Assimilation of GRACE Terrestrial Water Storage into a Land Surface Model: Evaluation and Potential Value for Drought Monitoring in Western and Central Europe

Bailing Li [a,b], Matthew Rodell [b], Benjamin F. Zaitchik [c], Rolf H. Reichle [d], Randal D. Koster [d], Tonie M. van Dam [e]

[a] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA

[b] Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[c] Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD 21218, USA

[d] Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[e] Department of Physics and Material Sciences, University of Luxembourg, L-1359 Luxembourg

Corresponding author:

Bailing Li
Bailing.li@nasa.gov
Telephone: 1-301-286-6020
Fax: 1-301-614-5808

Abstract

A land surface model's ability to simulate states (e.g., soil moisture) and fluxes (e.g., runoff) is limited by uncertainties in meteorological forcing and parameter inputs as well as inadequacies in model physics. In this study, anomalies of terrestrial water storage (TWS) observed by the Gravity Recovery and Climate Experiment (GRACE) satellite mission were assimilated into the NASA Catchment land surface model in western and central Europe for a 7-year period, using a previously developed ensemble Kalman smoother. GRACE data assimilation led to improved runoff correlations with gauge data in 17 out of 18 hydrological basins, even in basins smaller than the effective resolution of GRACE. Improvements in root zone soil moisture were less conclusive, partly due to the shortness of the in situ data record. In addition to improving temporal correlations, GRACE data assimilation also reduced increasing trends in simulated monthly TWS and runoff associated with increasing rates of precipitation. GRACE assimilated root zone soil moisture and TWS fields exhibited significant changes in their dryness rankings relative to those without data assimilation, suggesting that GRACE data assimilation could have a substantial impact on drought monitoring. Signals of drought in GRACE TWS correlated well with MODIS Normalized Difference Vegetation Index (NDVI) data in most areas. Although they detected the same droughts during warm seasons, drought signatures in GRACE derived TWS exhibited greater persistence than those in NDVI throughout all seasons, in part due to limitations associated with the seasonality of vegetation.

1. Introduction

Seasonal and interannual variability in terrestrial water storage (TWS) is of critical interest in water resource analysis and seasonal hydrological forecasts because TWS—which includes soil moisture, groundwater, surface water and snow—is an important hydrological indicator in its own right: volume of water stored in snowpack or groundwater, for example, reflects present hydrological conditions and can be used to infer the potential for future hydrological stress. TWS is also important because of its role in other aspects of the hydrological cycle. Its status can affect infiltration rates and subsurface flow, with associated impacts on runoff and recharge rates. TWS anomalies can also affect the hydrological cycle through soil moisture feedbacks on the atmosphere. One of the important aspects of TWS is its unique dynamics. Soil moisture and groundwater are low-pass filters on the terrestrial hydrological cycle that gradually remove high frequency variability associated with atmospheric forcing as depth increases (Eltahir and Yeh, 1999; Wu et al. 2002). This dynamic means that TWS acts as a “memory” component of the terrestrial hydrological cycle, with implications for land-atmosphere interactions (Koster and Suarez, 2001) and predictability in certain regions (Dirmeyer 2000; Dirmeyer et al., 2009; Koster et al., 2000b; Koster et al., 2010a).

Interactions among components of TWS not only re-distribute water spatially but also increase the complexity of the hydrological cycle. Groundwater, which accounts for a major part of TWS (Rodell and Famiglietti, 2001; Rodell et al., 2007; Yeh et al. 2006), can contribute substantially to stream flow in wet climates (Eltahir and Yeh, 1999). This connection, combined with the long memory of groundwater variability, means that accurate information on groundwater can contribute significant skills to seasonal river discharge forecasts (Birkens and Van Beek, 2009). Groundwater can also move upward to increase soil wetness through capillary lift or act as a sink to receive excess soil moisture from the land surface (Schaller and Fan, 2009). As appreciation for these processes has grown, an increasing number of land surface models have been developed to account for the impact of groundwater on near surface processes (e.g., Koster et al., 2000a; Niu et al., 2007; Miguez-Macho et al., 2007; Yeh and Eltahir, 2005). Including groundwater in a land surface model enables a more complete simulation of the terrestrial water cycle, but it also subjects the modeled states to additional uncertainties associated with the added physical processes and parameters. For instance, due to lack of global-scale groundwater measurements, most models depend on calibration to obtain the temporal variability and dynamic range of groundwater tables, which may not represent the interactions realistically, especially under extreme wet or dry conditions.

Precipitation data sets are a major source of uncertainty for land surface modeling, and their impacts on modeled states and fluxes may differ depending on seasons and climates (Fekete et al., 2004; Gottschalck et al., 2005). Great uncertainty also exists in model physics such as surface runoff algorithms which are often derived from empirical relationships (Koster et al., 2000a; Niu et al., 2005; Schaake et al., 1996). Stream flow is governed in varying degrees by topography, rainfall intensity, and soil wetness, making it a difficult process to simulate efficiently. Due to differences in model physics and parameter values, estimates by various land surface models exhibit large discrepancies even when models are run using identical forcing data (Mitchell et al., 2004). The combination of uncertainties in forcing, input parameters and model physics has led to dramatically different predictions for runoff trends in response to future climate changes (Hoerling et al., 2009).

The ambiguity in model estimates also complicates drought monitoring, which increasingly relies on model estimated soil moisture due to the current lack of accurate global soil moisture measurements (Mo, 2008). Although Koster et al. (2010b) provided a more optimistic assessment on soil moisture estimates by various models, Mo (2008) indicated that while drought indices derived from different models show stronger correlation in the eastern US, their correlation is so low in the western US that model based drought indices cannot be used for drought monitoring. Drought monitoring is also complicated by the interaction between soil moisture and groundwater. Through numerical simulations, Peters et al. (2005) showed that groundwater can provide moisture to reduce the impact of short-term droughts, but due to its long recovery time groundwater will also act to lengthen and increase the frequency of droughts. The importance of groundwater for drought monitoring has been recognized (Houborg et al., 2011; Svoboda et al., 2002) and efforts are underway to combine information about groundwater variability as well as surface vegetation conditions with model estimated soil moisture to form comprehensive drought indices (http://www.drought.unl.edu/dm/monitor.html). Nevertheless, such efforts are hindered by the lack of systematic groundwater measurements at continental scales, in addition to lack of accurate model based soil moisture estimates.

In order to capture the unique characteristics of TWS and reduce the uncertainty in model estimates, observations are needed to nudge model output towards reality. The GRACE satellite system detects temporal water storage changes in the entire vertical profile, including snow mass, surface water, vegetation, soil moisture and groundwater (Tapley et al., 2004). It is the only remote sensing platform that provides consistent monitoring of the Earth's terrestrial water storage, including groundwater. Recognizing the potential for GRACE data to improve the simulation of land surface processes, Zaitchik et al. (2008) developed an ensemble Kalman smoother (EnKS) to assimilate GRACE into the NASA Catchment model in the Mississippi basin, with promising results. The EnKS provides a systematic and dynamic way to disaggregate

GRACE-derived TWS anomaly estimates into snow, soil moisture, and groundwater components, so that the simulation of each component of TWS can be positively influenced.

In this study, the EnKS and the Catchment model are applied in western and central Europe where climate and hydrological conditions differ significantly from the Mississippi area studied by Zaitchik et al. (2008). As droughts are common in Europe, the unique ability of GRACE TWS to detect droughts and its potential for drought monitoring are considered in some detail. The paper is organized as follows: Sections 2 and 3 describe the study domain, ground based validation data and the land surface model. Section 4 briefly outlines the EnKS method and filter parameters. Section 5 presents the model simulation results and comparisons with independent datasets. Comparisons of anomalies of GRACE TWS with those of MODIS NDVI are also presented. Section 6 concludes with a summary and discussion.

2. Experiment site, GRACE and validation data

Figure 1 shows the simulation domain in western and central Europe. For GRACE data assimilation, major hydrological watersheds were combined into nine major “basins” at the scale of GRACE observations, to accommodate the spatial resolution of GRACE TWS, which is about 150,000 km^2 at best (Rowlands et al., 2005; Swenson et al., 2006). Table 1 lists the area of these basins, ranging from 300,000 to 800,000 km^2. Several islands and peninsulas such as Great Britain and Sweden/Norway were not included because GRACE TWS yielded much smaller dynamic ranges than model estimates, possibly due to the interference of ocean signals.

GRACE TWS used in this study were processed by University of Texas Center for Space Research (CSR, Release CSR_RL04) using a Gaussian filter with a 300 km smoothing radius to remove the stripes seen in the spherical harmonic coefficient fields (Swenson and Wahr, 2006).

The anomalies of GRACE TWS were obtained by removing the temporal mean of the gravity field (including the solid earth and the atmosphere) in 2003-2007 and converted to equivalent water heights. The 1^{o} gridded GRACE TWS anomalies were mapped to the nine major basins using area-weighted averaging, and these values were converted to absolute TWS by adding the 2003 – 2007 mean TWS from an open loop (no data assimilation) integration of the model.

Figure 1 also shows the locations of in situ measurements used for validating data assimilation results, including 18 stream flow stations along three major rivers (Danube, Elbe and Rhine) and 12 soil moisture sites from the Soil Moisture Observing System - Meteorological Automatic Network Integrated Application (SMOSMANIA, Calvet et al., 2007) project. The streamflow stations (station ids and drainage areas are given in Table 2) were chosen from Global Runoff Data Center (GRDC) for their length of records. Soil moisture measurements (started in 2007) are taken at 5, 10, 20 and 30 cm depths and every 30 minutes using impedance probes. Monthly averaged stream flow and root zone soil moisture (vertically integrated using the four layer measurements) were used to validate model simulation results.

3. The Catchment model and forcing data

The NASA Catchment model was developed for global scale coupled land/atmosphere modeling (Koster et al., 2000a). It simulates water and energy balances on catchment tiles, with some catchments split by a 1.0^{o}x1.25^{o} atmospheric grid. For the study domain, which consists of nearly 6000 tiles, the average tile size is around 1500 km^2. To increase sub-grid heterogeneity, each catchment contains dynamically changing saturated, transpiring and wilting areas where different runoff and ET schemes are applied. The model contains three subsurface states for water balance calculation: surface excess (sfEx) and root zone excess (rtzEx), representing the excessive soil moisture relative to the hydrostatic state for the top 2 cm and 100 cm of soils, respectively, and catchment deficit (catDef) defined as the amount of water (kg/m^2, averaged over the catchment) needed to bring the catchment to saturation (assuming sfEx and rtzEx are zero). Although groundwater is not explicitly simulated, its behavior, i.e., its two dimensional distribution and associated flow rates, is directly diagnosed from the catDef variable. The model also has three snow layers for modeling snow water equivalent (SWE) and snow depth. Thus, modeled TWS can be determined from sfEx, rtzEx, catDef and SWE in conjunction with model parameters. Lakes and reservoirs are not directly included in simulated TWS because, over large scales at mid-latitudes, they only constitute a very small fraction of observed TWS variability (Rodell and Famiglietti, 2001). The impact of GRACE data assimilation on runoff is exerted through its relationship with modeled states: sfEx, rtzEx, catDef and SWE.

