Exploring Spatial Dependency of Meteorological Attributes for Multivariate Analysis: A Granger Causality Test Approach

Shrutilipi Bhattacharjee, Soumya K. Ghosh
School of Information Technology
Indian Institute of Technology Kharagpur, West Bengal-721302, India
Email: shrutilipi.2007@gmail.com, skg@iitkgp.ac.in

***Abstract*—Analysis of meteorological attributes often involves the investigation of other secondary attributes which exhibit high spatial correlation with the primary attribute and among themselves. For multivariate spatial analysis, one of the major research challenges is to choose the influential secondary attributes with respect to the primary one and rank them accordingly as per their degree of influence. This research work focuses on electing the most suitable secondary attribute(s) which is/are correlated with the primary and prioritize them individually or as a group, for the corresponding spatial analysis. Time-series forecasting of the meteorological attributes has been chosen as the candidate analysis technique for the case study. A *Granger causality* test based approach is carried out to determine how a group of secondary time-series is statistically significant to forecast the primary. Among several meteorological attributes, the *rainfall* is chosen as the primary time-series, and the *soil moisture*, *relative humidity*, and *land surface temperature* are considered as the secondary time-series to forecast the *rainfall*.**

***Keywords*—Multivariate spatial analysis, Time-series data, Forecasting, Granger causality test.**

I. Introduction

Prediction or forecasting in time-series data often involves the correlated attributes information to enhance the prediction accuracy. Following the law of geographic proximity in two dimensional space, the spatial attributes exhibit high spatial correlation among themselves [1]. Therefore, the forecasting of one meteorological attribute, in the field of *remote sensing* (RS) and *geographic information system* (*GIS*), certainly involves further investigations on other attributes, which influence the primary attribute of interest. Thus, an utmost research challenge is to select the attribute or a group of attributes which influence the primary one and to prioritize them according to their degree of influence. The involvement of these significant influencing attributes helps to yield better precision in actual analysis. Therefore, extracting the causal dependency between the meteorological attributes is mandatory for any multivariate spatial analysis, which can be regarded as a pre-processing step to carry out the actual spatial analysis. In this work, we have investigated the causal linkages between different meteorological attributes to forecast a primary attribute for multivariate time-series forecasting. To carry out the case study, the causal relationships between the *rainfall* (as primary time-series) and *soil moisture*, *relative humidity*, and *land surface temperature* (as secondary time-series) [2] [3] are investigated further.

This work presents a *Granger causality* (*GC*) [4] [5] test based data pre-processing framework for the dependency analysis of meteorological attributes. The *GC* test is a purely data driven approach, which initially developed for the applications in econometrics, by the British economist, Clive Granger. Further, it has been extended in different fields of study to explore the causal relationships among several stochastic variables. It mainly aims at the minimization of error in prediction and forecasting. For the bivariate analysis, Granger [4] has defined the causality hypothesis as: "if some other series y_t contains information in past terms that helps in the prediction of x_t and if this information is contained in no other series used in the predictor, then y_t is said to cause x_t". This causal relationship can be extended in multivariate scenario and different combinations of secondary attributes can be tested to check which combination is the most influential and statistically significant for the primary time-series, hence prioritizing them accordingly.

Background:
A substantial amount of scientific investigations have been reported in many literatures, in the field of meteorological analysis, involving the *Granger causality* testing for extracting causal dependency. Lozano *et al.* [6] have presented a data centric approach for spatio-temporal causal modeling of climate change attribution. For their own dataset, they have reported that the change in atmospheric *temperature* is heavily dependent on the presence of CO_2 in the environment and other green-house gases. Salvucci *et al.* [7] have tested the causal dependency between *soil moisture* and *precipitation* for the study region Illinois, USA. Attanasio *et al.* [8] have presented a review on Granger causality technique based approaches, in order to study the causes of global warming. Smirnov *et al.* [9] have proposed the notion of long-term *Granger causality* and found CO_2 to play a significant role for the rise in *temperature* over the last several decades. Sfetsos *et al.* [10] have applied *Granger causality* based technique to extract the causal relationships between daily PM_{10} exceedances with PM_{10} concentrations. Kodra *et al.* [11] have proposed

a reverse cumulative *Granger causality* test to check the causal relationship between the globally averaged *land surface temperature* and CO_2 in the atmosphere. Dutta *et al.* [12] have proposed an *online* feature selection technique for neural network based prediction of *rainfall*. They have identified the influential features of past *rainfall* data to forecast in future.

