Revision of manuscript **egusphere-2025-621**
"Gridded Intensity-Duration-Frequency (IDF) curves: understanding precipitation extremes in a drying climate"
Responses to Reviewer 01

Cristóbal Soto-Escobar, Mauricio Zambrano-Bigiarini,
Violeta Tolorza, and René Garreaud

August 27, 2025

We would first like to express our sincere gratitude to the handling editor, Prof. Thomas Kjeldsen, and to the three anonymous reviewers for their valuable comments and suggestions, which have helped us to substantially improve the quality and clarity of our manuscript.

We hereby submit our detailed responses to the comments provided by Reviewer 01 regarding our article entitled "*Developing Intensity–Duration–Frequency (IDF) curves using sub-daily gridded and in situ datasets: characterising precipitation extremes in a drying climate*".

The main revisions introduced in the manuscript are as follows:

1. **New title**: following a comment made by Reviewer 01, the title of the manuscript was changed from "*Gridded Intensity-Duration-Frequency (IDF) curves: understanding precipitation extremes in a drying climate*" to "*Developing Intensity-Duration-Frequency (IDF) curves using sub-daily gridded and in situ datasets: characterising precipitation extremes in a drying climate*".

2. **Incorporation of reviewers' suggestions**: We implemented several of the proposed improvements, with the aim of enhancing clarity without unnecessarily increasing the manuscript length.

3. **Revisions to figures to improve clarity and conciseness and reduce redundancy**:

    - Figure 3: Number of subplots reduced from 25 to 5.
    - Figure 5: Combines the previous Figures 5 and 7, retaining only the most representative results.
    - Figure 6: Merges the previous Figures 6 and 8, reducing the number of subplots from 40 to 8.
    - Figure 8: Integrates the previous Figures 9 and 10, decreasing the total number of subplots from 16 to 8.
    - New Figure 7: Introduced to present a scatter plot comparing intensities derived from stationary and non-stationary models.

4. **Expanded bibliography**: Additional references have been included, following the reviewers' suggestions as well as our own evaluation, to strengthen the scientific context of the study.

5. **Updated supplementary material**: The supplementary files were revised to include metadata of the rain gauge network and several new figures. Furthermore, some material originally in the supplementary section has been incorporated into the main manuscript to improve accessibility of key information. The new supplementary material can be found on `https://doi.org/10.5281/zenodo.16956066`.

In the following sections, we provide a point-by-point response to all comments raised by Reviewer 01. We hope that our detailed explanations will satisfactorily address all concerns.

REVIEWER 1

R1C0: In this study, the Authors propose an IDF model for Chile by estimating the spatial maximum precipitation intensities, in varying climatic conditions and topography, through gridded (of 5 datasets) and 161 gauged hourly precipitation data by using stationary and non-stationary Gumbel distribution models. Please see several major and minor issues that I hope they ca be of help to the Authors:

We thank all the thorough comments made by Reviewer 1. In the following paragraphs we provide detailed replies to each one of them.

R1C1: Regarding the traditional methods in the literature, please see a recent and the most advanced stochaatic framework for the stationary IDF curves with application in the entire country of Greece by Koutsoyiannis et al. (2024; `https://doi.org/10.1080/02626667.2024.2345813`) and Iliopoulou et al. (2024; `https://doi:10.1080/02626667.2024.2345814`). In there, multiple sources have been used, like re-analysis and satellite data as well as rain-gauges, and they have been spatially combined following the regional model described in Iliopoulou et al. (2024; `https://doi.org/10.3390/hydrology9050067`). I would recommend the Authors to discuss these methods and to provide the differences presented in theirs as part of their literature review.

We appreciate this comment. In the new version of the manuscript we have added the following text to the literature review of our article:

While the aforementioned methods rely on stationary or non-stationary statistical models to summarise historical data, Koutsoyiannis et al., (2024) introduced a stochastic framework that models rainfall as a random process, providing a probabilistic and theoretically grounded approach to estimating intensity-duration relationships across time scales and return periods. This method accounts for the inherent variability and uncertainty of rainfall events without relying on stationarity assumptions, offering potentially more robust estimates under changing climate conditions (Iliopoulou et al., 2024).

R1C2: In the analysis, it is mentioned that the bias-correction factors are applied to match the gridded maximum intensity values with the in situ ones by implementing the modified Mann-Kendall test for the trends; however, there are also maps with high statistical significance (as 0.05 and 0.10), which may be considered quite large. I would recommend showing just the ones with significance lower than 0.05, while the rest can be showed in a supplementary material. Also, please include in the Conclusions whether there are any trends with significance 0.01 or lower, which I would consider the most important ones to report.

