-

Study the Space –time variations of Indus River Flow Propagation

Abstract: River flow is the most important geophysical process of the earth and its ecosystem. Being very unpredictable in nature, at several times its ensemble flow propagation along the network becomes hazardous flood, that causing enormous damage and loses. Pakistan is an agricultural country with a long network of rivers and canals and making one of the world's largest rivers, the well-known Indus River system. It is severely affected by such situation in comparison to other countries of the world. The system undergoes great climatic, strong seasonal and inter-annual variability dominated by both monsoon and snow-glacier dynamics of the Himalayan and Karakoram regions. A better understanding of the variability may provide insight into the problems associated with unpredictable variations in river flow and their propagation. This study is an attempt to enhance the current knowledge of the Indus River flow dynamics along the network. For this purpose linear and nonlinear methods of propagation analysis of the mean 10daily river flow (TDF) are utilised. The linear and nonlinear schemes employ cross-correlation and the normalized average of cross-mutual information (MI) method respectively. The overall dominant mechanism of the system shows a linear behaviour, however, some stations demonstrate nonlinear propagation in the network. The results of these analysis may also provide the viability to study the behaviour of regional and global climatic parameters related to the type of information propagated.

Introduction

Pakistan is an agriculturally dominant country demanding full utilization of its water resources based on one of the world's largest Indus River System (IRS). They nourished by melting of snow and glaciers in most parts of the year, orographic rainfall, during monsoon (June -to-September) [1, 2] and western monsoon in winter-spring (February-to- May) seasons [3]. The increase in water level begins in April and by June, July it rises to the maximum level. On several occasions, it produces heavy flooding; not only due to rainfall and melting of glacier snouts and/or ice bridges, sometimes the impact of temperature triggers disastrous snowstorm avalanches. Thus, the system is always threatened by severe catastrophic floods, especially, during the monsoon season. The flooding situation arrives every year but on some occasion it becomes disastrous causing loss of lives and destruction of property, agricultural, buildings, watercourses and etc. Moreover, due to poor water control- arrangements the excessive water is uselessly drained out to the sea.

The impact of local climatic analysis on the Indus River system (IRS) network is significantly dominating on the river flow system (Hassan **&** Ansari, 2015). It may show some considerable nonlinear behaviour of seasonality and stochasticity in the Indus River flow (Hassan & Ansari 2010). To investigate the influential role of the regional climatic parameters on the river flows along the IRS network this paper introduces linear and nonlinear methods of river flow propagation analysis. For this purpose river flow data from Kotri (consider as the lowest reference station) with all six upstream stations (Sukkur, Guddu, Taunsa, Chasma, Kalabagh and Tarbela,) mean 10daily flow (TDF) along the Indus River (Fig. 1) is analysed. This procedure is also repeated for Sukkur and Guddu (second and third last stations) as reference

UNDER PEER REVIEW

stations. The linear method evaluates cross-correlation of two data arrays at different lag values where, cross-correlation is a simple Pearson's correlation techniques. The nonlinear estimation considers normalized average of cross-mutual information (MI) method between the two data series. The MI among the two data series is a way to extract the nonlinear information (correlation) between them (Fraser & Swinney, 1986; Fraser, 1989; Pineda & Sommerer, 1993).

This paper investigates the existence of linear propagation behaviour in the TDF along the Indus River. The overall dominancy of the TDF propagation along the Indus River is linear. Results of this study will also lay the foundations for some of our future studies related to the behaviour of regional and global climatic parameters and the type of information they propagate. The next section describes the material and method. Section 3 presented results and discussion and finally section 4 concludes the paper.

2 Material and Method

This section analyses the linear and nonlinear variations in the Indus River flow from each of the lower three stations (Kotri, Sukkur, and Guddu) to all their upstream stations.

