

Temporal Analysis Of Temperature Variability In Pakistan Using The Method Of Empirical Mode Decomposition



Mariam Sultana^{1,2}, Faheem Zaidi¹, Javed K.Afridi¹, Muhammad Shoaib³

¹Department of Mathematical Sciences, Federal Urdu University, Karachi

²mariam.sultana@fuuast.edu.pk

³Department of Physics, Federal Urdu University, Karachi

Abstract

Pakistan is a country with altitudes that ranges from sea level to the world's second-highest mountain peak in the world. This distinctive characteristic renders a significant variation in the countries' climate, for instance, temperature differences and spatial distribution of rainfall. These variations in the corresponding temperature and precipitation are best described and predicted by well-defined classifications of climatic variables and their corresponding spatial distributions across Pakistan. This paper demonstrates the techniques that allow us to predict the average temperature of Pakistan using empirical model decomposing with auto regressive integrated moving average (EMD-ARIMA) and empirical model decomposing with auto regressive integrated moving average with neural networks (EMD-ARIMA-NN) models. This research was applied to a lengthy series of average monthly temperature (report in °C) of Pakistan from January 1901 to December 2016 with 1392 data points. The precision of these models is checked with the avail of statistical analysis and error tests. Comparative analysis shows that the empirical model decomposing with auto regressive integrated moving average with neural networks has a more preponderant forecast.

Key words: Empirical Mode Decomposition (EMD), Auto-Regressive Integrated Moving Average (ARIMA), Neural Network (NN), Average temperature, Global Warming

1. INTRODUCTION

The climate of Pakistan is comparatively temperate and varies along with its topography which is generally hot and dry around the coastal areas and all along the lowland plain of the Indus River. However, it becomes increasingly cooler in the northern parts of the country near the Himalayas. Generally speaking, four-season is acknowledged: (a) the season from March to May, which is hot and dry spring; (b) the season of rainy summer, generally known as the southwest monsoon, which runs through June to September; (c) the season from October to November, also known as the retreating monsoon; and finally, (d) the season of dry and cool winter which runs

through December till February. Besides the Northern parts of the country, where monsoon can cause around 200 mm rainfall per month from July to September, the rest of the country is predominantly less susceptible to receive rainfall. However, sequential patterns of droughts and floods are often the outcome of significantly varying inter-annual rainfall. In correlation with the El Niño cycle, anomalous behaviors are commonly observed corresponding to the frequencies in floods and temperature and their climatic impacts. In this connection, the surface temperature at the global level has been seen as an important indicator for explaining and predicting climatic changes at the global level such that they may be used for making well-estimated forecasts for helping form the sound foundation for planning the environmental public policies.

There have been recent research interests for the examination of patterns and trends in temperature and precipitation data and records. In this context, authors in [1] conducted their research and found that a range of negative diurnal temperature was present as a trend in South Asia. In [2], authors have shown that extreme events with the highest and lowest temperatures are rapidly increasing in the country throughout the years 1965 till 2009. In [3], authors have concluded that a monthly, seasonal and annual pattern of increasing (i.e. positive) and decreasing (i.e. negative) precipitation is emerging. They reached the conclusion after a careful analysis of changeability in precipitation through 15 different locations of Pakistan such that the analysis of the data involved spans over the period of 51-years (1961-2011). Research with similar goals and methodology was also conducted by authors in [4], where the country analyzed was Iran in a similar context of highest precipitation over the years 1951-2007. There also, it was concluded in similar manner that a strong similar pattern of increase was present. This evidence and others are enough to make the case for the claim that the world is overall facing a similar upward trend in climatic extremities: such as temperature and precipitations etc. On the other hand, the model of SARIMA has also been used in the same context for forecasting the highest temperature in Karachi [5]. There have been other researchers as well, who have used similar methods and models for the analysis of temperature [6].

In this research, Auto-Regressive Integrated Moving Average (ARIMA), the Empirical Mode Decomposition (EMD), and the method of Neural Network (NN) has been used to analyze the temperature changes and processes inasmuch as they help obtain valuable information depending upon the corresponding historical data in Pakistan. In section 2, a description of the data along with an introduction to the methodological and theoretical aspects are presented with techniques of performance evaluation. In section 3, new proposed hybrid method is given. Rigorous discussion and results are presented in section 4. Section 5, which is also the final section, presents the conclusion.

