REGIONAL FLOOD FREQUENCY ANALYSIS USING SUPPORT VECTOR REGRESSION UNDER HISTORICAL AND FUTURE CLIMATE

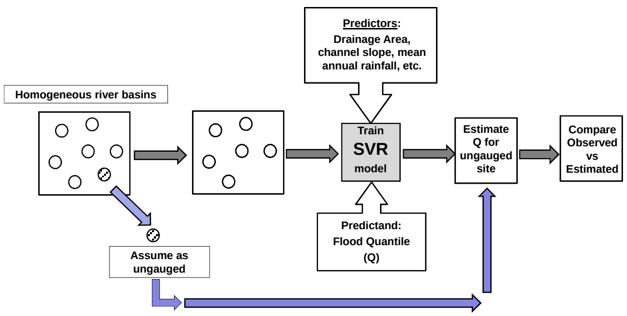
Mesgana Seyoum Gizaw[1] and Thian Yew Gan[1]

[1]Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Canada, T6G 1H9

1. INTRODUCTION

- Regional Flood Frequency Analysis (RFFA) are statistical methods that are widely used to estimate flood quantiles of catchments with limited or no streamflow data (Griffis and Stedinger 2007). Commonly used RFFA methods include rational method, index flood method and quantile regression techniques (QRT). In addition, recent studies using machine learning algorithms such as Artificial Neural Networks (ANNs) have been shown to be better than earlier RFFA methods such as QRT. Another machine learning algorithm that can potentially be used in RFFA is Support Vector Regression (SVR) developed from a kernel-based classification technique called Support Vector Machines (SVM). The objective of this study is to investigate the performance of SVR in RFFA and to compare its performance with ANN based RFFA models for two groups of river basins located in western and eastern Canada, British Columbia (BC) and Ontario (ON), respectively.

2. METHODOLOGY



3. SUPPORT VECTOR REGRESSION (SVR)

- SVR technique aims to find a function $f(x)$ that has an ε deviation from the observed targets for all training data x_i
- SVR uses structural risk minimization approach which is designed to minimize error more efficiently than empirical risk minimization which is used in ANNs
- Because of this SVR has better generalization ability than ANN when the available training data is limited
- In addition, unlike ANNs which are susceptible to finding local optimal solutions, the optimization approach of SVR is guaranteed to lead to global optimal solutions .

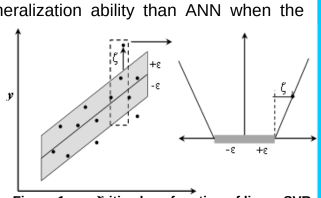


Figure 1 ε sensitive loss function of linear SVR

4. STUDY SITES

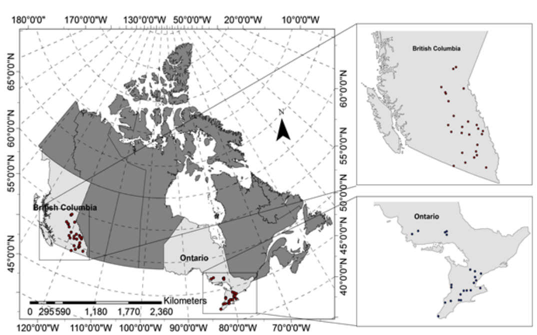


Figure 2 A group of 26 and 23 stations in BC and ON, respectively used in this study

5. RESULTS

- Physiographic, namely drainage area (*A*), time of concentration (*tc*), mainstream slope (*S*), and climatic, namely T-year return period precipitation (*I*) and mean annual precipitation (*M*), were used as predictors to estimate flood quantiles in the SVR-RFFA model. These individual predictor variables were combined to form thirteen sets of predictors, each having at least one physiographic and climatic predictor variables (Table 1).

Table 1 Thirteen sets of predictors used in SVR-RFFA model

	AI	AIM	AItc	AIS	AIMtc	AIMS	AItcS	AIMtcS	A(I+M)tcS	AM	AMtc	AMS	AMtcS
ID	1	2	3	4	5	6	7	8	9	10	11	12	13

- The performance of the SVR-RFFA model was tested using Nash-Sutcliffe coefficient (*Ef*), coefficient of determination (R^2), root mean squared error (*RMSE*), relative *RMSE* (*RMSEr*), mean bias (*BIAS*) and relative *BIAS* (*BIASr*). Among the thirteen sets of predictors in Table 1, predictor #9 (*A(I+M)tcS*) resulted in the best flood quantile estimate for both BC and ON study sites (Figure 3).

