Received: 30 October 2023 | Revised: 12 March 2024 | Accepted: 22 March 2024

DOI: 10.1002/joc.8456

RESEARCH ARTICLE

International Journal of Climatology RMetS

Evaluating the performance of key ERA-Interim, ERA5 and ERA5-Land climate variables across Siberia

Andrew A. Clelland[1,2] | **Gareth J. Marshall**[1] | **Robert Baxter**[2]

[1]British Antarctic Survey, Natural Environment Research Council, Cambridge, UK

[2]Department of Biosciences, Durham University, Durham, UK

Correspondence
Andrew A. Clelland, British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK.
Email: andell84@bas.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/S007431/1

Abstract

Reanalysis datasets provide a continuous picture of the past climate for every point on Earth. They are especially useful in areas with few direct observations, such as Siberia. However, to ensure these datasets are sufficiently accurate they need to be validated against readings from meteorological stations. Here, we analyse how values of six climate variables—the minimum, mean and maximum 2-metre air temperature, snow depth (SD), total precipitation and wind speed (WSP)—from three reanalysis datasets—ERA-Interim, ERA5 and ERA5-Land—compare against observations from 29 meteorological stations across Siberia and the Russian Far East on a daily timescale from 1979 to 2019. All three reanalyses produce values of the mean and maximum daily 2-metre air temperature that are close to those observed, with the average absolute bias not exceeding 1.54°C. However, care should be taken for the minimum 2-metre air temperature during the summer months—there are nine stations where correlation values are <0.60 due to inadequate night-time cooling. The reanalysis values of SD are generally close to those observed after 1992, especially ERA5, when data from some of the meteorological stations began to be assimilated, but the reanalysis SD should be used with caution (if at all) before 1992 as the lack of assimilation leads to large overestimations. For low daily precipitation values the reanalyses provide good approximations, however they struggle to attain the extreme high values. Similarly, for the 10-metre WSP; the reanalyses perform well with speeds up to 2.5 ms^{-1} but struggle with those above 5.0 ms^{-1}. For these variables, we recommend using ERA5 over ERA-Interim and ERA5-Land in future research. ERA5 shows minor improvements over ERA-Interim, and, despite an increased spatial resolution, there is no advantage to using ERA5-Land.

KEYWORDS

ERA5, ERA5-Land, ERA-Interim, reanalysis, Russian Far East, Siberia, snow depth, validation

---

1 | INTRODUCTION

Siberia can be defined as the vast region of Russia and Northern Kazakhstan that extends over 13 million square kilometres, bounded to the west by the Ural Mountains and to the east by the Pacific Ocean. Much of this region comprises boreal forest, known as the taiga, which is the world's largest forest, comprising over 20% of the planet's total forested area and 60% of all boreal forests in the Northern Hemisphere (Zyryanova et al., 2008). Siberia is projected to experience extreme weather events more frequently during this century than the last due to the warming climate. Such events include flooding, an increase in the number and severity of forest fires, and permafrost and glacial melting, leading to biodiversity changes and increased carbon dioxide and methane emissions (Callaghan et al., 2021). Therefore, major changes to the Siberian taiga ecosystem and its ability to operate as a carbon sink will itself have a considerable effect on the magnitude of future global warming (Olsson, 2009). Such changes are already apparent: for example, in recent decades, forested areas of Northern Russia with significantly positive cumulative near-surface air temperature trends have shown the most pronounced deterioration (Shabanov et al., 2021), leading to an increase in the amount of understorey fuel (Gale et al., 2021) and thus a greater propensity to suffer fire activity.

In order to understand the drivers of Siberian climate change and to provide a baseline against which to compare future changes, we need an accurate dataset that provides coverage across the entire region. Atmospheric reanalyses are the standard tool for providing such data in remote areas such as Siberia, where there are relatively few meteorological observations. They provide a uniform multivariate record of the atmosphere and hydrological cycle. Reanalyses employ a numerical weather prediction forecast model to assimilate an historical archive of meteorological data, derived from a range of ground-based observations, radiosondes and multifarious types of satellite data. By providing gridded global fields of meteorological parameters, reanalysis data give a continuous picture of the climate. The forecast and assimilation models are kept constant in these datasets to maintain temporal homogeneity for climate change studies, unlike operational forecast systems that are regularly updated. Nonetheless, it is known that reanalyses can be affected by variations in the coverage and type of assimilated data, especially in more remote areas where there are fewer other data to constrain the reanalysis model.

New versions of reanalyses are produced regularly, following improvements in model physics and data assimilation in the operational forecast models. The current European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis is their fifth-generation atmospheric reanalysis product (ERA5), with a 0.25° (31 km) horizontal resolution with output at an hourly timescale (Hersbach et al., 2020). ERA5 superseded ERA-Interim, where the data were produced at a reduced 80 km horizontal resolution on a 6-h timescale (Dee et al., 2011). Following the release of ERA5, ERA5-Land was created over the land component with the same temporal resolution but a higher spatial resolution of 0.1° (9 km) (Muñoz-Sabater et al., 2021).

It is especially important that the accuracy of reanalyses is properly evaluated to ensure the scientific validity of climate research that is based solely on these data. Several previous works have conducted validations of ERA-Interim with observed data for climate variables in remote regions of Russia.

The mean 2-metre air temperature is the most accurately reproduced climate variable in ERA-Interim, irrespective of study location, with correlation values consistently over 0.90 and mean biases of 0.80°C–1.50°C (e.g., Demchev et al., 2020; Grankina et al., 2019; Lan et al., 2023). In these studies, the validations were conducted on a daily timescale before the mean was taken either seasonally (Demchev et al., 2020; Grankina et al., 2019) or annually (Lan et al., 2023). Note that we expect the correlation values to increase as the timescale increases as daily extreme maxima/minima have less overall impact on the correlations of longer periods. Klehmet et al. (2013) found that snow depth (SD) in South Siberia was generally overestimated by 20%–40%, and up to 60% in April. ERA-Interim consistently significantly underestimated total precipitation (TP): Riazanova et al. (2016) found that precipitation values are underestimated by 56% in South Siberia during summer months, where the extreme high precipitation events occur associated with localized convective activity, and Dyakonov et al. (2020) stated that at the Caspian Sea the precipitation can be underestimated by up to 54% per year.

Previous recent studies of ERA5's performance have also been conducted in remote regions of Russia: Matveeva and Sidorchuk (2020) validated the daily mean 2-metre air temperature, TP and SD on the Yamal Peninsula from 1985 to 2019, Voropay et al. (2021) evaluated the total monthly precipitation with station data across South Siberia from 1979 to 2015, and Lan et al. (2023) used daily temperature and precipitation variables grouped at an annual timescale to compare several reanalysis datasets, including ERA-Interim and ERA5, across Northern Siberia from 2000 to 2018.

In the abovementioned three studies, the reanalyses again performed best when considering the mean air temperature, especially during summer, with correlations greater than 0.97 for ERA5 and ERA-Interim. The bias in

ERA5 is consistently around 0.50°C, whereas in ERA-Interim this increases to 0.96°C, and ERA5 performs best during the summer months and worst during the winter months. When considering the maximum and minimum temperatures, Lan et al. (2023) found that both ERA-Interim and ERA5 perform better reproducing temperature maxima. For ERA-Interim the absolute bias of the maxima is 1.29°C less than the minima (−1.73°C/3.02°C) and the correlation is increased by 0.06 (0.90/0.84), whereas for ERA5 the absolute bias of the maxima is 0.12°C less than the minima (−2.35°C/2.47°C) and the correlation is increased by 0.08 (0.93/0.85).

There are significant variations in the skill of the reanalyses in reproducing TP between the studies, highlighting the regional differences in the quality of the reanalyses. ERA5 performs worse in the northernmost region, the Yamal Peninsula, with daily correlations only between 0.50 and 0.70. Here, ERA5 struggles to reproduce the extreme high precipitation events, especially during summer, where the RMSE is greatest, although the reanalyses tend to overestimate precipitation all year round. In South Siberia, the correlations are on average >0.70, although ERA5 again performs worse during summer than in other seasons and there are high biases and RMSE values all year round.

The new ERA5-Land dataset appears in few validation studies to date. Evaluations of ERA5-Land's performance have again been conducted on individual variables, such as temperature trends (Liu et al., 2021, Wang et al., 2022, Yilmaz, 2023) and soil temperatures (Cao et al., 2020), the latter focused on permafrost regions, including across Russia. Whereas Liu et al. (2021) and Wang et al. (2022) looked at global temperature trends, Yilmaz (2023) focussed on one specific region—Turkey—and found no significant difference between ERA5 and ERA5-Land; in fact, ERA5 tends to estimate the trends marginally better than does ERA5-Land.

The present article is, we believe, the first to compare ERA-Interim with ERA5 and ERA5-Land across a range of climate variables. In this study, we conducted a validation of the three ECMWF reanalyses with observed data from 29 meteorological stations across Siberia and the Russian Far East. The validation investigates six climate variables: the minimum, mean and maximum 2-metre air temperature (MN2T, T2M, MX2T), mean SD, TP and 10-metre wind speed (WSP). These variables were chosen as this work is part of a broader project investigating how climate change has influenced forest fires in the region. In Section 2, we describe the research areas and reanalyses, the results of the validation are presented in Section 3, and Section 4 discusses the strengths and weaknesses of using these reanalysis data in Siberia, as well as which reanalysis dataset performs best for the variables studied. The shortened variable names will be used throughout the rest of this paper.

2 | METHODOLOGY

2.1 | Research areas and meteorological stations

The research areas were chosen based upon incorporating as many different tree types, topographies, and regional climates as possible, as part of the authors' ongoing research into forest fires. As such the regions cover areas of Scots pine (*Pinus sylvestris*), Siberian dwarf pine (*Pinus pumila*), larch (*Larix* sp.), Siberian fir (*Abies sibirica*) and birch (*Betula* sp.), over mountainous, swampy and steppe land (Kharuk et al., 2021). The coordinates of the boundaries of the seven research areas are given in Table 1; the full list of the 29 meteorological stations used in the validation is given in Table A1 and a map showing the stations within each area is provided as Figure 1. Note that some meteorological stations within the research areas were not used due to a lack of available data (>10%) over the whole time period.

2.2 | ERA-Interim, ERA5 and ERA5-Land

ERA-Interim was announced in 2007 as the ECMWF's replacement for their previous reanalysis product, ERA-40 (Simmons et al., 2007). Data are available from January 1979 to August 2019 on a spectral T255 (79 km) horizontal resolution with a 6-h temporal resolution. ERA-Interim was produced sequentially every 12 h using a 4D-Var data assimilation system (Dee et al., 2011). ERA-Interim was widely used and has been considered a valid alternative to observations globally (Liu et al., 2018).

ERA5 is the fifth-generation atmospheric reanalysis produced by ECMWF, replacing ERA-Interim as their latest reanalysis product. Data are currently available from 1940 onwards (Hersbach et al., 2023). ERA5 has a 0.25° (31 km) horizontal resolution, an hourly temporal resolution, and is produced using an improved four-dimensional data assimilation in ECMWF's Integrated Forecast System (Hersbach et al., 2020). ERA5's 37 post-processed pressure levels are the same as those of ERA-Interim, although the data that ERA5 assimilates better reflect observed changes in climate forcing and many new or recently reprocessed observations are used (Noël et al., 2020). The real vertical resolution is determined by the model levels from the surface up to a height of

TABLE 1 List of study areas with N, W, S and E boundaries and the number of usable meteorological stations within them.

Area	Name	N latitude (°N)	W longitude (°E)	S latitude (°N)	E longitude (°E)	No. of usable stations
1	NW Siberia	64.5	82.0	58.0	95.5	3
2	W Yakutsk	67.0	111.5	61.0	125.0	6
3	E Yakutsk	64.0	127.5	57.5	141.5	5
4	Amur region	58.5	123.0	52.0	134.0	5
5	SW Siberia	56.5	92.5	50.5	102.5	4
6	Verkhoyansk region	70.5	126.0	64.0	143.0	2
7	NE Sakha	69.5	144.5	62.5	162.0	4

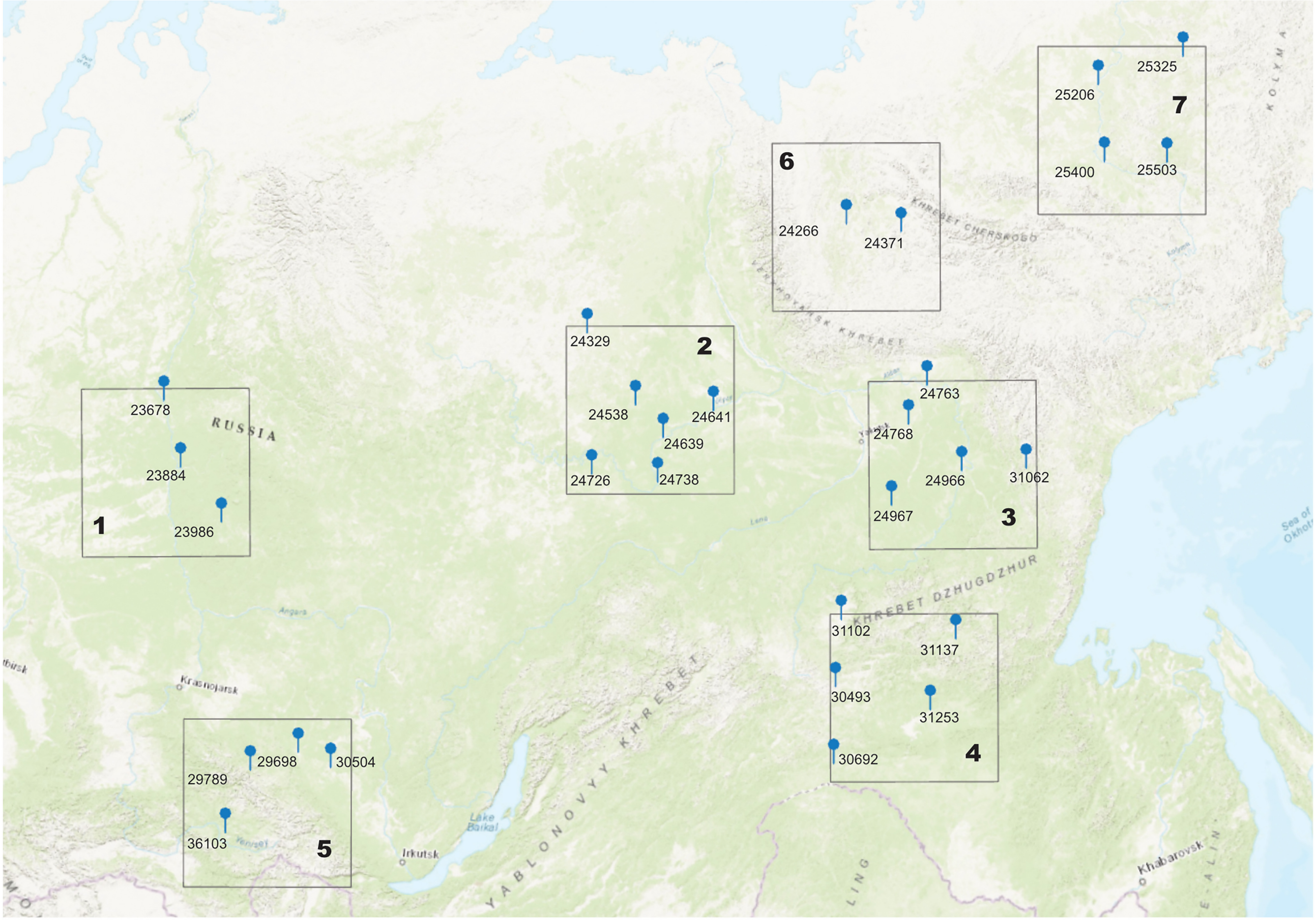

FIGURE 1 Map of the seven research areas in Siberia with pins representing the location of the 29 meteorological stations used in the validation.

