

## Anonymous Referee #2

Received and published: 12 May 2017

I regret to inform you that my personal opinion is that the paper needs a major revision in order to clarify the valuable work developed by its authors. To begin with, I really found confusing the description of the two objectives mentioned: the reconstruction of rainfall and runoff series. I would respectfully recommend writing two different papers in order to clarify different methodologies and perspectives unless you use runoff modelling results as criteria for rainfall estimation, which is a point that I couldn't clarify after reading your paper. I would highlight some other problems concerning the content of the paper reviewed:

### Authors:

A reliable description of the Ofanto river hydrograph goes through a proper description of the meteorological and soil processes, with the precipitation field playing the most relevant role. Details and references (Pappenberger et al., 2005; Zappa et al., 2010) about this approach are common in the literature and they are referenced in the Introduction.

Overall the causality dependence between a proper reconstruction of the meteorological and the hydrological events is the reason why we decided to consider both topics in the same paper. We have inserted a phrase in the Introduction that explicitly states:

*“In this paper we describe both the precipitation reconstruction and the hydrograph results since the former drive most of the quality of the hydrology of a river basin”.*

1. Considering the methodology. Don't you find quite dangerous the use of a huge amount of parameters to simulate meteorological, hydrological and hydraulic processes?. What controls the over parameterization effects on your model/system?

### Authors:

WRF-Hydro is one of the most advanced hydrological/hydraulics system currently available in the literature. We agree that this system includes a large number of tunable coefficients; on the other hand it ensures a good compromise between the number of described physical processes and the number of involved parameterizations.

We have now explained in more detail the preliminary step of the calibration procedure (based on PEST software, see section 4.3.1) which enabled us to reduce the original set of 25 to 7 tunable parameters. The coefficients showing a high correlation (i.e.  $|\text{corr}| > 0.9$ ) or the ones that preserved almost the original values after the PEST run, have been excluded from the second step, i.e. the manual calibration of the reduced set of parameters.

Details on the preliminary step of our calibration procedure have been added in section 4.3.1 and we modified the previous text as follows:

*“As a first step we adopted an automated calibration procedure, based on the PEST software (Doherty, 2002). This procedure minimizes an objective function, given by the sum of the mean squared differences between the modelled and observed river streamflow, using the Gauss-Marquardt-Levenberg non-linear least squares method. Several tests were carried out and we identified the most relevant parameters to be calibrated in our specific case study. The coefficients with a high correlation (i.e.  $|\text{corr}| > 0.9$ ) or the ones that preserved almost the original values after the PEST tests have been excluded. Thus we reduced the original set of 25 tunable parameters to 7 that are found to play a key role in the Ofanto basin. They are: the surface roughness scaling factor which controls the hydrograph shape and the timing of the peaks; the infiltration coefficient, the saturated hydraulic conductivity and the aquifer coefficients which control the total water volume.”*

Finally we added the following sentence in the Conclusions:

*“More research is still required on the groundwater modeling as it greatly impacts the overland waterflow and the river runoff but also the evapotranspiration and consequently the precipitation. We plan to evaluate different parameterizations of the aquifer recharge/discharge. Overall a reduction of the parameterizations involved in the WRF-Hydro system could be desirable”*

2. You have mentioned nothing about PET estimation, which may be an important process considering the duration of events. Have you discard its influence and why?. And considering the altitudes shown in the MDT figure and the mountainous characteristics of your catchment, I've also missed some comments about snow and melting processes.

**Authors:**

The NOAA-MP v2.7.1 model uses a modified Penman's relationship for the potential evapotranspiration, PET (Mahrt and Ek, 1984). The Penman's formula for the PET is widely used and we didn't perform any check on that.

As already detailed in section 3.1, we rather focused on the land use and topography datasets as we found that the low atmosphere and land surface fields including the evapotranspiration are strongly dependent on them, thus we replaced the default USGS data with the higher resolution and more recent data released by the European Environmental Agency.

