Scientific
Research
Publishing

Identification of Influential Sea Surface Temperature Locations and Predicting Streamflow for Six Months Using Bayesian Machine Learning Regression

N. K. Shrestha[1*], G. Urroz[2]

[1]Department of Agricultural and Biological Engineering, University of Florida, Immokalee, USA
[2]Department of Civil and Environmental Engineering, Utah State University, Logan, USA
Email: [*]nirojshrestha@ufl.edu

Abstract

Sea surface temperature (SST) has significant influence in the hydrological cycle and affects the discharge in the stream. SST is an atmospheric circulation indicator which provides the predictive information about the hydrologic variability in the region around the world. Use of right location of SST for a given location of stream gage can capture the effect of oceanic-atmospheric interaction, improving the predictive ability of the model. This study aims on identifying the best locations of SST at the selected stream gage in the state of Utah that spatially covers the state from south to north, and use them for next six-month streamflow volume predictions. The data-driven model derived from the statistical learning theory was used in this study. Using an appropriate location of SST together with local climatic conditions and state of basin, an accurate and reliable stream-flow was predicted for next six months. Influence of Pacific Ocean SST was observed to be stronger than that of Atlantic Ocean SST in the state of Utah. The SST of North Pacific developed the best model in most of the selected stream gages. Each model was ensured to be robust by the bootstrap analysis. The long-term streamflow prediction is important for water resource planning and management in the river basin scale and is a key step for successful water resource management in arid regions.

Keywords

Streamflow, Prediction, SWE, Temperature, RVM

---

[*]Corresponding author.

1. Introduction

Streamflow depends on the distribution of precipitation in time and space which further depends on the climatic conditions. Shivakumar [1] observed that the monthly and annual streamflow series are affected by long-term climate. The atmospheric and hydrologic sciences have recently used sea surface temperature (SST) to predict streamflow variability [2]-[4]. SST is an indicator of oceanic-atmospheric circulation and has important consequences on the weather around the globe. SST has strong link with the hydrology of the river basin which provides the predictive information about the hydrologic variability [5]. Identification of an appropriate location of SST is likely to improve the predictive ability of the model. Therefore input of climatic conditions in the model through SST has a significant importance.

The precise information about the quantity of water availability in the next season can be very useful for the agricultural planning, watershed management, and other decision making processes [6]. It can benefit the management of water resources, in particular allowing decision on water allocation for irrigations [7]-[9] and other purposes. Financial commitment made by the farmers early in the season can result in substantial economic losses if the resulting seasonal flow does not subsequently supply enough irrigation water. Forecast with long-lead time facilitates co-ordination between different system users that may be important in multiple-use water resource systems [10].

Machine learning regression model has been used as an alternative to physically based models. The complexities in the physically-based models and difficulties associated with the data acquisitions and corresponding expenses that these models would require have limited the application of such models. Machine learning models are good at capturing the underlying physics of the system by relating input and output. They are robust and are capable of making reasonable predictions using historical data [11]. Artificial Neural Network (ANN), Support Vector Machine (SVM), and Relevance Vector Machine (RVM) are few popular machine learning models. The disadvantage of the ANN model is that it may get stuck in local minima rather than global minima. The SVM model is a popular machine learning model [12] however this model makes unnecessary liberal use of the basis function. In the SVM model, the number of support vectors grows linearly with the size of data [13]. In addition, the SVM predictions are not probabilistic. The RVM is a Bayesian machine learning model. This is sparser than the SVM model and gives probabilistic output as well. Optimizing model parameter for the RVM model is relatively easier, however the performance is comparable. The RVM model has been successfully used by many past researches for water resources operation and management works [3] [14]. The RVM model is therefore proposed in this study. The objective of this study is to identify the best locations of SST for the selected locations of unimpaired stream gages (except one), and use them for the streamflow prediction.

2. Materials and Methods

2.1. Study Area

Five stream gages were selected at different locations of Utah which spatially covers the state from north to south (**Figure 1**). For most of the stream gages, the flows were not affected by diversion or regulation and the long year of systematic records were available. For Sixth Water Creek, the flow however, was partly affected by diversion until 2004. Two sites were chosen from northern region (Weber River near Oakley and Chalk Creek at Coalville), two from central region (Muddy Creek near Emery and Sixth Water Creek near Springville), and one from the southern region (Sevier River at Hatch) of the state. Snow accumulation and melt is a very significant feature in terms of annual hydrologic cycle for these streams [15]. **Table 1** shows the basin area, length of stream, and location of stream gages used in this study.

2.2. Relevance Vector Machine

Relevance Vector Machine is a supervised learning model based on sparse Bayesian learning. This is a model of identical functional form to the SVM developed by Vapnik [16] [17].

For the given input-target pair $\left\{x_n, t_n\right\}_{n=1}^{N}$ in training data set, the model learns a dependency of the targets (streamflow in this study) on the inputs (e.g. SST, snow and temperature data) with the objective of making accurate predictions of the target (t) for previously unseen values of input x [13] [18].

Target t_i is a sample from the model $\left(y_i\right)$ with additive noise $\left(\varepsilon_i\right)$ which has mean zero with variance σ^2 [13].
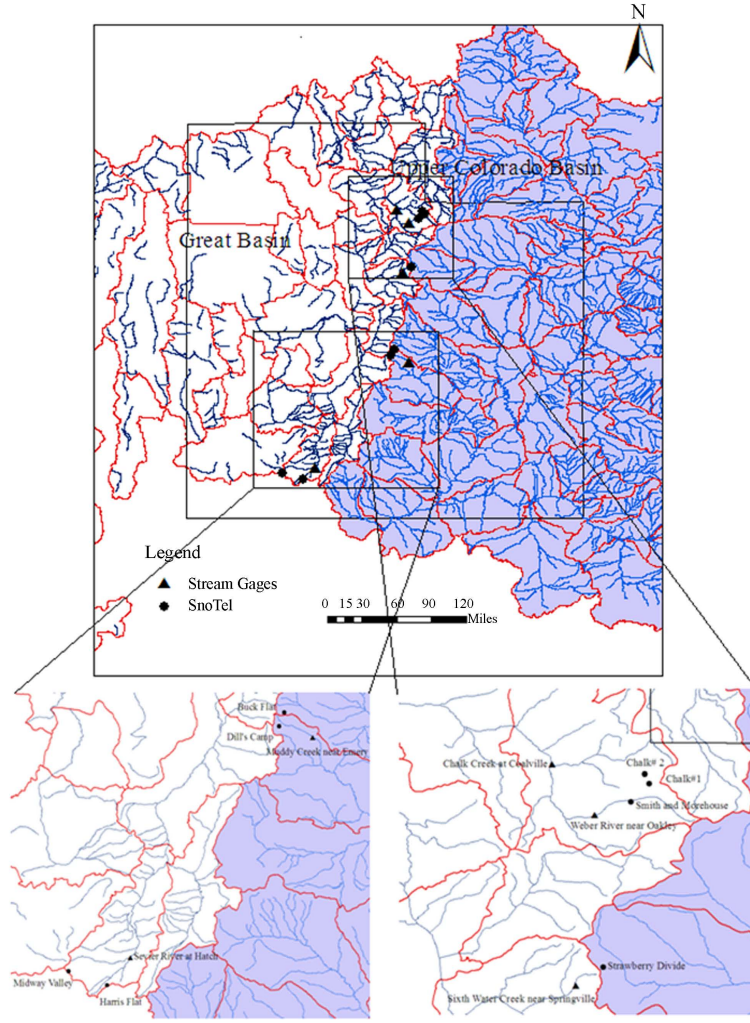
Figure 1. Location of the stream gages and SnoTel stations.

