Environmental Research Communications

LETTER

Forecast of seasonal water availability in Central Asia with near-real time GRACE water storage anomalies

Heiko Apel[1] , Ben Gouweleeuw[1], Abror Gafurov[1] and Andreas Güntner[1,2]

[1] GFZ German Research Centre for Geoscience, section 5.4 Hydrology, Potsdam, Germany
[2] University of Potsdam, Institute of Earth and Environmental Science, Potsdam, Germany

E-mail: heiko.apel@gfz-potsdam.de

Abstract

Water availability during summer in Central Asia is controlled by the snow melt in the surrounding mountains. Reliable forecasts of river discharge during this period are essential for the management of water resources. This study tests the predictive power of GRACE gravity-based water storage anomalies in a linear regression framework for two large catchments. The results show substantial improvements of the forecasts in the larger Amudarya catchment compared to forecasts using just climate, snow cover, and discharge data. In this catchment, GRACE water storage anomalies even provide the largest share of explained variance. This leads to the conclusion that GRACE data can improve the forecast of seasonal water availability for large basins in Central Asia. The GRACE-FO mission launched in May 2018 opens up the possibility of operational forecasts utilizing upcoming near-real time products from satellite gravimetry for Central Asia and similar environments.

1. Introduction

Central Asia (CA) is a semi-arid region spanning over Kazakhstan, Uzbekistan, Kyrgyzstan, Turkmenistan, Tajikistan, northern Afghanistan, and north-western regions of China. CA hosts two large mountain ranges (Tien Shan and Pamir). All the large rivers in the region (Amudarya, Sirdarya, Ili) originate from these mountains. The semi-arid and endorheic Aral-Sea basin, to which the mentioned river systems drain, shows a distinct seasonality in precipitation and river discharge. For large parts of CA, the bulk of precipitation occurs in the boreal winter months (Sorg *et al* 2012, Aizen *et al* 1996), with a particular spatial concentration in the mountain ranges. Precipitation during the summer months is, in contrast, practically zero in the lowland areas. Water originating from snowmelt and to a lesser extend from glacier melt is then the only source for generating river discharge. During summer river discharge is thus the only water resource available and thus of high importance for the agricultural systems and economies of the countries in the Aral-Sea basin. Particularly the agriculture with its intensive irrigation schemes crucially depend on the water provided by the river systems (Viviroli *et al* 2007). Therefore the forecast of river discharge during summer is of high importance for the regions, and in fact a mandatory task of all hydro-meteorological services in the area. Here forecasts of the seasonal river discharge during summer (April to September) are issued operationally on a monthly basis from January to June. For water resources planning the forecasts issued just before the summer period at the start of April is of highest importance. However, forecast methods and procedures in the CA countries are chronically outdated, as well as capacities are limited (Apel *et al* 2018). Recently, Apel *et al* (2018) developed a simple concept for forecast models for CA catchments based on precipitation, air temperature, river discharge that are readily available to the hydromet services, supplemented by operational satellite based snow cover data (Gafurov *et al* 2016). The concept proved to provide forecasts with high performance for lead times up to 3 months and for catchments of widely varying size and location in CA. This could be achieved because the accumulation period and runoff generation is well separated between the seasons. Snow is accumulating and stored in the mountains in winter, and released with the onset of snow melt in spring. The rational in Apel *et al* (2018) was that the

Figure 1. The Amudarya and Naryn basins in Central Asia considered in this study.
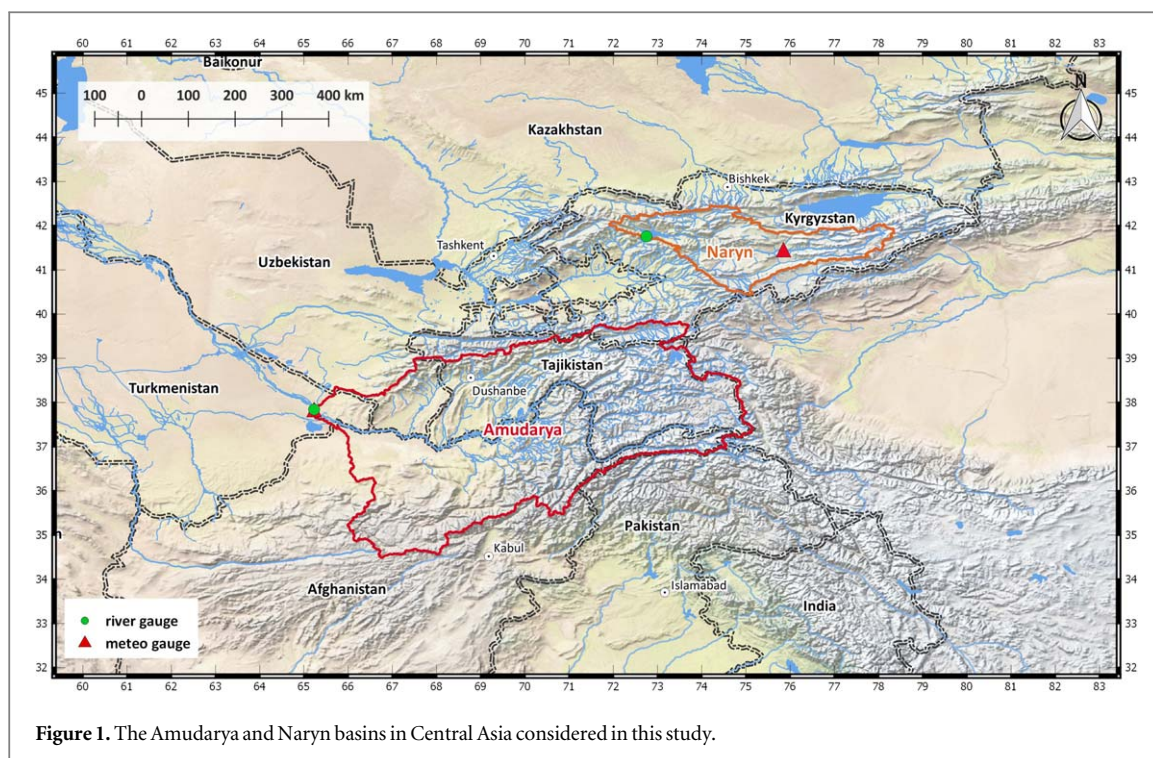
combination of snow cover information with temperature and precipitation may serve as a proxy for the water stored in the accumulated snow mass (the snow water equivalent SWE), and could thus be used for seasonal discharge prediction. The successful application of the model supported this rational.

