

Efficient SST prediction in the Red Sea using hybrid deep learning-based approach

M. M. HITTAWE S. Langodan O. Beya I. Hoteit and O. Knio
KAUST university KAUST university University of Angers KAUST university KAUST university
Thuwal, Saudi Arabia Thuwal, Saudi Arabia Angers, France Thuwal, Saudi Arabia Thuwal, Saudi Arabia
Environmental Protection
Saudi Aramco, Dhahran

Abstract—Prediction of Surface Sea Temperature (SST) is of great importance in seasonal forecasts in the region and beyond, mainly due to its significant role in global atmospheric circulation. On the other hand, SST predicting from given multivariate sequences using historical ocean variables is vital to investigate how SST physical phenomena generated. This paper seeks to significantly improve the prediction of Surface Sea Temperature (SST) by combining two machine learning methodologies: short-term memory networks (LSTM) added to Gaussian Process Regression (GPR). We developed a data-driven approach based on deep learning and GPR modeling to improve the prediction of SST levels in the red sea based on meteorological variables, including the hourly wind speed (WS), air temperature at 2m (T2), and relative humidity (RH) variables. The coupled GPR-LSTM model may potentially carry both flexibility and feature extraction capacity, which could describe temporal dependencies in SST time-series and improve the prediction accuracy of SST. It is necessary to indicate that these types of hybrid-based approach architectures have not used before in SST time-series prediction, so it is a new approach to deal with these types of problems. The results demonstrate a significant improvement when this hybrid model is compared to LSTM and the most frequently used ensemble learning models.

I. INTRODUCTION

Sea surface temperature (SST) is characterized as one of the most important physical properties of the world's oceans, playing major roles in different met-ocean aspects ranging from weather prediction to understanding marine ecosystems. Being a direct measure for the global warming, SST is widely used as climate indicator for studying the climate system and alerting persistent changes in weather patterns. SST variations can lead to the changes in the marine habitats - altering migration and breeding patterns of marine lives, threatening sensitive ocean life such as corals, and occurrences of harmful algal blooms, etc.

The prediction of SST can be achieved by different techniques, broadly divided into physics-based numerical methods and data-driven methods. The former one relies on mathematical model and are proven for estimating SST over a large domain by accounting all the physical processes modulating SST, but at the cost of large computational infrastructures. The accuracy of the prediction from data-driven techniques are largely depending on the amount of available data and the level of complexity involved in the methods.

AI technologies have evolved over the last decay. Machine learning and, especially, deep learning is a vital aspect of AI that can predict, forecast and extract deep features from big and complex datasets [1], [2], [3]. These aspects are considered as a promising advances used by various environmental researchers. The advancement in the deep learning methods allows the development of new algorithms to predict the SST, shifting from classical approaches to new models based on artificial intelligence ([4], [5], [6]). Recurrent Neural Network (RNN) becomes a popular choice for the prediction SST because of their flexibility in fitting to random data and their relatively simple development ([7]). Long Short-Term Memory (LSTM), an extension of the RNN is also introduced by making use of memory of past timeseries to improve their results ([8], [9], [10]).

The RNNs and LSTM are currently being used in SST prediction studies around the world. For example, previous studies over different regions such as eastern Pacific [11], Indian Ocean [12], [13], [14], the western Mediterranean [15], the Australian margin [16], and the South China Sea ([17]), etc. successfully demonstrated the capability of ANN in predicting the SST values over respective regions at different temporal scales. Previous studies suggest that the RNN has an attractive skill in predicting SST; however, the accuracy of the prediction becomes more demanding for the regions where the uncertainty of the results is sensitive to the marine habitats (e.g., coral reefs). In such cases, coupling the LSTM model with an uncertainty estimation model such as Gaussian Process Regression (GPR) is greatly desired.

LSTM networks, which are essentially deep in time, are usually used to model long-term temporal dependencies. On the other hand, Gaussian Process Regression (GPR) models are considered an attractive way of performing non-parametric Bayesian modeling in a supervised learning problem. GPR models showed suitable performance in capturing process nonlinearity in multivariate time-series data. Crucially, GPR models are characterized by their flexibility and efficiency in uncovering implicit correlations in multivariate data, making them especially relevant in handling nonlinear prediction problems. Also, these models are assumption-free; that is, no assumptions on the distribution underlying the data

are needed [18]. Accordingly, amalgamating the capacity of LSTM to efficiently learning temporal dependencies in the data with capability and superior nonlinear approximation of GPR will result in a more sophisticated and flexible model. Essentially, the proposed GPR-LSTM enables more flexibility to describe the temporal correlations involved and merged desirable properties of both GPR and LSTM to handle dependencies in time-series. Note that the only information needed to implement the GPR-LSTM model is the availability of data.

This paper introduces a hybrid GPR-LSTM model for SST prediction in the Red Sea based on meteorological parameters. The main contributions in this paper are listed in the following key points.

