Response to Referee #1

RC1: 'Comment on essd-2025-804', Anonymous Referee #1, 15 Apr 2026

The authors describe SMHIGridClim, a 2.5 km resolution gridded climatology dataset of hydrometeorological (near-)surface parameters of Fennoscandia.

It is DOI-index and hosted on a popular data repository.

The dataset is stored as CF-1.7-compliant NetCDF files in different temporal aggregation periods, which makes it easy to use and to integrate into existing workflows and data analyses.

The authors perform cross validation between background fields and the analyzed grid data, an impact analysis of various processing steps on the resulting grids, and a very concise comparison with other existing gridded datasets for the region.

This manuscript gives an overview of the input data, background fields, interpolation methodology, validation/evaluation, and a discussion about strengths and limitations of the described data.

While I feel that the dataset itself is certainly valuable, much work needs to go into improving the corresponding data description paper.

- *We thank the reviewer for these positive comments and for recognizing the value of the dataset presented in this paper.*

My major concerns with the manuscript are: a terse description of the methodology behind the dataset which limits the ability to gauge the data characteristics; limited independent evaluation of the data quality; numerous inconsistencies in naming, which make it hard to follow the content; not very well designed figures, both in terms of legibility and information content; and non-consideration of the ESSD author guidelines for English, figures, tables, and mathematical notation which puts a lot of effort on the reader to take in the presented information.

- *We agree with the reviewer on the importance of independent evaluation. However, constructing a multi-decade, high-resolution climatology presents a well-known constraint: to ensure maximum product quality, we must utilize all available in-situ observations rather than permanently withholding a subset. Furthermore, high quality, high-resolution independent spatial datasets (from radar or satellite) do not exist for most parts of this period. To resolve this, we employ a "leave-one-out" cross-validation framework within the gridpp software. This approach ensures strict statistical independence:*

1. *Out-of-sample testing: For every single analysis time step, when calculating error metrics at a specific station, that station's observation is strictly excluded from the optimal interpolation. The grid point value is reconstructed entirely from surrounding stations and the first-guess field.*
2. *Robust error metrics: The summary statistics presented for all parameters in Figure 9 (RMSE, standard deviation, and mean difference) are derived exclusively from these independent cross-validation estimates, representing a true independent error assessment.*

This strategy allows us to utilize our observational network to its full capacity while strictly adhering to the principles of independent validation. We have revised Section 4.2 on cross validation, to clarify the independent nature of our cross-validation workflow.

Furthermore, we have carefully addressed the comments concerning the text, formatting, and figures in the detailed responses provided below.

Next to these major points, I have the following detailed comments on the manuscript:

L12: gridpp is sometimes written in full uppercase (GRIDPP, e.g., Fig. 1). I suggest chosing one notation (ideally following the developers) throughout the paper.

- *Fig. 1 have been edited according to suggestion, using small letters for gridpp.*

L17: Please explain MEPS acronym.

- *The MEPS acronym was removed from the abstract. MEPS (Meteorological Ensemble Prediction System) is now introduced in section 2.3*

L51: It is important to stress the difference between grid resolution and resolution of the input data, i.e. the information content in the data. The authors rightly note the low station density for the NGCD dataset and this should be done for each of the discussed data products.
I suggest to name the grid resolution of the discussed datasets "horizontal grid resolution" to make it clear that even though the grids are produced at these scales, the information content may vary.

- *We agree, and changed throughout the text, now "horizontal grid resolution is used".*

L78: two-meter temperature is sometimes referred to as tas and sometimes as T2m. I suggest to adopt a consistent naming scheme (also for the other variables). Also two spellings (2m temperature, Fig. 1; 2 m temperature) exist, this should be consistent.

- *We apologise for this mistake, the abbreviation of the variables in the text was changed in the writing process in order to be consistent with the acronyms used in the uploaded data files, but we forgot to edit the varible names in the captions, they were inserted just before submission. Variable abbreviations are now standardised throughout text, tables and captions (tas, tasmin, tasmax, hurs, pr snd).*

L88: Sometimes AE is used, sometimes BE. This should be consistent throughout the article (https://www.earth-system-science-data.net/submission.html)

- *Spelling is now regularised to British English throughout (analyse, optimise, normalise, localise, recognise, centred, metre, etc.).*

L104: Please provide access date.

