Hydrol. Earth Syst. Sci. Discuss., 7, 3521–3550, 2010
www.hydrol-earth-syst-sci-discuss.net/7/3521/2010/
doi:10.5194/hessd-7-3521-2010

Hydrology and Earth System Sciences Discussions

Seasonal prediction of winter extreme precipitation over Canada by support vector regression

Z. Zeng[1,*], W. W. Hsieh[1], A. Shabbar[2], and W. R. Burrows[3]

[1]Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[2]Climate Research Division, Environment Canada, Toronto, Ontario, Canada
[3]Meteorological Research Division, Environment Canada, Edmonton, Alberta, Canada
[*]now at: COSMIC Project Office, UCAR, Boulder, CO 80301, USA

Received: 6 May 2010 – Accepted: 7 June 2010 – Published: 16 June 2010

Correspondence to: W. W. Hsieh (whsieh@eos.ubc.ca)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

For forecasting the maximum 5-d accumulated precipitation over the winter season at lead times of 3, 6, 9 and 12 months over Canada from 1950 to 2007, two nonlinear and two linear regression models were used, where the models were support vector regression (SVR) (nonlinear and linear versions), nonlinear Bayesian neural network (BNN) and multiple linear regression (MLR). The 118 stations were grouped into six geographic regions by K-means clustering. For each region, the leading principal components of the winter extreme precipitation were the predictands. Potential predictors included quasi-global sea surface temperature anomalies and 500 hPa geopotential height anomalies over the Northern Hemisphere, as well as six climate indices (the Niño-3.4 region sea surface temperature, the North Atlantic Oscillation, the Pacific-North American teleconnection, the Pacific Decadal Oscillation, the Scandinavia pattern, and the East Atlantic pattern). The results showed that in general the two robust SVR models tended to have better forecast skills than the two non-robust models (MLR and BNN), and the nonlinear SVR model tended to forecast slightly better than the linear SVR model. Among the six regions, the Eastern Prairies region displayed the highest forecast skills, and the Arctic region the second highest. The strongest nonlinearity was manifested over the Eastern Prairies and the weakest nonlinearity over the Arctic.

1 Introduction

Extreme precipitation events, usually responsible for major economic losses and ecological damage, have important impacts on agriculture, energy use and human activity. There has been enhanced interest in recent years on the apparent increase in the frequency and/or severity of extreme precipitation events for many regions, which might be related to the increasing concentrations of greenhouse gases (Easterling et al., 2000; Groisman et al., 2005). Though the long-term trend of extreme precipitation events

HESSD

7, 3521–3550, 2010

seems not so significant in most areas of Canada (Zhang et al., 2001; Kunkel, 2003), the establishment of an accurate and timely extreme event monitoring and prediction system is still of prime importance for circumventing the potential impacts posed by climate variations and extreme weather.

Numerous previous studies have shown that the El Niño-Southern Oscillation (ENSO), centered in the tropical Pacific, plays an important role in North American climate variability, especially during the winter season (Barnston, 1994; Shabbar and Barnston, 1996; Goddard et al., 2001; Wu et al., 2005; Shabbar, 2006). Besides ENSO, other circulation patterns, such as the North Atlantic Oscillation (NAO), Pacific-North American (PNA) teleconnection, Pacific Decadal Oscillation (PDO) etc. have been found to show influences on precipitation over the Northern Hemisphere (Hsieh et al., 2006; Wu et al., 2006a, Bonsai et al., 2006; Lorenzo et al., 2008; Lin et al., 2008), and may contribute skill in seasonal precipitation forecasts. Most seasonal forecasts focus on predicting the seasonal mean of the precipitation instead of seasonal statistics of extreme precipitation events. Such seasonal extreme statistics are potentially noisier than the seasonal mean, hence they may be even harder to predict.

One commonly used technique for seasonal predictions is the empirical or statistical approach, using linear statistical methods such as correlation, regression (Ward and Folland, 1991), and canonical correlation analysis (Shabbar and Barnston, 1996). More recently, machine learning methods such as neural networks (Haupt et al., 2008; Hsieh, 2009) have been introduced for nonlinear regression and nonlinear canonical correlation analysis (Wu et al., 2006a, b; Cannon and Hsieh, 2008). The advantage of nonlinear methods to linear methods is generally far less evident for climate applications than for weather applications, since averaging nonlinear daily relations produces near-linear seasonal relations as a consequence of the central limit theorem (Yuval and Hsieh, 2002; Hsieh and Cannon, 2008). A seasonal extreme statistic like the maximum amount of precipitation over 5 consecutive days in the winter season does not involve extensive averaging as in the computation of the seasonal mean, thereby avoiding the linearization effect of the central limit theorem. Hence despite their potentially higher

HESSD
7, 3521–3550, 2010

noise-to-signal level than the seasonal mean, seasonal extreme statistics may be more suited than the seasonal mean for nonlinear forecasting by machine learning methods.

Neural network (NN) methods, generally regarded as forming the first wave of breakthrough in machine learning, became popular in the late 1980s for nonlinear regression problems, whereas kernel methods (e.g. support vector regression, SVR) arrived in a second wave in the second half of the 1990s (Bishop, 2006; Hsieh, 2009). SVR has two advantages over NN models – it avoids the multiple minima problem associated with nonlinear optimization used in NN models, and robust error norms are used in SVR instead of the non-robust mean squared error (MSE) norm, allowing SVR to better handle datasets with outliers. The use of a suitable nonlinear kernel function in SVR allows it to be fully nonlinear, while the use of a linear kernel function restricts SVR to a linear model. Nevertheless, the linear SVR model is different from the multiple linear regression (MLR) model, since the robust error norm is used in SVR but not in MLR. Applications of SVR to hydrological problems include Dibike et al. (2001), Khan and Coulibaly (2006), Bürger et al. (2007) and Anandhi et al. (2008).

In this paper, we have a four-way comparison of forecast skills from nonlinear SVR, linear SVR, Bayesian NN (BNN) and MLR. The objective is to see how robust and non-robust structures as well as nonlinear and linear capability in the models affect forecast skills when the predictand is the very noisy and non-Gaussian winter extreme precipitation anomaly. The description of the data and the forecasting methods are given in Sects. 2 and 3, respectively. Section 4 presents the results of forecasting the winter extreme precipitation over Canada, followed by the conclusion in Sect. 5.

2 Data

Monthly extended reconstructed sea surface temperature (SST) data (ERSST version 3 (Smith et al., 2008)) were obtained from the National Oceanic and Atmospheric Administration (NOAA) with a spatial resolution of $2^\circ \times 2^\circ$ for the period 1950–2007; while monthly 500 hPa geopotential height (Z500) data with $2.5^\circ \times 2.5^\circ$ horizontal resolution

HESSD
7, 3521–3550, 2010
Seasonal prediction of winter extreme precipitation
Z. Zeng et al.

from the National Centers for Environmental Prediction (NCEP) reanalysis were used in this study for the same period (Kalnay et al., 1996). We only used SST data within the zonal band between 30° S and 70° N, and Z500 data over the North Hemisphere (20° N–90° N), quite similar to Shabbar and Barnston (1996). To reduce memory need, the SST data were averaged into 6°×4° grids with 1020 spatial points, and the Z500 data into 5°×5° grids with 1008 spatial points.

Seasonal SST and Z500 anomalies were obtained by removing the climatological cycle from the monthly mean data and filtering them using a 3-month running mean. After normalizing the anomalies, time-lagged copies of the data were stacked (i.e. the original copy, plus copies time-lagged by 3, 6 and 9 months were assembled together) and treated as a new enlarged dataset to be compacted by principal component analysis (PCA). This PCA process, called space-time PCA, singular spectrum analysis or extended empirical orthogonal function (EEOF) analysis, is performed on the SST and Z500 normalized anomalies separately, each having 5 leading principal components (PC) retained (and these will be referred to as the SSTPC and Z500PC below).

