LETTERS
PUBLISHED ONLINE: 6 JULY 2014 | DOI: 10.1038/NGEO2203

nature geoscience

River basin flood potential inferred using GRACE gravity observations at several months lead time

J. T. Reager[1*], B. F. Thomas[1] and J. S. Famiglietti[1,2,3]

The wetness of a watershed determines its response to precipitation[1–3], leading to variability in flood generation[4]. The importance of total water storage—which includes snow, surface water, soil moisture and groundwater—for the predisposition of a region to flooding is less clear, in part because such comprehensive observations are rarely available. Here we demonstrate that basin-scale estimates of water storage derived from satellite observations of time-variable gravity can be used to characterize regional flood potential and may ultimately result in longer lead times in flood warnings. We use a case study of the catastrophic 2011 Missouri River floods to establish a relationship between river discharge, as measured by gauge stations, and basin-wide water storage, as measured remotely by NASA's Gravity Recovery and Climate Experiment (GRACE) mission. Applying a time-lagged autoregressive model of river discharge, we show that the inclusion of GRACE-based total water storage information allows us to assess the predisposition of a river basin to flooding as much as 5–11 months in advance. Additional case studies of flood events in the Columbia River and Indus River basins further illustrate that longer lead-time flood prediction requires accurate information on the complete hydrologic state of a river basin.

In non-frozen catchments, the streamflow hydrograph can be partitioned into 'event flow', a transient response to increases in precipitation forcing that can be highly localized in time and space; and 'base flow', the underlying signal that is a function of the water stored in and slowly released from soils and groundwater within the watershed[5–8]. As a result of observational limitations, many flood prediction methodologies are able to deterministically approach only the first of these phenomena—storm event flow generation and the resultant flooding[9,10]—whereas groundwater-driven base flow variations and their impact on flooding are often neglected. Accordingly, operational hydrological simulations are typically forced with forecast precipitation, some incomplete measure of basin wetness (such as a model estimate of the water content of surface soils), and river level information. This can limit quantitative flood prediction capabilities to the timescales of a weather forecast (that is, in the range of three to ten days), or to probabilistic forecasts[11] motivated by the joint probability of rainfall distributions and hydrologic models. Such approaches present challenges for water management[9] because they may not include all of the components of storage that are important to better characterize regional flood potential[12,13] with greater lead times.

The GRACE satellite mission[14] provides a means to observe monthly variations in total water storage within large (>200,000 km^2) areas such as river basins, based on Earth's time-variable gravity field: when the amount of water stored in a region increases, the gravity signal in that region increases proportionately[15]. GRACE-based estimates of total terrestrial water storage anomalies (TWSA) have provided much insight for the field of hydrology in estimating large-basin river discharge[16] and in revealing large-scale groundwater depletion[17,18]. Importantly, previous studies have noted the existence of a capacity limitation on terrestrial water storage that is associated with regional flooding[12,19].

We present a case study for the Missouri River basin (Fig. 1), where a 500-year flooding event occurred from May to July of 2011 as two storm systems delivered high rainfall amounts to a region that was also being impacted by record snow melt, saturated soils and an elevated water table. The event was characterized as one of the most damaging floods in the past century[20]. Flooding within the Missouri basin and across the greater Mississippi basin resulted in the declaration of a federal disaster area in May[21].

Flows in the Missouri River were represented by 8 US Geological Survey (USGS) gauge stations (Supplementary Figs 1 and 2). These stations were used to estimate a relationship between total water storage from GRACE and the Missouri River response. Although it is logical to simply use the most downstream gauge station to represent total basin discharge, we selected a station that was free from the effects of management, such as sacrificial levee failure[21], which can subsequently impact river discharge. For instance, floodwater diversion practices at the Hermann station (USGS 06934500), downstream of Kansas City, resulted in minimal flood signal in the stream flow record during 2011, masking the underlying relationship between total water storage and river flow volumes. We selected the Missouri River at St Joseph, Missouri (USGS 06818000) as the most

