Itemized Replies to Dr. R. V. Donner's comments

This is the itemized response to the comments by Dr. R.V. Donner, Nonlin. Processes Geophys. Discuss., 1, C317–C322, 2014, on our manuscript NPGD, 1, 841–876, 2014, titled "An improved ARIMA model for hydrological simulations".

In their manuscript, the authors present a statistical model for describing and potentially predicting monthly precipitation values from available time series. Their approach is based on a combination of cluster analysis and ARIMA modeling for each calendar month, thereby accounting for seasonality in the underlying data. Their approach is demonstrated to be superior to classical seasonal ARIMA models in a case study of precipitation data from some Chinese meteorological station.

No Response needed.

Positively speaking, I don't find any significant scientific errors standing against a publication of this work. However, I am somewhat reluctant against recommending it for publication in NPG, given the aims and scopes of the journal as stated on its website: "...furthering knowledge on nonlinear processes in all branches of Earth, planetary and solar system sciences" and/or applying "nonlinear analysis methods to both models and data". I think that the present contribution fails to address any of these two aspects - the authors neither provide any discussion of nonlinear processes, nor they apply any nonlinear methods. What they do is using rather classical linear statistical methods to the problem of precipitation forecasting - the latter could be considered a nonlinear problem, but nonlinearity of specific processes do not play any role in the present manuscript. Therefore, I doubt that the contents of the manuscript fully meet the scope of this journal. In any case, this is a question that needs to be finally addressed by the editor.

Response: First of all, we do believe our manuscript fits the scope of NPG. As the reviewer Dr. Donner mentioned in his comment that precipitation process *"could be considered a nonlinear problem"*, actually the precipitation process and almost all the other hydrologic processes are nonlinear. Furthermore, although ARIMA models look like a *"classical linear statistical method"*, it does involve nonlinear analysis in its methodology, such as the nonlinear methods used for the parameter estimation. The data series transformation such as logarithmic transformation sometime involved in stationarization of data series is also a nonlinear transformation. It is hardly to separate the nonlinear component from ARIMA model. Therefore, the work presented in the manuscript is an attempt to further knowledge on the nonlinear hydrological processes through a nonlinear analysis approach. Last but not the least, there were some articles applying ARMIA model to hydrologic data series published in NPG, e.g., Kallache et al. (2005). We would like to ask the editor to consider that our manuscript does fit the scope of NPGD and NPG and the interest of some NPG readers.

Kallache, M., Rust, H.W., and Kropp,J.: Trend assessment: applications for hydrology and climate research, Nonlin. Processes Geophys., 12, 201-210, doi:10.5194/npg-12-201-2005, 2005.

Beyond this general impression, I have several concerns regarding the presentation of the material that would call for some major revision of the text before it could eventually be considered ready for publication.

Response: Below please see our point-to-point response to all the comments.

General Comments

1. The authors speak about "hydrological simulations" in the title of their manuscript, but exclusively deal with precipitation, which would be a meteorological rather than hydrological variable. Of course, rainfall is important for hydrology, but the present title is not sufficiently specific and does not clearly reflect the contents of the paper.

Response: We agree in part with the reviewer that precipitation process itself cannot represent all the hydrological processes. Many times, the analysis of a hydrological process start from precipitation and hydrologists always consider precipitation an important component of hydrology. The difference between the interest in precipitation of hydrologists and meteorologists may lie in the fact that hydrologists mainly investigate the results and impact of precipitation while meteorologists are discovering the cause of precipitation. However, both are interested in predicting the precipitation. We still would like to change the title our manuscript as the following in order to incorporate the reviewer's concern, "An improved ARIMA model for precipitation simulations".

2. Many aspects regarding the methods and data considered in this study are insufficiently explained, or statements are too vague to be actually convincing. For details, see below.

Response: Thanks for pointing out this issue. In addition to modified and corrected issues pointed out in the comments, we revised the whole manuscript significantly and tried our best to address all points clearly. I hope you will find the significant improvement of the manuscript when we have a chance to submit our revised version.

