Indian Journal of Geo Marine Sciences
Vol. 48 (09), September 2019, pp. 1427-1434

Operational use of machine learning models for sea-level modeling

Thendiyath Roshni[1*], Pijush Samui[2], & Drisya J[3]

[1]Assistant Professor, Department of Civil Engineering, National Institute of Technology Patna, India

[2]Associate Professor, Department of Civil Engineering, National Institute of Technology Patna, India

[3]Research Scholar, Department of Civil Engineering, National Institute of Technology Calicut, India

*[E-mail: roshni@nitp.ac.in]

Received 14 November 2017; revised 03 May 2018

Intense activity offshore warrants a temporal and accurate prediction of sea-level variability. Besides, the sea-level plays an important role in the groundwater level and quality of coastal aquifer. Climate change influences considerable change in all the hydrological parameters and apparently affects sea-level variability. For prediction, highly complex numerical models are usually generated. To address these challenges, the study proposes the use of machine learning (ML) models with the climate change predictands and sea-level predictors. Three ML models are employed in this study, viz., Regression Vector Machine (RVM), Extreme Learning Machine (ELM), and Gaussian Process Regression (GPR). The performance of the developed models is evaluated by visual comparison of predicted and observed datasets. Regression error curve plots, frequency of forecasting errors and Taylor diagram, along with statistical performance metrics were developed. Overall, it is found that the operational use of the selected ML algorithms was quite appealing for modeling studies. Among the three ML models, GPR performed slightly better than ELM and RVM.

[Keywords: RVM; GPR; ELM; Machine learning; Sea-level; Taylor diagram]

Introduction

As the cities are growing at alarming rate with the nearby ports elevated by intense offshore structures, which has spurred the ocean engineers and water resource experts to focus on modeling the sea-level variability. Sea-level change in turn affects the salinity of coastal aquifers, groundwater tables in low-lying coastal areas, as well as hydrological regimes of some coastal rivers[19,22]. Change in the climatic variables can also be an imperative concern for sea-level change[6,9]. Reliable modeling techniques of sea-level integrated with climate change scenarios help to monitor the sea-level variations.

models in sea-level modeling has rapidly increased. It is arguably correct that numerical-based models are data-driven models and are not easily available for different topographical conditions. Under such circumstances, data-driven models that utilize machine learning (ML) algorithms prove to be give promising results. These models utilise, assimilate and 'learn' from the evidence of past climate trends using observational dataset to predict the future development. Many such models are utilized for various applications in hydrology[4]. Such ML models are highly advantageous when there is a lack of data for performing process-based modelling, or when downscaled parameters from climate change models need to be linked to regional/or local scales. Another feature of modern ML models is that they can produce local scale forecasts without performing downscaling operations[23].

An ML technique is an algorithm that estimates (induces) a hitherto unknown mapping (or dependency) between a system's inputs and its outputs from the available data. As such, a dependency is discovered, which can be used to predict (or effectively deduce) the future system's output from the known input values. The ML technique, on the basis of input data,) the target function describing how the real system behaves[10]. Linear, non-linear or mixtures of both algorithms are used for the formulation of casual relationship between inputs and outputs. Linear models like multilinear regression, auto-regressive integrated moving average models perform with less accuracy than non-linear models[1]. Non-linear or hybrid models include artificial neural networks, support vector machines, genetic programming etc. Such soft computing approaches make the research fraternity to explore newer ideas for hydrological parameter modelling. For example, runoff occurred due to typhoon is modelled using artificial neural network models[3], whereas wavelet-support vector regression and wavelet-gene expression programming is used for river flow simulation[17]. Among the ML group of models, the most modern are Regression Vector Machines (RVM), Extreme Learning Machines (ELM) and Guassian Process Regression (GPR). Sea-level modelling has never been performed earlier with these three ML models.

