Responses to Referee #1:

Major point concerning the use of the Standardized Precipitation Index

1. The use of the SPI is never motivated and the focus on the 12-monthly timescale (page 1390) is not motivated either.

2. The SPI is based solely on precipitation. In a study on drought, evaporation should be discussed as well - or the title and focus of the study should be changed to "Precipitation variations in China...."

1. We have added a description, why the use of the SPI – especially with a focus on the 12-monthly timescale - is motivated.

2.2 Drought Indices

Drought is a complex phenomenon that mostly describes a persistent deficit of precipitation, usually of a long duration in a particular area. In general, indices of drought are based on precipitation amount and they may differ in other climatological parameters (such as temperature, pressure, and evapotranspiration and soil moisture) that could be taken into account (Bordi and Sutera, 2002). There are many drought indices such as the Palmer Drought Severity Index (PDSI), which has been widely used. Comparing with it, the SPI (Standardized Precipitation Index), which represents water deficit and surplus, derived from precipitation alone, can be easily computed and as the index is standardized, it is also suitable for comparing different regions or watersheds (Bordi et al., 2004). The SPI was developed by McKee et al. (1993) for the identification of dry/wet periods and to evaluate their severity (Moreira et al., 2006). Up to now, the SPI has been used e.g. for drought monitoring in Colorado, United States (McKee et al. 1993), for analyzing droughts in Western Iran (Raziei et al., 2009), for drought monitoring over Africa (Barbosa et al., 2009), and for assessing regional drought in the Marche region, Italy (Bordi et al., 2001). The SPI can be calculated for shorter or longer time scales, which may reflect lags in the response of different water resources to precipitation anomalies (Paulo et al., 2003). Originally, the SPI was calculated for 3-, 6-, 12-, 24- and 48-month time scales (McKee et al., 1993). A 3-month SPI reflects short and medium term moisture conditions and provides a seasonal estimation of precipitation, comparing the precipitation over a specific 3-month period with the precipitation totals from the same 3-month period for all the years included in the historical record (Tsakiris and Vangelis, 2004). As the time scale increases, the SPI responds more slowly to changes in precipitation, and results for the 12-month time scale identify dry periods of long duration, which relate with the global impact of drought on hydrologic regimes and water resources of a region (Moreira et al., 2008).

Due to this reason, in this study the 12-month time scale has been selected.

2. The SPI is known as a drought index. A meteorological drought can be defined as negative departure of precipitation from normal conditions. Based on this definition we would like to keep up the title of our manuscript.

Major point about the use and presentation of the model results

1. The ECHAM5/MPI-OM model is one of the best models available for this type of work (although it may have problems when the soils dry out). However, there is no information in the ms. that all available scenario runs are used - and this is essential to get a sound understanding of potential future changes in precipitation. For the A2 scenario, 4 runs are available, and for the A1b and B1 scenario, 3 runs are available. These simulations can readily be obtained via the KNMI Climate Explorer (http://climexp.knmi.nl). The information in these runs should be used to get an understanding of the variability in one ensemble, en allow the authors to test their derived trends against the spread in the ensemble. These features can all be tested in the Climate Explorer by the way.

This point directly relates to Fig. 5 and its discussion. This figure shows the results for the three greenhouse gas scenarios and there is considerable variation between the simulations. These variations concern the amplitude of the trend, but - more worryingly - the sign as well! The authors must make clear what part of this signal is robust and what part of the signal is likely to be related to climate variability.

If the authors want to excel, they could use a 17-member ensemble from the ESSENCE project, which is based on the same model and the same greenhouse gas scenarios, but with the added value of a much larger ensemble. These simulations are also available from the KNMI Climate Explorer.

