Australian Government
Bureau of Meteorology

Comparing the performance of ACCESS-A against ACCESS-C4 using gridded verification data

Meelis Zidikheri and Susan Rennie

June 2025

Bureau Research Report - 115

Comparing the performance of ACCESS-A against ACCESS-C4 using gridded verification data

Meelis Zidikheri and Susan Rennie

Bureau of Meteorology

Bureau Research Report No. 115

June 2025

National Library of Australia Cataloguing-in-Publication entry

Authors: Meelis Zidikheri and Susan Rennie

Comparing the performance of ACCESS-A against ACCESS-C4 using gridded verification data

ISBN: 978-1-923469-07-5

ISSN: 2206-3366

Series: Bureau Research Report – BRR115

Enquiries should be addressed to:

Lead Author: Meelis Zidikheri

Bureau of Meteorology

GPO Box 1289, Melbourne

Victoria 3001, Australia

meelis.zidikheri@bom.gov.au:

Copyright and Disclaimer

Contents

List of Figures

Executive Summary

The Bureau is currently developing a convection permitting regional model with complete coverage of Australia (ACCESS-A) that is intended to replace the multiple domains of its city-scale convection-permitting model (ACCESS-C4). ACCESS-A utilises the more advanced RAL3 science configuration while ACCESS-C4 is based on the older RAL1 science configuration. In this report, the performance of ACCESS-A during three-month long summer and winter trials is compared against ACCESS-C4 using gridded verification data from ERA5 reanalyses, GPM rainfall satellite data, and Rainfields radar rainfall data. Overall, it was found that ACCESS-A outperforms ACCESS-C4 forecasted wind fields from the surface up to 500 hPa when verified with ERA5. However, near-surface temperature errors were found to be somewhat larger in ACCESS-A when verified with ERA5. Rainfall forecasts were generally found to be better in ACCESS-A, especially light rainfall. When comparing the performance across each of the city-scale domains, it was found that near-surface temperature skill, as verified with ERA5, was somewhat degraded in ACCESS-A in some domains, except over south-east Australia. Winds fields were better in ACCESS-A in most domains. For rainfall, there were some differences between summer and winter. In summer, ACCESS-A errors were smaller in most domains. However, in winter, ACCESS-A had larger errors than ACCESS-C4 in the tropical north-Australian domains. When performing rainfall neighbourhood verification using the Fractions Skill Score (FSS), it was found that ACCESS-A generally outperformed ACCESS-C4 at small neighbourhood scales, even in the tropical domains in summer. However at large neighbourhood scales, the results were some somewhat mixed, although when averaged across all domains, ACCESS-A yielded better scores than ACCESS-C4. Subjective evaluation of precipitation fields during the Eastern Australia Christmas Day storms of 2023 also demonstrated the utility of ACCESS-A over ACCESS-C4 with the former exhibiting improved rainfall coverage during the event.

1. Introduction

The Bureau is currently developing a high-resolution (1.5 km) convection-permitting regional model covering all of Australia, named ACCESS-A. ACCESS-A is scheduled to replace the latest version of ACCESS-C (Rennie, et al., 2022), namely ACCESS-C4, in the next few years. ACCESS-C4, although at the same resolution, covers multiple, smaller domains that include major urban centres in Australia: Darwin (DN), North Queensland (NQ), Brisbane (BN), Sydney (SY), Victoria/Tasmania (VT), Adelaide (AD), and Perth (PH). Figure 1 shows the boundaries of the ACCESS-A and ACCESS-C4 domains.

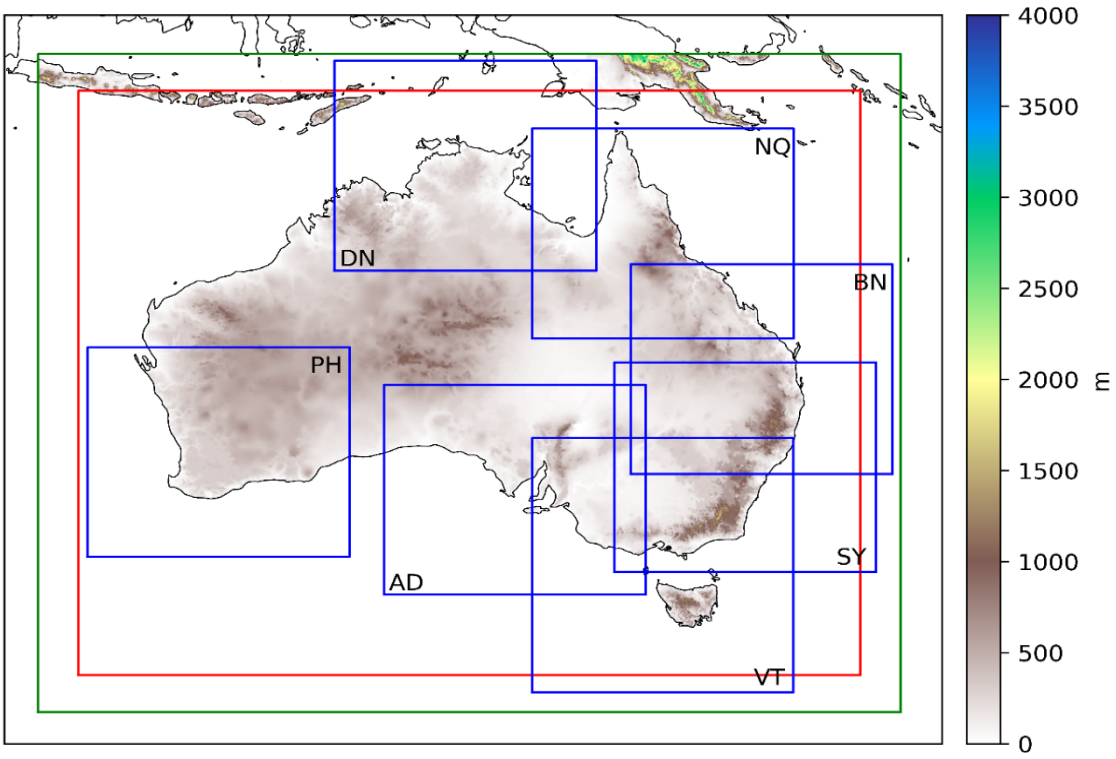

Figure 1: ACCESS-A fixed resolution boundary (red), variable-resolution boundary (green), and ACCESS-C multiple domains (blue).

Apart from differences in coverage, ACCESS-A and ACCESS-C4 also differ in model physics and model initialisation details. ACCESS-C4 uses the older RAL1 physics scheme (Bush, et al., 2020) while ACCESS-A employs the more advanced RAL3.2 physics scheme (Bush, et al., 2024). Although both ACCESS-A and ACCESS-C4 use rapid, 1 h cycling, 4DVar data assimilation (Milan, et al., 2020) for model initialisation, ACCESS-A assimilates additional observations including Himawari radiances and radar reflectivities. Moreover, ACCESS-A uses Large-Scale Blending (LSB) whereas ACCESS-C4 does not. LSB is a technique for including information from a parent global model within a regional model. It enables the regional model to take advantage of the parent model's global observation coverage to enhance forecasting skill at the large scales (Milan, et al., 2023). This is particularly important in ACCESS-A because of its

rather large Australia-wide domain. In ACCESS-A, the large scales from its parent model, ACCESS-G, are blended every 6 h, starting from the T+5 cycle (with T corresponding to the latest ACCESS-G cycle), after a spin-up period of 2 h (using the T+3 global model output). During the spin-up period, the global model adjusts to the higher ACCESS-A resolution using a downscaled forecast (In this context, this is just the ACCESS-G T+3 output field downscaled to 1.5 km and run forward for 2 h). Previous studies with ACCESS-A demonstrated significantly improved skill obtained using the LSB technique compared to not using LSB (Zidikheri, et al., 2025), which was consistent with the earlier results obtained with the United Kingdom Meteorological Office (UKMO) UKV regional model (Milan, et al., 2023).

In this study the performance of ACCESS-A is compared to the performance of ACCESS-C4 during two three-month trial periods using gridded verification data, namely the European Centre for Medium-range Weather Forecasts Reanalysis version 5 (ERA5), Global Precipitation Mission (GPM) rainfall satellite data, and Rainfields radar data. Section 2 describes the verification methodology in more detail. Section 3 describes the results. Section 4 presents a discussion of the results while Section 5 is the conclusion.

2. Trial details and verification methodology

2.1. Trial details

ACCESS-A was run for two trial periods, namely 1 December 2023 – 29 February 2024 (referred to as the summer period in this report) and 1 May 2024 – 31 July 2024 (referred to as the winter period in this report). ACCESS-C4 data was also obtained for these periods from pre-operational trial runs. ACCESS-A was run on the Bureau's Eastern Data Centre (EDC) machine while the ACCESS-C4 data was generated on the older Main Data Centre (MDC) machine.

2.2. Verifying data

The forecasts were verified against ERA5 reanalysis data (Hersbach, et al., 2020), GPM satellite rainfall data (Huffman, et al., 2019), and Rainfields radar rainfall composite data (Seed, et al., 2008) version 3.3.5. The verified fields were precipitation rate, screen-level temperature, mean sea-level pressure, and 10m winds. Unfortunately, the number of pressure-level fields was limited in the archived ACCESS-C4 dataset, so only 850 hPa and 500 hPa wind fields could be verified for ACCESS-C4. The ERA5 data was fairly coarse (0.25°) compared to the ACCESS-A and ACCESS-C4 model grids (0.0135°). However, given that ACCESS-A employs large-scale blending, it seemed to be important to understand how the larger scales compared in the two models, hence the use of ERA5, which makes use of the global observation network and is generally recognized as the most accurate global reanalysis product. The GPM data (late product) was at a resolution of 0.1° while the Rainfields data was regridded to 0.02°, which is of comparable resolution to the model data. The Rainfields data coverage over Australia is

shown in Figure 2. Prior to verification, the model data were re-gridded (using a conservative area-weighted re-gridding scheme) to match the verifying data. The verification was performed over each of the seven ACCESS-C4 domains, namely Darwin (DN), North Queensland (NQ), Brisbane (BN), Sydney (SY), Victoria/Tasmania (VT), Adelaide (AD), and Perth (PH).

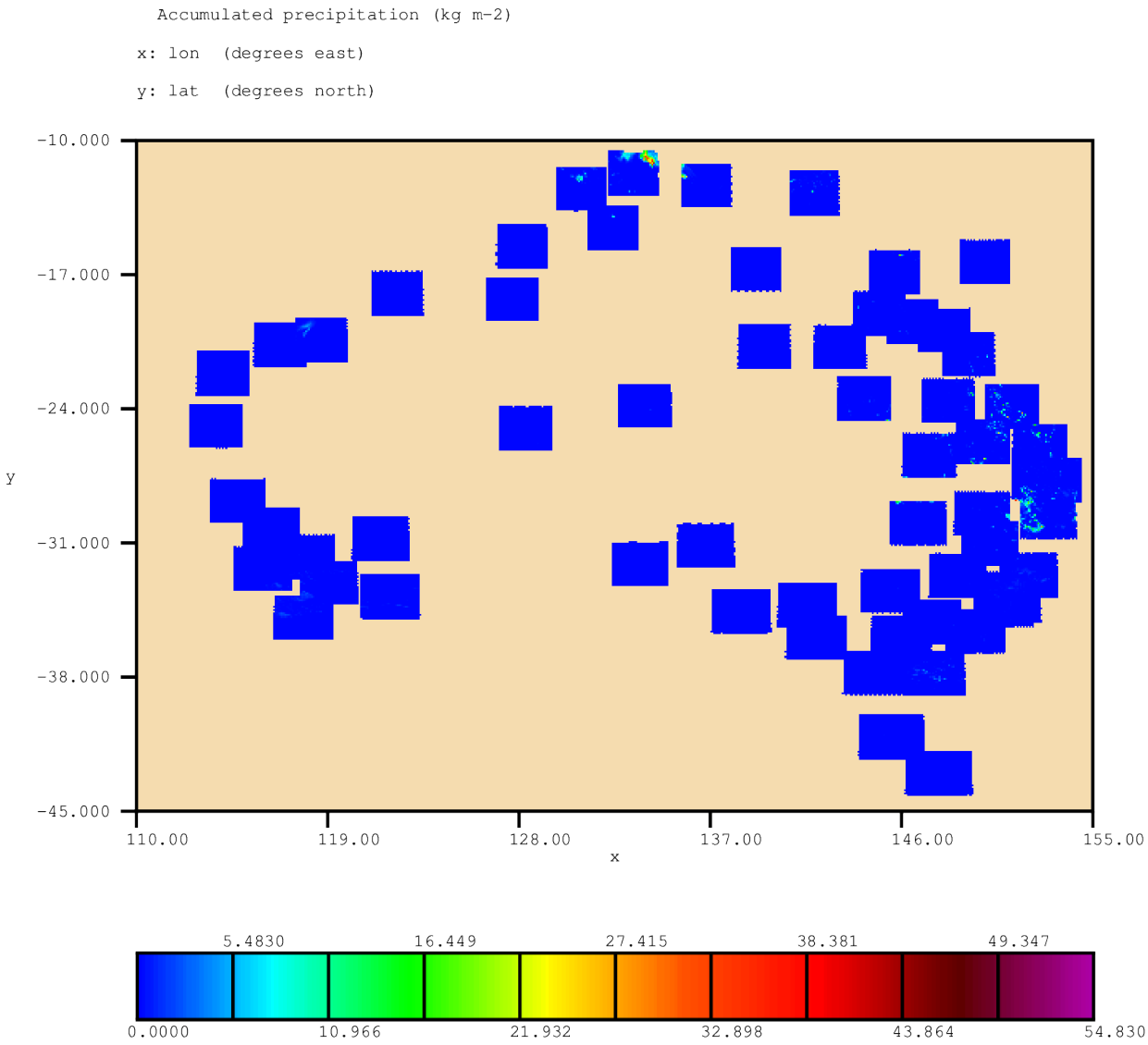

Figure 2: Rainfields radar data coverage over Australia. Regions not in the colour spectrum above (mostly oceans and interior regions) do not have coverage.

2.3. Verification methodology

The main metrics considered were mean error (which measures bias), error standard deviation, and root mean square error (RMSE). For precipitation, the fractions skill score (FSS) (Roberts & Lean, 2008), which measures how well the model and observations

agree over a specified neighbourhood distance scale, was also used. The FSS values were calculated for three precipitation rate thresholds: light (precipitation rate < 1 mm/h), moderate (1 mm/h < precipitation rate < 5 mm/h), and heavy (precipitation rate > 5 mm/h). When comparing forecasts from different modelling configurations, improved forecasts are indicated by reduced mean error, RMSE, and standard deviation as well as increased FSS and pattern correlation relative to the control forecasts. We shall also make use of relative (or percentage) differences defined as:

$$\text{metric \% difference} = 100\,\frac{\text{EXPERIMENT METRIC VALUE-CONTROL METRIC VALUE}}{\text{CONTROL METRIC VALUE}}.$$

Negative values of absolute mean error, RMSE, and standard deviation percentage differences indicate that the experiment is more skilful than the control while negative values of FSS and pattern correlation percentage difference indicate that the control is more skilful than the experiment. In the context of this study, ACCESS-A forecasts are the experiment and ACCESS-C4 forecasts are the control.

3. Results

3.1. Aggregated domain verification

In this section, we present the verification results aggregated over all seven ACCESS-C4 domains. This enables a synoptic view of the performance of ACCESS-A against ACCESS-C4. The verification results over each of the seven domains are discussed in Section 3.2. The aggregated verification metrics are computed as:

$$\text{agg. metric} = \frac{1}{N}\sum_{i=1}^{N}\text{metric}_i,$$

where metric_i is the value of metric over the domain i and N is the number of domains. Note that because the domains overlap and their sizes are not precisely the same, this is not exactly the same as calculating the metrics over the parts of Australia covered by the domains. The aggregated metric instead provides a convenient single metric to evaluate the relative performance of the two models given metrics computed over individual domains.

3.1.1. ERA5 Verification

Figure 3 shows the percentage difference RMSE values of winds, screen-level temperature, and mean sea-level pressure of ACCESS-A forecasts relative to ACCESS-C4 forecasts as a function of lead time. In this (and subsequent scorecards), blue colours indicate that ACCESS-A performs better than ACCESS-C4. Results for both the summer and winter trial periods are shown. The figure shows that for most fields, and lead times, ACCESS-A is more skilful than ACCESS-C4. The exceptions are the mean sea-level pressure forecasts during summer, which shows some degradation at longer (> 8 h) forecast lead times. The screen-level temperature, while not clearly degraded at longer

lead times, is at best neutral. Figure 4 shows the percentage difference of the absolute mean error (i.e., absolute value of the bias) for the same fields. This figure indicates that the magnitude of the mean sea-level pressure bias is higher in ACCESS-A compared to ACCESS-C4. The 10 m meridional wind bias also appears to increase, although this has less of an effect on the RMSE value in this case (the RMSE values decrease except for longer lead times in summer, as shown in Figure 3). The screen-level temperature bias, on the other hand, appears to improve (i.e., get smaller) in ACCESS-A. Figure 5 shows the error standard deviation percentage difference. The standard deviation quantifies the average size of the error component that deviates from the mean, herein referred to as the fluctuating error component. This reveals that the slight degradation of screen-level temperature RMSE in Figure 3 is mostly due to the increase in fluctuating error magnitude in ACCESS-A (as indicated by increased standard deviation values). Fluctuating errors are reduced in all the other fields that are shown. Pattern correlation percentage differences are shown in Figure 6. This figure reveals that pattern correlations are increased in ACCESS-A for all the indicated fields, including screen-level temperature.

