RESEARCH ARTICLE

The accuracy of climate variability and trends across Arctic Fennoscandia in four reanalyses

Gareth J. Marshall[1] | Sonja Kivinen[2] | Kirsti Jylhä[3] | Rebecca M. Vignols[1,4] | W. G. Rees[4]

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

[2]Department of Geographical and Historical Studies, University of Eastern Finland, Joensuu, Finland

[3]Weather and Climate Change Impact Research, Finnish Meteorological Institute, Helsinki, Finland

[4]Scott Polar Research Institute, University of Cambridge, Cambridge, UK

Correspondence
Gareth J. Marshall, British Antarctic Survey, Natural Environment Research Council, Cambridge, UK.
Email: GJMA@bas.ac.uk

Arctic Fennoscandia has undergone significant climate change over recent decades. Reanalysis data sets allow us to understand the atmospheric processes driving such changes. Here we evaluate four reanalyses against observations of near-surface air temperature (SAT) and precipitation (PPN) from 35 meteorological stations across the region for the 35-year period from 1979 to 2013. The reanalyses compared are the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR), the European Centre for Medium-Range Weather Forecast (ECMWF) Interim reanalysis (ERA-Interim), the Japanese Meteorological Agency (JMA) 55-year reanalysis (JRA-55) and National Aeronautics and Space Administration (NASA)'s Modern-Era Retrospective Analysis for Research and Applications (MERRA).

All four reanalyses have an overall small cool bias across Arctic Fennoscandia, with MERRA typically ~1 °C cooler than the others. They generally reproduce the broad spatial patterns of mean SAT across the region, although less well in areas of complex orography. Observations reveal a statistically significant warming across Arctic Fennoscandia, with the majority of trends significant at $p < .01$. Three reanalyses show similar regional warming but of smaller magnitude while CFSR is anomalous, even having a slight cooling in some areas. In general, the other reanalyses are sufficiently accurate to reproduce the varying significance of observed seasonal trends.

There are much greater differences between the reanalyses when comparing PPN to observations. MERRA-Land, which merges a gauge-based data set, is clearly the best: CFSR is the least successful, with a significant wet bias. The smoothed reanalysis orography means that the high PPN associated with the western side of the Scandinavian Mountains extends too far inland. Spatial patterns of PPN trends across the region differ markedly between the reanalyses, which have varying success at matching observations and generally fail to replicate sites with significant observed trends. Therefore, using reanalyses to analyse PPN changes in Arctic Fennoscandia should be undertaken with caution.

KEYWORDS
Arctic, climate, climate change, Fennoscandia, precipitation, reanalysis, temperature

1 | INTRODUCTION

Recent increases in near-surface Arctic temperatures are significantly greater than the global average (e.g., Bindoff *et al.,* 2013) and, since 1980, are more than twice that of

the Northern Hemisphere average (Overland *et al.,* 2016). This disproportionate warming relative to lower latitudes, termed Arctic amplification, results from several distinct radiative and advective processes (Cohen *et al.,* 2014, and references therein) and is outside the range of expected

internal climate variability: it is thus thought to be at least partially driven by anthropogenic forcing (e.g., Gillett *et al.,* 2008; Bindoff *et al.,* 2013; Chylek *et al.,* 2014). Temperatures have risen the most over the Arctic Ocean, where there has been a significant and well publicized decrease in sea ice, especially in the Barents Sea (e.g., Årthun *et al.,* 2012; Matishov *et al.,* 2012). This loss has been linked to recent cooler Eurasian winters through enhanced autumn snow cover and resultant large-scale circulation changes (e.g., Cohen *et al.,* 2012; Tang *et al.,* 2013; Cohen *et al.,* 2014; Mori *et al.,* 2014; Overland *et al.,* 2016), and slightly greater Eurasian winter climate variability in general (Li *et al.,* 2015). However, Overland *et al.* (2015) noted that northern Europe itself is outside the area affected directly by broad-scale Arctic changes: multiple factors, including jet stream position and internal climate "noise," drive the winter weather variability in this region.

Arctic Fennoscandia, which comprises northern areas of Norway, Sweden, Finland and the Kola Peninsula region of Russia, is located immediately south of the Barents Sea and has seen some marked changes in climate over recent decades (e.g., Førland *et al.,* 2009; Irannezhad *et al.,* 2015; Aalto *et al.,* 2016; Marshall *et al.,* 2016; Irannezhad *et al.,* 2017; Kivinen *et al.,* 2017; Swedish Meteorological and Hydrological Institute, 2017). Regional temperatures have increased, in particular during the 21st century. Although similarly high temperatures were observed in the 1930s, they have been more consistently warm in recent decades. Temperature rises of ~0.3–0.5 °C/decade between 1961 and 2010/2011 have been reported in northern Finland (Irannezhad *et al.,* 2015; Aalto *et al.,* 2016) while Marshall *et al.* (2016) described a mean warming of ~2.3 °C from 10 stations in the Kola Peninsula from 1966–2015 (0.46 °C/decade). Seasonally, both these studies indicated that spring was the period with the greatest, statistically significant warming over the past 50 years. The rate of winter warming in northern Finland was actually greater but the previously mentioned enhanced variability in this season means that such trends are not significant. Precipitation changes appear to be more variable across Fennoscandia. For example, Førland *et al.* (2009) and Aalto *et al.* (2016) demonstrated annual precipitation increases across much of northern Norway and Finland, respectively, for 1961–2010, although the rate of change and its statistical significance appear uncertain. However, Marshall *et al.* (2016) found no such trend in the Kola Peninsula during 1966–2015, but did describe significant wetting and drying trends in spring and autumn, respectively. Changes in the spatial and temporal distribution of heat and moisture, such as the frequency of temperatures close to zero and strong precipitation events, have the potential to trigger natural hazards in Fennoscandia such as avalanches, rock slides or flooding (e.g., Dyrrdal *et al.,* 2012). In addition, these climatic changes have led to major behavioural changes in the region's vegetation, including a

general lengthening of the growing season and both increases and decreases in plant productivity (e.g., Høgda *et al.,* 2013; Barichi*v*ich *et al.,* 2014; Bjerke *et al.,* 2014; Blinova and Chmielewski, 2015).

In order to understand the drivers of such observed changes, there is a need to link them to physical processes throughout the atmosphere. Atmospheric reanalyses provide an excellent tool for such studies. Reanalyses use a numerical weather prediction forecast model to assimilate an historical archive of meteorological data, derived from ground-based observations, radiosondes and satellite data. Their output is a uniform multivariate record of the atmospheric circulation and hydrological cycle including surface parameters. In terms of analysing climate change, reanalyses have the advantage over operational forecast systems of using a single version of both model and assimilation scheme, thus removing spurious changes due to alteration in the model physics and assimilation methodology. Nevertheless, their temporal coherence will be affected by changes to, and any time varying biases in, the observing system. Unfortunately, such impacts are manifested most strongly in regions where ground-based observations are sparsest and thus where reanalyses are most useful, such as the polar regions (e.g., Bromwich *et al.,* 2007; Rapaić *et al.,* 2015). Reanalyses are also being increasingly used to provide lateral boundary conditions for high-resolution regional model studies within the Arctic (e.g., Heikkilä *et al.,* 2011; Lenaerts *et al.,* 2013; Bieniek *et al.,* 2016) and for the pan-Arctic regional Arctic System Reanalysis (e.g., Bromwich *et al.,* 2016).

Here, we analyse the accuracy of four modern reanalyses in describing the surface climate of Arctic Fennoscandia: these are the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR), the European Centre for Medium-Range Weather Forecast (ECMWF) Interim reanalysis (ERA-Interim), the Japanese Meteorological Agency (JMA) 55-year reanalysis (JRA-55) and National Aeronautics and Space Administration (NASA)'s Modern-Era Retrospective Analysis for Research and Applications (MERRA). These reanalyses, together with their forerunners, have many thousands of users (Gregow *et al.,* 2016). The four reanalyses analysed here share a number of advances over their predecessors including (a) assimilation of more complete sets of observations, (b) higher spatial and vertical resolutions, (c) better stratospheric dynamics, (d) improved assimilation schemes that includes variational bias correction (VarBC) for satellite radiances and (e) greater emphasis on accurately representing the hydrological cycle. Further details of these reanalyses are given in section 2.2. We study the climatological means, variability and trends for near-surface temperature (hereinafter SAT) and precipitation (hereinafter PPN) from 1979 to 2013. Three of the reanalyses begin in 1979 and

JRA-55 finishes in 2013, thus defining the period available for comparison.

A number of previous studies have compared SAT and/or PPN from one or more of these reanalyses against ground-based observations and/or satellite-derived products in the Arctic. For example, Simmons and Poli (2014) examined Arctic SAT trends from 1979 to 2012 for ERA-Interim, JRA-55 and MERRA. The latter reanalysis has slightly less warming due to a shift in temperature from a warm bias prior to ~2008 to values similar to the other two reanalyses thereafter. Chung et al. (2013) found similar trends between CFSR, ERA-Interim and MERRA for 1979–2011, with the latter again having the smallest. However, both these papers focused on the region from 70°N to 90°N, which only includes the very north of the Fennoscandian region studied here. Lader et al. (2016) undertook a broadly similar study to this one for Alaska and included CFSR, ERA-Interim and MERRA among the reanalyses they examined. They showed that different reanalyses gave the best representation of SAT in different regions of Alaska while CFSR and ERA-Interim had the smallest PPN biases across the state.

Of greater relevance is the study by Lindsay et al. (2014), who analysed seven reanalyses across the whole Arctic for 1981–2010, including CFSR, ERA-Interim and MERRA, but used the older JRA-25 Japanese reanalysis rather than JRA-55. It is worth noting that these authors concluded that CFSR, ERA-Interim and MERRA stood out as being the reanalyses most consistent with independent observations. For the Arctic as a whole, Lindsay et al. (2014) established that all the reanalyses showed a significant positive annual SAT trend while much of the region did not have significant trends in PPN.

In addition, Behrangi et al. (2016) demonstrated that PPN across the Arctic from MERRA and ERA-Interim is broadly similar to each other for 2007–2010, and both compare favourably with the product produced by the CloudSat satellite and reasonably well with gauge-based observations. Finally, Behrangi et al. (2016) compared ERA-Interim to Arctic System Reanalysis data. Their analysis included a comparison of SAT against station observations for the 12-month period from December 2006 to November 2007. A mixture of biases was apparent in ERA-Interim across Arctic Fennoscandia: while cool biases occurred along the northern and western coasts, warm biases, generally less than 0.5 °C but with some values greater than 1.5 °C, arose inland (c.f., their fig. 2b). Equivalent PPN biases, based on only three stations, were positive, between 10 and 25%. (c.f., their fig. 6b). The authors noted that ERA-Interim was generally poorer at representing the surface climate than the Arctic System Reanalysis, which has a higher-resolution terrain and a detailed land-surface description.

Our motivation for this paper is to provide the first detailed validation of reanalyses across the region of Arctic Fennoscandia: while the aforementioned studies analysed the entire Arctic, they were unable to provide the detailed regional information on SAT and PPN we do here. This is important because the region has areas of complex orography, such as the Norwegian fjords and Scandinavian Mountains, while the climate changes markedly from marine to continental from west to east. Thus, replicating the regional SAT and PPN of Arctic Fennoscandia accurately represents a significant challenge for the reanalyses.

The remainder of this paper is set out as follows. In section 2 we describe the observations, reanalyses and statistical methodologies that we use. Results are given in section 3 while in section 4 we summarize and discuss our principal conclusions about the accuracy of the four reanalyses in describing SAT and PPN across Arctic Fennoscandia compared to observations.

2 | DATA AND METHODS

2.1 | Observations

We compare the four reanalyses with observations for the 35-year period of 1979–2013 for SAT and 1980–2013 for PPN. This difference results simply from PPN being calculated as a forecast from daily data in some reanalyses and thus unavailable for the first day of 1979 for those reanalyses that begin in this year. We use monthly SAT and PPN data from 35 stations within the Arctic Fennoscandian region. Their locations are shown in Figure 1 and further details are provided in Table 1. These particular stations were chosen based on a combination of having near-complete time series of data across 1979–2013 and providing a reasonably uniform spatial distribution of data across the region. The proportion of available *monthly* data is shown in Table 1. This reveals that in some cases the percentage of data for one or other of SAT or PPN is far from complete: when availability falls below 95% then these are not included in the comparison for that parameter. Therefore, 32 stations are included in the *monthly* analysis of SAT and 28 for PPN, and are differentiated in Figure 1. Similarly, when calculating the mean, variability and trends for the *annual* data only stations with at least 95% of the values for 1979/1980–2013, as derived from the mean of the 12 monthly values, are included in the relevant figures: 31 stations are shown for SAT and 26 for PPN.

