Toward A Practical Approach for Ergodicity Analysis

Hongrui Wang[1], Cheng Wang[1,†,*], Yan Zhao[1], Xin Lin[2], Chen Yu[1]

1. College of Water Sciences, Beijing Normal University--Key Laboratory for Water and Sediment Sciences, Ministry of Education, Beijing 100875, China.

2. School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China.

[*] Corresponding to: chengw@knights.ucf.edu

[†] Current affiliation: Environmental Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

Toward A Practical Approach for Ergodicity Analysis

Abstract: It is of importance to perform hydrological forecast using a finite hydrological time series. Most time series analysis approaches presume a data series to be ergodic without justifying this assumption. This paper presents a practical approach to analyze the mean ergodic property of hydrological processes by means of autocorrelation function evaluation and Augmented Dickey Fuller test, a radial basis function neural network, and the definition of mean ergodicity. The mean ergodicity of precipitation processes at the Lanzhou Rain Gauge Station in the Yellow River basin, the Ankang Rain Gauge Station in Han River, both in China, and at Newberry, MI, USA are analyzed using the proposed approach. The results indicate that the precipitations of March, July, and August in Lanzhou, and of May, June, and August in Ankang have mean ergodicity, whereas, the precipitation of any other calendar month in these two rain gauge stations do not have mean ergodicity. The precipitation of February, May, July, and December in Newberry show ergodic property, although the precipitation of each month shows a clear increasing or decreasing trend.

Introduction

A hydrological process can be usually regarded as a stochastic process and any observation is just a realization of a random variable representing the stochastic process. A realization of a stochastic process is defined as the outcome of an experiment in which the process is observed (Shahin et al., 1993). For example, a time series of observed precipitation

1

20   data at a gauge station is a realization of the precipitation process at the area the gauge station

covers. A collection of all possible realizations of a stochastic process, i.e., the ensemble, are

used to represent the process.

Given that the variations of a hydrological variable representing certain hydrological

process are usually very complicated and affected by random factors, statistical properties of

25   the process, such as the phase mean function of a data series $\{\xi(t); t = 1, 2, ..., N\}$,

$m(t) = E\{\xi(t)\}$, and the correlation function $R(s, t) = Cov(\xi(s), \xi(t))$, etc., must be known in

order to describe and thus analyze this stochastic process. In reality, however, only finite

realizations, i.e. a finite set of records $\{\xi(t); t = 1, 2, ..., N\}$ of the random series $\{\xi(t)\}$, in most

time only one single realization, are available from observation. It will be significantly

30   helpful and meaningful if the statistic properties of a process such as its mean and standard

deviation variance can be estimated from the observation of a realization. In fact, it is

common in practice for hydrologists to use the statistical properties of a single realization

from a hydrologic process as the statistical properties of this process. For example, the mean

value and the probability distribution of a certain number years of observation data at a flow

35   rate gauge station are often used for further hydrological process analysis, e.g., the flood

frequency analysis. By doing this, we actually assume that the statistic properties of the flow

rate at this station can be estimated from the finite number of observations and their values

are identical, or more strictly, close enough to the corresponding values of the statistic

properties of the flow process. Is it possible to estimate the statistical properties of the

40   population process using finite observations? How reliable is this estimate approach? These

2

questions arose and were described as the asymptotic convergence of the average over time (sample average) to the phase mean (population mean). This property is called the ergodic property or ergodicity (Chick et al., 1996). Ergodicity is the property by which each realization of a given process is a complete and independent representative of all possible realizations of the process (Shahin et al., 1993). Thus, the ergodic properties allow scientists and researchers to determine the statistical properties of a process from a single realization. In this sense, the current common practice actually assumes that hydrological processes have ergodicity. This then brings up a question: do the hydrological processes really have ergodicity, just as being assumed in practice? How can one justify a process having ergodicity? The desirability of doing this has been realized for a long time, but rigorous and practical approaches have yet been available.

To date, only limited discussions about the application of time series ergodicity (Domowitz and El-Gamal, 2001; MORVAI and WEISS, 2005) have been reported. Most studies of time series applications, such as in the fields of hydrology, hydrodynamics, and noise (Jiang and Zheng, 2005; Oliveira et al., 2006; Veneziano and Tabaei, 2004), discuss statistic characteristics simply by assuming time series having ergodicity without justifying this assumption with a rigorous approach. There have been a few discussions concerning ergodicity in the field of hydrological research. (Liu, 1998) assumed that ergodicity exists between the spatial distribution and the temporal propagation of hydrological factors of a water exchange system, i.e., these processes are restricted by ergodicity. (Xia, 2005) used power-weighted Markov chains to predict "plum rain" intensity (an East Asian rainy season

usually lasting from June to July) and concluded that this process has ergodicity. In general, the ergodicity of time series refers to the ergodicity of stationary processes, which means that the process averaged over time behaves identical to the process averaged over space.

65 Until recently, particular studies on ergodic property analysis for hydrological processes have not yet been performed. However, the study of the ergodicity itself is not only significant but also indispensable because it is a fundamental presumption for many time series problems (Ding and Deng, 1988; Fiori and Janković, 2005; Hsu, 2003; Liu, 1998; Mitosek, 2000; Wang et al., 2004). This study proposes a practical approach for mean ergodic

70 property analysis using autocorrelation function (ACF) or Augmented Dickey Fuller (ADF) test and a radial basis function (RBF) neural network. The term ergodic and ergodic property or ergodicity are used mainly in mathematical physics, e.g., dynamics, and the theory of stationary stochastic processes. This study focuses on the ergodicity analysis for stationary stochastic processes which are commonly applied in hydrology.