The RVM model, a high-performance machine learning model has grown from support vector machines. The difference is that it allows for probabilistic regression within a Bayesian context. It has excellent generation skills because it combines the strengths of kernel-based methods and Bayesian theory. RVM has been utilized in hydrological applications for ground water quality monitoring network analysis[8], long-term flow predictions for Caglayan Dam in Turkey[12] and evaporative loss estimation[5]. Another high performance ML algorithm, ELM model[7] is relatively faster three-step method with single layer feed forward neural network convergence. It is capable to solve the problems of back propagation algorithms and can be used for non-stationary time series prediction[21]. In hydrology, ELM is used for remote sensing image classification[14], drought prediction[4], monthly evaporation loss estimation[5], and stream flow prediction[24]. The third category of ML algorithm, GPR has more effective use in hydrological applications. It is a generalization of a multivariate Guassian distribution to infinitely many variables. It is used in stream flow modelling[18], and groundwater level forecasting[15]. These studies show that high performance machine learning models can thus be used for any multivariate processes in hydrology. Hence, an attempt was made for sea-level modeling with three ML models, viz., RVM, ELM and GPR. The operational use of these three advanced ML models is explained with the monthly sea-level data at Haldia Port, India. Climate change statistical predictors comprise air temperature, surface pressure, humidity, rainfall, and geopotential height. Performances of the developed models were compared by analyzing different statistical metrics. The study was undertaken as a novel approach in sea-level modelling.

Theoretical background of models

Regression Vector Machine (RVM)

Relevance vector machines is a Bayesian treatment of support vector machine output formulation. RVM[20] seek to predict y for any x according to $y = f(x) + \varepsilon_n$, where the error term has a zero mean and square of the standard deviation as variance from a Gaussian process. In RVM, a fully probabilistic framework is adopted and introduced a priori over the model weights governed by a set of hyper parameters, associated with weights, whose most probable values are iteratively estimated from the data.

The likelihood of the data sets can be defined as

$$p\left(y \middle| w, \sigma^2\right) = (2\Pi\sigma^2)^{-N/2} \exp\left(\frac{-\left\|y - \phi_w\right\|^2}{2\sigma}\right) \quad \ldots (1)$$

where $\varphi(x_k) = [1, K(x_k, x_1) \ldots\ldots K(x_k, x_N)]^T$

Penalty of some prior constraints, W is imposed by adding a complexity penalty to the likelihood or the error function. An explicit zero-mean Gaussian prior probability distribution over the weights, W with diagonal covariance of a, which is a vector of hyper parameters, described as follows:

$$p\left(w \middle| \alpha\right) = (\prod_{K=0}^{N} N\left(W_k \middle| 0\right), \alpha_k^{-1} \quad \ldots (2)$$

Where a is a hyper parameter vector that controls how far from zero each weight is allowed to deviate. Using Bayes rule, the posterior overall unknowns could be computed given the defined non-informative prior distributions as

$$p\left(w, \alpha, \sigma^2 \middle| y\right) = \frac{p\left(y \middle| w, \alpha, \sigma^2\right) p(w, \alpha, \sigma^2)}{\int p\left(y \middle| w, \alpha, \sigma^2\right) p\left(w, \alpha, \sigma^2\right) dw d\, \alpha d\sigma^2} \quad \ldots (3)$$

The posterior distribution over the weights is given by

$$p\left(w \middle| y, \alpha, \sigma^2\right) = (2\Pi)^{-N/2} \left|\sum\nolimits^{-0.5}\right| * \exp\left[-0.5(w-\mu)^T \sum\nolimits^{-1} w - \mu\right] \quad \ldots (4)$$

with mean and covariance,

$$\sum = \left(\sigma^{-2}\varphi^T\varphi + A\right)^{-1} \quad \ldots (5)$$

$$\mu = \sigma^{-2} \sum \phi^T y \quad \ldots (6)$$

Where A is the diagonal of $\alpha_0 \ldots \alpha_l$

The optimization involved in RVM is maximization of most probable hyper parameters, i.e.,

$$p\left(y\middle|\alpha,\sigma^2\right)=\int p\left(y\middle|w,\sigma^2\right)p\left(w\middle|\alpha\right)dw \quad \ldots (7)$$

It is equivalent to

$$p\left(y\middle|\alpha,\sigma^2\right)=\left(\frac{(2\Pi)^{-N/2}}{\left|\sigma^2+\varphi A^{-1}\varphi^T\right|^{0.5}}\right) * \exp\left[-0.5y^T\left(\sigma^2+\varphi A^{-1}\varphi^T\right)^{-1}y\right] \quad \ldots (8)$$

The result of this optimization is that many elements of *α* go to infinity such that *w* will have only a few non-zero weights that will be considered as relevant vectors. In SVM model, these relevant vectors are termed as support vectors. Similar to SVM-based models, selection of the convenient kernel function is also quite important in RVM modeling. The Gaussian radial basis function was preferred in RVM modeling.

