Response to Anonymous Referee 3

REPLY TO GENERAL COMMENTS:

"Thank you for inviting me to review this paper. The paper describes an interesting approach that could be useful also in the context of engineering practice. Although it is not very clearly structured, I think it presents an interesting method that will contribute to the scientific community."

We appreciate Anonymous Reviewer 3 for his/her detailed evaluation and constructive comments. We particularly thank him/her for finding our approach interesting and useful for practitioners, which was a key objective of this work. We also recognize that, regarding the structural organization of the article, there is a lot of space for improvement, and, taking advantage of the valuable recommendations by all reviewers, we are planning to make significant changes in the revised document.

1. STRUCTURE OF THE MANUSCRIPT

"I was a bit confused by the overall structure of your manuscript. The distinction between introduction and your methods is not clear. I think that some parts of section 3 belong to introduction and others to methods. The title of section 4 is redundant and somehow misleading, since 4.1 and 4.2 refer to the CN method in general. In 4.2, you describe an approach for creating raster CN maps, the application of which goes beyond HRU delineation (as you mention in page 27 line 20). Thus, I think it should not be under the title: "CN approach for HRU delineation". I suggest that you merge the parts of sections 3 and 4 that refer to your proposed methods in a "Methods" section. The scope of research (if necessary) should come right after the introduction and before the methods, but anyway you need to rephrase it (for example, the paragraph in page 5, line 6-11 is redundant)."

"I think that the summary before your conclusion (section 7) is too extensive: you may consider rephrasing or even omitting the paragraph in page 27, lines 11-19. Your conclusions should be more laconic. I would suggest you have one paragraph for each conclusion. For example, you could split the paragraph starting at page 27, line 30 in two: one for your conclusion regarding subjectivity vs objectivity and one for the correspondence of HRU response with the CN values."

Thank you for all recommendations. In the revised document we will substantially re-organise the text, including removal of redundant material, to improve the readability of the paper.

2. MAIN POINTS OF RESEARCH

"assign as many parameters as can be supported by the available hydrological information"

In page 27, line 27-28 you recommend that the number of HRUs should equal the number of the available hydrological stations, but it is not clear how this conclusion emerges from your results. You should show what the results of the proposed CN approach would be if the method was applied with fewer or more HRUs.

In order to support the general principle that the number of HRUs should equal the number of the available hydrological records, in the revised version, we will performe a calibration experiment with one to five HRUs calibrations. In this vein, the CN parameter map of Nedontas river basin, initially containing 34 classes (Fig. 6, left), is used to delineate 1, 2, 3 (the number of HRUs proposed and used in the initial simulations), 4 and 5 HRUs. We remind that in HYDROGEIOS, 7 parameters are assigned to each HRU; thus, for each configuration the total number of control variables is 7*N*, where *N* is the number of HRUs. Therefore, the simplest parameterization, the lumped (homogenous) basin, comprises 7 parameters, while the scenario with 5 HRUs involves the estimation of 35 parameters. In all cases, the same objective function will be optimized, comprising efficiency metrics at the 3 monitoring stations, and other terms that allow providing realistic responses for the remaining processes.

It is important to note that in this computational experiment we will not allow manual interventions within calibration, as done in the context of the scenarios examined in the initially submitted manuscript. We will simply run the global optimization algorithm assuming the full set of parameters (7, 14, 21, 28 or 35) within their physical bounds. In order to take into account the increasing computational burden of optimization against the number of control variables, we set a maximum budget/limit of 2000*N* trials, where *N* is the number of HRUs. Thus, for the lumped approach we allowed up to 2000 function evaluations in each calibration scenario, while for the more complex configuration ($N = 5$) the budget/limit was increased to 10000 function evaluations.

To ensure unbiased results as much as possible, all calibrations will be carried out by considering the same initial conditions and by assigning the same parameter values for the other model components (groundwater conductivities, sub-basin routing rates). Moreover, to avoid getting trapped in a local optimum, thus resulting in sub-optimal model performance, we will employ several independent optimizations for each parameterization, and kept the best solution.

As will be demonstrated in the revised article, the above experiment will further confirm our fundamental hypothesis that the "optimal" balance, in terms of model performance against computational effort, is ensured by considering the parameterization with three HRUs, which equals the number of available hydrographs.

"reduce the effort for model calibration, simultaneously ensuring good predictive capacity"

Since time and effort are important for your study (page 25 line 1-2), I think you need to be more specific regarding Table 3 (especially because the union of layers and the CN approaches do not differ a lot).

Actually, due to the hybrid calibration approach, involving subsequent optimizations of sub-sets of parameters, changes of parameter bounds as well as other manual interventions, it is very difficult, if not impossible, to quantify the effort made for the parameter estimation procedure. Thus, instead of a quantitative measure, we compare the effort expended with the proposed (i.e. CN-based) strategy against the other two approaches. Regarding the comparison with the "union of layers" approach, in terms of computational effort the differences are actually limited. This is not surprising, since the two approaches differ by only one HRU, and thus by only 7 parameters. However, as shown in Table 4, the CN-approach clearly outperforms the "union of layers" approach, in terms of model performance in calibration and validation, particularly in the two upstream stations (Karveliotis, Alagonia).