Norwegian SAT and PPN data for 14 stations (1–14 in Figure 1 and Table 1) were obtained from the Norwegian Meteorological Institute via their eKlima website. The Finnish data (for five stations; 15–19) were acquired from two sources. The first is the European Climate Assessment and Dataset (ECA&D, available at www.ecad.eu/download/millenium/millenium.php) (Klein Tank et al., 2002), which was used for data from Inari Ivalo Airport, Rovaniemi Airport and Sodankylä Arctic Research Centre. The second is the Spanish OGIMET website (www.ogimet.com), which
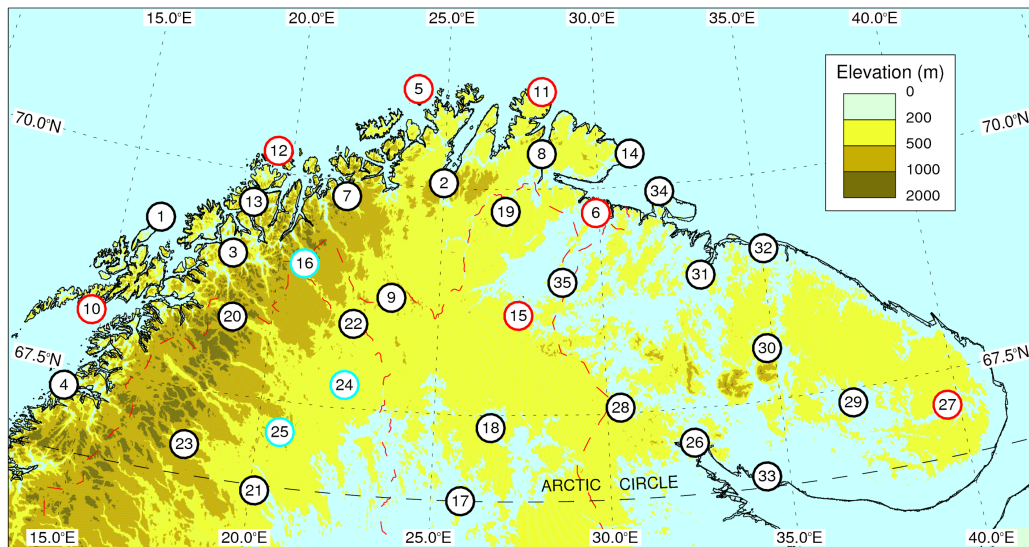
FIGURE 1 Spatial distribution of the 35 meteorological stations used to validate the reanalyses: The numbers are linked to station names in Table 1. Stations with black circles are used for both SAT and PPN, stations with red circles SAT only and cyan circles PPN only. The background shows the regional orography. International borders are shown as a dashed red line [Colour figure can be viewed at wileyonlinelibrary.com]

takes data from the monthly CLIMAT summaries, transmitted via the global telecommunication system (GTS), and makes them available in a convenient format. This provided data for Kilpisjärvi and Utsjoki Kevo. Data from six Swedish stations (20–25) were obtained from the Swedish Meteorological and Hydrological Institute (SMHI) website at http://opendata-download-metobs.smhi.se/explore/. Finally, we utilized data from 10 Russian stations in the Kola Peninsula (26–35). The Russian SAT data were obtained primarily from the Russian Research Institute of Hydrometeorological Information World Data Centre (RIHMI-WDC) as 6-hourly synoptic data. Some additional data were acquired from the UK Met Office Integrated Data System (MIDAS) (Met Office, 2012) and from the Weather Underground website (wunderground.com). These data were quality controlled and gross errors removed. A monthly mean value was produced if at least 95% of the 6-hourly data were available at the standard synoptic hours. To maximize the available monthly data, there were some instances when values were interpolated from adjacent 3-hourly values (see Marshall et al., 2016 for further details). The principal source for Russian PPN data was the ECA&D, with some gaps, particularly in the last few years, filled with data from the RIHMI-WDC. These data have been corrected for biases due to PPN measurement deficiencies and changes in instrumentation (Groisman et al., 2014). In addition, monthly SAT and PPN data for Lovozero prior to 1985 were kindly provided by Dr. Valery Demin. For the Russian station names, we use the anglicized forms, as employed by the World Meteorological Organization (WMO).

The station number used in Figure 1 and Table 1 is given in parentheses after the station name when it is being discussed in the text.

2.2 | Reanalyses

A summary of the four reanalyses analysed in this study—CFSR, ERA-Interim, JRA-55 and MERRA—including key characteristics, details and the online archives from which the reanalysis data were downloaded are given in Table 2. Further pertinent details on each are provided below.

2.2.1 | Climate forecast system reanalysis

In the CFSR reanalysis SAT is not analysed directly and was computed as the monthly mean of 1 hr forecasts taken at the standard synoptic hours (four per day). PPN was derived as the sum of the 6-hourly forecasts from +0 (initialization time) to +6 hr (four per day). Zhang et al. (2012) demonstrated that during this 6-hr cycle the model drifts towards a mean state that is both colder and drier, leading to excess moisture being removed as PPN. These authors noted that this anomalous PPN was enhanced after 1998 following the assimilation of the Advanced TIROS Operational Vertical Sounder (ATOVS) data. The CFSR data are available at a resolution of ~0.34° (38 km at the equator). Beginning on January 1, 2011 the CFSR was extended using the operational Climate Forecast System Version 2 (CFSv2): Saha et al. (2014) described improvements over the original CFSR. Here, the higher spatial resolution of CFSv2 (~0.2°) was downgraded to be consistent with the earlier CFSR data. Across the Northern Hemisphere land as a whole, PPN in CFSv2 is significantly reduced compared to CFSR, especially in summer (Saha et al., 2014). We note that the CFSv2 data more closely matched PPN observations than the earlier CFSR data, although they were generally still the poorest of the four reanalyses analysed here (c.f., section 3.2). For the remainder of this paper the combined CFSR and CFSv2 data will be considered as one single reanalysis and described simply as the CFSR.

TABLE 1 Station locations, elevation and proportion of monthly data available for SAT and PPN for the 1979/1980–2013 period

Station number in Figure 1	Station name	WMO number	Latitude (°N)	Longitude (°E)	Elevation (m a.s.l.)	Proportion of 1979/1980–2013 monthly data available	
						SAT (%)	PPN (%)
1	Andøya	01010	69.31	16.13	13	100.0	99.5
2	Banak	01059	70.06	24.99	8	97.4	96.8
3	Bardufoss	01023	69.06	18.54	77	100.0	99.5
4	Bodø VI	01152	67.27	14.36	13	100.0	100.0
5	Fruholmen Fyr	01055	71.09	23.98	14	100.0	70.8
6	Kirkenes Lufthavn	01089	69.73	29.90	86	99.8	75.0
7	Nordstraum I Kvaenangen	01045	69.84	21.89	6	99.8	100.0
8	Rustefjelbma	01075	70.40	28.19	11	98.3	98.3
9	Šihččajávri	01199	68.76	23.54	384	99.3	99.5
10	Skrova Fyr	01160	68.15	14.65	11	100.0	77.2
11	Slettnes Fyr	01078	71.09	28.22	8	100.0	72.3
12	Torsvag Fyr	01033	70.25	19.50	24	99.3	77.2
13	Tromsø	01026	69.65	18.94	115	100.0	99.8
14	Vardø Radio	01098	70.37	31.10	15	100.0	99.0
15	Inari Ivalo Airport	02807	68.61	27.42	147	100.0	79.4
16	Kilpisjärvi	02801	69.05	20.78	476	73.3	99.5
17	Rovaniemi Airport	02845	66.56	25.82	201	100.0	100.0
18	Sodankylä ARC	02836	67.37	26.63	179	100.0	100.0
19	Utsjoki Kevo	02805	69.76	27.01	101	100.0	99.8
20	Abisko	02022	68.36	18.82	394	100.0	100.0
21	Jokkmokk	02142	66.50	20.13	264	95.2	97.3
22	Karesuando	02080	68.44	22.45	331	100.0	100.0
23	Kvikkjokk–Årrenjarka	02120	66.89	18.02	315	96.9	99.0
24	Lainio	02086	67.76	22.35	317	94.3	100.0
25	Malmberget	02053	67.17	20.64	365	92.4	96.8
26	Kandalaksa	22217	67.13	32.43	26	100.0	100.0
27	Kanevka	22249	67.13	39.67	151	97.9	93.6
28	Kovdor	22204	67.57	30.38	246	98.6	98.5
29	Krasnoscel'e	22235	67.35	37.05	155	100.0	99.8
30	Lovozero	22127	68.08	34.80	162	100.0	100.0
31	Murmansk	22113	68.97	33.05	51	100.0	100.0
32	Teriberka	22028	69.18	35.08	30	99.8	97.3
33	Umba	22324	66.68	34.35	39	100.0	99.8
34	Vaida Guba Bay	22003	69.93	31.98	8	98.1	98.0
35	Yaniskoski	22101	68.97	28.78	101	98.6	95.1

2.2.2 | ERA-Interim

Model resolution is ~0.54° (60 km at the equator). SAT is analysed directly within ERA-Interim whereas PPN is derived as an accumulation from the model forecast. Surface land observations are assimilated but only used in a separate analysis for the near-surface that has little impact on the subsequent background forecast over the Arctic (Simmons and Poli, 2014). Thus, these authors found that any biases did not shift appreciably over time and that such changes are small compared to actual observed temperature changes.

Following on from advice given when using ERA-40 (e.g., Marshall, 2009), to avoid potential errors associated

with model spin-up in this reanalysis each daily PPN comprises a combination of four separate forecasts. The PPN for the first 12 hr of a given day is calculated as the difference between the +12 and +24 hr forecasts from 12Z on the previous day. Similarly, PPN for the second half of the day is the difference between the +12 and +24 hr forecasts starting at 00Z on the day itself. These two values are then added to produce a daily total, which are summed to give a monthly value.

2.2.3 | JMA 55-year reanalysis

Key improvements in JRA-55 compared to the previous JMA reanalysis included the assimilation of some new

TABLE 2 Details of the reanalyses used in the study (left justified)

Reanalysis	Reference	Model horizontal grid	Horizontal grid of downloaded data	Assimilation method	Data link
CFSR	Saha et al. (2010)	T382	Regular 1,152 × 576	3D VAR	https://rda.ucar.edu/datasets/ds093.2/
CFSv2	Saha et al. (2014)	T572	Regular 1,760 × 880	3D VAR	https://rda.ucar.edu/datasets/ds094.2/
ERA-Interim	Dee et al. (2011)	T255	Gaussian N128	4D VAR	http://apps.ecmwf.int/mars-catalogue/
JRA-55	Kobayashi et al. (2015)	T319	Regular 640 × 320	4D VAR	https://rda.ucar.edu/datasets/ds628.1/
MERRA	Rienecker et al. (2011)	0.50 × 0.67°	Regular 540 × 361	3D VAR	https://disc.gsfc.nasa.gov/uui/datasets/MAT1NXSLV_V5.2.0/summary
MERRA-Land	Reichle et al. (2011)	0.50 × 0.67°	Regular 540 × 361	3D VAR	https://disc.gsfc.nasa.gov/uui/datasets/MAT1NXMLD_V5.2.0/summary

observational data sets (Kobayashi et al., 2015): of particular note for this study is the inclusion of Russian snow depths from the RIHMI-WDC. The monthly SAT data are the mean of four daily analyses while the PPN data are the sum of four 6-hr forecasts per day, each being from the initialization time to 6 hr ahead, similar to CFSR. The model resolution is ~0.49° (55 km at the equator).

2.2.4 | Modern-era retrospective analysis for research and applications

For this study we use output from MERRA-Land for PPN. This is an "off-line" rerun of the land model component that has two primary changes from the original MERRA reanalysis (Reichle et al., 2011). These are (a) that the PPN is based on merging the MERRA PPN with gauge-based data and (b) it uses an updated catchment land surface model. Significant improvements over the original MERRA have been detected in various aspects of the hydrological cycle with the assimilation of the gauge observations having the larger impact. Note that as MERRA-Land is a land-only simulation, data are limited to those areas defined as land within this particular reanalysis land-sea mask. The monthly mean SAT (PPN) figures are computed as the mean (sum) of all the hourly means for that month from MERRA (MERRA-Land). The model resolution of both MERRA and MERRA-Land is 0.50° latitude by 0.67° longitude (~75 km at the equator).

2.3 | Methodology

The reanalysis data are interpolated to the station location to the nearest 0.1° latitude/longitude. Trends are calculated using standard least squares methodology with the effects of autocorrelation accounted for when calculating the significance assuming an autoregressive first-order process (e.g., Santer et al., 2000). Correlations are derived from the residuals of detrended data, assuming no a priori link between the two data sets. Four statistical parameters are calculated when validating the reanalyses against station observations. These are the mean difference, the root-mean-square error (RMSE), the ratio of the standard deviations and the correlation. The ratio of the standard deviations

(hereinafter σ ratio) is defined as $\sigma_{\text{reanalysis}}/\sigma_{\text{observations}}$. The four seasons are defined as spring (March–April–May), summer (June–July–August), autumn (September–October–November) and winter (December–January–February).

3 | RESULTS

The results of our analysis are presented in three ways. First, we compare the reanalyses in terms of the four types of error statistics calculated as mean values of all the stations for all months: these data are provided as summary tables. Second, we show plots of the annual mean, standard deviation and trends for each reanalysis across Arctic Fennoscandia, with the equivalent observations superimposed. Third, we choose four stations, which represent different climatic regions of Fennoscandia and have complete or near-complete time series of both SAT and PPN, to examine in greater detail how well the reanalyses reproduce the varying climates in terms of their mean, variability and change, both annually and across the four seasons. The stations chosen are Andøya (1), Sodankylä Arctic Research Centre (18), Kvikkjokk-Årrenjarka (23) and Krasnoscel'e (29) and their locations and climate are described in detail in Appendix S1, Supporting Information.