Extreme Learning Machine

Extreme learning machine[7] is the state-of-art novel machine learning algorithm for single layer feed forward neural network (SLFN). ELM allows for the analytical determination of the output weights using least-squares by reducing SLFNs to a linear system. These simplifications drastically increase the speed of the learning process, which also grants ELMs a better generalization than traditional SLFNs trained with gradient-based algorithms. From a topological point of view, the ELM resembles an SLFN with no biases on the output layer, for an ELM with one output neuron. For a given pattern *x* of p input variables, the output of an ELM with L hidden nodes and q output nodes is given by:

$$H\beta = T \quad \ldots (9)$$

Where

$$H=\begin{bmatrix} h(x_1) \\ : \\ h(x_N) \end{bmatrix}=\begin{bmatrix} G(a_1,b_1,x_1) & \ldots & G(a_1,b_1,x_1) \\ : & \ldots & : \\ G(a_1,b_1,x_N) & \ldots & G(a_L,b_L,x_N) \end{bmatrix}_{N\times L} \quad \ldots (10)$$

$$\beta=\begin{bmatrix} \beta_1^T \\ : \\ \beta_c^T \end{bmatrix} \quad \ldots (11)$$

$$T=\begin{bmatrix} t_1^T \\ : \\ t_N^T \end{bmatrix} \quad \ldots (12)$$

$$f_L(x)=\sum_{i=1}^{L}\beta_i G(a_i,b_i,x) \quad \ldots (13)$$

Where, $a_i \in R^p, b_i \in R, \beta_i \in R^q$ and $G: R^{p+1} \longrightarrow R$

$$\hat{\beta} = H^{+}T \quad \ldots (14)$$

$\hat{\beta}$ is smallest norm least-square solution, *H* is called the hidden layer output matrix of the ELM, and H^{+} is the Moore-Penrose generalized inverse of *H*. After the output weights have been determined, the ELM can be employed for prediction on a test dataset. The overall complexity *m* of an ELM is given by the sum of input layer (a_i,b_i) and hidden layer β parameters, for which an ELM with one output neuron is equal to *m=(NI+1).NH+NH. NI* is the number of model inputs and *NH* is the number of hidden neurons.

Gaussian Process Regression

Gaussian process regression is a probabilistic model that generalizes multivariate distribution of input data to infinite-dimensional space using a tractable Bayesian framework to infer posterior distributions[16]

$$y(k)=f[x(k)]+\zeta(k) \quad \ldots (15)$$

Where $y(k)$ =objective variable, $x(k)$ are regressors, and $\zeta(k)\sim N(0,\sigma^2)$ is white Guassian noise. A positive semidefinite matrix *K* is generated by the covariance deduced between random variable and input vectors[2,13].

$$K_{i,j}=Cov\left[f(x_i),f(x_j)\right]= \upsilon\exp\left[-\frac{1}{2}\sum_{d=1}^{D}w_d\left(x_i^d-x_j^d\right)^2\right]+\delta i_j\upsilon_o \quad \ldots (16)$$

Where, *D*=dimension of input (*x*); υ_o is the noise of the training data. In the case of a fixed Gaussian noise, the model is trained by applying maximizing the marginal likelihood, which minimizes the negative log-posterior, namely,

$$p\left(\sigma^2, k\right)=\frac{1}{2} y^T\left(K+\sigma^2 I\right)^{-1}$$
$$y+\frac{1}{2}\log\left|K+\sigma^2 I\right|-\log\ p\left(\sigma^2\right)-\log\ p\left(k\right) \quad \ldots (17)$$

Data used

Haldia Port, situated near Kolkata, India, is emerging as the fastest growing port in India. The monthly variations of sea-level data from 1985-2012 is obtained from Permanent Service for Mean Sea Level (PSMSL), Government of India, and is taken as predict and variable for this study. Climate variables or the predictor variables comprise air temperature, surface pressure, humidity, rainfall and geopotential height at different pressure levels. These hydroclimatic variables are selected because of their importance in sea-level dynamics studies[11], available for a period of 1985-2012 from NCEP/ NCAR Reanalysis data.

Model development

For developing RVM, radial basis function was used as kernel function. For best performance of RVM, width of radial basis function was set to 0.45. Radial basis function was also adopted for developing the GPR model. The optimum values of error and width of radial basis function are 0.001 and 0.3, respectively. For developing ELM, radial basis function was taken as activation function. The number of hidden neurons was kept 8 for optimum performance of ELM.