Monthly climate indices for the Niño-3.4 region SST (NINO), the North Atlantic Oscillation (NAO), the Pacific-North American (PNA) teleconnection, the Scandinavia (SCA) pattern, and the East Atlantic (EA) pattern – were downloaded from the website of Climate Prediction Center (CPC), NOAA. The description of listed indices can also be found from the CPC site (http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml). Monthly values of the Pacific Decadal Oscillation (PDO) were obtained from the Joint Institute for the Study of the Atmosphere and Ocean, University of Washington (http://jisao.washington.edu/pdo/PDO.latest).

Daily 5-d total precipitation records were obtained from 461 climate stations in Canada for the 1900–2007 period. Only stations with data covering at least the period of 1950–2007 were considered as candidates for the analysis. This period was selected to maximize the number of stations while attempting to maintain the longest possible records. In addition, stations with more than 5% missing data over 1950–2007 were not used. Under these conditions, only 118 stations qualified for further study. For

HESSD
7, 3521–3550, 2010

each station, its monthly maximum was first calculated from the daily 5-d accumulated precipitation data, which suggested the heaviest precipitation event during that month. The climatological seasonal cycle of 5-d precipitation was then removed, and the 3-month maximum was identified as the seasonal extreme precipitation anomaly. Only winter (December to February) data from 1950/51 to 2006/07 (57 winters) were analyzed here. The reason that the maximum 5-d total precipitation instead of the daily extreme is used here because this study focuses on the extreme events related to low-frequency signals of large-scale variations in the atmosphere-ocean system. In addition, larger-scale impacts, such as floods from heavy precipitation are mostly due to multi-day episodes. Maximum 5-d precipitation has been also chosen as one of the standard seasonal extreme precipitation indices by the European Union STARDEX project (STAtistical and Regional dynamical Downscaling of Extremes for European regions).

In view of the diversity of the Canadian climate, we classified the 118 stations into six groups using K-means clustering (Zhang et al., 2001; Whitfield et al., 2002). The 118×118 elements of the intercorrelation matrix among station precipitation, which assumes the internal spatial coherence of precipitation variability does not change with time, and the 118×6 elements of the correlation matrix between station precipitation and the six climate indices, which reflects the relationship between seasonal extreme precipitation and large-scale atmospheric teleconnection and SST indices, were taken as inputs to the K-means algorithm. The Euclidean distance was used in cluster analysis to measure dissimilarity between stations.

Figure 1 presents the spatial distribution of the Canadian stations, with their membership in the six clusters shown by different symbols. Hence the cluster analysis has divided the Canadian domain into six geographic regions. The Pacific coastal region (R1), under the influence of warm ocean currents and moisture-laden winds, receives the most rain and snow during winter. In the Cordilleran region (R2), the warm, moist Pacific air is forced to rise over the mountains, cools and falls on the western slopes in sizeable amounts of precipitation as rain at lower altitudes and snow at higher ones;

HESSD
7, 3521–3550, 2010

however, the eastern slopes and central plateau region are arid. Eastern Prairies (R3) receive considerably less precipitation than most other parts of Canada, often being dry for long periods. For the Arctic region (R4), it is extremely cold with very low precipitation. The Great Lakes region (R5) receives rather uniform precipitation through the year with heavy snowfalls in winter. In the Atlantic coast (R6), extremely cold air masses are modified by oceanic influences, which also cause considerable snow and precipitation in winter. The number of stations for each cluster/region and the corresponding mean precipitation, i.e. the 3-month mean of the 5-d total precipitation over all winters and over all stations in each region, are shown in Table 1, where the mean precipitation was 78.5 mm over the west coast and 63.8 mm over the east coast, much larger than the 8.8 mm over the Arctic region in winter.

For each region, we applied PCA to the seasonal extreme precipitation anomalies, and preserved the leading PCs. Table 1 summarized the explained variance by the first few PCs retained for each region in column 4. For example, the 7 leading PCs for the Pacific coastal region (R1) account for 85% of total variance of the precipitation anomalies. Each PC was then chosen as the predictand for a forecast model. The seasonal extreme precipitation anomaly forecasts were reconstructed by summing the forecasted PCs multiplied by their corresponding empirical orthogonal function (EOF) spatial patterns.

3 Methodology

3.1 Support vector regression

Support vector machines were originally designed for classification problems (Vapnik, 1995). They were then extended to nonlinear regression problems (Vapnik et al., 1997; Bishop, 2006). Here we describe the essence of support vector regression (SVR).

Let $\boldsymbol{x}$ denote the m inputs or predictors and y denote the single output variable or predictand. By introducing a nonlinear mapping function $\boldsymbol{\phi}$, the nonlinear regression problem between $\boldsymbol{x}$ and y can be converted to a linear regression problem between $\boldsymbol{\phi}$ and y, i.e.

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

where $\langle , \rangle$ denotes the inner product, and $\boldsymbol{w}$ and b are the regression coefficients obtained by minimizing the error between f and the observed values of y. To measure this error, instead of the commonly used mean squared error norm, SVR uses the ϵ-insensitive error norm defined by

$$|f(\boldsymbol{x},\boldsymbol{w}) - y|_\epsilon = \begin{cases} 0, & \text{if } |f-y| < \epsilon \\ |f-y|-\epsilon, & \text{otherwise,} \end{cases} \tag{2}$$

i.e. when the difference between f and y is smaller than ϵ, the error is ignored, whereas when the difference between f and y is large, the error approximates the mean absolute error, which unlike the mean squared error, is robust to outliers in the data.

The $\boldsymbol{w}$ and b coefficients are estimated by minimizing the regularized error function R using sample data $(\boldsymbol{x}_i, y_i)$, where

$$R = \frac{C}{N}\sum_{i=1}^{N} |f(\boldsymbol{x}_i,\boldsymbol{w}) - y_i|_\epsilon + \frac{1}{2}\|\boldsymbol{w}\|^2, \tag{3}$$

with C and ϵ prescribed parameters (commonly referred to as hyperparameters), and N the sample size. The second term is called the regularization (or weight penalty) term, and when a small value of C is used, the regularization term becomes prominent relative to the first term, and the minimization of R forces the $\boldsymbol{w}$ coefficients to have small magnitude, thereby limiting model complexity.

The conversion of a nonlinear regression problem to a linear regression problem (Eq. 1) eliminates the need for nonlinear optimization, which has to deal with the presence of multiple local minima in the error function, as in the case of NN methods. However, $\boldsymbol{\phi}(\boldsymbol{x})$ may be a very high (or even infinite) dimensional vector, hence solving the linear regression problem may be prohibitively expensive. In SVR, a kernel trick

HESSD
7, 3521–3550, 2010

Seasonal prediction of winter extreme precipitation

Z. Zeng et al.

is used, which is to replace the inner product $\langle \boldsymbol{\phi}(\boldsymbol{x}), \boldsymbol{\phi}(\boldsymbol{x}') \rangle$ in the solution algorithm by a kernel function $K(\boldsymbol{x}, \boldsymbol{x}')$, which does not involve handling the unwieldy $\boldsymbol{\phi}(\boldsymbol{x})$. The minimization of Eq. (3) involves Lagrange multipliers, and the final regression estimate can be expressed as in the form (Bishop 2006)

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} a_i K(\boldsymbol{x}, \boldsymbol{x}_i) + b. \quad (4)$$

Performance of the SVR model depends on the choice of the kernel function and the hyperparameters. In this study, we used the linear kernel, $K(\boldsymbol{x}, \boldsymbol{x}_i) = \langle \boldsymbol{x}, \boldsymbol{x}_i \rangle$, and the Gaussian or radial basis function (RBF) kernel, $K(\boldsymbol{x}, \boldsymbol{x}_i) = \exp\left(-\|\boldsymbol{x} - \boldsymbol{x}_i\|^2 / (2\sigma^2)\right)$, with the hyperparameter σ controlling the width of the Gaussian function. When the linear kernel is used, the SVR performs robust linear regression, whereas with the RBF kernel, the SVR model performs robust nonlinear regression. We used the SVR codes by Chang and Lin (2001), downloadable from the LibSVM website (http://www.csie.ntu.edu.tw/~cjlin/libsvm). The hyperparameters, C, ϵ, and σ (for the RBF kernel) can be tuned instead of predefined subjectively.