Table 1. Geometric characteristic of stream gages.

Site ID	Name	Basin	Stream		Gage Location	
		Area (Km^2)	Length (Km)	Slope	Latitude (°)	Longitude (°)
10128500	Weber River near Oakley	419.8	40.7	0.020	40.737	−111.247
10131000	Chalk Creek at Coalville	643.1	60.4	0.010	40.921	−111.401
10174500	Sevier River at Hatch	880.6	50.1	0.007	37.651	−112.430
09330500	Muddy Creek near Emery	272.0	32.3	0.004	38.982	−111.249
10149000	Sixth Water Creek near Springville	38.9	1.6	0.048	40.118	−111.314

$$t_i = y(x_i; w) + \varepsilon_i \tag{1}$$

The unknown function y is the product of design matrix (Φ) and weight parameter (w). In the vector form, Equation 1 can be written as,

$$t = \Phi w + \varepsilon$$

where the target and weight vector are expressed as $t = (t_1, \cdots, t_N)^{\text{T}}$ and $w = (w_1, \cdots, w_N)^{\text{T}}$, respectively. An

independent Gaussian noise is assumed. Thus, $p(t_i|x) \sim N(y(x_i;w),\sigma^2)$ and the likelihood of complete dataset is written as,

$$p(t|w,\sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{\frac{1}{2\sigma^2}\|t-\Phi w\|^2\right\}$$ (2)

The maximum likelihood estimate of w and σ^2 in Equation 2 may suffer from over fitting [13]. To avoid this, w is constrained with mean zero Gaussian prior probability, which results in majority of w being zero. This constrain makes the RVM model sparser than the SVM model [13]. The posterior covariance and mean of w, estimated from Bayes' rule [13] are $\Sigma = (\sigma^{-2}\Phi^T\Phi + A)^{-1}$ and $\mu = \sigma^{-2}\Sigma\Phi^T t$, respectively, and $A = \text{diag}(\alpha_0,\alpha_1,\cdots,\alpha_N)$, where α is uniform hyperpriors and diag (...) is a diagonal matrix. The α and σ^2 are estimated from an iterative re-estimation formula [13] given by,

$$\alpha_i^{\text{new}} = \frac{\gamma_i}{\mu_i^2} \text{ and } (\sigma^2)^{\text{new}} = \frac{\|t-\Phi\mu\|^2}{N-\sum_i\gamma_i},$$ (3)

where $\gamma_i = 1-\alpha_i\Sigma_{ii}$. The term μ_i is the ith posterior mean weight and N is the number of data examples (length of data set). The Σ_{ii} is i^{th} diagonal element of the posterior weight covariance computed with the current α and σ^2. The learning algorithm proceeds by iterative process of Equation (3) together with updating the posterior statistics Σ and μ, until suitable convergence criteria is satisfied [13].

The predictions for new input (x_*) are made based on the posterior distribution over the weights, conditioned on the maximizing values α and σ^2.

$$p(t_*|t,\alpha,\sigma^2) = N(t_*|y_*,\sigma_*^2),$$

where $y_* = \mu^T\phi(x_*)$, $\sigma_*^2 = \sigma^2 + \phi(x_*)^T\Sigma\phi(x_*)$, and ϕ is a basis function. Further details of the RVM model can be found in Tipping [13] [18]. The model used in this study is one introduced by Thayananthan [19]. This is a Bayesian regression tool extension of the RVM algorithm developed by Tipping and Faul [20]. Gaussian kernel was used in this study as it has shown to perform better than other kernels [21] [22].

2.3. Model Formulation

The model (Equation (4)) consists of predicting total volume of water passing the stream gage for next six months. Inputs to the model are past streamflow data, snow water equivalent and SnoTel temperature of nearby SnoTel stations and SST. The input variables are selected based on the underlying physical processes and climatic factors that influence the generation of streamflow.

$$Q_{t+6} = f\left(Q_{t-6}, \bar{S}_{t-12}, \bar{T}_{t-12}, \overline{SST}_{t-12}\right)$$ (4)

where Q_{t-6} is a total volume of water flowing through the gage in last six months, $\bar{S}_{t-12}$ and $\bar{T}_{t-12}$ are the average SWE and SnoTel temperature calculated over the last twelve months, $\overline{SST}_{t-12}$ represents 12 months previous monthly average sea surface temperature value. The output Q_{t+6} is the volume of water passing the stream gage for next six months.

Smith and Reynolds SST were used in this study which covers majority of world's ocean by 2° by 2° grid [23]. Six locations were selected from the Pacific and Atlantic oceans. They were North Pacific (NP), Central Pacific (CP), Tropical Pacific (TP), East Atlantic (EA), Middle Atlantic (MA) and Tropical Atlantic (TA) (Figure 2).

