly driven by the deviations between large-scale hydrological models with $3.8\,\mathrm{nm/s}^2$, as compared to $2.6\,\mathrm{nm/s}^2$, $1.7\,\mathrm{nm/s}^2$ and $0.6\,\mathrm{nm/s}^2$ for global atmospheric, non-tidal ocean loading and tides correction, respectively. For many studies focused on quantifying mass changes over time, the uncertainty of gravity differences in time are of interest. We have studied differences between 1 hour up to 12 months and found that the hourly differences show the lowest uncertainty in terms of root-mean-square error equal to $0.6\,\mathrm{nm/s}^2$. The analysis using global models shows that the uncertainty of differences longer than 5 hours exceeds on average $2\,\mathrm{nm/s}^2$ and reach a maximum of $6.7\,\mathrm{nm/s}^2$ for intervals of six months. This implies that observing seasonal mass variations by terrestrial gravimetry is burdened with highest correction uncertainty. The overall and the atmospheric correction uncertainty are reduced to $4.4\,\mathrm{nm/s}^2$ and $1.4\,\mathrm{nm/s}^2$ respectively if global models are combined with actual in situ pressure observations. Such approach also affects the uncertainty of gravity differences, reducing the root-mean-square error to $0.2\,\mathrm{nm/s}^2$ for hourly estimates, for instance.

Small to negligible systematic uncertainties on the long-term trend were found. The ratio between the amplitude of the correction and the root-mean-square error (equivalent to standard deviation σ) shows that the peak-to-peak amplitudes are in most cases at least 5 times higher than the uncertainty. This indicates that the corrections should be, from statistical point of view, taken into consideration despite their uncertainty. Overall, the results of this study demonstrate that substantial mass changes associated with a diversity of geophysical processes can be confidently resolved by terrestrial gravimetry even though necessary signal corrections impose non-negligible uncertainties. Thus, any direct resolution of subtle gravity changes in the order of few $\mathrm{nm/s}^2$ might be hindered by the total uncertainty composed of the instrumental and correction precision. Consequently, advancing to higher-precision gravimetry application is not only subject of further technological development of the gravimeters themselves, but also of improving the representation of mass transport processes in various components of the Earth system and at various spatial scales.

Chapter 6

Synthesis and outlook

6.1 Synthesis

The four studies presented contribute in different aspects to the overall objective of examining and evaluating the gravimetric method as a suitable tool for field hydrologists. To this end and directly related to the stated research aim and questions in chapter 1.4, the studies are discussed in the context of (i) what is necessary to deploy a gravimeter directly at a field site and what is the added value (chapter 6.1.1), (ii) what are advantages and disadvantages when using a gravimeter for applied process investigations and what is its potential (chapter 6.1.2) and (iii) how can hydrological signals be extracted from gravimetric observations related to different spatio-temporal scales and precisions with respect to magnitudes of hydrological processes (chapter 6.1.3).

6.1.1 From observations in buildings towards field deployment

Stationary gravimeters are traditionally located in dedicated buildings with pillars isolated from the building itself. This provides a stable device-surrounding temperature regime and also protects the instrument from environmental influences such as precipitation. These buildings usually serve multiple purposes (e.g. gravimeter comparisons, instrument maintenance and calibration, hosting other equipment) wherefore its size is not optimized for solely protecting the gravimeter itself. Building footprints for current installations vary between 10 and 400 m^2 [Reich et al., 2019]. This footprint is a sealed surface which impedes natural exchange of water mass (infiltration, evaporation) and consequently the direct vicinity of the gravimeter does not reflect natural conditions. As most of the gravimetric signal originates in this near field around the instrument [Creutzfeldt et al., 2008], the observed gravity changes include a signal mix (natural, non-natural) of local hydrological mass variations which is hard to disentangle. Depending on the research objective (e.g. near-surface mass variations, deep mass variations, spatial extent of interest, etc.), this can dramatically limit the usefulness for hydrological investigations. It can furthermore be problematic if the gravity signal should be corrected for this local hydrological component to focus investigations on different signal phenomena (e.g. in volcanology or geothermal energy). There are two options to overcome this limitation, both addressed in this thesis: (i) measuring soil moisture, precipitation,

evapotranspiration and groundwater via sensor installations, and (ii) minimizing the footprint by moving the gravimeter to a an outdoor field enclosure.

The first option focuses on quantifying the impact the shielding of the building constitutes, this is called *umbrella effect* or *masking effect*. Prior to this study, the recommendation was to set all mass variations below a building equal to zero. Reich et al. [2019] showed that this could not be confirmed as the best option but instead quantifying them in an indirect way by means of hydrological measurements. **The approach described in detail in chapter 2 showed that measurements of soil moisture taken next to the shielding area can be used jointly with a classification of soil texture, climate regime and the building dimensions to deduce mass variations beneath it**. The study was based on time series of soil moisture and groundwater measurements directly below as well as next to a gravimeter building. **The presented and validated approach underlines the possibility to retrospectively quantify the *umbrella effect* and use it for correction purposes, also in long-existing gravimeter installations with no dedicated hydrological measurements below the building. The study concluded that the proposed method minimizes possible errors (1.9 to 7.8 nm/s^2) in comparison to the previously recommended *zero mass change approach* (4.1 to 12.8 nm/s^2)**. Although the conceptual understanding and quantification of the *umbrella effect* was largely improved, as a consequence also leading to an increased awareness of station operators for this topic, the problem of partial blindness of gravimeters within observatory buildings to near surface mass variations remains.

The second option is described in chapter 3. The main motivation was to increase the sensibility of the gravimeter to near-surface water mass variations by reducing the extent of impervious material below the gravimeter to a minimum. This minimum consists of a single 1-m diameter pillar, necessary to minimize tilting and elevation change of the instrument. **Theoretical experiments showed that the sensitivity depends strongly on the size of the shielding pillar as well as on the height of the sensor above the surface**. For the flat terrain at the installation site, a compromise between high sensitivity, stability of the pillar monument and feasible access for maintenance was found around 1.0 m sensor height. This was proven a good choice by looking at the post-installation observed gravity data. **Comparing observations from the field deployment (iGrav) with a SG in the adjacent building verified the expected higher sensitivity to near surface water storage variations, while noise levels from the field instrument over a range of periods from 1 min to 3 hours were equivalently low as from the indoor instrument**. Water storage changes at different depths within the soil and with varying horizontal integration radii resulted in depth-independent, well-resolved gravity changes at the iGrav, whereas the gravity changes in response to water-content changes in the upper 2 m were lower and with different dynamics for the SG. This is a direct consequence of the umbrella effect. **While the SG shows a delayed and changing response to water infiltrating the soil, the gravity effect at the outdoor iGrav is the same regardless of the depth of the mass change, once the water entered its footprint**. This is true for sufficiently large integration radii (e.g. precipitation events) and considering no changes in lateral water flux at the footprint boundary (e.g. increased groundwater flow) or vertical mass extraction (e.g. evapotranspiration). **Observing gravity by instru-**

ments installed outdoors is consequently highly beneficial for direct and
unbiased water storage change monitoring. It enables a more precise pos-
sibility to observe hydrological dynamics at the crucial land-atmosphere
interface and investigations about dominant processes of the water cycle
such as precipitation, evapotranspiration, surface runoff and soil water
storage.
      Challenges of a field deployment are two-fold. First, the gravimeter still has
to be operated under stable conditions. This requires a housing able to maintain
a steady temperature, shielded from wind, direct sun, and humidity. Second, a
gravimeter demands a continuous power supply. This can be a limiting requirement
for the selection of an installation site. There is a broad range of power consump-
tion depending on the gravimeter type. Superconducting gravimeters (SG, iGrav)
need 1.5 to 2 kW, absolute gravimeters from 500 W (AQG, FG-5) to 300 W (A10)
and spring gravimeters between 150W (gPhoneX) and 5 W (CG-6) [GWR, 2013;
Muquans, 2019; Micro-g, 2015, 2008, 2013; Scintrex Ltd., 2017]. Even though some
devices can be operated by batteries, this is no option for long-term field instal-
lations. **The experiences laid out and results shown in chapter 3 clearly
demonstrate the feasibility of operating a superconducting gravimeter in
the field, fulfilling all requirements with respect to housing and power
supply**. A successfully proven field enclosure (for continuous installations) by of-
ficial manufacturers exists only for the iGrav, discussed in this work. Nevertheless
individually-built field enclosure solutions or equivalent field setups fulfilling the re-
quirement on stable environmental conditions were tested by various research groups
[Wilson et al., 2012; Kennedy et al., 2013, 2014; Reich et al., 2020], allowing to profit
from the added value of deploying a gravimeter directly in the field.

6.1.2   New approaches for experimental hydrological process investigations

With the intention to explore in more detail the enhanced capability to resolve near-
surface water mass variations as one of many possible show-cases for hydrological
field research, a sprinkling experiment was designed at the field deployed iGrav site.
Framing, equipment, setup and scope is discussed in detail in chapter 4. This setup
took advantage of the increased near-surface sensitivity of the iGrav provided by the
outdoor installation. At the same time, **the previously discussed feature of a
field setup to be depth-insensitive with respect to water mass variations
once entered the footprint does not apply in this case. On the contrary,
it is taken advantage of all mass changes occurring within a spatially
limited and controlled domain, the sprinkling domain. This enables to
use the recorded gravity changes as the integral system response to this
known mass increase**. The remaining problem, part of all integrative monitoring
methods, is about equifinality: infinite solutions can be found to match the observed
signal. This can only be overcome by reasonable fixed boundary conditions and/or
independent monitoring of the same system by other methods. Type, values and
settings of these vary depending on the experiment, investigated quantities and sur-
rounding environmental conditions. For the dedicated sprinkling experiments the
possibilities could be knowledge about soil porosity, soil type, infiltration capacity,
depth of the groundwater level, possible impermeable layers, local topography, and

surface-runoff behaviour. As shown in chapter 4, the approach was chosen to complement gravity observations on one hand with additional, independent monitoring and on the other hand with optimized, inverse hydrological process modelling. This framing and combined evaluation worked well to show that **theoretically modelled shapes and magnitudes of gravity signals, resulting from the mass increase and characteristic for individual subsurface water distribution processes, were identified also in the actual experimental data. The characteristic shape of the gravity response curve was the dominant factor for a successful process identification**.

Maybe even more important, however, were the limitations of the applied approach. **First, the hydrological model describing subsurface water distribution processes was too simple and did not reflect the high spatial heterogeneity of the soil. Consequentially, some spatial dynamics could not be investigated with respect to their resulting gravity signal**. In conclusion it is recommended to use a finely-discretized physically-based model in order to have a different spatially representation of natural processes, given soil parameter and hydrological flux data availability to constrain the model. Second, **there was still the necessity of additional measurements (ERT and soil moisture), including invasive sensor installations**. This roots again in the problem of equifinality and contradicts the strong motivation for gravimetry as a non-invasive field method. One of the downsides of traditional field hydrology setups is that installations require digging holes and trenches, drilling wells, injecting tracers, etc. The desire to establish gravimetry as an applicable field method is to a certain extent based on its non-invasiveness, making non-reversible modification of the research environment obsolete (except the installation of a pillar). **To benefit from the non-invasive nature of gravimetry and to overcome the issue of non-uniqueness, networks of gravimeters for such experiments should be pursued more intensively**. Many different network implementations are possible, depending on spatial extents of the study area, the desired temporal resolution, expected magnitudes of target processes in combination with accuracy of individual gravimeter types and most of all the objective of the research itself.

Applied hydrological field research could profit for spatially large extents in areas such as catchment dynamics and states, landscape behaviour understanding for flood generating process investigations and water budgets also in combination with long-term trends relevant for global climate change. In smaller spatial extents its benefit would be for investigating forest hydrology, vadose zone monitoring including groundwater recharge in the context of critical zone processes up to infiltration patterns and processes relevant for agricultural studies as well as in dedicated projects for geothermal fluid monitoring or carbon capture and storage.

6.1.3   Resolving hydrological signals at different spatio-temporal scales

Independent of the actual signal of interest of gravimetric data-based investigations, isolating the target signal component(s) from the observed gravity time series always follows the same processing schema: all signals that are of no interest for the study have to be subtracted from the observations. This process is called *reduction* and is based on the general composition of all represented gravity effects, shown in Eq. 1.7.

The local hydrological component can be split up into individual signals as shown in Eq. 1.8.

From restructuring equations 1.7 and 1.8 then follows:

$$g_{targetSignal} = g_{obs} - g_{nonDesiredSignals} \qquad (6.1)$$

Hence, all non-desired signals have to be quantified to correct for them. In general, all signals can be categorized as either possible to measure, or requiring modelling. Some signals have both options, e.g. groundwater-storage changes can be measured on-site in a well (if the aquifer storage coefficient is known), or extracted as information from a hydrological model. For the atmospheric component, there is the option to use global models, an admittance approach based on local pressure measurements, or a combination of both. When using model data, the applied model has to be chosen carefully to ensure that all effects expected to impact the observations, and which are important for the focus of the research questions, are included in the model in a sufficient spatio-temporal resolution. This is especially true for global models. Consequentially the choice of which model to use highly depends on biased experiences of station operators or the person responsible for the reduction processing. Chapter 5 addresses this topic in details, examining correction-induced uncertainties for large scale gravity effects (tides, atmosphere, non-tidal ocean loading and global hydrology) for different time periods. **The average total uncertainty found, expressed as the average root mean square error over all 11 investigated sites and all effect-types, was of 5.1 nm/s^2 (corrections of local hydrology and in-situ pressure not considered). Whereof uncertainties of the individual effects contributed in the order: global hydrology, atmosphere, non-tidal ocean loading and tides (declining mean RMSE values). The range between sites was spread from 3.5 nm/s^2 to 8.0 nm/s^2, pointing to the influence of the distance to the ocean (affecting non-tidal ocean loading correction) and the latitude coordinate of individual sites (affecting global hydrological correction)**. When conducting research on specific (hydrological) signal components, one important aspect is the frequency of this process of interest, also with respect to the study objective. For evapotranspiration, for example, one could be interested in its daily dynamics and/or its annual variation. **The overall reduction uncertainty stated above, encompassing all frequencies of the individual effects, will most likely be insufficient for capturing possible errors in a specific frequency range. These frequency dependent errors of individual effects are directly linked to the frequencies of the target signal to be investigated (e.g. as in the given example of evapotranspiration). Consequentially, an error quantification based on individual time periods is necessary. For the periods considered in the analysis within chapter 2, hourly uncertainties were smallest (0.6 nm/s^2) while 6 monthly frequencies showed highest values (6.7 nm/s^2). It has to be mentioned that the high correction-induced uncertainties are strongly driven by global hydrology**. Generally and in summary, these findings demonstrate that: (i) research of a phenomena of interest implies reducing the observed gravity time series for all non-related signal components, (ii) this reduction introduces a correction-induced uncertainty, (iii) the uncertainty is influenced partially by site location and (iv) the magnitude of the uncertainty highly depends on the frequency of the target signal. Consequentially, the possibility exists that the de-

sired target signal could be masked by the reduction uncertainty. Preparation of investigations using terrestrial gravimetry data should therefore consider not only the instrumental accuracy but additionally the magnitude and frequency of the signal of interest and how respective correction-induced uncertainties will influence the study setup. **With respect to hydrogravimetric studies, processes generally small in gravimetric magnitude (e.g. evapotranspiration, precipitation events depending on its temporal scale) or processes around frequencies of 6 months with a magnitude small in relation to its spatial extent, are most prone to be masked by the reduction process**.

