Nonstationary Modelling of Annual Discharge over the Tarim River Headstream Catchment, China

Yunbiao Wu¹, ², * and Lianqing Xue¹, ², ³

¹College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China
²Hohai University Wentian College, Maanshan 243031, China
³College of Water Conservancy and Architectural, Shehezi University, Shehezi 832003, China

*Corresponding author: wyb_0018@163.com

Abstract. In the changing environment, the hydrological time series no longer satisfies the assumption of stationarity. In this study, the changes in annual discharge at five hydrological stations over the Tarim River (TR) headstream catchment are analysed based on the hydrologic data during last 50 years, and the Generalized Additive Models in Location, Scale and Shape (GAMLSS) are used as a tool for modeling annual discharge series under no stationary condition with time and meteorological variables. The results show that: (1) There is an evident influence of meteorological variables on the mean and variance of annual discharge in the Tarim headstream catchment; (2) Compare to the stationary model, the no stationary models that incorporate time and meteorological variables have a significant improvement in description of the changes in the annual discharge series; (3) The stationary model may lead to underestimate or overestimate annual discharge under changing environment, and it is wrong to assume stationarity for statistical modeling of estimating water resource. The estimating water resource should be described by the no stationary models with more explanatory variables, such as time and meteorological variables. This study can serve as a reference for regional planning and management of water resources in the Tarim River Basin (TRB).

1. Introduction

Hydrological frequency analysis usually requires statistical analysis of hydrological variables. Traditionally, this statistical analysis relies on the assumption that the hydrologic data are independent and identically distributed [1]. However, the assumption of stationarity has been widely questioned due to the climate change and intensifying human activities, such as the construction of water reservoirs and urbanization [1-2]. Under this condition, the conclusions drawn based on the stationarity assumption were questionable. Milly et al. [2] asserted that stationarity is dead and should not serve as a default assumption in water resource risk assessment and planning.

Many studies have shown that annual average temperature and precipitation in the TRB have increased significantly [3-5], and stream flow in three headstreams of the TRB also significantly increased, which might be attributed to the increased snow-melt water caused by rising temperature [5-
6]. Gu et al. [7] found that the influence of the occurrence rates of floods by temperature and precipitation changes more sensitive than the climate indices in the TRB. The literature above indicates that annual discharge in the TR headstream catchment shows nonstationarity characteristics, but few studies have taken into consideration of this no stationary in modeling. This paper is to address the no stationary modeling of annual discharge in the TR headstream catchment, and demonstrate that the incorporation of meteorological variables may result in appropriate covariates to describe changes in annual discharge. Furthermore, we will also show the differences over time in estimated quantiles by considering and excluding nonstationarity with the purpose of stating the importance of using no stationary models.

2. Materials and methods

2.1. Study area and Data sources

The TRB is the largest inland river basin in China. It is a typical arid and semi-arid area with annual precipitation less than 50 mm, but the potential evaporation is more than 2000 mm per annum. Water of the TR is only supplied by discharge from precipitation and glacial melt water in the mountain headstreams [4]. Three headstreams (Aksu, Yarkant and Hotan River) converge at the Alar hydrological station. Daily streamflow data of five hydrological stations were collected from the Tarim River Basin Management Bureau. Meteorological data were collected from the China Meteorological Data Sharing Service System (http://data.cma.cn). Table 1 shows the information on the hydrological stations considered in this study in the headstream catchment of TRB. Table 2 shows the information on the meteorological stations considered in this study in the headstream catchment of TRB.

