ORIGINAL PAPER

Prediction of the Caspian Sea level using ECMWF seasonal forecasts and reanalysis

K. Arpe · S. A. G. Leroy · F. Wetterhall · V. Khan ·
S. Hagemann · H. Lahijani

Abstract The hydrological budget of the Caspian Sea (CS) is investigated using the European Centre for Medium-Range Weather Forecasts interim reanalysis (ERAi) and seasonal forecast (FCST) data with the aim of predicting the Caspian Sea Level (CSL) some months ahead. Precipitation and evaporation are used. After precipitation events over the Volga River, the discharge (Volga River discharge (VRD)) follows with delays, which are parameterized. The components of the water budget from ERAi and FCSTs are integrated to obtain time series of the CSL. Observations of the CSL and the VRD are used for comparison and tuning. The quality of ERAi data is sufficiently good to calculate the time variability of the CSL with a satisfactory accuracy. Already the storage of water within the Volga Basin allows forecasts of the CSL a few months ahead, and using the FCSTs of precipitation improves the CSL forecasts. The evaporation in the seasonal forecasts is deficient due to unrealistic sea surface temperatures over the CS. Impacts of different water budget terms on the CSL variability are shown by a variety of validation tools. The importance of precipitation anomalies over the catchment of the Volga River is confirmed, but also impacts from the two southern rivers (Sefidrud and Kura River) and the evaporation over the CS become obvious for some periods. When pushing the FCSTs beyond the limits of the seasonal FCSTs to 1 year, considerable forecast skill can still be found. Validating only FCSTs by the present approach, which show the same trend as one based on a statistical method, significantly enhances the skill scores.
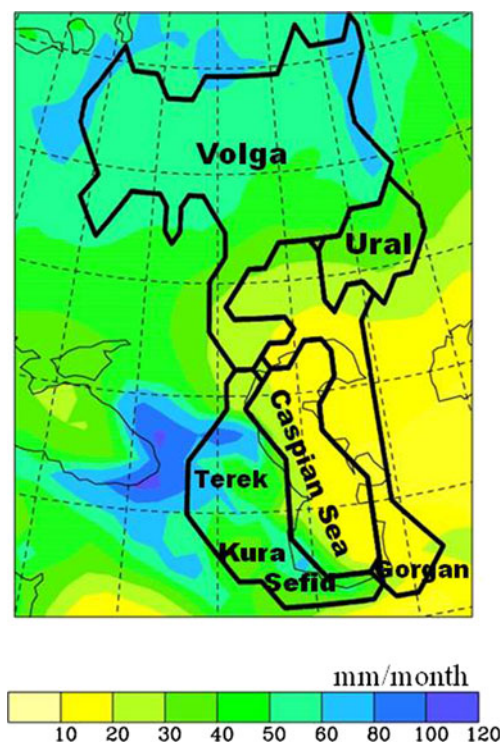
K. Arpe · S. Hagemann
Max Plank Institute for Meteorology, Hamburg, Germany

K. Arpe · S. A. G. Leroy (✉)
Institute for the Environment, Brunel University,
Kingston Lane, Uxbridge,
UB8 3PH London, UK
e-mail: suzanne.leroy@brunel.ac.uk

F. Wetterhall
European Centre for Medium-Range Weather Forecasts, Reading, UK

V. Khan
Hydrometeorological Research Center of the Russian Federation,
Moscow, Russia

H. Lahijani
Iranian National Institute for Oceanography (INIO), Tehran, Iran

1 Introduction

The Caspian Sea (CS) (36–47° N, 47–54° E) is a closed basin without any outlet. Its sea level lies below the mean sea level of the ocean and has varied between −25 and −29 m in the last 150 years. Water level changes occurred 100 times faster in comparison to global sea level changes over the last century. The main water source is the Volga River whose catchment area reaches well into the humid mid-latitudes. Rivers like the Ural, Kura, Terek, and Sefidrud have to be considered as well (Fig. 1). The water inflow is balanced by evaporation over the CS itself including the Kara Bogaz Gol (KBG), a bay on the eastern shore of the CS, which is connected with the main CS only by a channel (Giralt et al., 2003). The eastern coast of the CS is mostly desert with very low precipitation (Fig. 1).

The Caspian Sea Level (CSL) variability and the water budget of the CS have been investigated in previous studies, e.g., Rodionov (1994), Golitsyn (1995), Arpe et al. (1999, 2000), and Arpe and Leroy (2007). Because of the large socio-economic impacts of CSL changes, several attempts at forecasting them have been carried out. A CSL forecasting method based on water balance approximation calculation was developed in the early 1940s by Kalinin (1941). This method was successfully implemented at the Hydrometeorological Center of Russia (HMRC). Modified versions of the Kalinin method have been used until now to issue operational monthly CSL forecasts with up to 1 year lead time.

Fig. 1 Different catchment areas in *heavy black lines* and annual mean precipitation in *thin lines*. Contours at 10, 20, 30, 40, 50, 60, and 80 mm/month; *light shading* <20 mm/month, *darker shading* >50 mm/month

Since the 1960s, the process of artificial regulation of the inflow of the main rivers feeding the Caspian Sea has intensified. As a consequence, the natural river input to the water budget of the CSL has changed. Therefore, at HMRC, the operational CSL forecasting method was modified and adjusted to the new conditions of regulated river inflow (e.g., Sheremetevskaya and Smirnova 1966; Kalinin et al. 1968). Several other practical approaches were proposed (e.g., Smirnova 1974; Lunyakova 1991; Abuzyarov and Krasyuk 2003; Abuzyarov 2009) to calculate components of the water balance on the basis of information taken from hydrometeorological data, river inflow to the CS, effective evaporation (evaporation minus precipitation), discharge of waters to the KGB, estimated undergroundwater inflow, and volume expansion of the sea due to changes in water density. The last version of operational CSL annual forecast (Abuzyarov, 2009) was introduced in 2008. The skill score of CSL forecasts calculated by the improved method has increased by 6 %. Operational CSL forecasts along with the overview of the state of the CSL and water balance for the previous year are published in a special bulletin once a year at the beginning of May at the HMRC and disseminated to interested organizations, with updates in June or July. Monthly and annual averages of the sea level at seven shore locations of the CS are disseminated (Makhachkala, Baku, Zhiloy, Fort-Shevchenko, Kuuly-Mayak, Turkmenbashi, and KBG).

Meshcherskaya et al. (1997) described a method of forecasting the CSL based on teleconnections between global circulation indices, like NAO or ENSO, and the CSL. This became an operational forecast product at the Main Geophysical Observatory in St. Petersburg (MGO).

Forecasts have been developed, which are based on hidden cycles in the CSL curve. In this method, authors tried to find frequencies and wavelengths in the CSL curve and to extend them into the future as a forecast (Doganovskii 1982; Meshcherskaya and Aleksandrova 1993). They ignored elements of water balance that affect the CSL curve. Moreover, the mathematical character of the curve (deterministic, random, or something between the two) is not clear. Many authors have applied stochastic nonlinear equations for the CSL predictions. Numerical analysis through software was used to solve the equations (Vaziri 1997; Makarenko et al. 2004; Mozraji et al. 2010). The reliability of the prediction depended on the prediction of the main elements of the water balance. Long-term forecasting of the water balance needs reliable forecasts of the future ocean and atmospheric circulations.

Recently, Arpe et al. (2012) (referred below as AR2012) investigated the CSL change using data from the European Centre for Medium-Range Weather Forecasts (ECMWF), which provides reanalysis data including quantities, which are difficult to measure, such as precipitation (P) and evaporation (E) over continents and oceans. AR2012 suggested scope for predicting the CSL a few months ahead. In the present study, the quality of such forecasts is investigated. However, it was decided to use not only the ECMWF interim reanalysis (ERAi) data but also ECMWF seasonal forecasts, which are carried out every month with a lead time of 7 months. The data used are described in Section 2, and the methods are described in Section 3. First hindcast results are shown in Section 4; especially the choice of components of the hydrological budget, which should be used for a forecast, is a main issue. This investigation is assumed to be a pilot study and is intended to persuade operational centres like ECMWF or its member states to provide such forecasts operationally.

2 Observational, analysis, and forecast data

The recent CSL observations of altimetry observations by the satellites TOPEX, POSEIDON, and JASON are taken from USDA1 (2011). These data are available from 1993 onwards; before that date, gauge data from different sources are used. Lebedev and Kostianoy (2006, 2008) compared gauge observations around the CS with altimetry observations from satellite. They point to many uncertainties in both data sets, e.g., large differences in the CSL in different basins and the existence of surges, which can be as large as 3 m in the northern basin, which makes the definition of a mean

CSL very difficult. Their data were made available to this study in order to compare them with satellite data.

The gauge observations at the different sites are adjusted to each other for overlapping periods after removing the mean annual cycle for each data set separately. The means of the single gauge observations after ignoring the highest and lowest values agree quite well with the satellite estimate (Appendix 1), though from 2002 onwards a bias occurred, which is discussed below. The satellite observations have recently been revised (USDA2 2013), which do not show this bias anymore.

The precipitation (P) and evaporation (E) over continents and seas are available from the ECMWF interim reanalysis (ERAi). P and E in this data set are prognostic output from the analysis (ECMWF 2011; Dee et al. 2011). ERAi use observational data from a large range of sources; however, precipitation gauge observations are not used. They are produced twice daily, at 00UTC and 12UTC by 12-h forecasts. The ERAi analyses are carried out with a T255 resolution (0.703° grid) from 1979 up to near real time. The calculations in this study are carried out at T106 resolution (1.125° grid).

With such a coarse resolution it is not possible to separate the CS from the KBG. Both are therefore treated here as one unit and referred to below only as the CS. Section 3 describes a simple method to parameterize the delay of the water flow in the Volga River due to ice and snow. This can also be improved by incorporating surface temperatures, and this is indeed done if one uses the run-off data, which are provided in the ERAi dataset. We found, however, a few problems with the run-off data (Appendix 2), which prevented us using them. ECMWF also provides seasonal (7 months) forecasts every month. Here, their ensemble means are used (referred below as FCST). Data from version 3 (Anderson et al. 2007) and version 4 (Molteni et al., 2011) are available. As version 4 uses the more recent model, it is expected that, in general, the forecasts with version 4 should be superior to those using version 3. We investigated the precipitation over the Volga Basin because it is the main driver for the CSL variability and we found better results with version 4 only for winter, while for summer, version 3 is superior. For summer precipitation, the 3-month forecast with version 3 is as good as the 2-month forecast with version 4 (Appendix 2). The forecasts of evaporation over the CS are shown in Appendix 2 to be unrealistic and will not be used. A main problem with the evaporation over the CS is that the FCSTs do not have an appropriate sea surface temperature (SST), only a climatological mean.

The precipitation over continents up to 2010 is also available from the Global Precipitation Climate Center (GPCC), see Schneider et al. (2011) and GPCC (2011). These precipitation estimates are based solely on gauge observations and AR2012 has shown that they are mostly very similar to those by ERAi for the period investigated, i.e., 1993–2009. Examples are also shown in Fig. 15. As the estimates are based on completely different data and using completely different methods, their similarity suggests that they are both good.

Monthly mean Volga River discharge (VRD) data are taken from Dümenil Gates et al. (2000), which were complemented for the recent period with data provided by the Hydrological Forecasts Department of HMRC. Furthermore, the river-discharge data of the Kura and Ural rivers, and the Sefidrud and Gorganrud were collected.

3 Methods

For calculating the water budget of the CS, the input data were those available from ERAi and FCST. AR2012 have shown that over continents the evaporation is most likely to be overestimated, while over the CS, the evaporation is underestimated in this dataset. Therefore, precipitation minus evaporation (P-E) was not calculated directly, but the evaporation was first reduced by 5 % over the continents and increased by 20 % over the CS to get the mean values given by AR2012 into better balance. P-E over the CS is affecting the CSL directly, while P-E over the Volga Basin affect the CSL with some delays (see next section and AR2012). Other rivers entering the CS, the largest being the Ural, Kura, Sefidrud, and Gorganrud, could have an impact on the CSL as well. Their impact on the CS is investigated in Section 4.3. To reduce the uncertainty due to errors of absolute values, it is mostly the results of anomalies (taking away the mean annual cycle) that are shown and used for validation. The inclusion of the annual cycles in the data would have reduced the clarity of the plots and enhanced the anomaly correlations artificially.

