Article

Seasonal River Discharge Forecasting Using Support Vector Regression: A Case Study in the Italian Alps

Mattia Callegari [1,2], Paolo Mazzoli [3], Ludovica de Gregorio [1], Claudia Notarnicola [1], Luca Pasolli [4], Marcello Petitta [1] and Alberto Pistocchi [5,*]

[1] EURAC Research, European Academy of Bozen/Bolzano, Institute for Applied Remote Sensing, viale Druso, Bolzano 1-39100, Italy; E-Mails: mattia.callegari@eurac.edu (M.C.); ludovica.degregorio@eurac.edu (L.G.); claudia.notarnicola@eurac.edu (C.N.); marcello.petitta@eurac.edu (M.P.)

[2] Department of Earth and Environmental Science, University of Pavia, via Ferrata 1-27100 Pavia, Italy

[3] Research and development (R&D) Unit Suedtirol, Geographic Environmental COnsulting (GECO) Sistema—srl, via Maso della Pieve 60, Bolzano 39100, Italy; E-Mail: paolo.mazzoli@gecosistema.it

[4] Informatica Trentina Spa, via G. Gilli 2, Trento 38121, Italy; E-Mail: luca.pasolli@yahoo.it

[5] European Commission, Directorate-General Joint Research Centre (DG JRC), via E.Fermi 2749, Ispra (VA) 21027, Italy

* Author to whom correspondence should be addressed; E-Mail: alberto.pistocchi@jrc.ec.europa.eu; Tel.: +39-033-278-3658.

Abstract: In this contribution we analyze the performance of a monthly river discharge forecasting model with a Support Vector Regression (SVR) technique in a European alpine area. We considered as predictors the discharges of the antecedent months, snow-covered area (SCA), and meteorological and climatic variables for 14 catchments in South Tyrol (Northern Italy), as well as the long-term average discharge of the month of prediction, also regarded as a benchmark. Forecasts at a six-month lead time tend to perform no better than the benchmark, with an average 33% relative root mean square error (RMSE%) on test samples. However, at one month lead time, RMSE% was 22%, a non-negligible improvement over the benchmark; moreover, the SVR model reduces the frequency of higher errors associated with anomalous months. Predictions with a lead time of three months show an intermediate performance between those at one and six months lead time. Among the considered predictors,

SCA alone reduces RMSE% to 6% and 5% compared to using monthly discharges only, for a lead time equal to one and three months, respectively, whereas meteorological parameters bring only minor improvements. The model also outperformed a simpler linear autoregressive model, and yielded the lowest volume error in forecasting with one month lead time, while at longer lead times the differences compared to the benchmarks are negligible. Our results suggest that although an SVR model may deliver better forecasts than its simpler linear alternatives, long lead-time hydrological forecasting in Alpine catchments remains a challenge. Catchment state variables may play a bigger role than catchment input variables; hence a focus on characterizing seasonal catchment storage—Rather than seasonal weather forecasting—Could be key for improving our predictive capacity.

1. Introduction

The prediction of water availability is a key element for effective water storage management [1]. This is particularly true when the available water storage capacity cannot compensate for reduced discharges over a full season, so that a more precise regulation needs to be planned.

While "quick response" hydrological events such as floods cannot be predicted, in the best cases, with a lead time longer than a few days as they critically depend on catchment input (daily or even hourly precipitation), available water volumes over extended periods (e.g., months) may depend on the rate of depletion of the catchments and are therefore driven by the catchment state, which is much more smoothly varying over time and can be arguably predicted with longer lead time.

In principle, it is possible to obtain monthly streamflow forecasts with long lead time based on catchment water balance models; however, these depend on the availability and reliability of long-term (at least one season [1–3]) forecasts of weather variables. Yang *et al.* [4] showed that the forecasting of streamflow is generally more accurate when catchment storage dominates over catchment input.

Moreover, the data collection, fieldwork, calibration, and validation for the setup of detailed, catchment-specific water balance models may be long and expensive; this poses practical constraints to application in many mountainous areas of the world, where hydrologic information is scarce and sparse, sites are difficult to access, and the operational management of a model for decision support may be difficult due to economic, organizational, and technical limitations (as discussed in e.g., [5]).

An alternative to water balance models is given by data-driven ("black box") models based on statistical relationships between time series of a target (e.g., monthly water discharge) and predictors.

Traditional black box models such as linear approaches [6] are being challenged recently by non-linear machine-learning methods such as the support vector regression (SVR) [7], reported by several authors to be more effective than previous methods in hydrologic time series forecasting [8]. The better performance of non-linear models may be related to the non-linear behavior shown by discharge as a function of catchment input and state variables.

Among the available retrieval algorithms, SVR is often adopted in the field of bio-physical parameter retrieval for the following characteristics: SVR is able to handle complex and non-linear problems, can manage different kinds of inputs, and can reach high level performance even in cases where few data are available [9,10].

The first two characteristics are in common with other retrieval strategies. Both non-parametric methods, such as Artificial Neural Network (ANN) and Genetic Algorithm, and parametric methods like Bayesian approaches can handle the non-linearity between input and output variables. However, the main limitation is related to the availability of a large training database to perform a robust retrieval. SVR overcomes this limitation because it is based on a geometrical concept. In this context, for ANN and statistical approaches, the main aim is to populate the feature space with as much data as possible, so that all the possible combinations of input and output features are available to build the mapping functions. On the other side, a retrieval-approach-based SVR identifies the margins of the tolerance tube around the input data. Once the margins have been identified, the mapping function can be identified without the need for more data [7].

In recent years, this type of model has been applied in several cases [11–15], particularly with reference to reservoirs [14], indicating a potential role in practical water resources management. Most applications, however, predict the streamflow at a given month on the basis of streamflows in previous months without using catchment state variables [16–18]. Dehghani *et al.* [19] highlighted the importance of snow water content for the prediction of streamflow in the Karun basin in Iran, but made use of basin air temperature as a proxy for snowmelt instead of referring specifically to snow-related variables. Wang and Fu [20] made use of soil moisture as a catchment variable for the improvement of adaptive monthly and sub-monthly streamflow predictions in the Yangtze River, and found that this variable appreciably improved the forecasts.

Examples of analyses focused on the catchment state, rather than the catchment input, have been published with reference to the summer response of spring to autumn precipitation [21,22], or to the summer runoff and vegetation indexes' response to groundwater storage [23], to some extent enabling long lead-time forecasting.

It has long been recognized that, in Alpine and similar settings, summer flows may be controlled by the winter accumulation of snow in the catchment [24]. The role of mountain snow and glacier water in sustaining runoff, even in the presence of trends of reduced precipitation, has been observed in many regions in the last years [25], and is increasingly included in hydrologic models applied to mountain areas [26,27].

This contribution examines the performance of SVR models for the forecasting of monthly streamflow in catchments of a European Alpine region with a lead time of one, three, or six months. Unlike many similar forecasting studies that make use of discharges only, we attempt to improve the model's predictive accuracy by introducing catchment input and state variables (climatic signals, and the snow-covered area) as predictors, as for a water balance model. Arguably, the SVR technique may be able to extract from the data the unknown, non-linear relationships between catchment input, state, and output, when a proper water balance model is difficult to identify. In the remainder of the paper, after introducing the study region and providing an intuitive introduction to the SVR technique, we discuss the results obtained in the prediction of streamflow with one, three, or six months lead time and the potentials and limitations of the approach, and provide some concluding perspectives.

2. Materials and Methods

2.1. Study Area and Available Discharge Data

The Alps, as well as other mountain regions in the world, are of strategic importance for water resources [28], given that they are a climate change hot spot and a key context where adaptation is required [29].

