Hydrol. Earth Syst. Sci. Discuss., 5, S2647–S2650, 2009
www.hydrol-earth-syst-sci-discuss.net/5/S2647/2009/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Mapping model behaviour using Self-Organizing Maps" *by* M. Herbst et al.

M. Herbst et al.

Received and published: 7 March 2009

The authors greatly acknowledge the constructive work of the reviewers which allowed us to submit an improved and partly more comprehensible version of the original manuscript. In the following we briefly discuss their main issues of concern:

Referee#2 asks a) whether the study area was simulated as a lumped or distributed catchment and b) how the parameters of the distributed conceptual model NASIM fit into the distributed modelling framework and how they were assigned (to the spatial units of the model) in the Monte-Carlo simulation.

Authors: We admit that some more information could be provided here. It is indicated briefly (on p. 3523 line 17ff) that "the model simulates the soil water balance for spatially homogeneous units with respect to soil and land use. Each spatial unit is vertically subdivided into several soil layers.”

NASIM uses a spatial discretization based on sub-catchments. For the test watershed pre-processing of spatial data resulted in 71 sub-catchments with a mean size of approximately 1.8 km^2. These are further subdivided into spatially homogeneous units with respect to soil and land use. Each of these elementary spatial units is again vertically divided into soil layers. All lateral flow components that result from the processes within the elementary units are aggregated on the sub-catchment scale each passing an individual linear storage. Two of them, the interflow and the surface flow, are in a prior step transformed by convolution with the time-area relationship to integrate sub catchment characteristics into the process of flow accumulation. The NASIM parameters examined in this study (Tab. 1) are unit less factors that modify internal parameter values that are either based on global default values or have been determined individually for each sub-basin in the course of the spatial data pre-processing: The internal values modified by *RetOf* are determined in the course of the pre-processing depending on the slope in each sub-basin, while the internal *RetInf*, *RetBas*, *StFFRet* are set to global values. The internal values of *maxInf* as well as well as *vL* are determined according to soil type. Because appropriate prior information on factor distributions was missing uniform distributions were assumed. The variation of these factors during the Monte-Carlo simulation was performed with global values for all sub-catchments. Thus, parameter (factor) variation was carried out uniformly for the entire catchment. However, the local effects of the factors are scaled depending on physical and spatial characteristics of the sub-catchments and elementary units respectively.

Referee#2 contends that the notation in Eq. (7) and (12) is not clear, especially with respect to the meaning of the variables $\boldsymbol{x}$ and $\boldsymbol{y}$.

Authors: Although great care was taken regarding the completeness of the equations, we admit that the descriptions in Sect. 2.3 are a bit vulnerable to confusion, especially with respect to the meaning of variable $\boldsymbol{x}$ and $\boldsymbol{y}$. Therefore, some pertinent additions and modifications were made to the end of Sect. 2.2, Sect. 2.3 (now explicitly defining $\boldsymbol{x}$) as well as Sect. 2.4 (regarding variable $\boldsymbol{y}$). All variables and indices of Eq. (7) and Eq. (12) are completely described in the text, however, the content of variable $\boldsymbol{x}$ and $\boldsymbol{y}$ is now rendered more precisely through small additions made to the text.

Referee#2 asks how the best matching unit is determined and whether all five Signature Indices are used in this process.

Authors: The BMU of a Signature Index vector $\boldsymbol{v}$ (which could e.g. represent the time series of measured discharges or any other data item with the same dimensionality as the reference vectors) is identified by determining the reference vector (i.e. the node) that has the smallest Euclidean distance to $\boldsymbol{v}$. This process is expressed in Eqs. (12) and (7) of Sect. 2.3. All five Signature Indices are used to determine the BMU.

Referee#2 suggests exploring the stability of the SOM with respect to the initial values of the reference vectors $\boldsymbol{m}$ and supposes that a different or more favourable distribution of parameter values in Fig. 5 could result from a different initialization of the reference vectors.

Authors: According to the literature (e.g. Kohonen, 2001) the effect of the initial values of the reference vectors on the final organization of the map can be neglected. This is because already after the first iteration the reference vectors will be modified drastically so as to move closer to the input data. However, the initialization can influence the number of iterations that is required for the convergence of the map. In order to assure faster and more reliable convergence an initialization along the two greatest principal component eigenvectors of the data (see p. 3527) has proven to be a successful method.

Referee#2 notes that the numeration of some equations is incorrect.

Authors: Referee#2 is right. The second version of the paper will have the correct numeration.

HESSD

5, S2647–S2650, 2009

Interactive Comment

Referee#2 contends that the information provided by Fig. 1 and 2 is similar and that Fig. 1 may be removed.

Authors: We disagree in this point. The figures are based on the same data, however both show different meaningful aspects of it that help understanding the organisation of the data on the map. Figure 1 “highlights the trade-offs between different Signature Indices and therefore reveals which capabilities in reproducing the characteristics of the measured time series are mutually exclusive.” (p. 3532). As each of the components in Fig. 1 is scaled individually (see the scale bar) it is not very easy to spot the location or region on the map where the Signatures are closest to zero (i.e. the individual optima). From the additional Fig. 2 it becomes immediately evident that the Signature Index optima are in part mutually exclusive, however, Fig. 2 does not distinguish between negative and positive biases. Most importantly, Fig. 1 adopts a crucial role when it comes to interpreting the model parameters in Fig. 5 because “conformities or similarities with the patterns of Fig. 1 will point at high correlations between parameter values and Signature Indices. Because the Signature Indices represent meaningful hydrological characteristics, these figures indirectly reveal the function of each parameter in the hydrological context” (p. 3533). Thus, it would not be able to present one of the most interesting aspects of this SOM application if Fig. 1 was removed (see Sect. 3.1).

HESSD

5, S2647–S2650, 2009

Interactive Comment