Figure 7 shows the precipitation rate RMSE as a function of lead time for both ACCESS-A (red) and ACCESS-C4 (blue). This reveals that ACCESS-A has lower RMSE values at all lead times for both summer and winter, and is therefore more skilful, than ACCESS-C4 at these verification length scales. Figure 8 shows the light precipitation FSS scores as a function of grid length for summer and winter. In this, and all subsequent FSS plots, the forecast lead time is 6 h. ACCESS-A demonstrates substantial improvement in skill (higher FSS values) over all neighbourhood grid lengths for this precipitation intensity. Improvements in FSS are also seen for moderate rainfall rates although these are smaller in size (Figure 9). Improvements are less clear for heavy precipitation (Figure 10), and in fact FSS values appear to reduce at longer grid lengths during the summer period when employing ERA5 data for verification.

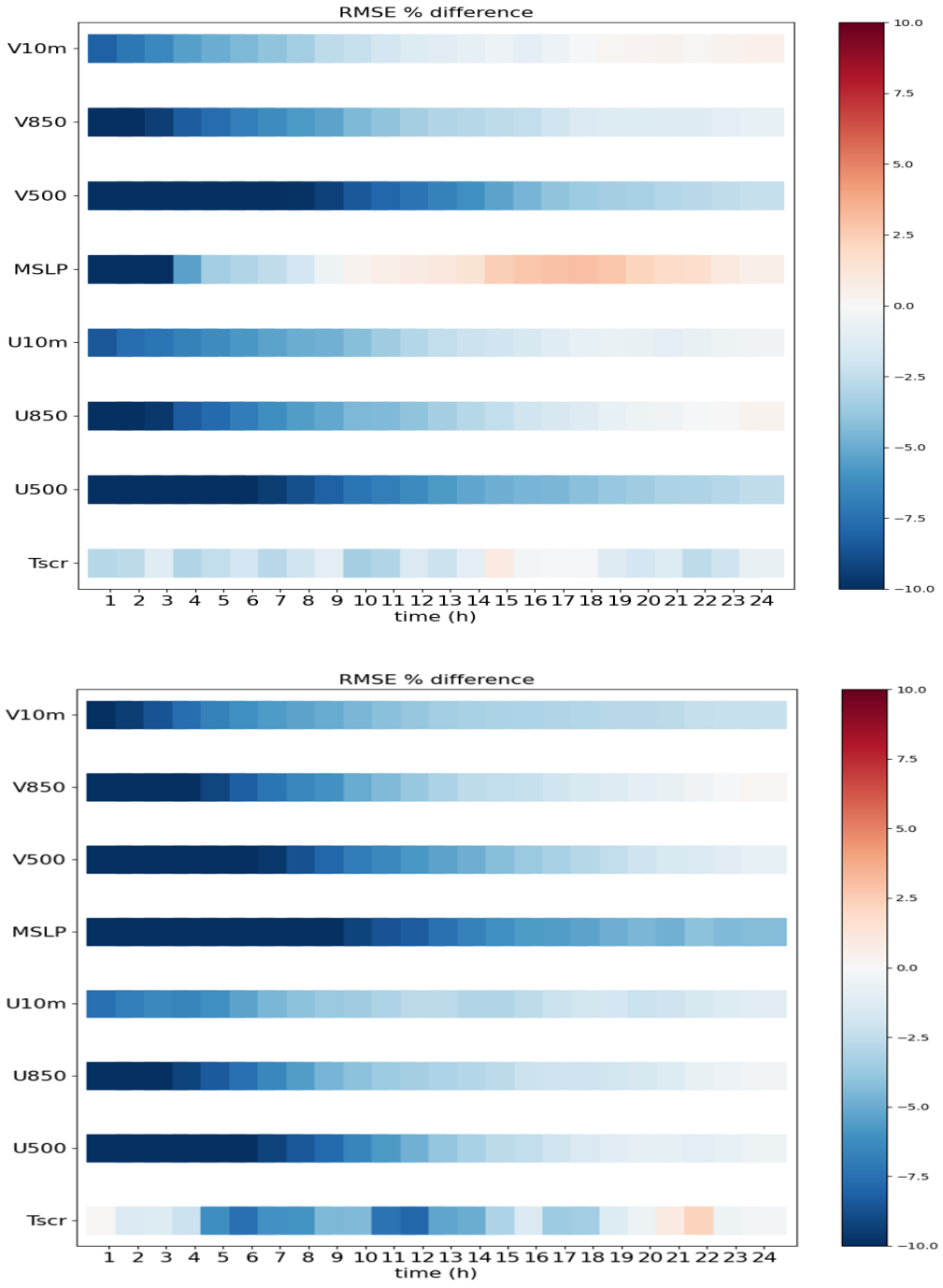

Figure 3: RMSE percentage difference for different fields as a function of lead time as obtained using ERA5 verifying data. The top panel shows the results for summer while the bottom panel shows the results for winter.

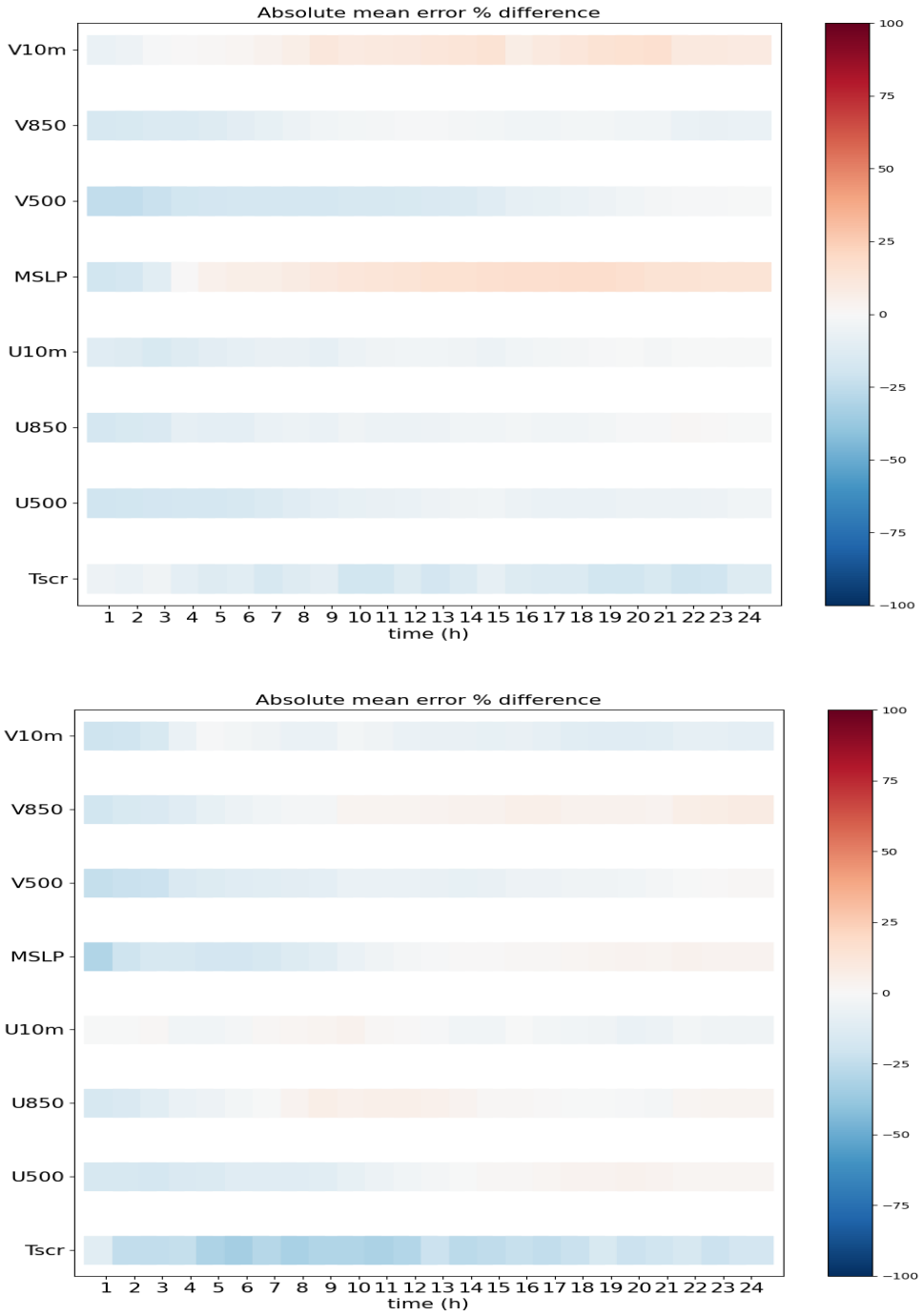

Figure 4: Absolute mean error percentage difference for different fields as a function of lead time as obtained using ERA5 verifying data. The top panel shows the results for summer while the bottom panel shows the results for winter.

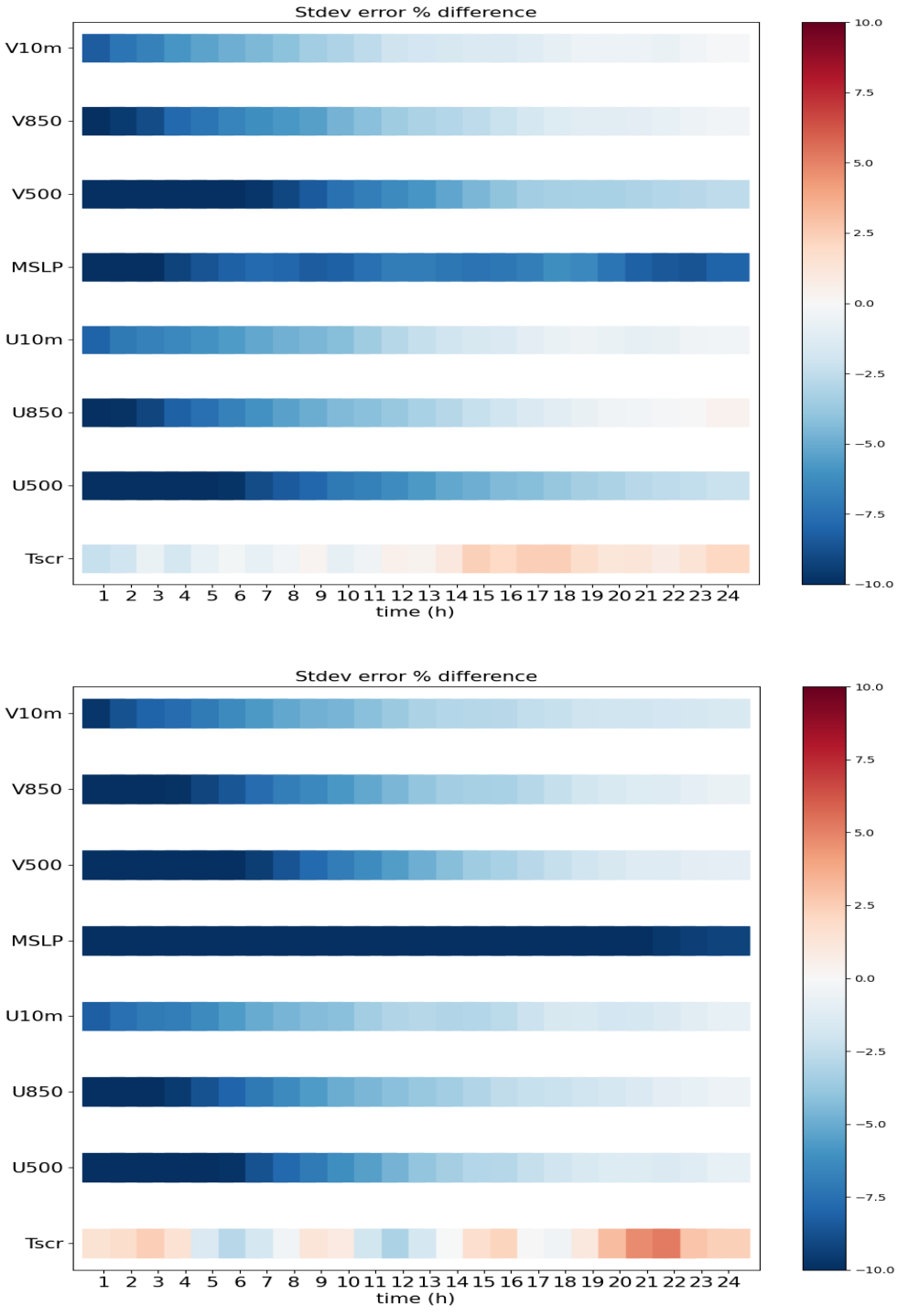

Figure 5: Error standard deviation percentage difference for different fields as a function of lead time as obtained using ERA5 verifying data. The top panel shows the results for summer while the bottom panel shows the results for winter.

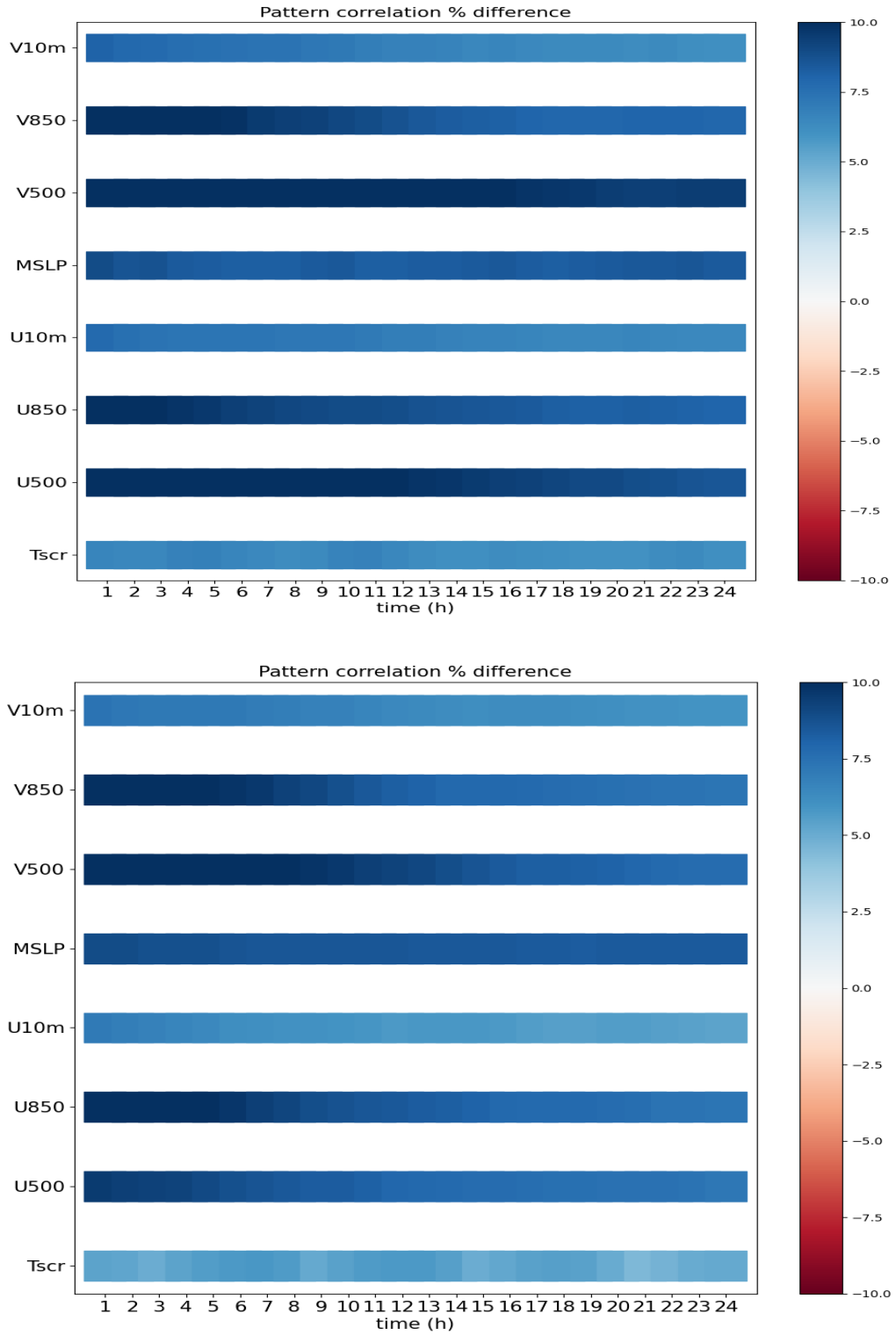

Figure 6: Pattern correlation percentage difference for different fields as a function of lead time as obtained using ERA5 verifying data. The top panel shows the results for summer while the bottom panel shows the results for winter.

