Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-118-AC2, 2018

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Interactive comment on "Discharge hydrograph estimation at upstream-ungauged sections by coupling a Bayesian methodology and a 2D GPU Shallow Water model" *by* A. Ferrari et al.

A. Ferrari et al.

alessia.ferrari@unipr.it

Received and published: 29 June 2018

The authors gratefully acknowledge the positive and constructive review of the anonymous Referee. In this document the comments provided by the Referee are reported in italic, whereas the authors' response and indications about the original paper modifications are marked in bold fonts.

General comments
The manuscript applies a Bayesian geostatistical methodology to the solution of the inverse problem aiming to estimate the upstream flood hydrograph at an un-gauged

river section. The downstream routing of the hydrograph is pursued by means of a 2D shallow water model. This leads to a computationally intensive problem, for which a parallel implementation is designed. The most computationally intensive operation (i.e.: the evaluation of the Jacobian matrix) is demanded to a multi-GPU HPC, and also the forward model exploits the opportunities of GPU-parallelization.
The adoption of two-dimensional hydraulic model represents a step forward compared with both the previous research developed by the Authors and with the state-of-the-art. The resulting complication arising from the increased computational effort is handled properly. Therefore, the research described in the paper appears to be sufficiently innovative, well-designed and of interest to the readers of HESS.
I am rather supportive of the publication of the manuscript, provided that the Authors put some additional effort in improving the quality of the presentation (especially of the English) and in addressing some issues in order to make their outcomes more conclusive. I provide in the following few specific comments to be considered in the revision, as well as some minor issues that could contribute to improve the quality of the manuscript.

The authors wish to thank the anonymous Referee for his positive overview about the manuscript.

Specific comments

- *I appreciate that the presentation of the Bayesian Geostatistical Approach (BGA) is concise but complete of every detail: however I found it not very clear at some points, detailed below:*

  1. *The "prior mean" defined in eq. (9) should be better commented, explaining why the vector* β *reduces to "a single value" (do the Authors mean the same*

*value for each parameter?), and why the matrix **X** reduces to "a single vector of ones".*

We appreciate this comment and we agree with the Referee that more information about the prior mean is needed to facilitate the readers in figuring out the Bayesian Geostatistical Approach philosophy.
As a result, in the revised version of the paper, we will reword the involved paragraph commenting the terms that form the prior mean as follows: "The prior mean is defined as *E*$[\mathbf{s}] = \mathbf{X}\beta$ where *E* is the expected value, β is the vector of drift coefficients, and X is a known matrix of basis functions.
In our case β is a single unknown scalar, but different drift coefficients can be used to introduce discontinuities in the stochastic function to be estimated (e.g. when the unknown parameters are likely to form distinct populations). For example, in the context of reverse flow routing problems, multiple values of β are adopted if more than one inflow hydrograph must be estimated at the same time (e.g. the inflow on both the upstream branches of a river confluence). The matrix of basis function, X, links each unknown parameter with the corresponding element of β and, at the same time, specifies the model of the mean (e.g. constant mean, mean with a trend, etc.); in our case the mean is constant and therefore X is a single vector of ones (Fienen et al. (2008))."

2. *The separation distance **d** should be defined explicitly.*

   We really appreciate this comment and we acknowledge the potential confusion that arises from the use of the term separation distance.

This is a legacy from the fact that geostatistics is manly used in estimating spatial parameter fields rather than time functions. We will add in the revised version of the paper that d represents the vector of the separation times between all the parameter pairs: $\mathbf{d}_{i,j} = \mathbf{t}_i - \mathbf{t}_j$ with *i,j*=1,...,N_p, where *t* represents the time associated with each parameter and N_p the total number of unknowns.

3. *I wonder about the opportunity of defining Q_{ss} as $Q_{ss}(\theta)$ since the r.h.s. of eq. (6) does not contain θ.*

   We acknowledge the mistake in the original version of our manuscript. The prior covariance matrix in Eq. (6) is not influenced by the slope parameter θ but by the variance σ_s^2 and the integral scale l; we will correct $\mathbf{Q}_{ss}(\theta)$ as $\mathbf{Q}_{ss}(\sigma_s^2, l)$.

