The quantification of internal variability and model uncertainty sources in Multi-scenario Multi-model Ensembles of climate experiments (MMEs) is a key issue. It is expected to both help decision makers to identify robust adaptation measures and scientists to identify where their efforts are needed to narrow uncertainty and give more robust projections. A number of publications have been devoted to this uncertainty issue in the recent years. In hydrological impact studies, the contribution of model uncertainty due to hydrological models has been considered in a few works. Hydrological modelling uncertainty has been found to be a major uncertainty sources for a number of hydrological variables, such as low flows (e. Vidal et al. 2016; Guintoli et al. 2018, Alder et al. 2019). The results presented in previous works however focused on a few specific regions with different specific hydrological behaviours and hydrological processes. Other such studies for other contexts are thus definitively required to gain a better knowledge of the configurations where hydrological modelling is a key contributor to uncertainty in projections.

The present work participates to this effort considering 28 small catchments in California. Its objective is thus really relevant. However it lacks from major limitations and need major improvements before it may be accepted for publication. In short they are :

- It considers as uncertainty source related to hydrological modelling 2 components of hydrological models: uncertainty due to the “runoff production scheme” and to the “river discharge routing function”. The second one is likely not really relevant. A major uncertainty source in hydrological model is on the other hand not accounted for: the one related to the representation of evapotranspiration losses within the catchment

- The uncertainty analysis framework is not relevant as it disregards the internal variability in climate projections. In addition, the Vetter et al. (2015) subsampling scheme for the ANOVA is not relevant.

- Some choices retained for hydrological modelling are not really convincing. Among others :

o The Hydrological Model based on the RunoffCoefficient concept cannot be consider as a relevant hydrological model. Such a representation never corresponds to a modern and state-to-the art representation of hydrological processes.

o The kinematic wave used to model the subsurface flow in the VIC model is spurious.

- The Bayesian Model Averaging method to produce pdf”s of change is based on weak assumptions.

The overall description of the work should be otherwise largely improved, especially what is related to the description of hydrological models which have been already described in a number of previous works.

My detailed comments are given below.

Hydrological models and contribution to uncertainty

Low performance models

The performance of the three models is not acceptable. (cf. results obtained for 1/3 of simulated years (figure 2/d)). The representation of the annual water balance is not relevant and has to be improved in all models. How are represented the evapotranspiration losses ? What is the uncertainty related to precipitation inputs? How many precip. stations are available for each catchment ? This may be a serious limitation of the different models considered here.

A low performance model (for the reproduction of observed time series when meteorological forcing data are observed ones) can obviously not be integrated in the analysis. The performance of the RCM model (based on the runoff coefficient) is too low. This is also the case for the VIC Model parametrized here. In such a configuration, as 2 models are known / found to be very poor in the representation of the target system, the Bayesian Model Averaging cannot be considered as a solution to weights the outputs of the different models.

Non relevant RCM model. The Runoff Coefficient Model used for the simulation of the "runoff production" process is based on the simplistic and non-relevant hypothesis that the runoff can be estimated for any time based on a constant runoff coefficient This representation is obviously non-appropriate. (the runoff coefficient actually takes two values depending on the level of relative moisture of a said upper layer but this is obvioulsy not a satisfying / state to the art approach). This model is much too simplistic and cannot be integrated in a state-to-the art uncertainty analysis. If climatologist may consider such a model as acceptable (probably the reason why the authors could publish this model in the "climatic change" journal), no serious hydrologist would consider such a model relevant for the long-term simulation (continuous simulation over multiple years) of the runoff generation processes and then discharge generation processes. This RCM should therefore be either removed from the analysis or adapted (the "runoff coefficient" should at least be a continuous increasing function of the 'relative moisture" variable).

Description of models The descriptions of the models have to be synthetized, have to refer to the original papers which first described the models. In the present analysis context : it is then important to highlight what key representations are different from one model to the other.