3.1 | Surface air temperature

Table 3 reveals that all four reanalyses have a similar, slightly cool, mean SAT bias across all months and stations, ranging from −0.10 (JRA-55) to −0.22 °C (ERA-Interim), the latter result perhaps being surprising given that ERA-Interim directly assimilates the SAT observations. However, this reanalysis has the lowest RMSE of 1.17 °C, indicating

TABLE 3 Summary of reanalysis SAT validation statistics across all months during 1979–2013 for 32 stations

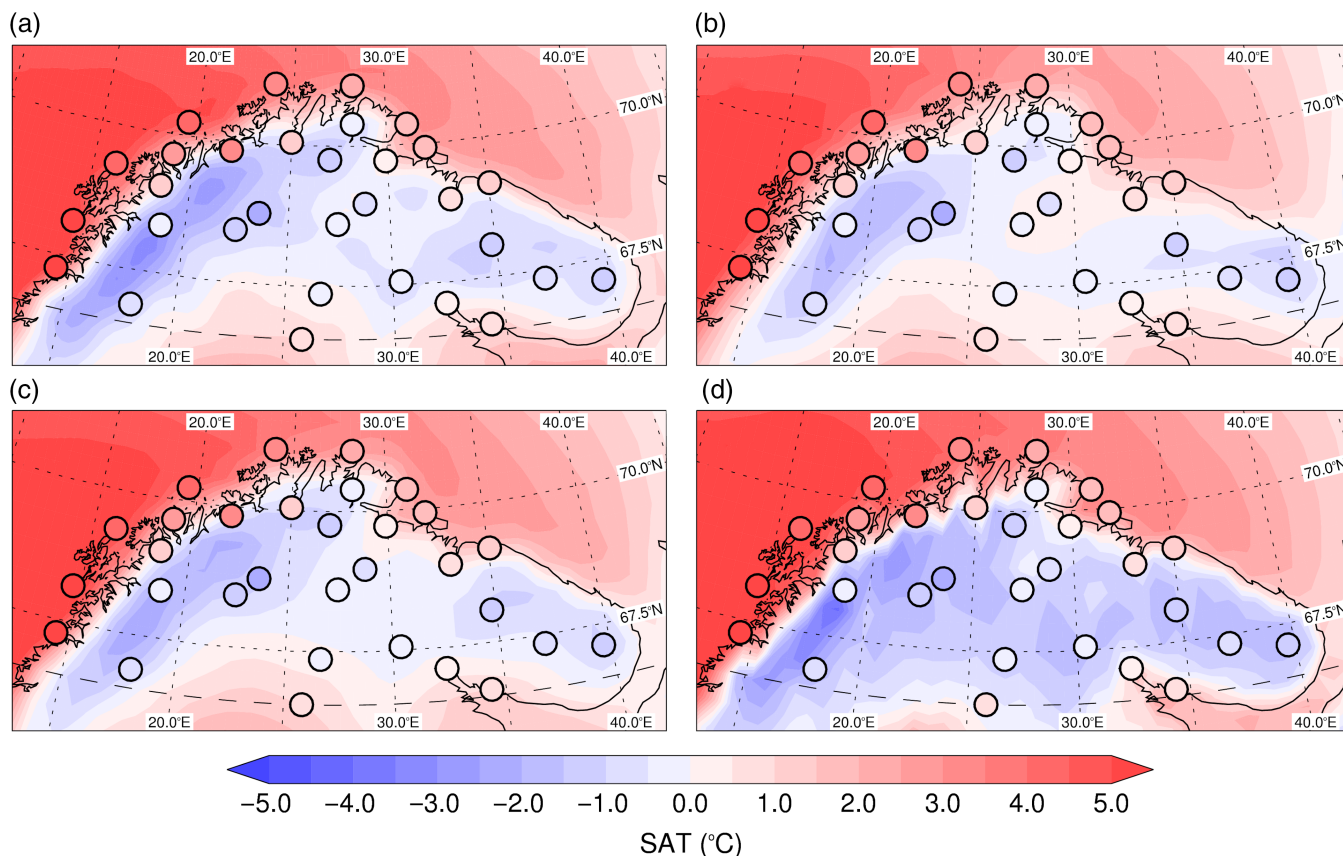|	CFSR	ERA-Interim	JRA-55	MERRA
ΔSAT (°C)	−0.19	−0.22	−0.10	−0.18
RMSE (°C)	1.64	1.17	1.33	1.45
σ ratio	0.9087	1.0087	0.9435	0.9673
Correlation	0.9908	0.9970	0.9959	0.9924

FIGURE 2 Mean annual SAT for 1979–2013 in (a) CFSR, (b) ERA-Interim, (c) JRA-55, and (d) MERRA. Circles represent equivalent station observations [Colour figure can be viewed at wileyonlinelibrary.com]

that on average the magnitude of any bias is likely to be smaller than the other reanalyses. Moreover, ERA-Interim also has the σ ratio closest to unity (1.0087) and the highest correlation value (0.9970) of the four reanalyses, suggesting that overall it best represents SAT across the Arctic Fennoscandian region. By the same measures, CFSR is the least successful at reproducing the SAT.

Mean annual temperatures from the four reanalyses are shown in Figure 2 together with equivalent station observations from stations with 95% available data in the annual time series. In general, the reanalyses pick up the much warmer annual temperatures in the fjord region of northern Norway but do less well at stations located at the southern end of the northern fjords (e.g., Banak (2)), where they are several degrees too cold, likely because there is significant spatial SAT variability due to the steep and complex orography (Eilertsen and Skardhamar, 2006). The cooler temperatures of the Scandinavian Mountains are clearly seen in the reanalyses although there are some common issues at individual stations: for example the alpine site of Abisko (20), a region with a highly complex microclimate (Yang *et al.*, 2012), is significantly too cold in all four reanalyses whereas Šihččajávri (9) on the lee side is too warm. Elsewhere, inland across the region, Figure 2 indicates that SAT is typically 1 °C cooler in MERRA than the other reanalyses and observations, especially in Finland and the

western Kola Peninsula. For example, at Kovdor (28) the bias in MERRA is −1.47 °C compared with −0.05 °C in ERA-Interim. The other three reanalyses also better reproduce the region of warmer SATs in the south of the study region, which includes Rovaniemi Airport (17). We also note that there is evidence to support the positive SAT bias that ERA-Interim has across northwest Fennoscandia shown in Bromwich *et al.* (2016, fig. 2b).

As a measure of inter-annual SAT variability, Figure 3 shows maps of the standard deviation of mean annual SAT from the reanalyses and observations. The reanalyses all do well in reproducing the low SAT variability in northern coastal Norway. Elsewhere, they tend to underestimate SAT variability, particularly in the inland areas of Norway and at the Swedish and Finnish stations. For example, the standard deviation at Inari Ivalo Airport (15) is 1.15 °C while the reanalysis values range from 0.86 to 0.97 °C. ERA-Interim and, to a lesser extent, JRA-55 do manage to replicate the higher inter-annual SAT variability observed in the Kola Peninsula while the other reanalyses are less successful; for example, at Kanevka (27), the observed standard deviation of mean annual SAT is 1.26 °C, while in ERA-Interim it is 1.16 °C and in CFSR it is 0.94 °C.

In Figure 4 we look at the annual and seasonal skill of the reanalyses at the four stations mentioned previously. For Andøya (1), on the Norwegian coast, all four reanalyses
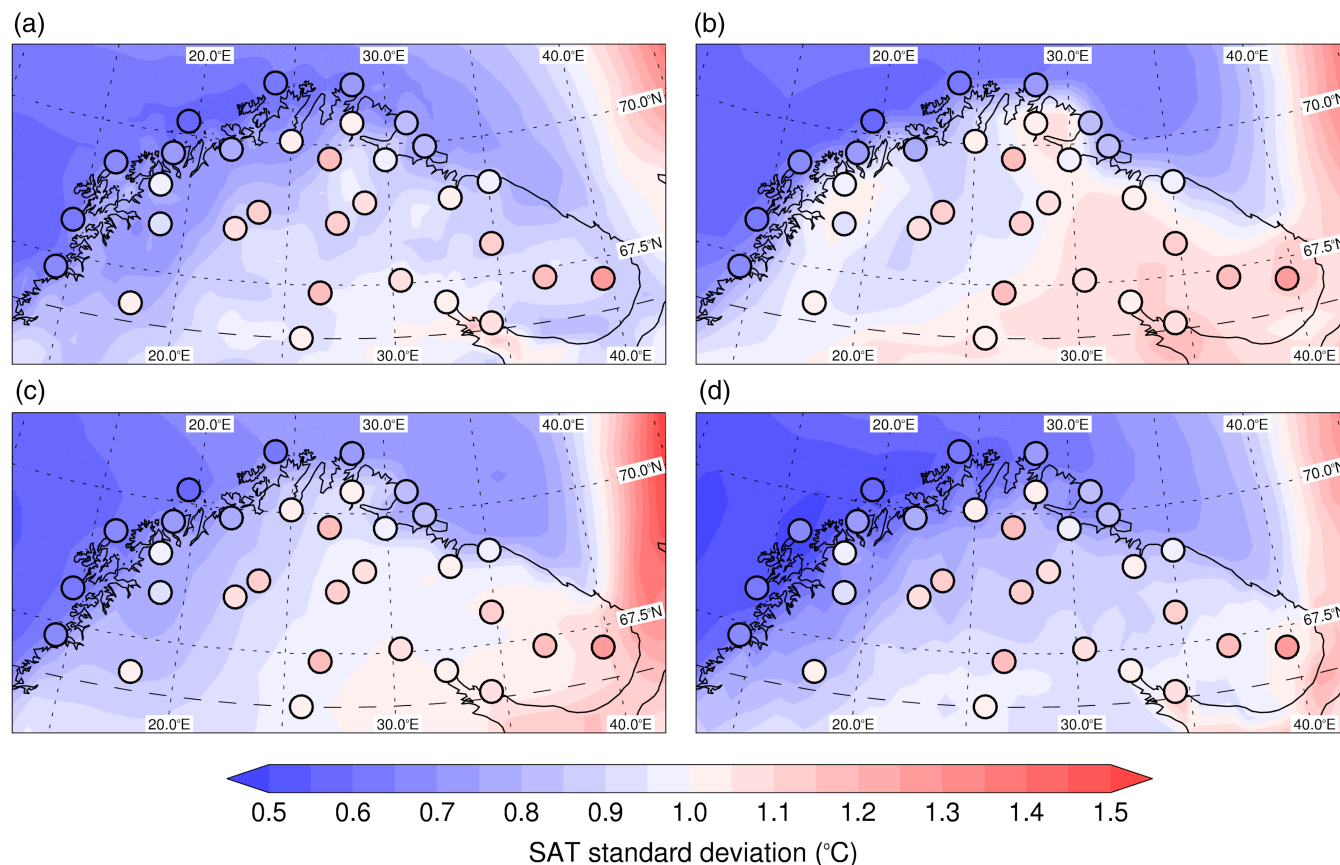
FIGURE 3 As Figure 2, but for mean annual SAT standard deviation [Colour figure can be viewed at wileyonlinelibrary.com]

have a positive SAT bias in all seasons (Figure 4a). This is such that the interquartile range of JRA-55 data does not overlap that of the observations in the annual data and in autumn and winter. The same is true for MERRA in the annual data and all seasons other than summer: in the annual and winter data the mean MERRA SAT is actually higher than any of the observations. However, we note that all the reanalyses correctly show greater SAT variability in autumn and winter and demonstrate a pronounced positive skewness in the SAT distribution of the former season.

At Sodankylä Arctic Research Centre (18) the biases vary between the different reanalyses and across the seasons but are generally small with the interquartile ranges always overlapping those of the observations (Figure 4b). In spring and autumn MERRA is the poorest reanalysis yet it does best at capturing the low winter SATs at this station, which CFSR in particular fails to replicate. All four reanalyses accurately show a much greater range of SATs in this season.

For Kvikkjokk-Årrenjarka (23) the reanalyses are all biased cold across all seasons, other than winter, when they are all biased warm (Figure 4c). We note the markedly incorrect elevation of the station in the smoothed orography of the reanalyses (c.f., Figures 1 and S1): the station altitude in the models is ~650–700 m a.s.l., as compared to the actual height of 315 m a.s.l. Assuming a lapse rate of 4.5 °C/km (Jonsell et al., 2013) an elevation difference of

350 m would cause the reanalyses to be ~1.6 °C too cold. As the actual cold bias varies between 0.6 and 1.4 °C, the marked error in reanalysis elevation at Kvikkjokk–Årrenjarka is likely to be a significant contribution to the annual SAT bias. In spring the cold bias in CFSR and MERRA is sufficiently large for the interquartile ranges not to overlap with that of the observations. In summer this is true for all four reanalyses while in autumn only for MERRA, which actually best reproduces mean winter SAT at Kvikkjokk–Årrenjarka. Similar to Sodankylä Arctic Research Centre, all the reanalyses correctly demonstrate a greater SAT variability in winter but here fail to reproduce the skewed distribution. The seasonal warm bias may be because the reanalyses fail to reproduce near-surface temperature inversions.

