

Interactive comment on “Stochastic reconstruction of spatio-temporal rainfall pattern by inverse hydrologic modelling” by Jens Grundmann et al.

Jens Grundmann et al.

jens.grundmann@tu-dresden.de

Received and published: 24 September 2018

article [utf8]inputenc geometry verbose,tmargin=3cm,bmargin=3cm,lmargin=3cm color
textcomp

Printer-friendly version

Discussion paper



Authors final comments to referee 1

September 24, 2018

Journal: HESS

Title: Stochastic reconstruction of spatio-temporal rainfall pattern by inverse hydrologic modelling

Author(s): Jens Grundmann, Sebastian Hörning, András Bárdossy

MS No.: hess-2018-350

We would like to thank the referee for his/her time to review the manuscript. Our reply is organized as follows: (1) comments from Referee are in black color, (2) author's response is marked in blue color and placed within referee comments whenever it's needed, (3) author's changes in manuscript based on comments of both referees are summarized at the end of this document.

REFeree 1:

My comments are in the order I read the paper:

C2

[Printer-friendly version](#)

[Discussion paper](#)



p213 - spelling mistake - 'generall'

Thanks, the spelling mistakes will be fixed in the revised manuscript.

p313 - some grammar issue - please check

Grammar will be improved throughout the whole paper. In general spelling and grammar will be double checked by a native speaker prior to resubmission.

p6 - in my understanding, the approach is roughly along the following lines - first transformed empirical CDFs are ascertained and used to create equivalent Normal observed rainfalls denoted as w .

We are not quite sure what you mean by transformed empirical cdfs are ascertained. What is done is that a cdf (and in general this cdf can be any type of cdf, i.e. parametric or non-parametric or a combination of both) is fitted to the observed precipitation values. The distribution used in this work is described in Eq. 3. It is a combination of a discrete probability for zero precipitation values and an exponential distribution for values greater than zero. Thus the parameters that need to be estimated are p_0 (the discrete probability of zero precipitation) and λ (the parameter of the exponential distribution). Subsequently, using this fitted cdf the observed precipitation values are transformed to standard normal values according to Eq. 4.

line 11 says Gaussian copula is fitted to describe spatio-temporal dependence structure - a few lines of what this entails should be provided for completeness

We are going to add more information on copulas in general, the Gaussian copula and the fitting process to the revised manuscript.

- in my reading this step seems independent of what is described next in the paper but

[Printer-friendly version](#)

[Discussion paper](#)



if I am wrong this should be corrected. I am guessing the $w_{j,i}$ is not from this copula but from equation 4 for each site and time step. So you have L sequences and the aim is to find α_l such that there is some minimal deviation with the transformed normal rainfall at each location and time step. So I guess the idea here is to keep generating fields until they match the observed rainfalls transformed to Normal. If that happens then you will have $L = J$ and all the alpha's being equal to $1/L$. And since there is spatial dependence, you would kind of expect $L < J$ if this works fine. Am I correct? May be good to spell this out a bit more.

Yes, to some extent but there also seems to be some misunderstanding. The $w_{j,i}$ -s (which represent the transformed precipitation observations at locations x and time steps t) are derived by Eq. 4 (see your third comment). The random fields V_l (which are independent standard normal spatial random fields) are simulated such that they all have the same spatial structure which is described by this Gaussian copula (this information is missing in the paper). Eq. 6 says that we want to find a linear combination of these independent standard normal random fields V_l such that this linear combination results in the values $w_{j,i}$ at locations x and time steps t . Thus Eq. 6 describes a linear equation system with the weights α_l being the unknowns (the values $V_l(x_j, t_i)$ are known). This equation system can be solved for $L \geq J$, the bigger L is the smaller the $\sum \alpha_l^2$ sum gets - if the solution is calculated using SVD.

P6123 - what are homogenous conditions? I didn't understand what is meant by $U_k(x_j, t_i) = 0$ - please clarify what this is and why is it needed.

The homogeneous conditions are $U_k(x_j, t_i) = 0$ (a system of linear equations with the right hand side being all zeros is called a homogeneous equation system). This means that now we want to find a linear combination of random fields which fulfills $U_k(x_j, t_i) = 0$, i.e. a linear combination that results in zeros at locations x and time steps t . This is done the same way as constructing the field W^* , i.e. by setting up an equation system using independent standard normal random fields $\sum_{k=1}^K \beta_k V_k(x_j, t_i) = 0$ (we didn't put this equation in the paper as it's basically the

Printer-friendly version

Discussion paper



same as Eq. 6 with the right-hand side being zero). The explanation why this is needed is actually given in the following sentences. Line 24: 'The advantage of these fields U_k is that they form a vector space (they are closed for multiplication and addition)...' This means when adding such a field U_k (or k of them) to W^* , the resulting field W_λ will exhibit the correct values $w_{j,i}$ at locations x and times t because the zeros in the field U_k do not affect these values. However, the rest of the field is affected (as fields U_k are conditional random fields) which enables modifying the final field W_λ without changing the conditioning values. By changing the arbitrary weights λ one can modify the field such that it represents the observed runoff (and therefore the procedure needs to be coupled with the rainfall runoff model) to a certain degree.

p712 - this is starting to become confusing now. Where did the covariance matrix come from? covariance of what?