75 **Definition of ergodicity**

A process is said to be ergodic if its statistical properties (such as its mean and variance) can be deduced from a single, sufficiently long sample (realization) of the process. A stochastic process shows ergodicity when its mean and covariance functions are ergodic, i.e., mean erogdicity and covariance ergodicity. Since the ergodicity of covariance function which

80 usually relates to forth-order moments of the process is difficult to verify, only the mean

4

ergodic property of the process (or sequence) is discussed in this paper. For a given stochastic sequence $\{\xi(t); t = 1, 2, ...\}$, and its sample mean sequence $\{M_T, T = 1, 2, ...\}$, $M_T = \frac{1}{T} \sum_{t=1}^{T} \xi(t)$, if the variance series of the sample mean sequence $\lim_{T \to \infty} D(M_T) = 0$, the sample mean series is ergodic and the process $\{\xi(t)\}$ is said to have mean ergodic property, where

85
$$D(M_T) = \frac{\sum_{t=1}^{T} (M_t - \bar{M}_t)^2}{T} \tag{1}$$

where $\bar{M}_t$ is the mean of M_T.

It has been proved that only stationary processes could have ergodicity (Davis et al, 1994). Stationarity implies that the statistical parameters of the series computed from different samples do not change except due to sampling variations. A time series is said to be

90   strictly stationary if its statistical properties do not vary with changes of time origin. A less strict type of stationarity is called weak stationarity or second-order stationarity where the first- and second- order moments depend only on time differences (Chen and Rao, 2002). In nature, strictly stationary time series does not exist, and weakly stationary time series is practically considered as stationary time series. In addition to the stationarity, another

95   necessary condition for ergodicity analysis is that the samples from the single realization should be taken from a large enough period of time.

A practical approach to ergodic property analysis

Currently there are no particular statistic tests designed for ergodic property analysis; we, therefore, perform the mean ergodicity analysis based on its definition and demonstrate a

100 practical approach with a series of case studies using monthly precipitation data series

collected from two rain gauge stations located in China and one in the US respectively.

Whereas the definition of mean ergodicity simples and straightforward, the practical analysis

of mean ergodicity can be complicated. As discussed in the last section that a stochastic

process is not ergodic unless stationary, a stationary test for the data series representing a

105 stochastic process is then necessary as a prerequisite for further ergodicity analysis. Another

challenge lies in the fact that the infinitely long data series required by the definition of mean

ergodicity cannot be achieved in reality. This challenge can be overcome by extending the

data series using approaches such as a reliable artificial intelligence approach. We propose to

solve this challenge by predicting the $D(M_T)$ series using artificial intelligence approach, for

110 example, Radius Basis Function (RBF) neural network, assuming that the predicted $D(M_T)$

series represents the characteristics of the population data series.

Stationarity test

Given that the commonly used statistical inference is no longer valid for a non-stationary

data series, it is necessary to examine the stationarity of a data series. The standard method

115 for stationarity test is a unit root test, i.e., a time series is stationary if there exists a unit root.

The stationarity of a stochastic process is determined by the roots of its characteristic

function. If all the characteristic roots are located outside of the unit element, then the process

is stationary, whereas, the process is non-stationary if one or more roots are on or within the

unit element or circle with unit radius. If a characteristic root has a value of unit, it is called

120 unit root. Dickey Fuller (DF) test and Augmented Dickey Fuller (ADF) test are two

commonly used unit root test. ADF test, actually an extension of DF test, eliminates the autocorrelation of residues by increasing the lags of the variable of a time series.

Without loss of generality, DF test can be illustrated with a simple AR (1) process. Consider a stochastic process,

125

$$x_t = \rho x_{t-1} + \mu_t \tag{2}$$

where μ_t is the white noise. If $|\rho| < 1$, the data series is stationary, while when $|\rho| > 1$, the series is non-stationary. In practice, Equation (2) can be rewritten as

$$\Delta x_t = \gamma x_{t-1} + \mu_t \tag{3}$$

where Δ representing differential, and $\gamma = \rho - 1$. The DF test is to test the following null
130   hypothesis,

$$H_0 : \gamma = 0 \quad H_1 : \gamma < 0 \tag{4}$$

For a time series where the random disturbance might be destructed by its high-order time lag, e.g., a $AR(p)$ process, ADF test is then to test the null hypothesis described by Equation (4) for the $AR(p)$ process as follows,

135

$$\Delta x_t = \gamma x_{t-1} + \xi_1 \Delta x_{t-1} + \xi_2 \Delta x_{t-2} + ... + \xi_{p-1} \Delta x_{t-p+1} + \mu_t \tag{5}$$

Besides DF or ADF test, the stationarity of a data series can also be determined by evaluating its autocorrelation. For a series variables $\{\xi(t); t = 1, 2, ... n\}$, its autocorrelation coefficient function (ACF) is defined as,

$$r_k = \frac{\sum_{t=1}^{n-k} \left(\xi_t - \bar{\xi} \right) \left(\xi_{t+k} - \bar{\xi} \right)}{\sum_{t=1}^{n} \left(\xi_t - \bar{\xi} \right)^2} = \frac{Cov[\xi_t, \xi_{t+k}]}{Var[\xi_t]} \tag{6}$$

140 where $\bar{\xi}$ is the mean. If the ACF rapidly approaches 0 (i.e. falls into the stochastic domain),

the time series is stationary; otherwise it is non-stationary (Cline and Pu, 1998, 1999).

RBF neural network

Since the $D(M_T)$ series is nonlinear, we adopt the RBF neural network approach (See

(Nørgaard, 2000) for detail) to extend the data series for predicting the trend of $D(M_T)$. The

145 RBF neural network is a well-performing forward neural network model. It has high

computational simplicity and extrapolation capacity, and can provide the network with strong

nonlinear projecting capability (Alp and Cigizoglu, 2007). The calculation required by RBF

neural network is relatively small; even with no more than a few cells, one can get a good

approximation as long as the center is properly selected. On the other hand, because data with

150 ergodic properties are unlikely to oscillate dramatically (Zhou et al., 2001), the RBF neural

network is a good choice to predict the future changes of ergodic properties.

