Dear Editors and Reviewers,

We have closely considered the comments of both reviewers and have undertaken a major rewrite of the text to improve our presentation and make the scientific novelty more clear. We thank both reviewers for their comments, which have greatly improved the manuscript. Pending approval from the editors, we will upload the updated manuscript. In the mean-time, we have outlined our edits and include details about specific improvements.

As described in the detailed responses below, the introduction is now more succinct and focusses more clearly on broad relevance and scientific questions. Much of the locally specific material has been moved to either section 2.1 (Methods>Study Area) or to a supplement. The methods section has been reduced by more than 25%, both by careful editing but also by moving supporting information to a supplement. The results section has been streamlined to more clearly focus on a description of results, following scientific convention. The discussion has been reorganized and interpretation and analysis has been moved there from other sections. The reorganization more clearly focuses on interpretation (section 4.1), implications (section 4.2), and limitations (section 4.3). Where requested, additional explanations and specificity have been added. Overall, the text is more than 100 lines shorter.

Through the careful rewrite, the modeling approach and novel contributions are more clearly findable and understandable, and the methods are more succinctly and clearly explained. The concerns of both reviewers concerns about the length and understandability of the manuscript have therefore been addressed. Additional comments and details are given in the response to the first reviewer.

Best regards, also on behalf of co-authors,

Stefan Talke

Response to Reviewer 2

We thank the reviewer for their positive and constructive comments; they have improved the text. Replies to comments are shown in blue below.

General Comments:

The manuscript gives a thorough investigation of changes in stream temperature in the Willamette river from the 1850s to the present. The authors compile air temperature, discharge, and stream temperature data from a variety of sources throughout the region, and use these data to construct statistical models that give insight into the magnitude of change in stream temperature throughout three different historical periods. The authors then investigate the seasonal and interannual changes in observed stream temperature records and validate their models by comparing model accuracy (including a comparison to other stream temperature predictions in the same region.) The models are then used to quantify the importance of climatic and system changes (notably reservoirs, loss of shading, and landscape alterations) over time, with good quantification of how the two compare in magnitude and seasonality using sensitivity experiments.

The results of this paper give a clear insight into how and why stream temperatures are changing in the Willamette river, investigate the causes of these changes, and draw connections to the ecological impacts of rising stream temperatures.

Thanks for the positive comments.

While the paper is well framed, the extensive methods section makes it difficult for the reader to keep track of the different data sources and to get a clear and concise understanding of how the models presented in the results were constructed. Thus, the manuscript could benefit from making the methods section more concise, contextualizing the model with other models that have been used for stream temperature modeling, with additional information made available in supplemental materials.

We agree that the Methods section was too long. As also described in the response to the other reviewer, we have made the suggested changes. We have reduced the introduction and methods section by more than 130 lines, moving some material to a supplement and eliminating it altogether. We have also edited the text for clarity. The remaining text reads more succinctly and clearly, in our opinion. As suggested below, we have added more references and included a discussion of the typical RSME found in other statistical and data-driven models, to help contextualize our results. We note that many temperature studies were already referenced in the original submission, including Caissie et al., 1998; Benyaha et al., 2007, Scott (2020), Moore (1967), Donato (2002), Bottom et al. (2011), Mayer (2012) and many others. The more compact methods section helps highlight these references better.

Overall, the model results and analysis substantiate the manuscript's conclusions about changing temperatures in the Willamette river over time. Finally, the digitization of historical observations since the mid-19th century adds value as data are sparsely available for these time periods, but are not yet archived for public use, detracting from the overall impact of the paper.

We agree that the archival measurements are valuable, and one of the potential broad impacts of this research effort. They are potentially valuable for future process-based modeling efforts, and for comparisons with other systems. Our intention has always been to make the measurements available, pending the review process. They are now available: https://doi.org/10.15760/cee-data.06

Specific Comments:

1. The authors provide extensive detail about the specific climate and long-term changes in the region. Because the detail of different regions is so extensive, it is difficult for a reader who is not familiar with the study area to distill the historical changes of the region to understand how system impacts have changed over time.

This is a good point. We aimed to be thorough and to show that much is already known about land-use changes, but may have inadvertently obscured the big picture. To understand why water temperatures are shifting, we still think it is important to show how, why, and when shifts occurred to the riparian corridor, the geometry of the river, and water surface area. But we have now moved most of these details to the "Methods>Study Area" area section or to a supplement, and focus the introduction more on the 'big picture' and scientific relevance (see below and the comments to the other reviewer). This way, readers who are interested in the detail can find it, but it is not essential to the motivation and 'flow' of the paper.

A more concise description would help the reader to focus on key points which are integrated into the model methodology. Additionally, contextualizing the changes in the Willamette river with other river basins in the continental U.S. (e.g. other basins with similar snowmelt influence) in the introduction (rather than just referring to similar studies in the discussion) could help a broader audience understand the study area before getting into methods and results.