Though many research works have applied *Granger causality* test based approach to infer the causal relationship between attributes, this study attempts to involve multiple attributes in order to identify the best possible combination, for the prediction of a primary time-series. In multivariate analysis, it might happen that a group of attributes is the most suitable to cause another, other than causing it individually. Though it is not statistically significant when tested as individual, but causes the primary attribute when grouped with one or more than one time-series. Therefore, a bottom-up approach is proposed, staring from the individual attribute to each type of combinations, and prioritizing them with respect to their influence on the primary attribute. The spatial autocorrelation property can also be verified from this test, such that the nearby locations would yield similar causal dependency than the distant pair. Similarly, for temporal analysis, it can be verified whether inter annual climate change has affected the causal dependencies for different spatial locations. We have followed the classical *F-test* methodology for the *GC* test. So, the main objectives of this work can be stated as follows:

- proposing a bottom-up approach for exhaustive bivariate and multivariate *Granger causality* checking, in different spatial locations
- ranking each group of attributes according to their degree of influence on the primary attribute, depending on the *F statistic* metric
- checking of spatial autocorrelation and its temporal comparison with time-series data

The rest of the paper is organized as follows: Section II gives a brief overview of *Granger causality* testing method, along with the description of statistical *F-test*. Section III presents the description of the proposed bottom-up approach to check and rank the attributes as per their *F-test* measures. The empirical experimentation with real meteorological data is bring forth in Section IV. Finally, the conclusion is summarized in section V.

II. Granger Causality Testing

The fundamental hypothesis of *Granger causality* is based on the notion of single directional causality in time dimension, i.e., the future value can be determined by the past measures, but the reverse is not true. Similarly, if the change of attribute Y is caused by the change in attribute X, then including X as an independent influencing attribute for the analysis of Y, should increase the accuracy in result. Based on this notion, the following hypotheses can be formulated:

- X causes Y implies that the future value of Y is dependent on the past values of both X and Y
- Y causes X implies that the future value of X is dependent on the past values of both X and Y
- X causes Y and Y causes X both imply that the future value of both X and Y are dependent on the past values of both X and Y
- X does not cause Y and Y does not cause X imply that X and Y both are independent of each other

The univariate autoregressive representation of Y (order: N) is given as follows:

$$\begin{aligned} y_t &= \theta_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_N y_{t-N} + \epsilon_t \\ &= \theta_0 + \sum_{i=1}^{N} a_i y_{t-i} + \epsilon_t \end{aligned}$$

where, y_i is the time-series value of Y at time i; θ_0, a_i are the constants and ϵ_t is the residual at time t. Similarly, the bivariate autoregressive representation of Y, including the lagged values of X is given as follows:

$$\begin{aligned} y_t &= \phi_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_N y_{t-N} + b_1 x_{t-1} + \\ & \quad b_2 x_{t-2} + \cdots + b_N x_{t-N} + \xi_t \\ &= \phi_0 + \sum_{i=1}^{N} a_i y_{t-i} + \sum_{i=1}^{N} b_i x_{t-i} + \xi_t \end{aligned}$$

where, x_i is the time-series value of X at time i; ϕ_0, b_i are the constants and ξ_t is the residual at time t. If X Granger causes Y, vector b is a non-zero vector of dimension N and bivariate autoregressive representation of Y should produce better result than univariate autoregression. *Granger causality* can be checked through a series of statistical hypothesis testing methods. In this work, we have chosen statistical *F-test* for null hypothesis checking.

