

# Evaluation of global and regional reanalyses performance over New Zealand

A.A.S. Pirooz<sup>1</sup>, S. Moore<sup>1</sup>, T. Carey-Smith<sup>1</sup>, R. Turner<sup>1</sup>, C.-H. Su<sup>2</sup>

<sup>1</sup>NIWA-Taihoru Nukurangi, 301 Evans Bay Parade, Ākautangi, Hātaítai, Wellington, Aotearoa New Zealand

<sup>2</sup>Bureau of Meteorology, Docklands, Victoria 3008, Australia

Correspondence: stuart.moore@niwa.co.nz, amir.pirooz@niwa.co.nz

**Key words:** Regional reanalysis, numerical weather prediction, BARRA, ERA, verification

---

## Abstract

Atmospheric reanalyses provide long-term physically consistent and spatially complete records of meteorological variables that can be used to study recent past climate trends and establish extreme event climatologies. In this work, three global reanalysis products, namely ERA-Interim, ERA5 and ERA5-Land along with a 12-km resolution regional reanalysis, BARRA-R, are compared against point-based and gridded observational data across New Zealand for the period 2014 – 2018 for four variables; mean and gust wind speeds, total precipitation, and surface air temperature. It is generally found that the skill of the reanalyses decays over the mountainous regions of the South Island, in high wind regions such as the Cook and Foveaux Straits, and in high rainfall regions like the west coast of the South Island. The study demonstrates that the BARRA-R reanalysis performs better for precipitation and temperature than ERA reanalysis products over New Zealand. BARRA-R outperforms ERA-Interim in predicting gust wind speeds. However, unlike ERA5, BARRA-R does not capture the frequency of high gust wind speeds over the more southern mountains of the South Island but does produce higher gust speeds in the lower North Island.

## 1. Introduction

Although meteorological observing stations generally provide valuable and reliable data, they are usually located far from one another, inhibiting high quality spatial analyses. Such data sparsity, the effects of local orography and regional climate variability can make interpolation techniques unreliable and diminish their value (Tait et al., 2006). Accurate spatial estimation of climate variables is essential for many applications and studies, such as extreme value analysis and engineering designs (Mo et al., 2015; Xu et al., 2020), energy resources assessment (Frank et al., 2018; Miao et al., 2020), hydrological models (Essou et al., 2017), enhancement of climate projection models and evaluation of climate

change effects (Di Virgilio et al., 2019; Avila-Diaz et al., 2020). In addition, historical observational data are often available only for a limited period or are not continuous. Therefore, alternative datasets, such as numerical model-derived reanalysis products, which can provide physically consistent and long-term spatially complete records of climate variables, can be employed to fill the gaps in observations time series (Dee et al., 2011; Tetzner et al., 2019).

Reanalysis products are developed utilising numerical weather prediction (NWP) models and data assimilation techniques to generate observation-constrained model estimates of climate variables (Su et al., 2019). Several global and regional reanalysis datasets with different

spatial and temporal resolutions are currently available. Early global reanalysis products, such as European ReAnalysis-40 (ERA40) (Uppala et al., 2005) employ a coarse resolution model that is unable to capture well the sub-grid variations of variables and small-scale processes, particularly over complex terrain (Su et al., 2019). Regional reanalyses are often developed by downscaling from a coarser resolution global-scale model to a higher resolution limited domain and therefore provide better temporal and spatial representations of meteorological fields, improving variability and the representation of extremes and frequency distributions.

Recently, Su et al. (2019) developed the first regional reanalysis covering a large region of Oceania, including Australia, New Zealand, and southeast Asia, which has a horizontal resolution of 12 km and uses ERA-Interim as the driving model. Compared to global reanalyses, the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis (BARRA-R) showed better statistical scores and agreement with respect to observations over Australia, particularly in estimating temperature, wind speed, surface pressure, precipitation and frequency of heavy rain events. Later, Su et al. (2021) created the BARRA-C reanalysis dataset by downscaling BARRA-R to 1.5 km horizontal resolution over four major Australian cities. At this scale wind and temperature performance, particularly over complex and coastal regions, was further enhanced.

As complementary datasets to observations, reanalyses provide invaluable information for meteorological and

climatological studies, particularly where observations are scarce. However, the performance of each reanalysis needs to be assessed to fully understand their limitations and uncertainties associated with their resolution or NWP science.

In this study, we evaluate the performance over New Zealand of the BARRA-R regional reanalysis (Su et al., 2019) and that of three commonly used global reanalysis products, namely European ReAnalyses (ERA) products of the European Centre for Medium-Range Weather Forecasts (ECMWF): ERA-Interim (Dee et al., 2011), ERA5 (Hersbach et al., 2020) and ERA5-Land (Muñoz Sabater, 2019). In Section 2, we outline the observation and reanalysis datasets used and the statistical methods applied in our evaluation. We present results in Section 3, focusing on precipitation, air temperature, and mean and gust wind speed performance over relatively flat and complex mountainous regions before summarising our findings in Section 4.

## 2. Methodology and data

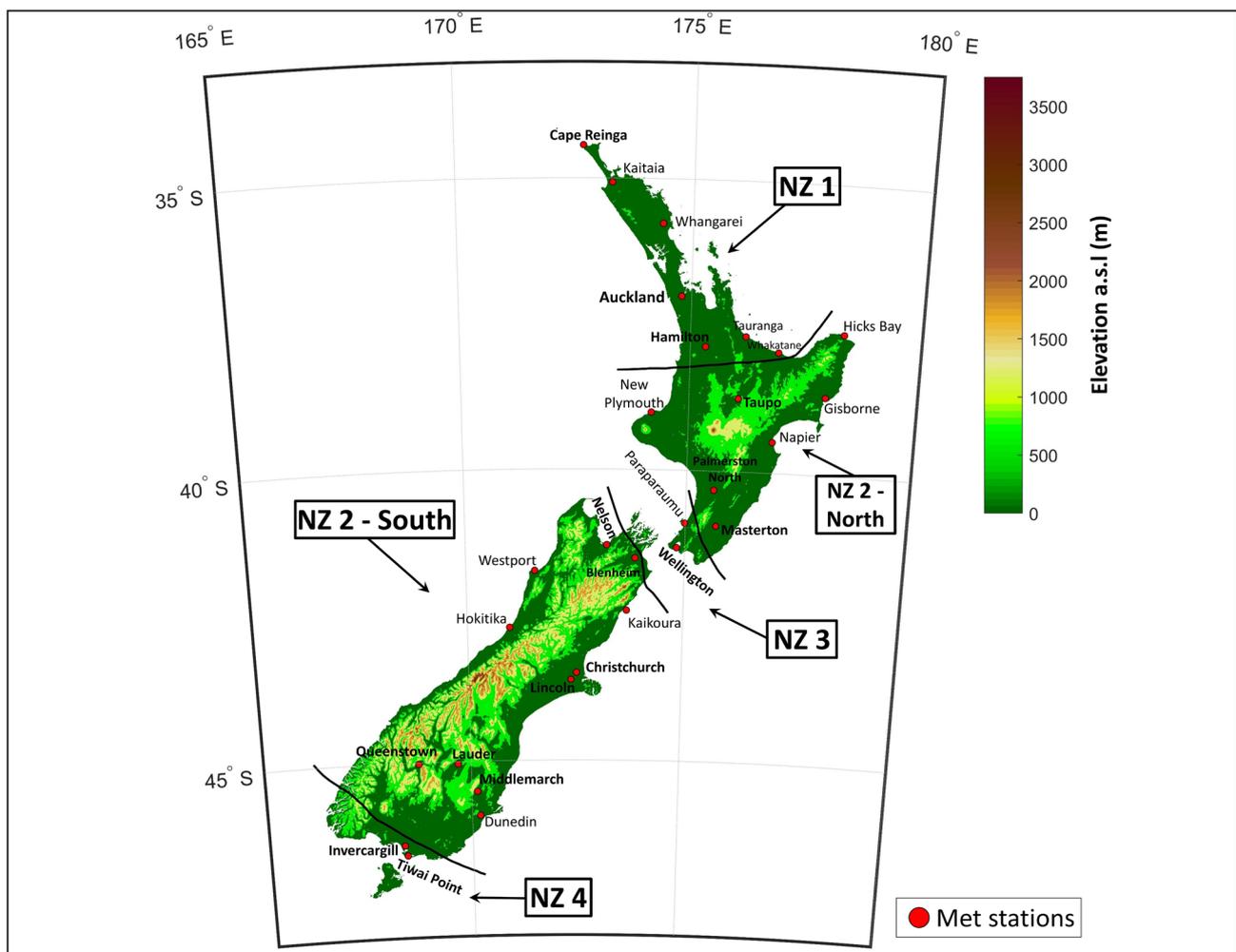
In this work we concentrate on the five-year period, 2014 to 2018, using data from ERA-Interim, ERA5, ERA5-Land and BARRA-R. Table 1 summarises the main specifications of the reanalysis products used in this study. ERA5 is the ECMWF's latest climate reanalysis product replacing ERA-Interim. Compared with ERA-Interim, ERA5 benefits from 10 years of developments in the model physics and data assimilation, including the use of satellite data in NWP and atmospheric modelling (Hersbach et

**Table 1:** Global- and regional- scale reanalysis products used in this study.

	Name	Reference	Spatial Resolution	Temporal Resolution	Temporal Availability	Model Cycle
Global	ERA-Interim	(Dee et al, 2011)	80 km	3-hourly	1979 – present	Cy31r2 (2006)
	ERA5	(Hersbach et al, 2020)	31 km	Hourly	1950 – present	Cy41r2 (2016)
	ERA5-Land	(Muñoz-Sabater, 2019)	9 km	Hourly	1981 – present	Cy45r1 (2018)
Regional	BARRA-R	(Su et al, 2019)	12 km	Hourly	1990 – present	UM 10.2

al., 2020; Muñoz-Sabater et al., 2021). The ERA products are based on the ECMWF Integrated Forecasting System (IFS) models, as outlined in Table 1. ERA5-Land is an enhanced global dataset for the land component of ERA5, and shares most of its parametrisations with ERA5 (Muñoz-Sabater, 2019; Muñoz-Sabater et al., 2021). The higher horizontal resolution along with the non-linear dynamical downscaling with corrected thermodynamic input are the main advantage and improvement of ERA5-Land over its driving model, ERA5. Unlike the ERA products, BARRA-R uses the Unified Model (UM) (Davies et al., 2005) and is initialised using ERA-Interim and is forced with lateral boundary conditions derived from ERA-Interim; for more details see Su et al. (2019). All the reanalyses considered in this study, except

for ERA5-Land, use a 4-dimensional variational data assimilation technique. Although ERA5-Land does not assimilate observational data directly, the observations affect the atmospheric forcing derived from ERA5, which drives ERA5-Land. In addition, BARRA-R has the benefit of assimilating additional observations over New Zealand from the National Climate Database (Cliflo, 2020), which may not be assimilated in other reanalyses. In BARRA-R, at each 6-hr analysis, observations valid from  $\pm 3$  hr were used in the data assimilation scheme. However, the validation in this study is done using BARRA-R forecast data valid from 3 hr to 9 hr after the analysis time. Thus, the validating observations are independent of those used in the data assimilation.



