Supplement of Hydrol. Earth Syst. Sci., 20, 659–667, 2016
http://www.hydrol-earth-syst-sci.net/20/659/2016/
doi:10.5194/hess-20-659-2016-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Technical Note: Initial assessment of a multi-method approach to spring-flood forecasting in Sweden

J. Olsson et al.

Correspondence to: J. Olsson (jonas.olsson@smhi.se)

Supplement: Methods

Three new approaches to seasonal hydrological forecasting are presented and compared to the current climatological ensemble procedure currently applied at SMHI: analogue ensemble, dynamical modelling and statistical downscaling. All methods are described in this Supplement.

S1 Climatological ensemble (CE; baseline forecasts)

The current spring flood forecasting practice at SMHI is a climatological ensemble approach based on the HBV hydrological model (e.g. Arheimer et al., 2011). The forecast procedure follows three steps:

1) A set-up of the HBV model, well-calibrated for the specific river basin and location, is run using observed meteorological data (T, P) as input for a period of not less than 12-24 months up to the forecast issue date, typically sometime in February. The state of the HBV model at the forecast issue date will thus reflect the current hydrological conditions in the basin with respect to e.g. streamflow, snow pack and soil moisture.

2) The resulting HBV state from step 1) is then used as the initial state for forecast runs. The input data for the forecast runs are catchment time series of T and P from all available historical years prior to the current one, which covers the period from the forecast issue date until the end of the spring flood period (Figure 3). The time series of each historical year represents one possible weather evolution and results in one possible spring-flood volume (SFV) estimate.

3) The results from all historical years make up a climatological forecast ensemble, which may be expressed in terms of percentiles with different probabilities. In current practice, as well as in this study, the median value of SFV is considered as the spring flood forecast.

In this experiment, the CE is thus made up of all historical years from 1961 to “present”. This means that the CE has 40 members in 2000 and increases in size by one member for each year thereafter (Table S1). The spin-up period used is from 01-01-1961 to “present”. As each new forecast is made, the initial conditions (i.e. model state) are saved and these are used when spin-up for the next forecast date is performed.

S2 Analogue ensemble (AE)

The analogue method has been widely used as a downscaling methodology since Zorita and von Storch (1999). This approach is based on the assumption that similar large-scale atmospheric patterns can result in similar meteorological conditions. The objective is to identify historical years with similar large-scale circulation conditions as the current year, up to the forecast issue date, and then assume that their subsequent weather evolutions are likely realisations over the coming forecast period. In this study we use a period of 1 to 6 months prior to the forecast issue date to identify the analogue years. This choice is motivated by the fact that it will cover the period when snow is accumulated in the catchments and that similar climate behaviour during this period could induce similar snow accumulation. Each identified individual year is an analogue year and a group of them will compose an analogue ensemble.

As compared with the CE method, this approach aims at identifying a reduced ensemble that will provide input to the hydrological model and thus generate the SFV forecast ensemble. To restrict the large number of degrees of freedom of the atmospheric circulation, that would require an unreasonable number of years in the historical data set, two methods are used for the selection of the analogue ensemble. The first one is based on teleconnection climate indices (TCI; Sect. S2.1) and the second on circulation patterns (CP; Sect. S2.2). After selection, the procedure described in Sect. S1 is followed but with the analogue instead of the full historical ensemble in step 2.

The performance of the analogue method is heavily affected by whether the climatic features in the forecasting data were encountered in the training period dataset. This prospect is related to both the archive length available and the presence of non-stationarities. In this study, the archive consists of gridded daily T and P time series in the catchment from 1961 to 1999. This 40-year period is limited in this context, but it is what we have available. Concerning non-stationarities, we here assume climatic stationarity, i.e. that all years in the historical period (1961-1999) are equally representative of the climate in the study period (2000-2010). Any historical trends would imply that years in the latter part of the historical period are more representative than years in the former part. Analyses of P, T and Q data in Vindelälven, however, fail to reveal any significant trends in the historical period. Further, it has previously been shown that using a more recent period in the CE method does not improve performance as compared with using the full historical period (Carlsson and Sjögren, 2003).

Concerning the size of the analogue ensemble, if a large fraction of the total number of years in the archive is selected the AE method will converge toward the CE baseline method and no additional skill will be attained. If only a few years are selected, on the other hand, the uncertainty of the reduced ensemble will be very large. In this study we have not put any restrictions on the ensemble size but used what the methods found. On average ~15% of the archived years were identified as analogues, although it varied between 0% and almost 50% (Table S1).