80 km, which is 137 for ERA5, increased from 60 in ERA-Interim (Hersbach et al., 2020).

ERA5-Land was created by forcing the HTESSEL land surface component with the atmospheric model (Muñoz-Sabater et al., 2021), but without coupling them. Observations, such as those from meteorological stations, are assimilated into ERA5 but not ERA5-Land, so, despite the fact that the snow model is the same in both ERA5 and ERA5-Land, for instance, snow observations are not directly assimilated in the latter (Hersbach et al., 2020; Muñoz-Sabater et al., 2021; Urraca & Gobron, 2023). ERA5-Land has the same temporal resolution as ERA5 but at a higher spatial resolution of 0.1° (9 km). It is available from 1950 and is updated to the present time with little delay.

ERA5 is generally considered the best reanalysis product currently available (e.g., Ramon et al., 2019; Tarek et al., 2019), and as such this work assesses how ERA5 compares to its predecessor, ERA-Interim, as well as its higher resolution land-based counterpart, ERA5-Land. Daily means for the climate variables are computed for all three reanalyses using values taken at each output temporal resolution (6-h for ERA-Interim and hourly for ERA5 and ERA5-Land). The daily minimum and maximum values are recorded directly as the lowest/highest values over the previous 24-h period (Hersbach et al., 2020).

This validation was conducted using the daily data, which were then analysed on a seasonal basis. The reanalysis data were linearly interpolated to the points of each meteorological station so that a direct analysis could take place.

2.3 | Missing and erroneous data from meteorological stations

The full evaluation was conducted using data from 1979 to coincide with the availability of data from all three reanalysis datasets, with an additional section on how ERA5 and ERA5-Land perform from 1959 to 1978. Time periods were excluded from statistical calculations if more than 10% of meteorological station data during that period were missing. As such, a month was excluded if there were four or more days of missing data, a year was excluded if there were two or more missing months and decadal trends were calculated subject to having 9- or 10-years' worth of data. There are more missing data from the meteorological stations before 1979 than after.

The use of ERA5, on occasion, highlighted where there were significant errors in data from a meteorological station. For example, ERA5 highlighted curious T2M values in the years 1966–1976 at station 23986 (Figure 2a). Although the original reason behind this is unknown, when multiplying the positive values by 10 (Figure 2b), the adjusted readings appear much more realistic, and these adjusted values are used in subsequent statistical calculations.

Similarly, ERA5 highlighted issues with the average SD at station 24371. Before 1984, ERA5 drastically underestimates the values from the station, although the peak in 1978 is represented, as seen in Figure 3. This is one of only two stations where ERA5 generally underestimates SD (along with 23986) and the meteorological station data at 24371 are never assimilated into ERA5. From 1982 to 1985 there is a steep drop in the meteorological station values, and from 1987 onwards the readings are as expected. Note that there are gaps in the meteorological station data where years have been excluded due to missing >10% of the data and the data only begin in 1966. Lobanov and Kirillina (2019) state that although the station has not changed latitude or longitude from 1942, there have been inhomogeneities in the average monthly precipitation in April and August. This leads to the possibility that there could be calibration errors at the station producing inaccurate readings.

2.4 | Statistics used in the validation

Five statistics were used in the validation: the Pearson correlation coefficient (r), the Spearman rank correlation coefficient (ρ), variance ratio (vr), root-mean-square error (RMSE) and bias. r is used for T2M, MX2T and MN2T as these variables can attain values below zero and form normal distributions, whereas the non-parametric ρ is used for SD, TP and WSP as these values cannot attain non-positive values and do not form normal distributions. The line of best fit shown on the figures for the results is determined by the regression model $y \sim x$. The equations for calculating r, ρ, vr, RMSE and bias, respectively, are given in Equations 1–5 below.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{1}$$

where x_i, y_i are the values of the x/y-variable in a sample and $\bar{x}$ and are the mean of the values of the x, y-variable.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{2}$$

where d_i is the difference between the two ranks of each observation and n is the number of observations.

$$S_i^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \tag{3}$$

where S_i is the sample standard deviation of variable i, x_i is the value of one observation, $\bar{x}$ is the mean value of all observations, and n is the number of observations. The variance ratio (vr) is then the variance of the reanalysis dataset divided by the variance of the observational data. If, the vr is >1, the variance in the reanalysis is

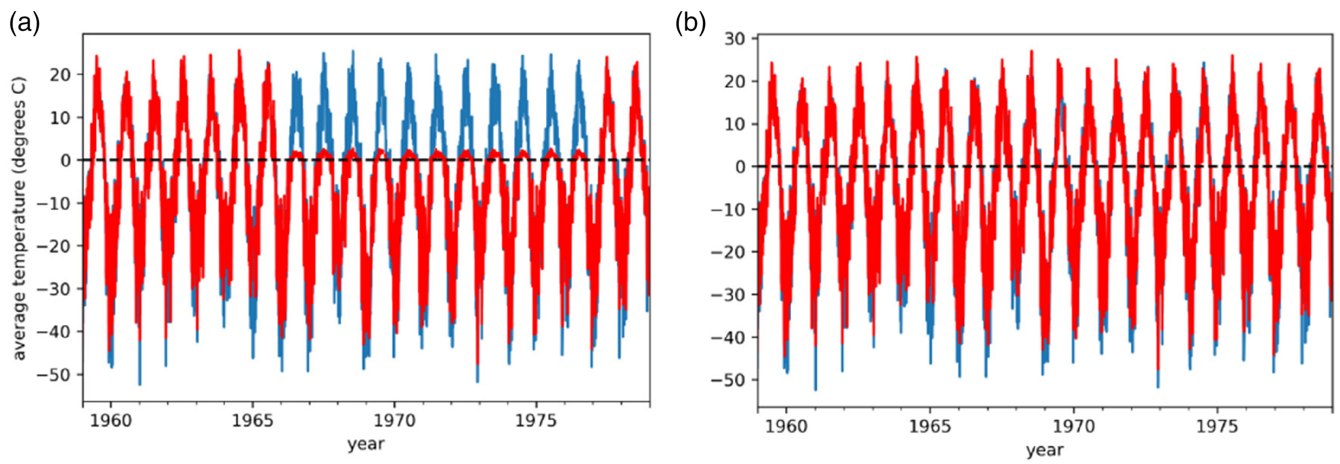

FIGURE 2 Average monthly mean 2-metre air temperature (T2M) values at station 23986 (see Figure 1) for the years 1959–1978 before (a) and after (b) the adjusted meteorological station readings. The blue and red lines show the ERA5 and station values, respectively.

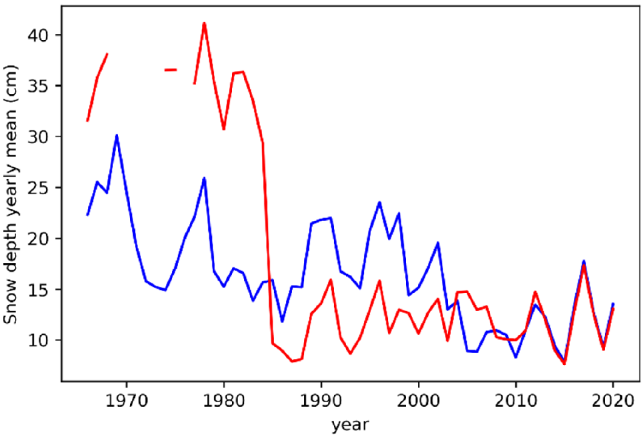

FIGURE 3 The mean yearly snow depth (SD) at station 24371 for the years 1966–2020. The blue and red lines show the ERA5 and station values, respectively.

overestimated, whereas a *vr* <1 indicates that the variance is underestimated.

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{N}\left(\widehat{x}_i-x_i\right)^2}{N}} \tag{4}$$

where N is the number of non-missing data points, $\hat{x_i}$ is the estimated time series, and x_i is the actual observations time series. Similarly:

$$\mathrm{Bias}=\frac{\sum_{i=1}^{N}\left(\widehat{x}_i-x_i\right)}{N} \tag{5}$$

The bias indicates whether the reanalyses overestimate or underestimate the observed parameter and by how much, whereas the RMSE finds the average error, irrespective of its sign.

3 | RESULTS

3.1 | 2-metre air temperature

3.1.1 | Mean 2-metre air temperature

T2M is the parameter studied that is most accurately reproduced by all three reanalyses. The validation statistics are given in Table 2.

The correlations on a daily timescale are ~0.990, and across all the stations there is a small average warm bias, ranging from 0.16°C in ERA5 to 0.50°C in ERA5-Land. The average RMSE is just under 3°C and the variance is consistently slightly underestimated across all three reanalyses. The differences in the correlations and variance ratios are negligible, with a 0.34°C difference between both the smallest and largest biases and RMSEs. Figure 4 shows the performance of ERA-Interim, ERA5 and ERA5-Land against meteorological station data for daily T2M.

The graphs in Figure 4 show that the same relationships between the biases and observed daily T2M values are seen across the three reanalyses: the reanalyses have the least deviation from the $y=x$ line when the temperature is around 0°C, with the deviation increasing above the $y=x$ line as the temperature decreases, leading to a small warm bias as the reanalyses fail to attain the extreme cold temperatures. This is further highlighted in Figure 5, which shows that the reanalyses consistently reproduce T2M less well in winter (DJF).

The season has the lowest average correlation of 0.903, and the greatest mean bias and RMSE values, 0.90°C and 3.79°C, respectively (Figure 5). However, the average variance ratio is the closest to 1.00 of all the seasons. The reanalyses perform similarly strongly in spring (MAM) and autumn (SON), with a correlation over 0.975, an average absolute bias <0.55°C, RMSE ∼2.75°C and variance ratio around 0.93. Summer (JJA) has the lowest average bias and RMSE values of 0.01°C and 1.71°C, respectively. Outliers lie both above and below the whiskers in winter for all three analyses for the T2M bias (Figure 5a), which supports the conclusions drawn from Figure 4 in that the maximum deviation both and above and below the $y = x$ line occurs at temperatures between −20 and −40°C.

Between the three reanalyses the differences between ERA-Interim and ERA5 are negligible, however ERA5-Land consistently performs marginally less well than the others. On a daily timescale, the reanalyses also perform equally strongly across all the study regions, as no area stands out ahead of the others, although Verkhoyansk region has the greatest average bias and RMSE values across the three reanalyses, at 0.87°C and 3.38°C, respectively. In winter, spring and autumn, the reanalyses reproduce T2M best in NW Siberia and worst in Verkhoyansk region, whereas in summer there are no significant differences across the study areas.

TABLE 2 Validation statistics for daily T2M.

Reanalysis	Bias (°C)	RMSE (°C)	*vr*	*r*
ERA-Interim	0.43	2.66	0.96	0.991
ERA5	0.16	2.89	0.98	0.989
ERA5-Land	0.50	3.00	0.94	0.989

Abbreviations: *r*, Pearson correlation coefficient; RMSE, root-mean-square error; T2M, mean 2-metre air temperature; *vr*, variance ratio.

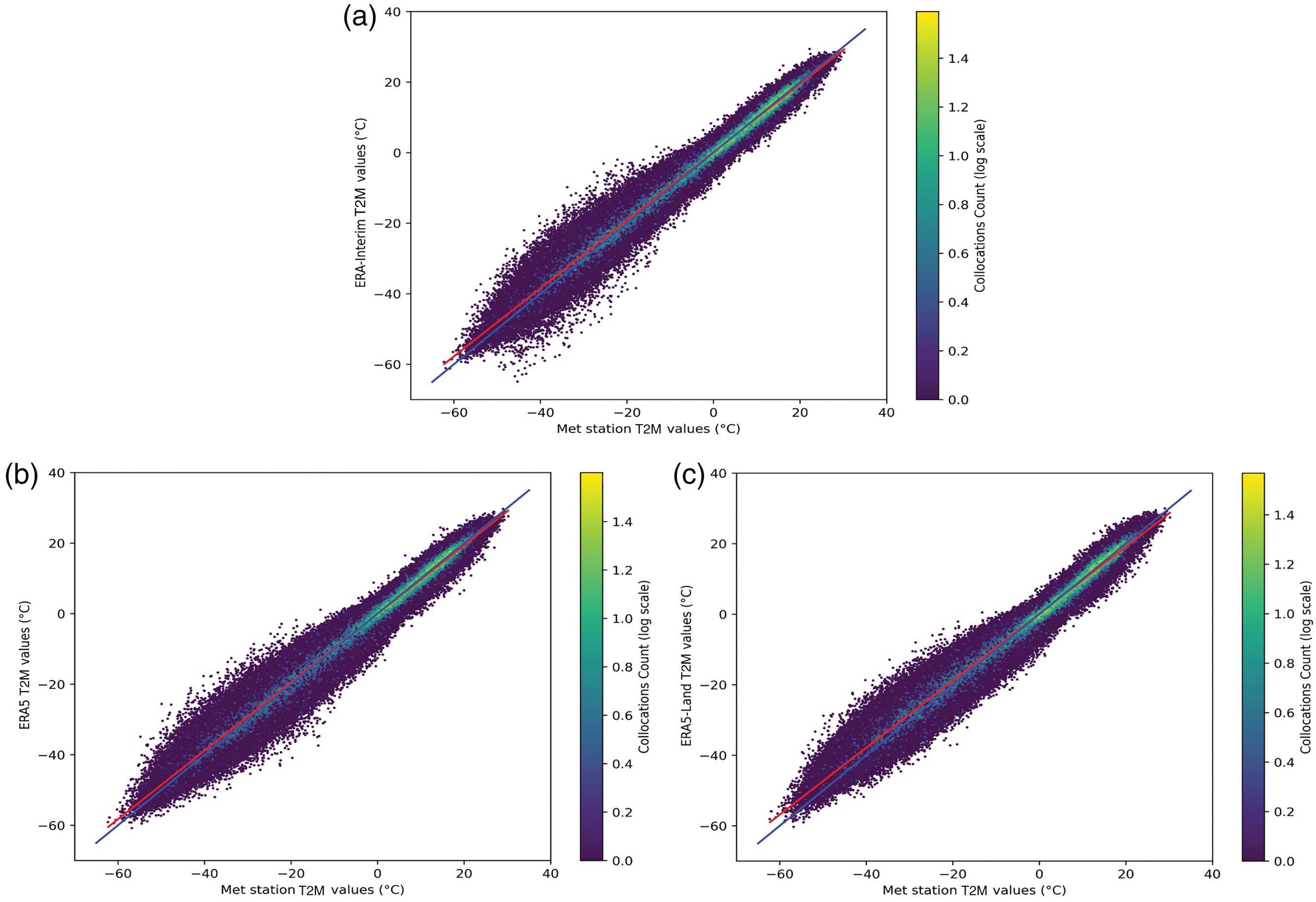

FIGURE 4 Daily mean 2-metre air temperature (T2M) values from ERA-Interim (a), ERA5 (b) and ERA5-Land (c) with respect to the corresponding values from each station. The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