2. DATA AND THEORETICAL ASPECT OF FORECASTING MODELS

The monthly average temperature of Pakistan data from January 1901 to December 2016 has been collected from World Bank Group, Climate change knowledge portal (<https://data.worldbank.org/country/pakistan>). As far as the geographical grid is concerned, Pakistan's coordinates correspond to 60 degrees 50 minutes to 77 degrees 50 minutes with Eastern longitudinal range, and 23 degrees 35 minutes to 37 degrees 05 minutes Northern latitudinal range. These 116 years of long series of data consists of 1392 data points of average monthly temperature (report in °C)

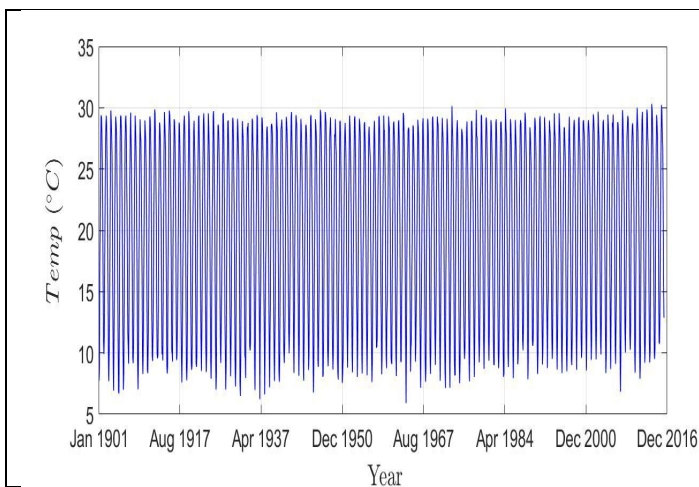


Figure 1: Monthly Average Temperature of Pakistan from Jan 1901-Dec 2016.

2.1 Empirical Mode Decomposition (EMD):

The Empirical Mode Decomposition algorithm was proposed and pioneered by Huang [11] as a method of signal transformation that was non-linear. The goal was to express a decomposition of a time series of some non-stationary kind into a sum of intrinsic mode function (or simply IMF). It involved the construction of smooth envelopes given by local minima and local maxima corresponding to a sequence. Furthermore, it was based upon subsequently subtracting the mean of these envelopes from the corresponding primary sequence. These local minima and local maxima, which produce both upper and lower envelop and are found to be attached by cubic splines are

ultimately considered and described by this method EMD.

From the initial sequence, we then subtract the mean so produced by the above two envelopes. We thus obtain the required first-order approximation with the empirical function. Obtaining IMF as the extraction from the EMD, one notes that the following conditions are necessarily supposed to be satisfied.

1. The number of zero-crossings must be either equal to the number of IMF minima and maxima; or if not equal, then the difference must be less than or equal to 1.
2. The mean value of the envelopes given by the local extrema must remain zero at every point of an IMF.

After subtracting the IMF, which is previously extracted, from the original signal, the second IMF is obtained as a result. This can be repeated continuously till all required IMF's are calculated. This procedure stops as soon as there remains no residue containing more than two extrema. This is how that final IMF is found or calculated by applying the same operation on the residue signal till the satisfaction of IMF properties. Following is a comprehensive summary of the overall algorithm.

1. Lower and upper envelopes are created with $\sigma(t)$, i.e. the given time-series signal, by interpolation of cubic-spline of the corresponding local extrema (i.e. local minima and local maxima)
2. Computation of the mean-value, i.e. ϵ_1 , of the lower and upper envelopes
3. First component $\gamma(t) = \sigma(t) + \epsilon(t)$ is obtained by making the subtraction of the mean from the original time-series.
4. Repetition of above three steps (in the same order) with taking $\gamma(t)$ as new $\sigma(t)$, up to the point where one of the above-mentioned terminating conditions is met.
5. Calculation of residual signal given by the formula: $\vartheta(t) = \sigma(t) - \gamma(t)$; by taking $\gamma(t)$ as an IMF.
6. Repetition of the above five steps (in the same order) up to the point where all of the IMF's are found by taking $\vartheta(t)$ (i.e. the residue-signal) as a new $\sigma(t)$.

2.2 Autoregressive Model Integrated Moving Average (ARIMA) :

On the basis of past patterns and behaviors, the autoregressive (AR) model helps in predicting future patterns and behaviors such that the modeling can be used in forecasting the preceding and succeeding values that correlate with the time serial values. This gives AR process as an exemplification of the stochastic process such that it possesses an in-built capacity for modeling uncertainty and randomness.