- *Access date has been inserted for all webblinks in the revised manuscript.*

L106: Is there a source for this equation?

- *Equation 1 is now referenced as "the Magnus-Tetens approximation formula", relating saturation vapor pressure to temperature"*

L107: Please see the guidelines for mathematical notation.

- *Equations in the text have been edited according to guidelines using the equation editor i MS Word .* https://www.earth-system-science-data.net/submission.html#math

L108: Td2m should be in math font face to be consistent with the equation.

L109: Same for tas and hurs.

- *Font of variables have been changed accordingly in revised manuscript*.

L109: On +5 degC: Remove whitespace before magnitude, add whitespace between magnitude and unit.

- *Changed accordingly (+5 °C).*

L112: Please move the BUFR acronym explaination to previous sentence.

- *BUFR acronym now expanded at first use.*

Figure 2:

The resolution of the figure is relatively poor which, together with the small font size, makes it difficult to parse the text and distinguish the marker symbols.
I also suggest checking the Figure using a color vision deficiency simulator (see https://www.earth-system-science-data.net/submission.html) to verify that the color scheme used is inclusive.

Concerning the legend, please increase the font size, align the numbers and add a unit directly to the magnitudes.

- *In the revised manuscript the figure has been changed according to suggestions.*

The variable naming is not consistent with the text.

- *Text and caption variable names are now consistent throughout the text.*

Figure 3:

The font size is too small. The y-axis should be aligned between the panels (if the range varies too much, please note this in the capition).
The panels can be arranged in a 2-by-2 grid to save space. The y-axis label is not correct. It should be "Number of observations per month per square km" as stated in the capition.

- *In the revised manuscript the figure has been changed according to suggestions.*

L148/L150: Please provide access date.

- *Access date is now inserted for all webblinks in the revised manuscript.*

L162: Is this standard deviation computed per time step or station-wise over time? Please provide more information as this make a difference. Computing the standard deviation per time step over all station differences results in a too optimistic bound for stations with large noise level.

- *The standard deviation is computed per time step over all station differences. As the reviewer correctly points out, this can lead to an optimistic error estimate, particularly for stations with higher intrinsic noise levels.*
  An important consequence of this choice is that the resulting error bounds may be too tight for such stations, potentially leading to an overly restrictive rejection of observations. A station-specific error characterization could mitigate this issue. However, making these error estimates dynamically dependent on the situation (as is currently done per analysis time) is non-trivial.
  Alternative approaches, such as prescribing static or seasonally varying (calendar-based) error estimates per station, could also be considered, although these come with their own limitations, including reduced adaptability to transient conditions.
  We thank the reviewer for highlighting this important point and will investigate improved strategies for handling station-dependent uncertainty in future versions of the analysis.

L178: I do not think that MEPS was explained so far.

- *MEPS acronym is now explained at first use.*

L179: horizontal grid resolution.

- *We changed to “horizontal grid resolution” throughout the text.*

Figure 4:

See comment on Fig. 2 concerning color vision deficiency.
Please move the explanations to a proper legend rather verbal descriptions in the title.

- *In the revised manuscript the figure has been changed according to suggestions.*

UERRA should be named UERRA-HARMONIE as in the main text.

- *‘Uerra-Harmonie' corrected to 'UERRA-HARMONIE' in the caption;*

L204: "surrounding 4x4 neighbouring points" could be explained more clearly, maybe by adding "(the 16 closest points/nearest neighbours)". This is done in the next paragraph anyway, but would be better here.

- *The section was rewritten accordingly.*

Figure 5:

Typo "Exampel", UERRA should be name UERRA-HARMONIE

Titles: please provide date/time in a proper format (e.g., ISO 8601). Please, also label each panel properly (https://www.earth-system-science-data.net/submission.html).
Colorbar: colorbar label missing or cut off.
Grid point panel: please add a legend with the explanations of each point/grid. Also consider using different markers (circles and squares) for UERRA-HARMONIE and SMHIGridClim for better readability.

- *Corrected: typo 'Exampel' -> 'Example'; 'UERRA' -> 'UERRA-HARMONIE';*
- *In the revised manuscript the figure has been changed according to suggestions.*

Figure 6:

Suggestion for renaming x-label: Time -> Time of Day, day number -> Day of Year
Please properly label panels.
Choose either "weights" or "weighting functions", not "weight functions".