This finding suggests that direct observations of water storage, here SWE, instead of proxies such as precipitation and snow cover area could also serve well as predictors for seasonal discharge in CA. Such observations are provided by the Gravity Recovery and Climate Experiment (GRACE) Satellite Mission (Tapley *et al* 2004). From the continuous measurements of the Earth's gravity field by GRACE, total water storage anomalies (TWSA) have been derived. GRACE TWSA encompasses storage variations in all terrestrial compartments, including glacier and snow accumulation and depletion. The feasibility of providing products derived from GRACE in near-real time has been demonstrated in the Horizon 2020 project European Gravity Service for Improved Emergency Management (EGSIEM). Against the background of the recently launched GRACE-Follow On (GRACE-FO) mission of which data products are expected to become available by mid 2019, this, in principle, opens up the possibility of using satellite gravity data in operational forecasts. Thus, the aim of this study is to include GRACE TWSA in the forecast models of Apel *et al* (2018) to test the applicability and performance of TWSA in forecasting seasonal discharge taking CA as a test region.

2. Study area and data

Out of the set of river basins tested by Apel *et al* (2018) for seasonal forecasts in Central Asia, the two largest catchments were selected for this study: the Amudarya catchment upstream of the gauge Kerky, close to the border of Tajikistan and Turkmenistan, and the Naryn catchment upstream of the Toktogul reservoir (figure 1). The Amudarya catchment covers an area of 287,714 km^2, the Naryn catchment an area of 51,926 km^2. The Naryn catchment is thus actually too small to be properly represented by the coarse spatial resolution of GRACE data in the order of 100 000 km^2. Below this size the signal-to-noise ratio tends to decrease markedly (e.g., Zhang *et al* (2016), Vishwakarma *et al* (2018)). Nevertheless, the Naryn catchment is included in this analysis in order to test if TWSA signals of the larger area, which are unavoidably included in the TWSA data extracted for Naryn, are representative for the storage in Naryn, and if TWSA can potentially be used consequently as a predictor also for smaller catchments, at least in this geographical context.

For Amudarya the river discharge (Q) time series of gauge Kerky was used to derive the seasonal discharge. For Naryn, the overall inflow into the Toktogul reservoir at the outlet of the catchment was used. Precipitation (P) and air temperature (T) data were obtained from the climate stations Kerky (Amudarya) and Naryn city (Naryn). All these data have monthly resolution. Monthly mean snow coverage (SC) of the catchments was derived from daily snow cover maps derived from MODIS satellite data by the ModSnow tool (Gafurov *et al* 2016).

The GRACE product used in this study was the daily GRACE gravity solutions using Kalman smoothing produced by the Technical University of Graz (ITGS-Grace2016_Kalman) based on the procedure described in Kurtenbach *et al* (2012). The predictions of the Kalman smoother in the daily GRACE product allow for filling observational data gaps, however, potentially at the expense of accuracy. The data was given as global 1 degree grids of TWSA from which daily area-mean values of all cells covering the two catchments were calculated, and further aggregated to monthly means.

The daily GRACE product was preferred to the standard monthly GRACE products (as, e.g., provided as Level-3 products from GFZ, CSR and JPL at http://grace.jpl.nasa.gov) because it provides data even for months with few GRACE observations, for which the standard monthly data have not been produced. Missing months would impair the testing of GRACE as predictor. In linear modelling, missing predictor data, which is equivalent to a lower number of observations to which the models are fitted, can lead to overfitting and thus spurious results. Short time series of observations or predictors can be much better fitted than longer time series, especially if the number of data points (here years with complete predictor data) to be fitted comes close to the number of predictors, thus reducing the degrees of freedom for model fitting. Overfitting reduces the robustness of the fits and thus the predictions. In the present study, the data gaps in the standard monthly GRACE data would reduce the number of years with GRACE predictors to 8, instead of 13 for the daily solution, thereby considerably increasing the chance of model overfitting and thus also a meaningful test of GRACE as seasonal discharge predictor. Therefore we did not use the standard monthly GRACE solutions nor the monthly GRACE mascon solutions which have the same temporal coverage as the standard monthly products, in order to minimize overfitting and spurious model fitting results and predictions.