The original data series $\{\xi(t)\}$ is first normalized by $\xi^* = \xi / \max(\xi)$, forming a new data

series $\{\xi^*(t)\}$. A 3-layer RBF network is then constructed following (Nørgaard, 2000), with

the *n1* neurons in the input layer and *m* neurons in the output layer. In this RBF neural

155 network model, the first *n1* data of $\{\xi^*(t)\}$ from the observation series can then be used to

predict the *n1*-th data; this process can be continued, predicting a longer series of ξ^* data.

A practical approach for ergodicity analysis

The following procedure for ergodicity analysis is then proposed for a practical analysis

of ergodicity of a data series: i) Perform the stationary analysis for the data series by

160   evaluating its autocorrelation function. A data series has no ergodicity unless it is stationary.

ii) Calculate the sample mean value series $\{M_T, T = 1,2,...\}$, $M_T = \frac{1}{T}\sum\limits_{t=1}^{T}\xi(t)$, and the variance

series $D(M_T)$ of mean value M_T; iii) Simulate $D(M_T)$ using approach such as Radius

Basis Function (RBF) neural network, predict the trend of $D(M_T)$ with time T approaching

∞, and iv) determine whether the original series $\{\xi(t)\}$ has mean ergodicity according to

165    $\lim\limits_{T\to\infty} D(M_T)$. It should be noted that using the proposed approach, we are trying to examine

whether the assumption of mean ergodicity is consistent with the time series rather than

mathematically prove the ergodicity of the time series.

Ergodicity analysis of precipitation process

Ergodicity analysis is performed for the monthly precipitation data series of three sites to

170   demonstrate the proposed ergodicity analysis approach, including Lanzhou of Gansu

Province and Ankang of Shan'xi Province, China, and Newberry Michigan, USA.

Ergodicity analysis of precipitation series of Newberry, USA

A mean ergodicity analysis is performed for each individual monthly precipitation data

series of the 121-year precipitation data (1893-2013) collected from NOAA Newberry

175   Correctional Facility, MI, USA (46.35°N 85.5°W, Elev. 240 m). The climatic region of

Newberry is typified by large seasonal temperature differences, with warm to hot (and often

humid) summers and cold (sometimes severely cold) winters. According to the Köppen

Climate Classification system, it has a humid continental climate. Newberry has an average

of 820.3 mm precipitation per year. Its driest weather is in February with an average of 40.1

180   mm of precipitation, and wettest weather is in August with an average of 94 mm precipitation.

The statistics of each individual monthly precipitation series of Newberry are given in Table 1. The ACF plots, as shown in <mark>Table 2 System clustering of the precipitation series at Ankang, China</mark>

Clusters	Calendar months	CV	Stationarity	Ergodicity
Group 1	$\begin{cases} Jan,\ Feb,\ Mar, \\ Nov,\ Dec \end{cases}$	1.0384	No	No
Group 2	$\begin{cases} Apr,\ May,\ Jun, \\ Jul,\ Oct \end{cases}$	0.6554	Yes	No
Group 3	$\{Aug\}$	0.5938	Yes	Yes
Group 4	$\{Sep\}$	0.6357	Yes	No

185   Figure 1, and the ADF test indicate that all the 12 individual monthly precipitation data

series at Newberry are stationary. The M_T and $D(M_T)$ for each monthly precipitation data

series are calculated and plotted as shown in Figure 2. Apparently, $T \to \infty$ cannot be

guaranteed due to the limited number of samples and thus the trend of $D(M_T)$ cannot

consequently be determined directly from the data series if it is not long enough. The monthly

190   mean precipitation data series of each data group is first normalized by $\xi^* = \xi / \max(\xi)$, and

then extended by using RBF neural network. The mean series of the new data series,

$M_T = \dfrac{1}{T}\sum_{t=1}^{T}\xi^*(t)$, and its variance series $D(M_T)$ can be calculated. The $D(M_T)$ for each

extended monthly data series are plotted as shown in Figure 3, which shows that only the

$D(M_T)$ of the monthly precipitation data series of February, May, July, and December, have

195   a clear trend approaching zero, thus these four months' precipitation being ergodic.

Ergodicity analysis of precipitation series of Lanzhou, China

Fifty years (1951-2000) of monthly precipitation data are collected from Lanzhou Rain

Gauge Station (103.70° E, 35.90° N) in the Yellow River basin of China. The statistics of

each individual monthly series are given in **Table 1**. The autocorrelation of each monthly

200   precipitation data series indicates that all the series data of Lanzhou rain gauge station are

stationary. The M_T and $D(M_T)$ of data series for Lanzhou rain gauge station are then

calculated; and the $D(M_T)$ of the predicted data series for each monthly precipitation data

series by RBF neural network shows that only the $D(M_T)$ of precipitations of March, July

and August, approach 0. Therefore, these monthly precipitation series has mean ergodicity.

Ergodicity analysis of precipitation series of Ankang, China

Similarly, the ergodicity analysis is performed for the monthly mean precipitation data series of each calendar month for the Ankang rain gauge station. Seventy years (1929-1998) of precipitation data from Ankang Rain Gauge Station (109.03° E, 32.72° N) in the Han River basin of China are collected for the ergodic property analysis. The statistics of each individual monthly series are given in Table 1. The stationarity analysis by evaluating the autocorrelation of each monthly data series shows that the monthly precipitation data series of each month is stationary. The mean ergodicity analysis is then performed for each monthly data series with extending the data series with RBF neuronal network. The results indicate that the monthly precipitation of May, June, and August is ergodic.

Results and Discussions

The coefficient of variance (CV) has been widely used to measure dispersion of a data series. The more concentrated the distribution of a random variable is, the more obvious is its regularity, and vice versa. The coefficients of variance for each monthly precipitation data series of Lanzhou rain gauge station, Ankang rain gauge station, and Newberry are calculated, as shown in Table 1, respectively. As all the monthly data series are stationary in our study, we synthesize a combined stationary and nonstationary data series by clustering the monthly data series of Ankang station into four classes, as shown in Table 2, in order to investigate the non/stationarity and ergodicity simultaneously. The stationarity test and mean ergodicity analysis are then performed to the clustered data using the proposed methodology.