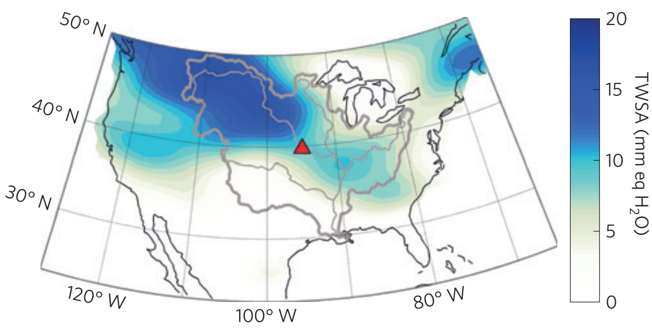

Figure 1 | GRACE water storage anomaly data for March 2011. The greater Mississippi River basin (grey outline) and the St Joseph, Missouri, gauge station USGS 06818000 (triangle) are shown. TWSA, terrestrial water storage anomalies.

[1]UC Center for Hydrologic Modeling, University of California, Irvine, California 92697, USA, [2]Department of Earth System Science, University of California, Irvine, California 92697, USA, [3]Department of Civil and Environmental Engineering, University of California, Irvine, California 92697, USA.
*e-mail: jreager@uci.edu

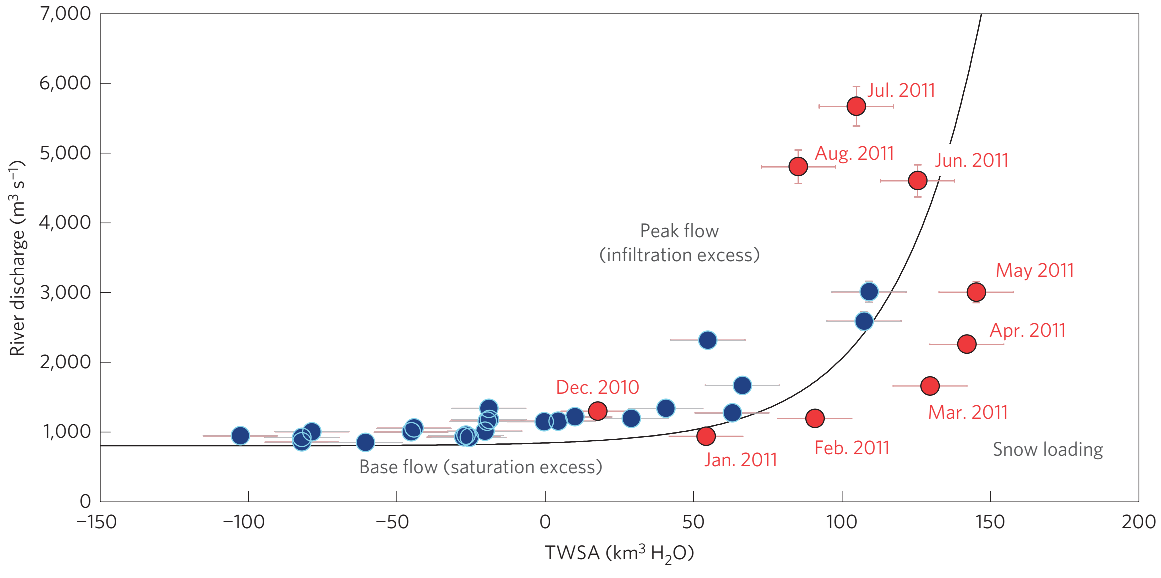

Figure 2 | Observed river discharge versus GRACE TWSA. Non-flood summer months (blue dots), showing the typical relationship between storage and discharge; and December 2010 to August 2011 months (red dots), showing flood event progression. A least-squares exponential regression for the summer months (shifted to intersect the lowest data point) represents a lower threshold on storage–river flow behaviour, as base flow drives increases in discharge with storage. Storage loading due to snow is a positive deviation from the curve in the storage dimension. Peak event flows are positive deviations from the curve in the discharge dimension. Observation errors are shown.

downstream station still preserving an unmanaged flood signal (Supplementary Information).

An averaging kernel approach[22] was applied to extract a time series of GRACE TWSA observations for the Missouri River basin (Supplementary Fig. 3). Monthly errors in the TWSA time series were estimated following previous methods[23].

