The following document provides a response to the comments provided by the Referees for the HESS Discussion paper:

GRACE storage-streamflow hystereses reveal the dynamics of regional watersheds

This single document contains:

The authors would like to thank both Referees for their comments, which we incorporated into our revised manuscript. This response section provides a detailed response to the comments, or provides the line numbers in the revised manuscript that address the comments. We have also numbered are responses for easier reference throughout the document.

Comments from the Referees are posted in italics.

And our responses are all in a normal font.

Anonymous Referee #1
Received and published: 5 December 2014

This paper addresses:

- The possibility to forecast runoff at certain times during summer from terrestrial water storage in spring measured by GRACE.

- The behavior of TWS and groundwater storage during the seasonal cycle

This study uses GRACE observations of terrestrial water storage observations to expand upon a fundamental concept in watershed hydrology – that the temporal relationship between storage and runoff can be used to quantify complex watershed behavior at broad scales, including groundwater recharge amounts and timing, baseflow recession characteristics, and long lead-time streamflow prediction. The methods is implemented over three catchments of minimum size (for the application of GRACE data) and similar climatic conditions

Interesting highlights in this study:
- Clock-wise behavior of groundwater hysteresis in the presence of a counter clock-wise behavior of the hysteresis for the GRACE signal.

- Prediction of seasonal runoff using GRACE data

1. *However, the focus of the presented results is not reflected in the title of the paper nor in the abstract. Both, title and abstract claim a much more general result with respect to the dynamic behavior of systems. This is not covered in the presented investigations.*

Thank you for your feedback. We have amended the abstract to incorporate your perspective, however we feel that the title is direct and reflects our actual results, and so have opted not to change it.

2. *The claimed methodology cannot be applied to other regional scale studies.*

We respectfully disagree with this statement, GRACE is a global dataset. The Referee is correct in stating that applying the methods we present to other regional watersheds would require some adaptation. But this is the norm for hydrological studies and research in general. Concerns with our methods provided by Referee #1 are included in the revised manuscript.

3. *As the authors mentioned in the text, the paper is based to a large extend on the work by Riegger & Tourian 2014, who have investigated and modeled the runoff–storage(R-S) relationship for large scale catchments in different climatic regions in detail.*

Here in this paper there is a lot of text claiming an explanation for system dynamics by hysteresis only in a qualitative way, mainly repeating the results of Riegger & Tourian yet not being supported by own investigations. There is also a lengthy text trying to explain the deviations from the expected results by anthropogenic management, pumping and local groundwater level measurements.

The work by Riegger and Tourian provided new insights into regional watershed dynamics, which we acknowledge several times throughout our manuscript. To further highlight the contributions of Riegger and Tourian we have added additional commentary in lines 91-95. We have also revised the methods, results, discussion, and conclusion section to address the Referee's suggestions.

We respectfully disagree with the Referee's opinion that we have repeated the results of Riegger and Tourian. Their paper included evapotranspiration as a focus of their research, where we focused on climate, topography, and geology. Topography and geology were not included in the Riegger and Tourian analysis.

The work of Riegger and Tourian (2014) was published prior to this manuscript, and we fully acknowledge their contributions throughout the paper. However, the methods that we applied in this paper were developed independently from the work of Riegger and Tourian. The conceptual idea for this research was presented at a GRACE workshop in 2013 and at the Fall meeting of the AGU later that year.

Sproles, E.A., Leibowitz, S.G., Wigington Jr., P.J., Patil, S., Reager, J.T., and Famiglietti, J.S.: Multi-scale analysis of terrestrial water storage and stream discharge in the Columbia River Basin, Using GRACE Data for Water Cycle Analysis and Climate Modeling Workshop Abstracts, 2013

Sproles, E.A., Leibowitz, S.G., Wigington Jr., P.J., Reager, J.T., Famiglietti, J.S., and Patil, S.D.: GRACE storage streamflow hystereses reveal the dynamics of regional watersheds, AGU Fall Meeting Abstracts, 2013

4. *The investigations presented here are not performed in a sound scientific way and are insufficiently documented, because for an assessment of the forecast potential a comparison of predicted and measured discharge is needed. This comparison should*

be supported by reporting appropriate metrics like RMSE, Nash-Sutcliffe and correlation (The claimed correlation coefficients seem to correspond to the fitting curves and not to the observed data).

The Referee provides an excellent point, and we have augmented our analysis in this revised manuscript. We divided the nine years into two sets of data and used a double pass approach to empirically fit the TWSA-Runoff power function. The first pass used one set of data to fit the model, which was then evaluated against the second set. The data sets were then switched for empirical fit and evaluation.

The results of all 9 years were then used to calculate RMSE, Nash-Sutcliffe, and R2.

We provide a more complete methodology (lines 237-248) and results (lines 325-357) in the updated manuscript (Tables 1 and A3).

5. *No reasons for the selection of points in time are given, neither for the use of TWSA (March) nor for the predictions of Qseason and QAug. No alternatives months are investigated or discussed.*

We have described the selection of months in greater detail in the manuscript (lines 120-123, 231-233, 481-482). In additional to testing for TWSA March and April we have also included February. We have also included runoff for the months of July and September.

6. *If there are anthropogenic impacts in the chosen catchments, why isn't other catchment or time periods used for which no anthropogenic effects occur?*

We appreciate the Referee's comment and agree that ideally regional watersheds that are not managed would benefit this research. However such regional watersheds do not exist in the western portion of the conterminous United States, where this study was focused, or in many areas of the world. The reality is that major dams on the Columbia River date back decades and the GRACE data record dates back to only 2002. Future research could include catchments in Alaska, but were beyond the scope of this study.

7. *Details in the calculation of GWSA are not given nor are the corresponding data sets (time series of mean GW-level, recharge, soil moisture, snow water equivalent, reservoir volumes, TWSA, discharge) displayed. Thus the calculation steps and conclusions with respect to the GWSA hysteresis cannot be retraced. An appropriate visualization of time series of different compartments is needed.*

The description of the methods used to isolate GWSA are provided in 176-180. Additionally, Figure 2 now includes a time series of the different components and the data as a supplemental file.

8. *In the comparison of TWSA vrs point specific well data (Fig.7) no explanation for the inconsistency of GWSA and mean groundwater level is given. Instead further detailed studies are proposed.*

We would like to thank the Referee for identifying this source of ambiguity. We have added additional discussion on lines 462-471.

9. *The authors should clearly describe the behavior of the groundwater system GWSA compared to TWSA and provide possible physical reason for the behavior. For a better understanding, one might think of showing the hysteresis of slow and fast discharge vrs precipitation and different storage compartments like soil moisture, snow water equivalent etc.*

We apologize if the original manuscript was not fully developed. We have added additional methods (GW + Soil Moisture), and the results are provided in lines 283-297). Additional discussion is provided in lines 429-461.

The separation of slow and fast discharge was not a focus of this study, and therefore were not included.

10. *Results and conclusions need to be related to the presented investigations.*

We apologize if the results and conclusions appeared not to be related to investigation. We have updated the manuscript with this goal in mind.

11. *Conclusions have to be explained in detail.*

The conclusions section was written as a brief summary of what was included in the research and their implications. The discussion section was written in a more thorough format, which has been added onto in this revised manuscript.

12. *Title and abstract have to reflect the actual investigations*

Please refer to comment and response #1.

Based on the aforementioned points, the paper can only be accepted after a major revision, in which the aforementioned points are taken into account. Specific Comments:

13. *12029 L14*
    The correlation values seem to correspond to data and fitting curves rather than to predicted and measured runoff.

Please refer to comment and response #4.

14. 12029 L15

This is very general sentence. In fact to apply the same methodology, one should characterize each basin individually. Indeed, this characterization sometimes becomes cumbersome due to the heterogeneous behavior of many large scale basins. In fact, the prerequisite of applying this method is not only availability of GRACE data. So please be precise in your statement.

Please refer to comment and response #2.

15. 12032 L8

There might be confusion in the citation of the respective publication of Reager et al. 2014: In the cited paper the relationship between regional water storage and specific streamflow is not addressed neither the corresponding hysteresis. In the cited paper GRACE TWSA is just used as additional forcing term in an auto-regression approach. Possibly it is from another publication, please cite the correct paper.

Thank you for introducing this topic. We have also included an additional citation (Reager et al., 2009). We still include the Reager et al (2014) citation as this paper uses GRACE as a major component of an analysis focused on specific streamflow event (Missouri River in 2011).

16. 12032 L10

In the paper of Riegger and Tourian 2014 the hysteresis between runoff and storage is described in detail for different climatic zones. For boreal regions they show that runoff is linear to coupled liquid storage. They claim that beside the time lag for runoff the uncoupled solid components of storage are responsible for the hysteresis. However their calculation of runoff from GRACE mass is based on a homogeneous distribution of aggregated snow mass and is not directly applicable to mountainous areas as in-vestigated here. In mountainous areas the snow mass distribution very much depends on local conditions like topography, elevation etc.

We have included a more detailed discussion of the Riegger and Tourian (2014) paper in the updated introduction (lines 95-103).

17. Please report the scale factors of studied basins.

These have been included in the appendix (Table A5).

18. 2035 L2

How the leakage is quantified. The scale factor would only deal with signal attenuation.

Leakage is much more complicated to be quantified by a simple scale factor. Please also describe the measurement error as well. Does the measurement error come from propagating the calibration error of spherical harmonic coefficients?

The scale factors that are applied at 1-degree resolution following the methodology outlined in Landerer and Swenson [2012] are designed to estimate and reduce the "leakage" error in the GRACE solutions. The following assumptions are made:

1) That GRACE estimates of water storage variation contain signal degradation due to the orbital configuration of the satellites and the nature of the observed quantity (i.e. the decorrelation length scale of regional water storage variability) [Wahr et al., 2006]. These "measurement" errors manifest as noise (i.e. random errors that increase in amplitude with increasing spherical harmonic degree) and as systematic errors that are correlated within a particular spectral order [Swenson and Wahr, 2006; Landerer and Swenson, 2012]. These errors are generally improved by truncation and filtering, though this compromises high-resolution signal. These errors should be maintained and propagated through a scaling approach.

2) That the signal attenuation effects of GRACE solution processing on an underlying hydrological signal result from the truncation of spatial harmonics at a degree and order selected to minimize noise and maintain spatially-variable signal for each monthly solution, and from the algorithmic filtering of the resulting harmonics to remove systematic correlated errors ("stripes" described in Swenson and Wahr [2006]) and also random errors (300km gaussian). This cumulative error is referred to as "leakage" and depends on the filtering process as well as the characteristics of the original signal. Here, a truncation to degree and order 60 was performed before application of the destriping and gaussian filters mentioned.

3) That there is some hypothetical "true" global hydrological state that can be estimated by a land surface model (LSM: a numerical simulation) that is forced by precipitation and solar radiation, and modulated by various and land surface characteristics (topographic variability, vegetation type, soil type, etc.). Here, this simulation is performed at 1-degree resolution by the NCAR CLM model as described on the TELLUS website (grace.jpl.nasa.gov).

4) That the errors resulting from these "leakage" effects can be modeled to 1-degree resolution by placing the monthly 1-degree LSM simulation outputs through processing equal to that performed on the GRACE spherical harmonic solutions (i.e. truncation and filtering). These errors can then be quantified by measuring the signal loss between the original simulations and those processed with "leakage" effects. Scale-factors can be calculated based on a least-squares minimization of the residual between the unprocessed and processed time series.

In summary, because the spatially-distributed scale factors have a sub-GRACE-resolution structure, and because scale factors can have a value of less than 1, they attempt to recover the best-case underlying hydrological signal according to the assumptions above. This is how the spatial scale factors estimate and correct signal attenuation caused by "leakage" errors. There is however some residual uncertainty associated with this process, and this process does not remove all of the "leakage" error in the GRACE observations — it only acts to reduce it. Conversely, the measurement error is actually amplified through the scaling process, along with the TWSA estimates. Of course these error estimates are both included in the regional application of GRACE TWSA data.

It is necessary to use some sort of regional scaling in the application of GRACE observations for hydrology [e.g., Famiglietti et al., 2011; Landerer et al., 2010; Swenson and Wahr,2007; Klees et al.,2007; Chen et al.,2007]. That is because GRACE observations are intrinsically more coarse than other hydrologic observations (e.g. stream gage), and a failure to account for and represent signal loss in TWSA would result in a mismatch of data [Swenson et al., 2003], leading to results that are based only on proportionality and are consequently of limited utility/accuracy [e.g. Riegger and Tourian, 2012; Tang et al., 2010].

All GRACE results, as the Referee has identified, can also possess complex errors, such as the existence of neighboring regions that are out-of-phase (e.g. the Orinoco and the Amazon) whose signals can interact. For this reason, any scaling approach needs to include some representation of measurement and leakage errors.

The 1-degree scale factors represent the opportunity to apply GRACE observations at a scale that is consistent with the hydrology simulations in global climate models (i.e. LSM's) and with many other global gridded hydrology data sets, and the error estimates that accompany them represent a more robust approach to quantifying potential limitations in hydrologic applications of GRACE.

The goal of the paper was not to provide more research on leakage and scale factors, and thus will include the same brief description with references as found in the initial manuscript.

Refs:

Chen, J. L., C. R. Wilson, J. S. Famiglietti, an d M. Rodell (2 007), Attenuation effect on seasonal basin-scale water storage changes from GRACE time-variable gravity, J. Geodesy, 81, 237–245, doi:1 0.1007 /s00190-006-0104-2.

Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed,S. C. Swenson, C. R. de Linage, and M. Rodell (2011), Satellites measurere rates of groundwater depletion in California's central valley, Geo-phys. Res. Lett., 38(3), L03403, doi:10.1029/2010GL046442.

Klees, R., E. A. Zapreeva, H. C. Winsemius, and H. H. G. Savenije (2007),The bias in GRACE estimates of continental water storage variations, Hydrol. Earth Syst. Sci., 11, 1227–1241.

Landerer, F. W., J. O. Dickey, and A. Guntner (2010), Terrestrial wate rbudget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003–2009, J. Geophys. Res., 115(D23), D23115, doi:10.1029/2010JD014584.

Tang, Q., H. Gao, P. Yeh, T. Oki, F. Su, and D. P. Lettenmaier (2010),Dynamics of terrestrial water storage change from satellite and surface observations and modeling, J. Hydrometeorol., 11(1), 156–170,doi:10.1175/2009JHM1152.1.

Swenson, S., and J. Wahr (2002), Methods for inferring regional surface-mass anomalies from gravity recovery and climate experiment (GRACE)measurements of time-variable gravity, J. Geophys. Res., 107(B9), 2193,doi:10.1029/2001JB000576.

Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33(8), L08402, doi:10.1029/2005GL025285.

Swenson, S., J. Wahr, and P. C. D. Milly (2003), Estimated accuracies of regional water storage variations inferred from the gravity recovery and climate experiment (GRACE), Water Resour. Res., 39(8), 1223,doi:10.1029/2002WR001808.

Wahr, J., S. Swenson, and I. Velicogna (2006), Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33, L06401, doi:10.1029/2005GL025305.

19. *12036 L9-17*
    The simultaneous display of TWSA and GWSA in one figure for each catchment might help to highlight the different dynamical behavior.

Thank you for the suggestion. We tried having the TWSA and GWSA in the same graphic, but found it visually confusing. We have also added subfigures for subsurface water, which highlight the combined signal of GW and soil moisture.

20. *12036 L22 Similar confusion in the citation of the respective publication of Reager et al. 2014 is seen as above (12032 L8).*

We would like to thank the Referee for their perspective regarding this article. Figure 2 in Reager et al., 2014 represents the relationship between TWSA and Q, and highlights the 2011 flood events in the Missouri River. Thus we have included it in the manuscript.

21. *12036 L22*
    In order to describe the systems independently from catchment area I propose to use runoff rather than discharge for a comparison

We would like to thank the Referee for this comment. We did use runoff throughout the study, and have clarified the labeling in the updated manuscript.

22. 12037 L25 – 12038 L8 To support this paragraph I suggest to show time series of soil moisture or distribution of snow coverage in time.

This suggestion has been incorporated into Figure 2 of the updated manuscript.

23. 12039 L14 Fig.3b shows the total runoff i.e. surface runoff and baseflow vrs. GWSA. Thus total runoff should be separated into its fast and slow components and each of them being displayed vrs. GWSA. The Clock-wise behavior of groundwater hysteresis is one of interesting finding of this study. Therefore, the authors should provide a physical explanation for that. Explanation of Fig 3b is not clear: Vertical branch Jun-Oct: How can runoff decrease with a nearly constant GW storage? Left branch Oct – Mar: how can runoff increase with a decreasing GW storage? Is it matter of surface runoff? For an understanding of the behavour it is essential to display time series of SM, SWE, RES, TWSA and RGW and RSW. A display of precipitation and evapotranspiration would be very helpful.

These suggestions have been incorporated into the updated manuscript. We have added additional methods (incorporating GW + Soil Moisture), and the results are provided in lines 283-297. Additional discussion describing these processes is provided in lines 429-461.

The separation of runoff into fast and slow components is an interesting idea, but fall outside the scope of this paper and therefore was not included.

24. 12039 L19
Confusion about Fig. numbering: possibly Fig 7 is meant

Thank you for identifying the mislabel. This has been corrected in the revised manuscript.

25. 12039 L21–23
To me taking TWSA_march for prediction looks like cherry picking. Why not March and not February for TWSA or June, July for discharge, for instance? Please provide reason for taking this month!

Please refer to comment and response #5.

26. Also, confusion about Fig. numbering: possibly Fig 6 is meant It is not clear from the text how stream flow is predicted and how this is related to Fig6: Qseason in Fig6: is it the mean observed discharge or predicted discharge? If it is the later, how is it calculated?

We apologize for any confusion. This is the observed discharge and text describing this has been added in the revised text caption.

27. *Are the correlations displayed in Fig6 and Table 1 the correlations between the measured Qseason and the fitted curves (power functions with which parameters?) or between the measured Qseason and TWSA (the high correlation value rather represents the curve fit than Qseason vrs TWSA).*

The correlations displayed in Figure 7 of the original manuscript represent the fitted curves, as do the metrics in the figure. Tables 1 and A3 provide a more detailed look at the results from testing the empirical fit using a double-pass calibration and validation described in the updated methods (lines 226-335). Table 2 and A4 provide the metrics for the complete data set.

28. *The fitting curves in Fig6 do not already represent predictions, yet are the basis for predictions using measured TWSA to determine forecast discharge (via the curve fits) (as Fig6 shows, that Qseason and QAug are not very much depending on TWSA for smaller values of TWSA, yet only for bigger values. The calculation scheme is essential for an assessment of the method). For an evaluation of the predictive potential of the investigated methods the parameters of the fitting curves determined on a training period should be used to calculate predic tions of discharge in an independent prediction period. The predicted discharge values should then be displayed versus the measured for the prediction period in a scatter plot and correlations should be calculated for forecasts vrs measured.*

    As a predicted value should always be better than a simple use of mean monthly values from the training period of forecasts, on top of conventional Nash Sutcliffe (NS) coefficient, in which the values are assessed w.r.t long term mean, the NS coefficient w.r.t. the seasonal signal (using the monthly residuals of the training period) should also be presented in table1 (NS_cycle: in the denominator instead of \bar{Qo} you should use monthly mean).

The Referee provides an excellent point, and we have augmented our analysis in this revised manuscript. We divided the nine years into two sets of data and used a double pass approach to empirically fit the TWSA-Runoff power function. The first pass used one set of data to fit the model, which was then evaluated against the second set. The data sets were then switched for empirical fit and evaluation.

The results of all 9 years were then used to calculate the standard hydrological metrics RMSE, Nash-Sutcliffe, and R^2.

We provide a more complete methodology (lines 224-248) and results (lines 325-357) in the updated manuscript (Tables 1, 2, 3, A3, A4).

12040 L1
Again why August and no other month? How a reader should follow the story here? What does a seasonal average/aggregation of discharge mean for possible applications?

Please refer to comment and response #5.

29. 12041 L5–29

Over boreal regions the R-S hysteresis is determined by (Riegger and Tourian 2014):
-Climatic impacts i.e. the relative importance of aggregated solid precipitation (repre-sented on the lower branch)
-The runoff time constant determining the slope of the linear part (upper branch)
-The time lag between mass and runoff being responsible for (a smaller) part of the hysteresis
The different forms of the hysteresis thus can be explained by the corresponding hydraulic time constants, which is shortest for steep slopes and fractured systems. This explains that the upper branch is steeper for the Upper Columbia than for the Snake River. This should be considered and discussed in this section.

We thank the Referee for this suggestion. We have addressed this concern using an alternative approach. We removed only the Snow and Reservoir signal to examine watershed hysteresis, and we tested the ability of SWE to predict Runoff.

The findings are included throughout the results and discussion (for example lines 375-399).

30. 12042 L10–17

See comment above on prediction accuracy in table 1 There are 9 years of measurements available. This period could be split into a training and a prediction period for a better estimation of prediction accuracy

We have augmented our analysis in this revised manuscript. We divided the nine years into two sets of data and used a double pass approach to empirically fit the TWSA-Runoff power function. The first pass used one set of data to fit the model, which was then evaluated against the second set. The data sets were then switched for empirical fit and evaluation.

The results of all 9 years were then used to calculate RMSE, Nash-Sutcliffe, and R2.

For more, please refer to comment and response # 28.

31. 12042 L18

Probably Fig 6, and Fig 7 is meant

We have addressed the mislabeling.

32. 12042 L21-26

The fact that Q is insensitive to TWSA < 100mm is reflected by the curves in Fig 6. This means that prediction could only be made for TWSA > 100mm. If these catchments are managed, a reliable prediction from TWSA cannot be made!! So either other, un-managed catchments have to be chosen or the authors should only consider the time periods with no management. The explanation of the Q – TWSA *relationship for TWSA < 100mm by water resources management is not sufficient to explain quantitative effects.*

We appreciate the Referee's comments and agree that ideally regional watersheds that are not managed would benefit this research. However such regional watersheds do not exist in the western portion of the conterminous United States, where this study was focused.

From a data perspective, we are also limited with regards to the length of the GRACE data record. However the available data demonstrates a threshold behavior for total seasonal runoff in each of the three regional watersheds when TWSA in March is less than 100 mm (lines 492-494). When TWSA in March is above 100 mm, GRACE measurements also provide a high degree of skill in predicting season total seasonal runoff.

From a scientific perspective these findings describe the measured relationship between storage and runoff at regional scales. These insights identify potential ways to apply in understanding watershed behavior and predicting streamflow.

From a management perspective, the skill of TWSA from a single month to predict seasonal streamflow provides value as well. It would allow managers to know how with a fairly high level of confidence how much water to expect in the system for the remainder of the water year. This is important in the Columbia River Basin where management balances protecting endangered species in a region with extensive industrial agriculture and hydropower generation. Since the stated goals of HESS are "the advancement of hydrologic science … to serve not only the community of hydrologists, but all earth and life scientists, water engineers and water managers," we think such an emphasis is appropriate.

Even in a managed basin such as the Columbia River you will have years of high and low flows, each of which can introduce management challenges (e.g., http://www.oregonlive.com/business/index.ssf/2011/05/bpa_curtails_wind_farm_generat.html; http://www.forbes.com/sites/jamesconca/2012/08/05/hydro-forced-to-take-a-dive-for-wind/, http://wdfw.wa.gov/news/aug2101a/).

The management of the Columbia River Basin is regulated by a treaty between the United States and Canada dating back to 1964, which is currently being renegotiated.

Our approach in this study was not to let what we cannot do limit what we can do.

33. 12042 L21-26 –12043 L11
Lengthy text, clear quantitative consequences are missing

Thank you for the suggestion, and we have shortened this text (lines 498-501).

34. 12043 L16-19

Explanation of GWSA-Q not understood (see also above)! During the winter period in snow covered areas discharge is released only from the groundwater system. The groundwater recharge from the surface in this case is zero, i.e. GWSA should decrease with Q. How is distinguished between surface runoff from snowmelt and runoff from groundwater. Possibly the behavior of the model-based calculation of soil water content helps.

Thank you for this comment, and we have included changes in subsurface water (soil moisture + groundwater) in our revised manuscript.

35. 12043 L25

GWSA is nearly constant from June to October. Why? This does not fit to the timing of pumping test! What is the purpose of the pumping tests mentioned in this paper?

We apologize if this was unclear. The purpose of including pumping tests is the timing of tests for management purposes corresponds to the highest groundwater levels. This could potentially lead to management strategies that do not capture periods of lower groundwater levels. We are publishing this paper from a government agency, and thus cannot include recommendations regarding changes to regulations. However we can highlight ways that understanding systems could potentially be improved.

36. 12044 L3

Probably Fig 7 instead of 6 Meaning and conclusion of Fig 7

Thank you, we have corrected the label.

37. 12044 L3–10

GWSA does not fit to the overall GW-levels from observations. So either the calculated GWSA is wrong or the selection of observation wells is not representative for the general GW level. GWSA correspond to the total volume of the groundwater system and not to groundwater levels!

The GW-storage coefficient determines the relationship between volume and level and is not mentioned here as it is probably not known on catchment scale. There is no scientific consequence from this observations mentioned in the paper. If this part is presented here, there should be a more detailed description of GW-level measurements, storage coefficients from hydrogeology, selection of observation points and a detailed discussion of the results and consequences for the message of the paper.

Thank you for introducing this discussion. We also understand that GWSA data (volume) do not equal well levels, and for this reason we normalized the standard deviation across the time series to compare the temporal signal. We concur with the Referee that the groundwater observation well levels provide no relationship with GWSA data. We debated whether or not to remove this component of the research. In the end we included it as the lack of correlation highlights the variability of site characteristics across a region, which is of scientific consequence.

The revised discussion (lines 454-471) addresses these points.

38. 12045 Conclusions
The conclusion should represent the conclusion of this work. In case the authors would bring in arguments from other published works the bridges between studies should be clear. The last paragraph of conclusion is too general and does not reflect the results of the study.

We have included updated results in the conclusion.

39. 12045 L10–13
Please report the prediction results here including RMSE, NS, NS_cycle, correlation

We have included model performance metrics in the conclusion.

40. 12045 L23 Please provide a citation for the background research in the text, otherwise please rephrase the sentence as this statement contradicts with your earlier statement in 12040 L22.

Thank you for pointing out any potential ambiguity. The statement in acknowledgements is factually accurate. The approach we present was developed concurrently, but independently from the research provided in Riegger and Tourian (2014). Because their work was published prior to this paper, we have cited their efforts in the research portion of the paper. Again our approach was developed independently, and implying that our approach was developed from Riegger and Tourian is not factually accurate.

Anonymous Referee #2
Received and published: 11 December 2014
General comments:
In its exploration of the use of GRACE gravity data for improved understanding of the hydrology of regional watersheds this paper presents intriguing results and points a way forward to further applications of this approach. To that extent it appears to merit publication.

This paper can be viewed within the general context of an increasing attention for changes of water storage in a watershed as the driver of streamflow. The classical rainfall-runoff models of hydrology avoid the obvious fact that streamflow is driven by storage in the watershed and not by precipitation as such, but the intermediate storage change step was skipped because there were no adequate means to observe storage changes in most regional watersheds. New observation techniques such as GRACE allow closer consideration of storage.

The prognostic ability demonstrated in this paper of the GRACE signal to predict seasonal runoff is impressive and indicates a potential for the use of GRACE results to enhance the reliability of seasonal water supply predictions.

Specific comments

41. *The distinction between soil moisture and groundwater appears rather arbitrary and re-quires more scrutiny. It might be preferable to combine the two as subsurface moisture storage because for much of these watersheds soil moisture changes on a monthly time step are likely to be closely linked to groundwater storage changes. The groundwater storage includes water storage in the capillary fringe above the water table and the top of the capillary fringe is likely to be above the 2000 mm below ground level over much of the time and space of the analysis. Thus it is possible that the soil moisture changes as estimated in this paper include much of the groundwater storage changes. That may be one of the reasons why the modeled groundwater storage changes appear to be small and have almost no correlation with the observation well records.*

Thank you for this suggestion. We have added a separate subsurface water signal (TWSAsub = TWSA – SWE – RES) into our revised manuscript.

42. *Fig 7. The almost total lack of correlation between the groundwater levels and TWSA serve to underline the questionable assumption that GWSA can be estimated from TWSA – SWE – SM. Likely the problem lies both in the uncertainty of the SM estimates and the high variability of groundwater dynamics across the whole basin from low to very high elevations. It is also likely that the groundwater observations are practically all for valley bottoms and do not represent the GW storage at higher elevations and on steep slopes. By contrast the TWSA is dominated by high-elevation snow.*

The paper should include plots, analysis and discussion of the changes of SWE and SM (or of SWE and SM+GW). These are just as critical an aspect of the components of the TWSA as the estimated groundwater changes. One would expect that the SWE can be validated fairly well on the basis of various point observations, at least much better than SM.

Thank you for this suggestion. We have included it into our revised manuscript as $TWSA_{sub}$.

43. *It would be intriguing to attempt a water balance for the watersheds by including pre-cipitation estimates. Since evaporation is relatively minor during the winter months P*
TWSA ~ Q for the winter and this would provide a test of the consistency of these components with the conservation of mass. However, this perhaps such analysis lies outside the scope of the present paper.

We agree on both parts of this comment. It would be interesting, but it does in fact lie outside the scope of this paper.

Technical comments:

44. *P. 12029 L 22. Topography is clearly a major watershed descriptor apart from climate and geology, as also implied in this paper by the contrast between the steep slopes of the Upper Columbia basin and the relative flatness of the Snake River watershed.*

We have added the role of topography as a watershed descriptor throughout the manuscript. Thank you for the suggestion, it helps frame our discussion all the better.