At Krasnoscel'e (29) the reanalyses all generally do a good job of replicating the observed SATs throughout the year, broadly similar to Sodankylä Arctic Research Centre, with MERRA having a negative bias while the other reanalyses have an overall positive bias in the annual data (c.f., Figure 2): in some cases the bias changes sign across different seasons (Figure 4d). In winter MERRA again has the smallest overall bias. We note that the results of Lindsay et al. (2014) indicate that MERRA has a negative SAT bias in winter across much of non-coastal Arctic Fennoscandia compared to the median values of seven reanalyses (c.f., their fig. 5). However, Figure 4b–d indicates that compared
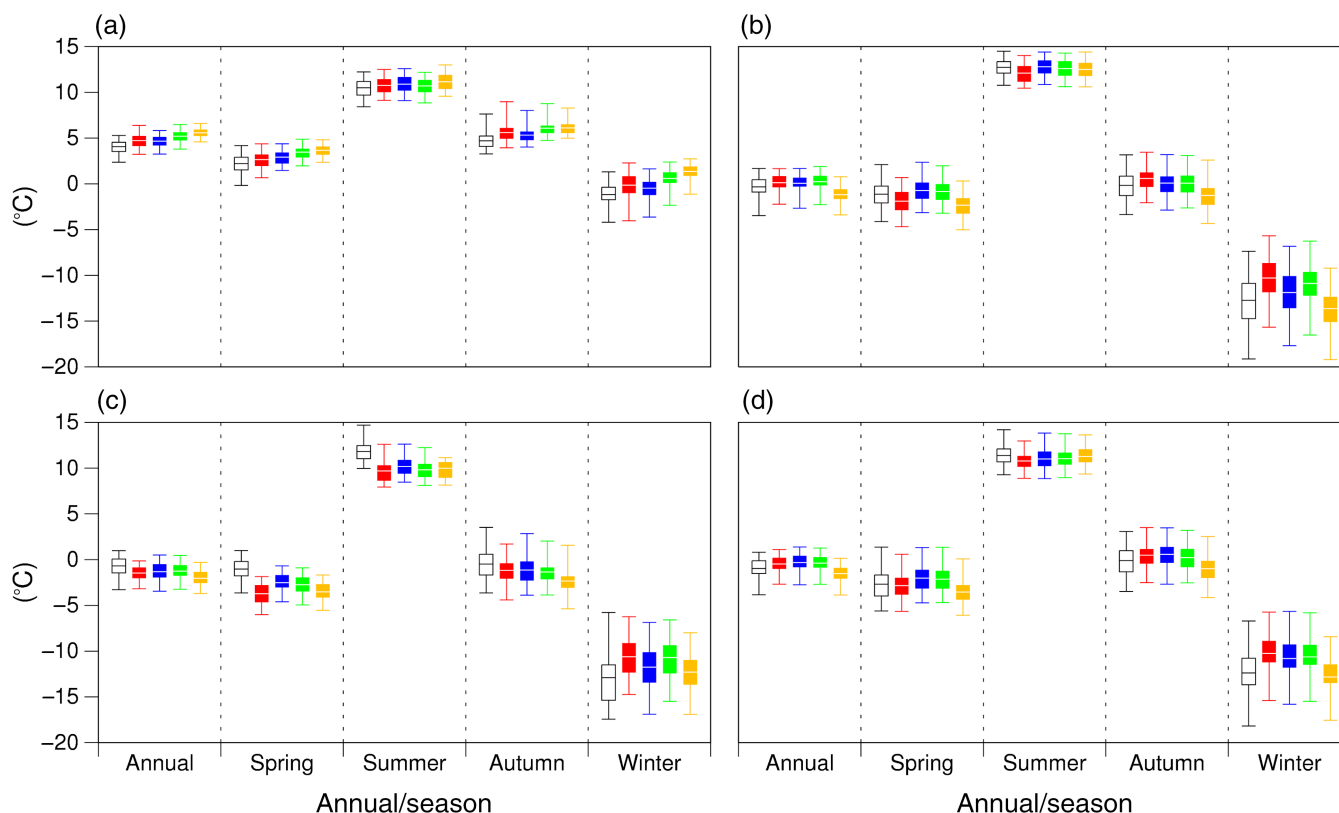
FIGURE 4 Box-whisker plots showing mean annual and seasonal SAT statistics from observations (black), CFSR (red), ERA-Interim (blue), JRA-55 (green) and MERRA (yellow) for (a) Andøya (1), (b) Sodankylä Arctic Research Centre (18), (c) Kvikkjokk–Årrenjarka (23) and (d) Krasnoscel'e (29). The box represents the upper and lower quartiles, with the mean value shown as a horizontal line, and the whiskers the range of data [Colour figure can be viewed at wileyonlinelibrary.com]

against observations, MERRA is actually the best of the four reanalyses examined here for reproducing winter SATs while the other three have a consistent warm bias.

Trends in mean annual SAT across northern Fennoscandia are shown in Figure 5. The observations reveal that a statistically significant warming occurred at all 31 stations in the analysis during the 1979–2013 period, 26 at $p < .01$ and 5 at $p < .05$. The latter are generally located in the west and north of the study region. Three reanalyses also show significant warming across the entire region although there are some spatial differences between them. However, CFSR is anomalous, showing much smaller positive SAT trends in the Scandinavian Mountains, such that there is even a slight negative trend in some parts. Both ERA-Interim and MERRA have areas where the significance of the warming trends is smaller (no longer at $p < .01$). Generally, these do not match spatially with the five stations having similarly less significant trends, although two of these (Bardufoss (3) and Rustefjelbma (8)) are located in the area of reduced significance in MERRA (c.f., Figure 5d).

As mentioned previously, Simmons and Poli (2014) demonstrated that for the Arctic region north of 70°N SAT trends in ERA-Interim and JRA-55 matched each other closely and were greater than in MERRA. There is corroborating evidence for this in Figure 5, with the warming in

that part of Fennoscandia (north of 70°N) in MERRA (~0.2 °C/decade) about half that in the other two reanalyses. The observations suggest that the warming at coastal stations (Fruholmen Fyr (5) and Slettnes Fyr (11)) is better reproduced in MERRA while stations further south have a stronger warming (Banak (2) and Rustefjelbma (8)), which is better simulated by ERA-Interim and JRA-55.

In Figure 6 we compare the observed annual and seasonal SAT trends at Andøya (1), Sodankylä Arctic Research Centre (18), Kvikkjokk-Årrenjarka (23) and Krasnoscel'e (29) from the reanalyses with observations. The reanalyses all correctly demonstrate statistically significant warming at Andøya annually and in autumn (Figure 6a), although the station lies in the region of reduced significance in MERRA in the former period (c.f., Figure 5d). In spring only JRA-55 correctly has a significant warming ($p < .05$) while in summer ERA-Interim erroneously has a significant warming. The greater uncertainty in the winter SAT trends at Andøya are accurately shown by all four reanalyses. For Sodankylä Arctic Research Centre the reanalyses also do well at representing the observed trends (Figure 6b), with all correctly having significant warming annually, in summer (apart from ERA-Interim) and autumn and with no false significant trends in the other seasons. At Kvikkjokk–Årrenjarka ERA-Interim and JRA-55 correctly have a significant warming
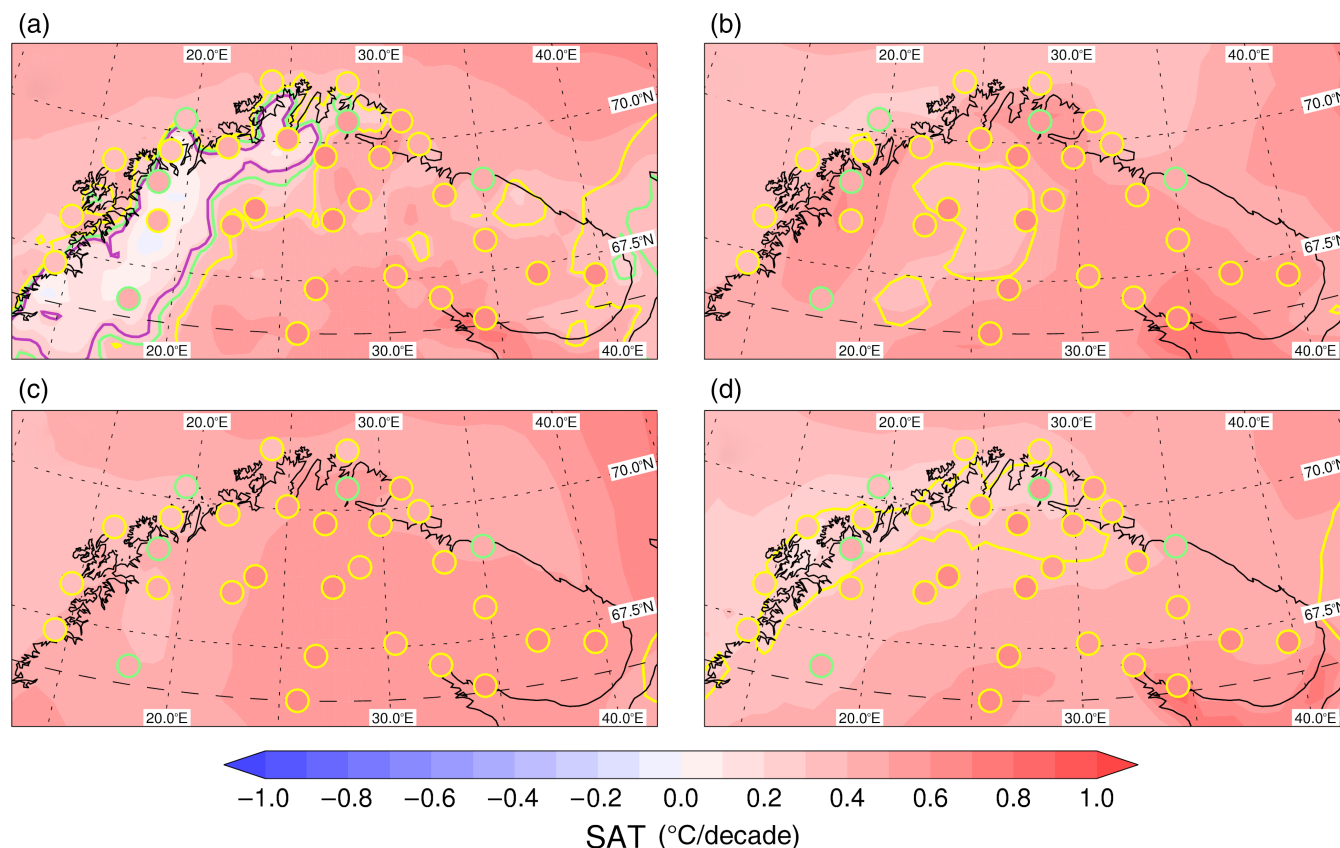
FIGURE 5 As Figure 2, but showing decadal SAT trends for the 1979–2013 period. Purple, green and yellow lines or circles represent significance levels of $p < .10$, $p < .05$ and $p < .01$, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

annually, and in summer and autumn. MERRA fails to reproduce the latter while CFSR has no significant seasonal warming trends and even incorrectly has a slight cooling in spring. Again, all four reanalyses have a large uncertainty in winter SAT trends that matches observations. The reanalyses all do well at reproducing the SAT trends at Krasnoscel'e with warming in the annual data and in summer and autumn all correctly exhibited (Figure 6d). However, both ERA-Interim and JRA-55 also have a significant warming in spring ($p < .10$), which is not apparent in the observations although the mean trends are very similar.

Lindsay *et al.* (2014) showed that all the reanalyses they examined, which included CFSR, ERA-Interim and MERRA, underestimated the annual SAT trend across Arctic Fennoscandia (c.f., their figs 9 and 10). Here, Figures 5 and 6 reveal that while this is indeed generally the case, there are local exceptions such as ERA-Interim having a stronger warming at Kvikkjokk–Årrenjarka (23) compared to observations. Figure 6 indicates that there is no clear seasonal pattern in the trend biases across the region.

3.2 | Precipitation

The skill of the reanalyses at reproducing PPN over northern Fennoscandia is summarized in Table 4. MERRA-Land, which merges a gauge-based data set, is significantly better than the other three reanalyses. The mean bias for all

stations and all months in this reanalysis is only −1.49 mm, whereas the best of the others (ERA-Interim) is 13.61 mm, all three having a positive bias. Unsurprisingly, MERRA-Land also has the lowest RMSE (14.71 mm) and highest correlation of the four reanalyses (0.8959). However, ERA-Interim has the σ ratio closest to unity (1.0244). MERRA-Land is the only reanalysis that has a lower inter-annual PPN variability than observed. CFSR is the least successful at replicating PPN across all four statistics: its marked positive bias is also seen in Lindsay *et al.* (2014, fig. 8).

