MAUSAM, **73**, 3 (July 2022), 537-554

MAUSAM

DOI : https://doi.org/10.54302/mausam.v73i3.5931
Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No. 551.509.314 : 551.524

A new neurocomputing approach for medium-range temperature prediction

PRAVAT RABI NASKAR and SOMNATH NASKAR*

Meteorological office Port Blair, India Meteorological Department, MoES, Port Blair – 744 106, India

**Department of Industrial Engineering, Maulana Abul Kalam Azad University of Technology,*

Kolkata – 700 064, India

(*Received 26 June 2020, Accepted 28 February 2022*)

e mail : pravat091@gmail.com

सार – **मध्यावधि तापमान का सटीक पूर्वानुमान लगाने के लिए यह अध्ययन किया गया है। हमने एक नई मशीन लर्निंग (न्यूरोकंप्यूटिंग) तकनीक की मदद से उच्चतम-न्यूनतम तापमान समय श्रृंखला (संकेत) 5, 7 और 9 दिन आगे का पूर्वानुमान लगाने की कोशिश की है। तापमान समय श्रृंखला को अपघटित करने के लिए विविधता विधि अपघटन (वीएमडी) का उपयोग किया गया है और वीएमडी से प्राप्त अपघटित विधियों का व्यक्तिगत रूप से लॉन्ग शॉर्ट टर्म मेमोरी (एलएसटीएम) के निवेश के रूप में उपयोग किया गया है। एलएसटीएम ने अलग-अलग रूप से विधियों का पूर्वानुमान लगाया और पूर्वानुमानित संकेत उत्पन्न करने के लिए इन पूर्वानुमानित विधियों को जोडा गया। इस पद्धति द्वारा अनुमानित संकेत वास्तविक परीक्षण संकेत के साथ निकटता से काफी मेल खाता है और संकेत का पूर्वानुमान लगाने में त्रुटि को कम करता है।**

ABSTRACT. To predict medium-range temperature with appreciable accuracy this study has been undertaken. We have tried to predict max-min temperature time series (signal) 5, 7 and 9 days ahead with the help of a new machine learning (neurocomputing) technique. Variational Mode Decomposition (VMD) has been used to decompose temperature time series and decomposed modes obtained from VMD have been individually used as input to Long Short Term Memory (LSTM). The LSTM predicted the modes individually and predicted modes are combined to generate the predicted signal. The signal predicted by this method closely matches with the actual test signal and minimizes the error in predicting the signal.

Key words – Neurocomputing, Long short term memory, Variational mode decomposition, Port Blair.

1. Introduction

Temperature is one of the most important meteorological parameters. It affects many sectors like human health, agriculture, energy, transport, weather forecasting and so on. Very early and accurate prediction of temperature will help these sectors. The exact prediction of temperature is a challenge and it has attracted the attention of many researchers. Presently short-range prediction (up to 72 hours) of meteorological parameters is nearly accurate. But the medium-range prediction (up to 10 days) of these parameters with certain appreciable accuracy is still a challenge.

Machine learning is the area of computational science that focuses on analyzing and interpreting patterns and structures in data to enable learning, reasoning and decision-making outside of human interactions. It has been improved a lot since its first use. It has now applications in industries including manufacturing, retail, healthcare and life sciences, travel and hospitality, financial services, energy, feedstock and utilities, etc.

Machine learning can also be used in predicting meteorological parameters such as temperature, precipitation, wind speed, wind direction, solar radiation, sea-surface temperature, etc. Many scientists have used machine learning to predict these parameters across the world like (Paniagua-Tineo *et al*., 2011) predicted daily maximum temperature using a support vector regression (SVR), (Abdel-Aal, 2004) used abductive networks to forecast hourly temperature, (Dombayc *et al*., 2009) used artificial neural networks (ANN) to predict daily means ambient temperature, (Radhika *et al*., 2009) used support vector machines (SVM) to predict atmospheric temperature, (Lin *et al*., 2019) applied convolution long short term memory (convLSTM) network in numerical temperature prediction, (Salcedo-Sanz *et al*., 2015)

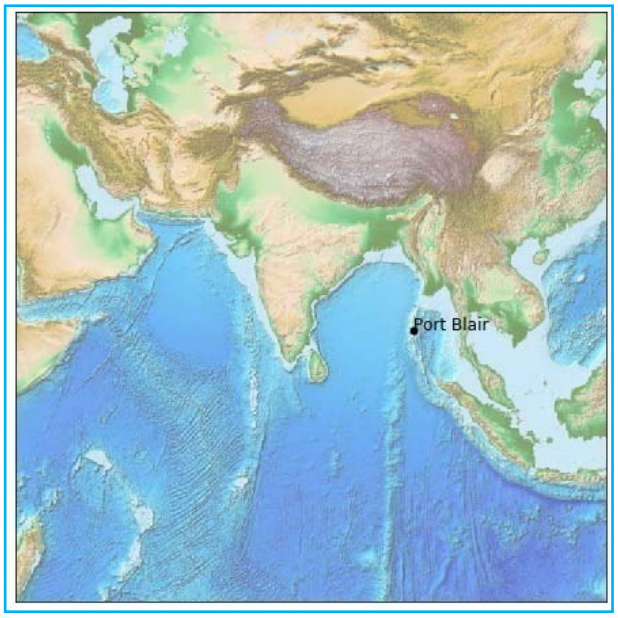

Fig. 1. Location map of Port Blair

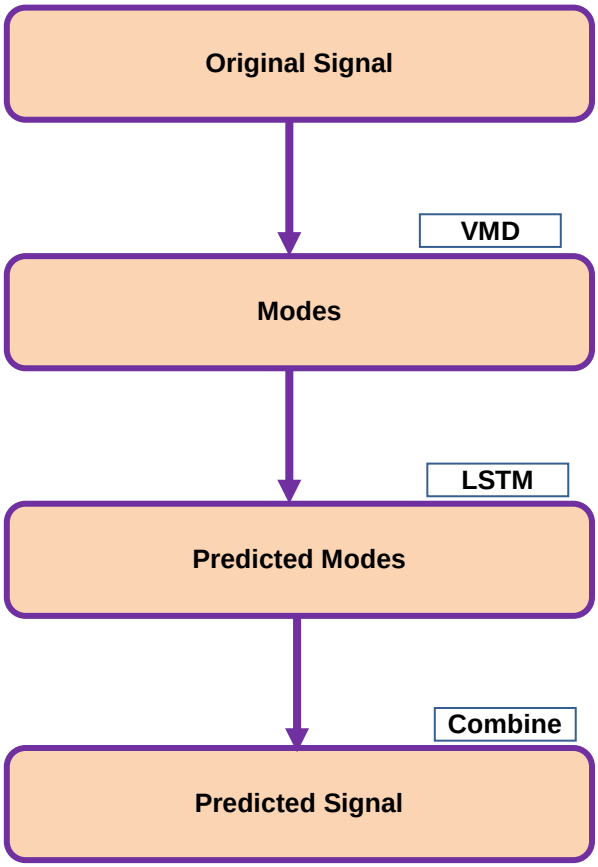

Fig. 2. Block diagram of medium-range temperature prediction

predicted monthly air temperature using machine learning algorithms, (Zhang and Dong, 2020) forecasted temperature combining Convolution Neural Network (CNN) and Recurrent Neural Network (RNN) based on time series data, (Poornima and Pushpalata, 2019) predicted rainfall using intensified long short term memory based recurrent neural network with weighted linear units, (Mislan *et al*., 2015) predicted monthly rainfall based on back propagation neural network (BPNN), (Dash *et al*., 2017) predicted rainfall of Kerala using single layer feed forward neural network (SLFN) and extreme learning machine (ELM) techniques, (Zhang *et al*., 2020) predicted short term wind speed using genetic algorithm and ANN model improved by variational mode decomposition (VMD), (Tian *et al*., 2019) used ensemble empirical mode decomposition-permutation entropy (EEMD-PE) and regularized extreme learning machine (RELM) to predict short term wind speed, (Khosravi *et al*., 2018) developed multilayer feed forward neural network, radial basis function neural network, SVR, fuzzy inference system, adaptive neuro-fuzzy inference system to predict hourly solar radiation, (Alzahrani *et al*., 2014) predicted solar irradiance using time series neural networks, (Patil *et al*., 2016) predicted sea surface temperature by combining numerical techniques and wavelet neural network, (Xu *et al*., 2020) predicted sea surface temperature by using multi-long short-term memory and CNN etc.

In this study, we have adopted a unique technique to predict the medium-range max-min temperature time series by using the long short-term memory network (LSTM) network and Variational Mode Decomposition (VMD). The block diagram of this technique is shown in Fig. 2. The LSTM shown in Fig. 3 has been used for its better performance over other neural networks such as ELM, SVR, Auto-Regressive Integrated Moving Average (ARIMA), Holt-Winters, RNN Silu, RNN with relu, etc. (Poornima and Pushpalata, 2019; Yu *et al*., 2019). The VMD has been considered over other mode decompositions such as Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD), etc. because it is adaptive quasi-orthogonal, completely non-recursive and effectively solves the problem of mode aliasing and endpoint effect.

2. Data and study area

Three groups of medium-range maximum and minimum temperature data sets are collected from Met Office Port Blair (India), the latitude and longitude of this observatory are 11°41′ N and 92°43′ E respectively. The location is shown in Fig. 1. The sampling period is 5, 7 and 9 days. The length of all the data sets is 1000.

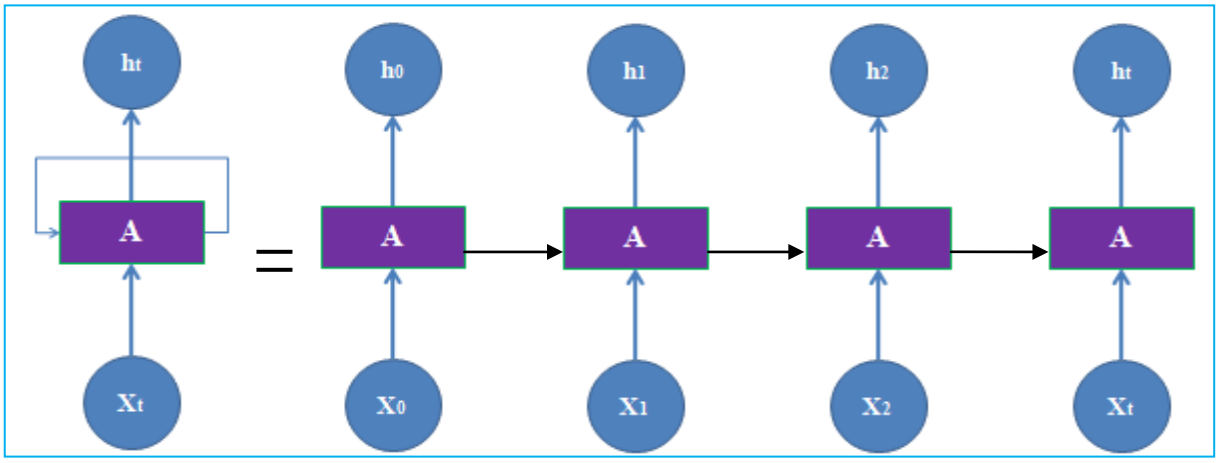

Fig. 3. Sequential processing of a recurrent neural network

3. Methodology

Here, first we have generated temperature time series then we have decomposed the time series spectrum into several modes using VMD and after that, each decomposed mode is separately used as an input to the LSTM network. After running the LSTM network we get the predicted mode for each input mode. Then the predicted modes are combined to generate the predicted signal. After that, the original test signal and the predicted signal is compared and we obtain the Mean Absolute Error (MAE), Mean Squared Error (MSE), R^2, relative root mean squared error (RRMSE), Theil Inequality Coefficient (TIC), Index of Agreement (IA) and Squared Sum Error (SSE). These values give an insight into how accurately the proposed approach has been able to predict the temperature time series. We have also compared the original test signal with that of the predicted signal without doing any mode decomposition.

