Reply to short comment on “Does the simple dynamical systems approach provide useful information about catchment hydrological functioning in a Mediterranean context? Application to the Ardèche catchment (France)” by Eric Gaume

M. Adamovic et al., January 2015

In the following, the reviewer comments appear in black italic and our answers are provided in blue.

1) *First of all, my deepest apologies to the authors and the editors for this extremely late review. The manuscript is interesting, clearly written and documented and in the scope of HESS, but has in its present form some defaults that have to be corrected before publication. It evaluates the performances of a simple conceptual global rainfall- runoff model based on a 3-parameter non-linear reservoir (eq. 12 of the manuscript) in simulating hourly discharge series of small watersheds. This model and its calibration procedure were initially introduced by Kirchner (2009) and used in several recent works (Krier et al., 2012, Brauer et al., 2013). The application of this approach to Mediterranean watersheds is the main originality of the manuscript according to its authors.*

   Answer: We thank Eric Gaume for his short comment of the paper content. We take it as an opportunity to make a few things clearer and to improve our paper. E. Gaume raises many issues in this comment and we tried to answer them as precisely as we could.

2) *First, the selected database appears to be of poor quality: the available measured series are short - less than 10 years - and the yearly water balances appear implausible for 3 out of the 4 considered test watersheds, indicating flux estimation errors.*
   These problems are acknowledged by the authors (p 10734) but their answers are moderately convincing. The authors suggest a correction of both - estimated actual evapo-transpiration and precipitation - to reach an annual balance. As a result, they work on artificial "scaled" data which limits their demonstration. A more in-depth critical analysis of their data would certainly have revealed estimation problems due to poor rating curves (according to published data, the streamflow of the Borne at Saint-Laurent-les-Bains (95 km^2) is equal to 880 mm/year, comparable to the other provided data). Likewise, the precipitation amount on the Altier Watershed (4) is surprisingly low if compared to the other available values. The whole work would have been much more convincing if based on good quality data.

   Answer: We agree with the Reviewer that data quality issues are important. This aspect was also pointed out by the students' reviews and some elements can be found in the reply to their comments[1]. Several answers can be given:

   - First, we agree on the interest of well monitored and controlled catchment for scientific studies. However, here we were specifically interested in testing the simple dynamical

[1] http://www.hydrol-earth-syst-sci-discuss.net/11/C6174/2015/hessd-11-C6174-2015.pdf

system approach for catchments, where only operational data are available, and that are more representative of the real world. This specific objective was also highlighted in the paper, but maybe not clearly enough.

- We chose to focus on the Ardèche catchment in this study because we wanted to be able to document site-specific conditions according to local knowledge, which was made possible in the framework of the Floodscale project (Braud et al., 2014). As a preliminary step to this study, we thoroughly analyzed the stations and their functioning with the help of the local authorities in charge of the network (SPC Grand Delta and EDF) who provided the rating curves and gaugings. The stations that are influenced by dam operations (Ardèche at Pont d'Ucel, Pont de Labeaume, Vogüé, Vallon Pont d'Arc, Sauze, Chassezac at Gravières) or present obvious rating curve or continuity problems (Beaume at Rosières, Volane at Vals-les-Bains) were discarded from our dataset. This explains why we ended up with only 4 stations. As indicated by the Reviewer, another strategy would have been to use research-grade data from experimental watersheds, but it is not the purpose of such a data-driven approach, which loses much of its interest if the contact with local knowledge is lost (see below for further development on this specific question). However, we would like to underline that the work of a Master student recently applied the same approach to a larger sample of catchments (20) in the Cévennes area (Coussot, 2015; see also reply to Referee#1[2]). Although preliminary, the results of this Master work confirm our results, with the same issues of data quality and catchment mass balance, and same conclusions on the applicability of the simple dynamical system approach on Mediterranean catchments. In conclusion, this study enabled us to identify the problems with measurement networks, on the one hand, and to better understand the catchment functioning taking many climate forcing uncertainties into account, on the other hand. We believe that our results are of interest, as they point out that, when provided data uncertainty is correctly handled, the simple dynamical system approach is applicable to Mediterranean type catchments.