3. Sometimes, the authors are not specific enough when using statistical terminology (e.g., regarding the use of ARMA vs. ARIMA, stationarity, cluster analysis, etc.). In several cases, additional explanations need to be added in order to allow for a fair and complete evaluation.

Response: As mentioned before, we incorporated the reviewer's suggestion and almost rewrite the manuscript. We will present this improvement in the revised manuscript to be submitted.

Specific Comments:

• p.843, ll.11-12: Do you mean an additive decomposition into stationary and nonstationary components here (e.g., obtaining stationary residuals after removing trends and seasonality in the mean)? I somehow doubt that such a decomposition always exists, especially in case of non-additive superpositions of different variability components.

Response: Sorry for the confusion resulted from our unclear statement. We corrected the statement as "A hydrological time series $\{y_t,\ t=1,2,\cdots,n\}$ could be either stationary or nonstationary." We were trying to say that some hydrological data series are stationary while most of others are nonstationary, rather than that hydrologic time series can be decomposed into stationary component and nonstionary component. Hydrological time series often shows nonstationary, only under very limited circumstances, it presents stationary, or weakly stationary, which is usually the basic assumption of most statistical analysis methods for hydrologic time series. However, if a nonstationary type of hydrological time series is identified, we can usually proceed to remove its nonstationarity to approach a stationary time series.

- *p.843, l.17: This is an ARMA model, not ARIMA.*

Response: Sorry for the mistake. Yes, it is ARMA model defined in Line 17 on Page 843. We corrected the manuscript, see Line..

• p.844, l.7: I recommend modifying the notation to differentiate between single observations and sequences, e.g., y_t *vs* $\{y_t\}$

Response: Thanks for the suggestion. We incorporated this suggestion into the manuscript.

• p.844, l.13: "quasi-periodic" has a clearly different meaning in the context of nonlinear geophysics; seasonal variation is by definition periodic, not quasi-periodic.

Response: The seasonal variation of hydrological processes can be considered periodic, but not strictly periodic, i.e., the time interval of periodic is not constant. For example, we can say sine function is periodic with a period of 2π; however, we may not be able to say the precipitation is periodic since we cannot identify its period even though we know a rough range of this period. Therefore, we use the same term "quasi-periodic" as Box (1967).

• p.844, ll.18-23: Please explain the meaning of S and D in equation (6).

Response: Sorry for missing the definition of S and D in Equation (6). We correct the manuscript. S is the period length (in month in this work) and D denotes the number of differencing passes by which the nonstationary time series might be described as a stationary ARMA process. .

• p.844, l.22: What do you mean by "seasonal autocorrelation coefficient" - this term is not clear.

Response: Thanks for pointing out this confusion. We correct the text using the standard definition of P, the number of seasonal autoregressive parameter.

• p.845, l.2: Statistical models can only be "estimated", not "calculated".

Response: Just corrected the text.

• p.845, ll.4-6: Which notion of stationarity is used here? Why don't the authors apply any statistical tests for stationarity if this is a crucial point, but restrict themselves to rather "qualitative" and "visual" evaluation of stationarity?

Response: Although a statistical test can be used to test the stationarity of a time series, plotting (autocorrelogram) approach is used more frequently in hydrologic application (Machiwal and Jha, 2012) as long as it can show clearly whether the time series is stationary.

Machiwal, D. and Jha, M.K.: Hydrologic Time Series Analysis: Theory and Practice. Springer, New Yor,. 2012.

• p.845, l.11: Reference Chen et al. (2004) seems to deal with ARFIMA models. Of course, ARIMA is a

subset of ARFIMA, but I wonder if there are no better references. In any case, some more details on the estimator should be given.

Response: We specified two references which focus on parameter estimation for ARMA models.

Mcleod, A.I. and Sales, P.R.H. An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models. Applied Statistics, 32:211-223. 1983.

Melard, G. A fast algorithm for the exact likelihood of autoregressive-moving average models. Applied Statistics, 33: 104-119. 1984.