If "p" denotes the order, then the model of AR(p) captures predictor variables in the form of α_i as specific lagged values.

These lags correspond to the results from one particular period of time to its effects in the following periods of time. For instance, following the preceding notations, the "first-order autoregressive process" would be denoted by AR(1). The time periods which correspond to "one unit-value apart" (i.e. with

variable-values at (t-1)) are related with outcome variable at a particular point in time “t” under this process of the first order; i.e. AR(1).

In Autoregression models, the shock and the finite linear combination involving preceded values of time, express the current value of time as follows:

$$\alpha_t = \beta_1\alpha_{t-1} + \beta_2\alpha_{t-2} + \dots + \beta_p\alpha_{t-p} + \mu_t$$

(1) with $\beta_1, \beta_2, \beta_3, \dots, \beta_p$ denoting the auto-regression parameters, p is the model order with white noise μ_t . If we let $\tilde{\chi} = \chi - \mu$ denote the deviations, then it admits the following expression:

$$\tilde{\chi}_t = \phi_1\tilde{\chi}_{t-1} + \phi_2\tilde{\chi}_{t-2} + \phi_3\tilde{\chi}_{t-3} + \dots + \phi_p\tilde{\chi}_{t-p} + a_t$$

(2) with $\mu, \phi_1, \phi_2, \phi_3, \dots, \phi_p, a_t$ corresponding to the unknown parameters which observation data can help approximately. Then, after an application of autoregressive operator, we obtain:

$$\phi(B) = 1 - \phi_1B - \phi_2B^2 - \phi_3B^3 - \dots - \phi_pB^p$$

(3) After straightforward manipulations, the autoregressive model admits the following concise expression:

$$\phi(B)\tilde{\chi}_t = a_t$$

(4) The Autoregressive (AR) models can be efficiently coupled with moving average (MA) models to form a more generalized model of the Box and Jenkins family i.e. autoregressive moving average (ARMA) models. As authors in [7] noted already, ARMA models involve expressions of linear aggregates of previous values denoted by ‘p’ and the weighted sum of ‘q’ previous deviations (obtained by subtracting the fitted value of previous data from the original value) plus a random parameter. However, ARMA models deal with stationary data only. If the data is non-stationary, it can be transformed into stationary by employing a commonly used logarithm transformation or standardization procedure, and then they could be modeled using ARMA. Another class of models called ARIMA models can be extended to nonstationary data also, which involves differencing of time series to attain stationarity. Moreover, ARMA (p,q) is the general case of ARIMA (p,d,q) where the differencing term ‘d’ is zero.

2.3 Artificial Neural Network (ANN)

Recognition and classification of features and patterns extraction corresponding to different fields have already been involving the Artificial Neural Networks (ANN), see [8]. The models of ANN are also increasingly utilized for predicting financial market trends [9]. These applications and utilizations are a necessary consequence of realizing the fact that these models have the capacity to generalize and learn from past historical or temporal events to identify future patterns and behaviors as events yet unseen [10]. An ANN model is a very efficient network of simple processing units which are well-interconnected. These units are called neurons. They resemble human biological neurons, present in a human brain, with

similar properties. These neurons from the ANN model form groups, also known as strata (or layers). There is no bonding between neurons within the same strata, however, they do form bonds with strata that are adjacent to them. In order to describe the intensity with which two neurons are associated in strata that are adjacent, the idea of ‘weight’ or ‘connection strength’ is used. There are generally three layers or strata in an ANN: an output stratum, an input stratum, and a hidden stratum. The output signal flow (denoted by ‘O’ below) is produced by the corresponding inputs given as $x_1, x_2, x_3 \dots x_n$ such that the production from the inputs is unidirectional. This phenomenon is mathematically captured by the following equation:

$$O = f(net) = f(\sum_{j=1}^n w_j x_j)$$

(5) with $f(net)$ corresponds to the activation function and w_j corresponds to the weight-vector. The scalar product of the input vectors and the weighted-vectors may be seen as describing the variable-net given by:

$$net = w^T x = w_1x_1 + w_2x_2 + \dots + w_nx_n$$

(6)