However, we evaluated the correlation between the monthly GRACE products (JPL RL05, GFZ RL05, CSR RL05, ITSG-Grace2016) and the daily Kalman smoothed solutions in order to get an insight into the potential applicability of different GRACE products for seasonal discharge forecasts. The analysis showed that predictors based on the monthly GRACE products are highly correlated to the predictors based on the daily ITSG solution in Amudarya, with linear correlation coefficients in the range of 0.96–0.97. For the Naryn basin, the correlations are lower but still high in the range of 0.77–0.87. This means that using one of the monthly GRACE solutions as predictor would result in similar to identical results compared to the predictors based on the daily GRACE products, because in linear modelling highly correlated predictors result in similar or even identical models. In linear modelling the variance of the predictors is of much higher importance than the actual value. The continuous time series of the GRACE predictors based on the daily ITSG solution is thus equivalent to the monthly solutions, with the additional benefit of more robust model fits, and is thus preferred in this study.
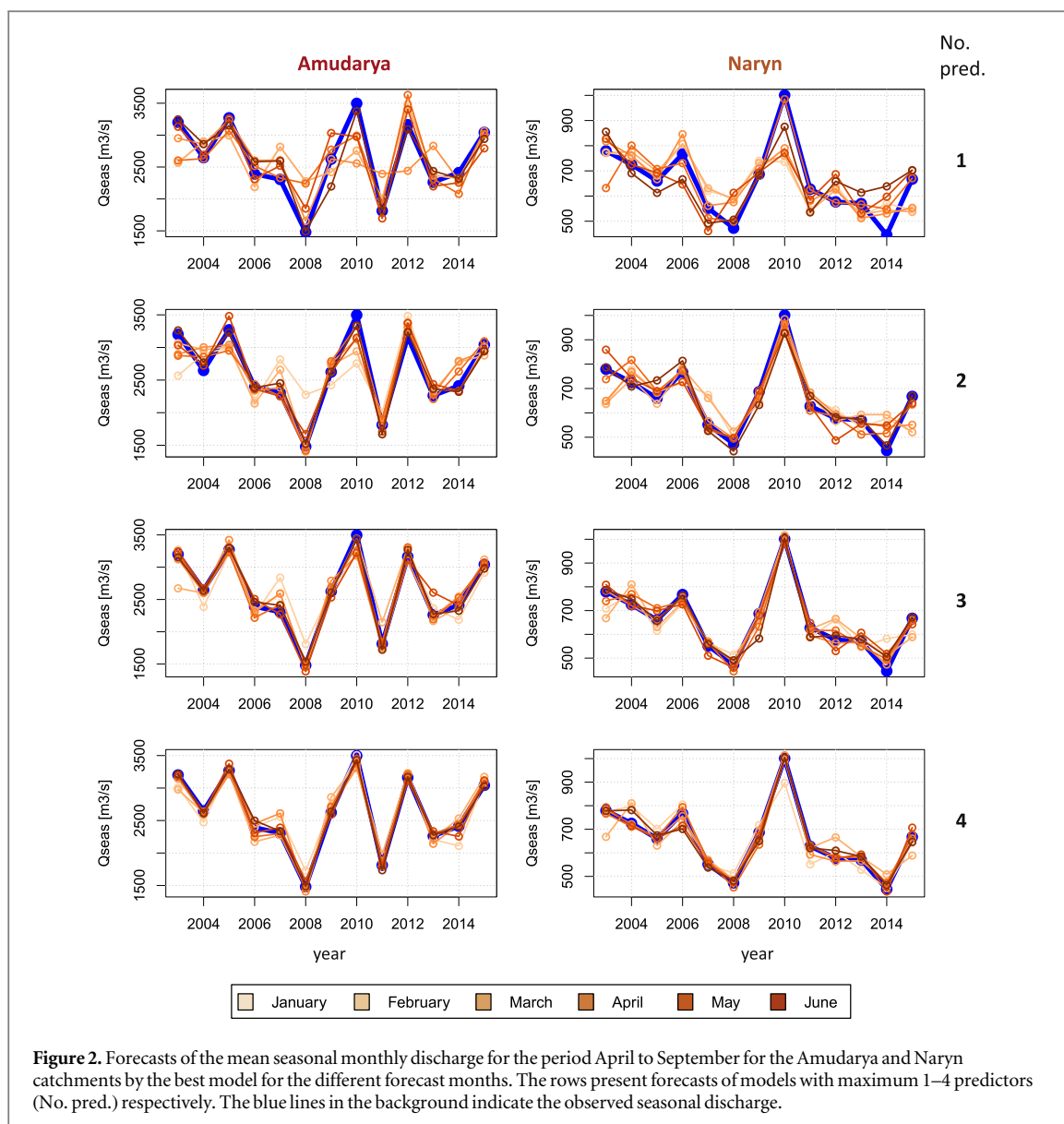
In order to properly map the water storage change over the accumulation period, differences between the monthly TWSA and the TWSA in September of the same season were calculated and used for the forecasts. By this procedure possible long term trends in the TWSA time series are eliminated and the TWSA signals of the particular accumulation periods are isolated. The time period of the analysis was 2003–2015 according to the availability of GRACE TWSA and discharge data.

