Total water storage dynamics in response to climate variability and extremes: Inference from long-term terrestrial gravity measurement

Benjamin Creutzfeldt,1 Ty Ferré,2 Peter Troch,2 Bruno Merz,1 Hartmut Wziontek,3 and Andreas Güntner1

Received 27 June 2011; revised 13 March 2012; accepted 15 March 2012; published 27 April 2012.

[1] Terrestrial water storage is a basic element of the hydrological cycle and a key state variable for land surface-atmosphere interaction. However, measuring water storage in a comprehensive way for different storage compartments and beyond the point scale is a challenge. In this study, we explore a 10-year time series of total water storage changes derived from high-precision superconducting gravimeter observations in a headwater catchment in Southern Germany. In combination with hydro-meteorological data, we examine the relationship between gravity-derived water storage changes, climate, and river discharge. Distinct seasonal water storage dynamics observed by the gravimeter are strongly related to the meteorological forcing, in particular evapotranspiration. Intra-annual water storage variations demonstrate that the simplifying assumption of water storage averaging to zero at the annual scale is not valid for this catchment. At the event-scale, gravimeters provide a measure of the available subsurface water storage capacity, which can be useful for runoff prediction. During the Central European drought in 2003, the gravimeter data show a strong depletion of water storage and a long-term recovery that extended over a period of several years. In comparison to point measurements or different environmental indices, our findings suggest that depth-integrated gravimeter measurements give a more complete picture of the dynamics of a hydrologic system in response to climate variability and extremes. In view of the considerable costs of gravimeters concerning the infrastructure and measurements, we suggest the strategic deployment of gravimeters at selected sites of hydro-meteorological monitoring networks.


1. Introduction

[2] Water storage is a key state variable of climate and the hydrological cycle, controlling water, energy, biogeochemical and geomorphic processes at and beneath the land surface. Water stored in the subsurface can influence the severity and persistence of hydrologic extremes, such as floods and droughts [Bonan and Stillwell-Soller, 1998; Hong and Kalnay, 2000; Pal and Eltahir, 2001; Schubert et al., 2004; Wu et al., 2002]. Hydrological partitioning at the land surface controls water storage in the subsurface. Water storage, in turn, controls drainage into the channel network, but also feeds back to the climate system and can control air temperature, precipitation and evapotranspiration [Dirmeyer et al., 2009; Entekhabi et al., 1996; Koster et al., 2004b]. In particular, near-surface soil moisture affects climate variables (see Seneviratne et al. [2010] for a review). For example, water stored in the near-surface exhibits a memory for wet and dry anomalies long after these conditions occurred in the atmosphere [Koster and Suarez, 2001; Seneviratne et al., 2006]. Near-surface soil moisture can be strongly coupled to deeper saturated storage through water fluxes by recharge and capillary rise, especially in humid catchments. Hence, there is an interaction in these cases between water storage at larger soil depths and the atmosphere, which might be of particular relevance to long time-scale weather processes [Anyah et al., 2008; Bogaert et al., 2008; Fan et al., 2007; Gutowski et al., 2002; Leung et al., 2011; Miguez-Macho et al., 2007]. For example, deeper storage can stabilize the temporal variations of near-surface soil moisture, giving a positive feedback to the near-surface soil moisture memory and stronger seasonal persistence in evapotranspiration rates, network.
specific humidity and air temperature [Lo and Famiglietti, 2010; Miguez-Macho et al., 2007]. When near-surface soil moisture and deep water storage become decoupled (e.g., during prolonged drought) the dampening effect of wet surfaces on climate variables can disappear. In the context of climate change, it is predicted that heat waves and droughts are likely to become more intense and more frequent [Meehl and Tebaldi, 2004; Schär et al., 2004; Sheffield and Wood, 2008]. Taking into account that water storage anomalies can have a positive feedback on heat waves, and that water storage deficits can amplify these extremes [Hirschi et al., 2011; Vautard et al., 2007], the role of water storage and its long-term recovery becomes critical.

[3] From a hydrological perspective, water storage is the state variable of the hydrological system. A direct relationship between storage and discharge exists and hence water stored in the subsurface is key to understanding the functioning of the catchment [Blöschl and Zehe, 2005; Sayama et al., 2011; Tromp-van Meerveld and McDonnell, 2006; Western et al., 2002] and to classifying catchments [Spence et al., 2010; Wagener et al., 2007]. Knowledge of water stored in the subsurface is important for initializing flood forecasting models [e.g., Aubert et al., 2003; Brocca et al., 2009; Crow et al., 2005; Merz and Plate, 1997; Pfister et al., 2003; Troch et al., 1993] and can improve the prediction of extreme events [e.g., Brocca et al., 2010b; Fennessey and Shukla, 1999; Goodrich et al., 1994; Koster et al., 2004a]. For sub-seasonal and seasonal climate prediction, water storage might be especially useful for regions with strong water storage-climate interaction and long water storage memory [Dirmeyer, 2000; Koster and Suarez, 2001; Seneviratne et al., 2010].

[4] Long-term data on water storage are essential to investigate the complex interaction of climate and hydrologic processes acting at different spatial and temporal scales [Robock et al., 2005; Vinnikov et al., 1996]. However, a strong discrepancy exists between the importance of water storage and the possibility to measure it directly. Traditional methods to measure water storage are point measurements, which are typically sparse, resulting in a low spatial resolution. High spatial variability at small scales [e.g., Teuling and Troch, 2005; Western et al., 2002] put into question the ability of these measurements to represent the hydrologic state of an entire area or watershed [Klemes, 1986]. Additionally, point measurements are limited to either the near surface (soil moisture probes) or the groundwater zone (groundwater level). Furthermore, complex subsurface conditions – fractured bedrock aquifer, high clay or stone content, deep vadose zone – limit the informative value and accuracy of such measurements. Remote sensing techniques provide spatially distributed information on water storage, but the information content is either limited to the very near-surface (techniques using the electromagnetic spectrum (see Wang and Qu [2009] for a review) or to continental scales (GRACE; see Ramilien et al. [2008] for a review). The development of new measurement techniques [Larson et al., 2008; Steele-Dunne et al., 2010; Zreda et al., 2008] (see also Robinson et al. [2008a, 2008b] for reviews) will overcome some limitations providing a more integrative measure of water storage at larger scales. However, the continuous and non-destructive observation of depth-integrated water storage change remains unresolved.