3.1.1. *Variational Mode Decomposition (VMD)*

This is a relatively new method to deal with unstable and non-linear signals based on the scale of the signal itself. It was proposed by (Dragomiretskiy *et al*., 2014). In this method, the signal is decomposed into a limited number of mono-components. Each component has specific sparsity properties in the frequency domain. One main assumption of this technique is that each mode has a limited bandwidth in the frequency domain and is compact around a center frequency. A mode of VMD $u_k(t)$ is considered as an amplitude-modulated-frequency-modulated signal, which is expressed as :

$$u_k(t) = A_k(t)\cos[\varphi_k(t)] \tag{1}$$

in which $\varphi_k(t)$ and $A_k(t)$ respectively denote the phase and envelope of the *k*-th mode. The instantaneous frequency $\omega^k(t) = \frac{\partial \varphi_k(t)}{\partial t}$ is nonnegative and varies much slower than the phase $\varphi_k(t)$.

The detailed theoretical background and applications in the simulated signal analysis of the VMD technique can be referred to (Dragomiretskiy *et al*., 2014). The overall framework is the variational problem and VMD is the solution to the problem. The objective function of the variational problem is:

$$\min_{\{u_k\},\{\bar{\omega}_k\}} \{\sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t) \right] e^{-j\bar{\omega}_k t} \right\|_2^2\},$$

subjected to $\sum_{k=1}^{K} u_k(t) = f(t)$ (2)

where, δ is the Dirac function, * denotes the convolution and $\{u_k\} = \{u_1, u_2, u_3, \ldots u_k\}$ and $\{\bar{\omega}_k\} = \{\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3, \ldots . \bar{\omega}_k\}$ are the modes and center frequencies respectively.

Now using the second penalty factor α and Lagrangian multipliers λ, the constraint variation problem is changed into a non constraint variation problem. When the quadratic multiplication factor can guarantee the signal reconstruction accuracy in the presence of Gaussian noise, the Lagrangian operator keeps the constraint condition strict. The extended Lagrange expression is as follows:

$$L[(u_k),(\bar{\omega}_k),\lambda] = \alpha \sum_{k=1}^{K} \left\| \partial_t \left\{ \left[\delta(t) + \frac{j}{\pi t} \right] * u_k(t) \right\} e^{-j\bar{\omega}_k(t)} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{3}$$

where, α depends on the fidelity constraint. To solve equation 3 Alternate Direction Method Multipliers (ADMM) are used. The different center frequencies and the corresponding modes can be obtained through a sequence of iterative sub optimizations.

Each mode is represented as:

$$u_k(\omega)=\frac{f(\omega)-\sum_{i\neq k} u_i(\omega)+[\lambda(\omega)/2]}{1+2\alpha(\omega-\overline{\omega}_k)^2}(k=1,2,...K) \tag{4}$$

where, $f(\omega)$ is the Fast Fourier Transform (FFT) of the signal $f(t)$.

VMD mainly consists of the following steps:

Step1. Intrinsic Mode Update. The mode $u_k^{n+1}(\omega)$ is updated with equations 5, 6. The Wiener filtering is embedded for updating the mode directly in the Fourier domain with a filter tuned to the current center frequency $\overline{\omega}_k^n$.

$$u_k^{n+1}(\omega)=\frac{f(w)-\sum_{i<k} u_i^{n+1}(\omega)-\sum_{i>k} u_i^n(\omega)+\left[\frac{\lambda^n(\omega)}{2}\right]}{1+2\alpha(\omega-\overline{\omega}_k^n)^2} \tag{5}$$

$$u_k^{n+1}(t)=\Re\left\{\text{ifft}\left[u_k^{n+1}(\omega)\right]\right\} \tag{6}$$

Where $\Re\{\ \}$ is the real part of an analytic signal and ifft() denotes the inverse FFT.

Step2. Center frequency update. The center frequency $\overline{\omega}_k^{n+1}$ is updated as the center of gravity of the corresponding mode's power spectrum, which is presented as:

$$\omega_k^{n+1}=\frac{\int_0^\infty \omega\left|u_k^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|u_k^{n+1}(\omega)\right| d\omega}(k=1,2...K) \tag{7}$$

Step3. Dual ascent. For all frequencies $\omega \geq 0$, the Lagrangian multiplier $\lambda^{n+1}(\omega)$ is obtained by equation 8 as the dual ascent to enforce the exact signal reconstruction until the convergence criteria as shown in equation 9 is satisfied.

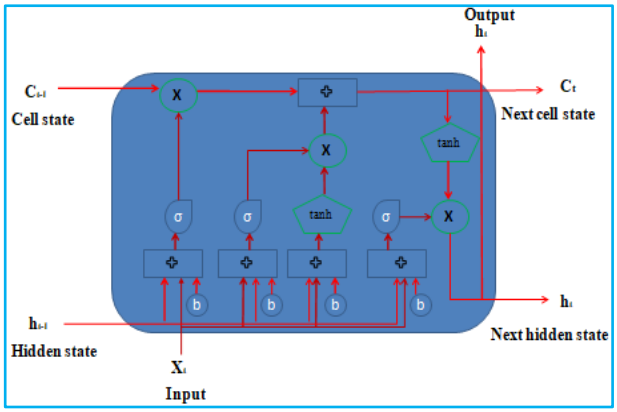

Fig. 4. Structure of an LSTM neural network

$$\lambda^{n+1}(\omega)=\lambda^n(\omega)+\tau\left[f(\omega)-\sum_k u_k^{n+1}\right] \tag{8}$$

$$\sum_{k=1}^{K}\frac{\left\|u_k^{n+1}-u_k^n\right\|_2^2}{\left\|u_k^n\right\|_2^2}\leq\varepsilon \tag{9}$$

3.1.2. *Long Short Term Memory (LSTM)*

LSTM is a specific architecture of RNN. It was intended to design for modeling temporal sequences and addressing the problems of back propagation through time (BPTT) in RNN architecture. It has a long-range dependency that makes it more accurate than conventional RNNs. A typical LSTM network is comprised of memory blocks called cells for remembering and propagating outputs explicitly in different time steps. The structure of the network is shown in Fig. 4. Here two states are being transferred to the next cell, the cell state and hidden state. The cell state is the main chain of data flow, which allows the data to flow forward essentially unchanged. However, some linear transformations may occur. The data can be added to or removed from the cell state via sigmoid gates. A gate is similar to a layer or a series of matrix operations, which contain different individual weights.

The first step in constructing an LSTM network is to identify information that is not required and will be omitted from the cell in the first step. This process of identifying and excluding data is decided by the sigmoid function, which takes the output of the last LSTM unit (h_{t-1}) at time t-1 and current input (X_t) at time t. Additionally, the sigmoid function determines which part from the old output should be eliminated. This gate is called forget gate (or f_t), where f_t is a vector with values

ranging from 0 to 1, corresponding to each number in the cell state, C_{t-1}.

$$f_t = \sigma W_f \left(h_{t-1}, X_t\right) + b_f \tag{10}$$

where in, σ is the sigmoid function, W_f and b_f are weight matrices and bias, respectively, of the forget gate.

The following step is deciding and storing information from the new input (X_t) in the cell state as well as to update the cell state. This step contains two parts, the sigmoid layer and tanh layer. The sigmoid layer decides whether the new information should be updated or ignored (0,1) and tanh function gives weight to the values which passed by, deciding their level of importance (-1 to 1). The two values are multiplied to update the new cell state. This new memory is then added to the old memory C_{t-1} resulting in C_t.

$$i_t = \sigma W_i \left(h_{t-1}, X_t\right) + b_i \tag{11}$$

$$N_t = \tanh W_n \left(h_{t-1}, X_t\right) + b_n \tag{12}$$

$$C_t = C_{t-1} f_t + N_t i_t \tag{13}$$

where, C_{t-1} and C_t are the cell states at time t-1 and t, while W and b are the weight matrices and bias, respectively, of the ell state.

In the final step, the output values (h_t) is based on the output cell state (O_t) but is a filtered version. Next, the output of the sigmoid gate (O_t) is multiplied by the new values created by tanh layer from the cell state (C_t), with a value ranging between -1 and 1.

$$O_t = \sigma\left[W_o \left(h_{t-1}, X_t\right) + b_o\right] \tag{14}$$

$$h_t = O_t \tanh\left(C_t\right) \tag{15}$$

where, W_o and b_o are the weight matrices and bias, respectively of the output gate.

3.1.3 *Error calculation*

To illustrate the effectiveness of the prediction approach for medium-range max-min temperature, the following 8 performance indicators are used to measure the prediction accuracy.

(*i*) *Mean Absolute Error (MAE)*

MAE can avoid the problem of mutual cancellation of errors, so it can accurately reflect the actual prediction error. The smaller the MAE the smaller the prediction error of the prediction model is.

It is expressed as

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| w(i) - \overline{w}(i) \right| \tag{16}$$

(*ii*) *Mean Squared Error (MSE)*

It can reflect the degree of dispersion of a data set. The smaller the MSE closer the sample value is to the average value.

It is expressed as

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left[w(i) - \overline{w}(i) \right]^2 \tag{17}$$

(*iii*) *Mean Absolute Percentile Error (MAPE)*

MAPE represents the percentage of the prediction error to the actual value. Therefore smaller the MAPE value the smaller the prediction error, *i.e.* the better the prediction effect.

It is expressed as

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| w(i) - \overline{w}(i) \right|}{w(i)} \times 100 \tag{18}$$

(*iv*) *Relative Root Mean Square Error (RRMSE)*

RRMSE represents the relative difference between the predicted and the actual value. The larger the RRMSE value the poorer the prediction accuracy.

It is expressed as :

$$\text{RRMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[\frac{w(i) - \overline{w}(i)}{w(i)} \right]^2} \tag{19}$$

(*v*) *Squared Sum Error (SSE)*

SSE calculates the sum of the error squares of the corresponding points between the predicted data and the actual data. The smaller the SSE the smaller the prediction error and the better the accuracy.

It is expressed as :

$$\text{SSE} = \sum_{i=1}^{N} \left[w(i) - \overline{w}(i) \right]^2 \tag{20}$$

Original spectrum and variational decomposed modes for maximum temperature (frequency 5 days) of Port Blair