Model performance assessment

Beyond developing ML models, another crucial task was assessing the model performances using statistical and graphical means. Statistical indicators like Nash–Sutcliffe model efficiency coefficient (*NSE*), coefficient of correlation (*R*), coefficient of determination (R^2), root mean square error (*RMSE)*, root square error (*RSR*), *BIAS* and index of agreement, *d* were used in this study. Visual comparison of observed and predicted values was performed through a time series plot. A scatter plot was also presented which shows the prediction interval band for all the models. Regression error characteristics curve (REC) plot and a Taylor diagram for the statistical analysis of the soft computing approaches were also considered for an easy and quick appraisal of the best performed model. Mathematically, the statistical metrics are defined as follows:

$$NSE=1-\frac{\sum_{i=1}^{N}((h_{oi}-h_{ci})^2}{\sum_{i=1}^{N}((h_{oi}-\overline{h_{oi}})^2} \quad \ldots (18)$$

$$R=\frac{\sum\left(h_{oi}\cdot h_{ci}\right)}{\sqrt{\sum h_{oi}{}^2\sum h_{ci}{}^2}} \quad \ldots (19)$$

$$R^2=\frac{\sum_{i=1}^{N}(h_{oi}-\overline{h_{oi}})^2-\sum_{i=0}^{N}(h_{oi}-h_{ci})^2}{\sum_{i=1}^{N}(h_{oi}-\overline{h_{oi}})^2} \quad \ldots (20)$$

$$RMSE=\sqrt{\left(\frac{1}{N}\sum_{i=1}^{N}(h_{oi}-h_{ci})^2\right)} \quad \ldots (21)$$

$$RSR=\frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N}((h_{oi}-\overline{h_{oi}})^2} \quad \ldots (22)$$

$$BIAS=\frac{1}{N}\sum_{i=1}^{N}\left(h_{oi}-h_{ci}\right) \quad \ldots (23)$$

$$\text{Index of agreement, } d=1-\frac{\sum\left(h_{oi}\cdot h_{ci}\right)^2}{\sum\left(\left|h_{ci}-\overline{h_{oi}}\right|\right)+\left\|\left(h_{oi}-\overline{h_{oi}}\right|\right)^2} \quad \ldots (24)$$

Where, h_{oi} = observed groundwater level

h_{ci} = calculated groundwater level,

$\overline{h_{oi}}|$ = mean of observed ground water level and

N = number of observations.

RMSE indicates the discrepancy between the observed and calculated values, prediction is more accurate with lower *RMSE* values. The *BIAS* is a measure of a model estimating the true values. A model with a positive bias will consistently underestimate and with a negative bias will consistently overestimate. The model is said to be unbiased, if *BIAS* = 0. *NSE* evaluates the predictive capability of the developed models.

Results and Discussion

Input Data Selection

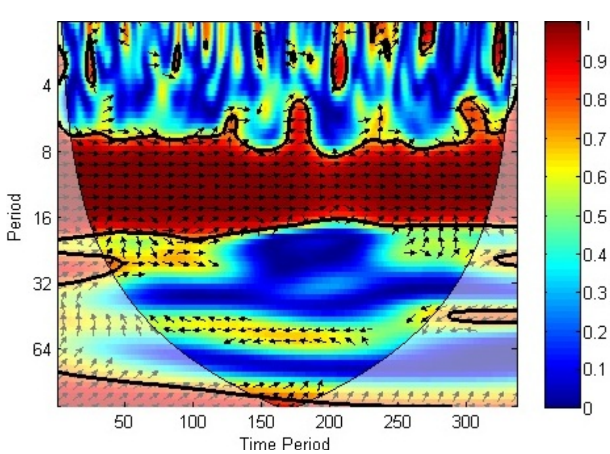
Potential predictors from NCEP/NCAR reanalysis data for the period 1985-2012 were selected based on correlation analysis using R studio. The air temperature at 500 mb pressure level, Geopotential height at 300 mb pressure level, specific humidity at 925 mb pressure level and relative humidity at 1000 mb pressure level were selected as potential predictors with correlation coefficient greater than 0.85. Table 1 shows the correlation of the input variables at different pressure levels. The correlation of the potential predictors with that of the predictand sea-level was also analyzed with wavelet coherence analysis shown in Figure 1. High correlation is visible in the corresponding plots of different potential predictors. Hence, these are used as the input nodes for all ML models.