F-test:
The statistical *F-test* [13] has been considered to determine whether a group of attributes or its subsets are jointly significant for the primary attribute. The *F-test* can be regarded as any statistical test where the test statistic has an *F-distribution* under the null hypothesis. The null hypothesis of *Granger causality* test for this work can be stated as follows: *for the univariate autoregressive representation of* Y, $y_i \in Y$ *does not cause* $y_j \in Y$. It is a variant of actual *Granger causality* hypothesis. This hypothesis can also be represented as all the coefficients a_i to be zero in vector a. The alternative hypothesis is given as $a_i \neq 0$ for at least one i in vector a.

For the bivariate regression analysis, the *F-test* method first determines the *F statistic* between two models. If there are n data points to estimate the parameters of both the models (let, first model has p_1 parameters, and second model has p_2 parameters; $p_2 > p_1$), then the *F statistic* is given as follows:

$$F = \frac{\left(\frac{\mathrm{SSE}_1 - \mathrm{SSE}_2}{p_2 - p_1}\right)}{\left(\frac{\mathrm{SSE}_2}{n - p_2}\right)}$$

where, SSE_i is the residual sum of squares of model i, simply given as $\sum(y_i - \bar{y_i})^2$; $\bar{y_i}$ is the mean of series Y. Though the group with more number of parameters will always be able to better fit the data than the model with less parameters, but under the null hypothesis, the former model would not provide any better fit than the latter. The null hypothesis is said to be rejected if the *F statistic* is greater than the critical value of the F-distribution, for some significance level s (usually 0.05). The critical value is defined as, $F_{crit}(m_1, m_2)$ where m_1 is the between-group degrees of freedom and m_2 is the within-group degrees of freedom. The p value is the probability of obtaining the *F statistic* at least as extreme as the one which is actually observed, by accepting the null hypothesis. To reject the hypothesis, the p value should be lesser than s.

III. Causality Testing Framework: CTF

In multivariate analysis, the ranking of the secondary attributes is necessary to decide which attribute or the group of attributes influence the primary attribute most. In this work, we attempt to identify and rank the group of secondary meteorological attributes in a bottom-up approach to forecast the primary one. The rejection of the null hypothesis in *F test* states that at least one of the coefficients in both a and b are non-zero, hence the model is statistically useful for forecasting Y. However, it cannot be concluded that the model is the best. In order to choose the best model, this work proposes an exhaustive hierarchical approach of grouping the attributes. First, each of the attributes is *Granger* tested individually with respect to the primary attribute. The *F statistic* is measured for each of the attributes and they are ranked accordingly with respect to this metric. As the model with higher *F statistic* is statistically more significant, it is assigned lower rank than the model with lower *F statistic*. If the null hypothesis with respect to the individual attribute (with lower *F statistic*) is accepted, still this attribute is not discarded at level 1 from further analysis. It may happen that this secondary time-series could be able to perform better when grouped together with other attributes. Then, each pair of them is clustered together for the second level of analysis. This step is continued until all the secondary attributes form a single group at level n (n is the number of secondary attributes).

Algorithm 1 presents a bottom-up approach of evaluating the *F statistic* for different combination of attributes. Each group of attributes is annotated with the corresponding *F statistic*. Considering the physical significance of this metric, i.e., higher value representing higher influence to the primary attribute, the corresponding ranking can be assigned to each group. In case of our multivariate analysis, the model with lower rank, for which the null hypothesis is rejected, is considered to be the best model.