3. Methods

Statistical forecast models for seasonal discharge were derived as in Apel *et al* (2018), based on station data (Precipitation P, air temperature T, antecedent discharge Q), satellite-based snow cover data (SC) (Gafurov *et al* 2016), and composites of P, T and SC. The core of the method is a Multiple Linear Regression (MLR), accompanied by a Leave-One-Out Cross Validation (LOOCV), testing all allowed combinations of predictors. The best performing 20 models were chosen according to the minimum of the Predictive Residual Mean of Squares (PREMS) between predicted and observed seasonal river discharge (mean monthly discharge April - September) of the LOOCV, whereby only models with overall significance at $p = 0.1$ and all predictors significant at $p = 0.1$ were selected. The predictors were grouped into 8 classes: T predictors, P predictors, SC predictors, Q predictors, and 4 composite predictors, combining T, P, and SC. Each model was allowed 1 predictor per group only, up to a maximum of 4 predictors (Apel *et al* 2018).

In order to exploit the specific value of TWSA as predictor, the maximum number of predictors per model was iteratively fixed to 1, 2, 3, and 4 predictors, respectively. This allows the investigation of TWSA as predictor in more detail, because the predictive power of TWSA alone and in combination with other predictors can be isolated. For comparison, the same procedure was applied to derive models excluding TWSA as predictor.

The performance of the models is reported as adjusted R^2 values of the best LOOCV performing model, and as mean adjusted R^2 of the best 20 models in the LOOCV. Additionally, the importance of the predictors was extracted as portions of the overall explained variance (adjusted R^2) in order to identify the contribution of the single predictors to the overall explained variance. For a better interpretation of the contribution of the predictors to model performance, the importance of the predictors was averaged over the set of best 20 models.

Figure 2. Forecasts of the mean seasonal monthly discharge for the period April to September for the Amudarya and Naryn catchments by the best model for the different forecast months. The rows present forecasts of models with maximum 1–4 predictors (No. pred.) respectively. The blue lines in the background indicate the observed seasonal discharge.

By this procedure potential spurious results in terms of predictor importance originating from a good model fit by chance can be avoided.

4. General results

Figure 2 illustrates the seasonal discharge forecasts for both Amudarya and Naryn by the best LOOCV model for the different forecast months. In general, the forecast improve with later forecast months (i.e. decreasing lead time) and with higher numbers of predictors, up to almost perfect forecasts for the late forecast months and 3–4 predictors. This high performance of the models is exceptional for seasonal forecasts, as already pointed out in Apel *et al* (2018) for the results without using GRACE data. Moreover, also the forecasts with only a single predictor provide reasonable forecasts of the seasonal discharge. The forecasts map the observed inter-annual variability of seasonal discharge quite well, even with 2–3 months lead time (cf figure 2). The visual impression given by figure 2 is substantiated by statistical model performance values (table 1). The mean performance follows the trend of the performance of the best model, thus indicating robust regressions with different predictor combinations. The performance is similarly high for both catchments for all lead times and predictor numbers when TWSA is included in the predictor set. Without TWSA, the performance of the Amudarya forecasts with 2–3 months lead time (January/February) is notably lower. For the Naryn basin, on the contrary, there is generally no marked improvement of the performance when TWSA is included.

The mean absolute error (MAE) and the root mean squared error (RMSE) follow the same trend as the adjusted R^2 (figure 3), with both measures being less than 10% of the mean seasonal discharge averaged over the