4. *I could not find the definition of ξ appearing in eq. (9) and eq. (13).*

   We agree with the Referee and the definition of ξ, which was missing in the original paper, will be included in the revised text. We will add: "In case a linear relationship between parameters and observations (linear forward model) holds, a computationally efficient method to find the best estimate $\hat{\mathbf{s}}$ of vector s (and $\hat{\beta}$ of β) is obtained introducing the vector $\xi = (\mathbf{H}\mathbf{Q}_{ss}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{X}\hat{\beta})$ and solving the following linear system of equations (Fienen et al. (2009)):"

5. *The Authors should better explain what they mean with "a flat solution".*

   We thank the Referee for this comment and we agree that the term

"flat" should be better explained. For this reason, we completely reworded the sentence making clear what we mean with "a flat solution". In the revised paper the involved sentence will be modified as follows: "The starting values for the structural parameters are assigned so that the variability between contiguous parameters is small (flat solution, with a high degree of correlation); complexity is then introduced during the optimization process if supported by the data. The variance of the epistemic errors is assumed close to the expected one."

- *In the scheme depicting the BGA in figure 3, I could not find the condition corresponding to the parameters convergence, which is claimed in the text. According to the scheme, the inner cycle terminates only when the maximum number of iterations N_i is reached. The Authors should clarify this point and modify accordingly the manuscript and/or the figure. Assuming that also convergence causes termination, the Authors should explain how did they check the convergence.*

  The Referee is right. We confirm that both the inner loop to estimate the model parameters and the outer one to estimate the structural parameters iterate until convergence or the assumed maximum number of iterations is reached. Therefore, the 2^{nd} (inner > N_i) and 3^{rd} (outer > N_o) decision blocks in Fig. 3 of the manuscript do not only check if the maximum number of iterations is reached, but also verify if convergence is achieved. The flow chart will be corrected in the revised paper. Additionally, we will include the definition about convergence. Recalling that the aim of the inverse procedure is to obtain the vector of the unknown parameters s, as well as to quantify the uncertainty in the estimation, the solution is found maximizing

  the posterior pdf or, more conveniently, minimizing its negative logarithm (objective function) (Fienen et al., 2013). The linearization process ends if the maximum number of iterations N_i is reached or if the improvement in the objective function (absolute difference between two successive iterations) is below a user defined value. The structural parameter iteration loop (outer loop) progresses until the maximum number of iterations N_o is reached or the norm of the differences between structural parameter values at consecutive iterations is below a user defined value (Fienen et al., 2013).

- *The Authors should explain how the credibility intervals may be evaluated based on the results of the BGA algorithm, or at least provide a reference to previous literature.*

  We really appreciate this suggestion and accordingly we will modify the involved paragraph in the revised manuscript. We clarify that, at the end of the parameter estimation, the linearized uncertainties of the unknowns can be evaluated in terms of the posterior covariance matrix of the estimated parameters, (Fienen at al., 2013). In addition, we will explain that the diagonal elements of this matrix represent the posterior variance (σ^2) of the estimated parameters and that the 95$\%$ credibility interval of the solution is approximately equal to $\pm 2\sigma^2$.

- *About the core of the research described in the manuscript, I am mostly concerned about three issues. They should hopefully be addressed in the revised version of the manuscript.*

  1. *Since the principal innovation comes from the adoption of a 2D forward*

hydraulic model, the improvement in terms of the quality of the estimated hydrograph deriving from the use of a more detailed (but also demanding) schematization of the hydraulic process should be explicitly assessed. For instance, how wrong is the estimated hydrograph if one uses a 1D model as the forward routing model in one of the presented examples?