Input: primary attribute $\{v_0\}$;
secondary attributes $\{\{v_1\}, \{v_2\}, \cdots, \{v_n\}\}$
Output: *F statistic*
$Model_1 = \{v_0\}$;
Hierarchy $H = \phi$;
$G = \{\{v_1\}, \{v_2\}, \cdots, \{v_n\}\}$
foreach $i \in 1$ *to* n **do**
    $H = H \cup G$
    $G = G \times v_i$
end
foreach $Model_i$ *in* H **do**
    $Fstatistic_{(Model_1, Model_i)} = \frac{\left(\frac{SSE_{group1} - SSE_{group2}}{p_2 - p_1}\right)}{\left(\frac{SSE_{group2}}{n - p_2}\right)}$
    if $Fstatistic_{(Model_1, Model_i)} < F_{crit}(m_1, m_2)$ **then**
        accept null hypothesis with $Model_i$
        discard $Model_i$
    end
    else
        reject null hypothesis with $Model_i$
        annotate $Model_i$ with $Fstatistic_{(Model_1, Model_i)}$
    end
end

Algorithm 1: Bottom-up *F statistic* evaluation

IV. Empirical Analysis

The empirical experimentation has been carried out for inferring the causal relationships between real meteorological attributes. The followings are the specifications of the experimental set-up considered for the case study.

Spatial region of interest:
The spatial region *Sundarbans*, India, (central coordinate: $21°56'59''N\ \ 89°10'59.99''E$), which is the largest tidal halophytic mangrove forest in the world, has been chosen for this study. This area is selected as its unpredictable climate change has a major impact in the whole eastern India and it has attracted a significant research interest for different climatological analysis. Due to the proximity of Bay of Bengal[1], the *rainfall* and *humidity* is very high in this region, and often cause a severe damage to the bordering regions during monsoon.

Area of application:
The time-series forecasting of *rainfall* (*RF*) is analysed as the candidate application to extract the causal dependency among meteorological attributes. The secondary time-series are *soil moisture* (*SM*), *relative humidity* (*RH*), and *land surface temperature* (*LST*). The causal dependencies of these attributes with *RF* are investigated for further.

Experimental datasets:
Due to inadequacy of real datasets, the meteorological data used for this study is retrieved from *FetchClimate*[2], by Microsoft-Research. It provides climate information service offering the past and present data of climate attributes. The meteorological attributes are fetched from this explorer for the past four years (2001, 2004, 2007 and 2010). As, *rainfall* is the primary attribute, the past data for rainy season (typically, during the month of July-September) is captured for each of the year, with an interval of twenty four hours. Five random locations have been chosen from the whole study region for testing the causal dependencies.

Statistical tool for analysis:
A predictive analytics software, IBM SPSS Statistics 15.0[3] has been used for this study. For bivariate analysis, *One-way Anova* test is carried out for each pair of attributes. The multivariate linear regression measures the *F statistic* for the combination of more than one attribute.

[1] http://www.britannica.com/EBchecked/topic/60740/Bay-of-Bengal; Accessed on August 2014
[2] http://fetchclimate.cloudapp.net; Accessed on August 2014
[3] http://www-01.ibm.com/software/in/analytics/spss/; Accessed on August 2014

The property of spatial autocorrelation is also investigated through this experimentation. For different types of spatial proximity, all the locations may not rank the secondary attributes in a similar manner. However, due to spatial autocorrelation, the nearby locations must yield similar results than the distant region. This hypothesis is tested in terms of similarity between *F statistic* for each of the groups among different regions. Due to brevity of space, the case study is reported for past four years (2001, 2004, 2007, and 2010) only, and the results are specified in Tables II, III, IV and V, respectively.

Discussion:
For each group <G>, the null hypothesis is: "G does not *Granger* cause *rainfall* (*RF*)". The *F statistic* is specified for each group, along with the critical value for that particular *F test* (refer Tables II, III, IV and V). The 'Reject?' column checks whether to reject (✓) or accept (×) the null hypothesis. In individual location, the groups are ranked with *F statistic* value. From all the tables (II, III, IV and V), it is observed that the group of all the secondary attributes, i.e., <*land surface temperature*, *soil moisture*, *relative humidity*> together proved to be a significant model for most of the years. These ranking can be analysed further for different spatial locations in the same year for spatial analysis, similarly for different years in the same location for temporal analysis. Though the temporal behaviours of the similar groups in the same location is not much similar, however, the results exhibit high spatial autocorrelation. For example, locations 1, 4, and 5 are nearby than location 2, and 3. For the spatial proximity, high autocorrelation within, and low correlation among the location clusters can be observed for each of the year in terms of *F statistic*. Another very important observation, from the empirical results: even each of the individual attributes are not *Granger causing* the primary attribute individually, they may do so when in a group. This is the reason for not discarding the insignificant attributes in level 1.