The mean annual PPN from the four reanalyses and observations is shown in Figure 7. The very high PPN along the western coast of northern Norway extends too far inland and, to a lesser extent, too far north for all the reanalyses other than MERRA-Land. As the high PPN also extends west over the ocean, smoothed model orography, with the Scandinavian Mountains stretched east–west, is highly likely to be the principal contributing factor behind this bias (c.f., Figure S1). This problem is especially pronounced in CFSR with, for example, the mean annual PPN at Abisko (20) being 1,326 mm compared to an observed value of 339 mm. The other reanalyses, including MERRA-Land, also struggle to accurately reproduce the lower PPN at this alpine location, perhaps because of local orographic effects: the Scandinavian Mountain Range immediately west of the site causes leewards conditions that reduce cloudiness and
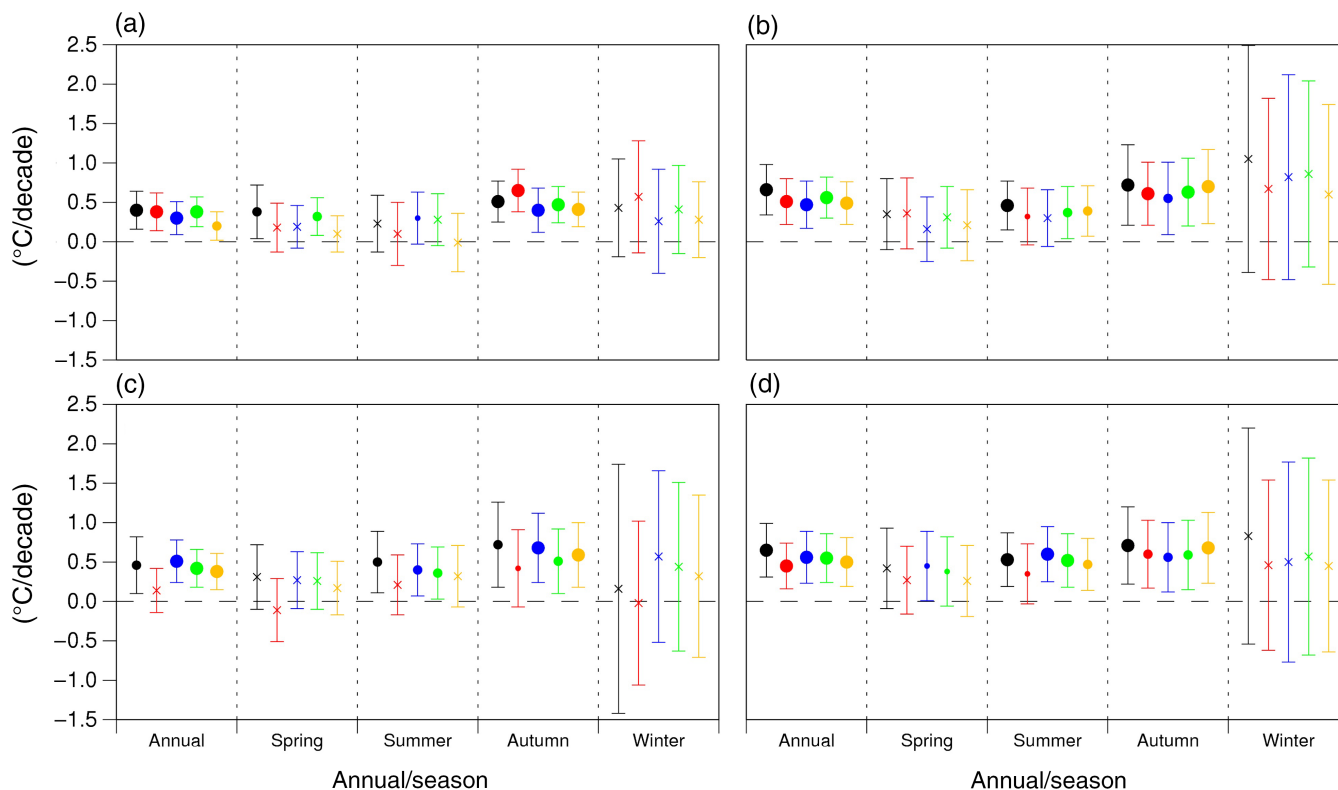
FIGURE 6 As Figure 4, but showing decadal trends in SAT. The whiskers represent the 95% confidence interval of the trend. Statistically significant trends are represented by circles rather than a cross: Small, medium and large circles represent significance levels of $p < .10$, $p < .05$ and $p < .01$, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

PPN (Barry et al., 1981). However, the PPN at other stations in the mountains, such as Bardufoss (3) and Kilpisjärvi (16), are also significantly exaggerated, especially in CFSR. We note that Johansson and Chen (2003) demonstrated that location with respect to these mountains was the single most important factor in determining PPN in Sweden. The higher spatial resolution of CFSR is apparent in it having greater spatial variability in mean PPN than the other reanalyses. Unfortunately, it has significantly too much PPN across almost all of northern Fennoscandia whereas MERRA-Land correctly demonstrates the lower PPN in the north of the region and in the northeast of the Kola Peninsula (Marshall et al., 2016).

The standard deviation of annual PPN is shown in Figure 8. All four reanalyses show a rapid reduction in the standard deviation away from the Norwegian coast. Values at the coastal stations are too high in CFSR and JRA-55 but are well reproduced in ERA-Interim and MERRA-Land. For example at Bodø VI (4), the observed standard deviation of mean annual PPN is 204.3 mm, in JRA-55 it is 287.8 mm while in MERRA-Land it is 197.6 mm. ERA-Interim and JRA-55 do well in replicating PPN variability across most of the rest of northern Fennoscandia, with the two reanalyses showing similar broad spatial patterns in PPN standard deviation. CFSR and MERRA-Land reveal more detail in their spatial patterns but this does not correspond to any increase in accuracy: Table 4 indicates that

these two reanalyses had the poorest σ ratios. For example, there are two areas of higher PPN variability to the south of the region in MERRA-Land (Figure 8d), but observations from Rovaniemi Airport (17) and Jokkmokk (21) suggest the reanalysis is incorrect.

In the annual PPN data of Andøya (1) there is a clear positive bias in CFSR and JRA-55, a negative bias in MERRA-Land with ERA-Interim being the best reanalysis at representing the mean state (Figure 9a). Similar patterns in the average biases are apparent in all four seasons. However, the variability is too large in ERA-Interim, much too small in MERRA-Land and markedly better in the other two reanalyses. The much smaller range of values in spring compared to the other three seasons is correctly shown in all four reanalyses as are the greater interquartile ranges in autumn and winter.

For Sodankylä Arctic Research Centre (18) three reanalyses are relatively poor: the annual PPN values of CFSR, ERA-Interim and JRA-55 are all substantially too large,

TABLE 4 Summary of reanalysis PPN validation statistics across all months during 1980–2013 for 28 stations

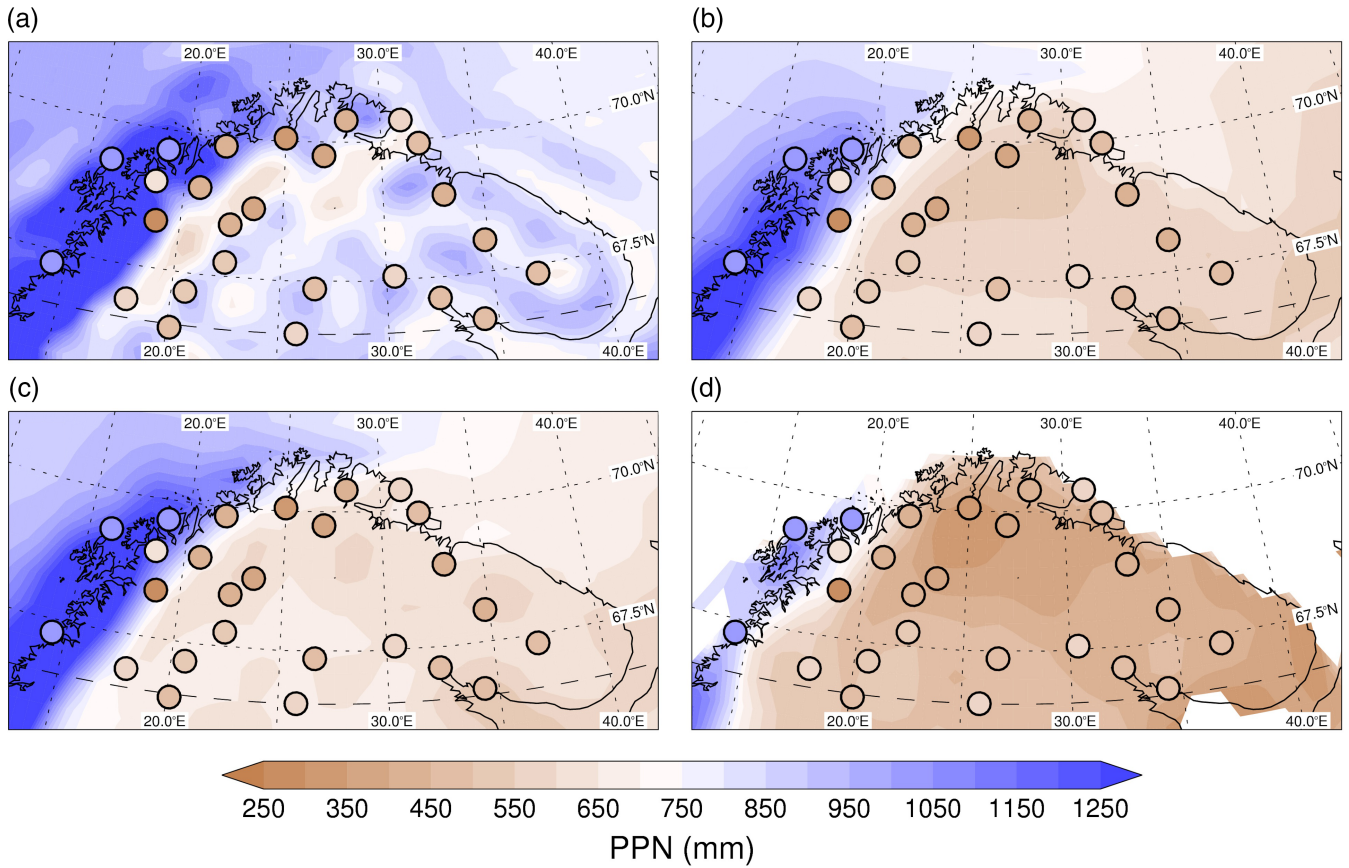|	CFSR	ERA-Interim	JRA-55	MERRA-Land
ΔPPN (mm)	33.04	13.61	19.01	−1.49
RMSE (mm)	39.44	22.61	26.37	14.71
σ ratio	1.1332	1.0244	1.0893	0.9034
Correlation	0.7882	0.8355	0.8423	0.8959

FIGURE 7 As Figure 2, but for PPN from 1980 to 2013 and (d) MERRA-Land [Colour figure can be viewed at wileyonlinelibrary.com]
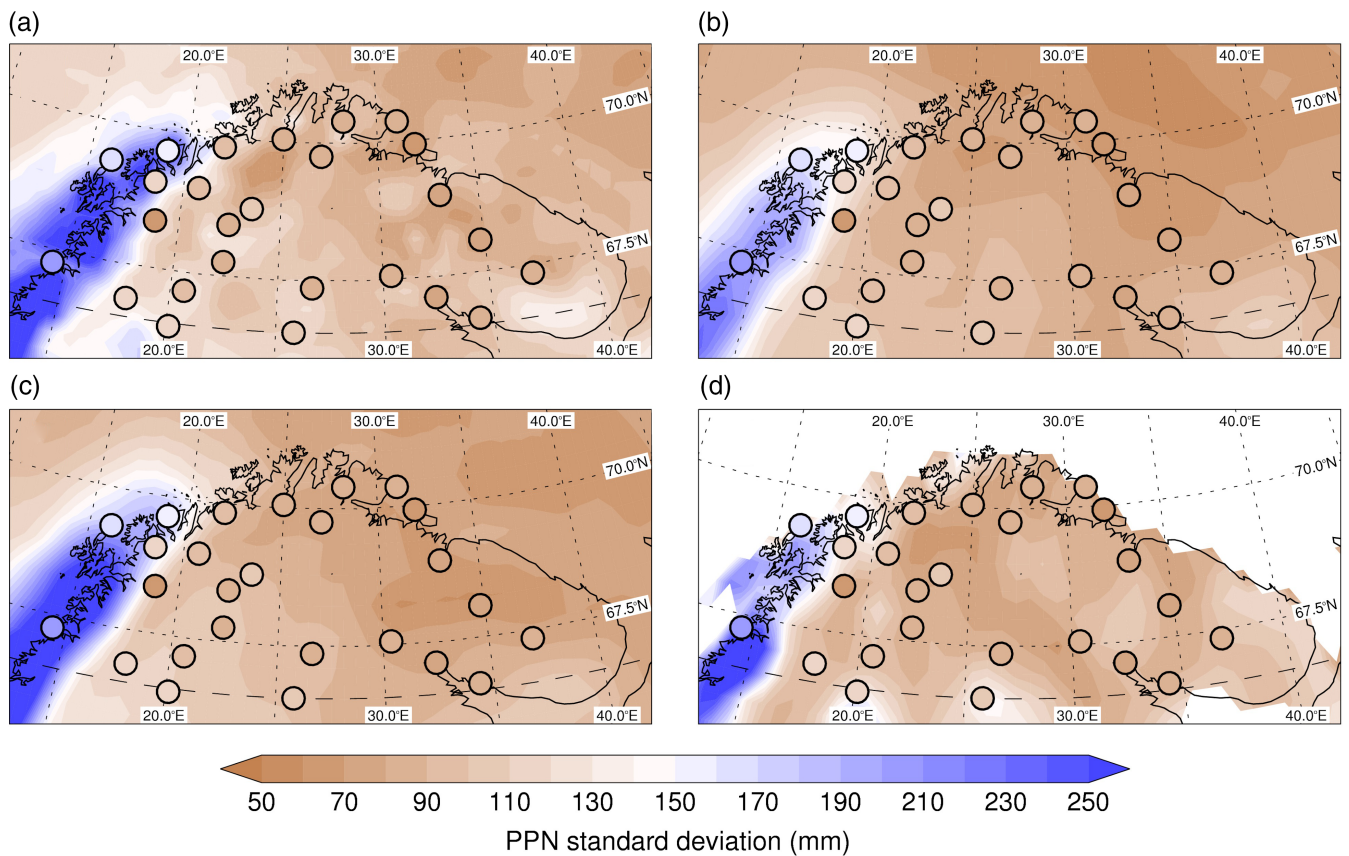

