

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2020-420-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2020-420

Anonymous Referee #1

Referee comment on "Convolutional conditional neural processes for local climate downscaling" by Anna Vaughan et al., Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2020-420-RC1>, 2021

The paper applies a convolutional conditional neural process (convCNP) to the task of statistical downscaling, in particular of temperature and precipitation.

Methodology falls within the realm of probabilistic deep learning, allowing to combine a bespoke statistical model with deep learning framework.

Extensive experiments are performed to compare the convCNP model to a range of benchmarks, with promising results.

While there are a number of modelling choices and experimental details in this paper that would benefit from a much better motivation and explanation, this is undoubtedly a good contribution to the literature.

28-29 "This is based on the assumption that while sub-grid-scale and parameterised processes are poorly represented in GCMs, the large scale flow is generally better resolved."

-- Can this assumption be elaborated further or an appropriate reference included?

83-84 "This is assumed to be Gaussian for maximum temperature and a Gamma-Bernoulli mixture for precipitation"

-- describe the reasoning for these particular choices?

99 should say "distributional parameters θ at each target location"

118 "parameterise a stochastic process over the output variable, in this case either temperature or precipitation"

-- Can it be made explicit how exactly does the model proposed here result in a stochastic process over the output variable since this is obscured by a large number of model components, and if there is anything in the proposed model that is in fact "nonparametric"? The overall model reads as a complicated, but a parametric model for the parameter of the conditional distribution. I found this paragraph confusing since it appears to contrast the proposed method to "parametric approaches". Also comment on what are the advantages of having such a stochastic process in this specific context?

154-155 "The VALUE experiment protocol does not specify which predictors are used in the downscaling model (i.e which gridded variables are included in Z)."

-- Can you clarify if all the baselines in comparisons use the same set of predictors? Do any of them use topographic predictors? If none of them use topographic predictors, would it be more fair to compare convCNP without topographic predictors to the baselines, and then consider the (additional) improvement due to including topographic information?

Fig.10 Plotting KDEs is highly problematic here due to bounded domain. It would be much more informative to simply plot a histogram of the evaluations of CDF at the true values. Perhaps perform a KS test to check if distribution is uniform in "well calibrated" cases?

Remarks on notation:

- * Some mentions of variables, e.g. f and Z are not in the inline math mode,
- * Certain variables follow inconsistent notation: sometimes boldface, sometimes not (x,h) ,
- * Arguments of θ swapped places in the figure,
- * ϕ_c takes x as input, then it doesn't,
- * Display mode equations are missing punctuation at the end,
- * Inconsistent usage of log vs. ln. Also, use \log or \ln in math mode.

Typos:

46 'avances'

83 'an distribution'

spearman -> Spearman

262 'application on'->'application of'