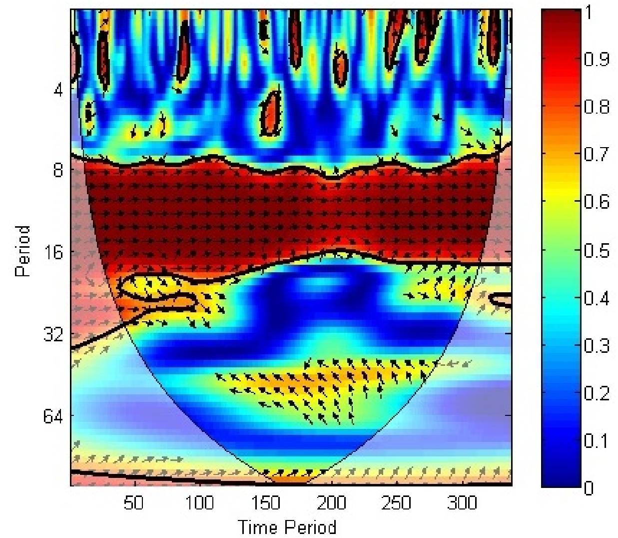
Model development results

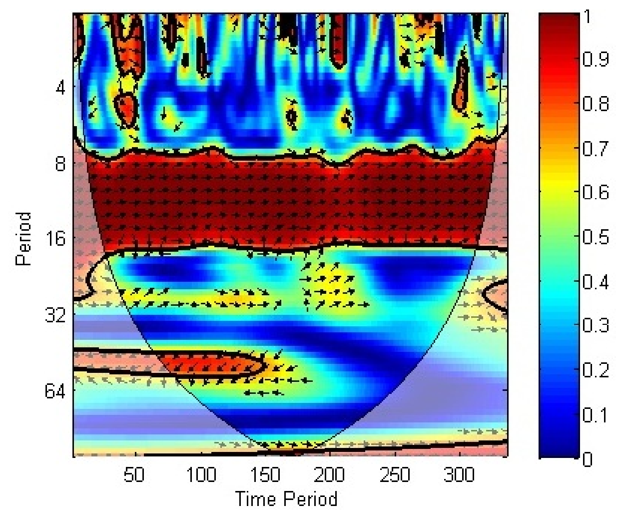
Three models; namely, RVM, ELM and GPR were developed. A visual comparison of the sea-level predicted from each model and observed is presented in Figure 2a. The three models predict the sea-level with more or less same accuracy. A scatter plot with

Table 1 — Potential predictor selection for sea-level modeling

Pressure Level	Air temperature	Geopotential height	Specific humidity	Relative humidity
300mb	0.83	0.87	0.60	0.46
400mb	0.84	0.83	0.57	0.41
500mb	0.85	0.54	0.57	0.48
600mb	0.83	-0.05	0.61	0.76
700mb	0.82	-0.5	0.79	0.78
850mb	0.50	-0.73	0.83	0.81
925mb	0.36	-0.73	0.89	0.89
1000mb	0.35	0.09	0.86	0.78







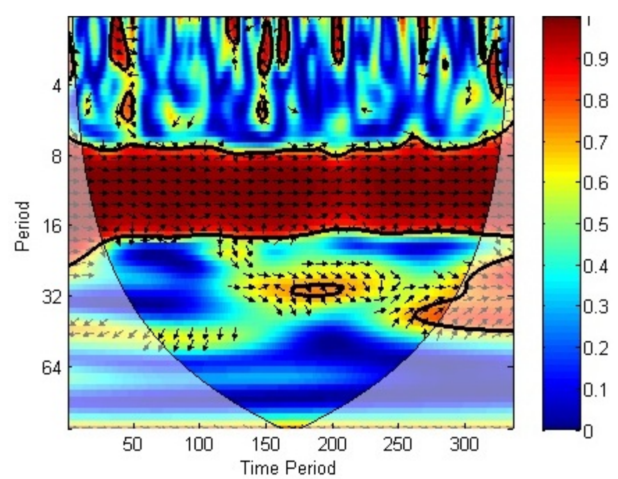


Fig. 1 — Wavelet coherence analysis plots between potential predictors and predictand sea-level (a) Air temperature at 500 mb pressure level, (b) Geopotential height at 300 mb pressure level, (c) Specific humidity at 925 mb pressure level and (d) Relative humidity at 1000 mb pressure level (period in months).