### 3.1 Delay of the water outflow from the Volga River

A time delay between precipitation anomalies over the Volga Basin and the discharge at the mouth of the river occurs due to storage in the ground, during winter as snow and ice and due to the long travel time down the river (AR2012). A further delay results from many dams, which store the water not only for generating electricity but also for irrigation and consumption by people and industries. Generally, deficient precipitation has less obvious impacts on the VRD than excessive precipitation, probably due to the existence of dams (AR2012). These dams let the water pass according to the demand for electricity in the country. Normally, the dams are full in July after the snow melt and with the increase in precipitation during summer. After that, the VRD responds to the precipitation more directly. Arpe et al. (1999, 2000), therefore, used annual means from July to June the following year in their investigation to reduce the effect of the dam storage.
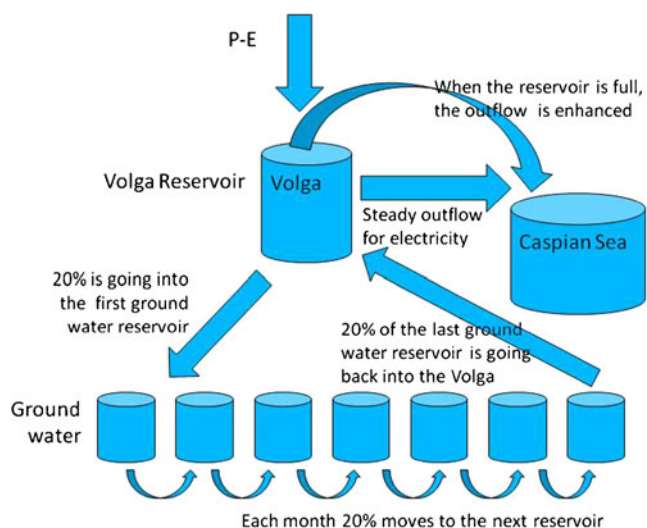
Fig. 2 Schematic of the model to parameterize the delay of the water in the Volga basin

A simple model tries to simulate the storage by ice and snow in winter; it uses P-E values of up to 4 of the preceding months. This input is fed into the Volga reservoirs (Fig. 2) from which a minimum nearly constant amount of water is released due to the need for electricity. When the amount of water in the reservoirs becomes too low, this released water is reduced to avoid negative amounts in the storage. If the amount of water in the storage exceeds a maximum threshold, more water is released for discharge until the storage is back to its maximum. In addition, a delay due to storage as groundwater is parameterized. The model is tuned by comparing the model output with the observed VRD data.

The effect of this delay can already be shown by comparing the Volga Basin P-E mean annual cycle with that of the observed and simulated VRD (Fig. 3). The original P-E (P-Ei) has a broad maximum from October to February, while the observed VRD (VRob) has a sharp maximum in May.
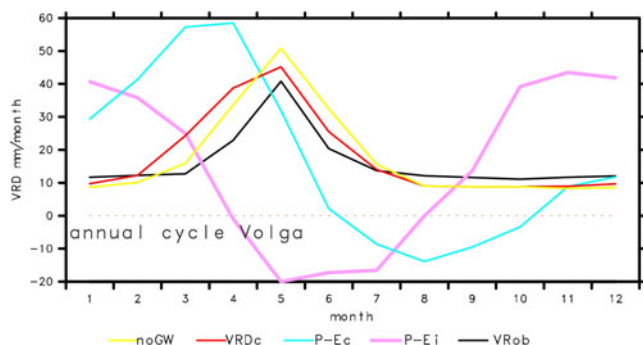
Fig. 3 Annual cycle of the Volga River discharge as observed (*VRob*) compared with P-E over the Volga Basin from ERAi (*P-Ei*) and after other calculations. A delay due to snow and ice for ERAi (*P-Ec*) was applied. After that, the effects of reservoirs and groundwater (*VRDc*) were parameterized. noGW refers to the delay without groundwater parameterization. VRD data are converted into mm/month over the catchment area

Thus, a clear delay of several months occurs mainly due to storage on the ground by ice and snow. For summer, the delay is shorter (AR2012).

### 3.2 Calculating the CSL from the water budget of the CS

The water budget for the CS is calculated from the VRD and P-E over the CS itself. Further inputs are from other rivers, especially the Ural and those on the southern and western coasts (Sefidrud and Kura River respectively), and perhaps to a smaller extent from underground inflow. The latter is not calculated explicitly but is being taken care of on average by investigating anomalies only. The VRD plus P-E over the CS and south-western coast are added to an initial CSL.

The FCST data are treated in the same way as the ERAi data. For calculating the VRD, P-E values for up to 4 months before the current months are needed. For a forecast, the ERAi data of the months before the initial date of the FCST are used and the FCST data for the months after that. At the beginning of each forecast, the CSL data and the storage of water in the Volga Basin are taken from the simulation driven by ERAi. Observed CSL data are only used once at the beginning of the integration of the water budget, i.e., 1979. Later, only P-E values from ERAi or FCST enter the calculations. The presentation and validation are mostly restricted to the period after 1986 because the storage terms need to be spun-up. The observed VRD and CSL data are used only for validation and for tuning.

4 Results

4.1 Estimating the Volga discharge from P-E

The first test of the parameterization of the delay of the water flow in the Volga Basin is to analyze the mean annual cycle of the VRD. The annual cycles of the Volga discharge (Fig. 3), as observed (VRob), is compared with the direct P-E from ERAi over the Volga Basin (P-Ei) and P-E after applying some delay due to snow and ice, water going into the soil and running down the river using ERAi data (P-Ec). The delay shifts the annual cycle of the direct P-E towards the one of the observed Volga discharge by several months, but a difference of 2 to 3 months still remains (Fig. 3). A main problem is the storage of water in the reservoirs along the Kama-Volga cascade. The water is let through according to the demand for electricity, only when the reservoirs are full, excessive rain will lead to an immediate discharge of water into the CS. A method to simulate this has been described in Section 3.1, and it shifts the mean annual cycle to a nearly perfect position (Fig. 3, noGW). In Fig. 4, it is shown that applying these methods leads to
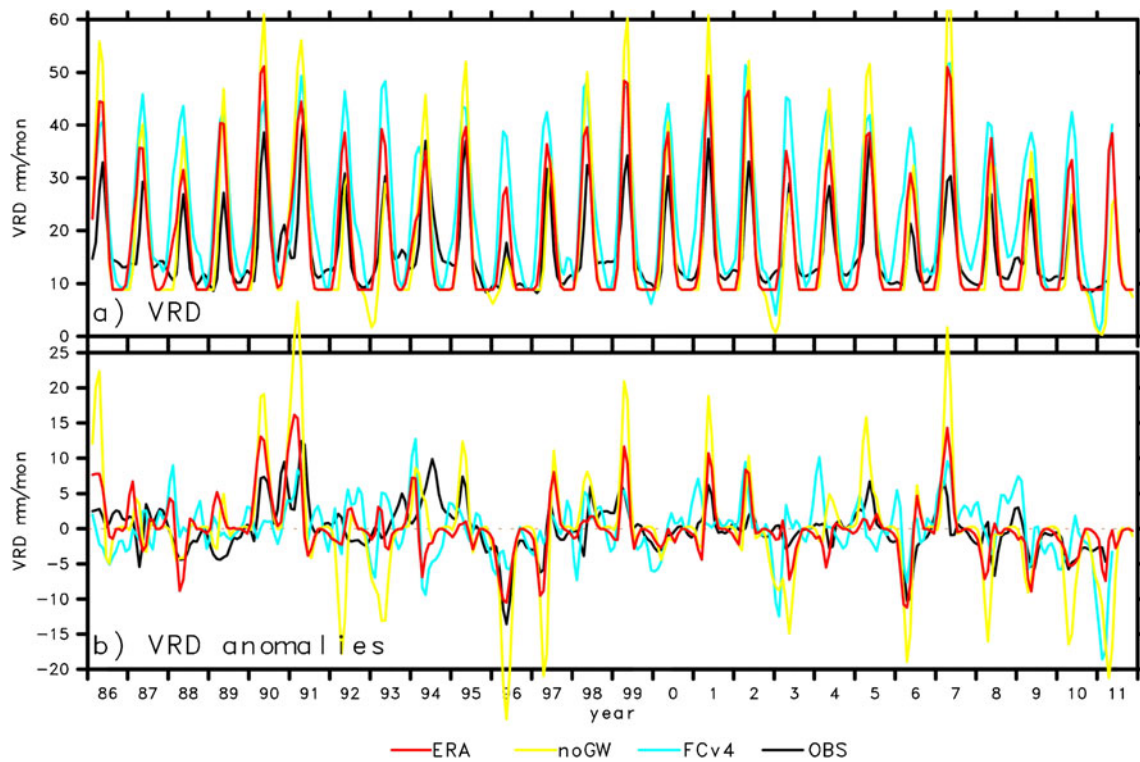
Fig. 4 **a** Monthly VRD as observed (*OBS*) or estimated from P-E in ERAi (*ERAi*), from P-E in ERAi but without using the groundwater parameterization (*noGW*) and FCST version 4 (*FCv4*). VRD data are converted into mm/month over the catchment area. **b** The same but with a mean annual cycle removed. All data sets are smoothed with a 1-2-1 filter

stronger extremes compared to that of the observed VRD. A groundwater reservoir with a weak transfer to the upper reservoir is therefore introduced. In Fig. 3, the mean annual cycle after applying this method becomes broader than the observed one, but the groundwater parameterization leads to better extremes in the time series (Fig. 4, ERAi).

The main features of the observed VRD variability are well reproduced by ERAi and FCST (Fig. 4). The variation of winter observed values (those lower values around 10 mm month^{-1} in Fig. 4a) are mainly influenced by the water management but are nevertheless partly reproduced by this method. Figure 4a shows as well that without the groundwater parameterization, the Volga River would occasionally run dry, which is unrealistic.

The anomaly correlation coefficients and root mean square (RMS) errors between OBS and ERAi or FCST are given in Table 1. The lower correlation coefficients for the forecasts are expected from viewing Fig. 4b. The slightly higher correlations in ERAi when not using the groundwater parameterization are probably due to the better annual cycle and because the anomaly correlation coefficients are not sensitive to exaggerated extremes but the parameterization of groundwater clearly reduces the RMS errors.

4.2 Integrating the water budget components of the CS to calculate the CSL

The sum of all components of the water budget for the CS needs to be integrated in time to get the variability of the

Table 1 Anomaly correlations and RMS (root mean square) errors (mm/month) of monthly VRD between observations and estimates by ERAi or FCST

Anomaly correlations/RMS errors	With annual cycle, no smoothing	No annual cycle, no smoothing	With annual cycle, with smoothing	No annual cycle, with smoothing
OBS-ERAi	0.79/8.2	0.35/5.5	0.86/6.3	0.54/3.7
OBS-ERAi, no ground water	0.83/9.9	0.48/7.9	0.89/8.1	0.61/6.1
OBS-FCSTv4	0.71/9.8	0.18/6.3	0.79/7.9	0.25/4.6

Values including or excluding the annual cycle and using no smoothing or a 1-2-1 filter are given

CSL. An initial value has to be given but is not important in this study because only anomalies are shown below. In addition, a mean annual cycle is removed to make the figures clearer. A mean annual cycle of the CSL with 40 cm lower values in winter than in summer (Fig. 5) has an amplitude, which is hardly exceeded by annual mean changes in the time-frame of few years, only in 1990–1992.

Figure 6 shows the time variability of the CSL as observed (OBS), simulated, or forecast. A simulation using the observed VRD together with P-E over the CS itself is included (VOLo). The latter gives the best similarity with the observed CSL as might be expected.

The anomaly correlations between the observed CSL (gauge before 1993 and satellite thereafter) and the calculations by the other methods are shown in Table 2. In addition to straight anomaly correlation, the correlation after calculating the time derivatives (tendency correlations) is also included. This reduces the impact from trends in the data. In both cases, the calculations using the observed VRD together with the evaporation over the CS give the highest values. However, during the period 2000–2003, the observed CSL drops by 30 cm and increases again (Fig. 6), which is not simulated with the observed VRD. This will be investigated further in Section 4.3. The estimate using no groundwater parameterization gives lower correlations than with such a parameterization (ERAi) showing the importance of this parameterization. The scores for the FCSTs are hardly worse than those for ERAi input and will be discussed in Section 4.4

Comparing the curve of the CSL estimate with the full ERAi input (Fig. 6, ERAi) to that having no parameterization of groundwater (noGW) indicates the beneficial effect of the

latter parameterization. Without it, the increase in the CSL before 1995 is much steeper with intermediate spurious decreases in 1992–1993, which are then smoothed out by this groundwater scheme, but this is still not sufficient. Further periods of extreme variability occur without the parameterization of ground water, which is then smoothed out when applying the scheme, which leads to improvements. In addition, Table 2 documents the positive impact of the groundwater parameterization.