45. *P 12033 LL 16-18. This characterization of aquifer storage capacity of the two watersheds is rather off-hand, and without any further explanation and references. The "well-developed soils" of the Snake River basin are perhaps relevant to soil moisture storage but not to aquifer storage which depends on the nature of the underlying subsoil and bedrock. Do the Snake River basalts in fact have much effective porosity at the water table (see also p. 12041, L 25)? The results shown later in the paper for groundwater storage changes in the Snake River watershed suggest very low groundwater storage capacity.*

Thank you for identifying any ambiguity. While the groundwater storage changes in the Snake River are more constrained, we interpreted this to represent that there is less change from max GWSA to min GWSA. This represents the fluxes of water moving through the system more slowly in the Snake. In part because 1) it is flat, 2) it is dry, 3) the aquifer can hold the water. This combination constrains the GWSA anomaly as compared to the Upper Columbia where 1) It is steep, 2) it is wet, 3) the aquifer cannot hold the water. It moves large fluxes of water through it, which expands the range of GWSA.

We have also added additional text in the manuscript to describe these processes in section 5.1 of the manuscript.

46. P 12035 LL 20-25. In view of all the uncertainties in measuring or estimating regional soil moisture, as summarized in the introduction to this paper, these GLDAS-derived estimates of soil moisture are surely highly tentative at best. This would appear to be a very uncertain foundation for estimating changes of groundwater storage. The error estimate for SM is not adequately estimated on the basis of the "monthly standard deviation" (p. 12036, L 11) because there are likely large biases in the GLDAS model algorithms.

We understand the Referee's concern that there may be issues with large-scale modeled soil moisture, but this comment does not offer a direction forward. The Referee seems critical of the state-of-the-science of hydrology and of an already-tested approach, and not of this specific work. Of course, there have been numerous (~30) studies published using GRACE to detect groundwater variability globally, and all present a similar, if not less rigorous, methodology than the current work. Many of those studies have used well data to validate their results (e.g. Scanlon et al., 2011).

The NASA LDAS simulations used here were driven by observed precipitation and radiative forcing. Whether the model does a good job with soil physics and the parameterization of sub-grid scale processes is another issue: as explained in the text, we attempt to address model structural error by using an ensemble of model outputs, to give us an estimate of the bias contained in a given model solution relative to others, which we then propagate through our calculations. We also represent a range of possible scale-variance uncertainty by considering the GLDAS simulations (at 1-degree resolution) and the NLDAS simulations (at 1/4 degree resolution). For large areas (i.e., an area equivalent to that of the GRACE observations) this should be entirely appropriate., and it is the best methodology we know to be available.

A note on bias: because we are dealing with terrestrial water storage variability (anomaly), and specifically not a mean-value, the typical pathway for model bias to manifest (e.g. through differing numbers of soil layers) is eliminated by removal of the model soil moisture mean. We are really only considering variability in soil moisture. The use of multiple models in an ensemble provides a large range of uncertainty that can be propagated through groundwater calculations. Also, the removal of soil moisture tends to reduce the variability in GRACE time series, especially at the seasonal period. This is a likely outcome considering soil moisture memory with depth.

Finally, there is really no alternative approach, as soil moisture observations across the domain are not available. If they were, then scaling or interpolating individual point measurements would arguably contribute as much or more error than a distributed model driven by observed forcing. While there is certainly space for continued work in hydrology on soil moisture variability and scale variance and invariance, it is really beyond the scope of the current manuscript to develop these analyses. Instead we rely on already proven, peer-reviewed methodologies to highlight new observations in the Columbia River basin. We humbly invite the Referee to collaborate on future refinement of these methods.

47. 12036, L 17 Insert "error" as in "individual ERROR components".

Thank you, and we have added this in the updated manuscript.

48. p. 12039 L 13. hardly "dramatic" since this is an obvious consequence of snow accumulation.

We agree and have changed the word to "distinct."

49. p. 12041, L 25. It is not obvious that the basalt provides excellent aquifer storage. The basalts provide excellent transmissivity for groundwater flow and discharge, but that is not the same as storage and in fact would go to counteract large changes of groundwater storage, as indeed is suggested by the analysis results of this paper (see Fig 2e).

Thank you for identifying this ambiguity and we have changed it in the updated manuscript (lines 400-403).

GRACE storage-runoff hystereses reveal the dynamics of regional watersheds

E.A. Sproles[1,2,3]; S.G. Leibowitz[4]; J.T. Reager[5]; P.J. Wigington Jr[4,6]; J.S. Famiglietti[5]; S.D. Patil[7]

*corresponding author: eric.sproles@gmail.com
[1] Oak Ridge Institute for Science and Technology, c/o National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR, 97333, USA
[2] Current address: Centro de Estudios en Zonas Áridas, Universidad de La Serena, Raul Bitran 1305, La Serena, Chile
[3] College of College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, 97331-5503, USA
[4] National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR, 97333, USA
[5] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, 91109, USA
[6] Retired
[7] School of Environment, Natural Resources and Geography, Bangor University, Bangor, LL57 2UW, Wales

Abstract

We characterize how regional watersheds function as simple, dynamic systems through a series of hysteresis loops using measurements from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites. These loops illustrate the temporal relationship between runoff and terrestrial water storage in three regional-scale watersheds (>150,000 km^2) of the Columbia River Basin, USA and Canada. The shape and size of the hysteresis loops are controlled by the climate, topography, and geology of the watershed. The direction of the hystereses for the GRACE signals move in opposite directions from the isolated groundwater hystereses, suggesting that regional scale watersheds require soil water storage to reach a certain threshold before groundwater recharge and peak runoff occur. The subsurface water (soil moisture and groundwater) hystereses more closely resemble the storage-runoff relationship of a soil matrix. While the physical processes underlying these hystereses are inherently complex, the vertical integration of terrestrial water in the GRACE signal encapsulates the processes that govern the non-linear function of regional-scale watersheds. We use this process-based understanding to test how GRACE data can be applied prognostically to predict seasonal runoff (mean Nash-Sutcliffe Efficiency of 0.91) and monthly runoff during the low flow/high demand month of August (mean Nash-Sutcliffe Efficiency of 0.77) in all three watersheds. The global nature of GRACE data allows this same methodology to be applied in other regional-scale studies, and could be particularly useful in regions with minimal data and in trans-boundary watersheds.

1. Introduction

At the most fundamental level, watershed processes can be described as the collection, storage, and release of water (Black, 1996; McDonnell et al., 2007). The runoff from these processes is governed by threshold mediated relationships across scales that result in storage—runoff hystereses (Spence, 2010). These threshold relationships between storage and runoff (*S*—*R*) are not uniform across a watershed, functioning as a series of discontinuous processes in soils and hillslopes that provide an integrated *S*—*R* relationship at the watershed scale (Spence, 2010). Kirchner (2009) described the *S*—*R* relationship to be non-linear and stated that watersheds typically function as dynamic systems governed by their unique climate and geology. These conceptual models of hydrologic behaviors help provide a process-based understanding of watersheds as dynamic environmental systems (Aspinall, 2010), and identify connections that advance hydrologic science and hydrologic prediction (Wagener et al., 2007).

At the local scale, *in situ* instrumentation can quantify the non-linear relationship between streamflow and water stored in a watershed as snow, soil moisture, groundwater and reservoirs (Appleby, 1970; Brutsaert, 2008; Kirchner, 2009; Sayama et al., 2011). These four primary storage components, along with climate, topography, and geology govern the fluxes of water through a catchment, and play an important role in the hysteretic nature of storage and runoff dynamics (McGlynn and McDonnell, 2003; McNamara et al., 2011). Knowledge of these processes is fundamental to developing an understanding of a watershed's hydrologic behavior. However, observations over larger regions can be technically challenging and costly, and *in situ* measurements from small basins do not necessarily represent the complexity inherent to watersheds at more broad scales (Spence, 2010). This scaling problem limits our capacity to understand and predict regional hydrologic processes, which is often the practical scale of watershed management (Blöschl, 2001; Western et al., 2002; Skøien et al., 2003; Peel and Blöschl, 2011; Thompson et al., 2011).

In the absence of broad-scale observations, past hydrological studies have typically relied on *in situ* measurements as a proxy for regional scale hydrological processes. For example, in higher latitude or mountainous regions measurements of snow water storage have provided a simple metric that has been used in water resource planning for decades (Cayan, 1996; United States Army Corps of Engineers, 2001), and are often correlated to streamflow gauged downstream (Dozier, 2011). While informative, this approach can often provide hydrological forecasts that are misleading, because point-based measurements do not fully represent the broad-scale variability of rugged mountain terrain (Dozier, 2011; Nolin, 2012; Webster et al., 2014; Ayala et al., 2014). Similarly, measurements of soil moisture in the upper 2000 mm of the soil rely on point-based data that are often distributed at the regional scale, but do not effectively represent the true variability of soil moisture found at the regional scale (Western et al., 2002; Brocca et al., 2010). A complete understanding of groundwater stores and fluxes (deeper than 2000 mm) at regional scales also remains elusive, despite its increasing importance in water resources management (Wagener et al., 2007; Gleeson et al., 2012; Famiglietti and Rodell, 2013; Barthel, 2014). In addition to contributing to runoff, groundwater serves as an important water resource for consumptive use (Gleeson et al., 2012).

While local-scale methods have been applied with moderate success in the past, current trends in climate and in consumptive water demand suggest that long-term changes in hydrological fluxes will have a major impact at the regional scale (Milly et al., 2008). As a result, the supply and demand of water is also expected to shift, especially at the regional scale (Wagener et al., 2010; Gleick, 2014a).

Hydrologic models can help address the questions of scale and bridge the gap between local scale observations and regional-scale processes by estimating the primary components of water storage (snow, soil moisture, reservoir, and groundwater) across a larger spatial grid. Regional-scale modeling approaches are integrated into water resource management operations for navigation, human consumptive use, irrigation, and hydropower (Payne et al., 2004; Rodell et al., 2004). Models can also be applied diagnostically to test scientific hypotheses and provide a better understanding of the physical processes that govern real world systems, such as the connections between snowmelt, streamflow, and groundwater (Beven, 2007, 2010; Moradkhani and Sorooshian, 2008; Kirchner, 2009; Clark et al., 2011; Capell et al., 2012). Despite their utility, developing and validating a model can be both time consuming and reliant on multiple data inputs, which even in the most well-instrumented basins provides sparse geographic coverage (Bales et al., 2006; Zang et al., 2012). The lack of an integrated measurement of water storage and streamflow has limited regional-scale hydrologic insights to model-based studies (Koster et al., 2010; Mahanama et al., 2011).

Since 2002, broad-scale measurements of changes in the amount of water stored across and through the earth have been available from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites (Tapley et al., 2004). GRACE measures monthly changes in the Earth's gravitational field that are proportional to regional changes in total water storage (Wahr et al., 2006). GRACE satellites provide a monthly record of terrestrial water storage anomalies (*TWSA*), which represent the changes in the vertical sum of water at the Earth's surface stored in snow, surface, soil and groundwater. Water losses to runoff and evapotranspiration are implicit in the GRACE storage signal, removing the added layer of complexity typically required to model the terrestrial water balance.

GRACE data, coupled with modeled and measured variations of water stored in snow, surface reservoirs and soils, have successfully been decomposed to quantify regional groundwater changes (Rodell et al., 2009; Famiglietti et al., 2011; Voss et al., 2013; Castle et al., 2014) and have contributed to improving water balance calculations (Zaitchik et al., 2008; Li et al., 2012). More recent efforts have quantified the relationship between regional water storage and specific streamflow events (Reager and Famiglietti, 2009; Reager et al., 2014). Riegger and Tourian (2014) coupled GRACE data using data-driven and model-based approaches to better understand the relationship between storage and runoff across climatic zones globally. Their study found that coupled liquid storage is linear to runoff, and that in climatic regions with snow and ice the relationship between storage and runoff is more hysteretic. These novel analyses, which are more diagnostic in nature, have provided new insights into regional watershed hydrology using GRACE measurements as a core data input. These studies have not explored how topography and geology can also help describe the *S*—*R* relationship of regional watersheds. Nor did these studies examine the ability of GRACE measurements to predict seasonal runoff.

In this paper, we use terrestrial water storage data from GRACE to better understand the hydrology of regional watersheds and the relationship between storage and runoff. The temporal relationships between coincident *TWSA* and discharge observations at three scales in the Columbia River Basin (CRB) of western North America are investigated using climate, topography, and geology as a framing principle to describe the shape of the storage-streamflow hysteresis. We associate regional and temporal differences in the hystereses with varying watershed dynamics. Finally, we compare the prognostic abilities of GRACE observations with individual modeled estimates of snow and soil moisture to predict seasonal streamflow at regional scales.

2. Study Area

Our study area is the Columbia River Basin (CRB; 41-53ºN and 110-122ºW; Fig. 1**)**. This basin has dry summers and wet winters. Up to 70% of annual precipitation falls between November and March, 50-60% of which occurs as snow (Serreze et al., 1999; Nolin et al., 2012). The spring months (April to June) are also wet, but warmer. Precipitation during the spring combines with snowmelt to swell rivers and potentially exacerbate flooding. Snowmelt also serves as a critical component of the hydrologic cycle, recharging aquifers and filling streams later in the year. These contributions bridge the temporal disconnect between wet winters and dry summers when demand is at its peak as farmers, fish, hydropower and municipal users vie for over-allocated water resources (United States Army Corps of Engineers, 2001; Oregon Water Supply and Conservation Initiative, 2008). However, concerns with winter surplus and summer scarcity are not uniform across the CRB, since climate and geology vary greatly. Two of the study watersheds, the Upper Columbia (155,000 km^2) and the Snake River basin (182,000 km^2), represent distinctly different climatic, topographic, and geologic provinces of the CRB (described and illustrated in Fig. 1). The Upper Columbia is wet and is characterized by steep topography of fractured rock and poor groundwater storage. In contrast, the arid Snake River basin is bowl-shaped with mountains on three sides. The interior of the Snake River basin is a broad plain with well-developed soils and expansive aquifer storage. The Columbia River at The Dalles (614,000 km^2) encompasses the Upper Columbia and the Snake River sub-basins, and its climate and geology are an integration of the two (Fig. 1). A distinct climatic feature of the Columbia River at The Dalles is the western slope of the Cascade Mountains, where over 3000 mm of mean annual precipitation at higher elevations sustains a considerable seasonal snowpack.

The scale of this study was constrained to watersheds larger than 150,000 km^2, the optimal minimum geographic limit of GRACE data (Yeh et al., 2006; Landerer and Swenson, 2012).

3. Methods and Data

We used 108 months of GRACE and streamflow data over nine water years (WY; Oct – Sep; 2004–2012). This data comprises positive, neutral, and negative phases of the El Niño-Southern Oscillation and negative and positive phases of the Pacific Decadal Oscillation (Feng et al., 2014; Iizumi et al., 2014). As a result, the data provides years of above- and below-average precipitation, snowpack, and streamflow for the region. The three watersheds were delineated upstream from United States Geological Survey (USGS) stream gages at 1° resolution, which is the resolution of GRACE data. In the CRB, these grid cells represent a dimension of approximately 80 km by 120 km. The Upper Columbia consists of the area upstream of the Columbia River at the International Boundary gage (USGS 12399500), just downstream of the confluence of the Columbia and Pend-Oreille Rivers. The Pend-Oreille is a major watershed in the upper portions of the CRB. The Snake River gage at Weiser (USGS 13269000) provides gauged streamflow data above Hell's Canyon Reservoir, the largest impoundment in the Snake River basin. The USGS gage at The Dalles (USGS 14105700) provides the most downstream streamflow data for the CRB. Monthly mean runoff (R; mm) was calculated for each of the three gages using the USGS streamflow data.

Measurements of *TWSA* were obtained from the GRACE RL-05 (Swenson and Wahr, 2006; Landerer and Swenson, 2012) data set from NASA's Tellus website (http://grace.jpl.nasa.gov). The errors present in the gridded GRACE data exist primarily as a result of truncation (i.e., a low number of harmonics) in the spherical harmonic solution, and smoothing and systematic noise removal (called “de-striping”) that is applied after GRACE level-2 processing to remove spatially correlated noise (called “stripes”) (Swenson and Wahr, 2006). This smoothing tends to smear adjacent signals together (within the radius of the filtering function), resulting in smaller signals being lost, and larger signals having a coarser footprint and a loss of spatial information.

To restore the GRACE signal lost during processing, the data were scaled using 1° Land-Grid Scale Factors produced by putting a 1° land surface model through identical processing (truncation and filtering) as the GRACE solutions, then measuring the decrease in the signal amplitude at each 1° grid. These procedures are described on the Tellus website and detailed in Landerer and Swenson (2012). Monthly 1° GRACE estimates of *TWSA*, and the associated 1° leakage and measurement errors, were spatially averaged over each of the three study watersheds following the procedures described in the Tellus website.

GRACE represents monthly storage anomalies relative to an arbitrary record-length mean value, analogous to the amount of water above or below the long-term mean storage of a bucket, and should balance with the equation:

$$\Delta Storage = TWSA = \Delta GW + \Delta SM + \Delta SWE + \Delta RES \qquad (1)$$

where all components are at monthly time steps; *GW* represents groundwater, *SM* represents soil moisture (from 0–2000 mm depth), *SWE* represents snow water equivalent (the equivalent depth of water held in snowpack), and *RES* represents reservoir storage. The Δ used here represents the anomaly from the study-period mean, rather than a monthly change. To isolate monthly groundwater storage anomalies ($\Delta GW = GWSA$) in the above equation, ΔSM, ΔSWE and ΔRES estimates were subtracted from the monthly *TWSA* data using methods described in Famiglietti et al. (2011). Similarly, the combined signal of water storage anomalies of subsurface moisture ($TWSA_{sub}$), *SM* and GW, was isolated by subtracting *SWE* and *RES* from *TWSA* values.

Monthly *SM* values over the study basins were obtained from the mean of the North American and Global Land Data Assimilation Systems (NLDAS at 1/8º resolution (Cosgrove et al., 2003) and GLDAS at 1/4º resolution (Rodell et al., 2004), respectively), and were spatially averaged over the three study watersheds. Monthly 1-km resolution *SWE* values were obtained from the mean of NLDAS and Snow Data Assimilation System (SNODAS; National Operational Hydrologic Remote Sensing Center, 2004) and were spatially averaged over the three watersheds. SNODAS data were used in place of the GLDAS data product, which considerably underestimated *SWE* in mountainous areas when compared to point-based measurements. Changes in monthly reservoir storage were calculated for the five largest reservoirs in the CRB (see Table A1). Other smaller reservoirs in the CRB were excluded when it was determined that fluctuations in their levels were below the detection limits of GRACE.

Like all measurements, estimates of *TWSA* from GRACE contain error. For all of the study basins, the range of error is well below the *TWSA* signal strength, approximately an order of magnitude below the annual amplitude (200 – 300 mm) of the *TWSA* signal in the CRB. The basin-averaged *TWSA* errors (time invariant) for the three study basins are 37 mm (Upper Columbia), 22 mm (Snake), and 25 mm (The Dalles).

The model data from LDAS and SNODAS simulations are driven by *in situ* measurements, and represents the best available data for broad scales. We address any structural error from an individual model by using an ensemble of outputs. Calculation of the error in individual terms followed standard methodologies (Famiglietti et al., 2011), where error in *SM* is the mean monthly standard deviation, and standard errors for *SWE* and *RES* are 15% of mean absolute changes. GWSA and $TWSA_{sub}$ anomaly errors are calculated as the sum of basin-averaged errors (added as variance) in the individual terms in each respective calculation (eq. 1), including the error in *TWSA* (Swenson et al., 2006). The basin-averaged error variance for *GWSA* (time invariant) in the three study basins are 45 mm (Upper Columbia), 26 mm (Snake), and 33 mm (The Dalles). For $TWSA_{sub}$ these values are 37 mm (Upper Columbia), 22 mm (Snake), and 25 mm (The Dalles). The individual error components (*SM*, *SWE*, *RES* respectively) for each basin are Upper Columbia (24 mm, 6 mm, 0.01 mm), Snake (14 mm, 3 mm, 0.01 mm), and The Dalles (21 mm, 4 mm, 0.01mm). Note that these error estimates are distributed across an entire regional watershed and do not represent the error at individual monitoring sites. A time series of these values and basin-averaged errors is provided in Fig. 2.

Based on an approach similar to Reager et al. (2014) and Riegger and Tourian (2014), we plotted the temporal relationship between *TWSA* and *R* to examine hysteresis relationships in all three of the study watersheds for each individual water year and for the monthly mean across all water years. Expanding from the integrated terrestrial component of water storage, we also plotted the relationships of $TWSA_{sub}$ and *GWSA* with *R*. We examined the branches of these hysteresis plots to better understand how the size, shape, and direction of the hystereses varied across years in each of the three regional watersheds.

In order to verify groundwater hysteresis, we compared the GRACE-derived *GWSA* to groundwater depths from well measurements at 33 sites throughout the study region (Fig. 1 and Table A2). These data were normalized by their standard deviation, and the mean of the 33 wells was calculated. The standard deviation of the GRACE-derived *GWSA* for The Dalles was normalized to provide a direct comparison of *GWSA* and *in situ* measurements.

We further hypothesized that because peak *SWE* accumulation occurs between February and April, that *TWSA* for these months could be used to predict *R* for an individual month and the cumulative seasonal runoff (R_{season}) that occurs after peak *SWE* accumulation. To test this prognostic hypothesis we used a two-parameter power function (The MathWorks, 2013):

$$R_{predicted} = a(TWSA_{month})^b + c \quad (2)$$

where $R_{predicted}$ is runoff for the predicted time interval; $TWSA_{\mathrm{month}}$ represents terrestrial water storage for an individual month, and a, b, and c are fittted parameters from the power function.

We tested this relationship for *TWSA* in February, March and April to predict R_{season} (April – September) and for the individual months of July (R_{July}), August (R_{Aug}), and September (R_{Sep}); these represent the lower-flow months when demand is near its peak. Additionally, we tested and compared the modeled-values of *SWE* from NLDAS and SNODAS and *SM* from NLDAS and GLDAS, and the model-derived values of $TWSA_{\mathrm{sub}}$ to predict R_{season} and for the individual months using the same power-function analysis.

Because our data set was constrained to nine water years, we used a double-pass approach to fit and test the empirical relationship between *S*—*R*. This approach allowed us double our data inputs for calculating standard hydrologic evaluation metrics such as Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE) and Coefficient of Determination (R^2); (Legates and McCabe, 1999). The nine years were divided into two sets (Set 1, even years 2004-2012; Set 2, odd years 2005-2011). The first pass calculated the power function of *S*—*R* to Set 1, and the parameters were then tested against Set 2. The roles of the datasets were then reversed, and the empirical model results of each pass were compiled into one data set and tested against measured values to calculate RMSE, NSE, and R^2. In order to maximize the limited data inputs, once we tested the two independent sets for model performance, we combined the data sets for a single power function curve. The observed data were tested against the simulated data from the complete, but limited data record. The final model curve was fit to these data.

4. Results

4.1. Storage-runoff hysteresis

The filling and emptying of the study basins at the regional-scale over the course of an individual WY results in a hysteretic relationship between storage and runoff (Fig. 3a). The hysteresis loops begin at the onset of the wet season in October, with *TWSA* increasing (Figs. 3a, 4a-c) as precipitation is stored as snow and soil moisture. An increase in storage that is not offset by an increase in discharge indicates a predominance of snow inputs and the freezing of soil water. The lower branch of the hysteresis plot (storage increase unmatched by runoff) can be used to estimate cumulative snow water equivalent and soil moisture in the basin. This is the water that later contributes to streamflow and groundwater recharge in the spring.

The hysteresis shifts direction from Feb-Apr (inflection 1, Fig. 3a) when saturated soils and snowmelt cause *R* to rapidly increase. Each hysteresis loop contains a vertical branch of the curve during which storage is relatively constant, but streamflow increases rapidly. This also represents the groundwater recharge branch of the loop. As snow melts and the ground thaws, runoff is generated, recharge into soils occurs, and basins tend to be at peak storage during this branch. Storage losses and additional precipitation inputs during this period are re-organized internally. A second shift (inflection 2, Fig. 3a) occurs from Apr-June when peak *TWSA* begins to decrease, representing spring snowmelt and a switch from precipitation that falls primarily as snow to rain; these combine to contribute to peak *R*.

Once peak *R* values are reached, the loop shifts direction a third time (inflection 3, Fig. 3a), receding on both axes as contributions from snowmelt diminish while presumably groundwater sustains streams and provides a source for irrigated agriculture. During this period, the relationship between *TWSA* and discharge is more linear, corresponding to baseflow-driven runoff processes in which each monthly change in storage causes a proportional change in the generation of streamflow.

The hysteresis plots of *TWSA*—*R* for an individual water year demonstrate that the timing and quantity of precipitation governs the size of a hysteresis loop for an individual WY (Figs. 3a, 4a-c, 5). For instance wet years (e.g., 2008) have bigger loops, while dry years (e.g., 2005) are more compressed along both axes. However, the general shape of the loops is distinct for each basin. Plotting multiple WYs provides a family of curves for each basin that helps describe how climate, topography, and geology govern the timing and magnitude of the relationship between *TWSA* and *R* (Figs. 3a, 5).

4.2. Subsurface water ($TWSA_{sub}$) – runoff hysteresis

The $TWSA_{sub}$ hysteresis curve contracts horizontally when the snow signal is removed from *TWSA* values for both the Upper Columbia and The Dalles (Figs. 3b, 4d-f), which collapses the loops and takes a form similar to a plot-scale hysteresis of soil. Peak $TWSA_{sub}$ occurs in June, which corresponds to the spring melt of mountain snowpack and the end of the wet season (Figs. 4d-f). However in the Snake River, the hysteresis curve still retains a loop, but the timing of maximum $TWSA_{sub}$ is also earlier, reaching its peak during March and April (Fig. 4e). It is noteworthy that in the Snake River the $TWSA_{sub}$—R hysteresis loop temporally progresses in the opposite direction, but stays in phase with precipitation inputs.

4.3. Groundwater-runoff hysteresis

The hysteresis loops describing the temporal relationship between *GWSA* and *R* are equally informative, with one distinct difference—they temporally progress in opposite directions of the hysteresis loops of *TWSA* and *R* (Fig. 3). For all three watersheds, *GWSA* decreases from Oct–Feb/Mar (Fig. 4h-j), which is out of phase with the onset of the wet season. *GWSA* does not shift towards positive gains until early spring and the initial stages of melt before reaching its maximum in June.

The 33 point-specific well data located across the CRB show considerable individual variability throughout a water year, and the mean of the normalized standard deviations of well levels was close to zero for all months. The temporal variability for the well data provides no discernable correlation with the derived *GWSA* signal (Fig. A1).

4.4. Individual basin hysteresis plots of *TWSA*, $TWSA_{sub}$, *GWSA* and *R*

Of the three study basins, the Upper Columbia is the most hydrologically active, showing the largest annual range for *TWSA*, $TWSA_{sub}$, *GWSA*, and *R* (Fig. 6). The *TWSA—R* hysteresis loops are more open (Fig. 4), corresponding to the fluxes of water moving through watershed. When *SWE* is removed and subsurface water is highlighted, the $TWSA_{sub}$—R hysteresis loops collapse horizontally and more closely resemble the hystereses associated with soil (Figs. 4d). However the inter-annual range (WY_{max} – WY_{min}) for $TWSA_{sub}$ in the Upper Columbia is considerably greater than the other two basins (median range = 234 mm; Fig. 6). As the hysteresis reverses directions for *GWSA*—*R*, the loops shift to a more open shape (Figs. 4d), but the inter-annual range remains similar.

In contrast to the rapid response of the Upper Columbia, the Snake River receives ~60% less annual precipitation, but has an annual *TWSA* range that is only 22% less (median annual range = 192 mm; *R*=7 mm; Figs. 4, 5, and 6). However, the *TWSA* hysteresis loops for the Snake River are collapsed vertically (Fig. 4b). In the more arid Snake River, removing the snow signal does not collapse the $TWSA_{sub}$—*R* hysteresis loops ($TWSA_{sub}$ = 89 mm). Similarly, the *GWSA* loops suggest that subsurface moisture plays a more prominent role in the Snake River.

The climate, topography, and geology of the Columbia River at The Dalles are an integration of the Upper Columbia and Snake River, seen in the shape of the hysteresis loops (Figs. 4, 5, 6; median annual range *TWSA*=195 mm; *R*=27 mm). The period from Feb–June more closely resembles the Snake River basin, with gradual increases in *TWSA* and sharp increases in *R*. The slope of the recession from June-Sept has the same general shape for The Dalles as the Upper Columbia (Figs. 4a, 4c), presumably from snowmelt-generated runoff.

4.5. Streamflow forecasting

We next present how *TWSA* was applied prognostically to predict streamflow. Using the double-pass calibration and validation approach, $TWSA_{Mar}$ provided the best overall predictive capabilities for R_{season} with a mean NSE ($\overline{\mathrm{NSE}}$) and mean R^2 ($\overline{\mathrm{R}}^2$) of 0.75 and 0.91, respectfully (Fig. 7a, Table 1), for all three basins. The Dalles had the highest NSE and R^2, and lowest RMSE values (0.98, 0.98, 6 mm; Table 1). The results in the Upper Columbia were also robust (0.82, 0.86, 33 mm; Table 1), while the Snake River performed with less skill (0.46, 0.59, and 14 mm, Table 1). Applying $TWSA_{April}$ also provided similar results, but with a lower degree of skill in predicting *R* ($\overline{\mathrm{NSE}}$ = 0.57, $\overline{\mathrm{R}}^2$= 0.69). $TWSA_{\mathrm{Apr}}$ provided improved predicted capabilities in the Upper Columbia (0.87, 0.88, and 28 mm, Table 1), but inferior results in the other two watersheds. $TWSA_{\mathrm{Feb}}$ had a low degree of skill in predicting *R* in all three watersheds (Table A3).