This goes back to the missing information that the fields V_i all have the same spatial structure which is described by the fitted Gaussian copula. The covariance we are referring to is the spatio-temporal covariance of the observations to which we have fitted the Gaussian copula. We are going to change the wording (as it isn't consistent) and add more information to the revised manuscript. As the field that fulfills the homogeneous conditions can be combined using arbitrary weights λ the scaling factor $k(\lambda)$ can be used to scale the final field such that the resulting field exhibits the spatio-temporal correlation/covariance of that copula. This is the case when the L^2 norm of the weights of the linear combination is equal to 1. As $\sum \alpha_i^2 \ll 1$ the weights λ need to be scaled (using the scaling factor $k(\lambda)$) such that $\sum \alpha_i^2 + \sum \lambda_k^2 = 1$. It's also worth mentioning that in this case the covariance is equal to the correlation as we are working in standard normal space (mean is zero and unit variance).

if W^* represents more or less the transformed observed precipitation field (from what I could gather), is this W_λ some randomised representation of that? If you are adding

Printer-friendly version

Discussion paper



positive random values to this, aren't you changing the probability distribution of W_λ from uniform to something shifted/tending to Gaussian?

W^* is already a random representation of a precipitation field that is conditioned on the available point observations. Due to the additional constraint $\sum \alpha_i^2 \ll 1$ it however is a very smooth field (like an interpolated field), i.e. it does not represent the observed spatial variability of the precipitation. By adding the fields U_1, \dots, U_k to W^* one can easily scale these fields to have the correct spatial dependence without modifying the observed values at the observation locations (because the weights λ are arbitrary as the fields U_1, \dots, U_k fulfill the homogeneous conditions) such that the final field W_λ exhibits the correct spatial variability. Further, each realization of W_λ (e.g. by taking different λ) is a conditional random field, i.e. a possible representation of the precipitation field. We are not adding positive random values and we are also not working with a uniform distribution but with a standard normal distribution (Eq. 4). This standard normal distribution doesn't change due to the linear combinations (zero mean will always remain zero mean in this case and the unit variance is ensured due to the scaling of the weights λ). Precipitation fields are obtained via back-transformation of these fields.

P717 - I presume this is a minimisation being performed which I think should attain a minimum value if the W^* is representing the observed precipitation field and the scaling weight $k(\lambda)$ equals zero. I am unclear about this approach - this is attempting to create the observed rainfall sequence instead of doing a stochastic generation as far as I can figure this out.

There seems to be another misunderstanding. The described approach is a stochastic procedure as all fields used are random fields. We do not try to create the observed rainfall sequence except that we want to represent point observations as well as the observed runoff. Thus we are working with conditional random fields. As described above the weights λ are arbitrary if we only intend to reproduce the observed precipitation at the observation locations and the spatial variability. From these λ weights we

Printer-friendly version

Discussion paper



identify those which also reproduce the discharge. Thus the optimization described here is a function of these λ (and because it is an unconstrained optimization it is straight forward). In simple words, the field W_λ (which is already conditioned on precipitation observations) is modified such that the resulting simulated runoff (by the RR-model) is close to the observed runoff.

P7112 - the authors are saying multiple sequences are created by generating new random fields V_i and enabling something called uncertainty quantification - please explain what this means. I am very curious how different the sequences end up being - and when they are really different, whether their probability distributions are consistent with the observed series that was used. Also - am I correct in stating that the timing of these sequences will be fairly similar to the observed sequence - hence the final sequences will be representing uncertainty about each observed value more than representing a stochastic system that is generating equally plausible sequences (a bit like a weather generator does conditional to exogenous inputs, compared to a stochastic generator where no two sequences have any exogenous binding variable). Yes this sentence should explained a bit more. It is mentioned in P3L7 that "... Our goal here is an event based reconstruction of possible realizations of spatio-temporal rainfall patterns which are conform with the measured point rainfall data and catchment runoff response at best. For that we are looking for potential candidates of three-dimensional (space-time) rainfall fields for sub daily time steps and spatial resolution of 1km² ... ". This means that each candidate (or sequence) reproduce the point observation of rainfall without any uncertainty (or deviation). Only the grid points between the observation differs within the 3D rainfall field and contain the stochasticity given by simulations conditioned on the observed values.

p7115 - Am I correct in interpreting that the rainfall is generated known the marginal distribution at each pixel of the 118km² catchment? Or is it based on the 6 hours of

[Printer-friendly version](#)

[Discussion paper](#)



rainfall at the 10 monitoring stations alone? If it is the latter, assumptions must have been made to spatially interpolate/extrapolate the rainfall to other pixels. Please state these. If it is the former, this is a limitation I believe as you need to be sure about the spatio-temporal structure of your storm to help refine it further using the flows.