4.3 Importance of different components of the CS water budget

The observed discharges of the rivers entering the CS give an overview about their importance for the CSL variability. The mean observed VRD would increase the CSL by 71 cm/year with extreme annual variations of 20 cm/year. The corresponding values for the Kura, Ural, and Sefidrud are 3.5 ± 2.4, 0.8 ± 1.3, and 1.0 ± 1.0 cm year^{-1}, respectively. The discharge of the Terek River is not known to us but may be important as well, while the Gorganrud has too low a discharge for having an impact on the CSL. To test if such impacts can be found as well with our scheme, we calculated the CSL variability by including different components of the CS water budget using observed VRD in combination with other components of the CS water budget from ERAi (Fig. 7). The observed VRD alone (VOon) already reproduces the main features of the observed CSL variability. Adding P-E over the CS (V+CS) improves the CSL variability considerably, especially before 1991.

The most difficult quantities to estimate in these budget terms are the evaporation over the CS itself (AR2012 pointed

Fig. 5 Annual cycle of the CSL as observed (*OBS*), simulated by ERAi or by ERAi/FCST 6 months ahead (*FCS6*)
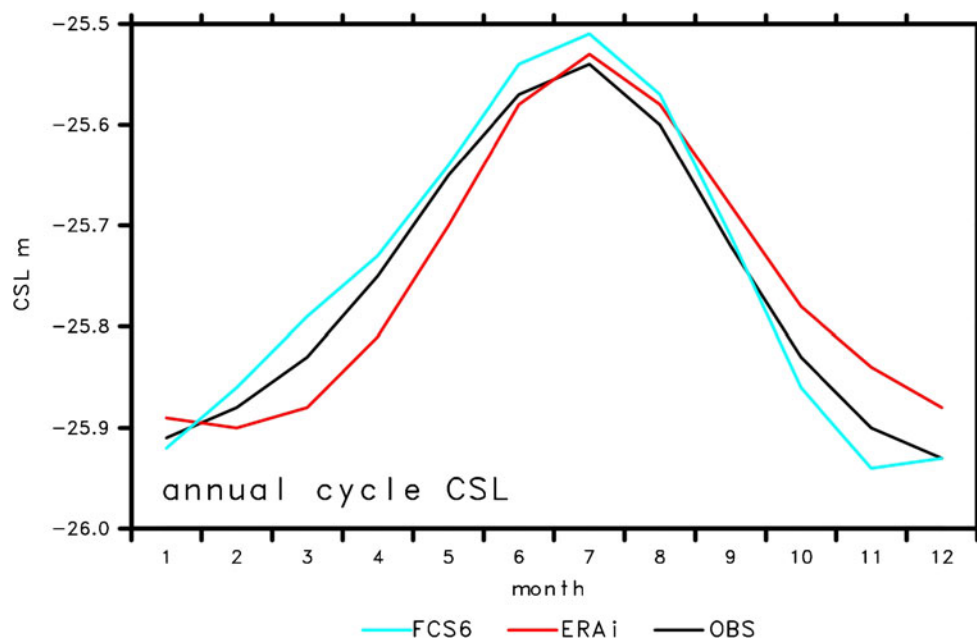
Fig. 6 Monthly mean Caspian Sea level (*CSL*) anomalies (mean annual cycle removed) as observed (*OBS*) and as estimated using observed Volga River discharge+P-E over the Caspian Sea (*VOLo*), using solely P-E values from ERAi (*ERAi*), using solely P-E (*VB+CS+SW*) values from ERAi but without groundwater parameterization (*noGW*) or mixed ERAi/FCST data with a lead time of 6 months (*FCv4*)
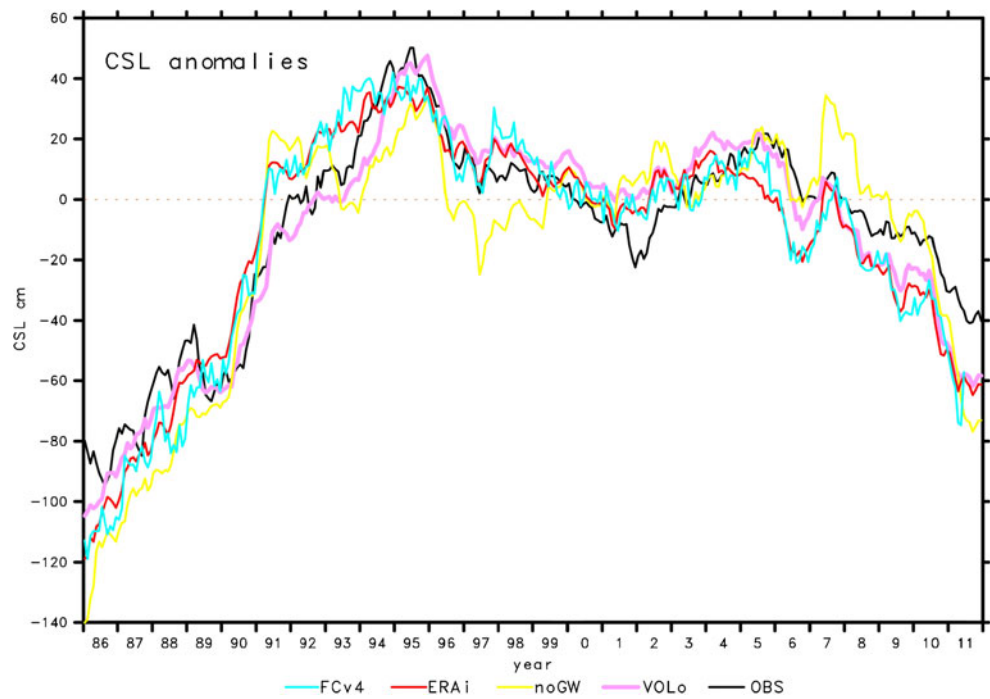
to a bias in the mean values). The precipitation over the southern and western shores of the CS is also difficult to estimate due to the high mountains there, i.e., the Alborz and Caucasus Mountains, which particularly affect the Sefidrud and Kura River, respectively. The south-west region (+SW and +SWg, the latter using GPCC precipitation data) shows the next important impact on the CSL variability. The +SW curve deviates substantially from the V+CS curve over several periods. Between 1996 and 2005, it gives higher values that mostly indicate less similarity with the observed ones. From 2007 onwards, the values are lower, again meaning deterioration. When using GPCC precipitation (+SWg), the differences are generally smaller and the estimates closer to the observations. Both SW data sets improve the CSL estimate for 1989/1990 and 1995. Such an impact from the south-western area is also confirmed from the river discharge observations, which were exceptionally large in 1988/1989 from the Kura River and the Sefidrud and low in 1990, large in 1993/1994 and low again in 1996. Generally, the precipitation data from GPCC (+SWg) give a CSL closest to that observed. The findings from inspecting the curves in Fig. 7 are confirmed by Table 3 or

made even more obvious. In particular, the improvements using GPCC precipitation over the SW area is very clear by all validation methods.

The impact from the Ural River turned out to be less important (Fig. 7 and Table 3) also when using GPCC data (not shown). Only the RMS errors show some improvements when using GPCC precipitation over the Ural River.

Figure 8, which repeats most of Fig. 7, but enlarged for the period 1997–2004, shows that the precipitation anomalies over the SW coast of the CS (+SW or +SWg) explains most of the decrease in the CSL in 2000/2001 and its increase in the following 2 years. Observations of the Kura River discharge in 2000/2001 were 47 % below and in 2002/2003 12 % above normal, while the observed discharge of the Sefidrud did not have marked anomalies in that period.

Another interesting CSL event at the end of 2002 can be seen in Fig. 7. The gauge data show a dip by 7 cm, which is not followed by the satellite observations. The estimates from the water budget reproduce this dip due to deficient VRD and to lower precipitation over the SW coast. AR2012 assigned this difference between the two observational data sets to a

Table 2 Anomaly correlation coefficients and RMS errors (cm) of the CSL as observed versus estimates by ERAi (VB+CS+SW), observed VRD+CS evaporation, ERAi but without groundwater parameterization and FCSTs with a lead time of 6 months

CSL OBS vs estimates by	ERAi	obs VRD+CS evaporation	ERAi no ground water	FCST V3 precipitation used	FCST V4 precipitation used	FCST no FCST data used
Anom corr/RMS	0.93/13.9	0.96/10.2	0.91/17.5	0.93/13.6	0.93/14.5	0.94/13.2
Correlation of tendency	0.89	0.93	0.87	0.87	0.86	0.89

In the second row, the time derivatives (tendencies) are correlated to reduce the impact of trends in the curves

Fig. 7 CSL (mean annual cycle removed) as observed (*OBS*) compared to that calculated only from the observed VRD (*VOon*), using observed VRD plus P-E over the Caspian Sea (*V+CS*), adding to V+CS impacts from the Ural River (*+UR*), adding to V+CS impacts from rivers at the SW coast (*+SW*), the same but using GPCC precipitation (*+SWg*)
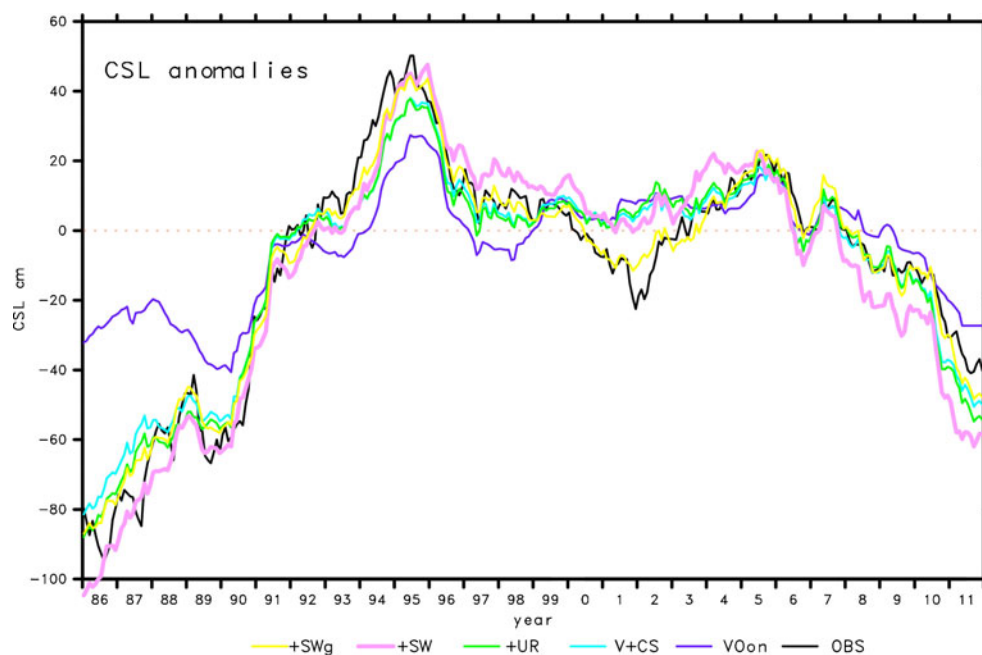
problem in the satellite data when the new JASON satellite data had to be adjusted to the previous TOPEX/POSEIDON data. Here, the water budget data suggest that the adjustment of the satellite data is not perfect. A more recent update of the satellite data (USDA2) does not show this bias anymore.

For this period, ERAi and GPCC over the SW area provide very similar results because the GPCC data have been shifted downwards by 10 cm (Fig. 8). For the whole period from 1986 to 2010 (Fig. 7), the differences are quite large due to accumulations of small uncertainties over a longer period. As our main aim is to make forecasts over 6 to 12 months, we will look below at time changes of the CSL, which act as a high-pass filter, which will suppress these uncertainties. The impacts from different components of the water budget of the CS agree with the estimates from observed discharge data given above, though in our estimate, we did not separate the impacts from the Kura River and the Sefidrud. For further investigations, the impacts from the Ural River will not be included as its influence is small.

