Comments on "Physical versus economic water footprints in crop production: a case study for China" by Yang et al.

The study assesses the WF of 14 crops from 2001 to 2016 in both physical and economic terms for 31 provinces of China. It also analyzes the spatial agglomeration and the temporal trends of the WFs. This provides valuable information as to which crops have a higher economic return per unit of water consumed. However, it is not clear why some of the analyzes are needed and there is a lack of in-depth discussion on the interpretation of the outcome. I suggest that the authors address the following comments before the paper got accepted:

#1 Subsection 2.2: The authors modeled crop yield using the AquaCrop model but used statistical output. The question is, why didn't you use the modeled yield? Don't you trust the outcome of your modeling? It is understandable that yield modeling has large uncertainty and may require a rigorous calibration of the model. Modeled ET and yield are consistent as both derive from the model, but to what degree is the ET consistent with the statistical yield? How reliable is the statistical yield?

On line 88, I see that the modeled yield was checked against the provincial statistical yield. How good was the modeled yield compared to the statistical yield? I suggest that you plot the modeled vs. statistical yield to show the fitting between the two results. You can add the graph for each crop as additional information. Please explain if you did some manipulation on the modeled yield so it matches the statistical yield.

#2. Line 179 – 183: The interpretation of the SI is confusing. It seems to suggest larger SI to be better as province with larger SI are deemed to have "… (less water consumption per yield and higher economic benefits per water consumption unit). The SI is derived by comparing provincial WF and national average values. The result would thus mean, the province performs better in terms of generating higher economic benefits per unit of water. I would expect the SI value would be different if you compare two high performing provinces or evaluate the SI against a benchmark value instead of the national average value.

#3. Please provide for each crop the yield, CWU, PWF, EWF for the irrigated and rainfed systems separately. This will help to see if there is difference in the economic WP of rainfed and irrigated systems.

#4. The purpose of some the analyses are not clear e.g. Mann-Kendall. For the current study we don't need this analysis as the positive trend is visually clear from the figures. The authors themselves have used the test result on only one sentence (line 218-219).

#5. The discussion is more on comparing the WF of the provinces and saying this WF is larger here and there (Line 226-247 and 371-400). The reader can read this fact from the figures. Please expand the discussion of the result and explain why the WF is large in one province and small in another. Is it climate, crop varieties, or what? Substantiate your argument with some references.

#6. Please explain why there is spatial agglomeration of the EWP for the grain crops while none for the cash crops? What are the reasons for spatial agglomeration and what does it explain? Generally, provinces in the same climate

region will have more or less similar WF. The price of the crops may also be dependent on the total production volume, demand for the crop, availability of market. Or are there other factors that play? Please discuss.

#7. Line 442-443: the statement seems to suggest that to improve the green water, rain water harvesting and storage should be improved. Is rainwater harvested green water or blue water? You need to be clear what you mean by the rainwater harvesting. If the farmer builds small retention pond to collect rainwater, the farmer is collecting blue water not green. But if a farmer manages his field to increase the water retention through tillage system and mulching, this is increasing the green water. Please clarify your suggestions.

#8. Line 444-445: the statement "As for northern China, green water (rain water) should be converted into blue water (irrigation water) as far as possible, so as to reduce blue water consumption while ensuring and increasing economic benefits." is not clear. What do you mean the green should be converted to blue? How do you convert green to blue? Do you mean, we need to increase irrigation?

#9. Line 446-448: the statement is an empty statement: " The necessary way to alleviate the contradiction between water resource consumption and economic value creation is to adjust the agricultural production mode and the irrigation method according to local conditions." What do you find from your study and what practical ways do you suggest? How do farmers or policy makers adjust the agricultural production mode?

#10. There is a statement in a number of places (lines 186, 371, 447, 468, 492) that reads, "contradiction between water consumption and economic value creation". There is no contradiction between water consumption and value creation. You cannot create value without water consumption. The issue should be how we optimize the value creation per unit of consumed water. Please rephrase your sentences.

#11. On Table 4, the $EWP_{g,ir}$ is almost half of the $EWP_{g,rf}$. Why is that? Is the equivalent rainfed yield under the irrigated condition double that of the rainfed yield? Or the $CWU_{g,ir}$ is half of the $CWU_{g,rf}$? Generally, $CWU_{g,ir}$ is slightly lower than $CWU_{g,rf}$ but cannot be close to half. Please explain.

Minor comments

Please provide the spatial scale of the analysis in the last paragraph of the introduction section. I see in the discussion section that the analysis was done at a meteorological station level. How many stations per province?

Please provide the definition of "synergy evaluation index", what it does and how to interpret the result.

Line 30-31: change occupation to consumption on the following sentence
 "The water footprint (WF) (Hoekstra, 2003) reveals the occupation and pollution of water in the process of production or consumption and …."

Lin 52: the sentence is not clear - what do you mean by "WF coordination"?
Lin 57: Remove the period before (Tilman et al., 2011;Gao and Bryan, 2017;Cui et al., 2018)

Line 100: add reference to Hoekstra (2019)

References

Hoekstra AY (2019) Green-blue water accounting in a soil water balance. AdWR 129:112-117. doi:https://doi.org/10.1016/j.advwatres.2019.05.012