225 The stationarity analysis by evaluating the autocorrelation of each group of data series shows

that the {Ankang: Jan, Feb, Mar, Nov, Dec} series is non-stationary; whereas all the other

three groups of data series are stationary. It can be seen that the coefficient of variance for

clustered monthly data series of Ankang that are stationary are smaller than that of non-

stationary monthly data series. This indicates that the monthly precipitation with small

230 coefficients of variance has more tendency to be stationary, or more regular.

Although a small coefficient of variance provides the representing data series with more

tendency to be stationary, it does not give a clear indication to the ergodicity of the data

series. The coefficient of variance of precipitation time series with ergodicity is not

necessarily smaller than those without ergodicity. Among the monthly precipitation data

235 series of Newberry, the data series of {Newberry: September} has the smallest coefficient of

variance but it does not have an ergodic property. Moreover, although the coefficients of

variance of {Newberry: September}, {Lanzhou: September}, and {Ankang: April}, are

smaller than {Newberry, July}, {Lanzhou: August}, and {Ankang: June}, respectively, the

latter ones have ergodicity while the former ones do not. Furthermore, an ergodic process is

240 stationary while the converse may not be true. The stationarity of a data series is the

prerequisite of its ergodicity, rather than a guarantee to ergodicity; there are stationary

processes which are not ergodic. In other words, a process with ergodicity is necessarily

stationary in the strict sense. Therefore, neither the coefficient of variance nor the stationarity

test, which is commonly performed in time series analysis, can take over the ergodicity

245 analysis in order to make sure the statistics of a reality, such as the mean, can be safely used

as those of its population.

Comparison of the ergodicity of the monthly precipitation data series of May and June of Ankang analyzed by their monthly data series and the clustered data series indicates that the ergodic property could change when new data is introduced. The analysis of May precipitation in Ankang shows it has ergodicity; whereas, the analysis using the clustered data series {Ankang: May, Jun, Sep} indicates that this data series does not have mean ergodic property. This difference can be explained that the {Ankang: May} monthly precipitation may become non-ergodic when new data, for example {Ankang: Jun, Sep}, is introduced, which actually can considered as new observations of {Ankang: May} and the variation of {Ankang: May} and {Ankang: Jun, Sep} could be owing to the change of natural or man-made factors that affect precipitation processes.

A linear trend analysis is also performed following (Vamos and Craciun, 2012) for some of the ergodic monthly precipitation data series of the three rain gauge stations, as shown in Figure 4. The August precipitation at neither Lanzhou nor Ankang station shows obviously periodic. The August precipitation at Lanzhou rain gauge station shows an overall decreasing trend while the August precipitation at Ankang rain gauge station shows remains almost stable around its mean value. In Newberry, only the precipitation of May shows a relatively stable trend, the precipitation of February has a clear decreasing trend since 1970s, and the precipitation of both July and December has an increasing trend. However, according to the ergodicity analysis, the precipitation of those months having ergodicity in the three rain

14

gauge stations, in the long run will fluctuate around their mean value rather than keep varying as shown in Figure 4.

In this study, the 50 years of monthly precipitation data series of Lanzhou rain gauge station, the 70 years of precipitation data series of Ankang rain gauge station are insufficient, and even the 121 years of monthly precipitation data from Newberry, MI, USA, is insufficient for direct analysis of ergodic. The BRF neural network is used to extend the data series for the practical analysis of ergodicity. The observation data on many other hydrological processes are expected not to be long enough for direct ergodicity analysis; we, therefore, suggest using some other reliable tools, such as the BRF neural network used in this study, to extend the data series and perform the ergodicity analysis based on the observation and prediction data series. Before a more rigorous method becomes available, such approach of using extended data series by a "black-box" type of method might be controversial, but, we still believe, helpful at least to give a sense whether the data series of analysis has erogdic property.

Conclusions

Ergodic property analysis for hydrological processes is difficult but worthy of discussion. One may argue that whether a data series representing a hydrological process is ergodic, it does not actually affect the practice of analysis of this hydrological process, therefore, the test of ergodicity can be completely neglected. Some researchers (Duan and

285 Goldys, 2001; Koutsoyiannis, 2005; Liu, 1998), however, have pointed out that hydrological

processes may have ergodic properties although no particular ergodicity analysis was

performed in these works. This study presents a practical approach to analyze the mean

ergodicity of hydrological processes, which bridges the concept of ergodicity and its

application in hydrological process analysis. This approach primarily includes the stationarity

290 test of the data series through its ACF or ADF test, avoiding the difficulty in analyzing the

stationarity of the data series directly from its definition, the extension of the length of the

data series, via the RBF network in this study, and the ergodicity analysis based on the

sample mean sequence and its variance series. Three case studies, the ergodicity analysis for

the monthly precipitation of Lanzhou in the Yellow River Basin of Chin, Ankang in the Han

295 River basin of China, and Newberry, MI, USA, are conducted using the proposed approach.

Our research reveals that the precipitations of March, July, and August in Lanzhou, and

May, June, and August in Ankang have ergodicity; therefore the stochastic and statistical

analysis of the precipitation of these months based on the observations (sample) in these two

stations are expected more reliable than the analysis for any other calendar months'

300 precipitation in the two stations. The ergodicity analysis of precipitation data series of each

individual month in Newberry, MI, USA, which has a relatively long observation history

indicates that the precipitation of February, May, July, and December show ergodic property,

although not all of the precipitation of these months has a tendency converging to its mean

value, respectively.

305   This study focuses mainly on the mean ergodicity analysis; approaches to the covariance

ergodicity analysis of hydrological processes need to be developed in the future, which would

provide us more useful information. In addition, as discussed, the application of ergodicity

seems still controversial although its concept and properties have been applied commonly in

hydrology by presuming hydrological processes automatically having ergodicity. More

310   discussion and methodologies on ergodicity analysis would certainly bridge the gap between

its concept and application.

Acknowledgements

References

Alp, M. and Cigizoglu, H.: Suspended sediment load simulation by two artificial neural
network methods using hydrometeorological data, Environ. Model. Softw., 22(1), 2–13,
doi:10.1016/j.envsoft.2005.09.009, 2007.