6.2   Conclusions and outlook

With a general view on the previously shown and discussed findings, the hygrogravimetric method has been investigated from an instrumental, methodological and applied point of view. Moving the gravimeter outdoors, where hydrological field studies are carried out, evades the *umbrella effect* and provides a substantial benefit for all kinds of research perspectives. This goes along with no major disadvantages, except for challenges in the technical setup which are manageable. The presented case study with a superconducting gravimeter can be conceptually transferred to other gravimeter types. They all have different requirements in terms of installation setup, maintenance, stability, accuracy and costs which should be considered in relation to the study objective. There is also no bound necessity to implement a permanent installation. While for many research targets long and continuous recordings of mass changes will be an asset, experiment-based research for looking at e.g. soil properties, infiltration capacities, biomass transfer or vadose zone processes could also be already satisfied with a more temporary setup.

It should be emphasized that there is a fundamental difference if gravity observations are planned within a stationary setup or based on field surveys. Different gravimeters with device specific properties are better suited for one or the other setup type. However, most important is a sufficient measurement precision to resolve the process or phenomena of interest. Additionally, dominant frequencies and the rate of change of the process(es) have to be taken into account. To give an example: Evapotranspiration has a rate of change which lies between one and several hours. This is generally well suited for relative gravimeters but not for absolute devices due to long measurement integration times (to achieve high accuracy). The actual mass change induced by this process is rather small, e.g. a maximum of 4.5 mm water during one daily cycle at the Wettzell observatory in South-Eastern Germany [Güntner et al., 2017]. When converting with Bouguer [Pasteka et al., 2017], this would lead to a gravitational attraction of 1.89 nm/s^2. Given the rate of change and the expected magnitude of the mass change, this could only be resolvable with an iGrav and therefore also imply a stationary setup. For processes like soil moisture or groundwater, the changes in magnitude are much higher, leading to mass attraction effects in the order of tens of nm/s^2. Their rate of change has a much broader range from hours (soil moisture increases after precipitation events) up to days, weeks or months (decreasing while drying). Consequentially, with respect to precision and measurement time, there are many different gravimeter types suitable for monitoring these example processes. Additionally or instead of stationary setups, gravity field surveys could be conducted if spatial features of these processes are of interest

(e.g. CG-6).

Absolute gravimeters are less suited for short-term dynamics but have the advantage to provide absolute gravity values. If several of these absolute gravity measurements are available at a location where continuous relative gravity measurements are conducted, the relative time series can be drift corrected. This enables the relative observations to act as a continuous reference for spatially distributed relative gravity surveys carried out at specific, limited time periods. During a single survey, measurements are taken at numerous points in the landscape. The reference station (also called base station), being one of those points, then serves to connect individual observations in a spatial context between each other. The continuous monitoring at the base station furthermore facilitates the combination of various, single surveys in time [Long and Kaufmann, 2013].

As can be seen, there is quite a high variety in gravimeter setups, devices and applications. This might complicate the choice for a desired research target but also provides a high potential of setups. Some years ago the concept of *hybrid gravimetry* was introduced [Hinderer et al., 2016]. By definition, continuous observations at one point (having the advantage of *time*) are combined with field surveys conducted at various points around this base station (advantage of *space*) [Pfeffer et al., 2013; Hinderer et al., 2015a; Quentin et al., 2021]. When the base station itself is located directly in the field, there is the additional benefit that it can sense hydrological mass variations in a direct and unbiased way (as explained above). Furthermore, if the base station is regularly evaluated by absolute gravity measurements, improved drift correction of the survey gravimeter and spatio-temporal connections between survey points lead to higher data quality. Many areas of hydrological research could highly profit from such a framing: riparian zone and hillslope hydrology with respect to changes in water availability during different seasons, spatial monitoring of pre-event states for an improved understanding of flood generation, identification and separation of hydrological research units (HRU) at spatial subsets within a catchment, monitoring of spatially distributed water storage changes within the vadose zone (normally difficult to access), amongst other things.

Multi-gravimeter setups, not necessarily within the concept of hybrid gravimetry, could also contribute to further investigate the aforementioned problematic topic of equifinality. As concluded from the sprinkling experiment (see chapter 4), several gravimeters installed at different distances to and different locations on the sprinkling footprint would enable the attribution of mass change to specific and locatable volumes. Preserving the intentions of non-invasive measurements would certainly also be possible via combination with other geophysical methods, thus also reducing equifinality. But there is a second benefit for using several gravimeters within a rather small domain: large scale effects are observed equally at both sensors. As a result, the discussed correction-induced uncertainties could be nearly eliminated by subtracting one observation from the other [Carrière et al., 2021].

From a theoretical point of view, the most efficient setup for gravity-based hydrological field studies would be a combination from all of the above: (i) two or more continuously operated relative gravimeter as base stations, deployed within a sufficiently small domain to avoid uncertainties from large-scale corrections, providing long-term observations, (ii) frequent surveys at numerous points around these base stations to evaluate spatially distributed mass changes, and (iii) repeated measurements with an absolute gravimeter at the base stations for instrument verification

(mostly drift correction). Certainly such a study design would need both sufficient financial budget and available manpower. From an instrumental point of view, future study designs will profit from developments such as the (early stage) Absolute Quantum Gravimeter [Ménoret et al., 2018], Differential Quantum Gravimeter [Janvier et al., 2022] or microelectromechanical systems (MEMS) [Middlemiss et al., 2016]. Their mobility and field-use capability of these devices is a direct consequence of repeatedly mentioned interest and need by the community for desired application type and area. AQG's could contribute with the advantages of continuous absolute gravity monitoring directly in the field in a portable setup, thus covering several of the above mentioned points in one single device. Being smaller and much cheaper than other gravimeters, MEMS have the potential to highly improve the spatial resolution of hydrogravimetric study setups.

The definition of the spatial extent of a gravity measurement is directly linked to the necessity of a hybrid setup. For rather small study areas (plot to field scale), the footprint of a single gravimeter could be sufficient for successfully monitoring all relevant mass changes. If target objectives of the study are spread in a larger spatial extent, a single gravimeter will not sufficiently capture all mass change processes and a hybrid setup should be considered. Expected frequencies and magnitudes of the expected gravity change also influence the choice of gravimeter, including considerations of instrument drift estimation, necessary power supply, and pillar/shelter requirements. A minimized housing or field enclosure is necessary if the target signals are expected near the surface. If correction-induced uncertainties are large multiple gravimeters may be necessary. If the design includes field surveys carried out periodically, reproducible measurement locations (especially with respect to their height coordinates) are required. Well trained staff with respect to device handling, in-situ field data quality judgement and their temporal availability for covering all planned field activities is a self-explanatory asset.

Nowadays, data organization is a substantial part of the preparation and should be considered in its complete extent, in order to facilitate data quality evaluation and data processing up to publishing final results. The chain involves data acquisition in the field, transfer of data to institutional IT-infrastructure, storage of data in different processing levels, sharing of data with project partners, optional visualization for quick looks, combination with device and field-setup specific metadata for streamless, DOI-based data publications up to a structure for reproducibility of all processing states for scientific publications. Additionally, software used for data processing is in the best-case verified by the community, shareable and in compliance with open-source standards.

The increased interest in gravimeters for hydrological research is reflected by numerous publications within the last years. A high potential for further development in different areas of this research field exists and is desired by many groups. This thesis showed that not only instrumentation improvements are necessary to enhance the usability of hydrogravimetry, but also deeper understanding of local hydrology at existing sites, modified and optimized installation setups for hydrological research orientation, study setup and implied constrains within in-situ based hydrological process investigations and a sensibility to and consideration of linking research target signals with respect to signal magnitude and frequency to possibly expected uncertainties within the reduction process. Complemented by experience-based recommendations about field setups, study design and its preparation, it was

demonstrated how these findings lead to added value for hydrological field research and advanced the hydrogravimetric method.

Code and Data availability

[Chapter 2] **Reducing gravity data for the influence of water storage variations beneath observatory buildings**

Code
The presented reduction routine is accessible and applicable via installing the R-package "UmbrellaEffect" (`https://github.com/marcianito/UmbrellaEffect`). An example script provided within the package can be easily adjusted to individual needs in order to carry out an automatized reduction of observed gravity data for a setup and location of choice. Sample data is also provided within the package.

Data
Upon request by the author.

[Chapter 3] **Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure**

Code
The code necessary for data processing and data analysis including the optimization models is provided in the form of Matlab and R scripts as supplementary material (repository) to this publication. The repository furthermore contains extensive explanatory files with all the instructions how to reproduce the results presented in this study.

Data
The data used in this study will be published via the IGETS (International Geodynamics and Earth Tide Service of the International Association of Geodesy) data base at GFZ Potsdam (`http://isdc.gfz-potsdam.de/igets-data-base/`). Güntner et al. [2017] provide the data of iGrav006, both raw 1 second gravity records and the gravity residuals, as well as the auxiliary hydro-meteorological time series (lysimeter evapotranspiration, precipitation, climate data for calculation of reference evapotranspiration, river discharge) and the spurious PCB temperature effect on gravity. Wziontek et al. [2017b] provide the raw gravity data of the two superconducting gravimeters in observatory buildings at Wettzell (SG029 and SG030).

[Chapter 4] **Field-scale subsurface flow processes inferred from continuous gravity monitoring during a sprinkling experiment**

Code

The hydro-gravimetric model, for both conversion and inversion mode, is accessible and applicable via installing the R-package "gravityInf" Reich [2021]. A therein supplied example script can be adjusted to individual needs in order to analyse sprinkling experiments in terms of infiltration modelling and comparison to measured gravity signal. Sample data is also provided within the package Reich [2021].

Data

A general repository containing processing scripts, input- and output data which where used in this study are available via GFZ Data Services Reich et al. [2021a].

[Chapter 5] **Resolving Geophysical Signals by Terrestrial Gravimetry: A Time Domain Assessment of the Correction-Induced Uncertainty**

Code

Repository containing processing scripts and outputs, as well as an explanatory read-me file with all instructions for reproducing the results presented in this study are available via GFZ Data Services [Mikolaj et al., 2019a].

Data

Repository containing processing scripts and outputs, as well as an explanatory read-me file with all instructions for reproducing the results presented in this study are available via GFZ Data Services [Mikolaj et al., 2019a].

Acknowledgements

First of all I want to thank my wonderful supervisor Andreas Güntner! His support, commitment, dedication, knowledge and advice was what always pushed me one step further and made me revisit and positively question many approaches, analysis, concepts and planned field setups. All the jointly carry out fieldwork demonstrated that he really participated actively in all of my tasks and also led to many interesting discussions away from the office. His overall positive attitude, openness for discussion in flat hierarchies and a tremendous sensitization for individual group member needs, created and supported a wonderful team spirit within the HyGra working group. Thank you for your great leadership in many aspects!

I am deeply grateful for Michal Mikolaj and all the positive influence he had on my academic forming. He was a wonderful colleague and became a friend, who introduced and taught me much of geodesy. Interchanging programming problems and actively sharing the office, not only resulted in a fun but highly efficient working atmosphere. So were the many field campaigns we went on. I miss working with you.

I am very thankful for Theresa Blume and the guidance she provided. With her perfectionism and eye for details I could highly improve many measurement and analyse concepts as well as manuscript texts. She would make me zoom-out of a problem to identify the big picture, which was very helpful.

I feel an extremely huge gratitude to all my dear colleagues from section hydrology. Especially those who were actively involved in my working environment. Short words cannot express what your dedication meant to me: Stefan Lüdtke (where would I be in Linux, computing and problem analysis without you! Thank you also for all the great coffee chats and your direct character), Heiko Thoss (your continuous disposition and helpfulness, days of numerous fieldwork trips, planing and brainstorming ideas as well as all the wonderful talks had a big positive impact on me), Stephan Schröder (thank you for all the diverse positive impacts and vibes, you are always helpful and going to the field with you was always a pure pleasure; it is great to see how our knowledge and skills complement and profit from each other thus leading to improved concepts and setups), Knut Günther (you always lend your ear and help! Your critical view from different perspectives leads to better results and tries to integrate everyone), Markus Morgner (thanks for being always a helping hand, in the field as well as at GFZ), Daniel Rasche (great sharing fieldwork with you and brainstorming about new ideas and concepts).

From the gravimetric community I want to mention some people who accompanied me throughout my PhD and whose diverse influences all jointly contributed to and impacted my scientific formation. Most of all Hartmut Wziontek (BKG) who patiently taught me a lot about tidal analysis but also in general about gravimetric concepts, detailed instrumental insights as well as processing concepts. I value a lot

all his honest and transparent critic and all the time he always took. I was always impressed by the high standards he kept and transmitted to others. Thomas Klügel (BKG) was always my contact for campaigns to the Wettzell observatory and I am so grateful and still astonished that he always finds time to jump in. His helpfulness, both on-site as well as remote saved us often. My gratitude also goes to Reinhard Falk (BKG) for sharing his knowledge and expertise about gravimeters and survey setup and planing. In general I want to thank the BKG for being a great project partner with all the wonderful people working there. Furthermore, I thank Ludger Timmen (LUH) for great and detailed discussions as well as much patience with everything. He always was available for questions, doubts or meter calibrations. I am also very thankful for the impact Christian Voigt (GFZ) had on my academic forming, both in the applied but also in the institutional context. He always lend his ear for problems and it felt really great to have such an honest, target-oriented colleague in this field at GFZ who always helped me to see the bigger picture. I am grateful for the valuable input and thoughts Michel Van Camp (Royal Observatory of Belgium) always gave me when we met. We had interesting discussions and his publications were inspiring for forming new research ideas. I also want to express my deep gratitude to Jeff Kennedy (USGS). He is an exceptional example of a researcher in the field of hydrogravimetry. I learned so much reading his eye-opening studies and am thankful for him sharing his experience and knowledge in gravity field surveys and instrument handling. Additionally, I am indebted to him for helping me tremendously in the final stretch of this work - Thank you!

Last but not least I am endlessly grateful for my family and friends and the support as well as the distractions they provided. Thank you for always being on my side in this long endeavour.