Here, T denotes the transpose of a matrix. This helps us define the following formula for the computation of the output O as:

$$O = f(net) = \begin{cases} 1 & \text{if } w^T x \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

(7) Here θ denotes the critical or threshold level. Given such a node, it is known as a unit of the linear threshold. On the other hand, the model of the internal activity of Neurons is expressed as follows:

$$v_k = \sum_{j=1}^p w_{kj}x_j$$

(8)

Here, y_k denotes the output as the outcome of the Neuron corresponding to some activation function applied on the value v_k

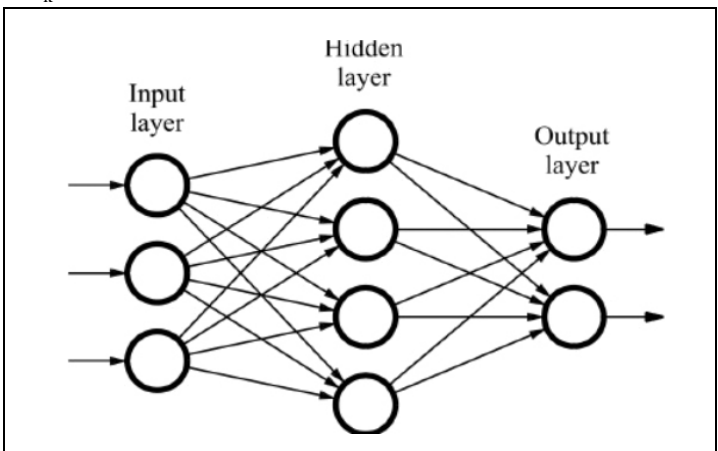


Figure 2:The basic structure of the neural network. In this basic structure, four hidden layers are used with three input and two output

2.4 Performance Estimation

We have used two methods that test the accuracy of the results. While testing the accuracy, our primary concern is the MSE (Mean Square Error). If by “correctness”, we mean the extent to which the computed values agree with the “true” physical values, then considerations about the accuracy or precision of the overall scheme become the assessment of its correctness. This is especially true when we assume the stability of the computation scheme. These methods are the Mean Absolute Error (or MAE) and Mean Square Error (or MSE). They respectively admit the following mathematical formulations:

$$Error_{MA} = \frac{1}{N} \sum_1^N |W_{real} - W_{forecasted}|$$

$$Error_{MS} = \frac{\sum_1^N (W_{real} - W_{forecasted})^2}{N}$$

where N is the number of data points and $W_{forecasted}$ is the forecasted price corresponding to the real price W_{real} .

3. PROPOSED METHODOLOGY

A very efficient way of handling a problem is by breaking down that problem into several sub-problems recursively such that these sub-problems are allowed to be either of the same types of the original problem or they can be seen as related to the original problem. This recursive process of breaking down can be repeated until sub-problems become simple enough so that they could be directly solved by some well-known efficient technique. A divide and conquer algorithm is exactly what helps achieve this recursive process of breaking down a problem into simpler manageable problems.

In the method proposed above, EMD helps decompose the temperature data into trend and several IMF's. To each IMF, an ARIMA is applied, whereas polynomial is applied on trend. The final prediction is obtained by either simply summing up or aggregating the prediction results by a single learning machine. Figure 3 shows the general schema of this ensemble method. This corresponds to the completion of the task of forecasting for all subseries. Following is a comprehensive summary of this overall process.

1. By using EMD, we decomposed the time-series data into a trend and several IMF's.
2. For every IMF extracted, the ARIMA(p, d, q) is trained, the selection of 'd' has done by using unit root test, for this study we used the Phillips–Perron test. In the case of AR and MA, the making of precise estimations requires optimal order selection. On contrary, lower-order modeling gives imprecise estimations on one hand, and higher-order generally leads towards the modeling of noise. In order to achieve the required selection of optimal order, several methods have been proposed. For the selection of the

value of parameters p and q, the Bayesian information criterion (BIC) is used.

3. Then by approximating the polynomial method, prediction about the trend can be made more efficiently.
4. By using the optimal order of ARIMA and polynomial selected in step 2 and 3 respectively, we forecast every IMF and trend one step ahead.
5. At the final step, a formulation of a time-series collective output is made by combining all forecasted results either simple addition or through the neural network.

