A coupled Bayesian and fault tree methodology to assess future groundwater conditions in light of climate change

J. J. Huang[1], M. Du[2], E. A. McBean[3], H. Wang[1] and J. Wang[1]

[1]{State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing,100038,China}

[2]{Polytech Nice-Sophia, 930 Route des Colles, 06903, Sophia Antipolis, France}

[3]{School of Engineering, University of Guelph, Guelph, ON, N1G 2W1, Canada}

Correspondence to: J. J. Huang (jeannehuang@gmail.com)

Abstract

Maintaining acceptable groundwater levels, particularly in arid areas, while protecting ecosystems, are key measures against desertification. Due to complicated hydrological processes and the uncertainty of the effects of anthropogenic activities, investigations of groundwater recharge conditions are challenging, particularly in arid areas under climate changing conditions. To assist the planning to protect against desertification, a fault tree methodology, in conjunction with fuzzy logic and Bayesian data mining, is applied to Minqin Oasis, a highly vulnerable region in northern China. A set of risk factors is employed within the fault tree framework, with fuzzy logic translating qualitative risk data into probabilities. This study provides a novel approach to utilize Bayesian data mining in conjunction with data transfer technology in validating the model with observation data. The framework of fault tree and Bayesian data mining approach have effectively quantify the causative effects between climate factors and human activities. A long series observation data ensures the model validated for a broad spectrum of changing conditions and therefore, it is satisfactory to be used in assessing the future projection of climate change on groundwater recharge conditions. Both the historical and future climate trends are employed for temperature, precipitation and potential evapotranspiration (PET) to assess water table changes under various future scenarios. The findings indicate that water table levels will continue to drop at the rate of 0.6m/yr in the future from climatic effects alone, if agricultural and industrial production

capacity remain at 2004 levels. It is necessary to plan water consumption of Minqin scientifically and effectively through management measures.

Keywords: Climate Change, Fault Tree, Fuzzy Logic, Bayesian Analysis, Groundwater, Data Mining

1 Introduction

The Shiyang River Basin is located in the eastern portion of the Hexi corridor in Gansu Province of northwest China, as shown in Fig. 1. Minqin Oasis lies at the lower reach of Shiyang River, which is the only oasis within both the Tenggeli and Badanjilin Deserts. This area is poised to become the largest desert in the world if the two deserts join together (Sun et al., 2006). Annual precipitation for Minqin Oasis has varied between 34~202 mm, with potential evaporation rates around 2600 mm, Runoff from upstream is the only source of renewable water supply for the Oasis to meet the increasing water needs due to population growth. Economic growth has led to the drilling of thousands of irrigation wells, resulting in withdrawals beyond recharge rates, causing groundwater mining (Tong et al., 2006). Falling water levels are adversely affecting land productivity. Further, ecological problems have intensified over recent decades, including soil desertification, due to the destruction of forestland and the groundwater mining (Zhang 2004). The water table at the edge of Minqin Oasis has dropped from 1~2m below ground surface (bgs) in 1950 to 16~24m bgs in 2008. Between 1998 and 2008, the groundwater table dropped 6.25 m on average, and even – dropped 3.25m/year in some locations (Hu et al., 2009). In addition to human-related water needs, the potential for climate change may worsen water security. For example, rising temperatures may boost evaporation rates and alter rainfall patterns, intensifying issues of water availability; additionally, rising temperatures may accelerate the glacier melting, which has been observed in nearby Qiling Mountain (Shi, et al., 2002 and 2007).

Investigation is needed to assess climate change impacts on water tables to provide a clearer picture of the potential for ecosystem failure and to provide insights into selection of optimal measures for improving the environment. This research estimates the likelihood of future climate change impacts on groundwater levels based on a lengthy record of historical data as well as using a general circulation model (GCMs), CGCM 3.1, for the projections of global climate change. The projection of future climate change is also used below in the risk assessment arising from water table changes in Minqin Oasis.

Since a declining groundwater level is an indication of potential ecosystem failure, a fault tree methodology for risk assessment of this failure is employed, to identify the contributing factors that impact water level changes. The risks of falling water table levels under changing climate conditions are assessed by applying the projected climate data along with the fault tree model.

Several studies have focused on the effects of climate change and the impacts of human activities on water resources in the Minqin Basin. For example, Wang et al.(2002) investigated environmental effects such as water quality deterioration, vegetation degradation, soil salinization and land desertification caused by human activities in the Basin. Kang et al. (2004) investigated the influence of human activities and global climate change on precipitation and runoff, together with a trend analysis of runoff in each tributary. Ma et al. (2008) showed that large-scale water resource development associated with dramatic population growth, led to large changes in groundwater regimes over the last 50 years. Li et al. (2007 and 2008) and Wang et al. (2009) explored the impact of land-use change on water resources. Huo et al. (2008) used a double cumulative curve and a multi-regression method to investigate the changes in streamflows under changing climates and human activities. These research findings indicate the water resources (surface and groundwater) in Minqin Basin are highly vulnerable to climate change. However, the quantification of the ineraction between changing climates and human activities have not been fully understood.

Due to complicated hydrologic processes and the uncertainty of the effects of human activities, investigations of groundwater recharge conditions are challenging, particularly in arid areas because of the interaction between climate factors and human activities. In order to identify and quantify the causes of the water table drop in Minqin, this study utilizes a modelling framework involving use of a coupled approaches of Fault tree and Bayesian data mining to quantify the effects of human activities and the influence of climate factors; the coupled model is validated by historical data and further was used in assessing the impact of climate change. A fuzzy number approach is very common in association with the use of a Fault tree in quantifying the risk associated with each risk factor as well as quantifying the consequence factor for each parent node. The fuzzy number in association to the fault tree provides an effective measure to quantify risk. However, it also raises the challenges in validating the model due to the lack of a mechanism for validating against historical data. This paper provides a methodology to use a data mining approach (WinBUGS) in conjunction with a data transfer technology to solve this validation problem. Section 3 explains the data transfer technology in details. This methodology allows the use of cumulative frequency to validate the model. Additionally, all the fuzzy numbers were calibrated by the cumulative frequency obtained from a long series of historical data, which makes the model valid for a wide spectrum of changing conditions.

2 Materials and methods

2.1 Fault tree methodology and fault tree model construction

The Fault Tree methodology is a graphical method for system reliability analysis. Among a variety of methodologies in the realm of probabilistic risk assessments, Fault Tree methodologies have been widely used and proven to be a versatile tool for modeling complex component behaviour (Lee et al., 2009; Singh et al., 2014;). For instance, Lee et al. (2009) assessed the risk of drinking water supply failure for small system by fault tree and fuzzy algorithm; Jurado et al.(2012) utilized fault tree methodology to assess the probability of groundwater related risk for the construction of underground structures; Singh et al.(2014) evaluated the seasonal effects and human activities on groundwater contamination using a modified fault tree methodology with ensemble machine learning approaches. The Fault Tree framework is a relatively straight-forward task; each risk item consists of a risk factor R and its “offspring” (contributing factors) Rc. Risk items with no further offspring are termed “basic risk items”, while the risk item without a ‘parent’ is the ‘top risk item’. Those between the top and bottom events are termed middle risk events. As a result, all critical paths for occurrence of an undesired state can be identified through the analysis and construction of the Fault Tree model.

There are two types of evaluations in Fault Tree Analyses, namely (i) qualitative evaluation and (ii) quantitative evaluation. The task of qualitative evaluation is to derive a logical structure amongst different events, while the objective of the quantitative evaluation is to assess the risk factor R(l, c) with the likelihood of a failure l and the consequence of failure c. In this paper, a quantitative evaluation is utilized for the risk assessment.

2.2 Bayesian data mining methodology

Bayesian Data Mining Methodology (BDMM) is a method of reasoning, based on a well-defined probabilistic theory, namely Bayes’ theorem. BDMM searches the results that best reflect the dependent relationships in a database of cases and provides a framework for handling probabilistic events. It has been proven to be a powerful formalism for expressing complex dependencies between random variables. In the presence of evidence, Bayes’ Theorem is used to compute the posterior probability distribution of a random variable and is here used to estimate the consequences of failure c for each risk factor R. Under the context of the Fault Tree Model, Bayes’ Theorem can be written in the following form:

$$P(c \mid R, R_c) = \frac{P(R_c, R \mid c) \times P(c)}{P(R_c, R)} \quad (1)$$

where,

R, R_c = observed risk factor data, which is the occurrence probability for each risk item;

c = a vector of parameters (c_1, c_2, c_3, … c_n) or consequence of a failure or a risk factor in this study under the fault tree model framework;

$P(c \mid R_c, R)$ = the conditional probability or so-called marginal probability, is in effect the consequence of a risk factor R_c in the matter of R, estimated based on observed R_c and R;

$P(c)$ = the prior probability, where the domain knowledge is applied;

$P(R_c, R \mid c) \times P(c)$ = the joint probability or so-called likelihood;

$P(R_c, R)$ = the density function of R_c and R.

where $P(c)$ is the prior distribution of the possible c values, $P(R_c)$ is the prior distribution of the observed data R_c, $P(R_c|c)$ is the probability of R_c, and $P(c|R_c)$ is the posterior distribution of c given the observed data R_c.

Due to improvements in sample-based Markov Chain Monte Carlo (MCMC) methods, recent work has led to sophisticated and efficient algorithms for computing and inferring probabilities in Bayesian analysis (Huang and McBean, 2007). The primary objective of the Bayesian parameter estimation method is to estimate the marginal probability $P(c \mid R_c, R)$ based on two components, the observed data R_c, R and the prior distribution gained from domain knowledge, $P(c)$. The results are given as the confidence intervals for model parameters based on predefined confidence levels. Sampling methods based on the Markov Chain principle incorporate the required search aspect in a framework where it can be proven that the correct distribution is generated at least in the limit, as the length of the chain grows. MCMC requires a large number of iterations (normally>10,000) to get convergence (Huang and McBean, 2008). This heavy computation requires a highly efficient program to undertake the computations. WinBUGS v1.4.3, open source software developed by MRC Biostatistics Unit, Cambridge, UK, was employed for this research to conduct the Bayesian data mining analysis utilizing the Markov Chain methodology.