Forcing fields were provided by the Global Land Data Assimilation System (GLDAS, Rodell et al. 2004). They are based on meteorological fields (temperature, humidity, wind speed and pressure) obtained from the NASA Global Modeling and Assimilation Office GEOS data assimilation system (Bloom et al., 2005), radiation fields from the U.S. Air Force Weather Agency, and precipitation prepared by spatially and temporally downscaling the 2.5^o x2.5^o, 5-day NOAA Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997). This GLDAS forcing data set, which has been used in previous data assimilation experiments (Reichle et al., 2007; Zaitchik et al., 2008), has a 3 hour temporal interval and a 2^o × 2.5^o spatial resolution.

A few adjustment and corrections were made in this study regarding the Catchment model and forcing fields. Zaitchik et al. (2008) found that Catchment sometimes does not provide a large enough dynamic range to match that of GRACE TWS. The same situation was observed in this study region as well. To mitigate this deficiency, following Houborg et al. (2011), the bedrock depth used for the model was uniformly increased by 2 m, which increased the dynamic range of catDef. To partially compensate for the increase in bedrock depth, a lower value of the decay factor for saturated conductivity was used for the base flow calculation (Ducharne et al., 2000). Longwave and shortwave radiation fields were further bias corrected based on NASA/GEWEX Surface Radiation Budget (SRB, Release-3.0) data by matching their spatial (for entire simulation area) and temporal averaged means with those of SRB. The goal of these adjustments and corrections was to achieve reasonable estimates of fluxes (ET and runoff).

Simulations were carried out from August 2002 to July 2009, which is the available GRACE data period at the start of this study. Since previous forcing data were not available, the model was first run through 2002 to 2009 and then spun up for 10 years using the forcing fields from 2002. A different initial condition, based on averaged model states from 2002-2009 on January 1 which yielded wetter soil moisture conditions than the one mentioned above, was also tested and the results (including runoff and soil moisture evaluations) were very similar to those presented here.

4. GRACE data assimilation method

Zaitchik et al. (2008) presented a detailed description of the ensemble Kalman Smoother (EnKS) developed specifically for assimilating GRACE TWS into the Catchment model. A brief outline of this assimilation method is presented here. Like an ensemble Kalman filter (EnKF), the EnKS consists of two steps: forecast and update. In the forecast step, the ensemble of the model runs forward in time with perturbations added to the states and forcing fields:

$$X_{T^-}^{i} = M(X_{(T-1)^+}^{i}, F^{i}, G) \tag{1}$$

where M is the model; F represents all the forcing fields and G represents all the static parameters; T is the time; superscripts (-) and (+) refer to results for the forecast and update, respectively; X is the vector containing updated states (rtzEx, catDef and SWE) for each catchment tile, and the superscript i indicates the i^{th} member of the ensemble. srfEx was not updated in the EnKS because of its very weak correlation with monthly TWS but was included in model simulated TWS for accuracy. Based on equation (1), the ensemble update equation can be written as:

$$X_{T^+}{}^{i} = X_{T^-}{}^{i} + K_T(\underline{Y}_T - H(\underline{X}_{T^-}{}^{i})) \tag{2}$$

where K is the ensemble gain matrix; Y represents observations (GRACE TWS) and H is the observation operator that converts predicted states to the observation.

The underscores in equation (2) indicate monthly TWS (observed or simulated) averaged for each major basin because the EnKS used here assimilates temporally integrated observations. To accommodate the monthly averaged nature of GRACE observations, the EnKS collects Catchment model predictions of TWS on a first pass through each simulated month (three collections per month, to mimic GRACE overpass characteristics), calculates the update at the end of the month, and then iterates through the month a second time, uniformly (for each state) applying increments to each daily value of model states for each ensemble member. Thus, X (without the underscore) in equation (2) represents daily estimates of model states on each catchment tile in month T.

All perturbation parameters and schemes were the same as Zaitchik et al. (2008) and Reichle et al. (2007), except that an observation (GRACE) error of 15 mm was used here, which is the average of the two GRACE errors (10 mm and 20 mm) tested by Zaitchik et al. (2008).

5. Results

Two model integrations were performed from August 2002 to July 2009: the open loop (OL) representing the model-only performance and the data assimilation (DA) with GRACE data assimilation using the EnKS outlined above. Since GRACE derived TWS values are anomalies only, simulation results were evaluated using time series correlations with in situ measurements.

5.1 TWS

Figure 2 presents the time series of daily simulated TWS and GRACE monthly observations for the nine major basins. The open loop run generally captured the seasonal variability and dynamic range of GRACE TWS. OL differs from GRACE mostly in interannual variability, especially in Finland, Loire/Seine and Rhone/Po where OL exhibits a marked increase in TWS in the later modeling period. While data assimilation checked that increase effectively, consistent with GRACE observations, it failed to reduce simulated TWS to the levels observed by GRACE in the Upper Danube in 2007 and 2008. This failure was possibly caused by negligible ensemble spread during dry conditions due to a lack of precipitation to perturb and the fact that direct perturbations to sfEx and catDef are small. Increasing the direct perturbations may enable TWS to go lower, but it may also lead to ensemble bias. Nevertheless, in most cases EnKS was effective in nudging the simulated TWS toward GRACE TWS.

To investigate the cause of the significant increase in OL TWS seen in the Finland, Vistula, Loire/Seine, and Rhone/Po basins, which was not observed as dramatically in the

GRACE observations (Figure 2), GLDAS/CMAP precipitation was compared with $1^{o} \times 1^{o}$ Global Precipitation Climatology Project (GPCP) precipitation data (Adler et al., 2003) mapped to the major basins following the same approach as that for GRACE. Figure 3 shows the comparison of annual (from August to July) precipitation totals in each basin. In general, GLDAS/CMAP has a negative bias against GPCP in all basins except Turkey. CMAP's low bias relative to other precipitation products stems from the fact that it does not correct for gauge under-catch (e.g., Yin et al., 2004). More importantly, the annual variations of GPCP and GLDAS/CMAP precipitation are well correlated, and both products indicate that precipitation in the four basins named above increased towards the end of the simulation period. Given that GRACE TWS also increased in those basins but to a lesser extent, we infer that either: (i) the model should have retained less water in the land and increased evapotranspiration (ET) and/or runoff instead, or (ii) the precipitation and GRACE datasets are inconsistent, due to errors in one or both.

The rate of long-term TWS changes can be more clearly illustrated using the slope of monthly TWS calculated using Sen's method (Helsel and Hirsch, 1992; Sen, 1968) as shown in Figure 4. Slopes with a 0.1 significance level were identified using the Mann-Kendal test (Helsel and Hirsch, 1992) and marked in bold symbols. These two methods have been widely used in analyzing trends in hydro-meteorological data sets (Mishra and Cherkauer, 2010; Lettenmaier et al., 1994; Yue and Wang, 2002). Figure 4 shows that the slope of TWS (modeled or observed) generally becomes smaller as the basin moves from north to south, which resembles the increasing rate of annual precipitation in each basin (Figure 3), suggesting the strong correlation of TWS with long-term precipitation. OL TWS generally exhibits larger rates of increase than GRACE-derived TWS, especially in Finland, Vistula, Loire/Seine and Upper

Danube, where larger increasing rates of precipitation were observed in the later modeling period (Figure 3).

5.2 Stream flow and soil moisture

Since stream flow is a product of upland surface runoff and subsurface runoff over a large area, gauged stream flow data are often used to evaluate model performances (Mishra and Cherkauer, 2010). For the same reason, stream flow measurements were used here to not only evaluate the impact of GRACE data assimilation on runoff but also provide overall assessment of the EnKS. Figure 5 shows the correlation of monthly simulated runoff with GRDC gauged data. Since Catchment does not have a routing scheme, the simulated stream flow is simply a spatial-aggregation of tile-space (individual land element) runoff over the drainage area. This is justifiable for monthly statistics, especially in smaller basins where the runoff response time is less than a month. GRACE data assimilation improved the correlation in all watersheds but one (D5), with more improvements observed in larger basins along Danube. Improvements in watersheds such as R6-R11, E1 and E2 (Table 2) with drainage areas smaller than their corresponding major basins (the scale at which GRACE TWS was generated) indicate that assimilation of GRACE TWS can influence simulation of land surface processes at sub-observation scale. The improvements shown in Figure 5 by DA all exceeded the 0.05 significance level based on the William-Hotelling t-test (Steiger, 1980; Van Sickle, 2005). It should be pointed out that many of the stream flow observations are not independent because they were measured at various points along the same river.

Improvements in runoff correlations are attributed to the close relationship between base flow and catDef, which is the model state most affected by assimilation of GRACE TWS. To illustrate this, Figure 6 shows the time series of simulated runoff in comparison with GRDC measurements in Lower Danube 6742800, a sub-basin of the Lower Danube major basin. DA significantly increased the runoff in the earlier period in accordance with changes in TWS, which helped improving the overall correlation and also lowered the increasing trend of runoff. Figure 7 shows the trend of runoff by OL, DA and GRDC gauge data in all GRDC basins. Similar to TWS, model estimates (OL) show higher trends than observed runoff with significant trends detected for most basins while observed runoff shows no significant trend in any basin. DA reduced trends in all basins, but did not change the significance of most trends.

An important role of the EnKS is to disaggregate GRACE so that each TWS component can be nudged towards its true state. To evaluate the vertical disaggregation, correlations of monthly root zone soil moisture estimated by OL and DA were calculated against in situ measurements from the SMOSMANIA sites and are given in Table 3. The statistics were calculated using in situ point data and model estimates at the tile containing the station. GRACE data assimilation generally did not have a significant impact on monthly correlations of soil moisture as the correlation of DA is not significantly different from OL at the 0.10 significance level, except at site URG. The coarser spatial representation of the model and the GRACE data may not capture the localized nature of station measurements. To alleviate the horizontal scale mismatch and obtain an overall impact on the entire SMOSMANIA area (about 4000 km^2), the area averaged statistics for OL and DA were also calculated against the averaged in situ measurements and are given in Table 3 (last row) which shows that GRACE data assimilation did not change the correlation of averaged soil moisture time series in the sampling area. The shorter SMOSMANIA data period (31 months) makes these statistics less conclusive because the confidence intervals are very large.