- *In the revised manuscript the figure has been changed according to suggestions. .*

L231 - L234: Please see the author guidenlies for math notation.

L238: Please see the author guidenlies for math notation.

changed acordingly.L255: Please provide access date; python -> Python

- Changed *'python' -> 'Python'. Access date was added to the revised manuscript.*

L257: optimal way -> statistically optimal way

- *Done: 'optimal way' -> 'statistically optimal way'.*

L258: bi-linear -> bilinear

- *Done: 'bi-linear' -> 'bilinear'.*

L261:

It might not be clear to everyone how the background and observation covariances are handled in gridpp. While this sentence hints at how this is performed, a more precise explanation would be helpful, maybe even with an equation.

Since the parameter determination and -interpolation of the covariances is discussed in the following paragraphs, it should be clear to the reader that this refers to a combined covariance information in the form of a single structure object and covariance ratios.

A sentence about the assumptions of the observation error covariance is also helpful (are they spatially correlated or not?) to understand the following explanations.

- *We agree with the reviewer that the description of the structure function and the associated error covarances is unclear and needs to be revised. We've now written a more detailed section on this following a notation in papers previously describing a very similar setup.*

L263 - L270: Please consistently highlight software methods and parameters in the text. What is the justification for using the default setting for decorrelation length? Did you conduct a sensitivity study?

- *We reformulated this paragraph to make it clear why we used the default setting for the maximum influence radius, corresponding to 3.64 times the horizontal decorrelation length. This choice implies that observations with a correlation smaller than approximately exp (−(3.64)^2)≈10^−6 are excluded. In practice, this is a very permissive threshold, meaning that nearly all observations with any meaningful correlation are retained.*
  The primary purpose of this parameter is to limit the number of assimilated observations in regions with

very high observation density, thereby reducing computational cost and potential redundancy. However, this situation is not relevant in our case, as the observation network is relatively sparse. As a result, the use of the default setting effectively does not impose a restrictive cutoff and is unlikely to significantly influence the results.
We did not conduct a dedicated sensitivity study for this parameter, as its role in our setup is limited. Nevertheless, we acknowledge that such a study could be useful in more observation-dense configurations or in applications where localization plays a stronger role.

L273: It would be interesting to see the results of the leave-on-out cross-validation. I feel like this should be a core validation metric of the dataset to really give users confidence in the data quality.

- *Indeed it is and the result of the cross validation for all parameters is what is shown in Figure 9. We have revised Section 4.2to clarify and emphasise this.*

L275: spatial scale -> spatial structure

- *Done: 'spatial scale' -> 'spatial structure'.*

L277: It may be helfpul here to reiterate about which structure function you are talking about here.

- *We agree and have clarified the text in Section 3.2. We now explicitly specify that we use the Barnes correlation formulation to define the spatial structure function (the background error correlation), which remains invariant to station placement. To make the mechanics clearer, we have re-written this section using standard notation to reference the underlying equation relating the error covariances of the first guess and the observations. We explicitly clarify that the structure function only dictates spatial correlations, whereas the error variance ratio is what dynamically adjusts the relative weighting of the background field versus the observation network as station density changes over time.*

L279: Please provide more information on the chosen splines (degree, nodal points, ...) as this determines the smoothness of the interpolated parameter time series.

- *We have expanded Section 3.2 to provide more technical detail on the spline setup. The parameter curves presented in Figure 7 were constructed using a quadratic smoothing spline (degree k=2) via the scipy.interpolate.UnivariateSpline framework. Rather than manually enforcing fixed nodal points, the number and locations of the internal knots (nodal points) are determined adaptively and automatically by the underlying FITPACK algorithm. The algorithm optimizes knot placement to satisfy a global smoothing constraint of s = N * 10^9 (where N is the number of data points). Outside the domain, a constant boundary extrapolation was used (ext=3)*

L283: Please see the author guidenlies for math notation.

L284: yr_k= ... should be properly formatted.

- *We have changed accordingly in the revised manuscript.*

L287: "the performance of the analysis at different times": How was this determined?