Thank you for your comment. First, we need to clarify that the modified Mann-Kendall test was not applied for carrying out the bias-correction of I_{max} values (as we understood from the comment), but to investigate whether using non-stationary IDF curves is justified or not for our study area, as mentioned in Section 3.3. Second, we have included the trend maps with high statistical significance ($\alpha = 0.01$) in Section S5.2.4 of the supplementary material, and we have added those results to our discussion:

[...] For all the gridded datasets, supplementary material S5.2 (Soto-Escobar, 2025) contains maps showing Kendall's τ values statistically significant at $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.10$; as well as maps with all the computed trends independent of their statistical significance. Although the trend areas are somewhat smaller at lower significance levels, such as $\alpha = 0.01$, the spatial distribution of areas with statistically significant trends remains the same. In general, for all durations the results of the trend analysis were similar between ERA5 and ERA5-Land, as well as for IMERGv06B and IMERGv07B. In addition, the trends obtained for ERA5 and ERA5-Land were also similar when using 20 (2001-2021) and 40 (1981-2021) years of data length, although with slightly smaller areas with significant trends in the latter case.

Also, the sixth point of our Conclusions section was rewritten as follows:

In Central Chile, all precipitation products revealed either significant decreases in I_{max} (at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$) or no detectable trends. While the extent of significant areas was smaller at the stricter level

($\alpha = 0.01$), their spatial distribution remained consistent across thresholds. For ERA5 and ERA5-Land, these declining trends were evident in both 1981–2021 and 2001–2021, whereas the other products were only available for the shorter 2001–2021 period. In contrast, regional differences emerged outside Central Chile: in the Near North and Far South, IMERGv07B displayed localised increases, ERA5-Land showed mostly decreases or no trends, and CMORPH-CDR consistently indicated widespread declines.

Finally, we kept Figure 4 showing trends significant at $\alpha = 0.05$. We hope we have correctly understood the last part of this reviewer's comment.

R1C3: Regarding the sentence *In addition, our results show that Imax reaches its maximum values in central and southern Chile, for all durations, in contrast to the mean annual precipitation, which increases steadily towards the south*, please separate the comparisons with extreme rainfall (I think it is better characterizing it like this instead of *maximum*, which could be confused with the empirical maximum values) and the mean rainfall, and perform comparisons for Chile for both.

We appreciate the comment, because the cited sentence could be misleading for the reader. To avoid confusion between the empirical maximum values and the extreme rainfall intensities obtained from the statistical models we have added the following text in te new version of the manuscript:

In this study, IDF curves are developed by fitting stationary and non-stationary statistical models to samples of annual maximum precipitation data. The annual maximum precipitation intensities corresponding to various durations d, as estimated by these models, are hereafter denoted as I_{max} (mm h^{-1}).

Considering that the difference between the spatial pattern of mean annual precipitation and the annual maximum intensities derived from stationary and non-stationary models is a finding that has not been previously reported in literature, we re-phrase the original text cited by the reviewers to avoid misinterpretations as follows:

In addition, our results show that the annual maximum values derived from stationary and non-stationary models (I_{max}) reached its highest values in central and southern Chile, for all durations and return periods, in contrast to the spatial pattern of mean annual precipitation, which increases steadily towards the south.

Finally, considering that Reviewer 2 asked to reduce the number of figures and the fact that a comparison between mean annual precipitation and annual maximum intensities (I_{max}) is not the main focus of this manuscript, we did not add a graphic comparison between I_{max} values with those of mean annual precipitation.

R1C4: I am confused with the comparison between *stationary and non-stationary* extreme rainfall; I think it is better to express it as *stationary and non-stationary model of extreme rainfall* since the data cannot be stationary or non-stationary but rather the model can be selected to be either stationary or non-stationary (see extended discussion and literature review in Koutsoyiannis 2024, `http://doi:10.57713/kallipos-1`; and application to extreme rainfall in Iliopoulou and Koutsoyiannis, 2020, `https://doi:10.1016/j.jhydrol.2020.125005`).

We greatly appreciate your valuable comment. We have carefully reviewed the text and made the suggested changes. We now avoid using phrases like "stationary I_{max} values" and instead use more precise expressions such as "I_{max} *derived from stationary models*" or "I_{max} *from stationary approaches*", following your recommendation to make reference to the values derived from the models rather than the values directly.

