Hydrol. Earth Syst. Sci. Discuss., 9, C5087–C5094, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C5087/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Predictability, stationarity, and classification of hydraulic responses to recharge in two karst aquifers" *by* A. J. Long and B. J. Mahler

A. J. Long and B. J. Mahler

ajlong@usgs.gov

Received and published: 2 November 2012

Authors' responses to review comments by J. Von Asmuth (Referee #3), RC C4953

Reply to opening comments and terminology: We see that some revisions need to be made regarding terminology and presenting convolution in terms of linear-systems theory. Stationary is not the common term here, but rather "time invariance." These are linear systems; some are time invariant and some are time variant, as far as the convolution part of this model. The recharge estimation part is nonlinear. We will present this in terms of its origins in time-series analysis and provide appropriate references.

Reply to comments on metrics and moments: The reviewer has brought to our attention several excellent references that are related to our metrics and their underlying moments of the impulse-response function. Von Asmuth and Knotters (2004) is a previous example of the use of moments for aquifer characterization. We plan to include a discussion of previous use of moments, how they relate to our metrics, and appropriate references in a revised version.

Reply to comments on predictability: Predictability is a general term used in the title for predictive strength, which will be formalized in a revised manuscript, as the reviewer suggests. The predictive strength will be rated as high, medium, and low for each site on the basis of the memory ratio and the percent difference in the Nash criterion for calibration and validation periods. This rating will be independent from any trial and error used to determine the best model to apply to each site; i.e., how many and what types of IRFs to use. As far as the identifiability of the metrics, principal component analysis (PCA) is useful for this. Metrics that plot close together on the biplot (fig. 4) are correlated. An even distribution of points on the plot indicates that nothing is highly correlated. We will discuss any highly correlated metrics and implications in a revised paper. Also, a correlation matrix for the metrics was provided with the supplemental information, and we can also quantify these correlations in our new discussion if useful.

Specific comments: A response from the authors follows each reviewer comment

p9579-21: There have been developments in the use of convolution for modelling groundwater head series, the use of characteristics or moments of impulse response functions and their physical interpretation. Although these methods are formulated in a time series analysis framework, they are closely related and may be relevant for the contents of this paper. See for instance: [see references in original comments]

Convolution is now widely used for groundwater heads also, see e.g.: [see references in original comments]

Authors' response: The reviewer has provided some excellent references, and we plan to cite some of these in a revised manuscript.

p9580-3: At this point, also a reference to the use of spectral analysis may be relevant, as also that is used for aquifer characterization and the methods are related. It would aid the reader if the advantages of convolution over alternatives like spectral analysis or their link are addressed. Please see for instance: [see references in original comments]

Authors' response: Indeed, a comparison of convolution and spectral analysis would be useful. We plan to add this, along with appropriate references, to a revised manuscript.

p9581-13: Is groundwater recharge the same as effective precipitation or precipitation surplus?

Authors' response: Yes, these are the same and will be clarified.

P9582-6: The sentence appears to list the assumptions when using convolution. As such, it may help to state that convolution may be used for linear, time-invariant systems in general.

Authors' response: Good point. We will add a statement to this effect.

P9582-12: In this formulation of a convolution integral, y(t) depends on u(ï′CÂÿt) which may be confusing.

Authors' response: I'm not sure if I understand this. If the confusion is that there are two different time variables, then we can clarify this.

P9582-13: system response function for y(t) may be a confusing term, as the impulse, step, block, and frequency response functions are also system responses. Why not use 'output' when u(t) is named input?

Authors' response: This makes good sense to us. We will simply use the term "output."

P9582-17: Please define N

Authors' response: N is number of days in the output record. This statement will be

added.

P9582-19: What is the length of the IRF? Isn't 'the time that the impulse persists' in principle and/or in equation (5) infinite?

Authors' response: Yes, mathematically it is infinite because we use asymptotic curves to approximate the true IRFs. However, the physical system probably does not have an infinite response, and so we define the system memory as the time at which 95% of the IRF area is to the left of the tail. This is explained on P9583-14 but will be moved up to P9582-19 for clarification. Even if one could argue that there is actually an infinite response in the physical system, the memory as defined here is useful for comparison of different IRFs.

P9582-20: Please share some thoughts on the use of exponential and/or lognormal IRFs. In: [see references in original comments]

the Gamma or PearsonIII distributions are used. In: Jury, W. A. and K. Roth, 1990. Transfer functions and solute movement through soil: theory and applications. Birkhäuser Verlag, Basel, Switzerland.

several are compared, also to an impulse response solution to the convection-dispersion equation. See also: Veling, E. J. M., 2010. Approximations of impulse response curves based on the generalized moving Gaussian distribution function. Advances in Water Resources. 33 546-561.

Authors' response: Yes, we can add some discussion here. Exponential and Pearson type III have been used previously as the reviewer points out (Von Asmuth et al. 2002). Also, Nash (1959; J. Geophys. Res., 64(1), 111–115) discussed possible parametric curves that might approximate IRFs, including exponential and lognormal curves. Berendrecht et al. (2003; J Hydrol, 278) used a lognormal IRF for groundwater applications. Chow (1954; Proc. Am. Soc. Civ. Eng.PACEAJ0097-417X, 80, 1–25) discussed application of the lognormal distribution for a wide variety of hydrological applications.