$TWSA_{\mathrm{Mar}}$ and $TWSA_{\mathrm{April}}$ also served as a good predictor of monthly runoff in July and August for the Upper Columbia and to a lesser degree in The Dalles (Tables 1 and A3). In the Snake River, *TWSA* did not serve as a good predictor for *R* in an individual month.

Snowpack and soil moisture play a considerable role in the hydrology of the CRB and are commonly used to help predict water demand and availability later in the year (Koster et al., 2010). We compared the capabilities of the modeled snow (*SWE*) and soil moisture (*SM*) products to predict *R* to the skill of measured GRACE *TWSA* data (Table 1). In the Upper Columbia and The Dalles, $TWSA_{\mathrm{Mar}}$ predicts seasonal and monthly runoff (July and August) with considerably more skill than *SWE* or *SM* (Figure 7, Table 1). In the Snake River, SM_{Mar} has a higher degree of skill than $TWSA_{\mathrm{Mar}}$ in predicting R_{season} and R_{Aug}. SWE_{Mar} provided inferior results in all three watersheds, but with some predictive skill in the Upper Columbia and The Dalles (NSE of 0.24 and 0.46 respectively, Table 1). In all three watersheds, $TWSA_{\mathrm{sub}}$ provided extremely poor predictions (Tables 1 and A3).

When the results of the empirical model using two independent sets of data proved robust for some of the storage metrics, the observed data were tested against the simulated data from the complete, but limited data record. The performance of the empirical model improved using the complete data set (Tables 2 and A4), with the same general results. $TWSA_{\mathrm{Mar}}$ provided the best model fit for seasonal runoff in the Upper Columbia (NSE = 0.93, RMSE = 19.8 mm) and The Dalles (NSE = 0.98, RMSE = 5.7 mm). In the Snake River, predictive capabilities improved more dramatically (NSE = 0.83, RMSE = 7.4 mm), but soil moisture still served as a better predictor of seasonal streamflow (NSE = 0.93, RMSE = 5.2 mm). Similarly, $TWSA_{Mar}$ provided the best model fit for runoff in August, one of the drier months when demand is at its peak (Tables 2 and A4).

5. Discussion

5.1. Storage-runoff hysteresis

Decades of data collection and monitoring at individual gage sites indicate that watersheds collect, store and release water. Using one integrated measurement from the GRACE satellites, our results show these same process at the regional scale in the hysteresis loops of storage (*TWSA*) and runoff (*R*). While hystereic processes have previously been identified in local-scale measurements (McDonnell, 2003; McGlynn and McDonnell, 2003), only recently has streamflow-storage hysteresis been identified at the regional scale (Riegger and Tourian, 2014).

Our work builds on Riegger and Tourian's (2014) results, and employs GRACE data to describe how regional watersheds function as integrated, non-linear systems governed by climate, topography, and geology. Climate controls the size of the hysteresis loops by providing a first-order control on hydrologic inputs and the storage of solid water, which in turn governs the ranges of *TWSA* and *R*. However, runoff response to precipitation and snowmelt does not act independently from topography and geology (Jefferson et al., 2008; Tague et al., 2008), which controls how liquid water is stored and routed through a watershed, even at the regional scale. The climatic, topographic, and geological characteristics of each watershed provide an explanation of the *S*—*R* relationship that helps govern the shape and size of its respective hysteresis curve. GRACE offers a single, integrated measurement of changes in water storage through and across a watershed that can be applied to predict regional streamflow using an empirical model. Where these predictive capabilities succeed and fail help better describe the climatic, topographic, and geological characteristics in each watershed.

For example, in the Upper Columbia, steep topography and wet climate fills subsurface storage quickly before reaching a threshold in April or May. After this watershed-scale threshold is reached, the steep topography moves snowmelt and rain quickly through the terrestrial system and into the river channel until cresting in June (Figs. 4, 5, and 6), followed by declines in *TWSA* and *R* from June-September. These large fluxes of water create a more open hysteresis loop, expanding non-linearly on both the horizontal and vertical axes.

The Upper Columbia also has the broadest range of annual $TWSA_{\text{sub}}$ and *GWSA* during the study period (Figs. 5 and 6), despite having limited aquifer capacity. Conceptually, this demonstrates that the upper limit of storage is greater than in the Snake River or The Dalles, but that it also loses the most water. Its minimums at the end of the WY are also the lowest (median $TWSA_{\text{Sep}}$ = -98mm; Figs. 5 and 6). This range across *TWSA*, $TWSA_{\text{sub}}$, and *GWSA* supports the conceptual model that the watershed fills during the wet season, and is then drained more quickly due to steep topography and limited water storage. The predictive capability of *TWSA* also strongly suggests that the components and temporal relationships of storage across this watershed are interconnected, and that incorporating April snowpack improves the model results.

In contrast, the arid Snake River basin provides a very different family of hysteresis curves (Figs. 4, 5) that identify groundwater and soil moisture as primary components of watershed function. The curves are compressed vertically (*R*) as compared to the Upper Columbia, and are more constrained horizontally (Fig. 6). The onset of spring melt runoff in

February does not deplete *TWSA* for the Snake River. Instead, *TWSA* continues to increase until May, when peak runoff occurs. As *TWSA* decreases to the end of the water year in September, the median $TWSA_{\mathrm{Sep}}$ measurement (-78 mm) is 20 mm greater than in the Upper Columbia. This indicates that the lower drainage threshold of the Snake River watershed is relatively greater than the Upper Columbia, potentially explained by a less severe topography and higher aquifer capacity.

The $TWSA_{\mathrm{sub}}$ hysteresis curves in the Snake River retain a similar shape to the *TWSA* signal. While they reverse direction they do stay temporally connected to the onset of the wet season in October, indicating that subsurface moisture is a central control on the filling of the watershed through May. The capabilities of *SM* to empirically predict *R* better than *TWSA* further highlight the importance of subsurface water in this watershed. The intra-annual range of *GWSA* in the Snake River is also more limited than in the more hydrologically responsive Upper Columbia. This more limited range of data supports the conceptual model of a watershed that retains comparatively more winter precipitation in soils and aquifers throughout the spring season, and that sustains flow later in the year and until the onset of melt the following winter.

The greater Columbia River Basin upstream from The Dalles integrates the climatic, topographic, and geologic characteristics of the Snake River and Upper Columbia as well as other areas within the CRB. The western slope of the Cascades (Fig. 1), which is outside of the Upper Columbia, accumulates up to several meters of *SWE* each winter. Due east of the Cascades, an expansive basalt plain that provides aquifer storage helps dampen the snowmelt pulse in the spring. The hysteresis loops for The Dalles reflect these combined characteristics.

Storage at The Dalles increases along the horizontal axis (*TWSA*) until peak storage is reached in March or April (Figs 3, 4, and 5). This *TWSA* threshold responds with an increase in *R* that continues through June. In July, the hysteresis begins to recede along both axes closing out the loop. The *GWSA* has the most limited range, potentially explained by the extensive basalt aquifer moderating the relationship between storage and runoff. In The Dalles, $TWSA_{\mathrm{Sep}}$ has a median value of -88mm (Fig. 6), between the lower drainage thresholds of the Upper Columbia and Snake River watersheds; indicating an integration of the contributing climate, topography, and geology.

5.2. Distinguishing the difference between $TWSA_{\mathrm{sub}}$ and *GWSA*

Conceptually $TWSA_{\mathrm{sub}}$ represents changes in the amount of water stored as soil moisture and groundwater, where as *GWSA* represents water changes greater than 2000mm below the soil surface. The goals of evaluating these metrics were to see if monthly changes in soil moisture were linked to changes in groundwater storage, and the role of snowpack in the *S*—*R* relationship.

The $TWSA_{\mathrm{sub}}$ hysteresis curves in the Upper Columbia and The Dalles collapse into a more linear relationship that is more commonly associated with the *S*—*R* relationship of a soil matrix (Fig. 3 and 4). This is in contrast to the *GWSA* hystereses that are represented by loops that show an out-of-phase relationship between precipitation and groundwater recharge from the start of the wet season in October until February or March. The $TWSA_{\mathrm{sub}}$ and *GWSA* hysteresis plots demonstrate that in these two basins changes in monthly soil moisture are not always temporally aligned with *GWSA*. This can be explained by the physical reality that soil moisture and groundwater are not always interconnected, and that there is not a fixed depth (i.e., 2000 mm) that separates the two components of water storage.

GRACE-derived calculations of *GWSA* also provide insights into the hydrological processes governing groundwater recharge and depletion, as evidenced in the *GWSA* hysteresis loops. The *GWSA—R* curves show an out-of-phase relationship between precipitation and groundwater recharge from the start of the wet season in October until February or March. This indicates that groundwater helps sustain stream flow during the wet fall and winter and that pore space in soils and geologic materials must fill to a certain threshold before groundwater begins to recharge and runoff is generated. The relationship between the *TWSA* and *GWSA* curves from Oct-Mar identifies how the onset of snowmelt also marks the beginning of groundwater recharge, and suggests that snowmelt inputs to groundwater are considerable. In the CRB this is critical as current climate trends are projected to reduce snowpack accumulation and exacerbate melt in the region (Wu et al., 2012; Rupp et al., 2013; Sproles et al., 2013).

Additionally, our analysis identifies summer as the time of peak groundwater storage in all three regional watersheds. This finding is of value for groundwater management and policy decisions, as peak groundwater levels in June correspond to the timing of groundwater pump tests that are used to develop groundwater withdrawal regulations (Jarvis, 2011, 2014). Our data suggest that groundwater pump tests should not be limited to an individual month, and should also include periods of reduced storage particularly during the winter months. The inclusion of multiple pump tests throughout the year could be particularly relevant as the population and water demand is projected to increase in the region.

The point-specific well data are not conclusive and show considerable variability with no consistent pattern regarding the timing of recharge and peak groundwater levels. This is presumably a function of how site characteristics (i.e., usage, depth, location, elevation) are extremely variable across a region. Rather than excluding these results or selecting individual wells that match GRACE data, we discuss the results from all 33 wells to help demonstrate the high variability that exists from well to well, and that measurements of groundwater changes at a fixed location does not represent watershed-scale characteristics (Jarvis, 2011, 2014). The disconnect between sites also highlights the concept brought forward by Spence (2010), that storage is not uniform across a watershed, and functions as a series of discontinuous processes at the watershed scale.

5.3. Applying the *S*—*R* relationship as a predictive tool

We applied these climatic, topographic, and geologic insights to develop and test the hypothesis that spring *TWSA* could predict *R* later in the year, based on two observations: First, the shapes of the hysteresis curves for each basin are similar (Figs. 4a-c, 5), but vary by magnitude of annual *TWSA*. Second, peak *TWSA* occurs before the peak runoff. We show that the integrated GRACE signal is a good baseline measurement to empirically predict seasonal streamflow across a range of water years with regards to precipitation and streamflow. In essence, our data suggest that the water stored across and through the Columbia River Basin in March describes the water available for the remainder of the water year.

In the CRB and in the northwestern United States, peak snowpack occurs in March or April, and is commonly used as a metric for predicting spring runoff. Despite the importance of snowpack to the hydrologic cycle of the region, measurements of $TWSA_{\mathrm{Mar}}$ from GRACE provide a better prediction of R_{season}, R_{July}, and R_{Aug} than model-derived estimates of snowpack. GRACE $TWSA_{Mar}$ also provided a better prediction for runoff than soil moisture, except for the Snake River watershed. There March soil moisture provided a better indicator of runoff for the rest of the year. $TWSA_{Feb}$ provided inferior predictive capacity, as the annual maximum *TWSA* values have not been reached.

These results are promising with regards to using GRACE as a predictive tool for water resources in both wet and dry years. Our limited data record represents a wide-range of conditions with regards to climate and streamflow, which is captured in our empirical models and is shown in the box plots to the right of Figs. 7a - b. These same results also indicate that *R* is insensitive to $TWSA_{Mar}$ values below 100 mm. This lower threshold describes with some certainty the amount of runoff that will be available for operations for the remainder of the year.

We recognize that all three of these regional watersheds are managed through a series of dams and reservoirs that create an altered runoff signal. Water resources managers use point-specific and model-based estimates of water storage in the region to optimize their operations for the water year. Additionally, in the fertile plains of the Snake River and lower CRB, broad-scale agriculture relies on both ground- and surface water for irrigation. Water withdrawals would be implicit in the *TWSA* signal and reduce *R*. However, a more detailed analysis of withdrawals lies outside the scope of this study.

Regardless of the length of record or anthropogenic influence, climate, topography, and geology still provide the first-order controls on water storage that are found in the hysteresis loops. GRACE encapsulates these hydrologic processes through measurements of *TWSA*. The hysteresis loops expand and contract accordingly during wet and dry years, as the intra-annual relationship between *TWSA* and Q represents the fluxes of water into and out of the watershed. Despite intra-annual differences, a family of hysteresis curves can describe each of the sub-regional watersheds. The predicative capability using *TWSA*, the vertical sum of water, as compared to snowpack and soil moisture further highlights the integrated nature of water storage in regional hydrology. These predictive capabilities highlights the potential of GRACE to improve upon seasonal forecast predictions and regional hydrological models.

5.4. GRACE as an analysis tool for regional watersheds

Where previous approaches to modeling watershed behavior have focused on separate storage compartments, new approaches should include the magnitude and direction of hysteresis (Spence, 2010). This integrated approach would provide new ways forward to classify watersheds not only by runoff, but also on the first-order controls that govern the non-linear hydrological processes.

Even though GRACE is somewhat of a blunt instrument with regards to temporal (monthly) and spatial (1°) resolution, this emerging technology provides a new dimension to regional watershed analysis by providing an integrated measurement of water stored across and through the Earth. These measurements continue to prove their value in retrospective analysis of regional hydrology (Rodell et al., 2009; Castle et al., 2014). However, the hysteresis loops presented by Riegger and Tourian (2014) and further developed in this paper demonstrate the ability of GRACE data to help develop a process-based understanding of how regional watersheds function as simple, dynamic systems. As the temporal record of GRACE continues to increase, its value as both a diagnostic and predictive tool will continue to grow. In the mean time, these data have value in augmenting existing management strategies.

Perhaps one of the most important facets of GRACE data is that it does not distinguish political boundaries. It is not linked to a specific *in situ* monitoring agency with limited data access and has the capacity to bridge sparse and inconsistent on-the-ground hydrologic monitoring networks that exist in many regions of the world. Previous GRACE-based analysis has shown its value in highlighting negative trends in terrestrial water storage in trans-boundary watersheds (Voss et al., 2013; Castle et al., 2014), and resulting regional conflict exacerbated by water shortages (Gleick, 2014b). GRACE provides an objective measurement of a region's water resources that can provide valuable insights into potential shortages or surpluses of water resources, and simple empirical predictions of seasonal and monthly runoff that are easily deployable in places with limited data.

6. Conclusions

We have shown how GRACE-based measurements of *TWSA* distill the complexity of regional hydrology into a simple, dynamic system. *TWSA* and derived estimates of *GWSA* reveal hysteretic behavior for regional watersheds, which is more commonly associated with hydrologic measurements at local scales. While the magnitude of the hysteresis curves vary across years, they retain the same general shape that is unique to each watershed. We demonstrated the utility of these hysteresis curves by showing how the complete *TWSA* record during March and April can be used to empirically predict *R* for the remainder for the water year ($TWSA_{\mathrm{Mar}}$, mean NSE = 0.91) and during the drier summer months ($TWSA_{\mathrm{Mar}}$, mean NSE for July = 0.76, August = 0.72; Tables 1 and 2).

Because GRACE *TWSA* can augment prediction, managers could start to interpret each year's hysteresis curve for the upcoming spring and summer, providing greater clarity and validation for model-based forecasts presently used by water resource managers. Our results demonstrate a way forward, expanding GRACE from a diagnostic tool, into a conceptual model and predictive resource.

Although this study focused on the CRB, which has a rich data record, GRACE data are available at a global scale and could be readily applied in areas with a paucity of data to understand how watersheds function and to improve streamflow forecasting capabilities. GRACE does not discern political boundaries and provides an integrated approach to understanding international watersheds (Voss et al., 2013). This resource could serve as a valuable tool for managers in forecasting surplus and scarcity, and in developing strategies that include changes in supply and demand due to human consumptive needs and current climate trends (Wagener et al., 2010; Gleick, 2014a).

Author Contributions

E.A.S., S.G.L., and P.J.W. developed the hysteresis concept based upon background research by J.R. and J.S.F. The data analysis was led by E.A.S., but represents a combined effort from all of the authors. J.R. provided expertise in the GRACE data product, groundwater, and error analysis. E.A.S. prepared the manuscript with contributions from all co-authors.

Acknowledgments

The authors would like to thank Matthew Rodell and Felix Landerer for their expertise in understanding GRACE data during the initial stages of the research. GRACE terrestrial data were processed by Sean Swenson, supported by the NASA MEaSUREs Program, and are available at http://grace.jpl.nasa.gov. The GLDAS and NLDAS data used in this study were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences Data and Information Services Center. We would like to thank the reviewers of the manuscript that have helped improve its overall quality. Additionally, Tim Kerr provided objective feedback and comments on the research findings. The information in this document has been funded entirely by the US Environmental Protection Agency, in part by an appointment to the Internship/Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and EPA. This manuscript has been subjected to Agency review and has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Appleby, F. V.: Recession and the baseflow problem, Water Resour. Res., 6(5), 1398–1403, doi:10.1029/WR006i005p01398, 1970.

Aspinall, R.: A Century of physical geography research in the Annals, Ann. Assoc. Am. Geogr., 100(5), 1049–1059, 2010.

Ayala, A., McPhee, J. and Vargas, X.: Altitudinal gradients, midwinter melt, and wind effects on snow accumulation in semiarid midlatitude Andes under La Niña conditions, Water Resour. Res., 2014.

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R. and Dozier, J.: Mountain hydrology of the western United States, Water Resour. Res., 42(8), W08432, doi:10.1029/2005wr004387, 2006.

Barthel, R.: A call for more fundamental science in regional hydrogeology, Hydrogeol. J., 22(3), 507–510, 2014.

Beven, K.: Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, Hydrol. Earth Syst. Sci., 11(1), 460–467, 2007.

Beven, K. J.: Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, Hydrol. Process., 24(12), 1537–1547, 2010.

Black, P. E.: Watershed hydrology, CRC Press, Boca Raton, FL., 1996.

Blöschl, G.: Scaling in hydrology, Hydrol. Process., 15(4), 709–711, 2001.

Brocca, L., Melone, F., Moramarco, T. and Morbidelli, R.: Spatial‐temporal variability of soil moisture and its estimation across scales, Water Resour. Res., 46(2), 2010.

Brutsaert, W.: Long-term groundwater storage trends estimated from streamflow records: Climatic perspective, Water Resour. Res., 44(2), W02409, doi:10.1029/2007WR006518, 2008.

Capell, R., Tetzlaff, D., Hartley, A. J. and Soulsby, C.: Linking metrics of hydrological function and transit times to landscape controls in a heterogeneous mesoscale catchment, Hydrol. Process., 26(3), 405–420, 2012.

Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C. and Famiglietti, J. S.: Groundwater Depletion During Drought Threatens Future Water Security of the Colorado River Basin, Geophys. Res. Lett., doi:10.1002/2014GL061055, 2014.

Cayan, D. R.: Interannual Climate Variability and Snowpack in the Western United States, J. Clim., 9(5), 928–948, doi:10.1175/1520-0442(1996)009<0928:ICVASI>2.0.CO;2, 1996.

Clark, M. P., Kavetski, D. and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47(9), 2011.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Marshall, C., Sheffield, J. and Duan, Q.: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, J. Geophys. Res., 108(D22), 8842, 2003.

Dozier, J.: Mountain hydrology, snow color, and the fourth paradigm, Eos, Trans. Am. Geophys. Union, 92(43), 373, doi:10.1029/2011EO430001, 2011.

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., de Linage, C. R. and Rodell, M.: Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophys. Res. Lett., 38(3), L03403, doi:10.1029/2010gl046442, 2011.

Famiglietti, J. S. and Rodell, M.: Water in the Balance, Science (80-.)., 340(6138), 1300–1301, 2013.

Feng, J., Wang, L. and Chen, W.: How does the East Asian summer monsoon behave in the decaying phase of El Niño during different PDO phases?, J. Clim., 27(7), 2682–2698, 2014.

Gleeson, T., Wada, Y., Bierkens, M. F. P. and van Beek, L. P. H.: Water balance of global aquifers revealed by groundwater footprint, Nature, 488(7410), 197–200, 2012.

Gleick, P.: Water, Drought, Climate Change, and Conflict in Syria, Weather. Clim. Soc., 6, 331–340, 2014a.

Gleick, P. H.: The World's Water Volume 8: The Biennial Report on Freshwater Resources, Island Press., 2014b.

Iizumi, T., Luo, J.-J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M. E. and Yamagata, T.: Impacts of El Niño Southern Oscillation on the global yields of major crops, Nat. Commun., 5, 2014.

Jarvis, W. T.: Unitization: a lesson in collective action from the oil industry for aquifer governance, Water Int., 36(5), 619–630, 2011.

Jarvis, W. T.: Contesting Hidden Waters: Conflict Resolution for Groundwater and Aquifers, Routledge., 2014.

Jefferson, A., Nolin, A., Lewis, S. and Tague, C.: Hydrogeologic controls on streamflow sensitivity to climate variation, Hydrol. Process., 22(22), 4371–4385, doi:10.1002/hyp.7041, 2008.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45(2), W02429, doi:10.1029/2008WR006912, 2009.

Koster, R. D., Mahanama, S. P. P., Livneh, B., Lettenmaier, D. P. and Reichle, R. H.: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, Nat. Geosci, 3(9), 613–616, doi:http://www.nature.com/ngeo/journal/v3/n9/suppinfo/ngeo944_S1.html, 2010.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48(4), W04531, doi:10.1029/2011wr011453, 2012.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, 1999.

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D. and van Dam, T. M.: Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe, J. Hydrol., 446, 103–115, 2012.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D. and Reichle, R.: Soil Moisture, Snow, and Seasonal Streamflow Forecasts in the United States, J. Hydrometeorol., 13(1), 189–203, doi:10.1175/JHM-D-11-046.1, 2011.

McDonnell, J. J.: Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response, Hydrol. Process., 17(9), 1869–1875, doi:10.1002/hyp.5132, 2003.

McDonnell, J. J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J. and Weiler, M.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43(7), doi:10.1029/2006wr005467, 2007.

McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and hillslope zones to catchment runoff, Water Resour. Res., 39(11), doi:10.1029/2003WR002091, 2003.

McNamara, J. P., Tetzlaff, D., Bishop, K., Soulsby, C., Seyfried, M., Peters, N. E., Aulenbach, B. T. and Hooper, R.: Storage as a Metric of Catchment Comparison, Hydrol. Process., 25(21), 3364–3371, doi:10.1002/hyp.8113, 2011.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science (80-.)., 319(5863), 573–574, doi:10.1126/science.1151915, 2008.

Moradkhani, H. and Sorooshian, S.: General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis, in Hydrological Modelling and the Water Cycle, pp. 1–24., 2008.

National Operational Hydrologic Remote Sensing Center: Snow Data Assimilation System (SNODAS) Data Products at NSIDC, [2003-2012], National Snow and Ice Data Center, Boulder, Colorado USA., 2004.

Nolin, A. W.: Perspectives on Climate Change, Mountain Hydrology, and Water Resources in the Oregon Cascades, USA, Mt. Res. Dev., 32(S1), S35–S46, doi:10.1659/mrd-journal-d-11-00038.s1, 2012.

Nolin, A. W., Sproles, E. A. and Brown, A.: Climate change impacts on snow and water resources in the Columbia, Willamette, and McKenzie River Basins, USA: A nested watershed study. , in Transboundary River Governance in the Face of Uncertainty: The Columbia River Treaty, edited by B. Cosens, Oregon State University Press, Corvallis, OR., 2012.

Oregon Water Supply and Conservation Initiative: Southern Willamette valley municipal water providers final report., 2008.

Payne, J. T., Wood, A. W., Hamlet, A. F., Palmer, R. N. and Lettenmaier, D. P.: Mitigating the Effects of Climate Change on the Water Resources of the Columbia River Basin, Clim. Change, 62(1), 233–256, 2004.

Peel, M. C. and Blöschl, G.: Hydrological modelling in a changing world, Prog. Phys. Geogr., 35(2), 249–261, 2011.

Reager, J. T. and Famiglietti, J. S.: Global terrestrial water storage capacity and flood potential using GRACE, Geophys. Res. Lett., 36(23), L23402, doi:10.1029/2009gl040826, 2009.

Reager, J. T., Thomas, B. F. and Famiglietti, J. S.: River basin flood potential inferred using GRACE gravity observations at several months lead time, Nat. Geosci., 7, 588–592, doi:10.1038/ngeo2203, 2014.

Riegger, J. and Tourian, M. J.: Characterization of runoff-storage relationships by satellite gravimetry and remote sensing, Water Resour. Res., 50(4), 3444–3466, 2014.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D. and Toll, D.: The Global Land Data Assimilation System, Bull. Am. Meteorol. Soc., 85(3), 381–394, doi:10.1175/bams-85-3-381, 2004.

Rodell, M., Velicogna, I. and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460(7258), 999–1002, 2009.

Rupp, D. E., Abatzoglou, J. T., Hegewisch, K. C. and Mote, P. W.: Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA, J. Geophys. Res. Atmos., 2013.

Sayama, T., McDonnell, J. J., Dhakal, A. and Sullivan, K.: How much water can a watershed store?, Hydrol. Process., 25(25), 3899–3908, doi:10.1002/hyp.8288, 2011.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A. and Pulwarty, R. S.: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, Water Resour. Res., 35(7), 2145–2160, doi:10.1029/1999wr900090, 1999.

Skøien, J. O., Blöschl, G. and Western, A. W.: Characteristic space scales and timescales in hydrology, Water Resour. Res., 39(10), 2003.

Spence, C.: A paradigm shift in hydrology: Storage thresholds across scales influence catchment runoff generation, Geogr. Compass, 4(7), 819–833, 2010.

Sproles, E. A., Nolin, A. W., Rittger, K. and Painter, T. H.: Climate change impacts on maritime mountain snowpack in the Oregon Cascades, Hydrol. Earth Syst. Sci., 17(7), 2581–2597, doi:10.5194/hess-17-2581-2013, 2013.

Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33(8), L08402, doi:10.1029/2005gl025285, 2006.

Swenson, S., Yeh, P. J. F., Wahr, J. and Famiglietti, J.: A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, Geophys. Res. Lett., 33(16), L16401, doi:10.1029/2006gl026962, 2006.

Tague, C., Grant, G., Farrell, M., Choate, J. and Jefferson, A.: Deep groundwater mediates streamflow response to climate warming in the Oregon Cascades, Clim. Change, 86(1), 189–210, 2008.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. and Watkins, M. M.: GRACE measurements of mass variability in the Earth system, Science (80-.)., 305(5683), 503–505, 2004.

The MathWorks: MATLAB and Statistics Toolbox Release 2013a, 2013.

Thompson, S. E., Harman, C. J., Schumer, R., Wilson, J. S., Basu, N. B., Brooks, P. D., Donner, S. D., Hassan, M. A., Packman, A. I. and Rao, P. S. C.: Patterns, puzzles and people: implementing hydrologic synthesis, Hydrol. Process., 25(20), 3256–3266, 2011.

United States Army Corps of Engineers: Comprehensive water supply study - An examination of current water supply issues, edited by T. M. Hillyer., 2001.

Voss, K. A., Famiglietti, J. S., Lo, M., Linage, C., Rodell, M. and Swenson, S. C.: Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region, Water Resour. Res., 2013.

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V, Kumar, P., Rao, P. S. C., Basu, N. B. and Wilson, J. S.: The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46(5), 2010.

Wagener, T., Sivapalan, M., Troch, P. and Woods, R.: Catchment classification and hydrologic similarity, Geogr. Compass, 1(4), 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.

Wahr, J., Swenson, S. and Velicogna, I.: Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33(6), L06401, doi:10.1029/2005gl025305, 2006.

Webster, C. S., Kingston, D. G. and Kerr, T.: Inter-annual variation in the topographic controls on catchment-scale snow distribution in a maritime alpine catchment, New Zealand, Hydrol. Process., 2014.

Western, A. W., Grayson, R. B. and Blöschl, G.: Scaling of soil moisture: A hydrologic perspective, Annu. Rev. Earth Planet. Sci., 30(1), 149–180, 2002.

Wu, H., Kimball, J. S., Elsner, M. M., Mantua, N., Adler, R. F. and Stanford, J.: Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers, Water Resour. Res., 48(11), W11530, doi:10.1029/2012wr012082, 2012.

Yeh, P. J. F., Swenson, S. C., Famiglietti, J. S. and Rodell, M.: Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), Water Resour. Res., 42(12), W12203, doi:10.1029/2006wr005374, 2006.

Zaitchik, B. F., Rodell, M. and Reichle, R. H.: Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River Basin, J. Hydrometeorol., 9(3), 535–548, doi:10.1175/2007JHM951.1, 2008.

Zang, C. F., Liu, J., Velde, M. and Kraxner, F.: Assessment of spatial and temporal patterns of green and blue water flows under natural conditions in inland river basins in Northwest China, Hydrol. Earth Syst. Sci., 16(8), 2859–2870, 2012.