- We developed a data-driven approach based on deep learning and GPR modeling to improve the prediction of SST levels in the red sea based on meteorological variables, including the hourly wind speed (WS), air temperature at 2m (T2), and relative humidity (RH) variables. The coupled GPR-LSTM model may potentially carry both flexibility and feature extraction capacity, which could describe temporal dependencies in SST time-series and improve the prediction accuracy of SST. The central characteristic of this coupled model is its flexibility and assumption-free design (i.e., no assumptions are made on the underlying distribution data).
- The coupled GPR-LSTM model is used to predict the daily SST values in the Red Sea. Specifically, we used SST data provided by the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, consisting of a combination of satellite data and in-situ observations. Here, we smooth the SST data via an exponentially weighted moving average (EWMA) to eliminate outliers and improve data quality.
- Furthermore, the performance of the GPR-LSTM model is verified via comparisons with the standalone models, namely SVR, GPR, RNN, and LSTM. The assessment is carried out based on three commonly used statistical indicators: coefficient of determination, R^2 , root mean square error (RMSPE), mean absolute percentage error (MAPE). Results testify the accuracy of the GPR-LSTM model for SST prediction.

The rest of the paper is organized as follows. Section 2 details the proposed GPR-LSTM for SST prediction. Section 3 presents an analysis of the experimental results. Section 4 concludes the work and discusses our future extensions and plans.

II. PROPOSED HYBRID GPR-LSTM REGRESSION MODEL

In this section, we present a hybrid deep-ensemble learning architecture called GPR-LSTM, this architecture is introduced to merges the benefits of the LSTM, which model and learn temporal dependencies embedded in data sequences, and GPR which can improve and the prediction of the series.

A. LSTM Regression

Recurrent Neural Network (RNN) is commonly used for languages models due to its capacity to memorize long terms dependencies [19], [20], [21]. Nevertheless, the gradients of RNN could vanish through unfolding RNN into a deep forward neural network in case of extending the time lags. Long short-Term Memory network (LSTM) is a particular case of RNN that solves the gradient vanishing problem, as it provides the forgetting units. This new structure can determine when to forget certain information by choosing the optimal time lags [22]. Additionally, LSTM is an excellent fit to be used for predicting long time-series based sequential data [23], [24].

LSTM is one of the commonly used RNN models for large time-series based data. It has mainly four interacting layers: cell state, forget gate, input gate and output gate. The new input is given to the input gate to be processed as new coming data. The output of the LSTM neural network in the last iteration will be allocated to the memory cell input gate. The forget gate is utilised to forget the output results by selecting the optimal time lags of the input sequence. Finally, all results will be processed and to generate the output of LSTM in the output gate. The outputs of the recurrent layer are the hidden states at each timestamp.

Accordingly, we used LSTM to perform accurate predictions according to the history of the time-series data and the memory units in the LSTM network.

B. Nonlinear Gaussian Process Regression (GPR)

GPR is a nonparametric Bayesian method that can be employed to solve nonlinear regression problems [25]. In GPR, the response y_i of a function g at the input is represented as

$$y_i = G(z_i) + \varepsilon_i \quad (1)$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. It should be noticed that this is related to the hypothesis in ordinary linear regression, where the output (dependent variable z) consists of independent inputs z contaminated by additive noise ε . GPR is a Bayesian regression designated by its adaptability and capability to give uncertainty estimates by assuming a Gaussian Process (GP) prior to the regression functions, $g(z)$.

Suppose $\mathcal{D} = \{(z_i, y_i)\}_{i=1}^n$, is provided input-output data and $g(\cdot)$ to be approximated is assumed to follow a Gaussian process. The prior Gaussian Process (GP) distribution is presented as:

$$g(z) \sim GP(m(z), k(z, z')) \quad (2)$$

Gaussian Process (GP) is defined by its mean function, $m(z)$, and covariance or kernel function: $k(z, z')$.

$$m(z) = \mathbb{E}[g(z)] \quad (3)$$

$$k(z, z') = \mathbb{E}[(g(z) - m(z))(g(z') - m(z'))] \quad (4)$$

The regression process can be managed and adjusted within the kernel and mean function, for modelling the evolution of the physical process. Mainly, the kernel function is used to