Appendix A

Extensive description of test site and measured datasets

For the observation-based part of this study, the test area is located in the Bavarian Forest in southeastern Germany at the Geodetic Observatory Wettzell, which is operated by the German Federal Agency for Cartography and Geodesy (BKG) [Schlüter et al., 2007]. This mid-mountainous region (altitude of the study site 606 m) is characterized by a temperate climate with mean annual precipitation of 863 mm, potential evapotranspiration of 433 mm and a mean annual temperature of 7°C (climate station Höllenstein-Kraftwerk, 1947–2005). Land cover around the station is dominated by grassland and forest, with some interlaced agricultural fields. The geology consists of gneiss as basement, followed by a fractured zone of the same rock, a saprolite layer of weathered gneiss and the mineral soil on top [Creutzfeldt et al., 2010a]. The soil type is classified as loamy-sandy brown Cambisol [Creutzfeldt et al., 2010b]. The hydrogeology is characterized by an unconfined, fractured aquifer [Bayerisches Landesamt für Umwelt, 2014]. Most of the groundwater is stored in the expansion joints of the fractured gneiss with low to very low hydraulic conductivities (mean of 9.5 * 10-6 m/s) [Büttner et al., 2003; Bayerisches Geologisches Landesamt, 1996]. The groundwater level varies significantly, both in time and space: groundwater tables within a radius of 350 m range between 3 to 14 m below the surface during the time period from 2010 to 2016.

A.1   Gravity data

The dual-sphere superconducting gravimeter SG30 was installed in the gravimeter building in Wettzell in June 2010. The gravity data used in this study is that of the lower sphere of the gravimeter. The gravimeter building has dimensions of 11 m x 8 m. The foundation of the building is made of a horizontal concrete baseplate with a thickness of 0.25 m and base walls 1.5 m deep with a thickness of 0.5 m. The SG30 is mounted on a pillar of 1.1 x 1.1 m in size and with a total depth of 2.0 m. A cross-section of the building is shown in Figure 2.1. The scale factor of SG30 was determined via comparison to measurements with an FG5 absolute gravimeter and an atom interferometer [Hauth et al., 2014]. Instrumental drift and steps were estimated using numerous FG5 absolute gravimeter measurements. Prior to the analysis, the gravity time series were decimated to one hour and corrected for all known processes besides local hydrology, i.e., Earth and ocean tides, polar motion,

atmosphere and global hydrology. The atmospheric correction was carried out using Atmacs, a 3D physically-based approach [Klügel and Wziontek, 2009]. The global hydrological effect was computed excluding the local zone with a radius of 5.5 km [Mikolaj et al., 2016], using simulated water storage variations of GLDAS/NOAH025 model [Rodell et al., 2004]. The derived residuals are from here on referred to as "gravity signal".

A.2   Hydrological data

Soil moisture was measured with Time Domain Reflectometry (TDR). The sensors were installed in 4 vertical profiles, 1 outside and 3 beneath the gravimeter building, up to a depth of 2 m. Their locations are shown in Figure 2.1. With respect to their vertical position in the profiles, the sensors are categorized as shallow, middle and deep (Figure 2.1). The sensor profiles are named from A to D, starting with profile A on grassland with a distance of 1.5 m to the building, profile B located directly inside the building footprint close to the wall, and profiles C and D 1.1 m respectively 2.1 m towards the center of the building (see Figure 2.1). Data was available in 15 min resolution since March 2010.

   Hourly data from well BK14 was used to determine groundwater fluctuations (located right next to the gravimeter building, see Figure 2.1). Data gaps in this time series were filled with data taken from well BK3 (operated by BKG, a little further than BK14 but also next to gravimeter building; see Figure 2.1), measuring at a 30 minute interval, which shows the same dynamics as BK14.

A.3   Meteorological data

A meteorological station at the site measured air temperature, wind speed, air humidity, global radiation and net radiation with a temporal resolution of 15 minutes. Precipitation was measured with a heated tipping bucket rain gauge with a temporal resolution of 30 minutes. Short data gaps in the time series were filled with data of a similar station, located nearby on the observatory site. A lysimeter (depth 1.5 m, surface area 1 m²) with undisturbed soil and grass cover is used to estimate actual evapotranspiration at 15 min intervals, making use of a simple water balance calculation [von Unold and Fank, 2007]. A large data gap of 8 weeks in 2011 and 10 months in 2012 was filled with reference evapotranspiration, calculated with the FAO (Food and Agriculture Organisation) Penman-Monteith equation [Allen et al., 1998], using global radiation data from the climate station Allmannsdorf located at a distance of 5 km (no.127, operated by Bavarian State Research Center for Agriculture).

A.4   Soil properties

Major soil layers at the site were distinguished based on data of the drill log of groundwater well BK14: organic top soil layer (Ah horizon), sandy-loamy (Bv) and silty-loamy (Cv) weathering horizons, followed by gneiss (partly fissured, fractured and weathered) and a basement of granite (starting at 12.8 m). Undisturbed soil

samples were analyzed for soil-hydraulic (van Genuchten) parameters (Table 2.1). As there were no samples taken from the fractured gneiss zone, its van Genuchten parameters are based on literature values [Lisboa et al., 2006] with a modification of the saturated moisture content following estimated values of specific yield by Creutzfeldt et al. [2008].

Appendix B

Evaluation of model results with example data from test site

The soil moisture time series simulated with HYDRUS with the parameters for the test site Wettzell were compared to observed sensor data (Figure B.1). All time series were normalized by subtracting their mean. Seasonal dynamics including peaks were similar with respect of timing and relative differences. Dynamics and amplitudes associated with major wet and dry spells were reasonably represented by the model. For shallow soil moisture beneath the gravimeter building (shallow sensors C and D) very small variations in the simulated time series corresponded to hardly any precipitation-driven variations seen in the observations. At profiles A and B, the shallow sensors showed quite high variations both in observed and simulated values.

In terms of performance measures, R^2 ranges from 0.09 to 0.59, root mean square error (RMSE) from 1 to 3 VWC in Vol% and the Nash-Sutcliffe-Efficiency from -1.77 to 0.56. The relatively low values of the performance measures can also be explained by the likely occurrence of preferential flow which is not implemented in the model.
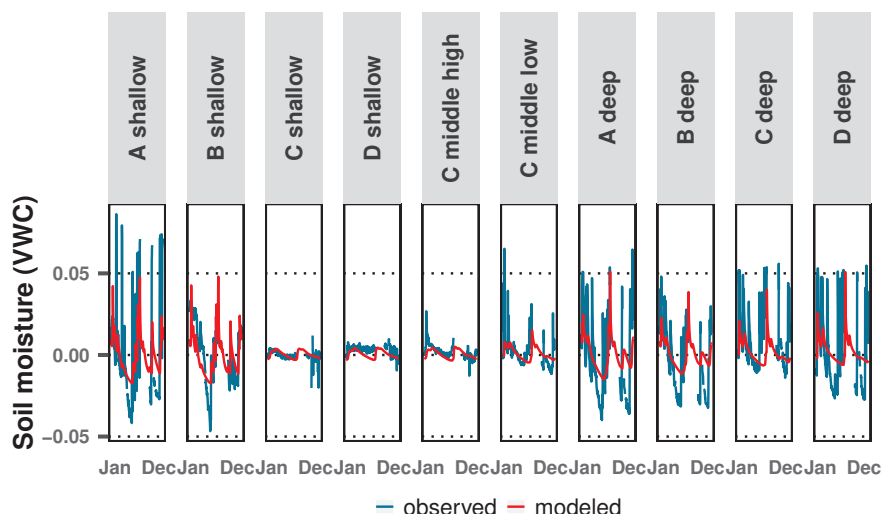
Figure B.1: Observed (blue) and simulated (red) soil moisture anomalies (deviations from mean soil moisture), for the different sensor positions (see Figure 2.2) and for the year 2011. Simulated values are the result from the HYDRUS model scenario "Wettzell".

We used the output of the hydrological model for scenario "Wettzell" to compute the corresponding gravity signal. This conversion was realized via forward modeling with a nested approach [Leirião et al., 2009]. 4 gravity component grids were set up, based on a local digital elevation model (DEM). Depending on the distance to the gravimeter, each grid has a distinct spatial discretization and method to calculate the gravity effect (Table B.1).

To calculate the gravity signal of a cell for a certain time period, the corresponding gravity effect of the cell was multiplied by its soil moisture content. For the hydrological model domain, soil moisture varies in all 3 dimensions. For cells outside of this domain, soil moisture was assumed to be uniform in the lateral dimension and to vary in the vertical only, following the model output of the area outside of the gravimeter building. The full gravity signal is the sum of all individual cell values of the 4 grids. The space of the gravimeter pillar was assumed to have no mass variations (Δ g = 0). Gravity effects were calculated for different positions of the gravimeter within the building (Figure 2.3).

Table B.1: Nested grid approach for gravity forward modeling [Leirião et al., 2009]: Spatial discretization and corresponding methods used to calculate gravity. Sizes of the grid shapes for each level (black shapes in the first column) are not true to scale.

Grid shape	Grid extent from center [m]	Lateral discretization [m]	Vertical discretization [m]	Gravity method
	Hydrus model domain: 18x21	0.1	0.1	Rectangular prisms (Forsberg 1984)
	Model domain - 50	0.1	0.1	Rectangular prisms (Forsberg 1984)
	50 - 300	1	0.1	Poin-mass approximation: MacMillan Formula(MacMillan 1958)
	300 - 5560	10	1	Point-mass (Telford 1990)

The comparison with the observed gravity signal of SG30 for the example of year 2011 is shown in Figure B.2. The overall seasonal dynamics as well as short-term dynamics related to larger rain events are well represented. The RMSE for hourly time series equals 7.6 nm/s^2, 40% of the variance of the observed gravity signal can be explained by the model. It is important to note that the gravity signal contains residual non-local contributions related to insufficient reduction of atmospheric as well as tidal and non-tidal ocean loadings. The high frequency part of the gravity data, however, is missing in the simulated time series. This is due to the inadequate representation of fast flow processes especially in the upper layers of the model. This aspect as well as some periods of over- and underestimation of the observed gravity variations, might be caused by inaccurate model forcing time series and deficient soil parameterization (e.g. for hydrological conductivity), which controls absolute soil moisture values and water flow processes.

Nevertheless, except for the high-frequency variations, the overall modeled gravity response is considered to be reasonable enough to qualify the underlying model for the further assessment of the umbrella effect on gravity time series, using the different model scenarios defined in the main text.
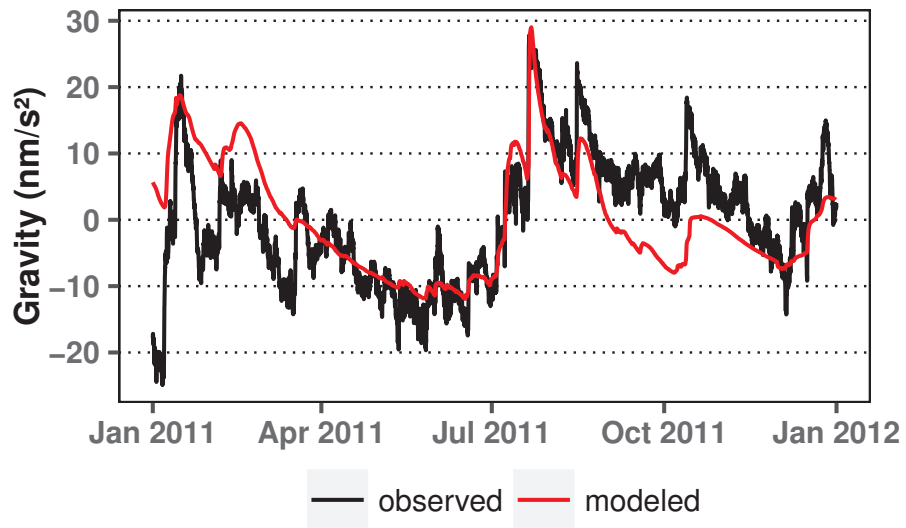
Figure B.2: Gravity signal of SG30 and the modeled gravity response based on the hydrological scenario "Wettzell".

Appendix C

Influence of the vertical extent of the model domain on the reduction ratios and their application for the reduction routine

For simplicity of presentation, all figures and results in the main text were based on the complete model domain with a vertical extent of 5 m. However, the mean soil moisture content, the modeled gravity response, and consequentially the estimated reduction ratios are a function of the depth, resp. the volume, of the model domain. In this section we show the results for varying soil depths and demonstrate that reduction ratios based on the maximum extent of our model domain (5 m) capture the overall system dynamics well enough to be a valid example for describing the consequences of the umbrella effect and for deriving the proposed reduction approach. Nevertheless, when applying our approach to real gravity data, it is highly recommended to use the specific parameters for the real soil depth at the observation site (Table C.1 to C.6).

In the first step of the reduction approach, gravity signals were modeled for both the real (WSC from below the building) and hypothetical (WSC taken from outside of building) case. The results are exemplary shown in the following for the hydrological model scenario "Soil type sandy loam" with the gravimeter located in the center of the building. For this setup and its corresponding vertical soil moisture distribution, gravity responses for both cases (below and outside building) increase in absolute values with an increasing vertical extent of the model domain (Figure C.1), as do the relative differences between gravity values of both cases.
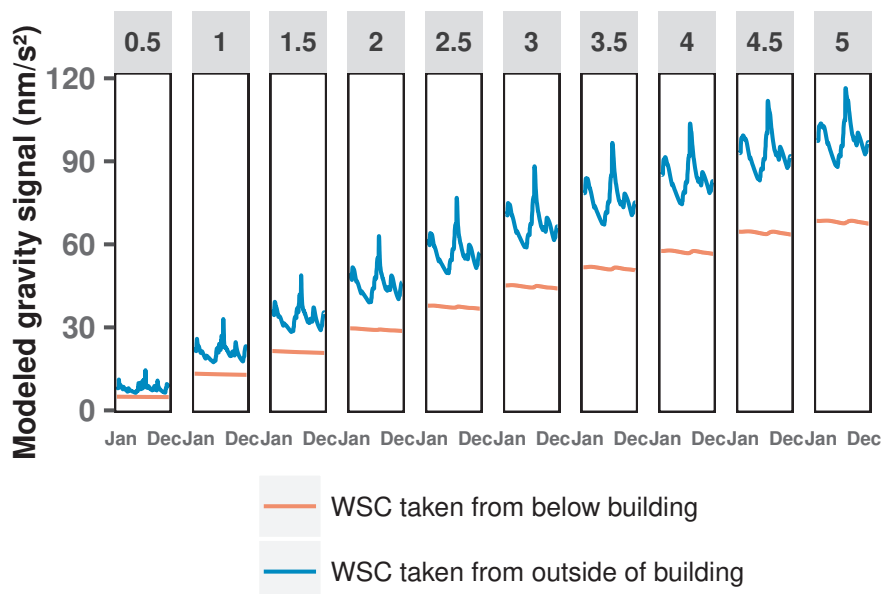
Figure C.1: Modeled gravity response for the hydrological scenario "Soil type sandy loam" throughout the second year of the simulation period. For a fixed position of the gravimeter (center) and different vertical extents of the model domain, the gravity response of WSC in the space below the building is shown, assuming the real WSC (red) and the hypothetical WSC taken from outside of the building (blue).