Table 1: Comparison of performance metrics using the dual-pass approach to apply GRACE TWSA data, model derived snow (SWE), and soil moisture (SM) products in predicting seasonal (R_{season}) and August (R_{Aug}) runoff by watershed. Average values for the three basins are also provided. RMSE values are in mm. Complete results can be found in Appendix table A3.

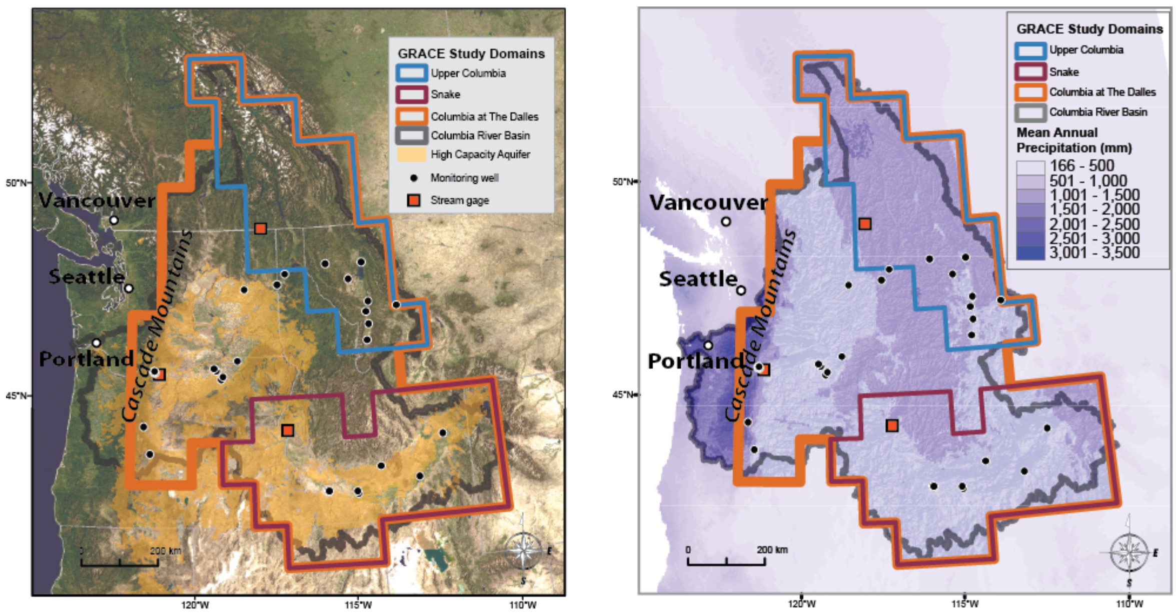
	R_{season}				R_{Aug}			
	$TWSA_{Mar}$	**$TWSA_{Apr}$**	**SWE_{Mar}**	**SM_{Mar}**	**$TWSA_{Mar}$**	**$TWSA_{Apr}$**	**SWE_{Mar}**	**SM_{Mar}**
Upper Columbia								
NSE	0.82	0.87	0.46	< 0	0.71	0.70	< 0	< 0
RMSE	33.06	27.62	56.10	> 1000	5.71	5.38	13.08	143.17
R^2	0.86	0.88	0.58	0.00	0.71	0.71	0.28	0.05
Snake River								
NSE	0.46	0.29	< 0	0.85	< 0	< 0	< 0	< 0
RMSE	14.03	15.71	21.53	7.38	13.59	0.76	0.78	0.72
R^2	0.59	0.47	0.08	0.86	0.15	0.08	0.27	0.29
The Dalles								
NSE	0.98	0.54	0.24	< 0	0.80	0.29	< 0	< 0
RMSE	6.01	26.50	26.48	122.88	1.86	3.31	18.91	22.10
R^2	0.98	0.71	0.39	0.00	0.82	0.71	0.03	0.02
Average								
NSE	0.75	0.57	0.35	0.85	0.76	0.50	< 0	< 0
RMSE	17.70	23.28	34.70	65.13	7.05	3.15	10.92	55.33
R^2	0.81	0.69	0.35	0.29	0.56	0.50	0.19	0.12

Table 2: Comparison of performance metrics from applying all nine water years of GRACE TWSA data, model derived snow (SWE), and soil moisture (SM) products in predicting seasonal (R_{season}) and August (R_{Aug}) runoff by watershed. Average values for the three basins are also provided. RMSE values are in mm. R^2 values are the same as NSE for this linear regression. Complete results can be found in Appendix table A4.

	R_{season}				R_{Aug}			
	$TWSA_{Mar}$	**$TWSA_{Apr}$**	**SWE_{Mar}**	**SM_{Mar}**	**$TWSA_{Mar}$**	**$TWSA_{Apr}$**	**SWE_{Mar}**	**SM_{Mar}**
Upper Columbia								
NSE	0.93	0.92	0.82	0.03	0.76	0.73	0.56	0.09
RMSE	22.18	23.18	36.19	82.90	6.60	6.90	8.92	12.79
Snake River								
NSE	0.83	0.75	0.34	0.93	0.68	0.52	0.62	0.76
RMSE	8.76	10.55	17.23	5.80	0.43	0.52	0.47	0.37
The Dalles								
NSE	0.98	0.91	0.67	0.00	0.88	0.91	0.46	0.02
RMSE	6.22	13.00	24.60	42.67	1.55	1.30	3.30	4.40
Average								
NSE	0.91	0.86	0.61	0.32	0.77	0.72	0.55	0.29
RMSE	12.39	15.58	26.01	43.79	2.86	2.91	4.23	5.85

Table 3: Parameters from the power function curves in each of the three watersheds using TWSA to predict streamflow. Figure 7 provides these results visually.

	Upper Columbia		**Snake River**		**The Dalles**	
	$TWSA_{Mar}$ R_{season}	**$TWSA_{Mar}$ R_{Aug}**	**$TWSA_{Mar}$ R_{season}**	**$TWSA_{Mar}$ R_{Aug}**	**$TWSA_{Mar}$ R_{season}**	**$TWSA_{Mar}$ R_{Aug}**
a	2.12E-10	4.83E-06	5.69E-05	2.26E-04	7.40E-10	3.61E-15
b	4.99	3.41	2.88	1.89	5.25	7.28
c	41.06	273.99	23.97	3.30	124.21	15.54



Watershed	Physical Characteristics	Climate
Upper Columbia (155,000 km^2)	- Steep topography - Low aquifer storage - Responsive runoff	Wet maritime conditions with 900 mm of annual precipitation, and substantial snowpack accumulation
Snake River (182,000 km^2)	- Low relief topography - High aquifer storage - Muted runoff	Dry continental climate with 350 mm of annual precipitation, falling primarily at upper elevations as snow
The Dalles (614,000 km^2)	- Includes the Upper Columbia and Snake River - Topography is a blend of the Upper Columbia and Snake River - High storage capacity - Large system that floods	Mix of dry continental climate plains and wet uplands. Mean annual precipitation is 625 mm, with considerable snowpack accumulation in the mountains

Fig. 1: Context map and descriptions of the three study watersheds and the locations of the groundwater wells used in the study.

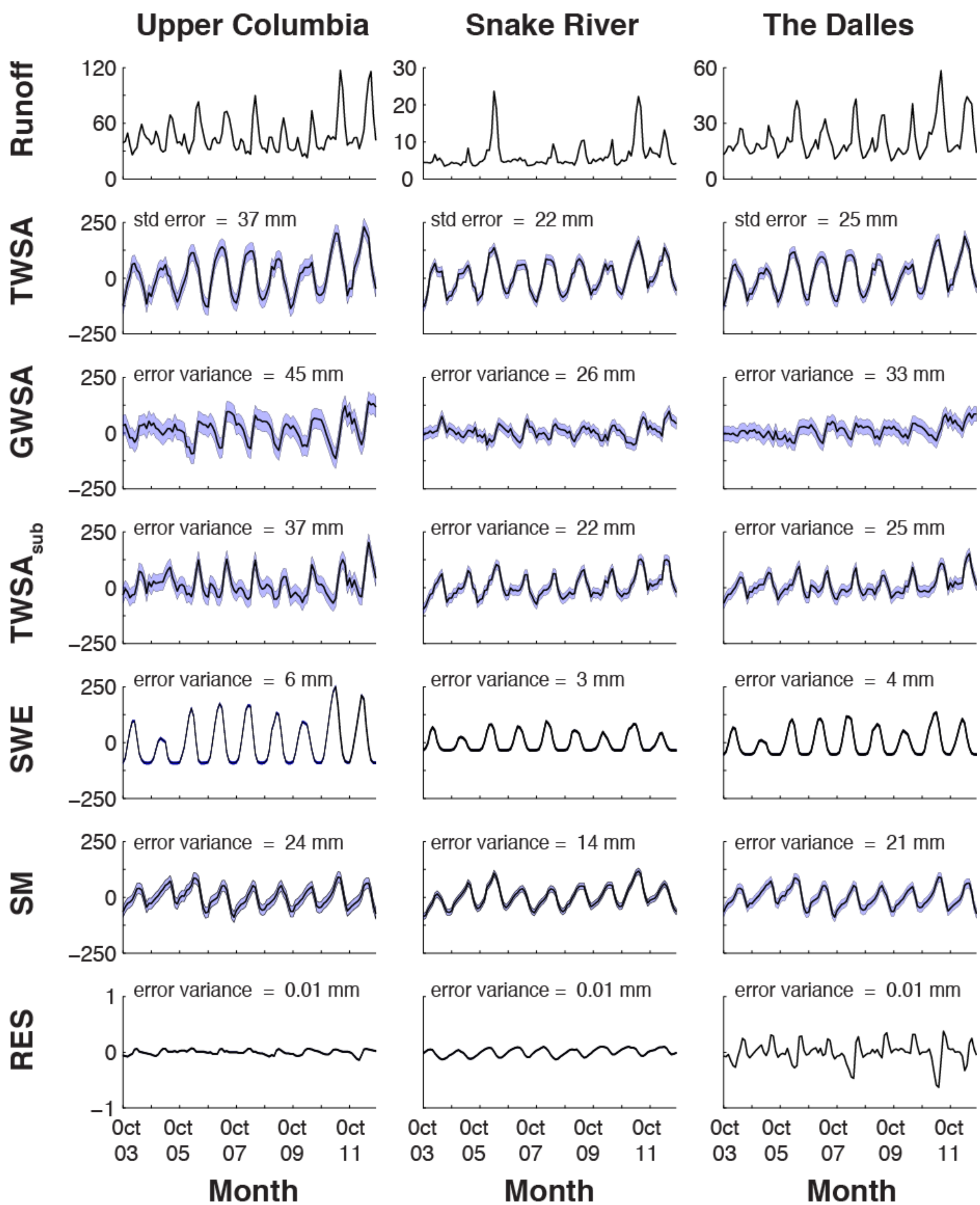


Fig. 2 Monthly storage anomalies for Runoff, *TWSA*, and the subcomponents of terrestrial water for the three watersheds. Standard errors and error variance for hydrological component are noted in each sub-figure, and represented by the blue shading. All units on the vertical axis are in mm. Note the different vertical scales for Runoff.

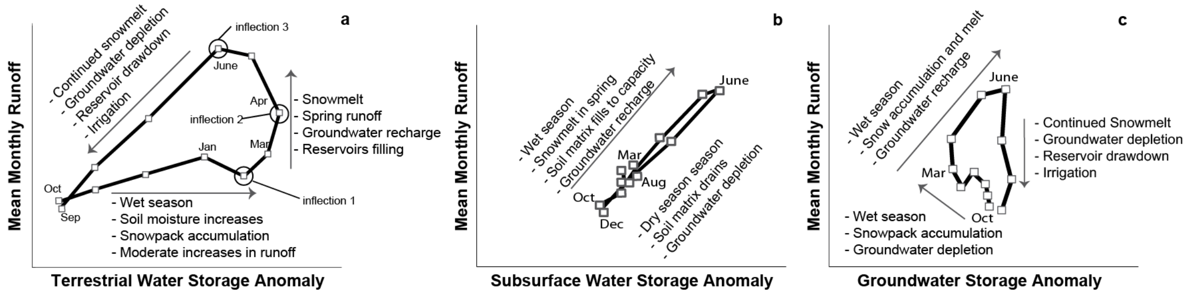


Fig. 3a-c: Annotated hysteresis curves of terrestrial water storage anomalies (a), the subsurface water storage anomalies ($TWSA_{sub}$; b), and groundwater storage anomalies (c) based upon the nine-year mean for the Columbia River at The Dalles. These curves describe the fluxes of water moving through the watershed.

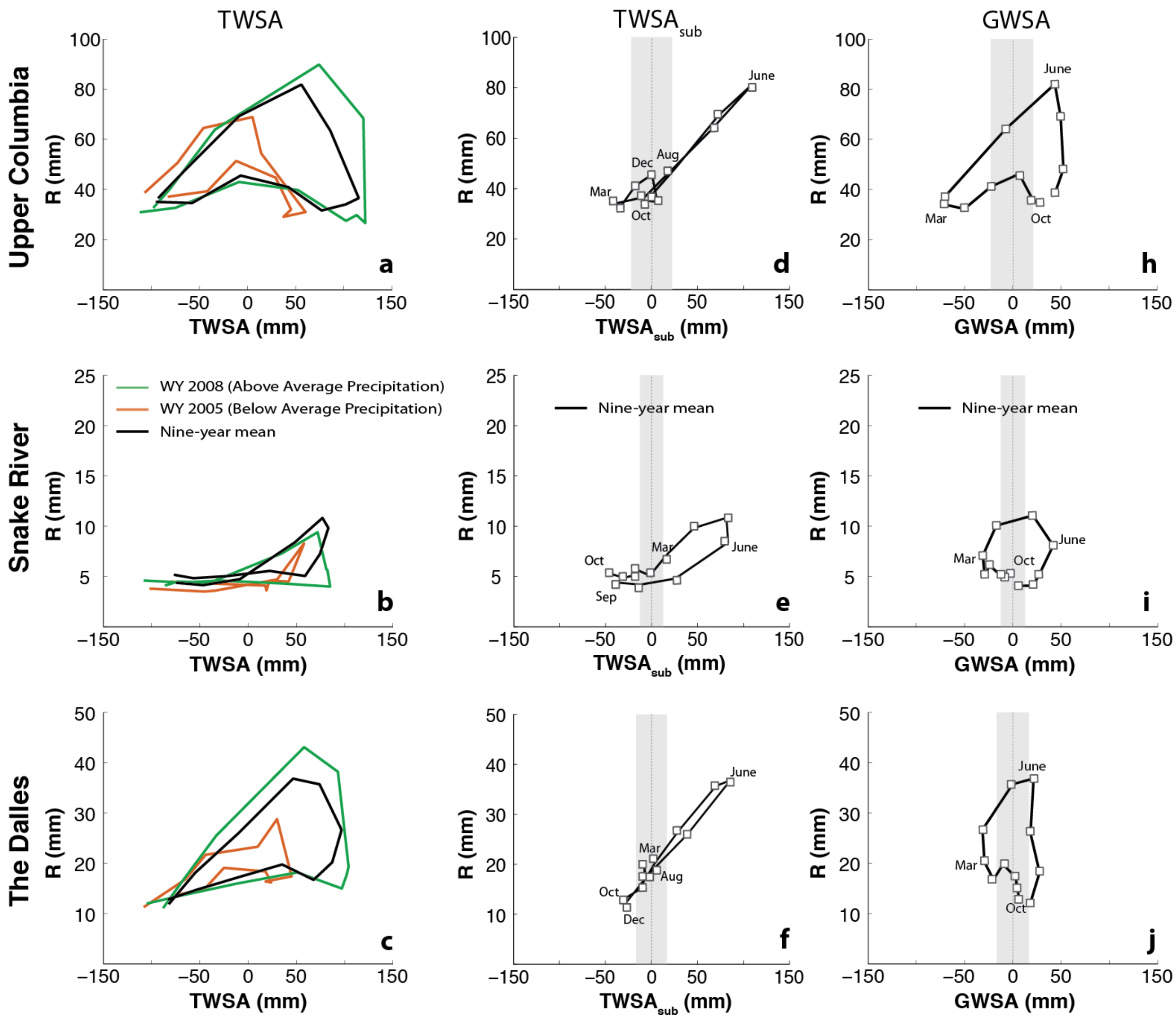


Fig. 4a-f: Individual hysteresis curves for the three study watersheds for *TWSA* (a-c), $TWSA_{sub}$ (d-f), and GWSA (h-j). $TWSA_{sub}$ in the Upper Columbia and The Dalles collapses to represent a shape more commonly associated with the hysteresis of a soil matrix. The Snake River retains a similar looping shape. The grey areas in the $TWSA_{sub}$ and *GWSA* plots provide a visual reference of the *TWSA* error variance for each watershed. The low topography and high storage capacity of the Snake aquifer provides a consistent groundwater signal, as compared to the limited aquifer of the Upper Columbia, which fills and drains quickly. Note the different scales on the y-axes.

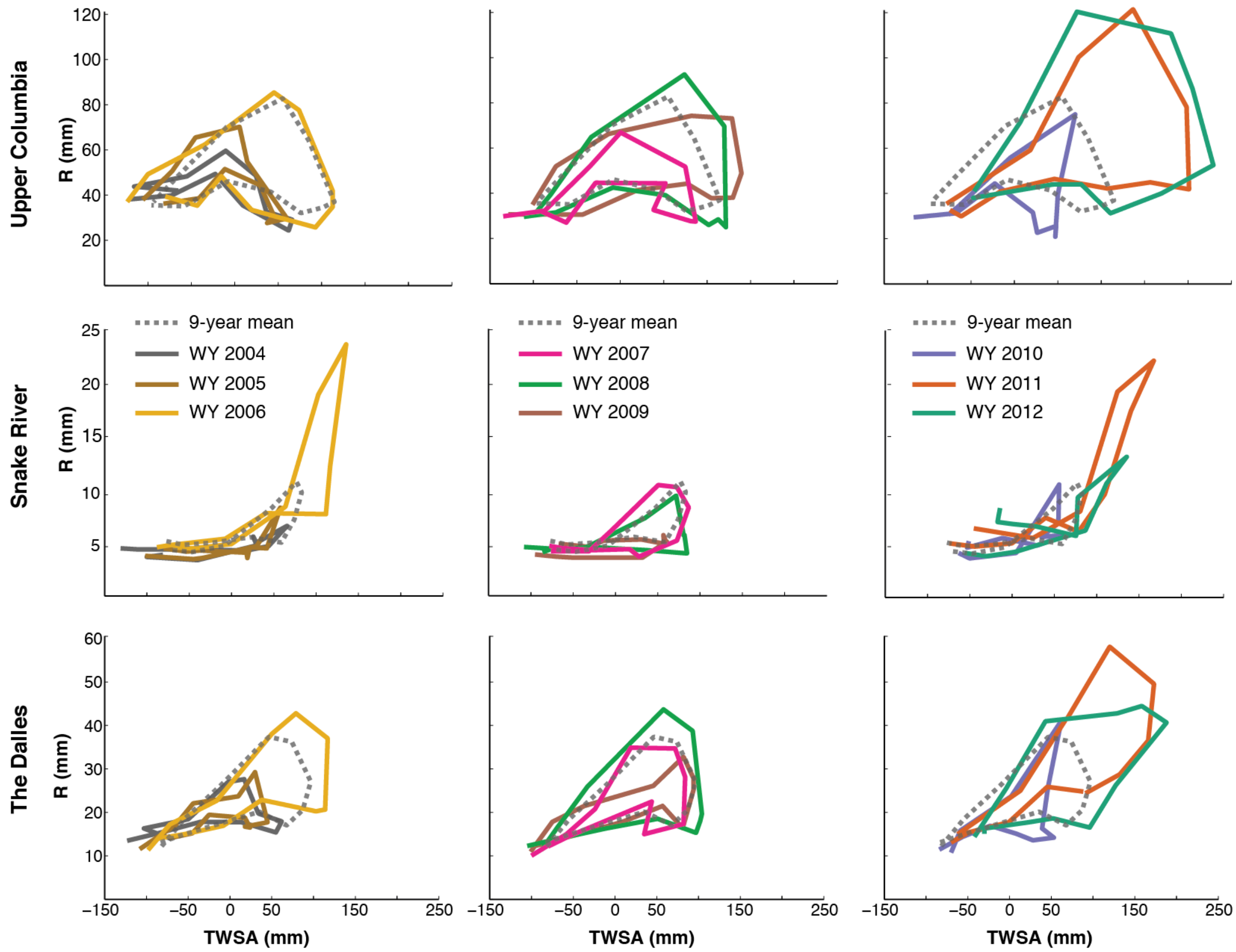


Fig. 5: Plots of the hysteresis curves for *TWSA* in each of the three study watersheds across all nine water years. For visual clarity, each plot contains three water years and the nine-year mean. Note the different scales on the y-axes for each basin.

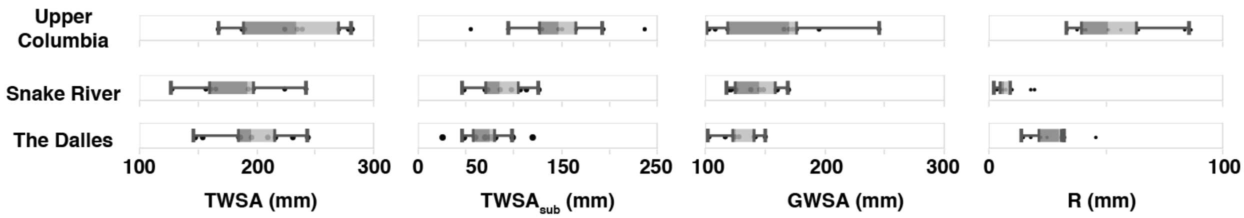


Fig. 6: The intra-annual range of *TWSA*, $TWSA_{sub}$, *GWSA*, and *R* for the nine water years of the study period.

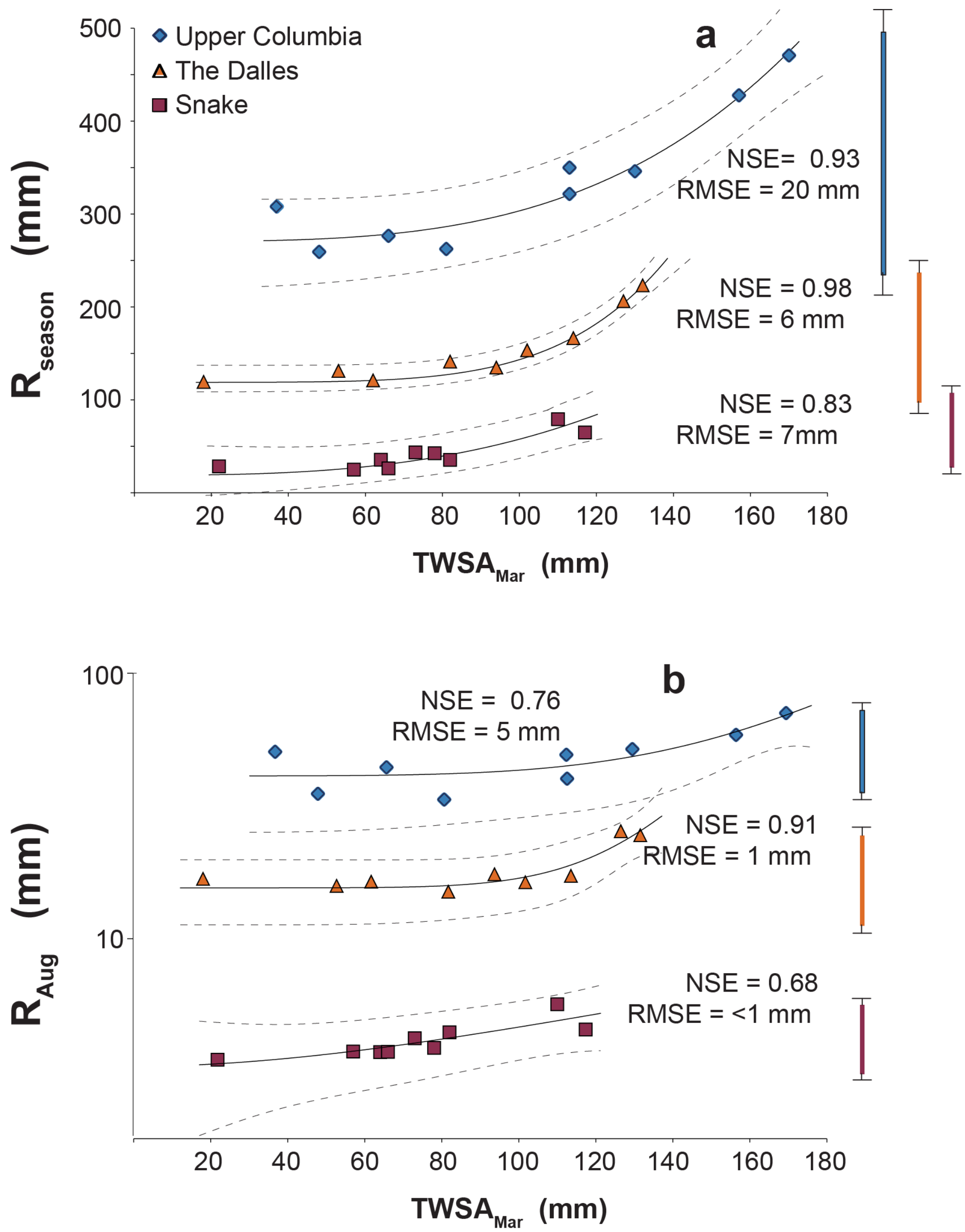


Fig. 7a-b: Measurements of terrestrial water storage anomalies in March ($TWSA_{\mathrm{Mar}}$) effectively predict the cumulative runoff for April – September (R_{season}; a), and help describe how these three regional watersheds function as simple non-linear systems. $\mathrm{TWSA}_{\mathrm{Mar}}$ also predicts mean runoff for August (R_{Aug}; b), one of the driest months of the year when demand for water is at its peak. The hashed lines represent the 95% confidence intervals. The box plots to the right of each plot represent the range of *R* for the respective watershed from WY's 1969 – 2012. Note the semi-log y-axis on (b). For complete results and parameters from the empirical model please refer to Tables 1, 2, 3, A3, and A4.

Table A1: The reservoirs used in the GRACE analysis.

Reservoir Name	Operating Agency	Normal Operating Capacity (m^3)
Grand Coulee	US Department of Interior	1.16×10^{10}
Libby	US Army Corps of Engineers	7.17×10^{9}
Hungry Horse	US Department of Interior	4.28×10^{9}
Dworsha	US Army Corps of Engineers	4.26×10^{9}
American Falls	US Department of Interior	2.10×10^{9}

Table A2: The groundwater wells used in the analysis that compares GRACE-derived groundwater with location-specifc wells. USGS is the United States Geological Survey and IDWR is the Idaho Department of Water Resources.

Well Number	Operating Agency
434400121275801	USGS
442242121405501	USGS
452855119064701	USGS
453239119031501	USGS
453845121191401	USGS
453937121215801	USGS
453944121211301	USGS
454013121225901	USGS
454027121212501	USGS
454040121222901	USGS
454047121203701	USGS
454100119164801	USGS
454416119212801	USGS
455418118333001	USGS
461518114090802	USGS
463750114033001	USGS
465520114074001	USGS
470049113035401	USGS
470946114013201	USGS
473442118162201	USGS
474011117072901	USGS
474251114385201	USGS
475439116503401	USGS
480519114091001	USGS
480621115244901	USGS
02S20E-01ACC2	IDWR
07S06E-29BBA1	IDWR
08S06E-03BDC1	IDWR
07S06E-34BCA1	IDWR
09S14E-03BAA1	IDWR
08S14E-16CBB1	IDWR
05S31E-27ABA1	IDWR
07N38E-23DBA1	IDWR

Table A3: Comparison of performance metrics using the dual-pass approach to apply GRACE *TWSA*, model derived snow (*SWE*), soil moisture (*SM*), and subsurface ($TWSA_{sub}$) data in predicting seasonal (R_{season}) and August (R_{Aug}) runoff by watershed. RMSE values are in mm.