Transfer functions. Transfer of runoff and subsurface flow: the kinematic wave is used to simulate both transfer component. It is likely not relevant for the subsurface flow. Is there any reference to give that uses such a model for subsurface flow ? Most models are either physically based, based on Darcy richards equations or are based on conceptual (linear on non linear) transfer reservoirs…. It is also spurious that it could be the official representation in VIC.

Routing model : In the introduction, the routing model is said to be a important contributor of hydrological uncertainty. This is likely not the case as the routing process (and not the runoff transfer processes) in small catchments has little chance to influence significantly flood regimes. It only produces a small distortion of runoff discharge series along the river network. The uncertainty due to the routing model is a priori not an issue. There is thus no reason to introduce this issue in the introduction (by the way, the authors conclude in the introduction they will not consider this issue in the paper, another reason to not introduce this issue there. In all cases, the justification the authors give for this omission (they say "there are less variants in the routing scheme") is not a good reason to disregard this issue. The good reason is mentioned above > in the hydrological behaviour of catchments, the main uncertainty sources are not due to routing but to the production/transfer processes and their representations.

Runoff representation. As mentioned in the introduction, hydrological models usually focus on a given runoff generation process (either Hortonian or Dunian). The runoff generation process is however not the main source of uncertainty. Runoff production is obviously also dependent on the initial saturation conditions of the catchment. Then, the representation of the water balance of the soil and the way its temporal dynamic is represented is important. Appropriate experiments have thus to be found. There are here two 2 majors issues : the type of runoff production (rainfall to rapid runoff) and the losses by evapotranspiration / evaporation that determine the state of soil storage/saturation… The second one is at least as important as the type of runoff production process (excess saturation / excess infiltration)…. This has to be discussed / integrated in appropriate experiment. …

The way Low Flow are simulated and underground storage is represented could be also an issue. There is likely a large impact on underground storage on the sensitivity of low flows to climatic changes (especially to different changes in different seasons and joint precipitation / temperature changes).

Uncertainty analysis.

- A deep review of all existing works focusing on the characterization of uncertainty sources in hydroclimate projections is definitively missing. The main conclusions of such works have to be clearly identified. The necessity/interest for additional work on this issue also. The contribution of the present work also. The usually large contribution of internal variability also.
- Methodology framework used for the Uncertainty Analysis. Authors argue they develop a framework to analyse uncertainty sources related to different sources. They just apply a methodology already presented elsewhere. The way this study could be used for the definition of a proper “methodological analysis framework” would have to be specified.
- The classical setup of Multimodel Multiscenario of climate experiments makes the number of scenarios, the number of climate models, the number of impact models very different. Vetter et al. 2015. proposed a resampling approach to apply the ANOVA on subsamples of the MME with same numbers for each uncertainty source. This approach is not really relevant as it does not use all available data for a unique / joint estimation of all uncertainty components. Classical ANOVA approaches account for these unbalanced configurations can be used. They typically propose unbiased estimators of all uncertainty components. Refer to any ANOVA handbook.
- Uncertainty sources accounted for in the analysis disregard internal variability. As in many hydrological impact studies unfortunately, internal variability in hydroclimate projections resulting from the internal variability in climate (leading to low-frequency fluctuations in climate projections) is disregarded. This makes the uncertainty analysis not sound (the GCM uncertainty estimated in the present work is a mix of GCM model uncertainty AND GCM internal variability). This internal variability component is one major uncertainty sources > the authors can not disregard this. The review given in the introduction has to discuss this issue and the framework used to characterize uncertainty sources has to account for this uncertainty source. The authors have especially to read and account for the papers of Hawkins and Sutton, 2011, Hingray and Said, 2014, Hingray et al. 2019 and Evin et al. 2019 for time series ANOVA approaches that are relevant for such configurations (The paper of Hingray et al. 2019 show that the classical Single Time approach of Yip et al. 2010 widley used in the recent years likely produces biased and wrong estimates of most uncertainty components).

