Submission of revised version of the manuscript titled “AutoQS v1: Automatic parameterization of QuickSampling based on training images analysis”

We would like to thank Ute Mueller and Reviewer 1 for their valuable feedbacks. Their inputs led to significant improvements in the manuscript. We have incorporated the suggestions made by the reviewers in the revised manuscript. We did a complete proofreading for the article, and added a the DOI of the code. We hope that with these change and all modifications discussed below, the revised manuscript is now ready for publication in Geoscientific Model Development.

Best regards,
Mathieu Gravey

Response to Reviewer 1

We thank Reviewer 1 for the feedback and interest in our work, as well as the rapidity of the review.

This very interesting paper would ease the use of multiple-point statistical simulation by optimizing the parameters that generally require manual tuning before acceptable results are obtained.
The main problem with the current version is that it has been written in a hurry, it contains too many typos, and, more importantly, the explanations are unclear, sometimes because of brevity, sometimes because of improper English usage, sometimes because the author presumes that the reader knows much more than he should.

Thank you again for the annotated pdf and we clarify the sections you highlighted, we hope that it will satisfy the quality require for a publication in GMD.

The paper needs a thorough review of the text to make it understandable to someone who is not an expert in MPS simulations and might not have worked with either QS or DS.
A few more sentences or paragraphs explaining some of the parameters or some of the technicisms used would provide a better understanding. A rearrangement of some of the sentences is also necessary to ensure that the flow of information is logical.

We introduced 59-67, the QS and DS, and their parameters.

Many sentences with interesting statements are thrown out in the middle of paragraphs with which they are unrelated without supporting evidence.
And most importantly, the authors must emphasize that their approach is valid for a specific training image. When new simulations are to be generated, a new optimization must be carried out.

We changed the highlight “Calibration depends only on the training image” to “Calibration depends on each training image” to highlight this particular point.
In addition, we think the title and the lines 343-344, make it clear.

Response to Prof. Ute Mueller

We thank Prof. Ute Mueller for her feedback and interest in our work!

The paper by Gravey and Mariethoz provides an interesting approach to determining a suitable parameter set for geostatistical multiple point simulation. The method is based on simulating individual pixels and using the prediction error as the sole metric to find optimal parameters.The approach is sufficiently novel to warrant publication ultimately, but at present it lacks clarity and the quality of the scientific writing needs to be improved.

Thanks again for the positive feedback

Comments on the Introduction

1. You state that the parameterisation depends on the specifics of the algorithm. This is undoubtedly true, but it would help the reader to have a quick synopsis of the two algorithms and associated parameters before going into details of selection of parameters.

We added a part line 59-67 to introduce the two algorithms and their parameters.

2. Line 45: I would not talk about philosophies but rather about approaches here.

We updated it accordingly.

3. Lines 47 to 49: It would help if you explained what kind of metrics were computed and I assume you mean that the corresponding metrics computed for the simulation matched as closely as possible. But this is not stated clearly.

We extended this part (lines 70-74), in particular we added information about the JS divergence and its use.

4. Line 60: You say "If both approaches show good results, then they are both related to optimization methods, and therefore the user has no control over the duration of the optimization process". It is not at all clear to me what you are trying to say here. Both approaches rely on optimization of some kind, irrespective of the quality of the results. So can you please be more precise?

We rephrased this part, line 84-86

5. Lines 70-71: I am not sure what you mean when you say "the underlying principle of our approach is that a sequence of well-simulated pixels converges to a good simulation overall". Please clarify! (what does well-simulated mean? How can a sequence of pixels converge to a good simulation? Good in what sense?)

We rephrased this all part, hoping to clarifying our meaning, lines 77-79

Comments on “Challenges related to inappropriate parameters”

Most of this section is about verbatim copy, should you maybe change the title of the section to a title that highlights this?