We have simplified and reduced the size of the introduction and have rewritten several paragraphs to more clearly frame that the major changes to flow regime, air temperature, and land use that have occurred in the Willamette have also occurred elsewhere. As described above, we have also made the Methods section more concise.

1. Section 2.3: While the discussion of the advection-diffusion equation provides important information about how physical understanding can help inform accurate statistical model architecture, the in-depth analysis detracts from the reader's ability to understand the equations used for the final statistical models that are run and used for reporting results. Shortening sections 2.3 and 2.4 and/or highlighting what equations relate to the final chosen models will greatly help the readability of the methods

We have greatly simplified section 2.3 and 2.4 and moved some of the material to a supplement. We believe the shortened section is much more readable and understandable.

1. Lines 509-510: The authors refer to the "total of 8 statistical models" that are developed (listed in Table 2). When the reader refers to Table 2, however, there are 7 different stations named, implying that there are not 8 but 7 different models. Furthermore, it is difficult to decipher the governing equations for these models, and if the only major difference is the data sources and time periods used for each. Clarification would help greatly in this section of the manuscript.

Thanks for pointing out this potential source of confusion and for catching the typo about the number of models (originally there were 8, but we removed a model based on 2000-2015 Vancouver measurements because the station was moved after 1966, leading to bias). We have also clarified which equation we use for all the models. Indeed all models use the same basis function and they diffes primarily in data source and calibration time period. We also now explain our naming convention, which should emphasize that the main difference is the data and time period used. The new description reads:

"A total of 7 statistical models are developed from Equation 7, using data from the 19th century (1881–1890), mid-20th century (1941–1952), and modern period (2000–2015) (see Table 2). The models differ in the location of air temperature data and time period used. These three calibration periods were chosen based on available data; they approximate (nearly) pre-development conditions, pre-flood control conditions, and modern conditions. The models are named based on the first year of calibration data and the first letter of the meteorological station used; for example, 1941V and 1941D are models trained with 1941–1952 data from Vancouver and Downtown Portland, respectively (Table 2). "

1. While there is extensive discussion of how the chosen statistical model was derived, there is no noted comparison to other statistical stream temperature methods making it difficult to put this model in the context of previous statistical models which are also derived from physics-based equations.

Thanks for noting this oversight. We have added the following sentence to the third paragraph of results section 4.1:

"Our results compare well with traditional linear regression and stochastic models, which have reported RMSE of ~0.6–1.9 ^oC, depending on model type, river size and location, and averaging period (e.g., Caissie 1998; see also review by Benyahya et al., 2007 and references therein). More recent statistical models, including air2stream (Toffolon and Piccolroaz, 2015) and machine learning approaches (e.g., Fiegl et al., 2021), report RMSE of 0.5-1 ^oC on a daily scale, similar to the results presented here (Table 2). Results are also comparable to numerical models that generally have an RMSE <1 ^oC (e.g., Dugdale et al., 2017).

Adding information on how this model compares to other statistical models (such as air2stream, ARIMA models etc.) would help give a better understanding and context to readers familiar with stream temperature modeling, and help make these methods more applicable to other systems.

We agree that framing our model skill with respect to other models and model types is important, and have added the sentences above. However, a detailed comparison to other model types and approaches is beyond the scope of this effort, and we have therefore kept our comparison general, since more specificity would also require introducing/explaining those models.

1. The authors mention 8 different models (line 509), however, the text does not clearly explain what differentiates these 8 different models. iTable 2 includes 7 different models. Including a table of the different models, and some measure of comparative accuracy for the summer vs. winter sub-models would be helpful.

Thanks again for catching the typo about the 8 different models. This has been changed, and we have expanded our explanation of how the models differ (see above). Note that we already included both daily and monthly-averaged RMSE in Table 2 for the summer, winter, and annual sub-models, so there is already some measure of comparative accuracy. However, we have rearranged the table so that the model name is on the left hand side of the table. In this way, it is now more obvious that each row refers to a model that is calibrated to a different location and/or set of years.

1. Line 969: The manuscript states that “The water temperature data used in the research is available upon request, and will be uploaded to a data repository upon acceptance of the manuscript.” With other meteorological and flow data available from other sources. Because the procurement of data used in this study is a large value add, publishing all data in a corresponding data package together (if possible) will greatly improve the open use of data from this study.

Our intention has always been to make the measurements available, pending the review process. These measurements are now available: https://doi.org/10.15760/cee-data.06.

Technical Corrections:

Line 245: Missing space– should read “1881– 2021 record”

This has been fixed.

Table 2: 10th column (RMSE Winter Calibration) Does not properly align/format with 11th column should be fixed.

Thanks for catching! This has been fixed.

Figure 3: the x-axis of the bottom right-hand figure (d) is cut off and should be fixed.

Fixed.

Citation: https://doi.org/10.5194/egusphere-2022-793-RC2