The linear analysis considers cross-correlation method and nonlinear scheme utilised mutual information method (Hassan, & Ansari, 2010 and some of references there in). These both techniques applied to TDF data series of two different stations. To find the cross mutual information (MI; relative nonlinear information) between two different data series, various techniques available for calculating MI are discussed in details in the Hassan *et al*., (Hassan *et al*., unpublished) (Tourassi et al., 2001). One of those techniques is the equal probable binning method (joint probability calculation) that will be used to estimate MI. Mutual information method measures the amount of information of one data series contained about the other. The MI is computed as

$$CMI(K,N) = \sum_{k,\,n} P_{KN}(k,n) \ln \frac{P_{KN}(k,n)}{P_K(k)\; P_N(n)}$$

where K and N are the two different data series with joint probability function $P_{KN}(k, n)$ and marginal probability functions $P_K(k)$ and $P_N(n)$. The MI is also defined as the relative entropy between the joint and product marginal distributions (Shannon & Weaver, 1949; Cover & Thomas, 1991).

2.1 Linear Analysis of TDF

To develop the linear method scheme Kotri is consider as reference station for evaluation of cross-correlation with all their upstream stations TDF data. Fig. 1.a shows the estimated cross-correlation values (height of different markers) between Kotri with self and all their upstream station's TDF from lag = –4 to lag = 4. Beyond lag = ± 4 the space time variation in the cross-correlation values makes no significant change. These values are now connected by dashed lines in such a way that each curve represents the linear relation from Kotri to Kotri (autocorrelation) and with all their upstream station. The curves depicts attenuation and shifting behaviour because all upstream stations have some (unequal) distances from Kotri. To represent their trend draw a thick line that starts from Kotri-Kotri (peak cross-correlation value = 1) to appropriate maximum of the lagged cross-correlation value curve of the Kotri-to-other stations. As we know that cross-correlation values represent the linear relation between the

UNDER PEER REVIEW

two data series, so this thick line can be considered as propagation of linear information of the river flow. It means that the linear information is at maximum = 1 at Kotri-Kotri cross-correlation value (at lag = 0) and it gradually decreases for Kotri-to other upstream station cross-correlation values along the Indus River, because of the water contribution from other resources and as well as their agricultural and domestic supplies, seepage, and evaporation, all are unevenly distributed. To see this decreasing trend more prominently Fig. 1.a transformed in to the space-time diagram (Fig. 2), for each lag value. Moreover, to represent the above discussed propagation of linear information behaviour draw a thick black line starting from the maximum value (at Kotri-Kotri) follow appropriate maximum lag cross-correlation values of Kotri-to other stations. This thick line exhibits a decreasing linear trend (Fig. 2) and is called as a liner information propagation roll-off (LR) curve. It is drawn by curve drawing facility of the Microsoft drawing software (Microsoft Word, 2010). This LR curve is based on cross-correlation vales obtained from TDF data series, they may also approximate the cross-correlation curves (at appropriate higher lags values) of the high resolution data series like the daily, hourly or at every minutes. Now repeat the same procedure of drawing LR curve from the other two reference stations of Sukkur and Guddu (Figs. 2.b & 2.c). These LR curves approximate as a part of the LR curve of Fig. 2.a.

Reviewing all cross-correlation curves in Fig. 2a it appears that around Taunsa station all of these become closer to each other while approaching to the Kalabagh station and beyond it they all spread away. This major change of cross-correlation are because of Kabul River (starts rising one month earlier than the Indus River) meet Indus main stream between Tarbela and Kalabagh station. However, after Taunsa station in the downstream there is no major river meeting the Indus main stream. Moreover, between Taunsa and Kalabagh the cross-correlation values have different structure of variation not only because of natural linear information roll-off behaviour but also because of the fact that these stations are much close to each other as compared to the other downstream stations.