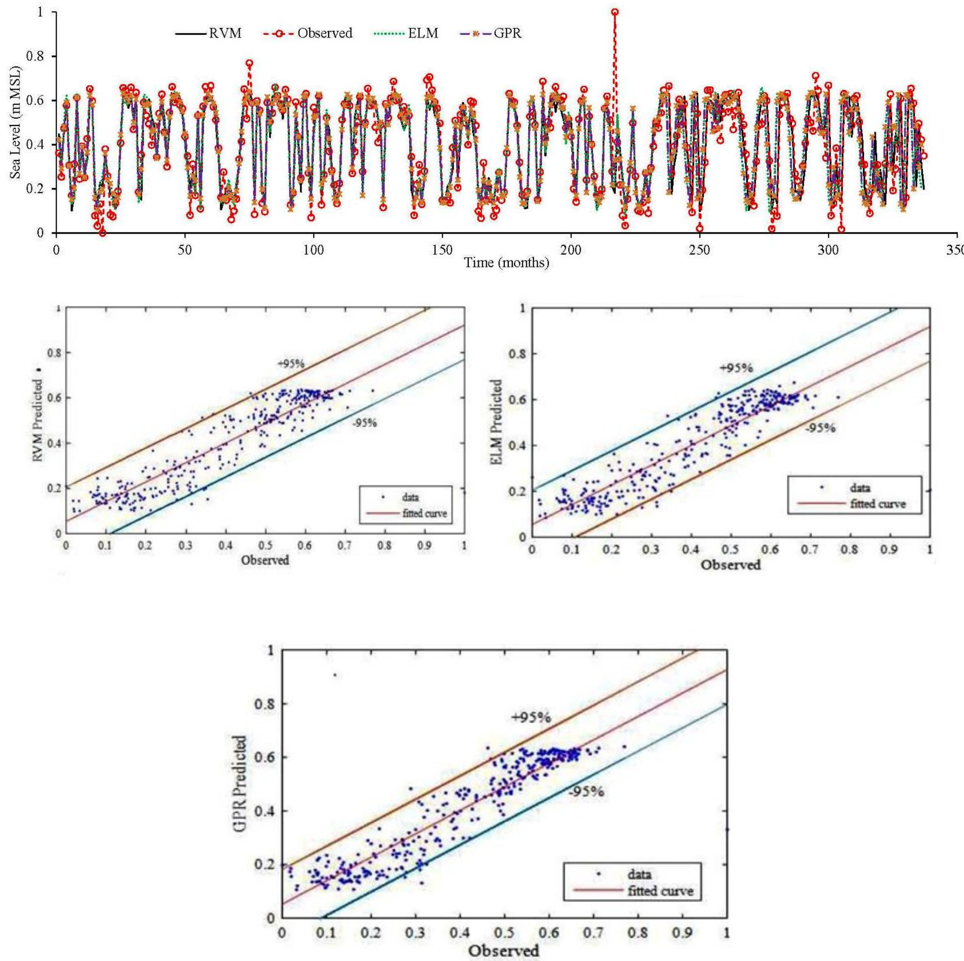


Fig. 2 (a) — Visual comparison of observed and predicted data models (b) Scatter plots between observed and predicted data at 95% prediction interval.

95% prediction interval bands is also shown (Figure 2b). Despite a very few scatter outside the 95% bands, it appears in a good agreement in overall training and testing phases for the three models developed. It also indicates that all the three models are able to predict with reasonable accuracy levels.

The demarcation visible in the training phase and testing phase is detailed in Table 2. During the training phase, when comparing all the statistical parameters, the best fit for sea-level model was found for GPR model. But in testing phase, best fit was found for RVM model. Although in training and testing phases, the percentage difference of best fit model with other two models are less than 0.1% only.

To find the best model, a detailed analysis and fine tuning is required. REC curves estimate the cumulative distribution function of error. In support of the above performance metrics, a REC curve is plotted in Figure 3a. REC curves plotted for three models for both training and testing phases indicates a superior performance for GPR, as the absolute deviation is less for GPR, than ELM and RVM. As an additional measure, a frequency plot of forecasting error (in m) is plotted in Figure 3b. It is evident from close examination of each model that the forecasting error frequency decreases for GPR than other models.

