A novel stochastic model based on echo state networks for hydrological time series forecasting

Anonymous Author(s)

Abstract

The Stochastic Streamflow Models (SSMS) are time series models for precise prediction of hydrological data useful in hydrologic risk management. Nowadays, deep learning networks get many considerations in time series forecasting. However, despite their theoretical benefits, they fail due to their drawbacks, such as complex architectures, slow convergence and the vanishing gradient problem. In order to alleviate these drawbacks, we propose a new stochastic model applied in problems that involve stochastic behavior and periodic characteristics. The new model has two components, the first one, a type of recurrent neural network embedding the echo-state (ESN) learning mechanism instead of conventional backpropagation. The last component adds the uncertainty associated with stationary processes. This model is called Stochastic Streamflow Model ESN (SSMESN). It was calibrated with time series of monthly discharge data from MOPEX data set. Preliminar results show that the SSMESN can achieve a significant prediction performance, learning speed. This model, can be considered a first attempt that applies the echo state network methodology to stochastic process.

1 Motivation

In probability theory, an stochastic process is defined as a set of models that allow the study of problems with random components. Natural phenomena such as precipitation and streamflow discharge have nonlinear, complex and chaotic characteristics. In order to model the behavior of these phenomena, initially linear approximation was used [2, 11]. Afterwards, were developed methods using self-correcting models such as the PAR(p) model [12, 4]. However, these models are statistical, linear and they cannot capture real chaotic characteristics of hydrometeorological time series, being they sometimes inadequate [13]. Currently, Deep learning (DL) approaches [17], attempt to model this complex non-linear behavior. In fact, DL have been widely used in the recent literature from simplest feedforward Neural Network (ANN) to the most complex Recurrent architecture LSTM. Studies on forecasting performance, show that Recurrent Neural Networks(RNN) [6] are better than their peers ANN, in virtually all tests [3]. However, literature on stochastic models shows preference to use feedforward ANN than RNN, because the last one generates greater complexity in the training process, slow convergence rate, as well as vanishing gradient problems [15]. All it is added to the complexity that uncertainty analysis and stochastic simulation requires [4]. This motivated the development of a stochastic process model using RNN and the echo-state (ESN) learning mechanism instead of conventional backpropagation, the interesting property of ESN is that only the readout layer is trained, whereas the recurrent topology has fixed connection weights [10]. ESN is a training approach attractive as simple and fast compared to other approaches used in RNN, all in order to reduce complexity, and leverage its proven ability to represents the characteristics of time series.

2 Proposal: Stochastic Streamflow model ESN (SSMESN)

The figure 1 details our model, it generates scenarios $Y_{v,t}$ of hydrological synthetic data , in terms of monthly intervals, and can be resume by

$$Y_{v,t} = f\left(R_{v,t} + E_{v,t}\right) \tag{1}$$

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

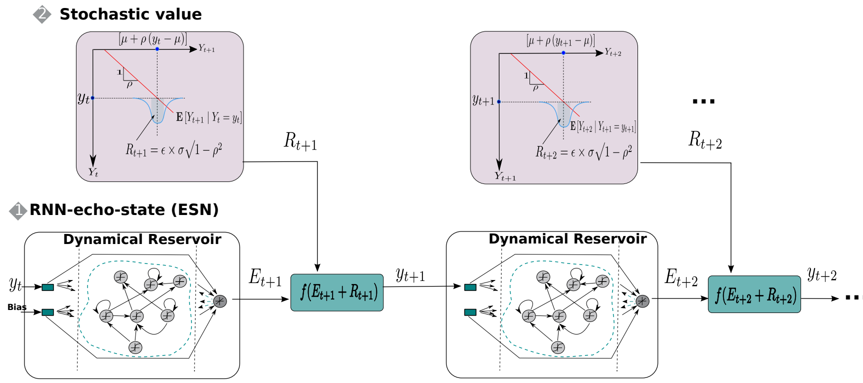

Figure 1: Generation of synthetic scenarios, the new SSMESN model.