5.3 Water budget

As hypothesized in section 5.1, elevated TWS by OL in Finland and Loire/Seine in the later modeling period were likely caused by either an underestimation of runoff and/or ET when precipitation rates increased or by improper increase in the precipitation rates themselves assuming GRACE data are accurate. When GRACE data assimilation reduced TWS in these basins, it also decreased ET and runoff estimates because of their positive correlations with TWS. As a result, the water budget of OL was not preserved by the simulation with GRACE data assimilation. Figure 8 features the annual (August to July) mass imbalance, defined as simulated water budget (sum of total fluxes and net change in TWS) minus precipitation, of OL and DA. As expected, OL has nearly zero mass imbalances throughout the entire period and in all basins while GRACE data assimilation disrupted the water budget, more so in Finland, Vistula, Loire/Seine and Rhone/Po, despite improving the simulation of TWS (assuming GRACE data are accurate). Since GLDAS precipitation generally has a negative bias against GPCP (Figure 3), positive imbalances (i.e., larger ET and runoff) would be preferable to the negative ones produced by GRACE data assimilation in this case. Unintended impacts of data assimilation on the water budget are always a danger, demanding the development of creative new assimilation techniques (e.g., Li et al., 2011; Pan and Wood, 2006; Zaitchik and Rodell, 2009).

5.4 Drought analysis

Droughts are common in Europe, and several episodes of severe droughts, including the 2003 drought (associated with the 2003 European heat wave, Rebetez et al., 2006; Zaitchik et al.,

2006) that spread across western and central Europe and the 2007/2008 droughts that affected southern and southwestern Europe (SOER Synthesis, 2010), were detected by GRACE TWS (Figure 2). Because droughts can be defined in a variety of ways depending on what indicators are taken into account, it can be instructive to compare a new drought observation with a more common indicator. Here we compare GRACE based TWS with monthly Normalized Difference Vegetation Index (NDVI) as recorded by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on NASA's Terra satellite. NDVI is strongly correlated with green biomass (Tucker, 1979), and is often used in satellite based drought-monitoring (e.g., Brown et al., 2008). Basin averaged NDVI was derived by averaging the Level-3 0.05^{o} MODIS NDVI monthly product (lpdaac.usgs.gov) across the same basins that were used to extract GRACE observations.

Figure 9 shows the averaged dryness ranks of NDVI and GRACE TWS in the summer season (April to September) for the 2003 to 2008 period (2002 and 2009 were excluded due to their incomplete summer seasons). To give equal weight to all monthly rankings, the averaged ranks in Figure 9 were obtained by first ranking each data set for each month and then averaging the ranks of summer months. GRACE TWS indicated 2003 as the driest condition in all basins except Loire/Seine, Lower Danube and Turkey, while NDVI only shows 2003 as the most severe drought in Rhone/Po, Upper Danube and Dnieper and a drought condition in Rhine/Elbe/Oder and Loire/Seine. The 2007/2008 droughts along the south and southwestern region (in Rhone/Po, Lower/upper Danube, Dnieper and Turkey) were indicated by both types of observations. The largest discrepancies between the two sources are in Finland and Vistula where, despite the increasing trend in precipitation (Figure 3), NDVI shows decreasing trends.

This is probably due to the fact that vegetation growth in the high latitude and high elevation regions is limited by energy availability, not by water availability (Karnieli et al., 2010).

Note that GRACE TWS characterized the 2003 drought in Loire/Seine as less severe than the 2005 drought (SOER Synthesis, 2010). According to GRACE, the land was very wet in early 2003 (Figure 2), and as a result dry meteorological conditions took longer to severely impact total TWS. In this situation, the effect of drought is less evident in the TWS anomaly than it is in the maximum decline of GRACE TWS from its early spring peak to the lowest value in the fall, which roughly measures the amount of water lost in the warm season. As seen in Figure 9, Loire/Seine and Upper Danube, which were at or near the center of the heat wave, experienced the most significant water loss in 2003. This is one of the advantages using a physical-based variable for drought monitoring because drought conditions can be evaluated from other aspects than anomalies.

The reason that we only compared the dryness rank of GRACE and NDVI during the warm season in Figure 9 is that NDVI is insensitive to water shortage when vegetation is senescent or when coverage is low (Karnieli et al., 2010). This can be seen in Figure 10 where the seasonal cycles of GRACE TWS and NDVI in the Lower Danube basin are presented. GRACE TWS shows signs of stress in 2007 very consistently over all seasons, in contrast with NDVI which indicated vegetation stress only after June. GRACE-derived TWS also exhibits large inter-annual variability and larger dynamic ranges that can provide more information on drought severity. These qualities, true in most areas (Rodell, 2011), are important both for drought monitoring and for early detection of drought onset and therefore make GRACE a useful complement to high-resolution NDVI-based measures of drought, especially in regions with low vegetation cover or where water is not a limiting factor for vegetation growth.

Figures 9 and 10 show the dryness ranks based on GRACE TWS data alone. To demonstrate the potential value of integrating GRACE and other data with a land surface model for drought monitoring, Figure 11 plots the dryness ranks (among 2002 to 2009) of OL and DA estimated root zone soil moisture (upper panels), which is of particular interest for monitoring agricultural droughts, and TWS (lower panels), which is an indicator of water depletion in the deeper subsurface, for November 2007. GRACE DA intensified the drought condition in Loire/Seine and Upper Danube relative to the open loop. The updates in both the root zone soil moisture and TWS demonstrate that data assimilation makes it possible to apply GRACE for monitoring both agricultural and hydrological droughts, and to do so with much greater spatial resolutions than with GRACE alone.

6. Summary and Discussions

This study demonstrated the value of GRACE TWS for correcting errors in model estimated TWS and its influence on related land surface processes. In particular, assimilation significantly improved runoff correlation in most basins, which attests to the overall robustness of the assimilation technique and the usefulness of GRACE TWS for runoff estimation. The improved runoff correlation in small watersheds also shows the potential of GRACE TWS to contribute to simulation of finer scale hydrological processes through data assimilation based downscaling. Assimilation of GRACE TWS did not improve the correlation of soil moisture with in situ measurements, perhaps due to the short in situ data record or insufficient spatial sampling. Although groundwater was not validated directly due to lack of in situ measurements, the improvements in stream flow estimates suggest more realistic estimates of subsurface water storage which controls baseflow.

GRACE data assimilation had a significant impact on reducing trends of modeled TWS and runoff. The original inconsistency between the GRACE and OL trends is caused by deficiencies in either the model's physics, the forcing data or the GRACE data themselves. The case presented here represents a relatively short period during which annual precipitation increased at a much higher rate in several basins than long term annual precipitation trends (Mishra and Cherkauer, 2010; Solomon et al., 2007). The fact that GRACE TWS was able to change the trend in runoff suggests that GRACE TWS data, if independently validated, may assist in model and forcing evaluation and calibration, which is an important part of climate prediction (Mishra and Cherkauer, 2010), especially in regions with scarce observation data. However, only models able to simulate groundwater storage can take full advantage of GRACE, because assimilation of GRACE TWS requires an analogous model state.

Monitoring of droughts has suffered from lack of reliable information on the water stored below the uppermost soil layer. Since GRACE measures the water storage changes in the entire profile, it provides valuable information on drought development beyond what can be seen at the surface. Its large dynamic range and inter-annual variability also provides better quantification of the severity of water depletion in the subsurface. The continued monitoring of dry conditions throughout all seasons, which cannot be achieved using vegetation based indicators, may further assist in tracking prolonged droughts and/or providing early signs of drought development.

While data assimilation of GRACE TWS helps to fill the need for regional to global scale information on deep moisture storage variability, it also presents some challenges. Since drought indices are derived based on the long term climatology of a given variable (Mo, 2008) and the GRACE observation period is not long enough to generate its own climatology, GRACE based drought indices must be linked to a model simulation that begins well prior to data assimilation.

This requires that the estimates from GRACE assimilation have the same dynamic range as GRACE, so that the anomalies from the assimilation period are comparable to the climatology. To accomplish this, it may be necessary to modify parameters such as the bedrock depth, which controls the amount of water available from storage to be lost during a drought (Houborg et al., 2011). The changing trends in DA TWS, as found in this study, may also reduce the dynamic range and the magnitude of anomalies and thus present a new challenge. Statistical techniques such as cumulative distribution function matching may also be used to ensure that the modeled and observed climatologies are consistent prior to generating drought indices (Houborg et al., 2011). Nevertheless, these challenges should not discourage the use of GRACE data assimilation for drought monitoring because the dryness information provided by GRACE TWS can lead to more objective and reliable drought indices (Rodell, 2011).

Water budget imbalance caused by GRACE data assimilation is an important issue for future research because existing flux biases may be exacerbated (assuming precipitation forcing data were accurately estimated). In this example, we speculate, without the benefit of ET and runoff observations in Finland and Loire/Seine regions, that a low bias in modeled ET and runoff might have caused the TWS anomaly to be elevated, which, when corrected by GRACE data assimilation, further reduced ET and/or runoff. This water budget imbalance might have been avoided, if observations of ET and runoff were available and assimilated simultaneously with GRACE TWS. Given that ET and runoff observations are rarely assimilated into land surface models, a more likely solution would be to remove excess TWS during the assimilation process in conjunction with increasing simulated ET and/or runoff. Exploring creative new data assimilation strategies such as this is recommended so that the benefits of GRACE DA can be realized while avoiding detrimental effects on modeled water budgets (Li et al., 2011; Pan and Wood, 2006; Zaitchik and Rodell, 2009).

Acknowledgements

This work was supported by funds from NASA's Terrestrial Hydrology Program. We thank GRDC and Meteo-France for providing validation data used in this study. Joseph Nigro provided assistance in creating the major basins. Barton Forman read a draft of this manuscript and provided useful comments.

References

Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, 2003. The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present), J. Hydrometero., 4, 1147-1167.

Bierkens, M. F. P. and L. P. H. van Beek, 2009. Seasonal predictability of European discharge. NAO and hydrological response time, J. Hydrometeor., 10, 953-968.