S2.1 Selection based on teleconnection climate indices (TCI)

The northern hemisphere teleconnection patterns are recurring air pressure and circulation anomalies identified by Barnston and Livezey (1987) using a Rotated Principle Component Analysis (RPCA) of standardised geopotential height anomalies. The prospect of using climate indices for identifying analogue years in a hydrological forecasting context has been previously explored by e.g. Hamlet and Lettenmaier (1999).

The three selected teleconnection indices can be characterized as follows.

- North Atlantic Oscillation (NAO): the positive phase of the NAO is associated with above average temperature (T) and precipitation (P) over Scandinavia during winter (mild, wet winters), while the negative phases tend to be associated with below average T and P (cold, dry winters) (Kushnir, 1999; Hurrell and Dreser, 2010; among many others)

– East Atlantic pattern (EA): the positive phase of the EA pattern is associated with above average winter T, below average winter P (mild, dry winters) in southern Scandinavia, and above average winter P (high snow accumulation) along the Scandinavian mountains and northern Scandinavia (Comas-Bru and McDermott, 2014).

– Scandinavia pattern (SCAND): the positive phase of the SCAND pattern is associated with below average winter P over Scandinavia (dry winters), except over the Scandinavian mountains, where little signal is present. For winter T, this phase is associated with below average in the southern (cold winters) and above average (mild winters) in the northern Scandinavia (Comas-Bru and McDermott, 2014).

The choice of climate indices is motivated by the fact that each of them represents a specific atmospheric circulation that is known to impact the T and P in Scandinavia, and so, its hydrology. As described before, different circulation patterns have different impacts on the

Scandinavian T and P. Patterns occurring concomitantly may increase or decrease their impact on T and P so that it is important to take into consideration the state of different circulation patterns, and thus different climate indices, at the same time.

The years are classified based on the indices' historical mean value TCI_M and standard deviation TCI_S. The current year, with its certain TCI-value, is classified as above normal if $TCI>TCI_M+TCI_S$, below normal if $TCI<TCI_M-TCI_S$ and normal if $TCI_M-TCI_S \leq TCI \leq TCI_M+TCI_S$. The same classification is done for the corresponding periods in each of the years in the historical archive. If the classification of the three different indices is in agreement with the index classification for the year in question for the forecast, the specific historical year is selected as an analogue year. Due to the restricted amount of years in the historical data, it is possible that no analogue years can be identified by this methodology. In this case, analogue years are sought using an agreement with two indices. The number of identified analogue years range between 1 and 19 with a mean value of 7 (Table S1).

S2.2 Selection based on circulation patterns (CP)

Circulation-pattern (CP) analysis is a commonly used tool in climatological and meteorological studies (Hay et al., 1991; Wilby and Wigley 1994). It was initially applied to explain climate variability at a large scale (Barry and Perry, 1973) and later on widely developed to downscale GCM output to local climate in e.g. climate change studies (Wetterhall et al., 2006; Yang et al., 2010).

The method is generally applied to reliable upper-air data at multi-grid, e.g. sea level pressure and geopotential height, to explain recorded observations of e.g. P and T. By differentiating historical observations into several representative CPs, each CP is supposed to represent specific climate conditions in the study area. The CPs are defined based on either professional knowledge of atmospheric motions (subjective classification) or statistical characteristics derived from the observations (objective classification). As the subjective classification is only available in a limited number of regions, the objective classification has been widely developed and used. The objective classification is a semi-automated or automated technique that pertains to mathematical approaches, e.g. hierarchical methods (Johnson, 1967), k-means methods (Mac-Queen, 1967), cluster analysis (Kyselý and Huth, 2005) and correlation methods (Yarnal, 1984). The method that is proposed and investigated here is based on fuzzy-rule logic.

Fuzzy-rule-based classification is built on the concept of fuzzy sets (Zadeh, 1965), using imprecise statements to describe a certain system, in this case the climate system. The classification scheme for CPs follows four steps: 1) transformation of large-scale data; 2) definition of the fuzzy rules; 3) optimisation of the fuzzy rules; and 4) classification of CPs. A detailed description of the methodology used here can be found in Bárdossy et al. (2002) and it is only summarised in the following.

In this work, the anomalies of daily mean sea level pressure (MSLP), $g(i,t)$, from reanalysis data (ERA40 or ERAINTERIM; Sect. 2.2), serves as a predictors according to

$$g(i,t) = \frac{h(i,t) - \mu(i,t')}{\sigma(i,t')} \qquad (1)$$

where $h(i,t)$ is daily MSLP at grid cell i and time t. Variables $\mu(i,t')$ and $\sigma(i,t')$ denote its climatological mean and standard deviation at grid cell i on Julian date t'. The anomaly $g(i,t)$ indicates the deviation of daily MSLP from the long-term climatology.