Before coming to conclude about the better performance of GPR, another close examination was done with the Taylor diagram for examining the performance of GPR, ELM and RVM models. Taylor diagram is plotted to compare the efficiency of models by correlation coefficient, RMSE and standard deviation (Figure 4). Again, GPR is found slightly superior to other models in Taylor diagram. However, the improvement percentage of GPR over ELM and RVM models is very less and in harmony with the other models.

Table 2 — Performance metrics for RVM, ELM and GPR models in training phase and testing phase

	Training			Testing		
	RVM	ELM	GPR	RVM	ELM	GPR
MAE	0.0539	0.0512	0.0418	0.0571	0.0625	0.0613
MSE	0.2892	0.2851	0.2584	0.2683	0.2754	0.2751
NSE	0.8258	0.8355	0.8891	0.8488	0.8319	0.8327
R	0.9088	0.9141	0.9435	0.9284	0.9181	0.917
R^2	0.8259	0.8355	0.8902	0.8619	0.8428	0.841
RMSE (m)	0.0836	0.0813	0.0667	0.072	0.0759	0.0757
RSR	0.4174	0.4055	0.3331	0.3888	0.41	0.409
BIAS	0.2092	0.003	-0.0833	-1.2809	-0.5245	-1.622
D	0.9511	0.9538	0.9689	0.9626	0.9573	0.9567

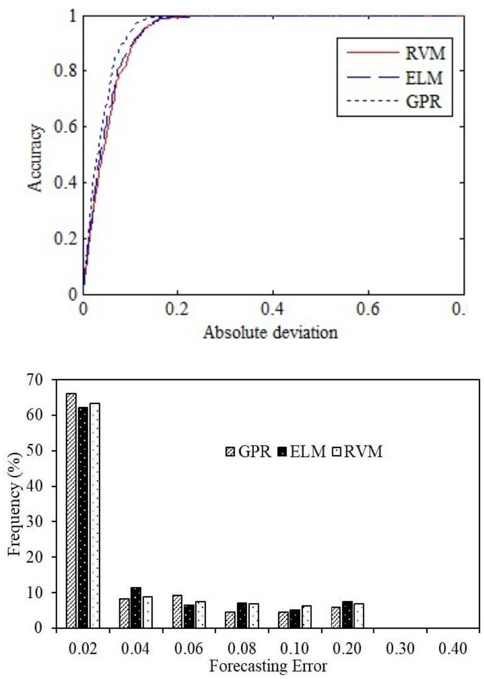


Fig. 3 — (a) REC plots (b) Frequency % for forecasting error for GPR, ELM and RVM models.

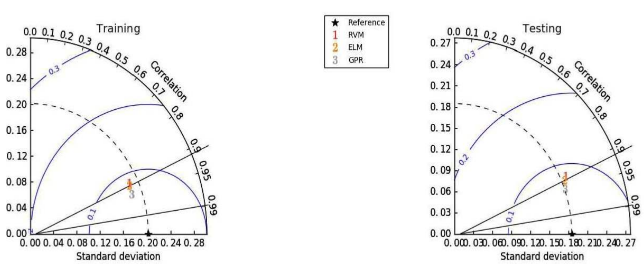


Fig. 4 — Taylor diagram for (a) Training and (b) Testing phase for RVM, ELM and GPR models.

Conclusion

The primary objective of the paper was the operational use of high performance ML models for sea-level modelling. Regression vector machine, extreme learning models and guassian process regression were the three models used for prediction of monthly sea-level data at Haldia Port, India. The potential predictors for statistical downscaling of sea-level variability were air temperature, geopotential height, specific humidity and relative humidity. The potentiality of the predictors with predictand sea-level was expressed as coefficients of correlation and the correlated variables were displayed with wavelet coherence analysis.

Sea-level simulation by three models shows the potential capability of machine learning algorithms. By comparing the performance measure metrics and visual comparison between observed and predicted shows very little demarcation of the three models. Variations among the models are less than 0.1%. It indicates the successful application of ML algorithms for sea-level estimations. A detailed analysis by REC plots, frequency of forecasting errors plots and a Taylor diagram also aids visually to find the better performance of the models. It is evidently proved by the plots that the GPR shows a slightly better approach among others and appears to be a promising soft computing approach for prediction. The viability of ML models for sea-level predictions is quite appealing for incorporating accurate forecasts in suppressing climate change models. However, ML models are highly data-demanding and perform better with longer datasets.