Bloom, S., A. da Silva, D. Dee, M. Bosilovich, J.-D. Chern, S. Pawson, S. Schubert, M. Sienkiewicz, I. Stajner, W.-W. Tan, and M.-L. Wu, 2005. Documentation and validation of the Goddard Earth Observing System (GEOS) data assimilation system. Version 4, tech. Rep. Ser.

Global Model. Data Assimil. 104606, 26, 187pp., Global Model. and Assimil. Off., NASA Goddard Space Flight Cent., Greenbelt, MD.

Brown, J.F., B.D. Wardlow, T. Tadesse, M.J. Hayes, and B.C. Reed, 2008. The vegetation drought response index (VegDRI): a new integrated approach for monitoring drought stress in vegetation, GIScience and Remote Sensing, 45(1):16-46.

Calvet, J.-C., N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B . Piguet, 2007. In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network, Geoscience and Remote Sensing Symposium, IGARSS 2007, IEEE International.

Dirmeyer P. A., 2000. Using a global soil wetness dataset to improve seasonal climate simulation, J. Climate, **13**, 2900-2922.

Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker, 2009. Precipitation, recycling, and land memory: an integrated analysis, J. Hydrometeor., **10**, 278-288.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz and P. Kumar, 2000. A catchment-based approach to modeling land surface processes in a general circulation model 2. Parameter estimation and model demonstration, J. Geophy. Res., **105**, D20, 24,823-24,838.

SOER Synthesis, 2010. The European environment – state and outlook 2010: synthesis, European Environment Agency, Copenhagen.

Eltahir, E. A. B. and P. J.-F. Yeh, 1999. On the asymmetric response of aquifer water level to floods and droughts in Illinois, Water Resour. Res., **35**(4), 1199-1217.

Fekete, B. M., C. J. Vorosmarty, J. Roads, and C. J. Willmott, 2004. Uncertainties in precipitation and their impacts on runoff estimates, J. Climate, **17**, 294-304.

Gottschalck, J., J. Meng, M. Rodell and P. Houser, 2005. Analysis of multiple precipitation products and preliminary assessment of their impact on Global Land Data Assimilation System land surface states, J. Hydrometeor., **6**, 573-598.

Helsel, D. R. and R.M. Hirsch, 1992. Statistical Methods in Water Resources, Elsevier, Amsterdam, 522 pp.

Hoerling, M., D. Lettenmaier D. Cayan and B. Udall, 2009. Reconciling Projections of Colorado River Streamflow, Southwest Hydrology, May/June, p.20-31.

Houborg, R., M. Rodell, B. Li, R. Reichle, and B. Zaitchik, Drought indicators based on model assimilated GRACE terrestrial water storage observations, Wat. Resour. Res., in review, 2011.

Karnieli, A., N. Agam, R. T. Pinker, M. Anderson, M. L. Imhoff, G. G. Gutman, N. Panov and A. Goldberg, 2010. Use of NDVI and land surface temperature for drought assessment: merits and limitations, J. Climate, **23**, 618-633.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000a. A catchment-based approach to modeling land surface processes in a general circulation model, 1, Model structure, J. Geophys. Res., **105**, 24809-24822.

Koster, R. D., M. J. Suarez and M. Heiser, 2000b. Variance and predictability of precipitation at seasonal-to-interannual timescales, J. Hydrometeor., **1**, 26-46.

Koster, R. D. and M. J. Suarez, 2001. Soil moisture memory in climate models, J. Hydrometeor., **2**, 558-570.

Koster, R. D. Koster, S. P. P. Mahanama, B. Livneh, D. P. Lettenmaier and R. H. Reichle, 2010a. Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, Nature Geos., **3**, 616-616.

Koster, R. D., Z. Guo, R. Yang, P. A. Dirmeyer, K. Mitchell, and M. J. Puma, 2010b. On the nature of soil moisture in land surface models. J. Climate, 22, 4322-4335.

Lettenmaier, D. P., E. F. Wood, and J. R. Wallis, 1994. Hydroclimatological trends in the continental United States, 1948–88. J. Climate, **7**, 586–607.

Li, B., D. Toll, X. Zhan, and B. Cosgrove, 2011. Improving simulated soil moisture fields through assimilation of AMSR-E soil moisture retrievals with an ensemble Kalman filter and a mass conservation constraint, Hydrol. Earth Syst. Sci. Discuss., **8**, 8131-8171, doi:10.5194/hessd-8-8131-2011.

Miguez-Macho, G., Y. Fan, C. P. Weaver, R. Walko, and A. Robock, 2007. Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation, J. Geophys. Res., **112**, doi:10.1029/2006JD008112.

Mishra, V. and K. A. Cherkauer, 2010. Assessment of drought due to historic climate variability and projected future climate change in the Midwestern United States, J. Hydrometeor., **11**, 46-68.

Mitchell, K. E., and Co-authors, 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, J. Geophys. Res., **109**, D07S90, doi:10.1029/2003JD003823.

Mo, K. C., 2008. Model-based drought indices over the United States, J. Hydrometeor., **9**, 1212-1230.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden, 2005. A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res., **110**, D21106, doi:10.1029/2005JD006111.

Niu, G.Y., Z. Yang, R. E. Dickinson, L. E. Gulden, and H. Su, 2007. Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data, J. Geophys. Res., **112**, D07103, doi:10.1029/2006JD007522.

Pan, M. and E. F. Wood, 2006. Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, J. Hydrometeor., **7**(3), 534-547.

Peters, E., H. A. J. van Lanen, P. J. J. F. Torfs, and G. Bier, 2005. Drought in groundwater—drought distribution and performance indicators, J. Hydrol., **306**, 302-317.

Rebetez, M., H. Mayer, O. Dupont, D. Schindler, K. Gartner, J. P. Kropp and A. Menzel, 2006. Heat and drought 2003 in Europe: a climate synthesis, Ann. For. Sci. **63,** 569–577.

Reichle, R. H., R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. Njoku, and M. Owe, 2007. Comparison and assimilation of global soil moisture retrievals from the Advanced Microware Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), J. Geophy. Res., **112**, D09108, doi:10.1029/2006JD008033.

Rodell, M. and J. S. Famiglietti, 2001. An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), Water Resour. Res., **37**(5), 1327-1339.

Rodell, M. and Coauthors, 2004. The global land data assimilation system, Bull. Amer. Meteor. Soc., **85**, 381-394.

Rodell, M., J. Chen, H. Kato, J. Famiglietti, J. Nigro, and C. Wilson, 2007. Estimating ground water storage changes in the Mississippi River basin(USA) using GRACE, Hydrogeol. J., **15**, 159-166, doi:10.1007/s10040-006-0103-7.

Rodell, M., 2011. "Satellite Gravimetry Applied to Drought Monitoring", in B. Wardlow, M. Anderson, and J. Verdin (Eds), Remote Sensing of Drought: Innovative Monitoring Approaches, CRC Press/Taylor and Francis, in press.

Rowlands, D. D., and Coauthors, 2005. Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, Geophy. Res. Lett., **32**, L04310, doi:10.1029/2004GL021908.

Schaake, J. C., V. Koren, Q. Duan, K. Mitchell, and F. Chen, 1996. Simple water balance model for estimating runoff at different spatial and temporal scales, J. Geophys. Res., **101**(D3), 7461-7475.

Schaller, M. F., and Y. Fan, 2009. River basins as groundwater exporters and importers: Implications for water cycle and climate modeling, J. Geophys. Res., **114**, D04103, doi:10.1029/2008JD010636.

Sen, P.K. 1968. Estimates of the regression coefficient based on Kendall's tau, J. American Statistical Association, **63**, 1379-1389.

Steiger, J. H., 1980. Tests for comparing elements of a correlation matrix, Psychological Bulletin **87**, 245-251.

SOER Synthesis, 2010. The European environment – state and outlook 2010: synthesis, European Environment Agency, Copenhagen.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.), 2007. IPCC Forth Assessment Report/Climate Change 2007: The Physical Science Basis, Cambridge University Press, 996p.

Svoboda, M., D. LeComte, M. Hayes, R. Heim, K. Gleason, J. Angel, B. Rippey, R. Tinker, M. Palecki, D. Stooksbury, D. Miskus, and S. Stephens, 2002. The Drought Monitor, Bull. Amer. Meteor. Soc., **83**(8):1181-1190.

Swenson, S. and J. Wahr, 2006. Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., **33**, L08402, doi:10.1029/2005GL025285.

Swenson, S. C., P. J.-F. Yeh, J. Wahr and J. S. Famiglietti, 2006. A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, Geophys. Res. Lett., 33, L16401, doi:10.1029/2006GL026962.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins, 2004. GRACE measurements of mass variability in the Earth system, Science, 305, 503-505.

Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation, Remote Sens. Environ., **8**, 127– 150.

Van Sickle, J., 2005. Errata:"Analyzing correlations between stream and watershed attributes", J. of the Ameri. Water Res. Assoc. (JAWRA), **41**(3), 741.

Wu, W., M. A. Geller, and R. E. Dickinson, 2002. The response of soil moisture to long-term variability of precipitation, J. Hydrometeor., **3**(12), 604-613.

Xie, P. and P. A. Arkin, 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs, Bull. Amer. Meteor. Soc., 78, 2539-2558.

Yeh, P. J. F., and E. A. B. Eltahir, 2005. Representation of water table dynamics in a land surface scheme, Part I: Model development, J. Clim., **18**(12), 1861–1880.

Yeh, P. J. F., S.C. Swenson, J.S. Famiglietti, and M. Rodell, 2006. Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), Water Resour. Res., **42**, W12203, doi:10.1029/2006WR005374.

Yin, X., A. Gruber, and P. A. Arkin, 2004. Comparison of the GPCP and CMAP merged gauge-satellite monthly precipitation products for the period 1979-2001. J. Hydrometeor., 5, 1207-1222.

Yue, S., and C. Y. Wang, 2002. Regional streamflow trend detection with consideration of both temporal and spatial correlation, Int. J. Climatol., **22**, 933–946.

Zaitchik, B. F. A. K. Macalady, L. R. Bonneau and R. B. Smith, 2006. Europe’s 2003 heat wave: a satellite view of impacts and land–atmosphere feedbacks, Int. J. Climatol. **26**, 743–769.

Zaitchik, B. F., M. Rodell, and R. H. Reichle, 2008. Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi river basin, J. Hydrometeor., **9**, 535-548.