For the empirical evidence of the proposed ranking strategy, the actual forecasting of *rainfall* is carried out for the year 2010 for each of the five locations, with past 3 years (2001, 2004, and 2007) data. In SPSS, the *Time Series* analysis is carried out with *ARIMA* model for forecasting. The corresponding *root mean square error* (*RMSE*) in prediction is specified in Table I. It is evident from the Table I that the multivariate time-series forecasting with the group <LST, RH, SM> reports the minimal error. This group is already found to be statistically most significant for the past scenarios using *GC* test.

Table I: RMSE IN TIME-SERIES FORECASTING OF *RAINFALL*

		Locations				
		1	**2**	**3**	**4**	**5**
Groups	$< \mathbf{LST}, \mathbf{RH} >$	1.583	1.440	1.249	1.446	1.334
	$< \mathbf{LST}, \mathbf{SM} >$	1.611	1.445	1.352	1.497	1.285
	$< \mathbf{RH}, \mathbf{SM} >$	1.758	2.259	1.661	2.754	2.064
	$< \mathbf{LST}, \mathbf{RH}, \mathbf{SM} >$	1.082	0.967	0.931	0.897	1.004

V. Conclusion

The relationships or the causal linkages between different time-series data play an important role in spatial analysis. Such analysis, like forecasting of one meteorological time-series, requires to infer the causal dependencies with other influential attributes. This work presents a data pre-processing framework for multivariate time-series forecasting to identify the influential secondary attributes and rank them as per their degree of influence. A hierarchical *Granger causality* test based approach is adopted to estimate the influence of the individual and the group of attributes on forecasting. Experimental study also demonstrates the efficacy of this framework. This method can be extended further to develop an optimized hierarchy through grouping and pruning of attributes.

References

[1] D. Gamerman and A. R. Moreira, "Multivariate spatial regression models," *Journal of multivariate analysis*, vol. 91, no. 2, pp. 262–281, 2004.

[2] J. Deardorff, "Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation," *Journal of Geophysical Research: Oceans (1978–2012)*, vol. 83, no. C4, pp. 1889–1903, 1978.

[3] K. L. Findell and E. A. Eltahir, "An analysis of the soil moisture-rainfall feedback, based on direct observations from illinois," *Water Resources Research*, vol. 33, no. 4, pp. 725–735, 1997.

[4] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[5] U. Triacca, "Is granger causality analysis appropriate to investigate the relationship between atmospheric concentration of carbon dioxide and global surface air temperature?" *Theoretical and applied climatology*, vol. 81, no. 3-4, pp. 133–135, 2005.

[6] A. C. Lozano, H. Li, A. Niculescu-Mizil, Y. Liu, C. Perlich, J. Hosking, and N. Abe, "Spatial-temporal causal modeling for climate change attribution," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2009, pp. 587–596.

[7] G. D. Salvucci, J. A. Saleem, and R. Kaufmann, "Investigating soil moisture feedbacks on precipitation with tests of granger causality," *Advances in water resources*, vol. 25, no. 8, pp. 1305–1312, 2002.

[8] A. Attanasio, A. Pasini, and U. Triacca, "Granger causality analyses for climatic attribution," *Atmospheric and Climate Sciences*, vol. 3, p. 515, 2013.

[9] D. A. Smirnov and I. I. Mokhov, "From granger causality to long-term causality: Application to climatic data," *Physical Review E*, vol. 80, no. 1, p. 016208, 2009.