2.3 Fuzzy logic methodology

In the traditional scientific view, uncertainty is regarded as undesirable and should be avoided. However, in many engineering systems, uncertainty is unavoidable and even essential (Lee et al., 2009 and Sadiq et al., 2004). Fuzzy set theory (Zadeh, 1965) was introduced to analyze objects that are not distinct and computed with words (Zadeh, 1996) as Fuzzy Logic is applied to explain reasoning linguistically rather than numerically (Sadiq 2004a). Fuzzy logic was initially implemented in control systems and programming by Bellman (1970) and is now extensively employed in engineering evaluations (e.g. Lee et al., 2009, Sadiq, 2004 b).

When risk items are evaluated, linguistic variables contain descriptive fuzzy terms such as high, medium, low, and so on. One linguistic variable can be defined by a term set, e.g., if there is a linguistic variable, ‘High’, can be defined more specifically by the term set ‘Extremely high’, ‘Very high’, ‘Moderately high’: each of which is a fuzzy set. The number of fuzzy sets is usually between three and seven (Lee et al., 2009).

Fuzzy sets are distributed as ‘*R*’, and each fuzzy set is mapped into a membership function, ‘*m*’, which ranges from 0 to 1. ‘*m*’ shows how strongly the ‘*R*’ is associated with the linguistic term, for example, a percentage of ‘*R*’ belongs to high. In the relationship between ‘*R*’ and ‘*m*’, a grade of membership is calculated using a membership function shape. To represent linguistic variables, Triangular Fuzzy Numbers (TFN) are used herein. There are three parametric variables for TFN, which are recorded as TFN= (*ma*， *mb*， *mc*). ‘*ma*’ represents the min. possible value of the fuzzy event, while ‘*mb*’ is the most likely value and ‘*mc*’ is the max possible values. The membership function of *l* and *c* to its respective granular are defined as:

$$m(r) = \begin{cases} 0, & r < ma \\ \frac{r-ma}{mb-ma}, & ma \leq r \leq mb \\ \frac{mc-r}{mc-mb}, & mb \leq r \leq mc \\ 0, & r > mc \end{cases} \tag{2}$$

Risk is assessed in seven grades here, namely, *extremely low, low , moderately low, medium, moderately high, high and extremely high* (Table 2). The membership function represents a means to convert a fuzzy number into a number, or vice versa. A crisp number differs from a fuzzy number such that the crisp number represents a real value and a fuzzy number represents only the relationship of the membership grades. For example, assuming the likelihood of a risk factor is between low (2), and moderately low (3) but much closer to 3 than 2, a crisp number of 0.31 may be chosen to represent the scenario.

Failure risk R is quantified as the product of risk likelihood and its consequence, which can be expressed as TFN_{lc}=TFN_l×TFN_c = ($ma_l \cdot ma_c$, $mb_l \cdot mb_c$, $mc_l \cdot mc_c$). When combining fuzzy sets, different fuzzy arithmetic mechanisms may be used within a rule-based system. Following this procedure is “defuzzification” which is a process to define a fuzzy set as a crisp number. The most common method is the Centre of Gravity approach, based on Bayesian probability.

Aggregation of fuzzy sets is performed by combining several fuzzy numbers to produce a single fuzzy number. Many different methods and operators may be used as an aggregation process, including fault tree analysis, means (e.g., arithmetic, geometric, or generalized), ordered-weighted averaging operators (OWA), and so on. The risk assessment starts from the ‘leaves’ of the tree (the ‘children’ risk items) and aggregates toward the ‘root’ (the top ‘parent’ risk item) through the “defuzzification” process.

2.4 Risk assessment for the water table decline for the Minqin Oasis construction of the fault tree model

During the 1950’s and 1960’s, prior to the introduction of large-scale groundwater withdrawal in Minqin Oasis, groundwater levels were near the surface. This allowed growth of significant natural vegetation. The annual water demand of natural vegetation is reported to be approximately 500 mm for healthy growth, while if constrained to 350 mm, growth would be inhibited. If decreasing groundwater levels occur, there will be vegetation die-off and reduced vegetation coverage, which will cause further soil erosion and intensified desertification. Therefore, the decline of groundwater levels provides an important indicator of ecosystem failure. The hierarchical fault tree shown in Fig. 2 indicates the top risk item of the Fault Tree model is identified as “decline of the water table”.

Since water consumption in the middle and upper portions of Minqin Basin has increased dramatically, the inflow to the lower basin (Minqin Oasis) has reduced substantially. The flow into the lower reaches of Hongyashang desert has decreased by 74%, although the discharge of the Shiyang River at the mouth of the mountain valley has remained at a level of 1.58 km^3/yr since the 1950s (Ma et al., 2005). In response, in order to maintain farmland production, groundwater has been exploited extensively since the 1970s. The annual groundwater exploitation has grown from 1.5×10^8 m^3 in the 1950s to 9.8×10^8 m^3 in the 1980s and to 11.16×10^8 m^3 in 1995 (Kang et al., 2004). Therefore, the fault tree components identified in the second layer of Fig. 2 are “total water consumption” and “supply from upper reaches”, the results of which drive the groundwater mining.

Water consumption is divided into three water use sectors, namely: “Agricultural Consumption”, “Domestic and Industrial Consumption”, and “Ecosystem Water

Consumption". Therefore, in the fourth layer of Fig. 2, there are three contributing risk factors identified as "Agricultural Consumption" including "Agriculture GDP", "PET" and "Precipitation". Similar to "Agricultural Consumption", the "Ecosystem Water Consumption" refers to water demand for the healthy growth of the natural vegetation, and is impacted by PET and precipitation. Since the primary goal herein is to investigate climate change impacts on the water balance, impacts of climate change on domestic and industrial water consumption are considered negligible and hence not further investigated.

3 Data analysis

Fault Tree Analysis was conducted based on fuzzy logic, which could provide an efficient way to quantify the risk likelihood and its consequence. The risk likelihood and its consequence were further validated against the data obtained from observation data by utilizing Bayesian Analysis and MCMC Analysis by WinBUGS. To assess the risk of the water table decline, the first step is to identify quantitative relationships between the variables. Although observation data of each variable are available, it is difficult to carry out the calculation due to different dimensions and magnitudes of the different data types. By applying the cumulative frequency of each variable as the quantitative criteria, each variable can be represented by normalized values between 0 and 1.

As an example of how this is accomplished, the mean value $\bar{x}$ is calculated for the observation of each variable, as the parameter estimate μ and standard deviation s as the parameter estimate of σ in the normal distribution. The cumulative frequency $F(x_i)$ for each x_i is the likelihood in Fault Tree. Precipitation influences on the decline of water table in a different manner than from that of other variables; precipitation is a negative relevant relation, as the greater precipitation, the less the water table will decline. Thus $1\text{-}F(x_i)$ is taken as the likelihood for precipitation.

Water-saving measures involving reducing cultivated land and abandoning groundwater wells have been taken place in Minqin since 2005. Thus, the year 2004 has been taken as the base year. Substituting precipitation data and PET between 2001 and 2100 into the Fault Tree model for the variance of water table, the risk parameter for the water table decline is calculated and compared with that of 2004. Data for agriculture GDP, water supply from upper reaches, and domestic & industrial water consumption of 2004 have been applied in the calculation.

The CGCM3.1 (Coupled Global Climate Model 3.1) has been applied to forecast the projected climate. In the model, Scenarios A2, A1B and B1 are chosen as three climate scenarios. Hargreaves and Samani (1982) proposed a methodology to estimate the potential evapotransporation (PET). The method was applied in this study to evaluate the PET based on the metrological data for Wushaoling and Minqin. The Gumbel Distribution has been demonstrated as the most effective distribution in the evaluation of extreme events (Wang and McBean, 2014). Hence, the return period is calculated assuming the Gumbel Distribution in this study. For the risk assessment for water table variance, Fault Tree Analysis (FTA) is introduced.

In the Fault Tree, data for agricultural GDP are from the *Almanac of Gansu*. Data of water supply from upper reaches, water table and consumption are from the Gansu Research Institute for Water Conservancy. The agricultural water consumption, domestic & industrial water consumption and the water supply from upper reaches between 2001 and 2008 were obtained from Gansu Research Institute for Water Conservancy together with the data of ecological water consumption as well as the data characterizing water table levels between 1998 and 2008.

Agricultural water consumption is the predominant factor of Minqin's total water consumption in 2000. Taking 2000 as the base year, the total water consumption before 2000 is calculated according to the annual gradient of agricultural water consumption. Thus, a complete series of consumption data between 1951 and 2008 is available. Table 1 lists the observed data used in the Fault Tree Analysis.

3.1 The model for climate change

The model utilized for climate projections is CGCM3.1 (Coupled Global Climate Model 3.1), developed by the Canadian Centre for Climate Modeling and Analysis (CCCMA). The climate scenarios that are most widely used are A2 (high degree of greenhouse-gas emission), middle degree A1B, and low degree B1, Committed degree (equal to that of the year 2000) and principal degree (on the premise of de-industrialization).