Figure 2 depicts the relationship between Missouri River discharge and GRACE terrestrial water storage anomalies. The figure shows river discharge at the USGS gauging station at St Joseph, Missouri (USGS 06818000, Fig. 1) and GRACE basin-average TWSA for the entire Missouri River basin (area: 1.37×10^6 km^2) during summer months (June–August) between 2003 and 2012. Summer months (plotted in blue) are used to examine the runoff response to basin storage changes without the influence of snow and frozen ground. Summer river discharge and total water storage generally show an exponential relationship. We include a lower envelope regression line on these points, and suggest that positive deviations from the regression line in the y-dimension result from high-intensity rainfall, localized runoff events and associated peak-flows (similar to infiltration excess runoff processes[6]), whereas high stream flow matched by high storage conditions (points closer to the regression line) are indicative of regional saturation-driven flow (similar to saturation excess runoff events[10]). We propose that these saturation-driven events have greater potential for widespread and damaging regional flooding because of the need for the basin to relieve its saturated state by discharging a greater volume of stored waters into the river channels and surrounding floodplains throughout the basin. We suggest further that the positive deviations from the lower envelope in the TWSA-dimension result mostly from an accumulation of snow and groundwater during winter months, given that increases in storage are not matched by increases in runoff. TWSA leads river discharge slightly in the winter and spring months, creating a hysteresis effect between the two time series. It is this effect that offers information on spring and summer streamflow and the predisposition for flooding.

The relationship between storage and discharge in the months leading up to the 2011 event (red points in Fig. 2) illustrates the impact of total water storage on flood potential, and the eventual need to discharge excess waters under saturated and snow-laden basin conditions. The relationship observed during these months is consistent with the conditions underlying the catastrophic flooding of 2011, namely record high snowmelt, frozen soils and high water table elevations. To the extent that these factors contribute to streamflow within a river basin, GRACE observations provide value in forecasting high discharge conditions; however, the occurrence of a significant flood event may still depend on the combined effects of high storage and strong precipitation.

Our methodology applies the information contained in storage anomalies from GRACE in a simple autoregressive model to determine its utility relative to traditional operational measures of basin wetness such as soil water content (Methods). Model forecasts of streamflow are shown in Fig. 3 for increasing lead times up to 6 months. Autoregressive model performance is measured by two metrics: the prediction of a river discharge greater than the historic (1980–2012) 99th percentile for monthly discharge during the forecast period (3,990 m^3 s^{-1} for USGS 06818000), which describes the ability of the model to recognize a potential high discharge event; and the mean absolute value error (MAE) between model and observations, which describes the model accuracy. MAE was calculated for all models, first for the time before the flood (Supplementary Information), then again for the entire forecast period including the flood. Both metrics are calculated for lead times from 1 to 22 months.

Figure 4 shows the summary of model performance, depicted as discharge maximum versus lead time (Fig. 4a) and MAE versus lead time (Fig. 4b). The addition of the TWSA term to the autoregressive model increases the magnitude of maximum river discharge forecasts at all lead times, relative to no additional indicator, and the ability to forecast an exceptional (greater than the 99th percentile of historic river discharge) event in the 2–5 month lead-time range is demonstrated by the GRACE TWSA model alone. The GRACE model forecasts river discharge near (but not over) the 99th percentile from 6 to 11 months, suggesting the possibility for even longer predictive potential. The addition of GRACE data improves model MAE more than any other tested data, by as much