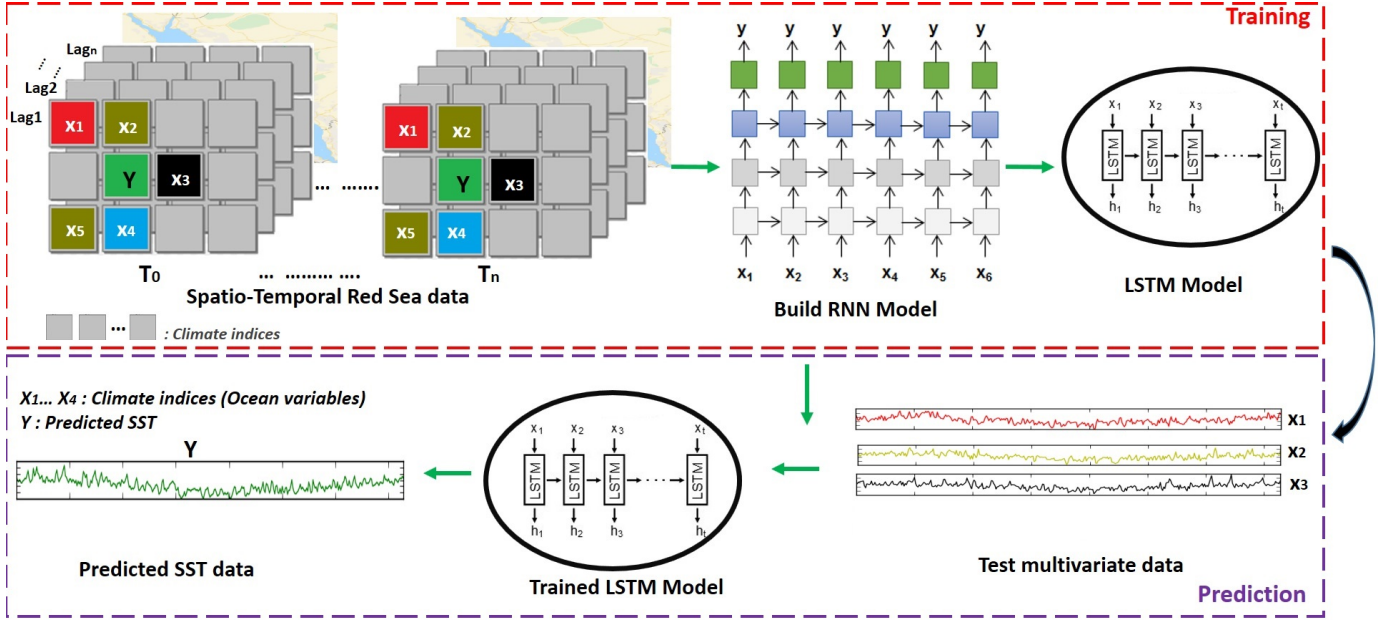


Fig. 1. Multivariate LSTM based machine learning framework for data prediction

describe the data structure and capture the correlation between data points in a data set.

Suppose that the observed y_i values, $[y_1, y_2, \dots, y_n]^T$ are measurable values of the tainted function $g(\cdot)$ with noises. Hence, y_i 's follow a joint Gaussian distribution :

$$\mathbf{Y} = [y_1, y_2, \dots, y_n]^T \sim \mathcal{N}(\mathbf{m}(z), \mathbf{K} + \sigma^2 \mathbf{I}) \quad (5)$$

where $\mathbf{m}(z) = [m(z_1), m(z_2), \dots, m(z_n)]^T$ denotes to the vector of mean $m(\cdot)$ which often are zero, \mathbf{I} refers the identity matrix, and \mathbf{K} is the $n \times n$ covariance matrix with $(i, j)^{th}$ element $\mathbf{K}_{ij} = k(z_i, z_j)$. For a GPR model, $k(z_i, z_j)$ is usually called a kernel function, which represents the generalization abilities of the model. Ordinarily, there is no intuitive alternative to choose a proper kernel. Selecting a suitable kernel is commonly based on suspicion of likely patterns and smoothness in the data. The most frequently kernel function used is a squared exponential (SE) and is expressed as follows:

$$k_{SE}(z_i, z_j) = \theta_1 \cdot e^{-\frac{(z_i - z_j)^2}{\theta_2}} \quad (6)$$

C. GPR-LSTM Regression

The input of our method $x_i = [rh, ws, t2]$ is multivariate of time-series ocean modelling. We demonstrate two main stages in the proposed cascade method. The first stage is using a regression LSTM as presented in Section 2.1. A trained regression LSTM model \hat{M} , where the input is multivariate time-series $x = [rh, ws, t2]$ and the output is a predicted univariate time-series $\hat{y} = [SST]$. The structure of the regression LSTM network is composed of four layers; the first input layer (sequence layer) has multiple input variables x_i , each $x_i = [rh, ws, t2]$ represents the daily observations from 2005 till 2015 in the training process. In the second layer, we used based on our experimentation 500 hidden memory units in

LSTM layers, followed by the third fully connected layer to do the mapping from the output of LSTM layer into desired output SST in the last regression output layers as shown in Fig 1. In the second stage, another training process is performed to construct the model \hat{M} of GPR-LSTM by using the resultant output \hat{y} of LSTM as an input to GPR regression where the output still the desired univariate $\hat{y} = [SST]$.