Upper Columbia

		TWSA			SM			SWE			$TWSA_{sub}$		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	< 0	0.82	0.87	< 0	< 0	< 0	< 0	0.46	< 0	< 0	< 0	< 0
	RMSE	84	33	28	>1000	>1000	134	110	56	309	>1000	>1000	354
	R^2	0.43	0.86	0.88	0.01	0.00	0.07	0.23	0.58	0.27	0.15	0.02	0.02
R_{July}	**NSE**	< 0	0.90	0.84	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0
	RMSE	32	7	8	>1000	71	56	28	25	108	>1000	>1000	123
	R^2	0.19	0.93	0.92	0.01	0.00	0.00	0.32	0.45	0.24	0.05	0.01	0.01
R_{Aug}	**NSE**	< 0	0.71	0.70	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0
	RMSE	228	6	5	>1000	143	32	12	13	51	>1000	>1000	30
	R^2	0.19	0.71	0.71	0.07	0.05	0.30	0.25	0.28	0.12	0.18	0.11	0.01
R_{Sept}	**NSE**	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0
	RMSE	2	21	104	4	28	10	>1000	3	50	20	587	6
	R^2	0.12	0.06	0.12	0.09	0.24	0.20	0.04	0.07	0.04	0.04	0.02	0.03

Snake River

		TWSA			SM			SWE			$TWSA_{sub}$		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	< 0	0.46	0.29	0.58	0.85	< 0	< 0	< 0	0.09	< 0	< 0	< 0
	RMSE	258	14	16	12	7	52	5	22	8	>1000	108	474
	R^2	0.21	0.59	0.47	0.64	0.86	0.29	0.00	0.08	0.13	0.04	0.11	0.01
R_{July}	**NSE**	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0
	RMSE	23	3	2	2	2	40	1	2	1	99	>1000	35
	R^2	0.00	0.05	0.01	0.01	0.09	0.11	0.15	0.00	0.04	0.00	0.06	0.02
R_{Aug}	**NSE**	< 0	< 0	-0.70	< 0	< 0	< 0	< 0	< 0	0.65	< 0	< 0	< 0
	RMSE	11	13.59	0.76	1	1	2	0	1	1	>1000	>1000	474
	R^2	0.05	0.15	0.08	0.06	0.29	0.10	0.00	0.27	0.67	0.04	0.11	0.01
R_{Sept}	**NSE**	< 0	< 0	-0.94	< 0	< 0	< 0	< 0	< 0	0.03	< 0	< 0	< 0
	RMSE	16	1	1	1	1	1	0	1	0	140	8	435
	R^2	0.01	0.04	0.03	0.07	0.15	0.11	0.03	0.00	0.11	0.00	0.00	0.01

The Dalles

		TWSA			SM			SWE			$TWSA_{sub}$		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	< 0	0.98	0.54	< 0	< 0	< 0	< 0	0.24	0.14	< 0	< 0	< 0
	RMSE	84	6	27	267	122	363	>1000	26	26	13	5231	737
	R^2	0.20	0.98	0.71	0.01	0.00	0.02	0.13	0.39	0.29	0.02	0.00	0.00
R_{July}	**NSE**	< 0	0.86	< 0	< 0	< 0	< 0	< 0	0.28	< 0	< 0	< 0	< 0
	RMSE	19	3	10	>1000	16	80	>1000	4	6	4	4	311
	R^2	0.05	0.86	0.64	0.00	0.00	0.02	0.03	0.30	0.10	0.00	0.00	0.00
R_{Aug}	**NSE**	< 0	0.80	0.29	< 0	< 0	< 0	< 0	< 0	0.05	< 0	< 0	< 0
	RMSE	9	2	3	>1000	22	16	>1000	19	2	2	1	3
	R^2	0.04	0.82	0.71	0.04	0.02	0.00	0.02	0.03	0.12	0.00	0.00	0.12
R_{Sept}	**NSE**	< 0	0.41	< 0	< 0	< 0	< 0	< 0	< 0	< 0	< 0	0.05	< 0
	RMSE	5	1	3	756	3	7	1	$5x10^9$	$7x10^{10}$	6	1	2
	R^2	0.00	0.42	0.28	0.03	0.01	0.03	0.06	0.02	0.02	0.22	0.06	0.14

Table A4: Comparison of performance metrics from applying all nine water years of GRACE *TWSA*, model derived snow (*SWE*), soil moisture (*SM*), and subsurface ($TWSA_{sub}$) data in predicting seasonal (R_{season}) and August (R_{Aug}) runoff by watershed. RMSE values are in mm. R^2 values are the same as NSE for this linear regression.

Upper Columbia

		TWSA			**SM**			**SWE**			**TWSA$_{sub}$**		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	0.84	0.93	0.92	0.01	0.03	0.33	0.63	0.82	0.62	0.15	0.22	0.22
	RMSE	28.62	19.81	20.72	8.38	14.30	36.80	37.78	30.27	37.85	28.22	32.50	32.50
R_{July}	**NSE**	0.75	0.95	0.96	0.01	0.00	0.18	0.53	0.79	0.60	0.05	0.22	0.22
	RMSE	10.38	5.00	4.74	2.16	1.34	9.10	11.95	9.80	11.73	5.38	9.86	9.86
R_{Aug}	**NSE**	0.62	0.76	0.73	0.07	0.09	0.44	0.37	0.56	0.34	0.18	0.11	0.23
	RMSE	6.02	5.31	5.48	3.12	3.50	6.15	6.00	6.16	5.87	4.80	3.95	5.22
R_{Sept}	**NSE**	0.20	0.07	0.13	0.31	0.28	0.40	0.04	0.04	0.10	0.39	0.15	0.51
	RMSE	1.60	1.05	1.32	1.85	1.80	1.96	0.80	0.80	1.22	1.95	1.42	2.00

Snake River

		TWSA			**SM**			**SWE**			**TWSA.sub**		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	0.39	0.83	0.75	0.84	0.93	0.91	0.09	0.34	0.60	0.35	0.39	0.42
	RMSE	9.59	7.39	8.48	7.15	5.16	5.64	5.60	9.37	9.65	9.39	9.63	9.71
R_{July}	**NSE**	0.07	0.43	0.43	0.41	0.63	0.51	0.09	0.21	0.70	0.05	0.19	0.23
	RMSE	0.41	0.80	0.80	0.79	0.78	0.81	0.46	0.66	0.74	0.34	0.63	0.68
R_{Aug}	**NSE**	0.35	0.68	0.52	0.56	0.76	0.61	0.24	0.62	0.91	0.13	0.09	0.12
	RMSE	0.34	0.33	0.35	0.35	0.30	0.34	0.30	0.34	0.21	0.24	0.20	0.22
R_{Sept}	**NSE**	0.18	0.53	0.58	0.60	0.88	0.66	0.08	0.30	0.91	0.16	0.18	0.18
	RMSE	0.34	0.44	0.44	0.43	0.29	0.42	0.25	0.41	0.25	0.32	0.34	0.34

The Dalles

		TWSA			**SM**			**SWE**			**TWSA.sub**		
		Feb	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**	**Feb**	**Mar**	**Apr**
R_{season}	**NSE**	0.48	0.98	0.91	0.00	0.01	0.22	0.21	0.67	0.65	0.19	0.23	0.27
	RMSE	19.82	5.70	11.43	2.10	3.59	16.53	16.06	18.65	18.95	15.43	16.74	17.61
R_{July}	**NSE**	0.27	0.89	0.89	0.04	0.03	0.09	0.07	0.52	0.51	0.20	0.38	0.40
	RMSE	4.05	2.90	2.87	1.73	1.52	2.64	2.27	4.55	4.55	3.66	4.43	4.47
R_{Aug}	**NSE**	0.29	0.88	0.91	0.04	0.02	0.24	0.05	0.45	0.42	0.34	0.44	0.49
	RMSE	1.89	1.34	1.22	0.77	0.65	1.78	0.88	2.07	2.05	1.96	2.06	2.08
R_{Sept}	**NSE**	0.20	0.57	0.53	0.03	0.03	0.13	0.02	0.29	0.34	0.37	0.15	0.35
	RMSE	0.75	0.94	0.94	0.34	0.31	0.63	0.28	0.86	0.90	0.92	0.67	0.90

Table A5: Scale factors for the GRACE data in each of the three watersheds.

		1.03	1.10								
			0.93	1.15	1.31						
		0.79	0.61	0.62	0.87	0.87	0.72				
	2.17	1.51	1.32	1.44	1.18	1.65	1.70				
	3.04	2.59	1.70	0.94	0.51	0.90	1.08	1.35			
	2.56	2.18	1.42	0.75	0.45	0.52	0.79	1.22			
1.93	3.25	2.00	0.96	0.82	0.60	0.60	0.80	0.90	0.84		
3.38	3.27	1.87	1.13	1.24	1.01	0.97	1.01	0.97			
2.70	2.59	1.65	1.30	1.53	1.15	1.15	1.16	1.10	0.94	0.74	1.01
3.12	2.49		1.28	1.28	0.89	0.94	1.22	1.26	0.94	0.58	1.21
				1.16	0.75	0.86	0.92	0.93	0.88	0.77	1.38
					0.79	0.83	0.78				

Upper Columbia

Snake

The Dalles (all cells)

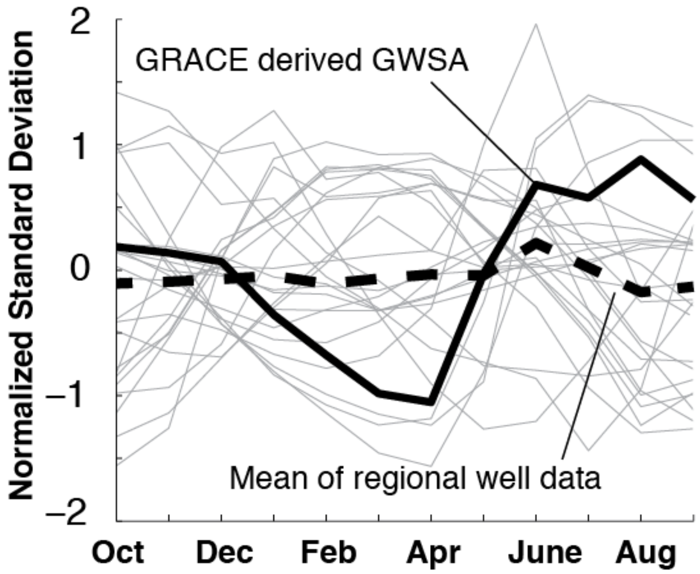


Fig A1: The normalized GRACE-derived groundwater anomaly compared to normalized well data over the study period. The grey lines in the background are the 33 individual wells, and the hashed line represents the mean of these wells. While some wells match the general GRACE-derived *GWSA*, variability across wells creates a muted mean signal.

GRACE storage-runoff hystereses reveal the dynamics of regional watersheds

E.A. Sproles[1,2,3]; S.G. Leibowitz[4]; J.T. Reager[5]; P.J. Wigington Jr[4,6]; J.S. Famiglietti[5]; S.D. Patil[7]

*corresponding author: eric.sproles@gmail.com

[1] Oak Ridge Institute for Science and Technology, c/o National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR, 97333, USA

[2] Current address: Centro de Estudios en Zonas Áridas, Universidad de La Serena, Raul Bitran 1305, La Serena, Chile

[3] College of College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, 97331-5503, USA

[4] National Health and Environmental Effects Research Laboratory, US Environmental Protection Agency, 200 SW 35th Street, Corvallis, OR, 97333, USA

[5] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, 91109, USA

[6] Retired

[7] School of Environment, Natural Resources and Geography, Bangor University, Bangor, LL57 2UW, Wales

E Sproles 2/20/15 11:07 AM
Eliminado: streamflow

E Sproles 2/20/15 11:07 AM
Con formato: Inglés (americano), No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Inglés (americano), No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Inglés (americano), No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: be both time consuming and reliant on multiple data inputs

E Sproles 2/20/15 11:07 AM
Con formato: español

Abstract

We characterize how regional watersheds function as simple, dynamic systems through a series of hysteresis loops using measurements from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites. These loops illustrate the temporal relationship between runoff and terrestrial water storage in three regional-scale watersheds (>150,000 km^2) of the Columbia River Basin, USA and Canada. The shape and size of the hysteresis loops are controlled by the climate, topography, and geology of the watershed. The direction of the hystereses for the GRACE signals move in opposite directions from the isolated groundwater hystereses, suggesting that regional scale watersheds require soil water storage to reach a certain threshold before groundwater recharge and peak runoff occur. The subsurface water (soil moisture and groundwater) hystereses more closely resemble the storage-runoff relationship of a soil matrix. While the physical processes underlying these hystereses are inherently complex, the vertical integration of terrestrial water in the GRACE signal encapsulates the processes that govern the non-linear function of regional-scale watersheds. We use this process-based understanding to test how GRACE data can be applied prognostically to predict seasonal runoff (mean Nash-Sutcliffe Efficiency of 0.91) and monthly runoff during the low flow/high demand month of August (mean Nash-Sutcliffe Efficiency of 0.77) in all three watersheds. The global nature of GRACE data allows this same methodology to be applied in other regional-scale studies, and could be particularly useful in regions with minimal data and in trans-boundary watersheds.

E Sproles 2/20/15 11:07 AM
Eliminado: .

E Sproles 2/20/15 11:07 AM
Eliminado: using measurements from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites

E Sproles 2/20/15 11:07 AM
Con formato: Inglés (americano)

E Sproles 2/20/15 11:07 AM
Eliminado: signal

E Sproles 2/20/15 11:07 AM
Eliminado: R^2

E Sproles 2/20/15 11:07 AM
Con formato: Sin Superíndice / Subíndice

E Sproles 2/20/15 11:07 AM
Eliminado: R^2

E Sproles 2/20/15 11:07 AM
Con formato: Sin Superíndice / Subíndice

1. Introduction

At the most fundamental level, watershed processes can be described as the collection, storage, and release of water (Black, 1996; McDonnell et al., 2007). The runoff from these processes is governed by threshold mediated relationships across scales that result in storage—runoff hystereses (Spence, 2010). These threshold relationships between storage and runoff (*S*—*R*) are not uniform across a watershed, functioning as a series of discontinuous processes in soils and hillslopes that provide an integrated *S*—*R* relationship at the watershed scale (Spence, 2010). Kirchner (2009) described the *S*—*R* relationship to be non-linear and stated that watersheds typically function as dynamic systems governed by their unique climate and geology. These conceptual models of hydrologic behaviors help provide a process-based understanding of watersheds as dynamic environmental systems (Aspinall, 2010), and identify connections that advance hydrologic science and hydrologic prediction (Wagener et al., 2007).

At the local scale, *in situ* instrumentation can quantify the non-linear relationship between streamflow and water stored in a watershed as snow, soil moisture, groundwater and reservoirs (Appleby, 1970; Brutsaert, 2008; Kirchner, 2009; Sayama et al., 2011). These four primary storage components, along with climate, topography, and geology govern the fluxes of water through a catchment, and play an important role in the hysteretic nature of storage and runoff dynamics (McGlynn and McDonnell, 2003; McNamara et al., 2011). Knowledge of these processes is fundamental to developing an understanding of a watershed's hydrologic behavior. However, observations over larger regions can be technically challenging and costly, and *in situ* measurements from small basins do not necessarily represent the complexity inherent to watersheds at more broad scales (Spence, 2010). This scaling problem limits our capacity to understand and predict regional hydrologic processes, which is often the practical scale of watershed management (Blöschl, 2001; Western et al., 2002; Skøien et al., 2003; Peel and Blöschl, 2011; Thompson et al., 2011).

In the absence of broad-scale observations, past hydrological studies have typically relied on *in situ* measurements as a proxy for regional scale hydrological processes. For example, in higher latitude or mountainous regions measurements of snow water storage have provided a simple metric that has been used in water resource planning for decades (Cayan, 1996; United States Army Corps of Engineers, 2001), and are often correlated to streamflow gauged downstream (Dozier, 2011). While informative, this approach can often provide hydrological forecasts that are misleading, because point-based measurements do not fully represent the broad-scale variability of rugged mountain terrain (Dozier, 2011; Nolin, 2012; Webster et al., 2014; Ayala et al., 2014). Similarly, measurements of soil moisture in the upper 2000 mm of the soil rely on point-based data that are often distributed at the regional scale, but do not effectively represent the true variability of soil moisture found at the regional scale (Western et al., 2002; Brocca et al., 2010). A complete understanding of groundwater stores and fluxes (deeper than 2000 mm) at regional scales also remains elusive, despite its increasing importance in water resources management (Wagener et al., 2007; Gleeson et al., 2012; Famiglietti and Rodell, 2013; Barthel, 2014). In addition to contributing to runoff, groundwater serves as an important water resource for consumptive use (Gleeson et al., 2012).

While local-scale methods have been applied with moderate success in the past, current trends in climate and in consumptive water demand suggest that long-term changes in hydrological fluxes will have a major impact at the regional scale (Milly et al., 2008). As a result, the supply and demand of water is also expected to shift, especially at the regional scale (Wagener et al., 2010; Gleick, 2014a).

E Sproles 2/20/15 11:07 AM

Eliminado: At the most fundamental level, watershed processes can be described as the collection, storage, and release of water (Black, 1996; McDonnell et al., 2007). At a more complex level, watersheds typically function as non-linear, dynamic systems governed by their unique climate and geology (Kirchner, 2009). Gaining insights into hydrologic processes and behaviors helps to provide a process-based understanding of watersheds as dynamic environmental systems (Aspinall, 2010), and to identify connections that advance hydrologic science and hydrologic prediction (Wagener et al., 2007). At the local scale, *in situ* instrumentation can quantify the non-linear relationship between streamflow and water stored in a watershed as snow, soil moisture, groundwater and reservoirs (Appleby, 1970; Brutsaert, 2008; Kirchner, 2009; Sayama et al., 2011). These four primary storage components, along with landscape and topography, govern the fluxes of water through a catchment, and play an important role in the hysteretic nature of storage and runoff dynamics (McGlynn and McDonnell, 2003; McNamara et al., 2011). Knowledge of these processes is fundamental to developing an understanding of a watershed's hydrologic behavior. However, observations over larger regions can be technically challenging and costly, and *in situ* measurements from small basins do not necessarily represent the complexity inherent to watersheds at more broad scales. This scaling problem limits our capacity to understand and predict regional hydrologic processes, which is often the practical scale of watershed management (Blöschl, 2001; Western et al., 2002; Skøien et al., 2003; Peel and Blöschl, 2011; Thompson et al., 2011).

E Sproles 2/20/15 11:07 AM

Eliminado: In the absence of large scale observations, past hydrological studies have typically relied on *in situ* measurements as a proxy for regional scale hydrological processes. For example, in higher latitude or mountainous regions measurements of snow water storage have provided a simple metric that has been used in water resource planning for decades (Cayan, 1996; United States Army Corps of Engineers, 2001), and are often correlated to streamflow gauged downstream (Dozier, 2011). While informative, this approach can often provide hydrological forecasts that are misleading, because point-based measurements do not fully represent the broad-scale variability of rugged mountain terrain (Dozier, 2011; Nolin, 2012; Webster et al., 2014; Ayala et al., 2014). Similarly, measurements of soil moisture in the upper 2000 mm of the soil rely on point-based data that are often distributed at the regional scale, but do not effectively represent the true variability of soil moisture found at the regional scale (Western et al., 2002; Brocca et al., 2010). A complete understanding of groundwater stores and fluxes (deeper than 2000 mm) at regional scales also remains elusive, despite its increasing importance in water resources management (Wagener et al., 2007; Gleeson et al., 2012; Famiglietti and Rodell, 2013; Barthel, 2014). In addition to contributing to runoff, groundwater serves as an important water resource for consumptive use (Gleeson et al., 2012). While local-scale methods have been applied with moderate success in the past, current trends in climate and in consumptive water demand su ... [1]

Hydrologic models can help address the questions of scale and bridge the gap between local scale observations and regional-scale processes by estimating the primary components of water storage (snow, soil moisture, reservoir, and groundwater) across a larger spatial grid. Regional-scale modeling approaches are integrated into water resource management operations for navigation, human consumptive use, irrigation, and hydropower (Payne et al., 2004; Rodell et al., 2004). Models can also be applied diagnostically to test scientific hypotheses and provide a better understanding of the physical processes that govern real world systems, such as the connections between snowmelt, streamflow, and groundwater (Beven, 2007, 2010; Moradkhani and Sorooshian, 2008; Kirchner, 2009; Clark et al., 2011; Capell et al., 2012). Despite their utility, developing and validating a model can be both time consuming and reliant on multiple data inputs, which even in the most well-instrumented basins provides sparse geographic coverage (Bales et al., 2006; Zang et al., 2012). The lack of an integrated measurement of water storage and streamflow has limited regional-scale hydrologic insights to model-based studies (Koster et al., 2010; Mahanama et al., 2011).

Since 2002, broad-scale measurements of changes in the amount of water stored across and through the earth have been available from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites (Tapley et al., 2004). GRACE measures monthly changes in the Earth's gravitational field that are proportional to regional changes in total water storage (Wahr et al., 2006). GRACE satellites provide a monthly record of terrestrial water storage anomalies (*TWSA*), which represent the changes in the vertical sum of water at the Earth's surface stored in snow, surface, soil and groundwater. Water losses to runoff and evapotranspiration are implicit in the GRACE storage signal, removing the added layer of complexity typically required to model the terrestrial water balance.

E Sproles 2/20/15 11:07 AM
Eliminado: Hydrologic models can help address the questions of scale and bridge the gap between local scale observations and regional-scale processes by estimating the primary components of water storage (snow, soil moisture, reservoir, and groundwater) across a larger spatial grid. Regional-scale modeling approaches are integrated into water resource management operations for navigation, human consumptive use, irrigation, and hydropower (Payne et al., 2004; Rodell et al., 2004). Models can also be applied diagnostically to test scientific hypotheses and provide a better understanding of the physical processes that govern real world systems, such as the connections between snowmelt, streamflow, and groundwater (Beven, 2007, 2010; Moradkhani and Sorooshian, 2008; Kirchner, 2009; Clark et al., 2011; Capell et al., 2012). Despite their utility, developing and validating a model can be both time consuming and reliant on multiple data inputs, which even in the most well-instrumented basins provides sparse geographic coverage (Bales et al., 2006; Zang et al., 2012). The lack of an integrated measurement of water storage and streamflow has limited regional-scale hydrologic insights to model-based studies (Koster et al., 2010; Mahanama et al., 2011).

E Sproles 2/20/15 11:07 AM
Eliminado: (Tapley et al., 2004)

E Sproles 2/20/15 11:07 AM
Eliminado: (Wahr et al., 2006)

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

GRACE data, coupled with modeled and measured variations of water stored in snow, surface reservoirs and soils, have successfully been decomposed to quantify regional groundwater changes (Rodell et al., 2009; Famiglietti et al., 2011; Voss et al., 2013; Castle et al., 2014) and have contributed to improving water balance calculations (Zaitchik et al., 2008; Li et al., 2012). More recent efforts have quantified the relationship between regional water storage and specific streamflow events (Reager and Famiglietti, 2009; Reager et al., 2014). Riegger and Tourian (2014) coupled GRACE data using data-driven and model-based approaches to better understand the relationship between storage and runoff across climatic zones globally. Their study found that coupled liquid storage is linear to runoff, and that in climatic regions with snow and ice the relationship between storage and runoff is more hysteretic. These novel analyses, which are more diagnostic in nature, have provided new insights into regional watershed hydrology using GRACE measurements as a core data input. These studies have not explored how topography and geology can also help describe the *S—R* relationship of regional watersheds. Nor did these studies examine the ability of GRACE measurements to predict seasonal runoff.

In this paper, we use terrestrial water storage data from GRACE to better understand the hydrology of regional watersheds and the relationship between storage and runoff. The temporal relationships between coincident *TWSA* and discharge observations at three scales in the Columbia River Basin (CRB) of western North America are investigated using climate, topography, and geology as a framing principle to describe the shape of the storage-streamflow hysteresis. We associate regional and temporal differences in the hystereses with varying watershed dynamics. Finally, we compare the prognostic abilities of GRACE observations with individual modeled estimates of snow and soil moisture to predict seasonal streamflow at regional scales.

E Sproles 2/20/15 11:07 AM

Eliminado: GRACE data, coupled with modeled and measured variations of water stored in snow, surface reservoirs and soils, have successfully been decomposed to quantify regional groundwater changes (Rodell et al., 2009; Famiglietti et al., 2011; Voss et al., 2013; Castle et al., 2014) and have contributed to improving water balance calculations (Zaitchik et al., 2008; Li et al., 2012). More recent efforts have quantified the relationship between regional water storage and specific streamflow events (Reager et al., 2014), and have described regional storage-streamflow hysteresis for large basins (Riegger and Tourian, 2014). Although these previous studies have provided new insights into regional watershed hydrology, their analyses are more diagnostic in nature and did not explore the processes behind the observed behavior.

E Sproles 2/20/15 11:07 AM

Con formato: Fuente: Sin Negrita, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM

Eliminado: In this paper, we use GRACE observations of terrestrial water storage observations to expand upon a fundamental concept in watershed hydrology—that the temporal relationship between storage and runoff can be used to quantify complex watershed behavior at broad scales, including groundwater recharge amounts and timing, baseflow recession characteristics, and long lead-time streamflow prediction (Brutsaert, 2008; Sayama et al., 2011; Reager et al., 2014; Riegger and Tourian, 2014). The temporal relationship between coincident TWSA and discharge observations at three scales in the Columbia River Basin (CRB) of western North America is investigated using climate and geology as a framing principle to describe the shape of the storage-streamflow hysteresis. We associate regional and temporal differences in the hystereses with varying watershed dynamics. Finally, we compare the prognostic abilities of GRACE observations to individual modeled estimates of snow and soil moisture to predict streamflow at regional scales.

2. Study Area

Our study area is the Columbia River Basin (CRB; 41-53ºN and 110-122ºW; Fig. 1). This basin has dry summers and wet winters. Up to 70% of annual precipitation falls between November and March, 50-60% of which occurs as snow (Serreze et al., 1999; Nolin et al., 2012). The spring months (April to June) are also wet, but warmer. Precipitation during the spring combines with snowmelt to swell rivers and potentially exacerbate flooding. Snowmelt also serves as a critical component of the hydrologic cycle, recharging aquifers and filling streams later in the year. These contributions bridge the temporal disconnect between wet winters and dry summers when demand is at its peak as farmers, fish, hydropower and municipal users vie for over-allocated water resources (United States Army Corps of Engineers, 2001; Oregon Water Supply and Conservation Initiative, 2008). However, concerns with winter surplus and summer scarcity are not uniform across the CRB, since climate and geology vary greatly. Two of the study watersheds, the Upper Columbia (155,000 km^2) and the Snake River basin (182,000 km^2), represent distinctly different climatic, topographic, and geologic provinces of the CRB (described and illustrated in Fig. 1). The Upper Columbia is wet and is characterized by steep topography of fractured rock and poor groundwater storage. In contrast, the arid Snake River basin is bowl-shaped with mountains on three sides. The interior of the Snake River basin is a broad plain with well-developed soils and expansive aquifer storage. The Columbia River at The Dalles (614,000 km^2) encompasses the Upper Columbia and the Snake River sub-basins, and its climate and geology are an integration of the two (Fig. 1). A distinct climatic feature of the Columbia River at The Dalles is the western slope of the Cascade Mountains, where over 3000 mm of mean annual precipitation at higher elevations sustains a considerable seasonal snowpack.

The scale of this study was constrained to watersheds larger than 150,000 km^2, the optimal minimum geographic limit of GRACE data (Yeh et al., 2006; Landerer and Swenson, 2012).

3. Methods and Data

We used 108 months of GRACE and streamflow data over nine water years (WY; Oct – Sep; 2004–2012). This data comprises positive, neutral, and negative phases of the El Niño-Southern Oscillation and negative and positive phases of the Pacific Decadal Oscillation (Feng et al., 2014; Iizumi et al., 2014). As a result, the data provides years of above- and below-average precipitation, snowpack, and streamflow for the region. The three watersheds were delineated upstream from United States Geological Survey (USGS) stream gages at 1° resolution, which is the resolution of GRACE data. In the CRB, these grid cells represent a dimension of approximately 80 km by 120 km. The Upper Columbia consists of the area upstream of the Columbia River at the International Boundary gage (USGS 12399500), just downstream of the confluence of the Columbia and Pend-Oreille Rivers. The Pend-Oreille is a major watershed in the upper portions of the CRB. The Snake River gage at Weiser (USGS 13269000) provides gauged streamflow data above Hell's Canyon Reservoir, the largest impoundment in the Snake River basin. The USGS gage at The Dalles (USGS 14105700) provides the most downstream streamflow data for the CRB. Monthly mean runoff (*R*; mm) was calculated for each of the three gages using the USGS streamflow data.

Measurements of *TWSA* were obtained from the GRACE RL-05 (Swenson and Wahr, 2006; Landerer and Swenson, 2012) data set from NASA's Tellus website (http://grace.jpl.nasa.gov). The errors present in the gridded GRACE data exist primarily as a result of truncation (i.e., a low number of harmonics) in the spherical harmonic solution, and

E Sproles 2/20/15 11:07 AM
Eliminado: Our study area is the Columbia River Basin (CRB; 41-53ºN and 110-122ºW). This basin has wet winters, with up to 70% of annual precipitation falling between November and March, 50-60% of which occurs as snow (Serreze et al., 1999; Nolin et al., 2012). The spring months (April to June) are also wet, but warmer. Precipitation during the spring combines with snowmelt to swell rivers and potentially exacerbate flooding. Snowmelt also serves as a critical component of the hydrologic cycle recharging aquifers and filling streams later in the year. These contributions bridge the temporal disconnect between wet winters and dry summers when demand is at its peak as farmers, fish, hydropower and municipal users vie for over-allocated water resources (United States Army Corps of Engineers, 2001; Oregon Water Supply and Conservation Initiative, 2008). However, concerns with winter surplus and summer scarcity are not uniform across the CRB, since climate and geology vary greatly. Two of the study watersheds, the Upper Columbia (155,000 km^2) and the Snake River basin (182,000 km^2), represent distinctly different climatic and geologic provinces of the CRB (described and illustrated in Fig. 1). The Upper Columbia is wet and is characterized by steep topography of fractured rock and poor groundwater storage. In contrast, the arid Snake River basin is bowl-shaped with mountains on three sides. The interior of Snake River basin is a broad plain with well-developed soils and high aquifer storage. The Columbia River at The Dalles (614,000 km^2) encompasses the Upper Columbia and the Snake River sub-basins, and its climate and geology are an integration of the two (Fig. 1). A distinct climatic feature of the Columbia River at The Dalles is the western slope of the Cascade Mountains where over 3000 mm of mean annual precipitation at higher elevations sustains a considerable seasonal snowpack. The scale of this study was constrained to watersheds larger than 150,000 km^2, the optimal minimum geographic limit of GRACE data (Yeh et al., 2006; Landerer and Swenson, 2012).

E Sproles 2/20/15 11:07 AM
Eliminado: (Feng et al., 2014; Iizumi et al., 2014)

E Sproles 2/20/15 11:07 AM
Eliminado: (described below).