4.4 CSL forecasts

In Fig. 6 and Table 2, the CSL variability calculated 6 months ahead is included. The differences between the estimates from

ERAi data and those from the forecasts are marginal. It has been shown in Section 4.1 that the water in the Volga Basin reaches the CS with a delay of a few months, and it is therefore a question of how much of the FCST quality is due to this storage provided by ERAi and then carried forward in the forecasts, and how much is due to the seasonal forecasts. Therefore, anomaly correlation coefficients and RMS errors are also calculated when the FCST depends solely on the storage in the Volga Basin and not using seasonal forecast data (Table 2). It turns out that the seasonal forecasts hardly improve the CSL forecast when measured in this way. However, both FCSTs are of similar quality to that of the ERAi simulation and one can hardly expect the FCSTs to provide better estimates of the CSL than when simulating the CSL with ERAi data.

On the other hand, it has been shown in Appendix 2 that the seasonal FCSTs do have some skill to predict the precipitation over the Volga Basin. Therefore, we extended the forecasts to 12 months and found correlation coefficients similar to that for ERAi and for 6 month FCSTs as shown below.

Above, the seasonal FCSTs with both versions are provided because version 3 provided a better precipitation forecast over the Volga Basin than version 4 (Appendix 2). Version 3 is, however, no longer operationally provided at ECMWF and therefore cannot be used in an operational

Table 3 Anomaly correlation coefficients and RMS errors (cm) of the CSL as observed versus estimates by VRD (observed)+CS evaporation and precipitation and estimates with adding further Ural or SW catchment precipitation

CSL OBS vs estimates by	V+CS	V+CS+UR	V+CS+UR(GPCC)	V+CS+SW	V+CS+SW(GPCC)
Anomaly correlation/RMS	0.97/8.9	0.96/8.9	0.97/8.2	0.96/10.2	0.99/5.3
Correlation of tendency	0.94	0.93	0.94	0.94	0.96

In the second row, the time derivatives (tendencies) are correlated to reduce the impact of trends in the curves

Fig. 8 The same as Fig. 7 but enlarged for the period 1997–2004 and with a reduced selection of experiments and adding the observations by gauges (GAUG, shifted upwards by 10 cm). The +SW data have been shifted downward by 10 cm to make them comparable with the GPCC data for that period
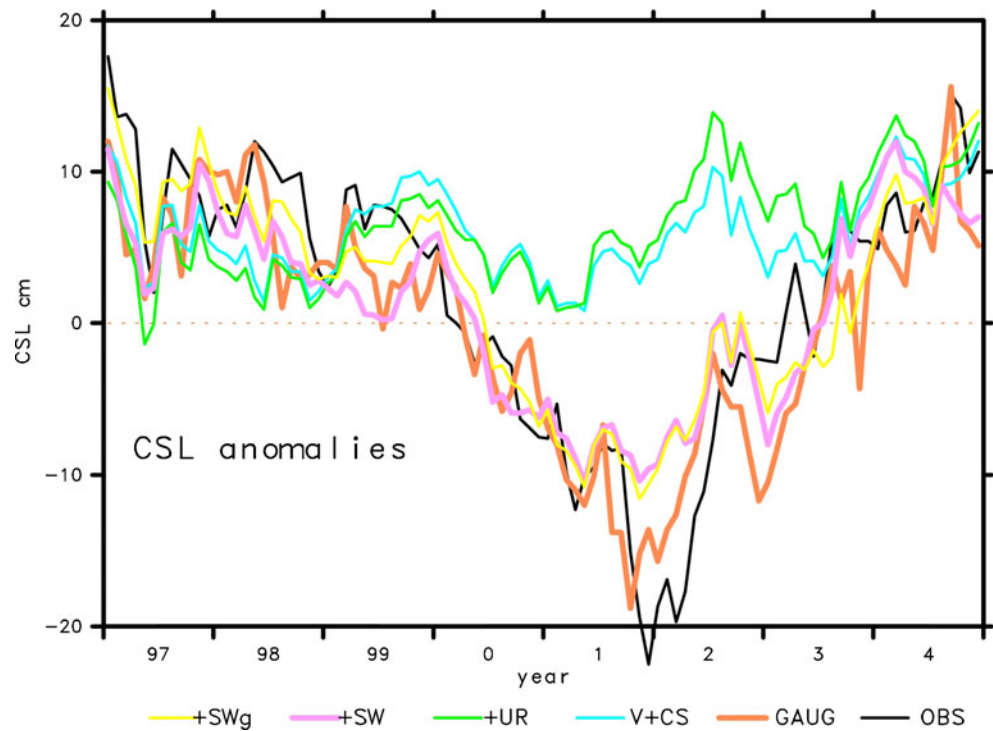

forecast of the CSL. As the summer precipitation forecasts over the Volga Basin are superior with version 3 compared to version 4, it is of no surprise that the FCSTs of the CSL are also superior with version 3. The anomaly correlations are 0.931 compared to 0.934; the RMS errors are 14.5 compared to 13.6 cm, and the anomaly corrections after calculating the time derivatives to filter out the trends are 0.864 compared to 0.870. Thus, in all scores, version 3 is superior to version 4 but only by a small amount.

4.5 CSL change in 6 months

The purpose of this study is to find out if there is scope for predicting CSL changes using ECMWF reanalysis and

seasonal forecast data. Figure 9 shows time series of the time change of the CSL over 6 months as observed, simulated, and forecast with the method described above. Dominant variabilities are similar in all curves. The extreme drop in 2010, connected with the Russian drought, occurred in ERAi 2 months before it actually happened. On the other hand, it is forecasted later because the early part of this drop is due to enhanced evaporation over the CS, as shown by AR2012, and this part of the water budget is not included in the FCSTs. The extreme observed decrease in the CSL in early 1989 is not at all captured, even not when using observed VRD (not shown). The general increase in the CSL in this period (Fig. 6) is, however, interrupted, strongest when using observed VRD together with the SW area P-E. The largest

Fig. 9 CSL change over 6 months as observed (*OBS*) or estimated from ERAi (*ERAi*) or ERAi/FCST (*FCv4* version4, *noFC* no FCST data used). Mean annual cycles have been removed
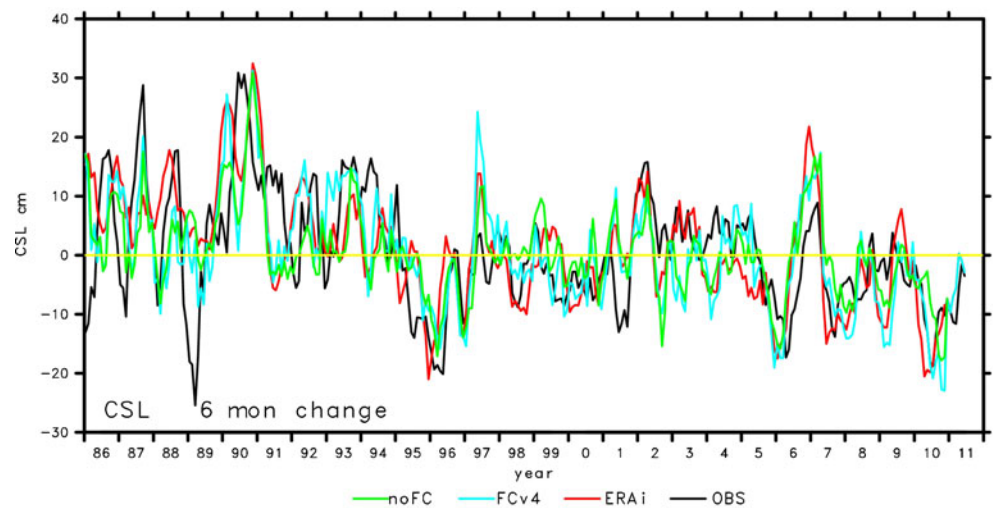
Table 4 Anomaly correlation coefficients and RMS differences (cm) between observed CSL changes over 6 months as observed by satellite or gauge and estimated from ERAi, ERAi/FCST with a lead time of 6 months, and observed VRD with ERAi

Correlation/RMS	1986–2010 satellite+gauge	1993–2005 satellite obs.	1993–2005 gauge obs.
OBS-ERAi	0.53/9.3	0.50/7.4	0.60/7.0
OBS-FCST VO V3	0.46/9.1	0.44/7.6	0.41/7.8
OBS-FCST VO+CS V3	0.50/8.9	0.49/7.3	0.45/7.7
OBS-FCST VO+CS+SW V3	0.53/9.0	0.55/7.0	0.50/7.4
OBS-FCST VO+CS+SW V4	0.54/9.1	0.54/7.5	0.53/7.6
OBS-FCST (noFCdata)	0.47/9.0	0.45/7.4	0.45/7.5
OBS-obs.VRD	0.67/7.4	0.69/5.9	0.80/4.9
OBS-persist. FCST	0.16/12.6	0.31/9.5	0.46/8.4
OBS sat-OBS gauge		0.73/5.9	
ERAi-FCST V3	0.82/5.6		
ERAi-FCST V4	0.79/6.2		

FCST skill scores not using any seasonal forecasts or using version 3 (V3) or version 4 (V4) are given. Version 3 also with different components of the water budget: precipitation over the Volga Basin (VO), over the Volga and the Caspian Sea (VO+CS), and over the Volga, the Caspian Sea, and the SW area (Vo+CS+SW). Persist. FCST means a forecast assuming that the trend of the last 6 months will persist also for the next 6 months

increase occurs in 1990, which is also simulated but with some time shifts. Some time shifts between observations and simulations can be seen for the other extreme events, e.g., 2005/2006 and 2006/2007. In 1995/1996, a shift even occurs between gauge and satellite observations. For 1993–2005, the observations by satellite USDA1 and gauge differ typically by 5 cm and occasionally reach 10 cm (not shown), i.e., a similar level of uncertainty as between observed CSLs

themselves (Appendix 1). Therefore, any changes of <5 cm will be regarded as noise.

When scoring the skill of weather forecasts, it is custom to validate or score them in comparison to a poor man's forecast, which can be a persistence forecast, i.e., the weather tomorrow will be the same as today, or a climate forecast, i.e., tomorrow's weather will be similar to a climate mean, better suited for longer range forecasts. The latter has already been applied

Table 5 Cases of events with CSL changes of more than 10 cm in the predictants

6 month change	Predictor/predictant	Correct <5 cm	Correct ±1 month	Correct with 1/2 change	Useful	Total events	Bad cases >20 cm
Increase	FCST (V3)/OBS	21	10	4	35/66 %	53	3/6 %
	FCST (V4)/OBS	22	7	8	37/70 %	53	2/4 %
	noFCST/OBS	16	8	3	28/53 %	53	3/6 %
	FCST (V3)/ERAi	34	6	3	43/88 %	49	0/0 %
	FCST (V4)/ERAi	33	6	5	44/90 %	49	1/2 %
	pers.trend/OBS	16	6	2	24/45 %	53	18/34 %
	OBS/ERAi	20	5	6	30/64 %	49	5/10 %
	ERAi/OBS	20	11	3	34/36 %	53	1/2 %
Decrease	FCST (V3)/OBS	11	5	7	23/58 %	40	4/10 %
	FCST (V4)/OBS	16	7	2	25/63 %	40	3/7 %
	noFCST/OBS	13	4	3	20/50 %	40	5/13 %
	FCST (V3)/ERAi	16	6	8	30/91 %	33	0/0 %
	FCST (V4)/ERAi	20	7	4	31/94 %	33	0/0 %
	pers.trend /OBS	10	4	4	18/45 %	40	9/22 %
	OBS/ERAi	17	6	2	25/76 %	33	0/0 %
	ERAi/OBS	17	5	2	24/57 %	42	5/14 %

Events occurring within ±1 month or changes smaller but still larger than half of the threshold of 10 cm are assumed to be useful. PERS.TREND means a forecast assuming that the trend of the last 6 months would persist also for the next 6 months. The last column gives the amount of bad cases, i.e., cases where the absolute error is >20 cm

Fig. 10 One-year forecasts of the change in 6 months means, annual cycle removed. The FCSTs with version 4 (*FCv4*) are compared with observed (*OBS*) or with ERAi (*ERAi*) simulated CSL changes. Data are smoothed with a 1-2-1 filter
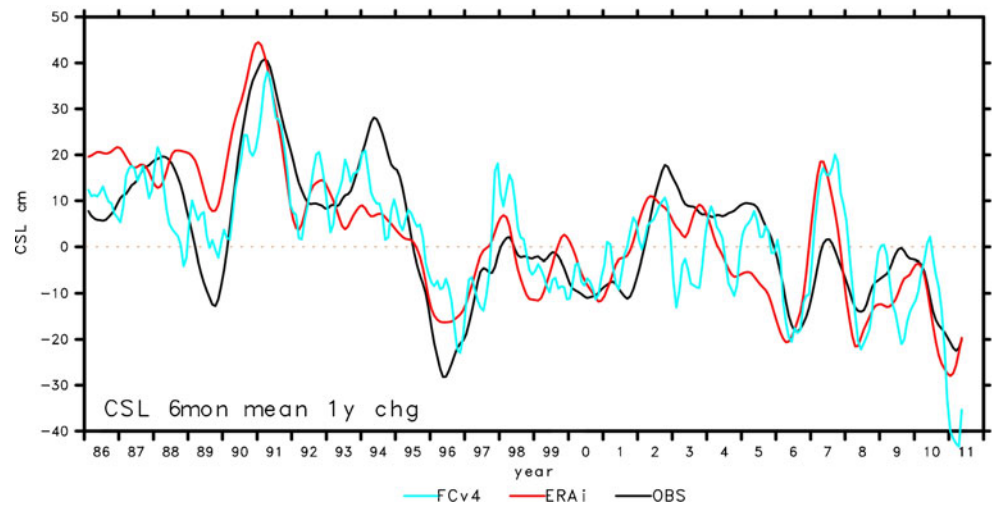
here using anomaly correlation coefficients as skill scores. The former would probably not have been used for the CSL forecast, especially not in cases when the CSL has been rising or sinking over a longer period and in such cases a better forecast would be needed. We will take here as a poor man's forecast the assumption that the trend, after removing the mean annual cycle, of the last 6 months will continue for the next 6 months. We call this a persistent trend forecast.

