A novel data-driven analytical framework on hierarchical water allocation integrated with blue and virtual water transfers

The manuscript has been reviewed by two reviewers, and the authors have, in my view, responded adequately to the issues raised.

I have, however, additional comments that were not raised by the reviewers, and which I like to share with the authors, and I invite them to react to these. My comments are quite significant. When the authors submit a revised manuscript I may decide to send it again out for review.

The paper addresses an important and interesting issue, namely how to deal with water scarcity taking both blue water transfers and virtual water transfers into account, in a 2-level decision setting, and also taking account of three different water using sectors. Quite a complex setting, and an ambitious undertaking. Given such complexity it is important that the argument is clearly presented and here the manuscript needs to be improved significantly (towards the end of this comment I give details what in my view should be addressed to improve the readability of the paper).

A first comment is that the paper makes no attempt whatsoever to validate the proposed model, or at least to show that model outcomes are plausible; this could have been done, for example by comparing model results with observed data, and discussing similarities and differences. Similarly, the authors could have conducted sensitivity analyses to verify how sensitive model results are to changes in input values of certain critical parameters; not for the authors to draw far reaching conclusions, but rather to re-assure the reader of the validity of the model. In the current manuscript the authors have indeed conducted several sensitivity analyses (on water availability, sectoral water demands, prices of import crops and water price), and they draw far reaching conclusions on the outcomes, without even trying to explain these outcomes (see below). But this is not convincing to me. It would be much more convincing to use sensitivity analysis to demonstrate the robustness and plausibility of the model.
The comparison of model results with scenarios omitting virtual water transfers and blue water transfers and with the two-staged model (section 5.2) is interesting, although it is not clear to me how the proposed model differs from the two-stage model. I would suggest to use this comparison to demonstrate the validity (and value) of the proposed model; rather than to formulate far-reaching conclusions. First the model must be validated before it is used to draw conclusions.

A second comment is that the entire section 5.4 on sensitivity analyses raises more questions than it answers. For example, consider the available water (Table 5): if I understand it well, when there is 10% more water available, less water is allocated to agriculture (X1), and when there is 15% more water than the base case, more water is allocated to agriculture, and when there is even more water (20% more than the base case) the volume allocated to agriculture is suddenly halved. This begs for an explanation. Similar questions may be raised concerning the water allocated to the domestic sector (X3). When considering changes in the water price, it is concluded that the price elasticity of water transaction is not linear, as is clearly displayed in Figure 12. But the authors fail to explain why this is so: is this because the model is working as it is supposed to, or may there be something wrong with the model? These are just a few examples on the problematic nature of this section 5.4. As stated in my first comment: I would prefer that the sensitivity analyses are used to validate the model rather than anything else.

This also means that the concluding section can be significantly shortened (as the current conclusions spend quite some words on findings from section 5.4).

The third comment: I find the manuscript not easy to read. In fact it was for me quite cumbersome to read.

First, the English is often grammatically incorrect, unclear or oddly formulated. This needs to be addressed. I advise the authors to engage a native English-speaking academic to carefully check the language used.

Second, some figures were for me impossible to interpret. Consider Figure 5: I guess this is a flow diagram indicating the source and destination of the three crops considered; but I couldn't understand it. Figures 8 and 9 I failed to understand at all. The vertical axis of Figure 11 remains a mystery to me. In all figures, the axis should not only have numbers but also the units declared. E.g. what does the negative values in Figure 10d mean?

Third, the use of significant numbers, and exponents: in this type of models, parameters may have a maximum of three significant numbers. So declaring values such as 1880616733 doesn't make any sense. Why not report it as 1,880 x10^6? Or even better, as 1.88 x10^9. Tables 2 and 3 are difficult to interpret because of different exponents used; why not adhere to the convention and use consistently 10^3, 10^6 and 10^9 and not anything in between?
Tables 4, 5, 7 and 8 could benefit if the first column would simply explain what the parameters (X1 etc.) actually mean, including their units. It is also not clear in these table what the final allocated amount of water is (is the volume of water finally and actually allocated to the agricultural sector X1-WTI-WTD; and for industry X2+WTI and for domestic X3+WTD?). Similar, what is the net import (imports minus exports of a certain crop) value of each of the three crops?

Fourth, not all variables/parameters are properly introduced in the text (and only later the reader is aware of the existence of annexes), and several parameters/variables have wrong or incomplete units/dimensions. For example, what does the variable theta θ (eq. 1) physically mean? What does it mean if its value increases from 0.5 to say 0.8 and if it decreases to 0.3? Similar for vulnerability F (section 5.4.2). And what does "destroying degree (caused by deficient water withdrawal)" (line 498) mean; and how does it compare/contrast with "economic loss (caused by excess water withdrawal)" (line 499)?
The correct unit/dimension of ERPk is not RMB/m^3 but I think it should be RMB/kg. The correct unit/dimension of Wk is not m^3 but m^3/hm^2. All fluxes should have a time dimension, such as Xi = m^3/yr; but also evaporation ET, rainfall R, irrigation w; and even crop consumption (kg per person per year!)). What is the difference between wk (water irrigation for crop k) and x1k (water irrigated to crop k in the agricultural sector)? What are beta β and Pk (eq. C5)? Why is effective rainfall crop dependent (Rk)? Equation C3 is the well-known Penman equation and need not be reproduced here. A reference would suffice.

Fifth, what do you mean with the following often repeated sentence: "inter-regional exports and international imports" (lines 475, 478, 504, 521); does this mean that only export between regions within China are considered and not internationally, and that only imports from outside China are considered? If so, why? If not, change the formulation.

Minor comments:

- The paper is very long. Some of the equations and Tables that now appear in the main text can go, together with the Appendices, to Supplementary Materials.

- The introduction is very long, and can in my view be shortened

- The paper states that wheat and sunflower need more water compared to maize. I agree that sunflower generally requires more water than maize and wheat; but generally wheat needs a similar volume of water as maize, or slightly less.

- In Figure 4, I guess that one horizontal arrow pointing from the lower level (left) to the upper level (right) is missing; without it there is no dynamic between both levels.

- Why didn't the authors refer to their own recent publication (Xu, Yao, Zhou, Moudi and Zhang, 2019, in the Journal of Hydrology), since it also uses the Stackelberg approach, but in a different manner? In what ways does it differ?