For the linear regression between the time-variable mean soil moisture content of the model domain and the dynamic reduction ratio, a close to linear trend with vertical extent can be observed in the regression parameters. Independent of the hydrological scenario (with an exception of the two driest climate scenarios), parameter intercept increases with depth of the model domain whereas parameter slope decreases (Figure C.2). The parameter range within specific model scenarios depends on the contrasts between soil moisture below and next the gravimeter building (compare Figure 2.7): larger contrasts result in a bigger range (e.g. anisotropy scenarios). On the one hand, as these regression parameters vary with the vertical model extent, the use of specific parameters for the actual soil depth at the observation site is recommended. On the other hand, the comparison of the hypothetical gravity signal after transformation to the corresponding signal below the building with (i)the reduction ratios based on the 5 m model extent and (ii) with depth-specific reduction ratios, shows minor differences in the resulting signals (Figure C.3).
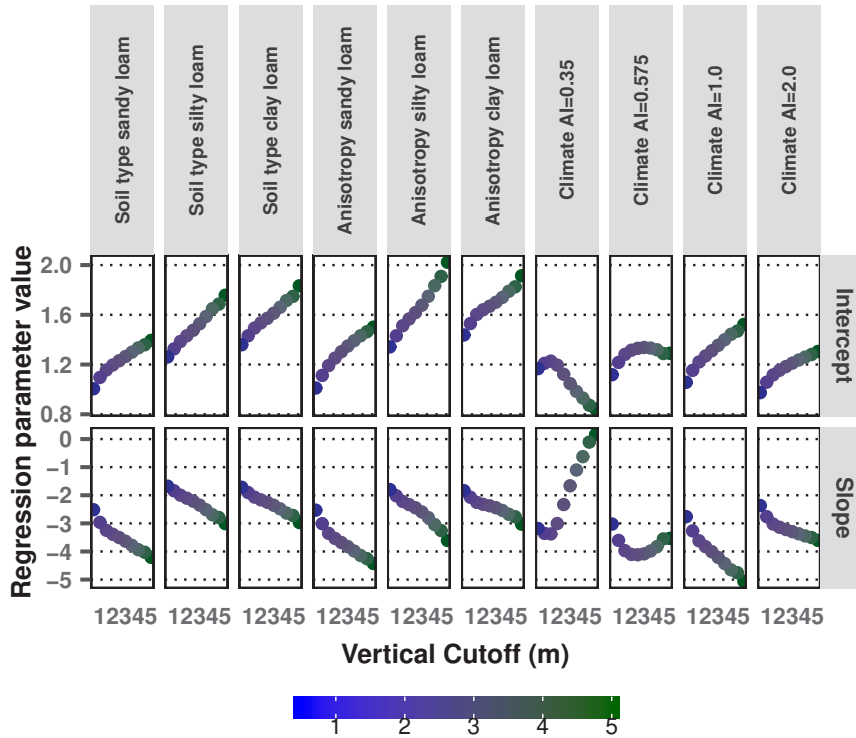
Figure C.2: Regression parameters intercept and slope, used in Eq.2.1 to calculate the dynamic reduction ratio for each hydrological scenario, here at a fixed gravimeter location (center, see Figure 2.3) and for different vertical extents of the model domain.

The differences range from 0.73 nm/s^2 (vertical extent of 4.5 m) to 3.4 nm/s^2 (vertical extent of 2 m). These small differences underline the robustness of our approach, even if site specific depth is neglected, for quantifying the influences of the umbrella effect.

For the reduction as a function of the size of the gravimeter building, the reduction ratios also vary with the chosen vertical model extent.

Except for the smallest building footprint, reduction ratios tend towards a value of 1 with decreasing depth of the model domain (Figure C.4). For the small building footprint, reduction ratios increase with decreasing depths of the model domain, ranging from 1.35 (5 m extent) to 1.64 (0.5 m extent), as do their error ranges. For a deceasing vertical model extent, the importance of the building footprint decreases. Parameters for the fitted exponential function for each vertical model extent are given in Table C.6.
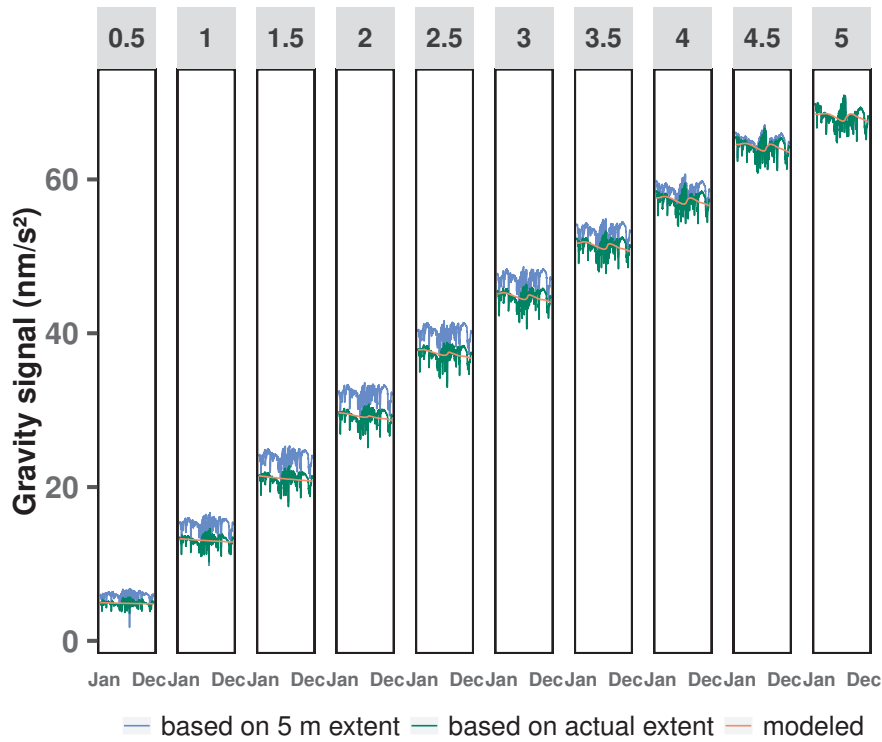
Figure C.3: Modeled gravity responses below the building for the hydrological scenario "Soil type sandy loam" throughout the second year of the simulation period. For a fixed position of the gravimeter (center) and different vertical extents of the model domain (vertical panels, labels indicate extent in m), the modeled gravity response of WSC in the space below the building is shown (red), next to a calculation carried out with parameter sets of the 5 m vertical extent (green) and with parameter sets corresponding to each actual extent (blue).
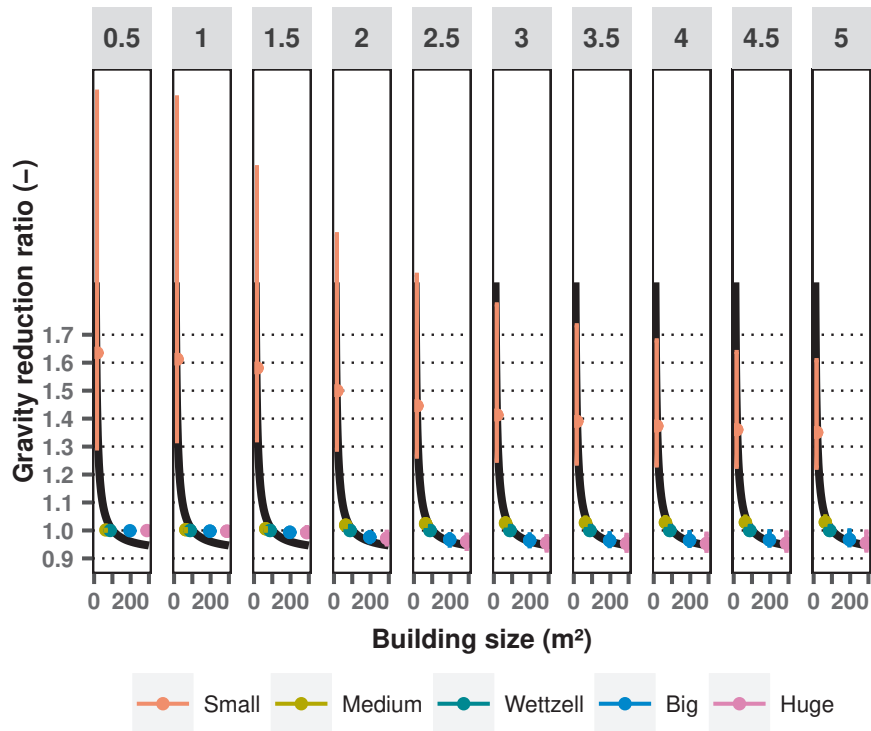
Figure C.4: Building size in relation to its mean gravity correction factor, normalized by the test site Wettzell building size scenario, shown for 5 different hydrological model building dimension scenarios and different vertical extents of the considered model domain (vertical panels, labels indicate extent in m). The black line represents the exponential function which was fitted to all data points. For each scenario, the range of its dynamic reduction ratios is shown in colored error bars. Upper end of error bar corresponds to the driest soil moisture state, the lower end to wettest state.

Table C.1: Linear regression parameters for calculating the reduction ratios corresponding to hydrological scenarios and the vertical extent of the model domain (see Eq.2.1) for the gravimeter at the center position.

SG position: Center

Vertical model extent [m]

Regression parameter	0.50		1		1.5		2		2.5		3		3.5		4		4.5		5	
	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
Anisotropy clay loam	1.44	-1.84	1.53	-2.07	1.60	-2.25	1.64	-2.32	1.67	-2.37	1.70	-2.44	1.74	-2.54	1.79	-2.67	1.82	-2.77	1.91	-3.03
Anisotropy sandy loam	1.01	-2.53	1.11	-3.01	1.19	-3.35	1.25	-3.54	1.30	-3.68	1.35	-3.83	1.39	-3.98	1.43	-4.14	1.46	-4.25	1.50	-4.43
Anisotropy silty loam	1.34	-1.80	1.43	-2.02	1.51	-2.22	1.57	-2.33	1.62	-2.44	1.68	-2.59	1.75	-2.79	1.83	-3.04	1.91	-3.26	2.02	-3.61
Climate AI=0.35	1.17	-3.18	1.21	-3.35	1.23	-3.38	1.20	-3.01	1.12	-2.33	1.05	-1.66	0.98	-1.10	0.93	-0.63	0.87	-0.11	0.84	0.16
Climate AI=0.575	1.12	-3.03	1.22	-3.61	1.28	-3.97	1.32	-4.10	1.33	-4.12	1.34	-4.08	1.33	-3.98	1.32	-3.83	1.29	-3.55	1.29	-3.53
Climate AI=1.0	1.06	-2.76	1.15	-3.27	1.22	-3.62	1.27	-3.83	1.31	-4.01	1.35	-4.21	1.40	-4.43	1.44	-4.65	1.47	-4.76	1.52	-5.06
Climate AI=2.0	0.97	-2.37	1.06	-2.76	1.12	-3.02	1.15	-3.14	1.18	-3.22	1.21	-3.30	1.24	-3.37	1.26	-3.45	1.28	-3.50	1.31	-3.60
Soil type clay loam	1.36	-1.71	1.43	-1.90	1.49	-2.07	1.54	-2.16	1.57	-2.25	1.61	-2.35	1.66	-2.48	1.71	-2.63	1.75	-2.72	1.83	-2.97
Soil type sandy loam	1.01	-2.52	1.10	-2.96	1.16	-3.25	1.20	-3.39	1.23	-3.50	1.26	-3.63	1.30	-3.78	1.33	-3.94	1.36	-4.03	1.39	-4.21
Soil type silty loam	1.26	-1.68	1.33	-1.84	1.39	-2.00	1.43	-2.10	1.48	-2.21	1.53	-2.35	1.59	-2.51	1.65	-2.69	1.69	-2.79	1.76	-3.01

Table C.2: Linear regression parameters for calculating the reduction ratios corresponding to hydrological scenarios and the vertical extent of the model domain (see Eq.2.2) for the gravimeter at the center position.

SG position: Corner

Vertical model extent [m]

Regression parameter	0.50		1		1.5		2		2.5		3		3.5		4		4.5		5	
	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
Anisotropy clay loam	1.44	-1.84	1.53	-2.05	1.58	-2.18	1.56	-2.03	1.54	-1.92	1.54	-1.89	1.54	-1.89	1.56	-1.91	1.55	-1.88	1.60	-2.03
Anisotropy sandy loam	1.01	-2.53	1.12	-3.02	1.20	-3.35	1.25	-3.27	1.28	-3.23	1.32	-3.28	1.35	-3.36	1.37	-3.45	1.40	-3.51	1.43	-3.65
Anisotropy silty loam	1.35	-1.80	1.43	-2.01	1.50	-2.16	1.49	-2.04	1.49	-1.97	1.51	-1.98	1.54	-2.05	1.59	-2.19	1.64	-2.30	1.73	-2.61
Climate AI=0.35	1.17	-3.18	1.20	-3.29	1.21	-3.25	1.15	-2.58	1.06	-1.71	0.97	-0.86	0.88	-0.08	0.79	0.63	0.71	1.34	0.66	1.75
Climate AI=0.575	1.12	-3.03	1.22	-3.58	1.28	-3.94	1.30	-3.87	1.31	-3.75	1.30	-3.60	1.28	-3.38	1.26	-3.13	1.21	-2.75	1.20	-2.63
Climate AI=1.0	1.06	-2.76	1.16	-3.26	1.23	-3.64	1.28	-3.72	1.33	-3.85	1.37	-4.02	1.41	-4.18	1.44	-4.32	1.44	-4.28	1.48	-4.47
Climate AI=2.0	0.98	-2.38	1.07	-2.77	1.13	-3.05	1.17	-3.01	1.21	-3.01	1.23	-3.06	1.26	-3.13	1.28	-3.20	1.30	-3.24	1.32	-3.32
Soil type clay loam	1.38	-1.75	1.46	-1.95	1.52	-2.11	1.53	-2.09	1.55	-2.09	1.57	-2.14	1.60	-2.20	1.62	-2.26	1.62	-2.23	1.66	-2.36
Soil type sandy loam	1.01	-2.52	1.10	-2.96	1.17	-3.27	1.21	-3.28	1.26	-3.36	1.30	-3.48	1.33	-3.60	1.36	-3.71	1.37	-3.73	1.39	-3.86
Soil type silty loam	1.28	-1.70	1.35	-1.88	1.42	-2.05	1.44	-2.04	1.46	-2.05	1.49	-2.10	1.51	-2.15	1.53	-2.19	1.51	-2.11	1.54	-2.19

Table C.3: Linear regression parameters for calculating the reduction ratios corresponding to hydrological scenarios and the vertical extent of the model domain (see Eq.2.2) for the gravimeter at the center position.