The main idea here is to reveal that we reduce the error of predicted result \hat{y} of model \hat{M} of $G(f(x))$ compared to the error of predicted result \bar{y} of model \hat{M} using f knowing that y is the real output of x . so, our mathematical proof will be as the following:

Notation:

We define: $\|x\|$ as $L2$ norm of x , G : Gaussian process regression. $f(x)$: LSTM regression of input x , $G(f(x))$: cascade of GPR and LSTM regression, also, we define \hat{x} :

$$\hat{x} = \underset{x \in \mathbb{R}}{\text{Argmin}} \|f(x)\|$$

$$\forall x \in \mathbb{R}, f(x) \geq f(\hat{x}),$$

Following to the earlier definition, we have:

Theorem:

$$\|\hat{y} - y\| \leq \|\bar{y} - y\|$$

where:

$$\hat{y} = G(f(x)), \bar{y} = f(x)$$

proof:

Knowing that G can be addressed as the following:

$$G = \underset{G \in \text{GPR}}{\text{Argmin}} \|\tilde{G}(\bar{y}) - y\|$$

$$\iff \|G(\bar{y}) - y\| \leq \|\tilde{G}(\bar{y}) - y\| \quad \forall G \in \text{GPR}$$

when $\tilde{G} = I$

$$\begin{aligned} &\iff \|G(\bar{y}) - y\| \leq \|\bar{y} - y\| \\ &\iff \|G(f(x)) - y\| \leq \|\bar{y} - y\| \\ &\iff \|\hat{y} - y\| \leq \|\bar{y} - y\| \end{aligned}$$

The proof can be summarized as the following:

G uses the argmin of \tilde{G} , its L2 norm is always smaller or equal to $f(\cdot)$.

III. RESULTS AND DISCUSSION

In this section, we present the studied SST dataset. Then, the performances of ensemble learning methods are demonstrated for SST prediction. Furthermore, the results from GPR-LSTM are compared to those obtained by SVR and GPR models. More details about SVR and GPR can be found in [25] and [26] respectively.

The performance of the proposed GPR-LSTM approach for SST Prediction is evaluated on the Red Sea SST time-series dataset. The SST dataset was obtained from the Operational Sea Surface Temperature, and Sea Ice Analysis (OSTIA) system, provided by the UK met office [27], [28]. The OSTIA SST dataset is a combination of two products: satellite data provided by international agencies via the Group for High-Resolution SST (GHRSSST), and in-situ observations from the International Comprehensive Ocean-Atmosphere Data Set (ICODAS) database. The OSTIA provides daily SST fields at 1/20 (approx. 5km) grid resolution spanning the period 1985 to present. In this study, we used external meteorological data for predicting SST in the Red Sea. The hourly wind speed (WS), air temperature at 2m (T2), relative humidity (RH) obtained from the fifth generation ECMWF reanalysis (ERA5, [29]) available at 0.25° resolution were used for training and predicting the SST in the proposed hybrid GPR-LSTM method.

A three measurement indexes have been adopted to quantify the accuracy of SST prediction: Root Mean Square Error (RM-SPE), Coefficient of determination (R^2), and Mean Absolute Percentage (MAPE). where y_t , and \hat{y}_t , are the i th measured SST and its predicted value, respectively, in the testing dataset, for $i=1, \dots, n_t$, where n_t is the whole number of the testing dataset as shown in Table I.

Metric	Definition
RMSPE	$\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$
MAPE	$\frac{100}{n} \sum_{t=1}^n \left \frac{y_t - \hat{y}_t}{y_t} \right \%$
R^2	$\frac{\sum_{t=1}^n [(y_t - \bar{y}) \cdot (\hat{y}_t - \bar{y})]^2}{\sqrt{\sum_{t=1}^n (y_t - \bar{y})^2} \cdot \sqrt{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}}$

TABLE I

DEFINITION OF MEASUREMENTS OF EFFECTIVENESS.

We present in Fig.2, some example of our extensive training progress on 516 spatial representative points of the Red Sea.; this figure shows the output of response, predicted vs

Actual, and residuals plots of GPR, LSTM and GPR-LSTM, respectively. We can observe clearly that the resultant error of regressed SST of GPR-LSTM is less than the ones in GPR or LSTM.

We conducted an experiment that consists of 15 years (2005 - 2020) daily SST, RH, WS, and T2 Time-series data. We divided these samples further into two sets: training and testing. The training dataset consists of 10 years of data (2005-2015), for a total of 3650 multivariate inputs, each day has three variables inputs (RH, T2 and WS), and the output of the training set is the corresponding SST as one output. The testing dataset also consists of 5 years of data (2016-2020), equivalent to 1825 time-series data.

In this study, we adopted the exponentially weighted moving average (EWMA) algorithm to smooth the SST time-series data and eliminate outliers:

$$\mathbf{z}_t = \alpha \mathbf{x}_t + (1 - \alpha) \mathbf{z}_{t-1}, \quad (7)$$

where \mathbf{x}_t denotes the SST observations, \mathbf{z}_t refers to the filter SST, \mathbf{z}_0 is selected to be the starting point \mathbf{x}_0 and $\alpha \in [0, 1]$ denotes the smoothing parameter that defines the memory depth of the exponential smoothing procedure. Here, we use a large value of α (i.e., $\alpha = 0.7$) to slightly smooth the data and keep the most variance in the original data.

Next, we normalize the smoothed data to get data with zero mean and unit variance.

$$\mathbf{z} = \frac{\mathbf{z} - \bar{\mathbf{z}}}{\sigma_{\mathbf{z}}}, \quad (8)$$

where \mathbf{z} denotes the smoothed data, $\bar{\mathbf{z}}$ refers to the average of the smoothed data, and $\sigma_{\mathbf{z}}$ is its standard deviation. This procedure will reversed after getting the prediction of SST.