Chen, H. and Rao, A.: Testing Hydrologic Time Series for Stationarity, J. Hydraul. Eng., 7(2),
129–136, 2002.

Chick, S., Shortle, J., Van Gelder, P. and Mendel, M. B.: A model for the frequency of
extreme river levels based on river dynamics, Struct. Saf., 18(4), 261–276,
doi:10.1016/S0167-4730(96)00020-3, 1996.

Cline, D. B. H. and Pu, H. H.: Verifying irreducibility and continuity of a nonlinear time
series, Stat. Probab. Lett., 40(2), 139–148, doi:10.1016/S0167-7152(98)00081-9, 1998.

Cline, D. B. H. and Pu, H. H.: Stability of nonlinear AR(1) time series with delay, Stoch.
Process. their Appl., 82(2), 307–333, doi:10.1016/S0304-4149(99)00042-3, 1999.

Ding, J. and Deng, Y.: Stochastic Hydrology, Press of Chengdu Technology University, Chengdu, China., 1988.

330   Domowitz, I. and El-Gamal, M. a.: A consistent nonparametric test of ergodicity for time series with applications, J. Econom., 102(2), 365–398, doi:10.1016/S0304-4076(01)00058-6, 2001.

Duan, J. and Goldys, B.: Ergodicity of Stochastically Forced Large Scale Geophysical Flows, , doi:10.1155/S0161171201012443, 2001.

335   Fiori, A. and Janković, I.: Can we determine the transverse macrodispersivity by using the method of moments?, Adv. Water Resour., 28(6), 589–599, doi:10.1016/j.advwatres.2004.09.009, 2005.

Hsu, K. C.: The influence of the log-conductivity autocovariance structure on macrodispersion coefficients, J. Contam. Hydrol., 65(1-2), 65–77, doi:10.1016/S0169-
340   7722(02)00231-0, 2003.

Jiang, G.-P. and Zheng, W. X.: A simple method of chaos control for a class of chaotic discrete-time systems, Chaos, Solitons & Fractals, 23(3), 843–849, doi:10.1016/j.chaos.2004.05.025, 2005.

Koutsoyiannis, D.: Stochastic simulation of hydrosystems, in Stochastic Simulation of
345   Hydrosystems: Surface Water Hydrology, edited by J. H. Lehr and J. Keeley, pp. 421–430, Wiley, New York., 2005.

Liu, C.: Some Problems of Water Interaction, in The experiment of water interaction and compute analysis, edited by C. Liu and H. Ren, p. 341, Science Press, Beijing, China., 1998.

Mitosek, H. T.: On stochastic properties of daily river flow processes, J. Hydrol., 228(3-4),
350   188–205, doi:10.1016/S0022-1694(00)00150-5, 2000.

MORVAI, G. and WEISS, B.: Forward estimation for ergodic time series, Ann. l'Institut Henri Poincare Probab. Stat., 41(5), 859–870, doi:10.1016/j.anihpb.2004.07.002, 2005.

Nørgaard, M.: Neural network based system identification toolbox., 2000.

Oliveira, T. F., Cunha, F. R. and Bobenrieth, R. F. M.: A stochastic analysis of a nonlinear
355   flow response, Probabilistic Eng. Mech., 21(4), 377–383, doi:10.1016/j.probengmech.2005.11.010, 2006.

Shahin, M., Oorschot, H. J. L. and Lange, S. J.: Statistical Analysis in Water Resources Engineering, Aa Balkema, The Netherlands, Rotterdam., 1993.

Vamos, C. and Craciun, M.: Automatic trend estimation, Springer Netherlands. [online]
360   Available from: https://books.google.com/books?id=LLYKSKkkfwYC, 2012.

Veneziano, D. and Tabaei, A.: Nonlinear spectral analysis of flow through porous media with
isotropic lognormal hydraulic conductivity, J. Hydrol., 294(1-3), 4–17,
doi:10.1016/j.jhydrol.2003.10.025, 2004.

Wang, H., Song, Y., Liu, C. and Chen, J.: Application and issues of chaos theory in
365   hydroscience, Adv. Water Sci., 15(3), 400–407, 2004.

Xia, L.: Prediction of plum rain intensity based on index weighted Markov chain, J. Hydraul.
Eng., 36(8), 988–993, 2005.

Zhou, P. L., Tao, X. L., Fu, Z. Q., Peng, H. and Wang, Y. X.: Precipitation prediction based
on improved radial basis function neural networks, J. Chinese Comput. Syst., 22(2), 244–246,
370   2001.

Table 1. Statistics of monthly precipitation data of each case study site

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Newberry, USA	Min. (mm)	0.0	0.0	0.0	0.0	0.0	0.0	17.8	3.8	19.3	6.4	2.5	0.0
	Max. (mm)	150.6	95.0	251.5	264.2	266.7	211.6	190.5	209.0	204.7	157.7	169.9	117.3
	Mean (mm)	48.7	33.9	45.7	49.2	68.5	77.7	73.5	77.4	87.4	75.0	62.5	50.2
	Std. (mm)	23.5	20.4	34.3	34.1	37.1	38.4	39.0	41.0	38.6	35.2	33.3	24.0
	CV	0.5	0.6	0.8	0.7	0.5	0.5	0.5	0.5	0.4	0.5	0.5	0.5
	Stationarity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Ergodicity	No	Yes	No	No	Yes	No	Yes	No	No	No	No	Yes
Lanzhou, China	Min. (mm)	0	0.0	0.0	0.6	1.2	11.0	14.3	7.1	8.5	0.0	0.0	0.0
	Max. (mm)	8.2	9.2	23.3	68.3	124.7	109.0	142.1	236.2	100.1	60.0	20.8	8.8
	Mean (mm)	1.5	2.3	8.3	16.6	35.4	39.5	63.3	78.3	45.7	23.1	3.5	1.1
	Std. (mm)	1.9	2.5	6.1	13.2	25.2	23.0	31.3	44.0	24.9	17.0	4.9	1.9
	CV	1.3	1.1	0.7	0.8	0.7	0.6	0.5	0.6	0.5	0.7	1.4	1.7
	Stationarity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Ergodicity	No	No	Yes	No	No	No	Yes	Yes	No	No	No	No
Ankang, China	Min. (mm)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Max. (mm)	26.0	39.0	104.0	131.0	168.0	319.0	391.0	326.0	391.0	185.0	88.0	71.0
	Mean (mm)	5.5	11.5	32.4	62.9	81.5	109.3	118.1	116.1	129.8	71.3	29.3	7.6
	Std. (mm)	4.8	10.1	20.1	32.0	38.8	67.8	81.6	69.9	83.3	43.8	18.7	10.2
	CV	0.9	0.9	0.6	0.5	0.5	0.6	0.7	0.6	0.6	0.6	0.6	1.3
	Stationarity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Ergodicity	No	No	No	No	Yes	Yes	No	Yes	No	No	No	No

