We thank the reviewer for very useful and generally positive comments. By addressing these comments, we believe the paper will be substantially improved, particularly with respect to refined focus and more detailed description of hydrological and thermal models. The reviewer comments are in italics and our responses are in normal font, the proposed text additions and modifications are in bold.

Please note that in the following, "P", "L" and "SM" stand for page number, line number and Supplementary Materials, respectively.

#Reviewer 2:

General comments:

The only main concern I have is some lack of detail on the description of the hydrological and thermal models (see below). I also find that the manuscript is a bit too long, and I believe that its readability would increase if some of the less relevant results (say e.g. Figs. 6, 7, 9 and respective paragraphs commenting them) were moved to the supplementary material.

We agree, and Figures 6 and 9 will be moved to SM.

Main comments:

L. 95-112: Many details on the implementation of the EROS model are missing. What are the free parameters of the model? What method for model calibration was used? How were Nash-Sutcliffe index and overall bias combined in the objective function? What was the rationale of choosing sqrt(Q) instead of Q in the computation of the Nash- Sutcliffe index? Was the output of this model validated with respect to Q time series not included in the calibration dataset?

We agree, and therefore the section "EROS hydrological model" (section 3.1.1), P4L95, will be modified as follows to include the principles of the EROS model, its input data, calibration process and variables, and applications in other studies:

"The EROS semi-distributed hydrological model simulates the daily streamflow at the outlet of 368 sub-basins (ranges between 40 and 1600 km^2; mean drainage area = 300 km^2), designed to be as homogeneous as possible with respect to land use and geology. On each of these sub-basins, the water balance is modelled by a lumped model using three reservoirs and a routing function for propagation across sub-basins (Thiéry, 1988; Thiéry and Moutzopoulos, 1995; Thiéry, 2018). This hydrological model has already been used in several other studies including climate change impact studies (Ducharne et al., 2011; Habets et al., 2013; Bustillo et al., 2014).

The EROS model uses daily Ta (°C), solid and liquid precipitation (mm), and reference evapotranspiration (ET0, mm). Ta and precipitation are provided by the 8 km gridded Safran surface reanalysis from Meteo-France over the 1958–2019 period (Quintana-Segui

et al., 2008; Vidal et al., 2010). ET0 was computed from Safran variables with the Penman-Monteith equation (Allen et al., 1998).

For calibration, 352 hydrometric stations with observed Q data were extracted from the French national Banque Hydro database (http://www.hydro.eaufrance.fr/) (Figure 1, right panel). Stations along the main Loire (14) and Allier (11) rivers are influenced by operations of 4 large dams for hydropower, flood control and low-flow management, notably through summer releases. Time series at these stations have been first naturalized by EDF (Électricité de France) based on variations in water storage in reservoirs due to operations based on target storage curves (Naussac, Villerest) and optimization of the hydropower energy generation for the electric grid needs (Grangent, Montpezat). EROS was then calibrated over the 1971-2018 period against daily Q at all 352 sub-basins with at least 10 years of daily observations. The calibration aimed at optimizing all unknown parameters (soil capacity, recession times, propagation times) through maximizing the Nash-Sutcliffe Efficiency (NSE) criterion on the square root of streamflow and minimizing the overall bias. Considering the NSE criterion on the square roots of the flows provides an estimate of model performance without favoring neither high flows nor low flows. The overall calibration criterion C is: C = mNSE – w.mRB where the weighting factor w is fixed at 0.05 .

A 3-year warm-up period (1971-1974) was discarded from the overall calibration period (1971-2018, which maximizes the number of streamflow observations). The calibrated model was then used to simulate streamflow over the whole 368 homogenous sub-basins in the Loire basin over the whole 1963-2019 period. Note that although meteorological variables are available from 1958 onwards, the first years (1958-1962) were discarded from the analysis to ensure the model convergence."