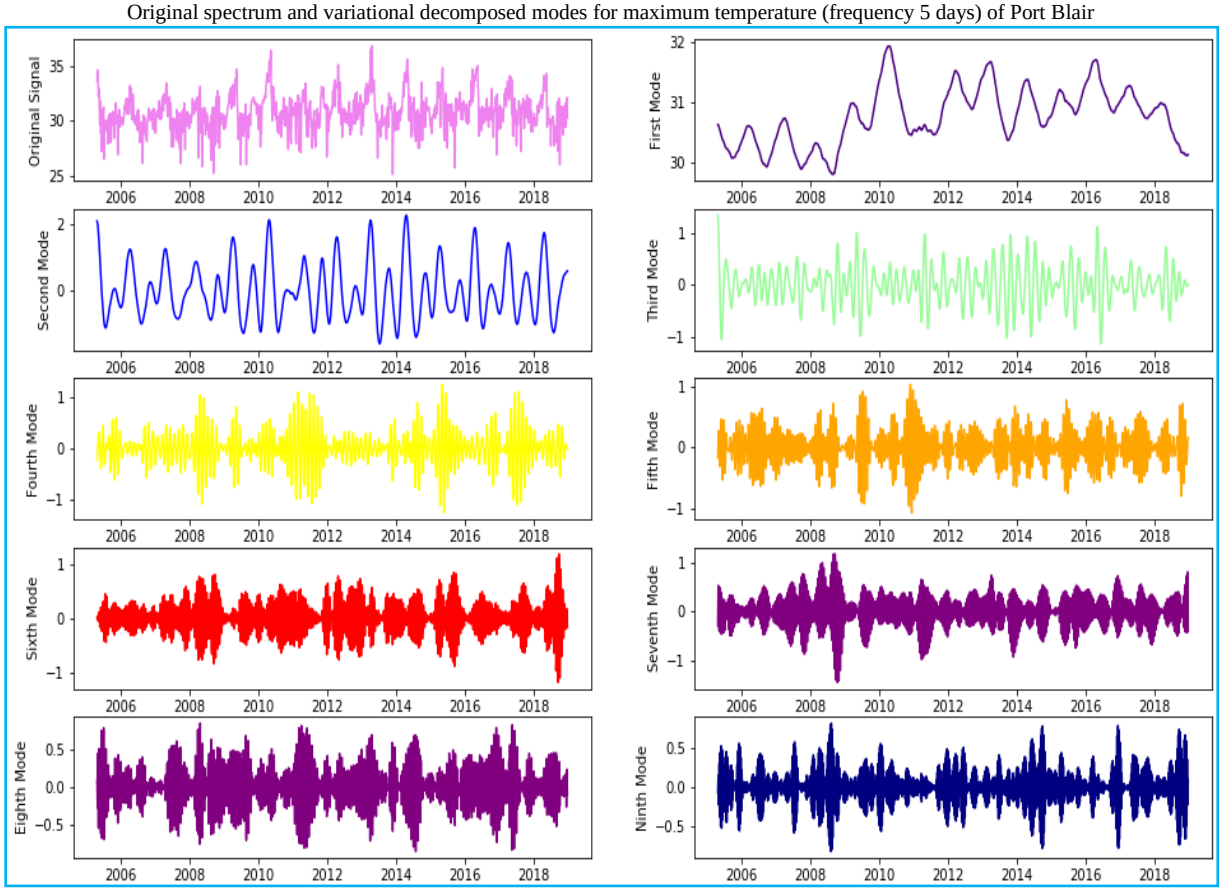

Fig. 5. Original maximum temperature spectrum of 5 days interval with its decomposed mode of Port Blair

(*vi*) R^2

It represents the fitness of a prediction model by the change of data. Its normal range is 0 to 1. The closer it approaches 1, the stronger the interpretation ability of the prediction model to the actual data, that is, the better the model fits the data. If $R^2 < 0$, the prediction model is inferior.

It is expressed as

$$R^2 = 1 - \frac{\sum_{i=1}^{N} [w(i) - \overline{w}(i)]^2}{\sum_{i=1}^{N} [w(i) - \overline{w}]^2} \qquad (21)$$

(*vii*) *Index of agreement*

Its physical meaning is that the ratio of the mean squared error and potential error of the predicted values to the actual value is subtracted by 1 and its range is 0 to 1. The larger the IA value, the higher the consistency between the actual and predicted values.

It is expressed as :

$$1 - \frac{\sum_{i=1}^{N} [w(i) - \overline{w}(i)]^2}{\sum_{i=1}^{N} [|w(i) - \overline{w}| + |w(i) + \overline{w}|]^2} \qquad (22)$$

(*viii*) *Theil Inequality Coefficient (TIC)*

Its value range is 0 to 1, the closer it is to 0, the smaller the root mean square of unit error and the actual value is closer to the predicted value.

It is expressed as :

$$\text{TIC} = \frac{\sqrt{\frac{\sum_{i=1}^{N} [w(i) - \overline{w}(i)]^2}{N}}}{\sqrt{\frac{\sum_{i=1}^{N} w(i)^2}{N}} + \sqrt{\frac{\sum_{i=1}^{N} \overline{w}(i)^2}{N}}} \qquad (23)$$

Variational decomposed modes for maximum temperature (frequency 5 days) for testing of Port Blair

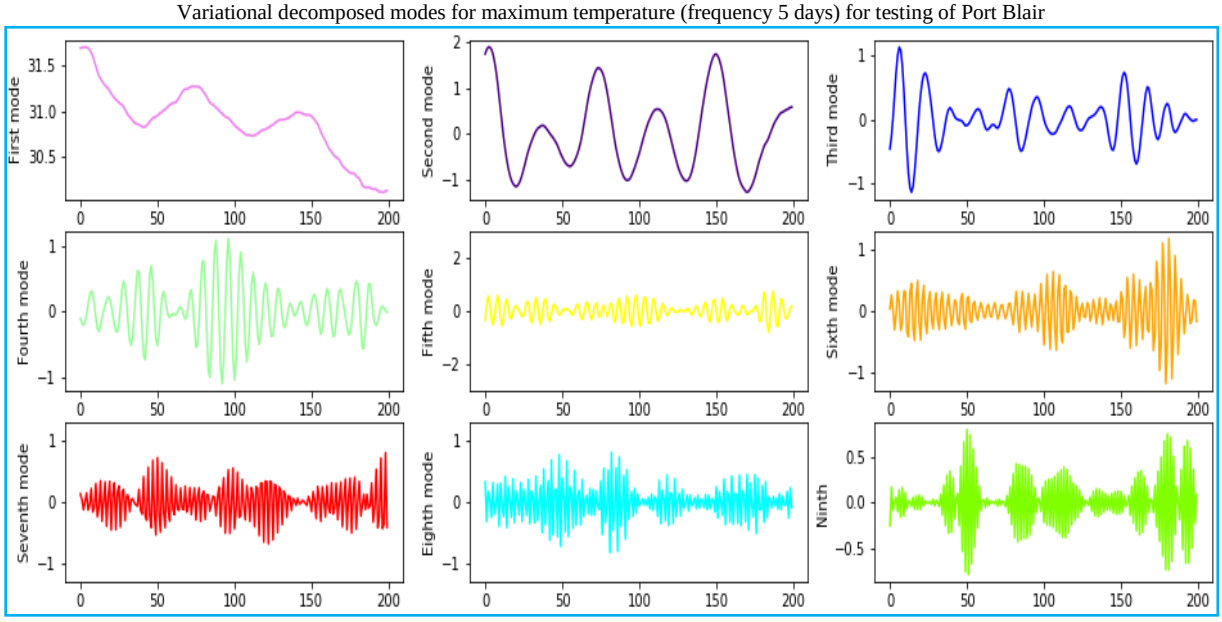

Fig. 6. The testing spectrum of the maximum temperature of 5 days interval with its decomposed mode of Port Blair

Variational decomposed modes for maximum temperature (frequency 5 days) for testing and predicted of Port Blair

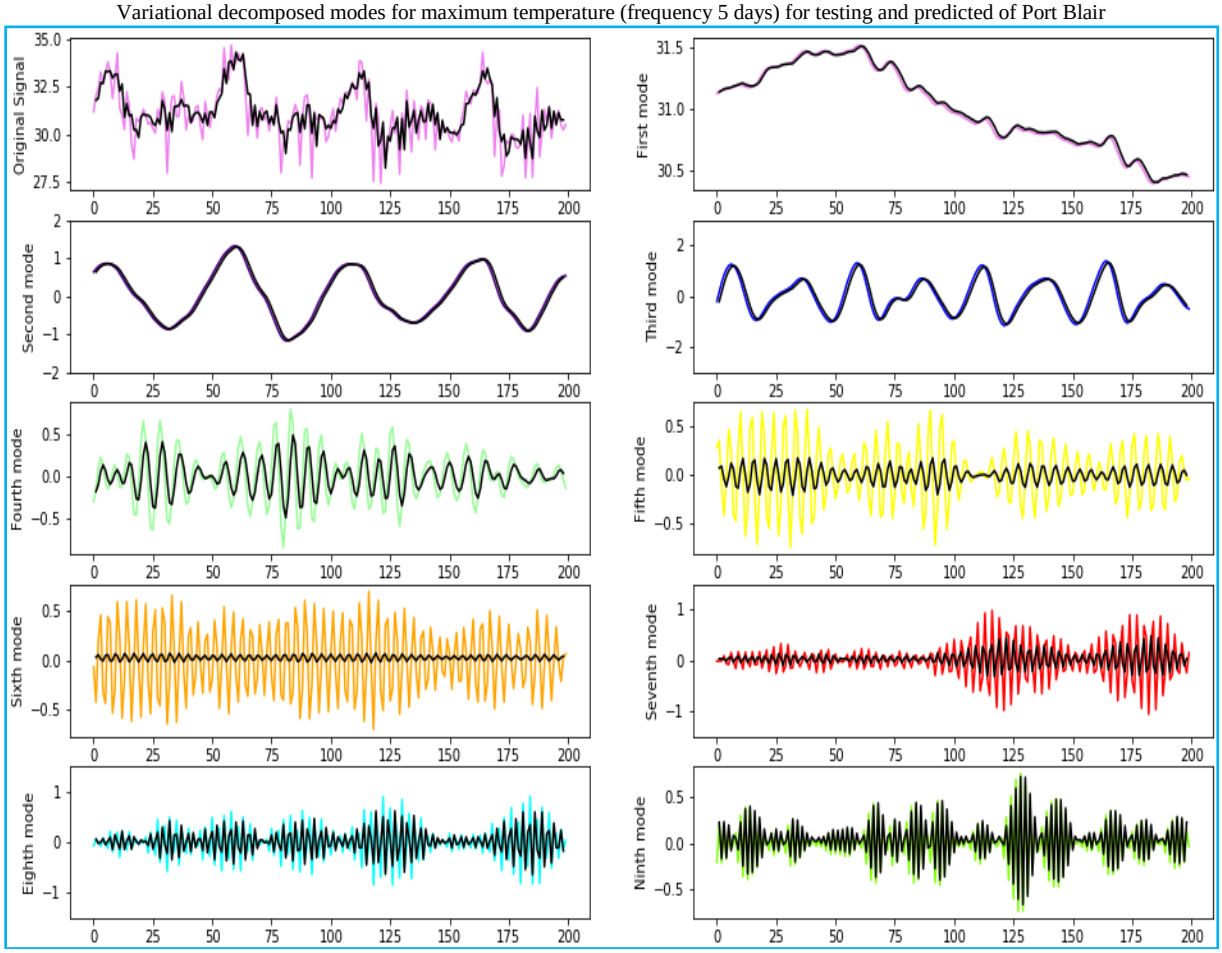

Fig. 7. Testing and predicted spectrum of the maximum temperature of 5 days interval with its decomposed mode of Port Blair

Original spectrum and variational decomposed modes for minimum temperature (frequency 5 days) of Port Blair