R1C5: The Authors present that a main drawback of a stationary model is that *simplifies the construction of IDF curves, it may not adequately capture climate change impacts or long-term variability in precipitation intensities.*, and that a non-stationary model *consider the time-dependent nature of distribution parameters and can capture existing trends in precipitation intensity.* Although I would not argue that a non-stationary model is more flexible, since it contains additional parameters than a stationary one, please consider discussing that if the long-term dependence is inserted in the stationary model then it can also capture several observed trends and clustering (see for example in Figures 8 and 11 in in Dimitriadis et al., 2021; https://doi:10.3390/hydrology8020059, how an observed trend can be actually well simulated with a stationary model but with a more flexible probability distribution than the Gumbel one, which there was selected the Pareto-Burr-Feller one, and with acquiring for the long-term persistence of rainfall). Therefore, I would recommend to check other distributions and correlation stationary structures than the Gumbel and uncorrelated ones, to check whether there is need for non-stationary ones.

We appreciate this valuable comment and the reference to the work of Dimitriadis et al. (2021), as well as the suggestion to explore stationary models with more flexible probability distributions. We have incorporated the following modifications and clarifications in the discussion of the manuscript:

In conclusion, our findings indicate that locations with a statistically significant trend in I_{max} do not necessarily exhibit significant differences between I_{max} values derived from stationary and non-stationary models. Therefore, while accounting for the non-stationarity of extreme precipitation is important, observed trends can also be captured by stationary models when using time-dependent parameters or flexible probability distributions (Dimitriadis et al., 2021), consistent with findings from Ganguli and Coulibaly (2017), Yilmaz et al. (2014), and Yilmaz and Perera (2014). In addition, Dimitriadis et al. (2021) and Dong et al. (2021) further showed that stationary models incorporating flexible distributions or temporal correlation can reproduce observed trends and long-term persistence in precipitation extremes.

We would also like to clarify that the main objective of our work is not to identify the "*optimal*" distribution to analyse extreme precipitation or to exhaustive ascertain whether stationary models can fully capture observed trends in extreme precipitation intensities. Instead, our goal was to compare widely used stationary and non-stationary modelling approaches to estimate annual maximum intensities derived from widely used statistical models, using the topographically and climatically diverse Chilean territory as case study. A comprehensive evaluation of alternative distributions and correlation structures is beyond the scope of this study but represents a promising direction for future work.

R1C6: Please see recent research by Koutsoyiannis et al. (2023; https://doi:10.3390/w15091711) for Greece rainfall trends and about how the IMERG satellite data underestimate the rainfall extremes, which is something also observed by the Authors, if I am not mistaken.

Thank you for addressing this point. We have added to our discussion some references showing IMERG's tendency to underestimate extreme precipitation:

[...]Nonetheless, in the Far North, where all the rain gauges are located above 3,000 m a.s.l., both IMERG products notoriously underestimated I_{max} for all durations. This underestimation is in agreement with previous studies (Xiong et al., 2025; Chen et al., 2023), specially in mountainous regions, where underestimations of up to 50% have been reported for IMERGv06B (Rojas et al., 2021).

- Chen, F., Wang, R., Liu, P., Yu, L., Feng, Y., Zheng, X., Gao, J. (2023). Evaluation of GPM IMERG and error sources for tropical cyclone precipitation over eastern China. Journal of Hydrology, 627, 130384. https://doi.org/10.1016/j.jhydrol.2023.130384
- Rojas, Y., Minder, J. R., Campbell, L. S., Massmann, A., Garreaud, R. (2021). Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central Chile. Atmospheric Research, 253, 105454. https://doi.org/10.1016/j.atmosres.2021.105454
- Xiong, J., Tang, G., Yang, Y. (2025). Continental evaluation of GPM IMERG V07B precipitation on a sub-daily scale. Remote Sensing of Environment, 321, 114690. https://doi.org/10.1016/j.rse.2025.114690

R1C7: Regarding the 4th questions made by the Authors, i.e. "4. What is the impact of the typical data length of P products used for estimating stationary and non-stationary IDF curves?", please note that due to the long-term persistence observed in regular rainfall or even rainfall extremes (see the work on this subject by Iliopoulou and Koutsoyiannis, 2019; `https://doi:10.1080/02626667.2019.1657578`), the impact of the length of the rainfall timeseries could be highly significant even when using a stationary model.