• *p.845, l.12: Which "white noise test" is used in the proposed procedure?*

Response: Sorry for the confusion. Actually the content Line 13 on Page 845 through Line 5 on Page 846 describes the white noise test through testing the null hypothesis. We reformate the manuscript to make it clearer to readers.

• *p.845, l.16: The mathematical statement should rather read:* $\exists k_0$*,* t *such that* $corr(...) \neq 0$.

Response: Just corrected the text.

• *p.845, l.20: Number of cases of what?*

Response: n is the number of cases of sample of series for white noise test.

• *p.845, ll.20-21: This appears just to be convention rather than a strict requirement. Is this correct? If so, please emphasize this fact more clearly.*

Response: Yes, these values of m are mainly used in practice rather than a strict requirement.

• *p.846, l.7: What is "linear least variance" - do you mean "linear least squares method"?*

Response: Thanks for the correction. Just corrected the text.

• *p.846, l.8: What is L?*

Response: L is the number of time steps for which the hydrologic variable can be predicted using the model. For example, $L=1$ month in the presented case studies. Please see the modification of is definition in the revised manuscript.

• *p.846, ll.14-15: Please explain why information on inter-monthly variation may be lost in seasonal ARIMA models. There is still the auto-regressive part of the model that should account for such dependencies.*

Response: Thanks for this good question. Let us illustrate this information loss through an example. For example, after an order of 12 of seasonal differencing (term S in general seasonal ARIMA model) of a monthly time series, the original monthly series has been migrated to a new time series. A nonseasonal ARIMA model is then fitted to the new time series where the inter-monthly variation of original monthly time series has also migrated to the inter-monthly variation of the new series after seasonal differencing. This transformation of inter-monthly variation of original monthly data to the new inter-monthly variation of seasonally differenced series may result in loss of accuracy of model performance. We addressed the explanation in the revised manuscript.

• p.847, ll.7-8: I suppose that the number of groups may depend on the specific cluster analysis approach utilized, as well as on the corresponding "model" selection criterion. Which methodological specifications are used in this work (should probably be addressed explicitly in Section 4.3).

Response: Clustering analysis with Euclidean distance as distance measurement is used in this study. Different clustering analysis methods other than Euclidean distance could be used, which may result in different clustering results and thus further affect the model selection in the later step. Actually, as discussed in the Section of "Results and discussion", using an appropriate clustering analysis method could be one way to improve the approach.

• p.847, ll.9-10: Be more specific: maximum, minimum, truncated mean of what? linear regression model of what as a function of what?

Response: Thanks for pointing out this unclearness. We re-addressed these points in a clearer way in the revised manuscript.

• p.847, ll.11-12: By using such a strict classification according to the results of cluster analysis, information on similarities between months from different classes is completely neglected. However, depending on the specific clustering algorithm, there can still be significant similarities. Wouldn't it make more sense to replace the clustering step at all by some alternative approach taking into account the degree of similarity between all pairs of calendar months as weight factors in estimating statistical models. Naively, I would expect that such an approach could further reduce the model error substantially.

Response: Thanks for the suggestion. We agree with the reviewer that the suggested method may improve the model performance and we would like to try this approach later. As shown in the case study results, the model missed the accuracy in prediction of September and July and October. Although we believe this mismatch is caused by the abnormal precipitation in those months in 2001 compared to the previous at least 10 years. The approach suggested by the reviewer may overcome the model's poor performance in those monthly to some extent.

• p.848, l.10: How did you determine that 1/3 is a proper order of the power-law transformation?

Response: The selection of power order of the power transformation depends mainly on the data series and the quality of the transferred data series. The order of ½, 1/3, and 1/4 are commonly used. In our case the order of 1/3 stabilizes the variation, thus we selected 1/3 as the power order.