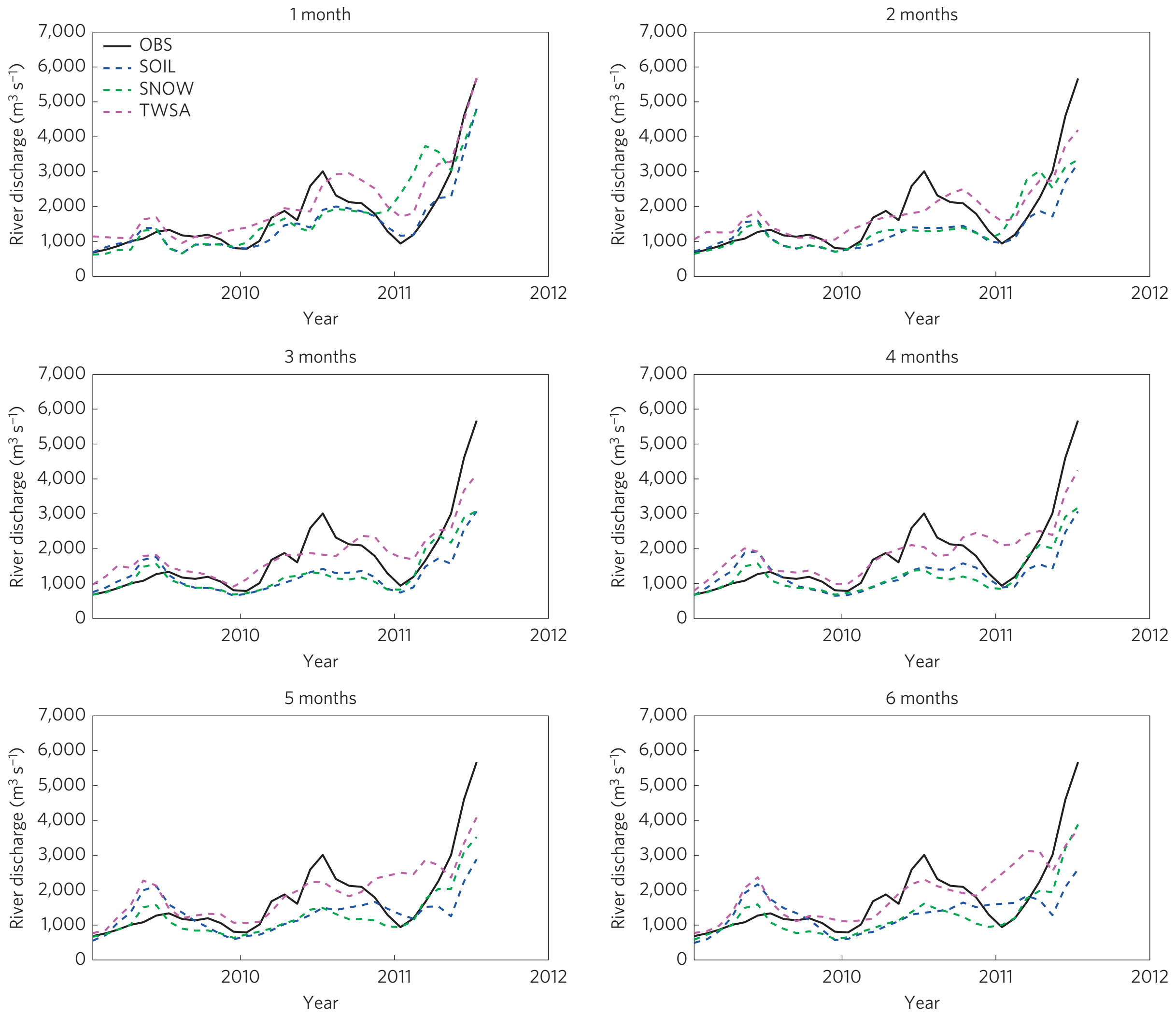

Figure 3 | Autoregressive model results. Results of the autoregressive models, driven by ensemble model soil moisture, model snow water equivalent and GRACE TWSA compared with USGS stream flow observations (OBS) for increasing predictive lead times, from 1 month to 6 months.

as 23% for lead times between 2 and 22 months. During the time closest to the flooding ($\tau = 1$), the TWSA model skill decreases as a result of early warning (over-prediction). During the pre-flood time period (2007–2010), the TWSA model generally has superior performance in predicting river discharge for lead times from 2 to 22 months (Supplementary Information).

Snowmelt was reportedly a large contributor to high streamflow during the 2011 Missouri River basin event[22], and here represents ~15% of the pre-flood, basin-wide water storage variability (Supplementary Fig. 3). The SNOW autoregressive model showed improved MAE over the basic model from 2 to 8 months lead time; however, both the SNOW and SOIL models were not able to predict exceptional river discharge from 1 to 5 months before the event. There is a window from 6 to 7 months lead-time in which the SNOW model predicts high discharge and improves MAE, indicating the importance of that storage component at a specific lead time.