**Figure 1:** Orography map of New Zealand (1-km resolution) along with the locations of meteorological stations considered in this study. The New Zealand wind regions are also labelled.

The evaluation is made against two sets of observation data:

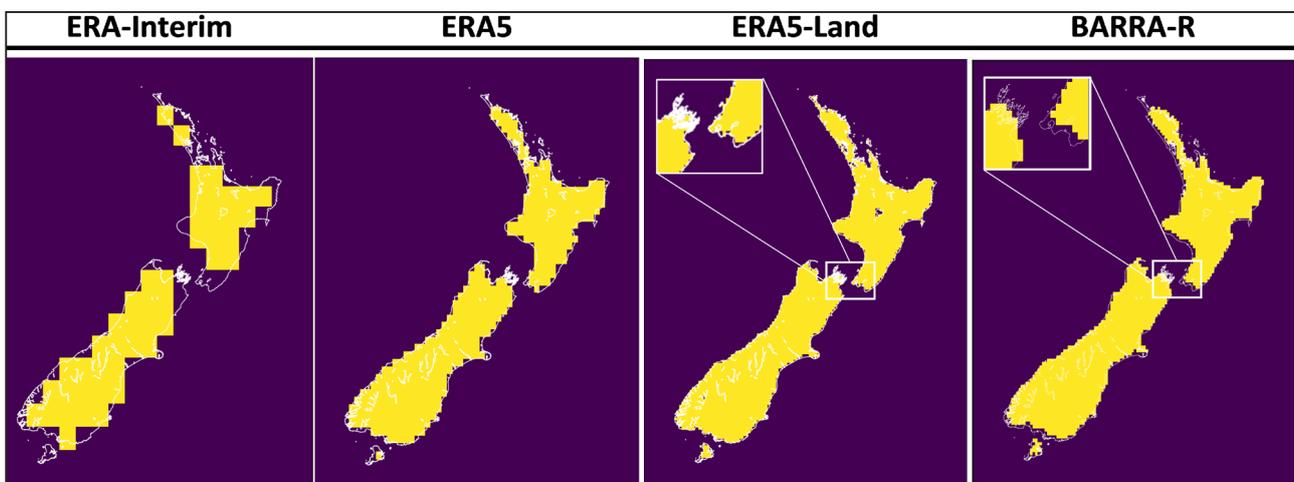
- Point-based meteorological station data recorded at selected stations around New Zealand and available from NIWA's online CliFlo climate database (Cliflo, 2020), and
- Virtual Climate Station Network (VCSN) data (Tait and Turner, 2005; Tait et al., 2006; Tait et al., 2012), which comprises 2D surfaces of climate variables (excluding gust wind speed) covering all of New Zealand on a 5-km grid based on the spatial interpolation of actual observation data at climate stations.

Station-based observations are available at an hourly temporal resolution, while VCSN only provides daily data. Further, VCSN provides two sets of rain data. The first, hereafter called VCSN, is the original rain field. More recently, a bias corrected rain field was added to VCSN, hereafter called VCSN\_BC, which was corrected by incorporating more regional council station data into the original VCSN. The statistical scores for total precipitation were calculated between the reanalysis data and VCSN\_BC.

In the case of mean and gust wind speeds, where often the measurements have not been taken following standard (World Meteorological Organisation, 2014), the retrieved observations were homogenised and converted to a common standard, namely 10-m height and 3-s gust duration, using the procedure proposed by Turner et al. (2019), Safaei Pirooz et al. (2020a) and Safaei Pirooz et al. (2020b).

Region NZ1 is relatively flat and covers the northern part of the North Island (NI). The NZ2 region consists of parts of both the North and South Islands, which have different orography. As shown in Figure 1, the South Island (SI) has steeper and more complex terrain compared to NI. NZ4 is exposed to strong westerly and south-westerly winds, and NZ3, due to the channelling effect of Cook Strait (Turner et al., 2019; Safaei Pirooz et al., 2020b), experiences strong westerly and north-westerly winds.

Land-sea masks of the reanalyses are shown in Figure 2 to better understand the performance of the models over different regions and coastal areas. It is clear that BARRA-R misses some land points in coastal regions, particularly in regions NZ3 and NZ4. In contrast, ERA5, and particularly ERA5-Land, product better represent the true coastline. By default, ERA products return a fractional land-sea mask. For the coastline comparison, the ERA's land-sea



**Figure 2:** Land-sea masks of the reanalysis products used in this study.

masks were processed to generate binary land-sea masks, according to the guidelines provided in <https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation>; that is where land-sea mask value is greater than 0.5, set to a value of 1 (land), and where land-sea mask value is equal to or less than 0.5, set to 0 (ocean).

In carrying out the assessment, the closest reanalysis grid point to the meteorological stations were determined. Then, three statistical tests are performed (Tetzner et al., 2019; Minola et al., 2020).

- Pearson's correlation coefficient, which measures the strength of a linear association between two time series and is calculated using Eq. 1, where  $X$  and  $Y$  are the reanalysis and observation data,  $cov()$  and  $var()$  are the calculated covariance and variance, respectively. The Pearson correlation coefficient can take values from  $-1$  to  $+1$ , with values closer to either  $-1$  or  $+1$ , depending on negative or positive relationship, showing a stronger linear association of the two variables. Values close to zero indicates there is no association between the two variables. Values between  $-1$  and  $+1$ , e.g. 0.7, show that there is variation around the line of best fit. To evaluate weather-generated variability, rather than climatological cycles, it is essential to remove periodicity in wind speed time series, which for the hourly and daily series can be done by subtracting mean diurnal and monthly series, respectively.

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X) \cdot var(Y)}} \quad (1)$$

- Root Mean Square Error (RMSE): shows how far or close the reanalysed values are from the observed values. RMSE is calculated by Eq. 2, in which  $X$ ,  $Y$  and  $T$  are the reanalysis, observed values, and number of data points, respectively.

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (X - Y)^2}{T}} \quad (2)$$

- Bias shows the deviation of the reanalysis data from the observations.

$$Bias = \frac{\sum_{t=1}^T (X - Y)}{T} \quad (3)$$

### 3. Results

#### 3.1 Total precipitation

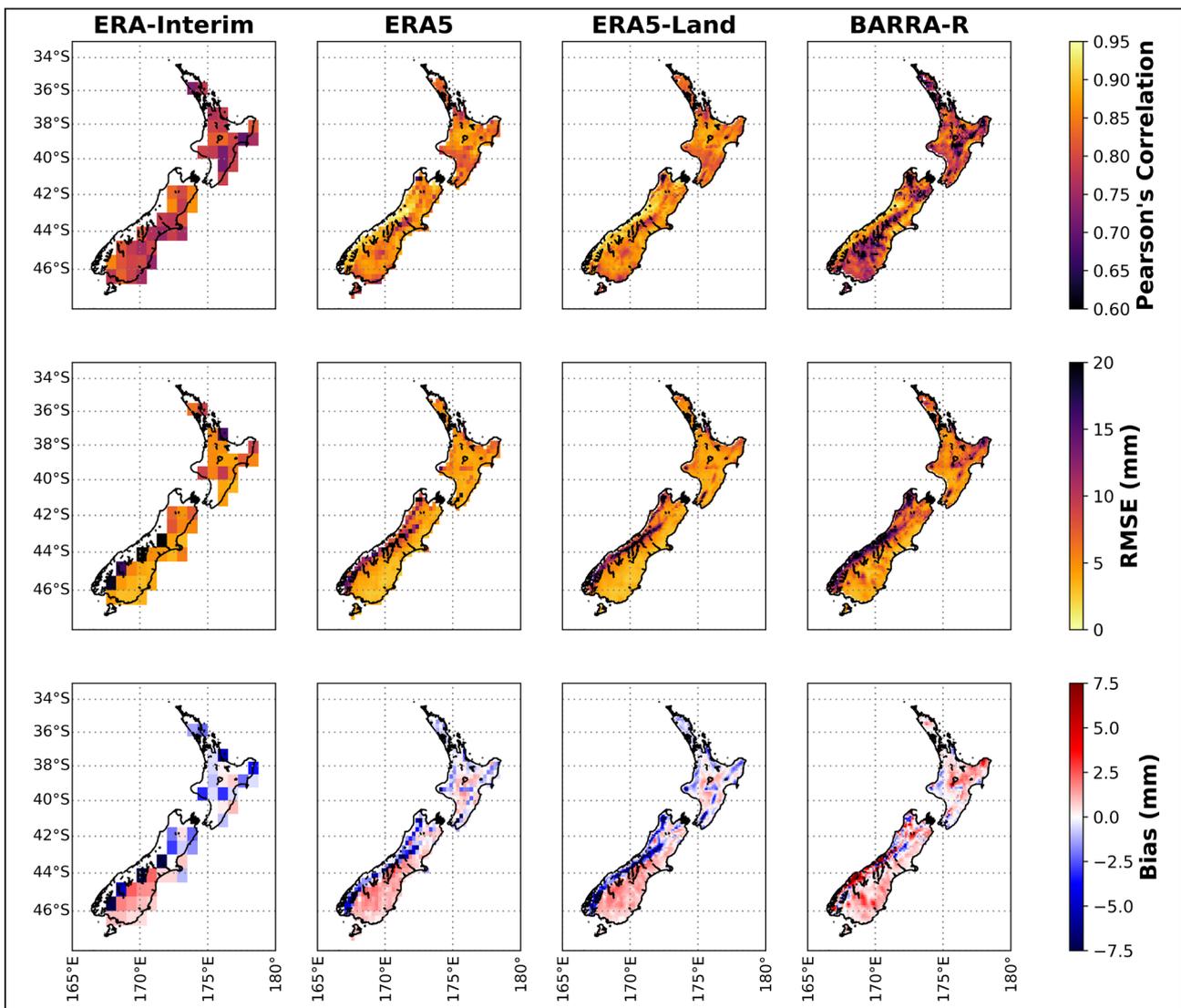
Total daily precipitation from each reanalysis dataset are assessed against VCSN\_BC data in Figure 3. ERA-Interim shows the lowest correlation, ranging from 0.7 – 0.8 across the country. ERA5 and ERA5-Land have correlations generally greater than 0.8, while the values drop over areas of complex terrain. BARRA-R depicts lower correlations, compared with ERA5 and ERA5-Land, of around 0.6 – 0.8 over the mountains and localised orography, which could be attributed to the different model physics and data assimilations used by these models. The precipitation is not corrected between ERA5 and ERA5-Land (Muñoz-Sabater et al., 2021). Thus, the performance of ERA5 and ERA5-Land is expected to be similar and marginal differences could be due to the grid resolution and the downscaling used in ERA5-Land.