SG position: Off-center

Vertical model extent [m]	0.50		1		1.5		2		2.5		2.5		3.5		4		4.5		5	
Regression parameter	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
Anisotropy clay loam	1.44	-1.84	1.53	-2.07	1.60	-2.25	1.63	-2.30	1.66	-2.34	1.69	-2.40	1.72	-2.49	1.77	-2.61	1.81	-2.70	1.89	-2.96
Anisotropy sandy loam	1.01	-2.53	1.11	-3.02	1.19	-3.36	1.25	-3.51	1.30	-3.64	1.35	-3.77	1.39	-3.91	1.42	-4.06	1.45	-4.16	1.49	-4.32
Anisotropy silty loam	1.34	-1.80	1.43	-2.02	1.51	-2.22	1.56	-2.31	1.61	-2.41	1.67	-2.55	1.74	-2.74	1.82	-2.99	1.89	-3.20	2.01	-3.55
Climate AI=0.35	1.17	-3.18	1.21	-3.36	1.23	-3.39	1.20	-3.02	1.13	-2.38	1.06	-1.73	1.00	-1.18	0.94	-0.71	0.89	-0.20	0.86	0.07
Climate AI=0.575	1.12	-3.02	1.22	-3.61	1.28	-3.98	1.32	-4.11	1.34	-4.14	1.35	-4.13	1.34	-4.04	1.33	-3.90	1.30	-3.64	1.31	-3.62
Climate AI=1.0	1.06	-2.76	1.15	-3.27	1.22	-3.63	1.27	-3.83	1.32	-4.02	1.36	-4.23	1.41	-4.45	1.45	-4.67	1.48	-4.77	1.53	-5.07
Climate AI=2.0	0.97	-2.37	1.06	-2.77	1.12	-3.03	1.16	-3.14	1.19	-3.21	1.22	-3.29	1.24	-3.36	1.27	-3.43	1.29	-3.49	1.31	-3.58
Soil type clay loam	1.36	-1.72	1.44	-1.91	1.50	-2.08	1.54	-2.17	1.58	-2.26	1.62	-2.37	1.67	-2.49	1.72	-2.64	1.75	-2.73	1.84	-2.97
Soil type sandy loam	1.01	-2.51	1.10	-2.97	1.16	-3.25	1.20	-3.39	1.24	-3.51	1.27	-3.64	1.31	-3.78	1.34	-3.93	1.36	-4.02	1.40	-4.19
Soil type silty loam	1.26	-1.68	1.33	-1.85	1.39	-2.01	1.44	-2.11	1.48	-2.23	1.54	-2.36	1.59	-2.53	1.66	-2.70	1.69	-2.79	1.76	-3.00

Table C.4: Linear regression parameters for calculating the reduction ratios corresponding to hydrological scenarios and the vertical extent of the model domain (see Eq.2.2) for the gravimeter at the center position.

SG position: Edge

Vertical model extent [m]	0.50		1		1.5		2		2.5		2.5		3.5		4		4.5		5	
Regression parameter	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope	Intercept	Slope
Anisotropy clay loam	1.44	-1.84	1.53	-2.06	1.59	-2.21	1.59	-2.16	1.59	-2.12	1.61	-2.12	1.63	-2.16	1.66	-2.24	1.67	-2.27	1.74	-2.47
Anisotropy sandy loam	1.01	-2.53	1.12	-3.03	1.20	-3.37	1.25	-3.40	1.30	-3.43	1.33	-3.51	1.36	-3.61	1.40	-3.72	1.42	-3.78	1.45	-3.92
Anisotropy silty loam	1.35	-1.80	1.43	-2.02	1.51	-2.19	1.53	-2.18	1.55	-2.18	1.59	-2.25	1.64	-2.39	1.70	-2.57	1.76	-2.74	1.87	-3.07
Climate AI=0.35	1.17	-3.18	1.21	-3.33	1.22	-3.34	1.18	-2.86	1.11	-2.12	1.02	-1.40	0.95	-0.75	0.88	-0.16	0.81	0.45	0.77	0.79
Climate AI=0.575	1.12	-3.03	1.22	-3.60	1.29	-3.98	1.32	-4.03	1.33	-4.01	1.34	-3.94	1.33	-3.81	1.31	-3.63	1.27	-3.32	1.27	-3.26
Climate AI=1.0	1.06	-2.76	1.16	-3.27	1.23	-3.64	1.28	-3.80	1.33	-3.97	1.38	-4.16	1.42	-4.36	1.46	-4.54	1.47	-4.58	1.52	-4.83
Climate AI=2.0	0.98	-2.37	1.07	-2.78	1.13	-3.05	1.17	-3.09	1.20	-3.13	1.23	-3.19	1.26	-3.26	1.28	-3.33	1.30	-3.37	1.32	-3.46
Soil type clay loam	1.38	-1.74	1.45	-1.94	1.52	-2.12	1.55	-2.16	1.58	-2.21	1.61	-2.29	1.65	-2.38	1.69	-2.50	1.70	-2.52	1.77	-2.71
Soil type sandy loam	1.01	-2.52	1.10	-2.97	1.17	-3.27	1.21	-3.35	1.25	-3.46	1.29	-3.59	1.32	-3.72	1.35	-3.85	1.37	-3.90	1.40	-4.04
Soil type silty loam	1.28	-1.70	1.35	-1.87	1.41	-2.05	1.45	-2.11	1.49	-2.18	1.53	-2.27	1.57	-2.38	1.61	-2.49	1.62	-2.49	1.66	-2.63

Table C.5: Linear regression parameters for calculating the reduction ratios corresponding to hydrological scenarios and the vertical extent of the model domain (see Eq.2.2) for the gravimeter at the center position.

SG position: Wettzell

Regression parameter	0.50 Intercept	0.50 Slope	1 Intercept	1 Slope	1.5 Intercept	1.5 Slope	2 Intercept	2 Slope	2.5 Intercept	2.5 Slope	3 Intercept	3 Slope	3.5 Intercept	3.5 Slope	4 Intercept	4 Slope	4.5 Intercept	4.5 Slope	5 Intercept	5 Slope
Anisotropy clay loam	1.44	-1.84	1.53	-2.08	1.60	-2.24	1.61	-2.22	1.62	-2.21	1.64	-2.24	1.67	-2.31	1.71	-2.41	1.74	-2.47	1.81	-2.70
Anisotropy sandy loam	1.01	-2.53	1.12	-3.04	1.20	-3.39	1.25	-3.44	1.30	-3.50	1.33	-3.58	1.37	-3.69	1.40	-3.80	1.42	-3.88	1.45	-4.01
Anisotropy silty loam	1.35	-1.80	1.44	-2.03	1.52	-2.22	1.55	-2.24	1.58	-2.29	1.62	-2.40	1.69	-2.56	1.76	-2.78	1.83	-2.98	1.94	-3.33
Climate AI=0.35	1.17	-3.18	1.22	-3.40	1.24	-3.49	1.22	-3.13	1.15	-2.51	1.08	-1.89	1.02	-1.35	0.97	-0.88	0.91	-0.36	0.88	-0.09
Climate AI=0.575	1.12	-3.03	1.22	-3.63	1.29	-4.05	1.33	-4.16	1.35	-4.19	1.37	-4.19	1.37	-4.11	1.35	-3.98	1.33	-3.72	1.33	-3.71
Climate AI=1.0	1.06	-2.77	1.16	-3.29	1.23	-3.67	1.28	-3.84	1.33	-4.02	1.38	-4.22	1.42	-4.43	1.46	-4.64	1.49	-4.72	1.54	-5.00
Climate AI=2.0	0.98	-2.37	1.07	-2.79	1.13	-3.07	1.17	-3.12	1.20	-3.17	1.23	-3.23	1.26	-3.30	1.28	-3.36	1.29	-3.41	1.32	-3.49
Soil type clay loam	1.37	-1.74	1.46	-1.95	1.52	-2.13	1.56	-2.19	1.59	-2.26	1.63	-2.36	1.67	-2.47	1.72	-2.61	1.75	-2.67	1.82	-2.90
Soil type sandy loam	1.01	-2.52	1.10	-2.99	1.17	-3.29	1.21	-3.38	1.25	-3.49	1.29	-3.62	1.32	-3.75	1.35	-3.89	1.37	-3.95	1.40	-4.09
Soil type silty loam	1.27	-1.69	1.35	-1.88	1.42	-2.07	1.46	-2.14	1.50	-2.24	1.55	-2.37	1.60	-2.51	1.65	-2.65	1.68	-2.70	1.73	-2.87

Table C.6: Parameters of the exponential function, fitted to mean gravity reduction ratios of different building sizes. For each vertical extent of the model domain, one set of parameters is obtained (see Eq.2.3).

Vertical model extent [m]	a	b
0.5	-10.10	11.01
1	-9.73	10.64
1.5	-9.23	10.14
2	-8.03	8.94
2.5	-7.21	8.13
3	-6.68	7.60
3.5	-6.30	7.22
4	-6.00	6.93
4.5	-5.76	6.69
5	-5.57	6.50

References

R. G. Allen, L. S. Pereira, D. Raes, M. Smith, et al. 1998. *Crop evapotranspiration - Guidelines for computing crop water requirements : FAO Irrigation and drainage paper 56.* FAO, Rome.

L. Angermann, C. Jackisch, N. Allroggen, M. Sprenger, E. Zehe, J. Tronicke, M. Weiler, and T. Blume. 2017. Form and function in hillslope hydrology: characterization of subsurface flow based on response observations. *Hydrology and Earth System Sciences*, 21(7):3727–3748. doi: 10.5194/hess-21-3727-2017.

J. Arnoso, M. Benavent, M. Bos, F. Montesinos, and R. Vieira. 2011. Verifying the body tide at the Canary Islands using tidal gravimetry observations. *Journal of Geodynamics*, 51(5):358 – 365. doi: https://doi.org/10.1016/j.jog.2010.10.004.

D. D. Baldocchi, B. B. Hincks, and T. P. Meyers. 1988. Measuring Biosphere-Atmosphere Exchanges of Biologically Related Gases with Micrometeorological Methods. *Ecology*, 69:1331–1340. doi: 10.2307/1941631.

D. Banka and D. Crossley. 10 1999. Noise levels of superconducting gravimeters at seismic frequencies. *Geophysical Journal International*, 139(1):87–97. doi: 10.1046/j.1365-246X.1999.00913.x.

C. J. Barrow. 1992. *World atlas of desertification (United nations environment programme)*, volume 3. John Wiley & Sons, Ltd.

Bayerisches Geologisches Landesamt. 1996. Grundlagen zum wasserwirtschaftlichen Rahmenplan NAAB-REGEN, Hydrogeologie. Technical report, Bayerisches Geologisches Landesamt.

Bayerisches Landesamt für Umwelt. 2014. Hydrogeologische Karte 1:100000, Blatt 1: Klassifikation der Hydrogeologische Einheiten. In *Geowissenschaftliche Landesaufnahme in der Planungsregion 11 Regensburg*, Bürgermeister-Ulrich-Strasse 160, 86179 Augsburg, Germany. Bayerisches Landesamt für Umwelt.

Bayerisches Landesamt für Umwelt, last access: 27 June 2017. URL http://www.hnd.bayern.de/pegel/donau_bis_passau/chamerau-15202300/statistik?days=1.

K. Beven. 2002. Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrological Processes*, 16(2): 189–206. doi: https://doi.org/10.1002/hyp.343.

K. Beven. 2006. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320 (1-2):18–36. doi: 10.1016/j.jhydrol.2005.07.007.

K. Beven and P. Germann. 2013. Macropores and water flow in soils revisited. *Water Resources Research*, 49:3071–3092. doi: 10.1002/wrcr.20156.

H. R. Bogena, J. A. Huisman, A. Güntner, C. Hübner, J. Kusche, F. Jonard, S. Vey, and H. Vereecken. 2015. Emerging methods for noninvasive sensing of soil moisture dynamics from field to catchment scale: a review. *WIREs Water*, 2(6):635–647. doi: https://doi.org/10.1002/wat2.1097.

M. Bonatz. 1967. Der Gravitationseinfluß der Bodenfeuchtigkeit. *Zeitschrift fü Vermessungswesen*, 92:135–139.

J. P. Boy and J. Hinderer. 2006. Study of the seasonal gravity signal in superconducting gravimeter data. *Journal of Geodynamics*, 41(1-3):227–233. doi: 10.1016/j.jog.2005.08.035.

J.-P. Boy, L. Longuevergne, F. Boudin, T. Jacob, F. Lyard, M. Llubes, N. Florsch, and M.-F. Esnoult. 2009. Modelling atmospheric and induced non-tidal oceanic loading contributions to surface gravity and tilt measurements. *Journal of Geodynamics*, 48(3–5):182–188. doi: http://dx.doi.org/10.1016/j.jog.2009.09.022.

J.-P. Boy, S. Rosat, J. Hinderer, and F. Littel, 2017. Superconducting gravimeter data from Djougou - Level 1. doi: http://doi.org/10.5880/igets.dj.l1.001.

G. Büttner, R. Pamer, and B. Wagner. 2003. Hydrogeologische Raumgliederung von Bayern. Technical Report 20, Bayerisches Geologisches Landesamt.

D. Carbone, M. P. Poland, M. Diament, and F. Greco. 2017. The added value of time-variable microgravimetry to the understanding of how volcanoes work. *Earth-Science Reviews*, 169:146–179. doi: https://doi.org/10.1016/j.earscirev.2017.04.014.

D. Carbone, F. Cannavò, F. Greco, R. Reineman, and R. J. Warburton. 2019. The Benefits of Using a Network of Superconducting Gravimeters to Monitor and Study Active Volcanoes. *Journal of Geophysical Research: Solid Earth*, 124(4): 4035–4050. doi: https://doi.org/10.1029/2018JB017204.

S. D. Carrière, B. Loiseau, C. Champollion, C. Ollivier, N. K. Martin-StPaul, N. Lesparre, A. Olioso, J. Hinderer, and D. Jougnot. 2021. First Evidence of Correlation Between Evapotranspiration and Gravity at a Daily Time Scale From Two Vertically Spaced Superconducting Gravimeters. *Geophysical Research Letters*, 48(24). doi: 10.1029/2021gl096579.

C. Champollion, S. Deville, J. Chéry, E. Doerflinger, N. Le Moigne, R. Bayer, P. Vernant, and N. Mazzilli. 2018. Estimating epikarst water storage by time-lapse surface-to-depth gravity measurements. *Hydrology and Earth System Sciences*, 22 (7):3825–3839. doi: 10.5194/hess-22-3825-2018.

D. Chapin. 1998. Gravity instruments: Past, present, future. *The Leading Edge*, 17 (1):100–100. doi: 10.1190/1.1437806.

M. S. Chauhan, F. Cannavó, D. Carbone, and F. Greco. 2020. Insights into mount etna december 2018 eruption from joint inversion of deformation and gravity data.

Geophysical Research Letters, 47(16):e2020GL087786. doi: https://doi.org/10.1029/2020GL087786.

H. Chen, L. Fan, W. Wu, and H.-B. Liu. 2017. Comparison of spatial interpolation methods for soil moisture and its application for monitoring drought. *Environmental Monitoring and Assessment*, 189(10):525. doi: 10.1007/s10661-017-6244-4.

L. Christiansen, E. B. Haarder, A. B. Hansen, M. C. Looms, P. J. Binning, D. Rosbjerg, O. B. Andersen, and P. Bauer-Gottwein. 2011a. Calibrating vadose zone models with time-lapse gravity data. *Vadose Zone Journal*, 10(3):1034–1044. doi: 10.2136/vzj2010.0127.