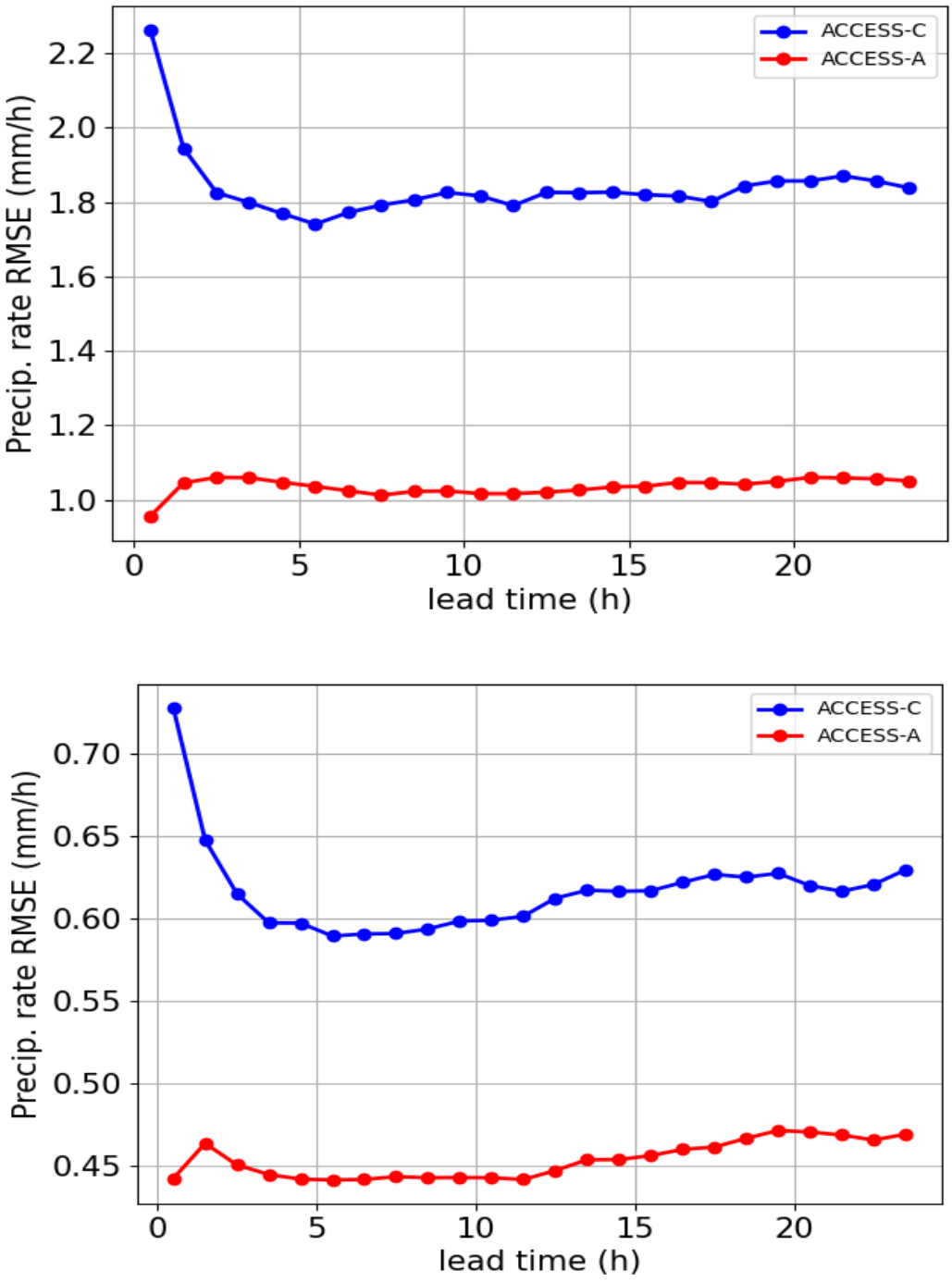

Figure 7: Precipitation rate RMSE as a function of lead time using ERA5 verifying data: summer (top) and winter (bottom).

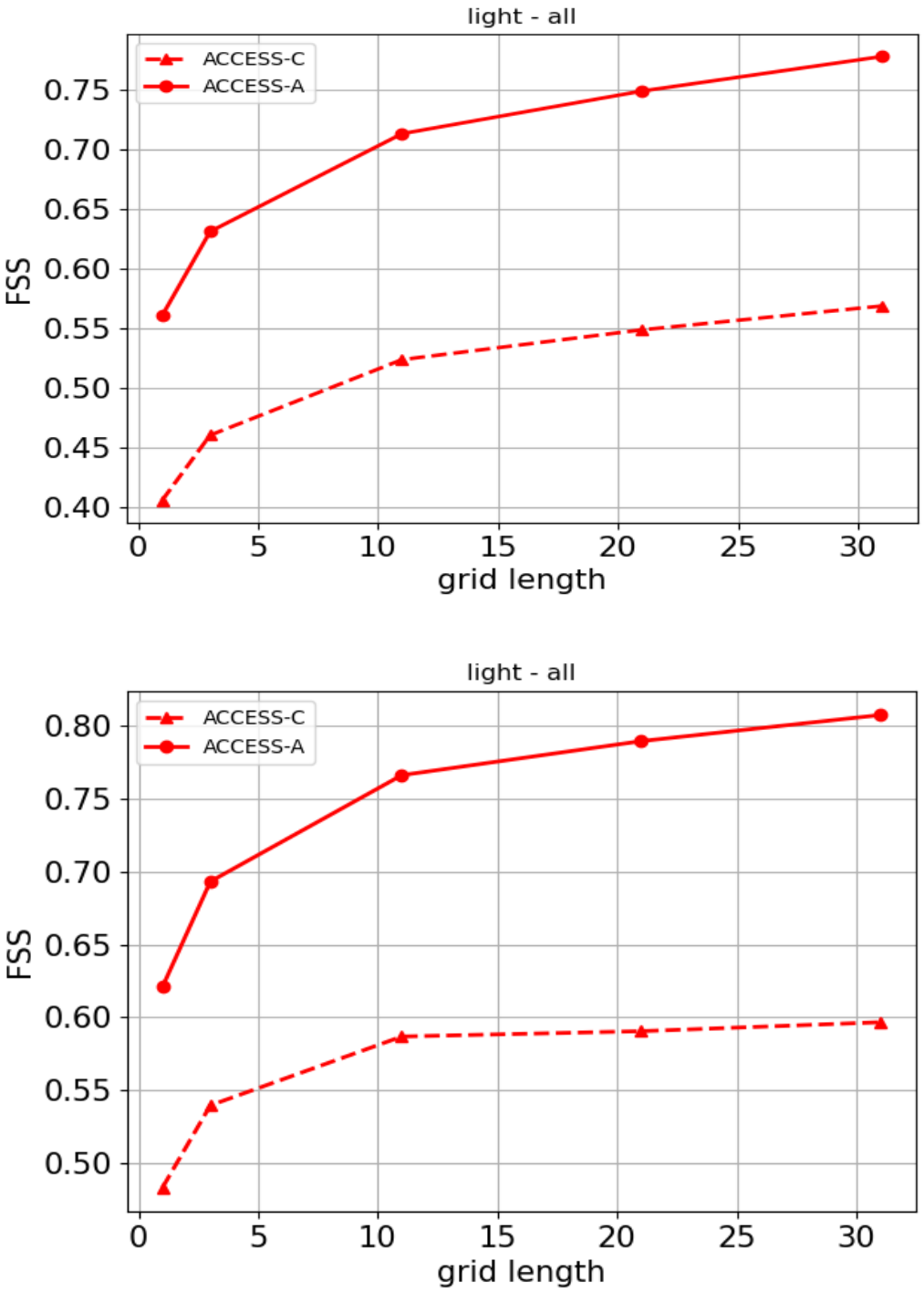

Figure 8: Light precipitation rate FSS as a function of neighbourhood grid length using ERA5 verifying data: summer (top) and winter (bottom).

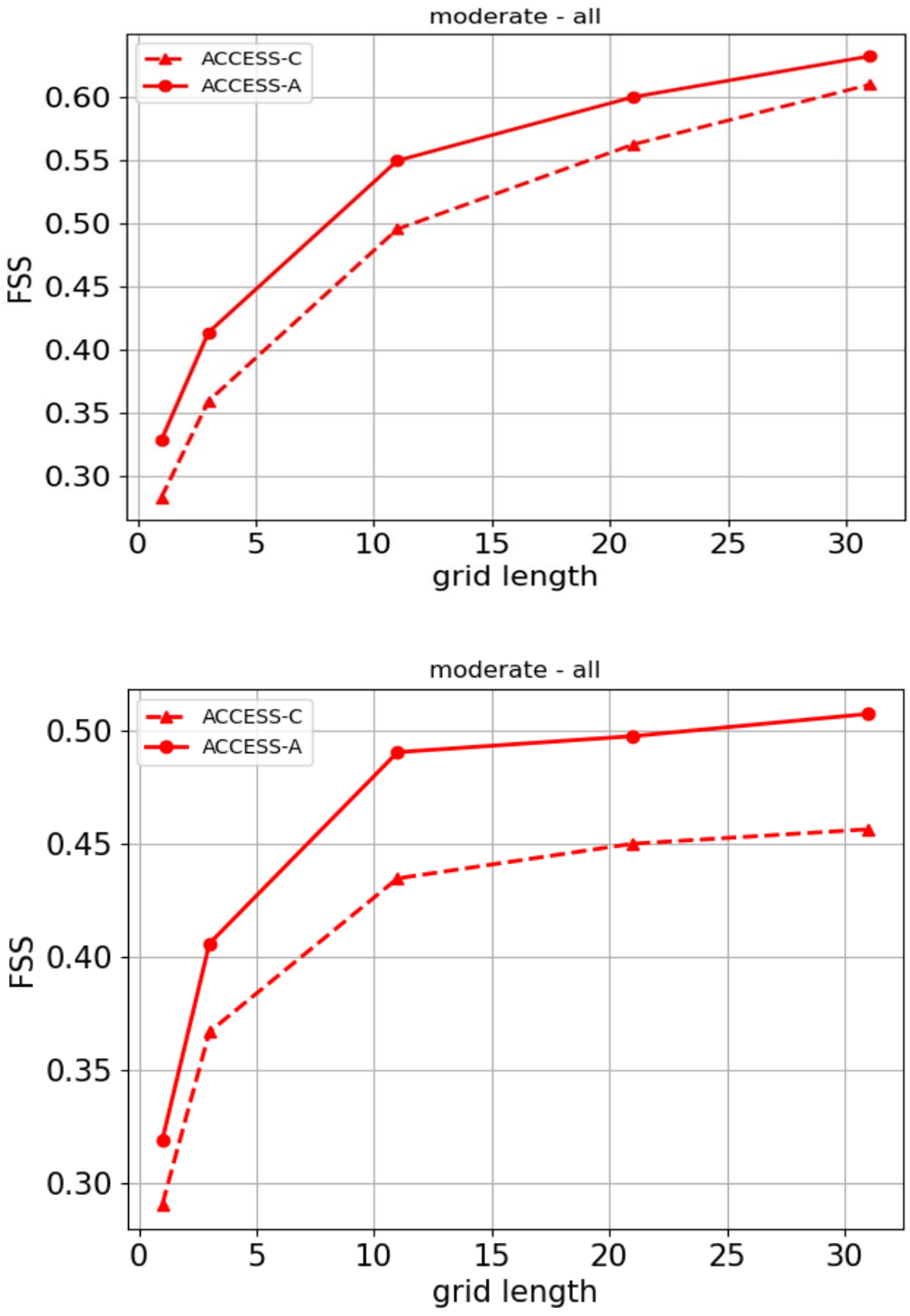

Figure 9: Moderate precipitation rate FSS as a function of neighbourhood grid length using ERA5 verifying data: summer (top) and winter (bottom).

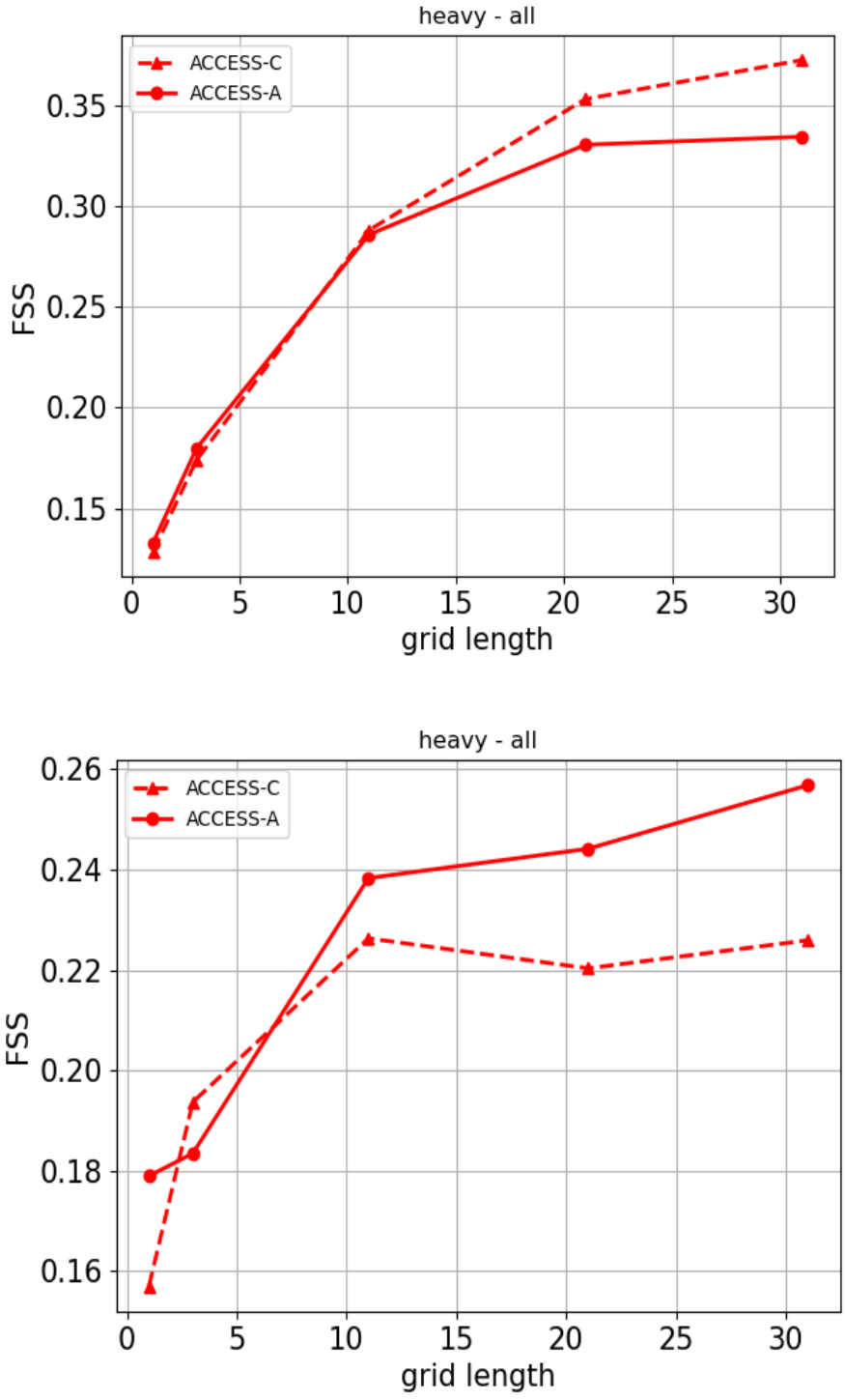

Figure 10: Heavy precipitation rate FSS as a function of neighbourhood grid length using ERA5 verifying data: summer (top) and winter (bottom).

3.1.2. GPM Verification

Figure 11 shows the precipitation rate RMSE values for ACCESS-A and ACCESS-C4 as a function of lead time during summer and winter using GPM satellite data for verification. These results are qualitatively the same as for the ERA5-verified data, namely that ACCESS-A has lower RMSE values than ACCESS-C4 at all lead times during summer and winter. For this observation type, however, the reduction in RMSE is more substantial during the summer period.

Figure 12, Figure 13, and Figure 14 show the FSS values as a function of grid length for light, moderate, and heavy precipitation, respectively. Light precipitation shows substantial improvement in skill for ACCESS-A, especially during summer. Moderate and heavy precipitation also show clear improvements in skill, albeit of smaller magnitudes.