We thank this important comment. First, we would like to mention that, in contrast to our study, the work by Iliopoulou and Koutsoyiannis (2019) was specifically focused on how data length affects the estimation of persistence-related indices. They concluded that "*identifiability of persistence from maxima depends foremost on the choice of the threshold for extremes, the skewness and kurtosis of the parent process, and less on sample size*". Additionally, their study differs importantly from ours in terms of data availability: their shortest record spans 150 years, a length currently unattainable in our study area, precluding a direct comparison between the two studies.

In our study, we observed changes in the parameters of the Gumbel distribution. As detailed in the manuscript, the scale parameter varied between -40% and +40%, while the location parameter ranged from -10% to +10%, as illustrated in Figure 9. Additionally, we assessed differences in I_{max} estimated from stationary and non-stationary models using 20- and 40-year time series. These differences were generally within the range of [-4, +4] mm h^{-1}, as shown in the new Figure 8.

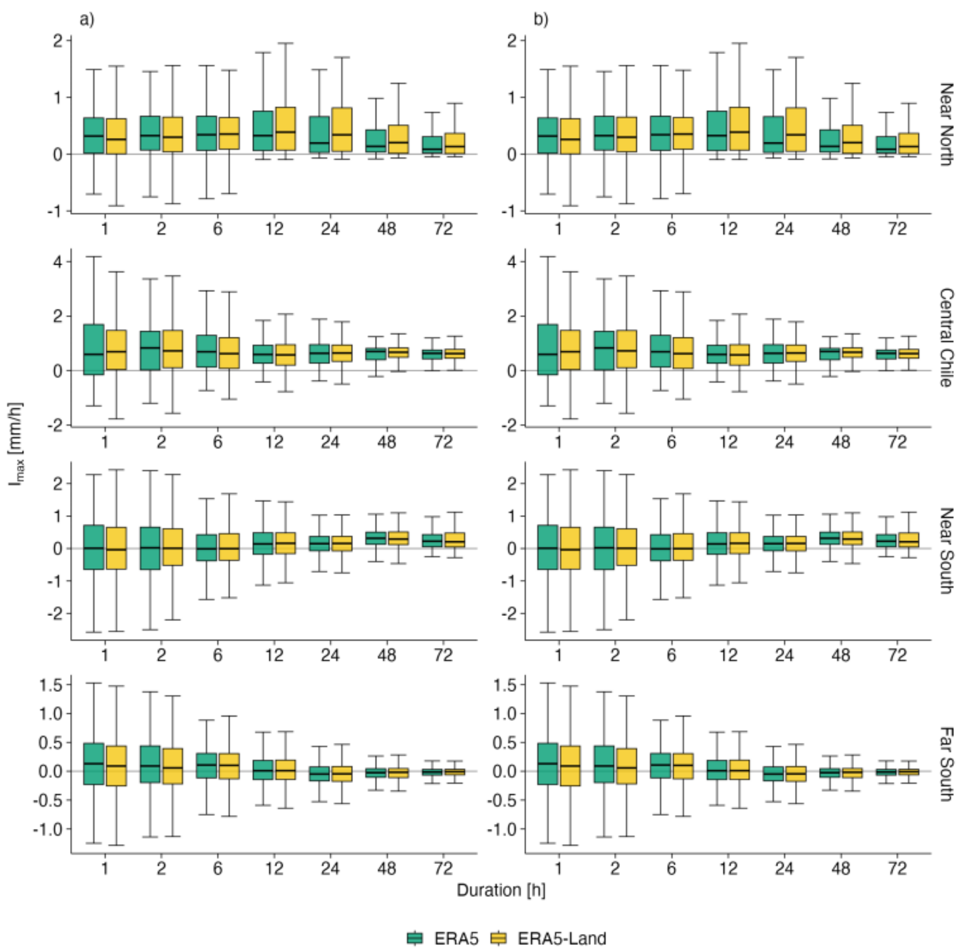

Figure 8. Panel of boxplots summarising differences between I_{max} derived from stationary and non-stationary models, estimated using data series of 40 and 20 years ($I_{max,40years} - I_{max,20years}$), for ERA5 (green colour) and ERA5-Land (yellow colour), considering a 50-years return period and durations of 1, 2, 6, 12, 24, 48, and 72 hours. From top to bottom, each panel corresponds to a different macroclimatic area. Column a) shows results obtained using the stationary model, while column b) shows results from the non-stationary model.