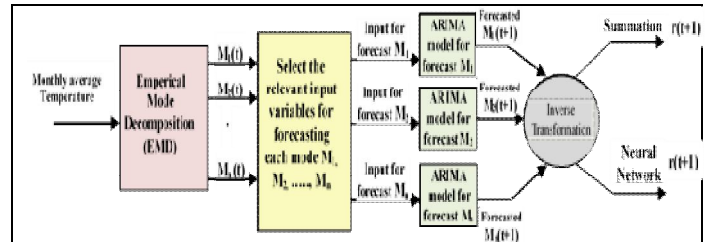


Figure 3: The general schema of this ensemble method. Monthly mean data of temperature of Pakistan is decomposed into IMF's and residual by EMD then these are used as input for forecasting by ARIMA and then again sum up with the help of summation and neural network after inverse transformation.

4. DISCUSSION AND RESULTS

By applying empirical model decomposition on the dataset of the average temperature of Pakistan, it decomposes the data set into 7 IFMs and residual (trend). Figure 4 shows the seven IFMs and residual of the decompose data of average temperature of Pakistan.

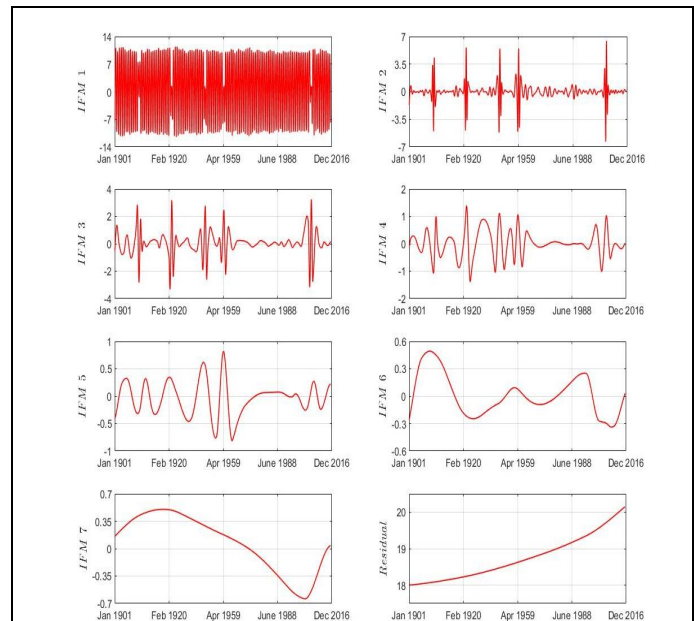


Figure 4: IFMs and residual of the decompose data of average temperature of Pakistan using empirical mode decomposition (EMD).

For every IFMs we used ARIMA(p, d, q) model to predict the one step ahead forecast, first we used the Phillips–Perron test to check the stationarity of each IFM. Table 1 shows the results of the P-P test, where D(IMF6) and D(IMF7,2) indicate the difference of IFM6 and double difference of IFM7 series respectively.

The series IFM1 to IFM5 are stationary as the test statistics of these are below the value of the test critical value for the 1% level. However, the IFM6 and IFM7 series are not stationary, although the transformation of difference of IFM6 and double difference of the IFM7 made them stationary. This allows us to apply the ARMA model to these stationary datasets.

To find the best ARMA lags, we fitted several models with different lag choices. We fitted all combinations of AR(1) to AR(4) and MA(1) to MA(4) (total of sixteen models) and computed the loglikelihood function for each model. As the number of ARMA's coefficients increases, it increases the value of the loglikelihood function. The BIC does not depend only on the value of the loglikelihood, it also pays the penalty of adding more parameters in the model. We calculated the BIC for each selected model by using the loglikelihood function and the number of parameters for each model. Appendix A shows the comparison of sixteen models for each stationary dataset, we selected the model for each stationary dataset that has minimum BIC. The tables show the comparison of BIC for different ARMA lag models, we highlighted the best ARMA lag models among them. Table 2 shows the coefficients and their t statistics of the selected ARMA model.

	Test Statistic	Prob.*	Test critical values:		
			1% level	5% level	10% level
IFM 1	-6.0182	0.000	-2.566	-1.94	-1.616
IFM 2	-5.1193	0.000			
IFM 3	-6.8289	0.000			
IFM 4	-5.7966	0.000			
IFM 5	-4.1237	0.000			
IFM 6	-1.7717	0.073			
IFM 7	-0.6335	0.443			
D(IFM6)	-2.6348	0.0082			
D(IFM7)	-1.5598	0.1118			
D(IFM7,2)	-20.104	0			

Table 1: Results of Phillips–Perron test of all IFMs, and difference and double difference of some IFMs.