In Utah, snow is an important variable affecting the discharge in the stream. When the precipitation falls as snow, it settles, compacts and melts several months later and is prominent source of streamflow [24]. Snow serves as storage of water especially in the western US which has major effect on the streamflow in the spring and early summer months. Snow Water Equivalent (SWE) is a common term used in the hydrological modeling which is defined as equivalent depth of water when snow completely melts. SWE data are collected from the nearby SnoTel stations. Use of SWE data from more than one SnoTel stations improves prediction as it incorporates SWE spatial variability [2]. Harris Flat and Midway Valley SnoTel stations were used for Sevier River at Hatch, Smith and Morehouse and Chalk#1 were used for Weber River near Oakley, Chalk#1 and
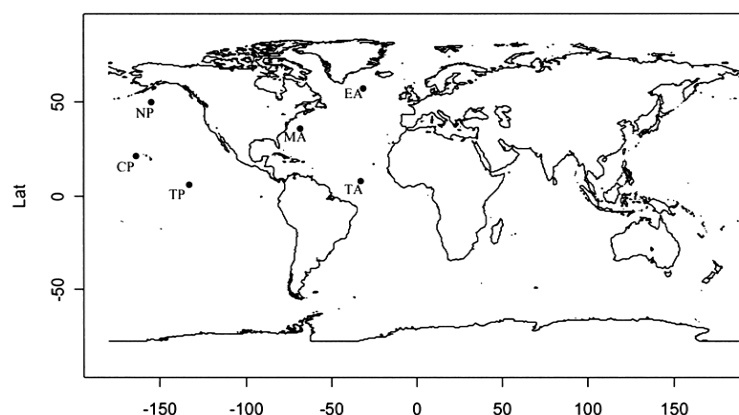
Figure 2. The locations for the sea surface temperature (from Khalil *et al.* [3]).

Chalk#2 were used for Chalk Creek at Coalville, Buck Flat and Dill's Camp were used for Muddy Creek near Emery and Strawberry Divide was used for Sixth Water Creek near Springville. Although some SnoTel sites were physically outside of the watershed, they were still included in the model due to their strong correlation with the streamflow processes. The SWE data were collected from Natural Resource Conservation Service (NRCS) (http://www.wcc.nrcs.usda.gov/snow). The period of 1980-2009 was used in this study because of the relative completeness of data in the basins for these years.

Temperature affects the melting rate of snow which consequently affects the discharge in the stream. The high discharge in the spring and early summer month is due to rising temperature when there is enough snowpack in the watershed. The temperature data were also collected from the SnoTel stations operated by NRCS and the period of data collection for local temperature was same as that of SWE.

The model was trained for 1980-2001 and tested on 2002-2009 for Weber River near Oakley, Chalk Creek at Coalville, and Muddy Creek near Emery. The Sevier River at Hatch was trained for 1982-2001 and tested for 2002-2009 while the Sixth Water Creek near Springville was trained for 2000 to 2006, and tested for 2007 to 2009. For the SST value, an individual as well as combinations of the SSTs were used for developing the best model. The best model was selected based on the test statistics (RMSE and Nash-Sutcliffe efficiency in the test phase).

3. Results and Discussion

3.1. Identification of Influential Sea Surface Temperature Locations and Prediction of the Volume of Water Passing the Gage for Next Six Months

The test statistics were computed for each individual SST for the volume of water passing through the stream gage for next six months. The SST locations that developed the best test statistics are shown in **Figure 3**.

A 95% confidence interval for the median RMSE is shown in **Table 2**. The test RMSE from the best identified SST locations is outside of the 95% confidence interval. This indicates that the test RMSE from the best SST location is significantly better than the test RMSE from the other SST locations. The summary results of the best SST locations for each stream gage are shown in **Table 3**.

Using the best SST locations, the volume of water passing through each selected stream gage was predicted (**Figure 4**). A good match between actual and predicted flow volume was obtained. The plot of predicted versus actual flow volume saturates about the bisector which indicates that the model prediction is close to actual values (**Figure 4**). The accuracy of the prediction was high for the unimpaired gages while it was relatively less for impaired gage (Sixth Water Creek near Springville). In general cases, the model has perfectly captured the high flow but the low flow was not captured accurately. Since the inputs representing the groundwater flow were not included in the model, this level of discrepancy is obvious. The overall prediction shows that the model is accurate and can be used for predicting six months ahead streamflow volume. The uncertainty of prediction was captured by confidence interval in the test phase for each gage.

The illustration about the best location of SST for the given location of stream gage is discussed below. When monthly data are used, the data consists of seasonal, annual, and inter-annual components. The effect of seasonal
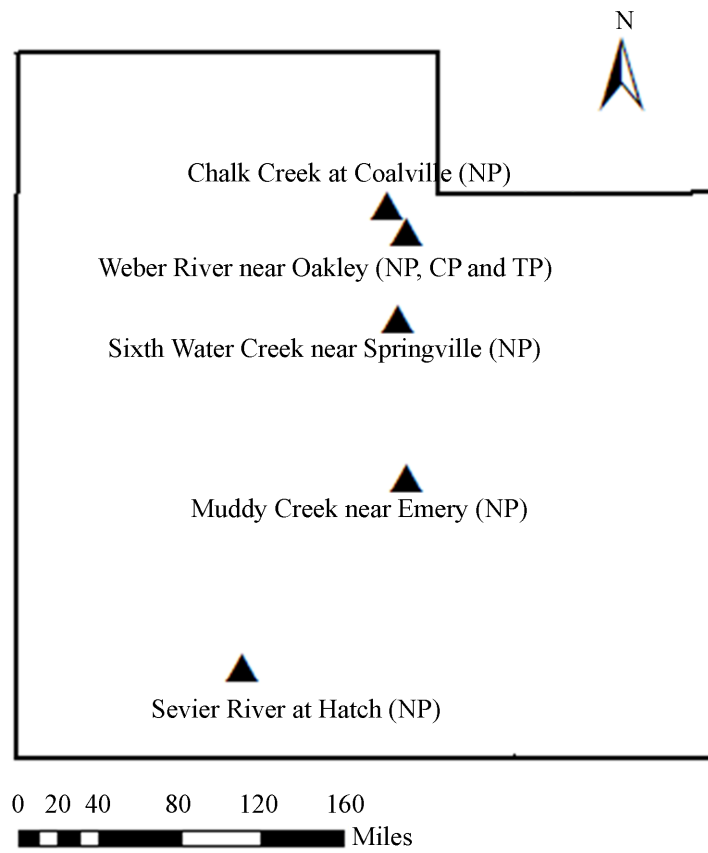
Figure 3. Locations of the sea surface temperature that developed the best test statistics for volume of water passing the stream gage for next six months.