Generally, the model results underscore the importance of water storage contributions from several watershed storage reservoirs, including groundwater, to flooding and flood prediction. To support the idea that groundwater variability is a major portion of the GRACE signal, groundwater observations were taken from 14 wells across the study basin and combined to estimate total basin groundwater variability during 2002–2011 (Supplementary Information). The results show good agreement with Missouri River basin TWSA, and a steady increase from 2007 to the flood event in 2011. Snow water and soil moisture observations are useful flood precursors in the near term (1–2 months), whereas the GRACE data represent all storages, including variations in groundwater storage, in a holistic manner that should ultimately lead to even longer predictive lead times.

GRACE observations are best realized as a flood predictive tool for cases of longer time-scale regional saturation and the accumulation of storage in snowpack and subsurface reservoirs, such as the Missouri River basin example shown. For comparison, we examined two additional case studies of flood events (Supplementary Information). First, in the Columbia River basin, minor flooding occurred in May–June 2011 attributable to late melting snow and high precipitation[24]. The same methodology was applied, resulting in lead times for high discharge of 3 months from GRACE data, outperforming soil and snow information times of 1–2 months. Second, in the Indus River basin, severe flooding occurred in 2010, as the result of sudden, intense storm events in which the basin received as much as 250 mm of rain in a 24 h period[25]. The predictability of monsoon-driven Pakistan floods and similar events that have little snow or groundwater input is not improved using the GRACE data owing to the fast onset of flash flooding.

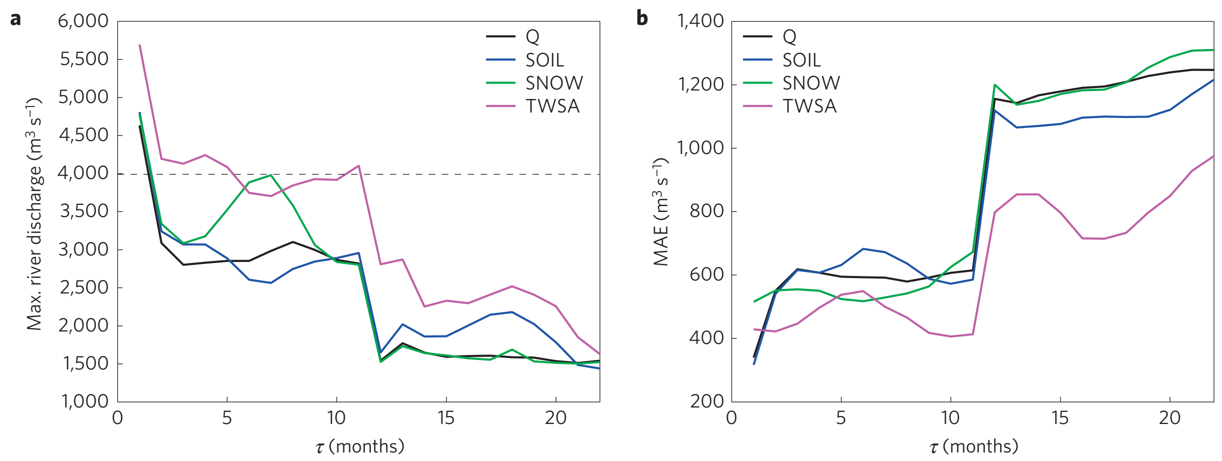

Figure 4 | Autoregressive model performance. a, Maximum predicted river discharge during the forecast period versus model lead time (τ), for the 2-term basic discharge model (Q, black line), and 3-term models driven by soil moisture (blue), snow water equivalent (green) and GRACE TWSA (magenta), with the 99th percentile of historic river discharge (dashed line). **b**, Model mean absolute error (MAE) over the entire forecast period (2009–2011) versus model lead time. The GRACE data improve model accuracy and predict high discharge 5 months before the 2011 flood.