All the considered reanalyses show high RMSE values, exceeding 15 mm, along the west coast of the SI. In terms of mean biases, ERA5 and ERA5-Land are generally negatively biased over complex terrain, particularly in the west part of the SI where the mean bias is about  $-7.5$  mm. On the other hand, BARRA-R mostly shows positive bias over the mountainous regions. VCSN has been determined to underestimate the total precipitation, particularly over high-elevation terrain (Tait et al., 2006; Ackerley et al., 2012). Thus, the apparent overestimation of precipitation by BARRA-R, especially on the west coast of SI, is not believed to necessarily indicate poor performance of the model. On the eastern side of the SI, all the reanalyses are positively biased, with values mostly range 0.0 to 2.5 mm. These scores are clearly tied to the

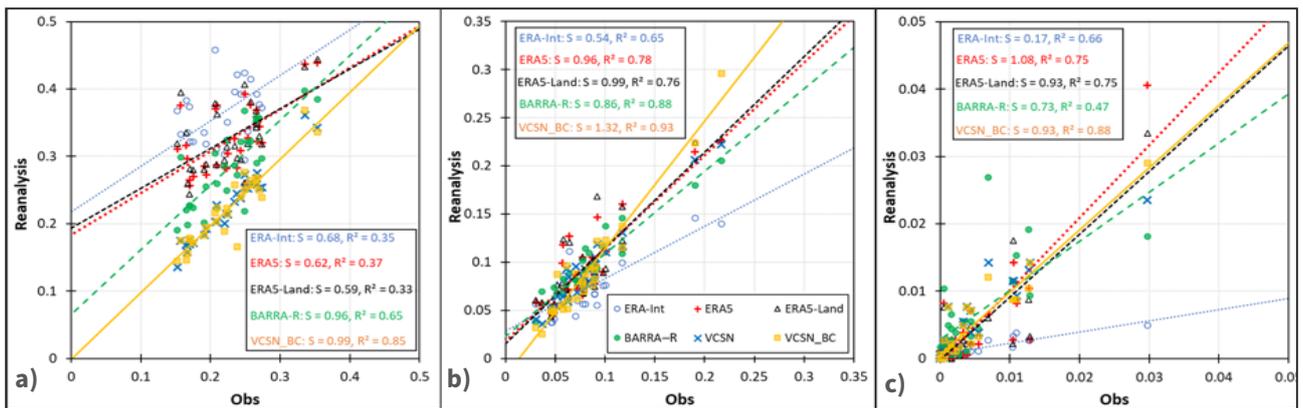
underlying terrain of the Southern Alps and is likely indicative of the reanalysis model's inability to correctly simulate the orographically enhanced rainfall in this region. In the NI, where the topography effects are less significant, the mean bias values are generally within  $-2.5$  mm to  $2.5$  mm.

The frequency of days with precipitation intensities between 1 mm to 10 mm, 10 mm to 50 mm, and greater than 50 mm, is shown in Figure 4. In these graphs, the X-axis shows the station observations and Y-axis is the

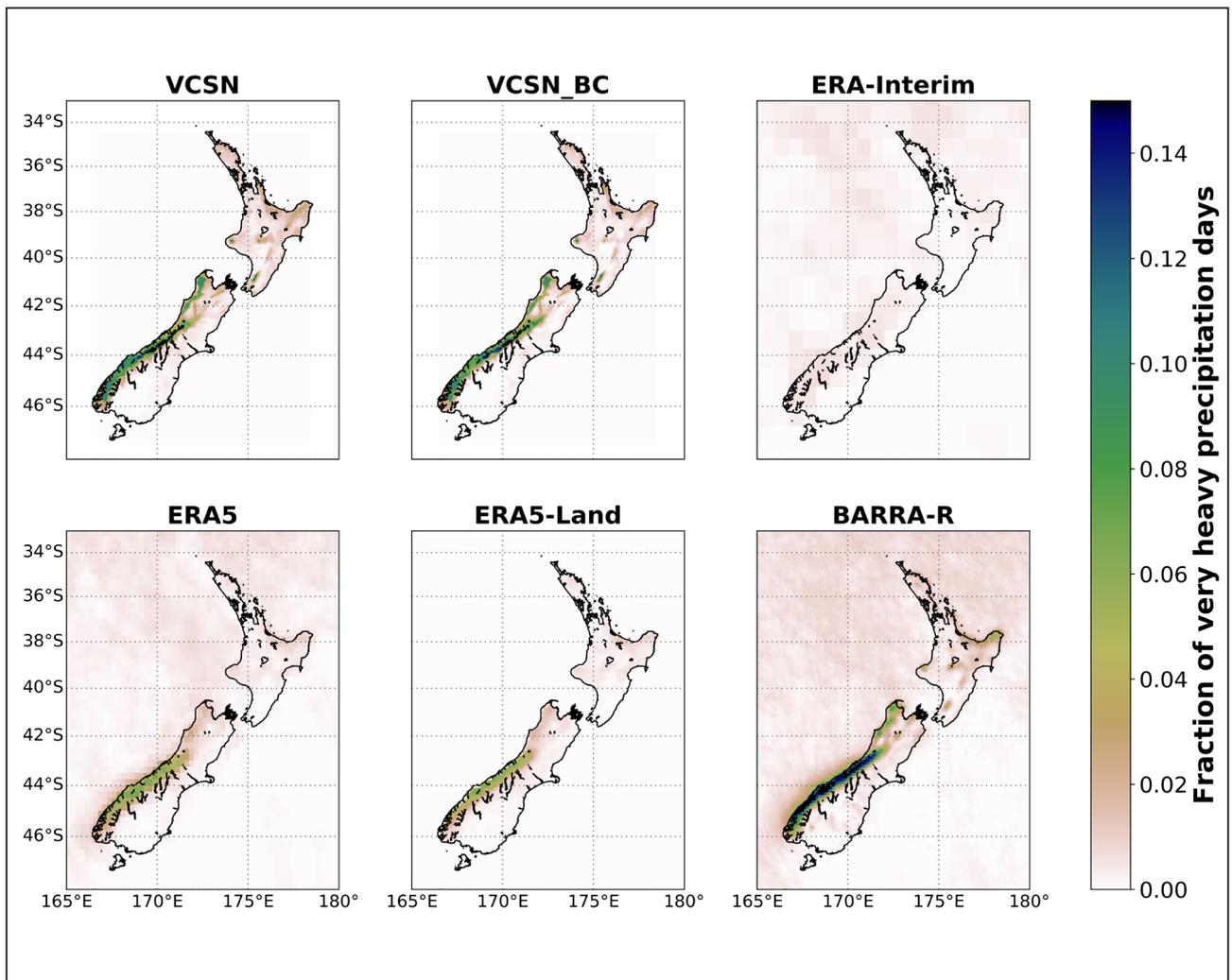
corresponding closest VCSN\_BC and reanalyses grid points values. The slope ( $S$ ) of the lines of best fit and correlation ( $R^2$ ) values are also shown in the figure. It can be seen that ERA5, ERA5-Land and ERA-Interim overestimate the fraction of days for which precipitation was between 1 – 10 mm. BARRA-R estimates the fraction of days with precipitation between 1 mm and 10 mm better than ERA5, with higher  $R^2$  and  $S$  closer to 1.0, but slightly overestimating the fraction compared with VCSN\_BC. ERA5, ERA5-Land and BARRA-R show similar results for 10 – 50 mm and over 50 mm precipitation ranges,



**Figure 3:** Statistical results calculated between the bias corrected VCSN and closest reanalysis grid point for total daily precipitation. First, second, and third rows are Pearson's correlation coefficients, RMSE and bias, respectively. Also, from left to right columns show the comparison between VCSN\_BC and ERA-Interim, ERA5, ERA5-Land, and BARRA-R, respectively.



**Figure 4:** Fraction of days that total daily precipitation is: (a) 1 – 10 mm; (b) 10 – 50 mm; (c) greater than 50 mm, for the station observation data (X-axis) and corresponding closest VCSN\_BC and reanalyses grid points (Y-axis) during 2014 to 2018. “S” and “R<sup>2</sup>” are the slope and correlation of the best fit line, respectively.