FIGURE 8 As Figure 3, but for PPN from 1980 to 2013 and (d) MERRA-Land [Colour figure can be viewed at wileyonlinelibrary.com]
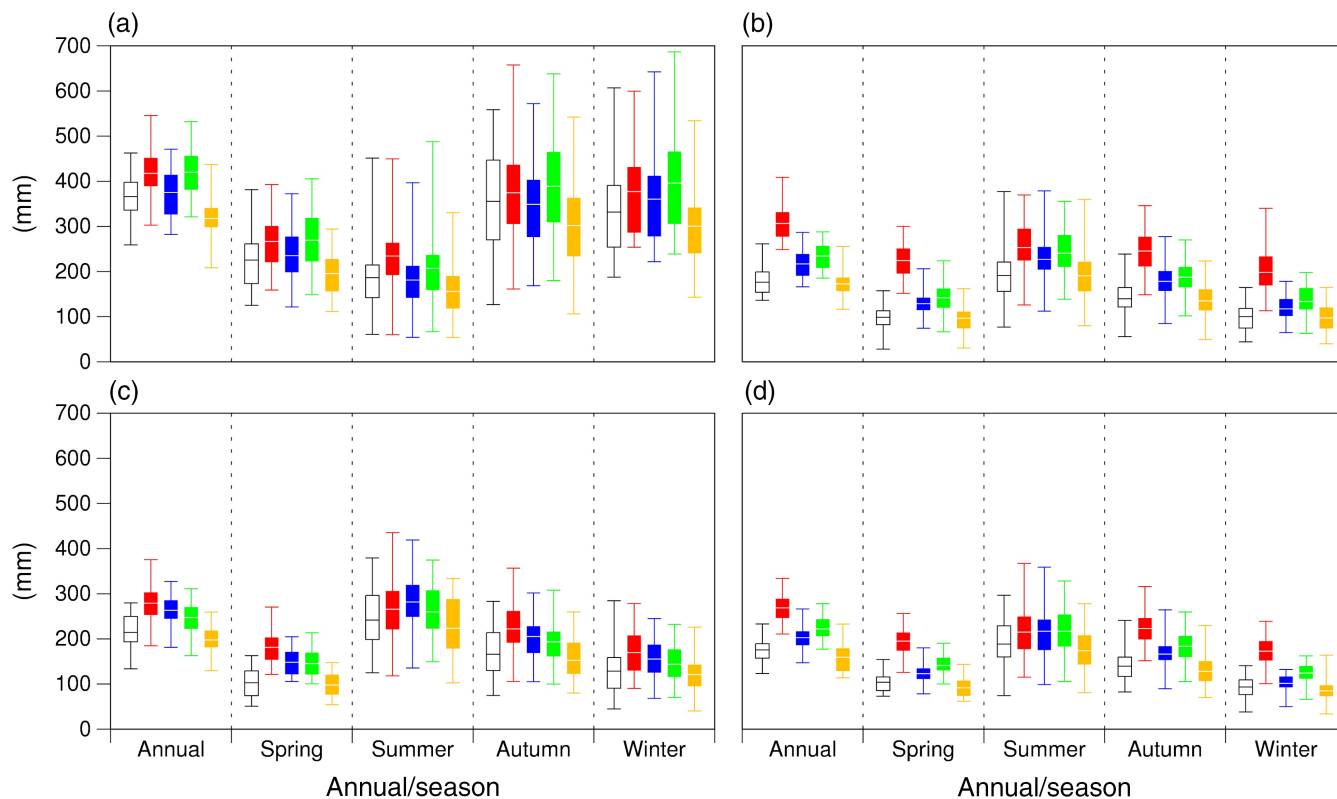
FIGURE 9 As Figure 4, but for PPN. Note that the annual PPN values are three times that shown on the *y*-axis scale [Colour figure can be viewed at wileyonlinelibrary.com]

with CFSR being especially so (Figure 9b). MERRA-Land is the most successful reanalysis and has a small negative bias. Once again, the sign and relative magnitude of the reanalysis biases are broadly comparable across all four seasons and thus are similar to the annual data. There is no overlap of the observations and CFSR interquartile ranges in any season and in spring there is barely any overlap between the two ranges of PPN values. JRA-55 and ERA-Interim show progressive improvements in accuracy, although neither has an overlapping interquartile range with observations in spring. Observations show a peak in variability in summer, which is most apparent in JRA-55 and MERRA-Land.

CFSR is also the least successful reanalysis at reproducing PPN at Kvikkjokk–Årrenjarka, (23) although it does better than at Sodankylä Arctic Research Centre. Mean values are biased high for all the reanalyses apart from MERRA-Land, with CFSR having non-overlapping interquartile ranges with the observations in the annual and spring data (Figure 9c). Although MERRA-Land does best at replicating the average PPN across all four seasons it underestimates the variability while CFSR overestimates it. All four reanalyses correctly show that spring (summer) is the season with least (greatest) PPN variability at Kvikkjokk–Årrenjarka.

The relative PPN biases in the individual reanalyses are broadly similar at Krasnoscel'e (29) to those at Sodankylä Arctic Research Centre (c.f., Figure 9b,d), although the

absolute magnitude of the biases is generally smaller. Of note is that all four reanalyses show the larger PPN variability in summer at this station and all except CFSR correctly indicate that this is the season of maximum PPN, associated with the Arctic frontal zone that appears over northern Eurasia at this time (Marshall *et al.*, 2016).

Trends in annual PPN for 1980–2013 from the reanalyses and observations from 26 stations are shown in Figure 10. There is relatively little agreement in the spatial pattern of PPN trends between the four reanalyses. The CFSR and MERRA-Land reanalyses display much greater variability in trends across the Arctic Fennoscandia region than either ERA-Interim or JRA-55. All the reanalyses do have some areas where the PPN trends are statistically significant (at *p* < .10 or below) but in the majority of cases they do not match the significance of observed trends.

CFSR has areas of significant drying in the Scandinavian Mountains and the eastern part of the Kola Peninsula, which, apart from the drying at Kvikkjokk–Årrenjarka (23), do not even match the sign of the observed trends (Figure 10a). Areas of significant wetting comprise the western and northern coasts of Norway and the southern Kola Peninsula. In the former (latter) this corresponds with the significant observed wetting at Bodø VI (4) (drying at Kandalaksa (26)). We note a similar spatial pattern of positive PPN trends is shown by Lindsay *et al.* (2014, fig. 12). However, the majority of the 11 stations showing a significant PPN trend (all wetting) lie outside the areas of
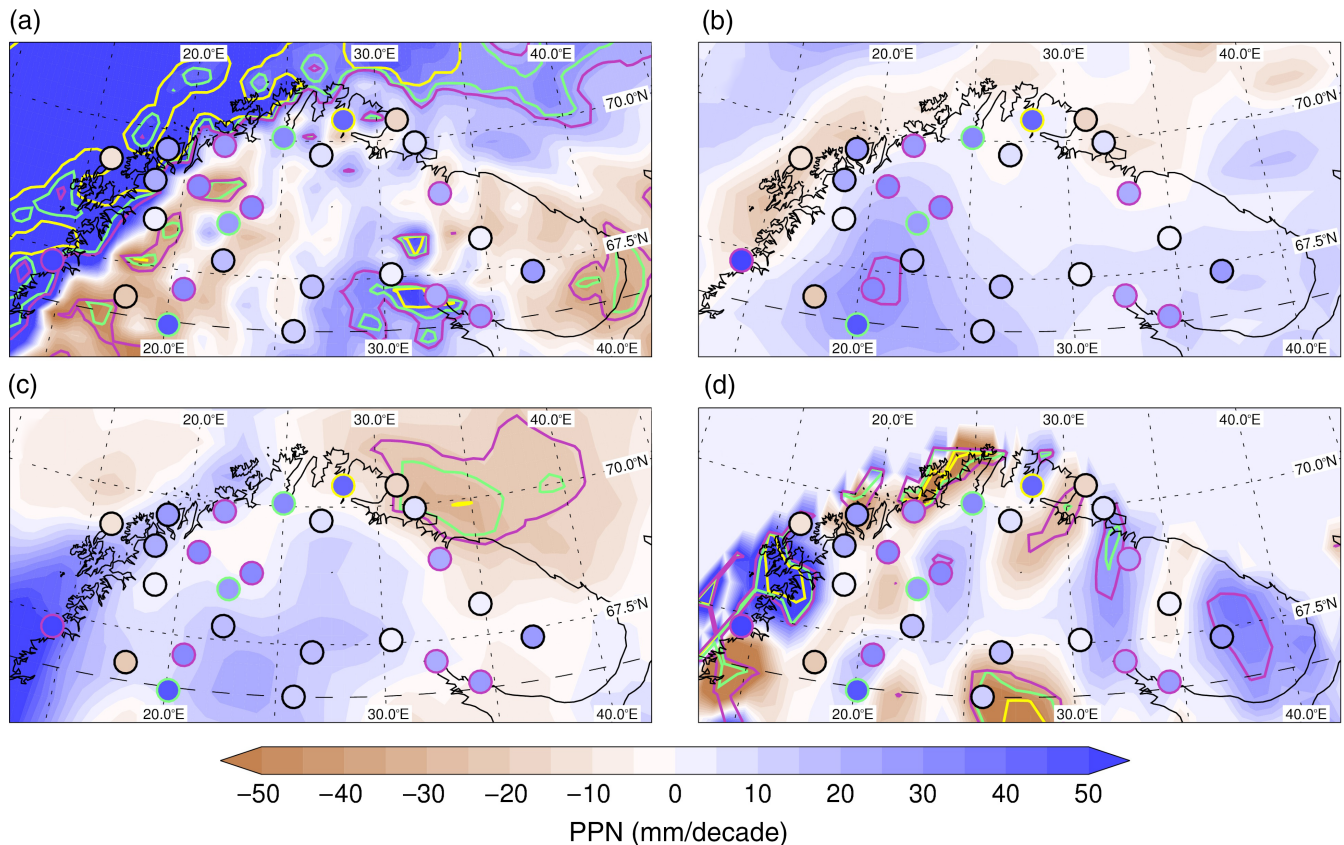
FIGURE 10   As Figure 5, but for PPN for 1980–2013 and (d) MERRA-Land [Colour figure can be viewed at wileyonlinelibrary.com]

significance in CFSR. Apart from some western and northern coastal areas, ERA-Interim shows wetting across the entire study region (Figure 10b). There is only one small area with a significant wetting trend ($p < .10$), which does match correctly the observed trend at Malmberget (25) that is located within it. As the majority of inland stations show a wetting, then ERA-Interim does get the sign of the general regional PPN trend correct. It is also the only reanalysis to reproduce the observed, although statistically insignificant, drying at Andøya (1), discussed below.

The JRA-55 reanalysis also displays a wetting across most of northern Fennoscandia (Figure 10c) although its only region where the trend in PPN is statistically significant is an area of drying along the northern coast at the Norwegian-Russian border that includes non-significant observed trends at Vardø Radio (14) and Vaida Guba Bay (34). This region of drying extends south into the Kola Peninsula but observations suggest it is erroneous. The spatial pattern of PPN trends in MERRA-Land demonstrates longitudinal banding of alternative wetting and drying (Figure 10d). In some cases these correctly match the observations but in others the sign of the trend is opposite. For example, MERRA-Land correctly shows the marked switch from a drying trend at Kvikkjokk–Årrenjarka (23) to the wetting trend at Jokkmokk (21) and Malmberget (25) to the east. It also accurately shows the pattern of PPN trends over the Kola Peninsula, with a north–south band of wetting that includes

Kandalaksa (26) and Murmansk (31), while to the east is a band of weak trends that contains Lovozero (30), and even further to the east there is another area of wetting that includes Krasnoscel'e (29). However, there are areas where MERRA-Land does less well. The reanalysis has a region of significant drying in Finland, to the south of the study region, which includes Rovaniemi Airport (17) where a slight wetting is observed. Moreover, it also indicates an area of drying in the northwest, whereas the stations located in it all have wetting trends, which are actually statistically significant at Nordstraum I Kvaenangen (7) and Kilpisjärvi (16).