- On our sample of stations, it seems that discharge estimation problems due to poor rating curves are not the main problem. Ongoing work focuses on the estimation of rating curves and their uncertainties (Le Coz et al., 2014; Branger et al., in preparation, see also Reply to Referee#2[3].). The application to stations in the Ardèche catchment shows that the uncertainty related to rating curves, although not negligible, especially for peak flow values, is not an explanation for the mass balance discrepancies that were found in the data. The main problem comes from the estimation of the precipitation and/or evapotranspiration. Evapotranspiration is not measured, and precipitation is difficult to measure in these mountainous areas. In the particular area of the Altier catchment (pointed out by the Reviewer), we used the SAFRAN reanalysis which was the only continuous rainfall data source for the upstream areas. SAFRAN has drawbacks: in particular it seems to underestimate the rainfall in the particular area of the Altier catchment. Other rainfall estimations are being developed in the framework of the

[2] http://www.hydrol-earth-syst-sci-discuss.net/11/C6170/2015/hessd-11-C6170-2015.pdf

[3] http://www.hydrol-earth-syst-sci-discuss.net/11/C6172/2015/hessd-11-C6172-2015.pdf

Floodscale project (see papers by Delrieu et al., 2014), but were not yet available at the time of our study.

- We must also point out that the values given by the Reviewer for the Borne at St Laurent les Bains – Pont de Nicoulaud catchment are erroneous, whereas the values in the paper are correct: the catchment surface is ~63 km² and not 95. We calculated the catchment surface based on the position of the station and the 25 m IGN DTM. The Banque Hydro database is in agreement (62.7 km²). However, the Wikipedia page of the Borne river[4] provides this 95 km² value, which may come from a previous erroneous publication. We will suggest a correction. Thus, a streamflow of 880 mm/year is also erroneous; the value presented in the paper (1579 mm/year) is correct considering the 2000-2008 period (the Banque Hydro database indicates 1357 mm/year for the 1969-2011 period).

- The purpose of the rescaling of rainfall and evapotranspiration input data is precisely to take into account these inconsistencies in the dataset that could not be solved using only available measurements. This explicit operation avoids, for example, having the model parameters compensate for the input data uncertainty, which is a common problem of conceptual hydrological models. It also allows for a more objective evaluation of the model performance, because a model that is based on mass conservation cannot work successfully on catchments with obvious mass balance problems.

3) *Moreover, the lengths of the available series does not allow for a validation of the calibrated models. To my opinion, validation (based on split-sample tests) is an absolutely necessary step of any model implementation work in hydrology. No work should be published without validation results. This is missing here and should absolutely be added.*

Answer: We do not agree with this remark. For example in the work of Melsen et al. (2014), the authors concluded that one winter season (November until March) can give reasonable results with two-parameter model in a small Alpine catchment (3.31 km^2). Considering that the Ardèche catchments are larger and more heterogeneous in terms of geology and land-use, we considered a nine year period which is sufficient to estimate the parameters of the g(Q) function in a robust manner.

The Reviewer also points out that the results are not validated using independent data. We cannot agree with that either. The g(Q) function is estimated using only a small part of the streamflow time series: only nighttime, rainless hours during the non-vegetation periods (November-March) of each year. The model is then run without additional calibration for the rest of the year. Therefore, although not a classical split-sample test, the model validation is performed on independent data.