- *We apologize that this was not sufficiently explained in the manuscript. The added value of temporally varying parameters was evaluated using the same leave-one-out cross-validation framework employed throughout the study for parameter optimization and assessment of GridPP analysis performance.*
We have clarified that the reported performance improvement refers to improved cross-validation scores relative to temporally constant parameters.

L288: Is a homogeneous (or more precisely stationary in time) background error covariance justified? The argument about varying observation density and -quality can be made for both the background field and the observations thus requiring temporally-varying covariance information anyway.

- *We agree with the reviewer that, ideally, the background-error covariance model should represent the underlying atmospheric process and remain stationary in time. However, the quality of the background field itself may vary over the study period. Earlier periods generally contain fewer observations in the data assimilation systems used to generate the background fields, which may lead to larger background errors*

than in more observation-rich periods. Consequently, temporal variation in the optimal OI parameters may reflect not only changes in the observation network and observation quality, but also changes in the background-error characteristics. This provides additional motivation for investigating temporally varying covariance parameters, even though a stationary covariance model would be preferable from a theoretical reanalysis perspective.

L292: "yielded relatively constant parameter values": How large was the variation?

- *We agree that this statement would benefit from quantification. Although the optimized parameters varied over time, the analysis performance was generally quite insensitive to substantial parameter changes, suggesting that part of the variability may reflect optimization uncertainty rather than a strong physical signal. As the intermediate optimization results were not retained, we are unfortunately unable to provide a quantitative estimate of the variability.*
  We have clarified this limitation in the manuscript and noted that future work could investigate regularization approaches to constrain unnecessary temporal variability.

L296: "gridpp.relative_humidity": See comment on highlighting software functions and parameters.

- *We have now highlithed software functions using font "Source Code Pro".”*

L297/L298: Please note here, that using dewpoint temperature implies a non-linear mixture of relative humidity and 2m temperature errors is diagnosed.

- *We agree with the reviewer that relative humidity errors obtained in this way are the result of a nonlinear combination of errors in the analyzed temperature and dew-point temperature fields. Our intention was not to imply that this removes all error-related complications, but rather that it avoids performing OI directly on a bounded variable whose error distribution departs more strongly from normality. The analysis is therefore performed on tas and Td2m, after which relative humidity is computed diagnostically from the analyzed fields. We have clarifed this point in section 2.2 in the revised manuscript.*

Figure 7:

Please properly label panels according to the author guidelines
Provide y-axis labels for each parameter with units.
Parameter descriptions can then be removed from the caption.

- *In the revised manuscript the figure has been changed according to suggestions. .*

Table 3:

Column headings should be horizontal and vertical correlation distance.

- *Changed accordingly.*

Please right-align numbers and provide consistent number of (significant) digits.

"snowd" should be snd or better: "Snow depth (snd)", similar for precip.

- *Changed accordingly.*

L333: Section numbers are out of order.

- *Cross-validation heading was renumbered 4.5 -> 4.2 and its cross-reference*

Figure 8:

Please add appropriate colorbar labels and panel titles.
The variable naming is not consistent with the text (Tn, Tx have not been introduced, RR should be pr maybe?) rh2m should be rhurs (maybe?)

- *Text and caption variable names are now consistent* throughout the text.

Figure 9:

This figure is very hard to parse and generally not well designed. Some suggestions to improve it:

- Color scheme: perform color vision deficiency simulation and adopt accordingly
- font sizes: far too small, figure may be split into two individual figures with less panels to increase the size of each subfigure a bit
- labels: everything is labelled as Error, but different metrics are shown. Please provide proper labels
- abbrevations: avoid "nbr obs" and the other overly terse abbrevations in the legend, the reader should not have to guess
- UERRA-HARMONIE is called UERRA-HARMONIE throughout the paper and should be here as well.

- *We apologise for the inconvenience reading this figure. In the revised manuscript the figure has been changed according to suggestions. The figure was split into three separate figures, fig. 9-11.*

L391: 95th

- *corrected: '95:th' -> '95th'.*

L398: bi-linear -> bilinear

- *corrected: 'bi-linear' -> 'bilinear'.*

Figure 10: Same as Fig. 8

Figure 11: Same as Fig. 8

- *We have change the figures according to suggestions in the revised manuscript.*
- *'[˚C]' -> '[°C]' in the caption; caption variable names now match the dataset naming.*

L445: Is this the wrong link? It leads to seNorge2

- *Link replaced by: https://doi.org/10.7910/DVN/ZFZL6K*

L447 - L453: This paragraph should be expanded and moved to section 4. Validation is a key component of the dataset description and should be given more weight, especially when independent datasets are used.