In Figure 11 we compare the observed and reanalysis annual and seasonal PPN trends at Andøya (1), Sodankylä Arctic Research Centre (18), Kvikkjokk–Årrenjarka (23) and Krasnoscel'e (29). At Andøya, CFSR has spurious wetting trends in the annual data, which result from similarly erroneous significant trends in spring and autumn with the summer PPN trend also being positive rather than negative (Figure 11a). As mentioned previously, ERA-Interim is the only reanalysis to replicate correctly the observed drying trend at this station. However, the seasonal data indicate this is through luck rather than skill: only in autumn is ERA-Interim the reanalysis that best matches the observed seasonal trend. For Sodankylä Arctic Research Centre there are no statistically significant observed PPN trends and all are close to zero; therefore, it is unsurprising that some of the reanalyses have the wrong
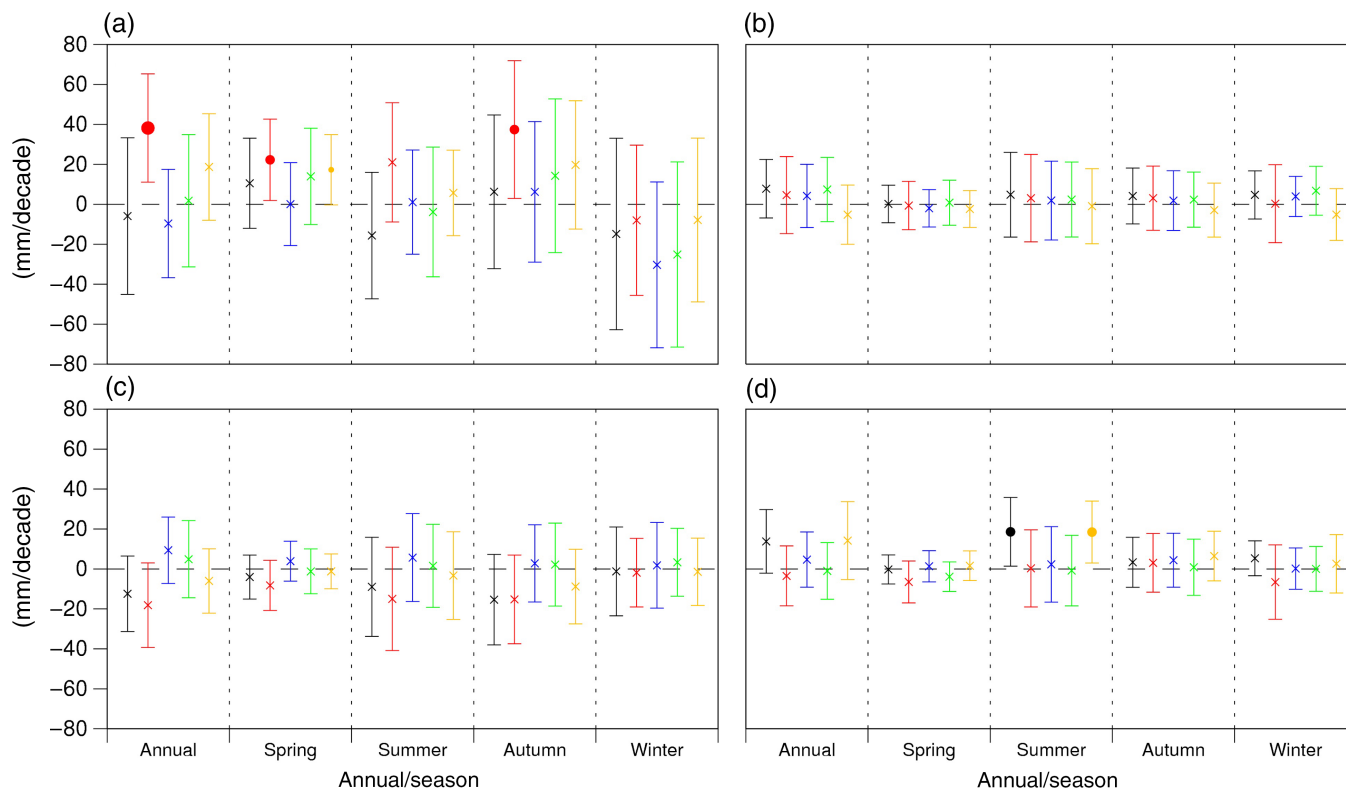
FIGURE 11 As Figure 6, but for PPN. Note that the annual PPN trends are twice that shown on the y-axis scale [Colour figure can be viewed at wileyonlinelibrary.com]

sign of the trend, with MERRA-Land showing a drying in each season (Figure 11b).

Similarly, at Kvikkjokk–Årrenjarka there are no significant annual or seasonal PPN trends in either the observations or any of the reanalyses. Figure 11c reveals that CFSR and MERRA-Land get the correct drying trend across all seasons whereas ERA-Interim and JRA-55 have small wetting trends, except for the latter in spring. The only statistically significant PPN trend at Krasnoscel'e occurs in summer and MERRA-Land is the only reanalysis to reproduce this (Figure 11d). The other seasons reveal that the four reanalyses all have broadly similar PPN trends, apart from CFSR in winter. Thus, it is the fact that MERRA-Land does uniquely well in summer that enables it to have a realistic annual wetting trend.

4 | CONCLUSIONS AND DISCUSSION

In this study we have validated SAT and PPN from four different reanalyses—CFSR, ERA-Interim, JRA-55 and MERRA—against observations from 35 meteorological stations across Arctic Fennoscandia for the 35-year period 1979–2013.

All four reanalyses have an overall small cool bias across this region, with MERRA typically about 1 °C cooler than the other reanalyses (Figure 2), although its validation statistics are not the poorest. Interestingly, the lower SAT in MERRA shown in this study contrasts with its warm bias

north of 70°N prior to 2008 (Simmons and Poli, 2014). The reanalyses also generally do well in reproducing the broad spatial patterns of mean SAT across Arctic Fennoscandia, although, not unexpectedly, they do less well in regions of steep and complex orography, such as the Scandinavian Mountains and inner Norwegian fjords. Clearly, the spatial resolution of modern reanalyses, although considerably improved over previous versions, is still too coarse to accurately represent such areas and the important near-surface processes that govern local SAT, such as temperature inversions and the amplified diurnal cycle in lapse rate, are missing from the reanalysis models. Inter-annual SAT variability in the coastal regions is well reproduced by the four reanalyses but they all tend to underestimate the variability inland (Figure 3).

A seasonal evaluation of SAT at four stations across the region indicates that the magnitude and even the sign of biases in individual reanalyses can vary both between seasons and at different locations. In general, winter biases tend to be larger and usually positive, although this is not always true for MERRA (Figure 4). Observations reveal a statistically significant warming across Arctic Fennoscandia for 1979–2013, with the majority of trends being significant at $p < .01$ (Figure 5). Three reanalyses show an overall similar regional warming pattern but of smaller magnitude, similar to Lindsay *et al.* (2014). However, CFSR is anomalous, having a much smaller warming in the Scandinavian Mountains and even a slight cooling in some areas. Seasonal

analysis indicates that this reduced warming in CFSR occurs across all seasons, but is particularly apparent in spring (Figure 6). The seasonal data also reveal that in general SATs in the other three reanalyses are sufficiently accurate to correctly reproduce the varying statistical significance of seasonal trends.

There are much greater differences between the four reanalyses for mean annual PPN than SAT across Arctic Fennoscandia. MERRA-Land, which merges a gauge-based data set, is distinctly better than the other three and is able to counteract the strong influence of local conditions on PPN (Table 4). It has a very small dry bias whereas the other reanalyses are too wet. CFSR is the least successful reanalysis at replicating the PPN observations, with a significant wet bias (Figure 7), also reported by Lindsay et al. (2014). The smoothed orography in the reanalyses means that the very high PPN associated with the western side of the Scandinavian Mountains extends too far inland in all but MERRA-Land. The wet bias in CFSR occurs across all seasons and is of such magnitude that the interquartile ranges of the observations and this reanalysis do not often overlap (Figure 9).

The spatial pattern of PPN trends across the region differs markedly between the four reanalyses, which demonstrate varying success at matching observations. Most of the latter reveal an increase in PPN from 1980 to 2013 with approximately a third of the stations having a statistically significant wetting. All four reanalyses also have small areas where the PPN trend is significant but in the majority of cases this does not align to the sites with significant observed trends. Despite MERRA-Land having the best validation statistics this does not necessarily translate into the most accurate PPN trends and there are some areas where it does particularly poorly, such as northern Finland (Figure 10). The pronounced divergence between the reanalyses means that any studies utilizing them to assess PPN changes in Arctic Fennoscandia should be undertaken with caution.

In their evaluation of reanalyses over the Canadian Arctic, Rapaić et al. (2015) found that the accuracy of some of the data sets varied considerably over time. For this study we have limited our analysis to 1979–2013, primarily because of the start date of the majority of the reanalyses. However, JRA-55 begins in 1958 so we can examine whether there are any significant temporal changes in the quality of this reanalysis. In Tables S1 and S2 we provide the validation statistics for SAT and PPN, respectively, at the four stations analysed in detail in this study, from four successive 14 year periods in the JRA-55 reanalysis, encompassing 1958–2013. Despite marked changes in the volume of data and the type of data sets assimilated over this period, there is no clear general improvement in the statistics over time. Although there are some examples, such as mean difference and RMSE of SAT at Sodankylä Arctic

Research Centre (18), where an improvement does occur (Table S1), there are also instances where the statistics actually get progressively worse over time, like the mean difference of PPN at Kvikkjokk–Årrenjarka (23) (Table S2). Therefore, we conclude that JRA-55 can be utilized with a similar degree of confidence prior to 1979–2013 as during this period itself.

In this analysis, we have focussed on annual and seasonal data but extreme SAT and PPN events are likely to be of greater importance to planning authorities given they can be potential triggers for natural hazards that may impact socio-economic activity (Dyrrdal et al., 2012). The spatial scale of modern global reanalyses, although considerably improved over previous versions, is still too coarse to accurately represent such events. This is especially the case in regions of steep and complex orography, such as the Scandinavian Mountains, where both SAT and PPN can vary markedly over short distances (e.g., Johansson and Chen, 2003; Yang et al., 2012; Pike et al., 2013; Aalto et al., 2017). The examples when the interquartile ranges of SAT and/or PPN from the gridded reanalysis data and observations fail to even overlap provide clear evidence of the spatial mismatch between the two data sets. Thus, in such regions, global reanalyses are probably most usefully employed in providing boundary conditions for either dynamical downscaling studies, using regional atmospheric models (e.g., Heikkilä et al., 2011; Lenaerts et al., 2013; Bieniek et al., 2016), or higher-resolution regional-scale reanalyses, such as the Arctic System Reanalysis (Bromwich et al., 2016) and the Uncertainties in Ensembles of Regional Reanalyses (UERRA) project that encompasses Europe (www.uerra.eu).

ORCID

Gareth J. Marshall 🄳 http://orcid.org/0000-0001-8887-7314

REFERENCES

Aalto, J., Pirinen, P. and Jylhä, K. (2016) New gridded daily climatology of Finland: permutation-based uncertainty estimates and temporal trends in climate. *Journal of Geophysical Research: Atmospheres*, 121, 3807–3823. https://doi.org/10.1002/2015JD024651.

Aalto, J., Riihimäki, H., Meineri, E., Hylander, K. and Luoto, M. (2017) Revealing topoclimatic heterogeneity using meteorological station data. *International Journal of Climatology*, 37, 544–556. https://doi.org/10.1002/joc.5020.

Årthun, M., Eldevik, T., Smedsrud, L.H., Skagseth, Ø. and Ingvaldsen, R.B. (2012) Quantifying the influence of Atlantic heat on Barents Sea ice variability and retreat. *Journal of Climate*, 25, 4736–4743. https://doi.org/10.1175/JCLI-D-11-00466.1.

Barichivich, J., Briffa, K.R., Myneni, R., van der Schrier, G., Dorigo, W., Tucker, C.J., Osborn, T.J. and Melvin, T.M. (2014) Temperature and snow-mediated moisture controls of summer photosynthetic activity in northern terrestrial ecosystems between 1982 and 2011. *Remote Sensing*, 6, 1390–1431. https://doi.org/10.3390/rs6021390.

Barry, R.G., Courtin, G.M. and Labine, C. (1981) Tundra climates. In: Bliss, L. C., Heal, A.W. and Moore, J. (Eds.) *Tundra Ecosystems: A Comparative Analysis*. Cambridge: Cambridge University Press, pp. 81–114.

Behrangi, A., Christensen, M., Richardson, M., Lebsock, M., Stephens, G., Huffman, G.J., Bolvin, D., Adler, R.F., Gardner, A., Lambrigtsen, B. and Fetzer, E. (2016) Status of high latitude precipitation estimates from observations and reanalyses. *Journal of Geophysical Research: Atmospheres*, 121, 4468–4486. https://doi.org/10.1002/2015JD024546.