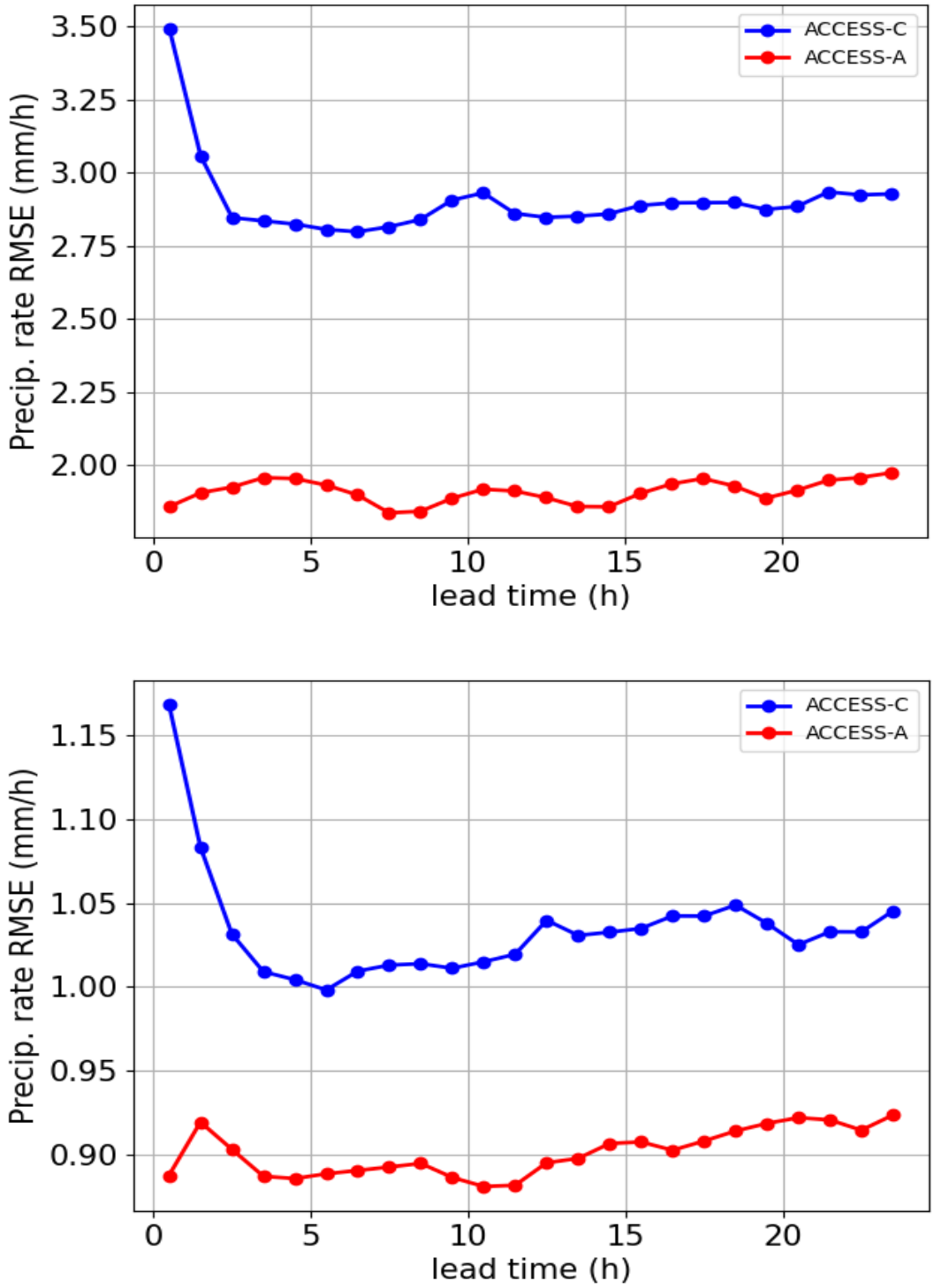

Figure 11: Precipitation rate RMSE as a function of lead time using GPM verifying data: summer (top) and winter (bottom).

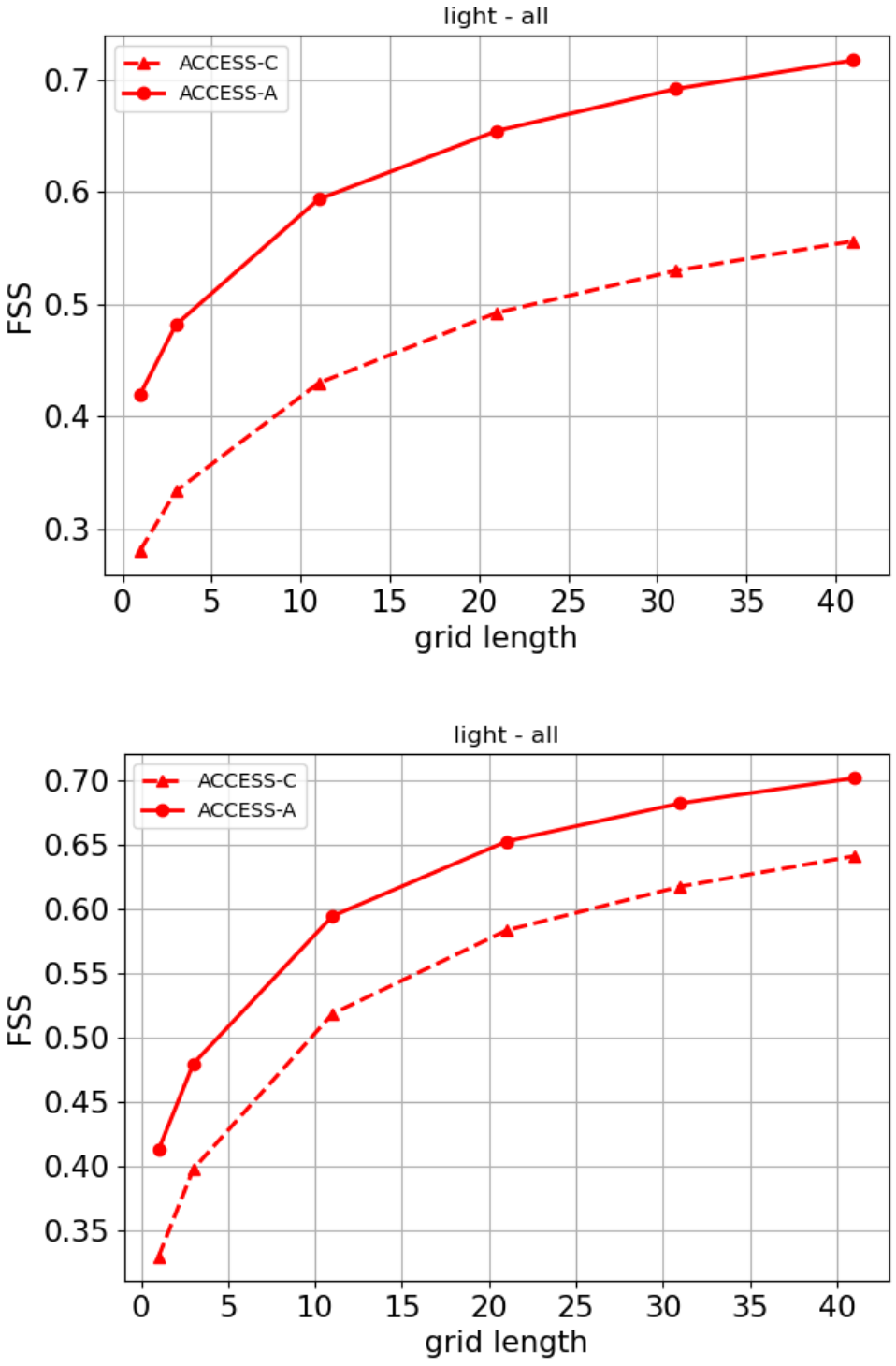

Figure 12: Light precipitation rate FSS as a function of neighbourhood grid length using GPM verifying data: summer (top) and winter (bottom).

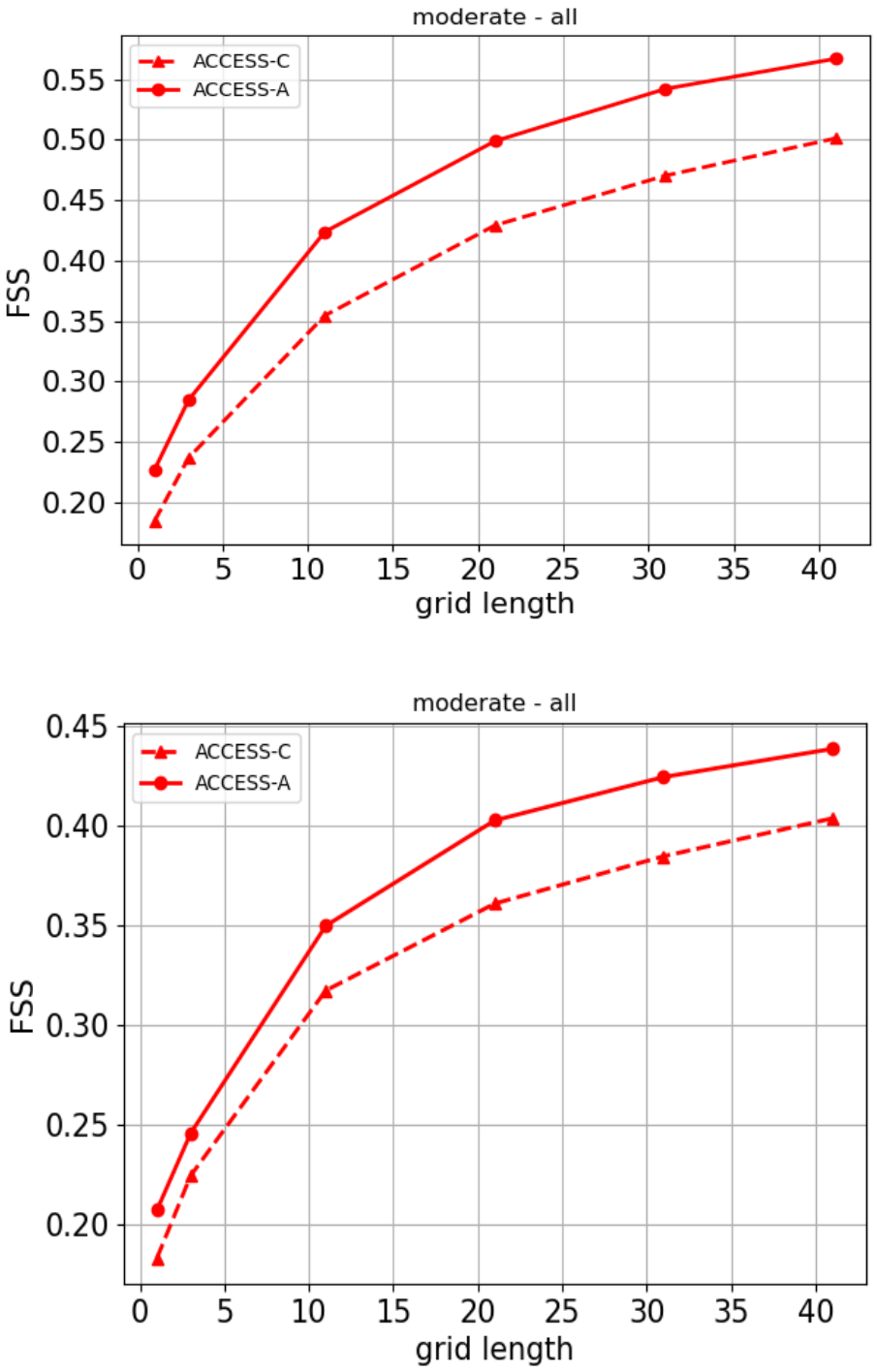

Figure 13: Moderate precipitation rate FSS as a function of neighbourhood grid length using GPM verifying data: summer (top) and winter (bottom).

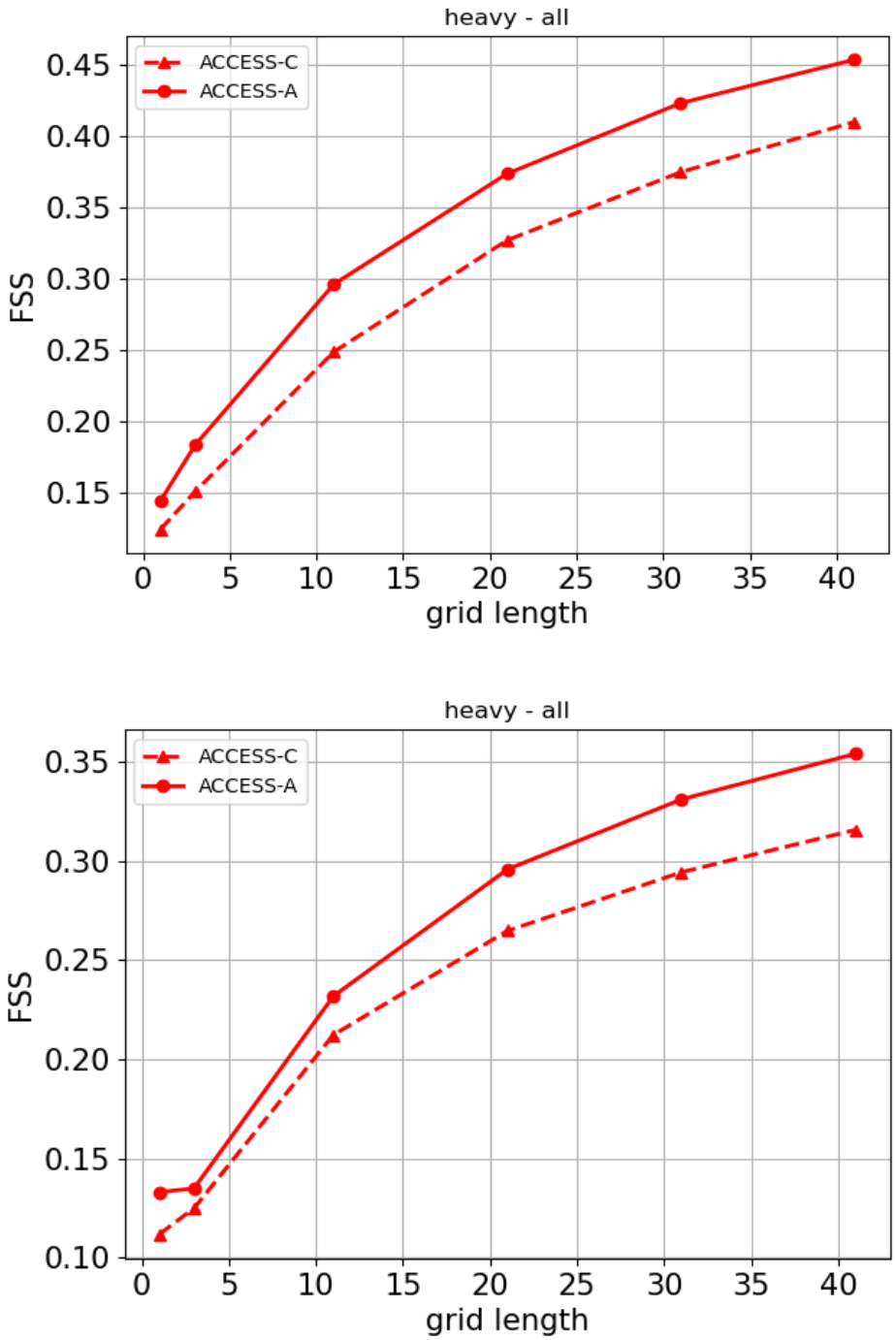

Figure 14: Heavy precipitation rate FSS as a function of neighbourhood grid length using GPM verifying data: summer (top) and winter (bottom).

3.1.3. Rainfields verification

Figure 15 shows the RMSE as a function of lead time during summer and winter. As found for the ERA5- and GPM-verified data, ACCESS-A has a substantially higher skill during summer and winter compared to ACCESS-C4.

FSS values as a function of lead-time for light, moderate, and heavy rainfall rates are shown in Figure 16, Figure 17, and Figure 18, respectively. Unlike the ERA5- and GPM-verified FSS values for light rain, the FSS values using Rainfields data for verification do not clearly indicate that ACCESS-A is substantially more skilful. During summer ACCESS-A is somewhat more skilful except for the largest neighbourhood size. During winter, ACCESS-A is only slightly more skilful for grid lengths less than about 15; at longer grid lengths, ACCESS-C4 is slightly more skilful. FSS values for moderate and heavy rainfall do appear to show that ACCESS-A is more skilful, however, in agreement with the GPM-verified results in particular.