L. Christiansen, S. Lund, O. Andersen, P. Binning, D. Rosbjerg, and P. Bauer-Gottwein. 2011b. Measuring gravity change caused by water storage variations: Performance assessment under controlled conditions. *Journal of Hydrology*, 402 (1-2):60–70. doi: 10.1016/j.jhydrol.2011.03.004.

B. Creutzfeldt, A. Güntner, T. Klügel, and H. Wziontek. 2008. Simulating the influence of water storage changes on the superconducting gravimeter of the Geodetic Observatory Wettzell, Germany. *Geophysics*, 73(6):WA95–WA104. doi: 10.1190/1.2992508.

B. Creutzfeldt, A. Güntner, H. Thoss, B. Merz, and H. Wziontek. 2010a. Measuring the effect of local water storage changes on in situ gravity observations: Case study of the Geodetic Observatory Wettzell, Germany. *Water Resources Research*, 46 (8). doi: 10.1029/2009wr008359.

B. Creutzfeldt, A. Güntner, S. Vorogushyn, and B. Merz. 2010b. The benefits of gravimeter observations for modelling water storage changes at the field scale. *Hydrology and Earth System Sciences*, 14(9):1715–1730. doi: 10.5194/hess-14-1715-2010.

B. Creutzfeldt, A. Güntner, H. Wziontek, and B. Merz. 2010c. Reducing local hydrology from high-precision gravity measurements: A lysimeter-based approach. *Geophysical Journal International*, 183(1):178–187.

B. Creutzfeldt, T. Ferré, P. Troch, B. Merz, H. Wziontek, and A. Güntner. 2012. Total water storage dynamics in response to climate variability and extremes: Inference from long-term terrestrial gravity measurement. *Journal of Geophysical Research: Atmospheres*, 117(D8). doi: https://doi.org/10.1029/2011JD016472.

B. Creutzfeldt, P. Troch, A. Güntner, T. Ferré, T. Graeff, and B. Merz. 2013. Storage-discharge relationships at different catchment scales based on local high-precision gravimetry. *Hydrological Processes*, 28(3):1465–1475. doi: https://doi.org/10.1002/hyp.9689.

B. Creutzfeldt, I. Heinrich, and B. Merz. 2015. Total water storage dynamics derived from tree-ring records and terrestrial gravity observations. *Journal of Hydrology*, 529:640–649. doi: 10.1016/j.jhydrol.2015.04.006.

D. Crossley, J. Hinderer, and U. Riccardi. 2013. The measurement of surface gravity. *Reports on Progress in Physics*, 76(4). doi: 10.1088/0034-4885/76/4/046101.

D. J. Crossley, O. G. Jensen, and J. Hinderer. 1995. Effective barometric admittance and gravity residuals. *Physics of the Earth and Planetary Interiors*, 90(3):221 – 241. doi: https://doi.org/10.1016/0031-9201(95)05086-Q.

J. A. C. Davies and K. Beven. 2015. Hysteresis and scale in catchment storage, flow and transport. *Hydrological Processes*, 29(16):3604–3615. doi: https://doi.org/10.1002/hyp.10511.

D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, and F. Vitart. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597. doi: 10.1002/qj.828.

S. Deville, T. Jacob, J. Chery, and C. Champollion. 2013. On the impact of topography and building mask on time varying gravity due to local hydrology. *Geophysical Journal International*, 192(1):82–93. doi: {10.1093/gji/ggs007}.

H. Dobslaw, I. Bergmann-Wolf, R. Dill, L. Poropat, M. Thomas, C. Dahle, S. Esselborn, R. König, and F. Flechtner. 2017. A new high-resolution model of non-tidal atmosphere and ocean mass variability for de-aliasing of satellite gravity observations: AOD1B RL06. *Geophysical Journal International*, 211(1):263–269. doi: 10.1093/gji/ggx302.

P. Dykowski, J. Krynski, and M. Sekowski. 2014. Testing the suitability of the A10-020 absolute gravimeter for the establishment of new gravity control in Poland. volume 141, pages 11–17. doi: 10.1007/978-3-319-10837-7_2.

G. D. Egbert and S. Y. Erofeeva. 2002. Efficient inverse modeling of barotropic ocean tides. *Journal of Atmospheric and Oceanic Technology*, 19(2):183–204. doi: 10.1175/1520-0426(2002)019⟨0183:EIMOBO⟩2.0.CO;2.

J. F. Ferguson, F. J. Klopping, T. Chen, J. E. Seibert, J. L. Hare, and J. L. Brady. Nov 2008. The 4D microgravity method for waterflood surveillance: Part 3 — 4D absolute microgravity surveys at Prudhoe Bay, Alaska. *Geophysics*, 73(6): WA163—-WA171. doi: 10.1190/1.2992510.

B. Fores, C. Champollion, N. Le Moigne, R. Bayer, and J. Chéry. Oct 2016. Assessing the precision of the igrav superconducting gravimeter for hydrological models and karstic hydrological process identification. *Geophysical Journal International*, 208 (1):269—-280. doi: 10.1093/gji/ggw396.

B. Fores, C. Champollion, N. L. Moigne, and J. Chery. 2017. Impact of ambient temperature on spring-based relative gravimeter measurements. *Journal of Geodesy*. doi: doi:10.1007/s00190-016-0961-2.

B. Fores, G. Klein, N. Le Moigne, and O. Francis. 2019. Long-term stability of tilt-controlled gphonex gravimeters. *Journal of Geophysical Research: Solid Earth*, 124(11):12264–12276. doi: https://doi.org/10.1029/2019JB018276.

R. Forsberg. 1984. A study of terrain reductions, density anomalies and geophysical inversion methods in gravity field modelling. Technical report, DTIC Document.

C. Förste, C. Voigt, M. Abe, C. Kroner, J. Neumeyer, H. Pflug, and P. Fourie, 2016. Superconducting gravimeter data from Sutherland - Level 1. doi: http://doi.org/10.5880/igets.su.l1.001.

O. Francis. 2021. Performance assessment of the relative gravimeter Scintrex CG-6. *Journal of Geodesy*, 95(10):116. doi: 10.1007/s00190-021-01572-y.

T. Francke, 2016. *ppso: Particle Swarm Optimization and Dynamically Dimensioned Search, optionally using parallel computing based on Rmpi*.

I. Fukumori. 2002. A partitioned kalman filter and smoother. *Monthly Weather Review*, 130(5):1370–1383. doi: 10.1175/1520-0493(2002)130⟨1370:APKFAS⟩2.0. CO;2.

C. L. Gehman, D. L. Harry, W. E. Sanford, J. D. Stednick, and N. A. Beckman. 2009. Estimating specific yield and storage change in an unconfined aquifer using temporal gravity surveys. *Water Resources Research*, 45(4). doi: 10.1029/2007wr006096.

R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A. M. da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J. E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S. D. Schubert, M. Sienkiewicz, and B. Zhao. 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454. doi: 10.1175/JCLI-D-16-0758.1.

P. Gillot, O. Francis, A. Landragin, F. P. Dos Santos, and S. Merlet. 2014. Stability comparison of two absolute gravimeters:optical versus atomic interferometers. *Metrologia*, 51(5):L15–L17. doi: 10.1088/0026-1394/51/5/l15.

F. Greco, G. Currenti, G. D'Agostino, A. Germak, R. Napoli, A. Pistorio, and C. Del Negro. 2012. Combining relative and absolute gravity measurements to enhance volcano monitoring. *Bulletin of Volcanology*, 74(7):1745–1756. doi: 10.1007/s00445-012-0630-0.

W. H. Green and G. A. Ampt. 1911. Studies on soil physics. *The Journal of Agricultural Science*, 4(1):1–24. doi: 10.1017/S0021859600001441.

A. Güntner, M. Reich, M. Mikolaj, B. Creutzfeldt, S. Schroeder, and H. Wziontek. 2017. Landscape-scale water balance monitoring with an iGrav superconducting gravimeter in a field enclosure. *Hydrology and Earth System Sciences*, 21(6): 3167–3182. doi: 10.5194/hess-21-3167-2017.

H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez. 2009. Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1):80 – 91. doi: https://doi.org/10.1016/j.jhydrol.2009.08.003.

GWR, 2013. *iGrav User Manual.*

A. Güntner, J. Stuck, S. Werth, P. Döll, K. Verzano, and B. Merz. 2007. A global analysis of temporal and spatial variations in continental water storage. *Water Resources Research*, 43(5). doi: https://doi.org/10.1029/2006WR005247.

M. Harnisch and G. Harnisch. 2002. Seasonal variations of hydrological influences on gravity measurements at Wettzell. *Marées Terrestres Bulletin d'Informations*, 137:10849–10861.

S. Hasan, P. Troch, J. Boll, and C. Kroner. 2006. Modeling the hydrological effect in local gravity at Moxa, Germany. *Journal of Hydrometeorology*, 7(3):346–354. doi: 10.1175/JHM488.1.

M. Hauth, C. Freier, V. Schkolnik, A. Peters, H. Wziontek, and M. Schilling. 2014. Atom interferometry for absolute measurements of local gravity. In *Proceedings of the International School of Physics "Enrico Fermi"*, volume 188, pages 557–586. doi: 10.3254/978-1-61499-448-0-557.

B. Hector, L. Seguis, J. Hinderer, M. Descloitres, J.-M. Vouillamoz, M. Wubda, J.-P. Boy, B. Luck, and N. Le Moigne. 2013. Gravity effect of water storage changes in a weathered hard-rock aquifer in West Africa: results from joint absolute gravity, hydrological monitoring and geophysical prospection. *Geophysical Journal International*, 194(2):737–750. doi: 10.1093/gji/ggt146.

B. Hector, J. Hinderer, L. Séguis, J.-P. Boy, M. Calvo, M. Descloitres, S. Rosat, S. Galle, and U. Riccardi. 2014. Hydro-gravimetry in West-Africa: First results from the Djougou (Benin) superconducting gravimeter. *Journal of Geodynamics*, 80:34–49. doi: http://dx.doi.org/10.1016/j.jog.2014.04.003.

B. Hector, L. Séguis, J. Hinderer, J.-M. Cohard, M. Wubda, M. Descloitres, N. Benarrosh, and J.-P. Boy. 2015. Water storage changes as a marker for base flow generation processes in a tropical humid basement catchment (Benin): Insights from hybrid gravimetry. *Water Resources Research*, 51(10):8331–8361. doi: 10.1002/2014WR015773.

A. Heim, 2010. Modelling the water dynamics at a hillslope in the bavarian forst, Germany. Master's thesis, Potsdam University.

W. A. Heiskanen and H. Moritz. 1967. *Physical geodesy*, volume 86. W.H. Freeman and Company, San Francisco and London.

B. Hemmings, J. Gottsmann, F. Whitaker, and A. Coco. 2016. Investigating hydrological contributions to volcano monitoring signals: A time-lapse gravity example. *Geophysical Journal International*, 207(1):259–273. doi: 10.1093/gji/ggw266.

H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730): 1999–2049. doi: https://doi.org/10.1002/qj.3803.

J. Hinderer, S. Rosat, D. Crossley, M. Amalvict, J.-P. Boy, and P. Gegout. 2002. Influence of different processing methods on the retrieval of gravity signals from GGP data. *Bulletin d'Information des Marées Terrestres*, (135).

J. Hinderer, B. Hector, J. P. Boy, U. Riccardi, S. Rosat, M. Calvo, and F. Littel. 2014. A search for atmospheric effects on gravity at different time and space scales. *Journal of Geodynamics*, 80(0):50–57. doi: http://dx.doi.org/10.1016/j.jog.2014.02.001.

J. Hinderer, M. Calvo, Y. Abdelfettah, B. Hector, U. Riccardi, G. Ferhat, and J.-D. Bernard. 2015a. Monitoring of a geothermal reservoir by hybrid gravimetry; feasibility study applied to the Soultz-sous-Forêts and Rittershoffen sites in the Rhine graben. *Geothermal Energy*, 3(1):1–19. doi: 10.1186/s40517-015-0035-3.

J. Hinderer, D. Crossley, and R. J. Warburton, 2015b. 3.04 - Gravimetric Methods - Superconducting Gravity Meters. In G. Schubert, editor, *Treatise on Geophysics*, pages 59–115. Elsevier, Amsterdam, Netherlands. doi: 10.1016/B978-0-444-53802-4.00062-2.

J. Hinderer, B. Hector, A. Mémin, and M. Calvo. 2016. Hybrid gravimetry as a tool to monitor surface and underground mass changes. volume 147, pages 123–130. doi: 10.1007/1345_2016_253.

M. Z. İmamoğlu and E. Sertel. 2016. Analysis of Different Interpolation Methods for Soil Moisture Mapping Using Field Measurements and Remotely Sensed Data. *International Journal of Environment and Geoinformatics*, 3:11 – 25. doi: 10.30897/ijegeo.306477.

Y. Imanishi, T. Sato, T. Higashi, W. Sun, and S. Okubo. 2004. A network of superconducting gravimeters detects submicrogal coseismic gravity changes. *Science*, 306(5695):476–478. doi: 10.1126/science.1101875.

Z. Iritz and A. Lindroth. 1994. Night-time evaporation from a short-rotation willow stand. *Journal of Hydrology*, 157(1):235–245. doi: https://doi.org/10.1016/0022-1694(94)90107-4.

C. Jackisch, L. Angermann, N. Allroggen, M. Sprenger, T. Blume, J. Tronicke, and E. Zehe. 2017. Form and function in hillslope hydrology: in situ imaging and characterization of flow-relevant structures. *Hydrology and Earth System Sciences*, 21(7):3749–3775. doi: 10.5194/hess-21-3749-2017.

C. R. Jackson, E. Du, J. Klaus, N. A. Griffiths, M. Bitew, and J. J. McDonnell. 2016. Interactions among hydraulic conductivity distributions, subsurface topography, and transport thresholds revealed by a multitracer hillslope irrigation experiment. *Water Resources Research*, 52(8):6186–6206. doi: 10.1002/2015wr018364.

T. Jacob, J. Chery, R. Bayer, N. Le Moigne, J.-P. Boy, P. Vernant, and F. Boudin. 2009. Time-lapse surface to depth gravity measurements on a karst system reveal the dominant role of the epikarst as a water storage entity. *Geophysical Journal International*, 177(2):347–360. doi: 10.1111/j.1365-246X.2009.04118.x.

C. Janvier, V. Ménoret, B. Desruelle, S. Merlet, A. Landragin, and F. P. dos Santos. 2022. Compact differential gravimeter at the quantum projection-noise limit. *Physical Review A*, 105(2):022801. doi: 10.1103/physreva.105.022801.

JCGM. 2008. Joint Committee for Guides in Metrology: Evaluation of measurement data–Guide to the expression of uncertainty in measurement. Technical report, JCGM.

S. Jomaa, D. Barry, A. Brovelli, G. Sander, J.-Y. Parlange, B. Heng, and H. T. van Meerveld. 2010. Effect of raindrop splash and transversal width on soil erosion: Laboratory flume experiments and analysis with the hairsine–rose model. *Journal of Hydrology*, 395(1-2):117–132. doi: 10.1016/j.jhydrol.2010.10.021.