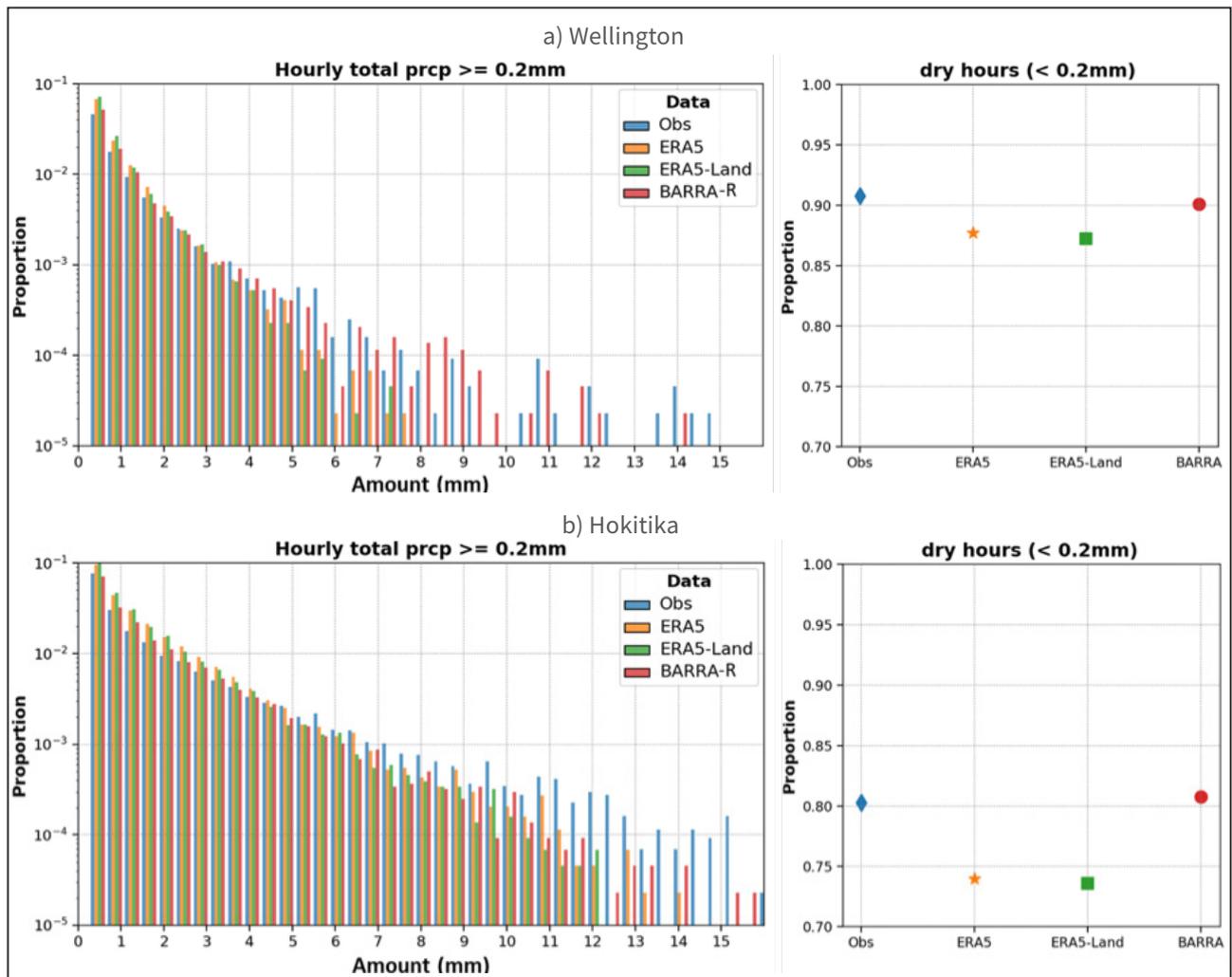
**Figure 5:** Fraction of days that precipitation is greater than 50 mm. Comparison between VCSN, VCSN\_BC and reanalysis products.

while ERA-Interim underestimates the fraction of rainy days for both these ranges. For over 50 mm precipitation ranges, BARRA-R shows lower  $R^2$  and S values compared with ERA5 and ERA5-Land.

In terms of fraction of rainy days in different seasons (not shown), unlike BARRA-R that agrees well with observations, ERA-Interim, ERA5 and ERA5-Land overestimate the numbers of days in almost all seasons and for the 1 – 10 mm range, and often show a poor correlation with the observation values (e.g. Jun – Aug (JJA) and Sep – Nov (SON) 1-10 mm). For 10 – 50 mm and > 50 mm precipitation ranges, BARRA-R, ERA5 and ERA5-Land depict good agreement with observations,

and BARRA-R slightly outperforms the other reanalysis products, except for high precipitation estimates in SON season (seasonal results not shown here).

For a better spatial assessment, Figure 5 compares the fraction of very heavy precipitation (i.e. > 50 mm accumulation) days across New Zealand obtained from the reanalyses and both VCSN datasets. BARRA-R is the only reanalysis dataset to provide similar estimates of the frequency of very heavy rains compared with VCSN, particularly on the west of the SI and over the NI topography. While ERA-Interim does not capture the frequency of very heavy rain days at all, ERA5 and ERA5-Land do exhibit signatures of increased heavy



**Figure 6:** Comparison of station data and reanalyses distributions of hourly wet (left), i.e.  $\geq 0.2$  mm precipitation, and dry (right), i.e.  $< 0.2$  mm precipitation, at: (a) Wellington; (b) Hokitika.

rainfall along the SI's west coast but both underestimate compared to VCSN and BARRA-R, suggesting these datasets do not capture rainfall totals representative of that region. Su et al. (2019) and Jerney and Renshaw (2016) also point out that higher-resolution reanalysis products better represent high-threshold events.

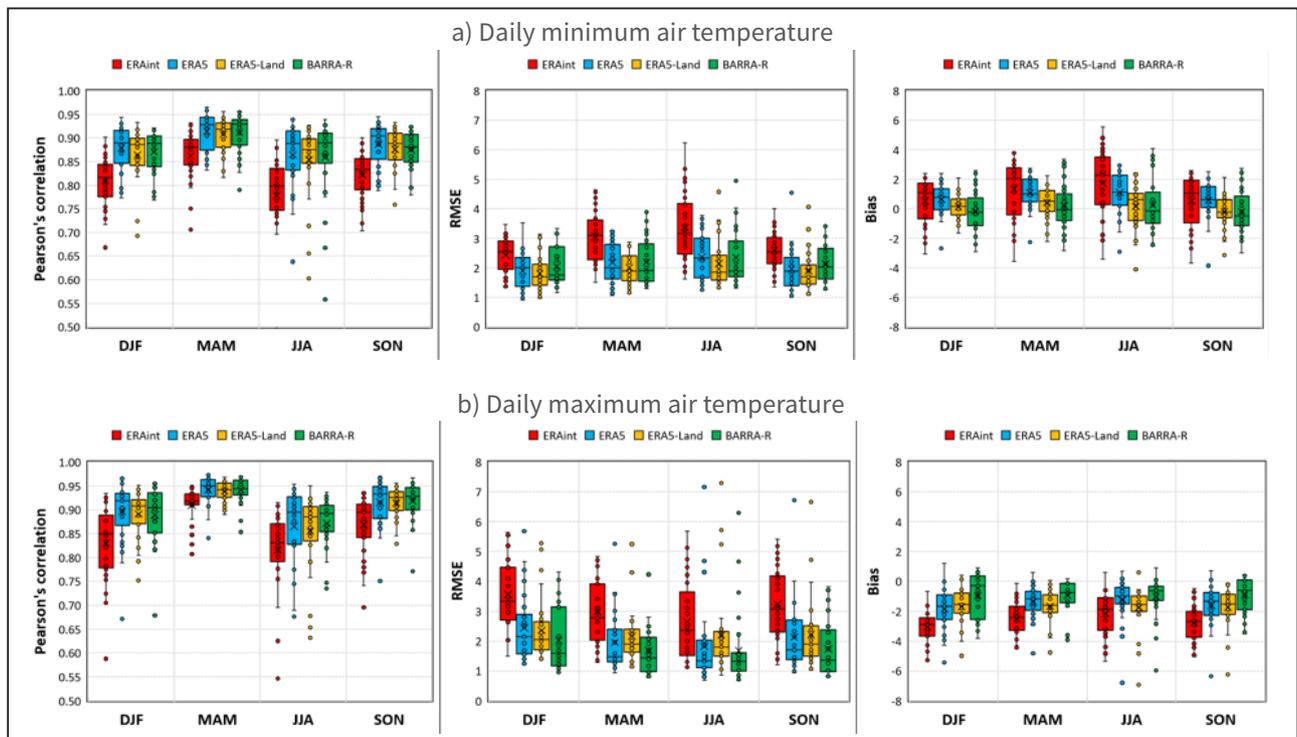
Figure 6 compares the distribution of hourly precipitation accumulations (i.e.  $\geq 0.2$  mm) and dry ( $< 0.2$ mm) times between the reanalysis datasets at two observing stations. Generally, the reanalysis products overestimate the hourly precipitation less than 3 – 3.4 mm and underestimates the precipitations greater than these values, except for BARRA-R, which estimates the high accumulations more accurately compared with observational data. This is particularly evident for Wellington (Figure 6a, left panel). At most locations ERA products do not capture well the frequency of high hourly and daily accumulations (daily results are not shown).

Regarding the frequency of dry ( $< 0.2$  mm) hours, the ERA

reanalyses predict fewer number of hours. BARRA-R, however, generally provides a much better estimation of dry periods across most sites (Figure 6, right panels).

### 3.2. Temperature

The statistical scores for minimum and maximum daily temperatures in each season over New Zealand between observations and reanalyses are shown in Figure 7. With the exception of ERA-Interim, the different reanalyses have similar performance with only minor differences. During winter (JJA), the reanalyses show slightly lower and more scattered correlation (Figure 7a), and higher RMSE and biases (Figure 7a) for the minimum daily temperature. In the case of maximum daily temperature (Figure 7b), during summer (DJF) and winter (JJA), the models show weaker and more scattered scores. The bias values for the minimum and maximum temperatures generally range from  $-2^{\circ}\text{C}$  to  $2^{\circ}\text{C}$  and  $-4^{\circ}\text{C}$  to  $0.5^{\circ}\text{C}$ , respectively. Compared to the ERA-5, BARRA-R generally exhibits better mean Pearson's Correlation and Bias scores across