**Table 2.** The 95% confidence interval of the median RMSE.

Streamflow Sites	95% Confidence Interval (1000 ac-ft)		Best RMSE (1000 ac-ft)	Remark
	Lower	Upper		
Weber River near Oakley	8.66	11.45	8.31	NP, CP and TP
Chalk Creek at Coalville	2.75	4.46	2.65	NP
Muddy Creek near Emery	2.52	4.11	2.44	NP
Sevier River at Hatch	5.36	7.72	5.04	NP
Sixth Water Creek near Springville	0.88	1.32	0.73	NP

**Table 3.** Summary of best test result for volume of water passing the gage for next six months.

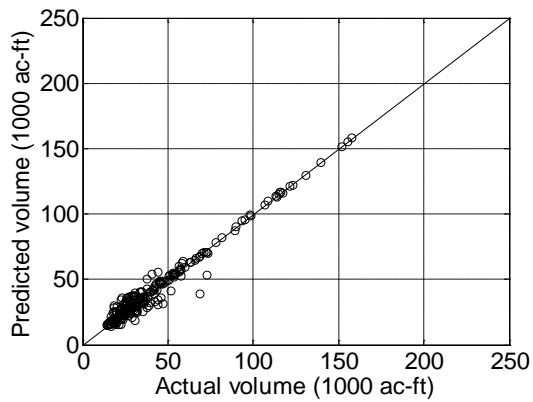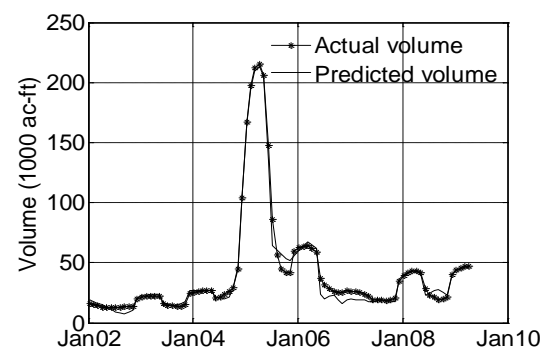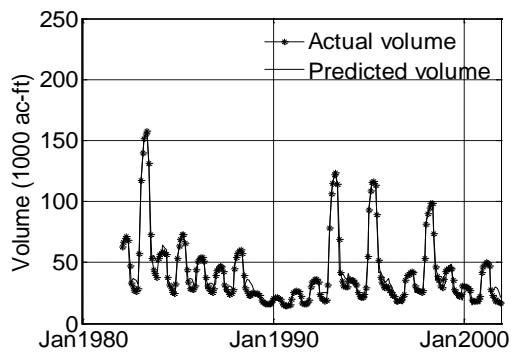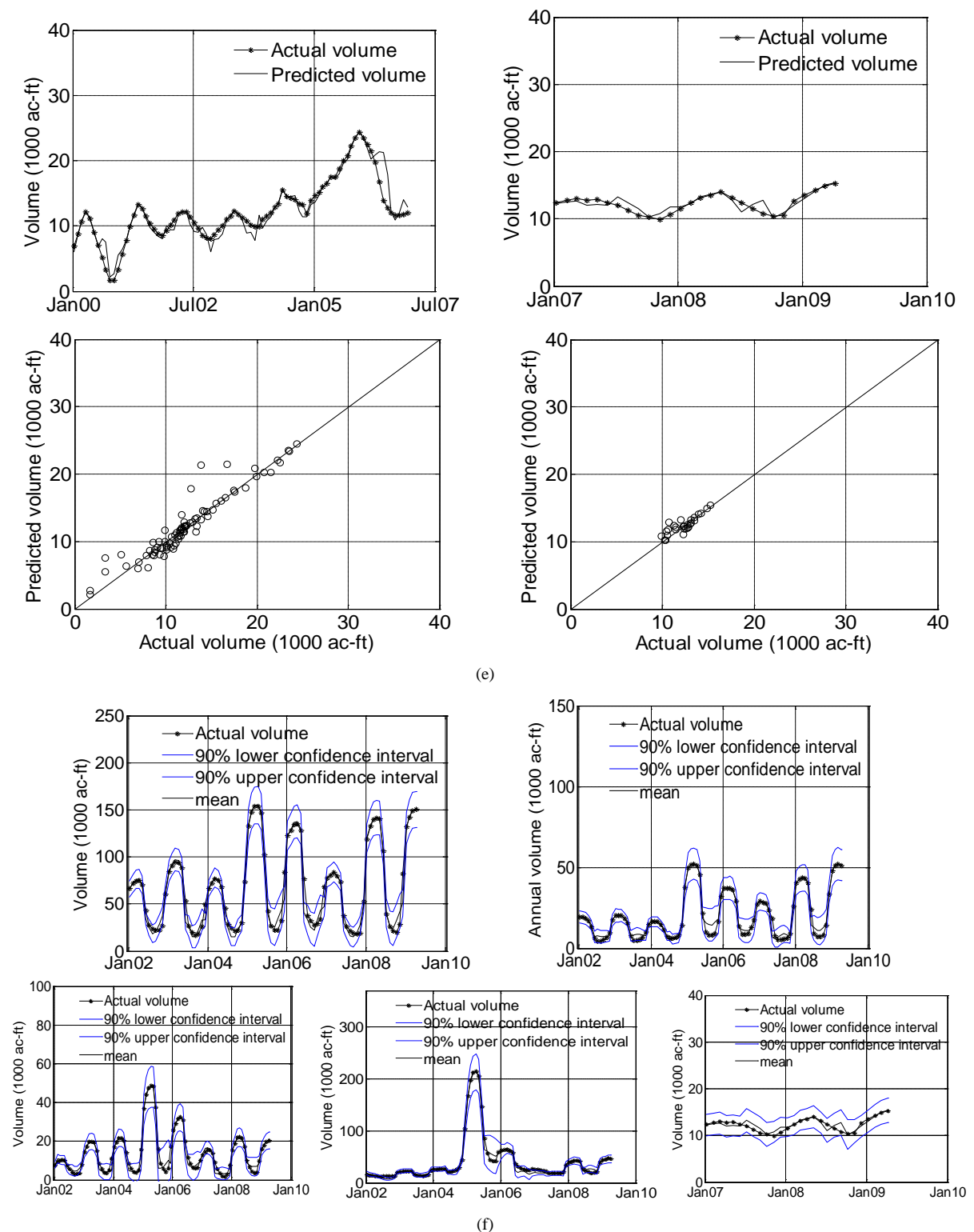| Stream