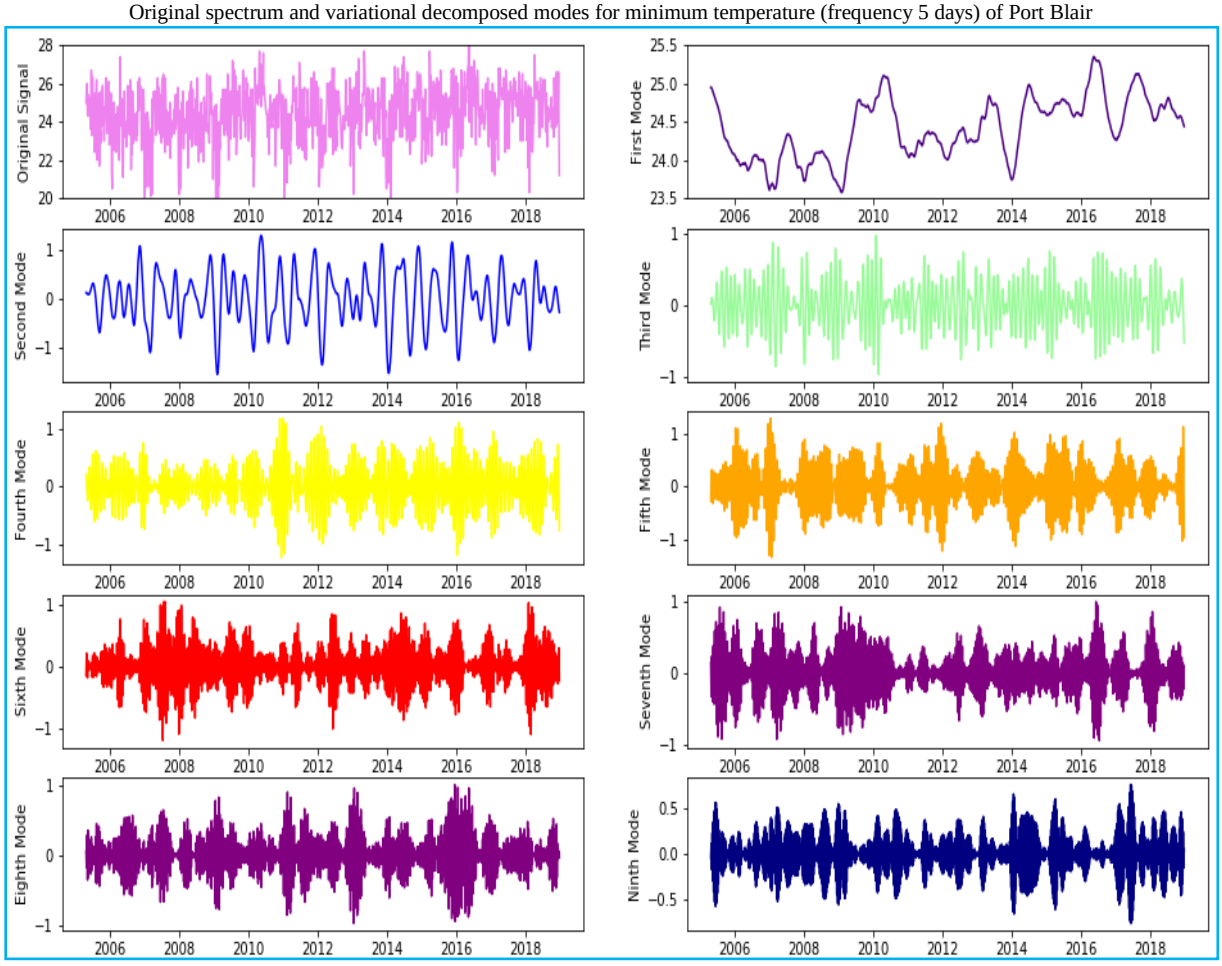

Fig. 8. Original minimum temperature spectrum of 5 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 5 days) for testing of Port Blair

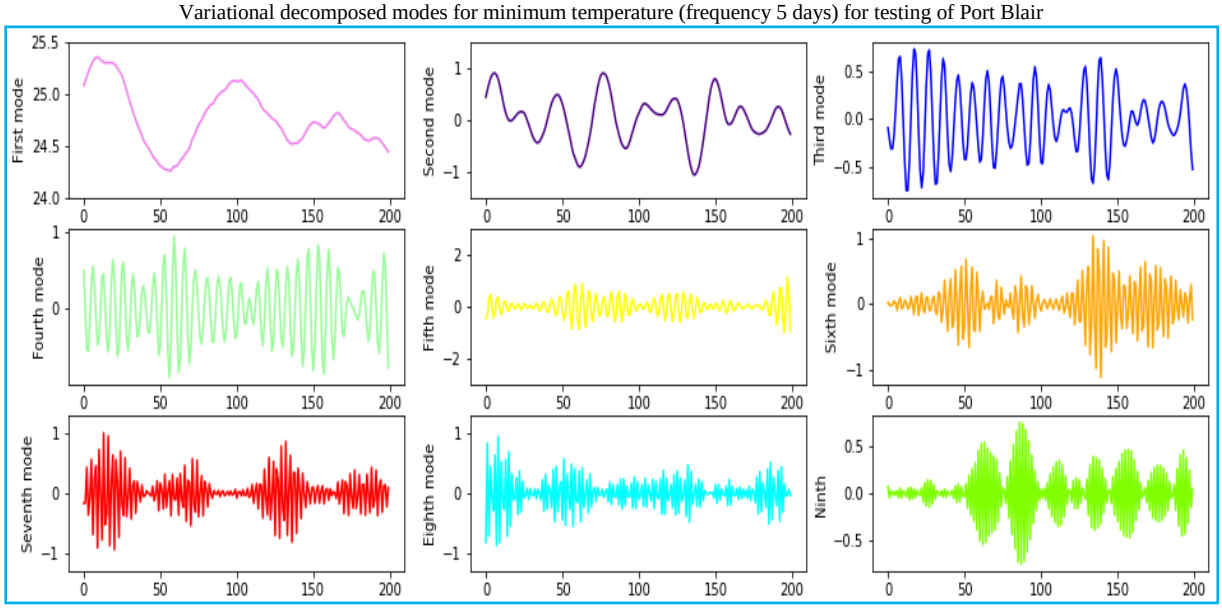

Fig. 9. The testing spectrum of the minimum temperature of 5 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 5 days) for testing and predicted of Port Blair

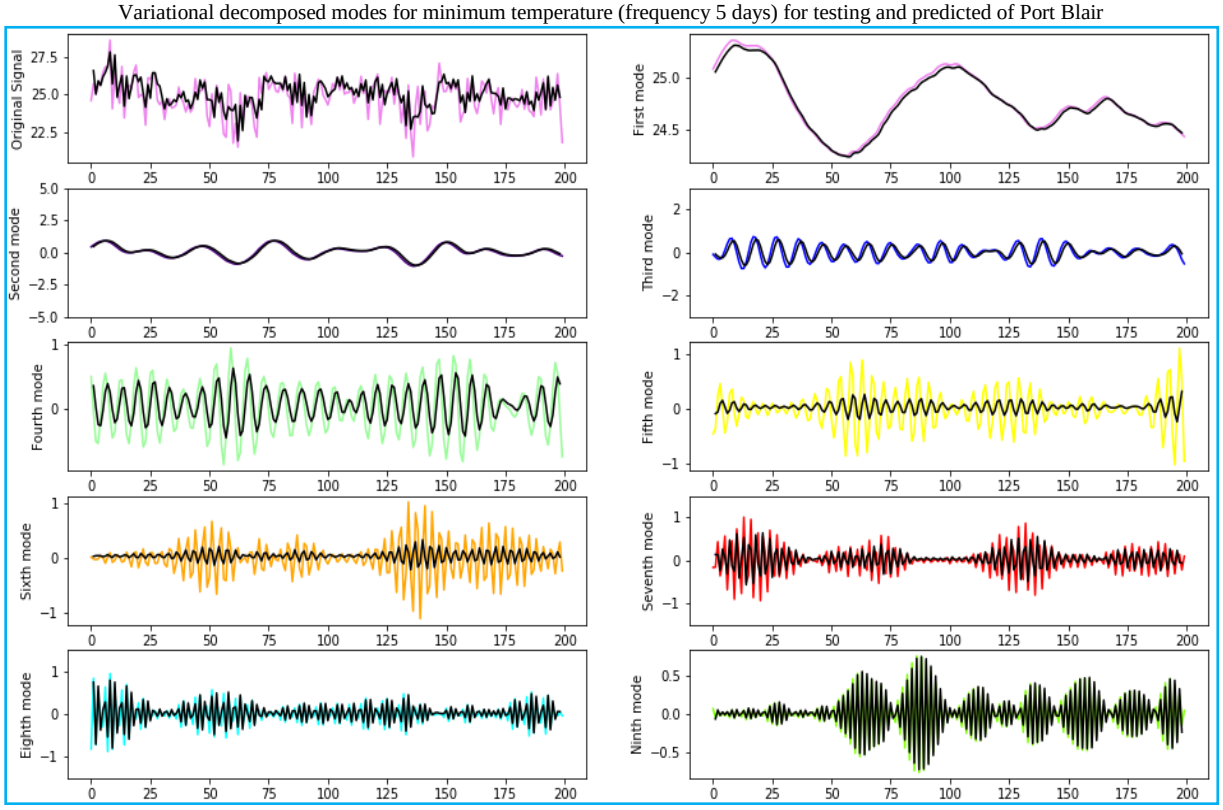

Fig. 10. Testing and predicted spectrum of the minimum temperature of 5 days interval with its decomposed mode of Port Blair

Original spectrum and variational decomposed modes for maximum temperature (frequency 7 days) of Port Blair

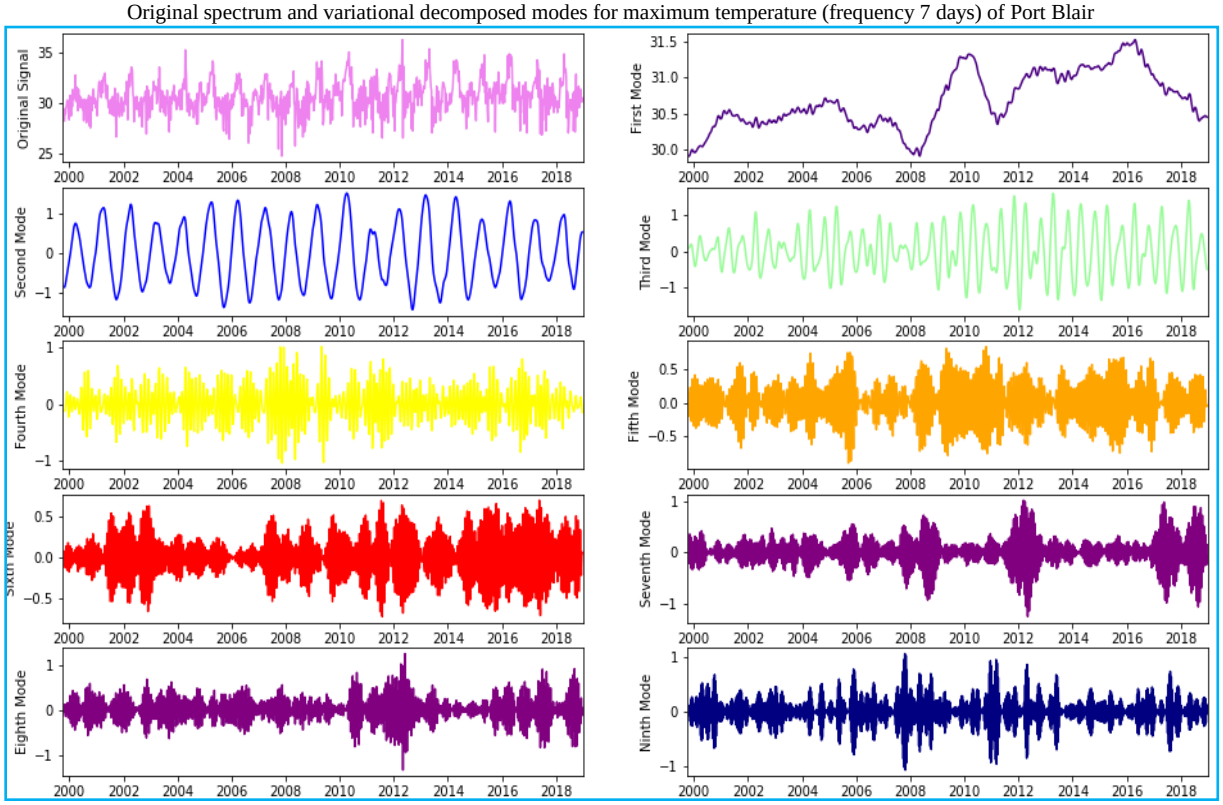

Fig. 11. Original maximum temperature spectrum of 7 days interval with its decomposed mode of Port Blair

Variational decomposed modes for maximum temperature (frequency 7 days) for testing of Port Blair