P. Jousset, S. Dwipa, F. Beauducel, T. Duquesnoy, and M. Diament. 2000. Temporal gravity at Merapi and during the and crisis: an insight and into the dynamical behaviour of volcanoes. *Journal of Volcanology and Geothermal Research*, 100 (100):289—-320. doi: 10.1016/s0377-0273(00)00141-4.

M. Kanamitsu, W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter. 2002. NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83(11):1631–1644. doi: 10.1175/BAMS-83-11-1631.

R. Kao, H. Kabirzadeh, J. W. Kim, J. Neumeyer, and M. G. Sideris. 2014. Detecting small gravity change in field measurement: simulations and experiments of the superconducting gravimeter–iGrav. *Journal of Geophysics and Engineering*, 11 (4):1–11.

T. Kazama, Y. Tamura, K. Asari, S. Manabe, and S. Okubo. 2012. Gravity changes associated with variations in local land-water distributions: Observations and hydrological modeling at Isawa Fan, northern Japan. *Earth, Planets and Space*, 64(4):309–331. doi: 10.5047/eps.2011.11.003.

T. Kazama, S. Okubo, T. Sugano, S. Matsumoto, W. Sun, Y. Tanaka, and E. Koyama. 2015. Absolute gravity change associated with magma mass movement in the conduit of Asama Volcano (Central Japan), revealed by physical modeling of hydrological gravity disturbances. *Journal of Geophysical Research: Solid Earth*, 120(2):1263—1287. doi: 10.1002/2014jb011563.

J. Kennedy, T. Ferré, B. Creutzfeldt, A. Guntner, J. Neumeyer, E. Brinton, and R. Warburton. 2013. Smaller is better: first experiences using the igrav superconducting gravimeter in a field enclosure. Warsaw, Poland. 17th International Symposium on Earth Tides.

J. Kennedy, T. P. A. Ferré, A. Güntner, M. Abe, and B. Creutzfeldt. 2014. Direct measurement of sub-surface mass changes using the variable-baseline gravity gradient method. *Geophysical Research Letters*, 41. doi: 10.1002/2014gl059673.

J. Kennedy, T. P. A. Ferré, and B. Creutzfeldt. 2016. Time-lapse gravity data for monitoring and modeling artificial recharge through a thick unsaturated zone. *Water Resources Research*, 52:7244–7261. doi: 10.1002/2016wr018770.

J. R. Kennedy and T. P. Ferré. 2015. Accounting for time- and space-varying changes in the gravity field to improve the network adjustment of relative-gravity data. *Geophysical Journal International*, 204(2):892–906. doi: 10.1093/gji/ggv493.

T. Klügel and H. Wziontek. 2009. Correcting gravimeters and tiltmeters for atmospheric mass attraction using operational weather models. *Journal of Geodynamics*, 48(3-5):204 – 210. doi: http://dx.doi.org/10.1016/j.jog.2009.09.010.

M. Kottek, J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3):259–263. doi: 10.1127/0941-2948/2006/0130.

P. Krause, M. Naujoks, M. Fink, and C. Kroner. 2009. The impact of soil moisture changes on gravity residuals obtained with a superconducting gravimeter. *Journal of Hydrology*, 373(1-2):151–163. doi: {10.1016/j.jhydrol.2009.04.019}.

C. Kroner and T. Jahr. 2006. Hydrological experiments around the superconducting gravimeter at Moxa Observatory. *Journal of Geodynamics*, 41(1-3):268–275. doi: 10.1016/j.jog.2005.08.012.

G. Laaha, T. Gauster, L. M. Tallaksen, J.-P. Vidal, K. Stahl, C. Prudhomme, B. Heudorfer, R. Vlnas, M. Ionita, H. A. J. Van Lanen, M.-J. Adler, L. Caillouet, C. Delus, M. Fendekova, S. Gailliez, J. Hannaford, D. Kingston, A. F. Van Loon, L. Mediero, M. Osuch, R. Romanowicz, E. Sauquet, J. H. Stagge, and W. K. Wong. 2017. The european 2015 drought from a hydrological perspective. *Hydrology and Earth System Sciences*, 21(6):3001–3024. doi: 10.5194/hess-21-3001-2017.

L. J. B. LaCoste. 1934. A New Type Long Period Vertical Seismograph. *Physics*, 5 (7):178–180. doi: 10.1063/1.1745248.

A. Lambert and C. Beaumont. 1977. Nano variations in gravity due to seasonal groundwater movements: Implications for the gravitational detection of tectonic movements. *Journal of Geophysical Research*, 82(2):297–306. doi: 10.1029/JB082i002p00297.

D. R. Legates and G. J. McCabe Jr. 1999. Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1):233–241. doi: https://doi.org/10.1029/1998WR900018.

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

J. Liard and C. Gagnon. 2002. The new A-10 absolute gravimeter at the 2001 International Comparison of Absolute Gravimeters. *Metrologia*, 39(5):477–483. doi: 10.1088/0026-1394/39/5/8.

R. Lisboa, R. De Azevedo, and R. Reis. 2006. Residual soil hydraulic conductivity determination using field and laboratory tests. In *Unsaturated Soils*, number 147, pages 1559–1570. doi: 10.1061/40802(189)130.

L. T. Long and R. D. Kaufmann. 2013. *Acquisition and Analysis of Terrestrial Gravity Data*. Cambridge University Press.

L. Longuevergne, J. Boy, N. Florsch, D. Viville, G. Ferhat, P. Ulrich, B. Luck, and J. Hinderer. 2009. Local and global hydrological contributions to gravity variations observed in strasbourg. *Journal of Geodynamics*, 48(3–5):189 – 194. doi: http://dx.doi.org/10.1016/j.jog.2009.09.008.

C. Loren and L. Florent. 2003. Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing-comparisons with observations. *Geophysical Research Letters*, 30(6). doi: 10.1029/2002GL016473.

W. Lowrie. 2007. *Fundamentals of Geophysics*. Cambridge University Press, 2 edition.

E. Malek. 1992. Night-time evapotranspiration vs. daytime and 24h evapotranspiration. *Journal of Hydrology*, 138(1):119–129. doi: https://doi.org/10.1016/0022-1694(92)90159-S.

J. J. McDonnell, M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, M. L. Roderick, and et al. 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research*, 43(7):1–6. doi: 10.1029/2006wr005467.

B. L. McGlynn, J. J. McDonnel, and D. D. Brammer. 2002. A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand. *Journal of Hydrology*, 257(1):1–26. doi: https://doi.org/10.1016/S0022-1694(01)00559-5.

D. Menemenlis, J. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlok, and H. Zhang. 2008. ECCO2: High Resolution Global Ocean and Sea Ice Data Synthesis. *Mercator Ocean Quarterly Newsletter*, (31):13–21.

V. Ménoret, P. Vermeulen, N. L. Moigne, S. Bonvalot, P. Bouyer, A. Landragin, and B. Desruelle. 2018. Gravity measurements below 10-9 g with a transportable absolute quantum gravimeter. *Scientific Reports*, 8(1). doi: 10.1038/s41598-018-30608-1.

J. B. Merriam. 1992. Atmospheric pressure and gravity. *Geophysical Journal International*, 109(3):488–500. doi: 10.1111/j.1365-246X.1992.tb00112.x.

B. Meurers. 2012. Superconducting gravimeter calibration by colocated gravity observations: Results from GWR C025. *International Journal of Geophysics*, 2012:1–12. doi: 10.1155/2012/954271.

B. Meurers, M. Van Camp, and T. Petermans. 2007. Correcting superconducting gravity time-series using rainfall modelling at the Vienna and Membach stations and application to Earth tide analysis. *Journal of Geodesy*, 81(11):703–712. doi: 10.1007/s00190-007-0137-1.

B. Meurers, M. Van Camp, O. Francis, and V. Pálinkáš. 2016. Temporal variation of tidal parameters in superconducting gravimeter time-series. *Geophysical Journal International*, 205(1):284–300. doi: 10.1093/gji/ggw017.

Micro-g, 2008. *A-10 User Manual*.

Micro-g, 2013. *gPhoneX User Manual*.

Micro-g, 2015. *FG5-X Absolute Gravimeter User's Manual*.

R. P. Middlemiss, A. Samarelli, D. J. Paul, J. Hough, S. Rowan, and G. D. Hammond. 2016. Measurement of the Earth tides with a MEMS gravimeter. *Nature*, 531(7596). doi: 10.1038/nature17397.

R. P. Middlemiss, P. Campsie, W. Cunningham, R. Douglas, V. McIvor, J. Hough, S. Rowan, D. J. Paul, A. Prasad, and G. D. Hammond. 2022. A MEMS gravimeter with multi-axis gravitational sensitivity. In *2022 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pages 1–4. doi: http://dx.doi.org/10.1109/INERTIAL53425.2022.9787754.

M. Mikolaj, B. Meurers, and M. Mojzeš. 2015. The reduction of hydrology-induced gravity variations at sites with insufficient hydrological instrumentation. *Studia Geophysica et Geodaetica*, 59(3):424–437. doi: 10.1007/s11200-014-0232-8.

M. Mikolaj, B. Meurers, and A. Güntner. 2016. Modelling of global mass effects in hydrology, atmosphere and oceans on surface gravity. *Computers & Geosciences*, 93:12—-20. doi: 10.1016/j.cageo.2016.04.014.

M. Mikolaj, M. Reich, and A. Güntner. 2019a. Supporting information to: Resolving geophysical signals by terrestrial gravimetry: a time domain assessment of the correction-induced uncertainty. *GFZ Data Services*. doi: http://doi.org/10.5880/GFZ.4.4.2019.001.

M. Mikolaj, M. Reich, and A. Güntner. 2019b. Resolving Geophysical Signals by Terrestrial Gravimetry: A Time Domain Assessment of the Correction-Induced Uncertainty. *Journal of Geophysical Research: Solid Earth*, 124:2153–2165. doi: 10.1029/2018jb016682.

P. Müller, 2010. Quantifizierung des Einflusses der Makroporeninfiltration auf den Bodenwasserfluss verschiedener Standorte im Bayerischen Wald. Master's thesis, Technical University Desden.

Muquans, 2019. *Absolute Quantum Gravimeter AQG B User Manual*.

D. Nagy. 1966. The gravitational attraction of a right rectangular prism. *Geophysics*, 31(2):362–371. doi: 10.1190/1.1439779.

M. Naujoks, C. Kroner, A. Weise, T. Jahr, P. Krause, and S. Eisner. 2010. Evaluating local hydrological modelling by temporal gravity observations and a gravimetric three-dimensional model. *Geophysical Journal International*, 182(1):233–249.

J. Neumeyer. 2010. *Sciences of Geodesy – I: Advances and Future Directions*, chapter Superconducting Gravimetry, pages 339 – 414. Springer, Heidelberg.

T. Niebauer, 2015. Absolute and Relative Gravity Meter: Instruments Concepts and Implementation. In G. Schubert, editor, *Treatise on Geophysics*, chapter 3.03 - Gravimetric Methods, pages 37–57. Elsevier, Oxford, 2nd edition. doi: https://doi.org/10.1016/B978-0-444-53802-4.00057-9.

T. M. Niebauer, G. S. Sasagawa, J. E. Faller, R. Hilt, and F. Klopping. 1995. A new generation of absolute gravimeters. *Metrologia*, 32(3):159–180. doi: 10.1088/0026-1394/32/3/004.

T. E. Ochsner, M. H. Cosh, R. H. Cuenca, W. A. Dorigo, C. S. Draper, Y. Hagimoto, Y. H. Kerr, K. M. Larson, E. G. Njoku, E. E. Small, and M. Zreda. 2013. State of the Art in Large-Scale Soil Moisture Monitoring. *Soil Science Society of America Journal*, 77(6):1888–1919. doi: https://doi.org/10.2136/sssaj2013.03.0093.

V. Ophaug, K. Breili, C. Gerlach, J. G. O. Gjevestad, D. I. Lysaker, O. C. D. Omang, and B. R. Pettersen. 2016. Absolute gravity observations in Norway (1993–2014) for glacial isostatic adjustment studies: The influence of gravitational loading effects on secular gravity trends. *Journal of Geodynamics*, 102:83–94. doi: https://doi.org/10.1016/j.jog.2016.09.001.

W. P., W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe. 2013. Generic mapping tools: Improved version released. *Eos, Transactions American Geophysical Union*, 94(45):409–410. doi: 10.1002/2013EO450001.

R. Pasteka, J. Mikuska, and B. Meurers. 2017. *Understanding the Bouguer Anomaly*. Elsevier.

J. E. Pendiuk, L. Guarracino, M. Reich, C. Brunini, and A. Güntner. 2020. Estimating the specific yield of the pampeano aquifer, argentina, using superconducting gravimeter data. *Hydrogeology Journal*, 28:2303–2313. doi: 10.1007/s10040-020-02212-z.

A. Peters, T. Nehls, H. Schonsky, and G. Wessolek. 2014. Separating precipitation and evapotranspiration from noise - a new filter routine for high-resolution lysimeter data. *Hydrology and Earth System Sciences*, 18(3):1189–1198. doi: 10.5194/hess-18-1189-2014.

G. Petit and B. Luzum. 2010. IERS Technical and Note No. 36. In *IERS Conventions*, page 179. Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.

J. Pfeffer, C. Champollion, G. Favreau, B. Cappelaere, J. Hinderer, M. d. Boucher, Y. Nazoumou, M. Oï, M. f. Mouyen, C. Henri, N. Le Moigne, S. Deroussi, J. Demarty, N. Boulain, N. Benarrosh, and O. Robert. 2013. Evaluating surface and subsurface water storage variations at small time and space scales from relative

gravity measurements in semiarid niger. *Water Resources Research*, 49(6):3276–3291. doi: 10.1002/wrcr.20235.

S. Piccolroaz, B. Majone, F. Palmieri, G. Cassiani, and A. Bellin. 2015. On the use of spatially distributed, time-lapse microgravity surveys to inform hydrological modeling. *Water Resources Research*, 51(9):7270–7288. doi: https://doi.org/10.1002/2015WR016994.

H.-P. Plag and M. Pearlman. 2009. *Global Geodetic Observing System*. Springer-Verlag Berlin Heidelberg.

M. P. Poland and D. Carbone. 2016. Insights into shallow magmatic processes at Kilauea Volcano, Hawaii, from a multiyear continuous gravity time series. *Journal of Geophysical Research: Solid Earth*, 121(7):5477–5492. doi: 10.1002/2016JB013057.

D. R. Pool. 2005. Variations in climate and ephemeral channel recharge in southeastern arizona, united states. *Water Resources Research*, 41. doi: 10.1029/2004wr003255.

D. R. Pool and J. H. Eychaner. 1995. Measurements of aquifer-storage change and specific yield using gravity surveys. *Ground Water*, 33(3):425–432. doi: 10.1111/j.1745-6584.1995.tb00299.x.

W. A. Prothero and J. M. Goodkind. 1968. A superconducting gravimeter. *Review of Scientific Instruments*, 39(9):1257–1262. doi: 10.1063/1.1683645.