• p.848, l.11: I disagree that the data become stationary due to a power-law transform, since this is a monotonous transform (in the case of the exponent used here). That is, if the original data are (non-)stationary, so are the transformed data. This is probably related to a previous comment of mine: it is not clear which concept of stationarity is considered here and if it is rigorously tested for (apparently not) or just visually inspected.

Response: We agree with the reviewer. Logarithmic transformation does not stationarize a time series just as does differencing. It is commonly used to stabilize the variance of a time series rather than directly stationarize a time series; however, the reduction in the variance of a time series is usually helpful to reduce the order of difference in order to make it stationary. We incorporated the reviewer's comment and revised the corresponding text in the manuscript.

• *p.848, l.15: What kind of model selection approach has been followed here to determine the order of the ARIMA model in equation (10)? This is crucial information and must be provided in order to allow for a fair assessment of this manuscript.*

Response: We detailed the methodology section of the manuscript and addressed the Stationary Identification and Identification of the Order of ARIMA Model in the revised manuscript as the response to this comments. While one could just try some different combinations of terms and see what works best strictly, the more systematic and common way is to tentatively identify the orders of the ARIMA model by looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the stationarized time series. In this study, we use the latter approach. Since we cannot upload the revised manuscript right now, we quote the text from the revised manuscript as following.

If a time series is identified nonstationary, differencing is usually made to stationarize the time series. In the differencing method, the correct amount of differencing is normally the lowest order of differencing that yields a time series which fluctuates around a well-defined mean value and whose autocorrelation function (ACF) plot decays fairly rapidly to zero, either from above or below. The time series is often transformed for stabilizing its variance through proper transformation, e.g., logarithmic transformation. Although logarithmic transformation is commonly used to stabilize the variance of a time series rather than directly stationarize a time series, the reduction in the variance of a time series is usually helpful to reduce the order of difference in order to make it stationary.

Identification of the order of ARIMA model. *After a time series has been stationarized, the next step is to determine the order terms of it ARIMA model, i.e., the order of differencing,* d *for nonstationay time series, the order of auto-regression,* p*, and the order of moving average,* q*. While one could just try some different combinations of terms and see what works best strictly, the more systematic and common way is to tentatively identify the orders of the ARIMA model by looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the sationarized time series. The ACF plot is merely a bar chart of the coefficients of correlation between a time series and lags of itself and the PACF plot present a plot of the partial correlation coefficients between the series and lags of itself. The detailed guidelines for identifying ARIMA model parameters based on ACF and PACF, can be found elsewhere, e.g, Pankratz (1983). It should be noted that, to be strict, the ARIMA model built in this step is actually an ARMA model with if the time series is stationary, which is actually a special case of ARIMA model with* $d = 0$.

• *p.849, ll.15-16: In fact, the models listed in Tab. 4 are ARMA, not ARIMA.*

Response: Yes, those models happen to be ARMA model since the data series is stationary after pretreatment. However, given that ARMA model is actually a special form of ARIMA model with 0 order of differencing, we still prefer to term them as ARIMA model in order to be consistent in the text and without losing the generality.

• *p.850, ll.4-5: Why is the offset from zero only considered for α≠0 ?*

Response: This is a standard Box-Cox transformation method using α≠0.

• *p.850, l.6: I disagree that the data in Fig. 11 are symmetric at all.*

Response: Sorry for this confusion by the improper statement we addressed in the text. We actually state "more symmetric" after the transformation in the text meaning the variance of data series has been reduced and stabilized compared with the original data series, which is helpful for the following clustering analysis.

• *p.850, ll.7-8: Which clustering approach is used here (see comment to p.847, ll.7-8)?*

Response: In this study, we use Euclidean distance as the distance measurement for clustering analysis.

• *p.850, l.12: Again - how has the presence of stationarity been tested for?*

Response: Although the stationary test, such as unit root test and KPSS test are used to identify if a time series is stationary, plotting approach based on autocorrelation plot and partial correlation plot is used in this study. Below please see the revised text in the manuscript.