3.2 Fault tree analysis

FTA is a graphical analytical method to evaluate system reliability. When there is an undesired state or a failure, all factors that directly lead to the undesired state or failure will be identified. Then, the identifying causes for the lower event are investigated until a root or controllable cause is obtained. At the end, all critical paths for the occurrence of the undesired state will have been identified. In FTA, the undesired event is called the “top event”, while the event where the failure rates and probabilities enter the Fault Tree is termed the “bottom event” or “base event”; those between the top and base events are the middle events. The relationship or logic of the Cause-Effect events is identified. By applying logic gates (AND and OR gates), a tree derivation is structured with these events, which are represented by standard logic symbols.

3.2.1 Fault Tree construction of water table level in Minqin

In the analysis, the decline of the water table in Minqin could be inferred to be an issue caused by the combination of human and natural factors. In the system, “decline of water table in Minqin” is classified as the top event. The causes for the decline are the increase in local water consumption and the decrease of water supply from the Hongyangshan Reservoir. Local water consumption consists of four parts, including agricultural water, domestic and industrial water, and consumption by the ecological system components.

Agricultural water consumption is influenced by the variability of the cultivated area, as well as local PET and the precipitation. Domestic and industrial water consumption are used as the “bottom” event because these are essentially immune from the natural factors indicated above. Ecological water is used for irrigation of local vegetation, in order to prevent further desertification. This water consumption is influenced by PET and precipitation. FT for groundwater in Minqin is shown in Fig. 2.

3.2.2 Transformation of observation data for variables in Fault Tree

In the risk assessment for decline of the water table, the first step is to establish the quantitative relationships of individual variables. Although observed data for each variable are available, it is difficult to carry out the calculation due to different dimensions and magnitudes. A new flexible approach is to rely upon the cumulative frequency of each variable as the quantitative criteria. In this way each variable can be represented by values between 0 and 1.

The experimental results follow the normal distribution based on the Central Limit Theorem (McBean and Rovers, 1998) if the results do not influence each other, despite different distributions for alternative random variables. The cumulative frequency $F(x_i)$ for each x_i is determined as the likelihood in the Fault Tree.

3.2.3 Bayesian Data Mining Methodology for Fault Tree Model Validation

The system for characterizing the decline of the water table in Minqin can be divided into four subsystems, including the subsystem “water supply from upper reaches/total water consumption in Minqin → decline of water table”, “agricultural water consumption/domestic & industrial water consumption/ecological water consumption → total water consumption”, “agricultural GDP/PET/precipitation → agricultural water consumption” and “PET/precipitation → ecological water consumption”.

In the subsystem “agricultural water consumption/domestic and industrial water consumption/ecological water consumption→ total water consumption”, values of *c* equal the percentage of each type of water consumption. In the other three subsystems, values of *c* are functions of position and are calculated from available data by WinBUGS.

The Bayesian models to represent the three subsystems are constructed using WinBUGS, respectively. These models apply the likelihoods derived from the cumulative frequency above. In each subsystem model, the length of the input data equals that of the shortest series. Comparison of the observed and calculated data (for the cumulative frequency) was conducted to verify the three sub models in WinBUGS.

4 Results and discussion

4.1 Climate change trends for the past several decades

4.1.1 Average annual temperature

The average temperature in Wushaoling has shown an upward trend (Fig. 3a). Applying trend line fitting, we can see an increase of 0.021 ℃/yr for the average temperature is defined, while after the 1980s, the rate of increase reached 0.062 ℃/yr. The average annual temperature in Minqin is also showing an upward trend (Fig. 3b). The average annual temperature increase is 0.059 ℃/yr, which is about the two times the rate of increase observed for Wushaoling for the entire time period. Minqin and Wushaoling, show a similar increasing trend in temperature after the 1980s, at rates of 0.06 ℃/yr. Temperature increases are at a lesser rate in Wushaoling for the past 50 years. In other words, increases in temperature along the lower reach of Shiyang River are more apparent than for the middle and upper reaches.

4.1.2 PET

PET data in Wushaoling and Minqin are shown in Fig. 4. PET in Wushaoling has a slight upward trend. The mean value of PET is 1580.3mm, with a standard deviation of 73.0mm and rate of change of 1.15mm/yr. PET in Minqin also has a slight upward trend. The mean value of PET is 2644.0mm with a standard deviation of 66.0mm and rate of change of 1.36mm/yr.

4.1.3 Precipitation

Precipitation data in Wushaoling and Minqin are shown in Fig. 5. The annual average precipitation in Wushaoling is 374mm with a standard deviation of 78.9mm. The maximum precipitation is 543mm in 2003 and the minimum is 176mm in 1962. Precipitation has shown a downward trend and then upward, not simply one-directional. In the 1960s, there was a reduction of precipitation, which lasted into the 1980s and then recovered to the previous level. The annual average precipitation in Minqin is 111mm with a standard deviation of 33.4mm. The maximum precipitation is 202mm in 1994 and the minimum is 38mm in 1962. It has shown a mild upward trend， with fluctuations around the mean value.

4.2 Projected climate change trends given by CGCM

Fig. 7 and 8 show the annual temperature change, PET and precipitation given by CGCM3.1. Comparing all of these curves, the average temperature, PET and precipitation in the three different scenarios all have fluctuating upward trends. The climate elements have different rates of increase. Except for the PET of the A1B scenario, the other rates of increase have all shown a relationship of A2>A1B>B1.

4.3 Risk assessment for water table variance

4.3.1 Results of Bayesian data mining methodology

Results of observed data-based methodology for Fault Tree calculation are shown in Tables 4. In the tables, CL is the unknown parameter *c*, and τ which is used to evaluate the closeness of the fit (the larger τ, the better the fit). Sigma represents parameter σ, another parameter to evaluate the closeness of fit and $\sigma=1/\sqrt{\tau}$. It can be seen from Table 3 through 6, the mean values of CL range between 0 and 1, and values of τ are 73.51, 14.63 and 84.4. The value of τ in the subsystem for "PET/precipitation→ ecological water consumption" is a little smaller, perhaps due to the short data series (only seven years). Risk probability of each node is listed in Table 4.

The values of cumulative frequency for the decline of the water table as well as for agricultural water consumption and ecological water consumption were calculated by WinBUGS. In order to demonstrate the validity of the models, the results from WinBUGs were plotted with the observation data. As shown in Fig. 9, it was observed that there are significant linear correlations between computed values and observed data for agricultural water consumption as well as for the decline of the water table. The correlation coefficients (R^2) are 0.894 and 0.875 for these two nodes. There is no obvious correlation between computed values and observed data for the cumulative frequency of ecological water consumption as a result of the short data series, the significance analysis is carried out at the confidence level of 90% where α=0.1. Values of R^2 and F were obtained from regression analysis as shown in Fig. 9.

For the decline of the water table, there are 49 observed data points, *the critical value* of $F_{1,51}\cong 2.81$, while for agricultural water consumption and ecological water consumption, there are 20 and 8 data points, respectively, with the corresponding critical value of F are 3.01 and 3.78 respectively. The F values shown on Fig. 9 are all far in excess of the critical values. Therefore, the correlations between computed values and observed values are statistically significant.

Therefore, the model can be regarded as valid and can be used for predictions for the future due to climate change.

4.3.2 Results of Fuzzy-based methodology

Importance degrees calculated according to fuzzy numbers are shown in Table 5, comparing the risk factor $g(l,c)$ for the decline of water table, ecological water consumption and agricultural water consumption with the cumulative frequency calculated by WinBUGS. The comparison results are shown in Fig. 10. Values of R^2 reach as high as 0.965, 0.996 and 0.987 respectively, which indicate an obvious correlation between the risk factor $g(l,c)$ and the cumulative frequencies. As the validity of cumulative frequencies has been proven, the importance degrees derived from the fuzzy numbers are also valid and can be used for the risk assessment.

4.3.3 Risk Assessment for the water table fault tree in Minqin

Risk parameters for the variance of water table calculated in the three climate scenarios are shown in Fig. 11, where the risk parameter of 2004 is represented by the dashed line and equals 0.835.

Comparing risk factors of the three climate scenarios, a fluctuation between 0.78 and 0.83 is observed. In scenarios of A2 and A1B, the magnitude of the fluctuation gradually declines, which indicates that the effect imposed by climate factors on the water table is diminishing. The risk parameters for most years are lower compared with that of 2004. Only the risk parameter of the year 2023 in scenario B1 is similar to that of 2004, equal to 0.824.

According to the definitions by IPCC (2007), the greenhouse gas emissions for scenario A2 are high, while they are lower for scenario A1B and the lowest for scenario B1. For the three scenarios, the mean risk of the water table drop is 0.799, 0.798 0.798 for scenarios A2, A1B and B1, respectively. Thus, there is a slightly higher risk of inducing water table decline in scenario A2 compared with the risks in scenarios A1B and B1, but the differences among the three scenarios are fairly negligible.

There are two reasons for risk factors in the three scenarios being lower than the year 2004. First, the year 2004 suffered a serious drought with the observation data of precipitation 100.2mm and PET 2808.8mm. The mean decline of water table levels reached as high as 0.835m. Second, the consequences of precipitation are greater than that of PET for the event of the same level in the Fault Tree model, which indicates that the effect of precipitation imposed on water table change is greater than that of PET when considering climatic conditions only. According to the features of precipitation in the three scenarios, the mean value of precipitation between 2001 and 2100 is high and showed an upward trend. Thus the water table decline risk is low despite the increasing PET.

Risk factors for the decline of water table levels in the three scenarios could be transformed to the rate of descent in water table levels by looking up the value at the cumulative frequency curve for the decline of the water table. It can be seen from Table 7 that the decline of the water table levels in Minqin will reach as high as about 63m in 96 years (between 2004 and 2100) under the assumption of no water saving measures being taken and that the agricultural and industrial production capacities remain at the same level with that of 2004. The impact of climatic conditions is the only factor being investigated in this study. It reveals a decline rate of 0.6m per year for the groundwater, which is an important index when planning water resources and allocating the domestic and industrial water.