We update the title to “Understanding and Addressing Verbatim Copy in Multiple Point Simulation”

Also, it seems the issue of “verbatim copy” is one of DS and QS only? It is also not clear how constraining the conditional probability distribution of Z(x) given an MPS algorithm is related to the issue of verbatim copy, (for ENESIM, SNESIM and IMAPLA this justifies the need of a large enough training image)

We added extra details in the section. However, from a technical standpoint, if we disallow data events with less than two replicates in SNESIM (and similarly in ENSESIM or Imapla), we guarantee to avoid verbatim copy in the resulting simulation. However, in practice, users often choose to allow such unique data events because they provide texture that can lead to visually satisfying results. There are therefore examples where verbatim copy does occur with algorithms such as SNESIM, although it is much less visible with categorical variables than with continuous ones.

We partially disagree with the idea that the training image plays a significant role in pattern variability. For example, in a chessboard pattern or any other repetitive texture, the image can be infinitely large, but the patterns remain the same. If the issue is not directly related to the size of the training image, then we agree that, in practice, using a large training image is the simplest solution to increase pattern variability, but this does not guarantee the absence of verbatim copy, even if the image is stationary.

6. Line 90-91: it would be useful to introduce the abbreviation for the threshold here

Done

7. Line 92 The definition of “verbatim copy” should be provided here rather than in lines 98 following.

We change the section to take this into account.

Comments on Method

8. Line 125 : The sentence “Binary variables are a particular case of continuous and categorical variables.” seems a throw-away comment whose purpose is entirely obscure.

After reconsideration on this point, we finally agree that the added value of this sentence is limited.

9. In the pseudocode for the QS algorithm it would help to have a definition of the entire parameter set

We added parametrization of QS as an example.

10. line 138: What do you mean by "find a candidate in T those matches N(x) using \theta?

In addition to the detail of θ (point 9), we update it to "Find a candidate in T those matches N(x) using the parametrization θ", to clarify that overall.

11. line 144: What do you mean by " A perfectly simulated pixel is a pixel that respects the conditional probability distribution" what probability distribution do you mean here? Presumably this is related to formula (1)?

We changed this part to clarify it, lines 156-158

12. line 159: please define the discretised parameter space \theta (it is not really defined in line 1 of algorithm 2)

We added the variable of this parameter space for the QS example. However with the change about theta in the Algoritm1, with think that now it seems logical.

13. line 161: you need to define what "th" stands for (I do realise you mean threshold, …..)

All these parameters are now introduced in the introduction.

14. in line 162 you talk about representative stages D of the simulation, but then in line 164 you say that D represents the density of a neighbourhood. What do you mean by density of a neighborhood? Also, is D part of the parametrerisation theta? If not, why not?

We added lines 186-187 to clarify this point.

15. Should algorithm 2 not have as a first step "randomly generate a set "V"

We change the input of algorithm 2 to clarify V, we think to have found a good compromise.

16. What do you mean when you say sample a neighborhood N(v) from T respecting D???, see also line 180

We added lines 186-187 to clarify this point.

Comments on "An efficient implementation"

Section 3.1 needs to be proofread carefully. You really should adhere to principles of good scientific writing and avoid starting sentences with symbols and also check word-order.

This part had a carful proofreading to adjust for this.

17. Line 198 The variables \theta_h and \theta_S represent sets. Sets cannot be added, but you can define their union. Thus \theta_h \cup \theta_S would be more appropriate than \theta_h + \theta_S, alternatively write \theta=(\theta_h, \theta_S)

Thanks to have highlighted this notation issue. In fact it is closer to a Cartesian product, we adjusted by replacing the **+** with a **x**, and an explanatory sentence.

18. the criterion you list in formula (4) needs further explanation and the sentence "… a given parameterisation is only further explored if the error is a range of a \sigma" does not make sense. To me the top line in 203 would make more sense written as

\epsilon(\theta,D,T)-\epsilon(\theta_{min}, D,T) > \frac{1}{2} (\sigma(\theta,D,T)-\sigma(\theta_{min}, D,T))

We adjusted to equation 4, and updated the text accordingly.