The lower RMSPE or MAPE and the higher R^2 value mean better performance and accurate prediction. Also, the distribution of prediction errors is investigated using the boxplot function.

Besides our GPR-LSTM method, we also predict SST via the Ensemble Trees (boosted and bagged trees), GPR, SVR (Support Vector Regression), and LSTM for the performance and model comparison. We consider four kernels for the GPR and six kernels for the SVR. Thus, we predict SST on a total of 14 methods, as shown in Table II.

Figure 3 shows some outputs of our method on the testing dataset at some different locations in the Red Sea; each sub figure presents the response plot of the predicted SST by GPR-LSTM method (blue) and the measured SST (red). We can see clearly that the Predicted SST follows the measured SST on different locations of the Red Sea.

Figure 4 demonstrates the scatter plot of some outputs by GPR-LSTM method, where the X-axis represents the measured SST, and the Y-axis represents the predicted SST by our method. This scatter plot confirms the concurrence between the predicted and measured SST at some different longitudes and latitudes of the Red Sea.

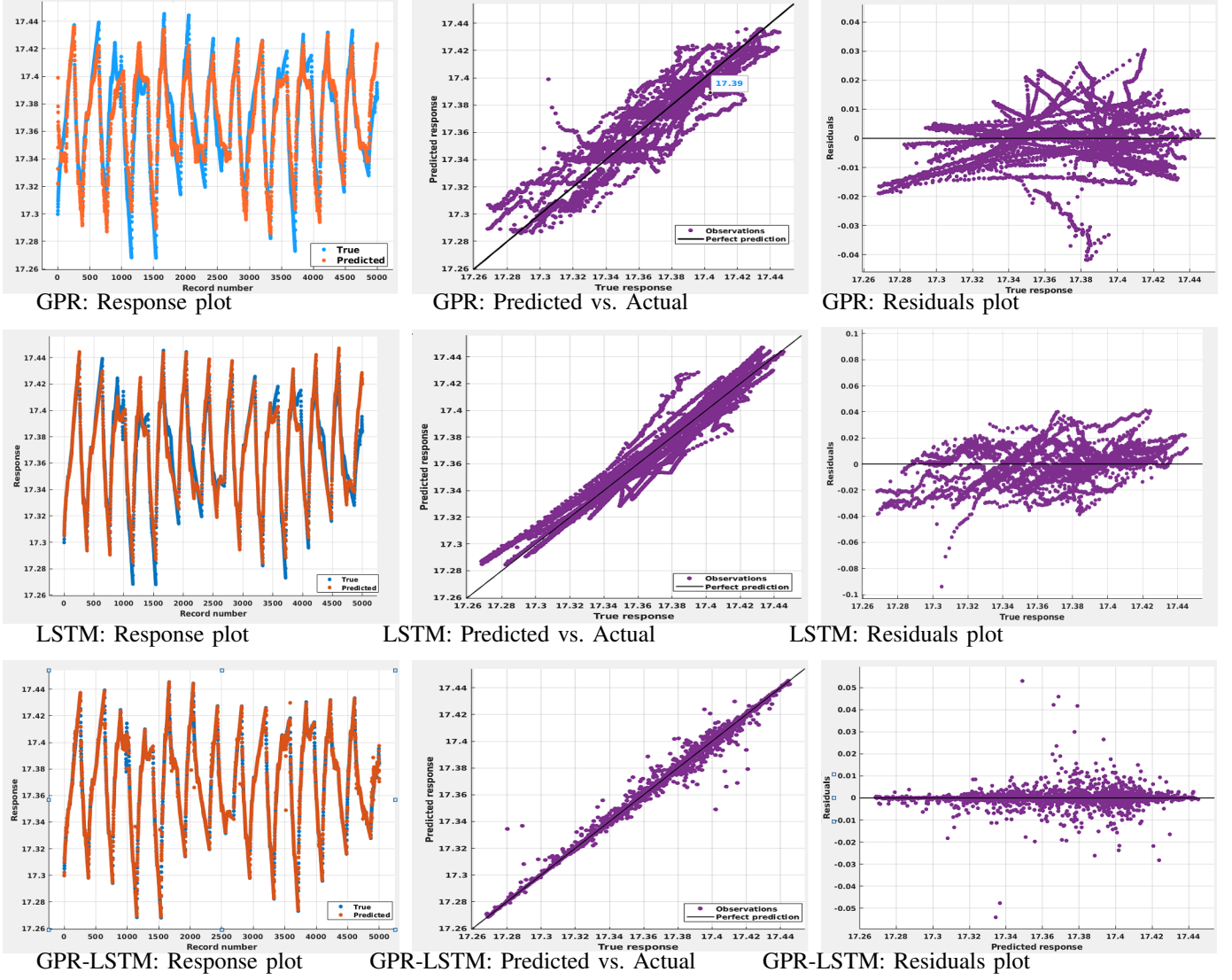


Fig. 2. Training Progress of daily SST (SQRT value) regression using GPR, LSTM and GPR-LSTM respectively