[10] A. Sfetsos and D. Vlachogiannis, "A new approach to discovering the causal relationship between meteorological patterns and PM_{10} exceedances," *Atmospheric Research*, vol. 98, no. 2, pp. 500–511, 2010.

[11] E. Kodra, S. Chatterjee, and A. R. Ganguly, "Exploring granger causality between global average observed time series of carbon dioxide and temperature," *Theoretical and applied climatology*, vol. 104, no. 3-4, pp. 325–335, 2011.

[12] B. Dutta, A. Ray, S. Pal, and D. C. Patranabis, "A connectionist model for rainfall prediction." *Neural Parallel & Scientific Comp.*, vol. 17, no. 1, pp. 47–58, 2009.

[13] J. C. Rogan and H. Keselman, "Is the anova f-test robust to variance heterogeneity when sample sizes are equal?: An investigation via a coefficient of variation," *American Educational Research Journal*, vol. 14, no. 4, pp. 493–498, 1977.

Table II: *F-TEST* FOR METEOROLOGICAL ATTRIBUTES OF *SUNDARBANS*, YEAR: 2001

Level	Attribute groups	*F statistic*	*p* value	Critical value	Reject?	Rank
Location 1 (Coordinate: 22.2541°N 89.7268°E)						
Level 1	<LST>	0.470	0.099		×	—
	<SM>	1.212	0.061	3.97	×	—
	<RH>	0.802	0.063		×	—
Level 2	<LST, SM>	7.578	0.015		✓	3
	<LST, RH>	4.007	0.020	3.12	✓	4
	<SM, RH>	13.94	0.011		✓	2
Level 3	<LST, SM, RH>	20.42	0.005	2.73	✓	1
Location 2 (Coordinate: 22.4523°N 89.9548°E)						
Level 1	<LST>	0.472	0.099		×	—
	<SM>	1.119	0.064	3.97	×	—
	<RH>	0.810	0.072		×	—
Level 2	<LST, SM>	8.969	0.001		✓	3
	<LST, RH>	3.245	0.041	3.12	✓	4
	<SM, RH>	13.50	0.033		✓	2
Level 3	<LST, SM, RH>	20.07	0.025	2.73	✓	1
Location 3 (Coordinate: 22.0990°N 89.9795°E)						
Level 1	<LST>	0.471	0.099		×	—
	<SM>	1.119	0.053	3.97	×	—
	<RH>	0.806	0.088		×	—
Level 2	<LST, SM>	8.503	0.021		✓	3
	<LST, RH>	3.576	0.007	3.12	✓	4
	<SM, RH>	13.78	0.002		✓	2
Level 3	<LST, SM, RH>	20.36	0.008	2.73	✓	1
Location 4 (Coordinate: 22.3507°N 89.5041°E)						
Level 1	<LST>	0.466	0.099		×	—
	<SM>	1.125	0.056	3.97	×	—
	<RH>	0.798	0.082		×	—
Level 2	<LST, SM>	7.127	0.023		✓	3
	<LST, RH>	4.494	0.012	3.12	✓	4
	<SM, RH>	14.55	0.021		✓	2
Level 3	<LST, SM, RH>	20.94	0.041	2.73	✓	1
Location 5 (Coordinate: 22.1728°N 89.4436°E)						
Level 1	<LST>	0.468	0.099		×	—
	<SM>	1.125	0.086	3.97	×	—
	<RH>	0.799	0.088		×	—
Level 2	<LST, SM>	6.665	0.042		✓	3
	<LST, RH>	4.459	0.013	3.12	✓	4
	<SM, RH>	14.08	0.019		✓	2
Level 3	<LST, SM, RH>	20.37	0.017	2.73	✓	1