5 Conclusions

5.1 Climate change trends

Changing trends of average temperature, PET and precipitation on the upper-middle reaches and lower reaches of Shiyang River are evident. The climate change trends of Scenarios A2, A1B and B1 in Minqin given by CGCM show significant increasing trends in average temperature and PET in Wushaoling on the upper-middle reaches and Minqin on the lower reaches. Meanwhile, there are also increasing trends in precipitation in Wushaoling and Minqin. The magnitude of increases for average temperature and PET are larger, while that of precipitation is smaller. The average temperature in scenario A2 increases at a rate of 0.061 ℃/yr, similar to the observed data in Minqin. The ascending speeds of PET in the three scenarios have all exceeded the observed data with a rate of 1.15mm/yr. As for the precipitation, the ascending speed of observation data equals 0.517mm/yr, between those in scenarios A2 and A1B.

5.2 Risk assessment of water table levels

The Fault Tree model as applied to the water table fluctuation in Minqin, demonstrates the risk factors for the water table fluctuation in the three projected climate scenarios, and demonstrates that precipitation has a greater effect on the water table than that of PET. The consequences approximate those calculated by the fuzzy algorithm and Bayesian data mining approach, providing evidence of the validity of the model.

The risk assessment shows that the declining rate of ground water levels will reach 0.6m/yr to 2100 considering the climatic effects only and under the assumption that the agricultural and industrial production capacity are maintained at the levels of 2004.

With climate change alone, the water table in Minqin may continue to decline, resulting in increasing challenges in dealing with ecological problems in the Minqin Oasis. It is necessary to plan water consumption of Minqin scientifically and effectively through management measures.

5.3 The Bayesian and fault tree methodology

This study provides the evidence that the coupled model with data transfer technology can not only assess the risk likelihood but also assess the interactive impacts of climate factors and anthropogenic activities. Additionally, the new method on the validation of the model against the long-series historical observation data makes the model more valuable. It provides an efficient and effective way in assessing the climate change impacts on groundwater recharging condition.

Acknowledgements

This paper arose from a three-year project of the Chinese Hydraulic Engineering Society (CHES), the Gansu Research Institute for Water Conservancy (GRIWC) and the Canadian Society for Civil Engineering (CSCE). Funding was provided by National Natural Science Foundation of China (40830637), State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin，China Institute of Water Resources and Hydropower Research (2014ZY01), the Chinese Ministry of Science and Technology (2007DFA70860) and Ministry of Education (NCET-09-0586). Gratitude is expressed to Dr. Gao Bo of the MWR and to Dr. Zantang Li, Secretary General, and Ms. Jing (Jenny) Gao of CHES. Special thanks are given to Mr. Yuanhong Li, Director, and other staff of the GRIWC for their support, particularly the arrangement of site visits and the provision of information.

References

Hargreaves, G.H. and Samani, Z.A.: Estimating potential evapotranspiration. ASCE Journal of Irrigation and Drainage Engineering, 108(IR3), 223-230, 1982.

Hu, L. T, Wang, Z.J., Tian, W., and Zhao, J.S.: Coupled surface water-groundwater model and its application in the Arid Shiyang River basin, China, Hydrological Processes, 23(14), 2033-2044, 2009.

Huang, J., and McBean, E. A.: Using Bayesian statistics to estimate the coefficients of a two-component second-order chlorine decay model for a water distribution system, Water Research, 41(2), 287-294, 2007.

Huang, J., and McBean, E. A.: Using Bayesian statistics to estimate the chlorine wall decay coefficients for a water supply system, ASCE Journal of Water Resources Planning and Management, 134(2), 129-137, 2008.

Huo, Z., Feng, S., Kang, S., Li, W., and Chen, H.: Effect of climate changes and water-related human activities on annual stream flows of the Shiyang river basin in arid north-west China. Hydrological Processes, 22, 3155-3167, 2008.

IPCC: Fourth assessment report climate change 2007, Paris, http://www.ipcc.ch/, 2007.

Jurado, A.; De Gaspari, D; Vilarrasa; V. ; Bolster, D. ; Sánchez-Vila, X.; Fernàndez-Garcia, D. and Tartakovsky, D.M.; (2012). Probabilistic analysis of groundwater-related risks at subsurface excavation sites; Engineering Geology,V125, pp 35-44

Kang., S. Sub, X. ,Tong, L., Shi, P., Yang, X., Abe, Y., Du, T., Shen, Q., and Zhang, J.: The impacts of human activities on the water-land environment of the Shiyang River basin, an arid region in northwest China. Hydrological Sciences Journal, 49(3) , 413-427, 2004.

Lee, M., McBean, E., Schuster, C., and Huang, J.: A fault tree and fuzzy logic methodology for risk assessment of drinking water supply failures, ASCE Journal of Water Resources Planning and Management, 135(6), 547-552, 2009.

Li, X., Xiao, D., He, X. Chen, W., and Song, D.: Factors associated with farmland area changes in arid regions: a case study of the Shiyang River basin, northwestern China; Front Ecol Environ 2007; 5(3), 139-144, 2007.

Li, Z. L., Xu, Z. X., Li, J. Y., and Li, Z. J.: Shift trend and step changes for runoff time series in the Shiyang River basin, northwest China. Hydrological Presses. 22, 4639-4646, 2008.

Ma, J., Li, D., Xie, Y., and Chen, F.: Impact of water saving on the desertification and pauperization in the Minqin Basin, Gansu Province. Proceedings of the International Symposium on Sustainable Water Resources Management and Oasis-hydrosphere-desert Interaction in Arid Regions, April 20-23, 2005, Beijing China. Chinese Academy of Sciences, Beijing China. 349-515, 2005.

Ma, Z., Kang, S., Zhang, L., Tong, L., and Su, X.: Analysis of impacts of climate variability and human activity on streamflow for a river basin in arid region of northwest China. Journal of Hydrology. 352(3), 239-249, 2008.

McBean, E.A., and Motiee, H.: Assessment of impacts of climate change on the water resources - Great Lakes of North America, Hydrology and Earth System Sciences, 12(1), 239-255, 2006.

McBean E. A. and Rovers, F. A.: Chapter 11, Nonparametric procedure, statistical procedures for analysis of environmental monitoring data and risk assessment, Prentice Hall PTR, New Jersey, ISBN 0-13-675018-4, 1998.

Sadiq, R., Rajani, B., and Kleiner, Y.: Risk analysis for water quality deterioration in distribution networks. Evaluation and Control of Water Loss in Urban Water Networks, 1-26, 2004a.

Sadiq, R., Kleiner, Y., and Rajani, B.: Aggregative risk analysis for water quality failure in distribution networks. Journal of Water Supply: Research and Technology - AQUA 53(4), 241-261, 2004b.

Singh, K.; Gupta, S and Mohan, D. (2014). Evaluating influences of seasonal variations and anthropogenic activities on alluvial groundwater hydrochemistry using ensemble learning approaches, Journal of Hydrology, v511, pp 254-266,

Shi, Ya-feng, Shen, Yong-ping, Hu, Ru-ji (2002). Preliminary Study on Signal, Impact and Foreground of Climatic Shift from Warm-Dry to Warm-Humid in Northwest China. JOURNAL OF GLACIOLOGY AND GEOCRYOLOGY, 2002 Vol. 24 (3): 219-226

Shi, Yafeng; Shen, Yongping; Kang, Ersi; Li, Dongliang; Ding,Yongjian; Zhang, Guowei and Ruji Hu. "Recent and future climate change in northwest China." Climatic change 80, no. 3-4 (2007): 379-393.

Sun D., Dawson, R., and Li, B.: Agricultural causes of desertification risk in Minqin, China. Journal of Environmental Management. 79(4), 348-356, 2006.

Tong, L., Kang, S., and Zhang, L.: Temporal and spatial variations of evapotranspiration for spring wheat in the Shiyang river basin in northwest China. Agricultural Water Management. 87(3), 241-250, 2007.

Wang, Q., Shi, J.A., Chen, G.J., and Xue, L.H.: Environmental effects induced by human activities in arid Shiyang River basin, Gansu province, northwest China. Environmental Geology. 43, 219-227, 2002.

Wang, Z.J., Zheng, H., and Wang, X.F.: A harmonious water rights allocation model for Shiyang River Basin, Gansu Province, China. International Journal of Water Resources Development, 25(2), 355-371, 2009.

Wang, Y., and McBean, E., Performance comparisons of partial duration and annual maximum series models for rainfall frequency of selected rainfall gauge records, Journal of Hydrologic Science, under review.

Wu, Q., and Wang, M.: A framework for risk assessment on soil erosion by water using an integrated and systematic approach. Journal of Hydrology. 337(1), 11-21, 2007.

Zhang, T.: Will the desert claim Minqin? http://www.china.org.cn/english/2004/Sep/105977. htm, 2004.

Table 1. Available data for Fault Tree Model.

Variables/unit	Time series(duration)	Notes
Meteorological data (historical Records)	1950-2009	China Metrological Data Sharing Service System
Agricultural GDP/10^8 RMB	1985~2004	Yearbook of Gansu Province
Water Supply from upstream/10^8 m^3	1956~2008	Gansu Research Institute for Water Conservancy
Precipitation	1960~2008	Gansu Research Institute for Water Conservancy
Total water consumption /10^8 m^3	1951~2008	Gansu Research Institute for Water Conservancy
Agricultural water consumption/10^8 m^3	1951~2008	Gansu Research Institute for Water Conservancy
Domestic & industrial water consumption/10^8 m^3	1951~2008	Gansu Research Institute for Water Conservancy
Ecology water consumption/10^8 m^3	2001~2008	Gansu Research Institute for Water Conservancy
Water table/m	1951~2008	Gansu Research Institute for Water Conservancy

Table 2. Triangular fuzzy number for granular.