Moreover, to avoid confusion (which happened for the other reviewer), P5L110 of the current manuscript will be removed and it will be added between P6L150 and L151 as following:

"- Streamflow: the daily streamflow simulated at the outlet of 368 homogeneous sub-basins by the EROS model (see section 3.1.1) are redistributed along the river network inside each sub-basin according to the reach drainage area for informing the T-NET model at the reach scale. The ratio of sum of the lengths of all reaches upstream of a reach to the sum of the lengths of all reaches located in a sub-basin is used as a proxy for the drainage area of a reach."

L. 113: If I understood correctly, the T-NET model does not have any free parameter, hence it does not require calibration. This should be stated explicitly.

We agree. In this regard, the following paragraph will be added L90 in Section 3:

"We used two models to calculate stream temperature in the Loire River basin according to the method developed by Beaufort et al (2016). The first model is the semi-distributed hydrologic model EROS, which estimates daily streamflow at sub-basin outlets. The second model is the fully distributed, mechanistic temperature model Temperature-NETwork (T-

NET; Beaufort et al. 2016 and Loicq et al. 2018) that uses streamflow from EROS and meteorological reanalysis data to estimate Tw at each reach in the Loire River basin. These models are briefly described below, and are detailed in Thiéry (1988), Thiéry and Moutzopoulos (1995), and Thiéry (2018) for EROS, and Beaufort et al. (2016) and Loicq et al. (2018) for T-NET."

LL. 115-119: The authors state that the first step for Tw estimation is the computation of the equilibrium temperature (Te). This is the stream temperature at which the net heat flux across the surface of the water body is null, and is a useful concept when one aims at producing a simplified, (semi-)mechanistic temperature model where Te is expressed e.g. as a linear or logistic function of Ta, which allows discarding the exact computation of the various non-advective heat fluxes (latent, short-wave radiation, etc..). However, here the non-advective heat fluxes are exactly computed (this is not explicitly mentioned here but is reported in Beaufort et al., 2016), and Te is calculated as the Tw value that nullifies the sum of non-advective heat fluxes. This is technically equivalent to a fully mechanistic model where non-advective heat fluxes are included in the energy budget at a reach (or cross-section) scale. Thus, I find it a bit confusing to invoke the concept of equilibrium temperature here.

We agree, and in this regard, we will improve description of the thermal model in P5L115-116 as following:

"TNET is a fully mechanistic, 1D model that simulates hourly $Tw_{i,j}$ at distance, i, along reach, j, by solving the local heat budget. The heat budget of each reach includes six fluxes ($W\ m^{-2}$): net solar radiation, atmospheric longwave radiation, longwave radiation emitted from the water surface, evaporative heat flux, convective heat flux, and groundwater heat inflow. Briefly, the model calculates the longitudinal change in Tw at time t (dTw/dx) for steady-state conditions in order to achieve thermal equilibrium (i.e., $\Sigma H_{i,j}=0$, where H is a heat flux), while accounting for confluence mixing. Detailed information about the T-NET model principles and calculation of six heat fluxes at the water-air and water-stream bed interfaces, thermal propagation were provided in Beaufort et al., 2016 and Loicq et al., 2018. Note that unlike the EROS hydrological model, the T-NET thermal model does not have any free parameter, hence, it does not require calibration and it is only validated."

L. 174: It is unclear why 72 stream temperature stations are mentioned here, but then validation is only performed on the 14 stations with long-term continuous data. Are the other 58 stations ever used in the analysis? If not, this information should be discarded.

Indeed, all 72 stations were used for assessing model performance in terms of bias and RMSE while 14 of these stations with long-term data where used to assess the ability of the model to capture long-term data. To avoid this confusion, we will do some modifications to the text as follows:

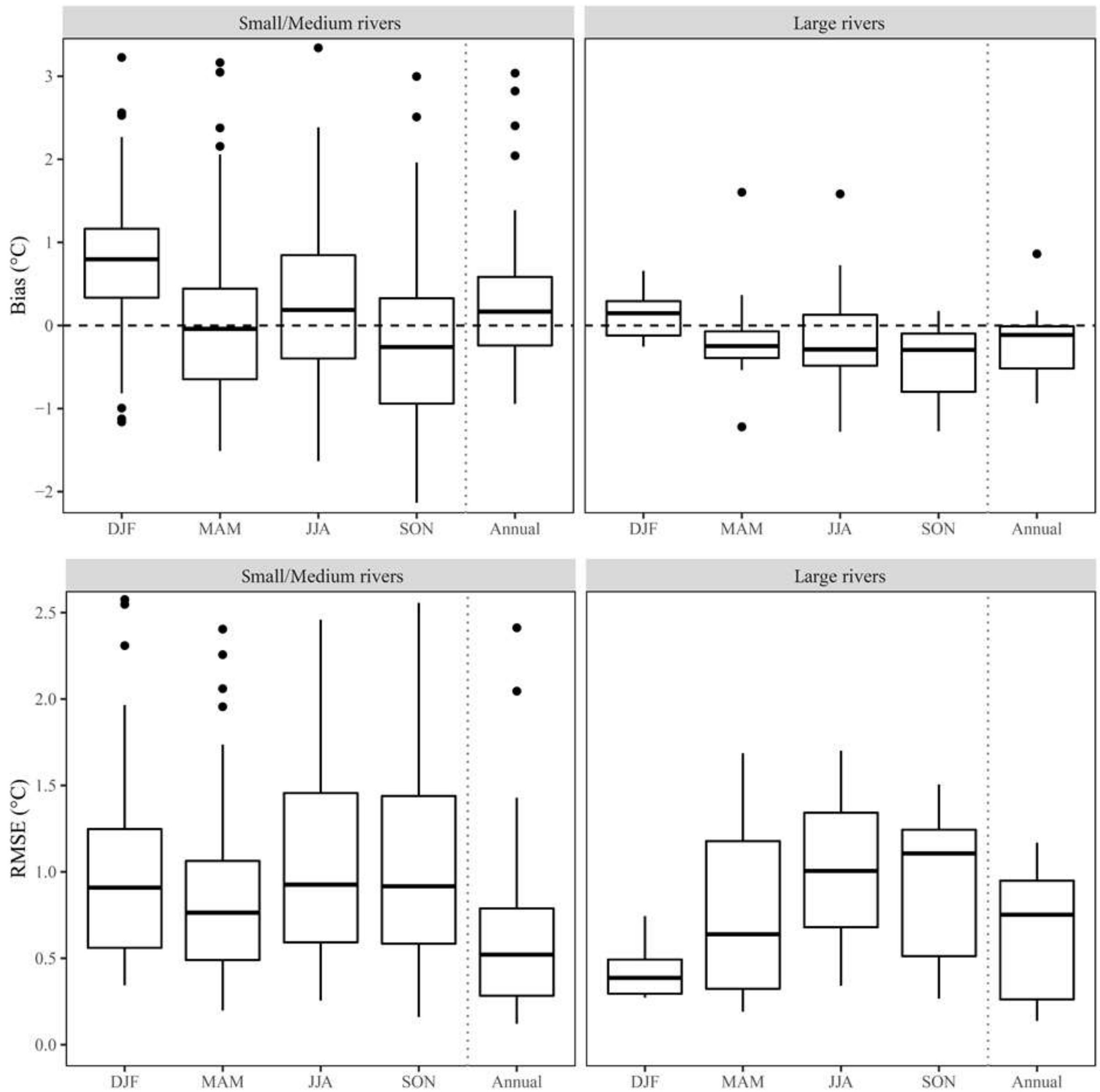
L. 173-175: **“Therefore, model performance was assessed on 72 near natural Tw stations, which are weakly influenced by impoundments and with continuous daily data over the 2010–2014 period (see Fig. 1). These stations were identified using the thermal signatures approach that allows distinguishing between natural and altered thermal regimes”**

L.180-183**: “Long-term continuous data was available at 14 of the 72 near-natural stations, including 9 stations with 8-13 years data and 5 stations with 20-40 years data. They were used as a validation dataset for the seasonal and annual trend assessment (see Table 1, red points in Fig.1 for the location and Fig.S1 for annual Tw time series)”.**

Figure 2. It would be interesting to have an estimation of the performance of the T-NET model in terms of RMSE (or MAE), in addition to the mean bias (as the latter can obviously be close to zero even when absolute errors are very high).