Table III: *F-TEST* FOR METEOROLOGICAL ATTRIBUTES OF *SUNDARBANS*, YEAR: 2004

Level	Attribute groups	*F statistic*	*p* value	Critical value	Reject?	Rank
Location 1 (Coordinate: 22.2541°N 89.7268°E)						
Level 1	<LST>	5.313	0.032		✓	6
	<SM>	18.67	0.021	4.04	✓	1
	<RH>	1.863	0.056		×	—
Level 2	<LST, SM>	10.81	0.011		✓	4
	<LST, RH>	11.80	0.001	3.19	✓	2
	<SM, RH>	9.161	0.002		✓	5
Level 3	<LST, SM, RH>	11.34	0.002	2.79	✓	3
Location 2 (Coordinate: 22.4523°N 89.9548°E)						
Level 1	<LST>	5.466	0.001		✓	6
	<SM>	18.82	0.001	4.04	✓	1
	<RH>	1.906	0.058		×	—
Level 2	<LST, SM>	10.20	0.001		✓	4
	<LST, RH>	10.88	0.040	3.19	✓	2
	<SM, RH>	8.527	0.022		✓	5
Level 3	<LST, SM, RH>	10.68	0.001	2.79	✓	3
Location 3 (Coordinate: 22.0990°N 89.9795°E)						
Level 1	<LST>	5.353	0.001		✓	6
	<SM>	18.80	0.002	4.04	✓	1
	<RH>	1.882	0.054		×	—
Level 2	<LST, SM>	10.55	0.002		✓	4
	<LST, RH>	11.19	0.004	3.19	✓	2
	<SM, RH>	8.680	0.001		✓	5
Level 3	<LST, SM, RH>	10.92	0.001	2.79	✓	3
Location 4 (Coordinate: 22.3507°N 89.5041°E)						
Level 1	<LST>	5.258	0.001		✓	6
	<SM>	18.41	0.001	4.04	✓	1
	<RH>	1.844	0.062		×	—
Level 2	<LST, SM>	11.29	0.001		✓	4
	<LST, RH>	12.45	0.001	3.19	✓	2
	<SM, RH>	9.620	0.001		✓	5
Level 3	<LST, SM, RH>	11.93	0.001	2.79	✓	3
Location 5 (Coordinate: 22.1728°N 89.4436°E)						
Level 1	<LST>	5.268	0.001		✓	6
	<SM>	18.52	0.001	4.04	✓	1
	<RH>	1.847	0.072		×	—
Level 2	<LST, SM>	11.07	0.001		✓	4
	<LST, RH>	12.39	0.001	3.19	✓	2
	<SM, RH>	9.626	0.001		✓	5
Level 3	<LST, SM, RH>	11.81	0.001	2.79	✓	3

Table IV: *F-TEST* FOR METEOROLOGICAL ATTRIBUTES OF *SUNDARBANS*, YEAR: 2007

Level	Attribute groups	*F statistic*	*p* value	Critical value	Reject?	Rank
Location 1 (Coordinate: 22.2541°N 89.7268°E)						
Level 1	<LST>	0.459	0.099	4.00	×	—
	<SM>	1.046	0.088		×	—
	<RH>	1.376	0.078		×	—
Level 2	<LST, SM>	6.244	0.002	3.15	✓	4
	<LST, RH>	9.542	0.001		✓	2
	<SM, RH>	6.334	0.002		✓	3
Level 3	<LST, SM, RH>	10.65	0.001	2.75	✓	1
Location 2 (Coordinate: 22.4523°N 89.9548°E)						
Level 1	<LST>	0.470	0.099	4.00	×	—
	<SM>	1.406	0.092		×	—
	<RH>	1.405	0.065		×	—
Level 2	<LST, SM>	6.466	0.002	3.15	✓	3
	<LST, RH>	8.468	0.001		✓	2
	<SM, RH>	5.309	0.006		✓	4
Level 3	<LST, SM, RH>	9.826	0.001	2.75	✓	1
Location 3 (Coordinate: 22.0990°N 89.9795°E)						
Level 1	<LST>	0.466	0.099	4.00	×	—
	<SM>	1.045	0.081		×	—
	<RH>	1.392	0.071		×	—
Level 2	<LST, SM>	6.406	0.002	3.15	✓	3
	<LST, RH>	8.959	0.001		✓	2
	<SM, RH>	5.774	0.004		✓	4
Level 3	<LST, SM, RH>	10.23	0.001	2.75	✓	1
Location 4 (Coordinate: 22.3507°N 89.5041°E)						
Level 1	<LST>	0.4582	0.099	4.00	×	—
	<SM>	1.046	0.067		×	—
	<RH>	1.362	0.085		×	—
Level 2	<LST, SM>	6.231	0.002	3.15	✓	4
	<LST, RH>	10.18	0.001		✓	2
	<SM, RH>	6.879	0.001		✓	3
Level 3	<LST, SM, RH>	11.04	0.001	2.75	✓	1
Location 5 (Coordinate: 22.1728°N 89.4436°E)						
Level 1	<LST>	0.453	0.099	4.00	×	—
	<SM>	1.048	0.095		×	—
	<RH>	1.360	0.085		×	—
Level 2	<LST, SM>	6.073	0.003	3.15	✓	4
	<LST, RH>	10.10	0.001		✓	2
	<SM, RH>	6.862	0.001		✓	3
Level 3	<LST, SM, RH>	11.01	0.001	2.75	✓	1