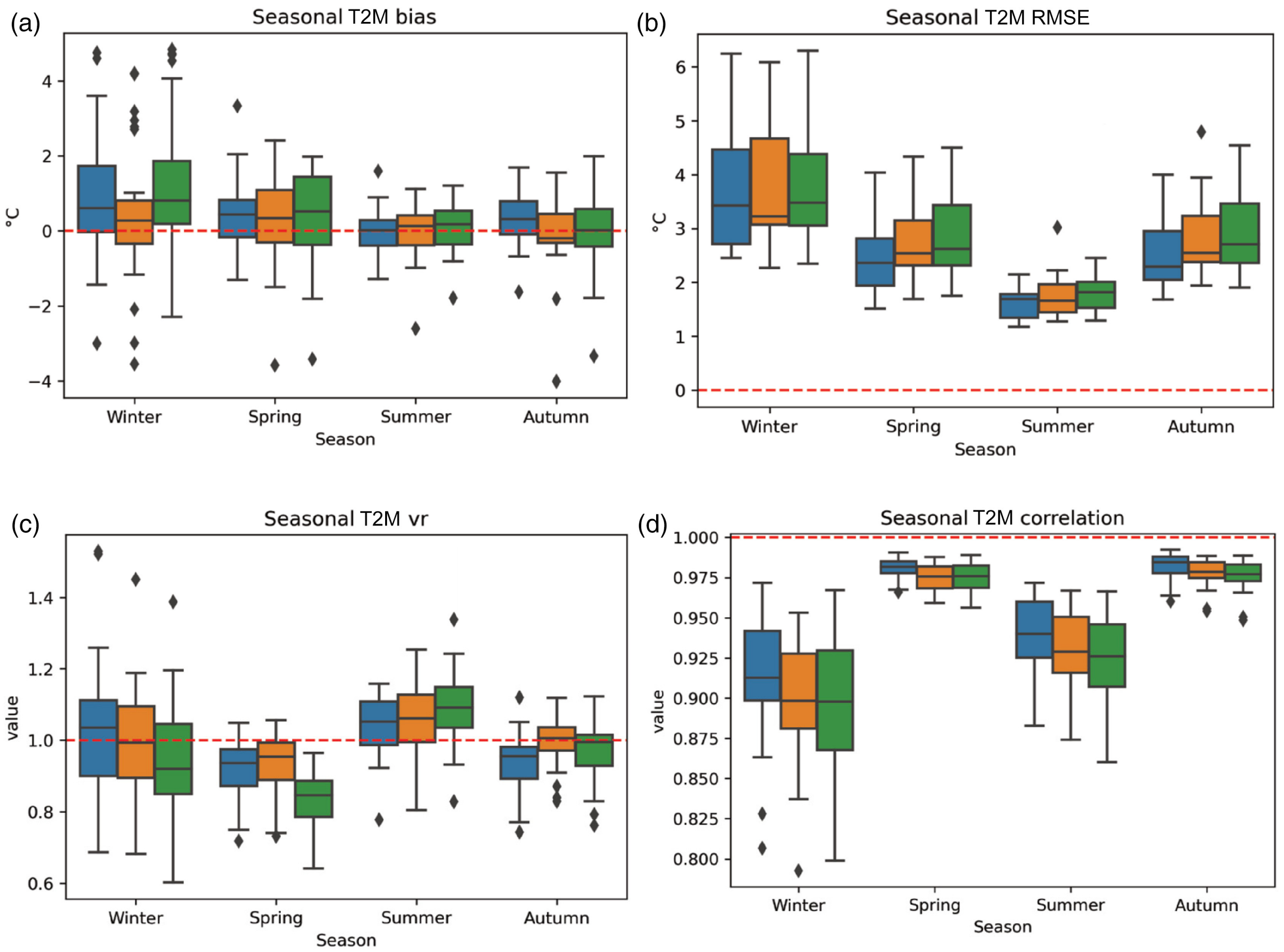

FIGURE 5 Seasonal values of the ERA-Interim (blue), ERA5 (orange) and ERA5-Land (green) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for mean 2-metre air temperature (T2M). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × interquartile range (IQR).

3.1.2 | Maximum 2-metre air temperature

When considering the daily temperature extremes, the three reanalyses tend to perform slightly better at reproducing the daily maxima than the daily minima. The validation statistics for MX2T are given in Table 3.

The reanalyses again all perform consistently well and there is little to separate them. The average correlation is 0.989, and the difference between the highest and lowest values for the bias, RMSE and *vr* are small. The correlation differences between the seven study areas are negligible, and the variance ratios are within 0.10 of each other. For the bias and RMSE values, the reanalyses all perform best in Amur region and least well in SW Siberia, with a 2.39°C and 1.93°C difference between them, respectively.

Figure 6 shows the performance of ERA-Interim, ERA5 and ERA5-Land against meteorological station data for MX2T. The graphs follow the same trends as T2M, although the average cold bias for positive temperatures and the warm bias for negative temperatures are slightly more pronounced here across all three reanalyses.

TABLE 3 Validation statistics for daily MX2T.

Reanalysis	Bias (°C)	RMSE (°C)	*vr*	*r*
ERA-Interim	−1.54	3.57	0.90	0.988
ERA5	−1.49	3.42	0.93	0.989
ERA5-Land	−1.52	3.34	0.94	0.989

Abbreviations: *r*, Pearson correlation coefficient; MX2T, maximum 2-metre air temperature; RMSE, root-mean-square error; *vr*, variance ratio.

There are more seasonal outliers in the MX2T statistics (Figure 7) than with the equivalent T2M statistics (Figure 5), this time throughout the year for both bias

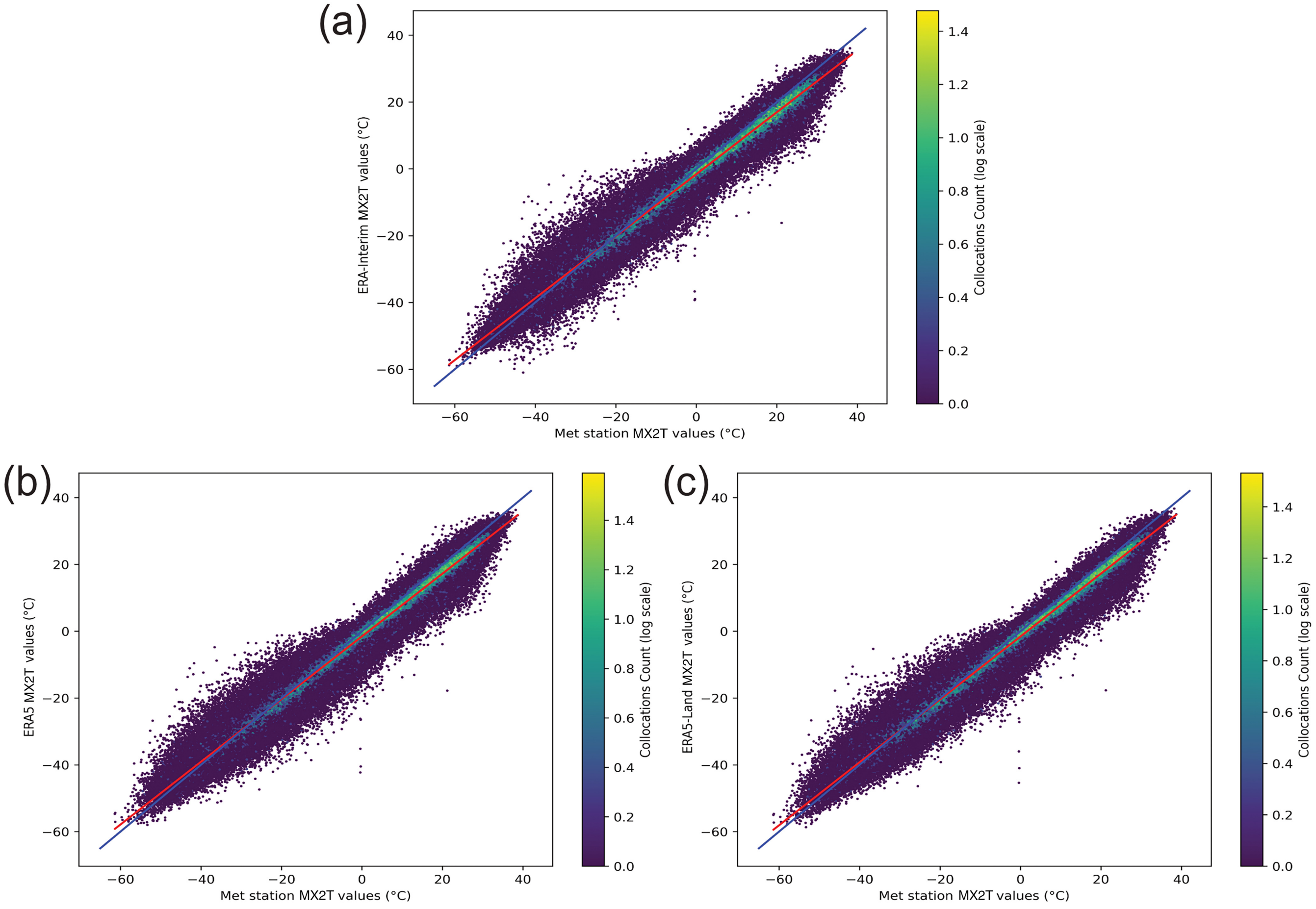

FIGURE 6 Daily maximum 2-metre air temperature (MX2T) values for ERA-Interim (a), ERA5 (b) and ERA5-Land (c) with respect to the corresponding values from each station. The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

and RMSE, and in winter for *vr* and *r*. The outliers occur when the maxima hit their extreme highs and lows as the reanalyses struggle to attain these values. The average correlation ranges from 0.894 in winter to 0.978 in both spring and autumn. In contrast to T2M, the bias is smallest in winter, −0.33°C, and largest, −2.45°C, in summer. However the RMSE is greatest in winter, 3.94°C, and smallest, 3.10°C, in spring. The variance ratio is consistent, at ~0.93, across the reanalyses and seasons. The variance is consistently underestimated, but we note a marked improvement in ERA5 and ERA5-Land over ERA-Interim in autumn (Figure 7c).

The seasonal variation in the validation statistics is inconsistent between study areas, with MX2T performing least well in Verkhoyansk region during summer and winter, where the extreme maximum and minimum temperatures are seen. In winter the region has the lowest average correlation and variance ratio values, and the greatest RMSE values, at 0.808°C, 1.29°C and 4.46°C, respectively, whereas in summer it has the largest mean bias and RMSE values, at −3.76°C and 4.44°C. For spring and autumn, the biases and RMSE values are greatest in SW Siberia, at −3.80°C and 4.63°C, respectively, in spring, and −3.14°C and 4.26°C in autumn. Amur region is the area where the reanalyses perform best in spring and summer, with average bias/RMSE values of −0.49°C/2.31°C in spring and −1.80°C/2.60°C in summer. There are no other significant seasonal spatial differences.

3.1.3 | Minimum 2-metre air temperature

The MN2T values are the least well reproduced temperature values by the reanalyses, as indicated especially in the bias and RMSE statistics; however, the values are still strong compared with the other climate variables. The validation statistics for MN2T are given in Table 4.

The correlation between the reanalyses and the meteorological station values is also the lowest for MN2T amongst the three temperature variables. This is due to unexpectedly low correlation values of 0.650, 0.661 and

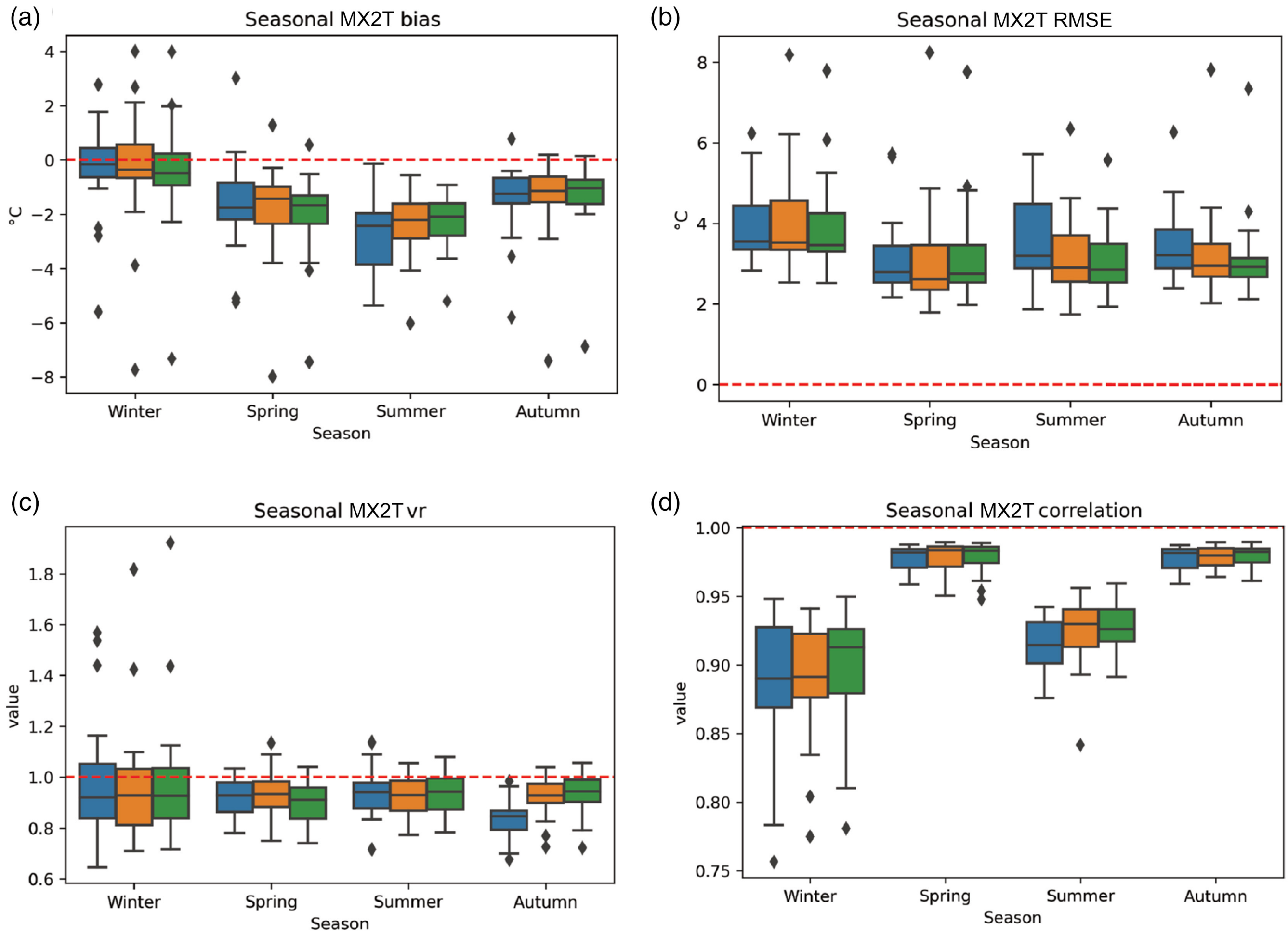

FIGURE 7 Seasonal values of the ERA-Interim (blue), ERA5 (orange) and ERA5-Land (green) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for maximum 2-metre air temperature (MX2T). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × interquartile range (IQR).

TABLE 4 Validation statistics for daily MN2T.

Reanalysis	Bias (°C)	RMSE (°C)	*vr*	*r*
ERA-Interim	2.16	4.88	0.89	0.972
ERA5	1.54	4.69	0.95	0.972
ERA5-Land	1.80	4.69	0.95	0.973

Abbreviations: *r*, Pearson correlation coefficient; MN2T, minimum 2-metre air temperature; RMSE, root-mean-square error; *vr*, variance ratio.

0.656, for ERA-Interim, ERA5 and ERA5-Land respectively, during the summer months, (Figure 8d). Freychet et al. (2017) used ERA-Interim to look at MX2T and MN2T events during summer heatwaves in China and suggested that during the summer months the night-time MN2T stays too high as a result of insufficient cooling by the atmospheric moisture.

The correlations for the other seasons are also the lowest of the three temperature variables, however the differences are markedly less than in summer, with average values of 0.853, 0.921 and 0.950 in winter, spring and autumn, respectively. As with T2M, the MN2T RMSE is lowest in summer, at 3.85°C, however the season with the highest RMSE is spring (5.49°C). Similarly, spring has the lowest variance ratio values across the three temperature variables, with an average of 0.75, and, as with MX2T, the ERA-Interim variance ratio is noticeably lower in autumn than both ERA5 and ERA5-Land. As expected, the most pronounced bias is in winter, 2.40°C, as the reanalyses struggle to attain the coldest minimum temperatures. The bias is smallest in autumn, at an average of 1.34°C. Most outliers occur in winter for the bias and RMSE, again highlighting the difficulty of the reanalyses in reproducing the coldest temperatures.