The study region considered here is the upper Adige catchment closed at Bronzolo/Branzoll, in the Autonomous Province of Bolzano/South Tyrol, Northern Italy (Figure 1). This is an alpine watershed with a drainage area of approximately 7400 km², with elevations ranging from about 200 m a.s.l. at the southern valley bottoms, to around 3900 m a.s.l. in the western upper ranges. The Adige is the second largest Italian river and is fed by a large tributary, the Isarco/Eisack, draining about half of the catchment. The area is covered by a relatively dense network of hydrographic and weather stations managed by the Province's Hydrographic Office (Figure 2). Among all the available water discharge stations, for the present study we considered just those in which the discharges have been continuously recorded for at least the last 20 years. Figure 1 shows the catchments closed at the selected stations, and Table 1 lists their main attributes.
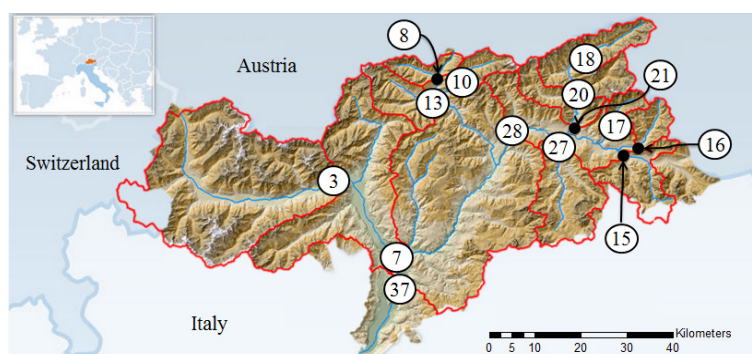

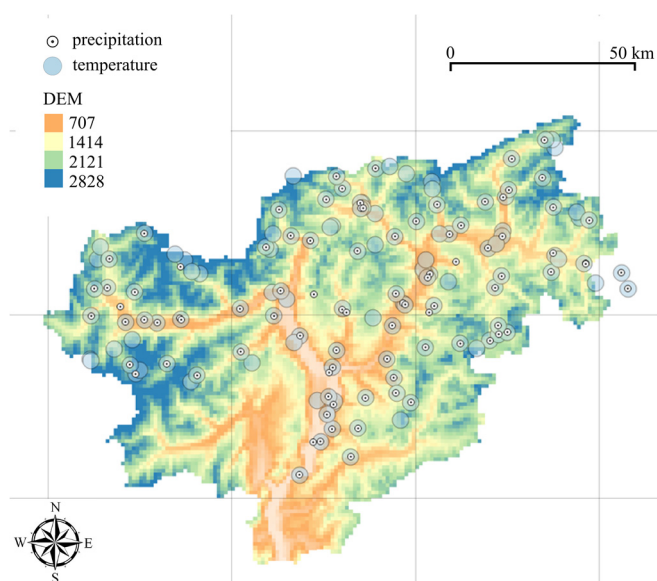**Figure 1.** South Tyrol in northern Italy, location of the analyzed catchments.

Figure 2. Available measurement stations in South Tyrol.

Table 1. Main characteristics of selected discharge measurement stations.

Station Name	Hydrographic Office Identifier	UTM 32N Coordinates (m)	Catchment Area (km²)	Min–Max–Mean Altitude (m a.s.l.)	Min–Max Discharge Q (m³/s)
Adige at Tel	3	$X = 659{,}075\ Y = 5{,}170{,}831$	1,676	510–3,893–2,011	11.50–84.39
Adige at Ponte Adige	7	$X = 676{,}882\ Y = 5{,}150{,}246$	2,705	240–3,893–1,907	22.68–114.96
Rio Fleres at Colle Isarco	8	$X = 685{,}584\ Y = 5{,}201{,}188$	149	1,047–3,479–2,167	0.61–9.64
Rio Vizze at Novale	10	$X = 692{,}269\ Y = 5{,}200{,}902$	108	1,375–3,500–2,186	0.76–13.03
Rio Ridanna at Vipiteno	13	$X = 685{,}246\ Y = 5{,}195{,}138$	207	939–3,456–1,927	1.41–20.37
Rienza at Monguelfo	15	$X = 737{,}906\ Y = 5{,}182{,}252$	264	1,096–3,217–1,827	2.16–12.50
Rio Casies at Colle	16	$X = 742{,}055\ Y = 5{,}184{,}156$	117	1,198–2,825–1,960	0.86–6.22
Rio Anterselva at Bagni Salomone	17	$X = 735{,}133\ Y = 5{,}190{,}077$	83	1,091–3,425–2,036	0.55–5.27
Aurino at Cadipietra	18	$X = 726{,}593\ Y = 5{,}208{,}875$	149	1,047–3,485–2,167	1.46–27.06
Aurino at Caminata	20	$X = 725{,}134\ Y = 5{,}198{,}678$	420	845–3,485–2,115	2.69–39.25
Aurino at San Giorgio	21	$X = 723{,}868\ Y = 5{,}189{,}015$	613	819–3,485–2,038	5.40–56.18
Gadera at Mantana	27	$X = 719{,}819\ Y = 5{,}184{,}323$	389	813–3,120–1,856	3.58–18.01
Rienza at Vandoies	28	$X = 706{,}592\ Y = 5{,}188{,}228$	1,920	735–3,217–1,861	14.32–100.07
Adige at Bronzolo	37	$X = 677{,}940\ Y = 5{,}142{,}627$	6,923	228–3,893–1,804	61.07–330.28

2.2. Introduction to Support Vector Regression (*SVR*)

2.2.1. Model Formulation

The statistical prediction of a variable $y(t + \Delta t)$ at an instant Δt from current time t may be framed as a particular case of regression problem, the goal of which is to identify a function $f(x)$ of a vector $x = (p_1, p_2, \ldots, p_m)$ of m (known) explanatory variables, so that $y(t + \Delta t) = f(x)$. The variables $p_1, \ldots, p_m$ forming vector **x** may include current and past data samples $[y(t), y(t-1), \ldots, y(t-q)]$ for $q \leq m$, as well as additional numerical variables of any kind.

The function is identified on the basis of reference samples for which both $y(t + \Delta t)$ and **x** are known. Let us refer to a set of N samples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ where each x_i and y_i, for $i = 1, \ldots,$ N, represent a particular value of the variable to be predicted, $y(t + \Delta t)$, and the corresponding vector of explanatory variables considered.

Traditional regression methods seek a function that minimizes prediction errors considering all N samples. The support vector regression (SVR) technique [30–34], instead, aims at finding the simplest function that can fit all the data while minimizing the sum of prediction errors above a predefined threshold. A review of the concepts and characteristics of these techniques with specific reference to hydrology is provided by [34]. For the sake of a quick illustration of the technique, let us first refer to the linear case, where the function to be estimated takes the form:

$$f(x) = \langle w, x \rangle + b , \tag{1}$$

w being a vector of weights of the same dimensionality of vector x of explanatory variables, and b a scalar. Here, the operator $\langle \cdot, \cdot \rangle$ refers to the scalar product of vectors. The "simplest function" is the function depending the least on input variables, and can be found by minimizing the norm $\|w\|^2 = \langle w, w \rangle$. As we want to minimize the sum of errors higher than a predefined threshold ε, we introduce the ε-insensitive loss function L_ε defined for each sample i as:

$$L_{\varepsilon i} := \begin{cases} 0 & if \ |f(x_i)-y_i| \leq \varepsilon \\ |f(x_i) - y_i| - \varepsilon & otherwise \end{cases}. \tag{2}$$

From the loss function computed for each of the *N* reference sample, the SVR problem can be written as the following convex optimization problem:

$$minimize \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} L_{\varepsilon i}, \tag{3}$$

where constant C is a model parameter, driving the tradeoff between the complexity of the function and tolerance of empirical errors. A graphical representation can be found in [31].

After appropriate mathematical manipulation, it can be shown that the weights can be written as a function of the reference sample vectors:

$$w = \sum_{i=1}^{N} \alpha_i x_i \tag{4}$$

and, consequently:

$$f(x) = \sum_{i=1}^{N} \alpha_i \langle x_i, x \rangle + b, \tag{5}$$

where α_i, $i = 1,...,$ N, are constant coefficients.

Coefficient *b* can be estimated based on the Karush–Kuhn–Tucker (KKT) conditions [31], while the two parameters C and ε represent "tuning" parameters. In general, the SVR does not assume *f*(*x*) to be linear, but stipulates it to be in the form:

$$f(x) = w \cdot \Phi(x) + b, \tag{6}$$

$\Phi(x)$ being a (non-linear) unknown function. In this case, it can be shown that the weights can be written as a function of the reference sample vectors:

$$w = \sum_{i=1}^{N} \alpha_i \Phi(x_i) \tag{7}$$

and, consequently:

$$f(x) = \sum_{i=1}^{N} \alpha_i \langle \Phi(x_i), \Phi(x) \rangle + b, \tag{8}$$

where α_i, $i = 1,..., N$, are constant coefficients. It should be noted that the identification of *f*(*x*) does not require us to know the function $\Phi(x)$, but only the scalar products $\langle \Phi(x_i), \Phi(x) \rangle$ for , $i = 1,..., N$. Under certain assumptions [31], the scalar product $\langle \Phi(x_i), \Phi(x) \rangle$ can be rewritten using non-linear functions (called "kernels") $k(x_i, x)$. A commonly used kernel is the Gaussian radial basis function, uniquely identified by the single parameter represented by the kernel width γ, which represents an additional model parameter. In this way, the optimization follows procedures similar to those in the linear case, but function $f(x)$ is allowed to be non-linear and of arbitrary complexity. Further details on the method can be found in [31], while a more intuitive description of the approach specifically targeted to hydrological applications is given in [34].