Model		Description	kernel
SVR	SVM_LI	Support Vector Machine with Linear kernel	$X_i^T X_j$
	SVM_CU	SVR with Cubic kernel	$(1 + X_i^T X_j)^3$
	SVM_QU	SVR with Quadratic kernel	$(1 + X_i^T X_j)^2$
	SVM_MG	SVR with Medium Gaussian kernel	$\exp(-\sqrt{\rho} \ X_i - X_j\ ^2)$
	SVM_FG	SVR with Fine Gaussian kernel	$\exp(-3\sqrt{\rho} \ X_i - X_j\ ^2)$
	SVM_CG	SVR with Coarse Gaussian kernel	$\exp(-4\sqrt{\rho} \ X_i - X_j\ ^2)$
Ensemble Trees	ENST_BS	Ensembles of Trees: Boosted Trees	-
	ENST_BG	Ensembles of Trees: Bagged Trees	-
GPR	GPR_MT	GPR with Mattern 5/2 kernel	$\sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2} \right) \exp\left(\frac{1}{2} \frac{r^2}{\sigma_l^2} \right)$
	GPR_QU	GPR with Rational Quadratic kernel	$\sigma_f^2 \left(1 + \frac{r^2}{2\alpha\sigma_l^2} \right)^{-a}$
	GPR_SQ	GPR with Squared kernel	$\sigma_f^2 \exp\left(\frac{1}{2} \frac{r^2}{\sigma_l^2} \right)$
	GPR_EX	GPR with Exponential kernel	$\sigma_f^2 \exp\left(\frac{r}{\sigma_l} \right)$
RNN	LSTM	RNN: LSTM	-
Ours	GPR-LSTM	GPR-LSTM with Exponential kernel	$\sigma_f^2 \exp\left(\frac{r}{\sigma_l} \right)$

TABLE II
ML BASED-APPROACH MODELS

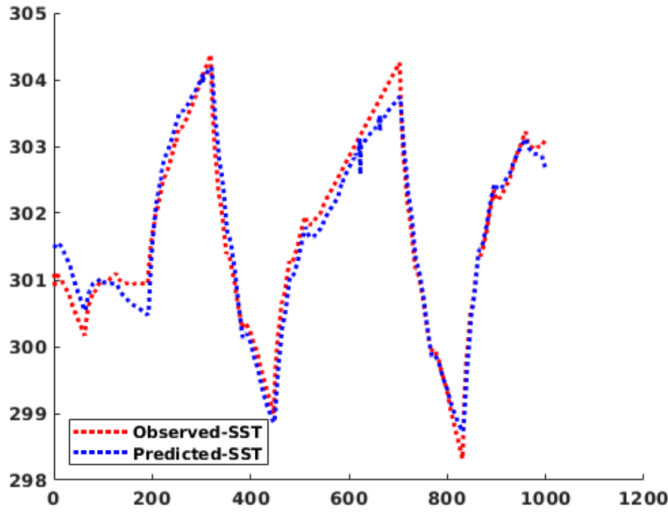


Fig. 3. Outputs of the measured and predicted daily SST (Kelvin) of GPR-LSTM

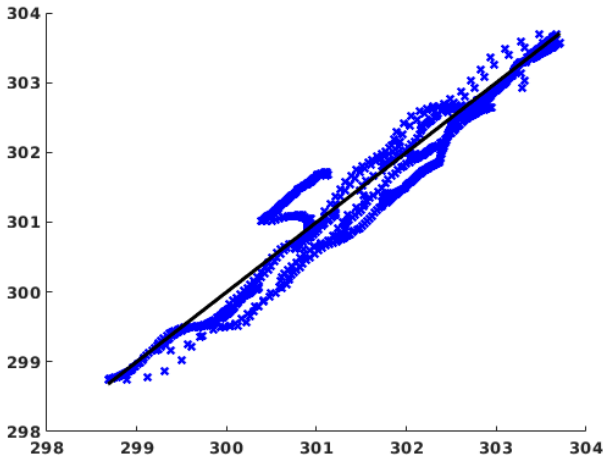


Fig. 4. Scatter plot of the measured and predicted daily SST in Kelvin of GPR-LSTM

Table III demonstrates the SST prediction comparison results of RMSPE and R^2 between SST \rightarrow SST (we predict SST from SST) and (WS,T2 and RH) \rightarrow SST (we predict SST from multivariate (WS,T2 and RH)). Also, these results are compared to ensemble models the three frequently used prediction ensemble learning methods: Ensembles Tress [30], Support Vector Regression [25], [31] and Gaussian Process Regression [25], [32]. We can observe clearly that Deep learning models and in specific the hybrid GPR-LSTM incorporating information from input variables (WS,T2 and RH) improves the prediction accuracy. This could be attributed to the cross correlation between the climate multivariate (WS,T2 and RH) and SST. These results represent the mean value of each of RMSPE, R^2 and MAPE over 516 different points in the Red Sea. We can observe clearly that our proposed GPR-LSTM method shows a better performance of RMSPE = 0.26, R^2 = 96.74 and MAPE = 17.59% for SST Prediction of the Red Sea.