Comments on Results

19. Line 213 the term "uniform" has specific connotations. I would suggest the term "fixed" might better capture what you mean here. Also, does the kernel you consider here also have radial shape and why do you use \omega, when later on you use w?

Here we believe that the term uniform is appropriate. It means that all neighbors have the same weight, which is the definition of uniform. We added lines 229-230 some more info to clarify "uniform"

In figure 3 you introduce the term "ignorance treshold" without provision of a definition.

We added the definition line 232

20. Lines 227-228: "we also note that even if the parameterisation is logical it is diffifult to predict" what do you mean?

The algorithms begin by simulating a large structure, followed by simulating smaller patterns. The simulation is done in a hierarchical manner, where each smaller structure is a part of the larger one, making the simulation logically consistent. We clarified this line 245.

21. I believe Figure 4 warrants more discussion. Also, in the discussion n line 238 you use the term "two-stage process" but that seems ill chosen for the behaviour you describe

The two-stage process refers to the hierarchical approach explain before. We have one side lines going down where quality improves with extra neighbors, and lines going up (for high density), where the quality decrease with each extra neighbor.

22. line 252: It is the neighbors furthest from the location considered that are allocated the negligible weights with large values of the shape parameter \alpha, and the sensitivity of n decreases, but n doesn't become insensitive

With large \alpha values, increasing n does not change the results as the weight of the additional neighbors becomes negligible. But indeed, one could indeed argue that there is still a sensitivity to n for small n values, where all neighbors are close to the center of the kernel. This rework the section lines 270-275 to clarify this.

23. in line 270 the word adaptative needs to be inserted

Done.

Figure 8: the title of the figure in row 3 column 1 needs to be corrected.

Done. Thanks to have spotted this mistake.

In the appendix, please provide a reference for each of the training images.

We added them in the title each time.

It is interesting to note that the variogram reproduction for Delta Lena shows the "right" shape, but the sill are too low and connectivity is not as good as one would hope. Any thoughts on why this happens?

The variance reduction in MPS is a generally under-discussed point that seems to be universal to all MPS solution, especially for continuous variables. We could not find a study that looks at the question thoroughly, and we believe that it should be the topic of further studies.

Rasera et al.(2020) already highlighted this we can read at the end of the discussion "
While the algorithm is able to generate realizations that depict the same type of variogram structures present in the reference fine-resolution DEMs, simulations tend to underestimate the variability of the reference data. This is a common problem for conditional MPS simulations. Straubhaar et al. (2016) reported the same phenomenon while running simulations constrained to block data, and Oriani et al. (2017) experienced a similar effect when simulating rainfall fields conditioned to weather state variables and DEMs. "

Interestingly, our results provide an intuition of why this happens. It seems due to the absence of compatible large-scale patterns in the training image. Essentially, if one wants to reproduce the large-scale connectivity observed in the training image, the only solution is to do a verbatim copy of the large structural elements found in the training image. The reason is that such large connected bodies are too few in the training image. Since our approach tends to avoid verbatim copy, these large-scale elements are not reproduced.

If taking a binary version of this training image (water/no water) and doing a simulation SENESIM/Impala, we will see that this difference really appears when we change the setting of a minimum number of replicated from 1 to 2. As suggest earlier maybe a larger training image could certainly help to have more pattern variability.

It would also be good to see some results for multivariable images. You allude to the algorithms working in this case also, but being slow

In fact, we tested it for multiple multivariable scenarios, and it is rapidly becoming slow, due to high dimensional parameter space. It is also hard to visually analyze or interpret the results. We didn't see the added value of just dumping the results at the end. The current solution is close to brute force, and we expect in the future to come up with an improved (mainly faster) strategy that takes into account the relative order in parameters (in particular in kernel weights) using a strategy closer to gradient descent.

Reference

Rasera, L.G., Gravey, M., Lane, S.N. *et al.* Downscaling Images with Trends Using Multiple-Point Statistics Simulation: An Application to Digital Elevation Models. *Math Geosci* **52**, 145–187 (2020). https://doi.org/10.1007/s11004-019-09818-4