2.2.2. Model Training and Feature Selection

No standard methods exist for calibrating C, ε, γ. The choice of these parameters is driven by the characteristics of the problem itself. Operationally, the SVR prediction process can be schematically divided into two parts, an off-line and an on-line procedure (Figure 3).
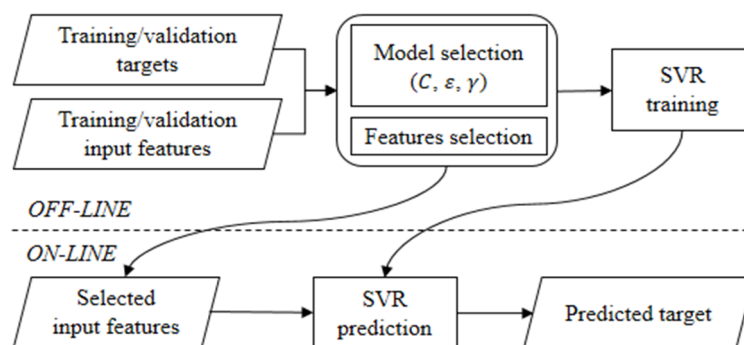

Figure 3. Support vector regression (SVR) general scheme separated between off-line and on-line phase. During the off-line phase the SVR is trained and the prediction function is estimated. The on-line phase refers to the operative part of the method, in which the prediction phase can accept a proper set of inputs for estimating the discharge in the target month.

The off-line part consists of SVR model training and supports the selection of relevant input variables. A common strategy for model training is to divide the available reference dataset into two subsets, one called "training set," which is used to find the optimal function for a fixed combination of parameters C, ε, γ, and one called "validation set," on which the performance in terms of a chosen empirical metric is computed (within this work, we used a set of metrics as explained in [35]). Instead of dividing the reference dataset into two subsets (*i.e.*, training and validation), we preferred to adopt a *t*-folds cross-validation procedure, which consists of randomly dividing the available reference samples into *t* subsets. Iteratively, $t − 1$ subsets are used for training, whereas the remaining subset is used for validation. In this way, all the samples are used for both training and validating the model. The cross-validation strategy may increase the representativeness of training and validation sets, especially in the case of limited samples being available [35].

The so-called "feature selection" operation aims at identifying those, among all explanatory variables ("features") considered, which are actually informative: The SVR model is trained iteratively with different combinations of features, and those combinations leading to the best model performances are eventually retained for further verification. Features that are found to be uninformative, or even worsen the performance of the model, are discarded.

2.2.3. Model Testing

Once the SVR model is properly trained with the selected input features, the SVR prediction can be used "on-line" for estimating a target variable using the set of selected input features. The estimates are compared to a set of known target variables (usually called "test set"), different from the training and validation ones, in order to independently assess the prediction accuracy of the method implemented.

The reference set can be built by exploring the whole time series, extracting at each instant the input feature vector by looking at the past and current values and the associated future target. For instance, monthly discharge at the instant $t + \Delta t$ may be predicted using monthly discharge at month t, $t - 1$, *etc.*, as well as the other explanatory variables at month t, $t - 1$, *etc.*

2.3. Input Data Considered

We want to predict monthly discharges at month $t + \Delta t$, with Δt equal to one, three, or six months, using as input features the discharges measured at the current (t) and past ($t - 1,\ldots, t - q$) months (q being the number of months in the past that we want to consider), as well as other known variables that we assume may allow predicting discharges. These were chosen among the following:

- Catchment state variables: Snow storage represented by snow-covered area (SCA) as a proxy;
- Climate signals related to possible catchment input (temperature and precipitation): Indexes North Atlantic Oscillation (NAO), Wave Amplitude Index (WAI), Baroclinic Amplitude Index (BAI), Standardized Precipitation Index (SPI), and air temperature, briefly introduced and discussed hereafter;
- Catchment characteristics: Area, minimum, maximum, and mean elevation.

SCA has been extracted from snow maps obtained with MODIS images, which are freely provided by NASA (http://modis.gsfc.nasa.gov/) and acquired through a dedicated antenna available at EURAC Research, Bolzano. The snow products have been derived with a specific algorithm adapted to mountain areas [36,37]. This algorithm takes 250 m resolution, atmospherically corrected surface reflectance MOD09GQ-MYD09GQ (Terra-Aqua satellites) products as input. Cloud detection is carried out using 500 m resolution MOD09GA-MYD09GA products and 1 km resolution MOD021KM-MYD021KM products. MOD03-MYD03 are used for correct geolocation. Daily snow maps at 250-m resolution obtained with the algorithm described in [36] allow improved snow cover mapping with respect to the 500 m standard MODIS product, especially in mountainous terrain where snow cover shows high spatial variability due to highly irregular topography [37].

Climate signals announcing dry and wet periods can be derived from atmospheric circulation patterns. As shown in several studies, dry or wet conditions of a region can be connected to some particular large-scale climatic indicators, like the El-Nino Southern Oscillation (ENSO) or NAO [24,38–40], although a certain indicator may have different relevance in predicting climate in separate, but bordering, regions [24].

Other authors have shown a clear relationship between the phenomenology of meso-scale atmospheric conditions and extreme wet and dry events, especially in the mid-latitudes (see [41]). The SPI [42–44] has been proposed for application at different time scales [42,45,46], because it quantifies relative weather dryness and wetness, allowing comparison of climatic conditions in areas governed by different hydrological regimes.

It has been observed that the SPI computed over one month may correlate with unregulated stream flow, while the SPI over a longer period may correlate with reservoir storage [1,2]. SPI prediction based on statistical models has been conducted e.g., in Sicily [47] and Iran [48].

The WAI [49–51] and the BAI [52] are two additional well-established climatological indexes related to large-scale atmospheric circulation. The WAI is used to describe regimes of circulation in

extra-tropical regions and can be used to identify atmospheric blocking events in Europe [49]. Atmospheric blocking leads to a stagnation of weather patterns, in turn possibly associated with floods, droughts, above- or below-normal temperatures, and other weather extremes. It is important to recognize a blocking pattern in its initial development. The BAI is related to heat fluxes and precipitation activity in the mid-latitudes [52], which may transport heat and momentum vertically and horizontally at synoptic scales in the mid-latitudes. These transfers may explain most of the high frequency variability in the extratropics [53]. The dominance of these climatic signals can be used to capture global trends that yield dry or wet periods [41].

Climatic signals may be related to catchment state and slow hydrological responses: Long lead-time forecasting has been successfully attempted on seasonal summer streamflows using ocean-atmospheric variables [54–56], suggesting that similar studies might be extended to a finer level of temporal detail (e.g., monthly). Studies on a two-century time series of discharges from the Po River, northern Italy, have highlighted a relationship between both winter precipitation and summer stream flow and the NAO [57].

In this case study, we considered SPI, WAI, BAI, NAO, and air temperature, the last deemed generally representative for snow melting.

Other variables may be in principle be added to the above list. For instance, cloud cover and catchment orientation (aspect) may be related to snowmelt, while vegetation and soil moisture indicators are representative of catchment state. In the catchments considered for the present analysis, the aspect varies at a relatively fine scale following the mountainous morphology, while the average aspect of each catchment is relatively constant. Vegetation and soil moisture were considered to be less representative than snow cover for the mountainous region of interest, as they account for relatively little water storage compared to snow. Finally, cloud cover was deemed a more indirect indicator than temperature. In a quest for balance between comprehensiveness of the model setup and complexity of model input preparation, we decided to limit the input variables to those discussed above.

In addition, a special input feature was also identified in the long-term average monthly discharge of the month of prediction. The latter was computed from the time series of discharges, considering the 10 years before the start of the prediction period. This choice, in general, puts an additional constraint on the selection of the hydrological stations, which need to be limited to those continuously monitored for at least 20 years.

2.4. Experimental Setup

In the study region, there are about 40 discharge gauging stations. However, many of these have not been monitored for a sufficiently long and continuous time period. In order to have a robust discharge prediction model, we chose to train a specialized SVR model only for the catchments with stations continuously monitored for at least 10 years, corresponding to a number of samples higher than 120; these correspond to the 14 measurement stations in Table 1.

The case of gauging stations with short or discontinuous monitoring is quite common in practice. Therefore, it is of interest to develop SVR models that may be applied to a catchment without specific calibration. A way to develop one such model is to consider all model training catchments together, while including catchment morphological descriptors as input features. In this way, we may hope to apply the model to any other catchment similar to the ones used for training. Such "regionalized" model

calibration does not need to be repeated for each new catchment. For this reason, we complemented the catchment-specific models in the study region with a catchment-independent model, as explained in Section 2.4.2.