375

380

Table 2 System clustering of the precipitation series at Ankang, China

Clusters	Calendar months	CV	Stationarity	Ergodicity
Group 1	$\left\{ \begin{array}{l} Jan,\ Feb,\ Mar, \\ Nov,\ Dec \end{array} \right\}$	1.0384	No	No
Group 2	$\left\{ \begin{array}{l} Apr,\ May,\ Jun, \\ Jul,\ Oct \end{array} \right\}$	0.6554	Yes	No
Group 3	$\{Aug\}$	0.5938	Yes	Yes
Group 4	$\{Sep\}$	0.6357	Yes	No

Figure 1. ACF of monthly precipitation data series of Newberry, MI, USA

Figure 2. M_T (mm) and $D(M_T)$ (mm) of precipitation data series of Newberry, MI, USA

Figure 3. Extended $D(M_T)$ (mm) series of precipitation by RBF neural network at Newberry, MI, USA

Figure 4. The precipitation data (dotted in figures) of ergodic months and their trend (straight lines) of Newberry
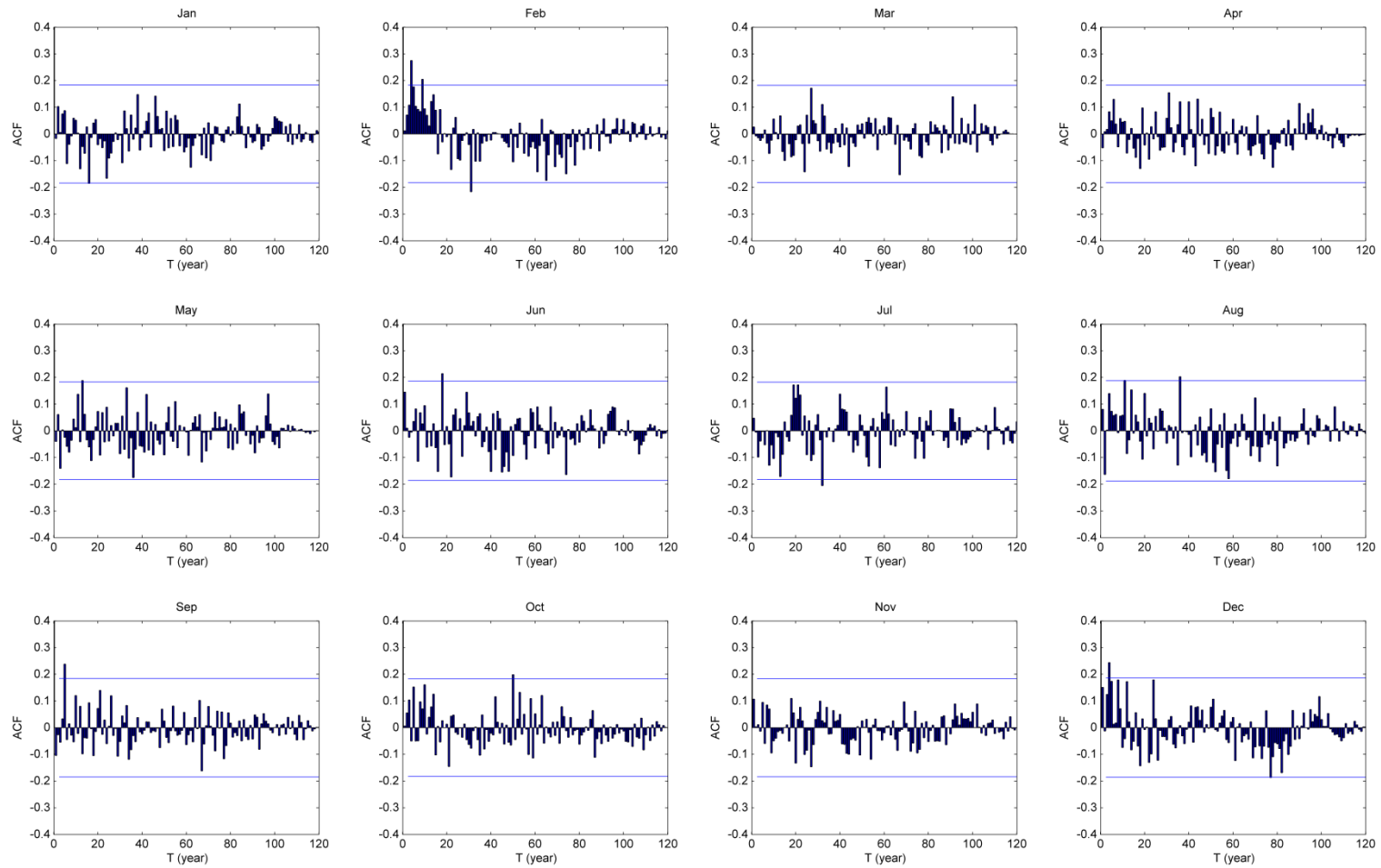
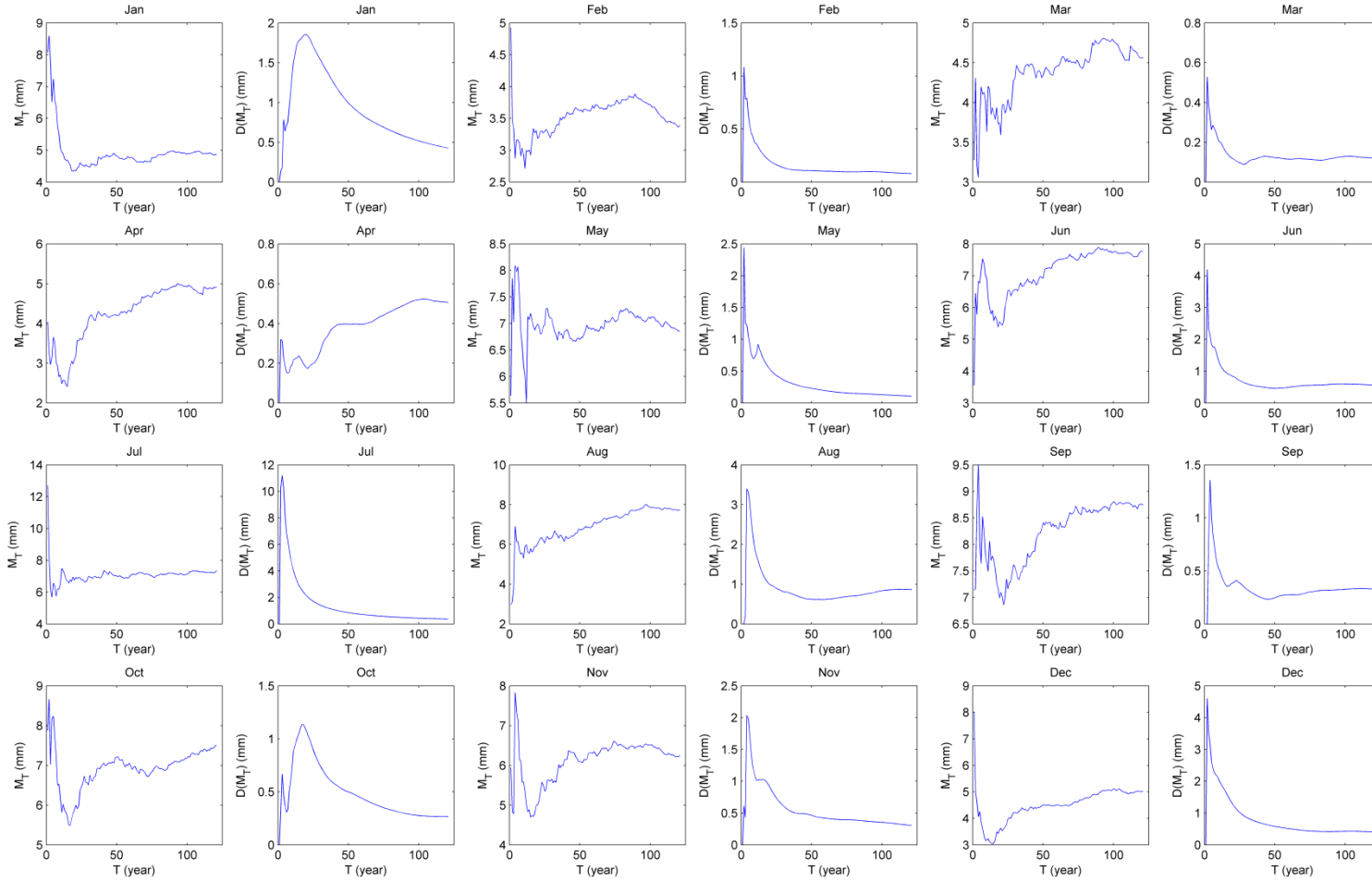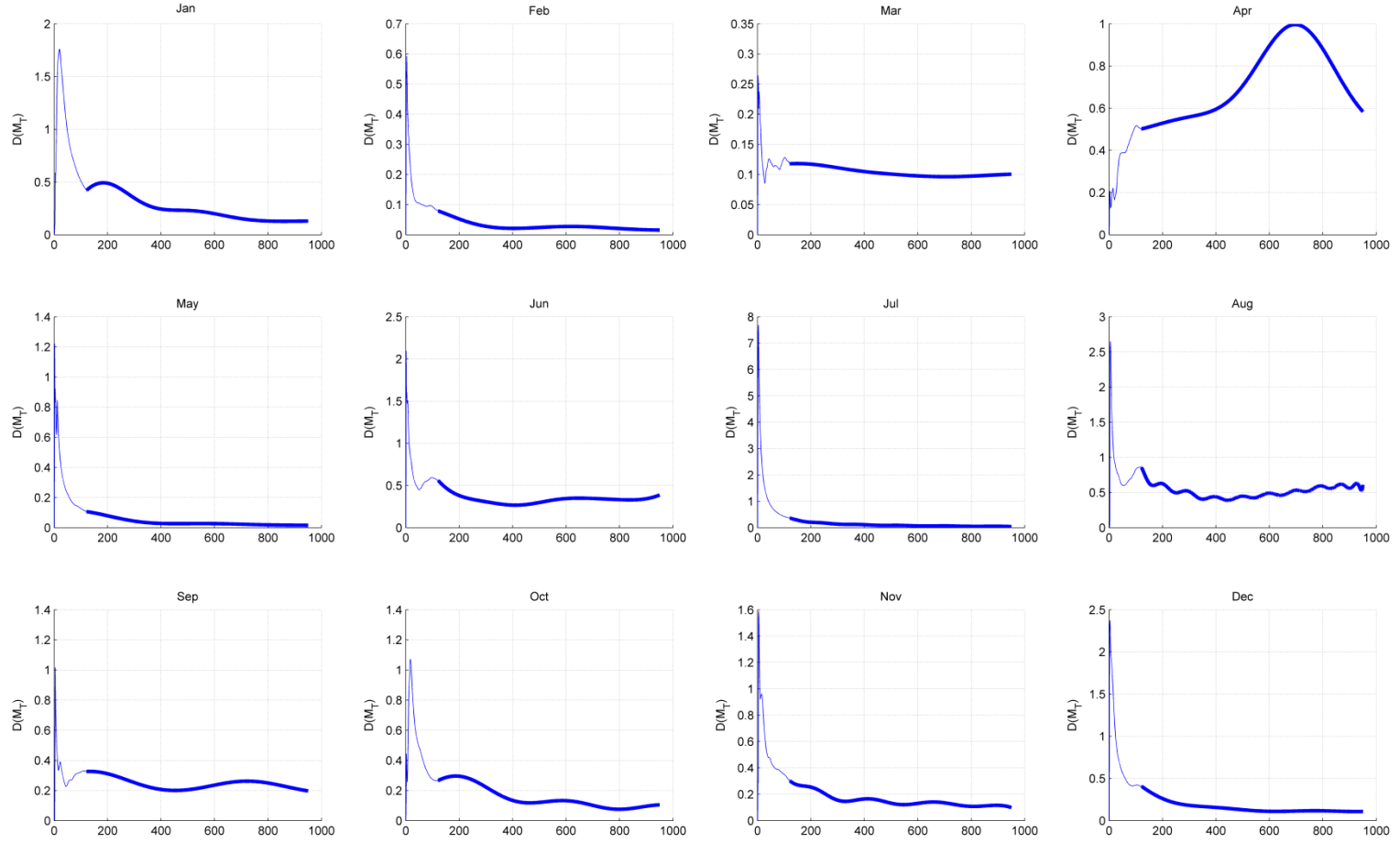**Figure 1.** ACF of monthly precipitation data series of Newberry, MI, USA