[5] Gravimeter measurements are influenced by water storage due to the Newtonian attraction of masses. Observations of temporal gravity changes provide a direct and depth-integrated measure of water storage change [Creutzfeldt et al., 2010c]. In this context, only the gravimeter residual signal is of interest, which is the gravity signal corrected for well-known processes (i.e., solid Earth tides, ocean tide loading, polar motion and atmospheric mass changes [Hinderer et al., 2007]). Different studies show a high correlation between temporal gravimeter measurements and water storage changes or the associated hydrological processes [e.g., Abe et al., 2006; Bonzat and Sperling, 1995; Boy and Hinderer, 2006; Crossley and Xu, 1998; Harnisch and Harnisch, 2006; Hasan et al., 2006; Kroner and Jahr, 2006; Longuevergne et al., 2009; Meurers et al., 2007; Van Camp et al., 2006]. Hence, temporal gravimeter measurements are used to estimate soil and aquifer properties [e.g., Christiansen et al., 2011; Hokkanen et al., 2007; Jacob et al., 2008; Montgomery, 1971; Pool and Eychaner, 1995; Van Camp et al., 2006]. In recent years, the emphasis has mainly been on the understanding of the physical relationship of gravimeter measurements and hydrology, motivated by the goal to evaluate hydrological models against gravimeter data [Christiansen, 2010; Creutzfeldt et al., 2010b; Hasan, 2009; Jacob, 2009; Longuevergne, 2008; Naujoks et al., 2010].

[6] Beyond the use of gravimeter measurements for hydrological modeling or parameter estimation, high-precision gravimeters can improve our understanding of the hydrological system by directly measuring water storage change without performing any modeling. However, only superconducting gravimeters (SG) [Goodkind, 1999], when combined with absolute gravimeter (AG) measurements [Niebauer et al., 1995], can continuously observe natural water storage changes with sufficient accuracy [Van Camp et al., 2005; Wziontek et al., 2009]. The additional advantage is that long-term data of more than a decade are available from SGS and AGs that have been deployed by geodesists to observe geodynamic processes [e.g., Francis et al., 2004; Nawa et al., 2009; Sato et al., 2006]. Therefore, these gravity time series offer unique opportunities to explore the interaction of climate and hydrologic response.

[7] In this study, we use a 10-year time series (2000–2009) measured by the SG of the Geodetic Observatory Wettzell, Germany. Based on previous studies that showed that gravity residuals are largely influenced by local water storage changes [Creutzfeldt et al., 2010a] and that water change in the aquifer, saprolite (deep vadose zone), soil, and topsoil storage are of similar importance for the signal of the SG Wettzell [Creutzfeldt et al., 2010c], we now use the SG residuals to study the interaction between weather (and its extremes) and the hydrological system. The long time series of SG residuals give us the unique opportunity to study total water storage-climate interactions and the role of water storage in the dynamic response of the hydrological system. Examples of both drought and flood conditions are shown, allowing us to examine the value of gravimeter measurements to understand the hydrologic response to these extreme events.

2. Study Area and Data
2.1. Study Area
The study is conducted in the area surrounding the Geodetic Observatory Wettzell, which is operated by the
The Geodetic Observatory Wettzell is located in the Bavarian Forest in the southeast of Germany. The observatory is located on a mountain ridge, which divides the watershed of the river Regen into the tributaries Weißer Regen and Schwarzer Regen. The observatory directly drains into the headwater catchments Höllensteinbach (catchment size: 5.5 km²) and Augraben (catchment size: 2.0 km²; Figure 1).

The area experiences a temperate climate with mean annual precipitation of 995 mm and a mean annual temperature of 7°C, which correspond to a mean annual potential evapotranspiration of 577 mm. At the observatory, the vegetation mainly consists of grassland, interspersed with a few groves. The land-use of the Höllensteinbach and Augraben catchments is predominantly grassland and forestry. The catchments are underlain by intact and impermeable Biotite-Gneiss of the moldanubian zone of variscan orogenesis [Bayerisches Landesamt für Umwelt, 2007; Raum, 2002]. Around the observatory, the basement is overlain by a zone of loosened crystalline bedrock. A highly fractured zone is covered by the saprolite zone of autochthonous weathered grus. The soils are Cambisols with an estimated depth of 1.25 m, whereby Stagnosols can also be found around the Geodetic Observatory Wettzell [Creutzfeldt et al., 2010c; Grams, 2010; Heim, 2010].

2.2. Hydro-meteorological Data

Precipitation, air temperature, snow height and streamflow were measured continuously since 1947 at the hydropower plant Höllensteinsee (2 km from the observatory), which has a catchment area of 981.0 km². In mid-2006, the climate station suspended its operation (see Table 1). From 2006 to 2009, the time series were extended using the climate data measured at the Geodetic Observatory Wettzell by correcting them for offset and scale. Offset and scale factors were determined based on the calibration period from 01 January 2000 to 31 July 2006. Double mass plots [Kohler, 1949] of cumulative precipitation between the two stations as well as of surrounding precipitation stations (Bühling, Kötzting) revealed homogeneity problems of the Geodetic Observatory Wettzell records during 2002, so this period was excluded from the calibration period. Precipitation records were corrected for the systematic undercatch of unshielded tipping bucket gauges [Richter, 1995]. Groundwater head has been measured in one borehole at the Geodetic Observatory Wettzell since 1998 at a distance of approximately 200 m from the SG. Additionally, one TRIME soil moisture sensor has recorded the soil moisture content at 0.5 m depth since December 2000 (Figure 1). The available data were processed to daily and monthly time series. Monthly potential evapotranspiration was derived based on the Thornthwaite equation [Thornthwaite, 1948]. The Standardized Precipitation Index (SPI) [McKee et al., 1993] was calculated for the time windows 3, 6 and 12 months to provide a simple but reliable drought index [Keyantash and Dracup, 2002]. The Palmer Drought Severity Index (PDSI) [Alley, 1984; Palmer, 1965] was derived from temperature and precipitation records, whereby the available water holding capacity was estimated based on soil water retention measurements [Grams, 2010]. Long-term averages and the associated standard deviations were derived.