Granular (p)	Qualitative scale for likelihood of risk (l)	Qualitative scale for peril of risk (c)	TFN_l or TFN_c
1	Extremely low	Extremely unimportant	(0.0, 0.0, 0.17)
2	Quite low	Quite unimportant	(0.0, 0.17, 0.33)
3	Low	Unimportant	(0.17, 0.33, 0.50)
4	Medium	Neutral	(0.33, 0.50, 0.67)
5	Quite high	Quite important	(0.50, 0.67, 0.83)
6	High	Important	(0.67, 0.83, 1.0)
7	Extremely high	Extremely important	(0.83, 1.0, 1.0)

Source: (Lee et al., 2009)

Table 3. Results of consequences for each node.

Nodes	Mean values	Notes
CL_{aw1}	0.9562	agriculture GDP → agricultural water consumption
CL_{aw2}	0.007455	PET → agricultural water consumption
CL_{aw3}	0.0403	Precipitation → agricultural water consumption
Tau_{gw}	73.51	
$sigma_{gw}$	0.1249	
CL_{ew1}	0.2833	Node of “PET → agricultural water consumption”
CL_{ew2}	0.673	Node of “precipitation → agricultural water consumption”
Tau_{gw}	14.63	
$sigma_{gw}$	0.3176	
CL_{gw1}	0.3616	Nodes of “water supply from upper reaches → the decline of water table”
CL_{gw2}	0.5758	Nodes of “total water consumption in Minqin → the decline of water table”
Tau_{gw}	84.4	
$sigma_{gw}$	0.1115	

Table 4. Risk probability of each node.

Numbers	Nodes	Risk Probability
4.1	Agriculture GDP→agricultural water consumption	0.9562
4.2	PET→ agricultural water consumption	0.007455
4.3	Precipitation→agricultural water consumption	0.0403
4.4	PET→ecological water consumption	0.2833
4.5	Precipitation→ ecological water consumption	0.673
3.1	agricultural water consumption→total water consumption	0.7898
3.2	Domestic and industrial water consumption→ total water consumption	0.0414
3.3	Ecological water consumption→ total water consumption	0.1302
2.1	Total water consumption→ decline of water table	0.3616
2.2	Water supply from upper reaches→ decline of water table	0.5758

Table 5. Fuzzy numbers for consequences.

Items	Nodes	Consequences
4.1	Agricultural GDP → agricultural water consumption	7
4.2	PET → agricultural water consumption	1
4.3	Precipitation → agricultural water consumption	1
4.4	PET → ecological water consumption	3
4.5	Precipitation → ecological water consumption	5
3.1	Agricultural water consumption → total water consumption	6
3.2	Domestic and industrial water consumption → total water consumption	1
3.3	Ecological water consumption → total water consumption	2
2.1	Total water consumption → decline of water table	3
2.2	Water supply from upper reaches → decline of water table	4

Table 6. Forecast for variance of water table in Minqin considering climatic conditions only.

	A2	A1B	B1
The decline of water table between 2004 and 2100 (m)	63.632	63.499	63.469
Water table in 2100 (m)	-86.583	-85.451	-86.421

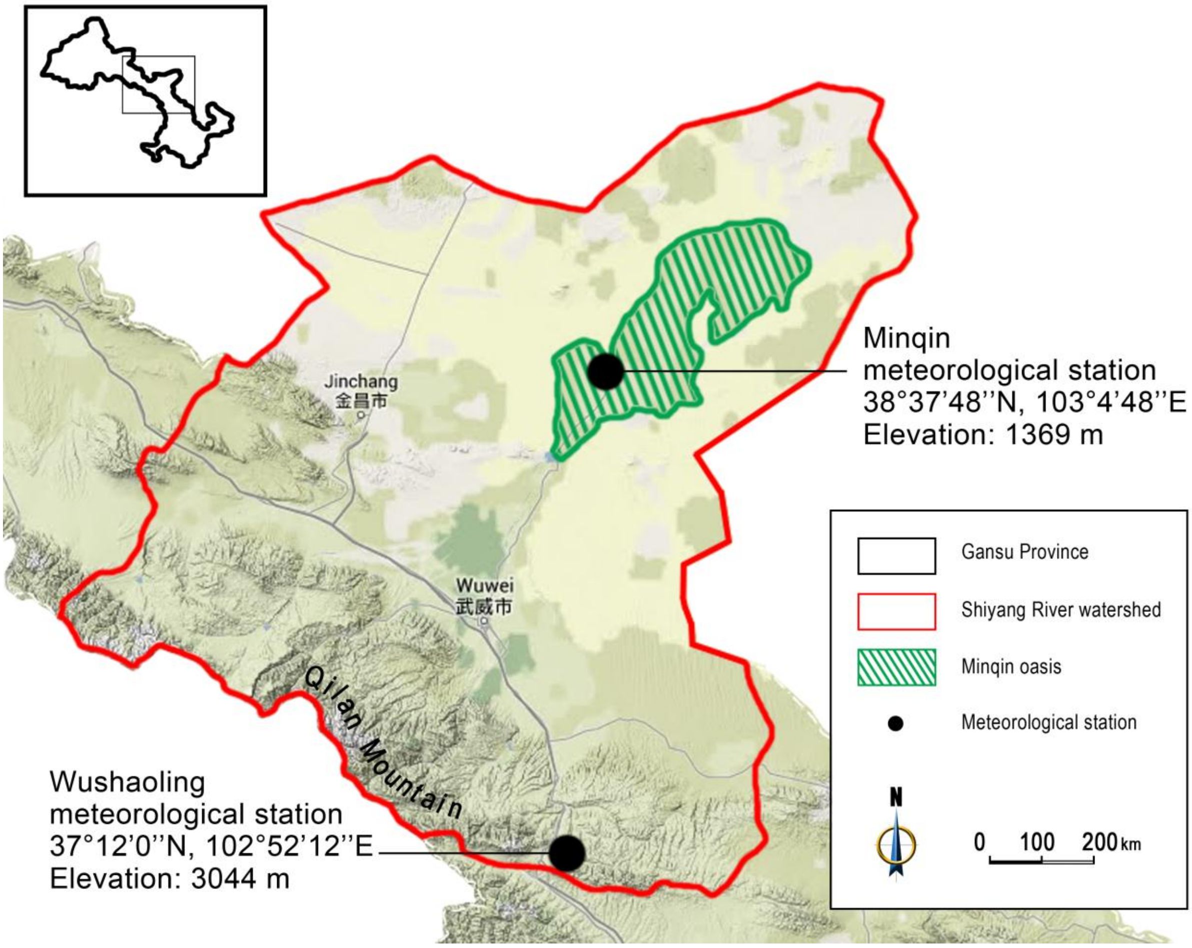


Figure 1. The Shiyang River Watershed and Minqin Oasis.

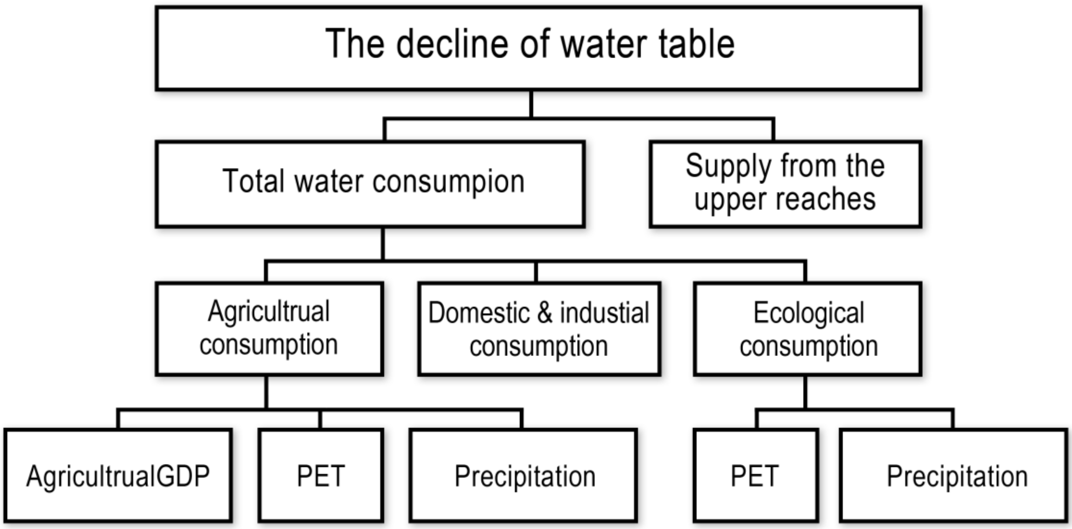
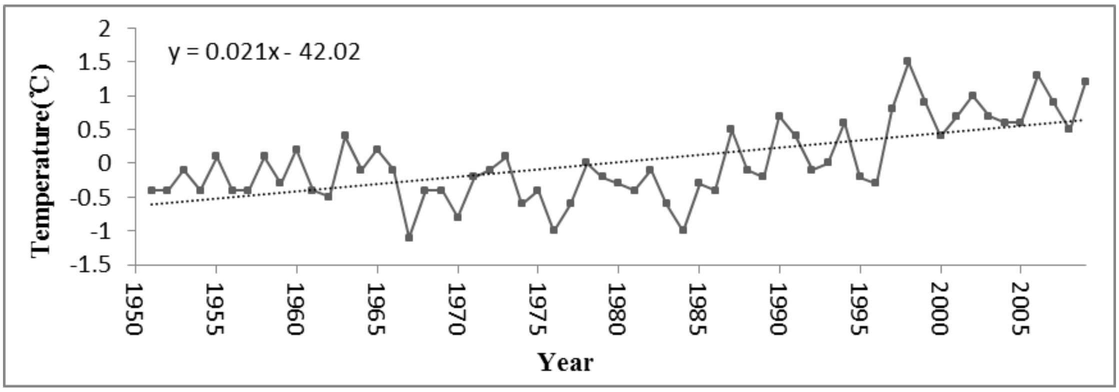
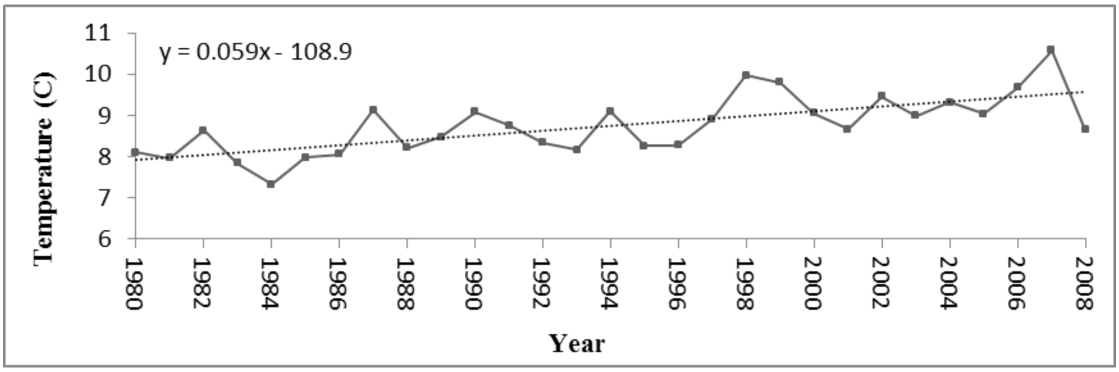