PEER REVIEW

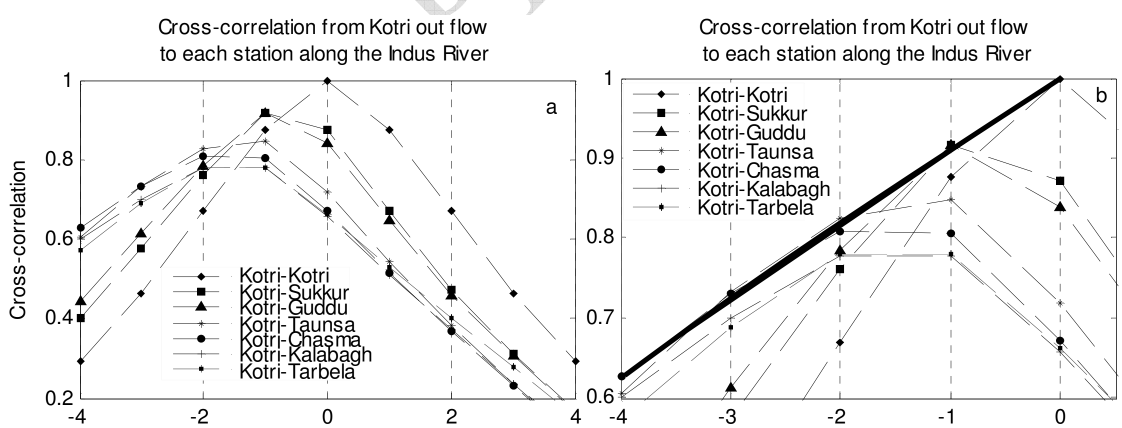

Fig. 1 Linear information propagation in TDF data series using cross-correlation analysis between Kotri with itself and with all its upstream station's along Indus River, vs lag values: (a) the appropriate stations cross-correlation values are connected with dashed lines (cross-correlation curves), and (b) the cross-correlation curves roll off (LR) behaviour, shows with thick black line, starts from Kotri-Kotri at lag = 0 follow appropriate maximums of the cross-correlation values of Kotri-to-Sukkur at lag = −1, Kotri-to-Taunsa at lag = −2 and end at Kotri-to-Chasma at lag = −1

The LR curve with reference to the Kotri station along the Indus River have their higher (maximum) value at Kotri station and gradually roll-off their values in upper stations and completely fall at Taunsa. This means that the TDF of the lower four stations are very well linearly correlated. The space-time LR curves in Figs. 2.b & 2.c have slightly sharp roll-off behaviour in comparison to that of Fig. 2.a. Considering other upstream stations as a reference station for the cross-correlation analysis same variations in cross-correlation values are observed at each lag. Moreover, The cross-correlation analyses with reference to the Tarbela station along the downstream of Indus River is not significantly shows of their linear roll-off.

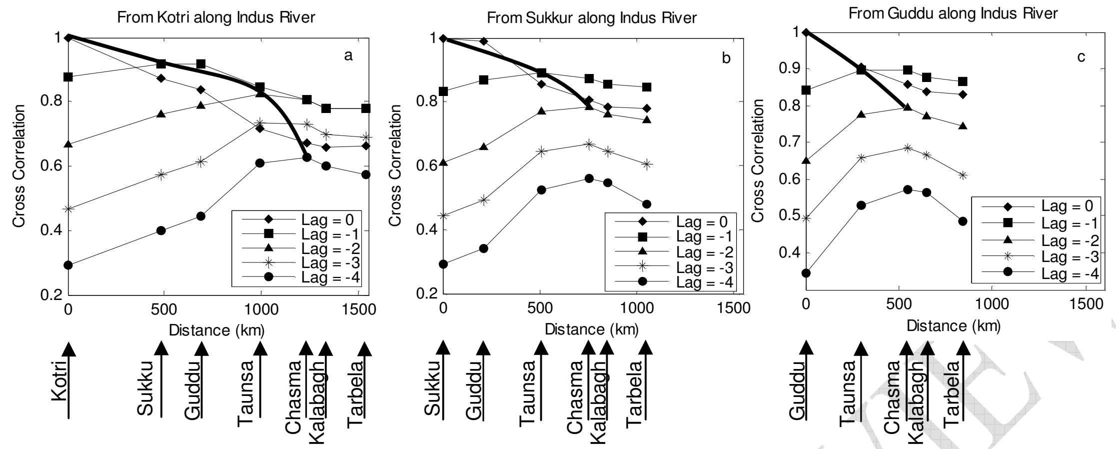

Fig. 2 The LR curve (Thick line) is prominently viewed over CROSS-CORRELATION value curves prominently (thin line) in space-time CROSS-CORRELATION analysis (method developed and transformed from Fig. 1), from lower three stations (a) Kotri, (b) Sukkur, and (c) Guddu with itself and with all its upstream station's along Indus River.