There are nine meteorological stations at which the correlations with the three reanalyses are less than 0.600 in summer. These are 24329 (W Yakutsk), 24967 and 31062 (E Yakutsk), 30493 and 31137 (Amur region),

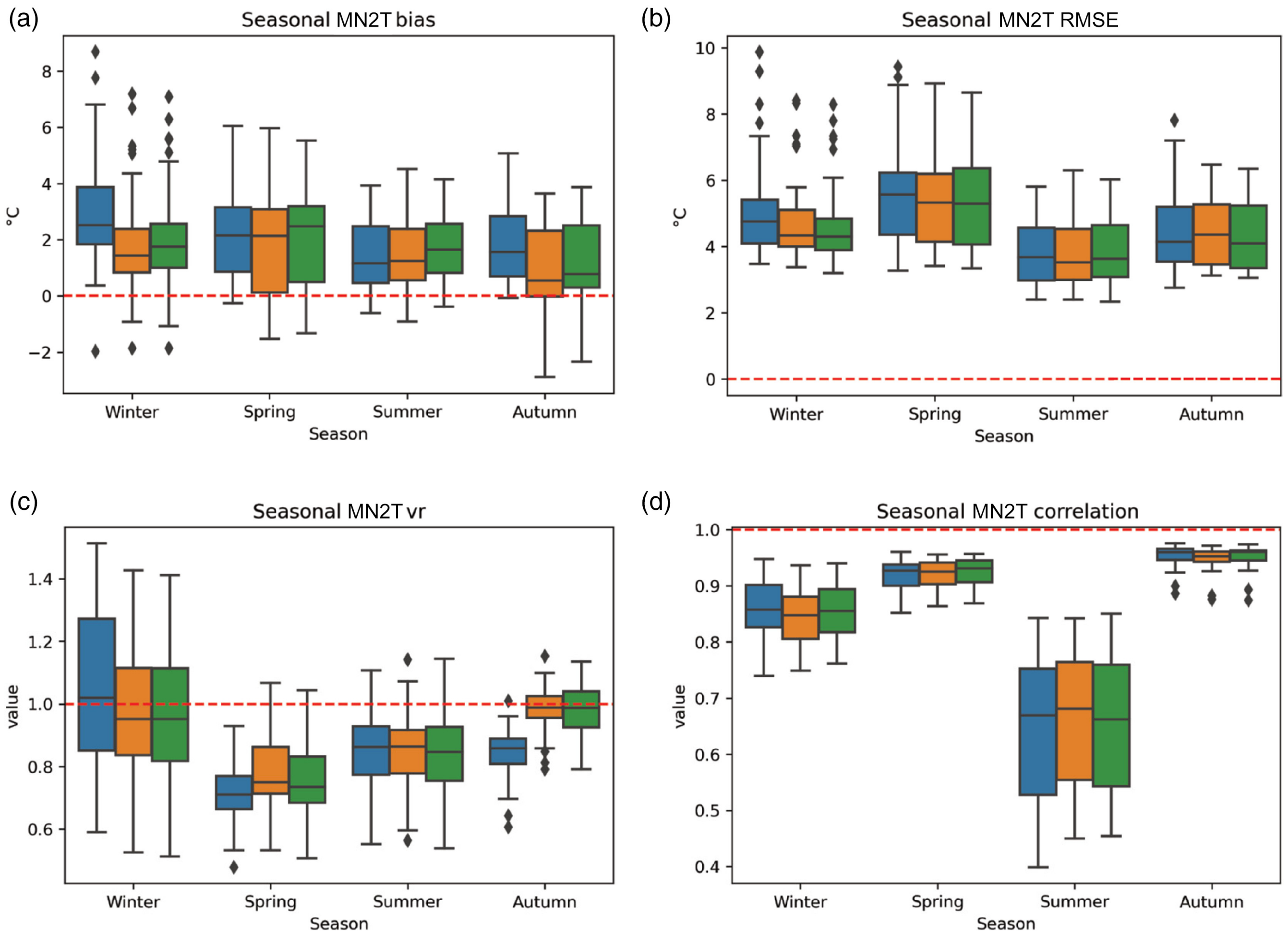

FIGURE 8 Seasonal values of the ERA-Interim (blue), ERA5 (orange) and ERA5-Land (green) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for minimum 2-metre air temperature (MN2T). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × interquartile range (IQR).

29789 and 36103 (SW Siberia), and 25325 and 25503 (NE Sakha). During summer these stations exhibit greater average biases (2.30°C), and RMSEs (4.86°C), and lower variance ratios (0.76) across the three reanalyses than the other meteorological stations, which have a mean bias, RMSE and variance ratio of 1.10°C, 3.35°C and 0.88, respectively. The stations are spread across five of the seven research areas and located at different altitudes (from 97 to 983 m above sea level). Most of these stations are surrounded by hills or mountains and at the bottom of a valley, however 24329 is located in a flat, swampy area. It is therefore likely local conditions play a significant, increased role at these stations, which cannot be picked up by the coarse spatial resolution of the reanalysis datasets.

Figure 9 shows the performance of ERA-Interim, ERA5 and ERA5-Land against meteorological station data for MN2T. The scatter plots are again similar to those for T2M and MX2T, however the greater number of points above the $y = x$ line below 0°C highlights the moderate warm bias, with the reanalyses struggling to attain the coldest daily minima. Similar to the other temperature variables, there is little to differentiate between the three reanalyses. However, as with T2M, Verkhoyansk region is the area where the reanalyses perform least well, most notably in spring and autumn, and NW Sakha has the lowest bias and RMSE values of the seven areas in winter, summer and autumn. Note that Verkhoyansk region is the area with the fewest number of usable stations (only two), therefore returns the fewest number of observations, and often appears as the "weakest" area in this validation for several of the climate variables.

When considering the three 2-metre air temperature variables together, ERA5 performs the best, although the advantage over ERA-Interim and ERA5-Land is negligible.

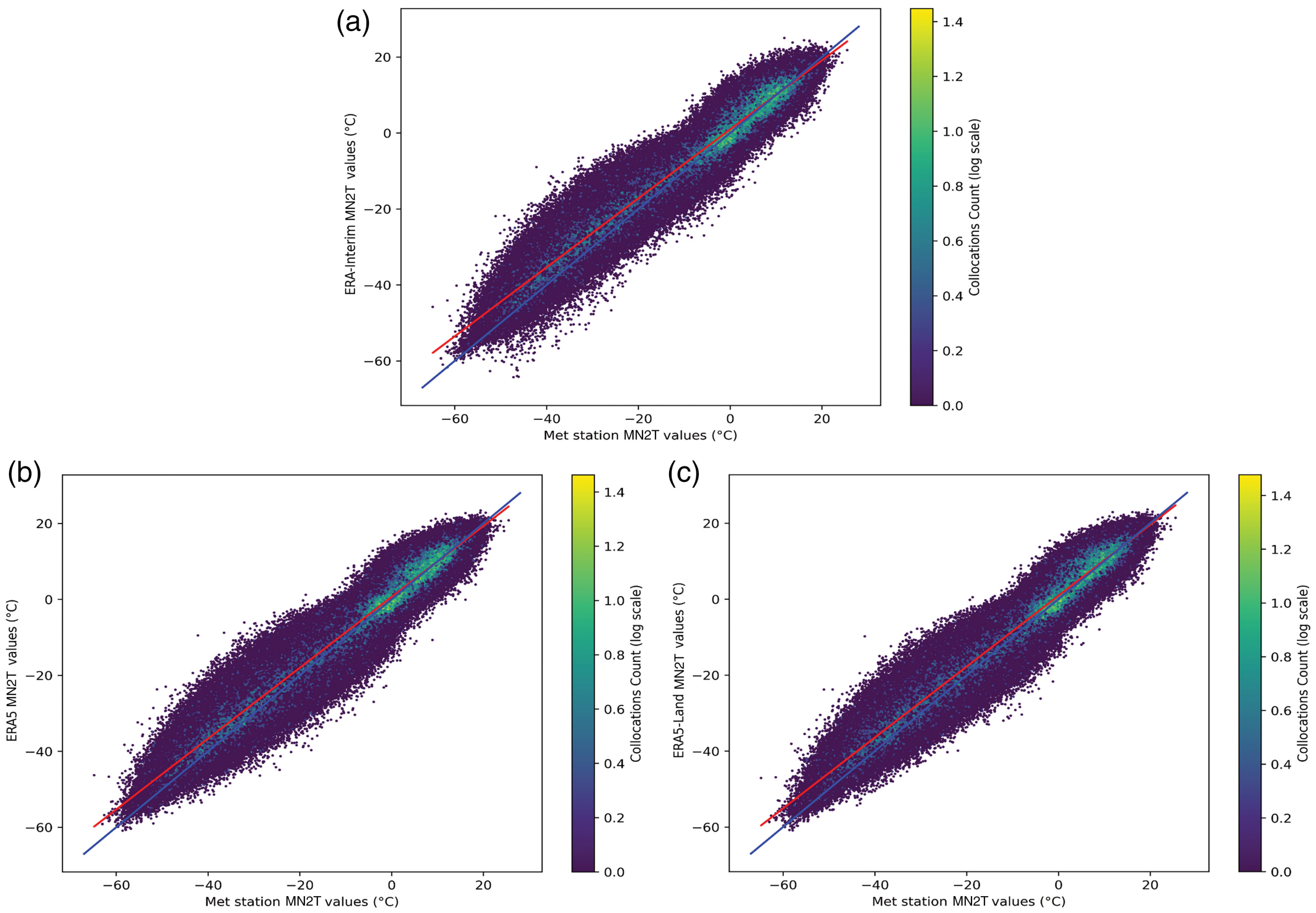

FIGURE 9 Daily minimum 2-metre air temperature (MN2T) values for ERA-Interim (a), ERA5 (b) and ERA5-Land (c) with respect to the corresponding values from each station. The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

TABLE 5 Validation statistics for the daily SD for the periods 1979–1991 and 1992–2019.

	Bias (cm)		RMSE (cm)		*vr*		*ρ*	
Reanalysis	**79–91**	**92–19**	**79–91**	**92–19**	**79–91**	**92–19**	**79–91**	**92–19**
ERA-Interim	1.84	3.19	10.83	11.80	1.21	2.39	0.883	0.910
ERA5	5.25	1.67	12.70	6.13	2.22	1.59	0.894	0.945
ERA5-Land	7.20	5.79	14.40	11.76	2.49	2.13	0.877	0.906

Abbreviations: *ρ*, The Spearman rank correlation coefficient; RMSE, root-mean-square error; SD, snow depth; *vr*, variance ratio.

3.2 | Snow depth

The daily SD provides the most interesting and variable results. In general, ERA5 only begins to correspond closely with in-situ observations from 1992: station data are available before this time but were not assimilated. Urraca and Gobron (2023) demonstrated an additional small correction to values in the mid-2000s associated with the introduction of the assimilation of a satellite-based binary snow cover product. Further commentary on this issue is given in the discussion. The number of assimilations at each station per year is given in Table S1. Where the data are not assimilated, ERA5 tends to vastly overestimate SD, albeit still picking up trends with a high correlation. ERA5-Land takes its boundary conditions from ERA5, so the link with values from meteorological stations is weakened (H. Hersbach, personal communication, 14 July 2022). Conversely, ERA-Interim appears to perform better in the earlier period, with a lower bias, RMSE and *vr*. The validation statistics, split into two periods, are provided in Table 5.

The daily SD values in the reanalyses against the meteorological station readings in the two time periods are shown in Figure 10. Although the trends across ERA5 and ERA5-Land are broadly similar, there is an upper limit of around 140 cm in ERA-Interim in the second time period (Figure 10d), which is even sometimes reached when the observations demonstrate there was no snow cover.

Whereas ERA5 can be used to find errors in mean near-surface temperature station data (Figure 2); here we found that station data can be used to find errors in ERA-Interim SD in the 1992–2019 period. Figure 11 highlights the four main stations where these errors occur. These graphs suggest why ERA-Interim performs better in the 1979–1991 period as the graphs spike most pronouncedly post-2000. As the SD for the reanalyses was calculated using both the ECMWF SD and snow density values, both values were checked for irregularities. We found that there were no irregularities with the snow density values; typical values during winter are between 285 and

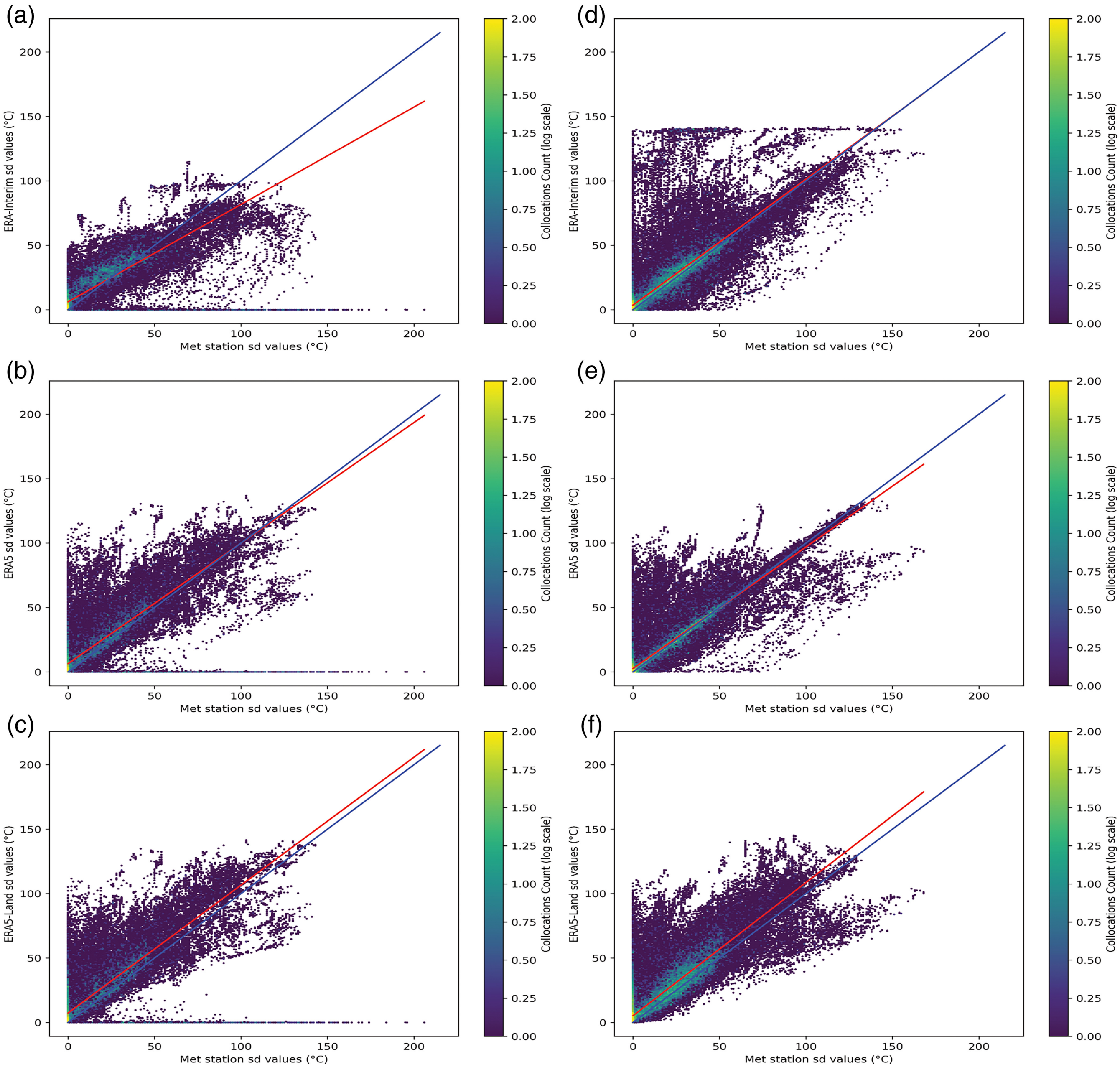

FIGURE 10 Daily snow depth (SD) values for ERA-Interim, ERA5 and ERA5-Land with respect to the corresponding values from each station, respectively, for the periods 1979–1991 (a–c) and 1992–2019 (d–f). The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