Snow modeling is active in the NOAA-MP model starting from the ECMWF accumulated snow depth data. However, over the whole watershed, the highest point is at 1100m a.s.l. thus the snowfall and the melting processes might be of minor importance in the Ofanto catchment. We have now modified Fig.2 to make evident the height of the terrain in the catchment. The new picture is attached below for convenience.

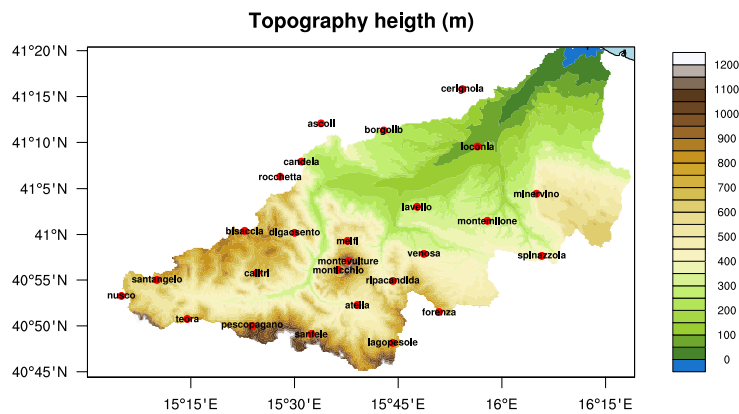


Figure 2. The Ofanto River Catchment. Top panel: Topography height (units of m) and location of 27 rain-gauge stations

We added more details in the manuscript :

*“The snow modeling is also active in NOAA-MP model: a multilayer snow pack, the snow albedo, the melting/refreezing capability are solved by NOAA-MP. Moreover the evaporation component coming from the snow sublimation is added and the evaporation component coming from the canopy water is split into the rainfall and the snowfall terms. The ECMWF analyses used for computing the initial and boundary conditions provide also the accumulated snow depth at the groundlevel. For our case studies, the snowfall and the melting processes do not seem to play a crucial role”.*

3. Are you using a continuous model or an event based one. It seems that the actual objective is not to represent a long time series, but a collation of convective events producing extremes.

**Authors:**

Our experimental design is based on a concatenation of 3-days hindcasting experiments. Details on the concatenation strategy and simulation period are provided below.

Literature proved that seasonal and sub-seasonal meteorological reconstructions benefit from a frequent reinitialization, i.e. few days (Qian et al., 2003; Koster et al., 2010; Lucas-Picher et al., 2013). On the other hand the hydrological-hydraulics reconstruction need a longer spin up period to distribute the overland and subsurface waterflow and to allow the river network to reach a steady state starting from dry at the initial state (Senatore et al., 2015), thus the WRF hydrological components are started from dry conditions and not restarted for the whole first month of concatenated meteorological forcing.

To make our strategy clear, we added few details at the end of section 3.1:

*“The hydraulics component of WRF-Hydro system is initialized with the NOAH MP overland and subsurface water flows that are dry at the initial time. Thus a spin-up period is required to laterally route the groundwater of the basin and to allow the river network to reach a steady state. Senatore et al. (2015) considered monthly spin-up for evaluating the WRF-Hydro results and we decided to follow the same strategy.”*

We have also added Koster et al. (2010) to the references about the WRF re-initialization strategy.

4. Would you say that long time response of aquifers is well represented in your model?. What are the physiographic characteristics of your basin in order to guess how is the aquifer response?. You also mentioned that the coefficients for the infiltration and saturated hydraulic conductivity are seasonally dependent. Don't you think that this is a problem of a lack of how your model works with water availability dependence of both infiltration and groundwater propagation laws or even more of the model capabilities but not of the parameters?. I mean, is there any structural problem in your model?. It seems that you've taken a practical approach, but not a science based one.