Figure 2. Fault tree of groundwater table drop for Minqin Oasis.

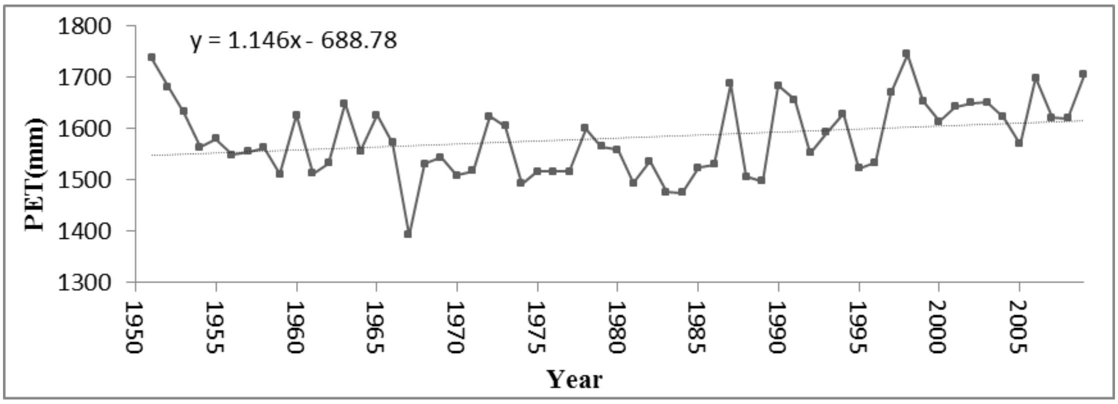


(a) Wushaoling

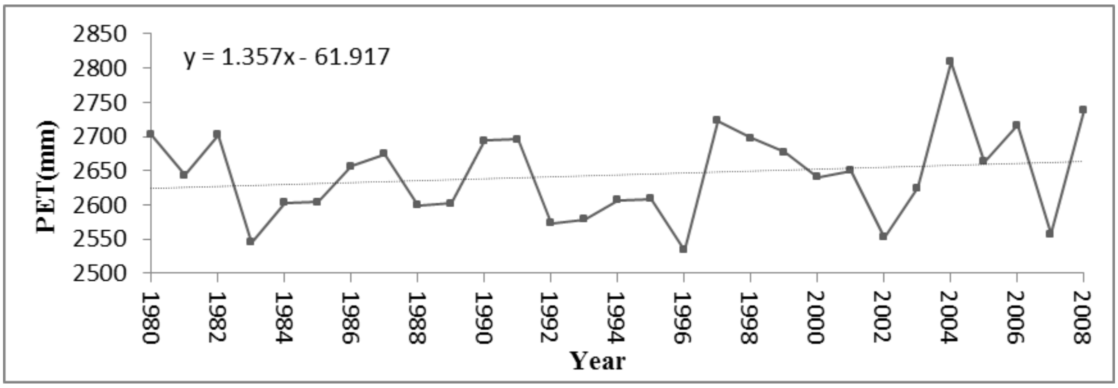


(b) Minqin

Figure 3. The trend of average annual temperature.

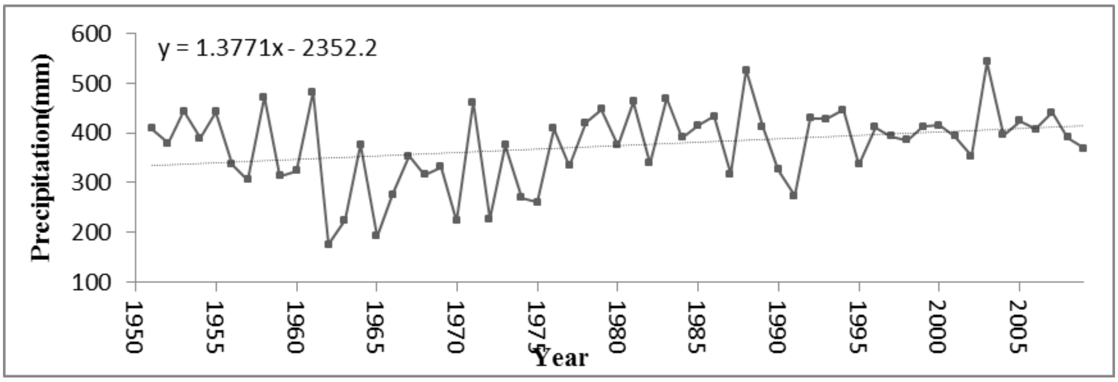


(a) Wushaoling

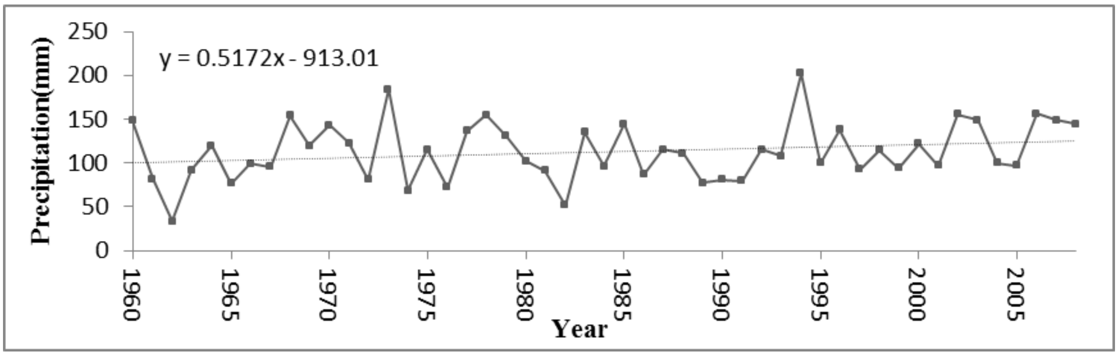


(b) Minqin

Figure 4. The trend of average annual PET.



(a) Wushaoling



(b) Minqin

Figure 5. The trend of annual precipitation.