3.2 Bayesian neural network (BNN)

As NN models are now commonly used in hydrology (Solomatine and Ostfeld, 2008), we will only briefly outline the approach used in our study. An NN model is trained from a data set $(\boldsymbol{x}, y)$, with $\boldsymbol{x}$ the predictors and y the predictand, by adjusting network parameters or weights $\boldsymbol{w}$ so as to minimize a regularized error function

$$E(\boldsymbol{w}) = \frac{C}{N} \sum_{i=1}^{N} (f(\mathbf{x}_i; \boldsymbol{w}) - y_i)^2 + \|\boldsymbol{w}\|^2, \quad (5)$$

where the first term is the parameter C times the mean squared error, while the second term is the regularization term. A small C will strongly suppress the magnitude of $\boldsymbol{w}$

HESSD
7, 3521–3550, 2010

found by the optimization process, thereby yielding a less complex (i.e. less nonlinear) model. The best value for C is commonly chosen upon validating the model performance over independent data not used in training the model. With the optimal C, the model should be neither overfitting nor underfitting the data.

An alternative to using validation to find the best value for C is BNN (MacKay, 1992), a neural network designed based on a Bayesian probabilistic formulation. The idea of BNN is to treat the network parameters or weights as random variables, obeying an assumed prior distribution. Once observed data are available, the prior distribution is updated to a posterior distribution using Bayes' theorem. BNN automatically determines the optimal value of C without the need of validation data (Bishop, 2006). In this study, the BNN model used was from the NETLAB toolbox (Nabney, 2002), with a standard mapping function f, i.e. a layer of hyperbolic tangent mapping followed by linear mapping. As NN suffers from multiple minima in E, an ensemble of 30 BNN models was built from random initial weights, and the ensemble mean was taken as the final forecast of the BNN model.

3.3 Double cross-validation

For seasonal forecasting, the sample size to the number of predictors is relatively small, since we have 5 SSTPCs, 5 Z500PCs and 6 climate indices as predictors. Hence PCA is again applied to these predictor time series to further reduce the number of predictors. An additional advantage of PCA is to produce uncorrelated predictors. To determine p, the optimal number of PCs to retain as predictors, cross-validation is needed. In an n-fold cross-validation procedure, the data record is divided into n segments, a segment is reserved as validation data, and the other segments as training data. The model is trained using the training data, then validated or tested on the independent data in the validation segment. By rotating the validation segments, the entire data record can be used for validation. As mentioned earlier, for each region (as determined by the cluster analysis), PCA was applied to the seasonal extreme precipitation anomalies for all the stations in that region, yielding the predictand PCs.

Cross-validation is also needed to determine *n*PC, the optimal number of predictand PCs to retain for each region (Table 1).

For the SVR model, we used the Cherkassky and Ma (2004) approach to estimate the value of the hyperparameters, and then use a finer grid search to pinpoint the optimal values of the hyperparameters under cross-validation. To use independent data to test or verify the model forecasts, a second round of cross-validation is needed, hence a double cross-validation procedure (Cannon and Hsieh, 2008).

The procedure involves two rounds of cross-validation, an outer round (CV1) and an inner round (CV2). For CV1, the first 5 yr of data were reserved for forecast testing, and the remaining data were used as training data. Forecast testing was only done on the middle 3 yr of the 5-yr data segment to alleviate the leakage of low-frequency signals from the training data to the adjacent test data. We repeated the above process by moving the 5-yr window of test data forward by 3 yr each time until the whole record was used for forecast testing.

On the training data, a 7-fold cross-validation (CV2) was implemented to determine the optimal values for *p*, *n*PC and the hyperparameters: First the Cherkassky and Ma (2004) estimates were used for the hyperparameters, and the optimal *p* was estimated in CV2. Then a finer grid search for the optimal hyperparameter values and for the optimal *n*PC was undertaken in CV2. The model trained with these optimal *p*, *n*PC and hyperparameters was then used to forecast the test data under CV1. For BNN, the optimal number of hidden neurons to use in a neural network model was found from CV2.

3.4 Forecast skill scores

To evaluate model performance on forecasting the seasonal extreme precipitation, we reported the Pearsons correlation coefficient (CORR), the Willmott index of agreement (IOA) between the observed and model-predicted values, the skill score based on the mean absolute error (MAE) of the forecast, and $\mathrm{Skill_V}=\mathrm{SD_p}/\mathrm{SD_o}$, the ratio of the standard deviation (SD) of the model predictions to that of the observations. All four skill

HESSD
7, 3521–3550, 2010

scores are used because they indicate different components of model error. While CORR is a common measure of the linear dependence between the forecast and the observation, it does not take forecast bias into account, thus it is possible for a forecast with large errors to still have a good CORR score. IOA is defined as (Willmott, 1982):

$$\mathrm{IOA} = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}\left(|P_i - \overline{O}| + |O_i - \overline{O}|\right)^2}, \tag{6}$$

where N is the number of samples at the station, O_i and P_i are, respectively, the observed and predicted values for the ith sample, $\overline{O}$ is the average of the observed values, and $0 \leq \mathrm{IOA} \leq 1$, with 1 being perfect score. IOA has been proposed as an alternative to CORR, but it is sensitive to the difference between the mean of P_i and $\overline{O}$ as well as the difference between the standard deviation of P_i and that of O_i. MAE measures the mean absolute error between the observed and predicted values, i.e.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|P_i - O_i|. \tag{7}$$

MAE is considered a more natural and superior measure of average error than the commonly used root mean squared error (Willmott and Matsuura, 2005). To compare forecasting performance across different regions, instead of MAE, we used the MAE skill score (MAESS), defined by $\mathrm{MAESS}=1-\mathrm{MAE}/\mathrm{MAE_c}$, where

$$\mathrm{MAE_c} = \frac{1}{N}\sum_{i=1}^{N}\left|\overline{O} - O_i\right|, \tag{8}$$

is the MAE of the climatological forecasts. The MAESS is positive (negative) when the accuracy of the forecasts is greater (less) than the accuracy of the climatological forecasts. The $\mathrm{Skill_V}$ score is used to measure how close the predicted standard deviation approaches the observed one, with the perfect score being 1.

HESSD
7, 3521–3550, 2010

4 Forecast results

The cross-validated forecast scores averaged over all stations in each region at lead times of 3, 6, 9 and 12 months using the MLR, SVR with linear kernel (SVR-L), nonlinear SVR with RBF kernel (SVR-R) and BNN models are shown in Figs. 2–7 for the six regions.

For the Pacific coastal area (Fig. 2), CORR, IOA and MAESS showed that in general the SVR-R model tended to do slightly better than the SVR-L model, and both did better than the MLR and BNN models. Only the SVR-R model attained slightly positive MAESS (Fig. 2c), while both linear models and BNN displayed negative MAESS, indicating that they underperformed climatological forecasts. The relatively poor performance in MAESS can be partly explained by the Skill_V score shown in Fig. 2d, which shows all models dramatically under-predicting the magnitude of the anomalies. Ironically, BNN had the best Skill_V scores as well as the worst MAESS among the four models. The reason is that BNN being a non-robust nonlinear model is easily overfitted to the very noisy data. Even with an ensemble average to alleviate overfitting, BNN generated relatively large amplitude forecasts compared to the other three models which are less prone to overfitting.