Methods

The GRACE data used here were produced by the Center for Space Research, University of Texas at Austin, and are available from the NASA Jet Propulsion Laboratory Physical Oceanography Distributed Active Archive Center (http://podaac.jpl.nasa.gov). Global Land Data Assimilation System (GLDAS) data are available from the Goddard Earth Sciences Data and Information Services Center (http://disc.gsfc.nasa.gov/hydrology), and stream gauge data from the USGS (http://water.usgs.gov).

Monthly time series were considered as possible regional flood indicators (Supplementary Fig. 5), each with its climatology removed and normalized by its standard deviation: river discharge from USGS 06818000; GRACE TWSA; and GLDAS (ref. 26) ensemble model estimates of basin-average snow water equivalent and soil water storage. All time series end in July 2011, the month of greatest river discharge at the St Joseph gauge. Positive deviations from the climatologies of GRACE TWSA and GLDAS snow-water-equivalent are evident before the 500-year flood event. These deviations point to the predictive quality of the GRACE observations, as they integrate the variability in all relevant water storage reservoirs (snow water equivalent, river storage, soil moisture and groundwater) within the basin and increase steadily for several months before the flood. Time series are divided into a 78-month training period (January 2003–July 2009) and a 24-month forecast period (August 2009–July 2011). Although the flood event began in May 2011, maximum stream flow occurs in July 2011.

We use an autoregressive model for river discharge with additional forcing terms at various lead times. The basic two-term autoregressive model for predicting river discharge as a function of river discharge at a previous time is:

$$Q(t)=a*Q(t-\tau)+b*Q(t-12)$$

where Q is the river discharge at time t, predicted with some lead time, τ (here in months), a climatological component representing discharge from the same month of the previous year (because of the strong seasonality in river discharge), and parameters a and b. To this model, we add a third term representing one of three potential regional flood predictors from those shown in Supplementary Fig. 5. These terms are:

$$Q(t)=a*Q(t-\tau)+b*Q(t-12)+c_{\mathrm{SOIL}}*\mathrm{SOIL}(t-\tau)$$

$$Q(t)=a*Q(t-\tau)+b*Q(t-12)+c_{\mathrm{SNOW}}*\mathrm{SNOW}(t-\tau)$$

$$Q(t)=a*Q(t-\tau)+b*Q(t-12)+c_{\mathrm{TWSA}}*\mathrm{TWSA}(t-\tau)$$

where SOIL represents the GLDAS soil moisture time series, SNOW represents the GLDAS snow-water equivalent time series, TWSA represents the GRACE storage anomaly time series, and c is a coefficient. The entire forecast time series are predicted using the autoregressive models, so that every point in the predicted time series was calculated using data at time τ months before it occurred.

Model coefficients are solved for using a least-squares linear regression during the training period and applied predictively during the forecast period. A nonlinear (exponential) regression yielded slightly better model MAE results, but it unduly complicated the model equations. As the linear GRACE-based autoregressive model showed significant improvement in model accuracy over the base autoregressive model, we deemed it sufficient (versus a nonlinear model) to show the importance of the information contained within the GRACE observations to flood prediction. The goal of this study was not to design the best performing or best optimal autoregressive model. The model framework adopted was used merely to illustrate the value of the GRACE observations, and is therefore not recommended for operational applications.