2 Nonlinear Analysis of TDF

To explore any possible nonlinearity inherited in the river flow propagation along the Indus River, this paper asses the MI behaviour of three reference stations and with their upstream stations. The MI values are normalized for each of the three set with their appropriate reference station's MI value at lag = 0 (Shannon entropy or mutual information at lag = 0). These MI values are now called normalized average of cross mutual information. The space-time MI value curves are obtained in a way similar to that used in obtaining space-time cross-correlation value curves as discuses in section 1

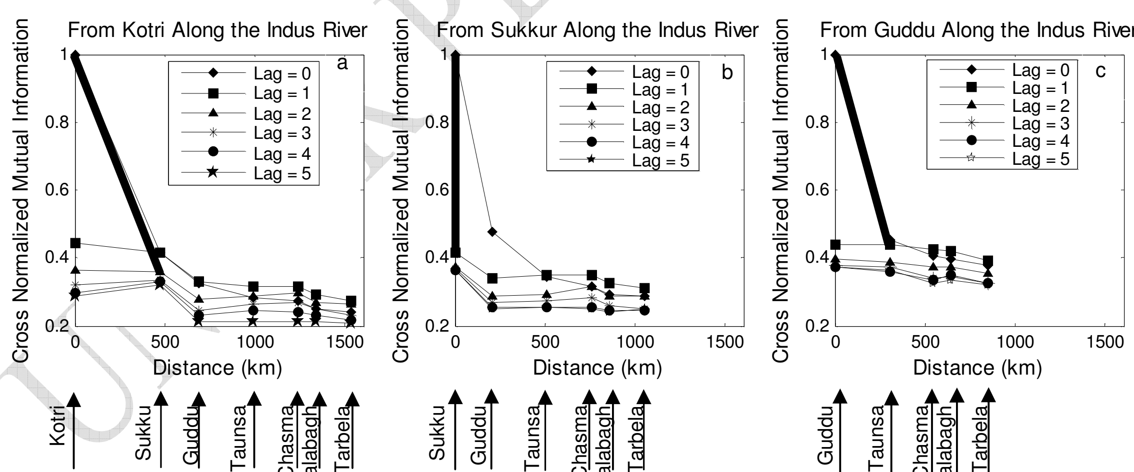

Fig. 3 The NR curve (Thick line) behaviour acquired from space-time MI analysis (obtained as the same procedure developed in Figs. 1.1 & 1.2), from lower three stations (a) Kotri, (b) Sukkur, and (c) Guddu with itself and with all its upstream station's along Indus River.

Using TDF data series all MI values are calculated from lower three reference stations (Figs. 3) separately, with itself, and with all its upstream station's along the Indus River for lag = 0, to lag = −5. To explore the possibility of propagation of nonlinear information draw a thick black line start from maximum MI value at reference station to appropriate maximum MI value at other upstream station. Because of the decreasing behaviour of this thick line we call it nonlinear information propagation roll-off (NR) curve. In Fig. 3.a the NR curve starts from Kotri-Kotri MI maximum value at lag = 0 to the next Kotri-Sukkur MI at lag = 2, where it ends

UNDER PEER REVIEW

(as there is no further MI maximum value appears afterwards). Fig. 3.b shows no proceeding NR curve because the upstream stations have no lagged maximum MI value and all maximum values appear at Sukkur station. Fig. 3.c represents the MI propagation starting from Guddu-Guddu MI maximum value at lag = 0 to its next maximum Guddu-Taunsa MI value at lag = 1 where it ends as there is no MI maximum value that appears further. Moreover, it is observed that beyond the lag = −5 the MI value curves undergo no change. The three figures (Figs. 3.a-c) indicate that the nonlinear information propagates only from Kotri and Guddu stations and prolong for only one station Sukkur at lag = 2 and Taunsa at lag = 1 respectively.