2.4.1. Data Preparation

The input dataset is formed by a time series of SCA, climatic signals (*i.e.*, NAO, WAI, BAI, SPI, and temperature), and measured monthly discharges.

The WAI and BAI indexes have been calculated from NCEP reanalysis. The NAO index has been taken by the Climate Prediction Center (CPC) of NOAA (National Oceanic and Atmospheric Administration) [58]. SPI and air temperature maps were obtained by interpolation of values measured at precipitation and temperature stations in the test area (Figure 2) over a 1-km resolution grid using ordinary kriging. A time series of the monthly mean SPI and temperature has been computed for each of the watersheds of the discharge stations by spatially averaging these maps (Figure 1).

We prepared the time series of all input features and target variables for each catchment for the period between November 2002 and May 2012, as the availability of the Aqua and Terra MODIS satellites, used to derive snow maps and thus extract the SCA, starts from 2002.

We selected three years to form the test set, while the remaining six years and seven months form the reference samples. This choice better ensures the independence between reference and test samples with respect to a random selection between the whole time series. Moreover, the choice of the reference set years as well as the choice of the test set years were made so as to ensure capturing as much inter-annual variability of discharges as possible. This is expected to yield a more reliable training of the SVR and a more realistic evaluation of the prediction accuracy on the test set.

The reference samples have been randomly divided into three subsets in order to set up the *t*-fold cross validation configuration explained in the above section. Then a multi-objective model selection procedure (as described in [35]) has been used to optimize with respect to more than one single figure of merit. In this approach both the percentage root mean square error (RMSE%) and mean square error (MSE) at the same time have been considered in the optimization.

2.4.2. Feature Selection

Feature selection is conducted together with model calibration during the off-line phase of SVR training (Figure 3). The feature selection procedure described hereafter was only applied to three representative catchments, a large one (n. 37, Adige at Bronzolo, 6923 km^2), a small one (n. 8, Rio Fleres at Colle Isarco, 149 km^2), and a medium-sized one (n. 28, Rienza at Vandoies, 1920 km^2), because of the time required for the examination of the high number of combinations of input features. With this choice, the procedure is faster and the representativeness of the selected features can be tested independently on the remaining 11 catchments of the study area.

The feature selection procedure was divided into three steps (Figure 4) and was based on the percentage root mean square error (RMSE%) of predicted *versus* observed discharges calculated on the validation samples.

The first step aims at selecting the input features related to catchment state, *i.e.*, the discharges of antecedent months and the snow accumulated in the catchment, described in our problem by the SCA

time series. We tested as input several combinations of discharges of one to 12 antecedent months, and SCA from one to four months before the time of prediction. The same simulations have been repeated including and excluding the long-term average discharge of the month of prediction as input feature. An SVR was trained and tested for several combinations of input features, so as to identify their possibly optimal combination.
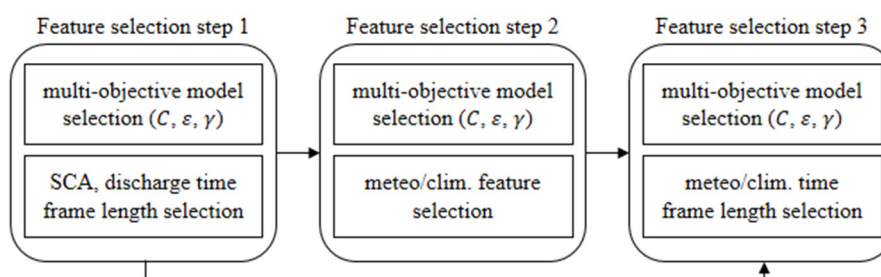

Figure 4. Feature selection scheme. The first step is relative to the selection of the time series length of the catchment state variables SCA and discharge. The second and the third step deal with the selection of the climatic and meteorological parameters that describe precipitation and temperature.

In the second step we select the climatic signals possibly relevant as input features. In order to better understand which parameter is actually informative, we added to the input feature selected in step 1 those variables that referred to the target month, thus mimicking an ideal forecast (*i.e.*, with no forecasting errors). The configuration that exhibits the highest performance on the validation set is selected.

In any operational prediction, the meteorological and climatic variables of the target month are not available, but they need to be predicted. Once we have identified those climatic signals relevant for the prediction, in step 3 of the feature selection process we replace the SVR input features that refer to the actual future value of the input features selected in step 2 with a combination of their values in a number of months before prediction; step 3 is therefore responsible for finding the optimal number of antecedent months for these variables.

2.4.3. Setup of a Catchment-Independent Support Vector Regression (SVR) Model

Following the approach described above, we have set up an SVR model for each individual catchment in the study region based on the features selected for three representative catchments and the parameters C, ε, and γ optimized individually for each catchment. In principle, any additional catchment in the area would require calibrating a specific SVR model. An SVR model calibrated on all catchments together on the basis of catchment features may be of practical interest as it would allow predictions on a generic catchment in the region. Therefore we have calibrated an SVR model for all catchments together, considering as target variables the normalized monthly discharges (discharges divided by watershed area), in order to have comparable quantities given the large variation of catchment sizes.

In the setup of an SVR model for all catchments together, catchment characteristics such as area and minimum, maximum, and average elevation were considered as input features along with the SCA and climate signals discussed above. On the contrary, the long-term average monthly discharge of the month of

prediction has not been used as an input feature because its use implies the availability of a long discharge time series, in contrast with the goal of the approach—To also address poorly gauged catchments.

We iteratively simulated one of the 14 catchments as the one without a long time series available and thus we trained an SVR using the remaining 13 catchments considering the discharges normalized by catchment area, and SCA.

It is anticipated that an SVR for all catchments together may suffer from lower prediction accuracy with respect to a specialized SVR trained on a single catchment. For the setup of an SVR model for all catchments together, we followed the same procedure described above, by iteratively excluding one catchment at a time from the training phase and using it as a test set, so as to assess the performance obtained for the test samples of all the 14 catchments considered in this study.

3. Results and Discussion

The long-term average monthly discharge of the month of prediction, also used as an input feature, is the expected value of the discharge itself, and may be regarded as a benchmark for the performances obtained with different SVR configurations. Indeed, any prediction performing worse than this benchmark is basically of no use. Table 2 reports the RMSE% and the coefficient of determination (R^2) computed on the whole set of discharges for the 14 stations, by comparing observations with the long-term average discharge of the target month (*i.e.*, the benchmark predictor). The RMSE% is defined as:

$$RMSE\% = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{A-P}{A}\right)^2} \times 100 \qquad (9$$

where A is the actual value and P the predicted value of the target. The "reference" RMSE% corresponding to the benchmark is equal to 33% ($R^2 = 0.7$) and sets the acceptance threshold for all other predictors. The RMSE% and R^2 of the long-term average predictor are obviously independent of the lead time. In the three-step feature selection procedure, any combination of features yielding a RMSE% higher than these values should be discarded.

Predictions at one month lead time with the features shown in Table 2 have a lower RMSE% with respect to the benchmark predictor (Figure 5). The mean RMSE% over the 14 catchments is consistently clearly reduced. With a prediction lead time of three months the improvement is less apparent, while with a lead time of six months the accuracy seems to be similar to the benchmark. The best performing SVR configuration for a lead time of one month reduces the RMSE% to 22% and increases R^2 to 0.76; for a lead time of three months, the corresponding figures are 28% and 0.68 (*i.e.*, a slight worsening), respectively; for a lead time of six months, we find RMSE% = 31% with no improvement on R^2.

The feature selection steps allow us to develop an understanding of the contribution of each single input feature of the SVR model. This is shown in Table 2. Step 1 of the feature selection procedure, in particular, allows us to identify the importance of antecedent monthly discharges, SCA, and the long-term average discharge of the target month.

SCA appears to be quite informative. Figure 6 shows the comparison between the RMSE% obtained on the individual catchments with SCA as input against the best configuration without SCA. As shown by the graph, the RMSE% decreases significantly when using SCA compared to using antecedent discharges only,

except for the lead time of six months, in which case differences vanish. This is somehow expected as, at a longer lead time, the catchment state before prediction described by the SCA assumes less relevance.

Table 2. Performances of different SVR model configurations for single stations. The fourth and fifth column refer to the mean RMSE% and R^2 computed over the test samples of the 14 watersheds. In the third column "disch" stand for discharge, "dischAvg10" for the 10-year average discharge of the target month, and "temp" for air temperature. The notation—*n*: 0 stands for a time series of length $n - 1$ starting from the n-tested or the previous month to the current month from where the discharge prediction is done. As a benchmark, we considered the method based on the average discharge calculated over the 10 previous years, whose performances are shown in the first row. In addition, for a prediction lag of one and three months the results obtained with the AR model are shown.

