Earth System Science Data

General comments: The core idea and content of this manuscript focus on applying machine learning algorithms to improve the prediction accuracy of heat stress in dairy cows, especially the prediction of temperature-humidity index (THI) under the background of climate change. It enables better understanding and predictions to mitigate the impacts of climate change on livestock farming, thereby aiding in the formulation of strategies to combat climate change. However, there are some problems that have to be addressed in this manuscript.

## 1. Introduction

Specific comments:
The authors did not clearly define the differences and novelty of its research compared to existing research. Why do you use the machine learning algorithm to downscale climatic data? What is its advantage in downscaling compared to other methods? The manuscript does not show the necessity of the method.

## 2. Methodology

Specific comments:
1) There are not enough details about model training. For example, when training model, how to build one-to-one relationships between the hourly training data from ERA5 dataset and the daily inference data from both ERA5 and NEX-GDDP CMIP6 data? These descriptions are critical because they will directly decide how to realize the downscaling from daily climate variables to hourly THI values.

2) Why do you use 12 climate models? What is the meaning of selecting so many climate models?

Other comments:

(1) I suggest to provide more explanations about performance metrics such as MAE, MSE, and how they can help assess model performance.

(2) Some contents in section Results should appear in Methodology. For example, how to test and valuate the model; needed computing resources; size of trained samples.

## Results

Specific comments:
(1) I suggest a discussion on the uncertainty of results and how to quantify this uncertainty.

(2) The manuscript does not discuss the credibility and potential limitations of long-term predictions that extend to the end of this century.

(3) I suggest clearly identify the limitations of the model, the impact of these limitations on the results, and discuss possible directions for improvement.

(4) For the results from Figure 4, the conclusion was made "Deviations from the ground truth are evident in various regions across the globe at these specific hourly time points". The analysis in Figure 4 is to further assess model performance and identify potential systematic errors. I am wondering what the potential systematic errors are from the evident deviations between the predictions and the ground truth. I suggest that the authors make some discussions about that.

(5) Authors indicated the absence of systematic errors because the mean difference between ground truth and prediction reveals minimal deviations from zero (Figure 5). This conclusion is arbitrary. It should be effect of temporal scale (grain). At finer temporal grain (such as at hourly grain), more fluctuations can be exhibited, while they become smoother at coarse temporal grain (such as at yearly grain) because delicate variations are masked.

Some descriptions are not clarified. For example,
(1) "Across these examples, the outputs from all three models closely followed the real THI fluctuations during the 10-day periods shown." This result is suitable for just two hot climate regions for the two bottom panels in Figure 2. However, in the text, it is not clarified.

(2) In the text, it is shown that "… at these specific hourly time points", however, in the caption of Figure 4, it is shown that "…for three randomly selected time points". I am wondering whether they have the same meaning.

(3) In Figure 5, it is not clear how to average the THI over the three-year period. Are all the hourly THI values (at all the time points) in the validation set (February 2018 - December 2020) were averaged?

Other comments:

It is not necessary to give the explanations in the whole names again and again after their simplified expressions have been made in the preceding text, for instance, Mean Absolute Error (MAE) and Mean Squared Error (MSE).

**Conclusion**

"The implicit assumption in this approach is that the diel cycle of the THI does not alter significantly under climate change scenarios." This assumption is not provided over the entire manuscript. Why does it appear in section Conclusion?