Every $g(i,t)$ is categorized into one of five groups using fuzzy logic: large positive deviation, small positive deviation, no deviation, small negative deviation and large negative deviation. To determine the fuzzy rule sets best describing the CPs, every rule is optimised with a local variable using a well-designed objective function that explains its statistics in a given region. In this study, P observations in Vindelälven during 1961-1990 are used as local observations. Finally, each CP can be described with a fuzzy rule represented by a vector $V(k) = (v_1(k), v_2(k), \ldots, v_i(k); i = 1, n)$, where n is the number of grid cells in the domain and k stands for the CP.

Two measures are considered as representative statistics, describing the difference from average conditions in terms of P probability (O_1) and P amount (O_2) according to

$$O_1 = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(PW_d\left(CP(n)\right) - \overline{PW_d}\right)^2} \qquad (2)$$

$$O_2 = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{Z(CP(n))}{\overline{Z}} - 1\right|^{1.5} \qquad (3)$$

where N is the total number of days used for the CP optimization. For a day n with a given circulation pattern $CP(n)$, PW_d denotes the probability of P exceeding depth d (generally 0.1 mm, which is used here, but also higher thresholds may be used) and Z denotes the mean P amount. Overbar represents the long-term climatological means of PW_d and Z, in practice calculated as the averages over all N days without regard to classification. The objective functions given by Eqs. (2) and (3) are combined in a weighted sum

$$O_3 = w_1 O_1 + w_2 O_2 \tag{4}$$

where the two weighting factors w_1 and w_2 are determined subjectively to adjust for relative differences in magnitude as well as importance. A higher value of O_3 indicates a better, more distinct classification.

Figure S1 illustrates how the best possible set of CPs is obtained. At each iteration, a set of fuzzy rules is randomly generated. They describe every CP by defining randomly selected membership functions (i.e. degree of daily MSLP anomaly deviated from long-term mean values) at randomly selected locations. Thereafter, the CP time series is generated in order to carry out the performance evaluation in which probability of precipitation and its amount, conditioned on classified CPs, are taken into account. The optimization procedure uses a simulated annealing algorithm. Steps (3) and (4) are repeated until the portion of accepted changes caused by introducing a new set of rules becomes smaller than a predefined criterion.

A successful CP classification should thus fulfil several requirements: 1) the classified CPs should be able to meaningfully explain large-scale climate conditions and their induced local weather phenomena; 2) each CP should be unique and as different from other CPs as possible. When the fuzzy rules that describe every CP have been optimized (see Fig. S1), daily CP time series are generated. The frequency of occurrence and persistence of individual CPs are calculated per month for all historical years as well as the current year to be forecasted. The two most frequently occurring CPs within a period of 1 up to 6 months prior to the forecast issue date are used as a criterion to select the analogue historical years to make forecasts (see Fig. S2). Using more than two CPs did not produce any additional skill.

In a very few cases the CP method, as implemented here, was not able to identify any analogue year (Table S1) and then it was replaced by the CE forecast to have a complete time series of forecasts for the multi-method. This has a negligible impact on the results. On average, 6 analogue years were identified by the CP method.

S3 Dynamical modelling (DM)

In this approach, the daily T and P ensemble of seasonal forecasts from ECMWF (Sect. 2.2) were converted into HBV input. This was done by mapping the daily forecasts from the IFS grid onto the HBV sub-catchments. The mapping was done by areal weighting, based on the catchment fractions covered by different IFS grid cells. The resulting sub-catchment average P and T values were then adjusted to represent different altitude fractions within the catchment. After conversion, the ECMWF forecasts were used to feed the HBV model from the same initial state as used in the current CE procedure, thus following the procedure in Sect. S1 but with forecasts instead of historical years in step 2. As in the CE procedure, the final forecast used in the evaluation is defined by the ensemble median.

S4 Statistical downscaling (SD)

Statistical downscaling is a widely accepted methodology used to connect coarse-scale climate data from GCM to local-scale climate. In this case, large-scale circulation variables are statistically connected to the SFV (e.g. Landman et al., 2001; Foster and Uvo, 2010). The method employed to establish the statistical relationship among the variables is the multivariate procedure known as Singular Value Decomposition (SVD) analysis (Bretherton et al. 1992). SVD analysis is a technique that isolates sets of mutually orthogonal pairs of spatial patterns that maximize the squared temporal covariance between two physical variables (e.g. Cheng and Dunkerton, 1995; Uvo et al., 1998; among many others). The SVD of the cross-covariance matrix of two fields yields two matrices of singular vectors and one set of singular values. A pair of singular vectors describes spatial patterns for each field that have overall covariance given by the corresponding singular value. This praxis has been recently re-named as Maximum Covariance Analysis (MCA).