We really thank the Referee for this comment since it allows us to discuss the motivations that led to enhance the serial Bayesian procedure introduced by D'Oria and Tanda (2012) for 1D cases, to 2D forward models. The choice between 1D and 2D models concerns the classical forward propagation rather than the Bayesian application. In fact, in literature the advantages of 2D-SWEs in comparison with 1D schematizations have been thoroughly discussed (e.g. Costabile et al., 2015), assessing that if river reaches present several floodable areas, meanders and floodplains, as it is typical for lowland streams, only 2D models can properly describe the flood propagation. As shown for example in Fig. 7 of the manuscript, in such rivers the low flow at the beginning of the event follows the meanders and water is contained in the main channel, whereas for high discharge the flow involves the river banks and a continuous mass and momentum exchange occurs between the main channel and the river banks and thus the assumptions of 1D models do not hold.
Therefore in our opinion, since the physical phenomena can be only accurately simulated by a 2D numerical scheme, no accurate upstream discharge hydrograph can be obtained by adopting 1D models. Finally, coupling the Bayesian approach with a fast, stable and accurate 2D forward model is the first step for reconstructing the discharge hydrograph during a levee failure and/or overtopping that causes the flooding of the nearest lowlands; the authors are also

working in this direction that clearly requires the adoption of a 2D model.

2. *Could the Authors discuss (hopefully with the aid of some additional results) the effects of the resolution of the DEM and/or of the values of the roughness parameters on the estimated hydrograph?*

We thank the Referee for this useful comment that allows us to clarify some further aspects of the forward numerical modelling.
The mesh design is an issue related to create an accurate forward model. As for every numerical method that aims at describing a physical phenomenon in a spatial domain, the mesh must be chosen considering both the needed accuracy and the required computational effort. Firstly, the mesh must be defined in such a way that the bathymetry of the rivers is adequately resolved. Figure 1 shows that the adopted mesh (with $\triangle$x=10 m inside the river) is able to accurately reproduce the river geometry. Secondly, the grid size must guarantee that the numerical solution is close to the "exact solution" of the SWEs. Convergence analysis can be proficiently performed for simple test cases, in which the mesh can be progressively halved many times with a reasonable computational effort. A similar analysis was done in a previous study conducted by some of the present authors (Aureli et al., 2008) and it is beyond the scope of this work. Anyhow, grid size, roughness estimation and numerical discretization of SWEs, all play an interlaced role on the solution results. First order accurate models, for example, intrinsically introduce more dissipation into the solution and this behaviour must be counterbalanced during the calibration phase, for example reducing the dissipation term due to friction, since a part of the dissipation is already embedded in the intrinsic numerical viscos-

ity of the model. Despite the calibration of the considered river (grid size, roughness and numerical discretization) was already assessed in previous studies (Vacondio et al., 2016), according to the Referee' suggestion we performed an additional inverse Bayesian estimation with a different roughness coefficient (please refer also to the next comment answer). Particularly, the Manning coefficient originally set equal to 0.05 s/m$^{1/3}$ was decreased by 15$\%$ and assumed equal to 0.0425 s/m$^{1/3}$, as for example can happen due to seasonal changes in vegetation. As shown in Figure 2, the estimated flood waves are similar and the highest difference, which is in correspondence with the second peak, is less than 6$\%$. Therefore, the influence of assuming a "wrong" roughness coefficient is less than linear in the discharge estimation. However, we want to stress that the same issue holds for any model setup. The revised version of the paper will include this analysis concerning the roughness values.

3. *I understand the role of the simulations based on synthetic data-sets, with or without accounting for measure corruption in the validation of the procedure. On the other hand, as far as the "real field application" is concerned, I think that a different test case should have been considered, namely one for which the measured hydrograph was available, in order to compare the estimated with the actual one. This not being the case, the evaluation of the procedure performance cannot go further than the "credibility" (in a statistical sense), and the claims by the Authors in the comments ("This real field application further confirms the capability of the proposed inverse procedure of estimating irregular inflow hydrographs in real rivers") may sound excessive and not fully supported. Could the Authors take into consideration the addition of such an example?*