Site	Test RMSE (1000 ac-ft)	Efficiency	Best Combination of SST Locations
Weber River near Oakley	8.307	0.965	NP, CP and TP
Chalk Creek at Coalville	2.653	0.968	NP
Muddy Creek near Emery	2.438	0.951	NP
Sevier River at Hatch	5.042	0.987	NP
Sixth Water Creek near Springville	0.732	0.739	NP
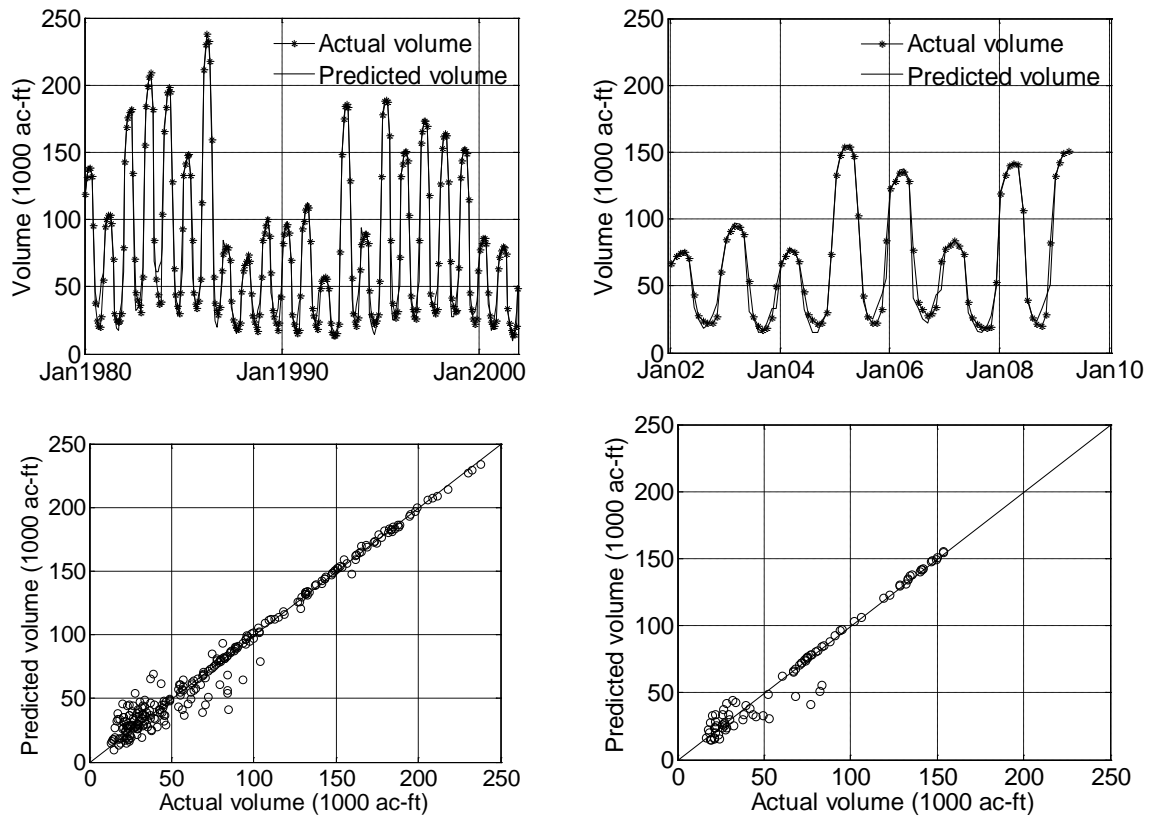
(a)
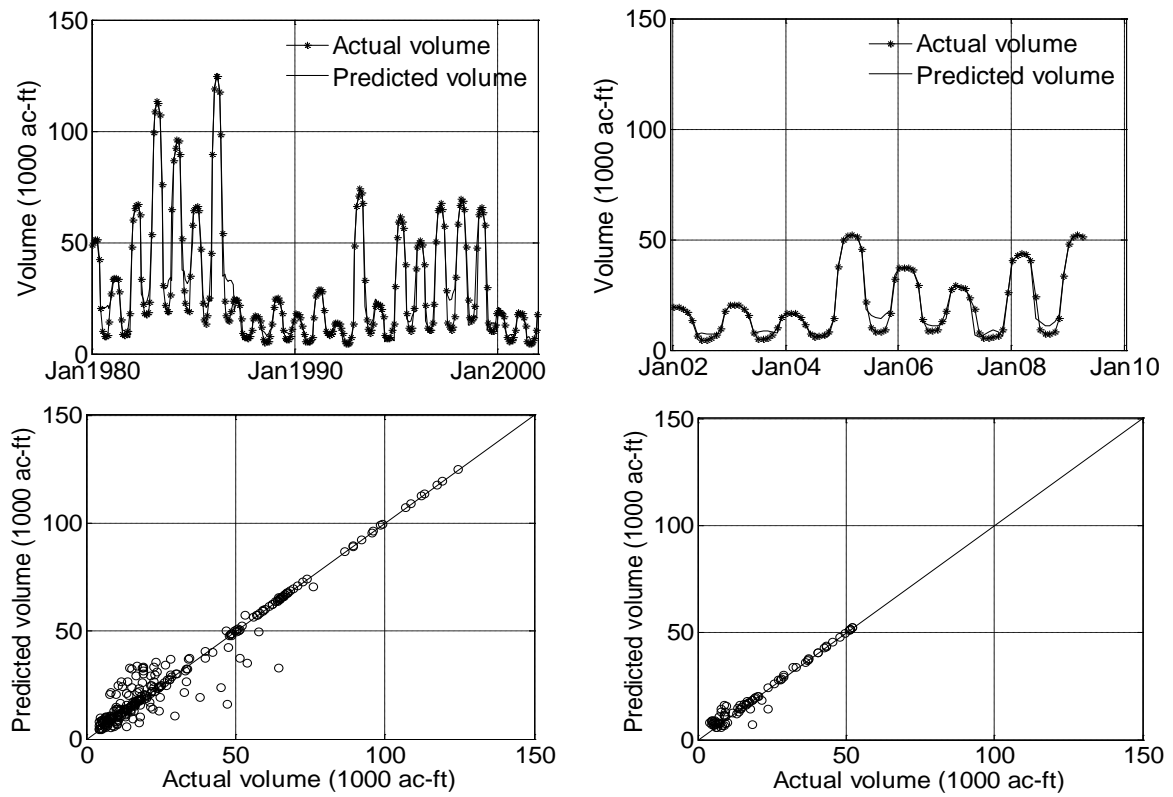
(b)
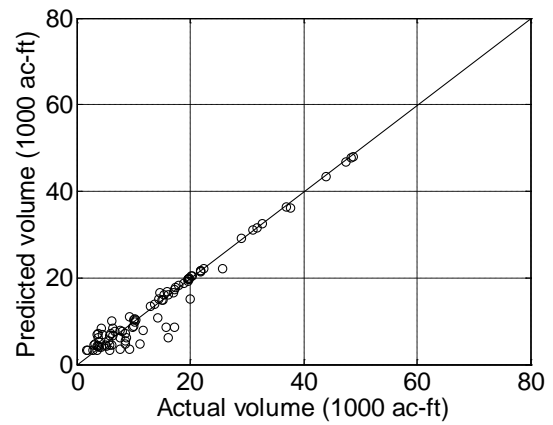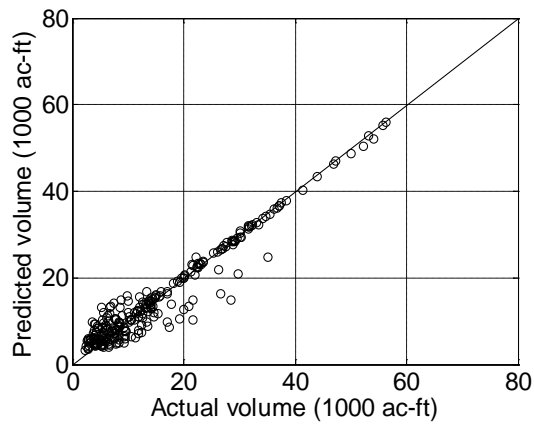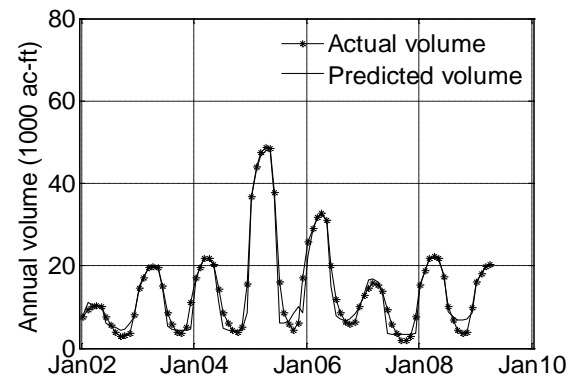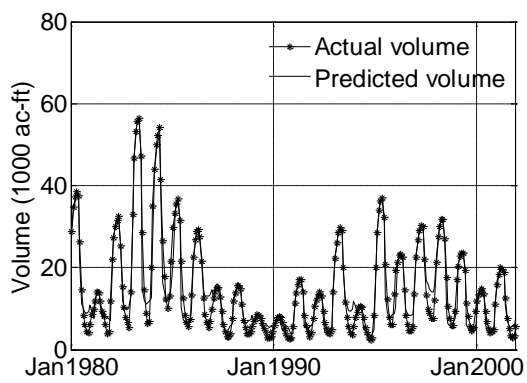
(c)
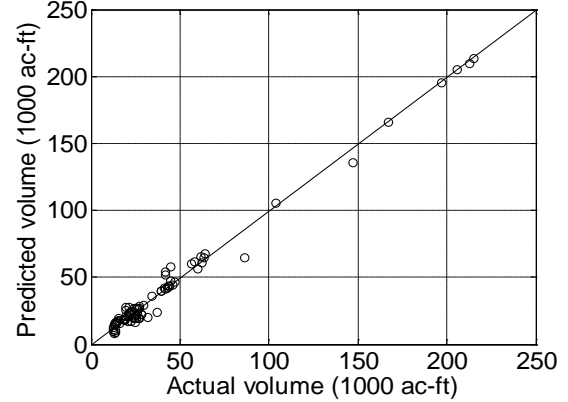
(d)