In 2007, an extensive hydrological monitoring system was installed around the Geodetic observatory Wettzell, consisting of piezometers, TDR soil moisture probes, a lysimeter, a permanent electrical resistivity profile, a snow
monitoring system and a sharp-crested 90-degree V notch weir [Creutzfeldt et al., 2010a, 2010c]. These data records started in 2007 and, hence, can only be used for demonstration purposes for single events or periods. At the Höllensteinbach headwater catchment (catchment size: 0.5 km²), the streamflow is derived from water level measurements using a piezometer [SEBA, 2008] based on the standard sharp-crested 90-degree triangle weir formula [Dodge, 2001]. These streamflow records with a temporal resolution of 30 min were used to calculate the runoff coefficient, dividing the direct runoff by the precipitation amount for each single runoff event. Events with more than 2 mm of rainfall were defined as events. Events were considered distinct if the dry spell between two events was larger than 2 h. All periods with snow were excluded from the analysis. The end of a runoff event was determined and direct runoff was separated from base flow as suggested by Blume et al. [2007]. In total, 76 runoff events were identified between December 2007 and June 2009.

2.3. Gravimeter Data and Water Storage Changes

[13] At the Geodetic Observatory Wettzell, the change of gravitational acceleration (referred to hereafter as gravity changes) are measured continuously by the SG 029, which is part of the Global Geodynamics Project (GGP) network [Crossley and Hinderer, 2009; Crossley et al., 1999]. Vertical gravity changes are recorded by the SG based on the current of the feedback coil required to levitate the superconducting sphere in a persistent magnetic field [Goodkind, 1999]. SGs are high-precision instruments measuring with a resolution of 0.01 μGal and a noise level of 5 [(nm s⁻²)Hz⁻¹], which corresponds to a noise level of 0.02 μGal during a period of 100 s [Van Camp et al., 2005]. Absolute gravity (AG) measurements were conducted with different FG5 gravimeters (FG5-101, FG5-227, FG5-301) approximately two times per year, resulting in 32 absolute gravity measurements for the whole study period. Based on the AG measurements, the scale factor of the SG was determined with a relative precision of better than 0.1%. The almost linear SG drift, which did not exceed 5.0 μGal/yr, was determined with an accuracy of better than 0.5 μGal/year by absolute gravity measurements [Wziontek et al., 2009]. The gravity effect of tides, polar motion and atmospheric mass changes were modeled and removed from the gravimeter signal [Hinderer et al., 2007; Klügel and Wziontek, 2009; Neumeyer, 2010]. The accuracy of the remaining signal – the gravity residuals – can be estimated roughly with 0.1 μGal for short term variations (1–30 days) and 0.5 μGal for interannual variations. The gravity residuals are considered to be largely influenced by hydrological mass variations. For example, for the Geodetic Observatory Wettzell, water storage changes were independently estimated based on lysimeter measurements and soil physical measurements. The calculated hydrological gravity response explained 97% of the variation of the gravity residuals measured by the SG, demonstrating that most of the gravity residuals are caused by local water storages [Creutzfeldt et al., 2010a].

[14] Different approaches exist to calculate the water storage change from gravity residuals. A simple and straightforward approach is the Bouguer approximation, where the gravitational acceleration g is calculated as 

\[ g = 2\pi G \Delta \rho z, \]

where \( G \) is the universal gravitational constant (Nm²/kg²), \( \Delta \rho \) is the density change (kg/m³) and \( z \) (m) the thickness of an infinitely horizontal slab. Applying the Bouguer approximation we find that a 1-m height change of water causes a gravity change of 42 μGal. However, topography determines the distribution of water masses in space and also the vertical (re-) distribution of mass changes above/below the sensor has an effect on gravity measurements. In this study, we therefore derived the total water storage change from gravity residuals based on the following scheme. A hydrological and geophysical model were coupled during the inversion process [Ferré et al., 2009; Himell et al., 2010] outlined in the study of Creutzfeldt et al. [2010b] by extending the model period to the study period of 10 years (blue line in Figure 2). In this framework, a conceptual hydrological model calculated the water storage change in the snow, soil, saprolite and groundwater storage using daily precipitation, daily reference evapotranspiration and snow height as input data. From these water storage changes...
changes the gravity response was calculated based on the extended point mass equation [Leirão et al., 2009] by considering the spatial distribution of water storage change along the topography in a nested discretization domain [Creutzfeldt et al., 2008]. The model parameters were automatically calibrated based on the Generalized Likelihood Uncertainty Estimation (GLUE) method [Beven and Binley, 1992] using the SG residuals as the only calibration constraint. This model can be used to simulate the water storage change but the results would be associated with uncertainty. To explore the SG data directly, i.e., to translate SG residuals into water storage change, we estimated the regression coefficient between measured SG residuals and modeled water storage change (Figure 3). The regression was then used to directly derive the water storage change from gravity residuals. The resulting time series is used throughout this study (green line in Figure 2). Calculating water storage changes directly from gravity residuals using the regression coefficient means that only the amplitude of the time series is influenced by model uncertainties but not the temporal characteristics of the measured time series. The slope

![Figure 2. Measured gravity residuals and corresponding water storage change for the study period from 2000 to 2009. The beginning of the study period is set as the reference point to express change. (top) Gravity residuals measured by the superconducting gravimeter (SG) and the absolute gravimeter (AG). AG measurements are displayed with the total uncertainty. (bottom) Water storage change derived from a hydrological model (blue line) with corresponding uncertainty band (gray bands [Creutzfeldt et al., 2010b]). Water storage change directly derived from the SG residuals (green line) estimated based on the regression coefficient between measured SG residuals and modeled water storage change (Figure 3).](image-url)