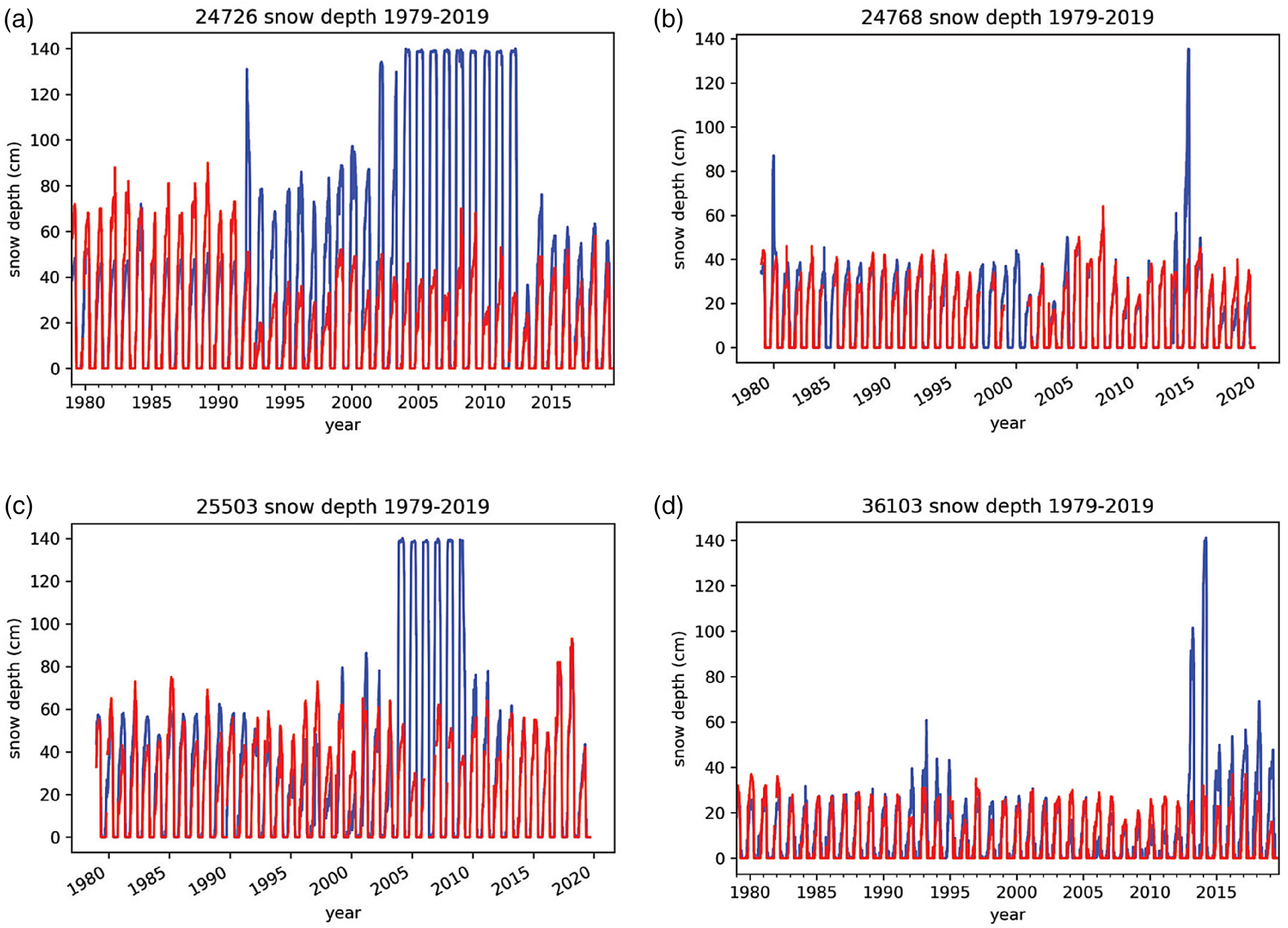

FIGURE 11 ERA-Interim (blue) and meteorological station (red) snow depth (SD) values, 1979–2019, for stations 24726 (a), 24768 (b), 25503 (c) and 36103 (d). Note that station 24768 has no data available in 1984, 1997, 1999 and 2000.

299 kg m^{-3}, and these values were present during the years where the SD values were unusual. However, there are irregular jumps in the original reanalysis SD values that are inconsistent with previous years. Typical values range between 0.06 and 0.25 m of water equivalent (mwe), and on several occasions, these jump to between 0.38 and 0.42 mwe, causing the irregularities in the calculated SD readings.

A previous issue has been highlighted (NCAR, 2023) where, between 1 July 2003 and 23 February 2010, data locations in ERA-Interim were shifted around 100 km to the south–east, as a result of the processing of National Environmental Satellite, Data, and Information Service snow cover data. This would explain most of the inaccurate ERA-Interim values at stations 24726 (Figure 11a) and 25503 (Figure 11c), although some of the former's erroneous values occur beyond 2010, as do the errors at stations 24768 and 36103. These errors do not occur in ERA5 (and therefore ERA5-Land) as SD readings are assimilated from station observations using 2D-Optimal interpolation in the data assimilation—as opposed to Cressman interpolation in ERA-Interim—and an enhanced snowpack parameterisation (Hersbach et al., 2020, P. Berrisford, personal communication, 30 May 2023). This helps to represent the snowmelt runoff timing more realistically and better align the albedo to satellite products (Dutra et al., 2010).

During the summer months SD is predominantly zero at all the study stations, and so this season was not included in the seasonal analysis (Figure 12). When comparing the periods 1979–1991 and 1992–2019, the statistics show the reanalyses have the same differences in the non-summer months, such that there are consistent improvements in all the statistics during the second time period.

The clear improvement of ERA5 in particular is highlighted in Figure 12, and ERA5 outperforms both ERA-Interim and ERA5-Land, especially after 1992. Highlighting ERA5, the greatest improvement in the

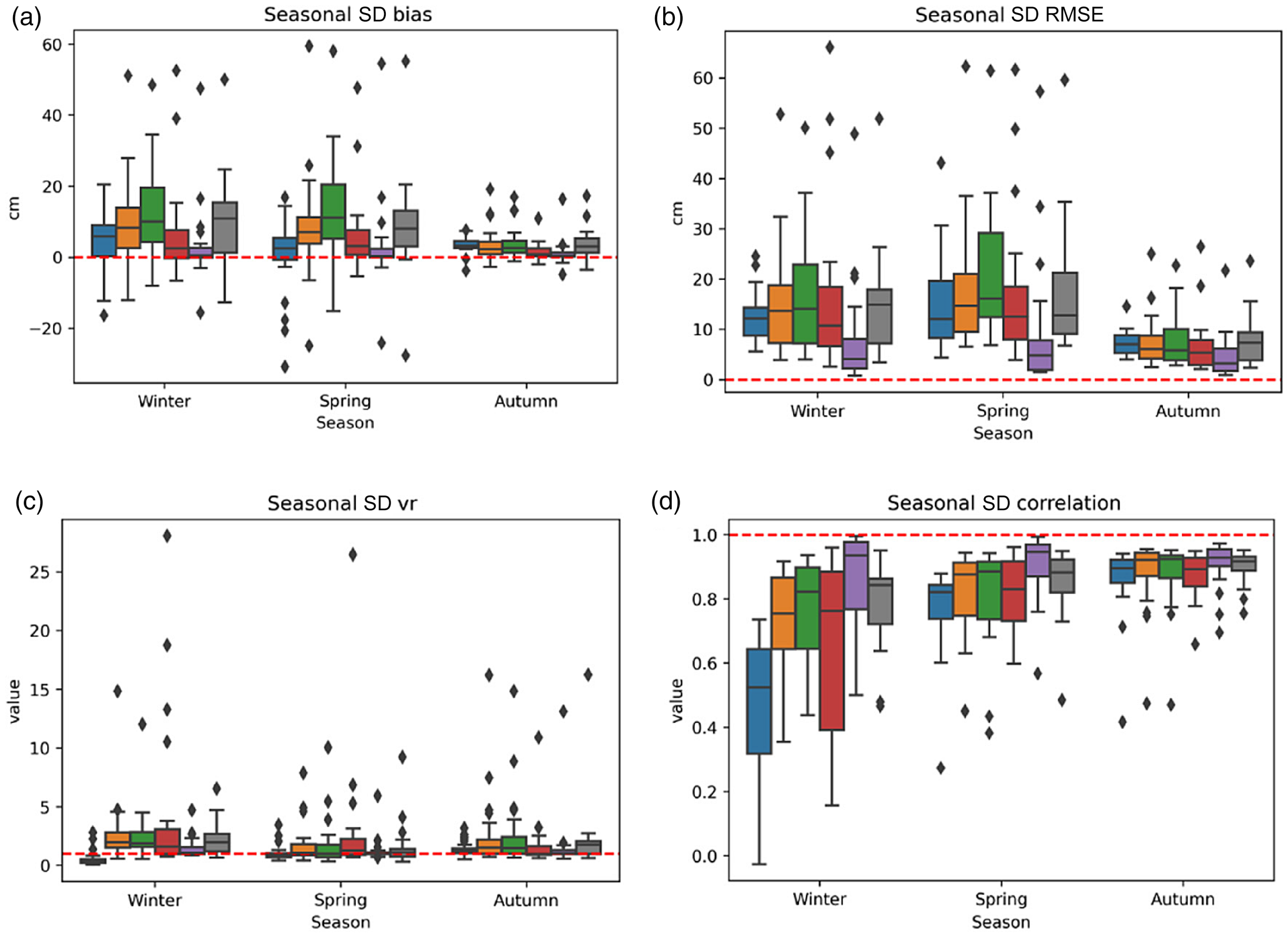

FIGURE 12 Seasonal values of the ERA-Interim 1979–1991 (blue), ERA5 1979–1991 (orange), ERA5-Land 1979–1991 (green), ERA-Interim 1992–2019 (red), ERA5 1992–2019 (purple) and ERA5-Land 1992–2019 (grey) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for snow depth (SD). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × IQR.

correlation, variance ratio and bias is in winter, with an increase of 0.132 and decrease of 1.10 and 6.65 cm, respectively, over the previous period. The greatest improvement in the RMSE is in spring, with a decrease of 8.61 cm, from 16.92 to 8.31 cm.

Figure 13 shows the yearly SD of ERA5 against the values from three different meteorological stations: 23884, 30493 and 30504. The difference before and after 1992 is clear. With the exception of 1983 and 1984, ERA5 assimilates only once or twice per year throughout the 1980s, then at least 110 times from 1992, increasing up to around 200 assimilations per year through the 2010s.

Similarly, Figure 14 shows three stations, 24371, 24763 and 29789, where the data from meteorological stations are almost never assimilated at all. The three stations are located in different areas of Siberia (Figure 1), highlighting the regional variability in the ERA5 SD as there is no consistency between them. Throughout the 1990s ERA5 vastly overestimates the values at station 24371 (Figure 14a), whereas from the early 2010s onwards the reanalysis data become very close to the station readings. Figure 14b shows that ERA5 performs well at station 24763 in the 1980s, 1990s and later 2010s, although underestimates SD during the 2000s. In contrast, ERA5 consistently significantly overestimates the values from station 29789 (Figure 14c). Station 24371 is in Verkhoyansk region with mountains rising 1000 m on three sides, and hills rising 300 m on the other, whereas 29789, in SW Siberia, is halfway down a steep valley, surrounded by mountains rising 500 m. Station 24763 is the anomaly here, located in E Yakutsk with flat, swampy surroundings.

The ECMWF's yearly statistics on the ERA5 usage of in-situ SD observations show that the coordinates of some of the studied meteorological stations are also altered in the ERA5 assimilation model over time (H. Hersbach, personal communication, 14 July 2022). In general, this means an improvement in the precision of

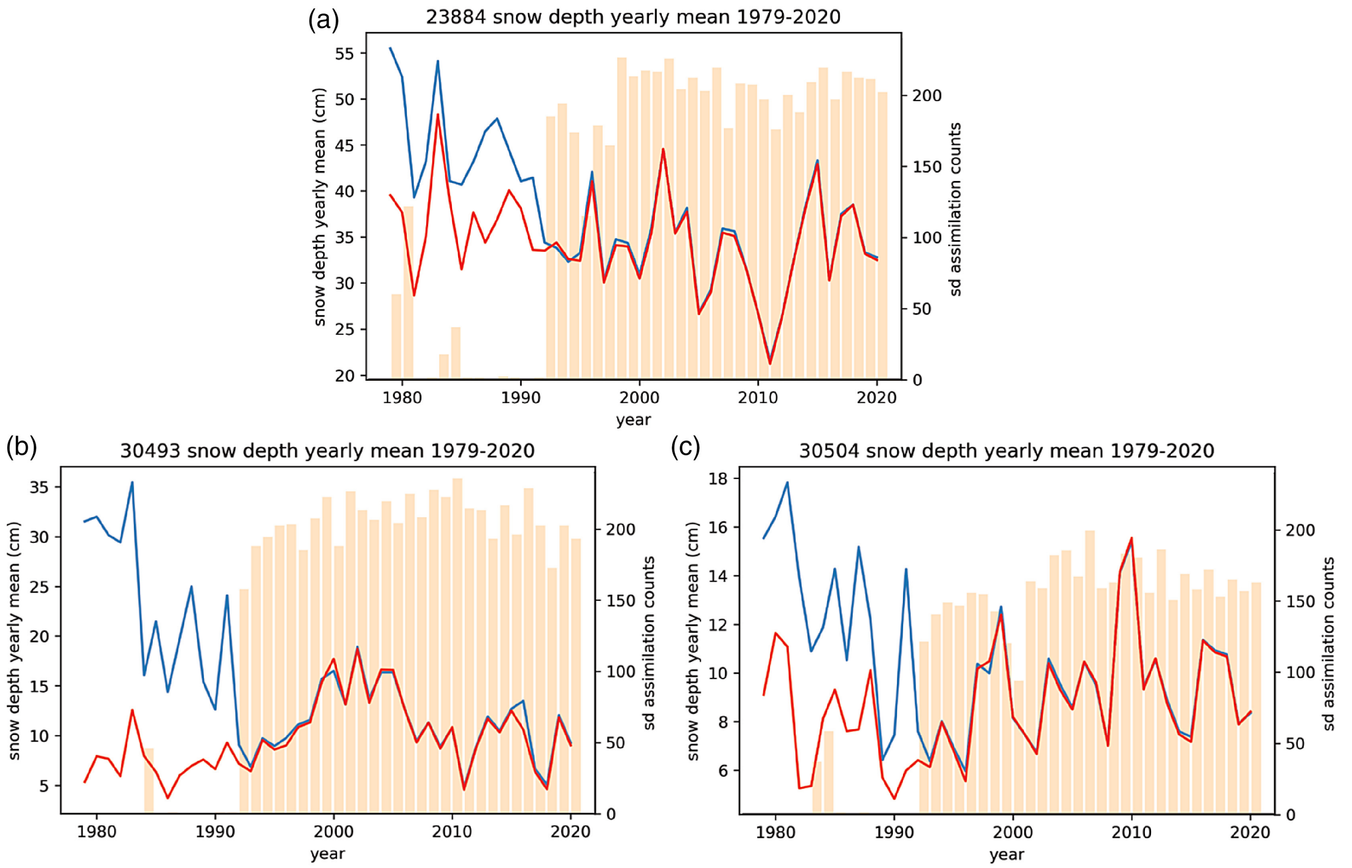

FIGURE 13 ERA5 (blue) and meteorological station (red) snow depth (SD) values per year, 1979–2020, for stations 23884 (a), 30493 (b) and 30504 (c), with their respective snow depth assimilation counts.

the meteorological station location, leading to an increased number of yearly assimilations. However, there are two stations where a change in the station coordinates does not benefit the SD reanalysis values. At 24266 (Verkhoyansk region), the change in coordinates in 2014 yields an improvement in the station location, however the number of assimilations drops to 0 and never recovers. In 2009, the coordinates of station 24967 (E Yakutsk) change from (60.17° N, 130.20 E) to (60.47° N, 130.00 E), implying that the station has moved from the outskirts of a town to a very remote area in the surrounding large hills. The current station location is disputed amongst sources (e.g., Gladstone, 2023; Klein Tank et al., 2002; Lobanov & Kirillina, 2019), although satellite imagery would suggest that the station location has not been changed. However, ERA5 does not assimilate 24967 before 2009 and the number of assimilations increases post-2009 to around 200 per year. 24967 is one of the few stations where the assimilation count makes little difference in reanalysis accuracy compared with the meteorological station values: ERA5 actually underestimates the recorded SD value peaks in both 2008 and 2012, and in general produces values close to those from the meteorological station pre- and post-2009.