Figure 4. Prediction for volume of water passing through the stream gages for next six months for (a) Weber River near Oakley; (b) Chalk Creek at Coalville; (c) Muddy Creek near Emery; (d) Sevier River at Hatch; (e) Sixth Water Creek near Springville; and (f) 90 percent confidence interval of prediction for (a) to (e). For each gage, the first figure is time series of actual and predicted flows in training phase, second figure is similar to first figure for test phase, third figure shows the plot of predicted volume versus actual volume in the training phase, and fourth column show similar plot for the test phase.

component is stronger than other components for the monthly data. However, when the variables were cumulative or averaged over the time for the model (Equation (2)), the seasonal component gets eliminated. The remaining components are annual to interannual components, which are low frequency components. North Pacific SST has low frequency component (annual, interannual to interdecadal) so it is obvious to have NP SST influencing more than any other SST locations for most of the streamflow sites in Utah for the volumetric predictions. This includes Chalk Creek at Coalville, Muddy Creek near Emery, Sixth Water Creek near Springville and Sevier River at Hatch. When monthly data were used, the best prediction was obtained from TP SST for Sevier River at Hatch, however, when predictions were made for the volume of water passing through the streamflow site, the variables were averaged or cumulative over the time. The seasonality effect was thus eliminated leaving low frequency components. These components were best represented by the NP region. Therefore, the best predictions were obtained from NP SST. This result is consistent with the result obtained by Asefa *et al.* [2].

For Weber River near Oakley, the combination of CP, NP, and TP developed the best model. However, this result was very close to prediction from the combination of NP and CP SST. The principal moisture source of this area is Pacific Ocean. In addition, this stream gage is outside of the ENSO dominance region. There is no seasonality component, therefore NP and CP SST appeared to be the most important inputs.

3.2. Generalization and Robustness of Model

The bootstrap analysis is a data-based simulation method for statistical inference [25] and gives the estimate of measure of variability of test statistics with the change in training data. We used bootstrap in this study to test the robustness and generalization ability of the model. For each stream gage, bootstrap analysis was performed and the test statistics were computed for each bootstrap sample. The narrow bound of histogram showed that the model was robust (**Figure 5** and **Figure 6**). The dotted red line in the figures shows the 2.5th percentile and 97.5th percentile values of the test statistics. This plot confirms that the model is robust and is accurate to use as a long-term streamflow prediction model.