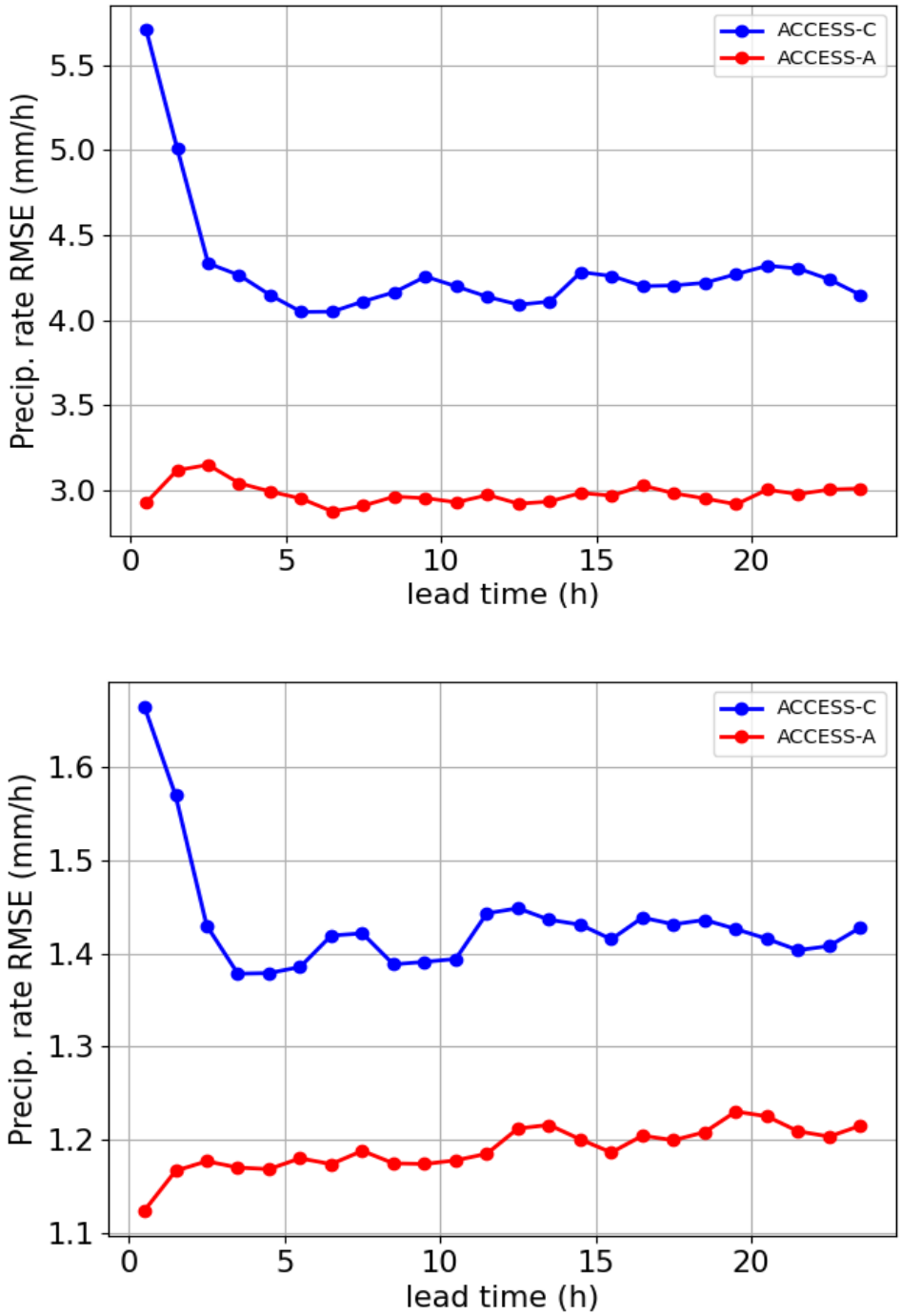

Figure 15: Precipitation rate RMSE as a function of lead time using Rainfields verifying data: summer (top) and winter (bottom).

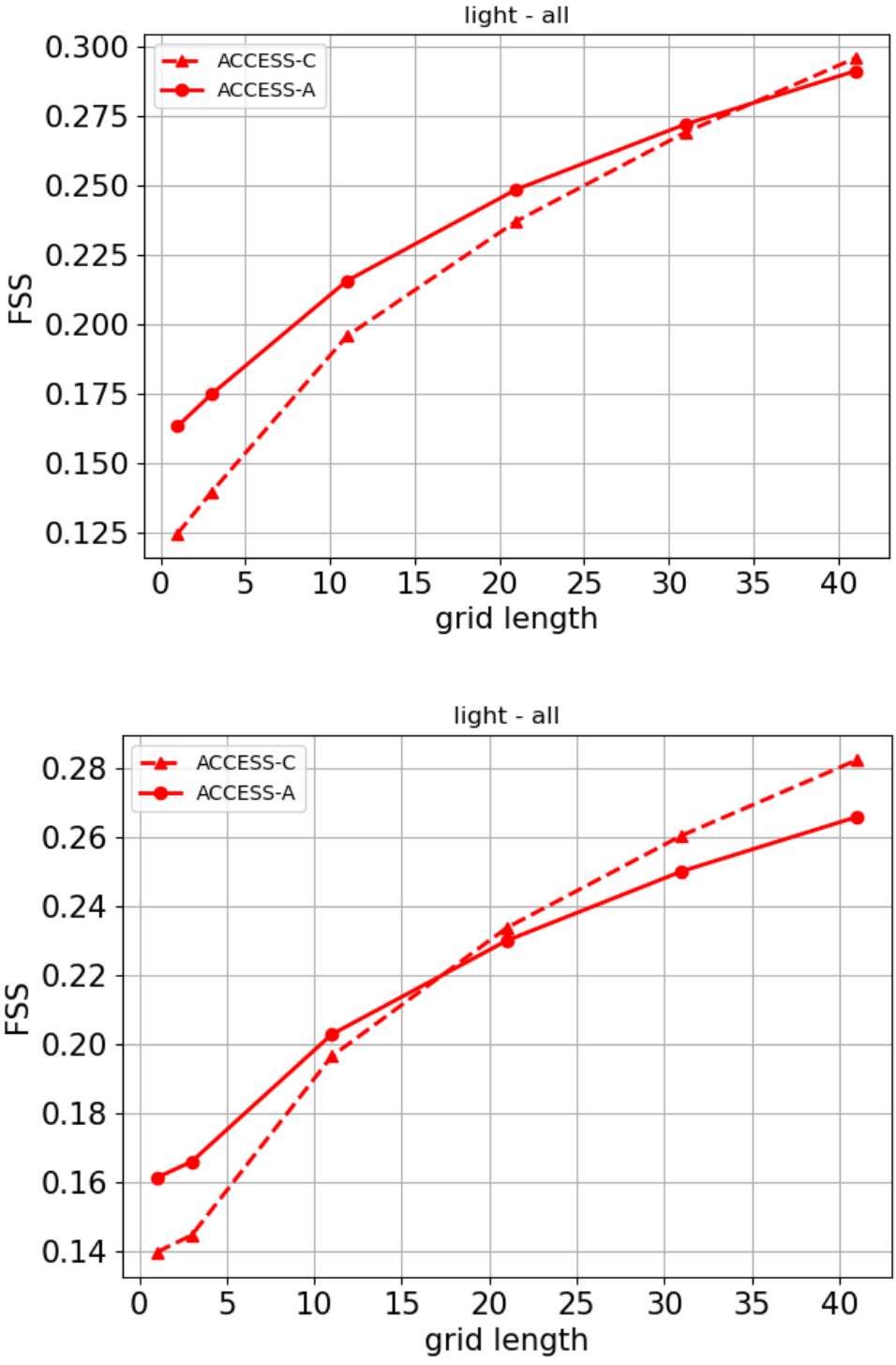

Figure 16: Light precipitation rate FSS as a function of neighbourhood grid length using Rainfields verifying data: summer (top) and winter (bottom).

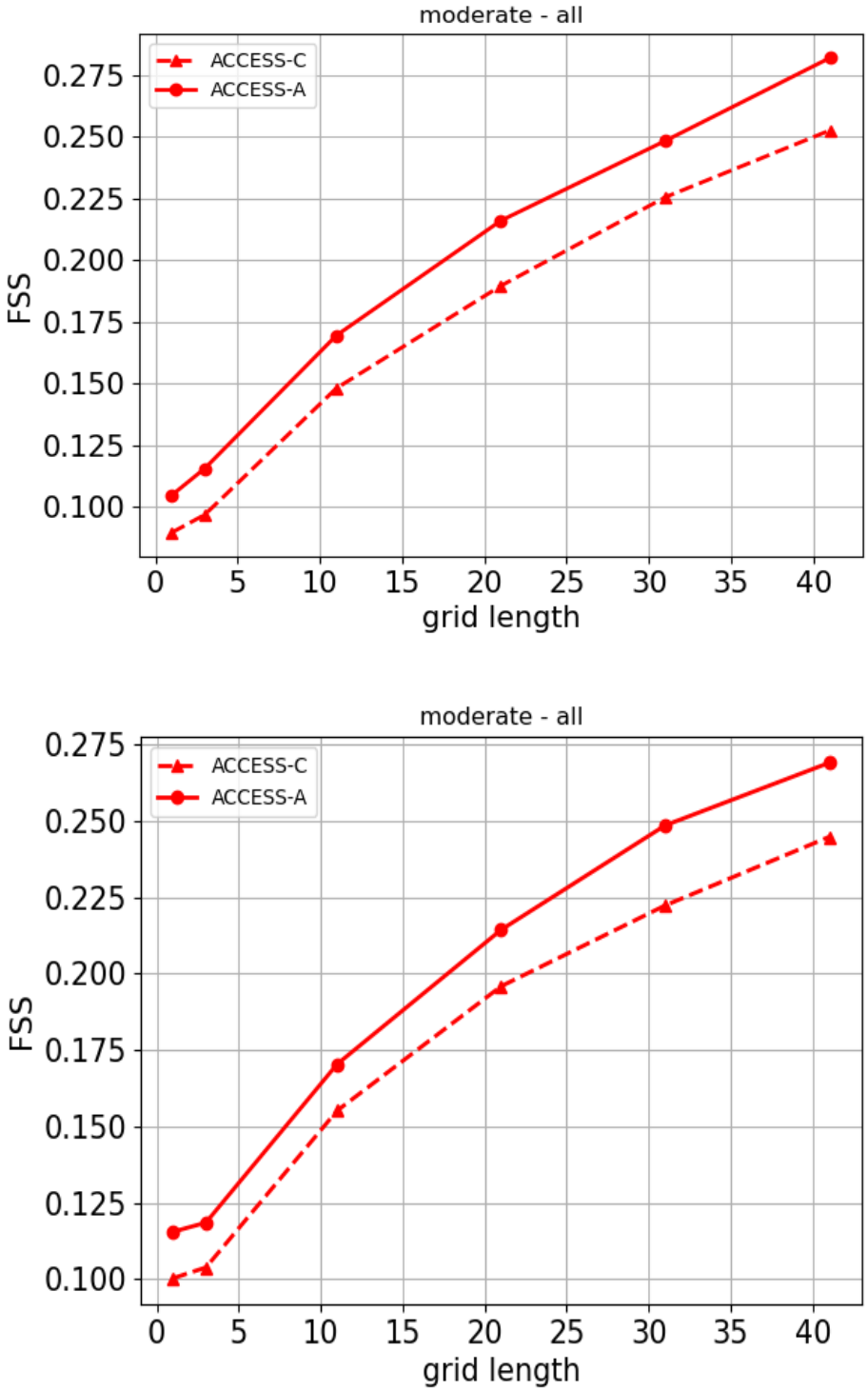

Figure 17: Moderate precipitation rate FSS as a function of neighbourhood grid length using Rainfields verifying data: summer (top) and winter (bottom).

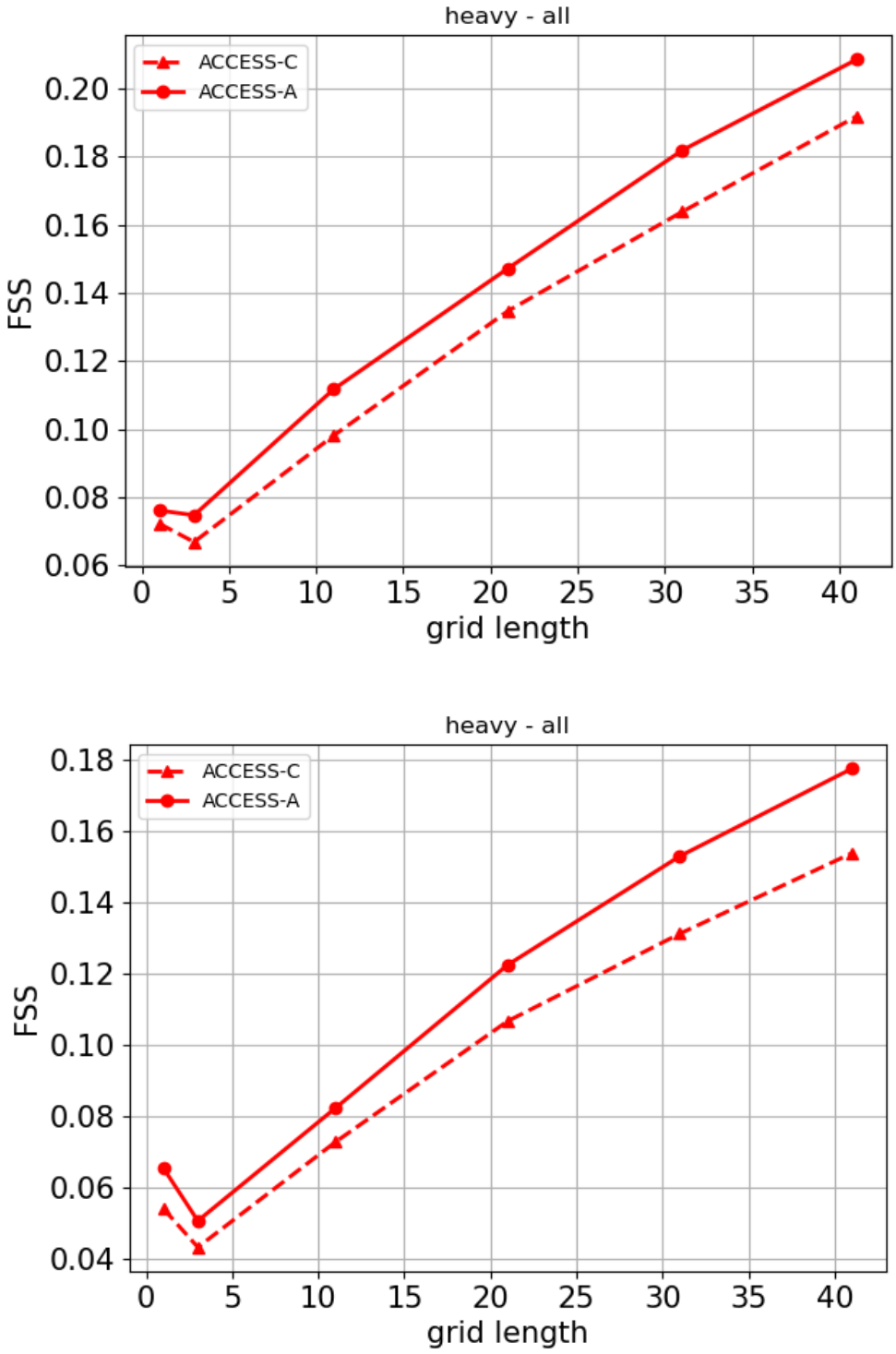

Figure 18: Heavy precipitation rate FSS as a function of neighbourhood grid length using Rainfields verifying data: summer (top) and winter (bottom).

3.2. Individual domain verification

In this section we discuss the verification results for each of the seven ACCESS-C domains. To compress the information, only the lead-time averaged verification results are presented here and plotted as scorecards. Figure 19 shows the absolute mean error percentage difference scores. The different fields verified are labelled on the vertical axis; for precipitation rates, the results obtained with the three different verifying data are all shown. The horizontal axis indicates the verification results for each of the seven ACCESS-C domains. Blue colours indicate that ACCESS-A is more skilful than ACCESS-C4 and red colours indicate that ACCESS-A is less skilful than ACCESS-C4 as measured by the verifying data. The absolute mean error (bias) scores are generally mixed across the different domains. However, the Sydney and Victoria/Tasmania domains generally have more reductions in bias (blue) compared to increases (red).

The standard deviation scorecards are shown in Figure 20. These have clearer patterns than the absolute bias scorecards. Overall, and as alluded to by the domain-aggregated results in Section 3.1, they indicate that ACCESS-A is more skilful than ACCESS-C4, particularly during summer. There are important exceptions, however. It does appear that the increase in ACCESS-A error noted for screen-level temperature (at longer lead times in Figure 5) is confined to Darwin, North Queensland, Brisbane, Perth, and Adelaide. In summer, the Sydney, Adelaide, and Victoria/Tasmania domains all have lower ACCESS-A screen-temperature errors. During winter, the Sydney, Victoria/Tasmania, and Brisbane domains have lower ACCESS-A screen temperature errors. The Darwin, North Queensland, and Perth domains have higher ACCESS-A screen-level temperatures in both periods. Precipitation rate verification displays two distinct anomalies. First, moderate precipitation in ACCESS-A has larger errors for the Brisbane, Perth, and Sydney domains in summer and winter. Second, all precipitation rates have larger errors in ACCESS-A for the tropical Darwin and North Queensland domains during winter.