In the first time period, there are no marked spatial differences in the quality of all three reanalyses, whereas in the second period the SD is best represented in NW Siberia and consistently least well represented in SW Siberia, mainly due to the performance of ERA5 and ERA5-Land at station 29789. In both time periods there are large inter-regional differences in all the SD statistical variables, the largest differences of all the studied climate variables. For example, the largest bias difference is between NW Siberia and SW Siberia (10.88 cm) in the early time period and these two regions have a variance ratio difference of 6.09 in the later period. The largest difference between the correlations occurs in the first time period between W Yakutsk and Verkhoyansk region (0.346) and the largest RMSE difference occurs in the second time period between SW Siberia and Verkhoyansk region (11.38 cm). Overall, the reanalyses represent SD better at more northerly latitudes. Post-1992, there is a considerable advantage to using ERA5

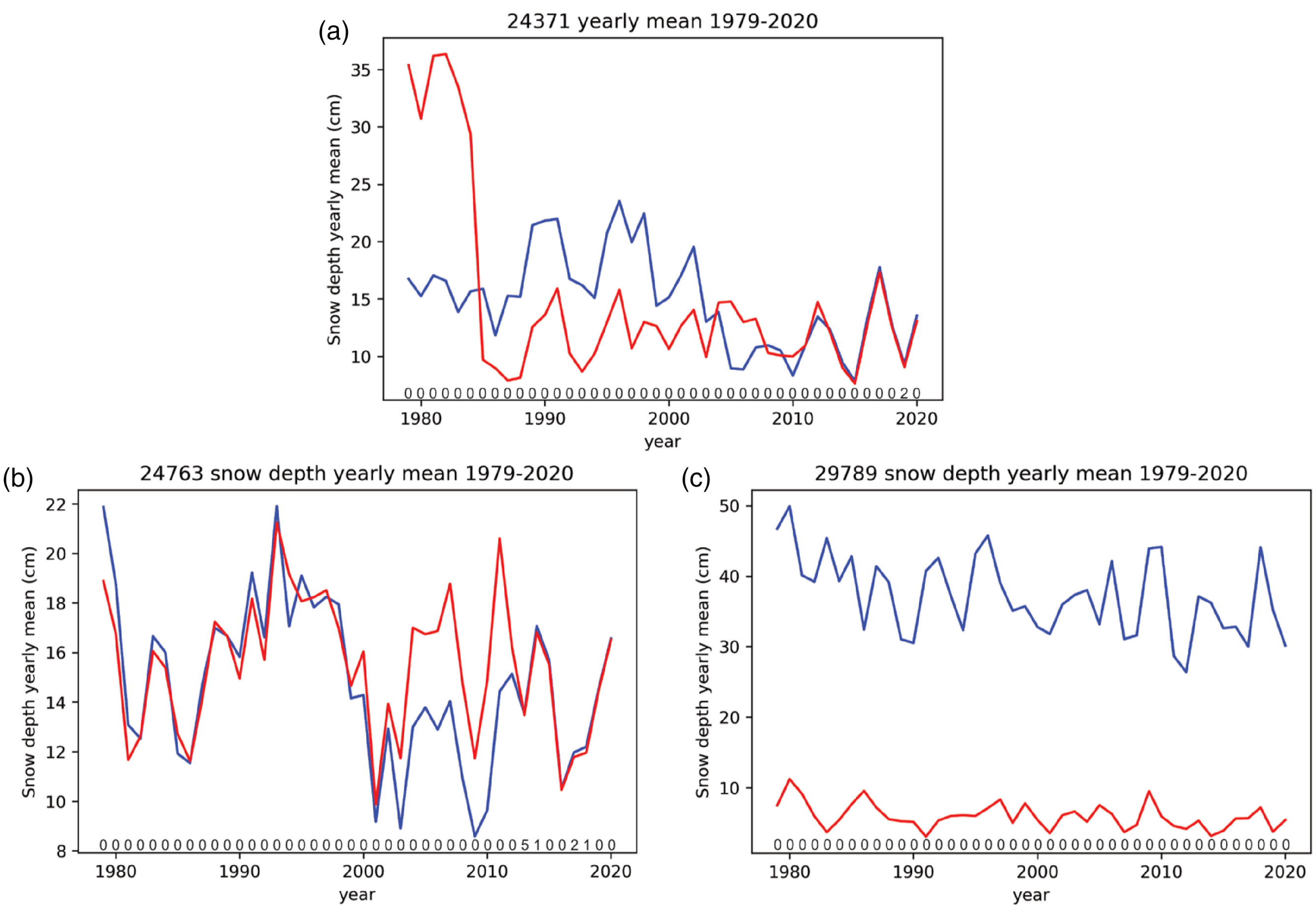

FIGURE 14 ERA5 (blue) and meteorological station (red) snow depth (SD) values per year, 1979–2020, for stations 24371 (a), 24763 (b) and 29789 (c) with the number of assimilations per year.

over both ERA-Interim and ERA5-Land; however, pre-1992 ERA-Interim and ERA5 perform similarly.

3.3 | Total precipitation

TP is calculated as the sum of the hourly/6-h precipitation and is the variable that the reanalyses reproduce least well out of those examined in this study, as shown in Table 6.

The reanalyses tend to marginally overestimate TP and there is little difference between them, with almost no difference at all between ERA5 and ERA5-Land. On a daily timescale there are no significant differences between the performances of the reanalyses in the seven study regions. As TP observed at the meteorological stations increases, the deviation from the $y = x$ line increases (Figure 15); ERA-Interim underestimates more than the other two reanalysis datasets as precipitation increases.

Unlike the temperature and SD variables, TP performs best in winter (Figure 16) as the majority of the extreme precipitation events, when the reanalyses perform worst, occur in summer. They are likely often associated with convective precipitation, which occurs on spatial scales too small for the reanalyses to represent accurately. The winter months have the lowest bias and RMSE (0.05 and 0.78 mm), and the second-lowest variance ratio (0.96) and correlation (0.554). On the contrary, summer has the greatest bias and the RMSE, 0.43 and 4.90 mm, respectively, whereas the variance ratio is furthest from 1.00 in spring (1.23).

Of the seven study areas, SW Siberia is the region where the reanalyses perform least well in winter and

TABLE 6 Validation statistics for daily TP.

Reanalysis	Bias (mm)	RMSE (mm)	*vr*	*ρ*
ERA-Interim	0.26	2.93	0.82	0.534
ERA5	0.30	2.97	0.96	0.549
ERA5-Land	0.30	2.97	0.96	0.552

Abbreviations: *ρ*, The Spearman rank correlation coefficient; RMSE, root-mean-square error; TP, total precipitation; *vr*, variance ratio.

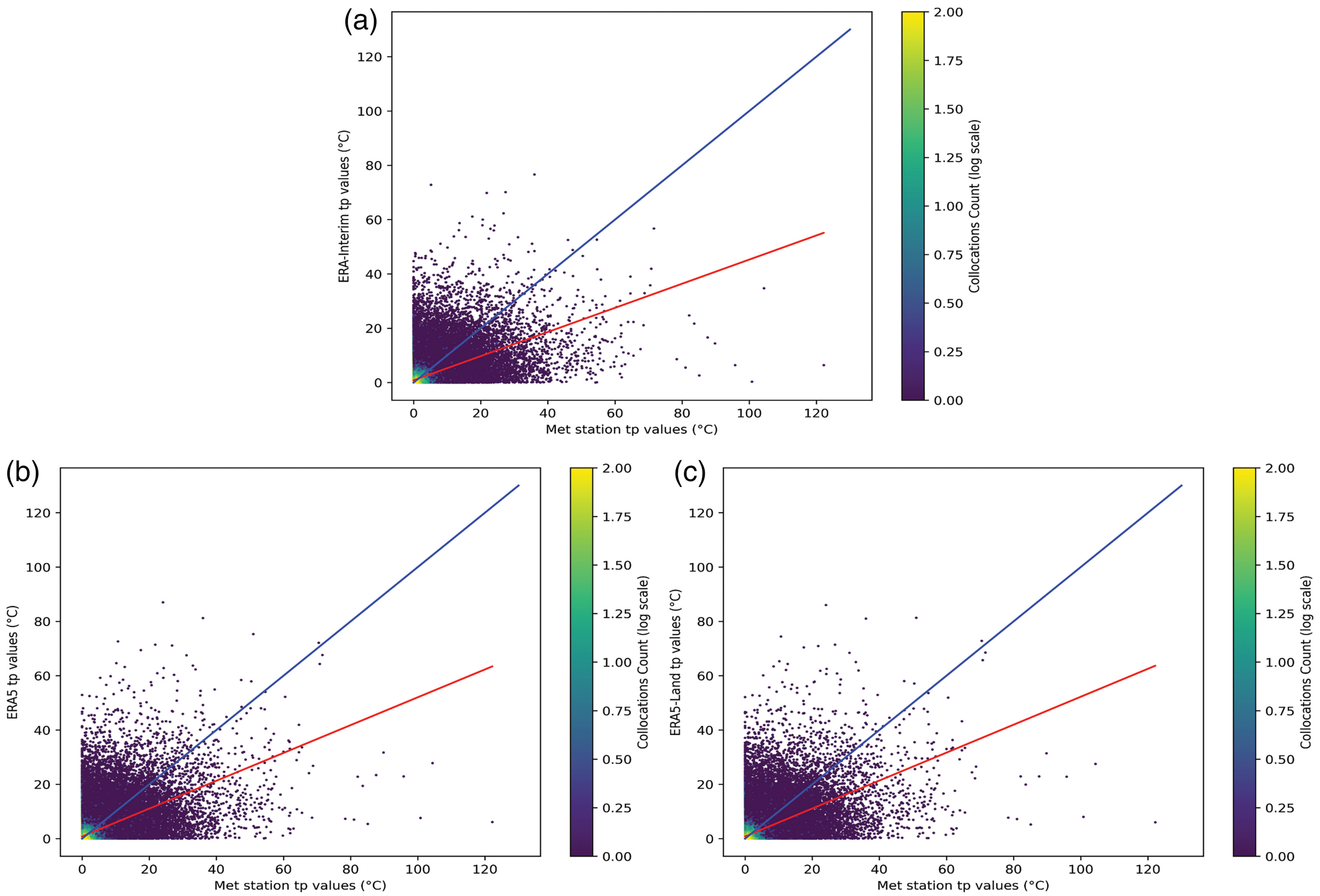

FIGURE 15 Daily total precipitation (TP) values for ERA-Interim (a), ERA5 (b) and ERA5-Land (c) with respect to the corresponding values from each station. The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

autumn, with the largest average biases of 0.31 and 0.66 mm, respectively, and the lowest correlations of 0.457 and 0.523, respectively. The area experiences an increased number of thunderstorms due to its geographical location close to the Altai Mountain range (Gorbatenko et al., 2019; Kocheeva et al., 2018). In autumn, NW Siberia is the region with the highest average correlation values (0.644) and lowest biases (0.108 mm), but there are no other significant seasonal spatial trends.

When considering TP, the difference between ERA5 and ERA5-Land is negligible, and the improvement of ERA5 (or ERA5-Land) over ERA-Interim is also minimal.

3.4 | Wind speed

WSP is the only variable studied where there is a noticeable difference between the performance of ERA5 and ERA5-Land. The validation statistics for WSP are given in Table 7. Both ERA-Interim and ERA5 have a positive bias and a variance ratio >1.00, whereas ERA5-Land has a negative bias and a variance ratio <1.00. ERA5-Land has the lowest RMSE and greatest ρ value, whereas ERA-Interim has the highest RMSE and lowest ρ value, although the difference between the maximum and minimum values are small.

Figure 17 shows the performance of the three reanalyses against observations from the meteorological stations. There is a significant difference with how the reanalysis datasets perform with WSPs above and below 5.0 ms^{-1}. At lower WSPs all three reanalyses tend to overestimate, most notably ERA-Interim, however they all greatly underestimate the highest recorded WSPs. ERA5 shows a noticeable improvement over ERA-Interim with reduced variability. ERA5-Land performs better than both ERA-Interim and ERA5 at lower speed; however, it is the weakest of the three at higher speeds, vastly underestimating the extreme highs. For instance, for speeds above 10.0 ms^{-1}, the ERA5-Land bias is $-6.48\ \text{ms}^{-1}$, compared with $-4.22\ \text{ms}^{-1}$ and $-4.49\ \text{ms}^{-1}$ for ERA-Interim and ERA5, respectively.

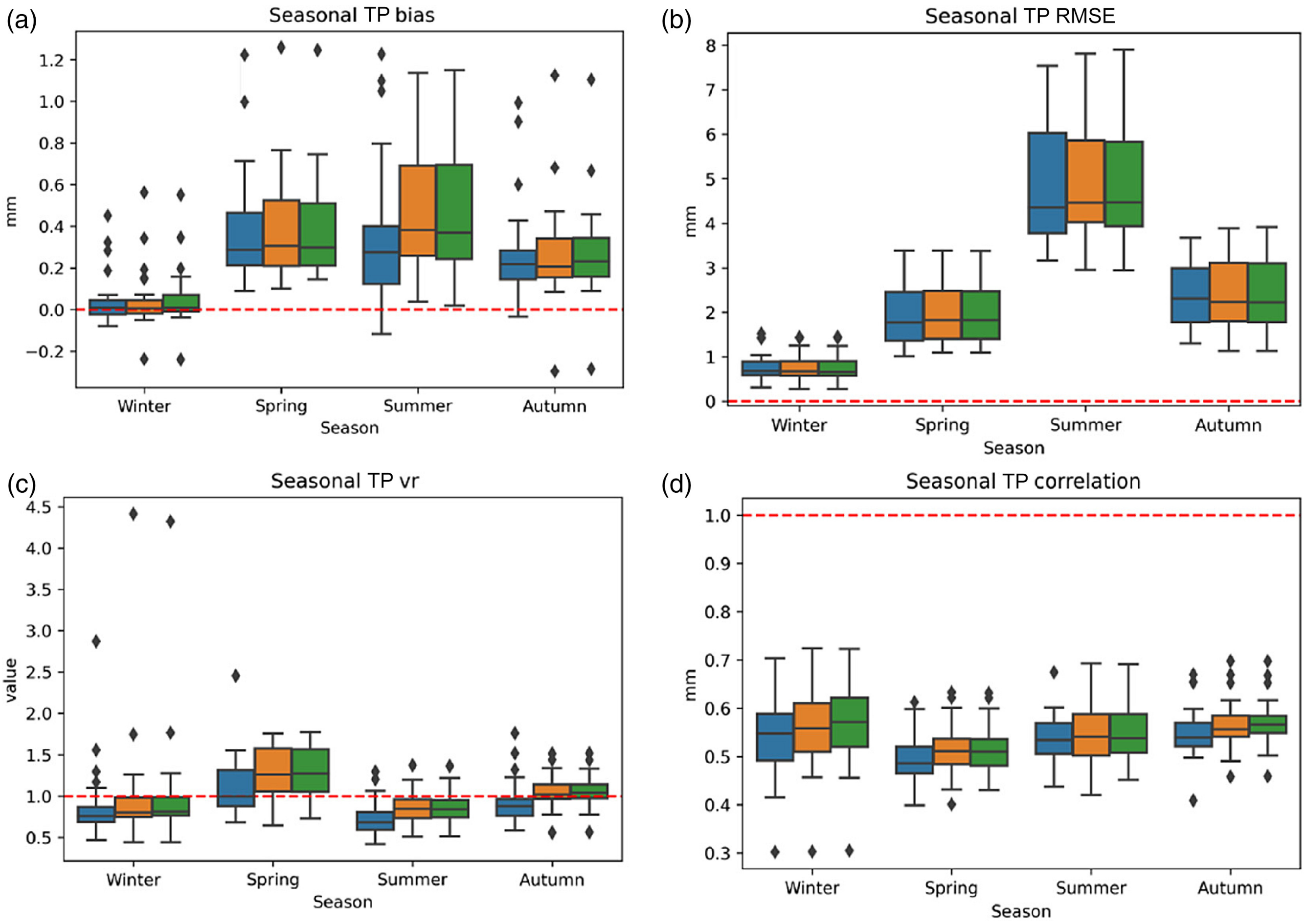

FIGURE 16 Seasonal values of the ERA-Interim (blue), ERA5 (orange) and ERA5-Land (green) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for total precipitation (TP). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × IQR.