**Table 1. Information on the hydrological stations and MK test of annual discharge.**

| River      | Hydrological station | time series | Annual mean discharge (m³/s) | ZMK |
|------------|----------------------|-------------|-----------------------------|-----|
| Hotan River| Tongguzlok           | 1960-2013   | 26457.85                    | 1.000          |
| Yarkand River| Yuzmenlek          | 1960-2011   | 10026.43                    | 4.017**        |
|            | Kaqung               | 1960-2013   | 77929.67                    | 2.164**        |
| Aksu River | Sharikilank          | 1960-2011   | 33489.03                    | 4.032**        |
| Tarim River| Alar                 | 1960-2014   | 52807                       | -1.525*        |

**Significant at the 0.05 level, * Significant at the 0.1 level. ZMK value is the test statistics of MK.**

**Table 2. Information on the meteorological stations and MK test of meteorological variables.**

| Meteorological station | time series | Pmean (mm) | ZMK | Tmean (°C) | ZMK |
|------------------------|-------------|------------|-----|------------|-----|
| Hotan                  | 1960-2013   | 39.42      | 0.962 | 12.82      | 5.595** |
| Pishan                 | 1960-2011   | 52.71      | 0.884 | 12.29      | 4.238** |
| Tashikurgan            | 1960-2013   | 75.06      | 2.343** | 3.69      | 3.551** |
| Ahqi                   | 1960-2011   | 216.88     | 3.007** | 6.62      | 4.238** |
| Alar                   | 1960-2014   | 49.06      | 1.067 | 10.79      | 1.278 |

**Significant at the 0.05 level, * Significant at the 0.1 level. Pmean represent mean annual precipitation, Tmean represent mean annual temperature. ZMK value is the test statistics of MK.**

2.2. Methods

2.2.1. Trend analysis. The non-parametric Mann-Kendall (MK) statistical test is firstly used to assess the trend of hydrology and meteorology data. Details can be found in references [8-9].

2.2.2. Generalized Additive Models for Location, Scale, and Shape. In a GAMLSS model [10], the observations \( y_i \) for \( i=1,2,...,n \) are assumed to be independent and have probability density function \( f_i(y_i | \theta^i) \) with \( \theta^i = (\theta_{1i}, \theta_{2i}, ..., \theta_{pi}) \) a vector of \( p \) distribution parameters accounting for location, scale
and shape. The distribution parameters are related to the design matrix of the selected covariate using the monotonic link function $g_k(\cdot)$ for $k = 1, 2, ..., p$. The formula is as follows:

$$g_k(\theta_k) = X_k \beta_k + \sum_{j=1}^{J_k} h_{jk}(x_{jk})$$

where $\theta_k$ are the parameter vectors of length $n$, $X_k$ is a matrix of explanatory variables (i.e. covariates) of order $n \times J_k$, $\beta_k$ is a parameter vector of length $J_k$ and $h_{jk}(\cdot)$ represents the functional dependence of the distribution parameters on explanatory variables $x_{jk}$.

Five widely used two-parameters distribution functions are employed to model discharge data: Weibull (WEI); Gumbel (GU); Gamma (GA); Logistic (LO) and Lognormal (LOGNO). The Akaike Information Criterion (AIC) was used to select the distribution function. Three different models were used for analyzing the annual discharge in the study stations: the stationary model (Model 0), where the distribution parameters do not depend on covariates (parameters are constant in time); the time-varying model (Model 1), where the distribution parameters vary as a function of time only; and the model (Model 2) that incorporates external covariates, where the distribution parameters can vary as a function of meteorological variables.

3. Results

3.1. Trend analysis

Annual discharge and annual meteorological data were tested by MK test. The results showed in Table 1 and Table 2. From Table 1 we can see the four hydrological stations in the mountain headstreams showed a positive trend of annual discharge, and three stations including Yuzmenlek, Kaqung, and Sharikilank had a significant increasing trend in annual discharge at the 0.05 level, but the station Alar which at the junction of the headstreams had a significant decreasing trend in annual discharge at the 0.1 level. From Table 2 we can see that all the five meteorological stations showed a positive trend of annual precipitation (P) and annual temperature (T), two stations including Tashikurgan and Ahqi had a significant increasing trend in annual precipitation at the 0.05 level, and four stations including Hotan, Pishan, Tashikurgan and Ahqi had a significant increasing trend in annual temperature at the 0.05 level. All these results showed that the hydrological and meteorological time series in this region had no stationary characteristics.