For the Cordilleran region (Fig. 3), CORR, IOA and MAESS again showed that in general the SVR-R model did slightly better than the SVR-L model, and both did better than the MLR and BNN models, which actually attained negative CORR scores for all lead times. Although all four models had negative MAESS (Fig. 3c), partly due to their under-predicting the magnitude of the anomalies (Fig. 3d), the SVR-R did slightly better than the SVR-L, and much better than MLR and BNN in terms of the MAESS. The forecast skills, even with the SVR-R model, are very modest, but interestingly, there is no significant decline in skills as the lead time increased, suggesting that the skills came from low-frequency signals in the climate system. Overall, the skills are lower in the Cordilleran region (R2) than in the Pacific coastal region (R1).

HESSD
7, 3521–3550, 2010

The overall skills in the dry Eastern Prairies region (R3) (Fig. 4) is in general higher than those in regions R1 and R2. For CORR (Fig. 4a), the SVR-R model shows the best forecast performance at all lead times, with values above 0.4 up to 6-month lead time. Both SVR models did much better than MLR and BNN, and up to a lead time of 9 months, the advantage of the SVR-R model over the SVR-L model is clearly manifested. In the MAESS, SVR-R did best among the four models, outperforming climatological forecasts due to its positive MAESS values. In $\mathrm{Skill_V}$ (Fig. 4d), again BNN forecasted anomalies with standard deviations most similar to those observed, however, the BNN forecasts are unimpressive in terms of the CORR, IOA and MAESS scores. Overall, among all 6 regions, this region shows the highest forecast scores and the clearest advantage of incorporating nonlinearity in the models. Shabbar and Barnston (1996) also found the highest winter seasonal mean skill in the Eastern Prairies in their canonical correlation analysis prediction model.

For the dry Arctic region (R4), the overall skills (Fig. 5) were about the second highest among the six regions. Here the SVR-L model clearly outperformed the MLR, but the SVR-R model did not improve on the SVR-L model, and in fact did slightly worse. Hence, in contrast to the Eastern Prairies, the Arctic region shows that incorporating nonlinearity was unnecessary; however, incorporating robustness in the linear regression was important as MLR could not match SVR-L. The Great Lake region (R5) (Fig. 6) and the Atlantic coastal region (R6) (Fig. 7) have the lowest skills among the six regions. In both regions, both the SVR-L and SVR-R models did better than MLR and BNN, suggesting that robustness helps. However, there is no significant difference between the SVR-R and SVR-L model scores, hence incorporating nonlinearity did not lead to significant improvement in these two regions.

As to the dependence of forecast skills on the lead time, the Eastern Prairies (R3) showed highest skills at 3-month lead, followed by gradual decline with longer leads, while the Pacific coast (R1) showed highest skill at 6-month lead. Surprisingly the other four regions all showed highest skills at 9-month lead, indicating that the signal is of low-frequency origin and forecasts of winter extreme precipitation made during

HESSD
7, 3521–3550, 2010

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

the summer-autumn seasons were actually less skillful than those made earlier in the year, which is not unprecedented as the ENSO system is also known to have a “spring barrier”, i.e. lower skills for forecasts issued during spring (Xue et al., 1994).

Instead of averaging forecast skills over each region, we next displayed the forecast correlation skills of the SVR-R model at each station (Fig. 8). Positive correlation occurred at the vast majority of stations at all lead times, with a few stations constantly showing good performance for all the lead times. The SVR-R model gave several of its best forecasts for the Arctic stations, with correlations around 0.7 – although over all stations in the Arctic region, the averaged correlation was lower than that in the Eastern Prairies region. Over most stations in the Eastern Prairies, the forecast correlation was around 0.3 to 0.5. For the Pacific coastal area, the correlation was always above 0.3 except at the 9-month lead. For the Western Cordillera, Great Lakes and Atlantic areas, the forecast skills were weaker.

The spatial distribution of the difference in the correlation scores between the SVR-R model and the SVR-L model (Fig. 9) shows that at most stations, the nonlinear SVR method has an advantage over the linear one, though its advantage decreases with increasing lead time. The advantage of the nonlinear SVR model is most prominent in the Eastern Prairies up to a lead time of 9 months. Nonlinearity is not advantageous over the Arctic region.

A similar plot displaying the difference between the SVR-R model and the MLR (Fig. 10) shows that with only a few exceptions, the SVR-R model is clearly superior to the MLR model over all areas and at all lead times, by at least 0.2 on average in terms of the correlation score.

5 Conclusions

SVR models, with linear and RBF kernels, have been applied to predict the seasonal extreme precipitation anomalies in winter over Canada. In general, the robust SVR models clearly outperformed the non-robust MLR and BNN models in terms of forecast

HESSD

7, 3521–3550, 2010

skills, thereby demonstrating the value of models with robust error norms for dealing with the very noisy and non-Gaussian winter extreme precipitation data. Meanwhile the performance of the nonlinear SVR model (SVR-R) tended to be slightly better than the linear SVR model (SVR-L), with the exception of the Arctic region, which seemed to lack a nonlinear signal.

The strongest nonlinearity was found over the Eastern Prairies according to the difference in the forecast performance between the SVR-R and SVR-L models. This indicates that in the Eastern Prairies, gains in forecast skills came not only from using a robust error norm, but also from the nonlinear influence of climate fluctuations such as ENSO and other teleconnections on the extreme precipitation.

Arctic winter precipitation is different from that in the other five regions as all of Arctic winter precipitation is snow with relatively low water content that is easily moved by strong winds. There are many occurrences of strong winds in the Arctic winter where snow is advected from other areas under clear skies, which causes biases in catchments. This may explain why Arctic winter precipitation appears more linear than the precipitation in other regions.

Comparing the skill levels of the six regions, we found highest skill in the Eastern Prairies, presumably due to the strong nonlinear signal there, followed by the Arctic (despite the lack of a nonlinear signal), then the Pacific coastal region, followed by the Cordillera region, and finally by the low skill regions of the Atlantic coast and the Great Lakes, where presumably the lack of a strong ENSO signal there contributed to the low skills (Shabbar et al., 1997; Shabbar, 2006).

A disadvantage of nonlinear methods such as SVR and BNN is that it is generally futile to determine the contribution of forecast skill from individual predictors when there are many predictors. The compression of predictors by PCA further made determining the contributions from individual predictors infeasible.

HESSD
7, 3521–3550, 2010

Acknowledgements. This work was supported by the Canadian Foundation for Climate and Atmospheric Sciences, and the Natural Sciences and Engineering Research Council of Canada. Alex Cannon and Aiming Wu provided helpful discussion.

References

Anandhi, A., Srinivas, V. V., Nanjundiah, R. S., and Kumar, D. N.: Downscaling precipitation to river basin in India for IPCCSRES scenarios using support vector machine, Int. J. Climatol., 28, 401–420, 2008.

Barnston, A. G.: Linear statistical short-term climate predictive skill in the Northern Hemisphere, J. Climate, 7, 1513–1564, 1994.

Bishop, C. M.: Pattern Recognition and Machine Learning, Springer, New York, 2006.

Bonsai, B. R., Prowse, T. D., Duguay, C. R., and Lacroix, M. P.: Impacts of large-scale teleconnections on freshwater-ice break/freeze-up dates over Canada, J. Hydrol., 33, 340–353, 2006.