4 *http://fr.wikipedia.org/wiki/Borne_%28rivi%C3%A8re_de_l%27Ard%C3%A8che%29*

4) *Second, the authors put forward the novelty of the proposed approach. This is also questionable. This approach is not uninteresting in its formulation, but far from new. What is proposed is a relatively standard method based on a non-linear reservoir for simulating recession curves. Such models exist since the very first hydrological model development works in the late sixties. The 3-parmeter non-linear reservoir drainage law (eq. 12) may be new. But by the way, the justification for the specific form of equation 12 is missing. Even, the retrieval of rainfall based on discharge measurement is not new: it was for instance the objective of the so-called DPFT method developed in France and that authors certainly know and should have cited (see for instance Sempere Torres et al, Natural Hazards, 1992). Finally, the proposed approach leads to the development of a 4-parameter conceptual rainfall-runoff model (3-parameters for the non-linear reservoir and 1 parameter for the rescaling of data ensuring mass- conservation), and this model only works in winter times. This is not particularly novel. Many conceptual models have been proposed and tested during the last 30 to 40 years in hydrology and it would be essential to evaluate the added value of the proposed model, comparing it to other existing models of the same type. This comparison should be added to my opinion in the proposed manuscript.*

Answer: We agree with the Reviewer that many recession models date back to the late sixties. However, there must be misunderstanding in the specific originality of the simple dynamical system approach. What is new in this approach is not the reservoir itself, but the manner to derive its structure and parameters from the data analysis: in particular, here the functional form of the storage-discharge relationship is not specified a priori, but determined directly from data (Kirchner, 2009). This is the very definition of the top-down or data-driven modelling approach, that is acknowledged to be a major paradigm shift in modelling by the hydrological community that occurred during the PUB decade (see for instance Sivapalan, 2003; Hrachowitz et al., 2013). Therefore we argue that testing this kind of approach on new datasets, for various climatic conditions, contributes to the advance of hydrological science in itself. We have also compared the model results with other models that are based on similar data-driven methodology (e.g. Brauer et al. (2013) and Melsen et al. (2014)) and obtained similar results. This mention was maybe not clear enough and will be added in the paper. The comparison with other more parametric models is not relevant for our study.

A few more detailed remarks:

- We are aware of the DPFT method proposed by Torres et al. (1992) and further revisited by Duband et al. (1993). However, the purpose and principles of the DPFT are different from the approach presented in our paper. The DPFT method is an event-based method where net precipitation and the unit hydrograph are identified at the same time, using optimization techniques between the simulated and observed discharge. In our application, the discharge sensitivity function g(Q) is estimated using non-vegetation periods (not only for selected events), and is derived from data analysis only (there is no optimization between measured and simulated discharge, which would make reproducing the hydrograph into a nearly trivial exercise). In our paper, the discharge simulation and the rainfall retrieval are two ways of assessing the relevance of the identified discharge sensitivity functions, which is estimated a priori, using only discharge fluctuations and discharge data (rainfall is only used for the selection of the points used in the g(Q) estimation).

- The SDSA model consists of 3 parameters and not 4. The rescaling is not systematic and performed independently from the model performance as explained before. It is just a way of ensuring mass balance in the catchment so that the model does not have to compensate for problems in the input data. Thus it cannot be considered as a calibration parameter (by the way, there is one rescaling coefficient for precipitation and one for evapotranspiration, which makes 2). In our study, no rescaling was done for the Ardèche at Meyras catchment. The SDSA was also used as a basis for the semi-distributed SIMPLEFLOOD model (Adamovic, 2014), and was applied to the whole Ardèche catchment without data rescaling, based on the SAFRAN forcing. The results pointed out systematic volume underestimation by the model. Further work will use improved rainfall forcing such as the radar/rain gauges reanalyses proposed by Delrieu et al. (2014) to see if those underestimation problems are linked to poor rainfall forcing.

- The judgment that the model works only in winter time is not correct. Our analysis showed that the model performs better during winter periods. However, NSE values calculated on the log of discharge are above 0.70 for 3 of the 4 catchments and the whole 2000-2008 period, which do not precisely indicate that the model “does not work”. Maybe the graphs presented in log scale were confusing on that point.