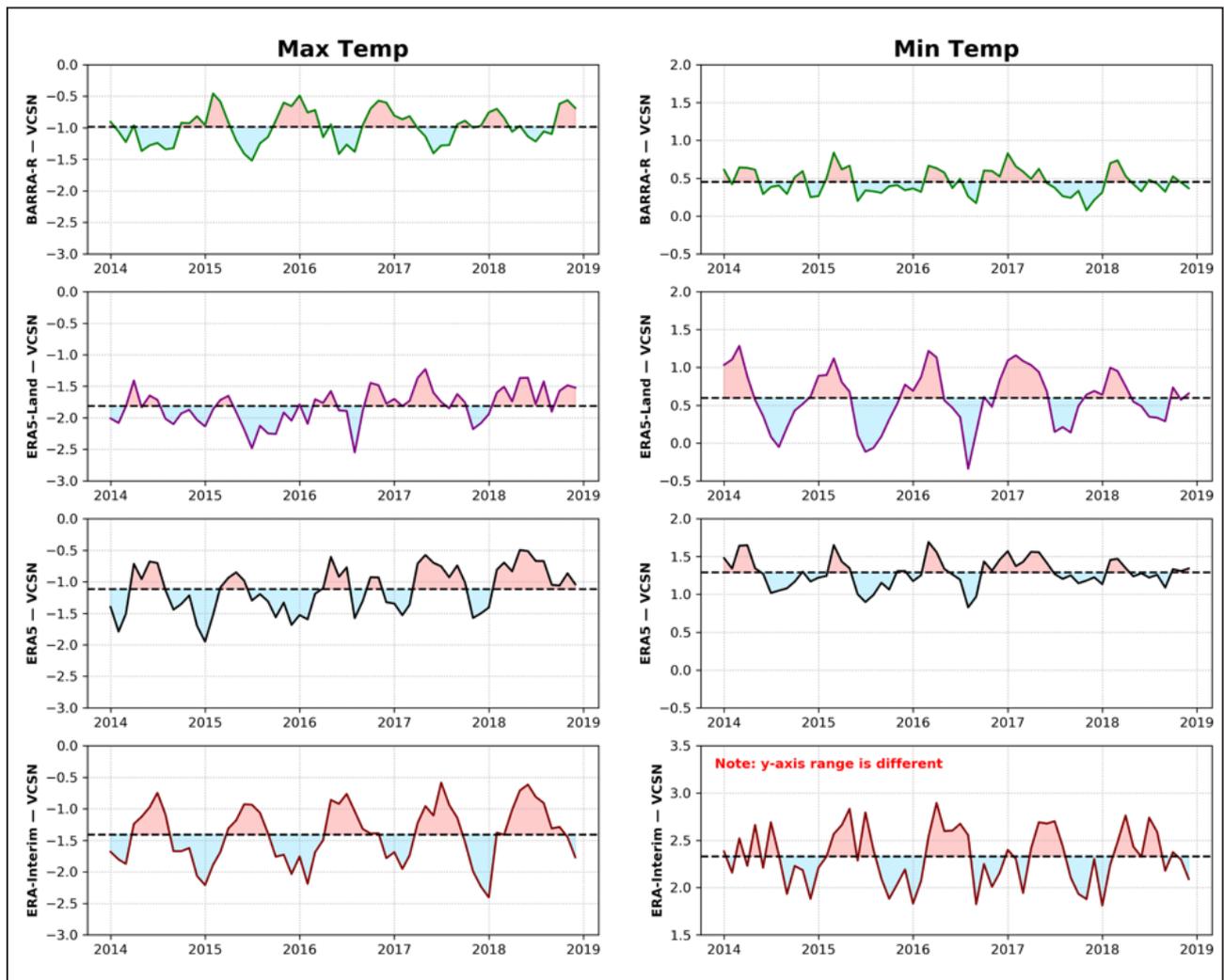


**Figure 7:** Box and Whisker plots showing the distribution of reanalysis products' evaluation scores relative to the stations in each season using observation data as reference for: (a) daily minimum air temperature; (b) daily maximum air temperature.

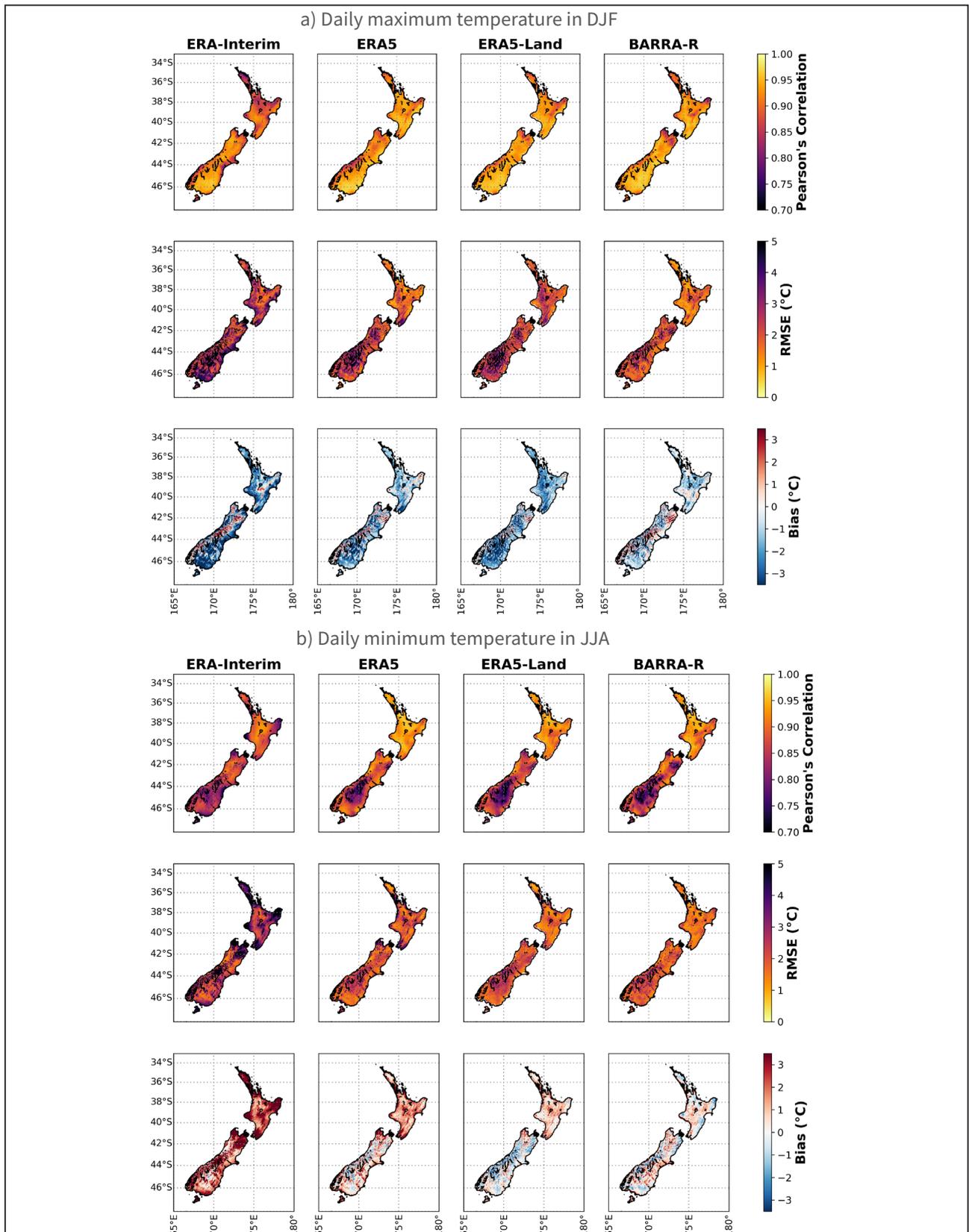
all seasons, but much more closely matched with ERA5-Land. This is most likely due to their similar horizontal resolutions of 12 and 9 km respectively, compared to the 31 km resolution of ERA5 and 80 km for ERA-Interim (Muñoz-Sabater et al., 2021).

Monthly mean differences of the reanalyses in daily maximum and minimum temperature with respect to VCSN over the period 2014 – 2018 averaged over all of New Zealand are shown in Figure 8. The highest inter-seasonal variations were seen for ERA-Interim and ERA5-Land maximum and minimum temperature, respectively, with a value of about 2°C. All the models show negative and

positive average biases for the maximum and minimum daily temperatures, respectively. BARRA-R depicts lowest average bias values (−1°C and 0.5°C for maximum and minimum temperatures, respectively) and variations of around 1°C, followed by ERA5, which performs equally well for maximum temperature but is almost a degree warmer than BARRA-R for the minimum temperature. For BARRA-R maximum daily temperature, biases are reduced (i.e. closer to zero) during warmer months (i.e. Nov – Mar) while during colder months (May – Sep) BARRA-R shows larger negative biases. This trend is almost reversed in ERA5 and ERA-Interim estimates.



**Figure 8:** Monthly mean difference in daily maximum (left column) and minimum (right column) temperatures averaged over New Zealand with respect to VCSN. Black dashed lines are the mean values during 2014 – 2018. Note that ERA-Interim daily minimum temperature has different y-axis range.



**Figure 9:** Statistical results calculated between VCSN and reanalyses for: (a) daily maximum temperature in DJF; (b) daily minimum temperature in JJA. First, 2nd, and 3rd rows are Pearson's correlation coefficients, RMSE and bias, respectively. The reanalyses are re-gridded onto the VCSN grid using the linear interpolation method.

The spatial variability of the reanalyses' performance, calculated with respect to VCSN, for maximum and minimum daily temperatures in summer (Dec to Feb, DJF) and winter (Jun to Aug, JJA) are shown in Figures 9a and 9b, respectively. For maximum daily temperature (Figure 9a), ERA-Interim has the lowest correlation overall and highest RMSE and bias (negative) values. BARRA-R generally outperforms the two ERA5 reanalyses at most locations, particularly over areas of complex terrain such as the Southern Alps and Central Plateau, but struggles in the Marlborough region. BARRA-R has the lowest biases ranging from  $-1^{\circ}\text{C}$  to  $1^{\circ}\text{C}$  over the NI and most parts of the SI. There is a generally cool bias in the maximum temperature over most New Zealand for all models with biases typically colder than  $-1^{\circ}\text{C}$ .

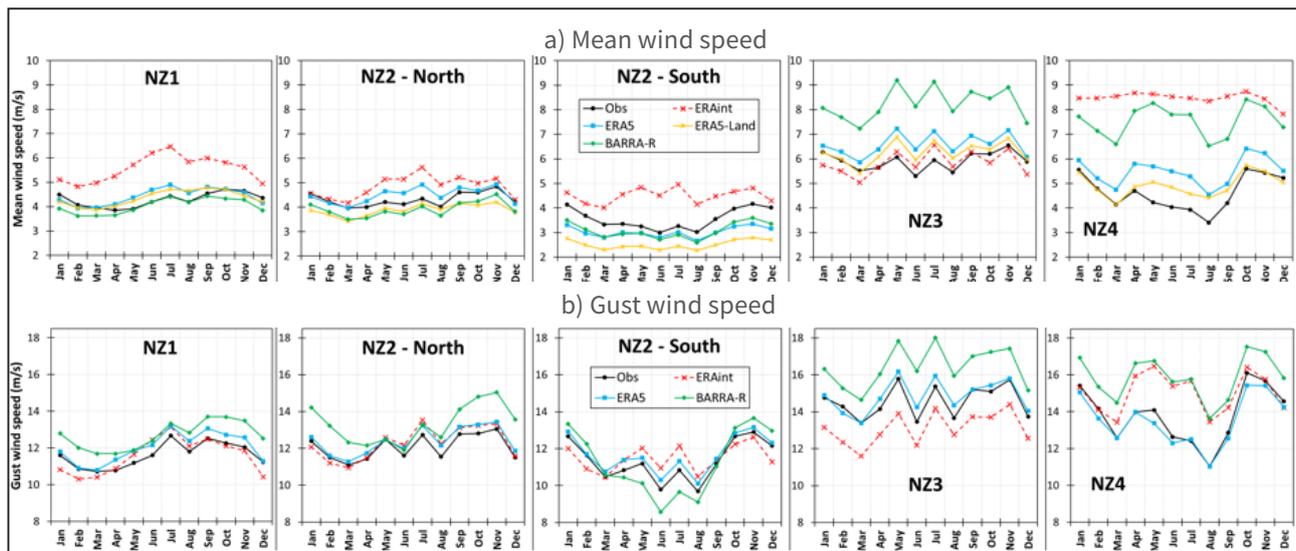
Pearson Correlation scores are significantly lower for all reanalysis products for minimum temperature (Figure 9b), especially over high orography such as the Southern Alps. Compared to the warm bias of BARRA-R, ERA5-Land is generally too cold over the SI and too warm over the NI. ERA5 and ERA5-Land display a more marked warm bias over the NI compared to the SI. ERA-Interim is too warm over all of NZ and coupled with its

maximum temperature bias (Figure 9a), it suggests that this reanalysis is not capturing the full range of observed seasonal temperature ranges. It is evident that BARRA-R outperforms its driving model, ERA-Interim, and shows considerably smaller negative (positive) bias values for maximum (minimum) daily temperatures.