We agree. The following figure (including RMSE) will be added to the bias figure in the SM (Fig. S2). The following results will be also added to P9L224:

“The median RMSE of the T-NET thermal model, for small and medium rivers, ranges between 0.52 °C (Annual) and 0.91 °C (DJF and JJA) across seasons (see Figure S2, bottom left panel). For large rivers, the median RMSE of the T-NET thermal model ranges between 0.38 °C (annual) and 1.11 °C (JJA and SON) across seasons (see Fig. S2, bottom right panel). Moreover, 53-83% stations (resp. 50-100%) on small and medium (resp. large) rivers have a RMSE<1 °C across seasons.”

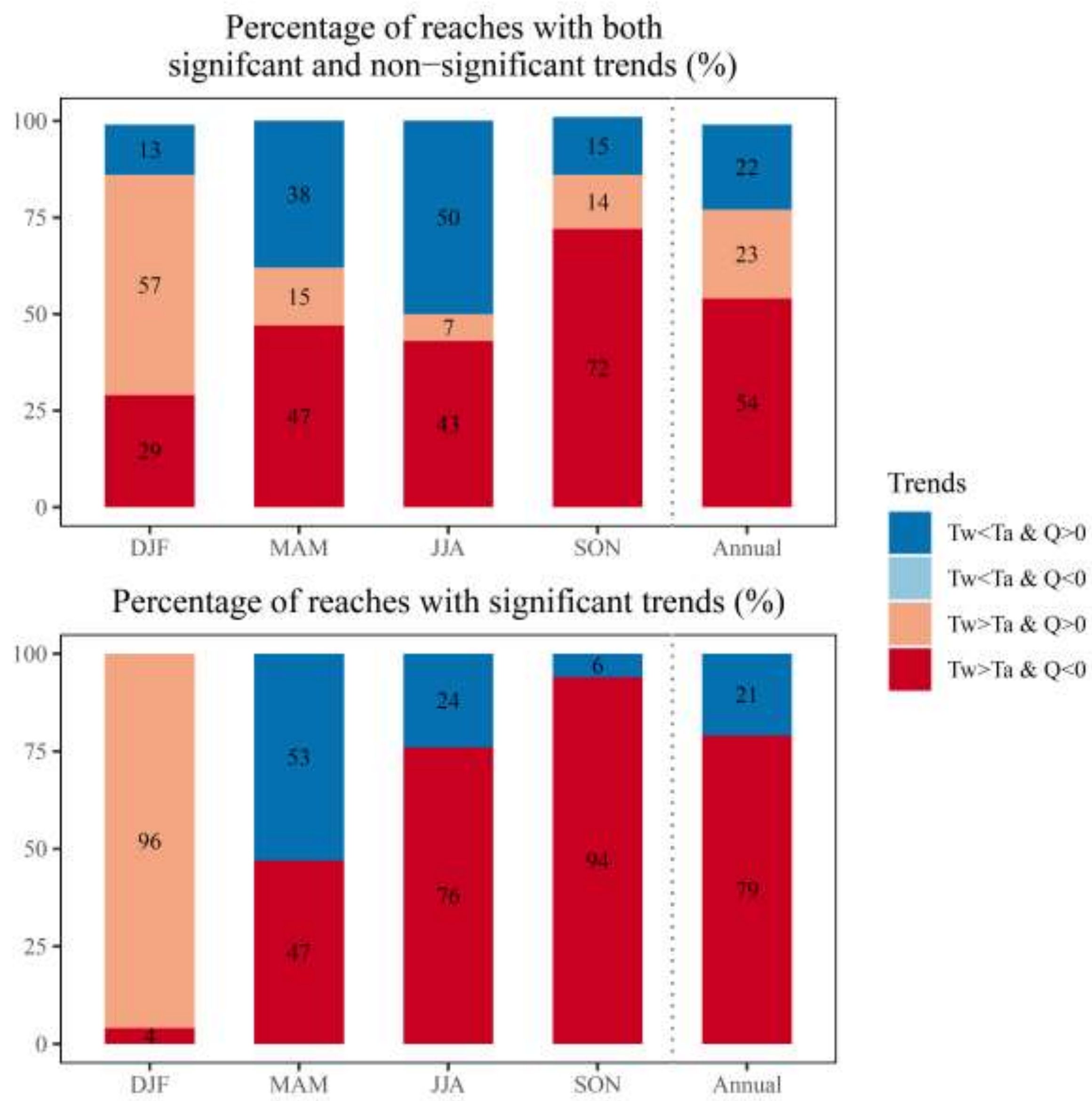


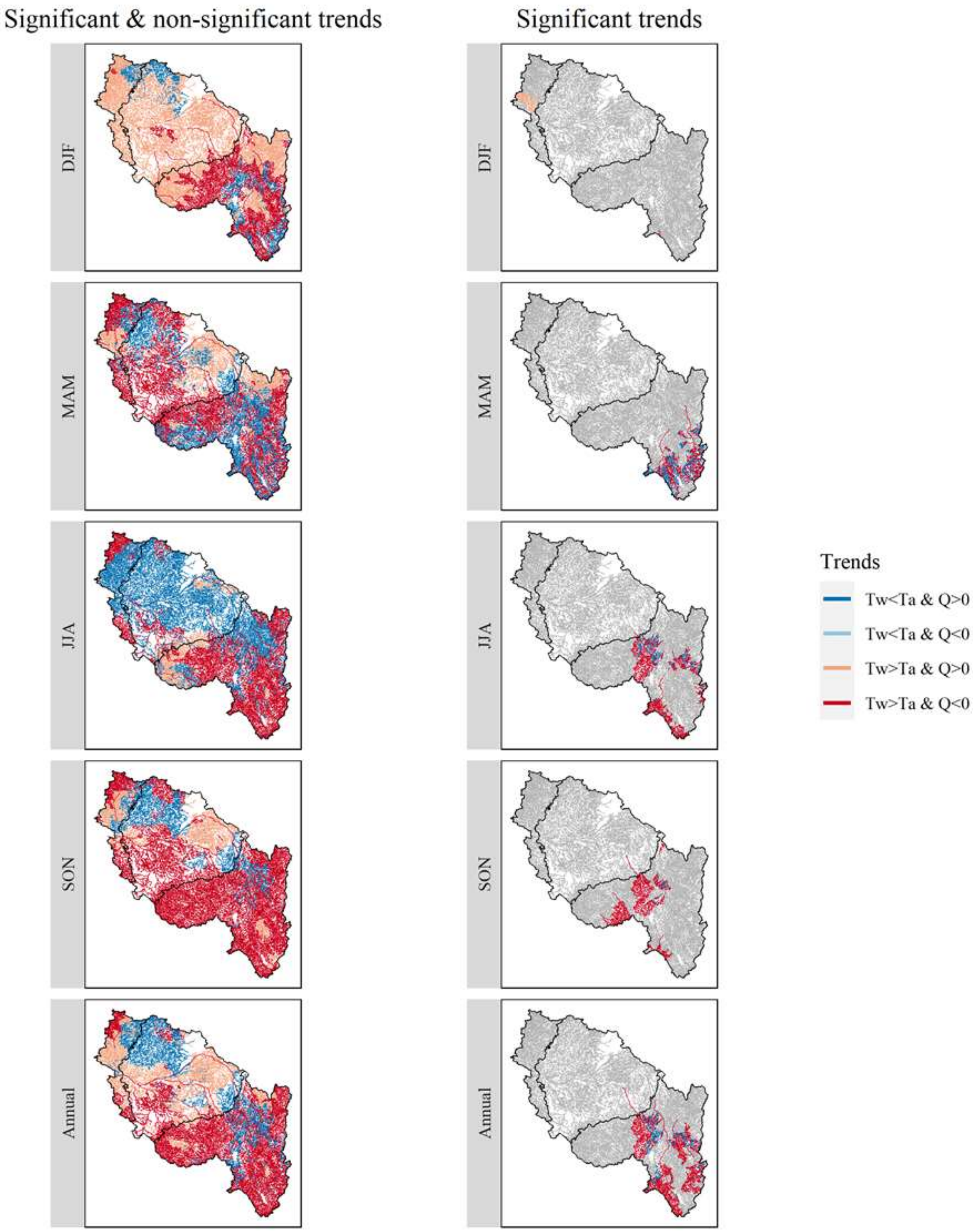
LL. 260-261: "Indeed, where … for all seasons". This is an interesting observation, but, for the sake of fairness, one would need to check whether the majority of reaches where Ta trend > Tw trend also showed a decreasing Q trend.

