Geoscientific
Model Development
Discussions

Interactive comment on "Climate projections of a multi-variate heat stress index: the role of downscaling and bias correction" *by* Ana Casanueva et al.

Anonymous Referee #1

Received and published: 22 March 2019

The manuscript investigates the role of downscaling and bias correction to capture the climate change signal of multi-variate heat stress index, by comparing GCM and RCM simulations at different spatial resolutions. The corrected heat stress index (WBGT in the shade conditions) is calculated from air temperature and dew point temperature, which were separately corrected using two BC methods; a) ISIMIP (parametric quantile mapping) and b) empirical quantile mapping. The bias-correction methods applied in the manuscript are not newly developed techniques. However, the application on a multi-variate index and the evaluation of the corrected index are a needed task in the topic of bias-correction on climate model simulations. The overall manuscript is well written, and most of the figures included are clearly stated.

Specific comments:

- Page 3, line 19: More explanation on "intensity-dependent biases" would help of the quantile mapping. Can you provide a reference for the term?

- Page 5, line 3: I am curious about the reasoning of using daily 'mean' dew point temperature, instead of using daily maximum dew point temperature, to calculate the daily maximum WBGT.

- Page 9, line 30: I like joint distributions of two input variables in Fig 5 to understand the characteristics of joint dependency for climate simulations better. However, it would be good to see some statistics like the correlation to show dependence between two input variables, maximum temperature and dew point temperature. In Fig 4d, it seems there exists a stronger negative correlation between two variables in the raw CCLM, compared to the correlation in Obs. If the negative relationship is stronger on extremes (e.g., above 95th percentile) of two variables, that might bring inaccurate bias adjustment in QM, leading to the underestimated negative biases?

- Page 11, line 24-25: I don't know how the conclusion is drawn. By comparing average Perkins scores?

- Page 15, line 6: If I understand correctly, you used a single ensemble (r1i1p1) of HadGEM2-ES. Do the biases relate to the biases across ensemble runs? If we use more ensemble members of the HadGEM2 simulation, do we expect the smaller biases?

- Fig 1a: I am a bit confused. Are the CDFs of the (historical and future) RAW from RCM? Or GCM?