Figure 21 shows the FSS percentage difference scorecard for precipitation calculated over a 3 grid-length neighbourhood. The FSS values across the different domains show that ACCESS-A precipitation forecasts are more skilful than ACCESS-C4 forecasts during summer. During winter, this is not always the case. However, generally the Perth, Sydney, Adelaide, and Victoria/Tasmania domains demonstrate higher ACCESS-A skill as measured by different rainfall rates and verifying data. The more tropical domains (Darwin, North Queensland, and Brisbane) although demonstrating higher ACCESS-A skill in general also have verification results that are poorer. For example, moderate precipitation skill as verified by GPM is lower in ACCESS-A over the Darwin domain in winter. Another example is moderate precipitation skill as measured by ERA5 during winter over the North Queensland domain. Sampling error may play a role here, particularly when the different rainfall rates are separated into light, moderate, and heavy. This may be especially true during dry seasons as is the case for the tropical domains in winter. When aggregating all rainfall rates, ACCESS-A is more skilful in all domains in summer (as verified by all data types). In winter, ACCESS-A is generally more skilful as well except for ERA5-based verification in the Brisbane domain and to a

lesser extent Rainfields verification during over the Adelaide and Victoria/Tasmania domains. It should be noted that the Rainfields data coverage is quite small compared to ERA5 and GPM, so sampling error would be more substantial for this data type. With longer neighbourhood sizes, such as 31, shown in Figure 22, it becomes harder to identify significant patterns in the verification scores over different domains, although as shown in Section 3.1, when the results are averaged over the different domains, ACCESS-A outperforms ACCESS-C4. It should be noted that the Rainfields data coverage is quite small compared to ERA5 and GPM, so sampling error would be more substantial for this data type. It should also be noted that because the different verifying data are at different resolutions, the FSS values at a particular grid length refer to different length scales. For example, 3 grid lengths in ERA 5 corresponds to 0.75° while for Rainfields, 3 grid lengths correspond to 0.06°.

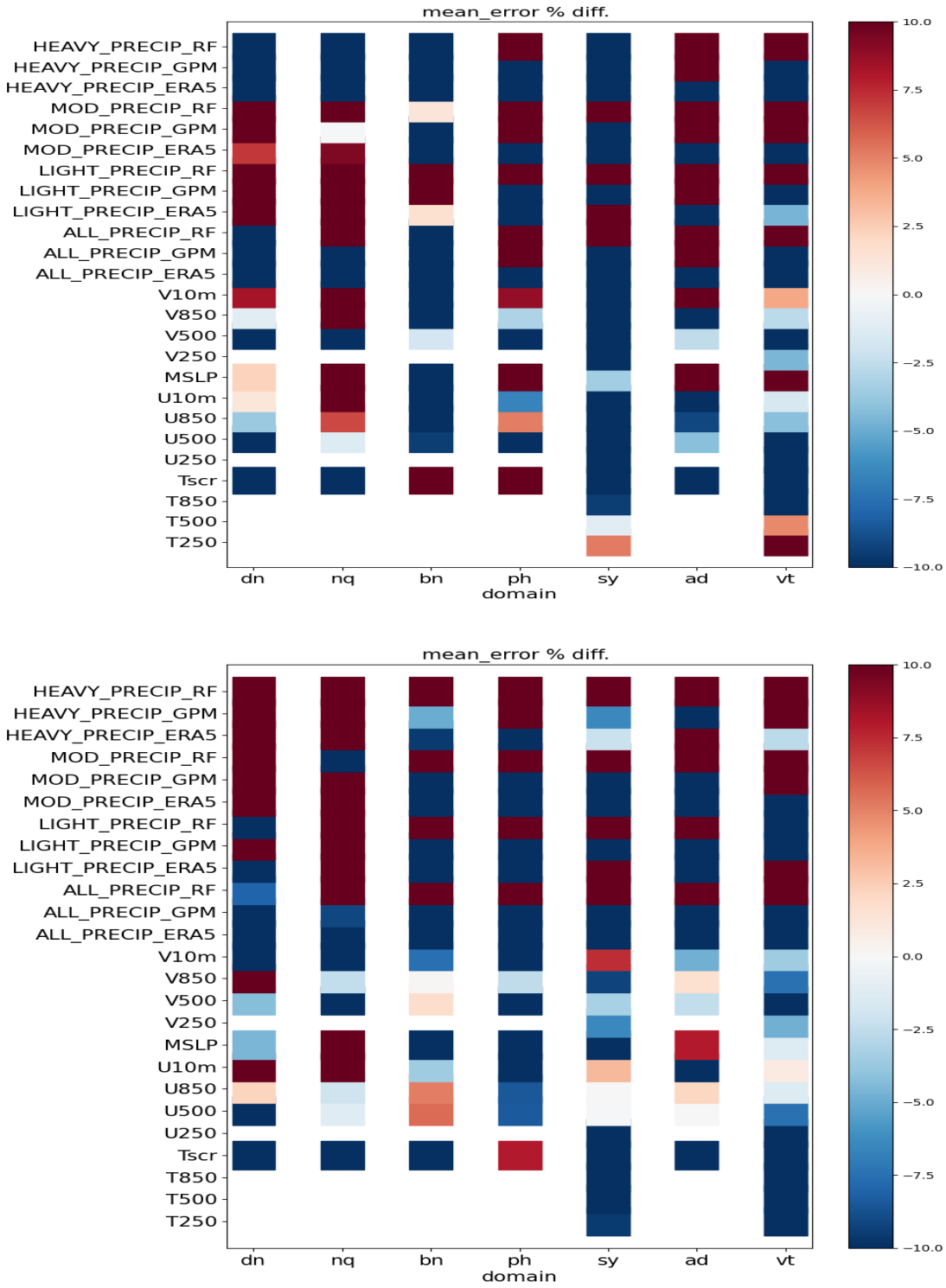

Figure 19: Lead-time averaged absolute mean error percentage difference for different fields and verifying data (vertical axis) and domains (horizontal axis) during summer (top panel) and winter (bottom panel).

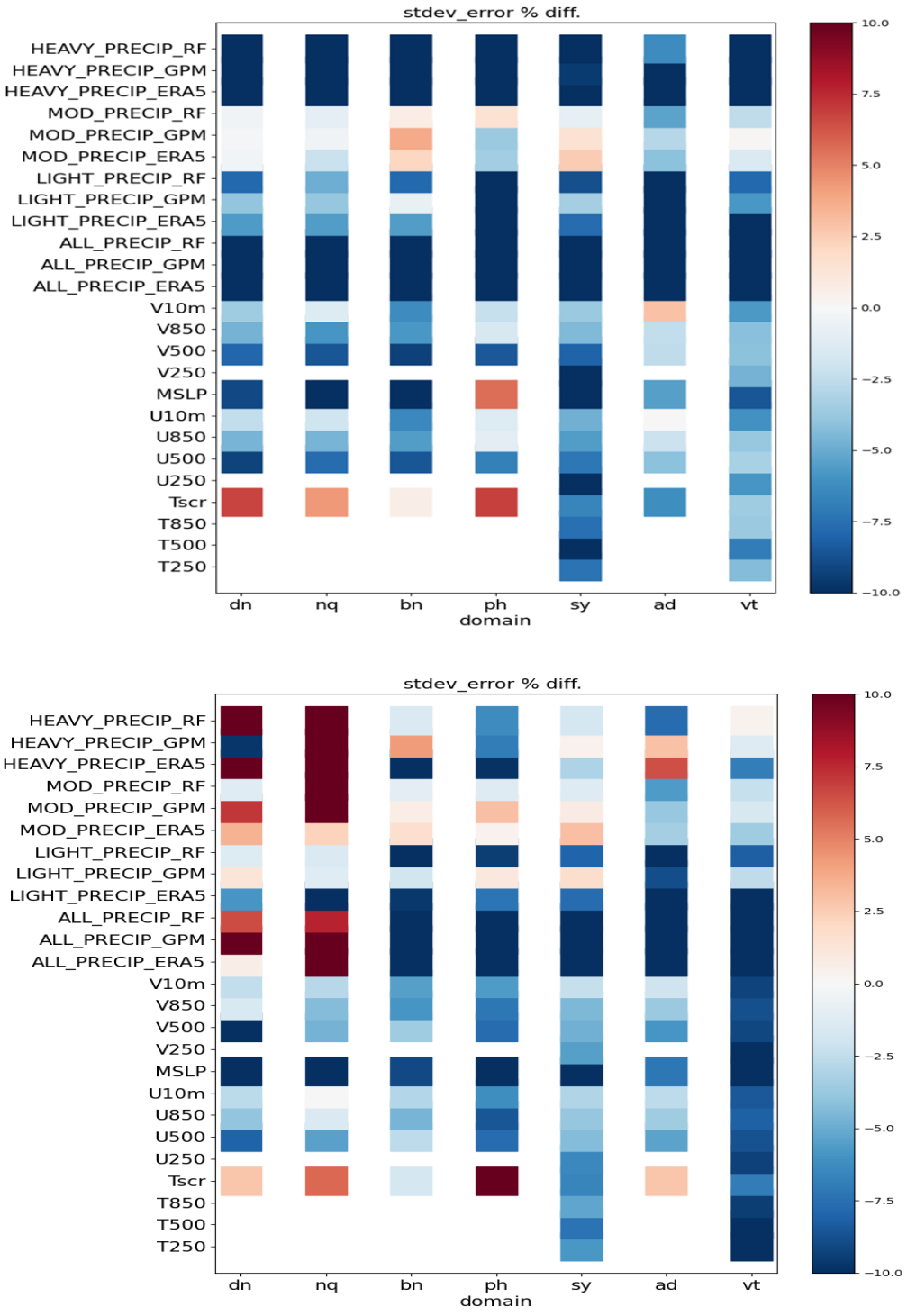

Figure 20: Lead-time averaged error standard deviation percentage difference for different fields and verifying data (vertical axis) and domains (horizontal axis) during summer (top panel) and winter (bottom panel).

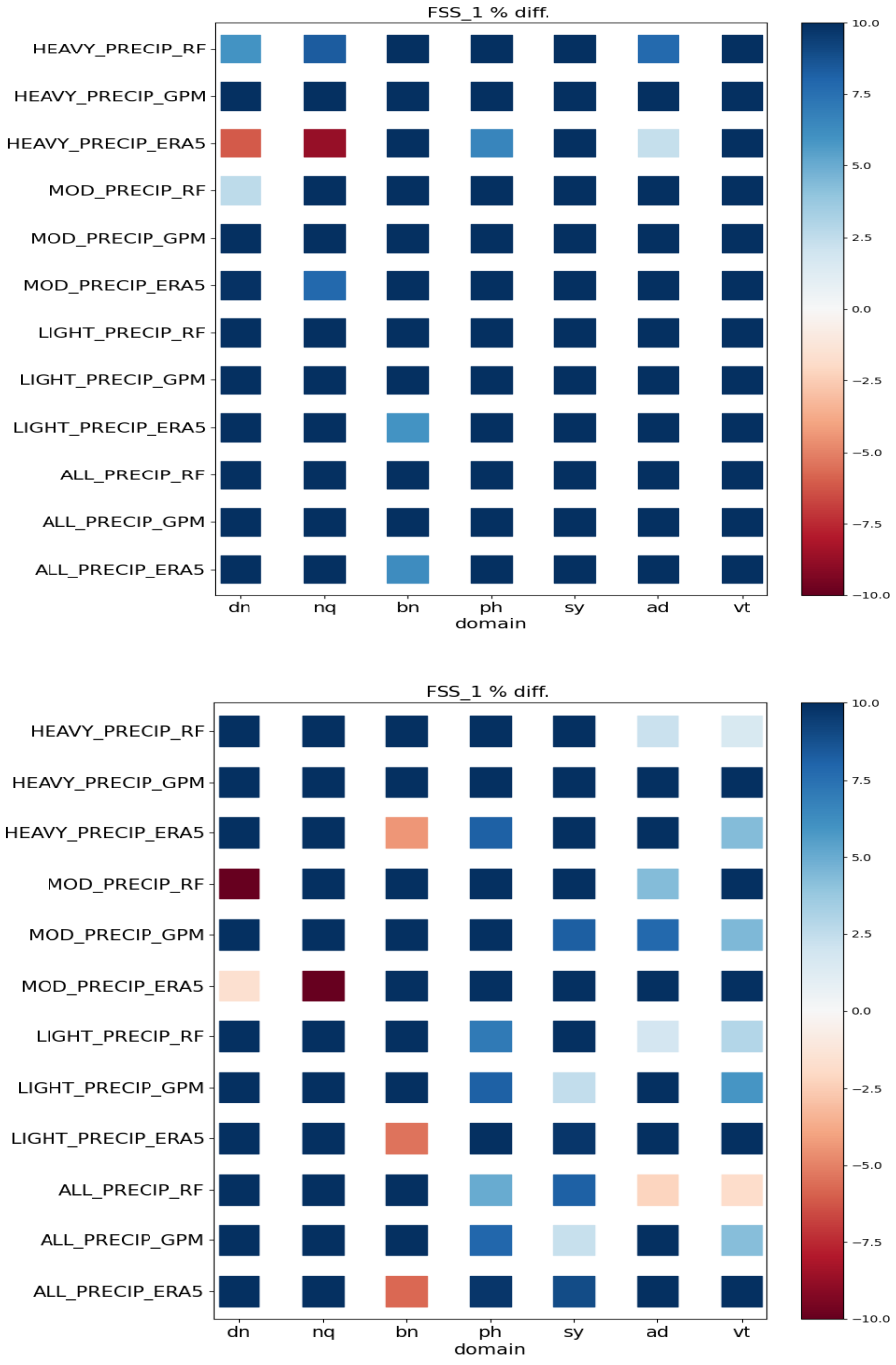

Figure 21: Lead-time averaged FSS calculated with 3 neighbouring grid lengths for different precipitation rates and verifying data (vertical axis) and domains (horizontal axis) during summer (top panel) and winter (bottom panel).

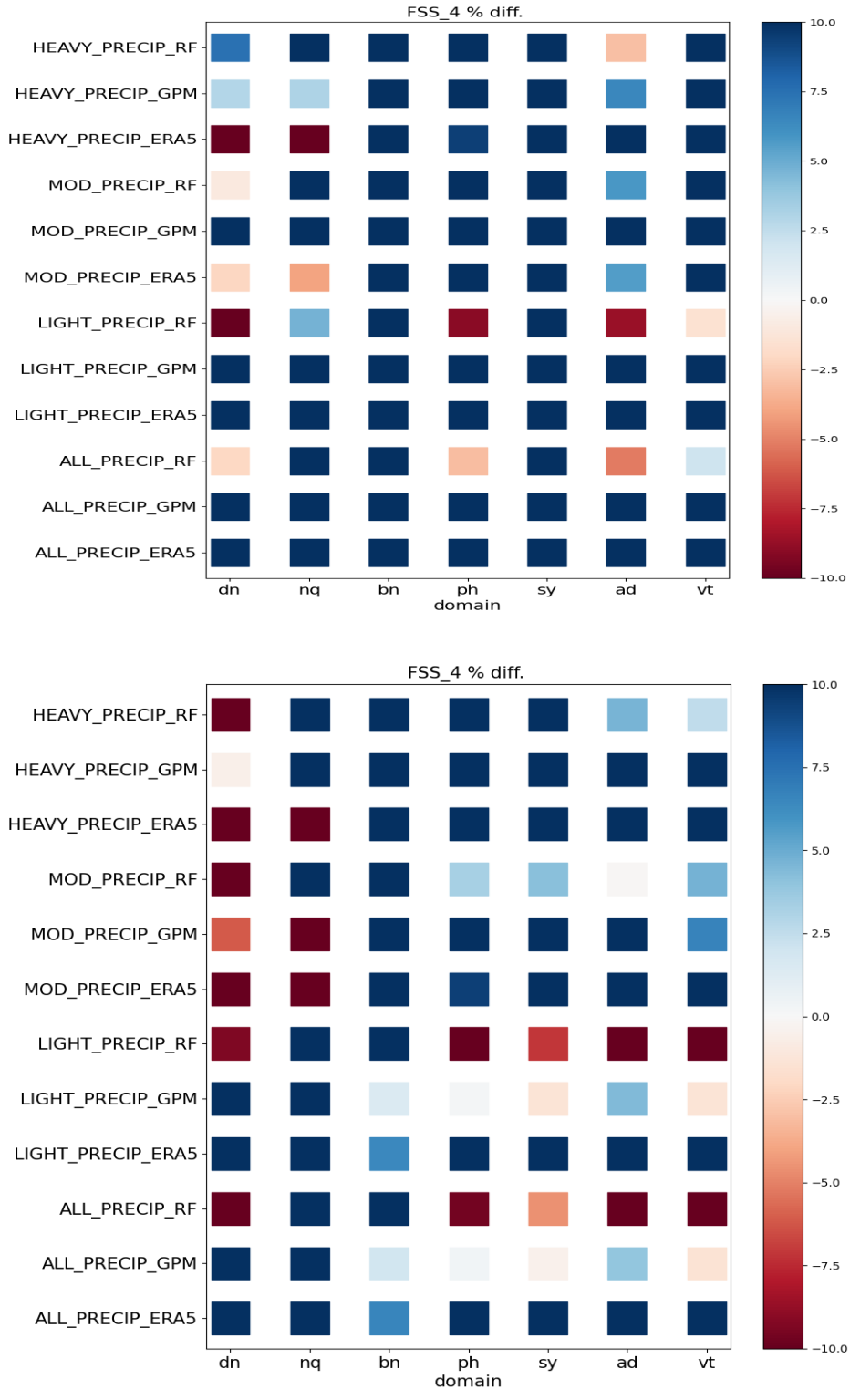

Figure 22: Lead-time averaged FSS calculated with 31 neighbouring grid lengths for different precipitation rates and verifying data (vertical axis) and domains (horizontal axis) during summer (top panel) and winter (bottom panel).