Finally, in order to recognise that there exist some cases where the the length of the precipitation time series could have an important impact on the maximum intensities when using a stationary model, we have added the following text to our discussion:

" *Although not directly comparable to our study due to substantial differences in data length and methodology, Iliopoulou and Koutsoyiannis (2019) highlight that the length of the precipitation time series can have an important influence on extreme precipitation when using stationary models, due to the long-term persistence observed in precipitation time series.* "

R1C8: Please include Tables about the rainfall gauges and gridded data that the Authors use in the analysis, and particularly, with primary information (e.g., length, zero values percentage, missing values, etc.), marginal statistics (e.g., mean, variance, skewness, kurtosis), and (cross-)correlation statistics (e.g., lag-1, 10, annual and 10 years of autocorrelation, cross-correlation among stations, etc.).

Thank you very much for your valuable comment. We have included a new CSV file in the supplementary material, which contains the basic information of the rain gauges (data length, percentage of zero values, amount of missing values) as well as marginal statistics (mean, variance, skewness, and kurtosis).

However, we did not include cross-correlation statistics in the supplementary material. Although this is an interesting analysis, it is not directly related to our expected results and the chosen methodology. Furthermore, due to the current length of this manuscript, the results of such analysis would not be covered in the Results, Discussion or Conclusion sections. Therefore, conducting this analysis was discarded because it is beyond the scope of this article.

R1C9: Please explain what the differences between the ERA5 and ERA5-Land datasets are, since I would expect to be similar for low elevation, since they are both coming from the same source.

We sincerely appreciate your comment. ERA5-Land precipitation is interpolated from ERA5 to a higher resolution of 9 km. This interpolation process is performed using a linear interpolation method based on a triangular mesh, and it does not include a bias-correction of ERA5 precipitation. In particular, Muñoz-Sabater et al. (2021) mention that "*Previous land reanalyses have included corrections to the precipitation forcing to address limitations of the precipitation fields of the atmospheric reanalysis. This is not the case in ERA5-Land, mainly due to the (1) enhanced quality of ERA5 precipitation when compared with previous atmospheric reanalyses (e.g. Beck et al., 2019; Tarek et al., 2020; Nogueira, 2020) and (2) reduced dependencies on external data that would limit the near-real-time data availability. However, air temperature, humidity, and pressure are corrected for the altitude differences between ERA5 and ERA5-Land grids*" (p. 4353).

We have added the following text to the description of ERA5-Land,:

" *Precipitation in ERA5-Land is obtained by linearly interpolating ERA5 forcing data onto a finer triangular mesh, without applying bias correction to the original ERA5 fields. Compared to ERA5, the input atmospheric variables (like temperature, humidity, and pressure) used for ERA5-Land are corrected for altitude differences between the ERA5 grid and the ERA5-Land grid to account for elevation effects, which improves the representation of land surface processes (Muñoz-Sabater et al., 2021)* "

Finally, we analysed whether differences between the annual maximum intensities derived from ERA5 and ERA5-Land datasets (without bias correction) have some relation with elevation in our study area for the 1981–2021 period (see Figure 1). We did not find any clear trend or relationship between the differences in $I_{\max}$ and elevation. Therefore, although similarities between the the annual maximum intensities derived from ERA5 and ERA5-Land products can be expected due to the way both products are produced, we observed slight differences within our study area, and those differences did not present a clear relationship with elevation. Consequently, this study provides important insights to future works interested in using both products.

This new figure has been included as Figure S12.1 in the new version of the updated supplementary material.