Stationary identification. *The input time series for an ARIMA model needs to be stationary, i.e., the time series should have a constant mean, variance, and autocorrelation through time. Therefore, the stationarity of the data series needs to be identified first. If not, the non-stationary time series is then required to be stationaried. Although the stationary test, such as unit root test and KPSS test are used to identify if a time series is stationary, plotting approaches based on scatter diagram, autocorrelation function diagram, and partial correlation function diagram are often used. The latter approach can usually provide not only the information whether the testing time series is stationary but indicate the order of the differencing which is needed to stationarize the time series. In this paper, we identify the stationarity of a time series from the autocorrelation function diagram, and partial correlation function diagram.*

If a time series is identified nonstationary, differencing is usually made to stationarize the time series. In the differencing method, the correct amount of differencing is normally the lowest order of differencing that yields a time series which fluctuates around a well-defined mean value and whose autocorrelation function (ACF) plot decays fairly rapidly to zero, either from above or below. The time series is often transformed for stabilizing its variance through proper transformation, e.g., logarithmic transformation. Although logarithmic transformation is commonly used to stabilize the variance of a time series rather than directly stationarize a time series, the reduction in the variance of a time series is usually helpful to reduce the order of difference in order to make it stationary.

• *p.850, ll.12-13: Please further justify why you expect to be able to extract more information from the data set of maxima. Stationarity alone appears no convincing justification.*

Response: We deleted this statement to avoid the potential confusion. We were trying to state that the characteristic information obtained from the clustering analysis show different behavior. Whereas fluctuations in the mean and minimum data series are relatively small, relatively larger variation are shown in the maximum data series. Therefore, we expected that the maxima series has more impacts on the final prediction of monthly precipitation than the other two classes.

• p.851, l.12: Please rephrase. If you would use information of the predicted month as well, this would not be a prediction anymore.

Response: The Section of Discussion and Conclusion has been rewritten and this mentioned text has been deleted. Due to the length of text, we hesitate to paste the whole section here. Please see the revised manuscript to be submitted.

• p.855, Tab.2: The models considered here are seasonal ARIMA, not ARIMA -please be more specific. The content of the last six columns is not understandable from the table header.

Response: The last six columns represent the autocorrelation coefficients corresponding to the autoregressive order specified in the rows and moving average order specified in the columns. We modified Table 2 incorporating this correction.

• p.857, Tab. 4: These models are ARMA, not ARIMA. Since you work with precipitation, this is probably not too surprising, since memory of precipitation is commonly relatively short (less than one year, and so an integration term is probably not needed).

Response: We agree with the reviewer. The models in Table 4 are actually ARMA models, or ARIMA models with 0 order of differencing. As we reply to another comment from the reviewer, we still prefer the term ARIMA models here just to be consistent with other part of the text without losing generality.

• p.858, Tab.5: From the caption and table header, it is not possible to infer the contents of the table completely.

Response: On equation, Eq. 10 is added (see below) in the manuscript and the coefficients in Table 5 is referred to Eq. 10.

Build linear regression models and determine the associated parameters for each monthly data series. For example, for the precipitation in the i-th month,

$$y_i = a_i y_{j,\max} + b_i y_{j,\min} + c_i y_{j,avg} + d_i \quad (1)$$

where a_i, b_i, c_i, and d_i are the coefficients in the model for the i-th month hydrologic parameter, e.g., precipitation, which need to be estimated, and $y_{j,\max}$, $y_{j,\min}$, and $y_{j,avg}$ are respectively the maximum, minimum, and truncated mean of the j-th class where the time series of the i-th month is identified in cluster analysis.

• pp.865-871: Terms like "line graph", "columnar section" and "function gram" (what is this?) are somewhat awkward in a caption and should be removed or rephrased.

Response: Thanks for pointing out these issues. Just rephrased all the captions.

• In general, careful proofreading is recommended. There are still some problems with the English, especially regarding missing articles and the improper use of participles.

Response: We have revised the manuscript significantly in both aspects of English grammars and structure and the manuscript. We would like to have the opportunity to submit and present it to reviewers very soon.