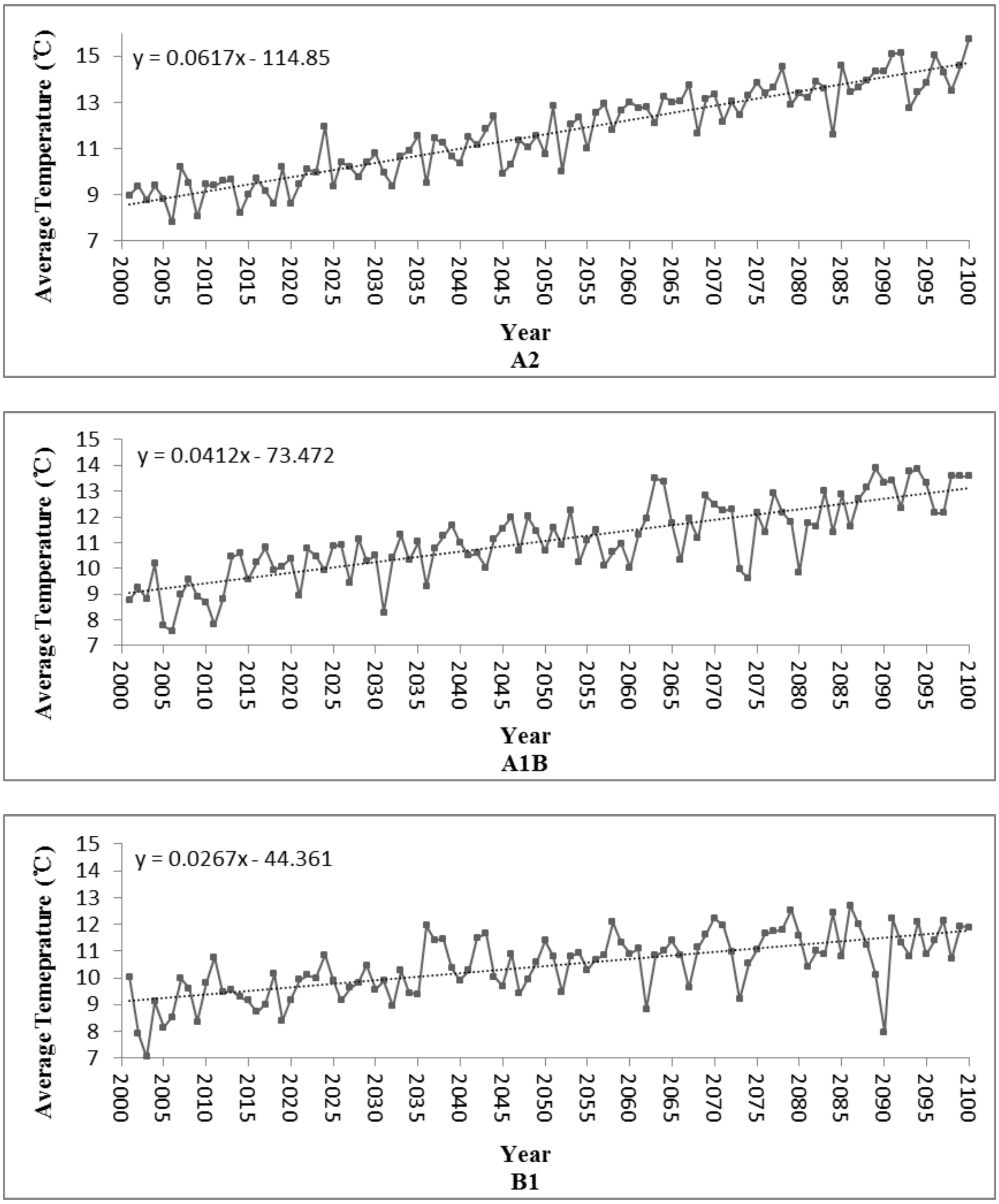


Figure 6. The trend of average temperature given by CGCM 3.1.

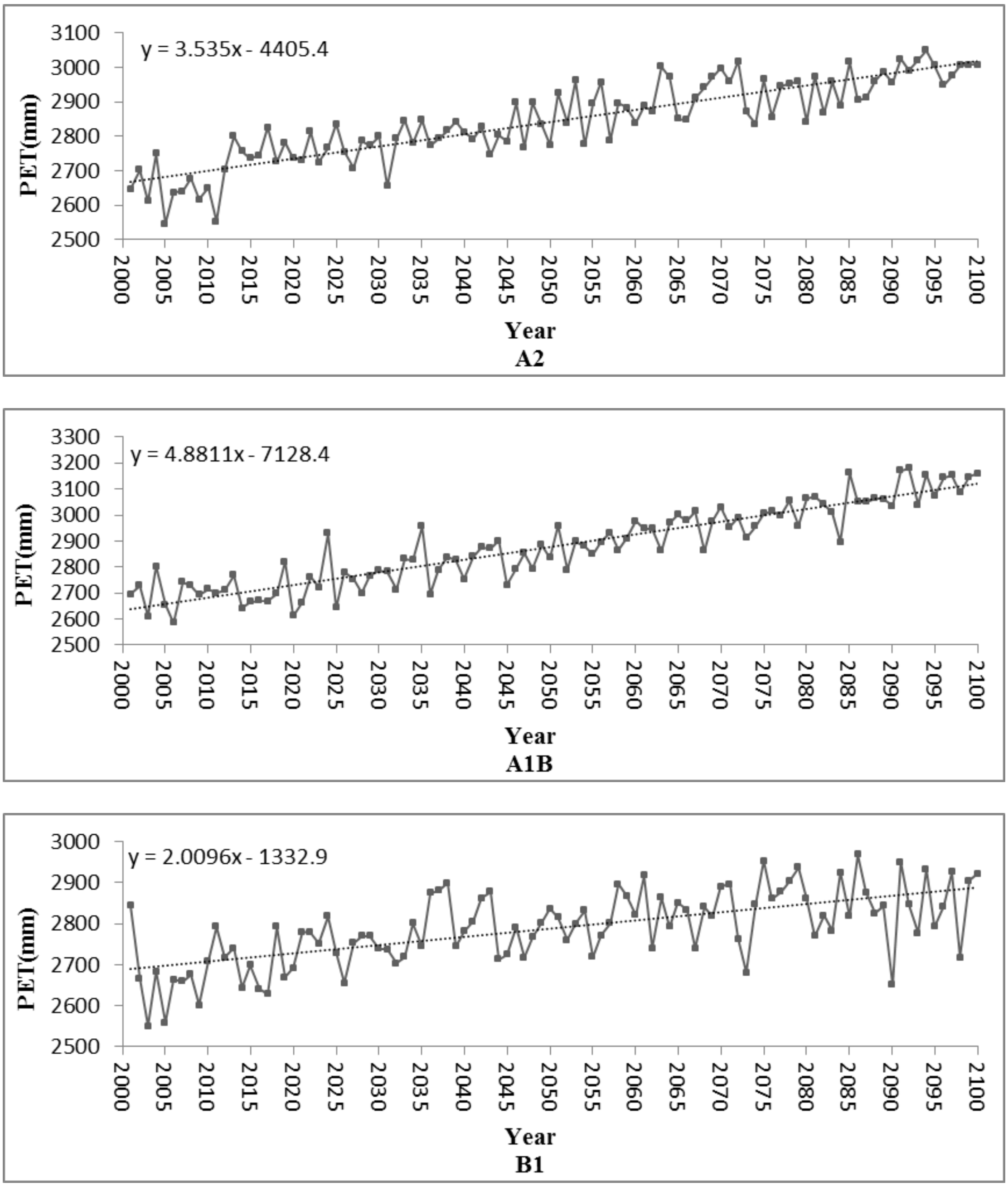
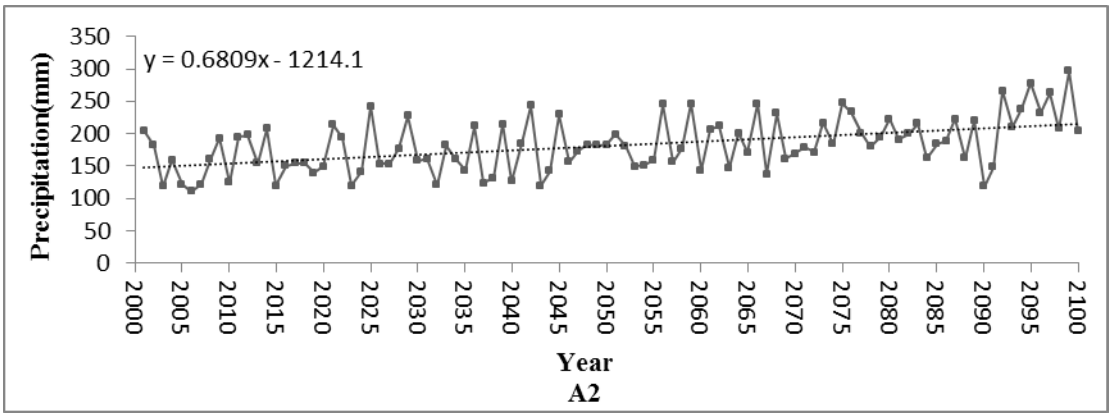
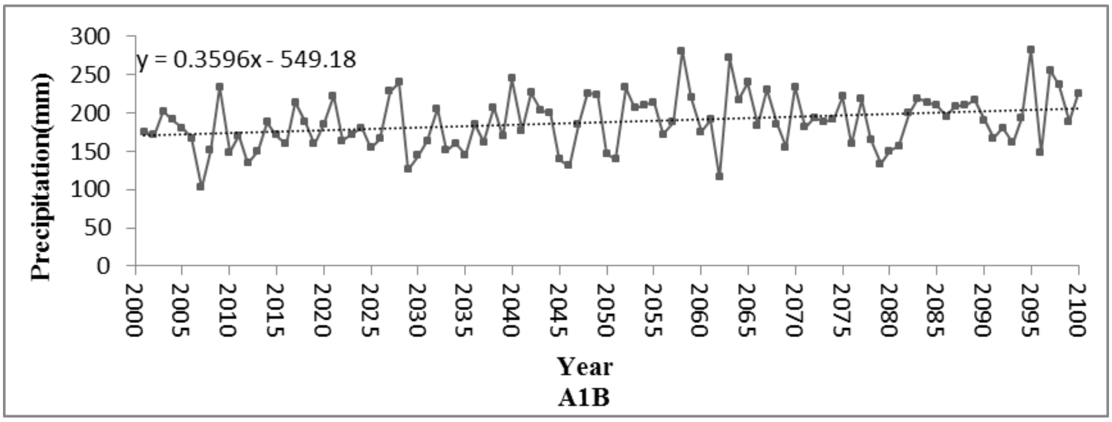


Figure 7. The trend of average PET given by CGCM 3.1.





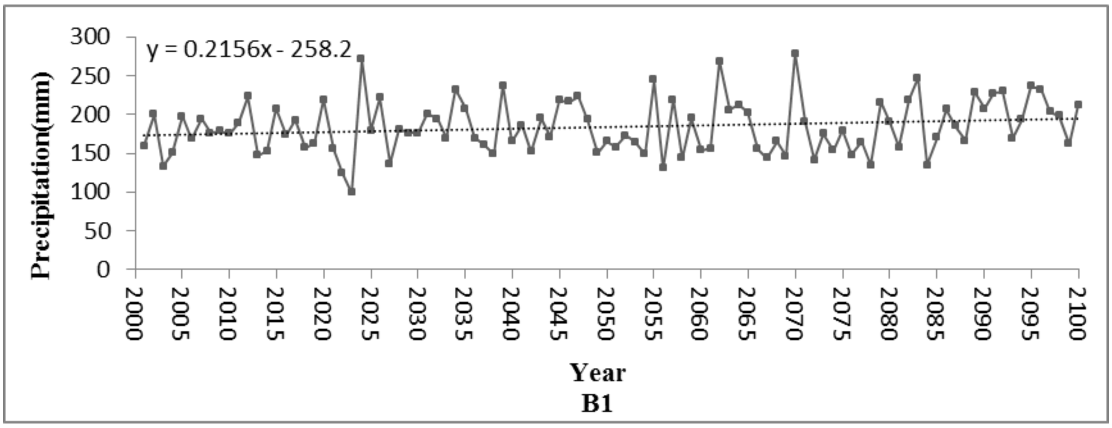


Figure 8. The trend of annual precipitation given by CGCM 3.1.