E Sproles 2/20/15 11:07 AM
Eliminado: Q

smoothing and systematic noise removal (called "de-striping") that is applied after GRACE level-2 processing to remove spatially correlated noise (called "stripes") (Swenson and Wahr, 2006). This smoothing tends to smear adjacent signals together (within the radius of the filtering function), resulting in smaller signals being lost, and larger signals having a coarser footprint and a loss of spatial information.

To restore the GRACE signal lost during processing, the data were scaled using 1° Land-Grid Scale Factors produced by putting a 1° land surface model through identical processing (truncation and filtering) as the GRACE solutions, then measuring the decrease in the signal amplitude at each 1° grid. These procedures are described on the Tellus website and detailed in Landerer and Swenson (2012). Monthly 1° GRACE estimates of *TWSA*, and the associated 1° leakage and measurement errors, were spatially averaged over each of the three study watersheds following the procedures described in the Tellus website.

GRACE represents monthly storage anomalies relative to an arbitrary record-length mean value, analogous to the amount of water above or below the long-term mean storage of a bucket, and should balance with the equation:

$$\Delta Storage = TWSA = \Delta GW + \Delta SM + \Delta SWE + \Delta RES \qquad (1)$$

where all components are at monthly time steps; *GW* represents groundwater, *SM* represents soil moisture (from 0–2000 mm depth), *SWE* represents snow water equivalent (the equivalent depth of water held in snowpack), and *RES* represents reservoir storage. The Δ used here represents the anomaly from the study-period mean, rather than a monthly change. To isolate monthly groundwater storage anomalies ($\Delta GW = GWSA$) in the above equation, ΔSM, ΔSWE and ΔRES estimates were subtracted from the monthly *TWSA* data using methods described in Famiglietti et

E Sproles 2/20/15 11:07 AM
Eliminado: Measurements of TWSA were obtained from the GRACE RL-05 (Swenson and Wahr, 2006; Landerer and Swenson, 2012) data set from NASA's Tellus website (http://grace.jpl.nasa.gov). The errors present in the gridded GRACE data exist primarily as a result of truncation (i.e. a low number of harmonics) in the spherical harmonic solution, and smoothing and systematic noise removal (called "de-striping") that is applied after GRACE level-2 processing to remove spatially correlated noise (called "stripes") (Swenson and Wahr, 2006). This smoothing tends to smear adjacent signals together (within the radius of the filtering function), resulting in smaller signals being lost, and larger signals having a coarser footprint and a loss of spatial information.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: or

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

al. (2011). Similarly, the combined signal of water storage anomalies of subsurface moisture ($TWSA_{sub}$), *SM* and GW, was isolated by subtracting *SWE* and *RES* from *TWSA* values.

Monthly *SM* values over the study basins were obtained from the mean of the North American and Global Land Data Assimilation Systems (NLDAS at 1/8º resolution (Cosgrove et al., 2003) and GLDAS at 1/4º resolution (Rodell et al., 2004), respectively), and were spatially averaged over the three study watersheds. Monthly 1-km resolution *SWE* values were obtained from the mean of NLDAS and Snow Data Assimilation System (SNODAS; National Operational Hydrologic Remote Sensing Center, 2004) and were spatially averaged over the three watersheds. SNODAS data were used in place of the GLDAS data product, which considerably underestimated *SWE* in mountainous areas when compared to point-based measurements. Changes in monthly reservoir storage were calculated for the five largest reservoirs in the CRB (see Table A1). Other smaller reservoirs in the CRB were excluded when it was determined that fluctuations in their levels were below the detection limits of GRACE.

Like all measurements, estimates of *TWSA* from GRACE contain error. For all of the study basins, the range of error is well below the *TWSA* signal strength, approximately an order of magnitude below the annual amplitude (200 – 300 mm) of the *TWSA* signal in the CRB. The basin-averaged *TWSA* errors (time invariant) for the three study basins are 37 mm (Upper Columbia), 22 mm (Snake), and 25 mm (The Dalles).

The model data from LDAS and SNODAS simulations are driven by *in situ* measurements, and represents the best available data for broad scales. We address any structural error from an individual model by using an ensemble of outputs. Calculation of the error in individual terms followed standard methodologies (Famiglietti et al., 2011), where error in *SM* is

E Sproles 2/20/15 11:07 AM
Eliminado: Famiglietti et al. (2011).

E Sproles 2/20/15 11:07 AM
Con formato: Sangría: Primera línea: 0.5"

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: (Cosgrove et al., 2003)

E Sproles 2/20/15 11:07 AM
Eliminado: and GLDAS at 1/4º resolution (Rodell et al., 2004), respectively), and were spatially averaged over the three study watersheds. Monthly 1-km resolution SWE values were obtained from the mean of NLDAS and Snow Data Assimilation System (SNODAS; National Operational Hydrologic Remote Sensing Center, 2004) and were spatially averaged over the three watersheds. SNODAS data were used in place of the GLDAS data product, which considerably underestimated SWE in mountainous areas when compared to point-based measurements. Changes in monthly reservoir storage were calculated for the five largest reservoirs in the CRB (see Appendix A). Other smaller reservoirs in the CRB were excluded when it was determined that fluctuations in their levels were below the detection limits of GRACE.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado:),

E Sproles 2/20/15 11:07 AM
Eliminado: plotted as bounds on

E Sproles 2/20/15 11:07 AM
Eliminado: TWSA time series in Fig. 2a-c.

E Sproles 2/20/15 11:07 AM
Eliminado: (Famiglietti et al., 2011)

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

the mean monthly standard deviation, and standard errors for *SWE* and *RES* are 15% of mean absolute changes. GWSA and $TWSA_{sub}$ anomaly errors are calculated as the sum of basin-averaged errors (added as variance) in the individual terms in each respective calculation (eq. 1), including the error in *TWSA* (Swenson et al., 2006). The basin-averaged error variance for *GWSA* (time invariant) in the three study basins are 45 mm (Upper Columbia), 26 mm (Snake), and 33 mm (The Dalles). For $TWSA_{sub}$ these values are 37 mm (Upper Columbia), 22 mm (Snake), and 25 mm (The Dalles). The individual error components (*SM*, *SWE*, *RES* respectively) for each basin are Upper Columbia (24 mm, 6 mm, 0.01 mm), Snake (14 mm, 3 mm, 0.01 mm), and The Dalles (21 mm, 4 mm, 0.01mm). Note that these error estimates are distributed across an entire regional watershed and do not represent the error at individual monitoring sites. A time series of these values and basin-averaged errors is provided in Fig. 2.

Based on an approach similar to Reager et al. (2014) and Riegger and Tourian (2014), we plotted the temporal relationship between *TWSA* and *R* to examine hysteresis relationships in all three of the study watersheds for each individual water year and for the monthly mean across all water years. Expanding from the integrated terrestrial component of water storage, we also plotted the relationships of $TWSA_{sub}$ and *GWSA* with *R*. We examined the branches of these hysteresis plots to better understand how the size, shape, and direction of the hystereses varied across years in each of the three regional watersheds.

In order to verify groundwater hysteresis, we compared the GRACE-derived *GWSA* to groundwater depths from well measurements at 33 sites throughout the study region (Fig. 1 and Table A2). These data were normalized by their standard deviation, and the mean of the 33 wells was calculated. The standard deviation of the GRACE-derived *GWSA* for The Dalles was normalized to provide a direct comparison of *GWSA* and *in situ* measurements.

E Sproles 2/20/15 11:07 AM
Con formato ... [2]
E Sproles 2/20/15 11:07 AM
Con formato ... [3]
E Sproles 2/20/15 11:07 AM
Eliminado: Groundwater
E Sproles 2/20/15 11:07 AM
Eliminado: error is the
E Sproles 2/20/15 11:07 AM
Eliminado: the
E Sproles 2/20/15 11:07 AM
Eliminado: of ΔGW
E Sproles 2/20/15 11:07 AM
Eliminado: (Swenson et al., 2006)
E Sproles 2/20/15 11:07 AM
Con formato ... [4]
E Sproles 2/20/15 11:07 AM
Con formato ... [5]
E Sproles 2/20/15 11:07 AM
Eliminado: are plotted as bounds on the GV ... [6]
E Sproles 2/20/15 11:07 AM
Con formato ... [7]
E Sproles 2/20/15 11:07 AM
Con formato ... [8]
E Sproles 2/20/15 11:07 AM
Con formato ... [9]
E Sproles 2/20/15 11:07 AM
Eliminado: (2014)
E Sproles 2/20/15 11:07 AM
Eliminado: (2014)
E Sproles 2/20/15 11:07 AM
Con formato ... [10]
E Sproles 2/20/15 11:07 AM
Eliminado: Q
E Sproles 2/20/15 11:07 AM
Con formato ... [11]
E Sproles 2/20/15 11:07 AM
Eliminado: relationship between
E Sproles 2/20/15 11:07 AM
Con formato ... [12]
E Sproles 2/20/15 11:07 AM
Eliminado: and Q
E Sproles 2/20/15 11:07 AM
Eliminado: measure basin groundwater re ... [13]
E Sproles 2/20/15 11:07 AM
Con formato ... [14]
E Sproles 2/20/15 11:07 AM
Eliminado: Appendix A
E Sproles 2/20/15 11:07 AM
Con formato ... [15]
E Sproles 2/20/15 11:07 AM
Con formato ... [16]

We further hypothesized that because peak *SWE* accumulation occurs between February and April, that *TWSA* for these months could be used to predict *R* for an individual month and the cumulative seasonal runoff (R_{season}) that occurs after peak *SWE* accumulation. To test this prognostic hypothesis we used a two-parameter power function (The MathWorks, 2013):

$$R_{predicted} = a(TWSA_{month})^b + c \qquad (2)$$

where $R_{predicted}$ is runoff for the predicted time interval; $TWSA_{month}$ represents terrestrial water storage for an individual month, and a, b, and c are fittted parameters from the power function.

We tested this relationship for *TWSA* in February, March and April to predict R_{season} (April – September) and for the individual months of July (R_{July}), August (R_{Aug}), and September (R_{Sep}); these represent the lower-flow months when demand is near its peak. Additionally, we tested and compared the modeled-values of *SWE* from NLDAS and SNODAS and *SM* from NLDAS and GLDAS, and the model-derived values of $TWSA_{sub}$ to predict R_{season} and for the individual months using the same power-function analysis.

Because our data set was constrained to nine water years, we used a double-pass approach to fit and test the empirical relationship between *S*—*R*. This approach allowed us double our data inputs for calculating standard hydrologic evaluation metrics such as Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE) and Coefficient of Determination (R^2); (Legates and McCabe, 1999). The nine years were divided into two sets (Set 1, even years 2004-2012; Set 2, odd years 2005-2011). The first pass calculated the power function of *S*—*R* to Set 1, and the parameters were then tested against Set 2. The roles of the datasets were then reversed, and the empirical model results of each pass were compiled into one data set and tested against measured values to calculate RMSE, NSE, and R^2. In order to maximize the limited data inputs,

E Sproles 2/20/15 11:07 AM
Eliminado: spring

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: Q later in the year

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: months and

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: (April – September).

E Sproles 2/20/15 11:07 AM
Eliminado: (The MathWorks, 2013) to evaluate

E Sproles 2/20/15 11:07 AM
Eliminado: ability of

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: cumulative Q from

E Sproles 2/20/15 11:07 AM
Eliminado: (Q_{season}

E Sproles 2/20/15 11:07 AM
Eliminado: Q_{Aug}), the low

E Sproles 2/20/15 11:07 AM
Eliminado: month

E Sproles 2/20/15 11:07 AM
Eliminado: and soil moisture

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: ,

E Sproles 2/20/15 11:07 AM
Eliminado: SNODAS

E Sproles 2/20/15 11:07 AM
Eliminado: Q_{season}

E Sproles 2/20/15 11:07 AM
Eliminado: Q_{Aug}

E Sproles 2/20/15 11:07 AM
Con formato: No revisar la ortografía ni la gramática, Sin Superíndice / Subíndice

once we tested the two independent sets for model performance, we combined the data sets for a single power function curve. The observed data were tested against the simulated data from the complete, but limited data record. The final model curve was fit to these data.

4. Results

4.1. Storage-runoff hysteresis

The filling and emptying of the study basins at the regional-scale over the course of an individual WY results in a hysteretic relationship between storage and runoff (Fig. 3a). The hysteresis loops begin at the onset of the wet season in October, with *TWSA* increasing (Figs. 3a, 4a-c) as precipitation is stored as snow and soil moisture. An increase in storage that is not offset by an increase in discharge indicates a predominance of snow inputs and the freezing of soil water. The lower branch of the hysteresis plot (storage increase unmatched by runoff) can be used to estimate cumulative snow water equivalent and soil moisture in the basin. This is the water that later contributes to streamflow and groundwater recharge in the spring.

The hysteresis shifts direction from Feb-Apr (inflection 1, Fig. 3a) when saturated soils and snowmelt cause *R* to rapidly increase. Each hysteresis loop contains a vertical branch of the curve during which storage is relatively constant, but streamflow increases rapidly. This also represents the groundwater recharge branch of the loop. As snow melts and the ground thaws, runoff is generated, recharge into soils occurs, and basins tend to be at peak storage during this branch. Storage losses and additional precipitation inputs during this period are re-organized internally. A second shift (inflection 2, Fig. 3a) occurs from Apr-June when peak *TWSA* begins

E Sproles 2/20/15 11:07 AM
Eliminado: discharge

E Sproles 2/20/15 11:07 AM
Eliminado: plots

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: streamflow

E Sproles 2/20/15 11:07 AM
Eliminado: snow melt

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

to decrease, representing spring snowmelt and a switch from precipitation that falls primarily as snow to rain; these combine to contribute to peak *R*.

Once peak *R* values are reached, the loop shifts direction a third time (inflection 3, Fig. 3a), receding on both axes as contributions from snowmelt diminish while presumably groundwater sustains streams and provides a source for irrigated agriculture. During this period, the relationship between *TWSA* and discharge is more linear, corresponding to baseflow-driven runoff processes in which each monthly change in storage causes a proportional change in the generation of streamflow.

The hysteresis plots of *TWSA*—*R* for an individual water year demonstrate that the timing and quantity of precipitation governs the size of a hysteresis loop for an individual WY (Figs. 3a, 4a-c, 5). For instance wet years (e.g., 2008) have bigger loops, while dry years (e.g., 2005) are more compressed along both axes. However, the general shape of the loops is distinct for each basin. Plotting multiple WYs provides a family of curves for each basin that helps describe how climate, topography, and geology govern the timing and magnitude of the relationship between *TWSA* and *R* (Figs. 3a, 5).

4.2. Subsurface water ($TWSA_{sub}$) – runoff hysteresis

The $TWSA_{sub}$ hysteresis curve contracts horizontally when the snow signal is removed from *TWSA* values for both the Upper Columbia and The Dalles (Figs. 3b, 4d-f), which collapses the loops and takes a form similar to a plot-scale hysteresis of soil. Peak $TWSA_{sub}$ occurs in June, which corresponds to the spring melt of mountain snowpack and the end of the wet season (Figs. 4d-f). However in the Snake River, the hysteresis curve still retains a loop, but the timing of maximum $TWSA_{sub}$ is also earlier, reaching its peak during March and April (Fig. 4e). It is

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: -Q

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: the

E Sproles 2/20/15 11:07 AM
Eliminado: governs

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: –c

E Sproles 2/20/15 11:07 AM
Eliminado: Individual basin

E Sproles 2/20/15 11:07 AM
Eliminado: plots

E Sproles 2/20/15 11:07 AM
Eliminado: Of the three study basins, the Upper Columbia is the most hydrologically responsive, with a mean range of 210 mm for TWSA and 50 mm for Q. The Upper Columbia's steep topography and wet climate fills the limited aquifer quickly, reaching a storage threshold and generating runoff. The steep topography moves snowmelt and rain quickly through the terrestrial system and into the river channel until cresting in June (Figs. 4 and 5), followed by declines in TWSA and Q from June-September. ... [17]

E Sproles 2/20/15 11:07 AM
Bajado [1]: . The slope of the recession from June-Sept has the same general shape for The Dalles as the Upper Columbia (Figs. 4a, 4c), presumably from snowmelt-generated runoff.

noteworthy that in the Snake River the $TWSA_{sub}$—R hysteresis loop temporally progresses in the opposite direction, but stays in phase with precipitation inputs.

4.3. Groundwater-runoff hysteresis

The hysteresis loops describing the temporal relationship between *GWSA* and *R* are equally informative, with one distinct difference—they temporally progress in opposite directions of the hysteresis loops of *TWSA* and *R* (Fig. 3). For all three watersheds, *GWSA* decreases from Oct–Feb/Mar (Fig. 4h-j), which is out of phase with the onset of the wet season. *GWSA* does not shift towards positive gains until early spring and the initial stages of melt before reaching its maximum in June.

The 33 point-specific well data located across the CRB show considerable individual variability throughout a water year, and the mean of the normalized standard deviations of well levels was close to zero for all months. The temporal variability for the well data provides no discernable correlation with the derived *GWSA* signal (Fig. A1).

4.4. Individual basin hysteresis plots of *TWSA*, $TWSA_{sub}$, *GWSA* and *R*

Of the three study basins, the Upper Columbia is the most hydrologically active, showing the largest annual range for *TWSA*, $TWSA_{sub}$, *GWSA*, and *R* (Fig. 6). The *TWSA*—*R* hysteresis loops are more open (Fig. 4), corresponding to the fluxes of water moving through watershed. When *SWE* is removed and subsurface water is highlighted, the $TWSA_{sub}$—R hysteresis loops collapse horizontally and more closely resemble the hystereses associated with soil (Figs. 4d). However the inter-annual range (WY_{max} – WY_{min}) for $TWSA_{sub}$ in the Upper Columbia is considerably greater than the other two basins (median range = 234 mm; Fig. 6). As the

E Sproles 2/20/15 11:07 AM
Eliminado: 3.
E Sproles 2/20/15 11:07 AM
Eliminado: discharge
E Sproles 2/20/15 11:07 AM
Eliminado: plots
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: Q
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: dramatic
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: Q
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: decreased
E Sproles 2/20/15 11:07 AM
Eliminado: 4d-f), and
E Sproles 2/20/15 11:07 AM
Eliminado: mean of the
E Sproles 2/20/15 11:07 AM
Bajado [2]: 6).
E Sproles 2/20/15 11:07 AM
Movido (inserción) [2]

hysteresis reverses directions for *GWSA*—*R*, the loops shift to a more open shape (Figs. 4d), but the inter-annual range remains similar.

In contrast to the rapid response of the Upper Columbia, the Snake River receives ~60% less annual precipitation, but has an annual *TWSA* range that is only 22% less (median annual range = 192 mm; *R*=7 mm; Figs. 4, 5, and 6). However, the *TWSA* hysteresis loops for the Snake River are collapsed vertically (Fig. 4b). In the more arid Snake River, removing the snow signal does not collapse the $TWSA_{sub}$—*R* hysteresis loops ($TWSA_{sub}$ = 89 mm). Similarly, the *GWSA* loops suggest that subsurface moisture plays a more prominent role in the Snake River.

The climate, topography, and geology of the Columbia River at The Dalles are an integration of the Upper Columbia and Snake River, seen in the shape of the hysteresis loops (Figs. 4, 5, 6; median annual range *TWSA*=195 mm; *R*=27 mm). The period from Feb–June more closely resembles the Snake River basin, with gradual increases in *TWSA* and sharp increases in *R*. The slope of the recession from June-Sept has the same general shape for The Dalles as the Upper Columbia (Figs. 4a, 4c), presumably from snowmelt-generated runoff.

4.5. Streamflow forecasting

We next present how *TWSA* was applied prognostically to predict streamflow. Using the double-pass calibration and validation approach, $TWSA_{Mar}$ provided the best overall predictive capabilities for R_{season} with a mean NSE ($\overline{NSE}$) and mean R^2 ($\overline{R}^2$) of 0.75 and 0.91, respectfully (Fig. 7a, Table 1), for all three basins. The Dalles had the highest NSE and R^2, and lowest RMSE values (0.98, 0.98, 6 mm; Table 1). The results in the Upper Columbia were also robust (0.82, 0.86, 33 mm; Table 1), while the Snake River performed with less skill (0.46, 0.59, and 14 mm, Table 1). Applying $TWSA_{April}$ also provided similar results, but with a lower degree of skill in

E Sproles 2/20/15 11:07 AM
Movido (inserción) [1]

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: can predict the total Q from April through September (Q_{season}) in all three basins

E Sproles 2/20/15 11:07 AM
Eliminado: an R^2 range of 0.83-0.98 and

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: $\overline{R}$

E Sproles 2/20/15 11:07 AM
Eliminado:).

E Sproles 2/20/15 11:07 AM
Eliminado: TWSA for April

predicting *R* (NSE = 0.57, $\overline{R}^2$= 0.69). *TWSA*$_{Apr}$ provided improved predicted capabilities in the Upper Columbia (0.87, 0.88, and 28 mm, Table 1), but inferior results in the other two watersheds. *TWSA*$_{Feb}$ had a low degree of skill in predicting *R* in all three watersheds (Table A3).

TWSA$_{Mar}$ and *TWSA*$_{April}$ also served as a good predictor of monthly runoff in July and August for the Upper Columbia and to a lesser degree in The Dalles (Tables 1 and A3). In the Snake River, *TWSA* did not serve as a good predictor for *R* in an individual month.

Snowpack and soil moisture play a considerable role in the hydrology of the CRB and are commonly used to help predict water demand and availability later in the year (Koster et al., 2010). We compared the capabilities of the modeled snow (*SWE*) and soil moisture (*SM*) products to predict *R* to the skill of measured GRACE *TWSA* data (Table 1). In the Upper Columbia and The Dalles, *TWSA*$_{Mar}$ predicts seasonal and monthly runoff (July and August) with considerably more skill than *SWE* or *SM* (Figure 7, Table 1). In the Snake River, *SM*$_{Mar}$ has a higher degree of skill than *TWSA*$_{Mar}$ in predicting *R*$_{season}$ and *R*$_{Aug}$. *SWE*$_{Mar}$ provided inferior results in all three watersheds, but with some predictive skill in the Upper Columbia and The Dalles (NSE of 0.24 and 0.46 respectively, Table 1). In all three watersheds, *TWSA*$_{sub}$ provided extremely poor predictions (Tables 1 and A3).

When the results of the empirical model using two independent sets of data proved robust for some of the storage metrics, the observed data were tested against the simulated data from the complete, but limited data record. The performance of the empirical model improved using the complete data set (Tables 2 and A4), with the same general results. *TWSA*$_{Mar}$ provided the best model fit for seasonal runoff in the Upper Columbia (NSE = 0.93, RMSE = 19.8 mm) and The Dalles (NSE = 0.98, RMSE = 5.7 mm). In the Snake River, predictive capabilities improved

E Sproles 2/20/15 11:07 AM
Eliminado: Q (range = 0.75-0.92, $\overline{R}^2$ = 0.86

E Sproles 2/20/15 11:07 AM
Eliminado: indicator

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: runoff (Q_{Aug}, range = 0.68-0.88, $\overline{R}^2$ = 0.77). While this overall mean is high,

E Sproles 2/20/15 11:07 AM
Eliminado: range of agreement between b ... [18]

E Sproles 2/20/15 11:07 AM
Eliminado: River at

E Sproles 2/20/15 11:07 AM
Eliminado: R^2 = 0.88).

E Sproles 2/20/15 11:07 AM
Eliminado: (Koster et al., 2010)

E Sproles 2/20/15 11:07 AM
Eliminado: predictive

E Sproles 2/20/15 11:07 AM
Con formato ... [19]

E Sproles 2/20/15 11:07 AM
Con formato ... [20]

E Sproles 2/20/15 11:07 AM
Eliminado: the

E Sproles 2/20/15 11:07 AM
Con formato ... [21]

E Sproles 2/20/15 11:07 AM
Eliminado: Compared to SWE

E Sproles 2/20/15 11:07 AM
Eliminado: provided a better indicator of

E Sproles 2/20/15 11:07 AM
Eliminado: August

E Sproles 2/20/15 11:07 AM
Con formato ... [22]

E Sproles 2/20/15 11:07 AM
Eliminado: in the Upper Columbia and at ... [23]

E Sproles 2/20/15 11:07 AM
Eliminado: 6

E Sproles 2/20/15 11:07 AM
Eliminado: SM provided

E Sproles 2/20/15 11:07 AM
Eliminado: slightly

E Sproles 2/20/15 11:07 AM
Con formato ... [24]

E Sproles 2/20/15 11:07 AM
Eliminado: Q_{season} and Q_{Aug}. However

E Sproles 2/20/15 11:07 AM
Eliminado: , SM provided inferior

more dramatically (NSE = 0.83, RMSE = 7.4 mm), but soil moisture still served as a better predictor of seasonal streamflow (NSE = 0.93, RMSE = 5.2 mm). Similarly, $TWSA_{Mar}$ provided the best model fit for runoff in August, one of the drier months when demand is at its peak (Tables 2 and A4).

5. Discussion

5.1. Storage-runoff hysteresis

Decades of data collection and monitoring at individual gage sites indicate that watersheds collect, store and release water. Using one integrated measurement from the GRACE satellites, our results show these same process at the regional scale in the hysteresis loops of storage (*TWSA*) and runoff (*R*). While hystereic processes have previously been identified in local-scale measurements (McDonnell, 2003; McGlynn and McDonnell, 2003), only recently has streamflow-storage hysteresis been identified at the regional scale (Riegger and Tourian, 2014).

Our work builds on Riegger and Tourian's (2014) results, and employs GRACE data to describe how regional watersheds function as integrated, non-linear systems governed by climate, topography, and geology. Climate controls the size of the hysteresis loops by providing a first-order control on hydrologic inputs and the storage of solid water, which in turn governs the ranges of *TWSA* and *R*. However, runoff response to precipitation and snowmelt does not act independently from topography and geology (Jefferson et al., 2008; Tague et al., 2008), which controls how liquid water is stored and routed through a watershed, even at the regional scale. The climatic, topographic, and geological characteristics of each watershed provide an explanation of the *S*—*R* relationship that helps govern the shape and size of its respective hysteresis curve. GRACE offers a single, integrated measurement of changes in water storage

E Sproles 2/20/15 11:07 AM
Eliminado: skill for Q_{season} and Q_{Aug} as compared to TWSA.

E Sproles 2/20/15 11:07 AM
Eliminado: .

E Sproles 2/20/15 11:07 AM
Con formato: Sangría: Primera línea: 0.25"

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Eliminado: (McDonnell, 2003; McGlynn and McDonnell, 2003)

E Sproles 2/20/15 11:07 AM
Eliminado: measured

E Sproles 2/20/15 11:07 AM
Eliminado: (Riegger and Tourian, 2014).

through and across a watershed that can be applied to predict regional streamflow using an empirical model. Where these predictive capabilities succeed and fail help better describe the climatic, topographic, and geological characteristics in each watershed.

For example, in the Upper Columbia, steep topography and wet climate fills subsurface storage quickly before reaching a threshold in April or May. After this watershed-scale threshold is reached, the steep topography moves snowmelt and rain quickly through the terrestrial system and into the river channel until cresting in June (Figs. 4, 5, and 6), followed by declines in *TWSA* and *R* from June-September. These large fluxes of water create a more open hysteresis loop, expanding non-linearly on both the horizontal and vertical axes.

The Upper Columbia also has the broadest range of annual $TWSA_{sub}$ and *GWSA* during the study period (Figs. 5 and 6), despite having limited aquifer capacity. Conceptually, this demonstrates that the upper limit of storage is greater than in the Snake River or The Dalles, but that it also loses the most water. Its minimums at the end of the WY are also the lowest (median $TWSA_{Sep}$ = -98mm; Figs. 5 and 6). This range across *TWSA*, $TWSA_{sub}$, and *GWSA* supports the conceptual model that the watershed fills during the wet season, and is then drained more quickly due to steep topography and limited water storage. The predictive capability of *TWSA* also strongly suggests that the components and temporal relationships of storage across this watershed are interconnected, and that incorporating April snowpack improves the model results.

In contrast, the arid Snake River basin provides a very different family of hysteresis curves (Figs. 4, 5) that identify groundwater and soil moisture as primary components of watershed function. The curves are compressed vertically (*R*) as compared to the Upper Columbia, and are more constrained horizontally (Fig. 6). The onset of spring melt runoff in

E Sproles 2/20/15 11:07 AM
Eliminado: and geology. Climate controls the size of the hysteresis loops by providing a first-order control on hydrologic inputs and the storage of solid water, which in turn governs the ranges of TWSA and Q. However, runoff response to precipitation and snowmelt does not act independently from geology (Jefferson et al., 2008; Tague et al., 2008), which controls how liquid water is stored and routed through a watershed, even at the regional scale. This in turn helps govern the shape of a watershed's hysteresis curve.

E Sproles 2/20/15 11:07 AM
Eliminado: basin steep slopes and fractured bedrock geology

E Sproles 2/20/15 11:07 AM
Eliminado: cause

E Sproles 2/20/15 11:07 AM
Eliminado: precipitation to run off

E Sproles 2/20/15 11:07 AM
Eliminado: ,

E Sproles 2/20/15 11:07 AM
Eliminado: a relatively small amount is retained in the soil. This phenomena is seen in the

E Sproles 2/20/15 11:07 AM
Eliminado: shape of the

E Sproles 2/20/15 11:07 AM
Eliminado: curve in

E Sproles 2/20/15 11:07 AM
Con formato: Sangría: Primera línea: 0.25"

E Sproles 2/20/15 11:07 AM
Eliminado: – once maximum TWSA is reached

E Sproles 2/20/15 11:07 AM
Eliminado: March,

E Sproles 2/20/15 11:07 AM
Eliminado: quickly transitions to runoff (Fig 4a). These data suggest that this is a watershed where snow storage is the primary component of water storage, which is also reflected in the comparison of SWE and SM data (not shown).