![Figure 3. Observed gravity residuals versus modeled water storage change. Black line is the linear regression function and points are the bin averages of three adjacent points. Note that 25.68 mm/μGal equals 39 μGal per 1 m water storage change.](image-url)
Figure 4. Meteorological forcing and corresponding water storages. The box and whisker plots of the seasonal variation of (a) air temperature (underlying time series 1947–2009), (c) precipitation (underlying time series 1947–2009), (e) snow height (underlying time series 1947–2009), (g) groundwater head (underlying time series 2000–2009), (i) total water storage change (underlying time series 2000–2009). Box and whisker plot: red line: median; blue box: 25th and 75th percentile; whiskers: lower and upper extreme values within 1.5 times the interquartile range (IQR); red cross: points drawn as outliers which are outside 1.5 × IQR. The monthly time series of (b) temperature and the long-term anomalies, (d) precipitation and the long-term anomalies, (f) snow, (h) soil moisture and groundwater, and (j) total water storage change. Gray bars in Figures 4b and 4d are anomalies of long-term monthly values.
between winter water storage change and the gravity residuals equals 39 \( \mu \)Gal per 1 m water storage change instead of 42 \( \mu \)Gal per 1 m water storage change as predicted by the Bouguer approximation. The difference arises because the Bouguer assumption of homogenous, flat lying and laterally infinite layer is violated by the mid-mountainous topography as well as the near-field settings (concrete foundation of the SG, the umbrella effect of the building) around the SG Wettzell [Creutzfeldt et al., 2008].

3. Water Storage Changes From 2000 to 2009

[15] Figure 4 shows the local water storage dynamics in terms of soil moisture, groundwater, and total water storage, in comparison to air temperature, precipitation and snow height. Seasonal variation of all measured variables as well as a clear response of water storages to meteorological forcing can be observed. Temperature and precipitation values peak in July and minimum temperature values are observed in January (long-term average monthly values). The time series of water storage change is relatively smooth in comparison to the groundwater level or the soil moisture, which is reflected in the absence of "outliers" of the box plot (red crosses). The water storage dynamics exhibit a sinusoidal character, whereas groundwater heads show the typical nonlinear recession after the sharp rise. The maximum total water storage change and groundwater head occur during March and April, following a maximum snow height in February. This coherence indicates that snowmelt is a controlling factor for subsurface storage recharge of this area.

[16] The SG derived water storage change does not show a strong reaction to the snow load because snow mass accumulates around the SG as well as above the SG sensor on the roof of the building, basically canceling out any effect on gravity residuals [Creutzfeldt et al., 2008]. When the snowmelts, water mass redistributes to deeper zones and becomes observable to the SG, causing an increase of total water storage change. For example, during the winter of 2005/06, with exceptional high snow loads damaging premises and buildings and even causing roofs to collapse [LWF, 2006; Pinto et al., 2007], snow was observed from December 2005 to April 2006 and the maximum recorded daily snow height was 0.9 m. However, only with the beginning of snowmelt, do water masses redistribute into the sub-surface and become observable to the gravimeter. This is reflected in a water storage increase of nearly 300 mm over the snowmelt period (Figure 5).

[17] Water storage changes of up to 284 mm between the beginnings of each consecutive year were observed (aWSC in Table 2). This shows that the simplifying assumption of water storage variations averaging to zero at the annual scale is not justified. Especially during extreme years, the storage term is a significant component of the annual water balance, which is in line with the study of Hudson [1988]. The temporal occurrence of the seasonal maximum/minimum values varies considerably between years (Table 2). The seasonal water storage change, i.e., the difference between the maximum and minimum water storage in each year (rWSC in Table 2), ranges between 193 and 399 mm, with 233 mm on average, highlighting the importance of the subsurface as a storage of water and energy.

[18] Table 3 compares the estimated water storage change to water storage dynamics of other humid temperate catchments in the context of their physio-geographical settings. Water storage changes were derived from many different methods, ranging from distributed point measurements and tracer studies to recession analyses. The water storage change in the Höllensteinbach catchment lies within the upper range of the reported storage changes. A stepwise multiple linear regression showed that only the potential evapotranspiration (PET) has a statistically significant predictive capability for the annual water storage dynamics. The correlation coefficient was estimated for the PET and the mean water storage change (mWSC) and for the PET and the range of the annual water storage (rWSC) to be 0.85 (\( p < 0.01 \)) and 0.84 (\( p < 0.01 \)), respectively. A high correlation of rWSC and PET indicates that the water storage variations are strongly dependent on the available energy. This is because all sites are located in mid latitudes with a strong seasonality of evapotranspiration. However, further conclusions from the inter-comparison are to be drawn with care as different methods were used for estimating water storage change and potential evapotranspiration.

[19] For the Geodetic Observatory Wettzell, we can also observe a strong relationship of available energy and water storage change. Evapotranspiration is significantly correlated to monthly change of water storage (correlation