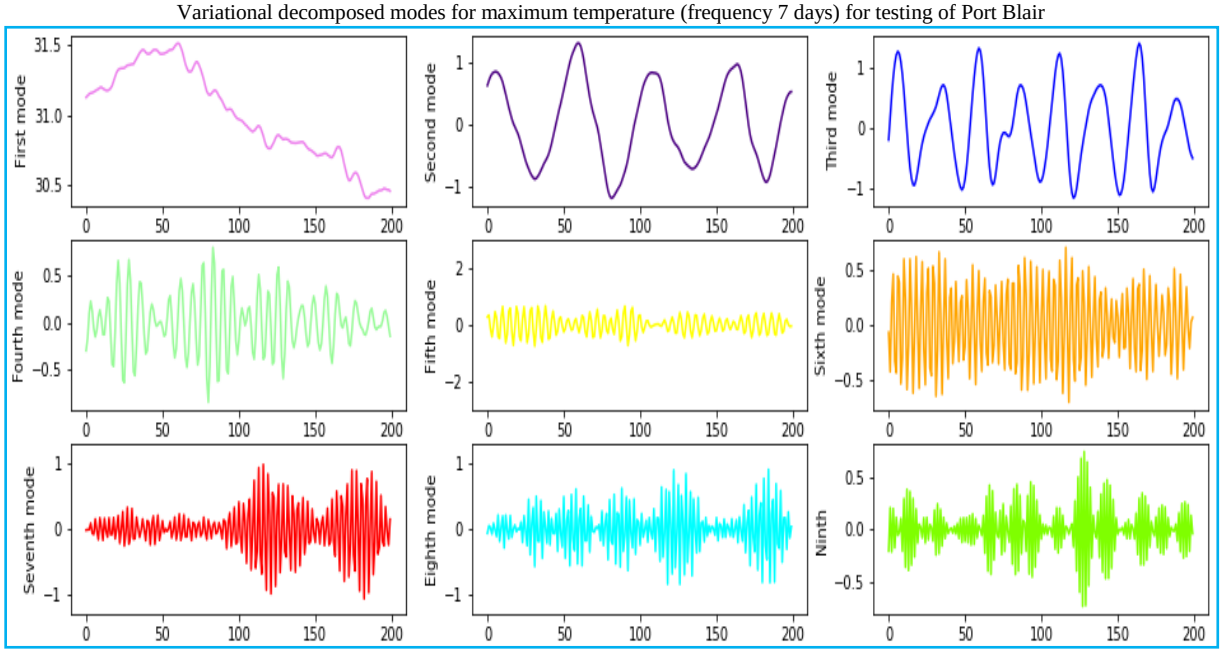

Fig. 12. The testing spectrum of the maximum temperature of 7 days interval with its decomposed mode of Port Blair

Variational decomposed modes for maximum temperature (frequency 7 days) for testing and predicted of Port Blair

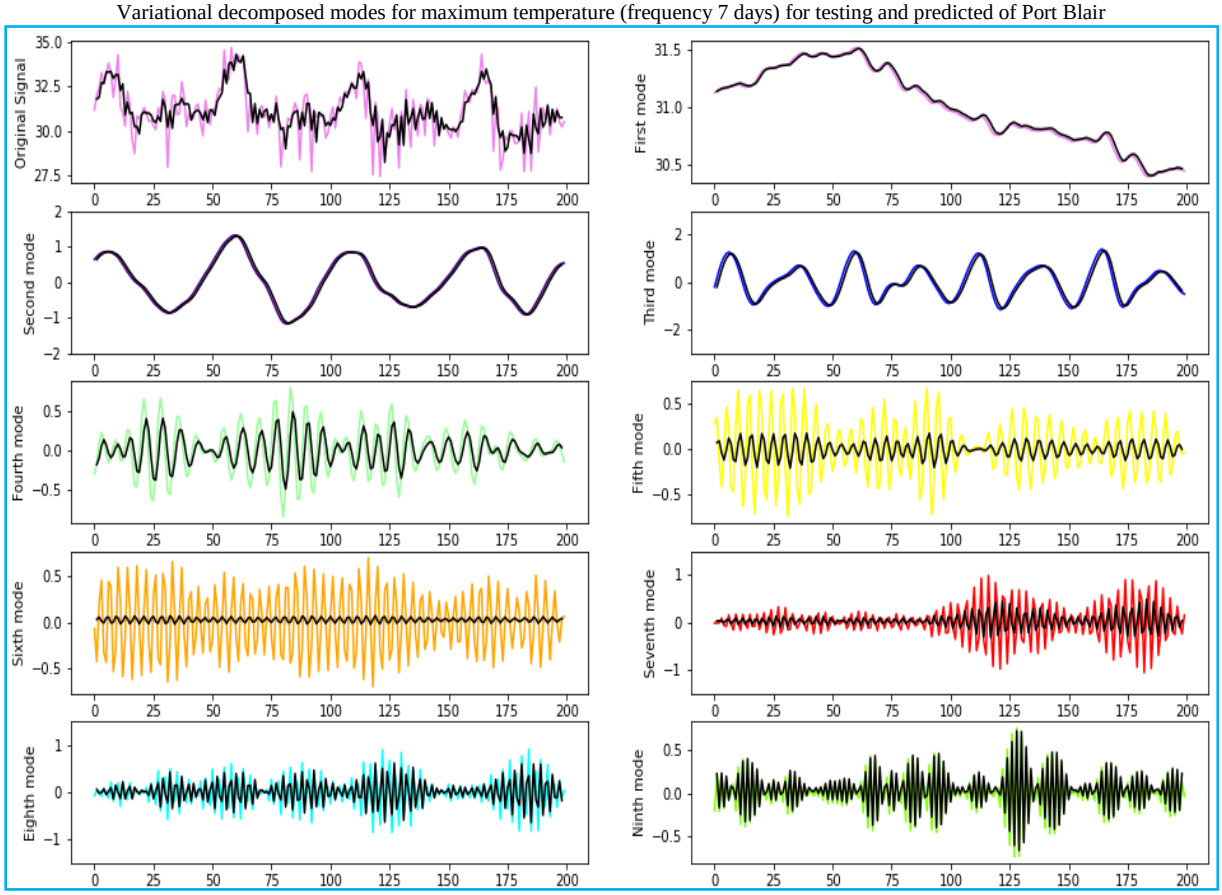

Fig. 13. Testing and predicted spectrum of the maximum temperature of 7 days interval with its decomposed mode of Port Blair

Original spectrum and variational decomposed modes for minimum temperature (frequency 7 days) of Port Blair

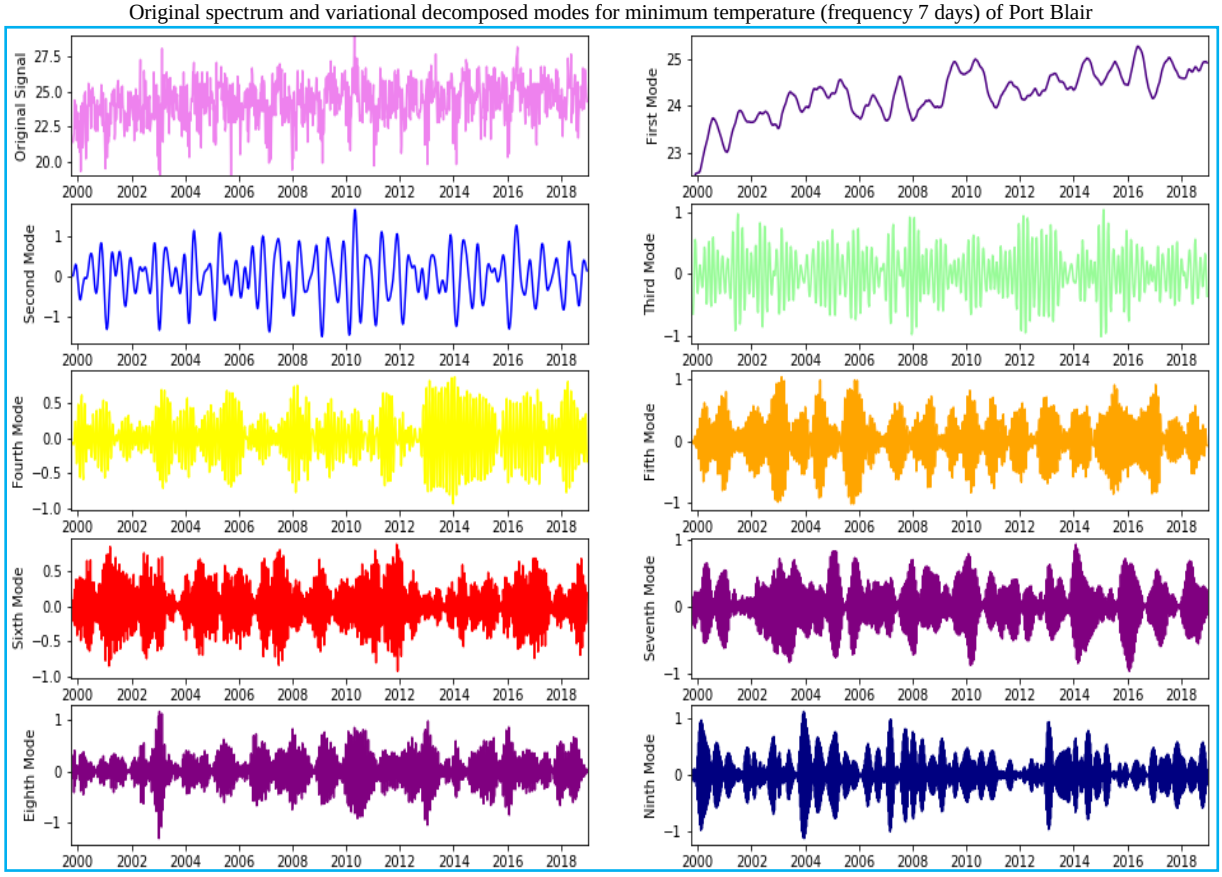

Fig. 14. Original minimum temperature spectrum of 7 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 7 days) for testing of Port Blair

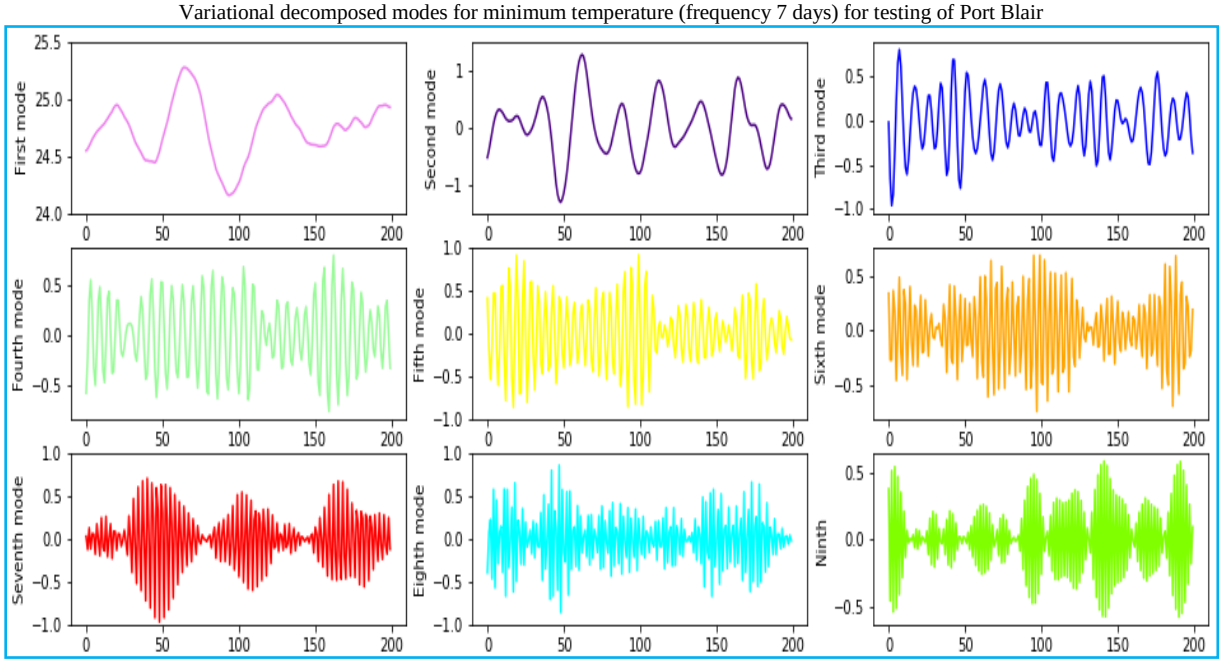

Fig. 15. The testing spectrum of the minimum temperature of 7 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 7 days) for testing and predicted of Port Blair