In addition, we can see clearly in table III, that the proposed GPR-LSTM outperforms the ensemble learning models. This is due to high capability of hybrid technique between Deep learning and GPR to extract robust features from our dataset.

In summary, the proposed GPR-LSTM algorithm enables analysts to formulate and address problems involving querying SST time-series prediction within a specific spatiotemporal range. The modular design supported by our methodology makes it suitable for use in the backend of interactive analytics systems as it does not require substantial computational resources because we are using a pre-trained LSTM, which was trained on Red Sea datasets. We demonstrate the suitability through the comparison of our results with 14 different state-of-the-art models. One significant advantage of the GPR-LSTM-based algorithm proposed in this paper is the flexibility to incorporate additional multivariate environmental or non-environmental variables for the prediction and forecasting that we plan to do in our future work, in addition, to add the spatial information to our method and analysis.

IV. CONCLUSION

This work proposed a novel approach GPR-LSTM deep learning-based approach through the combination of two well-known methods, GPR and LSTM. We addressed and tested the performance of the proposed model on SST prediction in the Red Sea. The proposed algorithm operates in two training stages: Multivariate inputs were given to RNN regression to extract the first phase of regression SST. Then, the resultant output is used as an input of the GPR model to find the final predicted SST. We performed extensive experiments utilizing daily SST datasets of the Red Sea over a period of 15 years. The experimentation results use three measurements: Root Mean Square Error (RMSPE), Mean Absolute Percentage (MAPE) and R^2 for validation SST prediction in the Red Sea. We compared the results of the proposed method with four well known ensembles learning methods to demonstrate the relevance of the proposed approach. Our results show a better performance in term of RMSPE, and MAPE for SST prediction in the Red Sea.

V. ACKNOWLEDGMENTS

The study was supported by King Abdullah University of Science and Technology (KAUST) under the “Virtual Red Sea Initiative”, Award Number REP/1/3268-01-01, and the Saudi ARAMCO Marine Environmental Research Center at KAUST (SAMERK).

REFERENCES

- [1] S. Bouindour, H. Snoussi, M. M. Hittawe, N. Tazi, and T. Wang, “An on-line and adaptive method for detecting abnormal events in videos using spatio-temporal convnet,” *Applied Sciences*, vol. 9, no. 4, p. 757, 2019.
- [2] M. M. Hittawe, S. Afzal, T. Jamil, H. Snoussi, I. Hoteit, and O. Knio, “Abnormal events detection using deep neural networks: application to extreme sea surface temperature detection in the red sea,” *Journal of Electronic Imaging*, vol. 28, no. 2, p. 021012, 2019.

Model		RMSPE		R ²	
		SST->SST	(T2, WS, RH)->SST	SST->SST	(T2, WS, RH)->SST
Ensemble Trees					
	ENST_BS	7.09	0.48	5.10	60.22
	ENST_BG	0.52	0.47	55.02	61.98
SVR					
	SVM_CG	0.33	0.34	89.31	89.33
	SVM_FG	0.40	0.37	81.80	69.31
	SVM_QU	0.37	0.40	69.13	81.23
	SVM_LI	0.36	0.36	90.21	89.21
	SVM_CU	0.32	0.33	89.31	89.28
	SVM_MG	0.34	0.36	75.84	83.24
GPR					
	GPR_MT	0.47	0.39	67.82	69.91
	GPR_QU	0.46	0.40	69.30	81.29
	GPR_SQ	0.44	0.39	81.80	67.20
	GPR_EX	0.39	0.31	88.11	90.01
Deep Learning					
	RNN_LSTM	0.41	0.30	89.20	91.35
	GPR_LSTM	0.29	0.26	95.21	96.74

