Hydrol. Earth Syst. Sci. Discuss., 11, C1147–C1154, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C1147/2014/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Hydrological model parameter dimensionality is a weak measure of prediction uncertainty" *by* S. Pande et al.

S. Pande et al.

s.pande@tudelft.nl

Received and published: 27 April 2014

We thank the referee for a detailed review of the manuscript. Please find in the following our response.

Comment: The manuscript is seen to be written in hurry. There is many simple mistake. E.g abbreviation etc. are not explained. The manuscript need a proper revision to make the message clear.

Response: We will revise the manuscript to ensure all possible abbreviations are explained and clarify the message along the lines of the referee's comments below.

Comment: This is true that only number of parameter should be not use to define the

predictive uncertainty of model. But as suggested by authors that only parameter range can define the predictive uncertainty of a model is not true.

Response: We agree with the referee that only the number of parameters should not be used to define the predictive uncertainty of a model. Our aim in using different parameter ranges is to demonstrate that complexity depends on the magnitude of parameters as well. Hence our claim that number of parameters is not a complete measure of predictive uncertainty.

We consider ranges of storage capacity and recession parameters in a structured manner such that the ranges of recession parameters are either smaller or larger than the median of the reference recession parameters ranges and/or the ranges of storage capacities are larger than the median of the reference ranges of storage capacities. The estimated complexities of models (for a given model structure) corresponding to these ranges would then in general bring out the effect of parameter magnitudes. By doing so, it also brings out the effect of parameter magnitudes on prediction uncertainty.

Pande et al (2012) have analytically shown, within the presented framework, for a class of linear reservoir models that the magnitudes of recession parameters and storage capacities influence model complexity. The authors found a similar relationship between model complexity and parameter magnitudes as the one presented here. This numerical study thus extends the analytical work to a general class of hydrological model.

We would further like to highlight that a relationship between parameter magnitudes and model complexity (within the predictive uncertainty framework presented here), in addition to the influence of parameter dimensionality on model complexity, has been demonstrated to hold for types of response functions other than the hydrological ones (examples being the theory underlying Support Vector Machines and ridge regression, i.e. penalized linear regression models).

Nonetheless, we are not suggesting that parameter dimensionality does not play a role in predictive uncertainty. It does. However, what we additionally want to suggest is that

there can be cases where highly parameterized models may be less complex than not so highly parameterized model because of the magnitudes of respective parameters.

Comment: As number of parameter increases, the parameter space increase, hence difficult to find the optimal parameters during calibration that ultimately add to the predictive uncertainty of the model. In other hand parameter range increases the volume of the parameter space. Hence, parameter range alone cannot define the source of the predictive uncertainty. In my opinion, both parameter range and number of parameter need to be consider together.

Response: We totally agree that both parameter ranges and number of parameters need to be considered together. The question remains how to isolate the effect of parameter magnitudes (and hence ranges in restricted sense) on model complexity and predictive uncertainty from the effect of parameter dimensionality (or the number of parameters). This paper takes a first step in this direction. The paper demonstrates the effect of parameter magnitudes on predictive uncertainty by demonstrating that model complexity, for a given model structure, varies as parameter ranges are varied in a selective manner. It also demonstrates that the effect of model structure (partly represented by parameter dimensionality) on model complexity can be isolated if the effect of parameter magnitudes on model complexity is controlled for. In order to demonstrate the latter, we considered two model structures: SAC-SMA and SIXPAR and controlled the effect of parameter magnitudes on model complexity by having a range for SAC-SMA that is equivalent to the reference range for SIXPAR.

We agree that it becomes difficult to find optimal parameter during calibration as the parameter space increases. This paper does not calibrate considered models. The estimation of model complexity is independent of observations of a variable of prediction interest. It only relies on the data of input (forcing) variables. Nonetheless difficulty in finding an optimum introduces another source of predictive uncertainty. Our only claim, as this referee also agrees, is that in addition to this source and parameter dimensionality, parameter magnitudes influence predictive uncertainty and should not be

ignored.

We also agree that the volume of output space increases as the parameter dimensionality increases. However that increase is also controlled by the magnitude of parameters involved. Therefore there may be a case where the volume of output space is larger for lower dimensional output space because of the magnitude of associated parameters.

Comment: Let say there is model A which have N number of parameters, and ranges are narrow, and there is model B with N-n number of parameters with wide ranges. Just by knowing these information it is very hard to tell which model will have less predictive uncertainty. It is simply because, more the number of parameters more the difficult to get optimal parameter set, same if the range of the parameter is wide. But dimensionality bring more complexity in model in term of calibration process (parameter interaction).