- *We moved the validation part into a new section 4.3: Comparison with other Nordic gridded datasets.*

L456 - L457: This is about the grid resolution and should be communicated as such. As you state in the next paragraphs the effective resolution of SMHIGridClim is far larger than 2.5 km, so the spatial information content of these datasets might not be too far apart. Please indicate this properly.

- *We thank the reviewer for this comment and agree that the distinction between grid spacing and effective resolution should be made more clearly. Although SMHIGridClim is provided on a 2.5 km grid, the spatial information content is limited by the effective resolution of the underlying background fields and the spatial density of the observations (figure 3). We have revised the text to explicitly distinguish between grid resolution and effective resolution and to clarify that the effective resolution is substantially coarser than the nominal 2.5 km grid spacing.*

Table 4: This table is not is not referenced. Also there is a mix in unit notations, e.g., [K] in figures (K) here.

- *Table 4 is now referenced in the text. Units are now expressed with [] in text and figures.*

L473: UERRA -> UERRA-HARMONIE

- *Changed 'UERRA' -> 'UERRA-HARMONIE'.*

L486: UERRA -> UERRA-HARMONIE

- *Changed 'UERRA' -> 'UERRA-HARMONIE'.*

Citation: https://doi.org/10.5194/essd-2025-804-RC1

Response to Referee #2

RC2: 'Comment on essd-2025-804', Anonymous Referee #2, 20 Apr 2026

This paper presents a nice and valuable application of downscaling of the UERRA-HARMONIE regional reanalysis that also incorporates observations by optimum interpolation to derive a 2.5x2.5km gridded dataset.

The paper has a good structure, is well written, and is mostly clearly presented. Even though the paper in general keeps a high standard I have a few concerns that could need some more consideration by the authors.

My main concern is the way the authors define their structure function for OI. In my philosophy this function should describe the physical process and be independent of the station density. I miss an analysis of the effect of various station densities on the structure function, e.g. by estimating it from random subsamples from the most observation rich periods. It is briefly discussed in the discussion, but I think it needs more attention. However, from another viewpoint can the approach taken by authors might be valid (but with the wrong motivation) The climate is non-stationary, and the structure function might vary over time due to climate variability. Can you comment upon that?

I also would have liked to see the real effects of different structure functions in the final grid and how sensitive the estimates are to the structure function.

- *We agree that, in the classical OI framework, the structure function is intended to describe the spatial correlation characteristics of the underlying field (or background error) and should therefore be independent of station density. In our implementation, the Barnes correlation formulation defines the spatial structure function itself, which is independent of station placement. The varying station density instead primarily affects the observational constraint provided to the analysis and the estimation of the associated parameters. We agree that it would be valuable to investigate the sensitivity of the estimated structure-function parameters to network density, for example through random subsampling experiments during observation-rich periods. Such an analysis could help distinguish between changes caused by the observing network and changes related to the underlying atmospheric variability.*
  We also agree that temporal variations in the estimated parameters need not solely be interpreted as artifacts of changing station density. Since the study spans a long period, genuine changes in the covariance characteristics of the atmospheric fields due to climate variability or other non-stationary processes may also contribute. The present study does not attempt to separate these effects, and the observed variability in the estimated parameters should therefore be interpreted as a combination of potential network- and climate-related influences.
  We have clarified this point in the manuscript and highlight the separation of observation-network effects from genuine temporal changes in covariance structure as an important topic for future work.

In ch. 2.2, line 161-3 you state you are using fixed lapse rates for tas and T2dm. Isn't there a risk that under certain weather situations many observations incorrectly will be excluded because the lapse rate is non-representative? Please comment/discuss.