E Sproles 2/20/15 11:07 AM
Eliminado: and groundwater dominated

E Sproles 2/20/15 11:07 AM
Eliminado: results provide

E Sproles 2/20/15 11:07 AM
Eliminado: 4, 5) that

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Eliminado: basin, despite showing a similar intra-annual range

E Sproles 2/20/15 11:07 AM
Eliminado: TWSA). Another distinction is that the

February does not deplete *TWSA* for the Snake River. Instead, *TWSA* continues to increase until May, when peak runoff occurs. As *TWSA* decreases to the end of the water year in September, the median $TWSA_{Sep}$ measurement (-78 mm) is 20 mm greater than in the Upper Columbia. This indicates that the lower drainage threshold of the Snake River watershed is relatively greater than the Upper Columbia, potentially explained by a less severe topography and higher aquifer capacity.

The $TWSA_{sub}$ hysteresis curves in the Snake River retain a similar shape to the *TWSA* signal. While they reverse direction they do stay temporally connected to the onset of the wet season in October, indicating that subsurface moisture is a central control on the filling of the watershed through May. The capabilities of *SM* to empirically predict *R* better than *TWSA* further highlight the importance of subsurface water in this watershed. The intra-annual range of *GWSA* in the Snake River is also more limited than in the more hydrologically responsive Upper Columbia. This more limited range of data supports the conceptual model of a watershed that retains comparatively more winter precipitation in soils and aquifers throughout the spring season, and that sustains flow later in the year and until the onset of melt the following winter.

The greater Columbia River Basin upstream from The Dalles integrates the climatic, topographic, and geologic characteristics of the Snake River and Upper Columbia as well as other areas within the CRB. The western slope of the Cascades (Fig. 1), which is outside of the Upper Columbia, accumulates up to several meters of *SWE* each winter. Due east of the Cascades, an expansive basalt plain that provides aquifer storage helps dampen the snowmelt pulse in the spring. The hysteresis loops for The Dalles reflect these combined characteristics.

E Sproles 2/20/15 11:07 AM
Eliminado: at

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: These data support

E Sproles 2/20/15 11:07 AM
Eliminado: drains to sustain

E Sproles 2/20/15 11:07 AM
Eliminado: ,

E Sproles 2/20/15 11:07 AM
Eliminado: The region due

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: is characterized by a

E Sproles 2/20/15 11:07 AM
Eliminado: underlain by basalt

E Sproles 2/20/15 11:07 AM
Eliminado: excellent

E Sproles 2/20/15 11:07 AM
Eliminado: and

E Sproles 2/20/15 11:07 AM
Eliminado: the CRB

Storage at The Dalles increases along the horizontal axis (*TWSA*) until peak storage is reached in March or April (Figs 3, 4, and 5). This *TWSA* threshold responds with an increase in *R* that continues through June. In July, the hysteresis begins to recede along both axes closing out the loop. The *GWSA* has the most limited range, potentially explained by the extensive basalt aquifer moderating the relationship between storage and runoff. In The Dalles, $TWSA_{Sep}$ has a median value of -88mm (Fig. 6), between the lower drainage thresholds of the Upper Columbia and Snake River watersheds; indicating an integration of the contributing climate, topography, and geology.

E Sproles 2/20/15 11:07 AM
Eliminado: , where the onset of melt in February produces a pronounced

5.2. Distinguishing the difference between $TWSA_{sub}$ and *GWSA*

Conceptually $TWSA_{sub}$ represents changes in the amount of water stored as soil moisture and groundwater, where as *GWSA* represents water changes greater than 2000mm below the soil surface. The goals of evaluating these metrics were to see if monthly changes in soil moisture were linked to changes in groundwater storage, and the role of snowpack in the *S—R* relationship.

E Sproles 2/20/15 11:07 AM
Eliminado: runoff similar to the Upper Columbia. However,

E Sproles 2/20/15 11:07 AM
Eliminado: with the Snake River, increases in measured TWSA are found in the CRB through April.

The $TWSA_{sub}$ hysteresis curves in the Upper Columbia and The Dalles collapse into a more linear relationship that is more commonly associated with the *S—R* relationship of a soil matrix (Fig. 3 and 4). This is in contrast to the *GWSA* hystereses that are represented by loops that show an out-of-phase relationship between precipitation and groundwater recharge from the start of the wet season in October until February or March. The $TWSA_{sub}$ and *GWSA* hysteresis plots demonstrate that in these two basins changes in monthly soil moisture are not always temporally aligned with *GWSA*. This can be explained by the physical reality that soil moisture and groundwater are not always interconnected, and that there is not a fixed depth (i.e., 2000 mm) that separates the two components of water storage.

GRACE-derived calculations of *GWSA* also provide insights into the hydrological processes governing groundwater recharge and depletion, as evidenced in the *GWSA* hysteresis loops. The *GWSA—R* curves show an out-of-phase relationship between precipitation and groundwater recharge from the start of the wet season in October until February or March. This indicates that groundwater helps sustain stream flow during the wet fall and winter and that pore space in soils and geologic materials must fill to a certain threshold before groundwater begins to recharge and runoff is generated. The relationship between the *TWSA* and *GWSA* curves from Oct-Mar identifies how the onset of snowmelt also marks the beginning of groundwater recharge, and suggests that snowmelt inputs to groundwater are considerable. In the CRB this is critical as current climate trends are projected to reduce snowpack accumulation and exacerbate melt in the region (Wu et al., 2012; Rupp et al., 2013; Sproles et al., 2013).

Additionally, our analysis identifies summer as the time of peak groundwater storage in all three regional watersheds. This finding is of value for groundwater management and policy decisions, as peak groundwater levels in June correspond to the timing of groundwater pump tests that are used to develop groundwater withdrawal regulations (Jarvis, 2011, 2014). Our data suggest that groundwater pump tests should not be limited to an individual month, and should also include periods of reduced storage particularly during the winter months. The inclusion of multiple pump tests throughout the year could be particularly relevant as the population and water demand is projected to increase in the region.

E Sproles 2/20/15 11:07 AM
Movido (inserción) [3]

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Movido (inserción) [4]

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Movido (inserción) [5]

E Sproles 2/20/15 11:07 AM
Movido (inserción) [6]

The point-specific well data are not conclusive and show considerable variability with no consistent pattern regarding the timing of recharge and peak groundwater levels. This is presumably a function of how site characteristics (i.e., usage, depth, location, elevation) are extremely variable across a region. Rather than excluding these results or selecting individual wells that match GRACE data, we discuss the results from all 33 wells to help demonstrate the high variability that exists from well to well, and that measurements of groundwater changes at a fixed location does not represent watershed-scale characteristics (Jarvis, 2011, 2014). The disconnect between sites also highlights the concept brought forward by Spence (2010), that storage is not uniform across a watershed, and functions as a series of discontinuous processes at the watershed scale.

5.3. Applying the *S—R* relationship as a predictive tool

We applied these climatic, topographic, and geologic insights to develop and test the hypothesis that spring *TWSA* could predict *R* later in the year, based on two observations: First, the shapes of the hysteresis curves for each basin are similar (Figs. 4a-c, 5), but vary by magnitude of annual *TWSA*. Second, peak *TWSA* occurs before the peak runoff. We show that the integrated GRACE signal is a good baseline measurement to empirically predict seasonal streamflow across a range of water years with regards to precipitation and streamflow. In essence, our data suggest that the water stored across and through the Columbia River Basin in March describes the water available for the remainder of the water year.

In the CRB and in the northwestern United States, peak snowpack occurs in March or April, and is commonly used as a metric for predicting spring runoff. Despite the importance of snowpack to the hydrologic cycle of the region, measurements of $TWSA_{Mar}$ from GRACE

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: Q
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática
E Sproles 2/20/15 11:07 AM
Eliminado: in predicting
E Sproles 2/20/15 11:07 AM
Eliminado: its
E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

provide a better prediction of R_{season}, R_{July}, and R_{Aug} than model-derived estimates of snowpack. GRACE $TWSA_{Mar}$ also provided a better prediction for runoff than soil moisture, except for the Snake River watershed. There March soil moisture provided a better indicator of runoff for the rest of the year. $TWSA_{Feb}$ provided inferior predictive capacity, as the annual maximum *TWSA* values have not been reached.

These results are promising with regards to using GRACE as a predictive tool for water resources in both wet and dry years. Our limited data record represents a wide-range of conditions with regards to climate and streamflow, which is captured in our empirical models and is shown in the box plots to the right of Figs. 7a - b. These same results also indicate that *R* is insensitive to $TWSA_{Mar}$ values below 100 mm. This lower threshold describes with some certainty the amount of runoff that will be available for operations for the remainder of the year.

We recognize that all three of these regional watersheds are managed through a series of dams and reservoirs that create an altered runoff signal. Water resources managers use point-specific and model-based estimates of water storage in the region to optimize their operations for the water year. Additionally, in the fertile plains of the Snake River and lower CRB, broad-scale agriculture relies on both ground- and surface water for irrigation. Water withdrawals would be implicit in the *TWSA* signal and reduce *R*. However, a more detailed analysis of withdrawals lies outside the scope of this study.

Regardless of the length of record or anthropogenic influence, climate, topography, and geology still provide the first-order controls on water storage that are found in the hysteresis loops. GRACE encapsulates these hydrologic processes through measurements of *TWSA*. The hysteresis loops expand and contract accordingly during wet and dry years, as the intra-annual

E Sproles 2/20/15 11:07 AM
Eliminado: seasonal

E Sproles 2/20/15 11:07 AM
Eliminado: August runoff

E Sproles 2/20/15 11:07 AM
Con formato: No revisar la ortografía ni la gramática, Subíndice

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: basin

E Sproles 2/20/15 11:07 AM
Eliminado: , although $TWSA_{Mar}$ had a similar accuracy (Table 1). Despite a relatively short

E Sproles 2/20/15 11:07 AM
Eliminado: , the years of our study represent

E Sproles 2/20/15 11:07 AM
Eliminado: 100mm.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: 100mm.

E Sproles 2/20/15 11:07 AM
Eliminado: complex

E Sproles 2/20/15 11:07 AM
Eliminado: .

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: Q. Subsequent,

E Sproles 2/20/15 11:07 AM
Eliminado: would help quantify these effects. However the compilation of a complete irrigation dataset

E Sproles 2/20/15 11:07 AM
Eliminado: We also acknowledge the length of record for this study would ideally be longer.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

relationship between *TWSA* and Q represents the fluxes of water into and out of the watershed. Despite intra-annual differences, a family of hysteresis curves can describe each of the sub-regional watersheds. The predicative capability using *TWSA*, the vertical sum of water, as compared to snowpack and soil moisture further highlights the integrated nature of water storage in regional hydrology. These predictive capabilities highlights the potential of GRACE to improve upon seasonal forecast predictions and regional hydrological models.

5.4. GRACE as an analysis tool for regional watersheds

Where previous approaches to modeling watershed behavior have focused on separate storage compartments, new approaches should include the magnitude and direction of hysteresis (Spence, 2010). This integrated approach would provide new ways forward to classify watersheds not only by runoff, but also on the first-order controls that govern the non-linear hydrological processes.

Even though GRACE is somewhat of a blunt instrument with regards to temporal (monthly) and spatial (1°) resolution, this emerging technology provides a new dimension to regional watershed analysis by providing an integrated measurement of water stored across and through the Earth. These measurements continue to prove their value in retrospective analysis of regional hydrology (Rodell et al., 2009; Castle et al., 2014). However, the hysteresis loops presented by Riegger and Tourian (2014) and further developed in this paper demonstrate the ability of GRACE data to help develop a process-based understanding of how regional watersheds function as simple, dynamic systems. As the temporal record of GRACE continues to increase, its value as both a diagnostic and predictive tool will continue to grow. In the mean time, these data have value in augmenting existing management strategies.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: can be described by a family of hysteresis curves.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: These integrated measurements of TWSA provide simple, but informative

E Sproles 2/20/15 11:07 AM
Eliminado: of seasonal

E Sproles 2/20/15 11:07 AM
Eliminado: monthly runoff

E Sproles 2/20/15 11:07 AM
Subido [3]: GRACE-derived calculations of *GWSA* also provide insights into the hydrological processes governing groundwater recharge and depletion, as evidenced in the *GWSA* hysteresis loops.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato ... [25]

E Sproles 2/20/15 11:07 AM
Eliminado: The GWSA-Q curves show an ... [26]

E Sproles 2/20/15 11:07 AM
Subido [4]: that groundwater helps sustai ... [27]

E Sproles 2/20/15 11:07 AM
Con formato ... [28]

E Sproles 2/20/15 11:07 AM
Con formato ... [29]

E Sproles 2/20/15 11:07 AM
Eliminado: (Wu et al., 2012; Rupp et al., 2 ... [30]

E Sproles 2/20/15 11:07 AM
Subido [5]: Additionally, our analysis ide ... [31]

E Sproles 2/20/15 11:07 AM
Eliminado: This peak corresponds to the t ... [32]

E Sproles 2/20/15 11:07 AM
Subido [6]: Our data suggest that ground ... [33]

E Sproles 2/20/15 11:07 AM
Eliminado: The point-specific well data (F ... [34]

E Sproles 2/20/15 11:07 AM
Eliminado: . However

E Sproles 2/20/15 11:07 AM
Eliminado: (Rodell et al., 2009; Castle et a ... [35]

E Sproles 2/20/15 11:07 AM
Eliminado: here and in

E Sproles 2/20/15 11:07 AM
Eliminado: develop

Perhaps one of the most important facets of GRACE data is that it does not distinguish political boundaries. It is not linked to a specific *in situ* monitoring agency with limited data access and has the capacity to bridge sparse and inconsistent on-the-ground hydrologic monitoring networks that exist in many regions of the world. Previous GRACE-based analysis has shown its value in highlighting negative trends in terrestrial water storage in trans-boundary watersheds (Voss et al., 2013; Castle et al., 2014), and resulting regional conflict exacerbated by water shortages (Gleick, 2014b). GRACE provides an objective measurement of a region's water resources that can provide valuable insights into potential shortages or surpluses of water resources, and simple empirical predictions of seasonal and monthly runoff that are easily deployable in places with limited data.

6. Conclusions

We have shown how GRACE-based measurements of *TWSA* distill the complexity of regional hydrology into a simple, dynamic system. *TWSA* and derived estimates of *GWSA* reveal hysteretic behavior for regional watersheds, which is more commonly associated with hydrologic measurements at local scales. While the magnitude of the hysteresis curves vary across years, they retain the same general shape that is unique to each watershed. We demonstrated the utility of these hysteresis curves by showing how the complete *TWSA* record during March and April can be used to empirically predict *R* for the remainder for the water year ($TWSA_{Mar}$, mean NSE = 0.91) and during the drier summer months ($TWSA_{Mar}$, mean NSE for July = 0.76, August = 0.72; Tables 1 and 2).

Because GRACE *TWSA* can augment prediction, managers could start to interpret each year's hysteresis curve for the upcoming spring and summer, providing greater clarity and

E Sproles 2/20/15 11:07 AM
Eliminado: Perhaps one of the most important facets of GRACE data is that it does not distinguish political boundaries, and it is not linked to a specific *in situ* monitoring agency with limited data access, and has the capacity to bridge sparse and inconsistent on-the-ground hydrologic monitoring networks that exist in many regions of the world. Previous GRACE-based analysis has shown its value in highlighting negative trends in terrestrial water storage in trans-boundary watersheds (Voss et al., 2013; Castle et al., 2014), and resulting regional conflict exacerbated by water shortages (Gleick, 2014b). GRACE provides an objective measurement of a region's water resources that can provide valuable insights into potential shortages or surpluses of water resources.

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: varies

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

E Sproles 2/20/15 11:07 AM
Eliminado: Q

E Sproles 2/20/15 11:07 AM
Eliminado: .

E Sproles 2/20/15 11:07 AM
Eliminado: -

E Sproles 2/20/15 11:07 AM
Con formato: Fuente: Cursiva, No revisar la ortografía ni la gramática

validation for model-based forecasts presently used by water resource managers. Our results demonstrate a way forward, expanding GRACE from a diagnostic tool, into a conceptual model and predictive resource.

Although this study focused on the CRB, which has a rich data record, GRACE data are available at a global scale and could be readily applied in areas with a paucity of data to understand how watersheds function and to improve streamflow forecasting capabilities. GRACE does not discern political boundaries and provides an integrated approach to understanding international watersheds (Voss et al., 2013). This resource could serve as a valuable tool for managers in forecasting surplus and scarcity, and in developing strategies that include changes in supply and demand due to human consumptive needs and current climate trends (Wagener et al., 2010; Gleick, 2014a).

Author Contributions

E.A.S., S.G.L., and P.J.W. developed the hysteresis concept based upon background research by J.R. and J.S.F. The data analysis was led by E.A.S., but represents a combined effort from all of the authors. J.R. provided expertise in the GRACE data product, groundwater, and error analysis. E.A.S. prepared the manuscript with contributions from all co-authors.

Acknowledgments

The authors would like to thank Matthew Rodell and Felix Landerer for their expertise in understanding GRACE data during the initial stages of the research. GRACE terrestrial data were processed by Sean Swenson, supported by the NASA MEaSUREs Program, and are available at http://grace.jpl.nasa.gov. The GLDAS and NLDAS data used in this study were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences Data and Information Services Center. We would like to thank the reviewers of the manuscript that have helped improve its overall quality. Additionally, Tim Kerr provided objective feedback and comments on the research findings. The information in this document has been funded entirely by the US Environmental Protection Agency, in part by an appointment to the Internship/Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and EPA. This manuscript has been subjected to Agency review and has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

E Sproles 2/20/15 11:07 AM

Eliminado: Although this study focused on the CRB, which has a rich data record, GRACE data are available at a global scale and could be readily applied in areas with a paucity of data to understand how watersheds function and to improve streamflow forecasting capabilities. GRACE does not discern political boundaries and provides an integrated approach to understanding international watersheds (Voss et al., 2013). This resource could serve as a valuable tool for managers in forecasting surplus and scarcity, and in developing strategies that include changes in supply and demand due to human consumptive needs and current climate trends (Wagener et al., 2010; Gleick, 2014a). ... [36]

References

Appleby, F. V.: Recession and the baseflow problem, Water Resour. Res., 6(5), 1398–1403, doi:10.1029/WR006i005p01398, 1970.

Aspinall, R.: A Century of physical geography research in the Annals, Ann. Assoc. Am. Geogr., 100(5), 1049–1059, 2010.

Ayala, A., McPhee, J. and Vargas, X.: Altitudinal gradients, midwinter melt, and wind effects on snow accumulation in semiarid midlatitude Andes under La Niña conditions, Water Resour. Res., 2014.

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R. and Dozier, J.: Mountain hydrology of the western United States, Water Resour. Res., 42(8), W08432, doi:10.1029/2005wr004387, 2006.

Barthel, R.: A call for more fundamental science in regional hydrogeology, Hydrogeol. J., 22(3), 507–510, 2014.

Beven, K.: Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process, Hydrol. Earth Syst. Sci., 11(1), 460–467, 2007.

Beven, K. J.: Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, Hydrol. Process., 24(12), 1537–1547, 2010.

Black, P. E.: Watershed hydrology, CRC Press, Boca Raton, FL., 1996.

Blöschl, G.: Scaling in hydrology, Hydrol. Process., 15(4), 709–711, 2001.

Brocca, L., Melone, F., Moramarco, T. and Morbidelli, R.: Spatial temporal variability of soil moisture and its estimation across scales, Water Resour. Res., 46(2), 2010.

Brutsaert, W.: Long-term groundwater storage trends estimated from streamflow records: Climatic perspective, Water Resour. Res., 44(2), W02409, doi:10.1029/2007WR006518, 2008.

Capell, R., Tetzlaff, D., Hartley, A. J. and Soulsby, C.: Linking metrics of hydrological function and transit times to landscape controls in a heterogeneous mesoscale catchment, Hydrol. Process., 26(3), 405–420, 2012.

Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C. and Famiglietti, J. S.: Groundwater Depletion During Drought Threatens Future Water Security of the Colorado River Basin, Geophys. Res. Lett., doi:10.1002/2014GL061055, 2014.

Cayan, D. R.: Interannual Climate Variability and Snowpack in the Western United States, J. Clim., 9(5), 928–948, doi:10.1175/1520-0442(1996)009<0928:ICVASI>2.0.CO;2, 1996.

Clark, M. P., Kavetski, D. and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47(9), 2011.

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Marshall, C., Sheffield, J. and Duan, Q.: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, J. Geophys. Res., 108(D22), 8842, 2003.

Dozier, J.: Mountain hydrology, snow color, and the fourth paradigm, Eos, Trans. Am. Geophys. Union, 92(43), 373, doi:10.1029/2011EO430001, 2011.

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., de Linage, C. R. and Rodell, M.: Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophys. Res. Lett., 38(3), L03403, doi:10.1029/2010gl046442, 2011.

Famiglietti, J. S. and Rodell, M.: Water in the Balance, Science (80-.)., 340(6138), 1300–1301, 2013.

Feng, J., Wang, L. and Chen, W.: How does the East Asian summer monsoon behave in the decaying phase of El Niño during different PDO phases?, J. Clim., 27(7), 2682–2698, 2014.

Gleeson, T., Wada, Y., Bierkens, M. F. P. and van Beek, L. P. H.: Water balance of global aquifers revealed by groundwater footprint, Nature, 488(7410), 197–200, 2012.

Gleick, P.: Water, Drought, Climate Change, and Conflict in Syria, Weather. Clim. Soc., 6, 331–340, 2014a.

Gleick, P. H.: The World's Water Volume 8: The Biennial Report on Freshwater Resources, Island Press., 2014b.

Iizumi, T., Luo, J.-J., Challinor, A. J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M. E. and Yamagata, T.: Impacts of El Niño Southern Oscillation on the global yields of major crops, Nat. Commun., 5, 2014.

Jarvis, W. T.: Unitization: a lesson in collective action from the oil industry for aquifer governance, Water Int., 36(5), 619–630, 2011.

Jarvis, W. T.: Contesting Hidden Waters: Conflict Resolution for Groundwater and Aquifers, Routledge., 2014.

Jefferson, A., Nolin, A., Lewis, S. and Tague, C.: Hydrogeologic controls on streamflow sensitivity to climate variation, Hydrol. Process., 22(22), 4371–4385, doi:10.1002/hyp.7041, 2008.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45(2), W02429, doi:10.1029/2008WR006912, 2009.

Koster, R. D., Mahanama, S. P. P., Livneh, B., Lettenmaier, D. P. and Reichle, R. H.: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow, Nat. Geosci, 3(9), 613–616, doi:http://www.nature.com/ngeo/journal/v3/n9/suppinfo/ngeo944_S1.html, 2010.

Landerer, F. W. and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resour. Res., 48(4), W04531, doi:10.1029/2011wr011453, 2012.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, 1999.

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D. and van Dam, T. M.: Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe, J. Hydrol., 446, 103–115, 2012.

Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D. and Reichle, R.: Soil Moisture, Snow, and Seasonal Streamflow Forecasts in the United States, J. Hydrometeorol., 13(1), 189–203, doi:10.1175/JHM-D-11-046.1, 2011.

McDonnell, J. J.: Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response, Hydrol. Process., 17(9), 1869–1875, doi:10.1002/hyp.5132, 2003.

McDonnell, J. J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J. and Weiler, M.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43(7), doi:10.1029/2006wr005467, 2007.

McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and hillslope zones to catchment runoff, Water Resour. Res., 39(11), doi:10.1029/2003WR002091, 2003.

McNamara, J. P., Tetzlaff, D., Bishop, K., Soulsby, C., Seyfried, M., Peters, N. E., Aulenbach, B. T. and Hooper, R.: Storage as a Metric of Catchment Comparison, Hydrol. Process., 25(21), 3364–3371, doi:10.1002/hyp.8113, 2011.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science (80-.)., 319(5863), 573–574, doi:10.1126/science.1151915, 2008.

Moradkhani, H. and Sorooshian, S.: General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis, in Hydrological Modelling and the Water Cycle, pp. 1–24., 2008.

National Operational Hydrologic Remote Sensing Center: Snow Data Assimilation System (SNODAS) Data Products at NSIDC, [2003-2012], National Snow and Ice Data Center, Boulder, Colorado USA., 2004.

Nolin, A. W.: Perspectives on Climate Change, Mountain Hydrology, and Water Resources in the Oregon Cascades, USA, Mt. Res. Dev., 32(S1), S35–S46, doi:10.1659/mrd-journal-d-11-00038.s1, 2012.

Nolin, A. W., Sproles, E. A. and Brown, A.: Climate change impacts on snow and water resources in the Columbia, Willamette, and McKenzie River Basins, USA: A nested watershed study. , in Transboundary River Governance in the Face of Uncertainty: The Columbia River Treaty, edited by B. Cosens, Oregon State University Press, Corvallis, OR., 2012.

Oregon Water Supply and Conservation Initiative: Southern Willamette valley municipal water providers final report., 2008.

Payne, J. T., Wood, A. W., Hamlet, A. F., Palmer, R. N. and Lettenmaier, D. P.: Mitigating the Effects of Climate Change on the Water Resources of the Columbia River Basin, Clim. Change, 62(1), 233–256, 2004.

Peel, M. C. and Blöschl, G.: Hydrological modelling in a changing world, Prog. Phys. Geogr., 35(2), 249–261, 2011.

Reager, J. T. and Famiglietti, J. S.: Global terrestrial water storage capacity and flood potential using GRACE, Geophys. Res. Lett., 36(23), L23402, doi:10.1029/2009gl040826, 2009.

Reager, J. T., Thomas, B. F. and Famiglietti, J. S.: River basin flood potential inferred using GRACE gravity observations at several months lead time, Nat. Geosci., 7, 588–592, doi:10.1038/ngeo2203, 2014.

Riegger, J. and Tourian, M. J.: Characterization of runoff storage relationships by satellite gravimetry and remote sensing, Water Resour. Res., 50(4), 3444–3466, 2014.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D. and Toll, D.: The Global Land Data Assimilation System, Bull. Am. Meteorol. Soc., 85(3), 381–394, doi:10.1175/bams-85-3-381, 2004.

Rodell, M., Velicogna, I. and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460(7258), 999–1002, 2009.

Rupp, D. E., Abatzoglou, J. T., Hegewisch, K. C. and Mote, P. W.: Evaluation of CMIP5 20th century climate simulations for the Pacific Northwest USA, J. Geophys. Res. Atmos., 2013.

Sayama, T., McDonnell, J. J., Dhakal, A. and Sullivan, K.: How much water can a watershed store?, Hydrol. Process., 25(25), 3899–3908, doi:10.1002/hyp.8288, 2011.

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A. and Pulwarty, R. S.: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, Water Resour. Res., 35(7), 2145–2160, doi:10.1029/1999wr900090, 1999.

Skøien, J. O., Blöschl, G. and Western, A. W.: Characteristic space scales and timescales in hydrology, Water Resour. Res., 39(10), 2003.

Spence, C.: A paradigm shift in hydrology: Storage thresholds across scales influence catchment runoff generation, Geogr. Compass, 4(7), 819–833, 2010.

Sproles, E. A., Nolin, A. W., Rittger, K. and Painter, T. H.: Climate change impacts on maritime mountain snowpack in the Oregon Cascades, Hydrol. Earth Syst. Sci., 17(7), 2581–2597, doi:10.5194/hess-17-2581-2013, 2013.

Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33(8), L08402, doi:10.1029/2005gl025285, 2006.

Swenson, S., Yeh, P. J. F., Wahr, J. and Famiglietti, J.: A comparison of terrestrial water storage variations from GRACE with in situ measurements from Illinois, Geophys. Res. Lett., 33(16), L16401, doi:10.1029/2006gl026962, 2006.

Tague, C., Grant, G., Farrell, M., Choate, J. and Jefferson, A.: Deep groundwater mediates streamflow response to climate warming in the Oregon Cascades, Clim. Change, 86(1), 189–210, 2008.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F. and Watkins, M. M.: GRACE measurements of mass variability in the Earth system, Science (80-.)., 305(5683), 503–505, 2004.

The MathWorks: MATLAB and Statistics Toolbox Release 2013a, 2013.

Thompson, S. E., Harman, C. J., Schumer, R., Wilson, J. S., Basu, N. B., Brooks, P. D., Donner, S. D., Hassan, M. A., Packman, A. I. and Rao, P. S. C.: Patterns, puzzles and people: implementing hydrologic synthesis, Hydrol. Process., 25(20), 3256–3266, 2011.

United States Army Corps of Engineers: Comprehensive water supply study - An examination of current water supply issues, edited by T. M. Hillyer., 2001.

Voss, K. A., Famiglietti, J. S., Lo, M., Linage, C., Rodell, M. and Swenson, S. C.: Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris Euphrates Western Iran region, Water Resour. Res., 2013.

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V, Kumar, P., Rao, P. S. C., Basu, N. B. and Wilson, J. S.: The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46(5), 2010.

Wagener, T., Sivapalan, M., Troch, P. and Woods, R.: Catchment classification and hydrologic similarity, Geogr. Compass, 1(4), 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.

Wahr, J., Swenson, S. and Velicogna, I.: Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33(6), L06401, doi:10.1029/2005gl025305, 2006.

Webster, C. S., Kingston, D. G. and Kerr, T.: Inter annual variation in the topographic controls on catchment scale snow distribution in a maritime alpine catchment, New Zealand, Hydrol. Process., 2014.