4. Conclusions

The Relevance Vector Machine successfully transformed the input variables (sea surface temperature, local
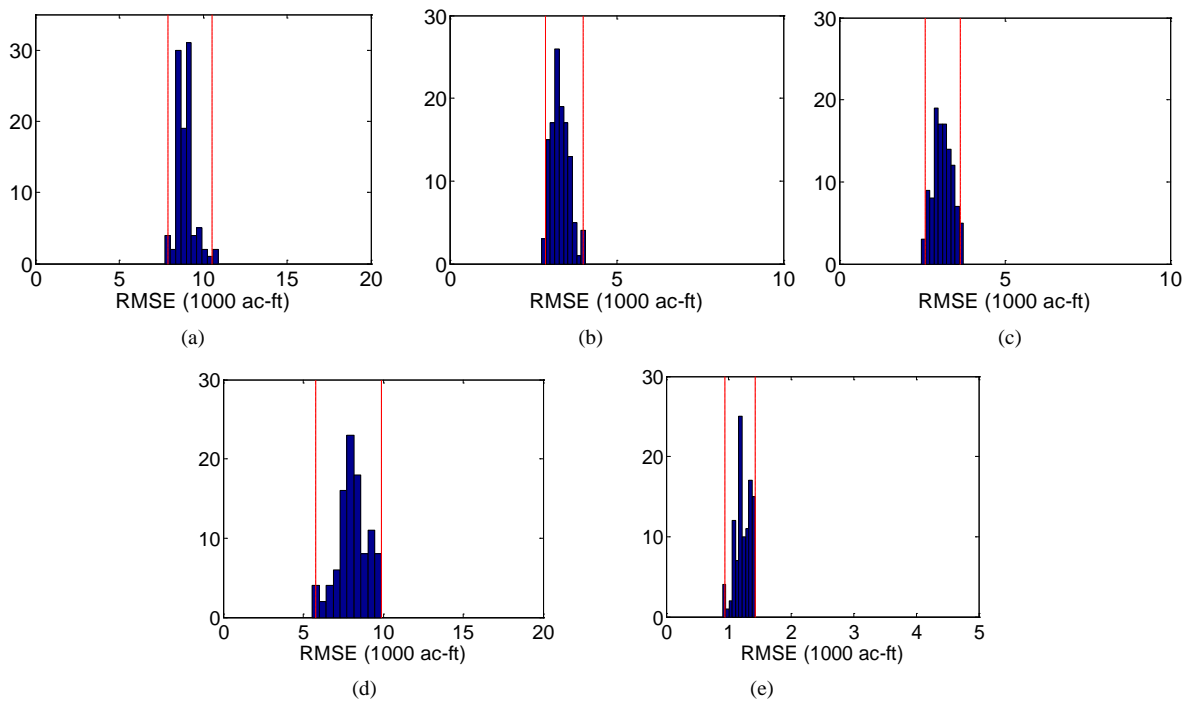

Figure 5. The RMSE from bootstrap analysis for volumetric prediction for (a) Weber River near Oakley; (b) Chalk Creek at Coalville; (c) Muddy Creek near Emery; (d) Sevier River at Hatch; and (e) Sixth Water Creek near Springville.
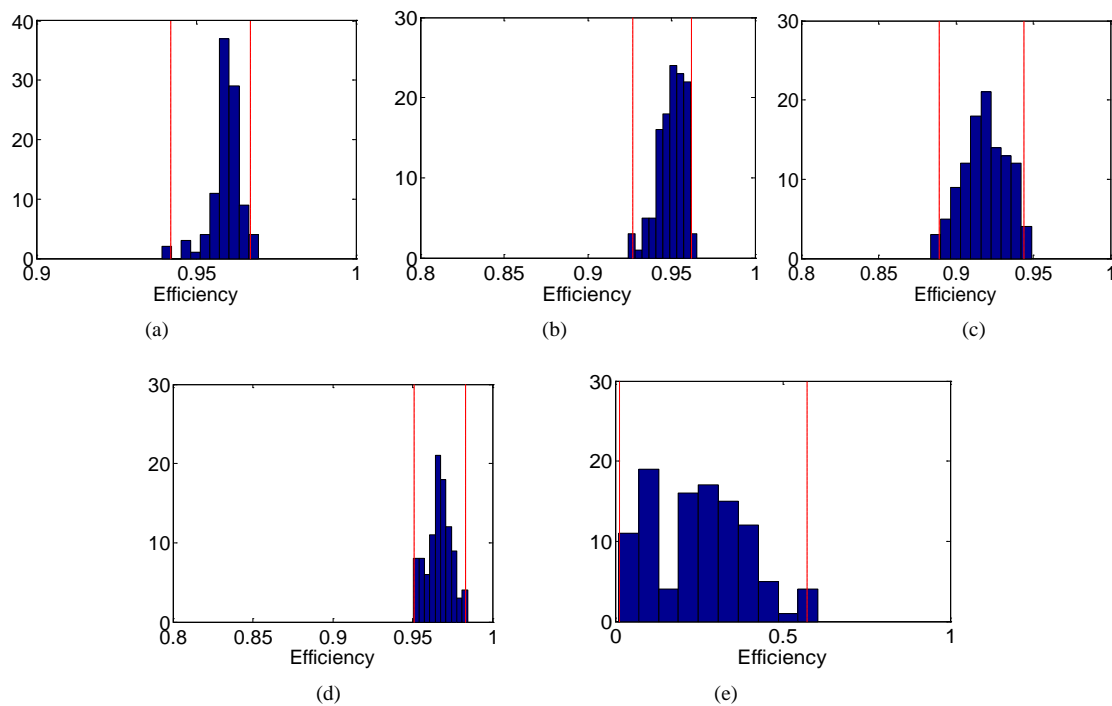
Figure 6. The efficiency from bootstrap analysis for volumetric prediction for (a) Weber River near Oakley; (b) Chalk Creek at Coalville; (c) Muddy Creek near Emery; (d) Sevier River at Hatch; and (e) Sixth Water Creek near Springville.

meteorological conditions, and SWE) into reasonably accurate forecasting of streamflows for next six months. For each gage, the best location of SST was identified. It was found that the SST of Pacific Ocean predicted better than that of Atlantic Ocean because this region represents the majority of Ocean-atmosphere climate influence in the western U.S. [26] [27]. NP SST was the best location of SST for most of the stream gages in Utah for the prediction of volume of water passing the gage for next six months.

The prediction results were highly accurate for unimpaired stream gages while the accuracy was satisfactory for the impaired gage (Sixth Water Creek near Springville). Since the human induced effects were not incorporated in the model for impaired gage, it is obvious to have less efficiency compared to the unimpaired gages. The model has predicted the streamflow perfectly for high flow but low flows were not captured perfectly. The overall predictions were, however, accurate and had good agreement with the observed streamflow values. The uncertainty of the predictions was also captured and presented by the confidence interval. The reliability and robustness of the model were tested from the bootstrap analysis. This analysis confirmed the good predictability and robustness of the model.

This study has demonstrated that with the use of appropriate input, the RVM model can be utilized for the successful forecast of the long-term streamflow. Accurate and reliable long-term streamflow prediction is crucial for the management of water resources in the basin scale. This information could help the water managers and stakeholders for the planning and decision making of the water resources which ultimately reduces the financial risk associated with the water users to future water shortages.