Figure 2. M_T (mm) and $D(M_T)$ (mm) of precipitation data series of Newberry, MI, USA

Figure 3. Extended $D(M_T)$ (mm) series of precipitation by RBF neural network at Newberry, MI, USA
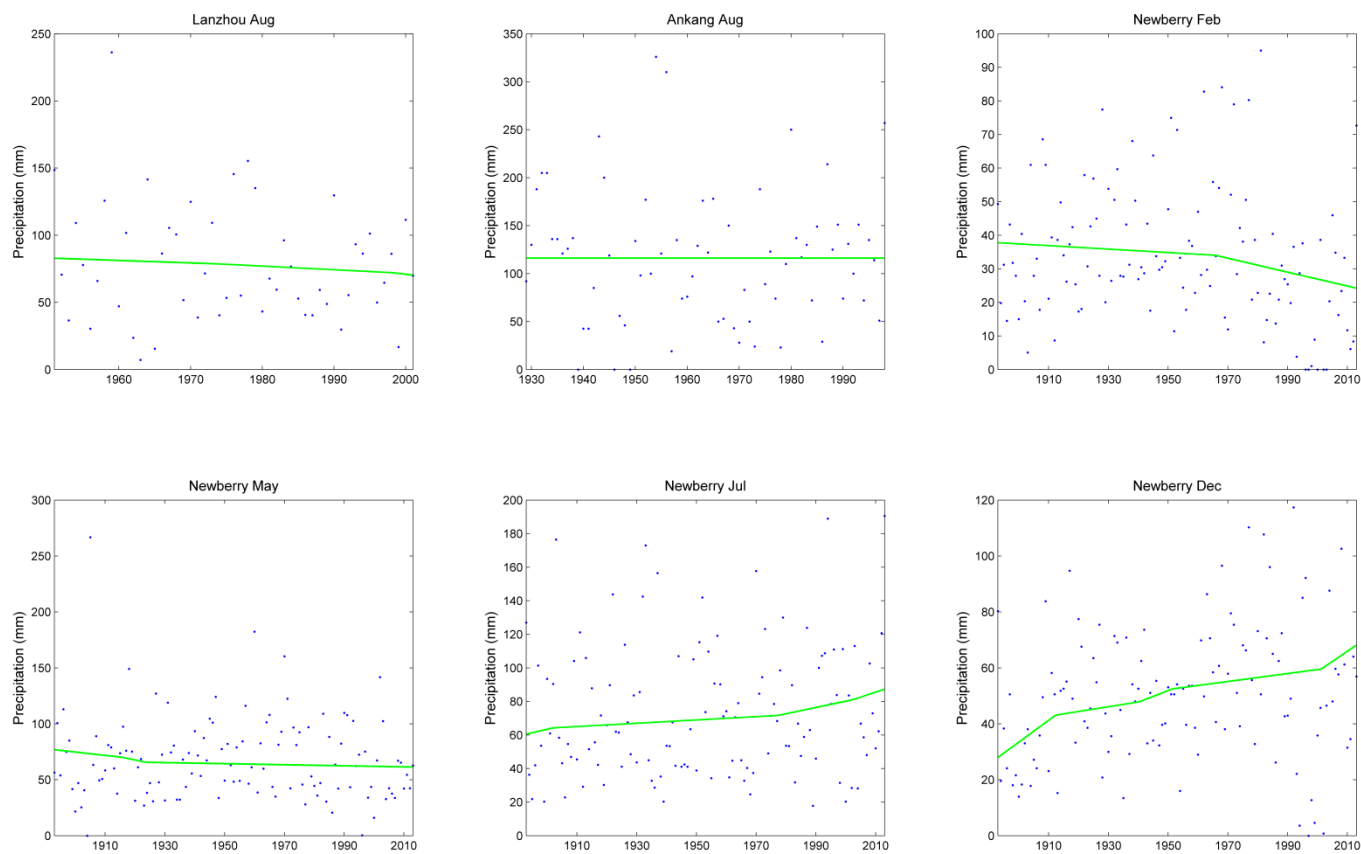
Figure 4. The precipitation data (dotted in figures) of some ergodic months and their trend (straight lines) in Lanzhou, Ankang, and Newberry

400