Bürger, C. M., Kolditz, O., Fowler, H. J., and Blenkinsop, S.: Future climate scenarios and rainfall-runoff modelling in the Upper Gallego catchment (Spain), Environ. Pollut., 148, 842–854, 2007.

Cannon, A. J. and Hsieh, W. W.: Robust nonlinear canonical correlation analysis: application to seasonal climate forecasting, Nonlinear Proc. Geoph., 12, 221–232, 2008.

Chang, C. C. and Lin, C. J.: LIBSVM: a library for support vector machines, Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm, last access: 21 March 2008, 2001.

Cherkassky, V. and Ma, Y.: Practical selection of SVM parameters and noise estimation for SVM regression, Neural Networks, 17, 113–126, 2004.

Dibike, Y. B., Velickov, S., Solomatine, D., and Abbott, M. B.: Model induction with support vector machines: introduction and applications, J. Comput. Civil Eng., 15, 208–216, 2001.

Easterling, D. R., Evans, J. L., Groisman, P. Y., Karl, T. R., Kunkel, K. E., and Ambenje, P.: Observed variability and trends in extreme climate events: A brief review, B. Am. Meteorol. Soc., 81, 417–425, 2000.

Goddard, L., Mason, S. J., Zebiak, S. E., Ropelewski, C. F., Basher, R., and Cane, M. A.: Current approaches to seasonal to interannual climate predictions, Int. J. Climatol., 21, 1–64, 2001.

HESSD
7, 3521–3550, 2010

Groisman, P. Y., Knight, R. W., Easterling, D. R., and Karl, T. R.: Trends in intense precipitation in the climate record, J. Climate, 18, 1326–1350, 2005.
Haupt, S. E., Pasini, A., and Marzban, C.: Artificial Intelligence Methods in the Environmental Sciences, Springer, Berlin, 2008.
Hsieh, W. W.: Machine Learning Methods in the Environmental Sciences. Cambridge Univ. Pr., Cambridge, 2009.
Hsieh, W. W. and Cannon, A. J.: Towards robust nonlinear multivariate analysis by neural network methods, in: Nonlinear Time Series Analysis in the Geosciences – Applications in Climatology, Geodynamics, and Solar-Terrestrial Physics, edited by: Donner, R. and Barbosa, S., Springer, Berlin, 97–124, 2008.
Hsieh, W. W., Wu, A., and Shabbar, A.: Nonlinear atmospheric teleconnections, Geophys. Res. Lett., 33, L07714, doi:10.1029/2005GL025471, 2006.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-yr reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, 1996.
Khan, M. S. and Coulibaly, P.: Application of support vector machine in lake water level prediction, J. Hydrol. Eng., 11, 199–205, 2006.
Kunkel, K. E.: North American trends in extreme precipitation, Nat. Hazards, 29, 291–305, 2003.
Lin, H., Brunet, G., and Derome, J.: Seasonal forecasts of Canadian winter precipitation by post-processing GCM integrations, Mon. Weather Rev., 136, 769–783, 2008.
Lorenzo, M. N., Taboada, J. J., and Gimeno, L.: Links between circulation weather types and teleconnection patterns and their influence on precipitation patterns in Galicia (NW Spain), Int. J. Climatol., 28, 1493–1505, 2008.
MacKay, D. J. C.: Bayesian interpolation, Neural Comput., 4, 415–447, 1992.
Nabney, I. T.: NETLAB: Algorithms for Pattern Recognition, Springer, London, 2002.
Shabbar, A. and Barnston, A. G.: Skill of seasonal climate forecasts in Canada using canonical correlation analysis, Mon. Weather Rev., 124, 2370–2385, 1996.
Shabbar, A., Bonsal, B., and Khandekar, M.: Canadian precipitation patterns associated with the Southern Oscillation, J. Climate, 10, 3016–3027, 1997.
Shabbar, A.: The impact of El Niño-Southern Oscillation on the Canadian climate, Adv. Geosci., 6, 149–153, doi:10.5194/adgeo-6-149-2006, 2006.

Smith, T. M., Reynolds, R. W., Peterson, T. C., and Lawrimon, J.: Improvements to NOAAs historical merged land-ocean surface temperature analysis (1880–2006), J. Climate, 21, 2283–2296, 2008.

Solomatine, D. P. and Ostfeld, A.: Data-driven modelling: some past experiences and new approaches, J. Hydroinform., 10, 3–22, 2008.

Vapnik, V.: The Nature of Statistical Learning Theory, Springer, New York, 1995.

Vapnik, V., Golowich, S., and Smola, A.: Support vector method for function approximation, regression estimation, and signal processing, in: Advances in Neural Information Processing Systems, 9, edited by: Mozer, M., Jordan, M., and Petsche, T., MIT Press, Cambridge, MA, 281–287, 1997.

Ward, N. M. and Folland, C. K.: Prediction of seasonal rainfall in the North Nordeste of Brazil using eigenvector of sea-surface temperatures, Int. J. Climate, 12, 3167–3191, 1991.

Whitfield, P. H., Bodtker, K., and Cannon, A. J.: Recent variations in seasonality of temperature and precipitation in Canada, 1976–95, Int. J. Climatol., 22, 1617–1644, 2002.

Willmott, C. J.: Some comments on the evaluation of model performance, B. Am. Meteorol. Soc., 63, 1309–1313, 1982.

Willmott, C. J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Climate Res., 30, 79–82, doi:10.3354/CR030079, 2005.

Wu, A., Hsieh, W. W., and Shabbar, A.: The nonlinear patterns of North American winter temperature and precipitation associated with ENSO, J. Climate, 18, 1736–1752, doi:10.1175/JCLI3372.1, 2005.

Wu, A., Hsieh, W. W., Shabbar, A., Boer, G. J., and Zwiers, F. W.: The nonlinear association between the Arctic Oscillation and North American winter climate, Clim. Dynam., 26, 865–879, doi:10.1007/s00382-006-0118-8, 2006a.

Wu, A., Hsieh, W. W. and Tang, B.: Neural network forecasts of the Tropical Pacific sea surface temperatures, Neural Networks, 19, 145–154, 2006b.

Yuval and Hsieh, W. W.: The impact of time-averaging on the detectability of nonlinear empirical relations, Q. J. Roy. Meteor. Soc., 128, 1609–1622, 2002.

Xue, Y., Cane, M. A., Zebiak, S. E., and Blumenthal, M. B.: On the prediction of ENSO – a study with a low-order Markov model, Tellus A, 46, 512–528, 1994.

Zhang, X., Hogg, W. D., and Mekis, E.: Spatial and temporal characteristics of heavy precipitation events in Canada, J. Climate, 14, 1923–1936, 2001.

HESSD
7, 3521–3550, 2010

Table 1. Number of stations, mean winter precipitation, and percentage variance of the winter extreme precipitation anomalies explained by the first several PCs (with *n*PC being the number of PCs chosen as predictands based on cross-validation), for each of the six regions over Canada.