<table>
<thead>
<tr>
<th>Year</th>
<th>P</th>
<th>PET</th>
<th>P-PET</th>
<th>aWSC</th>
<th>rWSC</th>
<th>MAX</th>
<th>MIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1218</td>
<td>641</td>
<td>577</td>
<td>-</td>
<td>210</td>
<td>Mar 00</td>
<td>Sep 00</td>
</tr>
<tr>
<td>2001</td>
<td>1248</td>
<td>605</td>
<td>643</td>
<td>92</td>
<td>193</td>
<td>Apr 01</td>
<td>Aug 01</td>
</tr>
<tr>
<td>2002</td>
<td>1298</td>
<td>638</td>
<td>660</td>
<td>19</td>
<td>237</td>
<td>Mar 02</td>
<td>Jul 02</td>
</tr>
<tr>
<td>2003</td>
<td>751</td>
<td>650</td>
<td>101</td>
<td>84</td>
<td>399</td>
<td>Jan 03</td>
<td>Sep 03</td>
</tr>
<tr>
<td>2004</td>
<td>931</td>
<td>588</td>
<td>343</td>
<td>-284</td>
<td>198</td>
<td>Apr 04</td>
<td>Sep 04</td>
</tr>
<tr>
<td>2005</td>
<td>999</td>
<td>595</td>
<td>404</td>
<td>143</td>
<td>236</td>
<td>Apr 05</td>
<td>Nov 05</td>
</tr>
<tr>
<td>2006</td>
<td>1098</td>
<td>652</td>
<td>446</td>
<td>-132</td>
<td>201</td>
<td>Apr 06</td>
<td>Dec 06</td>
</tr>
<tr>
<td>2007</td>
<td>1061</td>
<td>662</td>
<td>399</td>
<td>68</td>
<td>140</td>
<td>Mar 07</td>
<td>Aug 07</td>
</tr>
<tr>
<td>2008</td>
<td>929</td>
<td>639</td>
<td>290</td>
<td>157</td>
<td>193</td>
<td>Apr 08</td>
<td>Oct 08</td>
</tr>
<tr>
<td>2009</td>
<td>1013</td>
<td>653</td>
<td>360</td>
<td>-78</td>
<td>-</td>
<td>Apr 09</td>
<td>-</td>
</tr>
</tbody>
</table>
**Table 3. Water Storage Changes and Catchment Characteristics of the Headwater Catchment Höllensteinbach in Comparison to Other Catchments**

<table>
<thead>
<tr>
<th>Storage Estimation Method</th>
<th>Input-output dynamics of natural tracers</th>
<th>Distributed measurements of soil moisture and groundwater</th>
<th>Coupled modeling and distributed measurements of soil moisture</th>
<th>Distributed measurements of soil moisture</th>
<th>Water balance</th>
<th>Recession analysis</th>
<th>Recession analysis</th>
<th>Recession analysis</th>
<th>Temporal gravity measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (km²)</td>
<td>30</td>
<td>0.0063</td>
<td>0.38</td>
<td>0.02</td>
<td>0.41</td>
<td>3.31</td>
<td>8.7</td>
<td>10.55</td>
<td>0.5</td>
</tr>
<tr>
<td>Altitude (m)</td>
<td>405</td>
<td>133</td>
<td>2073</td>
<td>1500</td>
<td>222</td>
<td>816</td>
<td>529</td>
<td>540</td>
<td>592</td>
</tr>
<tr>
<td>Range (m)</td>
<td>620</td>
<td>20</td>
<td>118</td>
<td>100</td>
<td>57</td>
<td>268</td>
<td>419</td>
<td>397</td>
<td>71</td>
</tr>
<tr>
<td>Geology</td>
<td>Granite</td>
<td>Granodiorite</td>
<td>Basalt/Rhyolite</td>
<td>Granodiorite</td>
<td>Granodiorite</td>
<td>Sediments</td>
<td>Sediments</td>
<td>Sediments</td>
<td>Gneiss</td>
</tr>
<tr>
<td>P (mm)</td>
<td>1000</td>
<td>1050</td>
<td>866</td>
<td>641</td>
<td>1250</td>
<td>1459</td>
<td>2553</td>
<td>2599</td>
<td>1055</td>
</tr>
<tr>
<td>PET (mm)</td>
<td>442b</td>
<td>550f</td>
<td>795d</td>
<td>3653.7e</td>
<td>96c</td>
<td>501</td>
<td>566</td>
<td>488</td>
<td>632</td>
</tr>
<tr>
<td>Q (mm)</td>
<td>550</td>
<td>317</td>
<td>523</td>
<td>96</td>
<td>360</td>
<td>1063</td>
<td>1987</td>
<td>2111</td>
<td>667</td>
</tr>
<tr>
<td>mWSC (mm)</td>
<td>31</td>
<td>65</td>
<td>135</td>
<td>41</td>
<td>529</td>
<td>142f</td>
<td>1363e</td>
<td>42.2e</td>
<td>288</td>
</tr>
<tr>
<td>rWSC (mm)</td>
<td>62</td>
<td>120</td>
<td>205</td>
<td>66</td>
<td>928</td>
<td>205f</td>
<td>189h</td>
<td>75h,i</td>
<td>517</td>
</tr>
<tr>
<td>Ref.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

*aTable shows the size of the drainage area (A; km²), the approximated mean altitude (Altitude; m), the altitude range (Range; m), the dominant geology (Geology), the mean annual precipitation (P; mm), the mean potential evapotranspiration (PET; mm), the annual discharge (Q; mm), arithmetic averages of the daily storage changes (mWSC; mm), the range of water storage derived from the minimum and maximum observed daily storage values (rWSC; mm) and the corresponding reference (Ref.). (Table has been modified and extended according to McNamara et al. [2011].)

bAccording to Jenkins et al. [2007].
cAccording to Rodhe et al. [1996].
dAccording to Hanson et al. [2001].
eAccording to Kelleners et al. [2009].
fAccording to Falcone et al. [2010].
gEstimated from mean annual discharge.
hEstimated from maximum and minimum observed discharge.
iJ. W. Kirchner, personal communication, 2011.

References: 1, McNamara et al. [2011]; 2, Teuling et al. [2010], and 3, Kirchner [2009].
The coefficient of precipitation (P) has no significant predictive capability for the monthly change of water storage (Figures 6a and 6b). For precipitation minus evapotranspiration (P-PET) a high but positive correlation coefficient can be observed (correlation coefficient = 0.70; p < 0.01; Figure 6c). A wavelet-based semblance analysis [Cooper and Cowan, 2008] was performed to investigate the similarities of precipitation, evapotranspiration, precipitation minus evapotranspiration and water storage change as a function of both time and scale. In Figures 6d and 6e, red (blue) values indicate a high correlation (anti-correlation) between the respective time series. For monthly precipitation and change of water storage, we see a high positive correlation at scales larger than ~18 month, whereas at smaller scales the relationship is complex and for some scales and time periods both variables can even be anti-correlated (Figure 6d). Contrary to this, potential monthly evapotranspiration and change of water storage is anti-correlated at nearly all scales and times (Figure 6e), indicating a straightforward relationship of potential evapotranspiration and water storage dynamics. The same is true for precipitation minus evapotranspiration where we observe a positive correlation at nearly all scales and times (Figure 6f).