We really appreciate this comment and the suggestion pointed out by

the Referee. As claimed in the paper, the key point of the proposed procedure is to define a discharge hydrograph in an upstream river section that has no records (either in terms of water levels or discharges). However, to validate the methodology for a real field application, a measured reliable discharge hydrograph has to be available in the upstream section of the model. Usually, discharge hydrographs in natural rivers are obtained by converting registered water levels with an appropriate rating curve. Nevertheless, this procedure is affected by two different major uncertainties: (i) rating curves are usually calibrated only for small discharges (ii) the inertial terms of the SWEs, which cause a non-unique level-discharge relationship during floods, are neglected, despite the fact that they are not negligible in lowland rivers. Since in the real field application presented in Sect. 5 the upstream section of the river is located immediately downstream a flood control reservoir dam, the discharge hydrograph has been obtained by adopting the classical hydraulic theory of sluice gates and spillways, partially overcoming the previous issues.

With this aim, we recovered the dam geometrical data from the competent Authority (i.e. number and dimension of the bottom openings, crest length of the spillway, etc.). Please note that during flood events large wood debris accumulates in the reservoir reducing the outflow discharge from the bottom openings, especially during the depletion phase, and disturbing the overflow spillway (see Figure 3). Due to this issue and the uncertainty in evaluating the discharge coefficients, the calculated flood wave showed in Figure 2 has an envelope of different solutions obtained adopting equally likely coefficients.

As depicted in Figure 2, the estimated flood wave is in good agreement with the measured one; the main differences are after the highest peak, which is well reproduced, where the "observed" wave presents two

small peaks, whereas the inverse methodology provides a smoother solution. Despite all the involved approximations, this comparison confirms that the proposed inverse procedure is capable of estimating inflow hydrographs with multiple peaks and irregular shapes in real rivers. In addition to the flow hydrograph estimated with the Manning roughness coefficient equal to 0.05 s/m$^{1/3}$, Figure 2 reports the inverse estimation with a different roughness coefficient equal to 0.0425 s/m$^{1/3}$ as highlighted in the answer to the second Referee issue. Both these analyses will be included in the revised version of the paper.

- *English should be carefully revised throughout the entire manuscript to match the standards of scientific communication.*

  We thank the Reviewer for his suggestion. The entire manuscript will be carefully revised. Moreover, the language corrections kindly provided by Dr. A. D. Koussis (first Referee) will be considered in the revised manuscript.

technical corrections

- *Please refer to eq. (5) and (6) as to linear or Gaussian variogram, just the way you did in section 4.2*

  We thank the Reviewer for his technical corrections that we will include in the revised manuscript.

- *Probably in r.h.s. of eq. (14) a "+" sign is missing. Please check.*

The Referee is right: we will correct this in the revised paper.

- *Throughout the manuscript, "non linear" should better read "non-linear"*

  We thank the Reviewer for this suggestion: the correction will be included in the revised manuscript.

- *Please note that actually the r.h.s. of eq. (12) is not a fraction, therefore referring to "denominator of Eq. (12)" makes sense if you are considering the discrete approximation of the Jacobian.*

  We totally agree with the Referee and, as consequence, we will reformulate the involved paragraph in the revised version of the manuscript. In Sect. 2, where the theory of the Bayesian approach is described, Eq. (12) defines the Jacobian matrix formulation, which is not a fraction but a partial derivative. Therefore, in Sect. 3 of the revised paper we will properly refer to Eq. (12) as the formula to calculate the sensitivity matrix components. Moreover, we will clarify that the Jacobian matrix is approximated according to a finite difference scheme, and hence each element is evaluated as the ratio between the variation of each observation (numerator) for given variation of each parameter (denominator).

- *The description of fig. 6 and the figure itself refer to four cross-sections along the river: an upstream un-gauged one (A), two intermediate (B and C) where water levels are measured, and a fourth one (D) for downstream boundary condition assignment. However, in the presented examples, only a single intermediate measuring cross section is used, so maybe the description and the figure should be consistently simplified.*

We really thank the Referee for this comment and we acknowledge that the role of the section D was not clear in the manuscript. However, the presence of section D plays a specific role in setting up the synthetic case to use as benchmark for the inverse procedure. We will make it clear in the revised version of the paper adding that: "The information in section D is only preparatory for setting up the synthetic cases and it is not used in the inverse procedure. Imposing a rating curve in D allows to obtain water levels with a non-unique stage-discharge relationship in section C, which is more close to the real circumstances when applying the inverse procedure."