Table 4 provides an overview of the anomaly correlations and RMS differences between the observed CSL change over 6 months and that simulated or forecasts by ERAi, the FCSTs and the one when using the observed VRD. For the latter, one finds clearly higher correlations than with the other two as might have been expected. For the period 1993–2005, i.e., that with available gauge and satellite observations, the anomaly correlation coefficient between satellite and gauge observation (Table 4) is smaller than the correlation between the CSL as observed by gauges and estimated by the VRD observation and points to uncertainties in the observational data by satellite. Thus, the speculation by AR2012 that the satellite observations during 2002 may have some problems is supported here. The best results are achieved when using the ERAi P-E data over the Volga Basin, the CS, and the SW area. For the FCST (precipitation only), similar scores are achieved like those when simulating the CSL with ERAi, sometimes even better ones.

The two versions of FCST give similar results. Generally, the signals given by the anomaly correlation coefficients are supported by the RMS differences.

Table 5 gives an overview of the forecast skills of the CSL changes in the predictant of more than 10 cm. It is assumed that it is correctly simulated or forecast when the absolute error is <5 cm and a predictor would be still useful if the event occurs within ±1 month or if the change is smaller but still larger than 5 cm. Therefore, the numbers for the latter are included as well, but these cases are only counted if there are no other useful predictors in the other categories.

More than 60 % of such cases were correctly forecast or at least useful with the present scheme. The persistent trend forecast scored with only 45 % correct forecasts, clearly worse in this respect. When counting the cases when the forecasts were off by more than 20 cm, the scores are also in favor of the ERAi/FCST scheme as their typical occurrences of <10 % is by far exceeded by the persistent trend forecast of 22 and 34 %. The table shows that the ERAi scheme produces large changes less often than observed. Using the ERAi simulation as the truth, the scores for the FCST look much better with around 90 % useful cases and no bad cases at all.

It has been shown above that a good deal of the forecast skill is due to the storage of water in the Volga Basin, which is released with a delay. The seasonal FCST of precipitation is thus adding some more useful information. Because of

Table 6 Anomaly correlations and RMS errors (cm) for annual mean CSL forecasts in October and 6 months forecasts with the present scheme, also in October

FCST vs OBS corr/RMS	OBS MGO all	OBS MGO fcsts agree	OBS SAT all	OBS SAT fcsts agree	OBS SAT Apr-Sep all
MGO fcst	0.75/8.6	0.97/5.3	0.74/8.8	0.96/7.2	
ECM fcst (V3)	0.77/8.0	0.94/7.1	0.72/9.0	0.94/8.3	0.76/9.4
ECM fcst (V4)	0.82/7.2	0.97/5.5	0.79/8.1	0.98/6.4	0.67/11.4

Values are given for all forecasts (columns 2 and 4) and a subset where MGO and our FCST agree in predicting strong increase, little change, or strong decrease (columns 3–5). The last column shows forecasts in October for the following April to September means with the present scheme

Fig. 11 CSL anomalies as observed (*OBS*) or forecast 1 year ahead by HMRC and ECMWF version 4 (*FCv4*) and version 3 (*FCv3*). For 1996–2000, the forecasts start in May instead of April. For each forecast, two different alternating line styles are used to make the beginning and end of each forecast clearer
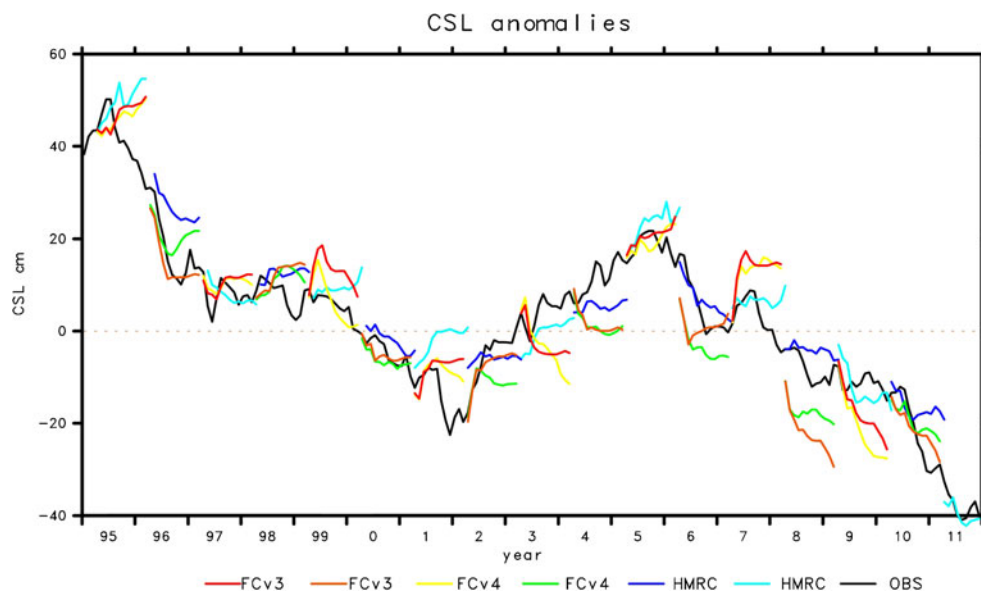
that, we pushed our scheme further to make 1 year forecasts using climatological mean precipitation values for the forecast range beyond the seasonal forecasts. In Fig. 10, the performance of 1 year FCSTs of 6 months means are shown.

The forecasts over 1 year show considerable skill with only slight differences when using the two versions of the seasonal FCSTs. The anomaly correlations are 0.78 for ERAi vs. OBS and 0.76 for FCST (both versions) vs. OBS. The correlations of the two FCST versions differ only by 0.005, though the RMS errors are clearly larger with version 4, 9.7 cm versus 9.3 cm.

4.6 Comparison with other forecasts

The CSL forecasts of annual means for the following year done each October by Meshcherskaya et al. (1997) at MGO, St. Petersburg, are available for comparison and are used here. These forecasts are based on statistical relations between the CSL and circulation indices, and is therefore a totally different approach from our study.

The MGO results are presented here as anomaly correlation coefficients and RMS errors comparing the 1-year

forecasts with the 1-year mean observed change, as provided by MGO (Table 6). Anomaly correlation coefficients of 74 % mean a high skill. In our study, we use the observed CSL data provided by satellite observations when available. If one uses this data set (column 4 in Table 6), the score remains the same though with slightly higher RMS errors. If one also uses the 6-month forecasts from our scheme and assumes that the 6-month forecasts in October would be an estimate of the annual means of the following year (row 3 and 4 in Table 6), one finds similar scores. If one puts the forecasts in three categories (strong decrease, small changes, and strong increases) and uses only those forecasts if they both fall into the same category (column 3 and 5 in Table 6), one finds substantially better scores, meaning that, using both approaches of forecasts, one can significantly increase the confidence in the forecasts. For these October FCSTs, version 4 gives higher scores than version 3 because this forecast range covers the winter season, for which the precipitation forecast is also better in version 4 than version 3.

One-year forecasts of half-year means have been presented in Fig. 10, and from that, the October forecasts are given in the last column (Table 6). These forecasts have

Table 7 Qualitative judgment of the quality of the forecasts in Fig. 11 by comparing the forecasts with the observations, especially for the winter period

	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10
HMRC	−	0	+	+	−	0	−	+	+	+	−	0	0	0	+	−
EC V4	−	+	−	+	+	+	+	+	−	−	+	0	−	+	−	+
EC V3	−	+	−	+	−	+	0	0	−	−	+	+	−	−	−	+
best	E	E	H	H	4	E	4	H	H	H	E	3	H	4	H	E

(+) good, (0) might be useful, (−) bad; best of the three forecasts: *H* HMRC, *E* version 3 and 4, *3* version 3, *4* version 4

skill scores similar to the others for all cases, but we believe that this is a more useful forecast because it provides the focused information for the following summer high stand. In this comparison, version 4 results are of lower quality.

In April, HMRC provides operationally 1 year forecasts of the CSL. Such a forecast can give guidance for the next low stand of the CSL in winter. From 1995 onward, these forecasts were available to us and we compared their forecasts with those by the present scheme for the same period and same initial date when extending them beyond the seasonal forecast length to one year (Fig. 11). The results are not conclusive as all three schemes seem to be of similar quality on the basis of 16 cases, which are too few for a statistical evaluation. However, in Table 7, a qualitative judgement on the three forecasts has been made for each year by viewing the difference between the forecasts and the observations. HMRC was best in 7 years and ECMWF version 4 in 8 years though the latter has to share the best place in 5 years with version 3. The difference becomes clearer when counting the good forecasts, nine for version 4 and six for HMRC. However, for bad forecasts, HMRC has five and version 4 has six cases. Thus, the present scheme using version 4 has only small advantages. The HMRC scheme has been developed and improved over more than 50 years, while the present scheme is just a first attempt, and with the many anticipated improvements to come, it is likely that the present scheme will become superior.

4.7 Three important events

Above, the anomalies (annual cycle removed) of the CSL and CSL changes have been investigated. The anomalies are used to increase the samples for a statistical validation without artificially enhancing the scores by including the annual cycle. The end user would probably take a different approach. They would be interested if the absolute CSL exceeds a threshold, perhaps surpassing the minimum or maximum of the last five or so years. For that, one has to take the annual cycle of the CS into account. There are only four events with such extreme values that can be investigated, i.e., the high stands in 1995 and 2005 and the low stands in 2002 and 2010–2011. The interruption of increase in 1989/1990 is also an interesting period when people might have hoped that the steady increase had stopped but then it started again. Too few cases are available for statistical scoring, but it can be looked at anyway. The reaching of a maximum CSL in 1995 was anticipated by ERAi and FCST 2 years earlier than observed (Fig. 6). The low level in 2002 was not caused by anomalously low VRD (observed or simulated) but by low precipitation over the mountains along the western and southern shores of the CS.

What would the advice have been from this study to the management during possible crisis meetings during events of long-lasting periods of increases or decreases of CSL? Below

all values are anomalies, i.e., after the removal of the mean annual cycle for clarity. The CSL is 40 cm higher in summer than in winter, observed and simulated as shown in Fig. 5. Therefore, such emergency meetings would be held only in cases of rising CSL in winter and in cases of decreasing CSL in summer in preparing for the next high or low stand, respectively. Three interesting events are shown here in detail. If the present forecast scheme is not available, a possible poor man's forecast would have been that the trend of the previous 6 months would continue for the next 6 months (persistent trend forecast).

(a) End of 1994: The CSL has been rising since 1977 with only a few breaks. It has increased in the previous 1, 2, and 3 years by 0.23, 0.34, and 0.42 cm, and so is a very scary scenario. Thus, the question would have been whether it will continue rising for the next summer high above the 40 cm due to the mean annual cycle (Fig. 5). The 6–9 months FCSTs in Fig. 12a (different line styles) all start from the estimates of the CSL with ERAi. The FCST for the month of interest (thick continuous) is shown as well shifted, so that it starts at the observed level. The forecast suggests only a slight increase—nearly perfect. The ERAi simulation suggests earlier that, due to excessive precipitation over the Volga Basin in 1990/1991 (Fig. 4), the CSL should have reached higher values in 1992 than observed, but it then stayed at that level for 3 years (Fig. 6). The reaching of the maximal CSL as observed in 1995 is correctly predicted by the present scheme (Fig. 12a), better with version 4 than version 3, while the persistent trend forecast would have given bad advice. If the emergency meeting had been held in January 1995, the error of the persistent trend forecast would have been less, but the present scheme would still have given better advice.