	IFM 1	IFM 2	IFM 3	IFM 4	IFM 5	D(IFM 6)	D(IFM 7,2)
C	1.1341 (14.37)	0.0170 (0.93)	0.0478 (0.65)	0.0074 (0.13)	-0.0417 (-1.58)	0.0060 (1.11)	0.0000 (-0.17)
AR(1)	1.7321 (35988.58)	2.9151 (174.45)	2.6802 (233.59)	2.8409 (287.02)	3.9364 (29286475)	0.9997 (2285.32)	1.9952 (2795.91)
AR(2)	-1.0000 (-80002)	-3.3992 (-97.31)	-2.4602 (-112.60)	-2.7026 (-137.55)	-5.8140 (-2798442)	-	-0.9955 (-1176.9)
AR(3)	-	1.8495 (58.60)	0.7689 (68.53)	0.8605 (86.67)	3.8186 (978811)	-	-
AR(4)	-	-0.3908 (-32.00)	-	-	-0.9410 (-497859)	-	-
MA(1)	-0.9927 (-23.57)	-0.1317 (-7.29)	0.9085 (50.96)	1.0567 (63.00)	0.0338 (24.29)	2.4074 (335.19)	0.0024 (0.08)
MA(2)	0.0658 (1.80)	-0.7486 (-43.67)	0.0108 (0.63)	0.0955 (6.35)	-0.8514 (-46.11)	2.6448 (155.36)	-0.9910 (-38.75)
MA(3)	0.1467 (3.80)	-	-	-	-0.0584 (-3.50)	1.7940 (92.89)	-
MA(4)	0.3411 (12.54)	-	-	-	-0.0945 (-5.32)	0.5644 (56.22)	-

Table 2: The coefficients and associated t-statistics for ARMA model for each IFMs selected based on Bayesian information criterion.

We have used two strategies to predict the average monthly temperature using predicted IFMs and trend. First, we compose back to the temperature by simple adding the IFMs and trend and secondly by using the neural networks. Figure 5 shows the comparison of the actual average monthly temperature of Pakistan with the prediction of these two strategies. Table 3 shows the comparison of models discussed above using MSE and MAE. These results show that the hybrid empirical mode decomposition with autoregressive integrated moving average predict more precisely the trends and details if we integrated it further with neural networks as the MSE and MAE are comparatively quite minimum than other methods, which is why they are made with very high precision.

depending upon the corresponding historical data in Pakistan. Techniques of performance evaluation along with rigorous discussion and results are presented which clearly show that EMDARNN has the least error. It is thus claimed that the method proposed (EMDARNN) in this paper is comparatively more reliable for both analysis and forecasting temperature trends.

REFERENCES

1. Frei, C and Schar, C.,(2001).Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region, *Journal of Climate*,14, 1568–1584
2. Tank, A.M.G. K., Peterson, T.C. , QuadirD.A.,Dorji,S., et al(2005). New and T.Spektorman, Changes in Daily Temperature and Precipitation Extreme in Central & South Asia, *Journal of Geophysical Research*, 1-18
3. Gwangyong, C., Collins,D., Ren,G.,Trewin,B.,(2009), Changes in means and extreme events of temperature and precipitation in the Asia-Pacific Network region, 1955–2007, *International Journal of Climatology*,29(13),1906-1925.
4. Zhang, F, Gao, H., and Cui, X.,(2008), Frequency of ExtremeHigh Temperature Days in China, 1961-2003, *Weather*,63(2), 46-49
5. Burney,A.,Mushtaq, A.K.B, and Steve E J.,(2017) Forecasting Monthly Maximum Temperature of Karachi City using Time Series Analysis, *Pakistan Journal of Engineering Technology and Science*, 7(2),125-135
6. Maida, Z., and Ghulam,R.,(2011) Frequency of extreme temperature & precipitation events in Pakistan 1965-2009, *Sci.Int. (Lahore)*, 23(4),313-319
7. Mishra,A.K.&Desai,V.R.,(2005), Drought forecasting using stochastic models, *Stoch Environ Res Ris Assess* 19, 326–339
8. Widrow, B., Rumelhart, D.E., Lehr, M.A., (1994). Neural networks: Applications in industry, business and science. *Communications of the ACM* 37 (3), 93–105
9. Sharda, R., (1994). Neural networks for the MS/OR analyst: An application bibliography. *Interfaces* 24 (2), 116–130.
10. Kecman, V., (2001), *Learning and soft computing*. The MIT Press Cambridge, Massachusetts London, England.
11. Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, et al,(1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis *Proc. R. Soc. Lond. A*.454(1971),903–995.