Prediction Lag	Prediction Method	Input Features	Mean RMSE%	Mean R^2
All	10-year average	--	33	0.70
1	SVRsingle noMeteo	disch0, SCA-2:0, dischAvg10	22	0.76
	SVRsingle noMeteo noSCA	disch-11:0, dischAvg10	28	0.72
	SVRsingle noMeteo noDischAvg	disch-11:0, SCA-2:0	31	0.72
	SVRsingle meteoIdeal	disch0, SCA-2:0, dischAvg10, SPI, temp	21	0.79
	SVRsingle meteo	disch0, SCA-2:0, dischAvg10, SPI-1:0, temp-11:0	25	0.76
	AR model	disch-11:0	33	0.69
	AR model	disch-6:0	36	0.48
	AR model	disch-3:0	33	0.50
	AR model	disch0	33	0.49
3	SVRsingle noMeteo	disch0, SCA-1:0, dischAvg10	28	0.68
	SVRsingle noMeteo noSCA	disch-10:0, dischAvg10	33	0.69
	SVRsingle noMeteo noDischAvg	disch-9:0, SCA-2:0	33	0.68
	AR model	disch-11:0	35	0.70
	AR model	disch-6:0	69	0.03
	AR model	disch-3:0	56	0.01
	AR model	disch0	61	0.01
6	SVRsingle noMeteo	disch-10:0, SCA0, dischAvg10	31	0.70
	SVRsingle noMeteo noSCA	disch-11:0, dischAvg10	32	0.70
	SVRsingle noMeteo noDischAvg	disch-11:0, SCA-3:0	32	0.69

In addition, the long-term average discharge of the target month improves the prediction accuracy for all lead times (one, three, or six months): The mean improvement of the RMSE% over all the catchments analyzed, with respect to the corresponding model configuration without this input, is equal to 9%, 5%, and 1%, respectively, for a prediction lag equal to one, three, and six months (see Table 2). This indicates the capacity of the SVR model to better exploit this input at shorter lead times. At six months' lead time,

the antecedent 12 months' discharges are shown to contain more or less the same information with respect to the long-term average discharge (see Table 2).
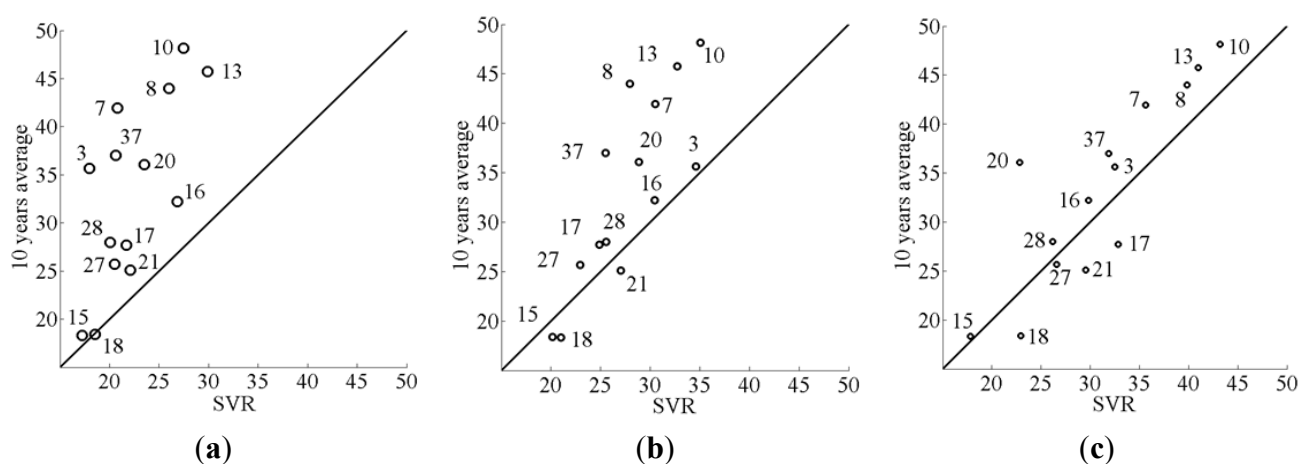

Figure 5. Comparison of the RMSE% obtained on the test set between the SVR method proposed (*x* axis) against the simple method based only on the average discharge of the target month computed over 10 previous years. In the scatterplots (**a**), (**b**), and (**c**) the prediction lag is equal to one, three, and six months, respectively.
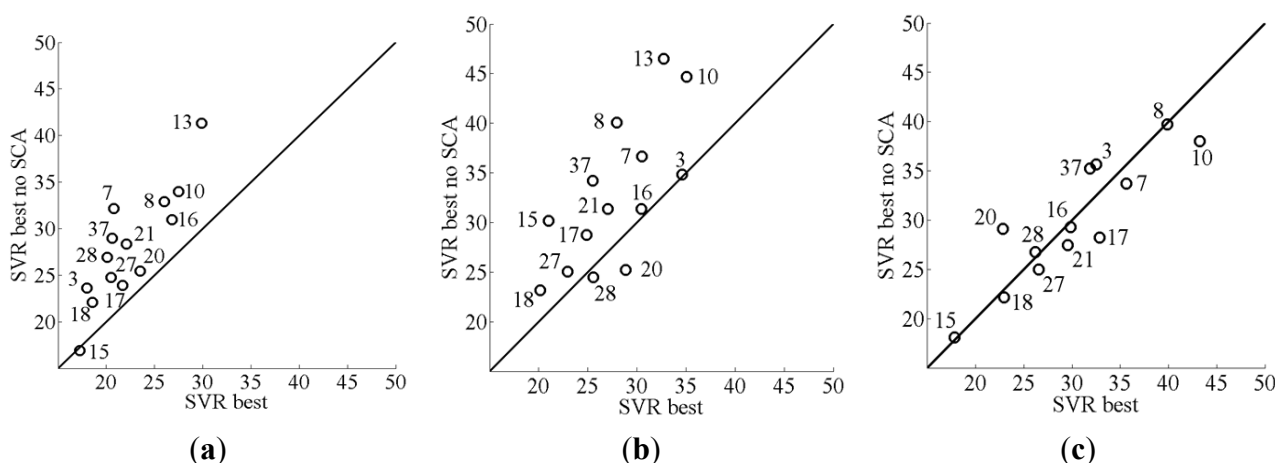

Figure 6. Comparison between the RMSE% obtained on the test set with (x axis) and without (y axis) SCA as SVR input feature. The scatterplots (**a**), (**b**), and (**c**) refer to the performances obtained by the SVR trained for the prediction with a lag equal to one, three, and six months, respectively.

SPI and temperature were the most significant input features resulting from the feature selection procedure on meteorological and climatic variables. However, their impact is appreciable only at one month lead time, in the case of an "ideal forecast," while using the values of antecedent months even increases the RMSE% (Figure 7). For a longer lead time, their inclusion as input features never improves the model. As an additional benchmark, we considered a linear autoregressive (AR) model set up with input ranging from one-month antecedent discharge to 12-month antecedent discharges. The statistical metrics of performance (RMSE% and R^2) of the AR model are shown in Table 2. These are only slightly better, and sometimes even worse, than the 10-year long-term average of the month of prediction. The