3.2. Modeling with GAMLSS

We examine GAMLSS for modeling no stationary annual discharge records. Table 3 summarizes the selected distributions and the type of dependence of distribution parameters as a function of time for Model 1. When turn to examine the no stationary models that incorporate external covariates (i.e. meteorological variables), The meteorological data from Hotan, Pishan, Tashikurgan, Ahqi and Alar meteorological station are used to represent meteorological conditions at Tongguzlok, Yuzmenlek, Kaqung, Sharikilank and Alar hydrological station respectively. Table 4 summarizes the selected distributions, the significant covariates for each distribution parameter and the type of dependence of distribution parameters as a function of external covariates for Model 2. From the two tables, it can be seen that the GA, LOGNO and WEI distributions offer the best results in modeling annual discharge. The observed results show that temporal trends and external forcing can affect the behavior of the mean and the variance of the annual discharge.

From Table 3 we can see that the parameters of Model 1 in all the five stations include time dependence, which indicates that the models with time as covariate perform better than the models with constant parameters. Among the five stations with time dependence, Model 1 shows that four of them
except Alar station, the parameter $\theta_1$ (relates to the mean value of annual discharge series) includes time dependence, and two of this dependence is via linear trend functions, the other two via non-parametric smoothing functions, while at Alar station, the parameter $\theta_1$ is constant. The parameter $\theta_2$ (relates to the variance of annual discharge series) is independent of the time trends in four models, only one (Alar station) with linear dependence.

**Table 3.** Summary for the fitted models with time as the covariate: cs (t) indicates the dependence is via the cubic splines; t means linear dependence; and ct refers to a parameter that is constant.

| Station       | Distribution | $\theta_1$ | $\theta_2$ |
|---------------|--------------|------------|------------|
| Tongguzlok    | GA           | cs(t)      | ct         |
| Yuzmenlek     | LOGNO        | t          | ct         |
| Kaqung        | GA           | t          | ct         |
| Sharikilank   | GA           | cs(t)      | ct         |
| Alar          | WEI          | ct         | t          |

From Table 4 we can see that T is a significant covariate in parameter $\theta_1$ in 4 stations except Sharikilank, the dependence is via linear trend functions at Tongguzlok, Yuzmenlek, Kaqung, but via non-parametric smoothing function at Alar station. T is a significant covariate for only Yuzmenlek station in the parameter $\theta_2$. The main reason is that glacial snowmelt water is the main source for river feeding, temperature is an important influencing factor to glacial snowmelt, and the higher temperature can result in the more glacial snowmelt water in the mountain headstreams. A weak statistical significance is observed for P, which is an explanatory covariate at Yuzmenlek and Sharikilank for the $\theta_1$ parameter and Alar for the $\theta_2$ parameter, and the dependence is via linear trend functions except a non-parametric smoothing function in parameter $\theta_1$ at Sharikilank. It is indicated that annual discharge of the headstream catchment is remarkably sensitive to variations of meteorological variables. Table 4 also shows that the impact of precipitation on stream flow is much greater than that of temperature in the Aksu River, but the impact of temperature on stream flow is much greater than that of precipitation in the Hotan River.

**Table 4.** Summary for the fitted models with meteorological variables (P and T) as the covariate: cs (·) indicates the dependence is via the cubic splines; P and T means linear dependence; and ct refers to a parameter that is constant.

| Station       | Distribution | $\theta_1$ | $\theta_2$ |
|---------------|--------------|------------|------------|
| Tongguzlok    | GA           | T          | ct         |
| Yuzmenlek     | LOGNO        | P+T        | T          |
| Kaqung        | LOGNO        | T          | ct         |
| Sharikilank   | LOGNO        | cs(P)      | ct         |
| Alar          | WEI          | cs(T)      | P          |