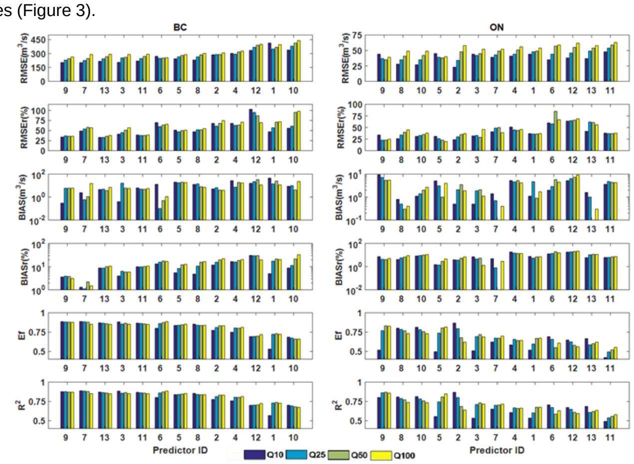


Figure 3 Predictor combinations used in SVR-RFFA model arranged in descending order of their overall performance

- Compared to ANN-based RFFA model (ANN-RFFA), the SVR-RFFA model performed relatively better on the bases of mean goodness-of-fit statistics for the 13 sets of predictors
- However, the coefficient of variation (CV) of the SVR-RFFA model was substantially lower than the CV of the ANN-RFFA model (Figure 4), which could be partly due to the different approaches ANN and SVR take in finding the optimal solution.
- ANNs generally use a backpropagation algorithm based on gradient descent or similar optimization methods. A pitfall of gradient descent approach is getting trapped in local optimal solution before reaching the global optimum.
- Because of this, the flood quantiles estimated by ANNs can vary over a relatively wide range, depending on the initial model parameters and search directions chosen by the optimization method
- On the other hand, SVR is designed to solve convex optimization problems which is generally guaranteed to find the global optimum solution (Chitralekha and Shah, 2010)

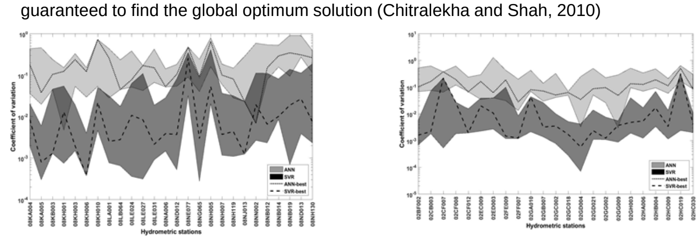


Figure 4 Range of CV of flood quantiles estimated by ANN and SVR based RFFA models for BC and ON study sites

6. FUTURE FLOOD QUANTILE PROJECTIONS

- The SVR-RFFA model that performed best with the overall goodness of fit statistics (#9) was used to estimate changes in flood quantiles in the two study areas.
- Statistically downscales projected climate data was derived from five CMIP5 GCMs from the Pacific climatic impacts consortium (PCIC) data base.
- For the 2041-2100 period, there was an increase of 22.6% (BC) and 6.2% (ON) in the T-year return period precipitation and 9.2% (BC) and 7.8% (ON) in the mean annual precipitation when compared to the 1971-2000 base period.
- This led to an average increase in flood quantiles of about 7.5% in BC, which constitutes river basins with larger drainage area, and 29% in ON which constitutes river basins with small drainage area (Figure 5).
- These results suggest that southeastern BC and southern ON would likely experience extreme flooding in the mid to late 21st century

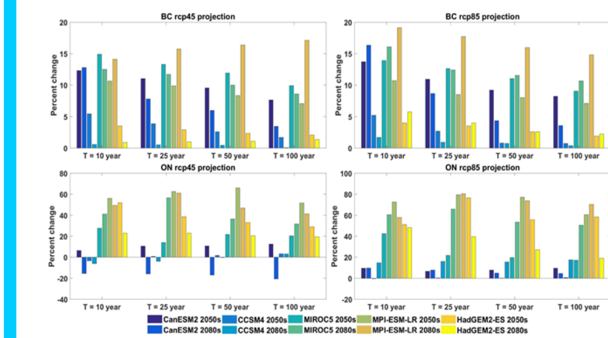


Figure 5 Changes in flood quantiles between the base period and 2041-2100 for BC and ON study sites

7. CONCLUSIONS

(1) The SVR based RFFA model showed satisfactory performance in estimating flood quantiles for a group of river basins located in different climatic regions, southeastern BC and southern ON.

(2) Compared to ANN-RFFA model, the SVR-RFFA model showed comparable and in most cases better performance in estimating the flood quantiles of the two study sites.

(3) The coefficient of variation (CV) of the SVR-RFFA model was substantially lower than that of the ANN-RFFA model which suggests that the flood quantiles estimated for a given station by the SVR-RFFA model vary over a small range when compared to flood quantiles estimated by ANN-RFFA model

(4) Due to the projected increase in precipitation intensity of T year return period and mean annual precipitation, the magnitude of flood quantiles is projected to increase by 7.5% and 29% during the 2041-2100 period in BC and ON study areas respectively. This suggests that climate change could result in extreme flooding in these regions during the mid to late 21st century.

6. REFERENCES

Chitralekha, S.B and Shah, S.L. 2010. 18th Mediterranean Conference on Control & Automation Congress Palace Hotel, Marrakech, Morocco, June 23-25, 2010

Gizaw, M.S., and Gan, T. Y. 2016. J of Hydrology, http://dx.doi.org/10.1016/j.jhydrol.2016.04.041

Griffis, V.W and Stedinger, J.R. 2007. J of Hydrology, doi:10.1016/j.jhydrol.2007.06.023