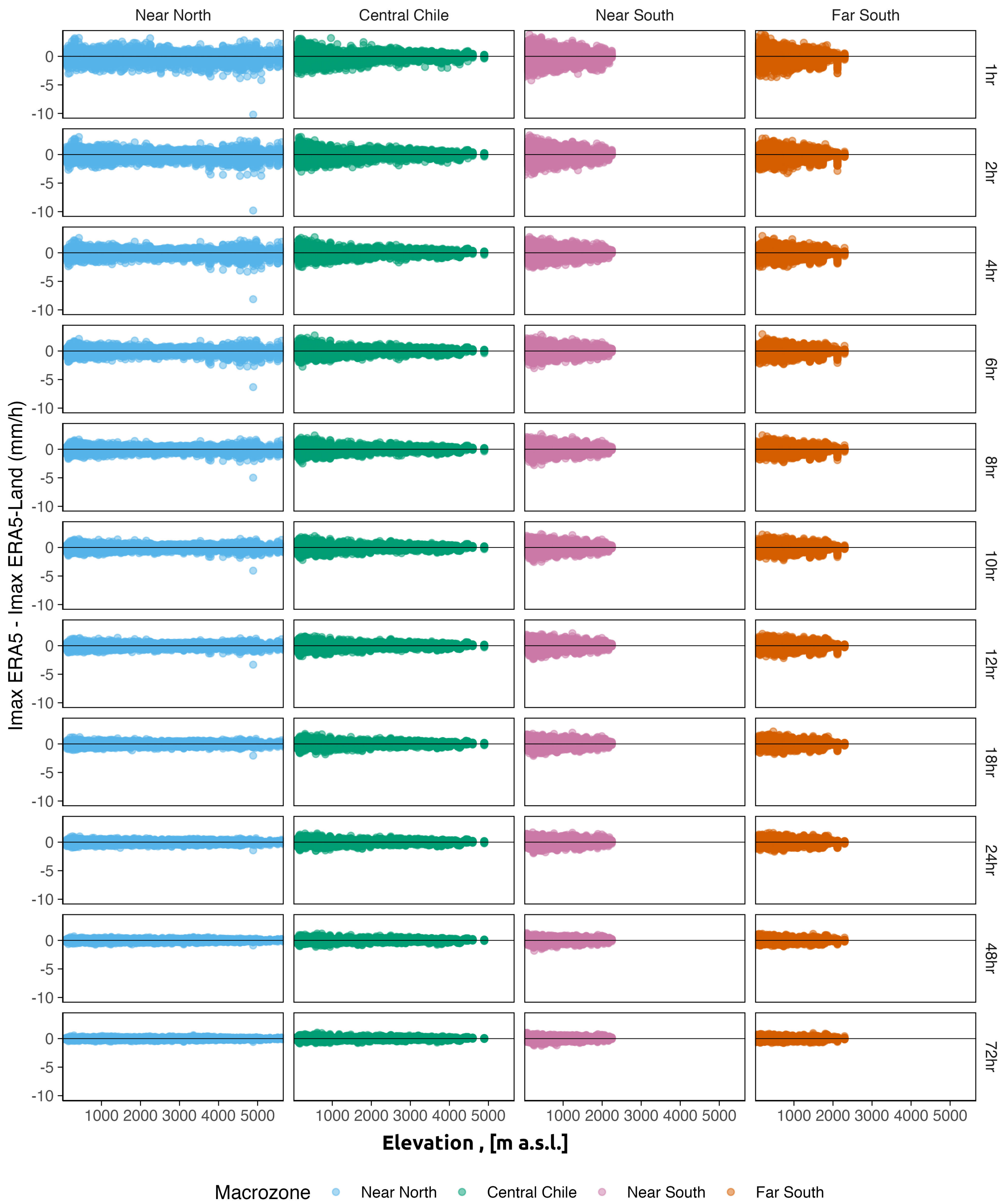

Figure 1: Differences in annual maximum intensities ($I_{\max}$) derived from ERA5 and ERA5-Land (without bias correction) versus elevation for all durations and macroclimatic regions.

R1C10: About the "bias correction discussed" in Figure 3 and "presenting such a large number of figures in a single document", I think pool graphs with all similar data on the same Figure could tackle these limitations of this study.

Thank you very much for your thoughtful comment. However, we did not get the specific observation or the exact modification being requested. Nevertheless, we have revisited Figure 3, reducing the original 25 subplots to only 5, and removing the red line that previously connected the median values of each boxplot across durations. This change aimed to improve the visualisation and analysis of our results.