SVR is able to reduce the RMSE% with respect to the best AR method of 9% and 7% for a prediction lag equal to one and three months, respectively.
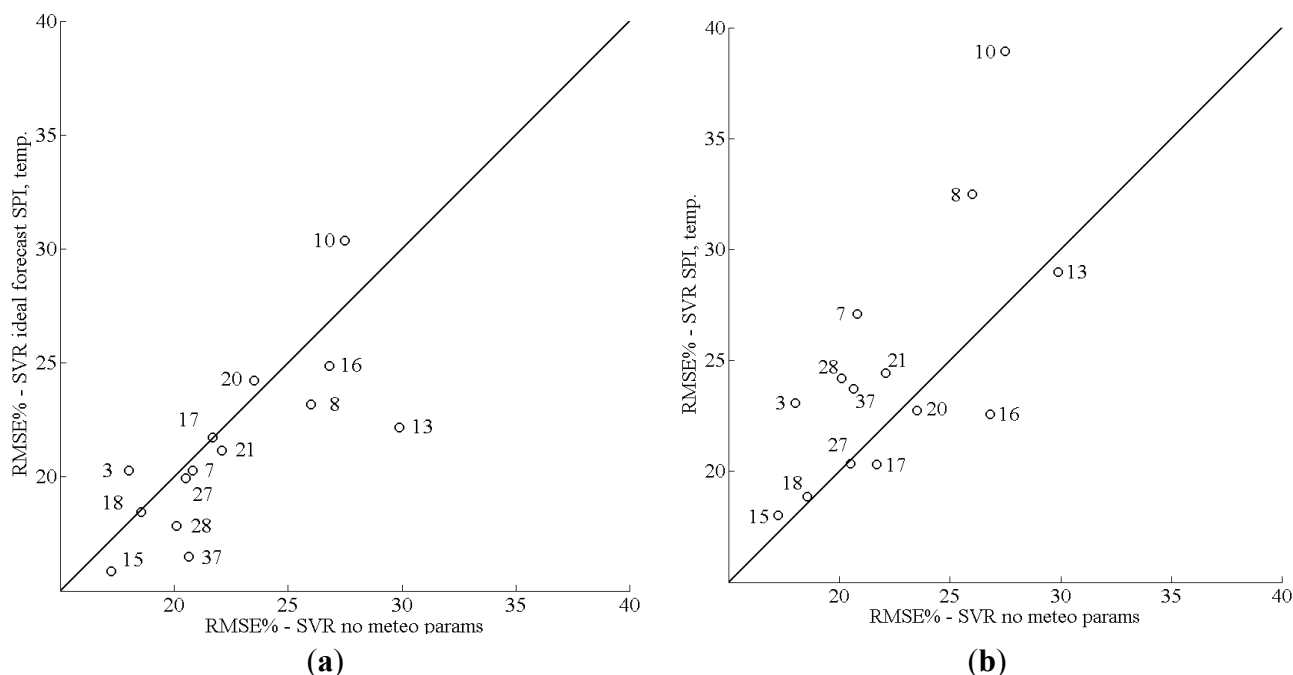

Figure 7. Comparison between the RMSE% obtained on the test set with (y axis) and without (x axis) SPI and temperature as SVR input feature. In (**a**) the actual target of SPI and temperature in the target month is employed (simulating an ideal forecast of this variable); In (**b**) the past time series of SPI and temperature is adopted. Both scatterplots refer to the performances obtained by the SVR trained for the prediction with a lag equal to one month.

The RMSE% is a metric of the general estimation accuracy for the different configurations tested and for the comparison with the benchmark method. However, a deeper error analysis on each catchment may yield useful additional insights. By looking at the error distribution of the Adige catchment closed at Ponte Adige (catchment identifier = 7 in Table 1), for instance, we have derived the error distribution on the test samples (shown in Figure 8a,b), for the most significant SVR configurations selected for predictions with one and three months lead time, together with the benchmarks. Besides yielding lower mean error values, the SVR models allow us to always keep error below 50%, while the benchmarks and SVR without SCA as input show a non-negligible fraction of higher errors for prediction lags equal to both one and three months.

Finally, it is of interest to inspect the cumulative volume errors: In many cases of practical interest, what matters is not necessarily the average error, but the total wrongly forecasted volume that may be associated with the monetary costs of water that was wrongly assumed to be available or lacking. Overestimation may, e.g., induce a hydropower plant manager to book unnecessary grid dispatching capacity, while underestimation may cause an irrigation manager to book unnecessary external supply. Figure 9 shows, as an example, a comparison of the volume forecasting errors for the benchmark of the long-term average discharge, for the SVR model with and without SCA as input, and for the best-performing AR model (the one with 12-month antecedent discharges), highlighting that the SVR model with SCA as input generates lower volume errors at a one-month lead time. At the same time, the

AR model and the SVR model without SCA introduce less appreciable advantages over the long-term average discharge. At three months' lead time, however, no advantage is shown by the SVR model, which performs no better than the benchmark (results not shown here for simplicity).
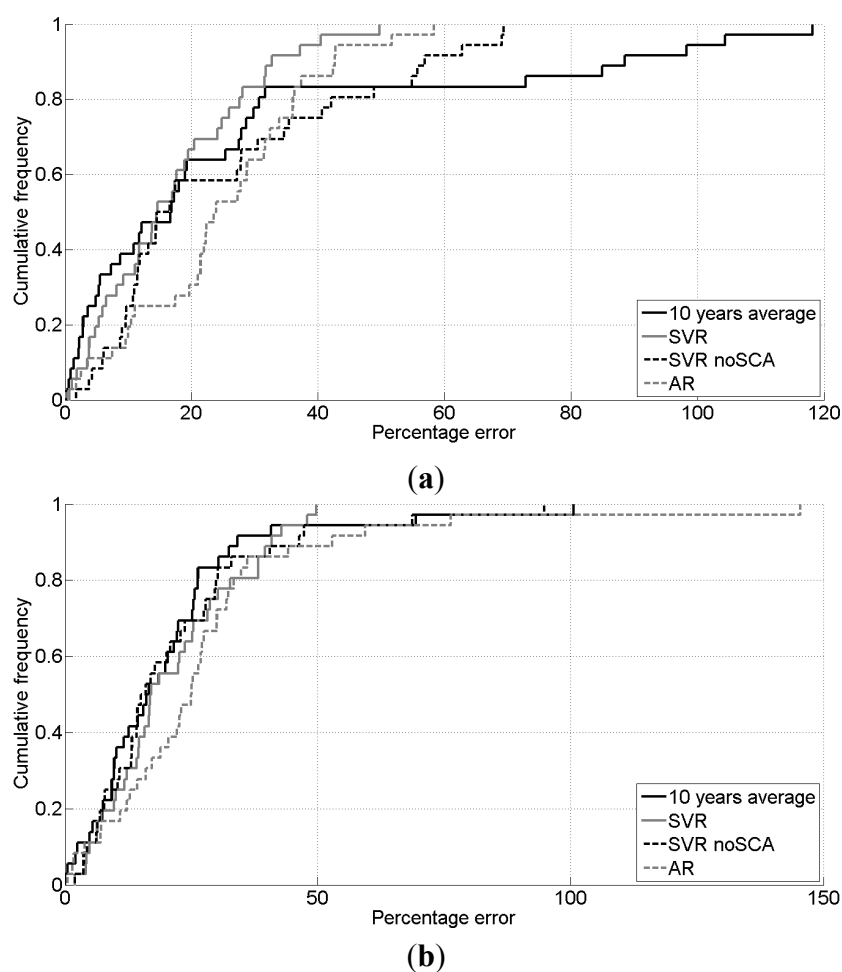

(a)

(b)

Figure 8. Cumulative frequency of the percentage error computed on the test set for watershed n. 7 (Adige at Ponte Adige), chosen as illustrative example. Each pdf shows the percentage error distribution for different input features configuration. The full gray line represents the selected SVR configuration. The benefit of the SVR approach with respect to the methods based on the 10-year average discharge (full black line) and to the AR method (dashed gray line) is clear as well as the improvement given by SCA (dashed black line). (**a**,**b**) show a prediction lag equal to one and three months, respectively.

When considering the SVR model calibrated for all catchments together, we observe that it generally yields less accurate predictions than a model set up for each individual catchment (see Table 3). At lead times of three and six months, none of the tested SVR configurations yields a better performance than the long-term catchment average, while at a lead time of one month the RMSE% may be reduced from 33% to 27%. Concerning the relevance of input features to prediction, the considerations expressed above for the SVR trained on individual catchments generally continue to hold: Also in this case the SCA and antecedent monthly discharges are retained as informative input features (step 1); the best performances are obtained using discharge data from the 12 months before prediction and the mean SCA

from the two antecedent months as input features. On the other hand, none of the considered watershed attributes proved to be informative in our tests apart from catchment area, which is only used for normalizing the discharges. At step 2 of the feature selection procedure, climatic signals reveal a minor impact on prediction only in the case of an "ideal forecast."
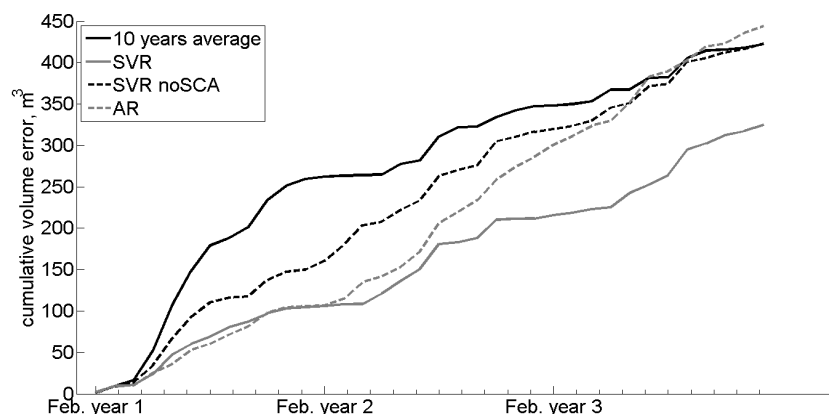

Figure 9. Cumulative volume errors of the one-month lead time forecast during the three test years (*i.e.*, 2005, 2010, 2011) relative to the catchment closed at Ponte Adige (*n.* 7).

