Reviewer's Report

Journal	:	Hydrol. Earth Syst. Sci.
Manuscript Number	:	doi:10.5194/hessd-7-3521-2010 (hess-2010-113)
Title	:	Seasonal prediction of winter extreme precipitation over Canada by support vector regression
Authors	:	Z. Zeng, W. W. Hsieh, A. Shabbar, and W. R. Burrows

Principal Criteria	**Excellent (1)**	**Good (2)**	**Fair (3)**	**Poor (4)**
Scientific Significance: Does the manuscript represent a substantial contribution to scientific progress within the scope of Hydrology and Earth System Sciences (substantial new concepts, ideas, methods, or data)?			**X**	
Scientific Quality: Are the scientific approach and applied methods valid? Are the results discussed in an appropriate and balanced way (consideration of related work, including appropriate references)?			**X**	
Presentation Quality: Are the scientific results and conclusions presented in a clear, concise, and well-structured way (number and quality of figures/tables, appropriate use of English language)?			**X**	

Access Review, Peer-Review & Interactive Public Discussion (HESSD)

1. Does the paper address relevant scientific questions within the scope of HESS? **Yes**
2. Does the paper present novel concepts, ideas, tools, or data? **No**
3. Are substantial conclusions reached? **No (See Comments)**
4. Are the scientific methods and assumptions valid and clearly outlined? **No**
5. Are the results sufficient to support the interpretations and conclusions? **No**
6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? **No**
7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? **No**

8. Does the title clearly reflect the contents of the paper? **No**
9. Does the abstract provide a concise and complete summary? **No**
10. Is the overall presentation well structured and clear? **No**
11. Is the language fluent and precise? **Yes**
12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Mathematical formulae are skipped at several locations, which if provided would have been helpful in better understanding the work.

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? **Yes (See Comments)**
14. Are the number and quality of references appropriate? **No**
15. Is the amount and quality of supplementary material appropriate? **No**

There is no supplementary material. The data used for the study should be made available from a public domain (at least to reviewers), to allow for verifying reproduction of the results.

General comments

The manuscript makes an attempt to address interesting problem in hydrology. However, description of models/methods considered in the study is incomplete, and the values of parameters estimated for each of the models are not given. Even results presented in the manuscript are inadequate and there are several others concerns that need to be addressed before the manuscript can be published in the Journal. Hence I recommend major revision. The authors should mention in replies how they addressed each review comment/suggestion.

Major Comments

1) Page 3522, Title of the manuscript: "…winter extreme precipitation…" can be modified as "…winter 5-day extreme precipitation…"

2) Abstract:
(a) Page 3522, abstract, line 8: "…the winter extreme precipitation were the predictands."
What is predictand in the study? Is it 'precipitation' or 'PCs extracted from 5-d precipitation anomaly'?

(b) Page 3522, Abstract, lines 13-15: "…two robust SVR models tended to have better forecast skills than the two non-robust models (MLR and BNN),…"
Is the result from the study unexpected, or is it obvious? Is the word 'robust' used in the right context? Can authors mention differences between robust, effective and efficient models? If it is already known that some models are robust and some models are non-robust, what is the need to make comparison between them?

(c) Page 3522, abstract, line 16: "…Among the six regions, the Eastern…"
Which six regions are being referred?

3) Where are the values of parameters estimated for each of the models considered in the study? They have to be given in the manuscript for obvious reasons.

4) Page 3526, lines 6-10: "The reason that the maximum 5-d total precipitation instead of the daily extreme is used here because … 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"

(a) The choice of 5-day precipitation for the study lacks proper reasoning. It is agreed that multi-day episodes might cause floods, but multi-day need not necessarily indicate 5-day period. Did the authors prepare any frequency plots for the data being analyzed to identify low-frequency signal? If so, the plot must be presented in the manuscript to justify considering analysis of only 5-day episode as adequate for the study area.