Zaitchik, B.F. and M. Rodell, 2009. Forward-looking assimilation of MODIS-derived snow covered area into a land surface model, J. Hydrometeor., **10** (1), 130-148.

Figures

Figure 1: Study area and major basin boundaries. The blue cross and red triangle represent locations of GRDC stream flow and SMOSMANIA soil moisture sites, respectively. Numbers 1 to 9 represent the nine major basins given in Table 1.

Figure 2: Time series of daily simulated TWS and monthly GRACE TWS in the nine major basins.

Figure 3: Comparisons of annual GLDAS and GPCP precipitation in the nine basins.

Figure 4: Slopes of trend for monthly TWS in the nine major basins. Trends with a 0.1 significance level are marked with bold symbols.

Figure 5: Correlations of monthly simulated runoff with GRDC stream flow. All improvements by DA exceed the 0.05 significance level. Station ids are given in Table 2.

Figure 6: Monthly time series of estimated runoff in comparison with GRDC gauge data.

Figure 7: Slopes of trend for monthly runoff at GRDC stations. Trends with a 0.1 significance level are marked with bold symbols.

Figure 8: Annual mass imbalance (simulated water budget minus precipitation) for OL and DA in the nine major basins.

Figure 9: Averaged dryness ranks of NDVI and GRACE TWS for the summer growing season (April to September) during the 2003 to 2008 period and maximum GRACE TWS declines from spring to fall in each year.

Figure 10: Seasonal cycles of GRACE TWS and NDVI in Lower Danube.

Figure 11: Dryness ranks of simulated root zone soil moisture and TWS for November 2007 in the 2002 to 2009 period.

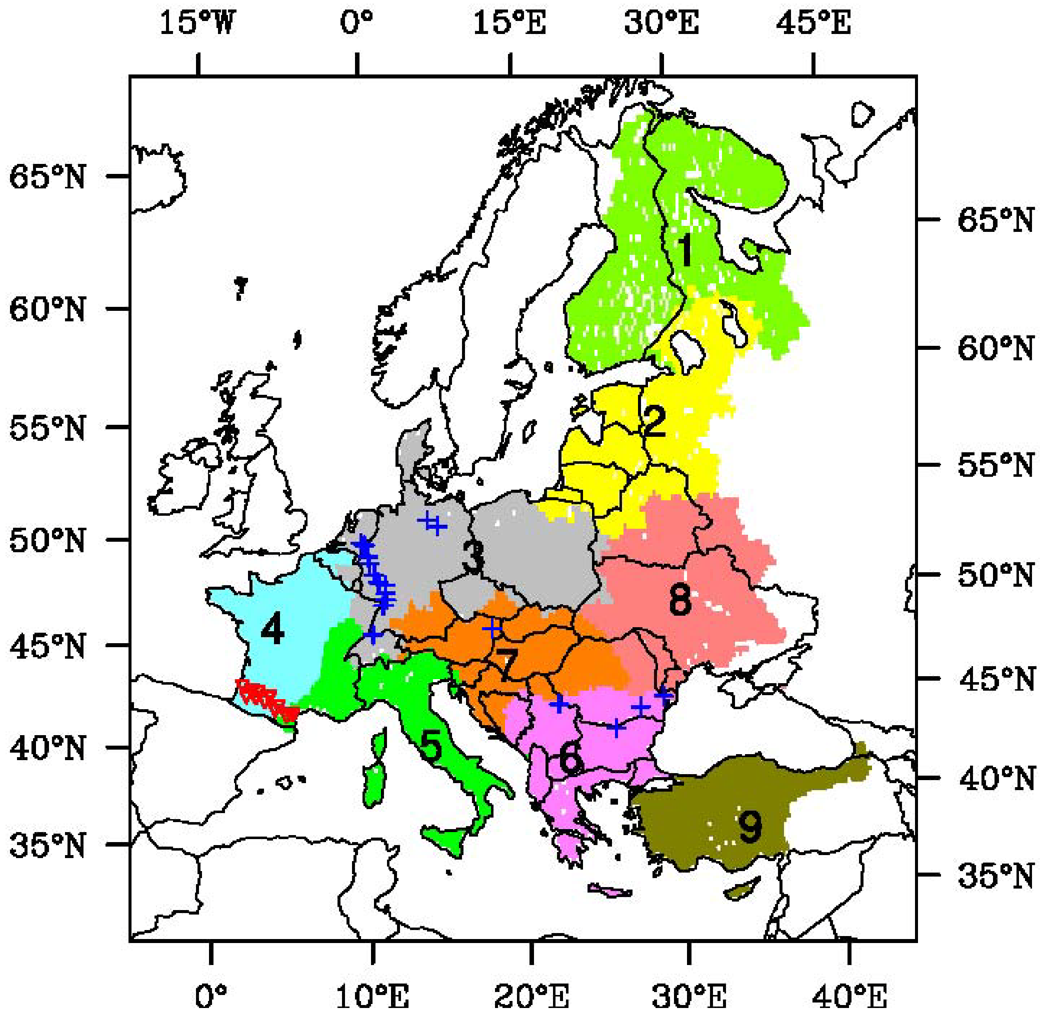

Figure 1. Study area and major basin boundaries. The blue cross and red triangle represent locations of GRDC stream flow and SMOSMANIA soil moisture sites, respectively. Numbers 1 to 9 represent the nine major basins given in Table 1.

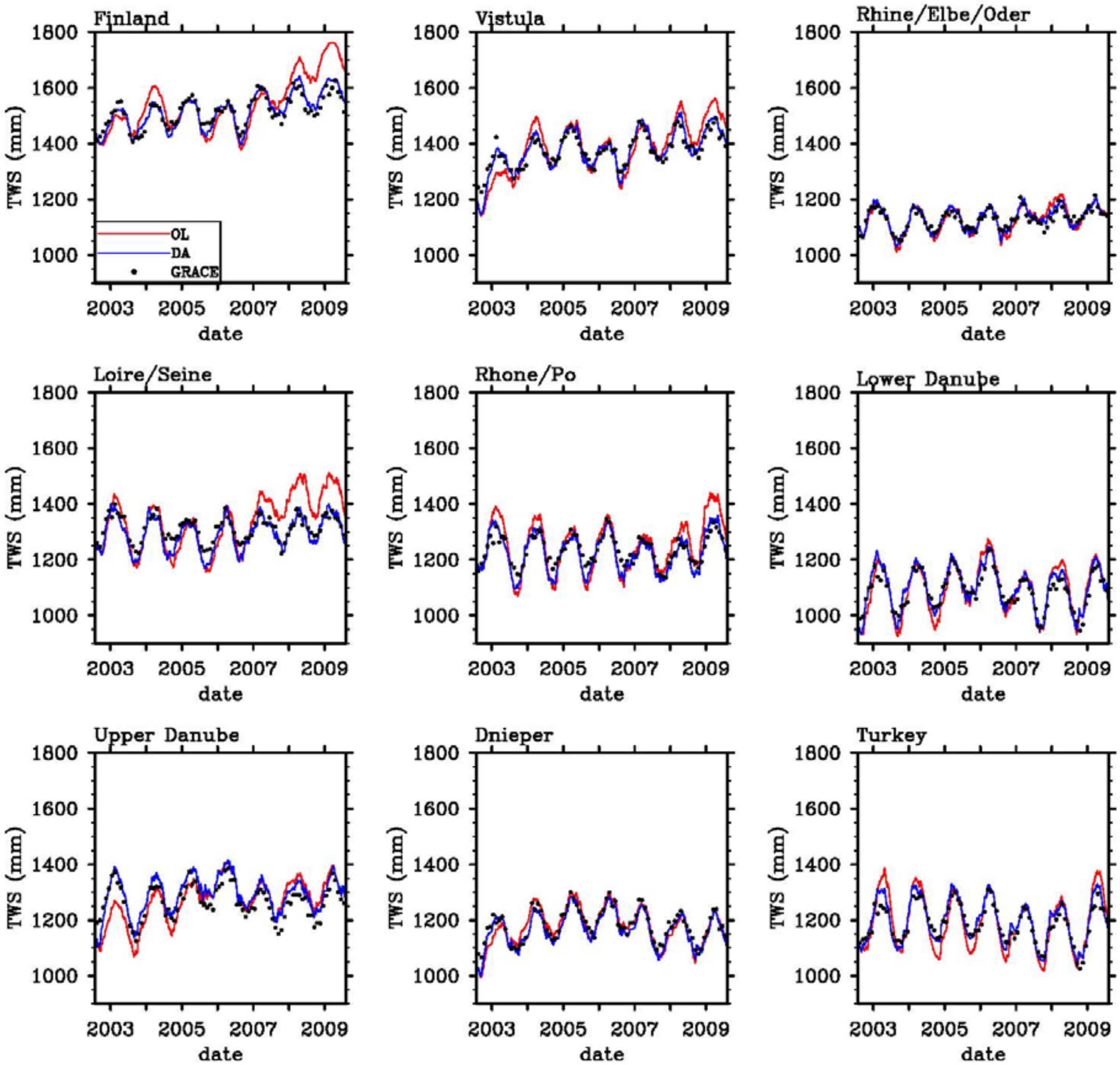

Figure 2. Time series of daily simulated TWS and monthly GRACE TWS in the nine major basins.