MCA can be used to derive specific prediction or specification models for a particular point in one variable's field (the predictand; SFV in this case) based on the spatial pattern and/or on the evolution patterns of the anomalous values in the other field (the predictor). From the singular vector pairs, the temporal expansion series of each field can be obtained by projecting the data onto the appropriate singular vector (Bretherton et al. 1992). The relationship between the variables is generated by calculating the matrix of regression coefficients which relates the values of the predictor singular mode temporal amplitudes to the individual points in the predictand field.

In this work, hindcasts for the predictors and historical observations for the predictands were used to define the statistical relationship between them i.e. to calibrate the model. To maximise the robustness of the forecast, multiple forecasts are made with different predictors resulting in an ensemble forecast. The predictors used were forecasted fields (ensemble mean) of large-scale circulation variables with a 2°×2° resolution from two different GCMs (Sect. 2.2) for the 3 months following the forecast issue date (Figure 3). The period used for developing the statistical model (that express the statistical relationship between predictors and predictand) was from 1982 until the year prior to the year being forecasted; thus the training period increased in length with each step forward through the study period.

The initial set of predictors to be evaluated were selected by an initial screening based on previous literature (Nilsson et al., 2008; Foster and Uvo, 2010) followed by an analysis of predictive skill in the historical period. The three best performing predictors for each station and forecast date were selected to comprise the SD ensemble (Table S2).

Figure S3 illustrates how the predictors are selected and the statistical model developed. Each SVD analysis calculates the heterogeneous correlations, how the spatial pattern of one field is correlated with the time series of the other (Fig. S4), and a matrix of regression coefficients relating both fields. The heterogeneous correlations are used as a selection metric and the three highest ranked predictors are chosen to be used in the SD model chain. Thereafter, seasonal forecasts of the selected predictors are downscaled (Fig. S5) using the applicable regression coefficients, obtained during the SVD analysis, where after they are combined to give the forecast of the SFV.

It should be noted that whereas the other methods generate daily discharge time series over the spring flood period, from which SFV is estimated, the SD method directly forecasts the SFV. Therefore forecasts from the SD method are of most interest in the early forecast issue dates and of less interest closer to the spring flood period, as they are not able to provide information about the flood profile.

References

Arheimer, B., Lindström, G., and Olsson, J.: A systematic review of sensitivities in the Swedish flood-forecasting system, Atmos. Res., 100, 275-284, 2011.

Bárdossy, A., Stehlík, J., and Caspary, H.: Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules, Clim. Res., 23, 11–22, 2002.

Barnston, G. and Livezey, R. E.: Classification, seasonality and low-frequency atmospheric circulation patterns, Mon. Weather Rev., 115, 1083-1126, 1987.

Barry, R. G. and Perry, A. H.: Synoptic climatology, methods and applications, Methuen, London, U.K., 1973.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns on climate data. J. Clim., 5, 541–560.

Carlsson, B. and Sjögren, J.: Updating and hydrological long-term forecasts. Report nr. 43 (hydrology), Swedish Meteorological and Hydrological Institute, 601 76 Norrköping, Sweden, 2003 (in Swedish).

Cheng, X., and Dunkerton, T. J.: Orthogonal rotation of spatial patterns derived from singular value decomposition analysis, J. Clim., 8, 2631-2643, 1995.

Comas-Bru, L. and McDermott, F.: Impacts of the EA and SCA patterns on the European twentieth century NAO–winter climate relationship, Q. J. Roy. Meteor. Soc., 140, 354–363, 2014.

Foster, K. L. and Uvo, C. B.: Seasonal streamflow forecast: a GCM multi-model downscaling approach, Hydrol. Res., 41,503-507, 2010.

Hamlet, A. and Lettenmaier, D.: Columbia River streamflow forecasting based on ENSO and PDO climate signals, J. Water Resour. Plann. Manage., 125, 333–341, 1999.

Hay, L., McCabe, G., Wolock, D., and Ayers, M.: Simulation of precipitation by weather type analysis, Water Resour. Res., 27, 493–501, 1991.

Hurrell, J. W. and Deser, C.: North Atlantic climate variability: The role of the North Atlantic Oscillation, J. Marine Syst., 79, 231–244, 2010.

Johnson, S.: Hierarchical clustering schemes, Psychometrika, 32, 141–154, 1967.

Kushnir, Y.: Europe’s winter prospects, Nature, 398, 289-291, 1999.

Kyselý, J., and Huth, R.: Changes in atmospheric circulation over Europe detected by objective and subjective methods. Theor. Appl. Climatol., 85, 19-36, 2005.