Table 3. Performances of different SVR model configurations for all stations. The fourth and fifth column refer to the mean RMSE% and R square computed over the test samples of the 14 watersheds. In the third column "disch" stands for discharge and "temp" for air temperature. The notation—*n*: 0 stands for a time series of length $n - 1$ starting from the *n*-th previous month to the current month from where the discharge prediction is done. As a benchmark we considered the method based on the average discharge calculated over the 10 previous years, whose performances are shown in the first row.

Prediction Lag	Prediction Method	Input Features	Mean RMSE%	Mean R^2
All	10-year average	--	33	0.70
1	SVRall noMeteo	disch-11:0, SCA-1:0	27	0.75
	SVRall meteoIdeal	disch-11:0, SCA-1:0, SPI, temp	24	0.81
	SVRall meteo	disch-11:0, SCA-1:0, SPI-3:0, temp-10:0	28	0.78
3	SVRall noMeteo	disch-10:0, SCA-1:0, SPI, temp	36	0.70
	SVRall meteo	disch-11:0, SCA-1:0, SPI-3:0, temp-9:0	36	0.74

The results obtained for South Tyrol are in line with findings in other contexts using similar techniques [8,11–15,54].

Purely statistical predictions as presented here usually perform better than deterministic predictions based on seasonal weather or climate forecasts. For instance, Bastola *et al.* [59] highlighted the shortcomings of such deterministic predictions in the southeastern United States, while in general there is substantial agreement on the importance of catchment state characterization for effective streamflow prediction [60–62].

4. Conclusions

We have presented an application of SVR models for predicting the monthly mean discharge over 14 catchments in the Alpine region of South Tyrol (Northern Italy) using monthly discharge and SCA of antecedent months as input features. The comparison between the SVR models and the long-term monthly mean discharge, as well as with a simpler linear autoregressive model, is encouraging, with a mean improvement of the RMSE% on the 14 catchments selected in this study of 11%, 5%, and 2% for a prediction lag equal to one, three, and six months, respectively (Table 2), and a reduction in frequency of larger errors (Figures 8 and 9). An SVR model proves to be useful compared to the long-term mean and a linear model, when a penalty is associated with over- or underestimation of monthly streamflow, as it allows reducing volume errors when predicting at a one month lead time, compared to the benchmarks. This advantage, however, is no longer apparent when considering lead times of three months.

The analysis also highlights the usefulness of SCA as a catchment state variable in the study region. Through appropriate feature selection, we showed that the SVR models trained with the SCA show an RMSE% on average 6% lower for a prediction lag equal to one month compared to the SVR models that consider only antecedent discharges as input features. On the contrary, catchment input (meteorological/climatological) variables, of which SPI and temperature have been shown to be the most informative in explaining monthly discharges, have been shown to bring in little, if any, prediction improvement. For the mountainous region considered in this study, even a perfect knowledge of precipitation (as SPI) and temperature during the month of prediction would not substantially improve discharge predictions, which suggests the importance of catchment state controls on flows, a (partial) descriptor of which is identified in snow-covered areas. This corroborates the observation, general in catchment hydrology, that discharge may be a complex and nonlinear response to precipitation and evapotranspiration, and can be more often related to the antecedent storage than to the input of water in a catchment. On the other hand, in models based on the catchment water balance, precipitation (and evapotranspiration) is as essential for prediction as antecedent storage; black-box models in this case study definitely "learn" more from the latter than from the former.

In order to predict monthly discharges for catchments where a sufficiently long discharge time series is not available, a general SVR trained by using the data of similar catchments, after normalizing discharges with the watershed area, still leads to a slight improvement for predictions at a lead time of one month (RMSE% reduced of 5%) while at longer lead times it yields predictions no better than the benchmark, or even slightly worse. This suggests that case-specific setup of statistical models as applied in this study may be difficult to generalize.

Machine learning techniques such as SVR may help improve our capacity for seasonal hydrological forecasting. Monthly discharge predictions at one to three months with RMSE% below 30% and an appreciable reduction in the frequency of high errors may already enable better management of water resources through optimization of storage at (and release from) reservoirs. However, further research is required, particularly in the description of catchment state (e.g., by using better proxies of snow water equivalent than simple SCA), in order to improve prediction performance in view of more effective practical applications, as catchment input seems relatively irrelevant even in the case of an ideal prediction, hence even less relevant after considering errors from real seasonal weather forecasts.

Acknowledgments

Author Contributions

All authors contributed to writing different parts of the paper. Alberto Pistocchi ideated the research building on discussions with Claudia Notarnicola, Paolo Mazzoli and Marcello Petitta and coordinated the writing of the paper. Mattia Callegari analyzed the data, developed the SVR calculations, and contributed to writing large parts of the paper; Ludovica De Gregorio contributed to the analysis of data and SVR calculations; Paolo Mazzoli supervised the selection and interpretation of hydrological data and largely contributed to the discussion of the results; Claudia Notarnicola. supervised the snow cover remote sensing data processing; Luca Pasolli contributed to the selection and calibration of the SVR model as well as the interpretation and discussion of the results; and Marcello Petitta selected and investigated the climate parameters used in the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Low Flow and Drought Research in Europe: Achievements and Future. Available online: http://webworld.unesco.org/water/ihp/pdf/conclusions_bratislava_lfdr_workshop08.pdf (accessed on 30 April 2015).
2. Mishra, A.K.; Desai, V.R. Drought forecasting using feed-forward recursive neural network. *Ecol. Model.* **2006**, *198*, 127–148.
3. Madadgar, S.; Moradkhani, H. A Bayesian framework for probabilistic seasonal drought forecasting. *J. Hydrometeorol.* **2013**, *14*, 1685–1705.
4. Yang, L.; Tian, F.; Sun, Y.; Hu, H. Attribution of hydrologic forecast uncertainty within scalable forecast windows. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 775–786.
5. Chalise, S.R.; Kansakar, S.R.; Rees, G.; Croker, K.; Zaidman, M. Management of water resources and low flow estimation for the Himalayan basins of Nepal. *J. Hydrol.* **2003**, *282*, 25–35.
6. Pandit, S.M.; Wu, S.-M. *Time Series and System Analysis with Applications*; John Wiley & Sons, Inc.: Hpboken, NJ, USA, 1983.
7. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.
8. Wang, W.; Chau, K.; Cheng, C.; Qiu, L. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol.* **2009**, *374*, 294–306.
9. Bruzzone, L.; Melgani, F. Robust multiple estimator system for the analysis of biophysical parameters from remotely sensed data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 159–174.

10. Pasolli, L.; Notarnicola, C.; Bruzzone, L. Estimating soil moisture with the support vector regression technique. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 1080–1084.

11. Shiri, J.; Kisi, O. Short-term and long-term streamflow forecasting using a wavelet and neuro-fuzzy conjunction model. *J. Hydrol.* **2010**, *394*, 486–493.

12. Guo, J.; Zhou, J.; Qin, H.; Zou, Q.; Li, Q. Monthly streamflow forecasting based on improved support vector machine model. *Expert Syst. Appl.* **2011**, *38*, 13073–13081.

13. Shiri, J.; Cimen, M. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J. Hydrol.* **2011**, *399*, 132–140.

14. Zhao, T.; Cai, X.; Yang, D. Effect of streamflow forecast uncertainty on real-time reservoir operation. *Adv. Water Resour.* **2011**, *34*, 495–504.

15. Lima, C.H.R.; Lall, U. Climate informed monthly streamflow forecasts for the Brazilian hydropower network using a periodic ridge regression model. *J. Hydrol.* **2010**, *380*, 438–449.

16. Liu, Z.; Zhou, P.; Chen, G.; Guo, L. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. *J. Hydrol.* **2014**, *519*, 2822–2831.

17. Huang, S.; Chang, J.; Huang, Q.; Chen, Y. Monthly streamflow prediction using modified EMD-based support vector machine. *J. Hydrol.* **2014**, *511*, 764–775.

18. Sun, A.Y.; Wang, D.; Xu, X. Monthly streamflow forecasting using Gaussian process regression. *J. Hydrol.* **2014**, *511*, 72–81.

19. Dehghani, M.; Saghafian, B.; Nasiri Saleh, F.; Farokhnia, A.; Noori, R. Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte Carlo simulation. *Int. J. Climatol.* **2014**, *34*, 1169–1180.

20. Wang, H.; Fu, X. Utilization of climate information and soil moisture estimates to provide monthly and sub-monthly streamflow forecasts. *Int. J. Climatol.* **2014**, *34*, 3515–3527.