We agree. In this regard, we will modify Fig.5 as follows, which includes your suggested class, and two types of data (1) significant-non significant trends, and (2) only significant trends as requested by the other reviewer. With this modified Figure, P11L260-264 will be modified as follows:

"Indeed, where Tw trends exceed Ta trends, decreasing Q trends occur coincidentally at the majority of reaches (43-94%) for all seasons (with the exception of winter) irrespective of whether all significant and non-significant trends are considered (Figure 5, top panel) or only significant trends of all three variables are considered (Figure 5, bottom panel). Of these specific reaches where all factors converge (trend in Tw higher than trend in Ta, and decreasing trend in Q), most are located in HER A irrespective of considering all significant and non-significant trends (52-90% of such reaches across seasons; see <u>spatial figure, left panel</u> in SM), or considering only significant trends of all three variables (100% of such reaches across seasons; see <u>spatial figure, right panel</u> in SM)."

Please note that the spatial figures showing location of these reaches will be put in SM (between Figure S7 and S8 of the current manuscript).





Spatial figures

L. 274: "shifting by approximatively +2 °C". I find this unclear. Is the shift observed by comparing current Tw and Ta values with those observed at the change-points (~1988)? Please state this more clearly.

We agree this was not clear. We will clarify the text on P13L273 to state that**:**

"Tw and Ta anomalies exhibited clear negative-to-positive change-points in the late 1980s at nearly all reaches, with median values increasing by approximately +2 °C after the change-point.

LL. 288-289: Strahler order is significantly and positively correlated with Tw only in Spring for all HERs (and in summer only for HER A). The way this sentence is written, one is led to think that this correlation is strong more often than it is the case by looking at Fig. 9.

We agree and it will be modified in the revised manuscript.

Minor comments

We agree with all comments except the ones to which we have responded.

L. 58: "in meeting these goals"

LL. 92-93: I think it would be clearer to mention "meteorological seasons", and then use the season names in lieu of the acronyms DJF, MAM etc. in the figures. This would make the figures much more intuitive.

L. 108: if 1971-1974 is the warm-up period, then the calibration period should be 1974 (or 1975)-2018.

L. 109: Why is the number of subcatchments here (368) higher than 352 (mentioned in L. 104)?

There were observed Q at these 352 stations but after calibrating the hydrological EROS model at these 352 stations, the EROS hydrological model simulated Q at 368 sub-basins. This will be added to the revised manuscript.

L. 112: "were discarded". Moreover, "(1958-1962)" is four (or five) years, not three years.

LL. 140-143: Eqs. (3-5) can be condensed in one equation: SF = max((W_shaded)_left * vc_left, (W_shaded)_right * vc_right)

L. 178: I suggest indicating the range rather than SD for drainage area values (when SD > mean, this does not make much sense)

L. 201: verb missing. Perhaps "used to detect/assess synchronicity".

L. 227: “The trends of both modeled and observed Q”

L. 247: “The highest (resp. lowest) Ta trend values”

L. 248: In Fig. S6, the season with lowest Ta values seems to be fall, not spring, with 0% of reaches showing values > 0.4 °C /decade.

L. 257: “The medians of Tw… than those of Ta”

L. 268: “they are either warm and wet”

L. 308: “strongly suggesting an effect on Tw” or similar

L. 319: “comparing trends… gives us a comprehensive view”

L. 364: “The warming effect… seems more significant”

L. 370: “an increase of >25%”. Increasing from 15% to 40% (this is what I believe the authors are referring to, see LL. 299-301) is actually a 267% increase. Perhaps it would be best to reformulate as “increasing riparian shading from 15% to 40%”.