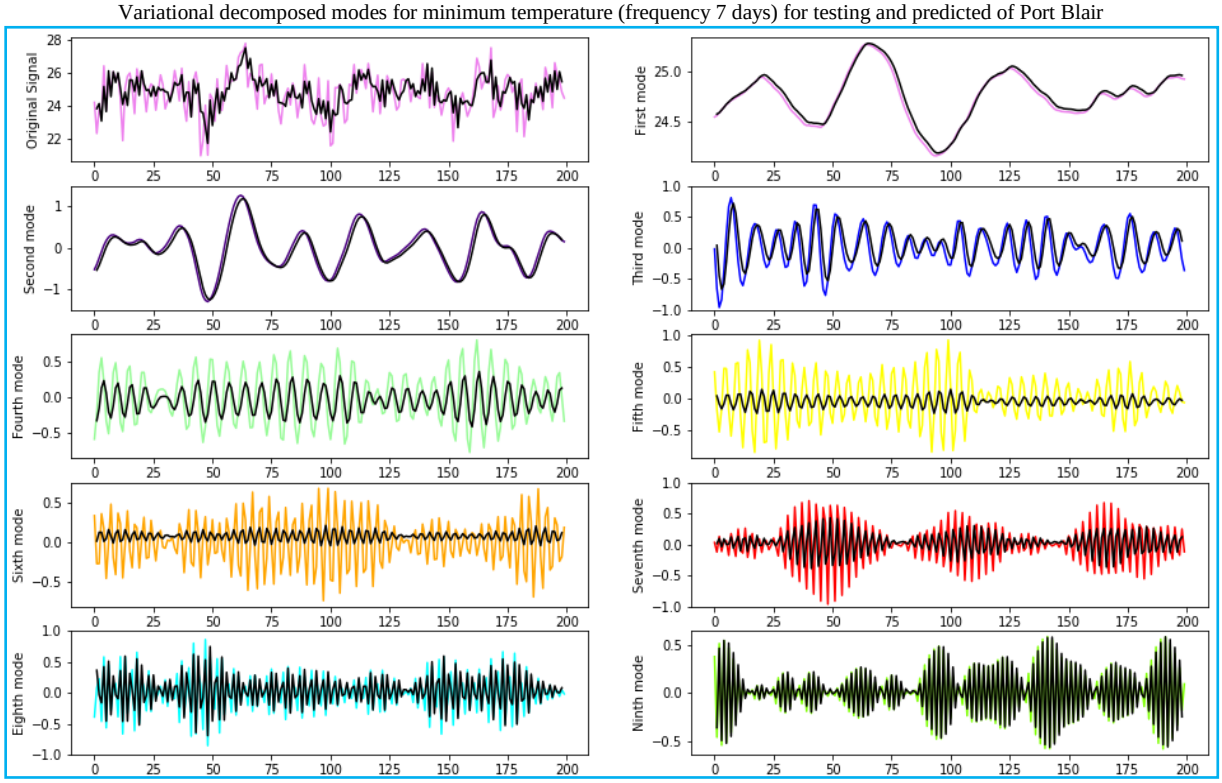

Fig. 16. Testing and predicted spectrum of the minimum temperature of 7 days interval with its decomposed mode of Port Blair

Original spectrum and variational decomposed modes for maximum temperature (frequency 9 days) of Port Blair

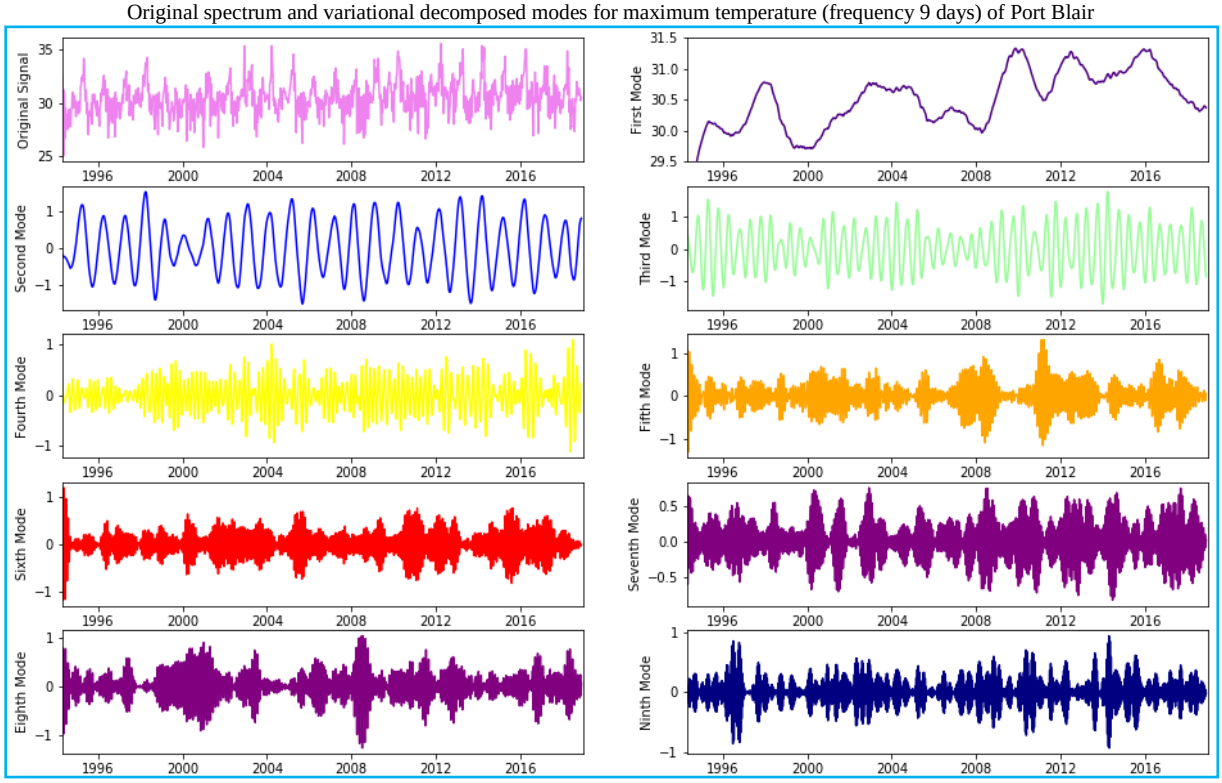

Fig. 17. Original maximum temperature spectrum of 9 days interval with its decomposed mode of Port Blair

Variational decomposed modes for maximum temperature (frequency 9 days) for testing of Port Blair

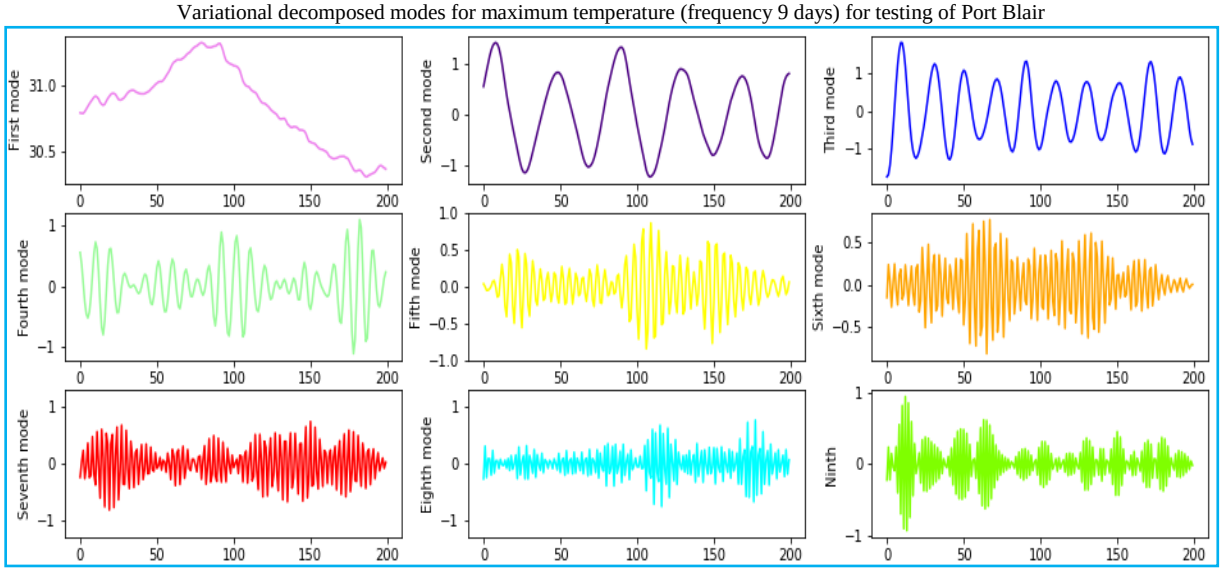

Fig. 18. The testing spectrum of the maximum temperature of 9 days interval with its decomposed mode of Port Blair

Variational decomposed modes for maximum temperature (frequency 9 days) for testing and predicted of Port Blair

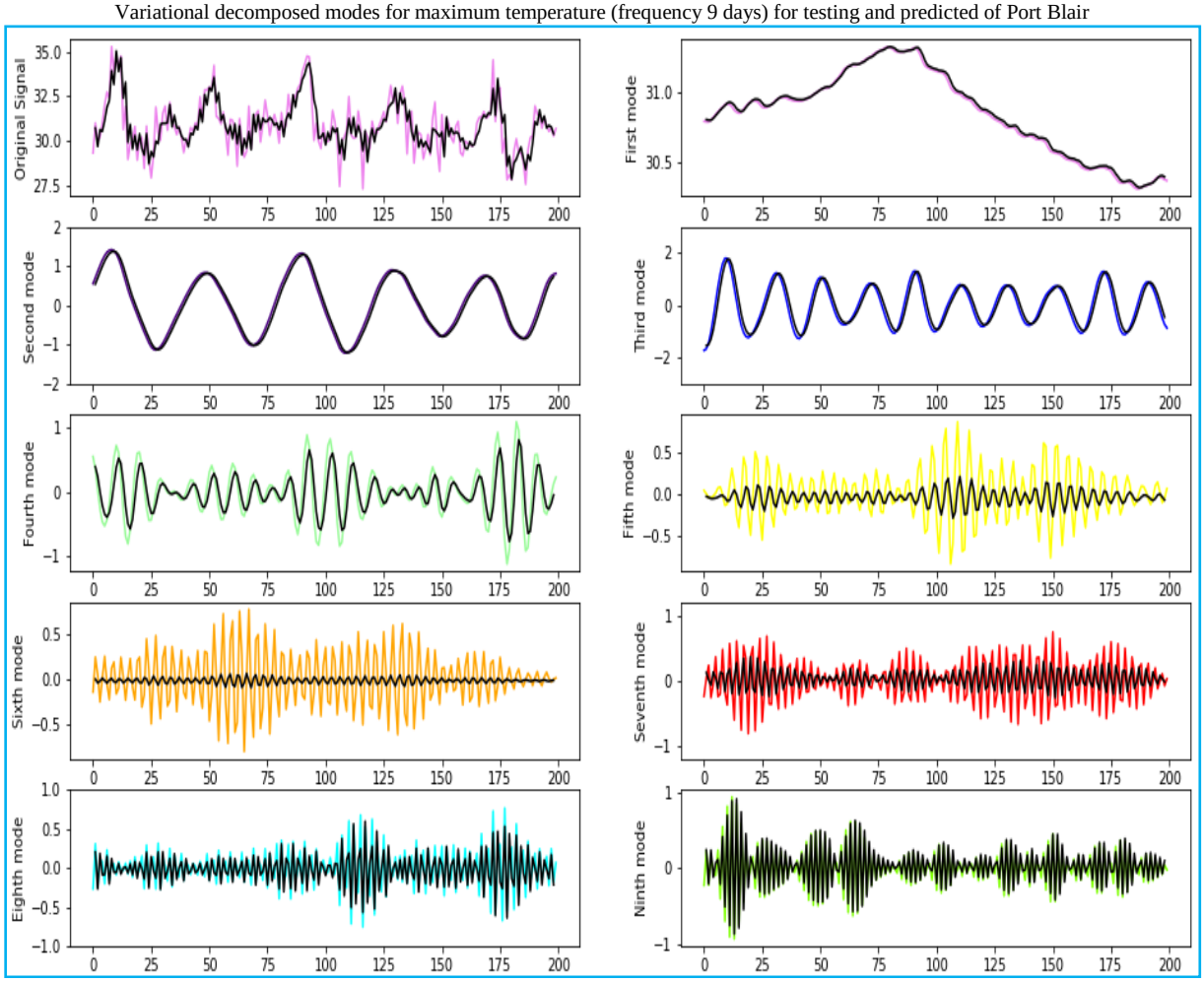

Fig. 19. Testing and predicted spectrum of the maximum temperature of 9 days interval with its decomposed mode of Port Blair