Response: We agree that it may be very hard to tell which model will have less predictive uncertainty unless we have a theory that can relate sample size, model complexity and prediction uncertainty and the measure of complexity is such that it can encapsulate both the effect of parameter dimensionality and parameter magnitude. We have proposed and numerically implemented such a theory for hydrological model complexity.

We thank the referee for her detailed comments. Please find our responses to her comments in the below.

Comment: Authors need to explain, is the conclusion made in this study is valid for all kind of model (e.g. conceptual model, physically based model, empirical model, data driven model etc.) Paper need restructuring. It need more clarity in methodology. Is the methodology is applicable to all kind of model (e.g. conceptual model, physically based model, empirical model, data driven model etc.).

Response: The theory is applicable to all kinds of models. Arkesteijn and Pande (2013) implemented the theory for a class of flexible model structures and used it for complexity regularized model selection. We will add additional details in our revised version.

Comment: Authors need to give more details about the study area and model structure.

Response: We will provide more details about the study area and model structures in our revised version.

Comment: Authors need to explain the utility of the current study. Will the results and conclusion will changes if we use other model and other catchment.

Response: The theory and algorithms presented in this paper are applicable to other models and catchments. We have used multiple catchments and found the same ordering of complexities for a given parameter range (parameter magnitudes). We will incorporate the above discussion in our revised version.

Comment: Authors have used Vapnik–Chervonenkis (VC) generalization theory to make the relationship between prediction uncertainty, sample size and model complexity, is there any other methods can be use, please explain, why this method was selected?

Response: We have implemented a frequentist approach to define a relationship between prediction uncertainty, sample size and model complexity. We have also discussed Bayesian approach in our paper that may define the above stated relationship in another fashion. We provided the reasons why we use the above stated theory in the introduction. In particular we emphasized that the presented approach can be applied without making any assumptions on the underlying distributions that generate hydrological observations. Bayesian approaches, though powerful, rely heavily on such assumptions. However we will expand upon it further in our revision.

Specific comments:

Comment: ..parameters) are considered less complex and hence are asso- ciated with

low prediction uncertainty." This statement need some references. Line 5 "One can envisage a case wherein a model has significantly higher model com- plexity than another model yet the tradeoff between model performance and complexity may deem the more complex model with lower prediction uncertainty." this statement need some references.

Response: We will add appropriate references.

Comment: Page 2559 Line 16 to 25 Please explain all the abbreviation before it uses. (e.g. ANN,SVM etc.) Page 2560 Line 6 "SACSMA" should be SAC-SMA Line 14. Authors have used some places short form of Section and other full, so please be consistent. Response: We will enter the full forms throughout the paper in the revised version.

Comment: Page 2562 What is time scale of the modelling?

Response: Time scale of modeling is daily.

Comment: Please explain why only 500 point.

Response: The algorithm to quantify complexity is computationally intensive. Hence the choice of 500 points. We are currently investigating possibilities to reduce its computational complexity.

Comment: Will the results change if we take more or less than 500 points?

Response: Larger sampling points are always desirable to sample from a parameter range. The 500 sampled points result in a distribution of quantified complexity. Our objective is to compare the distribution of complexities for different parameter ranges. If the distributions are sufficiently different for different parameter ranges, larger sampling of parameters from respective parameter ranges will not change such a conclusion. Such a logic underlies statistical tests such as t-statistics that estimates the significance of difference in means of two distributions for a given sample size.

We will incorporate the above discussion in our revised manuscript.

Comment: Page 2563 It has been mentioned that algorithm 2 and algorithm 1 has been used before explaining the algorithm. For readability, it will be nice to explain the algorithm first. For that reason I will suggest to move the algorithm in Methodology section.

Response: We will move the algorithm to the methodology section.

Comment: In many places it has been mentioned "The algorithm is obtained from Arkesteijn and Pande (2013)" so please avoid the replication.

Respone: We will avoid such a replication.

Comment: Page 2563 Authors mentioned about catchment that they are different, but fail to mention in what sense they are different. Please give more detail about the catchment.

Response: We will provide additional details about how the catchments are different in the revision.

Comment: Algorithm suddenly appear in text without any connectivity. Please move them in methodology section.

Response: We will do as suggested.

Comment: Page 2565 Please explain what the advantage of using these algorithm are.

Response: We will do so.

Comment: Table 3, please add flow information to understand the water balance in the catchment.

Response: We will do so.

References: Pande, S., Bastidas, L. A., Bhulai, S., McKee, M. (2012). Parameter dependent convergence bounds and complexity measure for a class of conceptual

hydrological models, Journal of Hydroinformatics, Vol 14, No 2 pp 443–463.

Arkesteijn, L., and Pande S. (2013), On hydrological model complexity, its geometrical interpretations and prediction uncertainty, Water Resour. Res., 49, 7048–7063, doi:10.1002/wrcr.20529.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 2555, 2014.