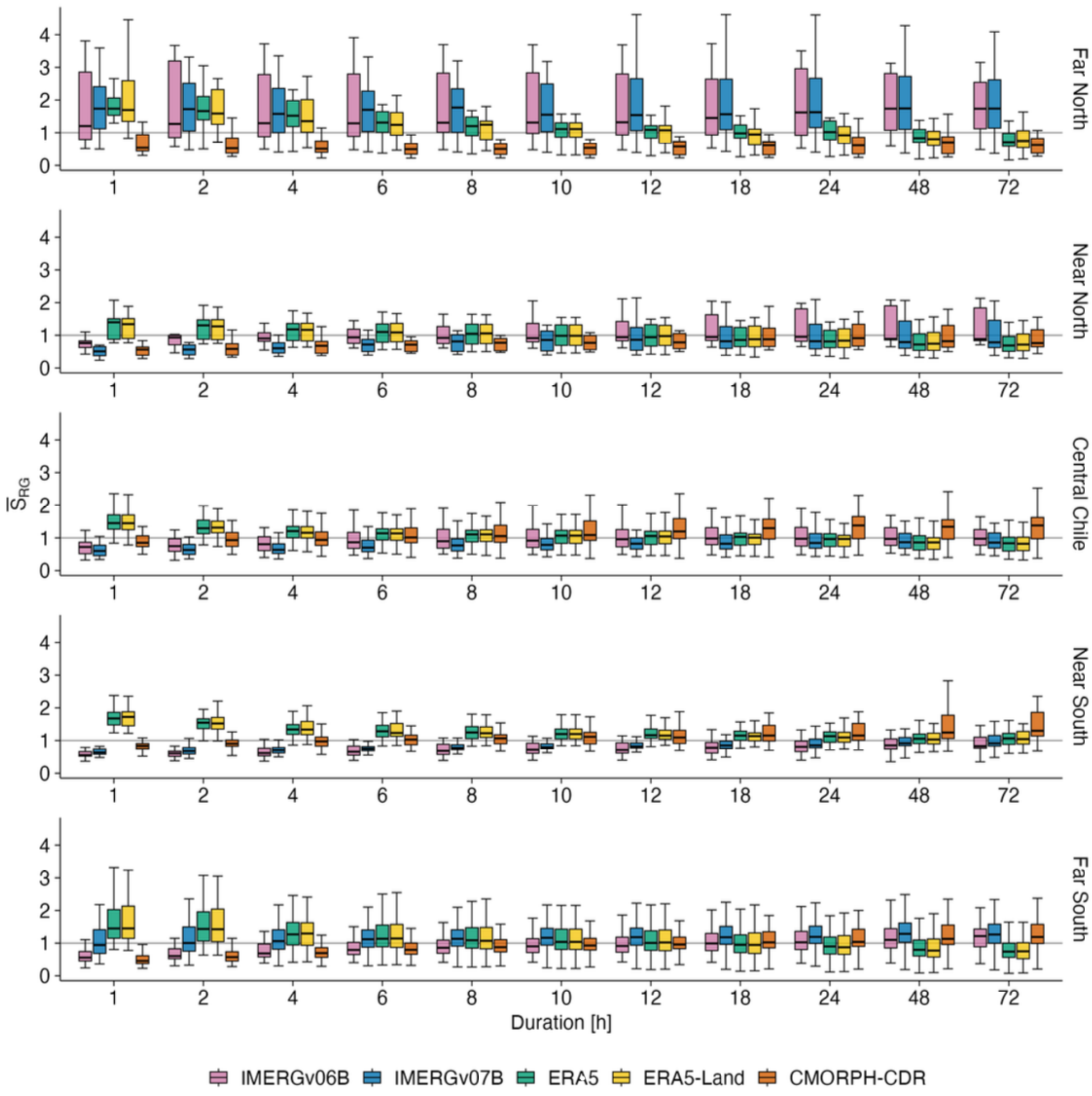

Figure 3. Boxplot of average annual bias-correction factors ($\overline{S}_{d,RG}$) for each gridded product per macroclimatic zone; for durations of 1, 2, 4, 6, 8, 10, 12, 18, 24, 48, and 72 hours.

R1C11: I think the strongest part of this research is the comparison among many satellite datasets and rain-gauges, and I think this should be the main point in the title and conclusions, and not so much the construction of the IDF curves, where, as discussed above, other stationary models or distributions could perform better.

We appreciate this comments. The original title of this manuscript was indeed focused on the usage of several gridded datasets: "*Intensity-Duration-Frequency (IDF) curves from sub-daily gridded datasets: Insights from Chile's diverse climate*". However, following an Editor's comment about revisiting that title to communicate more clearly the novelty and importance of the study we changed it to its current form. Considering that it is difficult to reconcile both requests, our new proposal for the title is:

"*Developing Intensity-Duration-Frequency (IDF) curves using sub-daily gridded and in situ datasets: characterising precipitation extremes in a drying climate*"

We would highly appreciate the explicit approval of both the Editor and Reviewer 1 about this new title, or specific suggestions about how to improve it.

Finally, be believe the first paragraph of the Conclusions is able to highlight the novel usage of several gridded datasets:

To overcome the limited availability of long time series of in situ sub-daily precipitation data, we use five state-of-the-art hourly precipitation datasets (IMERGv06B, IMERGv07B, ERA5, ERA5-Land, CMORPH-CDR) and 161 quality-checked hourly rain gauges to compute stationary and non-stationary annual maximum intensities (I_{max}) and IDF curves for the climatologically and topographically diverse Chilean territory (17-56°S).