3.3. **Results of no stationary models**

Figure 1 shows the observed values, the estimated median and the 5th and 95th percentiles for the 5 representative stations. The results obtained with no stationary models assuming temporal dependence only (Model 1) show a pattern of increasing trends in the median, the 5th and the 95th percentiles of the annual discharge at Yuzmenlek and Kaqung, while Tongguzlok presented a slight decreasing trend during the period 1960-1990, then a strong upward tendency is observed after 1990. At Sharikilank we observe an undulating behaviour, with two maxima (1970s and 2000s) and a minimum during the 1980s. At Alar station, we can notice a gradual decreasing trend. Obviously, Model 1 is able to capture the variability exhibited by the data. Figure 1 also shows the results under no stationary conditions by Model 2 with incorporation of meteorological variables. From Figure 1, we can see Model 2 can capture detailed and subtle changing properties of annual discharge variations and can mimic the influence of
climate variations on annual discharge. From the 5th and the 95th percentiles of the annual discharge we observe that the 90% confidence intervals of the annual discharge at Yuzmenlek and Alar stations obtained by Model 2 have a tendency to narrow, especially after 1990, this is different from Model 1. Furthermore, the median and the 5th and the 95th percentiles of annual discharge significantly increase near 1990 at Tongguzlok, Yuzmenlek and Sharikilank. These behaviors of runoff processes occur in the similar time intervals of temperature and precipitation which significantly increase near 1990. However, the temperature and precipitation of Alar station increase non-significantly, but the annual discharge at Alar has tended to significantly decrease. Obviously, both Model 1 and Model 2 show that the annual discharge in the mountain headstreams of the TRB have an increasing trend on the whole, especially after 1990. These results show that the annual discharge is sensitive to meteorological variables in the mountain headstreams.

Figure 1. Changes of estimates of the median, the 5th and 95th percentiles by Models 1 and 2. Blue (smooth) dashed curves denote the 5th and 95th percentiles and the blue (smooth) curves the median by Model 1; the red (saw tooth) dashed curves denote the 5th and 95th percentiles and the red (smooth) curves denote the median by Model 2.

3.4. Comparison between stationary and no stationary models

Table 5 shows the results of the comparison of the AIC between the three models for the five stations. According to the AIC criteria, the improvement after incorporating external covariates in the description of the changes in the annual discharge series is clear. It can be seen from Table 5 that the no stationary models (Model 1 and Model 2) have lower AIC values than the stationary model (Model 0), which shows the necessity of no stationary modeling. It can be seen from Table 5 that for the Kaqung and Alar stations, the AIC value is minimum for Model 2 which considers the influences of meteorological variables, while for the Tongguzlok, Yuzmenlek and Sharikilank stations, the AIC values are smaller.
for Model 1. At Kaqung and Alar stations, Model 2 captures more adequately the dispersion of annual discharge values than Model 1 and shows the effect of meteorological variables modulating of annual discharge, but at Tongguzlok, Yuzmenlek and Sharikilank stations, Model 1 seems work better than Model2.

**Table 5.** Comparison of the AIC between the three models for the 5 stations where the bold numbers denote the minimum AIC value for the three kinds of models.

| Station     | Model 0    | Model 1    | Model 2    |
|-------------|------------|------------|------------|
| Tongguzlok  | 1098.06    | 1096.697   | 1097.096   |
| Yuzmenlek   | 938.345    | 922.348    | 927.064    |
| Kaqung      | 1188.646   | 1186.003   | 1184.23    |
| Sharikilank | 1056.74    | 1031.192   | 1043.74    |
| Alar        | 1210.068   | 1209.416   | 1205.573   |

**Figure 2.** Changes of estimates of the median, the 5th and 95th percentiles by Models 0 and selected no stationary models. Red dashed lines denote the 5th and 95th.

Figure 2 shows the results of quantile estimates of the annual discharge under stationary and no stationary conditions. Model 0 based quantile estimates of the annual discharge are constant throughout the entire study time intervals and cannot reflect temporal variations under no stationary conditions, but the selected nonstationarity models indicate the annual discharge experiences significant variability. Figure 2 also shows that annual discharge is underestimated by Model 0 when compared to that by the selected nonstationarity models after 1990s, and the adverse is observed before 1990s, especially at Tongguzlok, Yuzmenlek and Sharikilank stations. However, different changing characteristics of
estimated annual discharge are observed at the Alar station when compared to that at the other stations, where it shows a gradual decreasing trend. These results confirm the suspicion of the stationary hypothesis, thus leading to our urgent request for no stationary modeling of annual discharge which can take this dynamic behavior into account.