Landman, W. A., Mason, S. J., Tyson, P. D., and Tennant, W. J.: Statistical downscaling of GCM simulations to streamflow, J. Hydrol., 252, 221-236, 2001.

MacQueen, J.: Some methods for classification and analysis of multivariate observations, in: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, p. 281–297, 1967.

Nilsson, P., Uvo, C. B., Landman, W., and Nguye, T. D.: Downscaling of GCM forecasts to streamflow over Scandinavia, Hydrol. Res., 39, 17-25, 2008.

Uvo, C.R.B., Repelli, C.A., Zebiak, S.E., Kushnir, Y.: The Influence of Tropical Pacific and Atlantic SST on Northeast Brazil Monthly Precipitation. J. Clim., 11, 551-562, 1998.

Wetterhall, F., Bárdossy, A., Chen, D., Halldin, S., and Xu, C.-Y.: Daily precipitation-downscaling techniques in three Chinese regions, Water Resour. Res., 42, W11423, doi:10.1029/2005WR004573, 2006.

Wilby, R. and Wigley, T.: Stochastic weather type simulation for regional climate change impact assessment, Water Resour. Res., 30, 3395-3403, 1994.

Yang, W., Bárdossy, A., and Caspary, H.-J.: Downscaling daily precipitation time series using a combined circulation- and regression-based approach, Theor. Appl. Climatol., 102, 439-454, 2010.

Yarnal, B.: A procedure for the classification of synoptic weather maps from gridded atmospheric pressure data, Comput. Geosci., 10, 397-410, 1984.

Zadeh, L.: Fuzzy sets, Inform. Control, 8, 338-353, 1965.

Zorita, E. and von Storch, H.: The analog method as a simple statistical downscaling technique: comparison with more complicated methods, J. Clim., 12, 2474-2489, 1999.

Table S1. Ensemble sizes generated by the different single methods. For the TCI and CP approaches, the numbers represent min / mean / max.

CE	TCI6	CP3	DM	SD
2000: 40 2001: 41 . . . 2010: 50	1/1: 1 / 6 / 16 1/3: 2 / 7 / 19 1/5: 2 / 7 / 19	1/1: 0 / 5 / 14 1/3: 0 / 5 / 13 1/5: 0 / 7 / 16	2000-2006: 11 2007-2010: 41	3

Table S2. Final predictors used in the SD method (with atmospheric model in parentheses).

Sorsele	1/1	850hPa temperature (ARPEGE)	850hPa zonal wind velocity (ARPEGE)	850hPa temperature (IFS)
	1/3	850hPa geopotential height (ARPEGE)	850hPa specific humidity (ARPEGE)	Zonal wind stress (IFS)
	1/5	850hPa temperature (IFS)	850hPa specific humidity (IFS)	2m temperature (IFS)
Vindeln	1/1	Surface latent heat flux (ARPEGE)	Surface latent heat flux (IFS)	Surface sensible heat flux (ARPEGE)
	1/3	Surface sensible heat flux (ARPEGE)	Total precipitation (ARPEGE)	850hPa temperature (IFS)
	1/5	Surface latent heat flux (ARPEGE)	Surface latent heat flux (IFS)	850hPa geopotential height (ARPEGE)