		Horizon											
		12 months						24 months					
	Model \ Metric	RMSE	MSE	MAD	NRMSE	MPE	NSE	RMSE	MSE	MAD	NRMSE	MPE	NSE
MOPEX basin 01541500	**TF**	0.68528	0.471	0.59818	0.74254	123.67	0.39673	0.64384	0.41501	0.53775	0.73835	98.612	0.43049
	SSMESN	**0.61774**	**0.3818**	**0.5057**	**0.66936**	**98.937**	**0.51098**	**0.55925**	**0.31282**	**0.44525**	**0.64134**	**76.176**	**0.57072**
	NSP	0.68959	0.47714	0.5934	0.74721	120.21	0.38887	0.69945	0.48996	0.58354	0.80212	106.03	0.32763
	LSTM	0.68863	0.47559	0.59665	0.74617	120.97	0.39085	0.64236	0.41309	0.5323	0.73665	96.057	0.43312
12413000	**TF**	1.4658	2.1549	1.0567	1.4564	156.42	-1.3207	1.3793	1.9048	0.99965	0.72012	160.2	0.45817
	SSMESN	**1.26**	**1.5903**	**0.93179**	**1.2519**	**154.03**	**-0.71271**	**1.306**	**1.7064**	**0.89157**	**0.68185**	**131.67**	**0.51463**
	NSP	1.3454	1.8235	0.98366	1.3368	150.53	-0.9639	1.365	1.8667	0.98498	0.71267	157.32	0.46902
	LSTM	1.2965	1.6952	0.94809	1.2881	142.2	-0.82572	1.3736	1.8903	0.99941	0.71717	162	0.4623
03054500	**TF**	1.1956	1.4316	1.045	0.69638	475.29	0.47014	1.1604	1.3471	0.94575	0.68206	312.91	0.51439
	SSMESN	**1.0688**	**1.1425**	**0.90358**	**0.62255**	**363.55**	**0.57716**	**1.1167**	**1.2471**	**0.87742**	**0.65638**	**225.15**	**0.55042**
	NSP	1.1793	1.3926	1.0242	0.68688	447.6	0.48458	1.1744	1.3796	0.95197	0.69026	316.54	0.50268
	LSTM	1.2063	1.457	1.0463	0.70264	457.65	0.46073	1.1611	1.3484	0.93855	0.68244	299.54	0.51391
01541000	**TF**	0.78917	0.6235	0.71423	0.77839	138.14	0.33828	0.7325	0.53715	0.62133	0.80215	97.899	0.32783
	SSMESN	**0.68499**	**0.46933**	**0.59049**	**0.67563**	**107.96**	**0.50191**	**0.6488**	**0.42102**	**0.52178**	**0.7105**	**76.572**	**0.47315**
	NSP	0.77286	0.59814	0.69586	0.76229	130.29	0.36519	0.73998	0.5481	0.62001	0.81034	96.087	0.31413
	LSTM	0.78557	0.61797	0.70885	0.77484	133.77	0.34415	0.73209	0.53651	0.6172	0.80171	95.867	0.32863

Table 1: Summary of results (RMSE, MSE, MAD, NRMSE, MPE and NSE) of all methods in four series datasets (discharge) MOPEX [7], 1) each column has the results of a specific stochastic model NSP[4][1], TF[16], LSTM [9] and our ***SSMESN***, in a particular metric; 2) each row compares the results of all the methods in a particular data set with a specific horizon value(Monthly Forecasts); 3) Bold rows indicate the best result of each column in a particular metric.

where $\boldsymbol{E_{v,t}}$ is the value produced by the RNN with the echo-state (ESN) learning mechanism, (see equation 2), where W^{out}, is the weight matrix between the internal states $x(t+1)$ added to the input signals y_t and the output neurons, $\delta(.)$ the activation function. $\boldsymbol{R_{v,t}}$ is the stochastic value (see equation 3), where σ_{t+1}, is the standard deviation in month $t+1$, the correlation coefficient between months $t+1$ and t is r_t and $\epsilon = N(0,1)$, a normally distributed random noise with zero mean and standard deviation one.

$$E_{t+1} = \delta \left(W^{out} \left(x(t+1) + y_t \right) + \theta_t \right) \tag{2}$$

$$R_{t+1} = \epsilon \times \sigma_{t+1} \times \sqrt{(1 - r_t^2)} \tag{3}$$

The above equations are concatenated, (2, 3), to obtain the extended equation of our model:

$$Y_{t+1} = f \left(\delta \left(W^{out} \times \left(\vartheta \left[W^{in} y_t + \theta_t + W x(t-1) \right] + y_t \right) + \theta_t \right) + R_t \right) \tag{4}$$