TABLE 7 Validation statistics for daily wind speed (WSP).

Reanalysis	Bias (ms^{-1})	RMSE (ms^{-1})	*vr*	*ρ*
ERA-Interim	0.72	1.31	1.19	0.525
ERA5	0.46	1.09	1.02	0.561
ERA5-Land	−0.05	0.94	0.54	0.577

Abbreviations: *ρ*, The Spearman rank correlation coefficient; RMSE, root-mean-square error; *vr*, variance ratio; WSP, wind speed.

Of the six studied variables, WSP has the greatest range in the statistical values between stations. The average correlation between the three reanalyses and the station values at station 36103 is 0.387, whereas at station 23986 it is 0.803. The variance ratio can be as high as 2.74 (station 29789) or as low as 0.34 (station 31102). This highlights that local factors can affect WSP readings more than any other studied variable.

On a regional scale, SW Siberia stands out as the area where the reanalyses reproduce WSP least well, especially in winter and autumn. WSP is the variable where the difference in regional correlation values is largest—the average correlation is greatest in NW Siberia (0.747), and lowest in SW Siberia (0.451). The average biases and RMSE values are relatively low—within 0.85 ms^{-1} of each other.

Figure 18 shows that, across all stations, the performance of the reanalyses in reproducing WSP is not affected by the season. The same broad trends are exhibited across all four seasons in all the statistics, with ERA-Interim having the most variance. ERA5-Land has the lowest bias and RMSE in winter, spring and autumn, and the highest correlations in the latter two seasons. The variance ratios are furthest from 1.00 (too much WSP variability) in summer.

When considering the whole picture, ERA5 and ERA5-Land are an improvement over ERA-Interim. ERA5-Land performs the best at lower WSPs, however ERA5 should be used for extreme high speeds, which are likely to be of most interest.

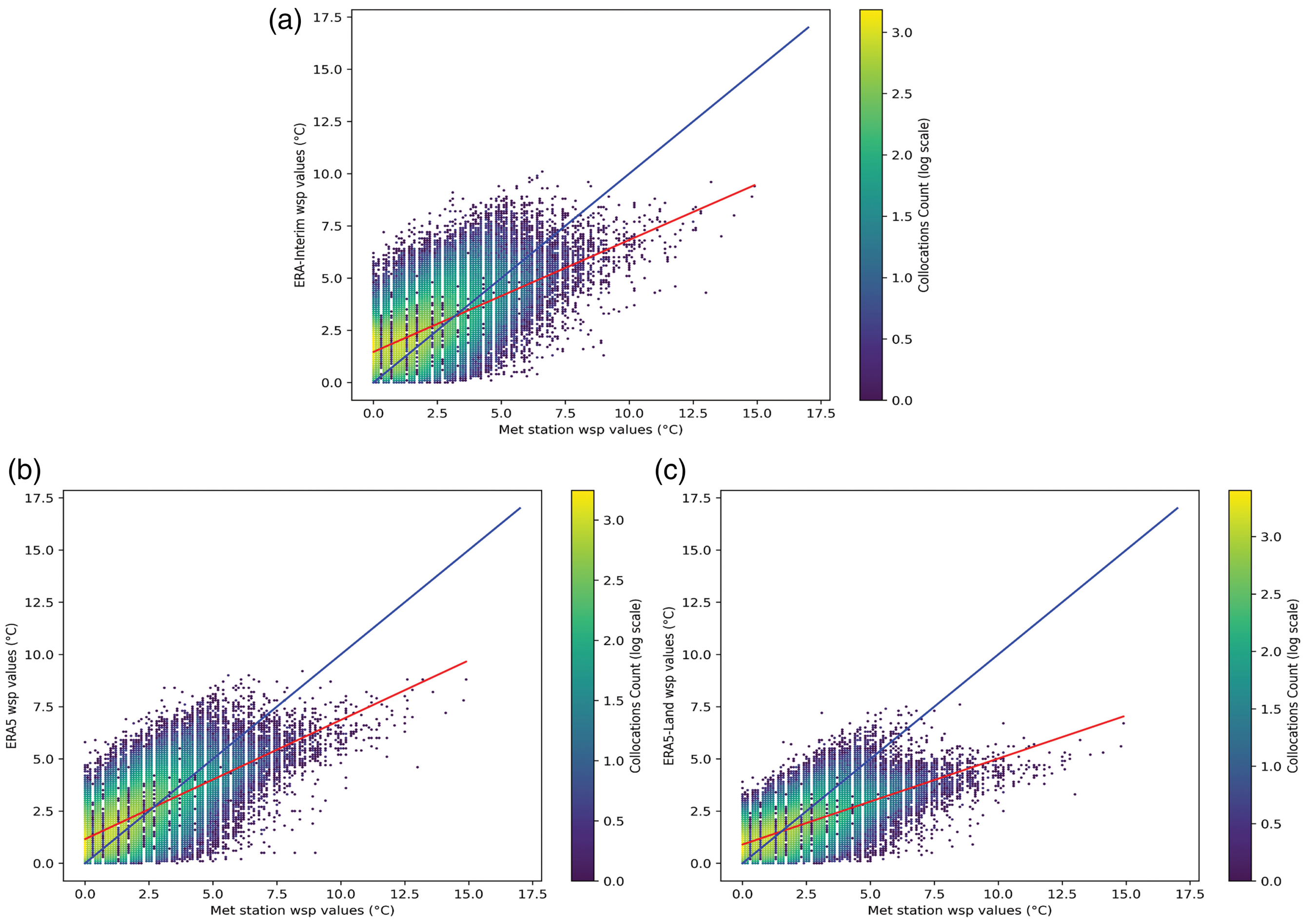

FIGURE 17 Daily wind speed (WSP) values for ERA-Interim (a), ERA5 (b) and ERA5-Land (c) with respect to the corresponding values from each station. The solid blue line indicates perfect agreement, and the line of best fit is in red. The pixel colour scale indicates the number of collocations of each pixel on a logarithmic scale.

3.5 | Analysis of data back to 1959

For ERA5 and ERA5-Land, the T2M, MX2T, MN2T, SD and TP validations were extended back to 1959, whereas the validations for WSP only start from 1966 due to the availability of data from the meteorological stations. Reanalyses have tended to struggle before 1979 due to reduced sounding data from instruments onboard polar orbiting satellites, fewer data from commercial aircraft and the lack of drifting buoys (Andersson, 2007; Minnett et al., 2019). Note that missing meteorological station data also becomes more of an issue prior to 1979, so only the key results and general trends are summarized here.

T2M, MX2T and MN2T are homogeneous back to 1959 in that the validation statistics exhibit the same seasonal variability, and the average bias and RMSE are within 0.5°C for both ERA5 and ERA5-Land across all three variables.

With TP, there is little difference on a daily basis from 1979, and the seasonal trends exhibit the same patterns as the earlier period. The differences in the biases and RMSEs are negligible. Although there are small changes in the validation statistics for WSP prior to 1979, they are not consistent across the two reanalysis datasets, leading to the conclusion that the accuracy of the reanalysis WSP is relatively homogeneous through time.

SD is the only variable for which the statistics are not temporally homogeneous. As discussed before, there are significant differences between the periods 1979–1991 and 1992–2019 (Table 5), however there is also a noticeable distinction between 1959–1978 and 1979–1991. The statistics for the first period are in Table 8.

Post 1979, the average correlation across the two reanalyses improves by 0.058, the mean daily RMSE decreases by 2.55 cm and the average bias is reduced by 1.57 cm. The seasonal trends remain consistent, which

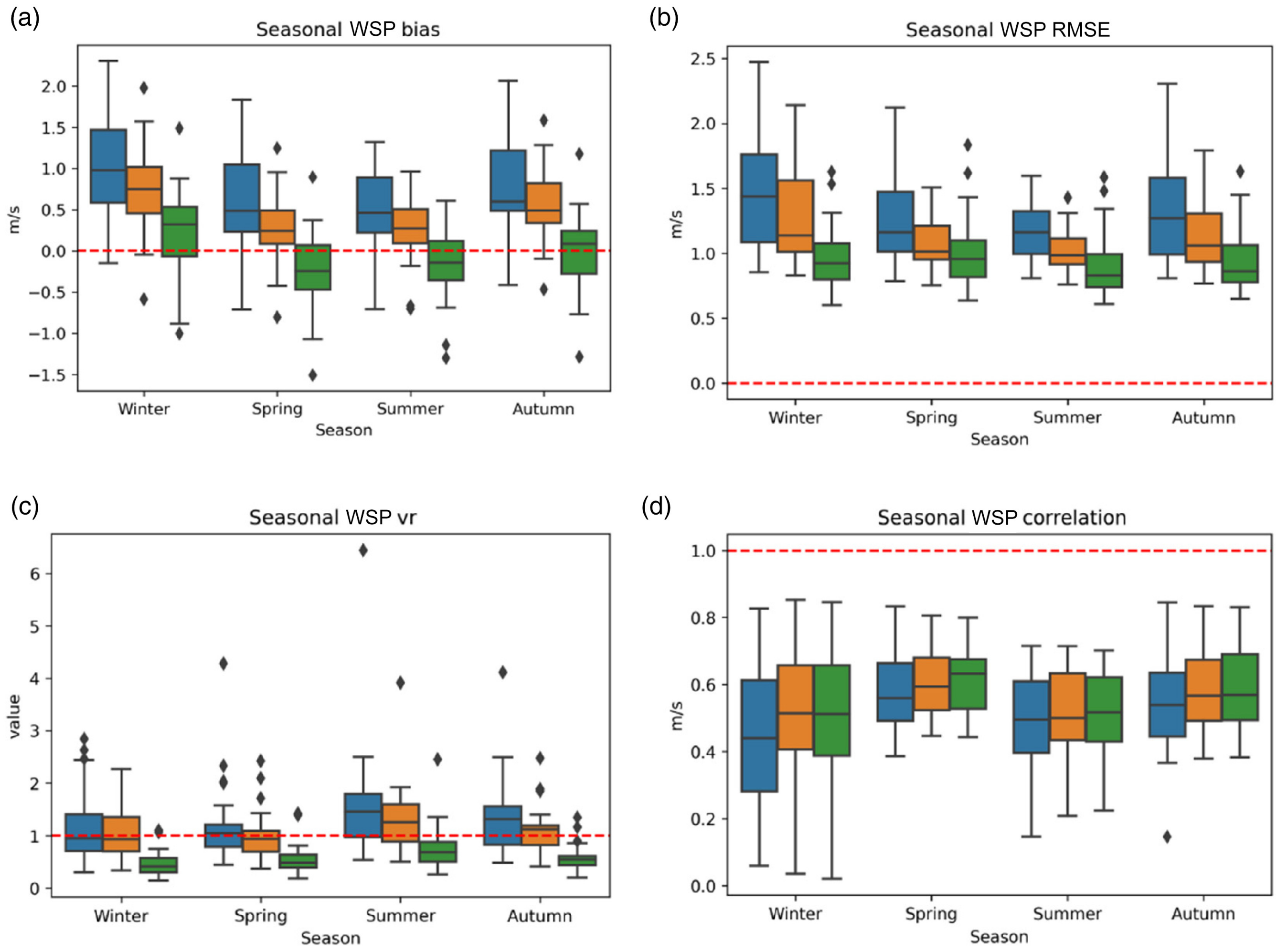

FIGURE 18 Seasonal values of the ERA-Interim (blue), ERA5 (orange) and ERA5-Land (green) bias (a), root-mean-square error (RMSE) (b), variance ratio (c) and correlation (d) statistics for wind speed (WSP). The red dashed line indicates the optimum value for each statistic; the whiskers extend to a maximum of 1.5 × IQR.

TABLE 8 Validation statistics for the daily SD for the period 1959–1978.

Reanalysis	Bias (cm)	RMSE (cm)	*vr*	*ρ*
ERA5	7.68	15.87	2.49	0.856
ERA5-Land	7.91	16.33	2.44	0.838

Abbreviations: *ρ*, The Spearman rank correlation coefficient; RMSE, root-mean-square error; SD, snow depth; *vr*, variance ratio.

again reaffirms the first discontinuity noted by Urraca and Gobron (2023).

4 | DISCUSSION

ERA-Interim, ERA5 and ERA5-Land can all be considered as excellent representations of T2M, MX2T and MN2T across Siberia and the Russian Far East. ERA5 and ERA5-Land are homogeneous from 1959, with correlations over 0.98 and small biases and RMSEs. There is only a marginal improvement with using ERA5 over ERA-Interim, and there is no additional benefit to using ERA5-Land over ERA5. The performance of the reanalyses at reproducing temperature is weakest in winter as the temperatures approach and reach their minima, although compared with the other studied variables the seasonal difference is minimal.

The results of our evaluation are consistent with previous validations of ERA-Interim and ERA5 against in-situ observations in comparable regions. As expected, when considering T2M, the reanalyses perform strongly with little variation across studies. The ERA5 T2M validation statistics here perform better than in Turkey (Yilmaz, 2023), and the Greenland Ice Sheet (Delhasse et al., 2020), and similar to results from the Yamal Peninsula (Matveeva & Sidorchuk, 2020), the Canadian Prairies (Betts et al., 2019), and the Arctic Gateway (Graham et al., 2019).

When considering TP, there is little difference between ERA5 and ERA5-Land, however both show some improvement over ERA-Interim. Nonetheless, all three reanalyses struggle with reproducing extreme precipitation events during the summer months, likely due to their inability to simulate the temporal and spatial distribution of convective precipitation. TP is relatively homogeneous through time in ERA5 and ERA5-Land.

Validations of TP in ERA-Interim and ERA5 have been conducted in various other studies, including in the Yamal Peninsula (Matveeva & Sidorchuk, 2020), and in South Siberia (Riazanova et al., 2016; Voropay et al., 2021). While there is a large amount of variability in the validation statistics between all the studies, depending on the research location, our results are consistent in that the correlations are significantly lower than temperature variables (only >0.5), the reanalyses perform worst during occurrences of extreme high precipitation in the summer, and the spatial variability in their quality is relatively small.