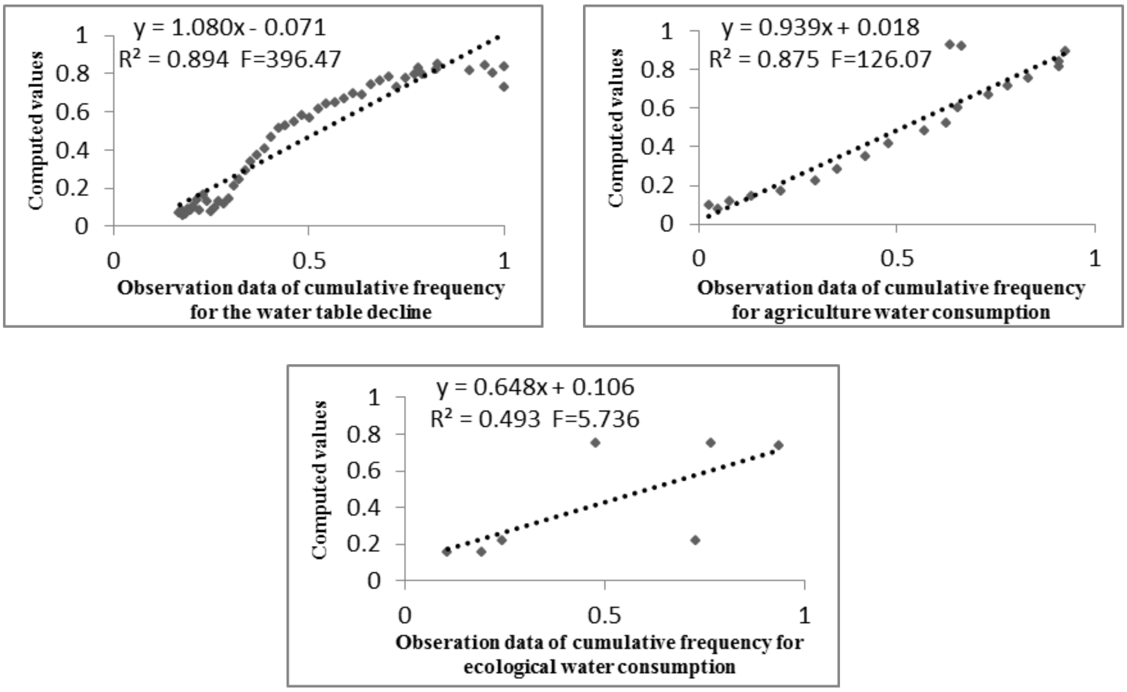


Figure 9. Comparison results of cumulative frequency between computed values and the observation data.

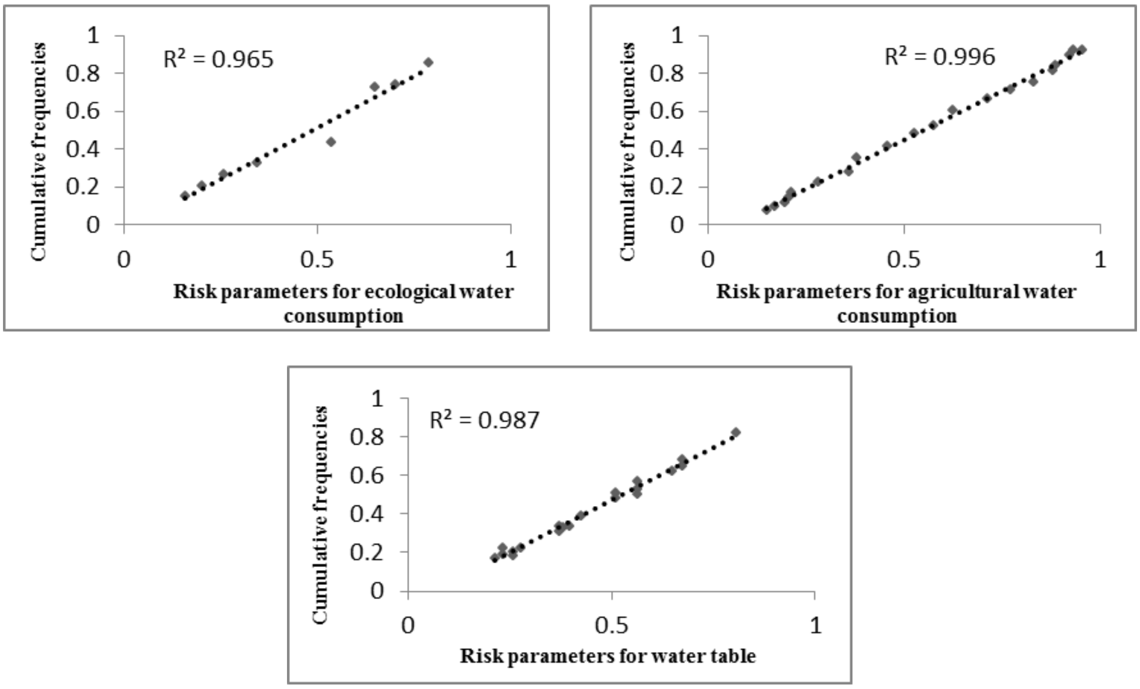


Figure 10. Comparison results between two methods.

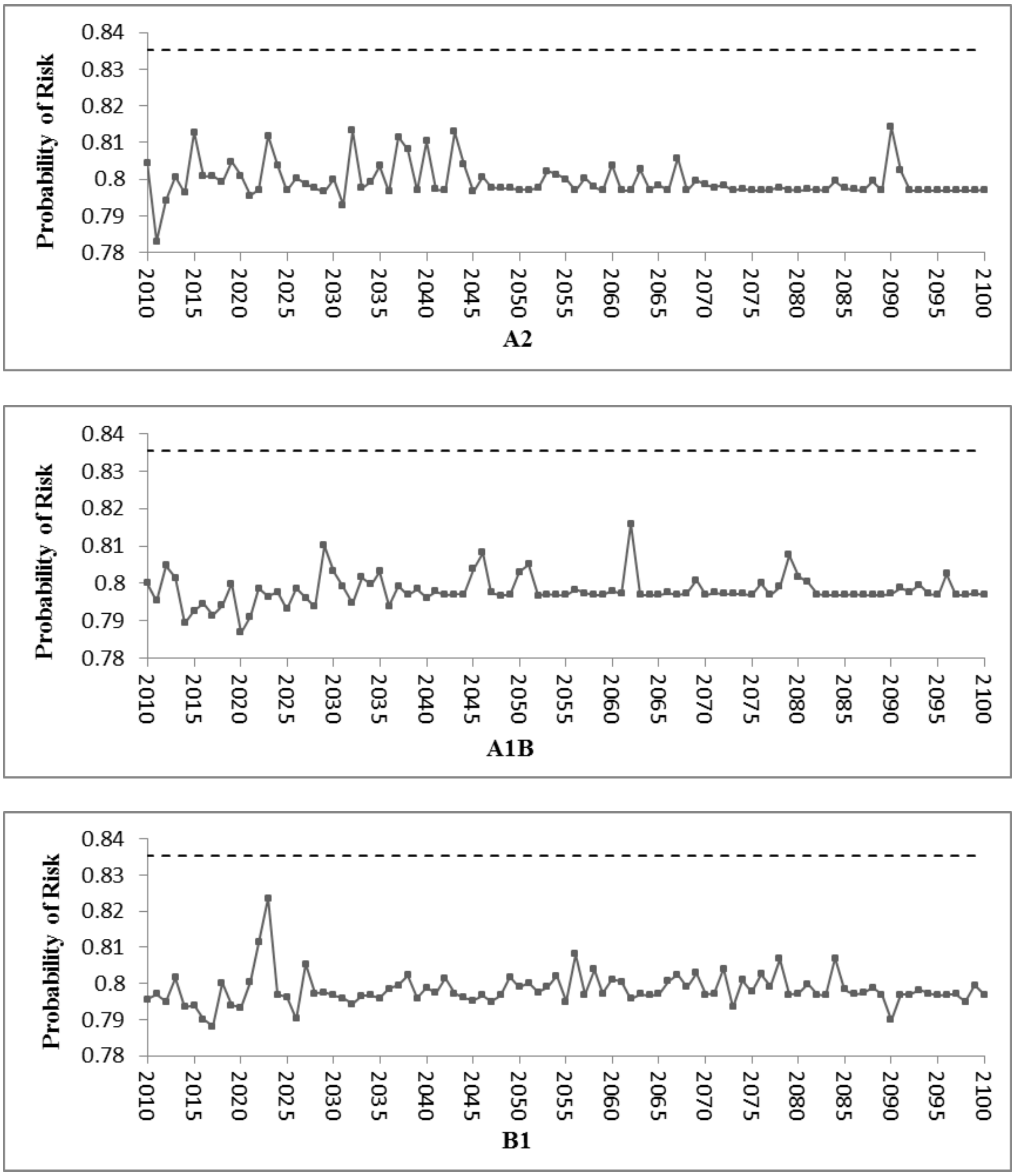


Figure 11. Comparison results for cumulative frequencies in different climate scenarios.