We are not quite sure what you mean with assumptions must have been made? Do you mean assumption must have been made to generate the synthetic reality? Or assumptions must have been made to generate possible realizations based on the 6 hours of rainfall at the 10 monitoring stations? If it is the latter then the assumptions that we made are that we can fit a marginal distribution and a spatial copula to these observations. Therefore only the values at the rainfall monitoring stations are used for the fitting etc. in order to make the synthetic test case a realistic scenario. But since this a synthetic test case all values at each pixel are known which enables comparison of the simulated results with the synthetic reality. We assume that the 6h precipitation distribution for the whole area is the same as the precipitation distribution derived from the observations (corresponding to the observation location).

P7I26 - some mention of the number of time steps in the observed record for rainfall and flows should be provided - there is a mention of 6 hours but I wasnt sure if that is the time step of the duration.

It is already mentioned in the manuscript nine lines above (P7L17: “A synthetic rainfall event of 6 hours duration with hourly time step ...”)

P9fig6 - I see all hydrographs are having roughly the same timing of the peak. So what I suspected about the time sequences of the rainfall is most likely correct. The differences across the storms would not be significant in terms of the spatial or the temporal pattern uncertainty that exists in real cases. I think this could be a limitation if the approach were being pitched as a stochastic generator - but could form an interesting way to generate alternate realisations of a storm sampled at specified point

[Printer-friendly version](#)

[Discussion paper](#)



locations alone.

The goal of the work is an stochastic “reconstruction” of spatio-temporal rainfall pattern ... (see title) which seems to be similar to what you called “to generate alternate realisations of a storm sampled at specified point locations alone”. We are not interested in exploring overall spatio-temporal pattern uncertainty (e.g. by performing unconditional stochastic simulations and considering measurement uncertainty too) since this was already done in research and has no benefit for the focus of this paper. Fig 6 shows the results of 200 simulated spatio-temporal rainfall pattern conditioned at rainfall point observations only, but containing the spatial uncertainty for the unobserved points. The hydrographs have to look a bit similar since all simulations used the same rainfall values at observation points transformed into runoff by the same hydrologic model (representing the hydrologic properties of the catchment).

And the need for having an accurate hydrologic model is a big limitation too as the uncertainty that arises from this can be significant.

Of course hydrologic model uncertainty plays an important role, but instead of changing the model to fit the observed discharge we estimate rainfall fields which fit the model and the discharge. As such plausible rainfall fields can be identified, the corresponding model and the rainfall field is plausible.

On the whole, I am unclear how I would use an approach such as this for my modelling application. I will need to have a fairly good idea of the spatio-temporal nature of the storm system to put this into use - along with having point rainfalls and modelled flow time series to help ascertain which sequences are good.

Some hints are given in the the summary section. (see P15L13 “... a reanalysis tool for rainfall-runoff events especially in regions where runoff generation and formation based on surface flow processes and catchments with wide ranges in arrival times at catchment outlet ...” or P15L22 “... where modelers are interested to explain the

extraordinary rainfall-runoff events ...”.) However, this section will be discussed more detailed in the revised manuscript.

I think the authors need to add more examples of this in their revision to establish a clear scenario how users will put their method into use. And some details of the tolerences etc that are used to make this stochastic should be added as I think they are not stated in the paper very clearly. Some indication of how this might perform over long storms/large catchments/very few point locations etc will really help readers. We are not sure, what you mean by “tolerences”. In general, the manuscript aims to present a new method and to show that it can deal with real world data. However, it is basic research and we are also very curious to explore the method further (see outlook P15L16). But this requires further developments (e.g. common interfaces for data, models, other types of copulas) which are not manageable within the next months. Among others we intended to show that models may be good even without any strong modification if we take the uncertainty of the precipitation into account. Thus models may help to improve precipitation estimation and one could consider model calibration under consideration of precipitation uncertainty.

author’s changes in manuscript

Some hints regarding author’s changes in manuscript have been already given in the comments section. Here, a summary of author’s changes in manuscript based on comments of both referees is given.

- chapter 1 “Motivation”: adding and discussion of literature regarding inverse hydrologic modeling, improved reasoning

HESSD

Interactive
comment

Printer-friendly version

Discussion paper



- chapter 2.2 “Rainfall runoff model”: description will be improved
- chapter 2.3 “Random Mixing for inverse hydrologic modeling”: description will be improved
- chapter 3.2 “Results and discussion”: discussion of the synthetic example will be enhanced and performed more precise.
- chapter 4.1 “Arid catchment test site”: figure of rain gauge measurements and additional information will be added
- chapter 4.2 “Results and discussion” discussion of the real case study will be enhanced and performed more precise
- chapter 5 “Summary and conclusion” reasoning will be enhanced and performed more precise
- revision of Figures 3,5,7,9
- all typos will be fixed and grammar will be improved

[Printer-friendly version](#)

[Discussion paper](#)