### 3.3. Mean and gust wind speed

For the five NZ regions (Figure 1), Figure 10 depicts the seasonal cycles, defined as the monthly averages of mean wind speed (Figure 10a) and maximum daily gust wind speeds (Figure 10b), for the point-based observations and all the considered reanalysis datasets.

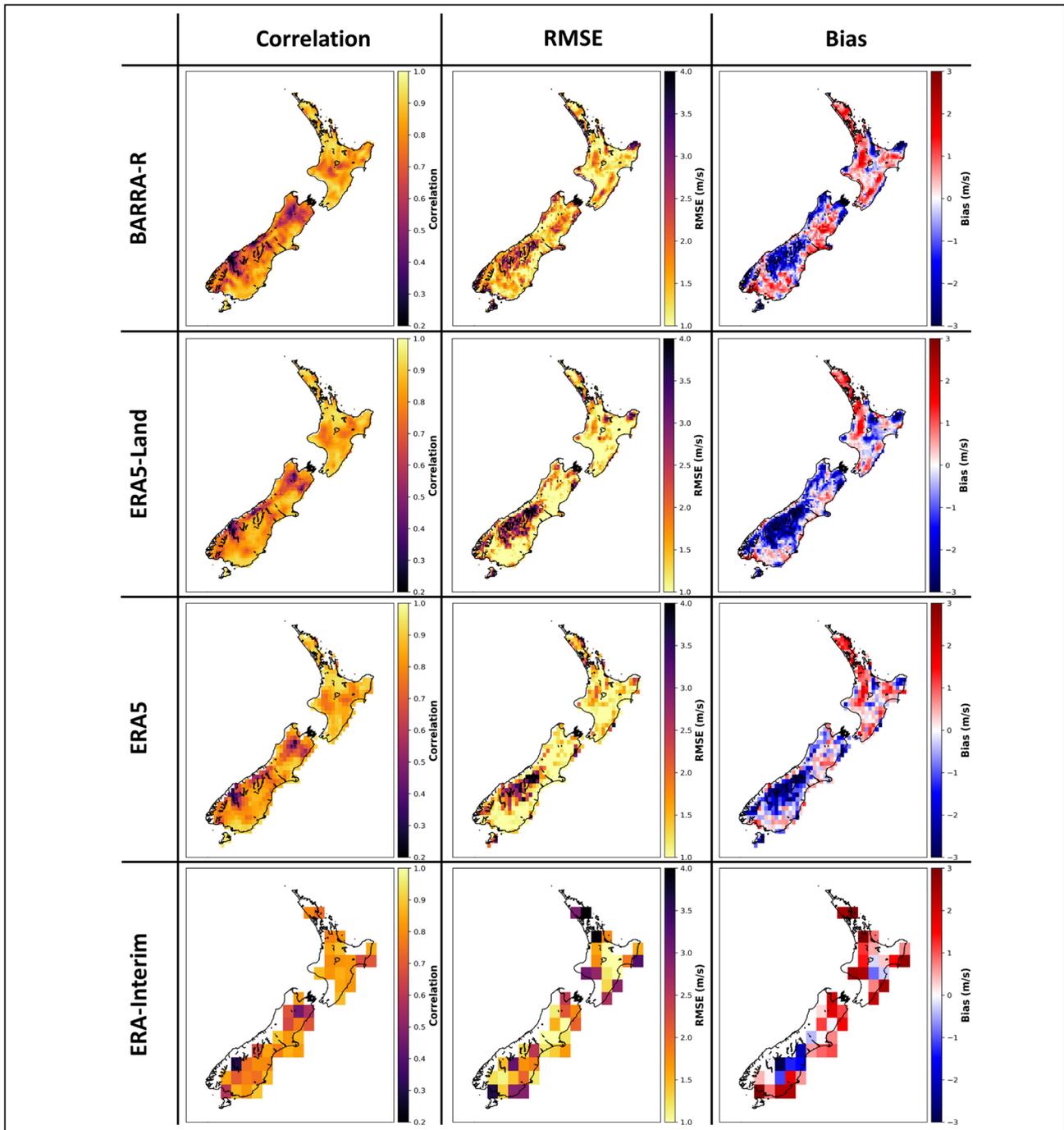
ERA-Interim considerably overestimates the seasonal cycles of mean wind speed (Figure 10a) in all the wind regions (except for NZ3), by up to about 2 m/s (NZ1 and NZ2) and 5 m/s (NZ4). For regions NZ1 and NZ2-North, ERA5, ERA5-Land and BARRA-R estimate monthly average mean wind speeds that are closer to the observations. In these two regions, ERA5 tends to slightly overestimate the seasonal cycle by about 0.5 m/s during Apr to Sep and during the rest of the months its predictions



**Figure 10:** Average monthly cycle of mean wind speeds computed from observations and different reanalysis products for the five New Zealand wind regions for 2014 – 2018.

match the observations. Over the NZ2-South region, the estimates of BARRA-R and ERA5 are very close to one another and both slightly (about 0.5 m/s) underestimate the monthly averages. For NZ3 and NZ4, BARRA-R significantly overestimates (up to 4 m/s) the monthly

averages of mean wind speed, while ERA5 and ERA5-Land provide better predictions, which are generally within 1 m/s of the observations. Overestimation of BARRA-R in this coastal region is likely to be because the closest grid point is not a land point as shown in Figure 2.



**Figure 11:** Statistical results calculated between VCSN and reanalysis data for mean wind speed for 2014 – 2018. First, 2nd, 3rd and 4th rows show the comparison between VCSN and BARRA-R, ERA5-Land, ERA5 and ERA-Interim, respectively. Also, the left, middle and right columns are Pearson's correlation coefficients, RMSE and bias, respectively.

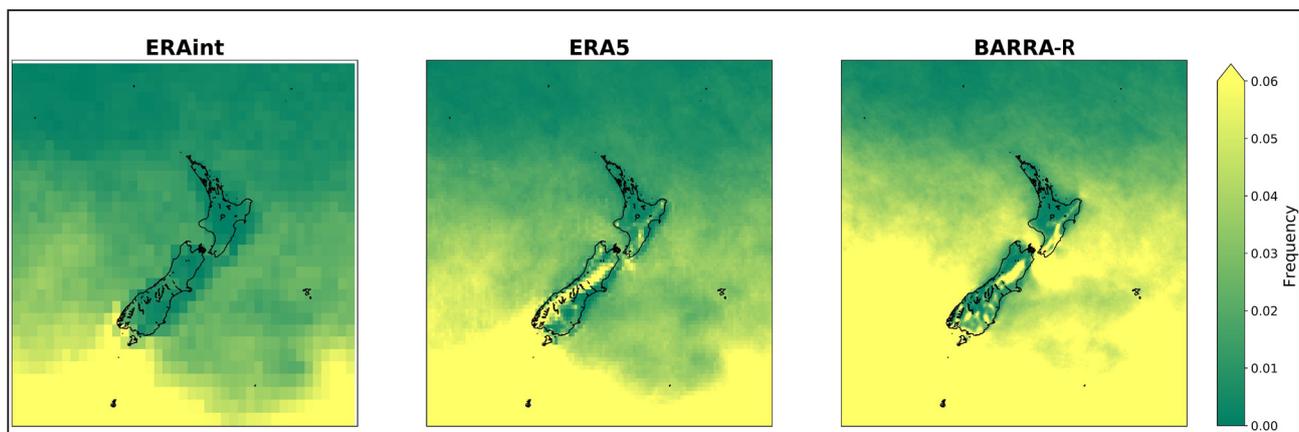
As can be seen in Figure 10b, in all the wind regions, ERA5 estimated the seasonal cycles of gust wind speeds more accurately. ERA-Interim also showed acceptable predictions of the seasonal cycles in regions NZ1, NZ2-North and NZ2-South, with differences being around 0.5 m/s reaching to 1 m/s in NZ2-South region during Apr to Jul. However, ERA-Interim performs poorly in NZ3 (underestimate) and NZ4 (overestimate) regions. BARRA-R continues to perform poorly in the coastal NZ3 and NZ4 regions, overestimating the seasonal gust wind speeds in all months by up to 2 – 2.5 m/s. In region NZ1, BARRA-R overestimates the monthly means by about 1 m/s and 1.5 – 2 m/s during Jan to Apr and Sep to Dec, respectively. Similar trends can also be seen for NZ2-North, with slightly larger differences. However, in NZ2-South this trend is reversed with BARRA-R agreeing well with observations during Apr to Aug, while for the other months the difference increases to 2 m/s.

As for precipitation and temperature above, VCSN provides a spatially complete observation-based records that can be used for the verification of the simulated wind speed in the reanalysis datasets. Figure 11 shows statistical results computed between VCSN and the reanalysis data for mean wind speed. Generally, higher correlation scores occur over areas of relatively flat terrain (e.g. north of NI, east and southeast regions of SI). This is very likely

a result of validating against VCSN which relies on New Zealand's observation station network with very few sites located in areas of high altitude, and is therefore heavily reliant on the interpolation and spline methods used to create a complete surface (Tait et al., 2012).