Western, A. W., Grayson, R. B. and Blöschl, G.: Scaling of soil moisture: A hydrologic perspective, Annu. Rev. Earth Planet. Sci., 30(1), 149–180, 2002.

Wu, H., Kimball, J. S., Elsner, M. M., Mantua, N., Adler, R. F. and Stanford, J.: Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers, Water Resour. Res., 48(11), W11530, doi:10.1029/2012wr012082, 2012.

Yeh, P. J. F., Swenson, S. C., Famiglietti, J. S. and Rodell, M.: Remote sensing of groundwater storage changes in Illinois using the Gravity Recovery and Climate Experiment (GRACE), Water Resour. Res., 42(12), W12203, doi:10.1029/2006wr005374, 2006.

Zaitchik, B. F., Rodell, M. and Reichle, R. H.: Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River Basin, J. Hydrometeorol., 9(3), 535–548, doi:10.1175/2007JHM951.1, 2008.

Zang, C. F., Liu, J., Velde, M. and Kraxner, F.: Assessment of spatial and temporal patterns of green and blue water flows under natural conditions in inland river basins in Northwest China, Hydrol. Earth Syst. Sci., 16(8), 2859–2870, 2012.

E Sproles 2/20/15 11:07 AM
Con formato: Normal,Nivel 1, Interlineado: doble

E Sproles 2/20/15 11:07 AM
Con formato: No revisar la ortografía ni la gramática

Date	DAL.R	DAL.TWS	DAL.GWSA	DAL.sub	DAL.SM	DAL.SWE	DAL.res
Oct-03	13.2	-123.0	-4.5	-69.7	-65.3	-53.2	-0.07
Nov-03	14.8	-88.0	-3.6	-46.4	-42.8	-41.5	-0.08
Dec-03	17.5	-35.0	-2.9	-29.9	-27.0	-5.0	-0.04
Jan-04	17.4	14.0	-6.5	-28.2	-21.7	42.4	-0.15
Feb-04	15.1	55.0	-0.8	-12.1	-11.3	67.3	-0.24
Mar-04	17.6	62.0	-9.5	2.0	11.5	60.3	-0.27
Apr-04	19.4	34.0	-13.7	24.8	38.5	9.2	-0.08
May-04	27.2	17.0	5.9	46.8	40.9	-29.9	0.06
Jun-04	26.6	-2.0	10.4	43.5	33.0	-45.7	0.25
Jul-04	18.1	-50.0	8.9	2.2	-6.7	-52.4	0.21
Aug-04	16.0	-103.0	-4.8	-50.1	-45.3	-52.9	-0.04
Sep-04	13.6	-70.0	11.0	-16.5	-27.5	-53.4	-0.10
Oct-04	13.8	-72.0	7.9	-20.3	-28.2	-51.7	-0.05
Nov-04	14.9	-45.0	14.9	-0.4	-15.3	-44.6	0.00
Dec-04	19.1	-25.0	-1.1	-3.6	-2.5	-21.4	0.05
Jan-05	18.4	18.0	4.8	12.3	7.5	5.6	0.09
Feb-05	16.2	24.0	-10.3	13.4	23.7	10.5	0.13
Mar-05	16.4	18.0	-21.2	14.4	35.6	3.6	0.00
Apr-05	17.4	45.0	-5.4	48.0	53.4	-3.0	-0.06
May-05	28.8	30.0	8.4	68.4	60.0	-38.4	0.04
Jun-05	23.3	10.0	-12.6	61.0	73.6	-51.3	0.31
Jul-05	21.7	-44.0	-31.8	9.1	40.9	-53.4	0.29
Aug-05	16.8	-65.0	9.2	-11.6	-20.8	-53.4	0.02
Sep-05	11.3	-108.0	-25.0	-54.5	-29.5	-53.4	-0.09
Oct-05	13.7	-86.0	-28.5	-33.2	-4.7	-52.7	-0.04
Nov-05	14.7	-54.0	-25.5	-18.9	6.6	-35.0	-0.02
Dec-05	16.6	-8.0	-9.6	-4.5	5.1	-3.5	-0.04
Jan-06	22.5	38.0	-39.0	-8.0	31.0	45.9	0.07
Feb-06	19.9	103.0	-23.6	14.2	37.8	88.9	-0.08
Mar-06	20.3	114.0	-42.4	10.4	52.8	103.8	-0.23
Apr-06	36.5	117.0	-46.0	41.2	87.2	76.1	-0.28
May-06	42.2	79.0	-6.3	77.7	83.9	1.6	-0.24
Jun-06	37.5	50.0	23.5	95.7	72.2	-46.0	0.28
Jul-06	22.3	-16.0	26.6	36.6	10.0	-52.9	0.29
Aug-06	16.9	-72.0	28.9	-18.6	-47.5	-53.4	0.03
Sep-06	11.1	-98.0	25.1	-44.6	-69.7	-53.4	-0.05
Oct-06	12.0	-96.0	21.9	-43.9	-65.8	-52.0	-0.03
Nov-06	15.5	-28.0	36.5	11.9	-24.6	-39.9	0.03
Dec-06	17.5	9.0	20.7	5.2	-15.5	3.8	0.06
Jan-07	21.1	57.0	5.4	-3.5	-8.9	60.5	0.08
Feb-07	16.1	81.0	-19.8	-21.1	-1.3	102.1	-0.01
Mar-07	24.8	94.0	-37.0	-12.4	24.6	106.3	0.07
Apr-07	27.4	94.0	-46.7	4.9	51.6	89.2	-0.06
May-07	32.3	82.0	3.3	55.6	52.3	26.4	-0.03
Jun-07	25.7	46.0	43.0	75.6	32.6	-29.8	0.24
Jul-07	21.1	-37.0	36.9	13.7	-23.2	-50.9	0.21
Aug-07	17.5	-75.0	51.2	-21.6	-72.8	-53.4	-0.03
Sep-07	10.7	-101.0	41.5	-47.5	-89.0	-53.4	-0.12
Oct-07	12.0	-105.0	6.4	-52.9	-59.2	-52.1	-0.07
Nov-07	13.6	-69.0	20.3	-25.5	-45.8	-43.4	-0.06
Dec-07	16.0	-11.0	23.9	-8.4	-32.2	-2.6	-0.02
Jan-08	18.2	51.0	18.6	-8.5	-27.1	59.5	-0.07
Feb-08	15.0	97.0	-1.4	-19.5	-18.1	116.7	-0.21
Mar-08	17.7	102.0	-21.2	-12.4	8.9	114.7	-0.31
Apr-08	19.3	104.0	-29.9	-0.3	29.6	104.7	-0.44
May-08	38.2	93.0	10.8	63.5	52.7	30.0	-0.47
Jun-08	43.1	58.0	44.1	88.6	44.5	-30.8	0.18
Jul-08	25.5	-33.0	25.2	17.6	-7.6	-50.8	0.28
Aug-08	16.2	-69.0	33.1	-15.7	-48.8	-53.4	0.06
Sep-08	11.1	-88.0	23.0	-34.5	-57.5	-53.4	-0.09
Oct-08	11.9	-80.0	26.2	-27.4	-53.6	-52.6	-0.05
Nov-08	14.4	-53.0	28.5	-5.7	-34.2	-47.3	-0.02
Dec-08	16.7	-24.0	22.7	-6.2	-28.9	-17.8	-0.02
Jan-09	22.1	44.0	7.8	-9.1	-17.0	53.1	0.05
Feb-09	14.7	35.0	-24.0	-31.2	-7.2	66.2	-0.01
Mar-09	17.0	82.0	-13.2	-0.8	12.4	82.7	0.05
Apr-09	27.7	84.0	-29.2	11.5	40.7	72.6	-0.04
May-09	34.2	72.0	0.8	53.6	52.8	18.5	-0.10
Jun-09	34.4	19.0	28.4	54.9	26.4	-36.2	0.30
Jul-09	20.4	-24.0	19.2	27.0	7.9	-51.4	0.35
Aug-09	14.7	-58.0	26.6	-4.8	-31.3	-53.4	0.14
Sep-09	9.8	-100.0	17.1	-46.6	-63.7	-53.4	-0.01
Oct-09	11.2	-84.0	29.6	-33.8	-63.4	-50.2	0.04
Nov-09	14.1	-56.0	14.4	-21.3	-35.7	-34.7	0.06
Dec-09	16.6	-23.0	10.8	-18.7	-29.5	-4.3	0.03
Jan-10	14.8	8.0	-7.8	-24.8	-16.9	32.7	0.07
Feb-10	13.3	28.0	-19.5	-25.8	-6.4	53.8	0.05
Mar-10	13.9	53.0	-5.1	7.9	13.0	45.1	0.00
Apr-10	16.0	39.0	-25.4	3.8	29.3	35.2	0.00
May-10	25.3	46.0	4.1	43.3	39.3	2.7	-0.02
Jun-10	40.5	62.0	42.6	90.7	48.1	-29.0	0.32
Jul-10	23.3	2.0	42.3	50.5	8.2	-48.8	0.32
Aug-10	15.5	-59.0	33.1	-6.2	-39.4	-52.9	0.11
Sep-10	10.5	-70.0	30.3	-16.6	-46.8	-53.4	-0.04
Oct-10	12.8	-70.0	25.3	-17.8	-43.2	-52.2	0.00
Nov-10	15.1	-49.0	15.8	-12.1	-27.9	-36.9	0.00
Dec-10	17.4	-1.0	2.6	-17.4	-20.1	16.4	0.01
Jan-11	25.2	45.0	-11.9	-18.6	-6.7	63.5	0.06
Feb-11	24.0	91.0	-9.4	-2.0	7.4	93.1	-0.09
Mar-11	28.3	132.0	-22.7	4.9	27.6	127.4	-0.35
Apr-11	36.0	166.0	-34.3	31.6	66.0	134.9	-0.59
May-11	48.9	173.0	-2.7	88.6	91.3	85.0	-0.63
Jun-11	58.3	120.0	21.7	109.7	88.0	10.4	-0.14
Jul-11	40.7	67.0	67.8	104.3	36.5	-37.7	0.38
Aug-11	24.6	13.0	83.8	63.1	-20.7	-50.4	0.28
Sep-11	14.7	-62.0	51.7	-9.2	-60.8	-52.9	0.07
Oct-11	14.8	-30.0	72.2	22.1	-50.1	-52.2	0.05
Nov-11	16.0	-30.0	41.6	2.8	-38.8	-32.8	0.06
Dec-11	17.4	23.0	57.0	27.8	-29.2	-4.8	0.00
Jan-12	18.3	52.0	36.6	15.7	-20.9	36.3	0.01
Feb-12	16.2	96.0	29.3	18.5	-10.7	77.5	0.00
Mar-12	25.8	127.0	9.6	20.9	11.4	106.1	-0.07
Apr-12	40.1	188.0	49.5	102.1	52.6	86.2	-0.33
May-12	44.3	158.5	74.2	136.9	62.7	21.8	-0.23
Jun-12	42.2	129.0	88.8	152.6	63.8	-23.7	0.16
Jul-12	40.4	43.0	65.8	90.6	24.8	-47.9	0.25
Aug-12	25.1	-4.0	85.2	48.8	-36.4	-52.9	0.06
Sep-12	14.1	-42.0	84.5	11.4	-73.1	-53.4	-0.05

Date	UPC.R	UPC.TWS	UPC.GWSA	UPC.sub	UPC.SM	UPC.SWE	UPC.res
Oct-03	38.9	-123	32.6	-32.6	-65.3	-90.4	-0.01
Nov-03	40.5	-74	38.1	-4.6	-42.8	-69.4	-0.02
Dec-03	49.4	-23	7.7	-19.3	-27.0	-3.7	-0.03
Jan-04	36.4	9	-23.3	-45.0	-21.7	54.0	-0.04
Feb-04	26.0	62	-21.4	-32.7	-11.3	94.7	-0.06
Mar-04	30.5	66	-41.2	-29.7	11.5	95.7	-0.07
Apr-04	33.8	36	-23.8	14.7	38.5	21.3	-0.04
May-04	48.7	24	32.8	73.7	40.9	-49.7	0.02
Jun-04	59.0	-11	33.7	66.7	33.0	-77.8	0.06
Jul-04	48.4	-56	41.1	34.4	-6.7	-90.4	0.06
Aug-04	44.2	-117	19.8	-25.6	-45.3	-91.4	0.01
Sep-04	42.5	-66	52.9	25.4	-27.5	-91.4	0.01
Oct-04	37.1	-83	33.3	5.1	-28.2	-88.1	0.01
Nov-04	39.3	-42	44.2	28.9	-15.3	-70.9	0.02
Dec-04	51.4	-12	21.7	19.2	-2.5	-31.2	0.00
Jan-05	44.7	29	15.3	22.8	7.5	6.1	0.02
Feb-05	32.0	45	2.3	26.0	23.7	19.0	0.03
Mar-05	29.1	37	-9.6	26.0	35.6	11.0	0.03
Apr-05	31.0	60	4.2	57.6	53.4	2.4	0.01
May-05	54.4	14	19.8	79.7	60.0	-65.8	0.04
Jun-05	68.9	5	18.7	92.3	73.6	-87.4	0.07
Jul-05	64.5	-46	4.5	45.4	40.9	-91.4	0.06
Aug-05	50.6	-73	39.3	18.4	-20.8	-91.4	0.01
Sep-05	38.7	-107	13.9	-15.6	-29.5	-91.4	0.01
Oct-05	40.1	-81	13.2	8.5	-4.7	-89.5	0.01
Nov-05	36.3	-44	8.3	14.9	6.6	-58.9	0.01
Dec-05	48.5	-16	0.2	5.4	5.1	-21.4	0.00
Jan-06	34.5	19	-62.9	-31.9	31.0	50.9	-0.01
Feb-06	27.3	93	-63.1	-25.3	37.8	118.3	0.00
Mar-06	35.8	113	-93.3	-40.4	52.8	153.5	-0.03
Apr-06	42.3	113	-91.2	-4.1	87.2	117.1	-0.03
May-06	75.8	74	-13.6	70.4	83.9	3.7	-0.03
Jun-06	83.1	45	53.0	125.2	72.2	-80.3	0.05
Jul-06	61.6	-35	46.4	56.4	10.0	-91.4	0.06
Aug-06	49.3	-101	38.0	-9.6	-47.5	-91.4	0.01
Sep-06	37.9	-125	35.6	-34.1	-69.7	-90.9	0.02
Oct-06	31.9	-130	24.0	-41.9	-65.8	-88.1	0.00
Nov-06	31.8	-44	43.4	18.8	-24.6	-62.8	0.00
Dec-06	41.9	19	26.3	10.8	-15.5	8.2	0.00
Jan-07	44.6	77	1.9	-7.0	-8.9	84.0	0.00
Feb-07	38.6	105	-39.9	-41.2	-1.3	146.2	-0.01
Mar-07	38.8	130	-68.6	-44.0	24.6	174.0	-0.01
Apr-07	48.9	140	-65.8	-14.2	51.6	154.2	-0.01
May-07	71.6	129	24.5	76.8	52.3	52.2	0.02
Jun-07	72.7	82	96.0	128.6	32.6	-46.6	0.06
Jul-07	65.2	-14	94.6	71.4	-23.2	-85.4	0.06
Aug-07	51.8	-75	88.7	15.9	-72.8	-90.9	0.01
Sep-07	35.9	-102	78.4	-10.6	-89.0	-91.4	0.01
Oct-07	30.9	-112	35.5	-23.7	-59.2	-88.3	-0.01
Nov-07	32.7	-75	42.2	-3.6	-45.8	-71.4	-0.02
Dec-07	43.0	-9	32.2	-0.1	-32.2	-8.9	-0.02
Jan-08	39.8	52	13.7	-13.4	-27.1	65.4	-0.04
Feb-08	27.5	102	-23.0	-41.1	-18.1	143.2	-0.06
Mar-08	29.8	113	-61.5	-52.6	8.9	165.7	-0.08
Apr-08	26.6	122	-70.3	-40.7	29.6	162.7	-0.05
May-08	68.4	120	14.4	67.1	52.7	53.0	-0.07
Jun-08	89.8	74	79.0	123.5	44.5	-49.6	0.03
Jul-08	63.9	-34	59.5	51.9	-7.6	-85.9	0.06
Aug-08	40.1	-81	58.7	9.9	-48.8	-90.9	0.03
Sep-08	32.7	-98	51.0	-6.6	-57.5	-91.4	0.00
Oct-08	33.0	-91	52.0	-1.6	-53.6	-89.4	-0.02
Nov-08	28.5	-63	49.5	15.3	-34.2	-78.3	-0.02
Dec-08	45.0	-29	36.7	7.7	-28.9	-36.7	-0.03
Jan-09	44.7	52	12.3	-4.6	-17.0	56.7	-0.03
Feb-09	33.7	38	-44.4	-51.6	-7.2	89.7	-0.06
Mar-09	29.1	81	-64.3	-51.9	12.4	132.9	-0.07
Apr-09	28.9	87	-77.5	-36.9	40.7	124.0	-0.06
May-09	51.6	75	-25.3	27.6	52.8	47.5	-0.04
Jun-09	65.8	0	22.7	49.2	26.4	-49.2	0.03
Jul-09	51.6	-36	41.0	48.9	7.9	-84.9	0.06
Aug-09	33.4	-88	34.7	3.4	-31.3	-91.4	0.06
Sep-09	31.1	-136	19.1	-44.6	-63.7	-91.4	0.05
Oct-09	30.8	-117	32.2	-31.2	-63.4	-85.8	0.03
Nov-09	32.5	-68	15.7	-20.0	-35.7	-48.0	0.02
Dec-09	44.8	-23	-3.3	-32.9	-29.5	9.9	0.01
Jan-10	32.8	21	-25.0	-41.9	-16.9	62.9	-0.01
Feb-10	24.4	26	-60.0	-66.4	-6.4	92.4	-0.03
Mar-10	27.2	48	-53.0	-40.0	13.0	88.1	-0.04
Apr-10	22.8	47	-61.1	-31.8	29.3	78.8	-0.04
May-10	40.5	50	-14.2	25.1	39.3	24.9	-0.01
Jun-10	73.4	70	55.2	103.3	48.1	-33.3	0.04
Jul-10	55.5	-5	65.7	73.9	8.2	-78.9	0.05
Aug-10	35.1	-66	63.8	24.4	-39.4	-90.4	0.04
Sep-10	31.9	-77	60.7	13.9	-46.8	-90.9	0.02
Oct-10	33.8	-75	58.2	15.1	-43.2	-90.1	0.01
Nov-10	31.2	-62	31.7	3.7	-27.9	-65.7	0.01
Dec-10	41.4	-17	-2.2	-22.3	-20.1	5.3	0.01
Jan-11	46.7	46	-32.5	-39.3	-6.7	84.8	0.00
Feb-11	42.7	108	-62.9	-55.5	7.4	163.5	-0.02
Mar-11	45.3	157	-96.7	-69.1	27.6	226.1	-0.07
Apr-11	42.4	202	-114.8	-48.8	66.0	250.9	-0.12
May-11	76.5	199	-57.3	34.0	91.3	165.1	-0.14
Jun-11	116.9	137	12.8	100.8	88.0	35.8	-0.05
Jul-11	97.1	74	89.4	125.9	36.5	-51.9	0.05
Aug-11	58.6	18	122.1	101.4	-20.7	-83.4	0.06
Sep-11	36.4	-79	71.2	10.4	-60.8	-89.4	0.05
Oct-11	39.1	-46	92.9	42.8	-50.1	-88.8	0.04
Nov-11	38.8	-48	40.3	1.5	-38.8	-49.5	0.03
Dec-11	44.4	44	63.4	34.2	-29.2	9.8	0.02
Jan-12	44.5	77	19.3	-1.6	-20.9	78.6	0.01
Feb-12	32.5	111	-24.6	-35.4	-10.7	146.4	-0.01
Mar-12	40.7	170	-52.6	-41.3	11.4	211.3	-0.02
Apr-12	52.4	230	-16.3	36.2	52.6	193.8	-0.03
May-12	84.0	206	70.8	133.5	62.7	72.0	0.00
Jun-12	106.9	181	140.0	203.8	63.8	-22.8	0.05
Jul-12	115.9	72	125.6	150.4	24.8	-78.4	0.06
Aug-12	70.0	6	132.3	95.9	-36.4	-89.9	0.06
Sep-12	41.6	-47	117.0	43.9	-73.1	-90.9	0.04

Date	SNK.R	SNK.TWS	SNK.GWSA	SNK.sub	SNK.SM	SNK.SWE	SNK.res
Oct-03	4.5	-131	-10.4	-96.4	-85.9	-34.6	-0.12
Nov-03	4.4	-110	-3.0	-84.3	-81.2	-25.7	-0.10
Dec-03	4.4	-59	0.0	-56.3	-56.2	-2.8	-0.07
Jan-04	4.3	18	9.0	-30.7	-39.5	48.6	-0.04
Feb-04	4.7	44	-1.2	-24.8	-23.4	68.6	-0.01
Mar-04	6.6	66	7.2	12.7	5.7	53.1	0.02
Apr-04	5.0	29	9.3	24.6	15.4	4.3	0.04
May-04	5.6	26	43.7	47.0	3.4	-21.2	0.01
Jun-04	4.8	31	74.5	62.4	-12.0	-31.5	-0.01
Jul-04	3.4	-40	38.1	-5.4	-43.4	-34.6	-0.06
Aug-04	3.7	-100	0.2	-65.4	-65.5	-34.6	-0.11
Sep-04	3.9	-81	21.7	-46.4	-67.9	-34.6	-0.13
Oct-04	4.5	-80	14.9	-46.7	-61.6	-33.3	-0.12
Nov-04	4.2	-53	7.7	-24.3	-31.8	-28.8	-0.08
Dec-04	4.3	-37	-7.6	-24.4	-16.7	-12.7	-0.05
Jan-05	4.1	18	3.1	-0.7	-3.7	18.6	-0.02
Feb-05	3.6	19	-14.4	-6.1	8.5	24.9	0.01
Mar-05	4.6	22	-21.0	2.4	23.6	19.4	0.04
Apr-05	4.5	42	-6.1	33.7	39.9	8.1	0.07
May-05	8.3	58	11.6	77.1	65.7	-19.3	0.07
Jun-05	4.7	30	10.3	62.8	52.6	-33.0	0.07
Jul-05	3.6	-34	0.3	0.5	0.4	-34.6	0.01
Aug-05	3.5	-45	25.0	-10.4	-35.3	-34.6	-0.06
Sep-05	3.8	-102	-16.0	-67.4	-51.3	-34.6	-0.11
Oct-05	4.6	-88	-9.1	-53.9	-44.7	-34.1	-0.10
Nov-05	4.8	-73	-21.5	-48.9	-27.3	-24.2	-0.07
Dec-05	5.4	-7	-5.3	-12.8	-7.3	5.7	-0.04
Jan-06	7.8	41	-53.3	-9.0	44.5	49.8	0.00
Feb-06	7.7	112	-26.2	32.2	58.6	79.6	0.04
Mar-06	12.2	117	-40.6	36.1	76.8	80.8	0.06
Apr-06	23.6	136	-30.5	77.5	108.1	58.3	0.08
May-06	18.9	103	13.5	104.2	90.8	-1.4	0.09
Jun-06	8.4	64	42.2	96.1	54.0	-32.3	0.08
Jul-06	4.9	0	32.9	34.5	1.7	-34.6	0.02
Aug-06	4.5	-46	23.4	-11.5	-34.7	-34.6	-0.05
Sep-06	4.6	-69	12.3	-34.4	-46.6	-34.6	-0.09
Oct-06	4.9	-67	2.8	-33.2	-35.9	-33.8	-0.09
Nov-06	4.6	-30	19.2	-1.9	-21.0	-28.2	-0.04
Dec-06	5.2	-14	-10.2	-12.7	-2.4	-1.4	0.00
Jan-07	5.3	31	-22.5	-10.5	12.1	41.4	0.02
Feb-07	4.9	58	-35.6	-14.2	21.5	72.0	0.05
Mar-07	5.7	57	-44.4	-10.1	34.5	66.9	0.08
Apr-07	4.9	61	-24.8	14.4	39.4	46.4	0.10
May-07	5.1	61	32.1	57.1	25.2	3.7	0.07
Jun-07	3.6	32	60.6	57.9	-2.6	-26.0	0.02
Jul-07	3.6	-49	26.3	-14.5	-40.7	-34.6	-0.05
Aug-07	3.7	-70	26.5	-35.4	-61.7	-34.6	-0.11
Sep-07	3.9	-95	8.7	-60.4	-69.0	-34.6	-0.13
Oct-07	4.6	-108	-21.8	-74.1	-52.1	-33.9	-0.12
Nov-07	4.4	-67	-1.9	-36.8	-34.8	-30.2	-0.09
Dec-07	4.6	-29	-4.2	-26.1	-21.8	-3.0	-0.05
Jan-08	4.3	37	1.3	-10.9	-12.1	47.8	-0.02
Feb-08	4.0	85	-11.9	-10.8	1.2	95.7	0.00
Mar-08	5.7	82	-30.8	-3.0	28.0	84.9	0.03
Apr-08	5.7	81	-23.6	15.1	38.8	65.7	0.05
May-08	9.4	72	29.0	63.4	34.6	8.4	0.04
Jun-08	7.3	35	41.0	61.0	20.2	-26.2	0.05
Jul-08	4.4	-39	21.4	-4.5	-25.8	-34.6	0.00
Aug-08	4.5	-73	15.4	-38.4	-53.8	-34.6	-0.06
Sep-08	4.1	-86	13.8	-51.4	-65.1	-34.6	-0.10
Oct-08	4.7	-78	11.8	-44.2	-55.9	-33.8	-0.09
Nov-08	4.5	-63	0.1	-33.6	-33.5	-29.5	-0.05
Dec-08	4.3	-26	7.7	-19.3	-26.9	-6.8	-0.02
Jan-09	4.4	17	-25.7	-29.7	-3.9	46.5	0.01
Feb-09	3.7	28	-39.4	-26.2	13.4	54.0	0.05
Mar-09	5.2	73	-14.2	22.4	36.7	50.4	0.08
Apr-09	8.3	87	-1.7	47.7	49.5	39.1	0.09
May-09	10.2	73	21.7	72.4	50.8	0.4	0.10
Jun-09	10.4	51	30.9	80.8	50.1	-30.1	0.09
Jul-09	6.1	-5	15.3	29.4	14.2	-34.6	0.07
Aug-09	4.2	-31	23.9	3.5	-20.3	-34.6	0.01
Sep-09	4.2	-76	0.7	-41.5	-42.0	-34.6	-0.04
Oct-09	5.3	-51	15.2	-19.8	-34.9	-31.3	-0.04
Nov-09	4.7	-53	0.1	-25.4	-25.4	-27.8	-0.01
Dec-09	5.1	-34	-5.3	-23.2	-17.8	-10.9	0.02
Jan-10	5.6	-11	-34.4	-35.1	-0.6	23.9	0.04
Feb-10	4.9	30	-34.7	-17.9	17.0	47.7	0.07
Mar-10	5.9	64	-7.3	32.6	40.0	31.2	0.09
Apr-10	5.7	35	-33.6	19.7	53.5	15.1	0.10
May-10	7.0	55	9.1	64.8	55.9	-10.1	0.08
Jun-10	10.6	56	32.6	84.8	52.3	-29.0	0.08
Jul-10	4.2	5	34.6	39.4	5.0	-34.6	0.05
Aug-10	3.7	-50	12.0	-15.5	-27.4	-34.6	-0.01
Sep-10	4.2	-63	13.1	-28.4	-41.4	-34.6	-0.06
Oct-10	5.1	-74	-0.2	-40.1	-39.8	-33.9	-0.08
Nov-10	4.8	-44	-10.1	-25.2	-15.0	-18.9	-0.04
Dec-10	5.1	0	-42.9	-28.5	14.5	28.4	-0.01
Jan-11	7.5	40	-44	-21.1	40.7	60.9	0.02
Feb-11	6.2	79	-45.2	11.6	57.0	67.2	0.03
Mar-11	9.7	110	-54.8	27.5	82.5	82.3	0.04
Apr-11	17.5	141	-47.5	60.3	107.9	80.5	0.07
May-11	22.2	168	4.6	122.1	117.6	45.7	0.08
Jun-11	19.3	125	47	128.6	104.3	-4.3	0.09
Jul-11	8.1	81	70.4	113.3	43.0	-32.6	0.10
Aug-11	5.6	25	60.2	59.4	-0.7	-34.6	0.07
Sep-11	6.5	-45	15.3	-10.5	-25.7	-34.6	0.00
Oct-11	8.4	-14	32.9	19.5	-13.3	-33.6	-0.04
Nov-11	7.1	-17	14.8	6.6	-8.1	-23.7	-0.03
Dec-11	6.9	2	19.0	14.7	-4.2	-12.8	-0.01
Jan-12	6.7	28	6.1	16.6	10.6	11.3	0.01
Feb-12	5.8	76	4.4	37.2	32.9	38.7	0.04
Mar-12	9.4	78	-14.5	35.1	49.7	42.7	0.07
Apr-12	13.2	136	60.6	122.3	61.8	13.5	0.09
May-12	10.7	112	76.3	126.0	49.8	-14.7	0.08
Jun-12	6.3	87	97.4	117.4	20.2	-30.6	0.03
Jul-12	4.3	4	58.7	38.5	-20.0	-34.6	-0.05
Aug-12	3.9	-30	53.6	4.6	-48.9	-34.6	-0.09
Sep-12	4.2	-56	42.9	-21.4	-64.2	-34.6	-0.11