MGO predicted in October 1994 for 1995 a 20-cm higher annual mean CSL than 1994 while the observed changed only by <10 cm. The present FCST from October 1994 gives an increase <2 cm for the 7 months, in line with observations.

(b) December 1989: The CSL had been rising since 1977 but not during the last year, so there was hope that the increase had stopped. However, after that month, the CSL started rising again, and this was well forecast by the present scheme, while a persistent trend forecast would have given bad advice (Fig. 12b).

MGO predicted in October 1989 the annual mean CSL for 1990 to be 33 cm higher than that of 1989, almost perfect, though the end user might initially have been disappointed because this increase did not occur for the next summer high stand but later. The present FCST scheme predicted in October 1989 only a 2 cm increase for the 7 months ahead, which is correct as well. The difference between the two
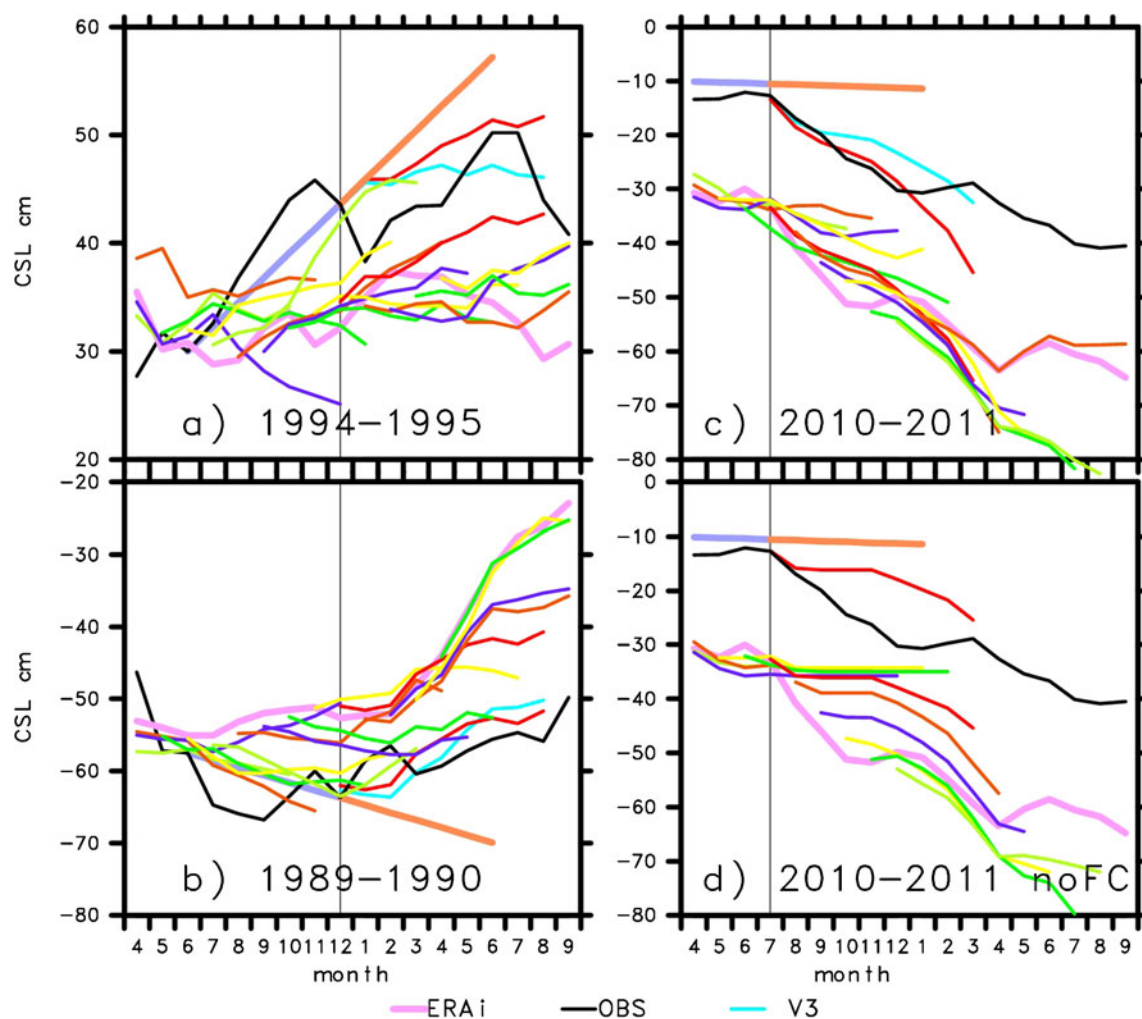
Fig. 12 Monthly CSL as observed (*OBS*), simulated by ERAi (*ERAi*) and by the FCSTs (*V4*) in different line styles for 9 months ahead. The FCSTs for the month of interest (December or July, indicated by a *vertical line*) is copied as well shifted, so it starts at the observed level. For that month also, the FCST with version 3 is included (*V3*). The

straight heavy line represents a forecast assuming that the trend of the last 6 months (*dotted*) will continue also for the next 6 months. **a** Years 1994–1995. **b** Years 1989–1990. **c** Years 2010–2011. **d** Years 2010–2011, but not using seasonal forecast data

forecasts is that the strong observed increase occurred only after August 1990, i.e., after the 7 months of the present forecast.

(c) July 2010: The CSL had been falling since 2005 with only a short break in 2007 by 25 cm though hardly during the previous 6 months. Because of the drought over European Russia a further drop had been anticipated by Lahijani et al. (2010) and a press release by Leroy et al. from Brunel University, which was picked up by the Financial Times on 27 August 2010 and by the Kazakhstan Today on 31 August 2010 (Financial Times 2010; Kazakhstan Today 2010) to warn for possible consequences. AR2012 discussed this case in detail. The drop in July–August of the observed CSL was due to extreme evaporation over the CS (AR2012), information that was not available to the forecast scheme

because of the bad evaporation forecasts (Appendix 2). Nevertheless, the ERAi/FCST scheme already forecast a drop of the CSL for the first 2 months due to deficient precipitation over the SW area of the CS.

From November onwards, because of the time delay in the Volga flow, the CSL drops further in the FCST (Fig. 12c). The fact that the version 4 FCST gives better results than version 3 up to 6 months is probably for the wrong reason, as the drop due to enhanced evaporation over the CS itself is not known to the present forecast scheme. The present scheme up to 9 months gives good guidance, while the persistent trend forecast would be bad guidance.

In spring 2010, HMRC predicted a drop of the CSL but too little particularly for the coming winter, while in April 2010, the present scheme had already predicted a much lower CSL for winter in line with observations (Fig. 11).

(d) No FCST used: Figure 12d shows the forecasts without using the information from the seasonal forecasts. The improvements are quite clear when using the seasonal forecasts.

The three examples with extreme scenarios showed that the present scheme would have been useful for planning ahead for next seasonal high or low stands. They also showed no clear advantage in using version 3 or 4 forecasts. With three cases, one cannot make a general statement, and therefore, below, an attempt is made to find some statistics on a broader data base.

Above, it was assumed that during long-lasting periods of increases or decreases of the CSL, emergency meetings would have been held in winter when the CSL was increasing and in summer when the CSL was decreasing to decide on measures to be taken for the coming high or low stand (respectively). In the three examples, it is demonstrated how the present scheme would have helped. Some objective measures of how often the availability of the presented scheme would have helped are shown in Table 8. From the 23 cases (June, July, December, and January are treated as separate events), our scheme, using the seasonal precipitation forecasts version 3, would have given seven forecasts with <5 cm error and one case with >15 cm errors, while it would have been six and seven cases, respectively, with a persistent trend forecast. Version 4 FCST performs clearly worse than version 3. It can be deduced that our scheme, using the seasonal precipitation forecasts version 3, gives the best results, though both versions are of similar quality when using RMS errors as a measure.

5 Discussion and conclusions

The impact of the Russian drought in 2010 on the CSL was studied by AR2012 and initiated the present investigation because of the good results obtained then. However, in that study, the timing of events was not a main issue. In the present study, the ERAi and FCST are used independently from observed CSL or VRD data for the whole 26–32 years of investigation.

The results are promising, but some better parameterizations, particularly for the storage of water in the Volga Basin, could be of advantage. The present study focuses on the forecast range of half a year, i.e., the length of the seasonal forecasts, but information is carried over by the water storage within the Volga Basin, which can store water for 2 years or more leading to some beneficial smoothing. It was shown that it is possible to make forecasts using only this storage of water in the Volga Basin without using the seasonal forecasts with similar skill but the seasonal forecasts lead to improvements. The positive impact of storage of water in the Volga Basin on the forecast skill gives the opportunity of carrying the forecasts beyond the length of the ECMWF seasonal FCSTs with some good results for half-year means 1 year ahead. The forecasts have skill scores, which are similar to those when comparing the observed CSL with estimates from ERAi and one cannot expect scores from the forecasts, which are better than those when using ERAi data.

In conclusion, this study has shown that there is scope for predicting the CSL some months ahead, but some problems have been pointed out that are believed to be solvable by ECMWF. Even forecasts for more than 6 months are possible because the VRD is the main driver of the CSL variability and the storage of water in the Volga Basin provides impacts for several months. This might be an essential contribution to the mitigation of societal impacts on and around the CS such as harbor accessibility and the petroleum and caviar industries. Restricting the FCST cases to those where the present scheme predicts a trend, which is the same as that with a completely different forecast scheme, like the one by MGO, improves the reliability of the forecasts considerably.

It seems to be the right time now to prepare for making such forecasts operationally. The use of run-off data instead of P-E needs to be tried and their deficiency needs to be understood. In addition, the use of a deficient SST in the seasonal forecasts needs to be corrected to provide even better forecasts. As the CS area increases with increasing CSL, the evaporation over the CS will increase, which provides a negative feedback. It might be useful to include this in a further development of this scheme though its effect remains small in the time range of 1 year.

Table 8 Guidance during assumed emergency meetings during long-lasting increases or decreases of more than 7 cm during the previous year and 13 cm during the previous 2 years

	Good	bad	RMS
Persist. trend	6	7	12.3
No FCST data	8	3	9.3
FCST V3	7	1	8.7
FCST V4	7	5	8.8

A guidance from a persistent trend forecast, which means that one assumes the trend from the last 6 months will continue also for the next 6 months, is shown in the first row while the other rows give values of the present scheme with or without using contributions from the seasonal forecasts. "Good" means an absolute error <5 cm and "bad" >15 cm. RMS errors between observed and forecast changes for these cases are shown in the last column

Appendix 1

Comparison between gauge and satellite observations of the Caspian Sea level

For the period 1993–2005, gauge as well as satellite observations (USDA1) of the CSL are available. The gauge observations at seven sites, adjusted to each other for overlapping periods after removing the mean annual cycle for each data set separately, are shown in Fig. 13. Single gauge observations differ considerably from each other, e.g., 1998 and 2005, which shows the uncertainty of using the gauge observations or for defining a mean CSL. Perhaps, these are periods with strong winds, which cause different CSLs at different sites. The means of the single gauge observations, after ignoring the highest and lowest values, agree quite well with the satellite estimate (Fig. 13, *top*) with anomaly correlation coefficients of 0.73 and RMS errors of 5.9 cm (Table 4). From 2002 onwards, a bias occurs which is discussed in Section 4.3 but has been corrected in a recent update of the satellite data (USDA2). Typical differences of the means are 10 cm, which can be used as a guide for uncertainties of CSL observations.