3.1. Water Storage and Runoff Response to Extreme Precipitation

[21] A catchment has a maximum storage capacity that is a function of the thickness of porous material and its porosity and compressibility. The amount of available storage is a critical factor when predicting, for example, catastrophic flooding. Therefore, it would be useful to be able to measure this storage capacity directly. We suggest that an estimation of this capacity is a recurrent, maximum storage range measured in a long time series of water storage change (Figures 2 and 4). A similar approach has been suggested at the global scale using water storage variations observed by the GRACE satellites [Reager and Famiglietti, 2009]. Based on the maximum storage capacity reached during the study period (SMAX) of 517 mm based on daily data, the daily storage deficit SDEF can be defined as

\[ S_{DEF}(t) = S_{MAX} - S(t) \]

where S(t) is the water storage change for each day t. From this storage deficit, the flood potential amount, FPOT, can be derived as

\[ F_{POT}(t + \Delta t) = \sum_{i=t}^{t+\Delta t} P(i) - S_{DEF}(t) \]

where P(i) is the daily precipitation volume. In order to use the storage deficit in a predictive mode for flood potential assessments, we set the time interval \( \Delta t \) in equation (2), i.e., the forecast period, to seven days assuming that the total precipitation amount can be predicted reasonably well for a one week period. Thus, the flood potential amount is based on the system state of the previous week and the predicted weekly precipitation input.

[22] A significant correlation can be observed between flood potential amount and the runoff coefficient observed in...
the headwater catchment of the Höllensteinbach (Figure 7). This positive correlation indicates that the proportion of direct runoff to total precipitation increases with an increasing flood potential amount. This relationship is stronger than using the precipitation amount only. Hence, the daily water storage deficit improves the prediction of effective rainfall for the week to follow. Viewed in this way, SG measurements could provide an additional piece of information that could be used together with real-time weather forecasts for short-to-medium term flood prediction.

[23] The critical role of the subsurface wetness condition on the runoff generation process was observed by other studies, using point measurements of soil moisture or groundwater head [Latron and Gallart, 2008; Peters et al., 2003; Troch et al., 1993]. Figure 8 shows that the runoff coefficient is related to pre-event water storage measured by soil moisture probes and the SG. For near-surface soil moisture, differences in the strength of the relationship between soil moisture and runoff coefficient can be observed between the three TDR arrays (Figures 8a-8c) even though topsoil moisture was estimated based on 21 to 45 TDR probes in each cluster. The relationship between the runoff coefficient and total water storage change measured by the gravimeter is stronger than that seen with near-surface soil moisture measurements (Figure 8d). This highlights the importance of deeper water storage for the runoff generation process for the humid catchment Höllensteinbach.

3.2. Water Storage Response to and Recovery From Drought

[24] During the summer of 2003, Western and Central Europe were severely affected by a heat wave and a corresponding drought [De Bono et al., 2004]. For the Bavarian forest, precipitation was up to 70% below average and monthly temperature was up to 5°C above average [Gietl, 2004]. According to the drought indices SPI and PDSI, 2003 can be classified as a severe to extreme drought depending on the index and time window used (Figure 9). We see good overall agreement between water storage change and the drought indices. For the comparison it has to be taken into account that PDSI is derived from precipitation and temperature data, whereas SPI was designed to reflect the precipitation anomalies for different time scales. The extreme drought in combination with the heat wave caused a decrease of total water storage by 400 mm, which is the largest decline observed over the 10-year period. The average annual depletion in stored water for the remaining 9 years studied was approximately 180 mm. Time differences between the different indices exist, e.g., the minimum value of the SPI 3-month is reached in May 2003, whereas the SPI 12-month shows its strongest decrease in December 2003. The lowest value of PDSI occurs between September and December 2003 and correlates well with the total water storage change (Figure 9).
Following 2003, the water storage required several years to recover to the pre-drought conditions (Figure 4). An Empirical Mode Decomposition (EMD) [Huang et al., 1998; Rilling et al., 2003] was performed to investigate the inter-annual impact of the 2003 drought on the hydrological system. EMD allows analyzing nonlinear and non-stationary time series. A time series is empirically and adaptively decomposed into a finite number of intrinsic mode functions based on the local characteristics of the data. The remaining residual value represents the long-term dynamics of meteorological forcing, system states, and outputs [Wu et al., 2007]. All of the system states and outputs (water storage, groundwater head, soil moisture and discharge) exhibit positive trends after the 2003 drought, implying an impact of the drought throughout the hydrological system (Figure 10). This is in agreement with the SPI 12-month average, which

**Figure 9.** (a) Standardized Precipitation Index (SPI) and (b) Palmer Drought Severity Index (PDSI) in comparison to total water storage change. Drought levels according to McKee et al. [1993] for SPI and according to Alley [1984] for PDSI.

