After careful consideration, we feel that it has merit, but is not suitable for publication as it currently stands. Therefore, my decision is "Major Revision."

General Content: There are numerous spelling mistakes and grammatical errors. This paper requires English editing and proof assistance. Authors should be careful using acronyms. If an acronym appears for the first time in the text, then it should be written in the full form.

Ln 28: What is "pollutant production" mean? Did authors mean by pollutant generation?

Ln 30: modulus? What does that mean, not sure why there is a mean and a modulus both? Again, could be something typo or just something else.

Ln 30-31: output intensities? Why not only output? Or load? Or Concentration? Or flux? Overall this section is also difficult to read.

Ln 31: intensities?

Ln 95: please do not refer readers to go and read literature from database. If there is literature pertinent to this paper, cite them; otherwise please do not direct readers. They can find any articles in Google scholar easily or other sites from libraries.

Ln 83-85: there is no link between these two sentences. The last sentence may need to be a starting sentence for this paragraph instead.

Ln 86: "The SWAT" of the present study"? Not sure what authors are trying to relate to?

Ln 123: Inconsistency is writing proper names? For example, Atrazine is written with capital A, but not other fertilizers/pesticides such as acetachlor and butachlor?

Ln 127: did the authors mean complete data? Statistics of what?

Model Description:

Ln 132: instead of "calculate", "predict" might be a better word.

Ln 142: the threshold of 0% creates large number of HRUs to capture all heterogeneities. The reason behind using 0% threshold is not well justified.

Model Inputs:

Ln 188: 1:250 000, there is an additional space in 250000, it should be a comma.

Calibration and Validation:

Ln 216-218: why only NSE? Using NSE alone, as a performance indicator is not sufficient. It will not indicate any bias in model output. I am assuming that there are more metrics used.

Ln 219: what is artificial parameter modification?

Ln 221-225: What is the difference between "real data" and "monitoring data"?

Aren't they both real?

Ln 243-245: If the authors used SWAT-Cup for automatic calibration, then what is this manual calibration? Sensitivity analysis is included in SWAT-CUP. Did authors conduct a separate sensitivity analysis outside of SWAT-CUP? If

so why? There is no need.

Ln 244-250: Did the authors separate baseflow from total streamflow and calibrate runoff and baseflow separately? If so, the method needs to be clearly stated on how it was done? Manual or SWAT CUP? What is the point of using LOADEST program during the calibration. I understand that LOADEST could be used to calculate pollutant load, but what is unclear it the progam used during calibration.

Ln283-299 In the preceding section, the authors mentioned that CN2 has no role. But it appears that runoff curve data was used for calibration. Please clarify.

Results and Discussions:

Similar observations could be made to TP and TN results and discussions. What are the common fertilizers used in the farmlands? Are there pastures and cattle feeding lots? What are the initial soil nutrients content? Did the authors use this information for model parameterization? It is unclear from the methods and discussion.

This paper could be improved if given time to rewrite everything. Paper could be published after revision.