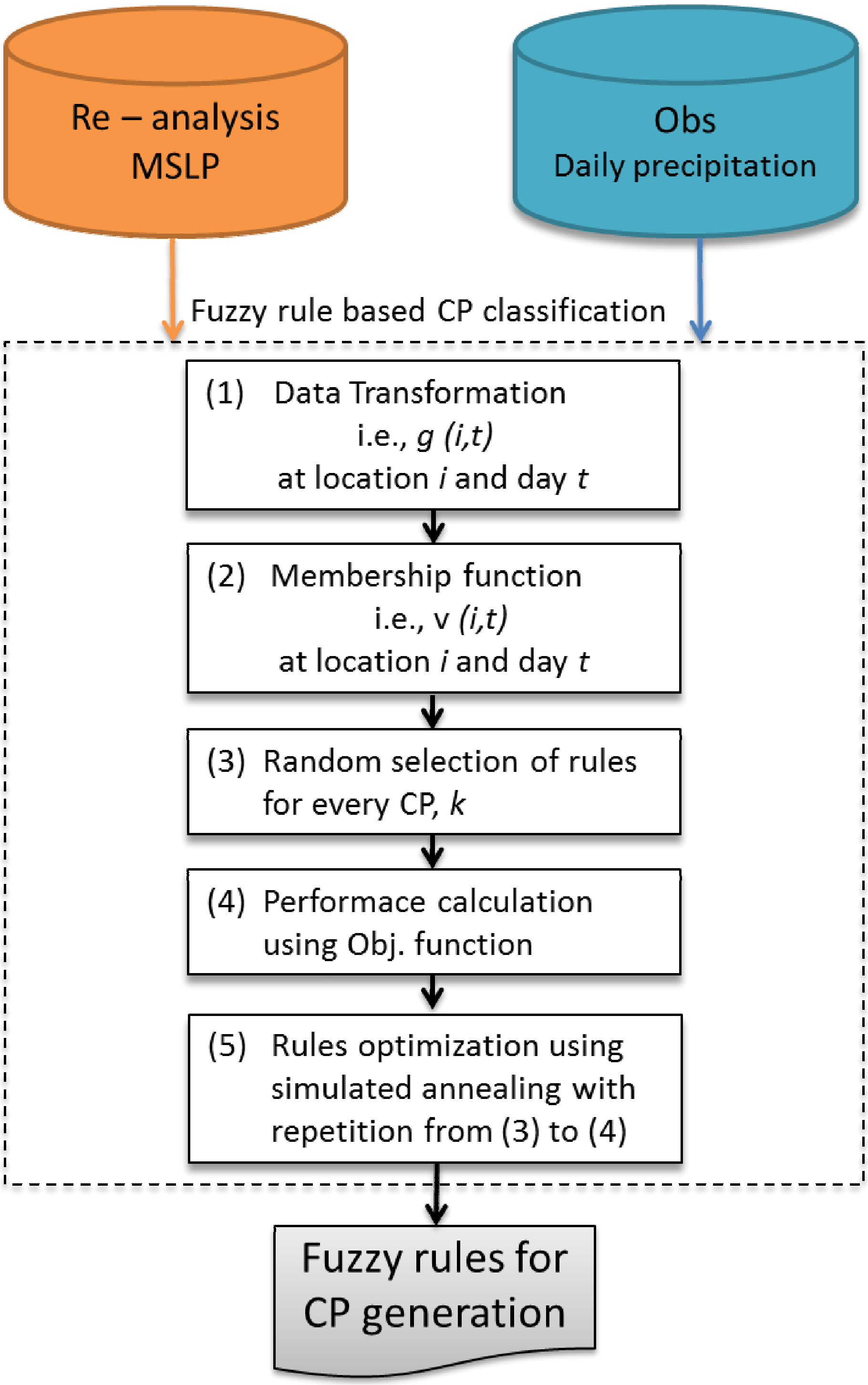

Figure S1. Schematic of the fuzzy-based CP classification process.