C. Quentin, J. Hinderer, F. Masson, D. Viville, J.-D. Bernard, S. Cotel, M.-C. Pierret, N. Lesparre, and B. Jeannot. 2020. *Continuous Monitoring with a Superconducting Gravimeter As a Proxy for Water Storage Changes in a Mountain Catchment*. doi: 10.1007/1345_2020_105.

C. Quentin, N. Lesparre, F. Masson, J. Hinderer, D. Viville, J.-D. Bernard, G. Ferhat, and S. Cotel. 2021. Hybrid gravimetry to map Water Storage Dynamics in a mountain catchment. *Frontiers in Water*. doi: https://doi.org/10.3389/frwa.2021.715298.

P. Quinn, K. Beven, P. Chevallier, and O. Planchon. 1991. The prediction and of hillslope and flow paths and for and distributed hydrological and modelling using and digital. *Hydrological processes*, 5:59–79. doi: https://doi.org/10.1002/hyp.3360050106.

R Core Team, 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

M. Reich, 2021. gravityInf - inverse hydrological modeling based on gravity observations to evaluate dominant infiltration and subsurface re-distribution processes. doi: https://doi.org/10.5880/GFZ.4.4.2021.002.

M. Reich, M. Mikolaj, T. Blume, and A. Güntner. 2019. Reducing gravity data for the influence of water storage variations beneath observatory buildings. *Geophysics*, 84:EN15–EN31. doi: 10.1190/geo2018-0301.1.

M. Reich, H. Thoss, S. Schröder, and A. Güntner, 2020. The gPhone-solar-cube: an energy self-sufficient mobile container for monitoring gravity and climate parameters at remote field sites. doi: 10.1002/essoar.10502918.1.

M. Reich, M. Mikolaj, T. Blume, and A. Güntner, 2021a. Gravity and ancillary monitoring data of a sprinkling experiment - complemented by model setups and model output. doi: https://doi.org/10.5880/GFZ.4.4.2021.001.

M. Reich, M. Mikolaj, T. Blume, and A. Güntner. 2021b. Field-Scale Subsurface Flow Processes Inferred From Continuous Gravity Monitoring During a Sprinkling Experiment. *Water Resources Research*, 57(10):e2021WR030044. doi: https://doi.org/10.1029/2021WR030044.

D. Richter. 1995. Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Meßfehlers des Hellmann-Niederschlagsmessers. Technical Report 194, Deutscher Wetterdienst, Offenbach.

M. Rodell, P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, and M. Bosilovich. 2004. The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, 85 (3):381—-394. doi: 10.1175/bams-85-3-381.

S. Rosat and J. Hinderer. 2011. Noise Levels of Superconducting Gravimeters: Updated Comparison and Time Stability. *Bulletin of the Seismological Society of America*, 101(3):1233. doi: 10.1785/0120100217.

S. Rosat and J. Hinderer. 2018. Limits of Detection of Gravimetric Signals on Earth. *Scientific Reports*, 8. doi: 10.1038/s41598-018-33717-z.

F. Rubel, K. Brugger, K. Haslinger, and I. Auer. 2017. The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800-2100. *Meteorologische Zeitschrift*, 26(2):115–125. doi: 10.1127/metz/2016/0816.

C. Rücker and T. Günther. 2011. The simulation of finite ERT electrodes using the complete electrode model. *Geophysics*, 76(4):F227–F238. doi: 10.1190/1.3581356.

A. Samouëlian, I. Cousin, A. Tabbagh, A. Bruand, and G. Richard. 2005. Electrical resistivity survey in soil science: a review. *Soil and Tillage Research*, 83(2):173–193. doi: https://doi.org/10.1016/j.still.2004.10.004.

W. Schlüter, N. Brandl, R. Dassing, H. Hase, T. Klügel, R. Kilger, P. Lauber, A. Neidhardt, C. Plötz, S. Riepl, and U. Schreiber. 2007. Fundamentalstation Wettzell –ein geodätisches Observatorium. *Zeitschrift für Vermessungswesen*, 132: 158 – 167.

M. Schumacher, E. Forootan, A. van Dijk, H. M. Schmied, R. Crosbie, J. Kusche, and P. Döll. 2018. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sensing of Environment*, 204:212–228. doi: 10.1016/j.rse.2017.10.029.

Scintrex Ltd., 2017. *CG-6 Operation Manual.*

M. Stock. 2013. Watt balance experiments for the determination of the planck constant and the redefinition of the kilogram. *Metrologia*, 50(1):R1.

M. Sugihara, Y. Nishi, H. Ikeda, K. Nawa, and T. Ishido. 2017. Monitoring CO2 injection at the tomakomai field using high-sensitivity continuous gravimetry. *Energy Procedia*, 114:4020 – 4027. doi: https://doi.org/10.1016/j.egypro.2017.03.1542.

Y. Tamura, T. Sato, M. Ooe, and M. Ishiguro. 1991. A procedure for tidal analysis with a Bayesian information criterion. *Geophysical Journal International*, 104(3):507–516. doi: 10.1111/j.1365-246X.1991.tb05697.x.

T. Tanaka, R. Miyajima, H. Asai, Y. Horiuchi, K. Kumada, Y. Asai, and H. Ishii. 2013. Hydrological gravity response detection using a gphone below- and above-ground. *Earth, Planets and Space*, 65(2):59–66. doi: 10.5047/eps.2012.06.012.

B. D. Tapley, S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins. 2004. Grace measurements of mass variability in the earth system. *Science (New York, N.Y.)*, 305:503–5.

D. K. Todd and L. W. Mays. 2004. *Groundwater Hydrology*. John Wiley & Sons.

C.-H. M. Tso, O. Kuras, and A. Binley. 2019. On the field estimation of moisture content using electrical geophysics: The impact of petrophysical model uncertainty. *Water Resources Research*, 55(8):7196–7211. doi: https://doi.org/10.1029/2019WR024964.

J. Šimůnek, M. Šejna, H. Saito, M. Sakai, and M. T. van Genuchten, 2013. *The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media*. Department of environmental sciences, University of California Riverside.

J. Šimůnek, M. T. van Genuchten, and M. Šejna. 2016. Recent Developments and Applications of the HYDRUS Computer Software Packages. *Vadose Zone Journal*, 15(7):vzj2016.04.0033. doi: https://doi.org/10.2136/vzj2016.04.0033.

M. Vallée, J. P. Ampuero, K. Juhel, P. Bernard, J.-P. Montagner, and M. Barsuglia. 2017. Observations and modeling of the elastogravity signals preceding direct seismic waves. *Science*, 358(6367):1164–1168. doi: 10.1126/science.aao0746.

M. Van Camp and P. Vauterin. 2005. Tsoft: graphical and interactive software for the analysis of time series and Earth tides. *Computers & Geosciences*, 31(5):631–640. doi: 10.1016/j.cageo.2004.11.015.

M. Van Camp, S. D. P. Williams, and O. Francis. 2005. Uncertainty of absolute gravity measurements. *Journal of Geophysical Research: Solid Earth*, 110(B5). doi: 10.1029/2004JB003497.

M. Van Camp, M. Vanclooster, O. Crommen, T. Petermans, K. Verbeeck, B. Meurers, T. van Dam, and A. Dassargues. 2006. Hydrogeological investigations at the membach station, belgium, and application to correct long periodic gravity variations. *Journal of Geophysical Research: Solid Earth*, 111(B10). doi: https://doi.org/10.1029/2006JB004405.

M. Van Camp, L. Métivier, O. de Viron, B. Meurers, and S. D. P. Williams. 2010. Characterizing long-time scale hydrological effects on gravity for improved distinction of tectonic signals. *Journal of Geophysical Research: Solid Earth*, 115 (B7). doi: 10.1029/2009JB006615.

M. Van Camp, O. de Viron, H.-G. Scherneck, K.-G. Hinzen, S. D. P. Williams, T. Lecocq, Y. Quinif, and T. Camelbeeck. 2011. Repeated absolute gravity measurements for monitoring slow intraplate vertical deformation in western Europe. *Journal of Geophysical Research: Solid Earth*, 116(B8). doi: 10.1029/2010JB008174.

M. Van Camp, O. de Viron, and J. Avouac. 2016a. Separating climate-induced mass transfers and instrumental effects from tectonic signal in repeated absolute gravity measurements. *Geophysical Research Letters*, 43(9):4313–4320. doi: 10. 1002/2016GL068648.

M. Van Camp, O. de Viron, G. Pajot-Métivier, F. Casenave, A. Watlet, A. Dassargues, and M. Vanclooster. 2016b. Direct measurement of evapotranspiration from a forest using a superconducting gravimeter. *Geophysical Research Letters*, 43(19):10225–10231. doi: 10.1002/2016GL070534.

M. Van Camp, B. Meurers, O. Viron, and T. Forbriger. 2016c. Optimized strategy for the calibration of superconducting gravimeters at the one per mille level. *Journal of Geodesy*, 90(1):91–99. doi: 10.1007/s00190-015-0856-7.

M. Van Camp, O. de Viron, A. Watlet, B. Meurers, O. Francis, and C. Caudron. 2017. Geophysics from terrestrial time-variable gravity measurements. *Reviews of Geophysics*, 55(4):938–992. doi: 10.1002/2017RG000566.

T. van Dam, O. Francis, J. Wahr, S. A. Khan, M. Bevis, and M. R. van den Broeke. 2017. Using GPS and absolute gravity observations to separate the effects of present-day and Pleistocene ice-mass changes in South East Greenland. *Earth and Planetary Science Letters*, 459:127 – 135. doi: https://doi.org/10.1016/j.epsl. 2016.11.014.

N. L. M. B. van Schaik, R. F. A. Hendriks, and J. C. van Dam. 2010. Parameterization of macropore flow using dye-tracer infiltration patterns in the SWAP model. *Vadose Zone Journal*, 9(1):95. doi: 10.2136/vzj2009.0031.

A. P. Venedikov, J. Arnoso, and R. Vieira. 2005. New version of program VAV for tidal data processing. *Computers & Geosciences*, 31(5):667 – 669. doi: https://doi.org/10.1016/j.cageo.2004.12.001.

H. Vereecken, A. Schnepf, J. Hopmans, M. Javaux, D. Or, T. Roose, J. Vanderborght, M. Young, W. Amelung, M. Aitkenhead, S. Allison, S. Assouline, P. Baveye, M. Berli, N. Brüggemann, P. Finke, M. Flury, T. Gaiser, G. Govers, T. Ghezzehei, P. Hallett, H. H. Franssen, J. Heppell, R. Horn, J. Huisman, D. Jacques, F. Jonard, S. Kollet, F. Lafolie, K. Lamorski, D. Leitner, A. McBratney, B. Minasny, C. Montzka, W. Nowak, Y. Pachepsky, J. Padarian, N. Romano, K. Roth, Y. Rothfuss, E. Rowe, A. Schwen, J. Šimůnek, A. Tiktak, J. V. Dam, S. van der Zee, H. Vogel, J. Vrugt, T. Wöhling, and I. Young. 2016. Modeling Soil

Processes: Review, Key Challenges, and New Perspectives. *Vadose Zone Journal*, 15(5):vzj2015.09.0131. doi: 10.2136/vzj2015.09.0131.

C. Voigt, C. Förste, H. Wziontek, D. Crossley, B. Meurers, V. Pálinkáš, J. Hinderer, J.-P. Boy, J.-P. Barriot, and H. Sun. 2016. Report on the Data Base of the International Geodynamics and Earth Tide Service (IGETS). Technical report, GFZ German Research Centre for Geosciences.

G. von Unold and J. Fank. 2007. Modular Design of Field Lysimeters for Specific Application Needs. *Water, Air, & Soil Pollution: Focus*, 8(2):233–242. doi: 10.1007/s11267-007-9172-4.

R. Warburton, H. Pillai, and R. Reineman. 2010. Initial results with the new GWR iGrav superconducting gravity meter. In *Extended Abstract Presented at 2nd Asia Workshop on Superconducting Gravimetry Taipei, Taiwan*.

R. J. Warburton and J. M. Goodkind. 1977. The influence of barometric-pressure variations on gravity. *Geophysical Journal International*, 48(3):281–292. doi: 10.1111/j.1365-246X.1977.tb03672.x.

A. Watlet, M. Van Camp, O. Francis, A. Poulain, G. Rochez, V. Hallet, Y. Quinif, and O. Kaufmann. 2020. Gravity monitoring of underground flash flood events to study their impact on groundwater recharge and the distribution of karst voids. *Water Resources Research*, 56(4):e2019WR026673. doi: https://doi.org/10.1029/2019WR026673.

H.-G. Wenzel. 1996. The nanogal software: Earth tide data processing package ETERNA 3.30. *Bulletin d'Informations Marées Terrestres*, 124:9425–9439.

H. Wickham. 2007. Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21(12):1–20.

H. Wickham. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

H. Wickham, R. François, L. Henry, and K. Müller, 2020. *dplyr: A Grammar of Data Manipulation*.

J. Wienhöfer and E. Zehe. 2014. Predicting subsurface stormflow response of a forested hillslope - the role of connected flow paths. *Hydrology and Earth System Sciences*, 18(1):121–138. doi: 10.5194/hess-18-121-2014.

J. Wienhöfer, K. Germer, F. Lindenmaier, A. Färber, and E. Zehe. 2009. Applied tracers for the observation of subsurface stormflow at the and hillslope scale. *Hydrology and Earth System Sciences*, 13:1145–1161. doi: 10.5194/hess-13-1145-2009.

C. R. Wilson, B. Scanlon, J. Sharp, L. Longuevergne, and H. Wu. 2012. Field test of the superconducting gravimeter as a hydrologic sensor. *Groundwater*, 50(3):442–449. doi: https://doi.org/10.1111/j.1745-6584.2011.00864.x.

H. Wziontek, P. Wolf, A. Böer, H. Hase, I. Nowak, B. Richter, and H. Wilmes, 2017a. Superconducting gravimeter data from TIGO/Concepción - Level 1. doi: http://doi.org/10.5880/igets.tc.l1.001.

H. Wziontek, P. Wolf, I. Nowak, B. Richter, A. Rülke, and H. Wilmes. 2017b. Superconducting Gravimeter Data from Wettzell - Level 1. *GFZ Data Services*. doi: https://doi.org/10.5880/igets.we.l1.001.

C. Xu. 2017. Evaluating mass loading products by comparison to GPS array daily solutions. *Geophysical Journal International*, 208(1):24–35. doi: 10.1093/gji/ggw385.

X. Yao, B. Fu, Y. Lü, F. Sun, S. Wang, and M. Liu. 2013. Comparison of Four Spatial Interpolation Methods for Estimating Soil Moisture in a Complex Terrain Catchment. *PLOS ONE*, 8(1):1–13. doi: 10.1371/journal.pone.0054660.

W. Zürn and R. Widmer. 1995. On noise reduction in vertical seismic records below 2 mHz using local barometric pressure. *Geophysical Research Letters*, 22 (24):3537–3540. doi: 10.1029/95GL03369.

Author's Declaration

I prepared this dissertation without illegal assistance. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. This dissertation has not been presented to any other University for examination, neither in Germany nor in another country.