Is the efficiency of page 20, line 15 the same with the one that you mention in page 14 line 13? I suggest you elaborate more on that.

In p. 14, the term "efficiency" is mentioned in line 9 (not line 13), and is the same metric with the one mentioned in p. 20, line 15.

In page 20, line 21-22 you refer to the terms unsatisfactory and acceptable. How do you define those and what are the thresholds?

It is well-known that in the hydrological literature there is a continuous discussion and many different opinions about performance metrics and evaluation techniques. According to the common practice, a rainfall-runoff model is generally judged as satisfactory if exhibiting efficiency values greater than 50% (e.g., Moriasi *et al*., 2007), while a value of 30% is considered as the limit to accept that the model can be representative of a basin's dynamics (e.g., Freer *et al*., 1996). In particular, Freer *et al.* (1996) have used this limit as threshold for distinguishing behavioral solutions within the GLUE method. In our context, we consider as "satisfactory" efficiency values above 75-80%, while efficiency values lower than 30% are characterized "unsatisfactory". This will be clarified in the revised text.

Looking Fig. 8 I do not think that your proposed method's results are really improved. The second last peak before the summer is not caught by any model version. This peak also does not appear in the other hydrometric stations. Could it be a measurement issue?

According to the results given in Table 4, the model performance was improved, as the efficiency and high flow efficiency values achieved by the simulation with the CN-based delineation were the highest for both calibration and validation. Although this increase is not "impressive", in terms of computational effort and time the model performance was significantly improved, and that was one of the main objectives of the proposed method. Another issue of improvement, which will be better highlighted in the revised version, is the better correspondence of parameter values to the physical characteristics of the HRUs. By virtue of the CN concept, the user can define a priori suitable parameter bounds, thus facilitating the calibration procedure and also ensuring physically consistent parameters.

Regarding your remark about the peak flow, this actually may be due to a local flood event, affecting a part of the basin where there was not any available rainfall information.

3. CASE STUDY

Since you try "to better represent the physiographic properties of the river basin", you should take into account that the basin was burned in 2007 (http://ocean.space.noa.gr/diachronic_bsm/, you may also check the 2011 google earth street view images of this area). Thus, the map in Fig. 5b (derived from 2000 land cover data) does not correspond to your simulation period. Did you apply the correction you mention in page 8, lines 27-28? It is not clear if this is your recommendation: in that case you need to explain it, otherwise you should provide a citation. Anyway, I believe that this approach is unsuitable for this case study. This is particularly important since evapotranspiration accounts for the 1/3 of your hydrological budget.

Thank you for this important remark. Actually, the classification of Nedontas river basin in terms of vegetation index, iVEG, was based both on Corine Land Cover 2000, as well as on a more recent map provided by the National Observatory of Athens, Institute for Astronomy, Astrophysics, Space Applications & Remote Sensing (C. Contoes, personal communication), which accounts for the burnt areas of the basin representing the vegetation state that corresponds to the simulation period. As explained in p. 8, lines 27-28, in our methodology we recommend that burned areas should be classified under one category with respect to their original condition, and this was also done for the Nedontas case. This important issue will be clarified in the revised version.

4. FIGURES

For Fig. 1 (and 2 maybe) please provide a color legend. I was a bit confused by Fig. 1c and 2a, in that 88 appears in red and in yellow and 64 and 49 are both green. I guess color expresses the CN classes, so probably these are just typos.

Yes, color expresses CN classes, and the CN classes you mentioned appearing in different colors should appear in the same color: 88 both in red, while 64 and 49 should have had different shades of green. Obviously, in the revised version, we will apply correct colors.

In Fig. 3 the Evapotranspiration from the Upper zone storage is denoted as Evapotranspiration from lower zone.

Thank you for the remark. The typo error will be corrected.

In Fig. 5 please enlarge the legend.

Thank you for the recommendation, which will be implemented in the revised version.

In the text, you cite Fig. 7 before citing Fig. 6.

We apologize for the inconvenience. In the revised text figures will be cited in increasing order.

In Fig. 6a, the legend shows 18 classes, and so is written in page 17, line 21. However, in the caption of Fig. 6a it is 34 classes.

Apologies: the caption is incorrect. It should have been: Figure 6: CN parameter map of the Nedontas river basin (a) initial map of 18 CN classes; (b) condensed map of three CN classes, using CN = 60 and 5 CN = 70 as thresholds. The typo error will be corrected in the revised article.

5. A FEW SUGGESTIONS

Page 7 line 10: I feel here is a good point to explain the meaning of high vs low CN values in this section rather than in page 25, lines 11-14.

Thank you for this useful recommendation. The aforementioned part will be moved in the description of the CN method.

Page 7 lines 24-25: Reference?

There are numerous research articles investigating the runoff dynamics in steep landscapes. As example, we will cite the work by Montgomery and Dietrich (2002), also containing an extended list of articles dealing with the above topic.