Table 1. Performance of the best LOOCV forecast models of seasonal discharge for the Amudarya and Naryn catchments for all forecast months, different numbers of predictors, and with and without GRACE TWSA as predictor. Performance is reported as adjusted R^2-values. The '*best*' columns show the performance of the single best LOOCV model, while '*mean*' indicates the mean adjusted R^2 of the best 20 LOOCV models. The symbols below the performance values indicate the significance of the model fits.
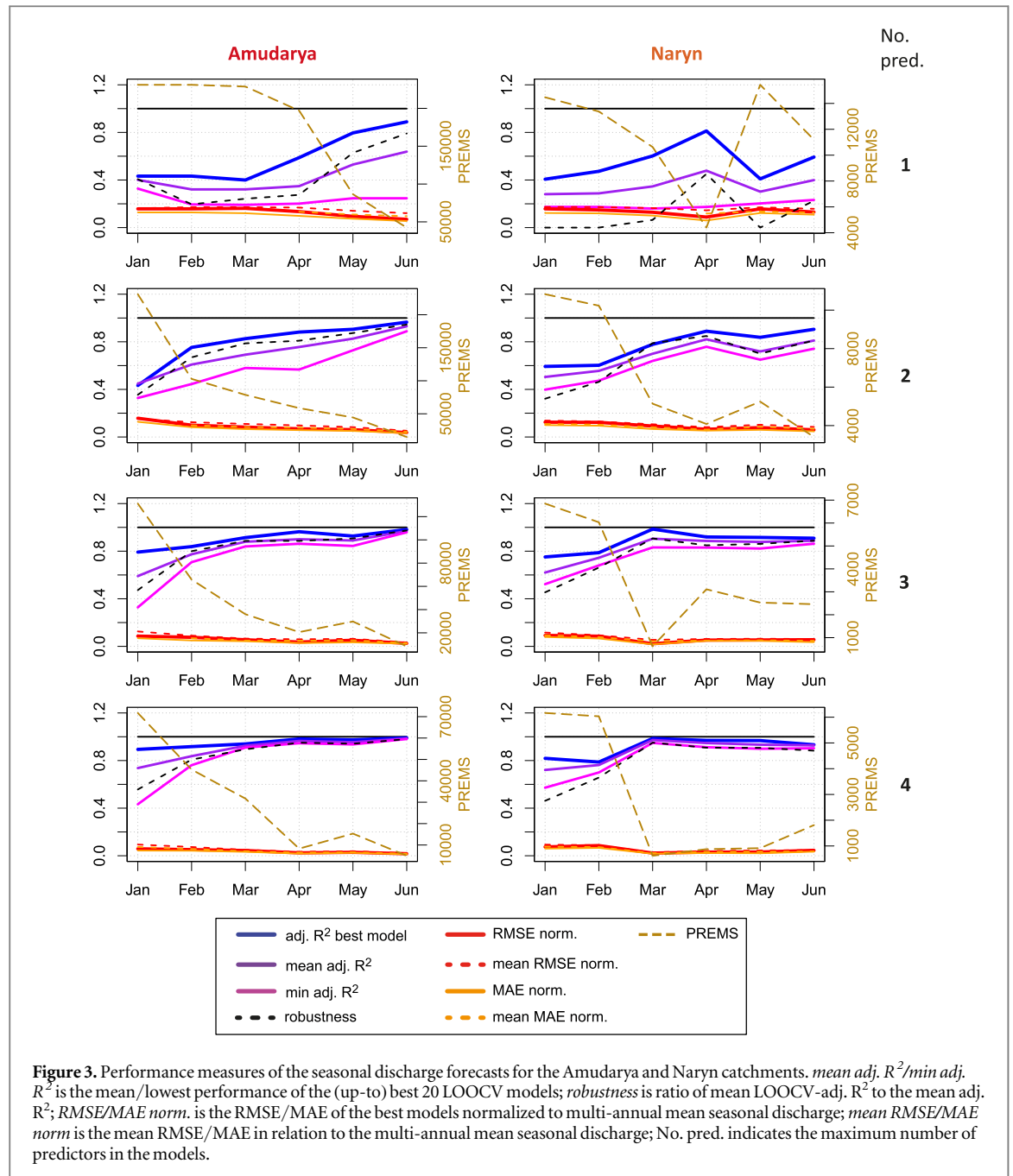
Amudarya		January		February		March		April		May		June	
predictors		best	*mean*	best	*mean*	best	*mean*	best	*mean*	best	*mean*	best	*mean*
1	with GRACE	0.433 +++	0.401 ++	0.433 +++	0.321 +	0.400 +++	0.321 +	0.588 +++	0.348 +	0.796 +++	0.530 ++	0.889 +++	0.638 ++
1	without GRACE	0.426 +++	0.386 ++	0.426 +++	0.281 +	0.400 ++	0.298 +	0.451 +++	0.256 +	0.721 +++	0.437	0.889 +++	0.521
2	with GRACE	0.433 +++	0.448 ++	0.753 +++	0.609 ++	0.826 +++	0.692 +++	0.882 +++	0.756 +++	0.905 +++	0.827 +++	0.965 +++	0.930 +++
2	without GRACE	0.426 +++	0.386 ++	0.456 ++	0.336 ++	0.732 +++	0.513 ++	0.567 +++	0.477 +++	0.794 +++	0.662 +++	0.948 +++	0.888 +++
3	with GRACE	0.792 +++	0.590 ++	0.838 +++	0.774 +++	0.914 +++	0.879 +++	0.963 +++	0.900 +++	0.927 +++	0.890 +++	0.982 +++	0.969 +++
3	without GRACE	0.426 +++	0.386 ++	0.456 ++	0.366 ++	0.826 +++	0.724 +++	0.802 +++	0.702 +++	0.927 +++	0.824 +++	0.964 +++	0.953 +++
4	with GRACE	0.894 +++	0.737 ++	0.917 +++	0.837 +++	0.940 +++	0.924 +++	0.981 +++	0.965 +++	0.974 +++	0.954 +++	0.991 +++	0.985 +++
4	without GRACE	0.426 +++	0.386 ++	0.456 ++	0.366 +	0.862 +++	0.796 +++	0.918 +++	0.843 +++	0.958 +++	0.905 +++	0.967 +++	0.967 +++
Naryn													
1	with GRACE	0.407 ++	0.281 +	0.473 +++	0.288 +	0.602 +++	0.346 +	0.812 +++	0.478 +	0.409 +++	0.302 +	0.593 +++	0.399 +
1	without GRACE	0.407 ++	0.281 +	0.473 +++	0.288 +	0.602 +++	0.355 +	0.812 +++	0.478 +	0.409 +++	0.297 +	0.498 +++	0.357 +
2	with GRACE	0.593 +++	0.505 ++	0.602 +++	0.556 ++	0.779 +++	0.700 +++	0.889 +++	0.821 +++	0.837 +++	0.719 +++	0.905 +++	0.811 +++
2	without GRACE	0.593 +++	0.501 ++	0.602 +++	0.555 ++	0.779 +++	0.698 +++	0.889 +++	0.823 +++	0.837 +++	0.719 +++	0.905 +++	0.798 +++
3	with GRACE	0.751 +++	0.620 ++	0.787 +++	0.744 +++	0.985 +++	0.905 +++	0.919 +++	0.886 +++	0.915 +++	0.878 +++	0.909 +++	0.888 +++
3	without GRACE	0.751 +++	0.612 ++	0.787 +++	0.740 +++	0.985 +++	0.904 +++	0.919 +++	0.887 +++	0.915 +++	0.885 +++	0.905 +++	0.871 +++
4	with GRACE	0.819 +++	0.721 ++	0.787 +++	0.762 +++	0.985 +++	0.973 +++	0.969 +++	0.947 +++	0.968 +++	0.933 +++	0.933 +++	0.928 +++
4	without GRACE	0.751 +++	0.703 ++	0.787 +++	0.756 +++	0.985 +++	0.973 +++	0.957 +++	0.942 +++	0.941 +++	0.933 +++	0.940 +++	0.899 +++

