

## ***Interactive comment on “A GCMs-based mathematic model for droughts prediction in the Haihe Basin, China: Multi-GCM Divide-Integration” by Dongmei Han et al.***

### **Anonymous Referee #2**

Received and published: 19 July 2016

Review of the Manuscript: ‘A GCMs-based mathematic model for droughts prediction in the Haihe Basin, China: Multi-GCM Divide-Integration’

by Dongmei Han, Denghua Yan<sup>2</sup>, Xinyi Xu, Zhongwen Yang, Yajing Lu

Not acceptable for publication in this form

The subject of the paper is quite relevant and potentially publishable, trying to combine independent GCM simulations of the SPI at the regional scale in order to get more skillful hindcasts. Then the authors apply the hybrid calibrated model to produce predictions of precipitation and number of droughts in the next future decades.

Main Comments:

1) Despite the potential interest of the manuscript, the English writing is quite poor and confuse with the presence of many syntactic grammar mistakes which puts serious problems concerning the correct understanding of the author's messages. In a future resubmission, it is suggested that authors ask for help from a speaking English native.

2) Beyond that, there are significant methodological errors thus making some of the authors' conclusions quite useless, in particular in what respects the decadal predictions. Moreover, there are relevant details in the regression model which remain unclear or absent. For instance, the regression is done point by point or for a whole area? The regression is done on a monthly or annual basis?

3) The simulation score evaluation from the MLRM must be done in cross validation mode (e.g. leave one year out strategy). That will reduce the strong over fitting effect by positively biasing the correlation skill.

4) The method is not detailed enough with some crucial points remaining unclear. For instance, does the linear regression 2 is performed in a point by point fashion with results summed over each basin or the converse is done with regression being computed for the regional SPI as a whole?

5) Throughout the paper, there is no statistical significance study present at all, both in calibration and validation period. Giving the overfitting effect, the regression coefficients may be affected by substantial levels of error. The confidence intervals for the issuing simulations (in the calibration period) and forecasts (in the future period) should be added. After applying that and looking for what is really statistically significant, maybe many of the author's conclusions regarding the SPI decadal trends and the expected number of droughts in the coming decades will become not plausible.

#### Small Comments

1) The 'Divide-Integration' model's attribute is not appropriate. 2) Authors shall mention the SPI references when it is referred for the first time in the paper. 3) Concerning the

[Printer-friendly version](#)[Discussion paper](#)

monthly MGDI model. Regression coefficients are computed for each month, rather than for the annual values? If yes, the over-fitting effect is still more drastic than when doing regression on a annual basis. 4) The style of references section is not the proper one. The body of individual references is not indented which makes difficult to distinguish each reference. 5) From the text, in line 150, authors refer to average correlation coefficient for each secondary basin. An average of correlation coefficients over an area has no sense. First, both observations and simulations are summed over the area, then the correlation coefficient is computed. 6) The Fig. 3 is not much informative, being somehow redundant (Figs 3c,d,e,f are equal). Climatological biases of the annual precipitation are preferable. It would be rather worth to give biases of the model ensemble mean.

---

[Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., doi:10.5194/nhess-2016-150, 2016.](#)

[Printer-friendly version](#)

[Discussion paper](#)