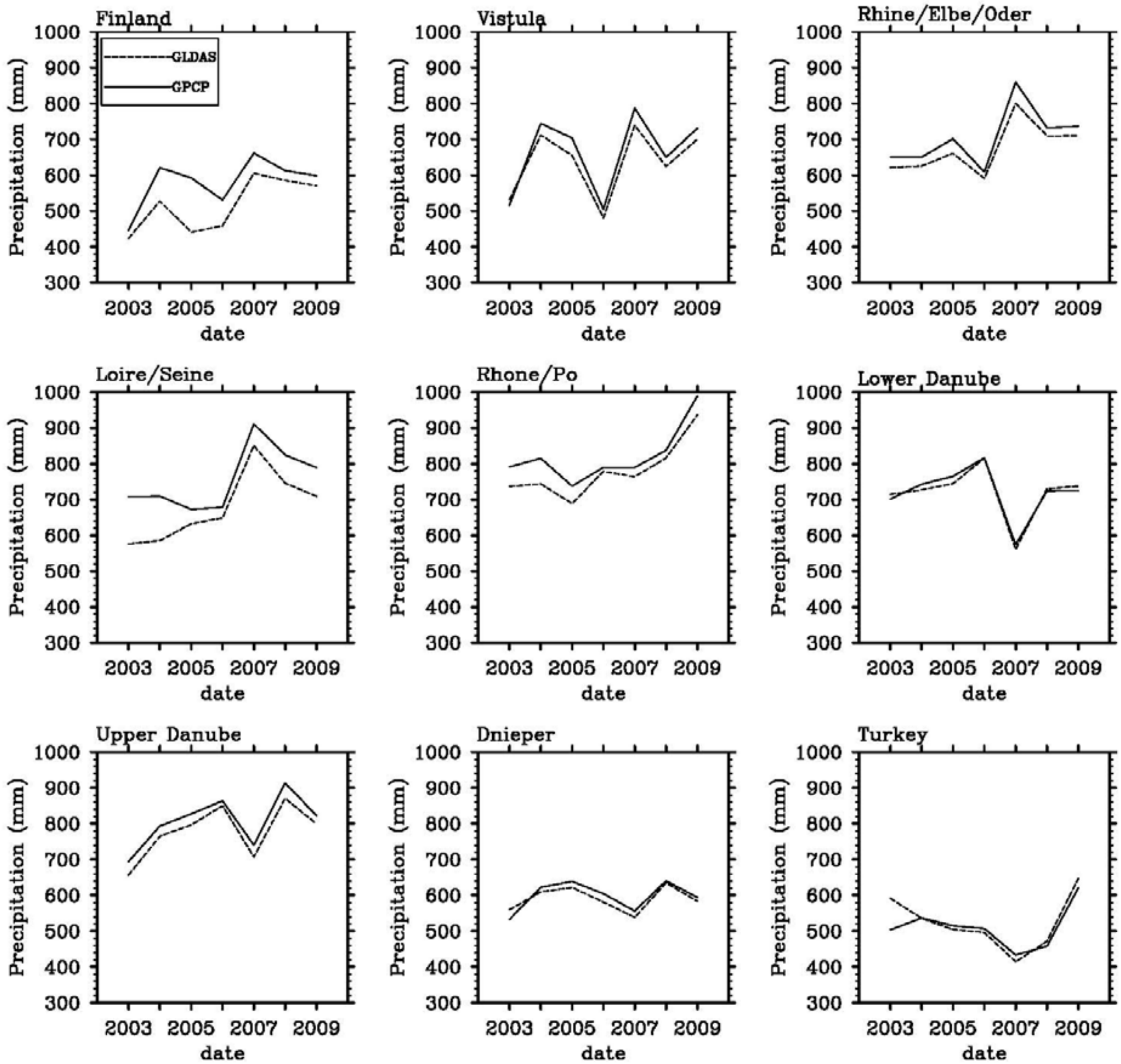

Figure 3. Comparisons of annual GLDAS and GPCP precipitation in the nine basins.

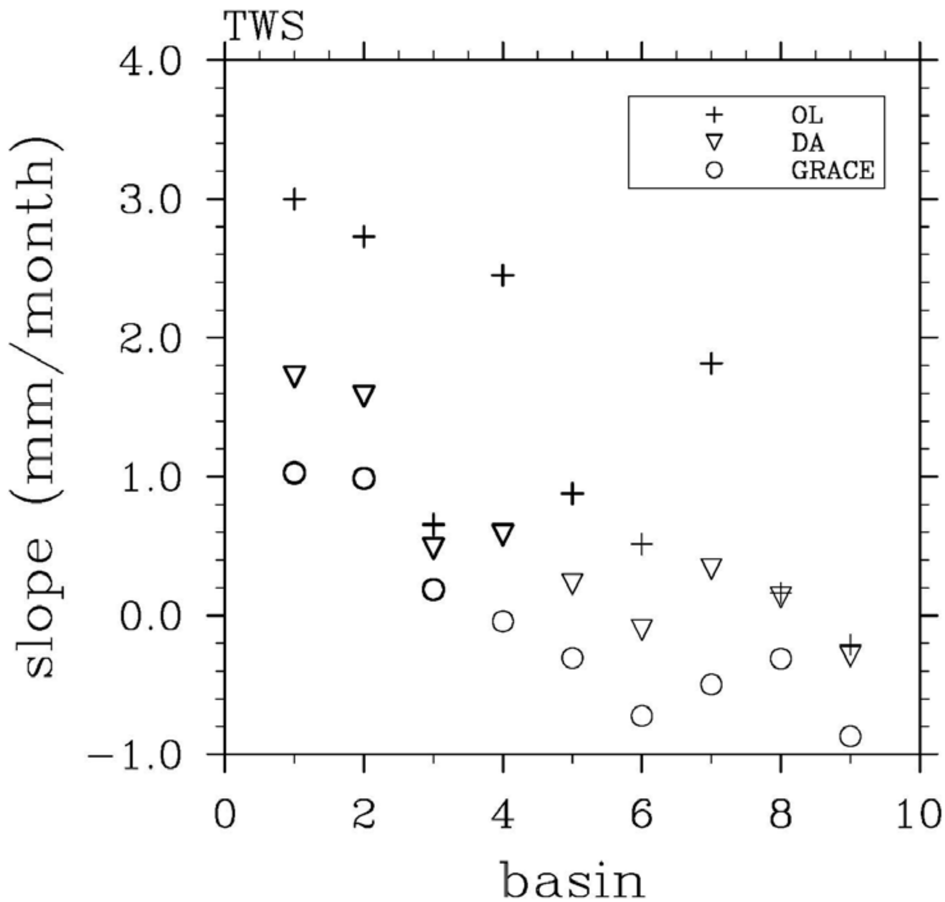

Figure 4. Slopes of trend for monthly TWS time series in the nine major basins. Trends with a 0.1 significance level are marked with bold symbols.

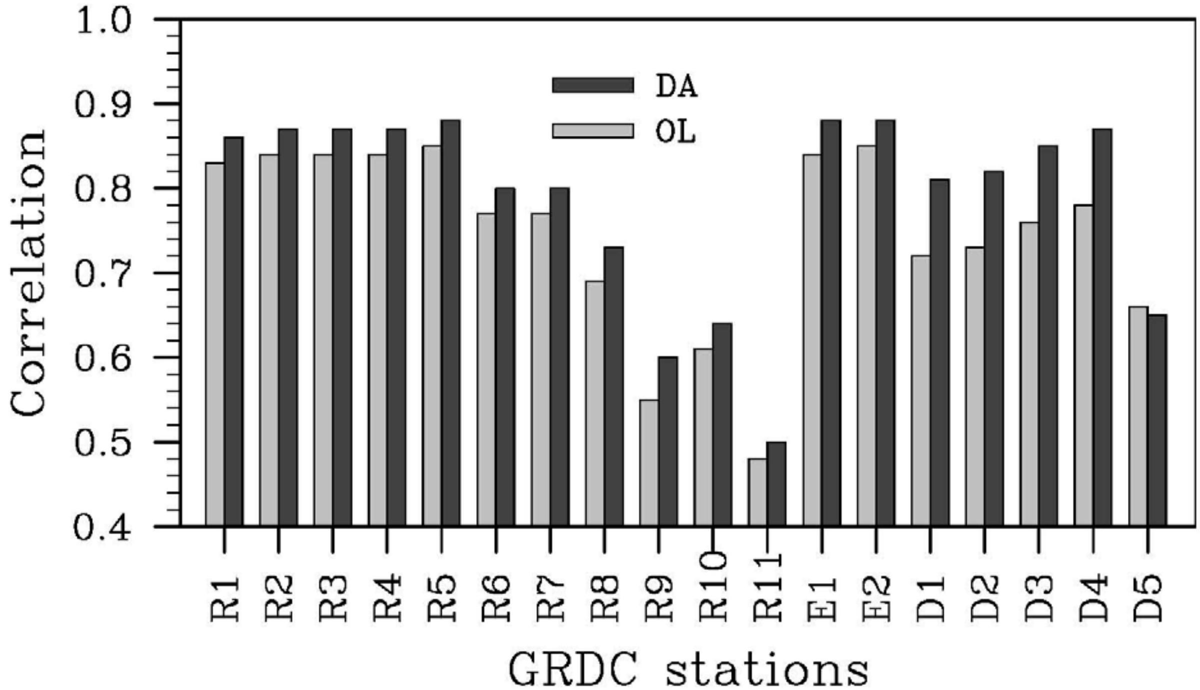

Figure 5. Correlations of monthly simulated runoff with GRDC stream flow. All improvements by DA exceed the 0.05 significance level. Station ids are given in Table 2.

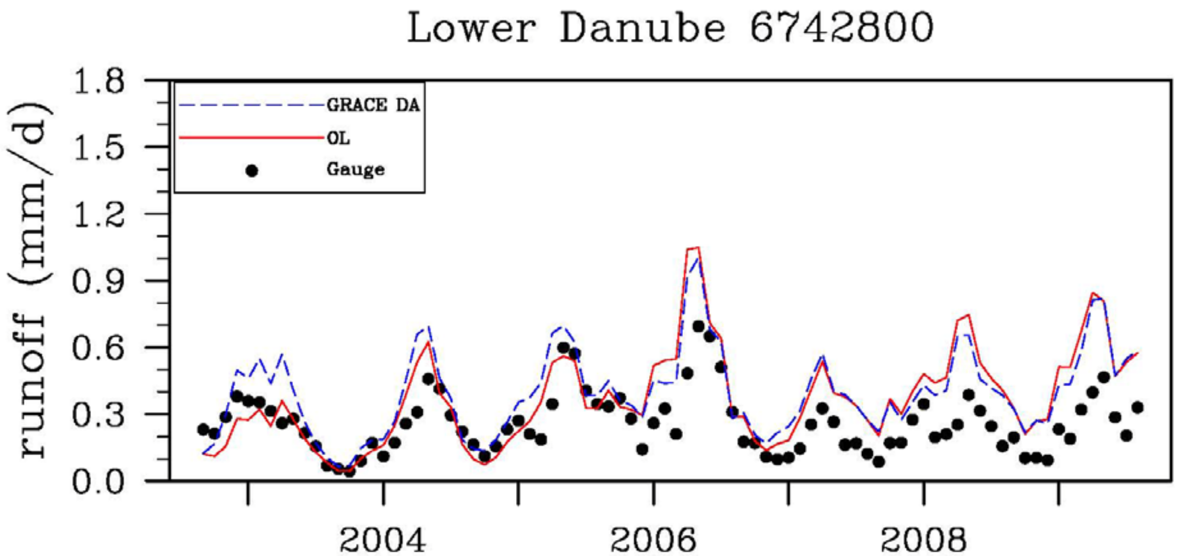

Figure 6. Monthly time series of estimated runoff in comparison with GRDC gauge data.