Significance level p: +++ = 0.01, ++ = 0.05, + = 0.1, o = 0.2, -- = >0.2; for *mean* the lowest significance of the model set is used.

investigation period 2003–2015. This holds true for the best model, but also on average for the model sets of 20 models. This indicates high reliability of the forecasts. The robustness of the forecasts is investigated by setting the LOOCV residuals in relation to the residuals of the model fitted to the complete time series (black dashed line in figure 3). This shows that the robustness of the January and February forecasts is comparatively moderate, but from March onwards very high.
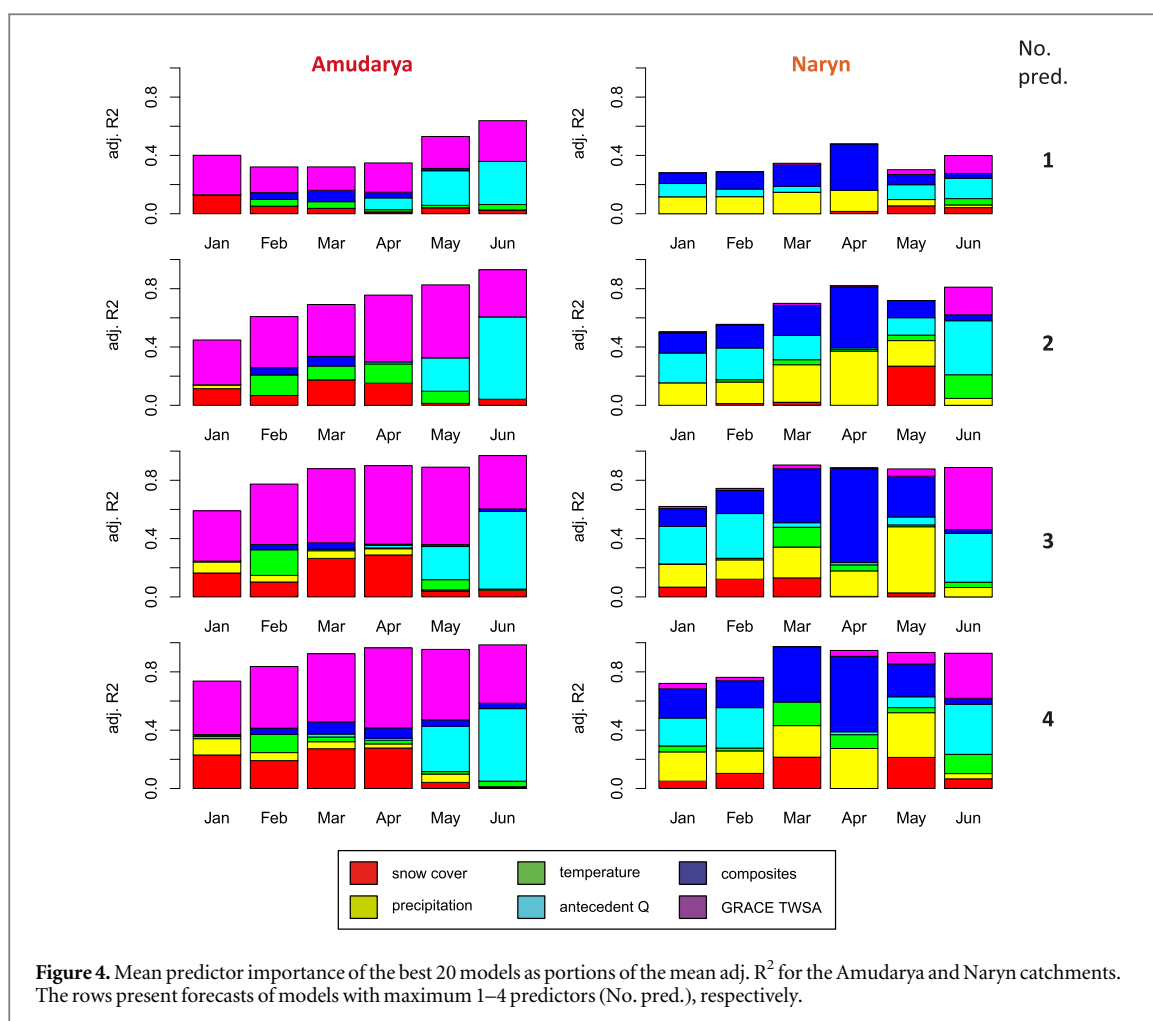
5. The role of GRACE TWSA as predictor