Table V: *F-TEST* FOR METEOROLOGICAL ATTRIBUTES OF *SUNDARBANS*, YEAR: 2010

Level	Attribute groups	*F statistic*	*p* value	Critical value	Reject?	Rank
Location 1 (Coordinate: 22.2541°N 89.7268°E)						
Level 1	<LST>	0.413	0.099	4.01	×	—
	<SM>	0.774	0.091		×	—
	<RH>	1.363	0.090		×	—
Level 2	<LST, SM>	14.55	0.001	3.16	✓	2
	<LST, RH>	12.27	0.001		✓	3
	<SM, RH>	7.764	0.001		✓	4
Level 3	<LST, SM, RH>	14.77	0.001	2.77	✓	1
Location 2 (Coordinate: 22.4523°N 89.9548°E)						
Level 1	<LST>	0.421	0.099	4.01	×	—
	<SM>	0.741	0.092		×	—
	<RH>	1.374	0.085		×	—
Level 2	<LST, SM>	13.97	0.001	3.16	✓	1
	<LST, RH>	11.63	0.001		✓	3
	<SM, RH>	7.131	0.001		✓	4
Level 3	<LST, SM, RH>	13.24	0.001	2.77	✓	2
Location 3 (Coordinate: 22.0990°N 89.9795°E)						
Level 1	<LST>	1.549	0.099	4.01	×	—
	<SM>	0.802	0.086		×	—
	<RH>	1.251	0.069		×	—
Level 2	<LST, SM>	13.64	0.004	3.16	✓	2
	<LST, RH>	12.43	0.001		✓	3
	<SM, RH>	7.672	0.001		✓	4
Level 3	<LST, SM, RH>	14.30	0.001	2.77	✓	1
Location 4 (Coordinate: 22.3507°N 89.5041°E)						
Level 1	<LST>	0.413	0.099	4.01	×	—
	<SM>	0.744	0.098		×	—
	<RH>	1.356	0.094		×	—
Level 2	<LST, SM>	14.92	0.001	3.16	✓	2
	<LST, RH>	12.53	0.001		✓	3
	<SM, RH>	8.058	0.001		✓	4
Level 3	<LST, SM, RH>	15.55	0.001	2.77	✓	1
Location 5 (Coordinate: 22.1728°N 89.4436°E)						
Level 1	<LST>	0.409	0.099	4.01	×	—
	<SM>	0.740	0.092		×	—
	<RH>	1.357	0.093		×	—
Level 2	<LST, SM>	14.63	0.001	3.16	✓	2
	<LST, RH>	12.57	0.001		✓	3
	<SM, RH>	8.090	0.001		✓	4
Level 3	<LST, SM, RH>	15.50	0.001	2.77	✓	1