Appendix 2

Choosing the input data

AR2012 chose to use precipitation and evaporation of the ERAi interim data set for their calculations. In Section 3.1,

Fig. 13 Comparison of CSL observations by satellite (*SAT*) and gauges (*mea*). Anomalies are shown, which have been adjusted to each other for overlapping periods. The *top panel* shows the means of seven gauge observational sites, leaving out the lowest and highest values, in comparison with the satellite observations. The *lower panel* shows the observations at seven single sites. Mean annual cycles are removed
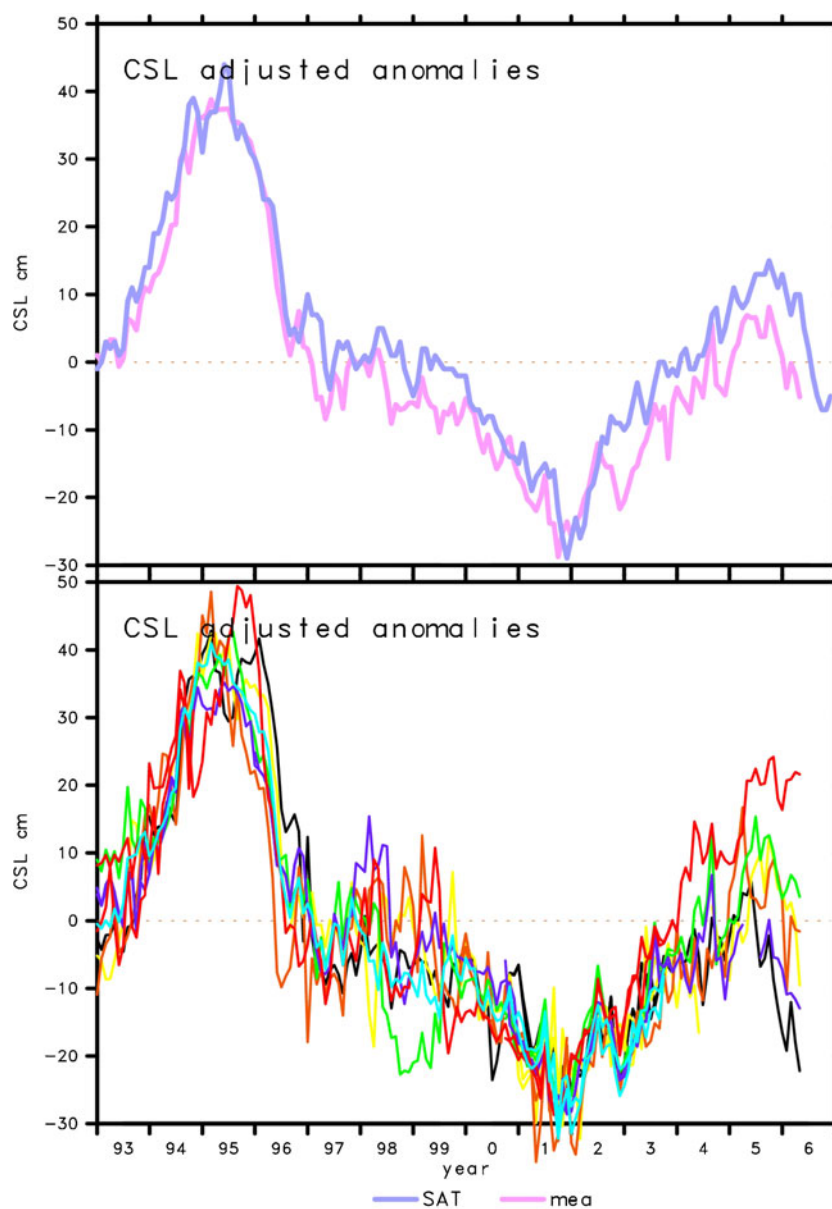
Fig. 14 Volga River discharge anomalies (annual cycle removed) as observed (*OBS*) or calculated from P-E (*P-E*) or from run-off data (*RUOF*) using ERAi data. Observed VRD data are converted into mm/month over the catchment area. A 9-month running mean has been applied for smoothing
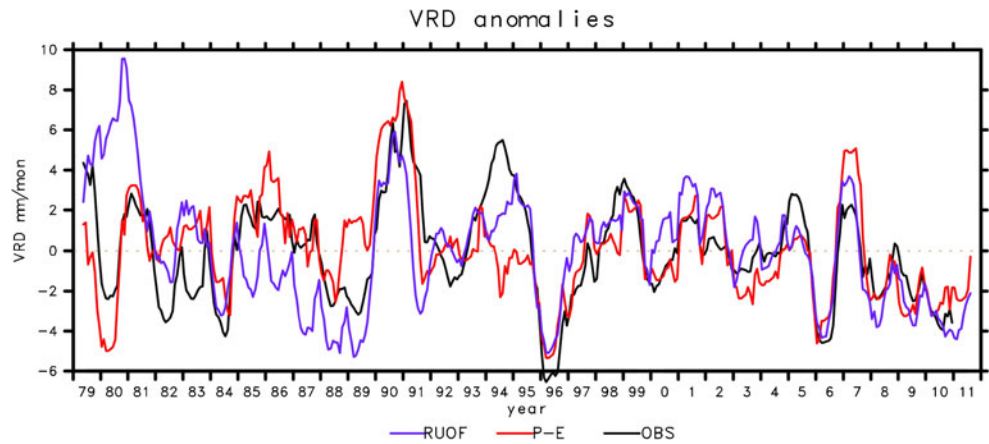
the method of estimating a delay between an event of precipitation over the Volga Basin and the time when the water reaches the mouth of the VR has been described. In particular, the delay, due to snow and ice on the ground described there, is quite simple and more sophisticated methods, which include the use of the temperature at the ground could be envisaged. The ERAi interim data set also provides run-off data, which deal with this sort of storage in a consistent way, at least considering the delay due to ice and snow.

A main problem with the hydrological budget in the analysis scheme is that it is not closed, i.e., the scheme might spuriously add or withdraw water perhaps due to the analysis of the soil moisture. Previously, AR2012 noticed this

problem when comparing precipitation and evaporation with observed VRD. Only by considering anomalies, this problem becomes less critical. Here, we tried to use the ERAi run-off data set and found that the analysis scheme introduced extra water sources, which are not physically explained. The run-off calculation not only uses P-E but also data from a soil moisture analysis. This analysis is not water conserving and may be the reason for this spurious water source. To obtain similar amounts of water discharge from the Volga River as observed or estimated from P-E with the run-off data, we had to reduce the values of the run-off by a factor of 0.7. This spurious water source may also be the reason why the evaporation over the Volga Basin appeared to be excessive.
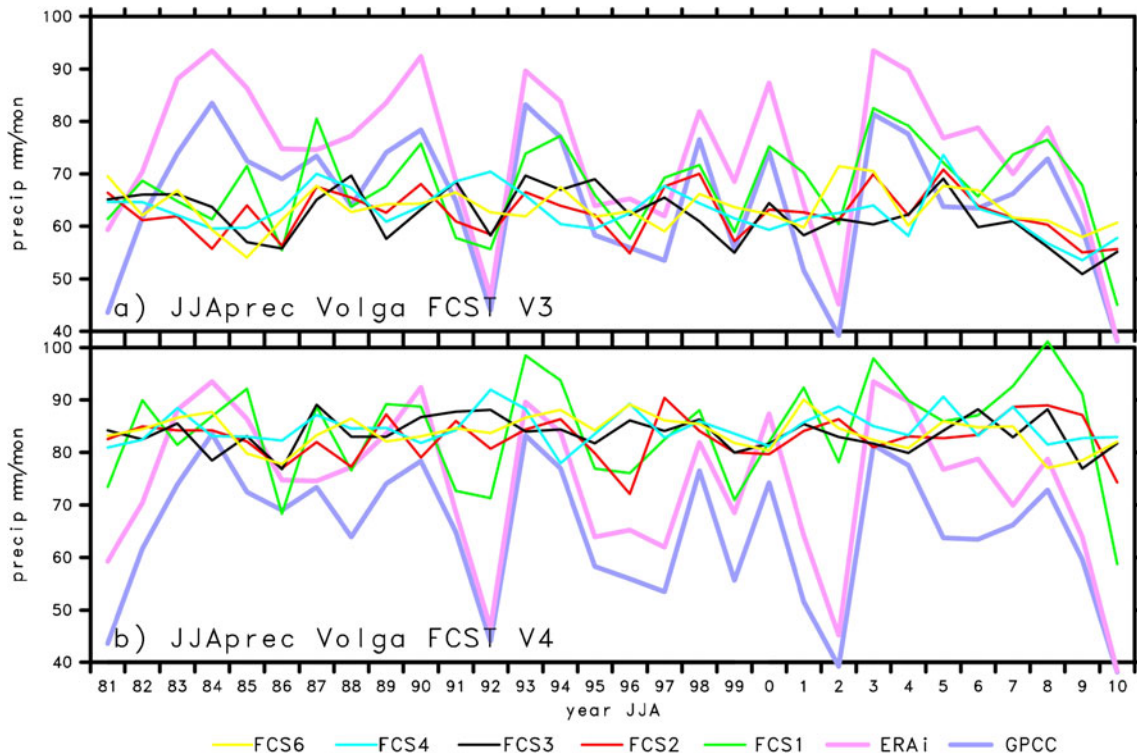
Fig. 15 Volga River precipitation for summer (*JJA*) as analyzed (ERAi or GPCC) or forecast by the two seasonal forecast versions V3 and V4, with lead times of 1, 2, 3, 4, and 6 months ahead

Table 9 Anomaly correlation coefficients of summer and winter precipitation over the Volga Basin between observations (analyses) by GPCC and ERAi or FCSTs of different numbers of months ahead

	GPCC-ERAi	GPCC-FCST1	GPCC-FCST2	GPCC-FCST3	GPCC-FCST4
Version 3 JJA	0.95	0.67	0.33	0.22	−0.09
Version 4 JJA	0.95	0.65	0.18	−0.05	−0.19
Version 3 DJF	0.98	0.35	0.20	0.01	0.11
Version 4 DJF	0.98	0.58	0.23	0.15	−0.02

The VRD as observed or calculated from P-E or run-off data show (Fig. 14) that, before 1990, the estimate using run-off data has variabilities, which are not seen in the other data. The analysis scheme is the same throughout the 33 years (1979–2012), but recently, new data sources for analyzing the soil moisture have become available, and it is speculated here that the spurious variabilities in the run-off data are caused by such changes in the availability of observational data. From 1995 onwards, the observed VRD and those calculated from P-E and run-off data show a high similarity. One would expect that such a similarity as P-E is the main input into the calculation of the run-off and the run-off may be a better input than P-E. Nevertheless, the factor of 0.7 has still to be understood.

Further data sets available at ECMWF are the seasonal forecasts in two different versions (version 3—Anderson et al. 2007 and version 4—Molteni et al. 2011). As version 4 is more recent, in general, one should expect the forecasts with version 4 to be superior to those with version 3. To test

if this is true for the present application, the summer precipitation over the Volga Basin of the different analyses and forecasts are compared (Fig. 15). Summer precipitation is chosen here because Arpe and Leroy (2007) showed that this is the most important season for the CSL variability.