In general, the results in terms of forecasts performance follow the findings of Apel *et al* (2018), but with some distinct differences. The comparison of the forecasts with and without GRACE TWSA shows substantial improvements for Amudarya for almost all lead times and number of predictors, while for Naryn the

Figure 3. Performance measures of the seasonal discharge forecasts for the Amudarya and Naryn catchments. *mean adj. R²/min adj. R²* is the mean/lowest performance of the (up-to) best 20 LOOCV models; *robustness* is ratio of mean LOOCV-adj. R² to the mean adj. R²; *RMSE/MAE norm.* is the RMSE/MAE of the best models normalized to multi-annual mean seasonal discharge; *mean RMSE/MAE norm* is the mean RMSE/MAE in relation to the multi-annual mean seasonal discharge; No. pred. indicates the maximum number of predictors in the models.

performance mostly remained identical. For Amudarya the smallest improvements are observed for the January to March forecasts with 1 predictor. However, for the most important forecasts for water management in April, improvements could be achieved even with a single predictor. This indicates that TWSA outperforms all other observables when used as the only predictor for Amudarya. If more than 1 predictor is used, the model performance increased substantially in Amudarya, particularly for the early forecasts. Although the later forecasts show a high performance even without using TWSA, an increase in performance is still achieved.

The quantification of predictor importance corroborates these findings (figure 4). For Amudarya, TWSA is responsible for the highest portion of the explained variance (= highest importance) for the forecast months January to April. This can be observed for models with a single and up to 4 predictors. For May and June, the antecedent discharge gains more importance and explains the highest share of variance, as already observed in Apel *et al* (2018). This is reasonable from a hydrological point of view, because snow melt already occurs in May and June and thus the observed discharge has a high indicative power for the overall seasonal discharge. This is expressed by a high correlation of sub-seasonal discharges spanning over shorter periods (e.g. May-September) to the discharge of the full summer period. In contrast to Amudarya, TWSA is of much lower importance in Naryn. It shows hardly any importance in the January to May forecasts. Only for the late June forecast, TWSA

Figure 4. Mean predictor importance of the best 20 models as portions of the mean adj. R^2 for the Amudarya and Naryn catchments. The rows present forecasts of models with maximum 1–4 predictors (No. pred.), respectively.

explains some variance, which is in line with the modest improvement of the mean performance for June reported in table 1.

The observed differences of the TWSA importance in the two catchments had to be expected due to the too small catchment size of Naryn for GRACE gravity signals, and can be explained by the spatial resolution of GRACE data. The Naryn catchment (about 50,000 km^2) is well below estimated limits of catchment sizes that GRACE is able to resolve (Vishwakarma *et al* 2018), given the noise in GRACE observations and signal deterioration due to subsequent filtering, causing not only the reduction of noise but also signal attenuation and leakage from neighboring areas. This effect could also explain the observed TWSA importance in the June forecasts in Naryn. Here, TWSA predictors show the difference between TWSA in May and September of the previous year, i.e. over the whole accumulation period where mass changes are likely to be better detectable by GRACE. This is particularly relevant for the Naryn catchment, which receives substantial precipitation—in contrast to Amudarya - also in late spring/early summer by westerly circulation patterns (Gerlitz *et al* 2016). Hence it can be argued that the additional precipitation received in May in Naryn provides the critical mass input that can be detected by GRACE in this catchment. However, in general the hypothesis that the TWSA signal for Naryn, which is influenced by the TWSA anomalies of the surrounding region, is actually representative for the water storage in Naryn, has to be rejected.

6. Conclusions

In the semi-arid study area of Central Asia, GRACE gravity based water storage anomalies TWSA showed high performance in forecasting river discharge in the vegetation period for the large Amudarya catchment. In this catchment, GRACE TWSA is the most important predictor. However, the combination of TWSA with other predictors even increased the importance of TWSA, along with the overall forecast performance. For the smaller Naryn catchment, which size is actually too small for GRACE, TWSA could not improve the forecasts based on climate data, snow cover and antecedent discharge alone. This finding supports the conclusion that TWSA actually maps the total water accumulation in the large Amudarya catchment, and is thus a meaningful predictor

for the subsequent seasonal discharge. Furthermore, it can be reasoned that TWSA might be used for forecasting seasonal discharge in other large regions with similar climatic and hydrological conditions, i.e., snow accumulation in winter and snow melt as most important factor for river discharge in summer, such as the Californian Sierra Nevada or the Chilean Andes, for instance.

The findings also suggest that direct observations of SWE would be an ideal predictor for seasonal forecasts, especially for smaller catchments. However, ground based SWE observations, particularly continuous time series spanning decades, are very scarce and limited to a few areas worldwide. This impairs a generalization of such an approach. On the other hand, state-of-the-art continental-scale, satellite observation-based SWE products, such as the Copernicus Global Land service SWE product (https://land.copernicus.eu/global/products/swe) based on the retrieval algorithm of Takala *et al* (2011), are restricted to non-mountainous areas. These products thus cannot deliver so far useful information for Central Asia and for similar regions where seasonal discharge is dominated by snow accumulation and melt in the mountains.