[25] Following 2003, the water storage required several years to recover to the pre-drought conditions (Figure 4). An Empirical Mode Decomposition (EMD) [Huang et al., 1998; Rilling et al., 2003] was performed to investigate the inter-annual impact of the 2003 drought on the hydrological system. EMD allows analyzing nonlinear and non-stationary time series. A time series is empirically and adaptively decomposed into a finite number of intrinsic mode functions based on the local characteristics of the data. The remaining residual value represents the long-term dynamics of meteorological forcing, system states, and outputs [Wu et al., 2007]. All of the system states and outputs (water storage, groundwater head, soil moisture and discharge) exhibit positive trends after the 2003 drought, implying an impact of the drought throughout the hydrological system (Figure 10). This is in agreement with the SPI 12-month average, which

**Figure 10.** The long-term dynamics of (a) precipitation and air temperature, (b) SPI 6 months and SPI 12 months, (c) soil moisture and groundwater head, and (d) total storage change and streamflow in response to the drought and heat wave of 2003. The long-term responses are the residuals derived from an Empirical Mode Decomposition.
showed that the drought started in August 2003 and lasted until March 2005 (Figure 9). Soil moisture in the upper soil layer also showed a strong decrease during the summer of 2003, but rainfall in the beginning of October 2003 (87 mm/10 days) brought soil moisture nearly back to pre-drought values (Figure 10c). From the beginning of 2005, the observed increase of 0.01% in soil moisture is below the detection accuracy of soil moisture probes. For groundwater, the maximum depletion occurred one year later in October 2004 (Figure 4). This decline delayed by one year is a characteristic response of groundwater storage to droughts [Entekhabi et al., 1996] and shows that groundwater head measurements might not be a good indicator for timely drought identification. The recovery of the groundwater took several years with the groundwater storage refilled in April 2007 (Figure 10c), whereas for total water storage the recovery lasted longer, approximately until the end of 2009. Surprisingly, the long-term characteristics of total water storage agree well with those of discharge (Figure 10d), where discharge can be thought of as the most integrative measure of the hydrological system. This suggests that total storage and groundwater head represent different drought recovery processes and that the vadose zone storage is an important storage component within the basin. Hence, it can be difficult to identify the impact of weather extremes on the hydrological state solely by soil moisture sensors and groundwater head measurements. Groundwater head showed a delayed response to the drought and near surface soil moisture recovered quickly.

[26] Another example of a water storage anomaly is given during the winter of 2006/07. From September 2006 to February 2007, the mean monthly temperature and evapotranspiration were well above and precipitation and snow height were below the long-term mean for most months (Figure 11). According to the SPI, this period can be classified as a mild to moderate drought, and according to the PDSI, this period was a moderate to severe drought (Figure 9). High evapotranspiration due to the early start of the vegetation period, in combination with the absence of snow as the major water recharge factor for this area, caused an increase of total water storage of only about 100 mm, where for the other years, the mean increase of water storage amounted to 185 mm. It is difficult to say whether this period is a drought due to a lack of precipitation or caused by the extraordinarily warm weather, but most likely, it is a combination of both. Interestingly, soil moisture and groundwater head characteristics do not deviate significantly from those seen in other winters (Figures 11 and 4). Furthermore, near-surface electromagnetic methods to estimate soil moisture usually fail during winter conditions because of soil frost. As a result, it would have been difficult to identify the winter storage deficit of 2006/07 without gravity data. Hence, we conclude that the integrative quantity provided by the gravi-meter is a uniquely valuable measurement to directly observe the effect of droughts on the hydrological system.

4. Discussion

[27] In this study, we analyzed the 10-year gravimeter time series of the Geodetic Observatory Wettzell to understand the link between water storage changes and hydro-meteorology with a particular focus on floods and droughts. We identified a maximum storage threshold and assumed that this threshold equals the water storage capacity of the area surrounding the gravimeter. This assumption allowed us to derive the time series of the flood potential from the SG measurements. We found a significant correlation between the flood potential and the runoff coefficient of the headwater catchment Höllensteinbach. Ten years is a relatively short time, and time series of several decades would be needed to draw statistically significant conclusions on the runoff generation during floods, so that these results await further confirmation at other sites. Additionally, the signal might ramp up at a similar yearly level because at the end of winter, water release through evapotranspiration pulls back the signal so that the observed threshold might be caused by the balance between wetting up the catchment and available energy. Over longer time spans, higher water storage may be encountered, leading to correspondingly higher storage capacity. However, depth-integrated gravity measurements offer the only known possibility to determine the storage capacity. In this study, we demonstrate the potential of gravity measurements to determine this storage capacity, which helps to understand the response of the hydrological system and to predict floods. The integration of water storage change to initialize state variables [e.g., Brocca et al., 2009; Bronstert et al., 2012; Goodrich et al., 1994] or to update state variables by assimilation techniques [e.g., Aubert et al., 2003; Crow et al., 2005] in a sophisticated modeling framework could be of high interest to improve flood forecasting. However, it has to be considered that, apart from the importance of the catchment state, the spatiotemporal patterns of rainfall mainly control the magnitude
of flood [Bronstert and Bárdossy, 2003; Merz and Plate, 1997; Wood et al., 1990].

[25] Water storage dynamics measured locally by a gravimeter are compared to the response of the headwater catchment Höllensteinbach. This raises the question whether measurement systems with such different support scales can be compared consistently. For example, the headwater catchment of the Höllensteinbach has a size of 0.5 km² (Figure 1). Assuming that the upper 10 m take an active role in the hydrological cycle, this would result in a volume of $5 \times 10^{-3}$ km$^3$ for which the statement is made. The response of a gravimeter is a function of mass (re-) distribution in space within its sampling volume and also depends on the topography and subsurface conditions. For the SG Wetzell, simulation showed that 80% of the signal is generated in a rectangle with a side length of approximately 100 m, assuming a mass change of 1 m in a layer distributed at a depth of 5 m along the topography [Creutzfeldt et al., 2008]. This would result in a sampling volume of $5 \times 10^{-3}$ km$^3$. The soil moisture values used in this study were obtained by 0.3 m long 3-rod TDR probes (sampling volume of $8 \times 10^{-11}$ km$^3$; e.g., Ferré et al., 1998). Using between 21 to 45 probes in an array, the sampling volume can be increased to roughly $3 \times 10^{-11}$ km$^3$. Comparing gravimeters and soil moisture probes, the sampling volume of the gravimeter is 6 orders of magnitude larger. Considered as a fraction of the size of the watershed, the gravimeter sample volume is two orders of magnitude smaller, whereas the sample volume of the soil moisture probe network is eight orders of magnitude smaller. For point measurements, the selection of a representative measurement location is critical to estimate the antecedent hydrological state of the system [Brocca et al., 2010a; Zehe et al., 2010]. This study indicates that increasing the measurement scale beyond the point scale can provide a more representative measure of the overall catchment state and may help to use storage measurements as valuable input for runoff prediction.