Figure 12 compares the frequency of daily gust wind speeds exceeding 25 m/s. Unlike ERA-Interim, ERA5 and BARRA-R, due to their higher spatial resolutions, better estimate the higher gust wind speeds, and consequently higher occurrence frequencies, which appear more realistic in comparison with the predictions of New Zealand's high-resolution (1.5 km) convective scale model (Safaei Pirooz et al., 2020b). This is particularly evident over the mountains of the SI and NI as well as high wind speeds in Cook Strait. It is noteworthy that BARRA-R, unlike ERA5, does not capture the high gust wind speeds over the more southern mountains of the SI, but does produce higher gust speeds in the lower NI.

To further investigate and verify the performances of the reanalysis products against observations, Figure 13 illustrates correlations between the reanalysed and observed maximum daily gust wind speeds as well as their distributions for a selection of stations. As can be seen, in most cases, BARRA-R overestimates the maximum daily gust wind speeds (see also Figure 10b) resulting



**Figure 12:** Proportion of gust wind speeds greater than 25 m/s over New Zealand for the period 2014 – 2018. (left) ERA-Interim, (middle) ERA5 and (right) BARRA-R.

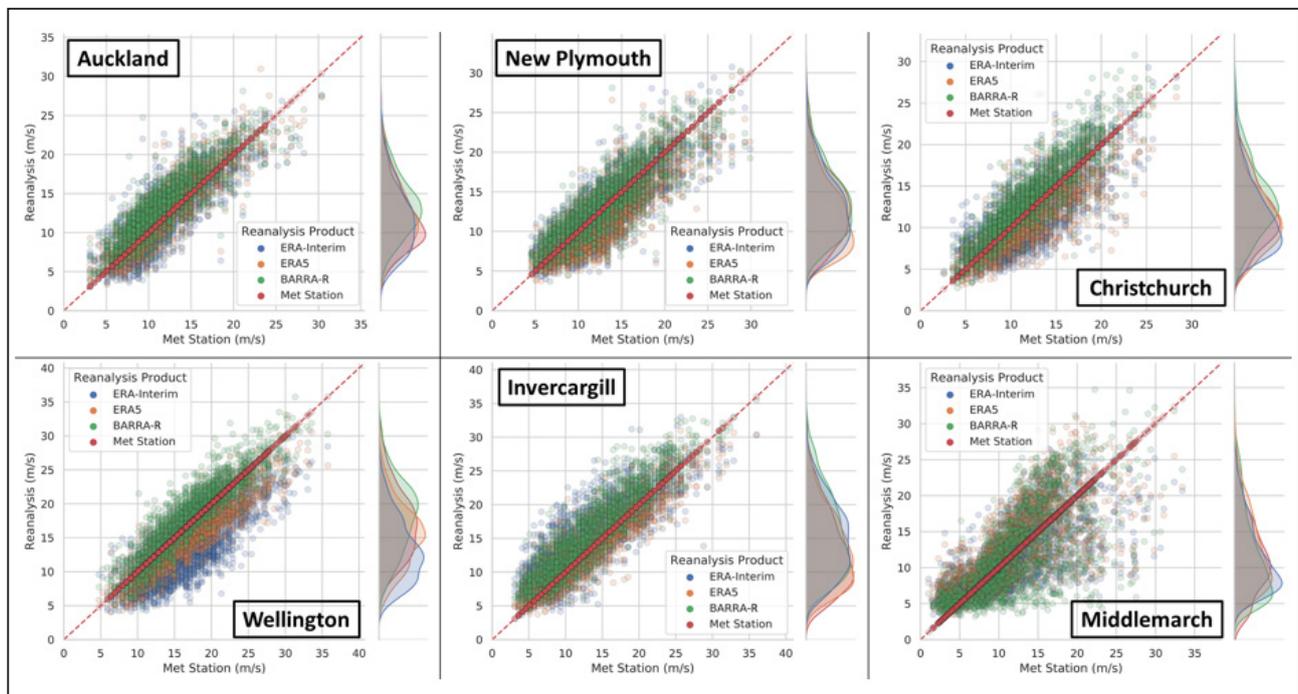
in skewness of BARRA-R distributions towards higher gust wind speeds. This shift in the distributions can be clearly seen in stations such as Auckland, Wellington, and Christchurch. This overestimation of gust wind speeds can influence future studies that are mainly based on the upper tail of the distribution, such as extreme value analysis for the estimation of design wind speeds, or the investigation of trends in frequency and magnitude of extreme winds.

On the other hand, ERA-Interim generally tends to underestimate the gust wind speeds. However, ERA5 seems to represent the distribution of gust wind speeds more accurately, particularly in terms of the location of the peak of the distribution. It should also be noted that in regions near complex terrain (e.g. Middlemarch), or high-wind locations (e.g. Cook and Foveaux Straits), the discrepancies between the observation and reanalyses increases, resulting in more scattered correlations.

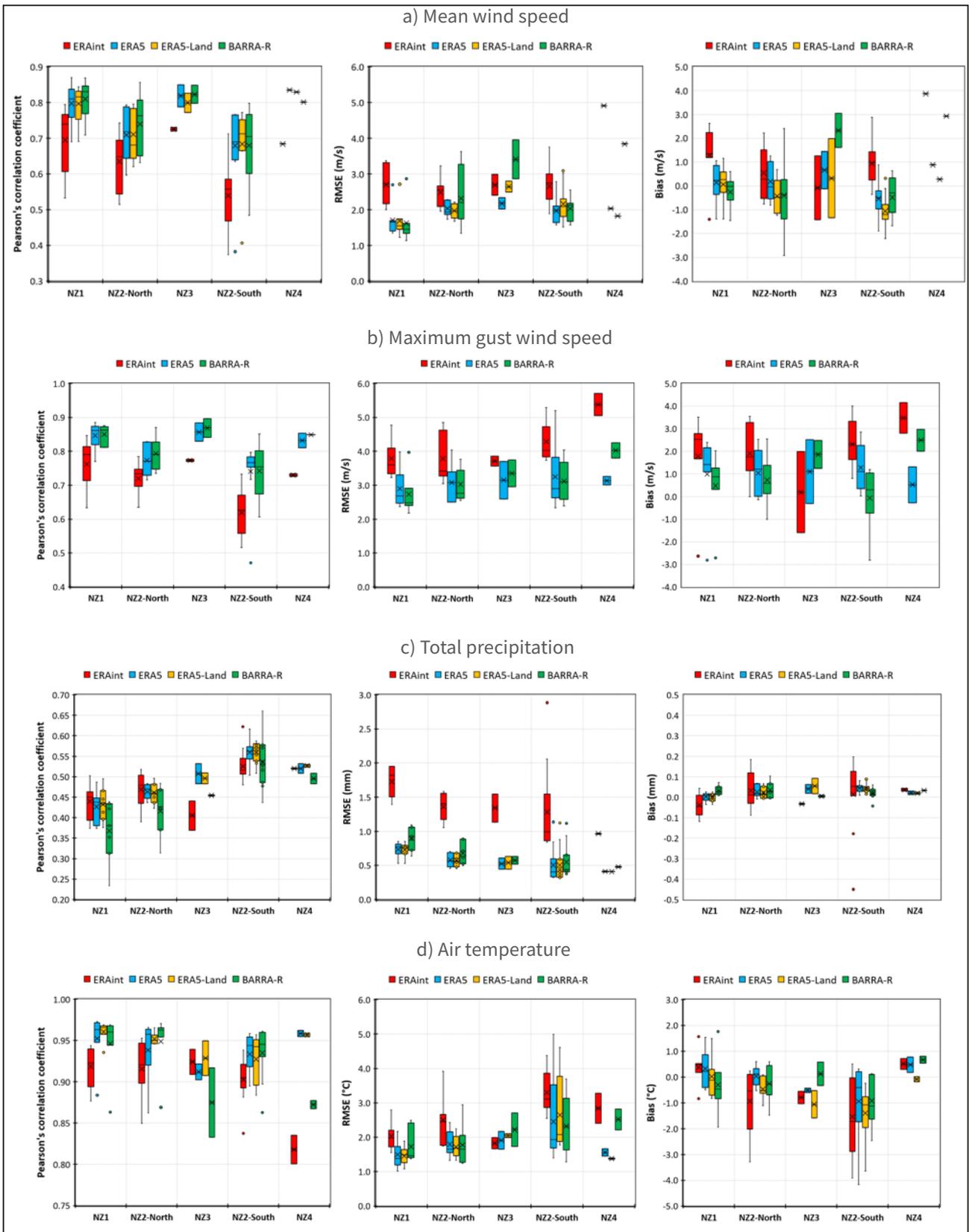
### 3.4. Summary point-based results

Figure 14 summarises the distribution of statistical metrics calculated for reanalysis variables against point-based station observations within each region. Note that in regions NZ3 and NZ4, the number of stations is limited, and stations are located close to each other. This leads to less scattered statistical scores and may make results in these regions less robust.

For mean and gust wind speeds, Figures 14a and 14b show that ERA-Interim generally has lower mean Pearson's coefficients, and higher RMSE and bias (generally positive) values at most station locations. BARRA-R generally shows better agreement with the observation data, except in regions NZ3 and NZ4, NZ's windiest regions climatologically, where BARRA-R's bias and RMSE values are higher than those of ERA5 and ERA5-Land. Also, in the NZ2-North region, BARRA-R shows a more scattered RMSE and bias distributions, while its average RMSE and bias values are close to ERA5 and



**Figure 13:** Plots of linear correlation between the observed and reanalysed maximum daily gust wind speeds along with the distribution of each dataset for a selected number of stations.



**Figure 14:** Box and Whisker plots showing the distribution of reanalysis products evaluation scores using hourly (3-hourly for ERA-Interim) data for: (a) mean wind speed; (b) daily gust wind speed; (c) total precipitation; (d) air temperature.

ERA5-Land. It should be noted that ERA5-Land does not provide gust wind speed data.