4. Conclusion
In this study we have analyzed the annual discharge in the TR headstream catchment under stationary and no stationary conditions. Both stationary and no stationary models were conducted using GAMLSS to fit annual discharge series and model the dependence of distribution parameters with respect to explanatory variables (time and meteorological variables). The mainly results can be summarized as follows:

(1) There is an evident influence of meteorological variables on annual discharge in the TR headstream catchment. The temperature has a linear influence on the mean of annual discharge changes at the Tongguzlok, Yuzmenlek and Kaqung stations and nonlinear influence on the mean of annual discharge at the Alar station. Moreover, the temperature also has a linear influence on the variance of annual discharge at the Yuzmenlek station. However, the precipitation has a linear influence at the Yuzmenlek and nonlinear influence at the Sharikilank station on the mean of annual discharge, and has a linear influence on the variance value of annual discharge at the Alar station. In addition, the impact of precipitation on stream flow is much greater than that of temperature in the Aksu River, but the impact of temperature on stream flow is much greater than that of precipitation in the Hotan River.

(2) Compare to the stationary model (Model 0), the no stationary models (Model 1 and Model 2) that incorporate external covariates have a significant improvement in description of the changes in the annual discharge series. At Kaqung and Alar stations, Model 2 captures more adequately the dispersion of discharge values than Model 1, but at Tongguzlok, Yuzmenlek and Sharikilank stations, Model 1 seems work better than Model 2.

(3) Model 0 cannot reflect temporal variations under changing environment, and can underestimate or overestimate annual discharge, it is wrong to assume stationarity for statistical modeling of estimating water resource. The estimating water resource should be described by the no stationary models (Model 1 and Model 2) with more explanatory variables, such as time and meteorological variables.

Acknowledgments
This study is supported by Program for Outstanding Young Talents in Colleges and Universities of Anhui Province (No. gxyq2018143).

References
[1] M. N. Khaliq, T.B.M.J. Ouarda, J.-C. Ondo, P. Gachon and B. Bobée, Frequency analysis of a sequence of dependent and/or nonstationary hydro-meteorological observations: A review, Journal of Hydrology. 329(2006) 534-52.
[2] P. C. D. Milly, J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.K. Kundzewicz, D.P. Lettenmaier and R. J. Stouffer, Stationarity Is dead: whither water management, Science. 319 (2008) 573-74.
[3] Y.T. Fan, Y.N. Chen, W.H. Li, H.J. Wang and X.G. Li, Impacts of temperature and precipitation on runoff in the Tarim River during the past 50 years, Journal of Arid Land. 3 (2011) 220 - 30.
[4] Y. Jiang, C.H. Zhou and W.M. Cheng, Streamflow trends and hydrological response to climatic change in Tarim headwater basin, Journal of Geographical Sciences. 17 (2007) 51 - 61.
[5] X.M. Hao, Y.N. Chen, C.C. Xu and W.H. Li, Impacts of Climate Change and Human Activities on the Surface Runoff in the Tarim River Basin over the Last Fifty Years, Water Resour Manage. 22 (2008) 1159 - 71.
[6] H. Y. Zhou, X.L. Zhang, H.L. Xu, H.B. Ling and P. J. Yu, Influences of climate change and human activities on Tarim River runoffs in China over the past half century, Environ Earth Sci. 67 (2012) 231 - 241.
[7] X. H. Gu, Q. Zhang, V.P. Singh, X. Chen, L. Liu, Nonstationarity in the occurrence rate of floods in the Tarim River basin, China and related impacts of climate indices, Global and Planetary Change. 142 (2016) 1 - 13.

[8] H. B. Mann. Non-parametric tests against trend, Econometrica. 13 (1945) 245 - 59.

[9] M. G. Kendall, Rank Correlation Methods, Charles Griffin, London, 1975.

[10] R. A. Rigby and D.M. Stasinopoulos, Generalized additive models for location, scale and shape, Appl. Statist. 54 (2005) 507 - 54.