The ERAi analysis suggests more precipitation than GPCC but with a similar variability (Fig. 15). The latter is documented by very high anomaly correlation coefficients between those two data sets of 0.95 (Table 9). The higher mean values in the ERAi data set compared to GPCC is probably due to the fact that the GPCC data should have been corrected for blow off of water from the gauges by the wind, which has not been done for the data used here. Such corrections can amount to 50 % in winter with snow fall but is much smaller in summer. Version 4 simulates on average more precipitation in summer than version 3, more in line with the analyses. Four years with extremely low precipitation amounts in the analyses stand out. The 1-month forecasts at least show indications of this for 1992 and 2010, better with version 3 than version 4. Also for
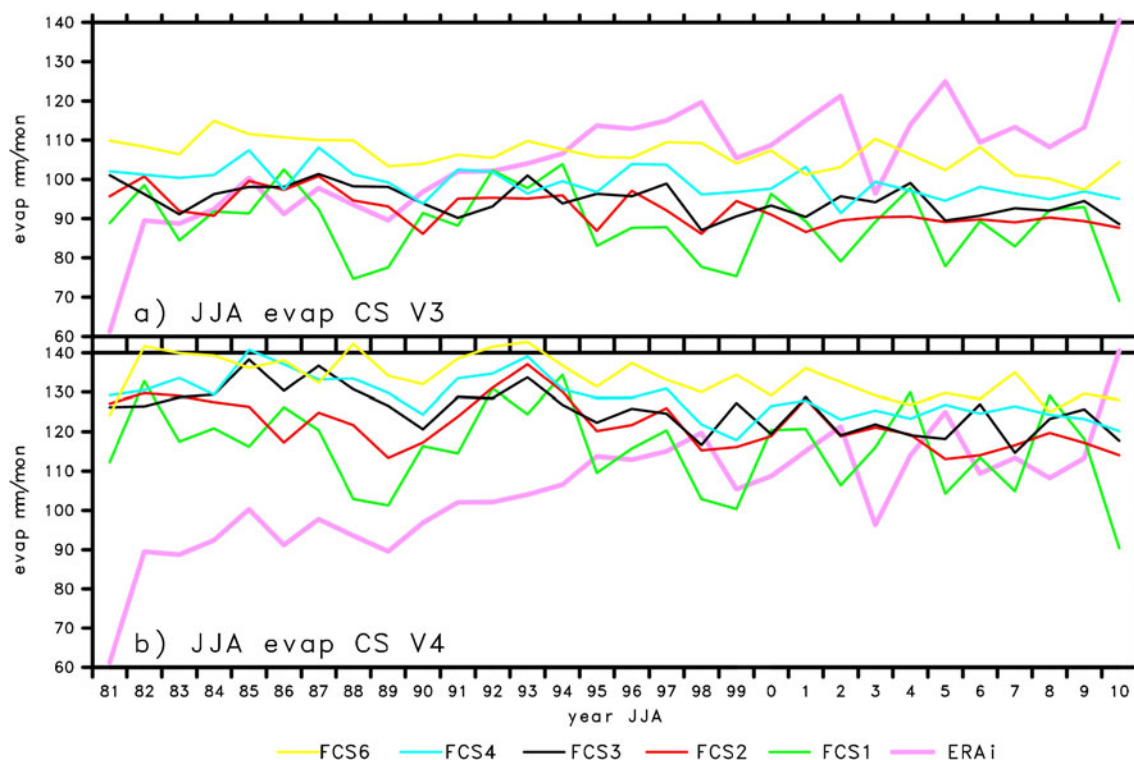
Fig. 16 Evaporation over the Caspian Sea for summer (*JJA*) as analyzed (*ERAi*) or forecast by the two seasonal forecast (versions V3 and V4), with lead times of 1, 2, 3, 4, and 6 months ahead

Fig. 17 Sea surface temperatures (*SSTs*) over the Caspian Sea for summer (*JJA*) and for annual means (*ann*) in the ERAi interim data set. From the SSTs for JJA, a constant value of 24 °C has been taken away and from those for the annual mean 15 °C, to make both lines fit in the same diagram
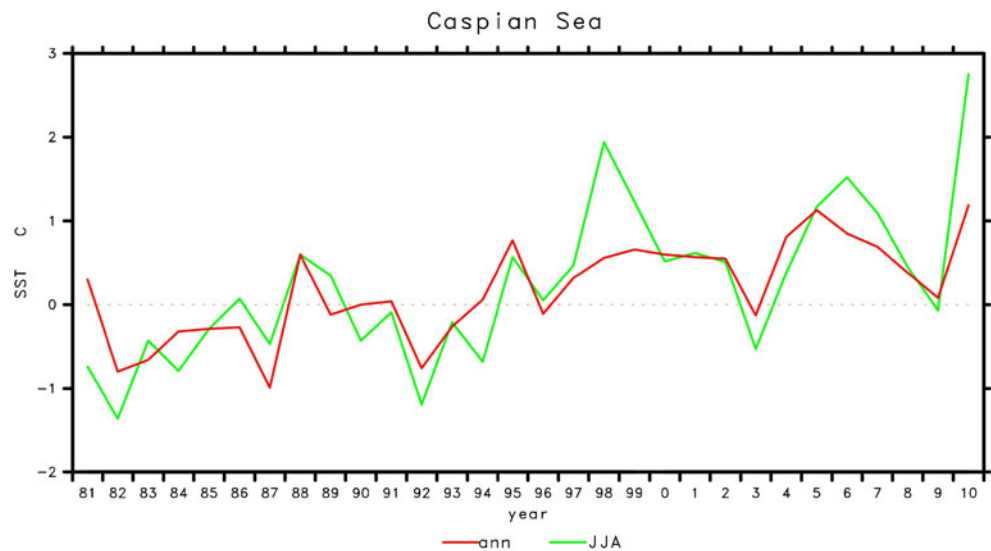
the 2002 event, version 3 is superior though far from perfect. Table 9 reveals further that the quality of the forecasts drops rapidly after a 1 month forecast, and more rapidly in version 4 than in version 3. The 3-month forecast of version 3 is as good as the 2-month forecast of version 4. Beyond this, the forecasts can be assumed to be useless (Fig. 15 and Table 9). The slight advantage of version 3 can be assigned to the events in 1992 and 1996 and the general drop from 2004 onwards. Above, the concentration was on summer precipitation because of its larger effect on the CSL. In addition, in Table 9, the scores for winter are shown, in which version 4 seasonal forecasts have some advantages compared to version 3.

The next important budget term for the CS and its variability is the evaporation over the sea itself. Figure 16 shows this quantity for the summer months. Again evaporation is much higher—more realistic—in version 4 forecasts than those of version 3. All the curves show trends, strongly upward in the ERAi data and slightly downward in the forecasts. This prevents the use of simple correlation coefficients. In Fig. 16, when comparing extreme events in the ERAi data with events in the forecasts, one cannot find any similarity. The high evaporation over the CS during the Russian drought in 2010, which is found in the ERAi data, is badly forecast as extremely low evaporation.

A major problem with the evaporation over the CS is that the FCSTs do not have an appropriate SST. They use a climatological mean value, which has been obtained by an off line simulation. Figure 17 shows the observed SST over the CS for summer and for the annual mean. A steady increase in the SST can be seen and the SSTs of summer 2010 were especially high, agreeing with the extremely high evaporation. Obviously without a realistic SST provided to the seasonal FCST, it is unlikely that they can provide a realistic evaporation level. The summer 2010 is an example

for that when ERAi provides very high evaporation rates, while the FCSTs simulate very low evaporation.

From these examples, one can deduct that the use of precipitation forecasts for the Volga Basin would be of advantage, especially those using version 3. The use of FCSTs of evaporation over the CS would cause deterioration and should not be used. Furthermore, it is found that the precipitation over the CS has a positive impact on the simulation of the CSL and is therefore also used. Using run-off data instead of P-E is probably better for the future (e.g., for operational forecasts of the CSL) or for the last 10 years but as our investigation concentrates on the last 30 years, it seems to be better to use P-E.

References

Abuzyarov ZK (2009) Improved method of CS level annual prediction. Trudy Gidrometcentra Rossii 341:48–66, in Russian

Abuzyarov ZK, Krasyuk VS (2003) Forecast of the annual variation of the Caspian Sea, based on an approximate calculation of the water balance equation. In: Shiklomanov IA, Vasiliev AS (eds) Hydrometeorological aspects of the Caspian Sea basin. Gidrometeoizdat, St. Petersburg, pp 310–322 (in Russian)

Anderson D, Stockdale T, Balmaseda M, Ferranti L, Vitart F, Molteni F, Doblas-Reyes F, Mogenson K, Vidard A (2007) Development of the ECMWF seasonal forecast System 3, ECMWF Technical Memorandum 503. ECMWF, Reading, UK, 56 pp

Arpe K, Leroy SAG (2007) The Caspian Sea level forced by the atmospheric circulation as observed and modelled. Quat Int 173–174:144–152

Arpe K, Bengtsson L, Golitsyn GS, Mokhov II, Semenov VA, Sporyshev PV (1999) Analysis and modeling of the hydrological

regime variations in the Caspian Sea basin. Doklady Earth Sciences 366(4):552–556

Arpe K, Bengtsson L, Golitsyn GS, Mokhov II, Semenov VA, Sporyshev PV (2000) Connection between Caspian Sea level variability and ENSO. Geophys Res Lett 27:2693–2696

Arpe K, Leroy SAG, Lahijani H, Khan V (2012) Impact of the European Russia drought in 2010 on the Caspian Sea level. Hydrol. Earth Syst Sci 16:19–27

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F (2011) The ERA -interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137(656):553–597. doi:10.1002/qj.828, Part A

Doganovskii AM (1982) Cyclic fluctuations in lakes level during the past hundred years. Geography and Natural Resources 3:152–156

Dümenil Gates L, Hagemann S, Golz C (2000) Observed historical discharge data from major rivers for climate model validation. Max-Planck Institute for Meteorology, Hamburg, Report no. 307, 93 pp

ECMWF: http://data-portal.ecmwf.int/data/d/interim_daily/. Last accessed 20 May 2011.

Financial Times (27 August 2010) Russia's drought threatens Caspian caviar

Giralt S, Julià R, Leroy S, Gasse F (2003) Cyclic water level oscillations of the Kara Bogaz Gol - Caspian Sea system. Earth and Planetary Science Letters 212(1–2):225–239

Golitsyn GS (1995) The Caspian Sea level as a problem of diagnosis and prognosis of the regional climate change. Izv. Russ. Acad. Sci. Atmos. Oceanic phys. Engl Trans 31:366–372

GPCC (2011), ftp://ftp-anon.dwd.de/pub/data/gpcc/html/download_gate.html. Last accessed 15 January 2011.

Kalinin GP (1941) Prediction of the Caspian Sea Level. Proc. NIU GUGMS. Ser.VIII.-issue 1, 72 (in Russian).

Kalinin GP, Smirnov KI, Sheremetevskaya OI (1968) Water balance calculations of the future state of the Caspian Sea level. Meteorologiya i Gidrologiya 9:45–52 (in Russian)

Kazakhstan Today (31 August 2010). Drought in Russia is dangerous to Caspian Caviar.

Lahijani H, Leroy S, Beni MN, Arpe K (2010) Blaze in Volga: is it a critical point for the fall in Caspian Sea level? News Report Episodes 33(3):208–208

Lebedev SA, Kostianoy AG (2006) The Caspian Sea level, dynamics, wind, waves and uplift of the Earth's crust derived from satellite altimetry. Proceeding of the International Symposium on Remote Sensing and the Pan Ocean Remote Sensing Conference (ISRS 2006 PORSEC), 2–4 November 2006, BEXCO, Busan, Korea, vol 2, pp 973–976 (CD-ROM ISSN 1598-6969).

Lebedev SA, Kostianoy AG (2008) Integrated use of satellite altimetry in the investigation of the meteorological, hydrological, and hydrodynamic regime of the Caspian Sea. Terr Atmos Ocean Sci 19(1–2):71–82

Lunyakova LG (1991) On the forecast average level of the caspian Sea in modern level rise. Trans (Doklady) USSR Acad Sci Earth Sci Sect 314:113–122

Makarenko NG, Karimova LM, Kyandykov YB, Novak MM (2004) Nonlinear dynamics and prediction of the Caspian Sea level. In: Novak MM (ed) Thinking in patterns: fractals and related phenomena in nature. World Scientific, Singapore, pp 91–101 (ISBN 981238822-2)

Meshcherskaya AV, Aleksandrova NA (1993) Caspian Sea level forecast from meteorological data. Russian Meteorology and Hydrology 3:52–60

Meshcherskaya AV, Golod MP, Mirvis VM, Belyankina IG (1997) A method for long-term forecasting of the Caspian Sea level variations with annual lead time from meteorological data. Russian Meteorology and Hydrology 7:69–81

Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitar F (2011) The new ECMWF seasonal forecast system (System 4). ECMWF Technical Memorandum 656, ECMWF, Reading, UK, 49 pp

Mozraji R, Yaghobi M, Kordkolaii JZ (2010) Caspian Sea level prediction by auto fuzzy regression. World Applied Sciences Journal 8(3):288–292

Rodionov SN (1994) Global and regional climate interaction: the Caspian Sea experience. Kluwer, Dordrecht, 241 pp

Schneider U, Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Ziese M (2011) GPCC full data reanalysis version 6.0 at 0.5°: monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. doi:10.5676/DWD_GPCC/FD_M_V6_050

Sheremetevskaya OI, Smirnova KI (1966) Effect of regulated river regime of Volga, Kura and Ural rivers on the annual variation of the Caspian Sea. Oceanology, Academy of Sciences Akad 5:788–793 (in Russian)

Smirnova KI (1974) Current and future outlook of the Caspian Sea. Meteorologiya i Gidrologiya 1:56–62 (in Russian)

USDA1: http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/gr_regional_chart.cfm?regionid=metu&region=&reservoir_name=Caspian. Accessed 20 May 2011

USDA2: http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/gr_regional_chart.aspx?regionid=stans&reservoir_name=Caspian. Accessed 2 May 2013

Vaziri M (1997) Predicting Caspian Sea surface water level. J Waterway Port Coastal Ocean Eng 123(4):158–162