3 Preliminary Results

Experiments were made using the well-known MOPEX data set [7]. Table 1 shows ***SSMESN*** as a promising stochastig model, it outperforms the feedforward models (NSP[4][1], LSTM [9]) and the shallow statistical model (TF[16]) in forecasting performance, learning speed and short-term memory capacity [15]. The main model component "RNN-Echo State Network (ESN)" has a highly inter-linked recurrent topology and random initialize. ESN has two interesting properties; the first is that only the last layer is trained, the second is thanks to its internal memory which is the result of recurrent connections it is not necessary to embed previous input signals (sliding windows). Apparently it may seem surprising that a recurrent neural network with random connections may be effective, but randomized parameters have been successful in several domains [8, 14, 5].

References

[1] Taymoor Awchi and D K Srivastava. Artificial neural network model application in stochastic generation of monthly streamflows for mula project. In *International Conference on Water and Environment, At Bhopal - India*, 10 1999.

[2] George Edward Pelham Box and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, Inc., San Francisco, CA, USA, 1990.

[3] D. Brezak, T. Bacek, D. Majetic, J. Kasac, and B. Novakovic. A comparison of feed-forward and recurrent neural networks in time series forecasting. In *2012 IEEE Conference on Computational Intelligence for Financial Engineering Economics (CIFEr)*, pages 1–6, March 2012.

[4] L. C. D. Campos, M. M. B. R. Vellasco, and J. G. L. Lazo. A stochastic model based on neural networks. In *The 2011 International Joint Conference on Neural Networks*, pages 1482–1488, July 2011.

[5] Mayur Datar, Nicole Immorlica, Piotr Indyk, and Vahab S. Mirrokni. Locality-sensitive hashing scheme based on p-stable distributions. In *Proceedings of the Twentieth Annual Symposium on Computational Geometry*, SCG 04, pages 25–262, New York, NY, USA, 2004. ACM.

[6] N. J. de Vos. Echo state networks as an alternative to traditional artificial neural networks in rainfall runoff modelling. *Hydrology and Earth System Sciences*, 17(1):253–267, 2013.

[7] Q. Duan, J. Schaake, V. Andréassian, S. Franks, G. Goteti, H.V. Gupta, Y.M. Gusev, F. Habets, A. Hall, L. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O.N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener, and E.F. Wood. Model parameter estimation experiment (mopex): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, 320(102):3 – 17, 2006. The model parameter estimation experimentMOPEXMOPEX workshop.

[8] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun. What is the best multi-stage architecture for object recognition? In *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153, Sept 2009.

[9] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. *arXiv preprint arXiv:1703.07015*, 2017.

[10] Mantas Lukoševičius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, August 2009.

[11] Ivette Luna, Rosangela Ballini, and Secundino Soares. Tecnicas identificacao de modelos lineares e no-lineares de series temporais. *Sba: Controle e Automacao Sociedade Brasileira de Automatica*, 17:245 – 256, 09 2006.

[12] D. D. J. . D. J.M. Maceira, M. E. P. ; Penna. Geracao de cenarios sinteticos de energia e vazao para o planejamento da operacao energética. *In: XVI Simposio Brasileiro de Recursos Hidricos,*, 2005.

[13] H. Raman and N. Sunilkumar. Multivariate modelling of water resources time series using artificial neural networks. *Hydrolog. Sci. J*, pages 145–163, 1995.

[14] Andrew Saxe, Pang W. Koh, Zhenghao Chen, Maneesh Bhand, Bipin Suresh, and Andrew Y. Ng. On random weights and unsupervised feature learning. In Lise Getoor and Tobias Scheffer, editors, *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pages 1089–1096, New York, NY, USA, 2011. ACM.

[15] Xiaochuan Sun, Tao Li, Qun Li, Yue Huang, and Yingqi Li. Deep belief echo-state network and its application to time series prediction. *Knowledge-Based Systems*, 130:17 – 29, 2017.

[16] M. B. Thomas, H. A.; Fiering. Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. *Design of Water Resource Systems*, pages 459–493, 1962.

[17] Wenchao Yu, Fuzhen Zhuang, Qing He, and Zhongzhi Shi. Learning deep representations via extreme learning machines. *Neurocomputing*, 149:308 – 315, 2015. Advances in neural networks Advances in Extreme Learning Machines.