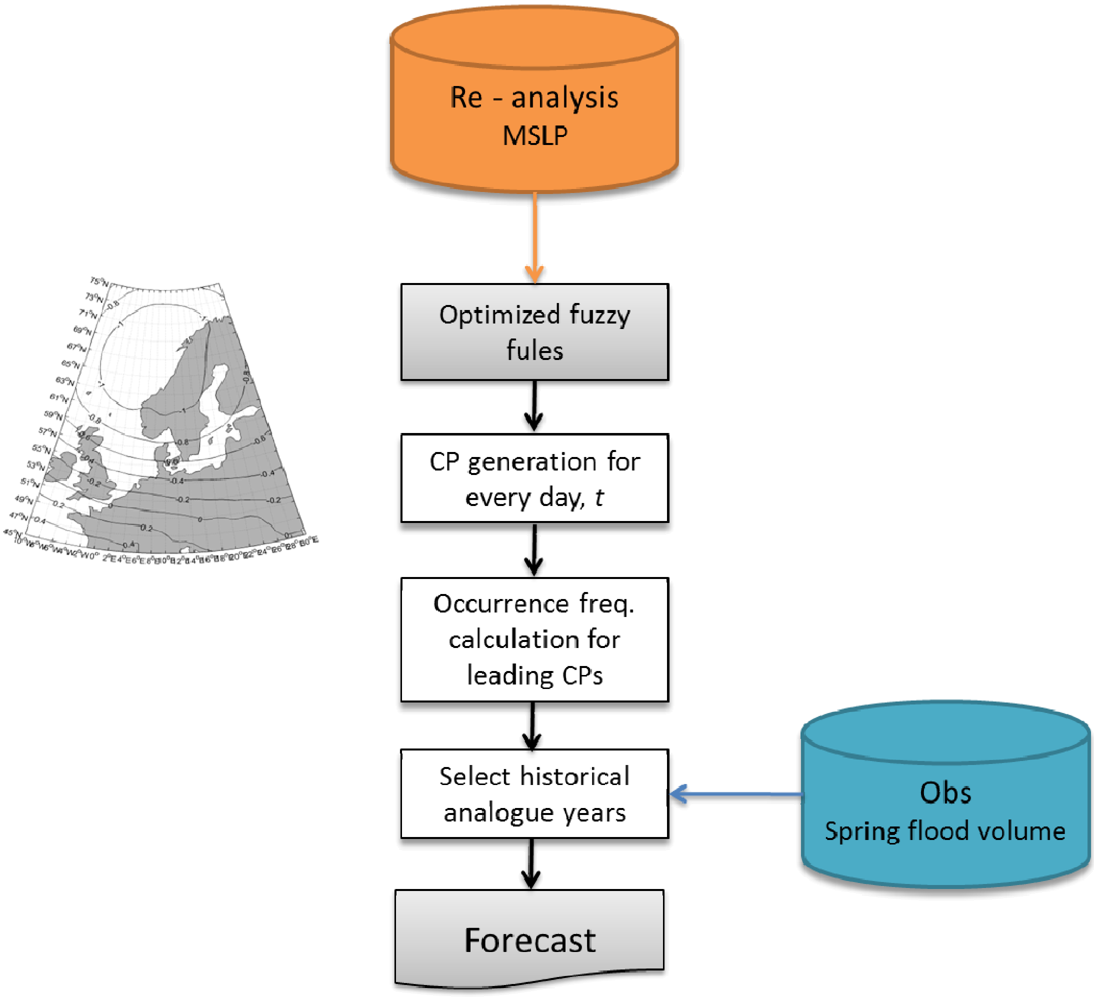

Figure S2. Schematic of the CP-based analogue ensemble forecasting approach.

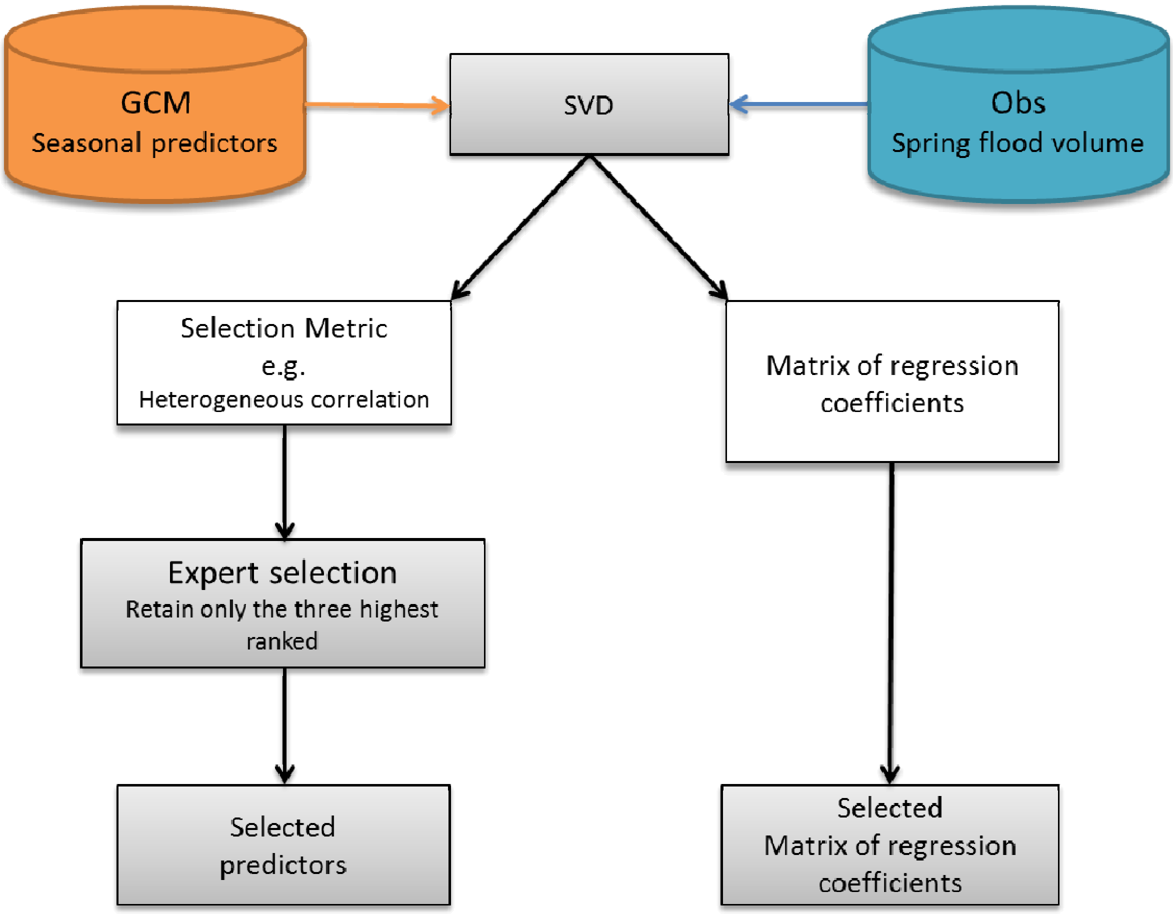

Figure S3. Schematic of the calibration process for the SD model. The three highest ranked predictors and their corresponding transform matrices are chosen for use in the SD forecast model.