To the best of our knowledge, this is the first work comparing annual maximum intensities derived from stationary and non-stationary statistical models and from two different families of state-of-the-art gridded precipitation datasets: the satellite products IMERGv06B/IMERGv07B and the reanalysis datasets ERA5/ERA5-Land. Moreover, this is the very first study providing intensity-duration-frequency curves at high spatial and temporal resolution using state-of-the-art gridded precipitation datasets for continental Chile. This constitutes an important contribution to advancing our knowledge about extreme precipitation events in mountainous areas where such information is generally unavailable.

R1C12: Regarding the conclusion "Given the convergence between Imax obtained from ERA5, ERA5-Land and IMERGv07B, we recommend using the highest value among them for designing climate-resilient infrastructure.", please mention by how much is the under/over-estimation.

Thank you very much for your insightful comment. While we observe a general convergence between the maximum intensities obtained from ERA5, ERA5-Land, and IMERGv07B, we provide the full Intensity-Duration-Frequency (IDF) curve results on our website `www.curvasidf.cl`. There, users can compare the intensity values across all products, durations, and return periods, allowing them to assess differences among the products and select the most appropriate dataset for their specific infrastructure design. Since we do not have a definitive "target" value that can be used to quantify precise over- or under-estimations by any product, it is preferable for users to evaluate the presented range of design intensity values to make informed decisions. This approach acknowledges the differences inherent to the estimation procedures used by each dataset and supports climate-resilient infrastructure design and planning.

R1C13: Similarly, please mention by how much "All gridded datasets – except CMORPH-CDR – show smaller biases for longer durations." and by how much "Both IMERG products overestimate Imax at shorter durations, while ERA5 and ERA5-Land underestimate it.", and compare these results with the ones indicated by other studies on IMERG (for example, the one in Greece).

To answer this specific request of the reviewer, the following paragraph was added to section 4.1:

In all macroclimatic zones but the Far North, most of the gridded datasets -except CMORPH-CDR- showed smaller biases for longer durations. In particular, both IMERG products tended to overestimate I_{max} at short durations, with median bias correction factors between 0.65 and 0.82 for 1–6 h, but this overestimation decreased for *longer durations, reaching values between 0.92 and 0.98 for 24–72 h. In contrast, ERA5 and ERA5-Land underestimated I_{max} at short durations (median bias correction factors in [1.16, 1.49] for 1–6 h). This underestimation decreased with increasing durations, reaching almost unbiased values for 10–12 h (median bias correction factors in [1.04, 1.07]), before shifting to an overestimation for longer durations (median bias correction factors in [0.83, 0.95] for 24–72 h).*

We believe the previous amount of details is too long to be included in the bullets points of the Conclusions. Therefore, we summarised as follows:

" *The biases in I_{max} varied depending on the gridded precipitation product, the macroclimatic zone and the duration considered in the analysis. In general, most gridded datasets –except CMORPH-CDR– showed smaller biases for longer durations. IMERG products consistently overestimated short-duration extremes (1-6 h) but improved toward near-unbiased estimates at longer durations (24-72 h), whereas ERA5 and ERA5-Land shifted from slight underestimation at short durations (1-6 h) to slight overestimation at longer durations (24-72 h). Bias variability is greater in the extreme Far North and Far South, as compared to the more central macroclimatic zones* ".

On the other hand, we could not find the magnitude of IMERG underestimation in Greece, because the source is in Greek language (`https://www.itia.ntua.gr/en/getfile/2273/1/documents/ntua_ombrian_reportF4.pdf`). However, in Section 4.1 we compared the bias obtained in our study with those obtained in previous studies.

R1C14: Regarding the conclusion that "Bias variability is greater in the extreme Far North and Far South, as compared to the more central macroclimatic zones.", please check whether this is due to other factors like the number of stations, or different climatic conditions, altitude, etc.

Thank you very much for this important comment. We have added the following hypothesis at the end of Section 4.1, in the Results section:

" *The higher variability in gridded precipitation biases in the Far North and Far South of Chile likely arises from the combination of complex orography, sparse observational networks, and the nature of precipitation processes in these regions. In the Far North, precipitation is highly sporadic and convective (e.g., Garreaud 1999), often associated with isolated storms and strong topographic gradients, which are challenging for coarse-resolution or satellite-based products to capture accurately. In the Far South, precipitation is dominated by frontal systems with cold cloud-tops and marked orographic enhancement over the austral Andes (Viale and Garreaud 2015), producing highly spatially variable precipitation that may not be fully resolved by the spatial resolution of ERA5 or IMERG. In contrast, Central Chile exhibits more frequent and spatially uniform precipitation events (Falvey and Garreaud 2007), which are easier for gridded products to represent, resulting in lower bias variability* ".