Hydrol. Earth Syst. Sci. Discuss., 8, C5731–C5735, 2012
www.hydrol-earth-syst-sci-discuss.net/8/C5731/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Interannual hydroclimatic variability and its influence on winter nutrients variability over the southeast United States" *by* J. Oh and A. Sankarasubramanian

Anonymous Referee #2

Received and published: 16 January 2012

In this paper the authors demonstrate how two predictive models can be used in order to forecast the winter nutrient loading (in January, February and March) with the information given by precipitation forecast.

In my opinion, the paper gives a good contribution to the research in water nutrient forecasting, but the author can do a better job in stressing and giving value to the originality of the paper. Too much effort is dedicated to the presentation of the two models – Principal Component Regression (PCR) and Canonical Correlation Analysis (CCA) – on their validation, and on data sources. Yet this is not merely the originality of the paper and the reader cannot find immediately the (good) contribution of this paper to research on nutrient forecasting. The original contribution of the paper to the nutrient forecasting could be better put in light by improving the readability of the paper itself by means of a better organization of ideas, and improve the comments on the results both in terms of comparison between the two used models, and in terms of different behavior of the models on different basins. I would recommend publication once these improvements are provided.

I really hope that my comments and suggestions are helpful to the authors.

I start from section 3. In this section before starting 3.1 the authors explain a logic scheme that is fundamental for their work. The logic scheme is the following:

1) Check correlation between observed precipitation and simulated JFM loadings with LOADEST → ok for all basins. 2) For each basin: find relevant grid points, i.e., the ones with best correlation between precipitation forecasts and observed precipitation in that basin. The variance explained by PC1 is high. 3) For each basin, among relevant grid points, the nearest are chosen. 4) Check correlation between TN loadings and PC1 of precipitation forecasts and between stream flow and PC1 of precipitation forecasts in all sites → ok for all but basin 18. 5) It is worthy to use only PC1 of precipitation forecasts.

I think that this logical scheme must be better explained. These are some considerations of mine for the authors:

(a) Point 1) is at the base of all the work: if there is no correlation between precipitation and simulated loadings, it is senseless to develop a prediction model based on precipitation forecasts. I think that this point must be emphasized. (b) Observed precipitations are not used in the two models as predictors. Thus why are relevant grid points chosen passing by correlation between observed and forecasted precipitation? Why don't you chose relevant grid points by selecting the grid points with PC1 most correlated with TN loadings? In this way there is no need to take into account the errors of precipita-

HESSD

8, C5731–C5735, 2012

Interactive Comment

tion predictions, because the calibration is done directly on forecasts. (c) Why do you check correlation between PC1 and stream flow if stream flow is poorly considered in the discussion? (d) I think that you should spend few words in section 2.4 in defining PCs components of ECHAM4.5, and define better what are the low-dimensional components of precipitation forecast. The reader who do not know ECHAM4.5 must know what are PCs component without going to the references, because they are a primary information for the article comprehension. (e) The sentence "To summarize, Fig. 2 and Table 3 provide the scope for using the low dimensional components for using .... " (line 10/13 page 10943) should be better explained, such as "Since variance explained by... is high, and correlation between... is high, as possible to see in Fig. 2 and Table 3, building predictive models with low dimensional components of climate forecasts can be a powerful tool for predictive TN loadings....."

Section 4.3 Role of antecedent flow condition. I think that the role of antecedent flow condition must be better put in light. It seems like a secondary result of the work, but it is very important. I think that this must be anticipated in the introduction, so that the reader can be better oriented reading the paper without finding the "surprise" of the antecedent flow conditions.

Section 4.4 Source of climatic information... Also this should be anticipated in the introduction, with emphasis of why it is important to consider ENSO.

In the paper the authors do mainly two things: illustrating two models and applying them two 18 basins. Nevertheless, the possible applications of the predictive models are stressed, but:

(1) Results and performances of the two models are not compared and poorly commented. (2)The performance of the two models on the different basins can be improved

Concerning this I make these considerations:

(a) Have you considered summarizing Fig.s 2, 4, 5, 6, 7 in one or more tables? This

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

can provide a synoptic view of the performance of the models site per site, and can put better in evidence sites where performances of models are better and where they are worse. This can also be a good tool to give guidelines to the comments on the sites. Maybe using letters A, B, C, D instead of symbols. (b) Are morphologic parameters (first of all, the area) of basins influent in model performances? Is there a correlation, at least qualitative, between the area and models performances? The same can be said also for years of observation in table 1. (c) Do the models behave well for the same basins or not? What can be inferred from this? Is one of the two absolutely the better, or the performance is dependent on the particular case? Why in your opinion? (d) Consider separating the validation procedure of the two models from the results of the model for all basins (section 4.1 & 4.2), so that a space is dedicated to a more schematic illustration of results.

The introduction provide a good motivation for the importance of predicting TN loadings, on the other hand, I think that that section must answer to these questions:

(a) Why do you predict TN loadings in winter? Is this season particularly important, or is it for data availability in order to test the goodness of the models? (b) Why are low dimensional models so important? This can strengthen the originality of the model. (c) Which are the main sources of nutrients in a watershed and how do they reach water courses? How do precipitation can influence nutrients concentration in JFM? (in other words, briefly a discursive phenomenology should be provided)

Section 2.3 Simulated nutrient database In my opinion something has to be said about a0, a1, a2. . . and other parameters of LOADEST of Table 2, without making the reader find the references to understand the table. Or better, are the numerical values of a0, a1, a2, strictly necessary?

Section 4.1 Line 2 page 10947 “The entire procedure (i) – (iv) is repeated 100 times. . .” Why 100 times? Is there a particular reason?

Abstract line 16 LOADEST and WQN are undefined before, who reads the only abstract may not understand the meaning.

Section 5 line 6 page 10952 “Thus, the intent of the study is to understand. . .rather than developing a skilful nutrient forecasts using low dimensional model”. I think that this must be corrected in something like: “The intent of the study is to develop a skilful nutrient forecasts using low dimensional model and, in addition understand how well climate and basin condition. . ..”. And I also think that this could be moved to the introduction of the paper. In fact most energies are spent in the paper to describe the models, and this is good, but few is spent on the results comment (as I written before). The sentence put in that way seems a bit in contradiction to how the paper is written. I think it can be transformed in the way I suggested, with adding, coherently, improvements to results comments for the basins and to comments and comparisons of the two models.

In my opinion the limits of these models have to be acknowledged: they can be calibrated once series of prediction forecasts and nutrient loadings are known. If they are calibrated for a basin, they cannot be exported to other basins, and, in addition, they cannot be predictive if anthropogenic actions (dams, use soil change. . .) are made in the watershed.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 10935, 2011.