Region	Stns.	Mean (mm)	% var. (*n*PC)
R1 (Pacific Coast)	20	78.51	85 (7)
R2 (Cordillera)	43	25.9	80 (8)
R3 (Eastern Prairies)	19	14.2	73 (4)
R4 (Arctic)	11	8.8	65 (4)
R5 (Great Lakes)	15	36.8	79 (7)
R6 (Atlantic Coast)	10	63.8	95 (8)

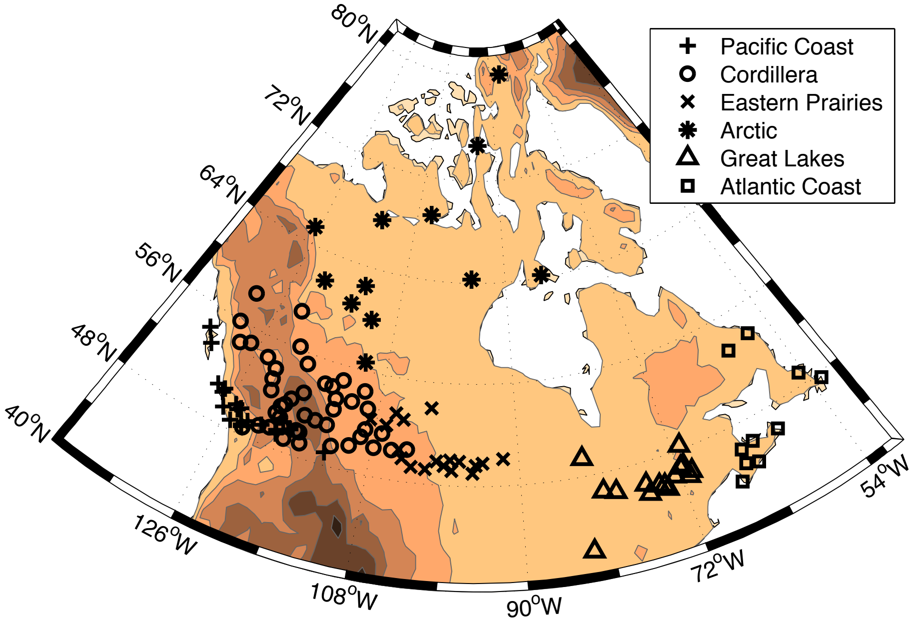

Fig. 1. Spatial distribution of the Canadian stations, with different symbols used to indicate the six geographic regions determined by a cluster analysis. The shading illustrates the Canadian topography.

HESSD

7, 3521–3550, 2010

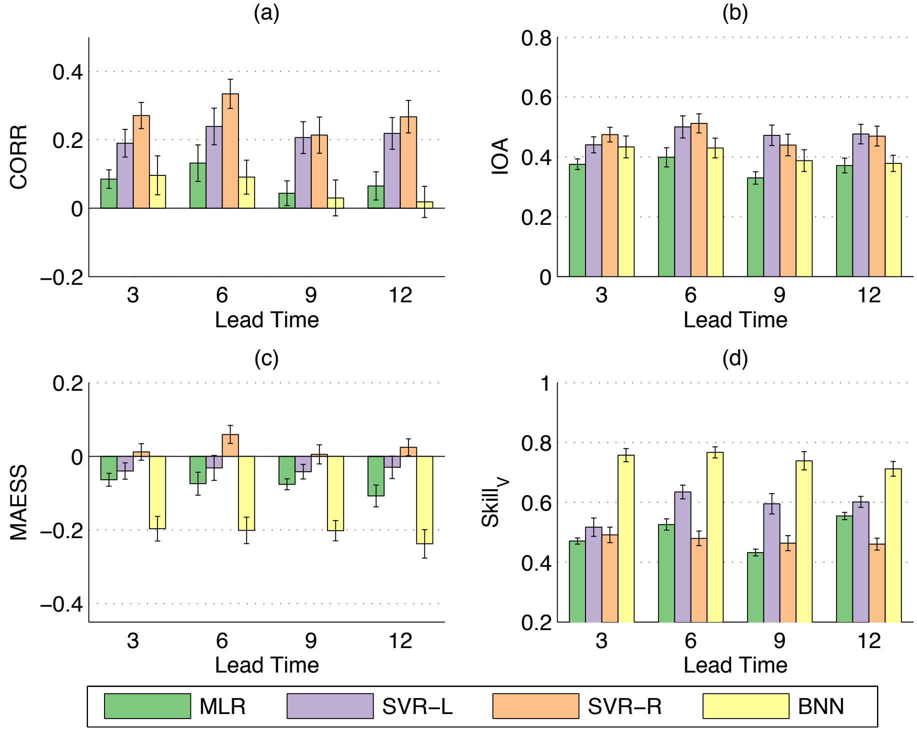

Fig. 2. Cross-validated forecast scores, **(a)** CORR, **(b)** IOA, **(c)** MAESS and **(d)** Skill_V, averaged over all stations in the Pacific coastal region (region R1) for the winter extreme precipitation at lead times of 3, 6, 9 and 12 months using the MLR, SVR with linear kernel (SVR-L), nonlinear SVR with RBF kernel (SVR-R) and BNN models. The error bars indicate ±1 standard error of the mean. Lead time of 3 months means that predictor data up till September–November were used to forecast the December–February extreme precipitation.

HESSD
7, 3521–3550, 2010

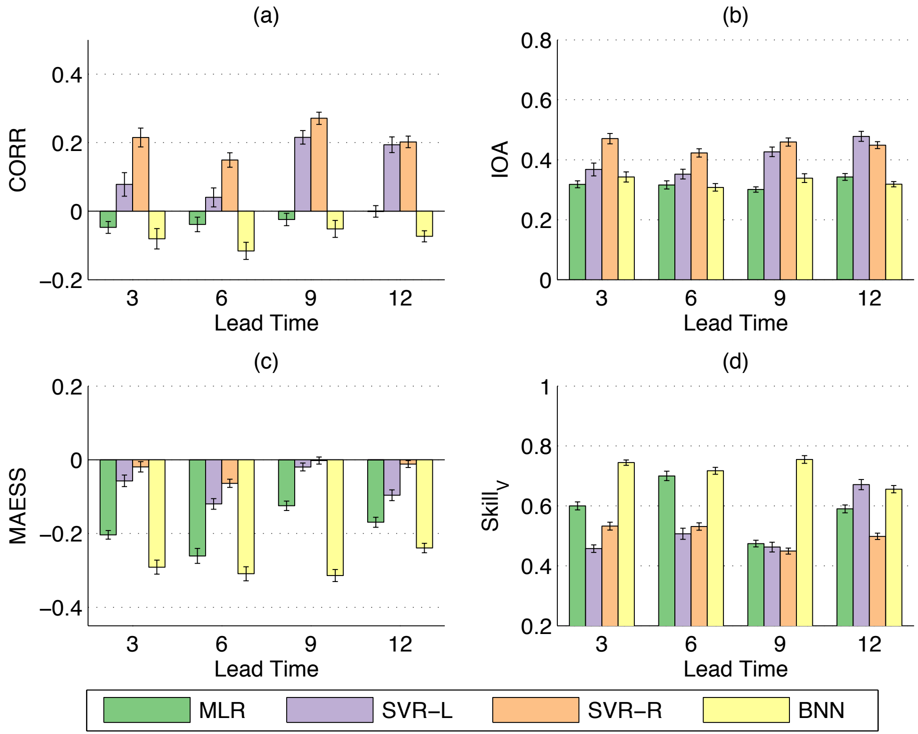

Fig. 3. Same as Fig. 2, except over the Cordillera (region R2).