References

[1] Sivakumar, B. (2003) Forecasting Monthly Streamflow Dynamics in the Western United States: A Nonlinear Dynamical Approach. *Environmental Modelling & Software*, **18**, 721-728. http://dx.doi.org/10.1016/S1364-8152(03)00074-4

[2] Asefa, T., Kemblowski, M., McKee, M. and Khalil, A. (2006) Multi-Time Scale Stream Flow Predictions: The Support Vector Machines Approach. *Journal of Hydrology*, **318**, 7-16. http://dx.doi.org/10.1016/j.jhydrol.2005.06.001

[3] Khalil, A.F., McKee, M., Kemblowski, M. and Asefa, T. (2005) Basin Scale Water Management and Forecasting Using Artificial Neural Networks. *JAWRA Journal of the American Water Resources Association*, **41**, 195-208. http://dx.doi.org/10.1111/j.1752-1688.2005.tb03728.x

[4] Kalra, A. and Ahmad, S. (2011) Improving Streamflow Forecast Using Predefined Seas Surface Temperature. *American Geophysical Meeting*, San Francisco, 5-9 December 2011.

[5] Tootle, G.A. and Piechota, T.C. (2006) Relationships between Pacific and Atlantic Ocean Sea Surface Temperatures and U.S. Streamflow Variability. *Water Resources Research*, **42**, W07411. http://dx.doi.org/10.1029/2005WR004184

[6] Shrestha, N.K. (2014) Long Lead-Time Streamflow Forecasting Using Oceanic-Atmospheric Oscillation Indices. *Journal of Water Resources and Protection*, **6**, 635-653. http://dx.doi.org/10.4236/jwarp.2014.66062

[7] Shrestha, N.K. and Shukla, S. (2014) Basal Crop Coefficient for Vine and Erect Crops with Plastic Mulch in a Sub-Tropical Region. *Agricultural Water Management*, **143**, 29-37. http://dx.doi.org/10.1016/j.agwat.2014.05.011

[8] Shukla, S., Shrestha, N.K. and Goswami, D. (2014) Evapotranspiration and Crop Coefficient for Seepage-Irrigated Watermelon with Plastic Mulch in a Sub-Tropical Region. *Transactions of the ASABE*, **57**, 1017-1028.

[9] Shukla, S., Shrestha, N.K., Jaber, F.H., Srivastava, S., Obreza, T.A. and Boman, B.J. (2014) Evapotranspiration and Crop Coefficient for Watermelon Grown under Plastic Mulched Conditions in Sub-Tropical Florida. *Agricultural Water Management*, **132**, 1-9. http://dx.doi.org/10.1016/j.agwat.2013.09.019

[10] Hamlet, A.F. and Lettenmaier, D.P. (1999) Columbia River Streamflow Forecasting Based on ENSO and PDO Climate Signals. *Journal of Water Resources Planning and Management*, **125**, 333-341. http://dx.doi.org/10.1061/(ASCE)0733-9496(1999)125:6(333)

[11] Khalil, A.F., McKee, M., Kemblowski, M., Asefa, T. and Bastidas, L. (2006) Multiobjective Analysis of Chaotic Dynamic Systems with Sparse Learning Machines. *Advances in Water Resources*, **29**, 72-88. http://dx.doi.org/10.1016/j.advwatres.2005.05.011

[12] Shrestha, N.K. and Shukla, S. (2015) Support Vector Machine Based Modeling of Evapotranspiration Using Hydro-Climatic Variables in a Sub-Tropical Environment. *Agricultural and Forest Meteorology*, **200**, 172-184. http://dx.doi.org/10.1016/j.agrformet.2014.09.025

[13] Tipping, M. (2001) Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, **1**, 211-244.

[14] Ticlavilca, A. (2010) Multivariate Bayesian Machine Learning Regression for Operation and Management of Multiple Reservoir, Irrigation Canal, and River Systems. Ph.D. Dissertation, Utah State University, Logan.

[15] Perica, S. and Stayner, M. (2004) Regional Flood Frequency Analysis for Selected Basins in Utah. Utah Department of Transportation Research and Development Division, Salt Lake City.

[16] Vapnik, V.N. (1995) The Nature of Statistical Learning Theory. Springer Verlag, New York. http://dx.doi.org/10.1007/978-1-4757-2440-0

[17] Vapnik, V.N. (1998) The Nature of Statistical Learning Theory. Springer Verlag, New York.

[18] Tipping, M. (2000) The Relevance Vector Machine. *Proceeding of Advances in Neural Information Processing Systems*, the MIT Press, 652-658.

[19] Thayananthan, A. (2005) Template-Based Pose Estimation and Tracking of 3D Hand Motion. PhD Dissertation, University of Cambridge, Cambridge.

[20] Tipping, M.E. and Faul, A.C. (2003) Fast Marginal Likelihood Maximization for Sparse Bayesian Models. *Proceedings of the* 9*th International Workshop on Artificial Intelligence and Statistics*, Key West, 3-6 January 2003.

[21] Dibike, Y., Velickov, S., Solomatine, D. and Abbott, M. (2001) Model Induction with Support Vector Machines: Introduction and Applications. *Journal of Computing in Civil Engineering*, **15**, 208-216. http://dx.doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208)

[22] Scholkopf, B. and Smola, A.J. (2002) Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, Cambridge.

[23] Smith, T.M. and Reynolds, R.W. (2003) Extended Reconstruction of Global Sea Surface Temperatures Based on COADS Data (1854-1997). *Journal of Climate*, **16**, 1495-1510. http://dx.doi.org/10.1175/1520-0442-16.10.1495

[24] Soukup, T.L., Aziz, O.A., Tootle, G.A., Piechota, T.C. and Wulff, S.S. (2009) Long Lead-Time Streamflow Forecasting of the North Platte River Incorporating Oceanic-Atmospheric Climate Variability. *Journal of Hydrology*, **368**, 131-142. http://dx.doi.org/10.1016/j.jhydrol.2008.11.047

[25] Efron, B. and Tibshirani, R.J. (1998) An Introduction of the Bootstrap, Monographs on Statistics and Applied Probability. CRC Press LLC, Boca Raton.

[26] Ting, M. and Wang, H. (1997) Summertime U.S. Precipitation Variability and Its Relation to Pacific Sea Surface Temperature. *Journal of Climate*, **10**, 1853-1873. http://dx.doi.org/10.1175/1520-0442(1997)010<1853:SUSPVA>2.0.CO;2

[27] Wang, H. and Ting, M. (2000) Covariabilities of Winter U.S. Precipitation and Pacific Sea Surface Temperatures. *Journal of Climate*, **13**, 3711-3719. http://dx.doi.org/10.1175/1520-0442(2000)013<3711:COWUSP>2.0.CO;2