Received 10 October 2013; accepted 12 June 2014; published online 6 July 2014

References

1. Sayama, T., McDonnell, J. J., Dhakal, A. & Sullivan, K. How much water can a watershed store? *Hydrol. Process.* **25,** 3899–3908 (2011).
2. Kirchner, J. W. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backwards. *Wat. Resour. Res.* **45,** W02429 (2009).
3. Brutsaert, W. Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. *Wat. Resour. Res.* **44,** W02409 (2008).
4. Li, L. & Simonovic, S. P. System dynamics model for predicting floods from snowmelt in North American prairie watersheds. *Hydrol. Process.* **16,** 2645–2666 (2002).
5. Beven, K. J. *Rainfall-Runoff Modeling: The Primer* 360 (John Wiley, 2001).
6. Horton, R. E. *Surface Runoff Phenomena, Part 1. Analysis of the Hydrograph* (Horton Hydrological Laboratory, Publication 101. Edward Bros, 1935).
7. Hall, F. R. Base flow recessions—A review. *Wat. Resour. Res.* **4,** 973–983 (1968).
8. Appleby, V. C. Recession and the baseflow problem. *Wat. Resour. Res.* **6,** 1398–1403 (1970).
9. Wetterhall, F. *et al.* Forecasters priorities for improving probabilistic flood forecasts. *Hydrol. Earth Syst. Sci. Discuss.* **10,** 2215–2242 (2013).
10. Guo, Y., Liu, S. & Baetz, B. W. Probabilistic rainfall-runoff transformation considering both infiltration and saturation excess runoff generation processes. *Wat. Resour. Res.* **48,** W06513 (2012).
11. Siccardi, F., Boni, G., Ferraris, L. & Rudari, R. A hydrometeorological approach for probabilistic flood forecast. *J. Geophys. Res.* **110,** D05101 (2005).
12. Reager, J. T. & Famiglietti, J. S. Global terrestrial water storage capacity and flood potential using GRACE. *Geophys. Res. Lett.* **36,** L23402 (2009).
13. Famigilietti, J. S. & Rodell, M. Water in the balance. *Science* **340,** 1300 (2013).
14. Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. & Watkinse, M. M. GRACE measurements of mass variability in the Earth system. *Science* **305,** 503–505 (2004).

15. Ramillien, G., Famiglietti, J. S. & Wahr, J. Detection of continental hydrology and glaciology signals from GRACE: A review. *Surv. Geophys.* **29,** 361–374 (2008).
16. Syed, T. H., Famiglietti, J. S. & Chambers, D. GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales. *J. Hydromet.* **10,** 22–40 (2009).
17. Rodell, M., Velicogna, I. & Famiglietti, J. Satellite-based estimates of groundwater depletion in India. *Nature* **460,** 999–1002 (2009).
18. Famiglietti, J. S. *et al.* Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophys. Res. Lett.* **38,** L03403 (2011).
19. Crowley, J. W., Mitrovica, J. X., Bailey, R. C., Tamisiea, M. E. & Davis, J. L. Land water storage within the Congo Basin inferred from GRACE satellite gravity data. *Geophys. Res. Lett.* **33,** L19402 (2006).
20. *Spring Flooding Underway, Expected to Worsen through April* (NOAA news website, 2011); http://www.noaanews.noaa.gov/stories2011/20110317_springoutlook.html
21. US Army Corps of Engineers, Post 2011 Flood Event Analysis of Missouri River (USACE, 2012).
22. Swenson, S. & Wahr, J. Methods for inferring regional surface-mass anomalies from Gravity recovery and Climate Experiment (GRACE) measurements of time-variable gravity. *J. Geophys. Res.* **107,** 2193 (2002).
23. Wahr, J., Swenson, S. & Velicogna, I. Accuracy of GRACE mass estimates. *Geophys. Res. Lett.* **33,** L06401 (2006).
24. *Advanced Hydrologic Prediction Service: Portland: Columbia River* (NWS, 2011); http://water.weather.gov/ahps2/hydrograph.php?wfo=pqr&gage=vapw1
25. *Natural Hazards: Unusually Intense Monsoon Rains* (NASA Earth Observatory, 2010); http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=45177
26. Rodell, M. *et al.* The global land data assimilation system. *Bull. Am. Meteorol. Soc.* **85,** 381–394 (2004).

Acknowledgements
This work was supported by grants from the NASA GRACE Science Team, from the NASA Earth and Space Science Fellowship program, from the University of California Office of the President, Multicampus Research Programs and Initiatives, and from the Gordon and Betty Moore Foundation (GBMF3269). We offer a special thanks to S. Swenson of the National Center for Atmospheric Research for GRACE data processing and early discussion.

Author contributions
J.T.R. and J.S.F. conceived the project; J.T.R. designed the study, performed the data analysis and wrote the manuscript; B.F.T. advised on methods and interpretation; B.F.T. and J.S.F. assisted in writing the manuscript.

Additional information
Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.T.R.

Competing financial interests
The authors declare no competing financial interests.