**Authors:**

In the presented study we focused on evaluating the model capability to reconstruct the precipitation and the river runoff. The aquifer representation is important in so far as it feeds the river baseflow and the numerical outcomes prove that the results are not extremely dependent on the aquifer specific parametrizations.

In general we followed a practical approach. The physiographic characteristics of the basin support our practical approach as we expect the soil types impact the capability of the aquifer to discharge to stored water (as already highlighted in section 4.3.1). Overall we agree that the empirical formula for the aquifer recharge/discharge and the high number of tunable coefficients affecting the groundwater are far from being satisfactory as we stressed at the end of section 4.3.1:

*“This study found that the soil infiltration and the aquifer water storage parametrizations should be seasonally dependent. This means that the present parameterizations of these processes are not capable to capture the complexity of the groundwater physical processes”.*

Additionally we have mentioned this problem in the Conclusions:

*“More research is required to establish a better groundwater modeling that at the moment considers seasonally dependent, ad hoc values of the soil infiltration and the aquifer water storage. We plan to evaluate different parameterizations of the aquifer recharge/discharge. Overall a reduction of the parameterizations involved in the WRF-Hydro system could be desirable”*

5. What is the main criteria to calibrate and validate rainfall and runoff? It seems that highest extremes had been used a reference of quality. If so, why not other kind of values considering that your objective is to represent long time series?

**Authors:**

For both precipitation and runoff, all the available observations covering the whole time windows of the experiments, including both extreme events and dry periods, have been used to calibrate the tunable coefficients and to validate the modeling results. We were interested both in the hydrograph baseflow and peak events which clearly are the most important for societal impact. Following the comments of the first reviewer we have modified Fig. 9 adding the experiments with and without the aquifer switched on since we were trying to simulate also the baseflow.

6. Furthermore, I would say that there are many more inconsistencies that make difficult the reading of the paper. Some others may be the following ones: a. Would you say that a reconstruction from winter 2011 to autumn 2013 is a long time series? b. How do you define your catchment?. Is it a small river catchment or a medium sized one?.

**Authors:**

We believe we have already answered these questions since:

a) we performed two seasonal experiments over winter 2011 and autumn 2013 and not a unique reconstruction from winter 2011 to autumn 2013. Thus our reconstructions produce relatively short-term timeseries. We hope our answer to point 3 has clarified the strategy adopted for the experimental design.

b) As we already stated in section 2, the Ofanto catchment is a medium size catchment (i.e. between 1000 and 10000 km<sup>2</sup>) because it is about 2790km<sup>2</sup>.

7. I would recommend reviewing the conclusion section too. What I found there is a review of some topics previously described as well as some evidence previously known as the key role played by aquifer discharge to affect the baseflow. Finally, I would recommend rewriting the whole paper (or papers) considering the interesting work developed by authors and the interesting topics they have assessed based on a global modelling approach.

**Authors:**

We reviewed the conclusions following the reviewer's suggestions and we entitled this section "*Summary, conclusions and future plans*" since it also includes an overview of the main modeling findings of the presented study. We list no more the aquifer influence on the river baseflow among our findings, as this is an expected outcome and the small storage capacity of the Ofanto aquifer makes the statistics on the aquifer influence unimportant.

The following sentences have been added aiming at making the purpose of this study more clear:

*"Overall we highlighted the 2-way feedback existing between a proper reconstruction of the meteorological events and the hydrological ones. A reliable description of the river hydrograph goes through a proper description of the meteorological and soil processes, with the precipitation field playing the most relevant role. At the same time the validation of the river hydrograph works as a effective post-processing tool to calibrate the water infiltration through the soil column and the aquifer recharge/discharge as well as to correct the modeled precipitation with the OA+LS method".*

*"More research is required to establish a better groundwater modeling that at the moment considers seasonally dependent, ad hoc values of the soil infiltration and the aquifer water storage. We plan to evaluate different parameterizations of the aquifer recharge/discharge. Overall a reduction of the parameterizations involved in the WRF-Hydro system could be desirable".*