3 A comparison of linear and nonlinear analyses

The linear information (LR curves; Fig. 2) sustain more lag value as compared to nonlinear method (NR curves; Fig. 3). Thus it can be concluded that for the ensemble TDF of lower five stations (Kotri, Sukkur, Guddu, Taunsa, and Chashma) effectively explains by linear method as comparison to nonlinear method. However, as there are some indications of nonlinear behaviour (Fig. 3.a & c) so, it can’t be completely ignored the existence of nonlinearity in this river flow network. It is a separate topic which consider out of the scope of this study. The Fig. 2.a shows that LR curve tracking cross-correlation values are higher at lower stations and gradually roll-off at their corresponding upper stations and completely vanish at Chashma. This means that the TDF of the lower five stations can be well explained through the linear stochastic modelling. However, the upper three station’s (Chashma, Kalabagh, and Tarbela) LR tracking cross-correlation values demonstrate the existing of strong relations among them at lag = 0. This means that these three stations TDF’s are very well explained by regression. The validity of nonlinear analyses is only effective from Kotri to Sukkur and from Guddu to Taunsa. This informs that the possibility of existence of nonlinearity inherited in the propagation of TDF along the Indus River in lower stations. However, their propagation is restricted to the next station and does not prolong further.

4 Conclusion

In view of the above discussion it follows that over all dominant mechanisms of the TDF propagation along the network is linear. However, some stations demonstrate existence of nonlinearity in the propagation of TDF of the network. These analyses not only provide the information regarding the type of propagation and their limits along the network but also provide the viability to study the behaviour of regional and global climatic parameters related to type of information propagation and its roll-off behaviour.

Reference

Cover, T. M. & Thomas, J.A. (1991) *Elements of information theory*. John Wiley & Sons, New York, NY.

Fraser, A. M. (1989) Reconstructing attractors from scalar time series: a comparison of singular system and redundancy criteria. *Physica D* 34, 391–404.

Fraser, A. M. & Swinney, H. L. (1986) Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* 33 (2), 1134–1140.

Hassan, S. A. & Ansari, M. R. K. (2010) Nonlinear analysis of seasonality and stochasticity of the Indus River. *Hydrol. Sci. J*. 55(2), 250–265.

UNDER PEER REVIEW

Hassan, S. A., Porporato, A. & Ansari, M. R. K. *How cross mutual Information method determine the nonlinear relation (Approximation of Cross Mutual Information methods)*, unpublished.

Microsoft Word, (2002) Microsoft Corporation

Hassan, S. A., & Ansari, M. R. K. Climate Forcing River Flow Process, unpublished.

Pineda, F. J. & Sommerer, J. C. (1993) Estimating generalized dimensions and choosing time delays: a fast algorithm. In: Time Series Prediction: Forecasting the Future and Understanding the Past (ed. by A. S. Weigend & N. A. Gershenfeld), 367–385. Addison-Wesely, Reading, Massachusetts, USA.

Shannon, C. E. & Weaver, W. (1949) *The mathematical theory of communication. University of Illinois Press*. Urbana, Illinois.

Tourassi, G.D., Frederick , E.D., Markey, M.K. & Floyd, Jr. C.E. (2001) Application of the mutual information criterion for feature selection, in computer-aided diagnosis. Med. Phys. 28 .12., 2394-2402

Box, G. E. & Jenkins, G. M. (1994) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, California, USA.

Hassan, S. A. & Ansari, M. R. K. (2015) Hydro-climatic Aspects of Indus River Flow Propagation. Arab. J. Geosci. 8:10977–10982. DOI 10.1007/s12517-015-1960-y

UNDER PEER REVIEW