21. Fiorillo, F.; Esposito, L.; Guadagno, F.M. Analysis and forecast of water resources in an ultra-centenarian spring discharge series from Serino (Southern Italy). *J. Hydrol.* **2007**, *336*, 125–138.

22. Fiorillo, F. Spring hydrographs as indicators of droughts in a karts environment. *J. Hydrol.* **2009**, *373*, 290–301.

23. Mendicino, G.; Senatore, A.; Versace, P. A Groundwater resource index (GRI) for drought monitoring and forecasting in a Mediterranean climate. *J. Hydrol.* **2008**, *357*, 282–302.

24. Bartolini, E.; Claps, P.; D'Odorico, P. Interannual variability of winter precipitation in the European Alps: Relations with the North Atlantic Oscillation. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 17–25.

25. Singh, P.; Bengtsson, L. Hydrological sensitivity of a large Himalayan basin to climate change. *Hydrol. Process.* **2004**, *18*, 2363–2385.

26. Parajka, J.; Blöschl, G. The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models. *J. Hydrol.* **2008**, *358*, 240–258.

27. Viviroli, D.; Gurtz, J.; Zappa, M. The Hydrological Modeling System PREVAH—Part II—Physical model description. In *Geographica Bernensia P40*; Institute of Geography, University of Berne: Berne, Switzerland, 2007, p. 86.

28. Viviroli, D.; Dürr, H.H.; Messerli, B.; Meybeck, M.; Weingartner, R. Mountains of the world, water towers for humanity: Typology, mapping and global significance. *Water. Resour. Res.* **2007**, *43*, doi:10.1029/2006WR005653.

29. European Environmental Agency. *Regional Climate Change and Adaptation: The Alps Facing the Challenge of Changing Water Resources*; EEA report No. 8/2009; European Environmental Agency: Copenhagen, Denmark, 2009.

30. Evgeniou, T.; Pontil, M.; Poggio, T. Regularization Network and support vector machines. *Adv. Comput. Math.* **2000**, *13*, 1–50.

31. Smola, A.J.; Sholkopf, B. A tutorial in support vector regressions. *Stat. Comput.* **2004**, *14*, 199–222.

32. Zahn, H.; Shi, P.; Chen, C. Retrieval of oceanic chlorophyll concentration using support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2947–2951.

33. Pai, P.F.; Lin, K.P.; Lin, C.S.; Chang, P.T. Time series forecasting by a seasonal support vector regression model. *Expert Syst. Appl.* **2010**, *37*, 4261–4265.

34. Raghavendra, N.S.; Deka, P.C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372–386.

35. Pasolli, L.; Notarnicola, C.; Bruzzone, L. Multi-objective parameter optimization in support vector regression: General formulation and application to the retrieval of soil moisture from remote sensing data. *IEEE J. Sel. Top. Appl.* **2012**, *5*, 1495–1508.

36. Notarnicola, C.; Duguay, M.; Moelg, N.; Schellenberger, T.; Tetzlaff, A.; Monsorno, R.; Costa, A.; Steurer, C.; Zebisch, M. Snow cover maps from MODIS images at 250 m resolution, Part 1: Algorithm description. *Remote Sens.* **2013**, *5*, 110–126.

37. Notarnicola, C.; Duguay, M.; Moelg, N.; Schellenberger, T.; Tetzlaff, A.; Monsorno, R.; Costa, A.; Steurer, C.; Zebisch, M. Snow Cover Maps from MODIS Images at 250 m Resolution, Part 2: Validation. *Remote Sens.* **2013**, *5*, 1568–1587.

38. Piechota, T.C.; Dracup, J.A. Drought and regional hydrologic variation in the United States: Associations with ElNino-Southern Oscillation. *Water Resour. Res.* **1996**, *32*, 1359–1373.

39. Tadesse, T.; Wilhite, D.A.; Harms, S.K.; Hayes, M.J.; Goddard, S. Drought monitoring using data mining techniques: A case study for Nebraska, USA. *Nat. Hazards* **2004**, *33*, 137–159.

40. Barros, A.P.; Bowden, G.J. Toward long-lead operational forecasts of drought: An experimental study in the Murray-Darling River Basin. *J. Hydrol.* **2008**, *357*, 349–367.

41. Bordi, I.; Fraedrich, K.; Petitta, M.; Sutera, A. Extreme value analysis of wet and dry periods in Sicily. *Theor. Appl. Climatol.* **2007**, *87*, 61–71.

42. McKee, T.; Doesken, N.; Kleist, J. The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology, Anaheim, CA, USA, 17–22 January 1993.

43. Hayes, M.J.; Svoboda, M.D.; Wilhite, D.A.; Vanyarkho, O.V. Monitoring the 1996 drought using the standardized precipitation index. *Bull. Am. Meteor. Soc.* **1999**, *80*, 429–438.

44. Vicente-Serrano, S.M.; Lopez-Moreno, J.I. Hydrological response to different time scales of climatological drought: An evaluation of the Standardized Precipitation Index in a mountainous Mediterranean basin. *Hydrol. Earth Syst. Sci.* **2005**, *9*, 523–533.

45. Keyantash, J.; Dracup, J.A. The quantification of drought: An evaluation of drought indices. *Bull. Am. Meteor. Soc.* **2002**, *83*, 1167–1180.

46. Bordi, I.; Sutera, A. Fifty years of precipitation: Some spatially remote teleconnections. *Water Resour. Manag.* **2001**, *15*, 247–280.

47. Bordi, I.; Fraedrich, K.; Petitta, M.; Sutera, A. Methods for predicting drought occurrences. In Proceedings of the EWRA (The European Water Resources Association) Conference, Menton, France, 7–10 September 2005.

48. Morid, S.; Smakhtin, V.; Bagherzadeh, K. Drought forecasting using artificial neural networks and time series of drought indices. *Int. J. Climatol.* **2007**, *27*, 2103–2111.

49. Sutera, A. Probability density distribution of large-scale atmospheric flow. *Adv. Geophys.* **1996**, *29*, 227–249.

50. Hansen, A.R.; Sutera, A. On the probability density distribution of planetary-scale atmospheric wave amplitude. *J. Atmos. Sci.* **1986**, *43*, 3250–3265.

51. Calmanti, S.; Dell'Aquila, A.; Lucarini, V.; Ruti, P.M.; Speranza, A. Statistical properties of mid-latitude atmospheric variability. In *Nonlinear Dynamics in Geosciences*; Springer: New York, NY, USA, 2007; pp. 369–391.

52. Dell'Aquila, A.; Ruti, P.M.; Sutera, A. Effects of the baroclinic adjustment on the tropopause in the NCEP-NCAR reanalysis. *Clim. Dyn.* **2007**, *28*, 2–3.

53. Fraedrich, K.; Bottger, H. A wavenumber frequency analysis of the 500 mb geopotential at 50°N. *J. Atmos. Sci.* **1978**, *35*, 745–750.

54. Soukup, T.L.; Oubeidillah, A.A.; Tootle, G.A.; Piechota, T.C.; Wulff, S.S. Long lead time streamflow forecasting of the North Platte River incorporating oceanic-atmospheric climate variability. *J. Hydrol.* **2009**, *368*, 131–142.

55. Chiew, F.H.S. Climate variability, seasonal streamflow forecast and water resources management. In *Climate Impact on Australia's Natural Resources: Current and Future Challenges*; Climate impact on Australia's Natural Resources: Queensland, Australia, 2003; pp. 21–23.

56. Chiew, F.H.S.; McMahon, T.A. Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrolog. Sci. J.* **2002**, *47*, 505–522.

57. Zanchettin, D.; Traverso, P.; Tomasino, M. Po river discharges: A preliminary analysis of a 200-year time series. *Clim. Chang.* **2008**, *89*, 411–433.

58. Nationa Weather Service (NWS) Climate Prediction Center: North Atlantic Oscillation. Available online: http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml (accessed on 30 April 2015).

59. Satish, B.; Misra, V.; Li, H. Seasonal hydrological forecasts for watersheds over the southeastern United States for the boreal summer and fall seasons. *Earth Interact.* **2013**, *17*, 1–22.

60. Yossef, N.C.; Winsemius, H.; Weerts, A.; van Beek, R.; Bierkens, M.F.P. Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.* **2013**, *49*, 4687–4699.

61. Shukla, S.; Sheffield, J.; Wood, E.F.; Lettenmaier, D.P. On the sources of global land surface hydrologic predictability. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 2781–2796.

62. Mahanama, S.; Livneh, B.; Koster, R.; Lettenmaier, D.; Reichle, R. Soil moisture, snow, and seasonal streamflow forecasts in the United States. *J. Hydrometeor.* **2012**, *13*, 189–203.