PDFs of climate projections

The “Bayesian Model Averaging” work in this work is not appropriate for at least two reasons:

- It is not relevant to apply it on a set of chains where some are very poorly performing (here hydro models, their performance to reproduce time series of observations from meteorological observation is very poor)

- Because of internal variability, the evaluation of chains is likely not relevant. I understand that : All combinations of hydrologic models, parameter sets and GCMs (3x3x10=90) seem to have been evaluated. What are the criterion / variables / data periods used for the evaluation ? Do you consider that the evaluation data to be used for each simulation chain are those obtained with this chain for a given control period ? Do you estimate for instance the ability of a given chain to reproduce over this control period the statistical distribution of the observed variable for this control (i.e. the same) control period ? If, yes, this evaluation is not really relevant. The observations are actually one realisation of the recent (control) climate (because of the internal variability) and the simulations are also one realisation of the recent (Control) Climate. So there is no reason why simulations from a given GCM/HydroModel should correspond to observation for a given Ctrl period.

Other comments :

Hydrological models

VIC : The description is not clear. For instance the following 2 sentences have to be clarified: “Ds is the fraction of Dm at which the non-linear base flow begins. Ws is the faction saturation of at which the non linear base flow occurs “ … What is the difference between Ds and Ws ? At least, both seem to be very correlated. What is the relation ?

Topmodel :

- The detailed description of the model is not required. One has to refer to the reference and original paper describing the principle of TOPMODEL (Beven et al .2000)

- In the current paper, it is not clear on which units TOPMODEL is applied. In the original version, the model is applied in a lumped way, with one single model for the whole basin. Is the model here applied in the same way or is it applied on each hydrological unit ? in the latter case, on which DEM data (resolution) are estimated the topographical index ?

Ln 231 : the water movement between soil layers is similar to that in the modified VIC. Is it for the Topmodel? In all cases, a schematic representation fo the different models (with identification of reservoirs could be added to have a clear idea of the structure of the different models). How many components of discharge are simulated ? runoff + subsurface flow + base flow ??? How is simulated the transfer of each component to the outlet of the catchment ?

Ln 231 – 241 : what is the model described here ?

Ln 258. Give a table with a list of parameters to be calibrated.

What are Ksall and KssAll. Please clarify their role. It is mentioned they allow to account for spatial variability of parameters within the catchment. How ?

How are estimated the parameters of the models for the 28 bassins shown in Figure 5

Equations 29/30 and 31/32 are the same than equations 27/28. To be simplified.

Figure 2a : what is described here ?

References to be considered / integrated:

Importance of hydrological processes in hydrological changes.

Folton, N., Martin, E., Arnaud, P., L'Hermite, P., and Tolsa, M.: A 50-year analysis of hydrological trends and processes in a Mediterranean catchment, Hydrol. Earth Syst. Sci., 23, 2699–2714, https://doi.org/10.5194/hess-23-2699-2019, 2019.

Robust ANOVA frameworks to partition uncertainty sources in MMEs

Hawkins et Sutton (2009). The potential to narrow uncertainty in regional climate predictions. Bulletin of the American Meteorological Society, 90(8), 1095–1108.

Hawkins et Sutton (2011). The potential to narrow uncertainty in projections of regional precipitation change. Climate Dynamics, 37(1-2), 407–418.

Hingray et Saïd 2014. Partitioning internal variability and model uncertainty components in a multimodel multireplicate ensemble of climate projections. J.Climate. 27(17); pp. 6779-6798.

Hingray, et al. 2019. Uncertainty components estimates in transient climate projections. Precision of estimators in the single time and time series approaches. Clim.Dyn. 53:2501-2516

Evin, et al. 2019. Partitioning uncertainty components of an incomplete ensemble of climate projections using data augmentation. J.Climate.

Applications with hydrological model uncertainty

Vidal, et al. 2016. Hierarchy of climate and hydrological uncertainties in transient low flow projections. Hydrol. Earth Syst. Sci.

Giuntoli et al. 2018. Uncertainties in projected runoff over the conterminous United States. Climatic Change.

Alder et Hostetler. 2019. The Dependence of Hydroclimate Projections in Snow-Dominated Regions of the Western United States on the Choice of Statistically Downscaled Climate Data. WRR. 2019.