Original spectrum and variational decomposed modes for minimum temperature (frequency 9 days) of Port Blair

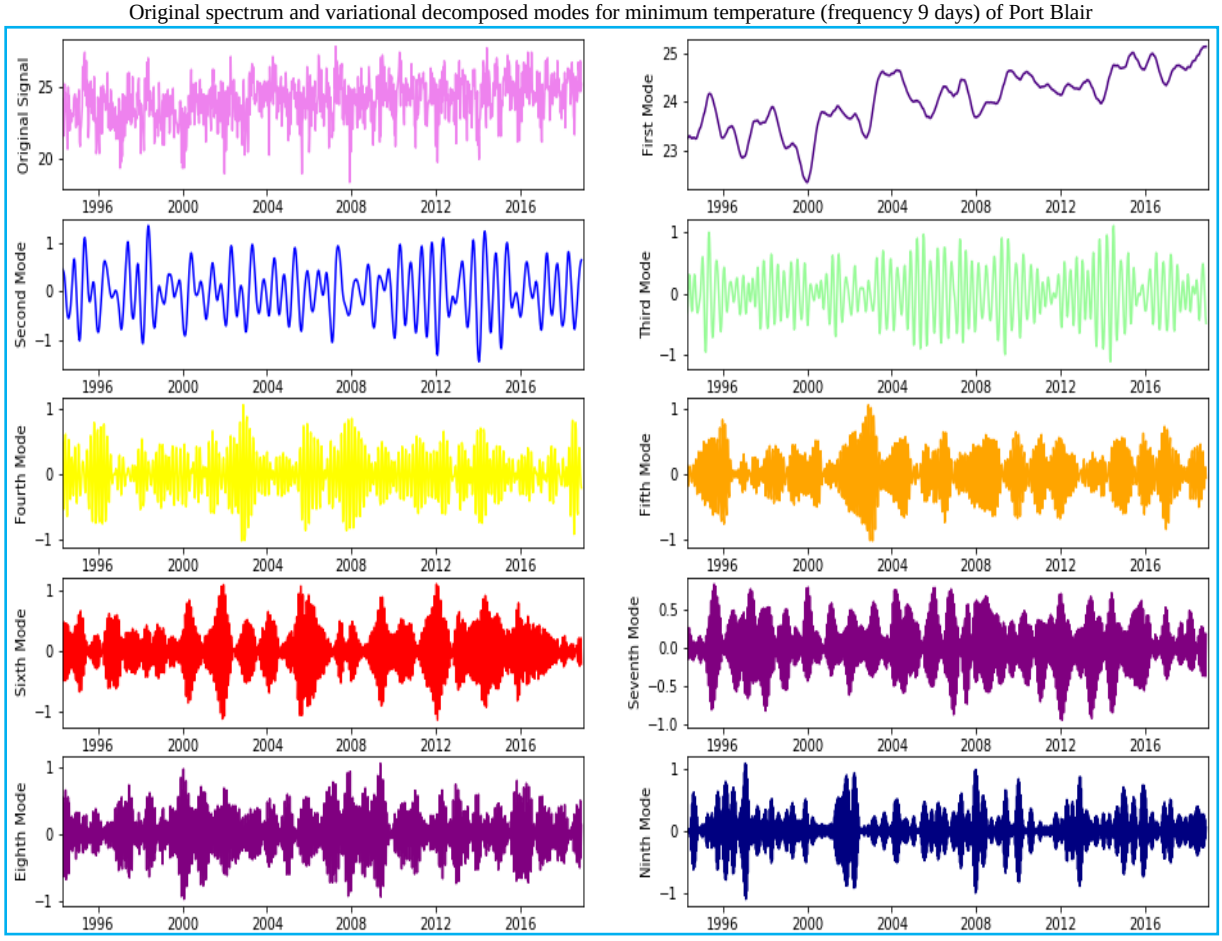

Fig. 20. Original minimum temperature spectrum of 9 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 9 days) for testing of Port Blair

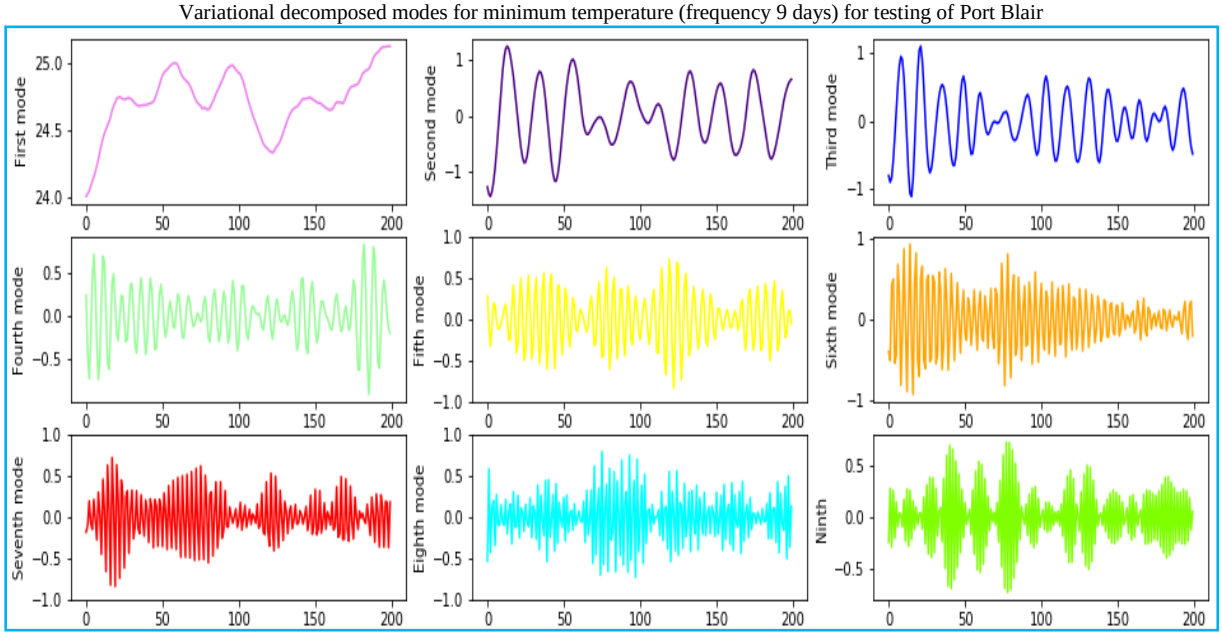

Fig. 21. Testing spectrum of minimum temperature of 9 days interval with its decomposed mode of Port Blair

Variational decomposed modes for minimum temperature (frequency 9 days) for testing and predicted of Port Blair

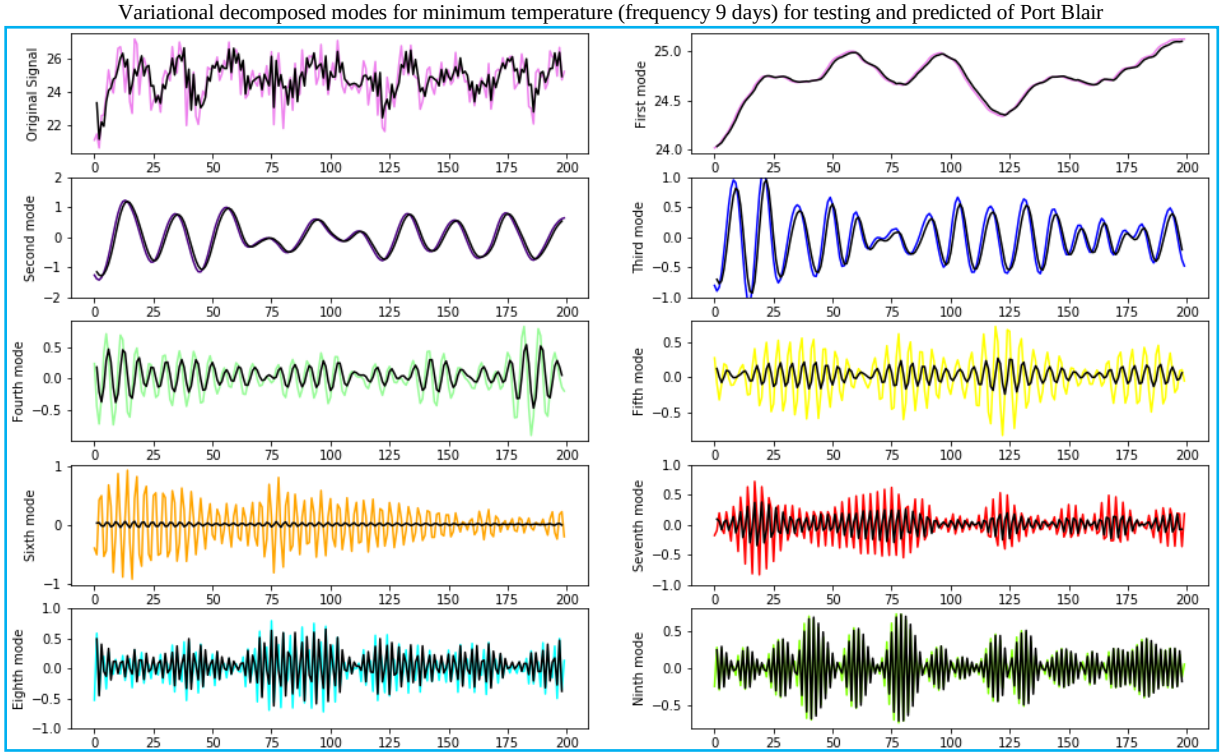

Fig. 22. Testing and predicted spectrum of the minimum temperature of 9 days interval with its decomposed mode of Port Blair

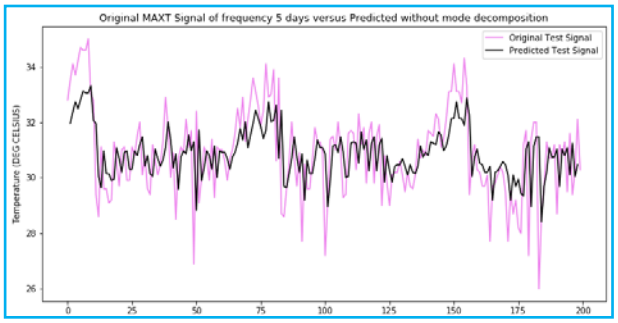

Fig. 23. Testing and predicted spectrum of the maximum temperature of 5 days interval of Port Blair without mode decomposition

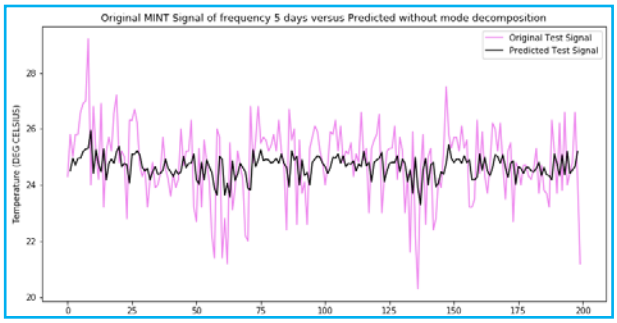

Fig. 24. Testing and predicted spectrum of the minimum temperature of 5 days interval of Port Blair without mode decomposition

where, N is the number of samples, $w(i)$ is the actual value, $\overline{w}(i)$ is the predicted value and $\overline{w}$ is the mean value.