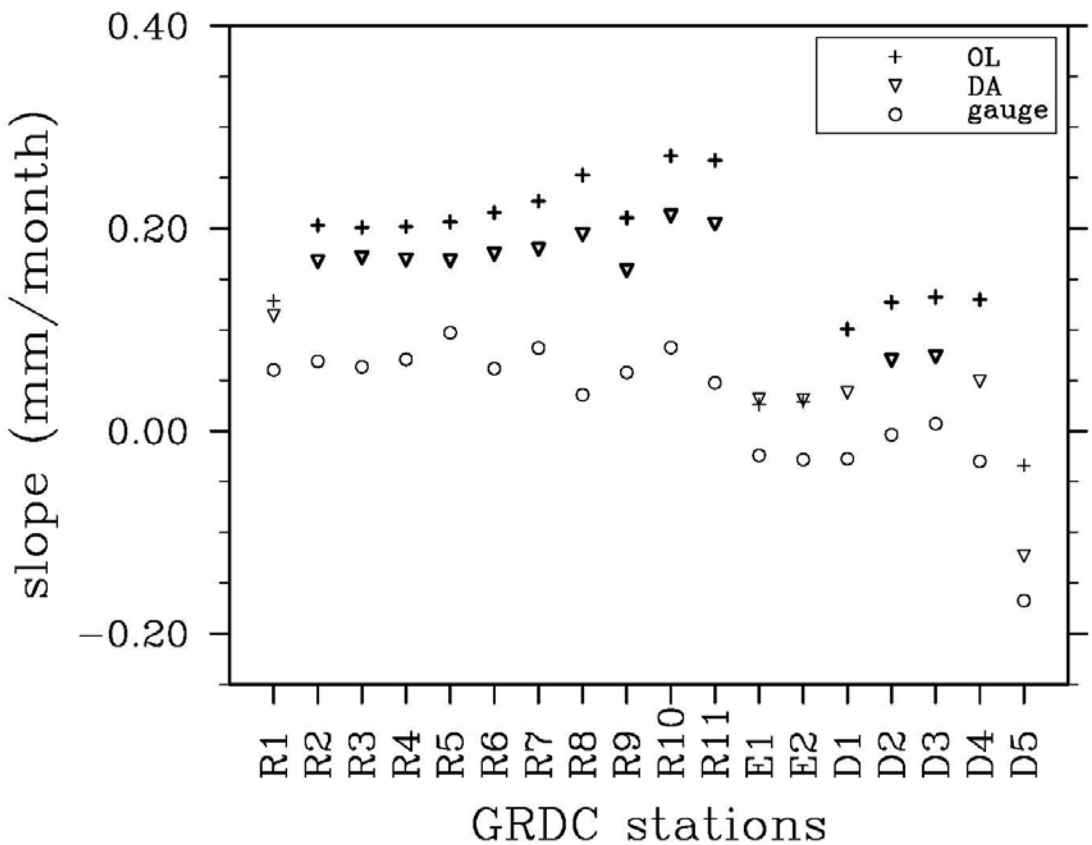

Figure 7. Slopes of trend for monthly runoff at GRDC stations. Trends with a 0.1 significance level are marked with bold symbols.

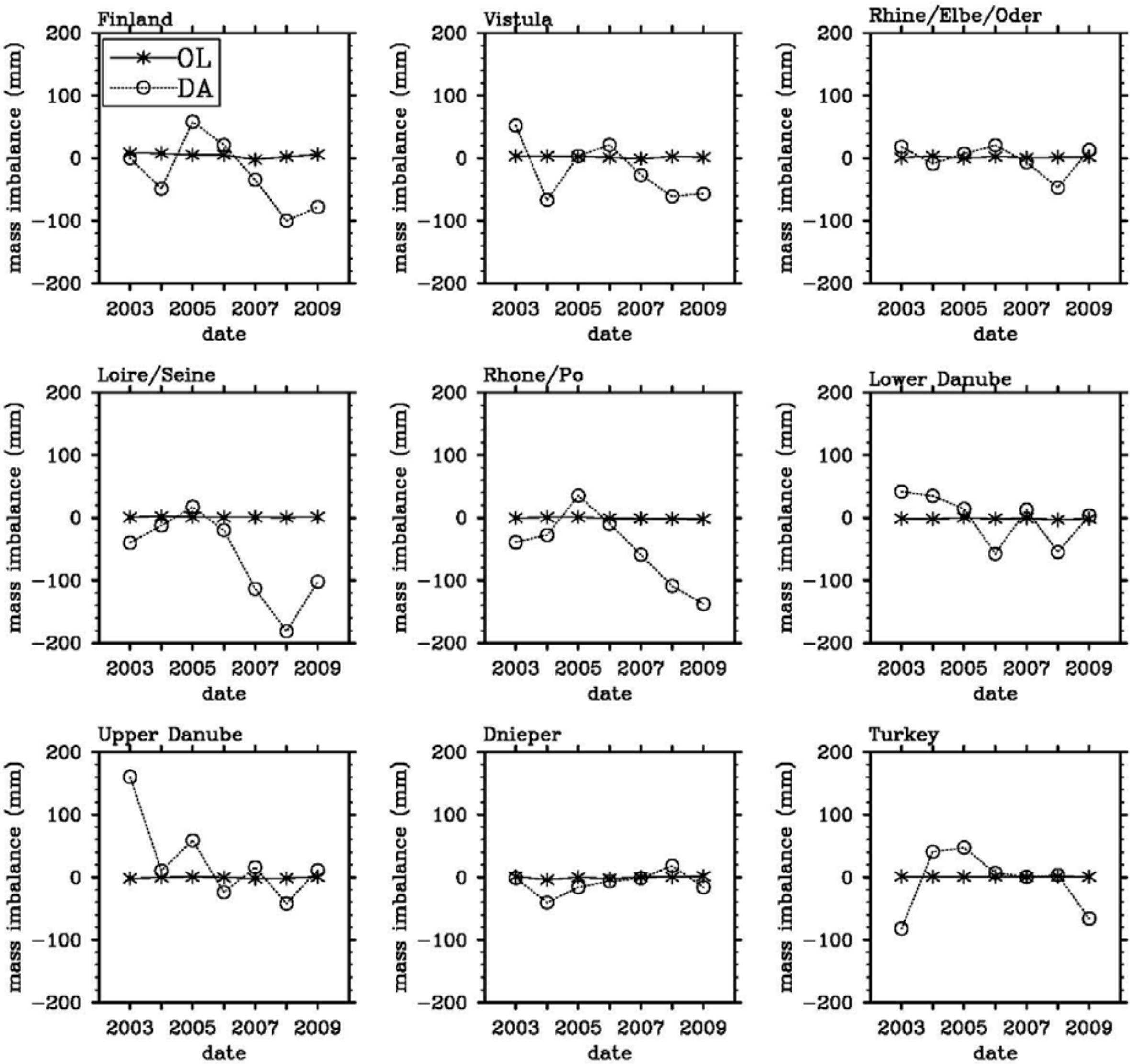

Figure 8. Annual mass imbalance (simulated water budget minus precipitation) for OL and DA in the nine major basins.

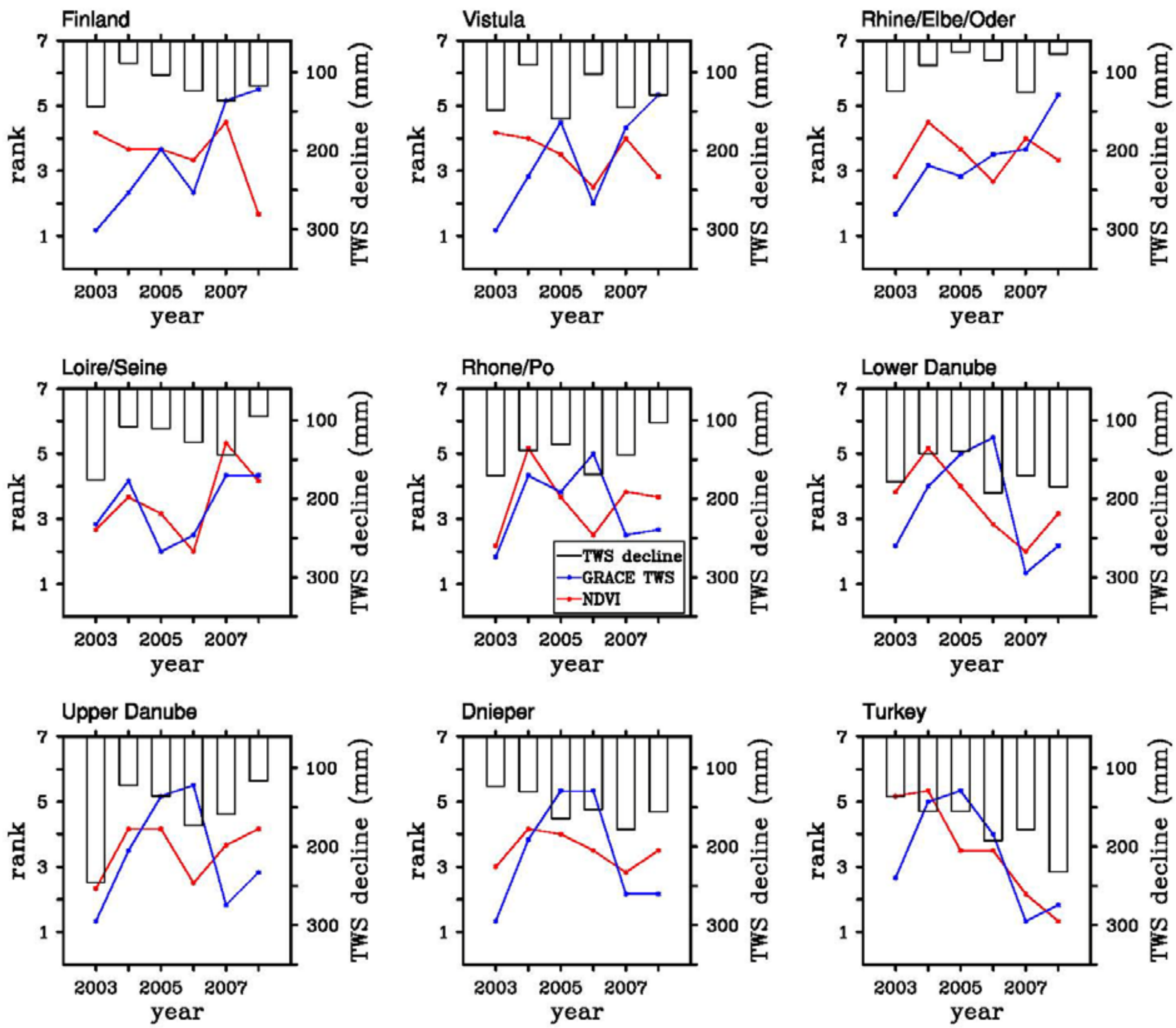

Figure 9. Averaged dryness ranks of NDVI and GRACE TWS for the summer growing season (April to September) during the 2003 to 2008 period and maximum GRACE TWS declines from spring to fall in each year.