(b) What are the other standard seasonal extreme precipitation indices chosen by STARDEX project? Why they are not considered in the present study?

5) Page 3522, last line: "…the long-term trend of extreme precipitation events seems not so significant in most areas of Canada (Zhang et al., 2001; Kunkel, 2003)…"
The statement 'seems not so significant' is subjective. Though there is growth in evidence of climate change on extreme precipitation in different parts of the world, it is surprising to note that there is no trend in the Canadian precipitation data, over the period (1950 to 2006) being considered for the analysis. What are the tests and significance level considered by authors to test the long-term trend?

6) Page 3523, lines 13-15: "…Most seasonal forecasts focus on predicting the seasonal mean of the precipitation instead of seasonal statistics of extreme precipitation events."
(a) What do authors mean by "seasonal mean of the precipitation"? Should it be 'seasonal cumulative precipitation'?
(b) Which seasonal statistics of extreme precipitation are being referred? Are they statistics such as mean, standard deviation, skew, kurtosis, or is the reference to accumulated precipitation over several days (3-day, 5-day, 7-day etc.). It would be better if authors can list potential uses of predicting the seasonal statistics of extreme precipitation events.

7) Page 3524, lines 19-20: "…the predictand is the very noisy and non-Gaussian winter extreme precipitation anomaly."
There is inconsistency in definition of predictand (see abstract, line 8: '…the winter extreme precipitation were the predictands'). Is the predictand maximum winter 5-day accumulated precipitation, or precipitation anomaly, or principal components (PCs) extracted from precipitation anomaly?

8) Page 3525, lines 7-8: "…removing the climatological cycle from the monthly mean data and filtering them using a 3-month running mean…"
The write-up lacks clarity. It would be better if equation is given to explain this part of the analysis. Does 'removing the climatological cycle' indicate deseasonalization (or removing periodicity) in

data? Or does it indicate removing long-term cyclicity in the data? What is the reason for choosing 3-month running mean for filtering?

9) Page 3525, line 9: "…After normalizing the anomalies, time-lagged copies of the data were stacked…"
How is the data normalized? and why is it normalized? If deseasonalization is done in the first step, then perhaps normalization is not necessary.

10) Page 3526, lines 3-4: "The climatological seasonal cycle of 5-d precipitation was then removed, and the 3-month maximum was identified as the seasonal extreme precipitation anomaly."
The statement lacks clarity, and it would be better if equation is given to explain this part of the analysis. How was the climatological seasonal cycle removed? Does it indicate deseasonalization?

11) Page 3525, line 14: "…This PCA process, …, is performed on the SST and Z500 normalized anomalies separately, each having 5 leading principal components (PC) retained..."

Principal components must be PCs, rather than PC. What is the logic behind choosing 5 PCs? How much variance did they preserve?

12) "…In view of the diversity of the Canadian climate, we classified the 118 stations into six groups using K-means clustering…"
(a) The description of K-means cluster analysis lacks clarity.
(b) How is it decided that there are six groups? Cluster validity measures have to be used. Authors can refer some of the latest works in hydrology (e.g., Rao and Srinivas, 2006) to know the procedure.
(c) How many feature vectors form input to K-means clustering algorithm, and what are the elements in each feature vector?

13) The data used for the study should be made available from a public domain (at least to reviewers), to allow for verifying reproduction of the results.

14) Figures
(a) Figure 1: What is the scale of the figure?

(b) Figures 2 to 7: The figures show 'average skill score' over all stations for each region. The 'average skill score' is not interesting. Instead, box-plots could have been presented for each region to draw correct inference about the range of skill scores (maximum, 95%, 75%, 50%(median), mean, 25%, 5% and minimum skill score) computed for each region. For each lead time, one box plot can be prepared for each model using estimates of skill scores for all the sites in the region.