Errors	EMD+ARIMA	EMD+ARIMA+NN
MSE	1.8646	1.1858
MAE	0.0628	0.0529

Table 3: Error analysis of mean temperature of Pakistan; prediction by the different forecasted method

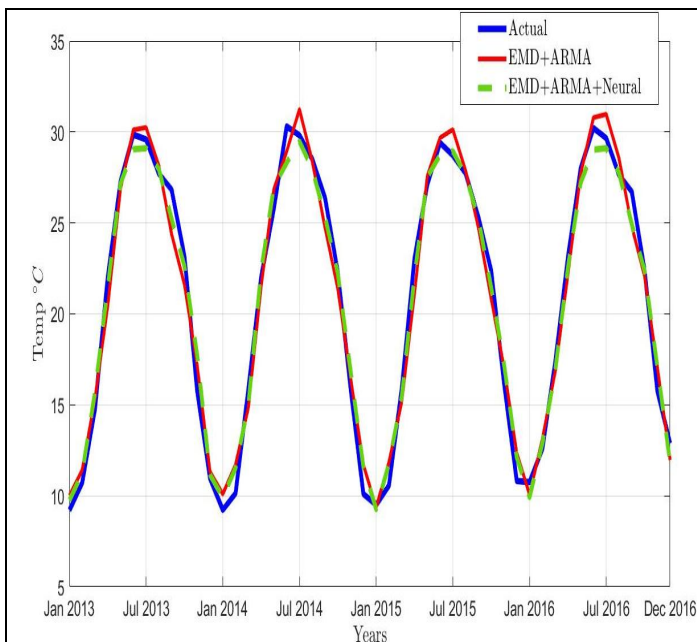


Figure 5: Comparison of the actual temperature with predicted results through EMD-ARIMA and EMD-ARIMA-NN for last three years of data set.

5. CONCLUSION

The present investigation involved 116 years of long series of data consists of 1392 data points of the average monthly temperature of Pakistan (report in °C). Auto-Regressive Integrated Moving Average (ARIMA), the Empirical Mode Decomposition (EMD), and the method of Neural Network (NN) have been used to analyze the temperature changes and processes inasmuch as they help obtain valuable information

Appendix A

Bayesian information criterion of various ARMA models for the selection of Auto regressive and moving average lags.

IFM 1		MA			
		1	2	3	4
AR	1	6877	6488	6270	6156
	2	5184	5190	5029	4889
	3	5190	5191	5197	4896
	4	5192	5198	5201	5209

IFM 2		MA			
		1	2	3	4
AR	1	193	-969	-1408	-1515
	2	-2048	-2193	-2186	-2184
	3	-2201	-2199	-2100	-2198
	4	-2205	-2235	2228	-2211

IFM 3		MA			
		1	2	3	4
AR	1	-2672	-4235	-5008	-5322
	2	-6976	-7695	-7927	-8036
	3	-7152	-8166	8158	-8152
	4	-8099	-8160	-8151	-8144

IFM 4		MA			
		1	2	3	4
AR	1	-6608	-8375	-9086	-9492
	2	-12868	-13854	-14401	-14609
	3	-12911	-14947	4937	-14886
	4	-13936	-14931	-14901	-14935

IFM 5		MA			
		1	2	3	4
AR	1	-9974	-11819	-12491	-12949
	2	-18506	-18737	-18789	-18797
	3	-18502	-9194	-9180	-9174
	4	-18748	-9180	-9280	-18812

D(IFM 6)		MA			
		1	2	3	4
AR	1	-20362	-20387	-20394	-20394
	2	-20320	-20308	-20301	-20296
	3	-20309	-20332	-18848	-19686
	4	20303.25	-18848	-19692	-19637

D(IFM 7,2)		MA			
		1	2	3	4
AR	1	-31656	-32404	-31639	-32393
	2	-31725	-32595	5054	-32543
	3	-31168	-25054	-25047	-32564
	4	-32197	-32474	-32571	-25032