5) *Finally, and line with this last comment, the whole manuscript gives the uncomfortable impression that the authors try to reinvent hydrology and hydrological modelling from scratch, without considering the past. One of the last comments of the paper on page 10756 is particularly illustrative of this state of mind. "Our result suggest the existence of another storage, probably more superficial than the "Kirchner" storage which could be used to supply evapotranspiration...". What a discovery ! This reservoir is called soil and taken into account in most of the RR models and the central concern of the SWAT models. This certainly false impression could easily corrected by a better formulation and putting less emphasis on the novelty of the proposed method.*

<u>Answer:</u> We are not sure to understand this remark very well. The data-driven approach that was proposed by Kirchner and was tested in this study is not just reinventing the wheel; it presents real advantages in terms of consistency between model structure, parameters and observed data, as already explained above, and it is amenable to much more rigorous testing than typical RR models, since one never optimizes any time-series fits to either precipitation or runoff. The models obtained through this approach are simple, with a limited number of parameters that can be estimated from the available data. The main hypothesis underlying the SDSA approach is that the major contribution to the flow can be approximated by drainage from a single non-linear reservoir, the form and parameters of which can be estimated directly from recession analysis.

Our analysis shows that for our catchments that have high evapotranspiration rates in summer, the simple assumption that was used so far (PET=AET) is not fully adequate, and that an alternative evapotranspiration model should be used. Adding a superficial storage to the existing one is one possible way of dealing with this. This could be seen as a superficial soil reservoir.

However, we interpret that the subsurface flow that is produced by the current model comes probably also from the soil. We don't think that attributing specifically a specific originating zone for each flow component is of particular interest for a top down approach.

Detail remark: why specifically the SWAT (Soil Water Assessment Tool) model? Or did you mean SVAT (Soil Vegetation Atmosphere Transfer) models?

Finally our results still confirm that the main mechanism we are speaking about is quick sub-surface flow which transits through the reservoir considered in the Simple Dynamical System Approach.

References

Adamovic, M.: Development of a data-driven distributed hydrological model for regional catchments prone to Mediterranean flash floods. Application to the Ardèche catchment (France). , Université de Grenoble, Grenoble, France, 05 December 2014, 298 pp., 2014.

Braud, I., Ayral, P. A., Bouvier, C., Branger, F., Delrieu, G., Le Coz, J., Nord, G., Vandervaere, J. P., Anquetin, S., Adamovic, M., Andrieu, J., Batiot, C., Boudevillain, B., Brunet, P., Carreau, J., Confoland, A., Didon-Lescot, J. F., Domergue, J. M., Douvinet, J., Dramais, G., Freydier, R., Gérard, S., Huza, J., Leblois, E., Le Bourgeois, O., Le Boursicaud, R., Marchand, P., Martin, P., Nottale, L., Patris, N., Renard, B., Seidel, J. L., Taupin, J. D., Vannier, O., Vincendon, B., and Wijbrans, A.: Multi-scale hydrometeorological observation and modelling for flash flood understanding, Hydrol. Earth Syst. Sci., 18, 3733-3761, 10.5194/hess-18-3733-2014, 2014.

Coussot, C.: Assessing and modelling hydrological behaviours of Mediterranean catchments using discharge recession analysis. Master Thesis, HydroHazards, University of Grenoble, France, 2015.

Delrieu, G., Wijbrans, A., Boudevillain, B., Faure, D., Bonnifait, L., and Kirstetter, P.-E.: Geostatistical radar–raingauge merging: A novel method for the quantification of rain estimation accuracy, Advances in Water Resources, 71, 110-124, http://dx.doi.org/10.1016/j.advwatres.2014.06.005, 2014.

Duband, D., Rodriguez-Hernandez, J. Y., and Obled, C.: Unit hydrograph revisited: an alternative iterative approach to UH and effective precipitation identification, Journal of Hydrology, 150, 115-149, 1993.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrological Sciences Journal, 1-58, 10.1080/02626667.2013.803183, 2013.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resour. Res., 45, W02429, 10.1029/2008wr006912, 2009.

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrological Processes, 17, 3163-3170, 2003.

Torres, D. S., Rodriguez, J. Y., and Obled, C.: Using the DPFT approach to improve flash flood forecasting models, Natural Hazards, 5, 17-41, 1992.