References

1 Adamowski, J., Fung Chan, H., Prasher, S.O., Ozga-Zielinski, B. and Sliusarieva, A. Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resources Research*, 28 (2012) 1-14

2 Ažman, K. and Kocijan, J. Application of Gaussian processes for black-box modelling of biosystems. *ISA Trans,* 46 (2007) 443-457.

3 Chen, S.M., Wang, Y.M. and Tsou, I. Using artificial neural network approach for modelling rainfall--runoff due to typhoon. *Journal of Earth System Science,* 122(2) (2013) 399-405.

4 Deo, R.C. and Şahin, M. Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmospheric Research,* 153 (2015) 512-525.

5 Deo, R.C, Samui, P. and Kim, D. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. *Stochastic Environmental Research Risk Assessment,* 30 (2016) 1769-1784

6 Goharnejad, H. and Eghbali, A.H. Forecasting the Sea Level Change in Strait of Hormuz. World Acad Sci Eng Technol *Int J Environ Chem Ecol Geol Geophys Eng,* 9 (2015) 1319-1322

7 Huang, G-B., Zhu, Q-Y. and Siew, C-K. Extreme learning machine: theory and applications. *Neurocomputing,* 70 (2006) 489-501

8 Khalil, A., Mac, M. and Jagath, K. Bayesian Method for Groundwater Quality Monitoring Network Analysis. *Journal of Water Resources Planning and Management,* 137 (2011) 51-61

9 Masciopinto, C. and Liso, I.S. Assessment of the impact of sea-level rise due to climate change on coastal groundwater discharge. *Science of the Total Environment,* 569 (2016) 672-680

10 Mitchell, T.M. Machine Learning. *McGraw-Hill International Editions Computer Science Series*, (1997)

11 Naren, A. and Maity, R. Hydroclimatic modelling of local sea level rise and its projection in future. *Theor Appl Climatol*. (2016)

12 Okkan, U., Serbes, Z.A. and Samui, P. Relevance vector machines approach for long-term flow prediction. *Neural Computational Applications*, 25 (2014) 1393-1405

13 Pal, M. and Deswal, S. Modelling pile capacity using Gaussian process regression. *Computational Geotech,* 37 (2010) 942-947.

14 Pal, M., Maxwell, A.E. and Warner, T.A. Kernel-based extreme learning machine for remote-sensing image classification. *Remote Sens Lett* 4 (2013)853-862.

15 Raghavendra, N.S. and Deka, P.C. Multistep ahead groundwater level time-series forecasting using gaussian process regression and ANFIS. *In: Advanced Computing and Systems for Security. Springer*, (2016) 289-302

16 Rasmussen, C.E. and Williams, C.K.I. Gaussian processes for machine learning. *MIT press Cambridge* (2006)

17 Solgi, A., Pourhaghi, A., Bahmani, R. and Zarei H. Pre-processing data using wavelet transform and PCA based on support vector regression and gene expression programming for river flow simulation. *Journal of Earth System Sciences,* 126 (2017) 65.

18 Sun, A.Y., Wang, D. and Xu, X. Monthly streamflow forecasting using Gaussian process regression. Journal of Hydrology 511 (2014) 72-81.

19 Thain, R.H., Priestley, A.D. and Davidson, M.A. The formation of a tidal intrusion front at the mouth of a macrotidal, partially mixed estuary: a field study of the Dart estuary, UK. *Estuar Coast Shelf Sci* 61(2004) 161-172.

20 Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1(2001) 211-244.

21 Wang, X. and Han, M. Online sequential extreme learning machine with kernels for nonstationary time series prediction. *Neurocomputing* 145 (2014) 90-97.

22 Werner, A.D. and Simmons, C.T. Impact of sealevel rise on sea water intrusion in coastal aquifers. *Groundwater* 47 (2009) 197-204.

23 Wu, C.L and Chau, K-W Data-driven models for monthly streamflow time series prediction. *Eng Appl Artif Intell,* 23(2010)1350-1367.

24 Yaseen, Z.M., Jaafar, O., Deo, R.C., et al Stream-flow forecasting using extreme learning machines: A case study in a semi-arid region in Iraq. *Journal of Hydrology* 542 (2016) 603-614