Bieniek, P.A., Bhatt, U.S., Walsh, J.E., Rupp, T.S., Zhang, J., Krieger, J.R. and Lader, R. (2016) Dynamical downscaling of ERA-Interim temperature and precipitation for Alaska. *Journal of Applied Meteorology and Climatology*, 55, 635–654. https://doi.org/10.1175/JAMC-D-15-0153.1.

Bindoff, N.L., Stott, P.A., AchutaRao, K.M., Allen, M.R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Mokhov, I.I., Overland, J., Perlwitz, J., Sebbari, R. and Zhang, X. (2013) Detection and attribution of climate change: from global to regional. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge; New York, NY: Cambridge University Press.

Bjerke, J.W., Karlsen, S.R., Høgda, K.A., Maines, E., Jepsen, J.U., Lovibond, S., Vikhamar-Schuler, D. and Tømmervik, H. (2014) Record-low primary productivity and high plant damage caused by multiple weather events and pest outbreaks. *Environmental Research Letters*, 9, 084006. https://doi.org/10.1088/1748-9326/9/8/084006.

Blinova, I. and Chmielewski, F.-M. (2015) Climatic warming above the Arctic circle: are there trends in timing and length of the thermal growing season in Murmansk region (Russia) between 1951 and 2012? *International Journal of Biometeorology*, 59, 693–705. https://doi.org/10.1007/s00484-014-0880-y.

Bromwich, D.H., Fogt, R.L., Hodges, K.I. and Walsh, J.E. (2007) A tropospheric assessment of the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *Journal of Geophysical Research*, 112, D10111. https://doi.org/10.1029/2006JD007859.

Bromwich, D.H., Wilson, A.B., Bai, L.-S., Moore, G.W.K. and Bauer, P. (2016) A comparison of the regional Arctic system reanalysis and the global ERA-Interim reanalysis for the Arctic. *Quarterly Journal of the Royal Meteorological Society*, 142, 644–658. https://doi.org/10.1002/qj.2527.

Chung, C.E., Cha, H., Vihma, T., Räisänen, P. and Scremer, D. (2013) On the possibilities to use atmospheric reanalyses to evaluate the warming structure in the Arctic. *Atmospheric Chemistry and Physics*, 13, 11209–11219. https://doi.org/10.5194/acp-13-11209-2013.

Chylek, P., Hengartner, N., Lesins, G., Klett, J.D., Humlum, O., Wyatt, M. and Dubey, M.K. (2014) Isolating the anthropogenic component of Arctic warming. *Geophysical Research Letters*, 41, 3569–3576. https://doi.org/10.1002/2014GL060184.

Cohen, J.L., Furtado, J.C., Barlow, M.A., Alexeev, V.A. and Cherry, J.E. (2012) Arctic warming, increasing snow cover and widespread boreal winter cooling. *Environmental Research Letters*, 7, 011004. https://doi.org/10.1088/1748-9326/7/1/014007.

Cohen, J., Screen, J.A., Furtado, J.C., Barlow, M., Whittleston, D., Coumou, D., Francis, J., Dethloff, K., Entekhabi, D., Overland, J. and Jones, J. (2014) Recent Arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, 7, 627–637. https://doi.org/10.1038/NGEO2234.

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, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and

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.

Dyrrdal, A.V., Isaksen, K., Hygen, H.O. and Meyer, N.K. (2012) Changes in meteorological variables that can trigger natural hazards in Norway. *Climate Research*, 55, 153–165. https://doi.org/10.3354/cr01125.

Eilertsen, H.C. and Skardhamar, J. (2006) Temperatures of north Norwegian fjords and coastal waters: variability, significance of local processes and air–sea heat exchange. *Estuarine, Coastal and Shelf Science*, 67, 530–538. https://doi.org/10.1016/j.ecss.2005.12.006.

Førland, E.J., Benestad, R.E., Flatøy, F., Hanssen-Bauer, I., Haugen, J.E., Isaksen, K., Sorteberg, A. and Ådlandsvik, B. (2009) *Climate development in north Norway and the Svalbard region during 1900–2100* (Report No. 128). Tromsø: Norsk Polarinstitutt.

Gillett, N.P., Stone, D.A., Stott, P.A., Nozawa, T., Karpechko, A.Y., Hegerl, G. C., Wehner, M.F. and Jones, P.D. (2008) Attribution of polar warming to human influence. *Nature Geoscience*, 1, 750–754. https://doi.org/10.1038/NGEO338.

Gregow, H., Jylhä, K., Mäkelä, M., Aalto, J., Manninen, T., Karlsson, P., Kaiser-Weiss, A.K., Kaspar, F., Poli, P., Tan, D.G.H., Obregon, A. and Su, Z. (2016) Worldwide survey of awareness and needs concerning reanalyses and respondents views on climate services. *Bulletin of the American Meteorological Society*, 97, 1461–1473. https://doi.org/10.1175/BAMS-D-14-00271.1.

Groisman, P.Y.A., Bogdanova, E.G., Alexeev, V.A., Cherry, J.E. and Bulygina, O.N. (2014) Impact of snowfall measurement deficiencies on quantification of precipitation and its trends over northern Eurasia. *Ice Snow*, 2, 29–43.

Heikkilä, U., Sandvik, A. and Sorteberg, A. (2011) Dynamical downscaling of ERA-40 in complex terrain using the WRF regional climate model. *Climate Dynamics*, 37, 1551–1564. https://doi.org/10.1007/s00382-010-0928-6.

Høgda, K.A., Tømmervik, H. and Karlsen, S.R. (2013) Trends in the start of the growing season in Fennoscandia 1982–2011. *Remote Sensing*, 5, 4304–4318. https://doi.org/10.3390/rs5094304.

Irannezhad, M., Chen, D. and Kløve, B. (2015) Interannual variations and trends in surface air temperature in Finland in relation to atmospheric circulation patterns, 1961–2011. *International Journal of Climatology*, 35, 3078–2092. https://doi.org/10.1002/joc.4193.

Irannezhad, M., Chen, D., Kløve, B. and Moradkhani, H. (2017) Analysing the variability and trends of precipitation extremes in Finland and their connections to atmospheric circulation patterns. *International Journal of Climatology*, 37, 1053–1066. https://doi.org/10.1002/joc.5059.

Johansson, B. and Chen, D. (2003) The influence of wind and topography on precipitation distribution in Sweden: statistical analysis and modelling. *International Journal of Climatology*, 23, 1523–1535. https://doi.org/10.1002/joc.951.

Jonsell, U., Hock, R. and Duguay, M. (2013) Recent air and ground temperature increases at Tarfala Research Station, Sweden. *Polar Research*, 32, 19807. https://doi.org/10.3042/polar.v3210.19807.

Kivinen, S., Rasmus, S., Jylhä, K. and Laapas, M. (2017) Long-term climate trends and extreme events in northern Fennoscandia (1914–2013). *Climate*, 5, 16. https://doi.org/10.3390/cli5010016.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Cocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeir, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V. and Petrovic, P. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *International Journal of Climatology*, 22, 1441–1453. https://doi.org/10.1002/joc.773.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onada, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K. and Takahashi, K. (2015) The JRA-55 reanalysis: general specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, 93, 5–48. https://doi.org/10.2151/jmsj.2015-001.

Lader, R., Bhatt, U.S., Walsh, J.E., Rupp, T.S. and Bieniek, P.A. (2016) Two-meter temperature and precipitation from atmospheric reanalysis

evaluated for Alaska. *Journal of Applied Meteorology and Climatology*, 55, 901–922. https://doi.org/10.1175/JAMC-D-15-0162.1.

Lenaerts, J.T.M., van Angelen, J.H., van den Broeke, M.R., Gardner, A.S., Wouters, B. and van Meijgaard, E. (2013) Irreversible mass loss of Canadian Arctic archipelago glaciers. *Geophysical Research Letters*, 40, 870–874. https://doi.org/10.1002/grl.50214.

Li, C., Stevens, B. and Marotzke, J. (2015) Eurasian winter cooling in the warming hiatus of 1998–2012. *Geophysical Research Letters*, 42, 8131–8139. https://doi.org/10.1002/2015GL065327.z.

Lindsay, R., Wensnahan, M., Schweiger, A. and Zhang, J. (2014) Evaluation of seven different atmospheric reanalysis products in the Arctic. *Journal of Climate*, 27, 2588–2606. https://doi.org/10.1175/JCLI-D-13-00014.1.

Marshall, G.J. (2009) On the annual and semi-annual cycles of precipitation across Antarctica. *International Journal of Climatology*, 29, 2298–2308. https://doi.org/10.1002/joc.1810.

Marshall, G.J., Vignols, R.M. and Rees, W.G. (2016) Kola Peninsula climate change during the last 50 years from meteorological observations. *Journal of Climate*, 29, 6823–6840. https://doi.org/10.1175/JCLI-D-16-0179.1.

Matishov, G., Moiseev, D., Lyubina, O., Zhichkin, A., Dzhenyuk, S., Karamushko, O. and Frolova, E. (2012) Climate and cyclic hydrobiological changes of the Barents Sea from the twentieth to twenty-first centuries. *Polar Biology*, 35, 1773–1790. https://doi.org/10.1007/s00300-012-1237-9.

Met Office. (2012) *Met office integrated data archive system (MIDAS) land and marine surface stations data (1853–current)*. NCAS British Atmospheric Data Centre. Available at: http://catalogue.ceda.ac.uk/uuid/220a65615218d5c9cc9e4785a3234bd0 [Accessed 25th February 2017].

Mori, M., Watanabe, M., Shiogama, H., Inoue, J. and Kimoto, M. (2014) Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. *Nature Geoscience*, 7, 869–873. https://doi.org/10.1038/ngeo2277.

Overland, J., Francis, J.A., Hall, R., Hanna, E., Kim, S.-J. and Vihma, T. (2015) The melting Arctic and midlatitude weather patterns: are they connected? *Journal of Climate*, 28, 7917–7932. https://doi.org/10.1175/JCLI-D-14-00822.1.

Overland, J.E., Dethloff, K., Francis, J.A., Hall, R.J., Hanna, E., Kim, S.-J., Screen, J.A., Shepherd, T.G. and Vihma, T. (2016) Nonlinear response of mid-latitude weather to the changing Arctic. *Nature Climate Change*, 6, 992–999. https://doi.org/10.1038/NCLIMATE3121.

Pike, G., Pepin, N.C. and Schaefer, M. (2013) High latitude local scale temperature complexity: the example of Kevo Valley, Finnish Lapland. *International Journal of Climatology*, 33, 2050–2067. https://doi.org/10.1002/joc.3573.

Rapaić, M., Brown, R., Markovic, M. and Chaumont, D. (2015) An evaluation of temperature and precipitation surface-based and reanalysis datasets for the Canadian Arctic, 1950–2010. *Atmosphere-Ocean*, 53, 283–303. https://doi.org/10.1080/07055900.2015.1045825.

Reichle, R.H., Koster, R.D., de Lannoy, G.J.M., Forman, B.A., Liu, Q., Mahanama, S.P.P. and Touré, A. (2011) Assessment and enhancement of MERRA land surface hydrology estimates. *Journal of Climate*, 24, 6322–6338. https://doi.org/10.1175/JCLI-D-10-05033.1.

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bachmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M. and Woollen, J. (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24, 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G. and Goldberg, M. (2010) The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057. https://doi.org/10.1175/2010BAMS3001.1.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-Y., Iredell, M., Ek, M., Meng, J., Yang, R., Peña Mendez, M., van den Dool, H., Zhang, Q., Wang, W., Chen, M. and Becker, E. (2014) The NCEP Climate Forecast System version 2. *Journal of Climate*, 27, 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1.

Santer, B.D., Wigley, T.M.L., Boyle, J.S., Gaffen, D.J., Hnilo, J.J., Nychka, D., Parker, D.E. and Taylor, K.E. (2000) Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research*, 105, 7337–7356. https://doi.org/10.1029/1999JD901105.

Simmons, A.J. and Poli, P. (2014) Arctic warming in ERA-Interim and other analyses. *Quarterly Journal of the Royal Meteorological Society*, 141, 1147–1162. https://doi.org/10.1002/qj.2422.

Swedish Meteorological and Hydrological Institute. (2017) *Meteorologiska Observationer*. Available at: https://www.smhi.se/en/climate/climate-indicators/climate-indicators-precipitation-1.91462.

Tang, Q., Zhang, X., Yang, X. and Francis, J.A. (2013) Cold winter extremes linked to Arctic sea ice loss. *Environmental Research Letters*, 8, 014036. https://doi.org/10.1088/1748-9326/8/1/014036.

Yang, Z., Hanna, E., Callaghan, T.V. and Jonasson, C. (2012) How can meteorological observations and microclimate simulations improve understanding of 1913–2010 climate change around Abisko, Swedish Lapland? *Meteorological Applications*, 19, 454–463. https://doi.org/10.1002/met.276.

Zhang, L., Kumar, A. and Wang, W. (2012) Influence of changes in observations on precipitation: a case study for the Climate Forecast System Reanalysis (CFSR). *Journal of Geophysical Research*, 117, D08105. https://doi.org/10.1029/2011JD017347.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.