The reanalysis validation statistics highlight the marked local variability of WSP. There is a benefit to using either ERA5 or ERA5-Land over ERA-Interim, with ERA5-Land performing better at lower WSPs and ERA5 performing better at the extreme high speeds. WSP shows the lowest seasonal variation of the studied variables amongst the validation statistics, although the station-to-station variability is the highest of all the variables. When considering WSP, the performance of ERA-Interim and ERA5 is again consistent with previous studies. The WSP statistics in this study are improved over the Greenland Ice Sheet area validation (Delhasse et al., 2020), and similar to results in the Black and Azov Seas (Amarouche et al., 2021; Grankina et al., 2019) and the Arctic (Dyakonov et al., 2020; Graham et al., 2019).

SD is the variable where most caution should be taken. ERA5 only consistently begins to assimilate the meteorological station data from 1992; before that the reanalysis vastly overestimates SD. Furthermore, ERA5 never assimilates data from some of the meteorological stations examined, and in these areas the reanalysis continues to overestimate SD. ERA5-Land takes its boundary conditions from ERA5, so the ERA5-Land SD values are further away from the observed values at meteorological stations. ERA-Interim appears to perform better before 1992; however, this is likely due to anomalous values at several stations post-2000, where there appears to be a consistent maximum upper limit of 140 cm. The reanalyses perform weakest in winter and spring when dealing with extreme values.

Only two other studies have conducted validations of SD in reanalyses across Russia, the findings of this study are consistent with results for ERA-Interim over Siberia (Klehmet et al., 2013), and the post-1992 results for ERA5 from the Yamal Peninsula (Matveeva & Sidorchuk, 2020). Urraca and Gobron (2023) identified potential discontinuities in the SD in 1977–1980, 1991–1992 and 2003–2004, with a negative trend between 1980 and 1991. This study reaffirms their findings that, unsurprisingly, the assimilations into ERA5, first in 1979 then 1992, significantly improved the SD values from the reanalysis. However, our study shows the third discontinuity to be in 2004–2005, as opposed to 2003–2004 although, as with the original findings, this discontinuity is relatively minor and likely associated with the start of the assimilation of a satellite snow cover product into ERA5. Graphs from our study showing the temporal evolution of the bias in ERA5 and ERA5-Land SD, equivalent to Figure 4 from Urraca and Gobron (2023), can be found in Figure S1a,b.

The validation could be refined if more meteorological stations were available. Individual stations may not always represent the region as a whole as they can be affected by local climate variables and regional topography. The coarse resolution of the reanalyses may also affect the results as they may be unable to reproduce localized weather activity or topographical influences, and interpolation cannot correct for this. Furthermore, when considering the line of best fit on each plot, the fit around the outliers may not always be well represented due to high-density values carrying more weight.

In this study, we assessed the performance of ERA Interim, ERA5 and ERA5-Land in Siberia and the Russian Far East. We believe this is the first work comparing these three ECMWF reanalyses across a range of climate variables in this region. The results of this validation have shown that for the studied climate variables there is a small benefit to using ERA5 over ERA-Interim in Siberia and the Russian Far East, most notably as regards SD. The improved spatial and temporal resolutions as well as the increased data availability mean that ERA5 should be the reanalysis dataset of choice for the core climate variables moving forward. Despite a further increase in spatial resolution, we found no consistent, significant benefit to using ERA5-Land over ERA5. Thus, we conclude that ERA5 is the most appropriate current ECMWF reanalysis to examine climate change in Siberia.

AUTHOR CONTRIBUTIONS

Andrew A. Clelland: Validation; formal analysis; writing – original draft; investigation; conceptualization; visualization. **Gareth J. Marshall:** Conceptualization;

supervision; writing – review and editing; methodology. **Robert Baxter:** Supervision; writing – review and editing.

ACKNOWLEDGEMENTS

The authors would like to thank Tony Phillips (British Antarctic Survey), Hans Hersbach and Paul Berrisford (ECMWF) for their contributions to this article.

FUNDING INFORMATION

The lead author was funded by the IAPETUS2 Doctoral Training Programme scheme through the Natural Environment Research Council (grant NE/S007431/1).

CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

ERA-Interim, ERA5 and ERA5-Land data are available from the Climate Data Store database (https://cds.climate.copernicus.eu/). The meteorological station data are available from non-Russian climate data archives, such as OGIMET (https://www.ogimet.com).

ORCID

Andrew A. Clelland https://orcid.org/0009-0002-7391-5204
Gareth J. Marshall https://orcid.org/0000-0001-8887-7314
Robert Baxter https://orcid.org/0000-0002-7504-6797

REFERENCES

Amarouche, K., Akpinar, A., Soran, M.B., Myslenkov, S., Majidi, A.G., Kankal, M. et al. (2021) Spatial calibration of an unstructured SWAN model forced with CFSR and ERA5 winds for the black and Azov seas. *Applied Ocean Research*, 117, 102962.

Andersson, E. (2007) Data assimilation in the polar regions. *ECMWF Newsletter*, 112, 10–15.

Betts, A.K., Chan, D.Z. & Desjardins, R.L. (2019) Near-surface biases in ERA5 over the Canadian prairies. *Frontiers in Environmental Science*, 7, 1–17.

Callaghan, T.V., Shaduyko, O., Kirpotin, S.N. & Gordov, E.P. (2021) Siberian environmental change: synthesis of recent studies and opportunities for networking. *Ambio*, 50, 2104–2127.

Cao, B., Gruber, S., Zheng, D. & Li, X. (2020) The ERA5-land soil temperature bias in permafrost regions. *The Cryosphere*, 14, 2581–2595.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S. et al. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597.

Delhasse, A., Kittel, C., Amory, C., Hofer, S., van As, D., Fausto, R.S. et al. (2020) Brief communication: evaluation of the near-surface climate in ERA5 over the Greenland ice sheet. *The Cryosphere*, 14, 957–965.

Demchev, D.M., Kulakov, M.Y., Makshtas, A.P., Makhotina, I.A., Fil'chuk, K.V. & Frolov, I.E. (2020) Verification of ERA-interim and ERA5 Reanalyses data on surface air temperature in the Arctic. *Russian Meteorology and Hydrology*, 45, 771–777.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P.M.A., Beljaars, A., Schär, C. et al. (2010) An improved snow scheme for the ECMWF land surface model: description and offline validation. *Journal of Hydrometeorology*, 11, 899–916.

Dyakonov, G.S., Ibrayev, R.A. & Shishkova, P.O. (2020) Assessment of ERA-interim reanalysis data Qualityfor the Caspian Sea area. *Russian Meteorology and Hydrology*, 45, 650–657.

Freychet, N., Tett, S., Wang, J. & Hegerl, G.C. (2017) Summer heat waves over eastern China: dynamical processes and trend attribution. *Environmental Research Letters*, 12, 24015.

Gale, M.G., Cary, G.J., van Dijk, A.I.J.M. & Yebra, M. (2021) Forest fire fuel through the lens of remote sensing: review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour. *Remote Sensing of Environment*, 255, 112282.

Gladstone, P. 2023. *Synop information for 24967 in Tegyul'tya|Tegjultje-Terde, SA, Russian Federation [online]*. Available from: https://weather.gladstonefamily.net/site/24967 [Accessed 7 August 2023].

Gorbatenko, V., Nechepurenko, O. & Ershova, T. (2019) Characteristics of atmosphere on days with thunderstorms in the southeast of Western Siberia. *IOP Conference Series: Materials Science and Engineering*, 698, 44043.

Graham, R.M., Hudson, S.R. & Maturilli, M. (2019) Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses. *Geophysical Research Letters*, 46, 6138–6147.

Grankina, T.B., Ibrayev, R.A. & Mogilnikov, P.A. (2019) Verification of the ERA-interim reanalysis data in the Azov-Black Sea Basin. *Physical Oceanography*, 26, 236–246.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz-Sabater, J. et al. (2023) ERA5 hourly data on pressure levels from 1940 to present [online]. *Copernicus Climate Change Service*, C3S Accessed 4 July 2023.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.

Kharuk, V.I., Ponomarev, E.I., Ivanova, G.A., Dvinskaya, M.L., Coogan, S.C.P. & Flannigan, M.D. (2021) Wildfires in the Siberian taiga. *Ambio*, 50, 1953–1974.

Klehmet, K., Geyer, B., & Rockel, B. (2013) A regional climate model hindcast for Siberia: analysis of snow water equivalent. *The Cryosphere*, 7, 1017–1034. https://doi.org/10.5194/tc-7-1017-2013

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A. et al. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *International Journal of Climatology*, 22, 1441–1453.

Kocheeva, N.A., Chankibaeva, M.H., Minaev, A.I., Sukhova, M.G. & Modorov, A.A. (2018) Peculiarities of thunderstorms' occurrences on the border of the Western Siberia

Plains and Altai Mountains. *Journal of Environmental Management and Tourism*, 8, 1061–1068.

Lan, H., Guo, D., Hua, W., Pepin, N. & Sun, J. (2023) Evaluation of reanalysis air temperature and precipitation in high-latitude Asia using ground-based observations. *International Journal of Climatology*, 43, 1621–1638.

Liu, J., Hagan, D.F.T. & Liu, Y. (2021) Global land surface temperature change (2003–2017) and its relationship with climate drivers: airs, Modis, and ERA5-land based analysis. *Remote Sensing*, 13, 44.

Liu, Z., Liu, Y., Wang, S., Yang, X., Wang, L., Baig, M.H.A. et al. (2018) Evaluation of spatial and temporal performances of ERA-interim precipitation and temperature in mainland China. *Journal of Climate*, 31, 4347–4365.

Lobanov, V.A. & Kirillina, K.S. (2019) *The modern and future climate changes in the republic of Sakha (Yakutia)*. St. Petersburg: RSHU.

Matveeva, T. & Sidorchuk, A. (2020) Modelling of surface runoff on the Yamal peninsula, Russia, using ERA5 reanalysis. *Water*, 12, 2099.

Minnett, P.J., Alvera-Azcárate, A., Chin, T.M., Corlett, G.K., Gentemann, C.L., Karagali, I. et al. (2019) Half a century of satellite remote sensing of sea-surface temperature. *Remote Sensing of Environment*, 233, 111366.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G. et al. (2021) ERA5-land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13, 4349–4383.

NCAR. (2023) *ERA-Interim [Online]*. UCAR. Available from: https://climatedataguide.ucar.edu/climate-data/era-interim [Accessed 31 May 2023 2023]

Noël, B., Kampenhout L, V., Van de Berg, W.J., Lenaerts, J.T.M., Wouters, B. & Van Den Broeke, M.R. (2020) Brief communication: CESM2 climate forcing (1950–2014) yields realistic Greenland ice sheet surface mass balance. *The Cryosphere*, 14, 1425–1435.

Olsson, R. (2009) *Boreal Forest and climate change*. Goteborg: Air Pollution & Climate Secretariat & Taiga Rescue Network.

Ramon, J., Lledó, L., Torralba, V., Soret, A. & Doblas-Reyes, F.J. (2019) What global reanalysis best represents near-surface winds? *Quarterly Journal of the Royal Meteorological Society*, 145, 3236–3251.

Riazanova, A.A., Voropay, N.N., Okladnikov, I.G. & Gordov, E.P. (2016) Development of computational module of regional aridity for web-GIS “climate”. *IOP Conference Series: Earth and Environmental Science*, 48, 12032.

Shabanov, N., Marshall, G.J., Rees, W.G., Bartalev, S.A., Tutubalina, O. & Golubeva, E. (2021) Climate-driven phenological changes in the Russian Arctic derived from MODIS LAI time series 2000-2019. *Environmental Research Letters*, 16, 1–14.

Simmons, A., Uppala, S., Dee, D. & Kobayashi, S. 2007. ERA-interim: new ECMWF reanalysis products from 1989 onwards. Newsletter number 110-winter 2006/07, pp. 25–35.

Tarek, M., Brissette, F. & Arsenault, R. (2019) Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modeling over North-America. *Hydrology and Earth System Sciences Discussions*, 24, 2527–2544.

Urraca, R. & Gobron, N. (2023) Temporal stability of long-term satellite and reanalysis products to monitor snow cover trends. *The Cryosphere*, 17, 1023–1052.

Voropay, N., Ryazanova, A. & Dyukarev, E. (2021) High-resolution bias-corrected precipitation data over South Siberia, Russia. *Atmospheric Research*, 254, 105528.

Wang, Y.-R., Hessen, D.O., Samset, B.H. & Stordal, F. (2022) Evaluating global and regional land warming trends in the past decades with both MODIS and ERA5-land land surface temperature data. *Remote Sensing of Environment*, 280, 113181.

Yilmaz, M. (2023) Accuracy assessment of temperature trends from ERA5 and ERA5-land. *Science of the Total Environment*, 856, 159182.

Zyryanova, O.A., Milyutin, L.I., Muratova, E.N., Ryzhkova, V.A., Larionova, A.Y., Sedelnikova, T.S. et al. (2008) Boreal forests of Siberia: genetic, species and ecosystem diversity. *Contemporary Problems of Ecology*, 1, 22–28.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Clelland, A. A., Marshall, G. J., & Baxter, R. (2024). Evaluating the performance of key ERA-Interim, ERA5 and ERA5-Land climate variables across Siberia. *International Journal of Climatology*, *44*(7), 2318–2342. https://doi.org/10.1002/joc.8456

APPENDIX A

A.1 | List of stations used in the validation

TABLE A1 List of meteorological stations used in the validation in this report with their World Meteorological Organization (WMO) number, latitude and longitude and the years for which data are available.

Area	WMO No.	Station name	N latitude (°N)	E longitude (°E)	Data availability
1	23678	Verkhneimbatsk	63.15	87.95	1950–2020
1	23884	Bor	61.60	90.02	1950–2020
1	23986	Severo-Yeniseysk	60.37	93.02	1950–2020
2	24329	Selagoncy	66.25	114.28	1950–2020
2	24538	Chuchukan	64.23	116.92	1955–2020
2	24639	Nyurba	63.28	118.33	1950–2020
2	24641	Vilyuysk	63.77	121.62	1950–2020
2	24726	Mirny	62.53	113.87	1959–2020
2	24738	Suntar	62.15	117.65	1950–2020
3	24763	Lippelyakh	62.82	134.43	1966–2020
3	24768	Curapca	62.03	132.60	1950–2020
3	24966	Ust-Maya	60.38	134.45	1950–2020
3	24967	Tegyultya	60.17	130.20	1950–2020
3	31062	Yugorenok	59.77	137.72	1950–2020
4	30493	Nagorniy	55.97	124.88	1950–2020
4	30692	Skovorodino	54.00	123.97	1950–2020
4	31102	Kankunskiy	57.65	125.97	1950–2020
4	31137	Toko	56.28	131.13	1962–2020
4	31253	Bomnak	54.72	128.93	1950–2020
5	29698	Nizhneudinsk	54.88	99.03	1950–2020
5	29789	Verkhnyaya Gutara	54.22	96.97	1950–2020
5	30504	Tulun	54.60	100.63	1950–2020
5	36103	Toora-Khem	52.47	96.12	1950–2020
6	24266	Verkhoyansk	67.55	133.38	1950–2020
6	24371	Ust-Charky	66.80	136.68	1966–2020
7	25206	Srednekolymsk	67.45	153.72	1950–2020
7	25325	Ust-Oloy	66.55	159.42	1950–2012
7	25400	Zyryanka	65.73	150.90	1950–2020
7	25503	Korkodon	64.75	153.97	1950–2019

A.2 | Variable conversions

Raw reanalysis data were converted to draw direct comparisons with meteorological station data and the conversion formulae are given below.

Temperature $\left({}^{\circ}\mathrm{C}\right) = {}^{\circ}K - 273.15.$
Total precipitation $(\mathrm{mm}) = \frac{m}{1000}.$
Wind speed $\left(\mathrm{ms}^{-1}\right) = \sqrt{u^2 + v^2}.$
Snow depth $(\mathrm{cm}) = \left(\frac{\text{snow depth} * 1000}{\text{snow density}}\right) * 100.$