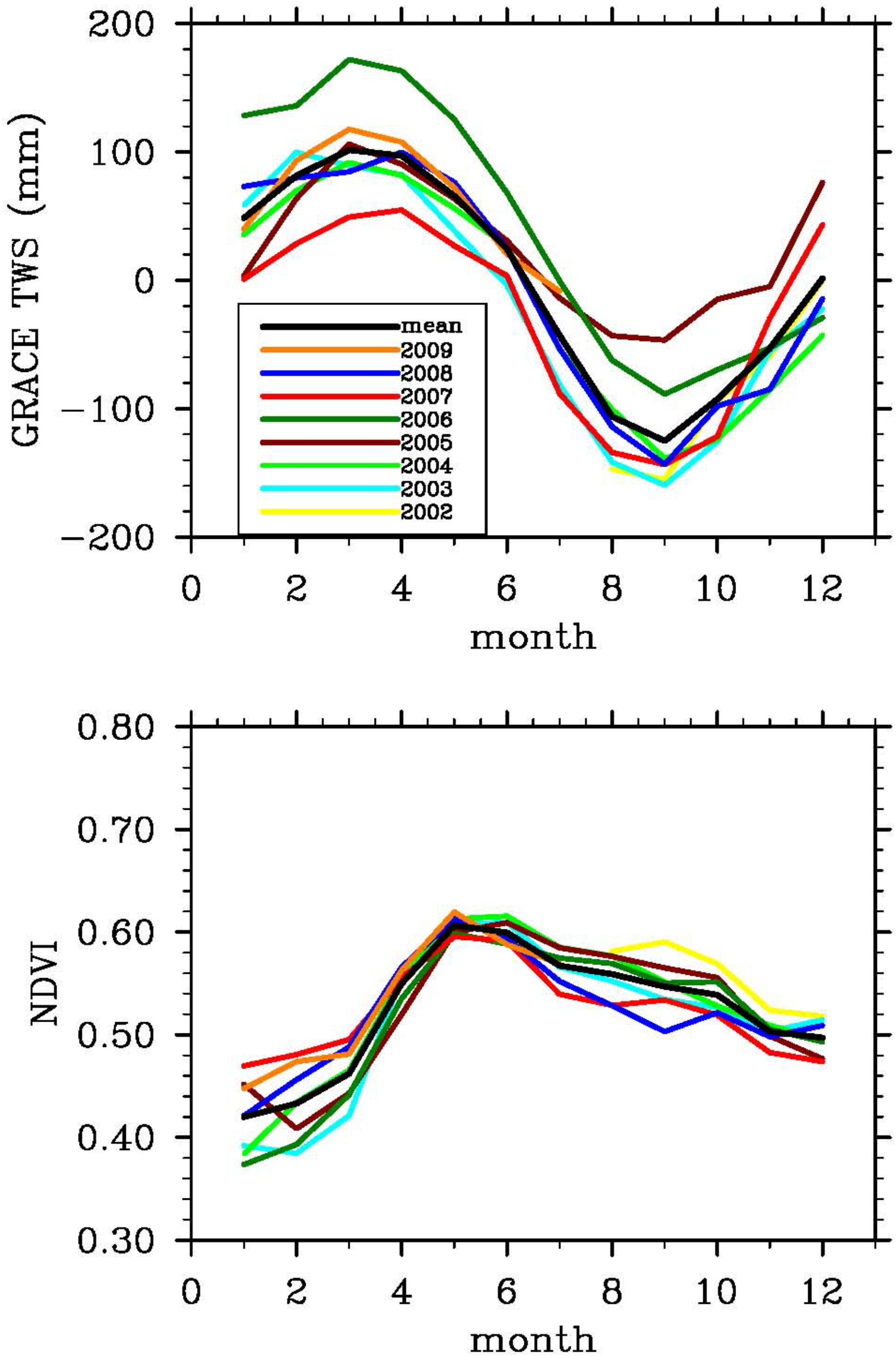

Figure 10. Seasonal cycles of GRACE TWS and NDVI in Lower Danube.

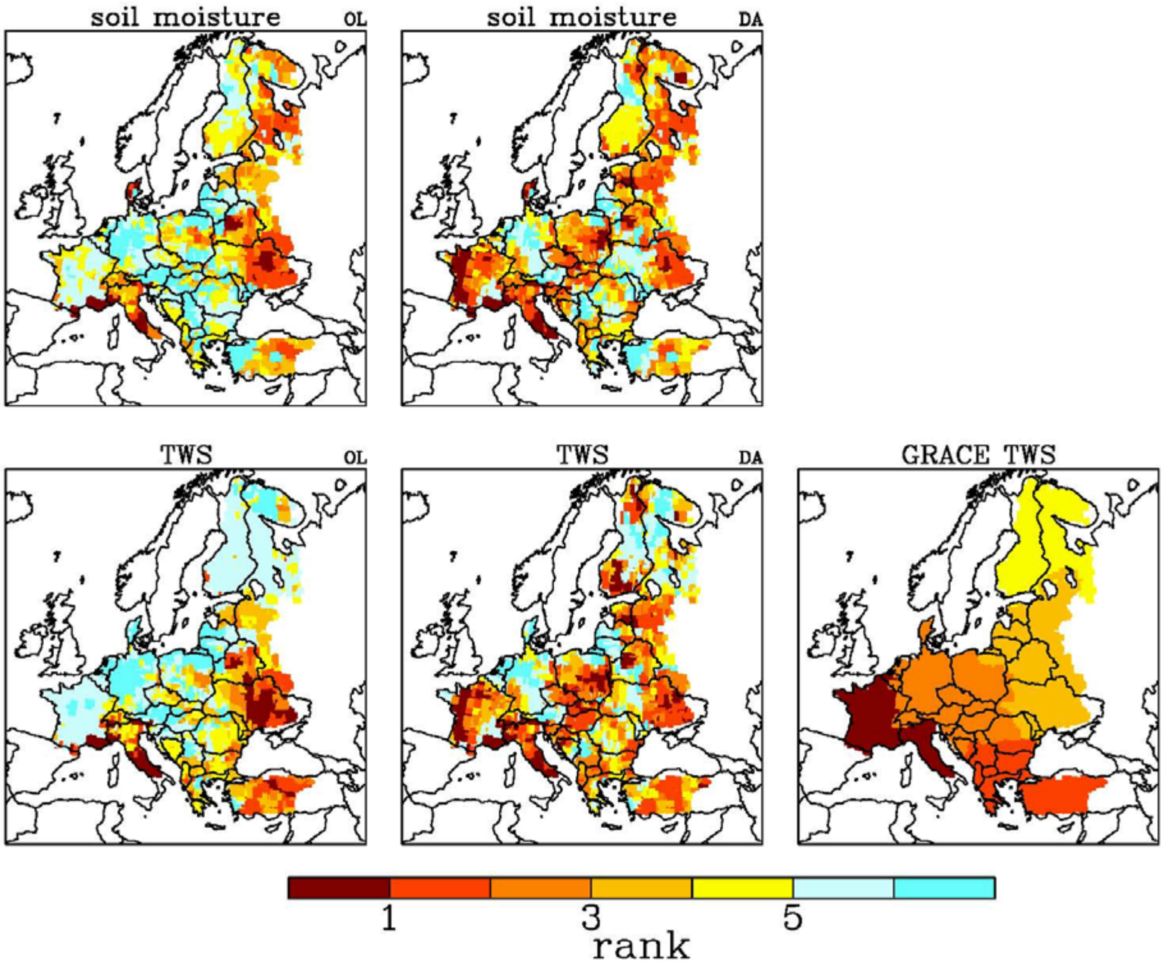

Figure 11. Dryness ranks of simulated root zone soil moisture and TWS for November 2007.

Table1: Major basins and their drainage areas.

Table 2: GRDC stations, drainage areas and record lengths.

Table 3: Correlations of monthly averaged simulated soil moisture with observations at SMOSMANIA sites. Except for the URG site, the OL and DA correlation values are not significantly different at the 0.10 significance level.

Table 1: Major basins and their drainage areas.

basin ID	basin name	Area (km^2)
1	Finland	498000
2	Vistula	547000
3	Rhine/Elbe/Oder	797000
4	Loire/Seine	393000
5	Rhone/Po	319000
6	Lower Danube	503000
7	Upper Danube	490000
8	Dnieper	721000
9	Turkey	403000

Table 2: GRDC stations, drainage areas and record lengths.

Station ID	GRDC number	drainage area (km^2)	record length (months)
	Rhine		
R1	6435060	160800	65
R2	6335020	159300	77
R3	6335050	147600	77
R4	6335060	144200	77
R5	6335070	139500	77
R6	6335100	103500	77
R7	6335150	98200	77
R8	6335180	68800	77
R9	6335170	53100	77
R10	6335200	50200	77
R11	6335400	34600	77
	Elbe		
E1	6340110	131900	77
E2	6340150	123500	77
	Danube		
D1	6742900	807000	77
D2	6742800	709100	84
D3	6742500	658400	84
D4	6742201	570900	77
D5	6242501	101500	53

Table 3: Correlations of monthly simulated soil moisture with observations at SMOSMANIA sites. Except for the URG site, the OL and DA correlation values are not significantly different at the 0.10 significance level.

site	record length (months)	correlation	
		OL	DA
CDM	31	0.75	0.67
CRD	29	0.71	0.76
LHS	26	0.62	0.51
LZC	28	0.68	0.67
MNT	29	0.90	0.90
MTM	22	0.67	0.66
NBN	28	0.44	0.36
PRG	25	0.80	0.77
SBR	31	0.83	0.83
SFL	31	0.67	0.72
SVN	28	0.65	0.56
URG	31	0.81	0.75
Average	31	0.84	0.84