HESSD
7, 3521–3550, 2010

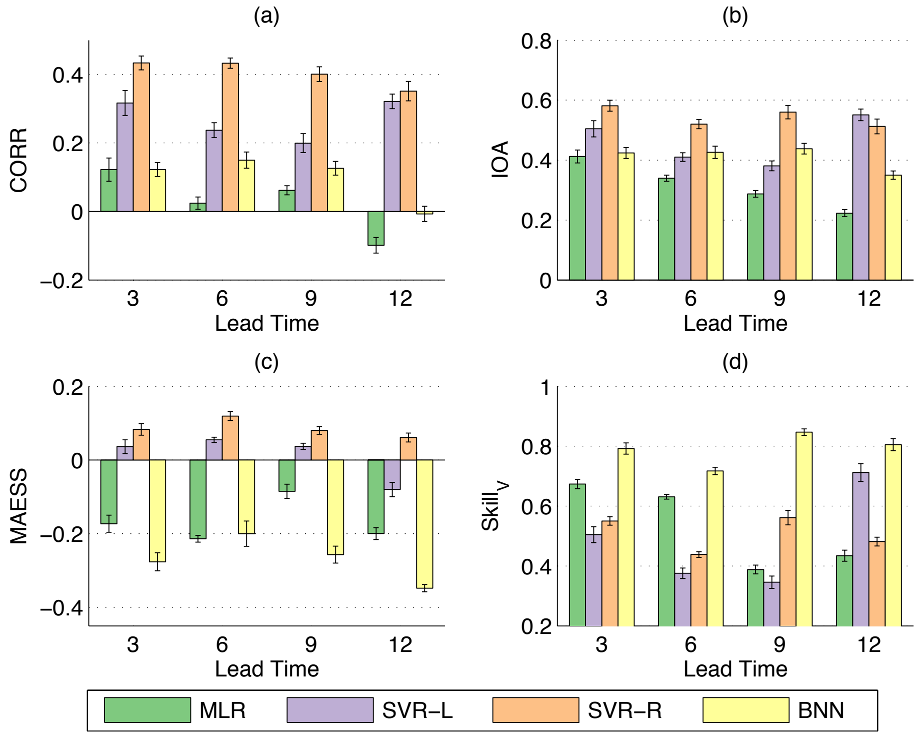

Fig. 4. Same as Fig. 2, except over the Eastern Prairies (region R3).

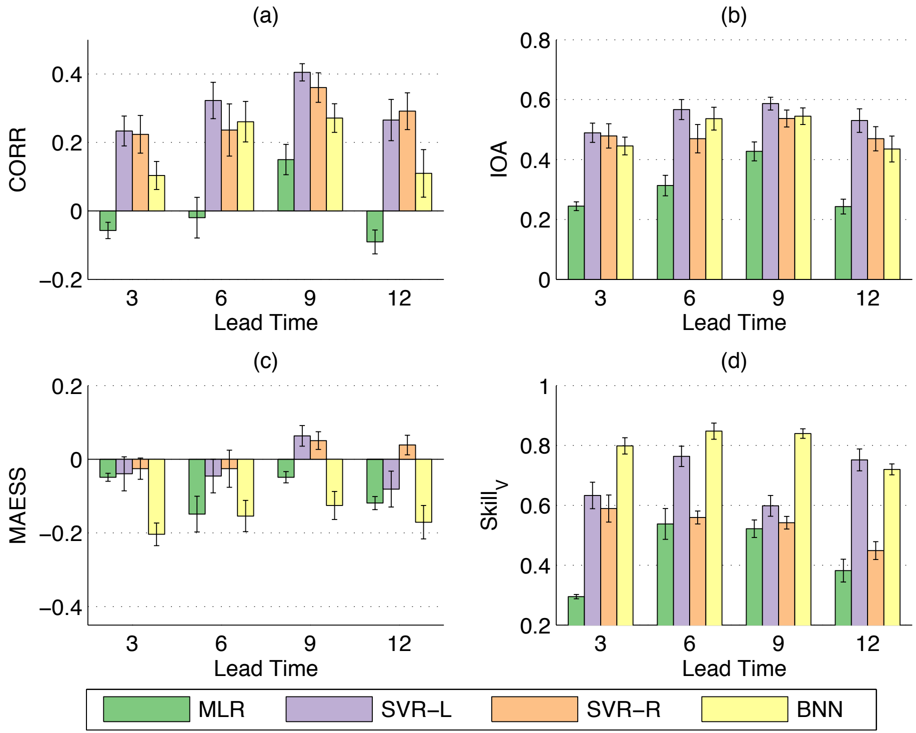

Fig. 5. Same as Fig. 2, except over the Arctic (region R4).

HESSD

7, 3521–3550, 2010

Seasonal prediction of winter extreme precipitation

Z. Zeng et al.

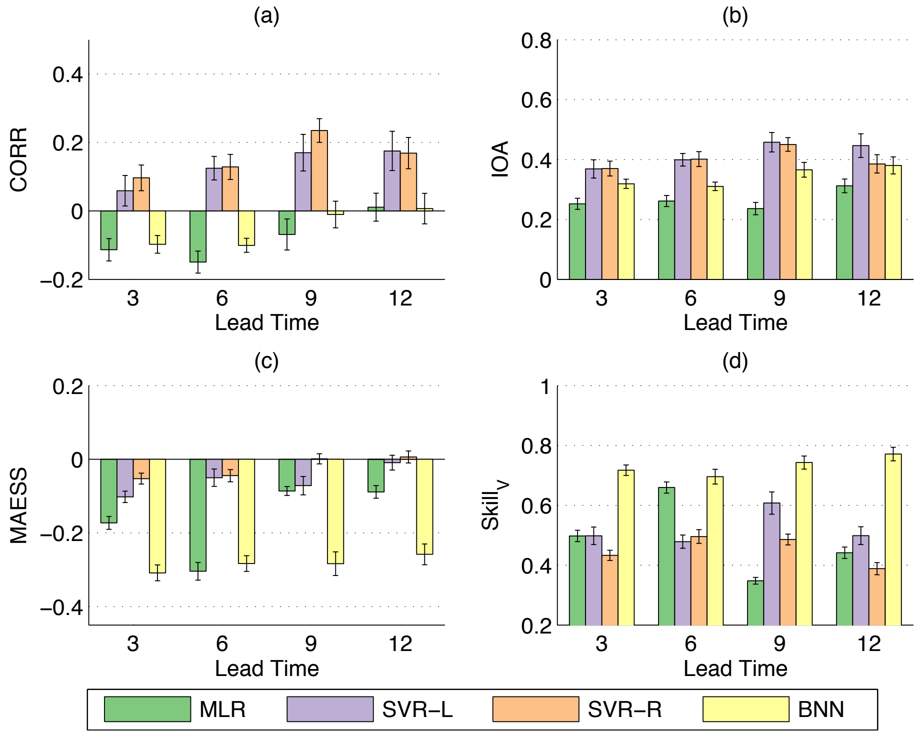

Fig. 6. Same as Fig. 2, except over the Great Lakes (region R5).

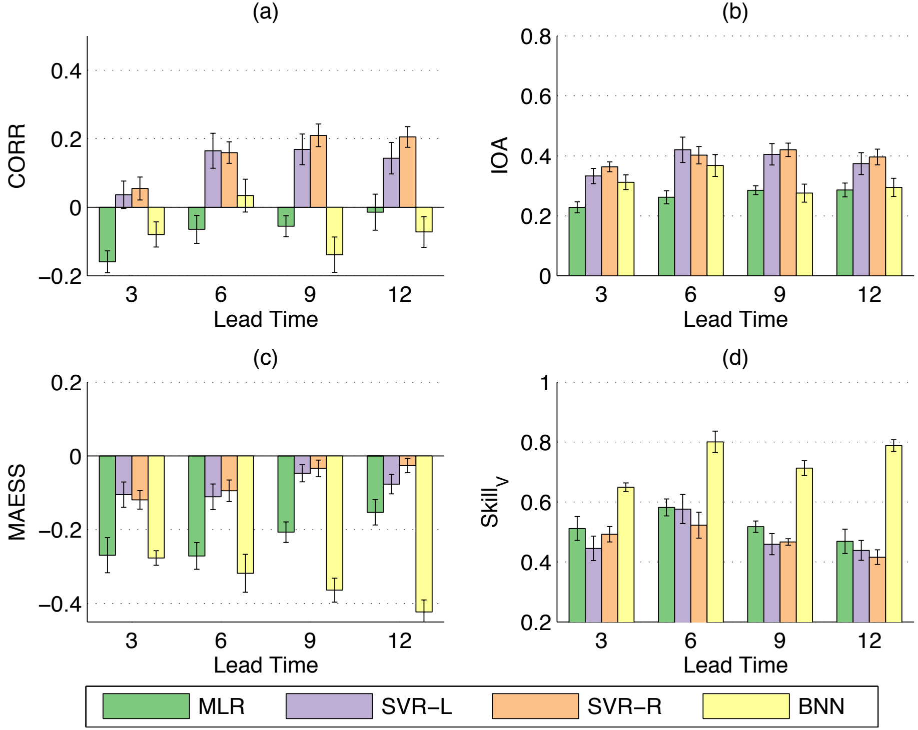

Fig. 7. Same as Fig. 2, except over the Atlantic coast (region R6).