References

Aureli, F., Maranzoni, A., Mignosa, P., and Ziveri, C. A weighted surface-depth gradient method for the numerical integration of the 2D shallow water equations with topography. Advances in Water Resources, 31(7), 962-974, 2008.

Costabile, P., Macchione, F., Natale, L., and Petaccia, G. Flood mapping using LIDAR DEM. Limitations of the 1-D modeling highlighted by the 2-D approach. Natural Hazards, 77(1), 181-204, 2015.

D′Oria, M. and Tanda, M. G.: Reverse flow routing in open channels: A Bayesian Geostatistical Approach, Journal of hydrology, 460, 130–135, 2012.

Fienen, M. N., Clemo, T., and Kitanidis, P. K.: An interactive Bayesian geostatistical inverse protocol for hydraulic tomography, Water Resources Research, 44, 2008.

Fienen, M., Hunt, R., Krabbenhoft, D., and Clemo, T.: Obtaining parsimo-

nious hydraulic conductivity fields using head and transport observations: A Bayesian geostatistical parameter estimation approach, Water resources research, 45, 2009.

Fienen, M. N., D′Oria, M., Doherty, J. E., and Hunt, R. J.: Approaches in highly parameterized inversion: bgaPEST, a Bayesian geostatistical approach implementation with PEST: documentation and instructions, Tech. rep., US Geological Survey, 2013.

Vacondio, R., Aureli, F., Ferrari, A., Mignosa, P., Dal Palù, A. Simulation of the January 2014 flood on the Secchia River using a fast and high-resolution 2D parallel shallow-water numerical scheme. Natural Hazards, 80(1), 103-125, 2016.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-118, 2018.

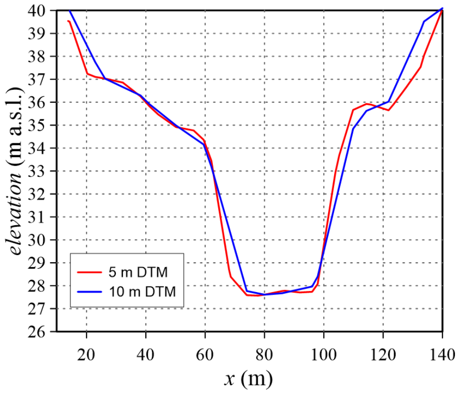

Fig. 1. Ponte Alto section on the Secchia River: comparison between the sections extracted from a 5 m and 10 m resolution DTM.

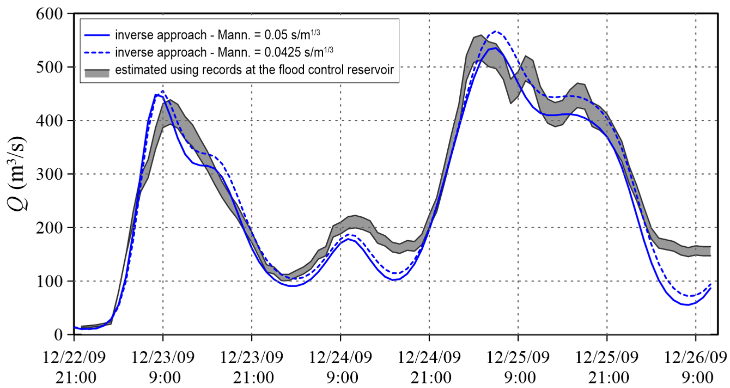

Fig. 2. Secchia 2009 event: comparison among the inflow hydrographs assuming two different Manning coefficients, and the one calculated using the records at the flood control reservoir.

Fig. 3. Large wood debris accumulation on the Secchia flood control reservoir during a flood event close to the bottom openings (left) and on the overflow spillway (right).