The lognormal curve, which we use, has a similarly skewed shape to the Pearson type III. The exponential curve has a mode (peak response) at zero, whereas the lognormal has a mode > 0. The two general types are appropriate for different system responses, and so we allow either curve or a combination of the two. Additional discussion with appropriate references will be given in a revision.

See: Von Asmuth, J. R., 2012. Groundwater System Identification, through Time Series Analysis. Ph.D. thesis, Delft University of Technology, Delft. for a treatment of physically based responses of several elementary groundwater systems.

P9584-23: Here, the issue of non-stationarity is addressed by citing references where non-stationary IRFs are used. As stationarity is one of the main issues of the paper, it would benefit by defining stationarity more clearly and/or discussing the term time invariance used elsewhere and/or the difference with non-linearity.

Authors' response: Yes, this is good point. Please see our response to general comments above.

P9585-6/7: These sentences postulate that the characteristics change and that is useful to separate the precipitation record, but this postulation is not really corroborated.

Authors' response: We were trying to say that the system output responds to recharge on a scale of as long as years to decades, but this was poorly worded. This will be revised along with the rest of this section.

P9585-10: Please define CDMP mathematically

Authors' response: This is the same thing as saying that wet and dry periods were defined as years when precipitation was above or below the long-term mean, respectively. This will be described in this alternative way for simplicity.

P9585-15: It is stated that this method is not previously used (for this, see also the general comments). Is it physically justified and what errors are made by doing so?

Authors' response: Here, the assumption is that the IRF is time invariant during wet and dry periods. The abrupt change from wet to dry periods certainly introduces errors, as well as does the time-invariance assumption within each period. A gradual change from wet to dry periods might make more physical sense in some cases, as referee #1 points out. However, in karst aquifers, the change from wet to dry periods might be more abrupt than gradual in some cases because of the large heterogeneities that can exist. For example, different parts of the aquifer are saturated during wet and dry periods, and these different parts can have very different conduit or fracture networks. We simply found that this was an approach that produced good results, both for calibration and validation, and this keeps the parameters to a minimum, while also allowing the shape of the IRF to change. The simplicity of this approach is an advantage, particularly for application to the classification approach. We will add a discussion of this in this section.

P9586-4: the paper would benefit if the authors refer to other publications where the use of metrics or characteristics of IRFs is treated. In the field of statistics, distribution functions are commonly characterized by their moments. Moments may provide a more general framework than the metrics presented here. For instance, what is and what isn't a ratio depends on the definition of the metrics. The use of moments is suggested in: [see references in original comments]

Authors' response: Yes, as previously stated we will include some previous publications for a more thorough discussion. We start with moments (i.e., mean, variance, skewness, and kurtosis), which are very useful. However, we decided to only use metrics that are independent of the overall IRF scale, which seemed to be appropriate for comparison of sites that are located at different distances from the recharge area, which could affect the mean and variance, even if aquifer characteristics are the same. Also, this approach is convenient for comparison of springs and wells, which have different units. Skewness and kurtosis are scale-independent, but mean and variance are not, so we use ratios. We also use ratios to compare wet- and dry-period IRFs. For the most part, the metrics are not redundant, which is seen in biplots (fig. 4) from principal component analysis.

P9587-8: 'There is less confidence:::..' Are or can the confidence intervals of the parameters and/or the simulation be used to quantify this?

Authors' response: This sentence will be replaced with a discussion of the predictive strength that will be quantitatively rated as high, medium, or low, as previously described.

p9587-13: What exactly were the selection rules?

Authors' response: This paragraph is merely an introduction to the concept and is discussed in much more detail in section 3.3 along with the specific rules. For example, a threshold value for the difference in the Nash criterion between a time-variant and time-invariant model is given to determine if time invariance can be assumed. Examples of specific sites also are given.

P9590-9: Are the parameters adjusted or estimated by trial and error or automatic parameter optimization? I think I don't really understand the process described

Authors' response: Parameter optimization doesn't work well unless the initial values are relatively close to optimized values. Therefore, trial and error was used to obtain rough approximations of parameter values that were then used as initial values for optimization. This sentence will be revised for clarity.

P9590-17: The Nash criterion ranges up to 1. This implies a model without error. How can that be? Please explain.

Authors' response: Yes, a value of unity would be a perfect fit and not achievable in reality. We will revise this sentence for clarity.

P9591-5: Trial and error is not really explicit. What were the criteria?

Authors' response: The criterion was this: which curve type or combination of curve

types resulted in the best Nash criterion? But other caveats also are given; e.g., a simpler model is preferred if it did not decrease Nash by more than 0.02. This will be explained more precisely in a revision.

P9596-4: Are the metrics novel, or are they comparable to the use of moments?

Authors' response: Yes, the metrics are comparable to the use of moments, which will be clarified. This particular set of metrics, which consists of moments and ratios of moments, have not been used before. But, the main aspect of our approach that is novel is the use of metrics in principal component analysis to characterize and classify aquifers, and this particular set of metrics is useful for this purpose. We will clarify this point in a revision.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 9577, 2012.