HESSD
7, 3521–3550, 2010

Seasonal prediction of winter extreme precipitation

Z. Zeng et al.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

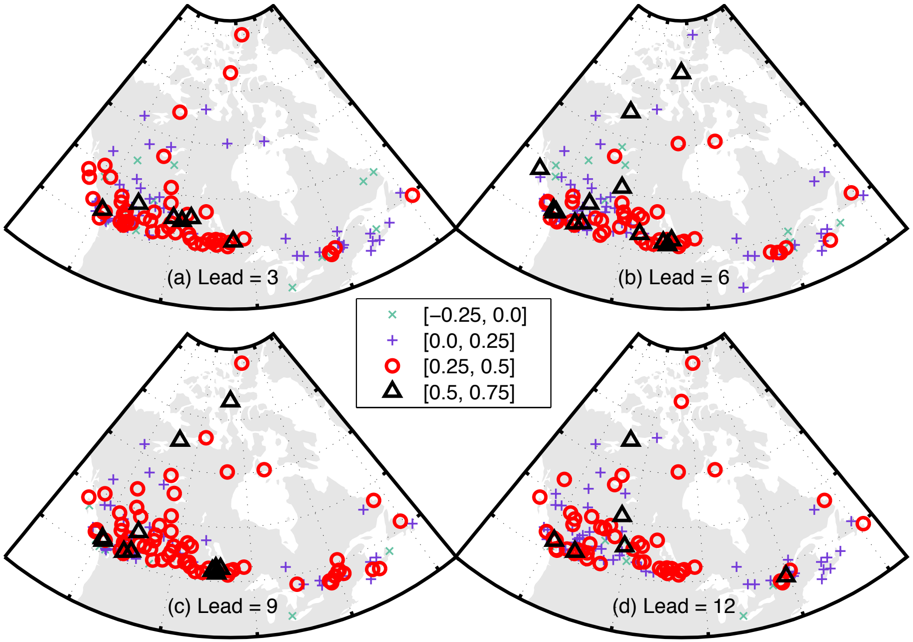

Fig. 8. Spatial distribution of the forecast correlation skills of the SVR-R model at individual stations over Canada at lead times of **(a)** 3, **(b)** 6, **(c)** 9 and **(d)** 12 months.

HESSD

7, 3521–3550, 2010

Seasonal prediction of winter extreme precipitation

Z. Zeng et al.

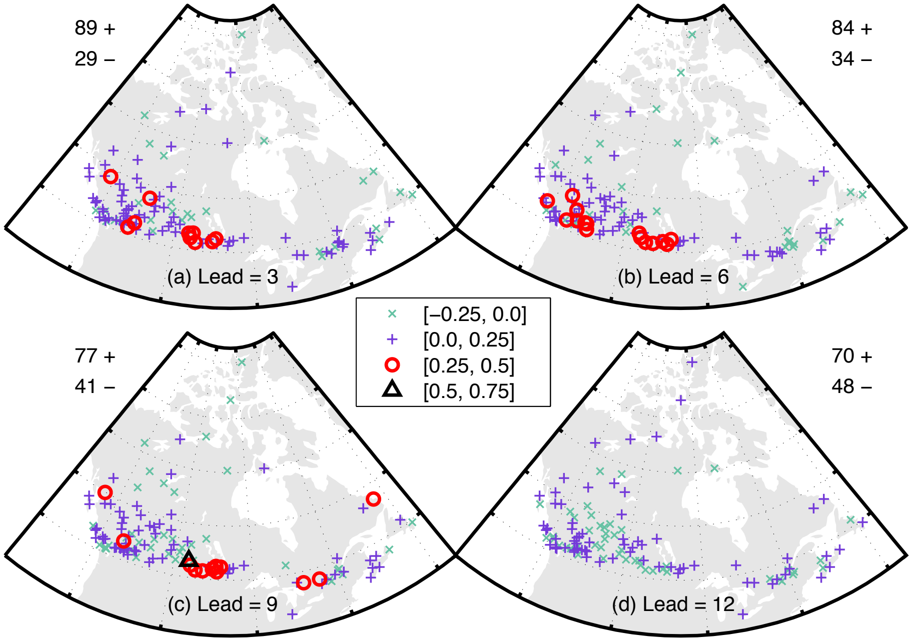

Fig. 9. Difference between the forecast correlation skills of the nonlinear SVR model (SVR-R) and that of the linear SVR model (SVR-L) at lead times of **(a)** 3, **(b)** 6, **(c)** 9 and **(d)** 12 months. The two numbers beside each panel give the number of stations where the SVR-R correlation is higher (lower) than that of the SVR-L model, as indicated by the +(−) sign.

HESSD
7, 3521–3550, 2010

Seasonal prediction of winter extreme precipitation

Z. Zeng et al.

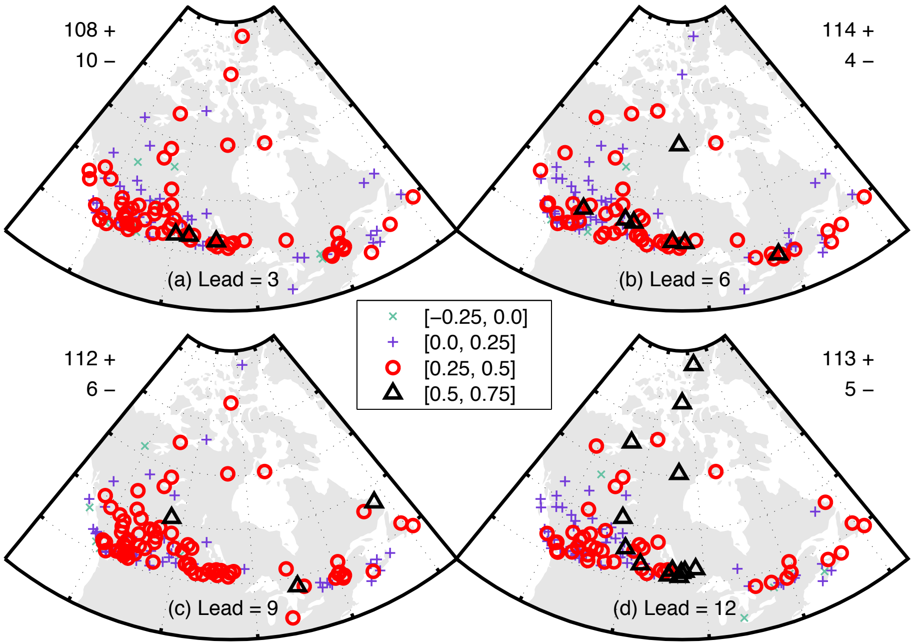

Fig. 10. Difference between the forecast correlation skills of the SVR-R model and that of the MLR model at lead times of **(a)** 3, **(b)** 6, **(c)** 9 and **(d)** 12 months. The two numbers beside each panel give the number of stations where the SVR-R correlation is higher (lower) than that of the MLR model, as indicated by the +(–) sign.

HESSD
7, 3521–3550, 2010