Page 8 line 1: What are the higher values that would be otherwise inferred?

The standard SCS approach provides too extended range of CN values in case of soils falling in hydrological soil group A. For instance, in forested areas CN ranges from 25 to 45, while for pasture areas the recommended range is even larger, from 30 to 68. Most practitioners prefer using either the average or the highest values (for safety), which are however not representative for the case of very permeable sub-surfaces (e.g., karst). The proposed split of HSG A into two categories allows better definition of the CN parameter across soils of low and very low permeability.

Page 11 lines 6-8: Again you use passive voice, thus making it confusing: who recommends? If it is you, then explain why, otherwise provide citations.

The recommendation is made by the authors, based on the information and bibliography mentioned in the previous lines, i.e. the classical articles by Jakeman and Hornberger (1993) and Wagener *et al.* (2001). In the revised paper, to avoid confusion, we will use “we propose” instead of “it is proposed”.

Page 11 line 8: You may comment the quality of your data later in the manuscript.

Details about the monitoring system and the data processing and management procedures, also including quality tests, are provided in the cited work by Efstratiadis *et al.* (2013).

Page 14 line 31: Which station do those temperatures refer to?

The meteorological information for the broader area was taken from the study by Koutsoyiannis *et al.* (2008), which will be included in the revised version.

Page 20 line 26-onwards: You should cite Table 4 when mentioning the efficiencies.

Citations of Table 4 are added in text.

6. TECHNICAL COMMENTS/SUGGESTIONS:

Thank you for your comments/suggestions; we will make the necessary corrections in the revised version of the manuscript.

Page 5 line 23: do you mean "compared the" instead of "compared to"?

Yes, it should be "compared the".

Page 5 line 24: "area" instead of "areas".

Will be corrected in the revised text.

Page 6 line 10: "increases" instead of "increase".

Will be corrected in the revised text.

Page 7 line 32: "in" instead of "is".

Will be corrected in the revised text.

Page 8 line 22 & 32: you may cite Table 1. You may also include the ranking in the supplementary Tables.

The citation is added in text and the ranking is now included in the supplementary tables. Please find the update supplement tables attached.

Page 9 line 3-4: Consider rephrasing, maybe: According to the above classifications, the dominant classes of permeability, land use/cover and drainage capacity, as well as the corresponding indices iperm, iveg, and islope (ranging from 1-5) are assigned for a given area.

To be corrected in the revised text.

Page 11 line 5: This makes it essential to. . .

Will be corrected in the revised text.

Page 12 line 8: replace "are" by "is".

Will corrected in the revised text.

Page 14 line 12-13: Consider rephrasing: In the last version of HYDROGEIOS, a modified efficiency index has also been introduced to account for. . .

Will be corrected in the revised text.

Page 19 line 13: Omit "of" between "most" and "parameters".

Will be corrected in the revised text.

Page 19 line 24: I think yu mean "combining" instead of "combing".

To be corrected in the revised text.

Page 27 line 28-29: It is not very clear, do you mean: "and also makes it possible to take advantage..."?

Will be corrected in the revised text.

Page 27 line 33: "thus" instead of "this".

Will be corrected in the revised text.

Page 28 line 6: "CN" instead of "CN's".

Will be corrected in the revised text.

Page 28 line 14: parameterisation: be consistent with your spelling.

Will be corrected in the revised text.

Page 29 line 13: physically-based.

Will be corrected in the revised text.

References

Efstratiadis, A., Koussis, A. D., Lykoudis, S., Koukouvinos, A. Christofides, A., Karavokiros, G., Kappos, N., Mamassis, N. and Koutsoyiannis, D.: Hydrometeorological network for flood monitoring and modeling, in *Proceedings of First International Conference on Remote Sensing and Geoinformation of Environment*, Society of Photo-Optical Instrumentation Engineers (SPIE), 8795, 10-1–10-10, Paphos, Cyprus, 2013.

Freer, J. E., Beven, K. J., and Ambroise, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Resour. Res.*, 32(7), 2161-2173, 1996.

Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model?, *Water Resour. Res.*, 29(8), 2637–2649, doi:10.1029/93WR00877, 1993.

Koutsoyiannis, D., Andreadakis, A., Mavrodimou, R., Christofides, A., Mamassis, N., Efstratiadis, A., Koukouvinos, A., Karavokiros, G., Kozanis, S., Mamais, D., and Noutsopoulos, K.: *National Programme for the Management and Protection of Water Resources*, 748 pages, Department of Water Resources and Environmental Engineering – National Technical University of Athens, Athens, February 2008.

Montgomery, D. R., and Dietrich, W. E.: Runoff generation in a steep, soil-mantled landscape, *Water Resour. Res.*, 38(9), 1168, doi:10.1029/2001WR000822, 2002.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *Transactions of the ASABE*, 50, 885-900, doi:10.13031/2013.23153, 2007.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. and Sorooshian, S.: A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, 5(1), 13–26, doi:10.5194/hess-5-13-2001, 2001.