

Review of “Three novel copula-based bias correction methods for daily ECMWF air temperature data” by Fakhreh Alidoost et al.

1 General comments

The idea of this paper is to use copulas in order to capture the dependence between site-measured and remotely-sensed data. Yet, there is a major problem with the approach suggested by the authors as described in this paper. Unless wrongly understood, marginal distributions and copulas are fitted on the basis of only four data points, as for each time step a copula and marginals are fitted on data of 4 stations. How valid can such an approach be?

Furthermore, notations are not straightforward and unambiguous. For instance, u_1 and u_2 are sometimes used to indicate a value (e.g. in equation 4), sometimes a variable (e.g. in equation 8). Similarly, p_{u_1} is sometimes used to indicate a variable (equation 4), sometimes a function (e.g. p.5, line 4: The conditional quantile $p_{u_1} \dots$). This hinders the reader to understand the approach followed by the authors.

2 Specific comments

- I would suggest to use capitals to indicate variables, and small letters to indicate values.
- I would suggest to remove the superscript t in the naming of the copulas. This is confusing, as it is only fairly late in the paper that it becomes clear where the t comes from.
- Why not using the general notations e.g. $C_{12}(u_1, u_2)$ and $F_{2|1}(u_2|u_1)$ for copulas and conditional distributions, respectively?
- What is the definition of a bivariate conditional copula? By taking the derivative of a copula to one of its variables, one ends up with a conditional distribution function, but is this function a copula?

- Equation (4): u_1 and u_2 are treated as values whereas they are treated as variables in Equation (8).
- line 16, page 3: s refers to a location: there is no s in Equation (4).
- line 21, page 3: Empirical marginal distributions u_1 and u_2 : not clear. Are u_1 and u_2 marginal distributions here?
- lines 24-26, page 3: not clear, what is meant.
- line 27, page 3: I doubt that fitting a polynomial to 4 data points to obtain a marginal distribution is correct. Furthermore, how is this done in the fitting as it requires a positive slope everywhere and its maximum value should be one.
- line 5, page 4: conditional quantile p_{u_1} : not clear what is meant.
- line 22, page 4: F_1, F_2 : not explained
- Equation (7): notations are not clear
- Equation(8): $c^t(u_1|u_2) = c^t(u_1, u_2)$: not clear why.
- line 5, page 5: marginal quantile \hat{u}_1 : not clear
- line 7, page 5: empirical marginal quantile \hat{u}_1 equals u_2 or $1 - u_2$: not clear why.
- line 20, page 5: generating independent variates u_2 : isn't u_2 known? This is after all the conditioning value, or not?
- paragraph 2.3.4: this paragraph is not clear
- page 6: the explanation of the methods BCQM-type I and BCQM-type II is not clear. What and how many copulas are fitted, what is then done with these copulas. What is meant with "Substituting the quantiles p_{u_2} for p_{u_1} ".
- page 6: How can the influence of land cover be captured by taking R into account? If the elevation (value of e^s) is kept constant, data with an equal value of R are located on a circle.
- equations (13) and (14): What is $C_{u_R}^t$?
- Wouldn't it be better (if enough data are available) to fit a multivariate copula or a vine to take into account the dependences between all these variables?
- page 7: Quantile search: I don't understand which parameter the authors want to optimize such that the fitness function is maximized.

- page 7: how are genetic algorithms applied here?
- lines 14-16, page 7: the authors refer to previous Sections (2.2, 2.3.5.1, 2.3.5.2) for an explanation. Yet, I don't see any explanation in these Sections.
- Equation (21): T?
- line 4, page 8: A minimum value of the error score indicates for the minimum SMAE. Not clear.
- lines 18-20, page 8: not clear
- line 24, page 8: not clear
- Section 3: more info is needed on the data (resolution, correlations, ...)
- Section 4.1: how do you define outliers?
- line 28, page 9: "As can be seen": this is not clear to me.
- line 2, page 10: there was no need to remove the outliers: what is the purpose of defining outliers if they don't need to be removed?
- lines 16-19, page 12: Is it necessary for the research carried out that observations are separated day by day, in order to remove autocorrelation. In this research, one is only interested in the dependence between observations of the same day. Hence, no dependence between data values at certain time lags is to be taken into account.
- lines 34 page 12, lines 1-2, page 13: not clear
- lines 11-12, page 13: "The new methods are beneficial for the local refinement..." : this is not shown in the manuscript
- lines 13-14, page 13: "The new methods are advantageous ..." : this is not shown in the manuscript