4. Results and discussion

To validate the effectiveness of the proposed method for predicting medium-range maximum and minimum temperature, we have generated the time series spectrum (signal) from the temperature data and it is shown in Figs. 5, 8, 11, 14, 17 and 20. Then the signal is decomposed into 9 modes by the VMD technique. Here to decompose the signal by VMD technique we have used the python vmdpy module (https://pypi.org/project/vmdpy/). Here sample parameters kept for this decomposition are alpha = 1000, tau = 0, K = 9, DC = 0, init = 1, $tol = 1 \times 10^{-7}$. The decomposed modes are also shown in Figs. 5, 8, 11, 14, 17 and 20. After obtaining the decomposed modes, each mode is used as the input of the LSTM network. Here the used LSTM network is trained using a Python package called 'keras' on top of the Tensor flow backend. The network has a visible layer with 1 input, a hidden layer with 10 LSTM blocks and an output layer that makes a single value prediction. The network is trained for 30 epochs and a batch size of 1 is

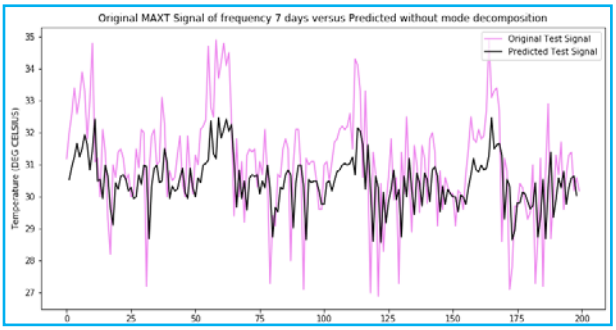

Fig. 25. Testing and predicted spectrum of the maximum temperature of 7 days interval of Port Blair without mode decomposition

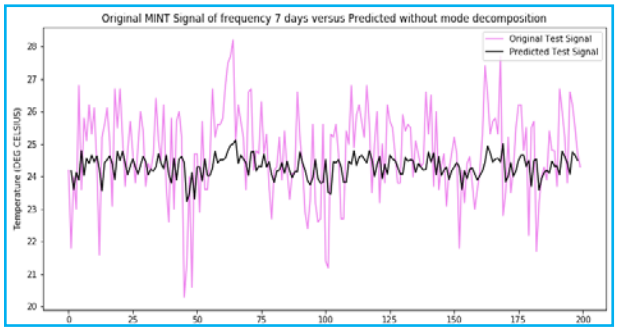

Fig. 26. Testing and predicted spectrum of the minimum temperature of 7 days interval of Port Blair without mode decomposition

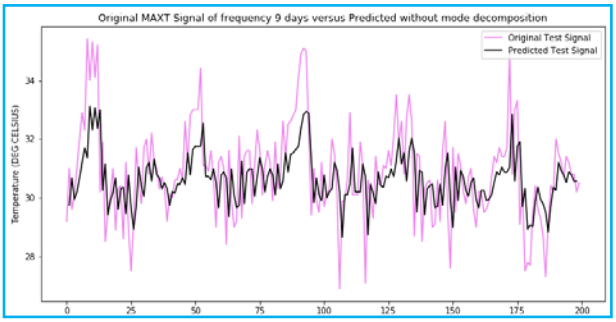

Fig. 27. Testing and predicted spectrum of the maximum temperature of 9 days interval of Port Blair without mode decomposition

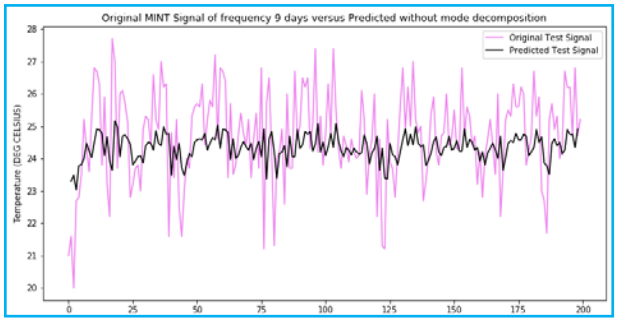

Fig. 28. Testing and predicted spectrum of the minimum temperature of 9 days interval of Port Blair without mode decomposition

TABLE 1

Error in predicting maximum temperature

Method	MAE			MSE			RRMSE			MAPE			SSE			R^2			IA			TIC		
	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days
Proposed	0.561	0.604	0.581	0.548	0.663	0.562	0.059	0.047	0.001	1.831	1.984	1.897	109.526	132.665	112.353	0.733	0.680	0.739	0.991	0.989	0.991	0.012	0.013	0.012
Only LSTM	1.022	1.235	1.102	1.834	2.355	1.943	0.012	0.230	0.094	3.348	3.977	3.558	366.736	471.069	388.582	0.262	0.081	0.236	0.971	0.963	0.969	0.022	0.025	0.023

TABLE 2

Error in predicting minimum temperature

Method	MAE			MSE			RRMSE			MAPE			SSE			R^2			IA			TIC		
	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days	5 days	7 days	9 days
Proposed	0.665	0.687	0.678	1.072	0.969	1.104	0.183	0.100	0.109	2.751	2.824	2.801	214.312	193.711	220.822	0.229	0.337	0.289	0.979	0.981	0.978	0.021	0.020	0.021
Only LSTM	1.119	1.201	1.185	2.298	2.318	2.473	0.007	0.171	0.130	4.579	4.857	4.822	459.526	463.518	494.607	-0.263	-0.204	-0.223	0.954	0.954	0.950	0.031	0.031	0.032

used. The activation function is tan h, the optimizer used is Adam. Here the input is split into train and test by 80% and 20% respectively. Figs. 6, 9, 12, 15, 18 and 21 show the decomposed modes for testing. These modes are used for testing the mentioned LSTM network to predict max-min temperature signals for 5, 7 and 9 days. After running

the LSTM network for each of the testing modes the network gives us the predicted modes. These predicted modes are combined to obtain the predicted max, min temperature signal. Figs. 7, 10, 13, 16, 19 and 22 show the test signal and predicted signal along with test and predicted modes. From these figures we can visually compare the test and predicted signals. It can be seen from Figs. 7, 10, 13, 16, 19 and 22 that the test and predicted signal closely match. But to measure the actual error and degree agreement with the original we have calculated MAE, MSE, RRMSE, MAPE, SSE, R^2, IA and TIC, details of which are given in Section 3. We have also predicted the max, min temperature signal without going through VMD. We have directly used the max, min signal as input to the LSTM to predict the signals. The predicted signal by LSTM along with original signals is shown in Figs. 23-28. We have also calculated MAE, MSE, RRMSE, MAPE, SSE, R^2, IA and TIC for the signal with only LSTM and then compared with our proposed approach and the results are shown in Tables 1-2. It is clear from the Tables that MAE, MSE, RRMSE, MAPE, SSE for the proposed approach is far better than those of only LSTM. The value of R^2, IA for the proposed approach is always higher than only the LSTM predicted signal.

5. Conclusions

This study aimed to predict medium-range temperature with appreciable accuracy. For this, we have used 5, 7 and 9 days max-min temperature data. From the results obtained in section 4, it is seen that the proposed approach has MAE, MSE, RRMSE and MAPE below 0.7, 1.1, 0.2 and 2.85 respectively in predicting max-min temperature 5, 7 and 9 days ahead, which is very less. The values of IA are always above 0.97 in this approach. The performance of only LSTM to predict the temperature is poor compared to the proposed approach. Hence it may be concluded that the proposed approach is suitable and may be used to predict medium range temperature.

Acknowledgments

Authors acknowledge IMD for encouraging and supporting to carry out this work and RMC Kolkata for supplying meteorological data.

Disclaimer : The contents and views expressed in this study are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Abdel-Aal, RE., 2004, "Hourly temperature forecasting using abductive networks", *Engineering Applications of Artificial Intelligence*, **17**, 543-556**.**

Alzahrani, A., Kimball, J. W. and Dagli, C., 2014, "Predicting solar irradiance using time series neural networks", *Proc Comput Sci*, **36(c)**, 623-628.

Dash, Y., Mishra, S. K. and Panigrahi, B. K., 2017, "Rainfall prediction of maritime state (Kerala), India using SLFN and ELM techniques", In the Proceedings of the 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 6-7 July 2017, 1714-1718.

Dombayc, O. A. and Gc, M., 2009, "Daily means ambient temperature prediction using artificial neural networks method : a case study of Turkey", *Renewable Energy*, **34**, 1158-1161.

Dragomiretskiy, K. and Zosso, D., 2014, "Vaiational mode decomposition", *IEEE Transactions on Signal Processing*, **62**, 3, 531-544.

https://doi.org/10.1145/3318299.3318381.

Khosravi, A., Koury, R. N. N., Machado, L. and Pabon, J. J. G., 2018, "Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms", *Journal of Cleaner Production*, **176**, 63-75.

Lin, H., Hua, Y., Ma, L. and Chen, L., 2019, "Application of convLSTM network in numerical temperature prediction interpretation" ICMLC' 19, February 22-24, 2019, Zhuai, China.

Mislan, H., Hardwinarto, S. and Sumaryono, M. A., 2015, "Rainfall monthly prediction based on artificial neural network: A case study in Tenggarong station, East Kaliamantan, Indonesia", *Procedia Comput Sci.*, **59**, 142-151.

Paniagua-Tineo, A., Salcedo-Sanz, S., Casanova-Mateo, C., Ortiz-García, E. G., Cony, M. A. and Hernández-Martín, E., 2011, "Prediction of daily maximum temperature using a support vector regression algorithm", *Renewable Energy*, **36**, 3054-3060.

Patil, K., Deo, M. C. and Ravichandran, M., August 2016, "Prediction of Sea surface temperature by combining numerical and neural techniques", *Journal of Atmospheric and Oceanic Technology*, **33**, 1715-1726.

Poornima, S. and Pushpalata, M., 2019, "Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units", *Atmosphere*, **10**, 668. https://doi:10.3390/atmos10110668.

Radhika, Y. and Shashi, M., April 2009, "Atmospheric temperature prediction using support vector machines", *International Journal of Computer Theory and Engineering*, **1**, 1, 1793-8201. https://.doi.org/10.7763/IJCTE.2009.V1.9.

Salcedo-Sanz, S., Deo, R.C., Carro-Calvo, L. and Saavedra-Moreno, B., 2015, "Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms", *Theor. Appl. Climatol*. https://doi.org/10.1007/s00704-015-1480-4.

Tian, Z., Li, S. and Wang, Y., 2019, "A prediction approach using ensemble empirical mode decomposition-permutation entropy and regularized extreme learning machine for short-term wind speed prediction", 2019, *Wind Energy*, **2019**, 1-30.

Xu, L., Li, Y., Yu, J., Li, Q. and Shi, S., 2020, "Prediction of Sea surface temperature using a multiscale deep combination neural network", *Remote Sensing Letters*, **11**, 7, 611- 619.

Yu, Y., Si, X., Hu, C. and Zhang, J., 2019, "A review of recurrent neural networks: LSTM cells and network architectures", *Neural Computation*, **31**, 1235-1270.

Zhang, Y., Pan, G., Chen, B., Han, J., Zhao, Y. and Zhang, C., 2020, "Short-term wind speed prediction model based on GA-ANN improved by VMD", *Renewable Energy*, **156**, 1373-1388.

Zhang, Z. and Dong, Y., 2020, "Temperature forecasting via convolution recurrent neural networks based on time series data", *Complexity*, **Vol 2020**, article id 3536572. https://doi.org/10.1155/2020/3536572.