- *We agree that a fixed lapse rate is a simplification and may not be representative under all weather situations. In particular, strong temperature inversions, stable boundary-layer conditions, or other situations with anomalous vertical temperature gradients could lead to errors in the elevation correction and consequently increase the difference between an observation and the interpolated first guess.*
  However, the lapse-rate correction is only used to account for elevation differences between the station location and the model grid prior to the gross-error check. Observations are only rejected when the adjusted observation-minus-background difference exceeds three standard deviations of all differences, which represents a relatively conservative threshold. Therefore, moderate deviations from the assumed lapse rate are generally unlikely to result in the rejection of otherwise valid observations.
  Nevertheless, we acknowledge that under certain meteorological conditions, particularly in regions with

large elevation differences, the use of fixed lapse rates may contribute to the erroneous rejection of some observations. This represents a limitation of the current approach and could potentially be improved in future versions through the use of dynamically estimated lapse rates or model-derived vertical temperature gradients.
We have added a discussion of this limitation to the manuscript.

In the introduction I would have appreciated a more complete introduction to the methodological framework (lines 76-80) to also introduce the downscaling. Now this section appears incomplete.

- *We thank the referee for this suggestion and have added few more lines with information to this section.*

Minor comments/issues.

L.39: What do you mean by microscale? Consider to use local scale instead.

- changed: 'microscale' -> 'local scale'.

L.40: You might delete “often”.

- 'often' deleted.

L.56: Change Bazile et al.,2017a to Bazile et al., 2017.

- corrected: 'Bazile et al., 2017a' -> 'Bazile et al., 2017'.

L.101: I think frost.met.no is an API, not the archive.

- We have reformulated the section somewhat and made this clear.

L107: Eq(1). Reformulate to solve Td2m (Td2m =…)

- *Eq (1) was reformulated as requested.*

Ch. 2.1.1: How do you treat redundant data from national vs. ECMWF archives?

- *When there was overlapping data from these two sources, we used only the national data.*

Fig 2. Add national borders, that would make the connection to the text better.

- *In the revised manuscript the figure has been changed according to suggestions.*

Fig 3. Increase the font size for readability.

- In the revised manuscript the figure has been changed according to suggestions.

L.251: Panofsky and Brie, 1968 is not in the reference list.

- Reference added.

Ch3.2 Consider a different title than “Analysis with gridpp”. gridpp is the tool, SMHIGridClim is your target.

– We suggest changing the title to “Optimal interpolation analysis”

Fig 7. Add axis titles and units.

- In the revised manuscript the figure has been changed according to suggestions.

L.322: You mean south-west part of Norway? In Figure 8 you should consider presenting the difference in precipitation as a ratio instead of difference. Consider also to add legend title and units.

– *Yes this was a mistake however when changing the figure using ratio instead for precipitation these features in the figure disappear.*

L.328-329: Can these differences also be a result of changes in the observation network over time?

– *Yes this is true, as pointed out in the discussion section, SMHIGridClim suffers from inhomogeneity that can cause false trends in the dataset. The included observation network for precipitation in Norway shows*

a decrease in number in the later period. However the decrease does not seem to be focused to the specific region, rather a decrease in general over the country. Further this is a region with difficult topography to represent in model data. According to ratios shown in the homogenised dataset KlimGrid (Lutz et al, MET rapport 2023 - Endringer nedbørnormaler KiN bakgrunnsrapport) this region does not have such a strong increase compared to surrounding regions. We have added text to the section to make this clear in the revised manuscript, although changing the plot to show ratio instead, this pattern is not evident.

L.373,374: rms? You mean RMSE? Correct or explain.

- *Corrected: 'rms' -> 'RMSE'.*

Figure 9. Sorry, but this figure is completely unreadable in the version I have. Increase the size and fonts to make it readable.

- *We apologise for the inconvenience. In the revised manuscript the figure has been changed according to suggestions. Figure 9 was split into 3 separate figures (fig. 9 to 11) to increase readability.*

In conclusion; a very nice paper. A more in-depth analysis and discussion concerning the OI structure function would earn the paper a lot. Figures need to be improved. The dataset is a useful supplement to existing gridded datasets for Fennoscandia, especially since it provides more variables and higher temporal resolution than many of the existing datasets. It needs to be updated and made available until present to keep its relevance.

- *We thank the referee for the positive evaluation and constructive suggestions, which have helped improve the manuscript and provided useful insights for further work with upcoming versions of the dataset.*