TABLE III
QUANTATIVE RESULTS OF RMSPE AND R²

- [3] Y. Zhang, M. Mazen Hittawe, K. Katterbauer, A. F. Marsala, O. M. Knio, and I. Hoteit, "Joint seismic and electromagnetic inversion for reservoir mapping using a deep learning aided feature-oriented approach," in *SEG Technical Program Expanded Abstracts 2020*. Society of Exploration Geophysicists, 2020, pp. 2186–2190.
- [4] J. L. Charco, T. Roque-Colt, K. Egas-Arizala, C. M. Pérez-Espinoza, and A. Cruz-Chóez, "Using multivariate time series data via long-short term memory network for temperature forecasting," in *Systems and Information Sciences*, M. Botto-Tobar, W. Zamora, J. Larrea Plúa, J. Bazurto Roldan, and A. Santamaría Philco, Eds. Cham: Springer International Publishing, 2021, pp. 38–47.
- [5] M. M. Hittawe, S. M. Muddamsetty, D. Sidibé, and F. Mériaudeau, "Multiple features extraction for timber defects detection and classification using svm," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 427–431.
- [6] O. Beya, M. Hittawe, D. Sidibé, and F. Meriaudeau, "Automatic detection and tracking of animal sperm cells in microscopy images," in *2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2015, pp. 155–159.
- [7] K. Patil and M. C. Deo, "Prediction of daily sea surface temperature using efficient neural networks," *Ocean Dyn.*, vol. 67, pp. 357–368, 2017.
- [8] F. Harrou, A. Zeroual, M. M. Hittawe, and Y. Sun, *Road Traffic Modeling and Management: Using Statistical Monitoring and Deep Learning*. Elsevier, 2021.
- [9] O. Beya, M. M. Hittawe, T. Alashkar, E. Fauvet, and O. Lalignant, "Applying non linear approach for ecg denoising and waves localization," in *2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2015, pp. 42–47.
- [10] O. Beya, M. M. Hittawe, N. Zegadi, E. Fauvet, and O. Lalignant, "Electrocardiogram signal analysing," in *Bio-inspired systems and signal processing (Bio-Signal 2016)*, 2016.
- [11] F. T. Tangang, W. W. Hsieh, and B. Tang, "Forecasting regional sea surface temperatures in the tropical pacific by neural network models, with wind stress and sea level pressure as predictors," *Journal of Geophysical Research: Oceans*, vol. 103, no. C4, pp. 7511–7522, 1998. [Online]. Available: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JC03414>
- [12] M. M. Hittawe, D. Sidibé, and F. Mériaudeau, "A machine vision based approach for timber knots detection," in *Twelfth International Conference on Quality Control by Artificial Vision 2015*, vol. 9534. SPIE, 2015, pp. 144–151.
- [13] M. M. Hittawe, D. Sidibé, O. Beya, and F. Mériaudeau, "Machine vision for timber grading singularities detection and applications," *Journal of Electronic Imaging*, vol. 26, no. 6, p. 063015, 2017.
- [14] K. Patil and M. C. Deo, "Basin-scale prediction of sea surface temperature with artificial neural networks," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 7, pp. 1441 – 1455, 01 Jul. 2018. [Online]. Available: <https://journals.ametsoc.org/view/journals/atot/35/7/jtech-d-17-0217.1.xml>
- [15] E. Garcia-Goriz and J. Garcia-Sanchez, "Prediction of sea surface temperatures in the western mediterranean sea by neural networks using satellite observations," *Geophysical Research Letters*, vol. 34, no. 11, 2007. [Online]. Available: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL029888>
- [16] M. Kim, H. Yang, and J. Kim, "Sea surface temperature and high water temperature occurrence prediction using a long short-term memory model," *Remote Sensing*, vol. 12, no. 21, 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/21/3654>
- [17] L. Wei, L. Guan, and L. Qu, "Prediction of sea surface temperature in the south china sea by artificial neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 4, pp. 558–562, 2020.
- [18] T. J. Sullivan, *Introduction to uncertainty quantification*. Springer, 2015, vol. 63.
- [19] A. Graves, "Generating sequences with recurrent neural networks," *arXiv preprint arXiv:1308.0850*, 2013.
- [20] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model," in *Eleventh annual conference of the international speech communication association*, 2010.
- [21] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *arXiv preprint arXiv:1409.3215*, 2014.
- [22] J. Schmidhuber and S. Hochreiter, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [23] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, "Learning precise timing with lstm recurrent networks," *Journal of machine learning research*, vol. 3, no. Aug, pp. 115–143, 2002.
- [24] L. Zhu and N. Laptev, "Deep and confident prediction for time series at uber," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2017, pp. 103–110.
- [25] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [26] T. Liu, H. Wei, and K. Zhang, "Wind power prediction with missing data using gaussian process regression and multiple imputation," *Applied Soft Computing*, vol. 71, pp. 905–916, 2018.
- [27] C. J. Donlon, M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, "The operational sea surface temperature and sea ice analysis (ostia) system," *Remote Sensing of Environment*, vol. 116, pp. 140–158, 2012.

- [28] J. D. Stark, C. J. Donlon, M. J. Martin, and M. E. McCulloch, "Ostia: An operational, high resolution, real time, global sea surface temperature analysis system," in *Oceans 2007-Europe*. IEEE, 2007, pp. 1–4.
- [29] H. Hersbach and Bell, "The era5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020. [Online]. Available: <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803>
- [30] T. G. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," *Machine learning*, vol. 40, no. 2, pp. 139–157, 2000.
- [31] M. M. Hittawe, D. Sidibé, and F. Mériaudeau, "Bag of words representation and svm classifier for timber knots detection on color images," in *2015 14th IAPR international conference on machine vision applications (MVA)*. IEEE, 2015, pp. 287–290.
- [32] H. Tolba, N. Dkhili, J. Nou, J. Eynard, S. Thil, and S. Grieu, "Multi-horizon forecasting of global horizontal irradiance using online gaussian process regression: A kernel study," *Energies*, vol. 13, no. 16, p. 4184, 2020.