2. Producing a projection of future changes in precipitation obviously requires a thorough comparison between model results and observations. This comparison relates directly to the relevance of the conclusions of the study. There are a few comments to be made on this aspect of the study. Annual averaged precipitation of model and observations are compared, the authors should change this to a comparison on a seasonal basis, given the strong seasonality of precipitation over China. Furthermore, the variations in precipitation, averaged over complete China, are compared between the observations and the model in fig. 3. The authors claim that the trend in precipitation towards wetter conditions from 1980 onwards is similar between model and observations. The resemblance is not too convincing, but the authors should make their comparison more quantitative.

Moreover, if the authors want to present results for the catchment areas for the major Chinese rivers, they should do a similar analysis for each of these catchment areas. The results shown in figs. 6,7 & 8 are only relevant if this analysis is included in the paper.

1. This paper just focus on the basic variation of dryness/wetness in the future in China, so we just have calculated the SPI based on the averaged results of the 3 runs for the A1B and B1 scenarios, and of the 4 runs for the A2 scenario. Thank you for the good suggestions. However, we feel that could be a topic for another paper.

2. According to the advice, we have calculated a comparison of precipitation on a seasonal basis (Tab. 2). Furthermore, we have calculated correlation coefficients

between the simulated and the observed data for each of the ten large river basins (Tab.3).

Tab. 2 Comparison of seasonal precipitations in China during 1961-2000 (mm)

Season	Category	Observed Data	Simulated Data	Simu.-Obser.
Spring	Average value	211.8	243.0	31.2
	Stan. Devi.	22.8	18.3	-4.5
Summer	Average value	413.8	356.5	-57.3
	Stan. Devi.	31.1	26.5	-4.7
Autumn	Average value	166.0	141.9	-24.2
	Stan. Devi.	23.5	21.5	**-2.1**
Winter	Average value	60.6	81.1	20.6
	Stan. Devi.	16.2	13.6	**-2.6**
Year	Average value	852.2	822.5	-29.7
	Stan. Devi.	47.7	38.8	-8.9

Tab. 3 Comparison of observed and simulated annual precipitations in ten large river basins in China during 1961-2000

River basins	Songhuajiang River Basin	Liaohe River Basin	Haihe River Basin	Yellow River Basin	Huaihe River Basin	Yangtze River Basin	Pearl River Basin	Southeast River Basins	Southwest River Basins	Northwest River Basins
Correlation coefficient	0.80	0.72	0.67	0.59	0.59	0.57	0.64	**0.15**	**0.17**	0.5

Other points

• §2.2: The authors may want to expand the explanation of the SPI to include a little more detail, so that readers unfamiliar with this index can understand it too. It is central in the analysis of this study; it should receive the attention it deserves.

A detailed description of the calculation procedure is now provided.

The method of computing the SPI is described by Bordi and Sutera (2002)as follows: If x is the random variable, the gamma distribution is defined by its probability density function as

$$g(x) = \frac{1}{\beta^{\alpha}\Gamma\ (\alpha)} x^{\alpha-1} e^{-x/\beta} \qquad \text{for } x > 0$$

where α >0 is a shape parameter, β >0 is a scale parameter and Γ (α) is the gamma- function.

The value of α and a can be get as following:

$$\tilde{\alpha} = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right), \quad \tilde{\beta} = \frac{\bar{x}}{\tilde{\alpha}},$$

where

$$A = \ln(\bar{x}) - \frac{\sum \ln(x)}{n},$$

with n the number of observations in which some precipitation has occurred. In *addition $\bar{x}$, given a particular month, is the mean of the cumulative precipitation computed for the same month for the different years in the record.*

The cumulative probability, letting $t=x/\widetilde{\beta}$, becomes the incomplete gamma-function:

$$G(x) = \int_0^x g(x)dx = \frac{1}{\Gamma\ (\widetilde{\alpha})}\int_0^x t^{\widetilde{\alpha}-1}e^{-t}dt\,.$$

Since the gamma-function is undefined for x=0 and the precipitation field may contain zeros, the cumulative probability becomes

$$H(x) = q + (1-q)G(x),$$

where q is the probability of zero precipitation.