For total precipitation (Figure 14c), although ERA-Interim depicts the highest RMSE values with respect to the point observations in all the regions, its Pearson's coefficients and bias values are relatively close to those of the other reanalyses. ERA5, ERA5-Land and BARRA-R show quite similar scores, such that the average and median values of the scores are close across the New Zealand regions. However, in some cases, ERA5 and ERA5-Land depict better performance. For instance, in NZ1 region, BARRA-R shows more scattered correlations and higher RMSE and bias values. Similarly, for NZ2-North region, BARRA-R has slightly higher RMSE and bias. On the other hand, BARRA-R depicts lower RMSE and bias values in regions NZ3 and NZ2-South. ERA5 and ERA5-Land generally show higher correlations and lower RMSE values compared with BARRA-R, particularly over NI. In terms of mean bias, all the three products are positively biased at most station locations. Over SI, BARRA-R depicts slightly lower bias values, while in NI, especially at Hicks Bay and northern part of the island, BARRA-R has higher bias values. The overall performances of BARRA-R, ERA5 and ERA5-Land are quite similar (as also seen in Figure 7 of Su et al., 2021). However, at some locations, for example Auckland, Taupo (centre of NI), Cape Reinga (North of the NI) and Hicks Bay, ERA5 and ERA5-Land show slightly better scores.

The statistical metrics calculated using hourly air temperature between station observations and the reanalyses for each region in New Zealand are shown in Figure 14d. ERA-Interim shows the lowest correlations and more scattered RMSE and bias values. ERA5, ERA5-Land and BARRA-R depict correlations of higher than 0.9 and around 0.95, except for in regions NZ3 and NZ4, where BARRA-R's average correlations drop to 0.87. However, a possible explanation for lower correlation in NZ3 by BARRA-R could be that the closest BARRA-R grid point to one of the two stations in this region (i.e.

Paraparaumu) is on water. In the NZ2-South region (mountainous region) all the reanalysis products show scattered RMSE and bias values. The bias and RMSE values of the reanalyses generally range from  $-1^{\circ}\text{C}$  to  $+1^{\circ}\text{C}$  and  $1^{\circ}\text{C}$  to  $3^{\circ}\text{C}$ , except in region NZ2-South, respectively.

## 4. Conclusion

The ability and accuracy of three global and one regional reanalysis products to capture the meteorological variability of New Zealand were evaluated using point-based and gridded observational data. ERA-Interim was outperformed by all the other reanalyses considered in this study. ERA5 and ERA5-Land generally showed a similar performance with slight discrepancies, which could be attributed to several factors. Although ERA5-Land has a higher spatial resolution, unlike ERA5, it uses indirect data assimilation of observational data and also lacks atmosphere coupling. The performance of all the models investigated here decayed over complex and mountainous regions.

The statistical and climatological metrics used in this study demonstrate that the higher resolution regional reanalysis model, BARRA-R, presents a significant improvement compared to the global reanalyses studied. Most notably, BARRA-R, unlike the ERA datasets, successfully captured the high rainfall characteristics over the western parts of the SI, including more accurately simulating the frequency of very high rain days. It was shown that the slightly weaker statistical scores of BARRA-R in coastal areas could be attributed to the land-sea masks and the representation of part of the land as sea points. The results demonstrate that BARRA-R not only outperforms its driving model, ERA-Interim, but generally it performs better than ERA-Interim's successor, i.e. ERA5.

Su et al. (2021) showed that further downscaling of BARRA-R to 1.5 km horizontal resolution adds significant value over its coarser resolution driving model. Such a long-term high-resolution dataset does not currently exist

for NZ. The National Institute of Water and Atmospheric Research (NIWA) Ltd is currently preparing the New Zealand Reanalysis (NZRA) to create just such a dataset. Forced by BARRA data the NZRA will be available for the period 1991 to 2018 at 1.5 km horizontal resolution and cover all of New Zealand's land mass and coastal waters.

## References

- Ackerley, D., Dean, S., Sood, A., Mullan, A.B., (2012). Regional Climate Modelling in New Zealand: Comparison to Gridded and Satellite Observations. *Weather and Climate*, 32, 3-22.
- Avila-Diaz, A., Benezoli, V., Justino, F., Torres, R., Wilson, A., (2020). Assessing current and future trends of climate extremes across Brazil based on reanalyses and earth system model projections. *Climate Dynamics*, 55, 1403-1426. <https://doi.org/10.1007/s00382-020-05333-z>
- Cliflo, (2020). New Zealand's National Climate Database. <https://cliflo.niwa.co.nz>. Accessed in Sep 2020.
- Davies, T., Cullen, M.J.P., Malcolm, A.J., Mawson, M.H., Staniforth, A., White, A.A., Wood, N., (2005). A new dynamical core of the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131, 1759-1782. <https://doi.org/10.1256/qj.04.101>
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., Vitart, F., (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553-597. <https://doi.org/10.1002/qj.828>
- Di Virgilio, G., Evans, J.P., Di Luca, A., Olson, R., Argüeso, D., Kala, J., Andrys, J., Hoffmann, P., Katzfey, J.J., Rockel, B., (2019). Evaluating reanalysis-driven CORDEX regional climate models over Australia: model performance and errors. *Climate Dynamics*, 53, 2985-3005. <https://doi.org/10.1007/s00382-019-04672-w>
- Essou, G.R.C., Brissette, F., Lucas-Picher, P., (2017). The Use of Reanalyses and Gridded Observations as Weather Input Data for a Hydrological Model: Comparison of Performances of Simulated River Flows Based on the Density of Weather Stations. *Journal of Hydrometeorology*, 18, 497-513. <https://doi.org/10.1175/jhm-d-16-0088.1>
- Frank, C.W., Wahl, S., Keller, J.D., Pospichal, B., Hense, A., Crewell, S., (2018). Bias correction of a novel European reanalysis data set for solar energy applications. *Solar Energy*, 164, 12-24. <https://doi.org/10.1016/j.solener.2018.02.012>
- GLOBE Task Team, (1999). The Global Land One-kilometre Base Elevation (GLOBE) Digital Elevation Model, Version 1.0, in: Hastings, D.A., Dunbar, P.K., Elphinstone, G.M., Bootz, M., Murakami, H., Maruyama, H., Masaharu, H., Holland, P., Payne, J., Bryant, N.A., Logan, T.L., Muller, J.P., Schreier, G., John, S.M. (Eds.). National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80305-3328, U.S.A.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999-2049. <https://doi.org/10.1002/qj.3803>
- Jermey, P.M., Renshaw, R.J., (2016). Precipitation representation over a two-year period in regional reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 142, 1300-1310. <https://doi.org/10.1002/qj.2733>
- Miao, H., Dong, D., Huang, G., Hu, K., Tian, Q., Gong, Y., (2020). Evaluation of Northern Hemisphere surface wind speed and wind power density in multiple reanalysis datasets. *Energy*, 200, 117382. <https://doi.org/10.1016/j.energy.2020.117382>
- Minola, L., Zhang, F., Azorin-Molina, C., Pirooz, A.A.S., Flay,

- R.G.J., Hersbach, H., Chen, D., (2020). Near-surface mean and gust wind speeds in ERA5 across Sweden: towards an improved gust parametrization. *Climate Dynamics*, 55, 887-907. <https://doi.org/10.1007/s00382-020-05302-6>
- Mo, H.M., Hong, H.P., Fan, F., (2015). Estimating the extreme wind speed for regions in China using surface wind observations and reanalysis data. *Journal of Wind Engineering and Industrial Aerodynamics*, 143, 19-33. <https://doi.org/10.1016/j.jweia.2015.04.005>
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.N., (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data Discussion*, 2021, 1-50. <https://doi.org/10.5194/essd-2021-82>
- Muñoz Sabater, J., (2019). ERA5-Land hourly data from 1981 to present. <https://doi.org/10.24381/cds.e2161bac>.
- Safaei Pirooz, A.A., Flay, R.G.J., Minola, L., Azorin-Molina, C., Chen, D., (2020a). Effects of sensor response and moving average filter duration on maximum wind gust measurements. *Journal of Wind Engineering and Industrial Aerodynamics*, 206, 104354. <https://doi.org/10.1016/j.jweia.2020.104354>
- Safaei Pirooz, A.A., Flay, R.G.J., Turner, R., (2020b). New Zealand design wind speeds, directional and lee-zone multipliers proposed for AS/NZS 1170.2:2021. *Journal of Wind Engineering and Industrial Aerodynamics*, 208. <https://doi.org/10.1016/j.jweia.2020.104412>
- Su, C.H., Eizenberg, N., Jakob, D., Fox-Hughes, P., Steinle, P., White, C.J., Franklin, C., (2021). BARRA v1.0: kilometre-scale downscaling of an Australian regional atmospheric reanalysis over four midlatitude domains. *Geoscientific Model Development*, 14, 4357-4378. <https://doi.org/10.5194/gmd-14-4357-2021>
- Su, C.H., Eizenberg, N., Steinle, P., Jakob, D., Fox-Hughes, P., White, C.J., Rennie, S., Franklin, C., Dharssi, I., Zhu, H., (2019). BARRA v1.0: the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia. *Geoscientific Model Development*, 12, 2049-2068. <https://doi.org/10.5194/gmd-12-2049-2019>
- Tait, A., Henderson, R., Turner, R., Zheng, X., (2006). Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *International Journal of Climatology*, 26, 2097-2115. <https://doi.org/10.1002/joc.1350>
- Tait, A., Sturman, J., Clark, M., (2012). An assessment of the accuracy of interpolated daily rainfall for New Zealand. *Journal of Hydrology (NZ)*, 51, 25-44.
- Tait, A., Turner, R., (2005). Generating Multiyear Gridded Daily Rainfall over New Zealand. *Journal of Applied Meteorology*, 44, 1315-1323.
- Tetzner, D., Thomas, E., Allen, C., (2019). A Validation of ERA5 Reanalysis Data in the Southern Antarctic Peninsula—Ellsworth Land Region, and Its Implications for Ice Core Studies. *Geosciences*, 9, 289. <https://doi.org/10.3390/geosciences9070289>
- Turner, R., Safaei Pirooz, A.A., Flay, R.G.J., Moore, S., Revell, M., (2019). Use of High-Resolution Numerical Models and Statistical Approaches to Understand New Zealand Historical Wind Speed and Gust Climatologies. *Journal of Applied Meteorology and Climatology*, 58, 1195-1218. <https://doi.org/10.1175/JAMC-D-18-0347.1>
- Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Berg, L.V.D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, I., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J., (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961-3012. <https://doi.org/10.1256/qj.04.176>
- World Meteorological Organisation, (2014). Guide to meteorological instruments and methods of observations, 8th ed. WMO, Geneva, Switzerland.
- Xu, H., Lin, N.A.m.a., Huang, M.A.m.a., Lou, W.P., (2020). Design Tropical Cyclone Wind Speed when Considering Climate Change. *Journal of Structural Engineering (United States)*, 146. [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0002585](https://doi.org/10.1061/(ASCE)ST.1943-541X.0002585)