3.3. Case study: Eastern Australia Christmas day 2023 storms

The south-east of Australia (Victoria, New South Wales, and Southern Queensland) experienced severe storms from 23 to 29 December 2023, resulting in fatalities and property damage in excess of one billion dollars by some estimates (Allen, N., 2024). In this section, the performance of ACCESS-A and ACCESS-C4 precipitation rate forecasts at 0000Z on Christmas day 2023 are compared to radar data from Rainfields at selected lead times over the Victoria/Tasmania, Sydney, and Brisbane domains in order to ascertain any subjective differences in these forecasts during the event.

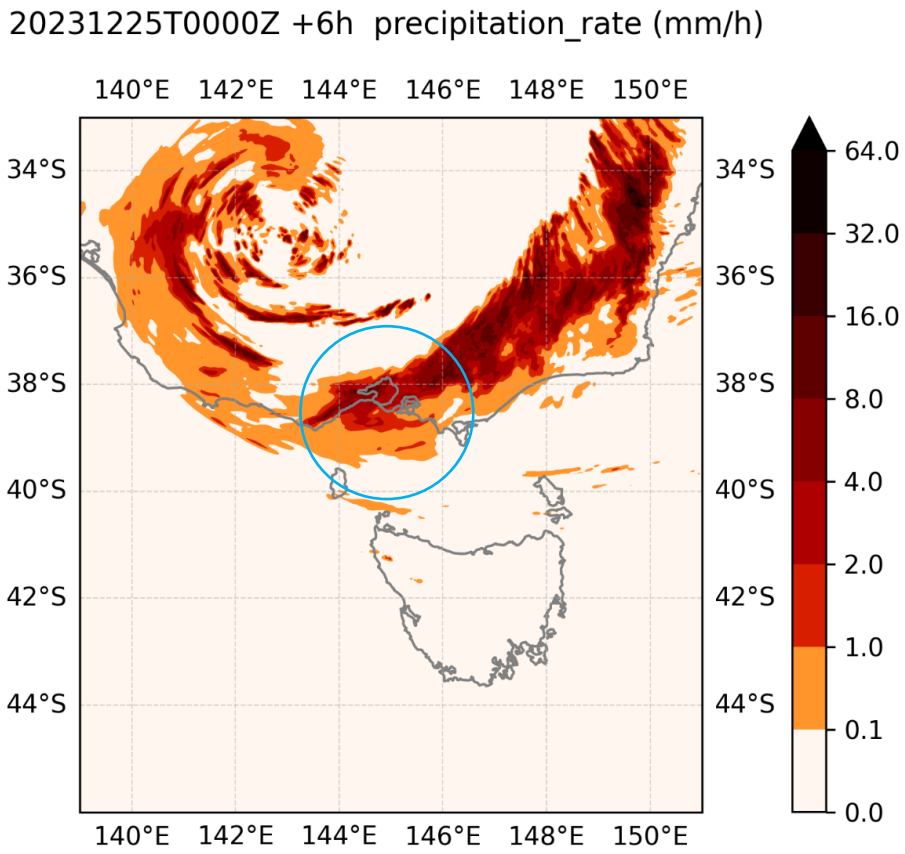

Figure 23: ACCESS-A 6h precipitation rate forecast at 0000Z on 25 December 2023 over the Victoria/Tasmania domain. The blue circle highlights an area of interest.

Figure 23 shows the ACCESS-A 6 h forecasted precipitation rate valid at 0600Z on 25 December 2023 over the Victoria/Tasmania domain. This shows a band of rainfall over southern and eastern Victoria, extending around a low-pressure system over Western Victoria. Comparing this forecast against the corresponding ACCESS-C4 forecast in Figure 24, a few differences stand out. First, the ACCESS-A forecast predicts rainfall over a somewhat larger area, with the difference particularly striking for light rainfall (< 1 mm/h). Also, ACCESS-C4 appears to predict somewhat larger rainfall rates particularly over eastern Victoria. Overall, it is not immediately clear which forecast better agrees

with the radar rainfall rates at the valid time of the forecasts (Figure 25) by visual inspection alone. However, one location that does appear to clear favour ACCESS-A is over Port Phillip Bay. ACCESS-A predicts significant rainfall over this region in agreement with the radar observations. In contrast, ACCESS-C4 does not predict significant rainfall at this location.

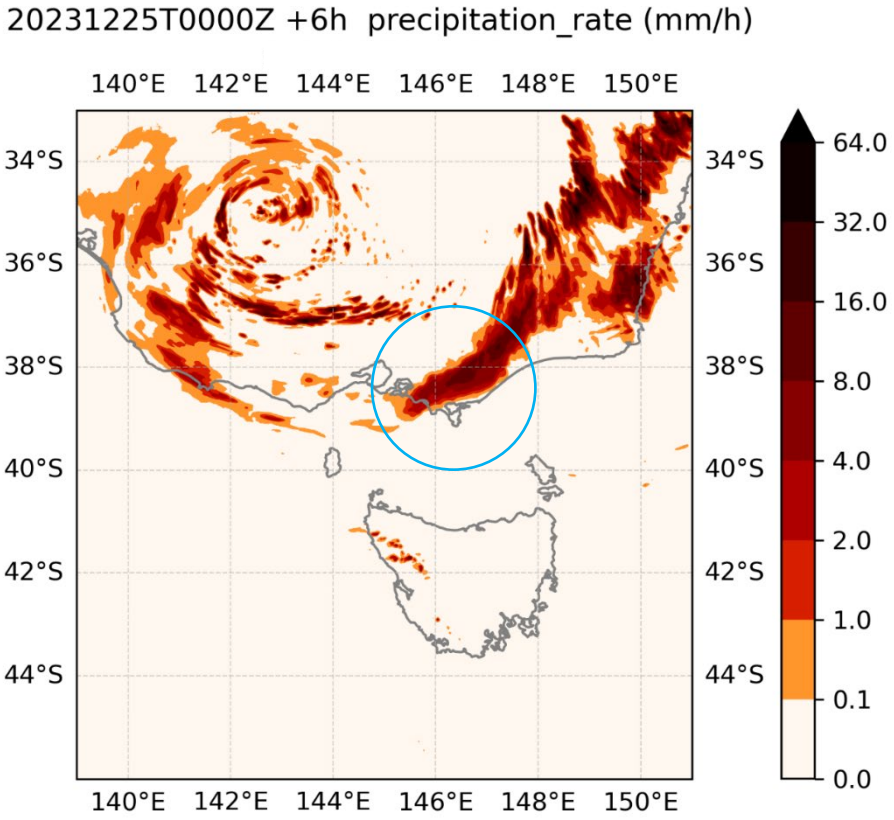

Figure 24: ACCESS-C4 6h precipitation rate forecast at 0000Z on 25 December 2023 over the Victoria/Tasmania domain. The blue circle highlights an area of interest.

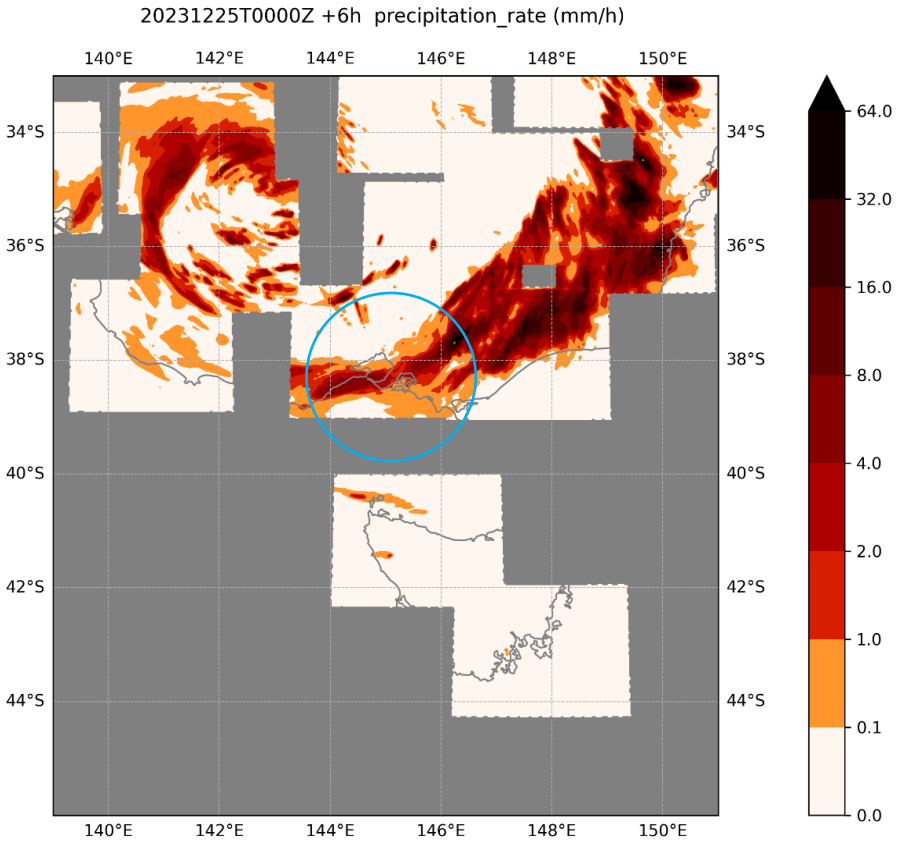

Figure 25: RainFields (Radar) observed precipitation rate at 0600Z on 25 December 2023 over the Victoria/Tasmania domain. The blue circle highlights an area of interest.

Figure 26 shows the 6 h ACCESS-A forecast (also valid at 0600Z on December 25, 2023) over the Sydney domain. This shows a continuous band of rainfall extending over south-eastern NSW, with more isolated storms north of the Sydney region. As for the Victoria/Tasmania domain, the ACCESS-C4 forecast (Figure 27), although broadly similar, predicts rainfall over a somewhat smaller region, particularly light rainfall. A striking difference is over south-eastern NSW (highlighted), where ACCESS-C4 does not appear to predict much rainfall. Radar observations at the valid time of the forecasts (Figure 28) shows that this location experienced significant rainfall in agreement with the ACCESS-A forecast.

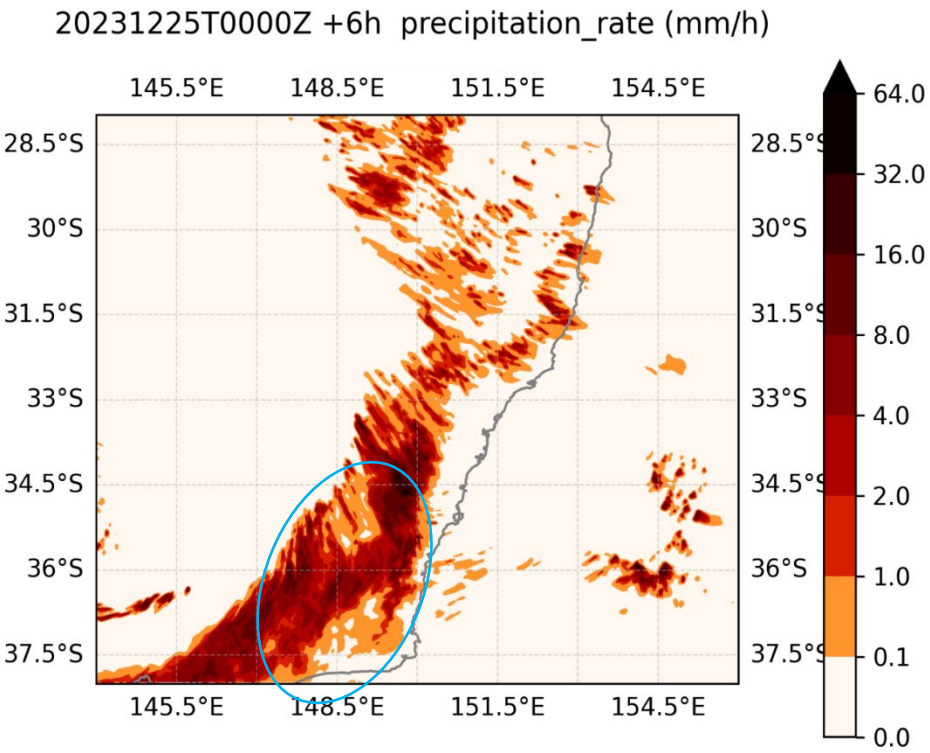

Figure 26: ACCESS-A 6h precipitation rate forecast at 0000Z on 25 December 2023 over the Sydney domain. The blue ellipse highlights an area of interest.

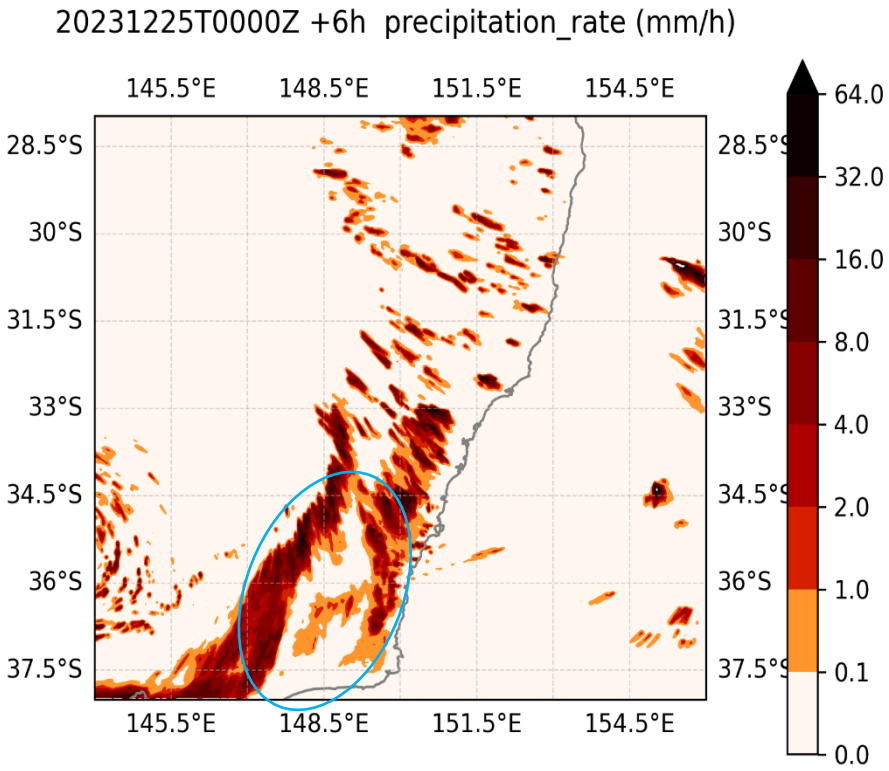

Figure 27: ACCESS-C4 6h precipitation rate forecast at 0000Z on 25 December 2023 over the Sydney domain. The blue ellipse highlights an area of interest.