15) Subsection 3.3: The description of double cross-validation procedure lacks clarity. It would be better if equations are given to explain this part of the analysis.
(a) Describe in detail the procedure followed to determine optimal number of preditand PCs.
(b) Complete results of the double cross-validation must be presented in the manuscript for obvious reasons.

16) Page 3530, lines 18-20: “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.”

(a) What is the maximum number for predictors that can be considered for the sample size being analyzed? Are there any guidelines fro deciding the number of predictors for given sample size?

(b) What is the reason for applying PCA to predictors comprising of 10PCs (apart from 6 climate indices)? If 10 PCs were excessive, why were so many PCs extracted? Is it logical to apply PCA to PCs? If the resulting PCs are also excessive do authors suggest applying PCA once more?

17) Page 3531, lines 10-12: “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.”
The write-up lacks clarity, and the procedure has to be explained in detail for the sake of readers. What is leakage of low-frequency signals?

18) Page 3531, lines 20-22: “For BNN, the optimal number of hidden neurons to use in a neural network model was found from CV2.”
The procedure must be explained in detail for obvious reasons.

19) Subsection 3.4: What is the reason for choosing only linear measure for estimating correlation?

20) Page 3533, lines 12-14: “Ironically, BNN had…”
Ironically Skill_v score of SVR-R model is least of all models, which according to author is worst performance. What is the reason for the worst performance?

21) Page 3533, lines 14-15: All the models are underpredicting standard deviation. Justify using the terminology ‘overfitted’ for BNN model. How to decide whether a model overfitted data? What can be concluded if Skill_v is nearly zero, less than one, and greater than one?

22) Latest references relevant to the study are not referred, whether it is forecasting or cluster analysis (e.g., Partal and Kisi, 2007).

23) Page 3534, lines 12-13: Why comparison is not presented with results from the canonical correlation analysis prediction model?

24) Section 4: There are 4 skill scores and 4 models. How is it decided whether a region showed highest forecast skill at a particular lead time. Is the judgment purely subjective? The details must be provided for obvious reasons.

25) Page 3536, lines 6-7: “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.”
Page 3536, lines 17-18: “…we found highest skill in the Eastern Prairies, presumably due to the strong nonlinear signal there,…”

Did authors use any standard procedures to detect the presence of nonlinearity in data? Or did they speculate the presence of nonlinearity in data based on the difference in the forecast performance between the linear and nonlinear models?

26) Page 3536, last paragraph: Even if the contribution to forecast skill (from individual predictors) cannot be quantified quantitatively, is it not possible to arrive at subjective conclusion based on analysis? It is necessary to state which predictor contributes more to forecast skill in each region?

Minor comments

1) Page 3522, line 22: "…have important impacts on agriculture…"
Does 'important impacts' indicate 'useful impacts'? What are they?

2) Page 3523, lines 15-16: "…seasonal extreme statistics are potentially noisier than the seasonal mean…"
What does 'potentially noisier' indicate? Are the predictions of extreme statistics noisier, or is it that the estimates of extreme statistics from historical record are noisier?

3) 'Sects.' and 'Sect.'
Do they refer to 'sections' and 'section'?

4) Page 3525, line 24: "Daily 5-d total precipitation records were obtained…"
Define daily 5-d total precipitation record. For any calendar day, does 5-d total precipitation denote cumulative precipitation on the calendar day and its four previous days?

5) Page 3527, line 10: Ensure consistency.
(a) "…west coast"
Is it 'Pacific coast'?
(b) '…east coast…"
Is it 'Atlantic coast'

6) Page 3528, line 22: NN is not defined.

7) Page 3529, line 17: Check English in "An NN model…"

8) Page 3528, line 1: Please check the write-up
"…linear regression problem between ϕ and y,…"

References

Partal, T. and O. Kisi (2007) Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, Journal of Hydrology, 342(1-2), 199-212.

Rao, A. R. and V. V. Srinivas (2006) Regionalization of watersheds by hybrid cluster analysis. Journal of Hydrology, 318 (1-4), 37-56.