

## ***Interactive comment on “A hybrid least squares support vector machines and GMDH approach for river flow forecasting” by R. Samsudin et al.***

**Anonymous Referee #1**

Received and published: 17 September 2010

Review report on the Manuscript “A hybrid least squares support vector machines and GMDH approach for river flow forecasting” by R. Samsudin, P. Saad, and A. Shabri

Summary of review The manuscript compares the performance of the GLSSVM model approach to more traditional approaches such as ANN, ARIMA, and LSSVM to forecast stream flow of two catchment systems in Malaysia. The paper quite extensively addresses the various model structures but lacks description on the hydrology of the basins. This causes problems to the reviewer since all simulations and forecasts are executed at monthly time steps. Considering that the basins only are of (sub-) regional scale (<1500 km<sup>2</sup>) it may be questioned whether the real world catchments response times are well presented. Analysis on the relation between rainfall and runoff are not

C2351

presented and the claim that selected model approaches are particular suitable for modeling non-linear and chaotic time series is not substantiated by the results. Actually, simulation of catchments responses at monthly time steps ignores much of the (highly) dynamic non-linear responses by rainfall. It is surprising that meteorologic (i.e., rainfall) data is not considered for use in this study. Also, there is no discussion/reflection how the increasing (large) number of weights/parameters could impact the modeling results. This specifically applies to my comment above on the (large) modeling time step. A further note on the manuscript is the writing and the presentation of figures. Besides that, there are (too) many typos also the grammar often is poor starting from the Abstract. Figures are not well readable by the small font sizes and lack explanation in the manuscript. In summary, I think the paper in its current form is not suitable for publication. Authors should aim at simulations at daily time step to make the work more attractive to the forecasting community.

Some detailed comments. – In the Introduction section, the actual objectives of this study are not well formulated and are not clear.

– The literature review does not address the time-space scales that apply to the various studies so it is difficult to become conclusive how effective approaches have been in forecasting ‘non-linear’ behavior.

– The Case study (section 3) should be interchanged with Section 2. A clear methodology is not presented but the description is diffusive over the manuscript.

– Section 3 case Study: Climatologic data should be discussed as well as characteristics such as steepness/flatness and land cover that all directly effect runoff production and response times.

– Page 3696 Line 15: the actual selection procedure is not well described.

– Page 3698 Line 27: Equation number is wrong.

– Section 4: Time lags presumably are in ‘month’. It is questionable why time lags as

C2352

large as 2, 6, 12 months are taken. One cannot claim that there is a relation between stream flows at such large time lags particularly when catchments only are of regional scale.

– Table 4: Units are unknown since actual information on the magnitude of monthly flow discharges is not presented. Model performance assessments in terms of R are low to my opinion and suggest poor model performance for all approaches. The meaning of bold numbers is not described in the manuscript and the Table header.

– Results in Figure 9 do not indicate that non-linear behavior is well represented. Actually, none of the higher peak values are well simulated by any of the approaches. To my opinion, any simple ANN should be able to simulate the hydrographs with much higher accuracy as presented in this study when better input signals are selected.

– For many tables: References to the numbers in the manuscript are not always clear and sometimes even wrong.

– Annotations of the figures need to be improved: e.g. lag is in ?

Some Editorial comments on the first few pages. I note that the list with comments actually is much longer for these pages only but it takes too much time to address all required improvements.

Page 3692 line 12-14: Modify sentence to improve grammar. Page 3693 line 3-4: The use of terms 'full of nonlinearity and chaotic' is not very common in hydrologic modeling. Page 3692 lines 16-19: This introduction is poor and importance of forecasting is overstated. Page 3692 line 24; ARMA and ARIMA are mixed. Page 3693 line 2: 'is only' needs to be replaced by 'belongs to' the class of. . . Page 3693 line 5: 'More advanced can be removed Page 3693 line 11: Modify sentence to improve grammar. Page 3693 line 20: The use of phrases like 'a hot topic of intensive study' is not very common in scientific literature Page 3692 lines 18-23: Sentence is too long, not well readable and requires improvement. Page 3707 line 9: Modify sentence to improve

C2353

grammar. Page 3707 line 16: Modify sentence to improve grammar. Page 3708 line 10: Modify sentence to improve grammar.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 3691, 2010.

C2354