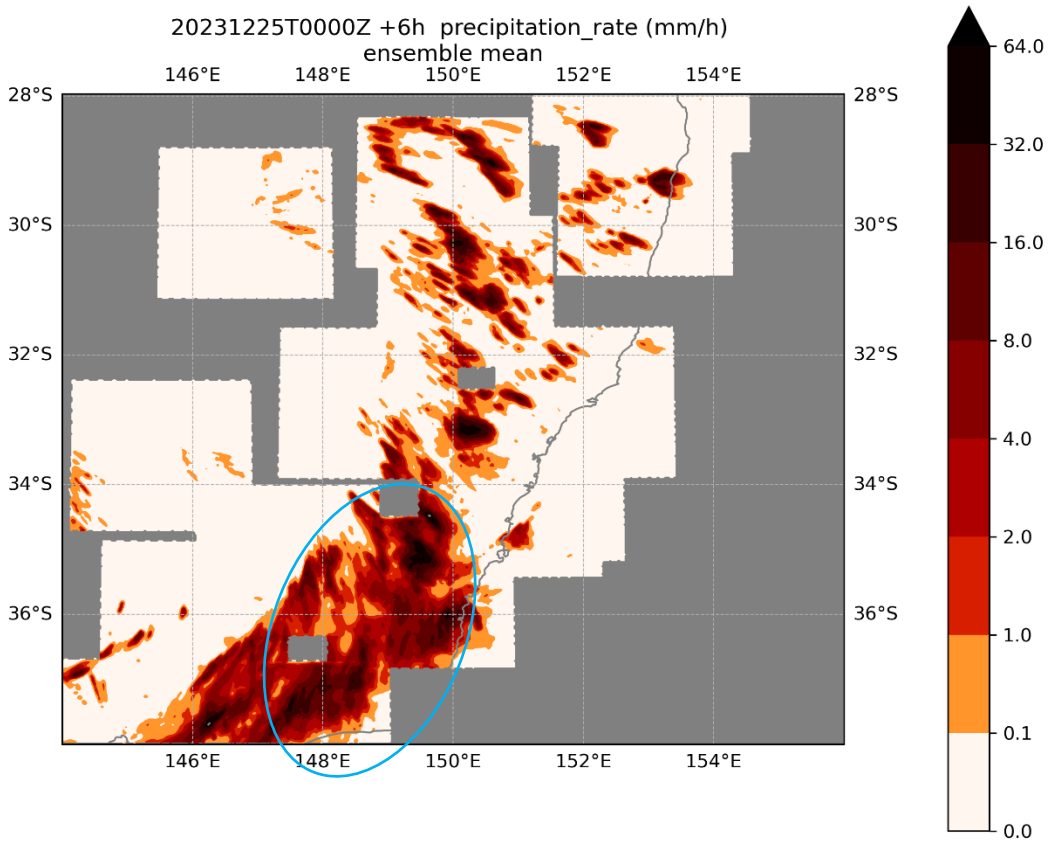

Figure 28: RainFields (Radar) observed precipitation rate at 0600Z on 25 December 2023 over the Sydney domain. The blue ellipse highlights an area of interest.

Figure 29 shows the 12 h ACCESS-A forecast valid at 1200Z on December 25, 2023, over the Brisbane domain. This shows a band of intense rainfall extending over most of eastern Queensland. In contrast, ACCESS-C4 (Figure 30) predicts a much less extensive bad of rainfall (albeit with higher intensity) around the Brisbane region. The radar observations (Figure 31) are unequivocally in better agreement with the ACCESS-A forecasts in this case as far as rainfall coverage is concerned. Significant rainfall is seen both north and south of the Brisbane city region in agreement with the ACCESS-A forecasts.

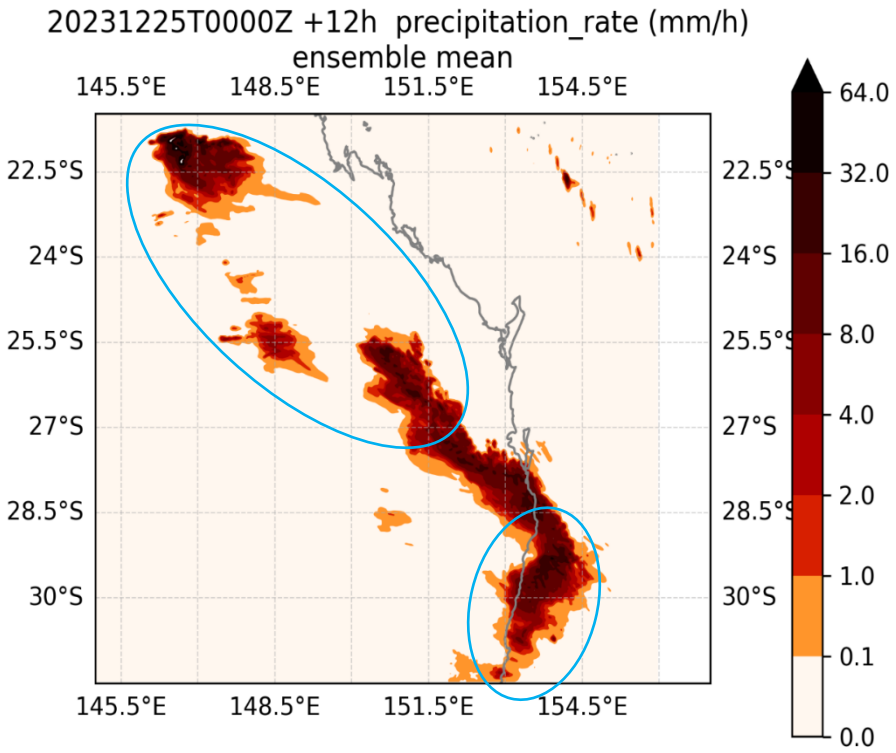

Figure 29: ACCESS-A 12 h precipitation rate forecast at 0000Z on 25 December 2023 over the Brisbane domain. The blue ellipses indicate areas of interest.

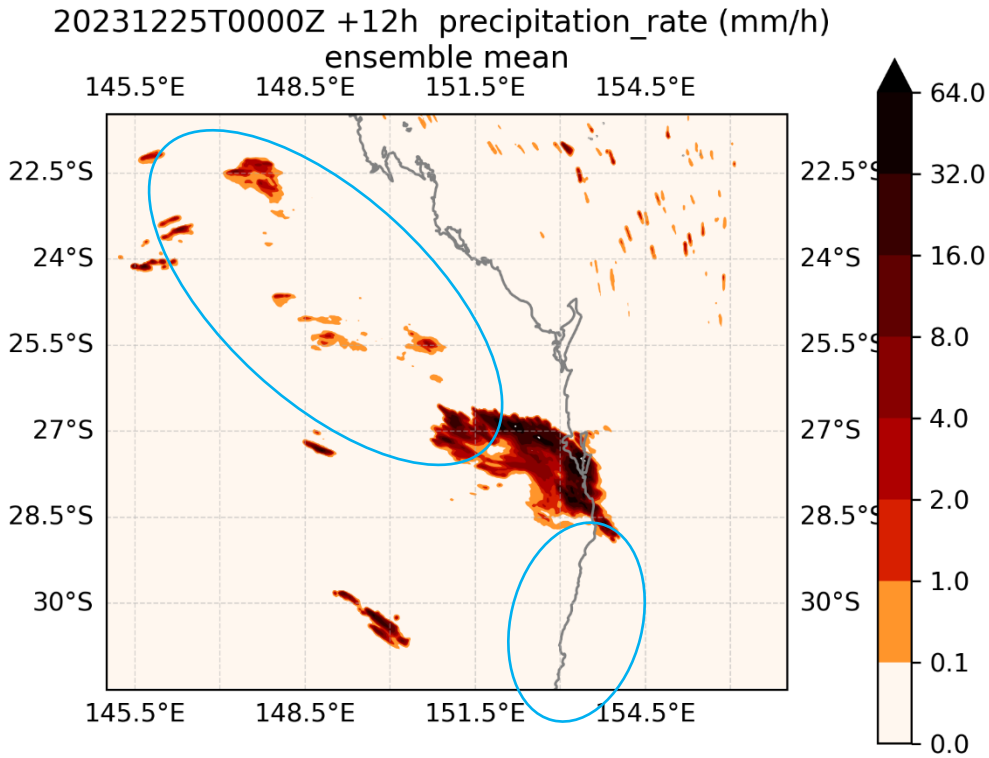

Figure 30: ACCESS-C4 12 h precipitation rate forecast at 0000Z on 25 December 2023 over the Brisbane domain. The blue ellipses indicate areas of interest.

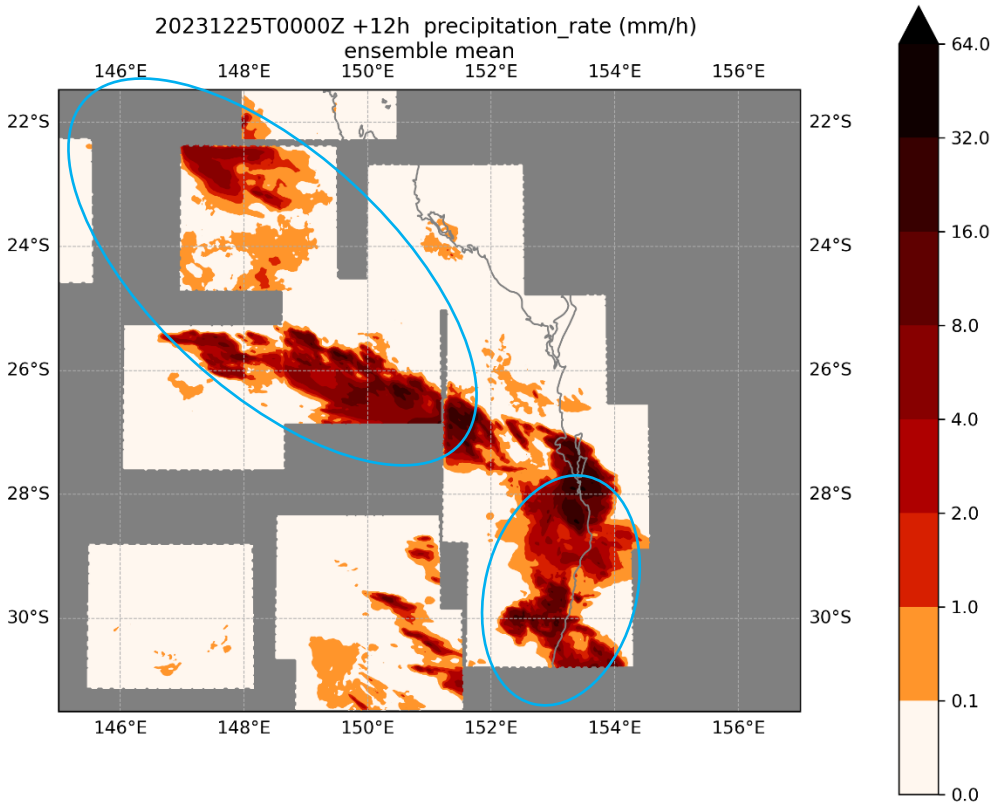

Figure 31: RainFields (Radar) observed precipitation rate at 1200Z on 25 December 2023 over the Brisbane domain. The blue ellipses indicate areas of interest.

4. Discussion

In this study the performance of ACCESS-A was compared to the performance of ACCESS-C4 using ERA5, GPM rainfall satellite, and Rainfields radar gridded data to verify both models within a 6-month long trial covering summer and winter periods. In a broad sense (aggregating the results for all fields and sub-domains), ACCESS-A clearly outperformed ACCESS-C4. This was demonstrated using a variety of metrics that measure bias, magnitude of error, and neighbourhood skill in the case of precipitation forecasts. Even though ACCESS-A verified better overall, there were some fields that were poorer or did not show much improvement across some metrics. One such field was the screen-level temperature, which displayed larger errors compared to ACCESS-C4 when verified against ERA5. Precipitation, however, appeared to verify better in ACCESS-A, particularly light precipitation. Also summer precipitation forecasts in ACCESS-A demonstrated a greater degree of improvement over ACCESS-C4 than winter precipitation forecasts, even though there were improvements in both periods. The improvement in ACCESS-A precipitation skill compared to ACCESS-C4 is consistent with the improved microphysics scheme in RAL3 (Bush, et al., 2024) and the use of large-scale blending (Zidikheri, et al., 2025).

When looking at the individual ACCESS-C4 domains, consistent with the domain-averaged results, ACCESS-A outperformed ACCESS-C4 across many domains. However, there were notable exceptions. For ACCESS-A screen-level temperature, the Sydney and Victoria/Tasmania domains had smaller errors in both seasons, but the other domains demonstrated larger errors, with Darwin, North Queensland, and Perth displaying this phenomenon in both seasons. Even precipitation, which was one of the better performing fields in ACCESS-A, demonstrated some anomalies. Errors were larger in ACCESS-A during the winter period over the tropical Darwin and North Queensland domain. This phenomenon was somewhat diminished when performing neighbourhood verification using the FSS score at small grid lengths (Figure 21), which demonstrated improved precipitation skill over most domains. The difference in performance between the tropical and midlatitude domains, particularly during winter, is possibly related to the replacement of different tropical and midlatitude physics by the unified physics scheme in RAL3 used in ACCESS-A.

Subjective evaluation of ACCESS-A precipitation forecasts against ACCESS-C4 precipitation forecasts was also performed for selected periods for the Eastern Australia Christmas Day Storms event on December 25, 2023. This demonstrated that ACCESS-A precipitation forecasts had more extensive coverage than ACCESS-C4 precipitation forecasts, particularly for light precipitation. Moreover, the precipitation coverage in ACCESS-A forecasts better agreed with the radar coverage at the valid time of the forecast. This was most striking in the Brisbane domain at 1200Z on 25 December, 2023 (Figure 29, Figure 30, and Figure 31), with ACCESS-C4 unable to pick up a significant portion of the rainfall over most of north-east and south-east Queensland. The ACCESS-A forecast, on the other hand, was able to predict the extent of the rainfall band much more successfully.

Finally, some limitations of this study should be mentioned. Apart from precipitation, ERA5 was the only dataset that was used to verify the output fields. Despite being recognized as a high-quality dataset, ERA5, being a reanalysis product, is a blend of model outputs and observations and may therefore partly reflect model biases, particularly in observation-sparse locations. In addition, the ERA5 resolution is only about 30 km, so it cannot be used to verify features below this scale. This is particularly relevant near the surface where small-scale topographic features are important. Ongoing efforts are focussed on better understanding the performance of ACCESS-A and ACCESS-C4 at these smaller scales using different observation types. It should be stressed, however, that precipitation was verified against different observation types in this study, so it is not subject to the above caveat. This study was primarily focussed on computing the relative forecast skill of ACCESS-A against ACCESS-C4. The question of whether the forecast skill improvement is 'useful' was not addressed quantitatively in the sense discussed by Antonio & Aitchson (2023). However, qualitative assessment of precipitation forecasts during the aforementioned severe weather event suggests that the improvements are indeed useful as discussed above.

5. Conclusion

This study evaluated the performance of the Bureau's Australia-wide model currently being developed, namely ACCESS-A, against the multi-domain city-scale, ACCESS-C4 model, which it is scheduled to replace operationally in the next few years. It was found that overall ACCESS-A is more skilful than ACCESS-C4 when verified against various gridded datasets. Fields such as precipitation and winds in ACCESS-A clearly outperformed corresponding fields in ACCESS-C4. However, this improvement was not demonstrated for all fields and across all sub-domains within the Australian region. For example, near-surface (screen-level) temperature was poorer in ACCESS-A over some sub-domains, especially in the tropics. Further investigation is required to understand the source of this discrepancy and whether it can be ameliorated by better tuning of the ACCESS-A system. A limitation of this study is that, apart from precipitation, forecast fields were verified against fairly coarse reanalysis data rather than observations. Ongoing efforts are focussed on verification against other data types, which will address these limitations.

References

Allen, N., 2024. *Perils.* [Online]
Available at: https://www.perils.org/files/News/2023/Loss-Announcements/Australia-Christmas-Storm/2024-04-03-PERILS-Press-Release-Australia-Christmas-Storms-23-29-Dec-2023.pdf
[Accessed 14 3 2025].

Antonio, B. & Aitchson, L., 2023. How to derive skill from the Fractions Skill Score. *arXiv preprint arXiv:2311.11985.*

Benson, S., Bannister, T. & Krysta, M., 2022. *Evaluating the advantages and challenges of the National Analysis System,* Melbourne: Australian Bureau of Meteorology.

Bush, M. et al., 2020. The first Met Office Unified Model–JULES regional atmosphere and land configuration, RAL1.. *Geoscientific Model Development ,* 13(4), pp. 1999-2029.

Bush, M. F. et al., 2024. Unifying mid-latitude and tropical regional model configurations: the third met Office unified model–JULES regional atmosphere and land configuration, RAL3. *Geoscientific Model Development Discussions,* pp. 1-58.

Hersbach, H. et al., 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society,* 146(730), pp. 1999-2049.

Huffman, G. et al., 2019. *The GPM Integrated Multi-satellitE Retrievals for GPM (IMERG) Final Run Version 06.,* s.l.: NASA Goddard Space Flight Center..

Milan, M. et al., 2023. Large-scale blending in an hourly 4D-Var framework for a numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society,* 149(755), pp. 2067-2090.

Milan, M. et al., 2020. Hourly 4D-Var in the Met Office UKV operational forecast model. *Quarterly Journal of the Royal Meteorological Society,* 146(728), pp. 1281-1301.

Rennie, S. et al., 2022. ACCESS-C: Australian convective-scale NWP with hourly 4D-Var data assimilation.. *Weather and Forecasting,* 37(7), pp. 1287-1303.

Roberts, N. & Lean, H., 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review,* 136(1), pp. 78-97.

Seed, A., Leahy, C., Dutie, E. & Chumchean, S., 2008. *Rainfields: the Australian Bureau of Meteorology system for quantitative precipitation estimation, and it's use in hydrological modelling..* s.l., Proceedings of Water Down Under..

Zidikheri, M., Krysta, M., Rennie, S. & Steinle, P., 2025. *Assessing the impact of large-scale blending on the Bureau's NAS and ACCESS-A regional forecasting models,* s.l.: Bureau of Meteorology.