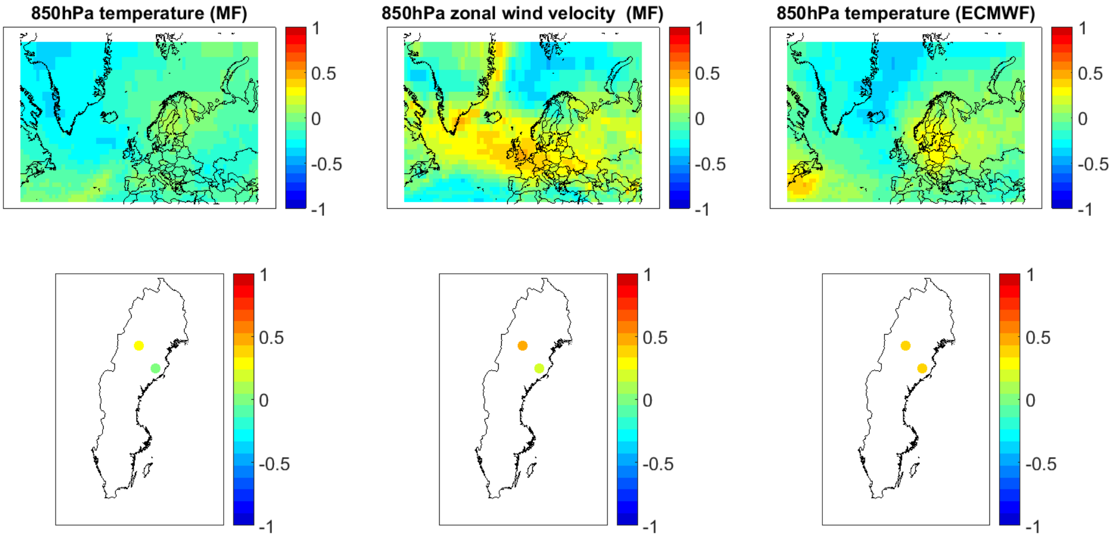

Figure S4. The heterogeneous correlation maps for the three highest ranked predictors (initialised in January) for the station Sorsele. The heterogeneous correlation is a measure of how the spatial pattern of one field is correlated with the time series of the other and is used as a metric for selecting the predictors to use in the statistical downscaling approach.

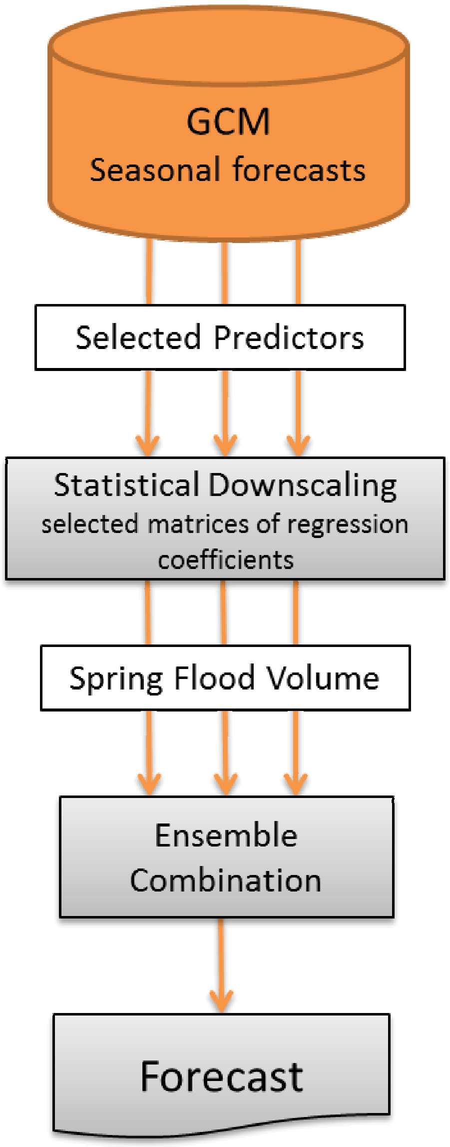

Figure S5. Schematic of the SD-based ensemble forecasting approach.