H(x) is then transformed to a normal variable Z by means of the following approximation:

$$Z = SPI = -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad \text{for } 0 < H(x) \le 0.5\,,$$

$$Z = SPI = +\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad \text{for } 0.5 < H(x) < 1\,,$$

where

$$t = \sqrt{\ln\left(\frac{1}{\left(H(x)\right)^2}\right)} \qquad \text{for } 0 < H(x) \le 0.5\,,$$

$$t = \sqrt{\ln\left(\frac{1}{(1.0 - H(x))^2}\right)} \qquad \text{for } 0.5 < H(x) < 1.0$$

and c_0, c_1, c_2, d_1, d_2, d_3 are the following constants:

$$c_0 = 2.515517, \qquad d_1 = 1.432788$$

$$c_1 = 0.802853, \qquad d_2 = 0.189269$$

$$c_2 = 0.010328, \qquad d_3 = 0.001308$$

So, the SPI represents a Z-score or the number of standard deviations that an event deviates from the mean. The classification of dry and wet event intensities resulting from SPI computation is shown in Tab.1.

• §2.2: The authors may want to spend a little more time on the introduction and use of the Mann-Kendall tests, and refer to standard references like Von Storch (1999) or Wilks (1995) for further explanation.

We have adopted the suggested change.

The Mann-Kendall method as described in Zhang et al. (2005),assumes that the time series under research is stable, independent and random with equal probability distribution. Assume the time series is x_1, x_2, x_3, ⋯, x_n, and m_i denotes the cumulative total of samples so that $x_i > x_j$ ($1 \leqslant j \leqslant i$), where n is the number of the sample.

The definition of the statistical parameter d_k is as follows:

$$d_k = \textstyle\sum_i^k m_i \qquad\qquad (2 \le k \le N)).$$

on the condition that the original time series is random and independent. The variance *and mean of d_k are defined as:*

$$E[d_k] = \frac{k(k-1)}{4}$$

and

$$var[d_k] = \frac{k(k-1)(2k+5)}{72} \qquad 2 \le k \le N$$

Under the assumption above, the definition of the statistical index of UF_k is as follows:

$$UF_k = \frac{d_k - E[d_k]}{\sqrt{var[d_k]}} \qquad (k = 1,2,3,\cdots,n)$$

§2.2: In my opinion, the use of a spatial interpolation method after a selection of areas with a high number of dry month to produce a fully covered grid is plain wrong. There is no need to use this technique at all: you can do the analysis directly. So, for each grid box, you count the number of months with a SPI value <-1, and on this grid you calculate trends. The method followed in the paper leads to erroneous results.
We used a spatial interpolation to show the figure more clearly, otherwise the figure dense with huge numbers of points and hardly to look. However, the results are the same.

• Fig. 1: Give the names and numbers of the river basins
The names and numbers have been given in the revised paper.

• Fig. 2a: put in the 483 stations as dots. Can you give an impression over what time interval these stations have data? Do they all commence in 1961?
We have redrawn Fig 2a and now the 483 stations are shown as dots. The establishment of a country wide network of climate monitoring stations started after the foundation of the People's Republic of China in 1949. Most climate stations were established in the mid-1950s. All in all, data from 752 climate stations have been provided for this study. Anyway, especially for the period of the Cultural Revolution there exist large data gaps in many of the climatic time series. Moreover, stations have been relocated or the surrounding has changed leading to inhomogeneities in the time series. In the context of this study, we used data of 483 stations with uninterrupted and homogeneous monthly precipitation data covering the period from 1961-2007.

Data from 752 meteorological stations in China were provided by the National Climate Center of the China Meteorological Administration. For this study only 483 stations have been chosen based on the following criteria: data quality, continuity, homogeneity and the length of the data record. All these stations have uninterrupted observational monthly precipitation data covering the period 1961-2007. Only periods during 1961-2000 were chosen for testing and verifying the simulation ability of the ECHAM5 Model.

Other minor points

All of the other minor points have been revised.