[29] Droughts affect ecosystems but they also have a strong impact on the economy. Many different drought definitions have been developed to describe the different impacts of the drought (see Heim [2002] for a review). No single index adequately captures the intensity and severity of a drought. As highlighted by Heim [2002], any drought monitoring system or drought index must ‘address the total environmental moisture status’. In comparison to point measurements and drought indices, our findings show that depth-integrated measurements of water storage change give a more complete view of the nature of the hydrologic system in response to the drought. The characteristics of the total storage can differ from single storage components and suggest that deeper water storage may be an important storage within the basin. Gravimeters provide a measure of the ‘total environmental moisture status’ which might be useful for agriculture and water resources planning. Beyond the point scale, the absolute amount of water stored in an area can be estimated by natural conservative tracers [e.g., McGuire and McDonnell, 2006; Soulsby et al., 2009] or magnetic resonance sounding [e.g., Lubczynski and Roy, 2004; Müller-Petke et al., 2011], so that the absolute water storage amount and its variation can be derived only by combining the different methods [McNamara et al., 2011; Pfeffer et al., 2011]. This study shows a good agreement between total water storage change and two frequently used drought indices, the SPI and the PDSI. The evaluation of gravimeter measurements as a “drought measure” is difficult due to the absence of a unique definition of drought and a universal drought index. Comparing gravimeter time series to different indices shows that beyond simply mimicking the temporal pattern, direct water storage measurements give insight into the hydrological system and help to understand and predict the system response.

[30] Based on the time series of water storage change, we observed a long-term recovery of water stored in the vadose and saturated zone in response to the drought. This long-term recovery of water storages lasted several years and might be critical when taking feedback mechanisms between water storage and drought into account. For example, low water storage reduces local evaportranspiration rates, which leads to a higher temperature and this can reduce precipitation, which in turn amplifies the drought [Atlas et al., 1993; Hong and Kallay, 2000; Shukla and Mintz, 1982]. Several studies showed that near-surface soil moisture “remembered” the wet or dry condition in response to weather anomalies over weeks to months [e.g., Koster and Suarez, 2001; Seneviratne et al., 2006]. Other studies focused on the near-surface memory and its relationship to groundwater level persistence [Miguez-Macho et al., 2007]. Based on model experiments, Lo and Famiglietti [2010] for example showed that feedback mechanisms between groundwater level and near-surface soil moisture could cause very different overall system responses depending on the depth of the groundwater below the land surface. Little is known about how deep soil moisture contributes and comes into play when groundwater is replaced, mainly because deep soil moisture is impossible to measure. High-precision gravimeters do not distinguish between near-surface and deep water storage changes but record an integral signal, so that also water storage changes in the deeper vadose zone can be monitored [Creutzfeldt et al., 2010b; Jacob et al., 2010]. We found that water storage was strongly depleted by the drought and heat wave of 2003 and that it showed an inter-annual long-term memory at the Geodetic Observatory Wetzell. Streamflow data indicate that the water storage recovery requires several years at the watershed scale, too. It is most likely that this long-term recovery of water storages also exists at other locations or at larger scales, but due to the lack of an appropriate monitoring technique, it has not been possible to observe this long-term impact.

[31] It would be useful to add high-precision gravimeters to our hydro-meteorological monitoring networks so that we can understand the impacts of climate extremes and water storage conditions on the hydrologic responses. Very high costs of acquisition and operation of SGs, as well as the need for a very extensive knowledge of installation, measurement and processing, hamper the widespread deployment of continuous gravity measurements. The strategic deployment at selected locations can help to improve our understanding of the environmental system; it can be particularly helpful to test hypotheses or to evaluate environmental indices or models. As highlighted by Legates et al. [2011], soil moisture acts as a unifying theme that integrates different disciplines like meteorology, climatology, geomorphology and pedology, biogeography, and hydrology. By integrating over the saturated and unsaturated zone, gravimetry adds
hydrogeology to this list of different disciplines. Geodesy and geophysics are needed to understand the gravimetric signal. Hence, gravimeters may not only provide a direct measure of water storage change by integrating over depth and over a larger area, but can be a unifying instrument and another building block for a truly interdisciplinary Earth science community.

5. Conclusion

[32] The 10-year time series of total water storage changes derived from high-precision temporal gravimeter observations, revealed distinct seasonal and inter-annual dynamics in response to the meteorological conditions, in particular potential evapotranspiration. The integrative nature of gravity observations, covering all water storage compartments at scales on the order of 1–100 m, is shown to be most appropriate to directly observe the effect of droughts on the hydrological system, including multiyear water storage depletion. The pre-event system state inferred from gravimetry improves the prediction of runoff coefficients and flood response compared to point-scale soil moisture measurements. Hence, high-precision gravimetry allows us to observe total water storage changes at the event, interannual and intra-annual time scales. High costs and sophisticated operation requirements hamper the widespread deployment of high-precision gravimeters. Therefore, we suggest a strategic deployment of continuous gravity measurements at selected locations as a reference for total water storage changes.

[33] Acknowledgments. The authors would like to thank Heiko Thoss and Thomas Klügel for their help in the installation and maintenance of the extensive hydrological measurement system. This work has been supported by Deutsche Forschungsgemeinschaft (German Research Foundation) within the Priority Program SPP 1257 “Mass transport and mass distribution in the system Earth,” project TASMAGOG. We want to thank the editors and reviewers for their comments that significantly improved the study.

References


Creutzfeldt, B., A. Güntner, T. Klügel, and H. Wziontek (2008), Simulating the influence of water storage changes on the seasonal gravity signal of the Geodetic Observatory Wettzell, Germany, Geophysics, 73(6), WA95, doi:10.1190/1.2929250.