Dedicated post-processing of GRACE data for small basins of interest as proposed, e.g.,. by Longuevergne *et al* (2010) or Khaki *et al* (2018), may reduce the magnitude of filter and leakage errors in the final storage time series for these areas, but such an approach will be of limited use for the present forecasting study due to the gaps in the time series, thus suffering the same constraints in linear modelling as the standard monthly solutions. Expectations, albeit of limited scope, of a slightly higher spatial resolution of GRACE-FollowOn data based on its laser ranging interferometer (Flechtner *et al* 2016) may open pathways for transferring the proposed seasonal discharge forecasting method with GRACE predictors to smaller catchments. It should be noted that the same argument applies for the use of multiplicative time-invariant scaling factors applied in post-processing of GRACE data to restore part of the signal amplitudes that have spuriously been lost during GRACE filtering operations before (e.g., Landerer and Swenson (2012)) is expected to have no effect on the forecast results. This sort of linear bias correction will result in GRACE-based predictors that are highly correlated to the original ones, and thus, as explained above, will have no effect on the linear forecasting model applied in the present study.

For an operational use of GRACE TWSA in seasonal forecasts, the data need to be provided in near-real time (NRT). Given the recent development of a NRT service for GRACE products within the European Gravity Service for Improved Emergency Management EGSIEM (Jäggi *et al* (2019)), such data are expected to be soon available based on the GRACE-Follow On mission, which was launched in May 2018. Our study thus paves the road for a novel application of satellite gravity data in hydrological forecasting, possibly in several regions worldwide, including chronically data scarce areas. Given the specific findings for Central Asia, GRACE-FO could be included in operational forecasts procedures by Central Asian water authorities for the summer season 2019.

Acknowledgments and Data

ORCID iDs

Heiko Apel ⓘ https://orcid.org/0000-0002-8852-652X
Abror Gafurov ⓘ https://orcid.org/0000-0003-0337-465X

References

Aizen V B, Aizen E M and Melack J M 1996 Precipitation, melt and runoff in the northern Tien Shan *J. Hydrol.* **186** 229–51

Apel H, Abdykerimova Z, Agalhanova M, Baimaganbetov A, Gavrilenko N, Gerlitz L, Kalashnikova O, Unger-Shayesteh K, Vorogushyn S and Gafurov A 2018 Statistical forecast of seasonal discharge in Central Asia using observational records: development of a generic linear modelling tool for operational water resource management *Hydrol. Earth Syst. Sci.* **22** 2225–54

Flechtner F, Neumayer K H, Dahle C, Dobslaw H, Fagiolini E, Raimondo J C and Guntner A 2016 What can be expected from the GRACE-FO laser ranging interferometer for Earth Science applications? *Surv. Geophys.* **37** 453–70

Gafurov A, Lüdtke S, Unger-Shayesteh K, Vorogushyn S, Schöne T, Schmidt S, Kalashnikova O and Merz B 2016 MODSNOW-Tool: an operational tool for daily snow cover monitoring using MODIS data *Environmental Earth Sciences* **75** 1–15

Gerlitz L, Vorogushyn S, Apel H, Gafurov A, Unger-Shayesteh K and Merz B 2016 A statistically based seasonal precipitation forecast model with automatic predictor selection and its application to central and south Asia *Hydrol. Earth Syst. Sci.* **20** 4605–23

Jäggi A *et al* 2019 European gravity service for improved emergency management (EGSIEM) - from concept to implementation *Unpublished to Geophysical Journal International (minor revision)*

Khaki M, Forootan E, Kuhn M, Awange J, Longuevergne L and Wada Y 2018 Efficient basin scale filtering of GRACE satellite products *Remote Sens. Environ.* **204** 76–93

Kurtenbach E, Eicker A, Mayer-Gurr T, Holschneider M, Hayn M, Fuhrmann M and Kusche J 2012 Improved daily GRACE gravity field solutions using a Kalman smoother *J. Geodyn.* **59-60** 39–48

Landerer F W and Swenson S C 2012 Accuracy of scaled GRACE terrestrial water storage estimates *Water Resour. Res.* **48**

Longuevergne L, Scanlon B R and Wilson C R 2010 GRACE hydrological estimates for small basins: evaluating processing approaches on the high plains aquifer, USA *Water Resour. Res.* **46**

Sorg A, Bolch T, Stoffel M, Solomina O and Beniston M 2012 Climate change impacts on glaciers and runoff in Tien Shan (Central Asia) *Nat. Clim. Change* **2** 725–31

Takala M, Luojus K, Pulliainen J, Derksen C, Lemmetyinen J, Kärnä J-P, Koskinen J and Bojkov B 2011 Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements *Remote Sens. Environ.* **115** 3517–29

Tapley B D, Bettadpur S, Ries J C, Thompson P F and Watkins M M 2004 GRACE measurements of mass variability in the Earth system *Science* **305** 503–5

Vishwakarma B, Devaraju B and Sneeuw N 2018 What is the spatial resolution of grace satellite products for hydrology? *Remote Sensing* **10** 852

Viviroli D, Durr H H, Messerli B, Meybeck M and Weingartner R 2007 Mountains of the world, water towers for humanity: typology, mapping, and global significance *Water Resour. Res.* **43**

Zhang L J, Dobslaw H and Thomas M 2016 Globally gridded terrestrial water storage variations from GRACE satellite gravimetry for hydrometeorological applications *Geophys. J. Int.* **206** 368–78