Evolution of observed and modelled temperatures in Finland in 1901–2018 and potential dynamical reasons for the differences

Kimmo Ruosteenoja | Jouni Räisänen

Abstract

Observed monthly and annual mean temperatures in Finland in 1901–2018 were compared with simulations performed with 28 global climate models (GCMs), and dynamical factors behind the emerging differences were studied by regression analysis. Observational temperatures were extracted from high-quality kriging analyses specifically tailored for Finland. Considering the entire time interval, the increase in the annual multi-GCM mean temperature agrees well with the observed warming, even though observations exhibit substantial inter-decadal fluctuations. After 2000, the mean temperatures have been higher than during any period in the 20th century. In the baseline regression model, the 10 leading EOFs of the European—Northeast Atlantic sea-level pressure (SLP) field were used to explain differences between the GCM-mean and observed evolution of temperature. The regression model is able to reduce the mean squared difference of the temporally-smoothed temperature by 58%. The performance is highest in winter and summer and lowest in April. For a sensitivity assessment, multiple alternative regression models were tested, for example, one using the local SLP, geostrophic wind and vorticity as predictors. These models mostly showed somewhat inferior performance. We specifically explored the trends of monthly temperatures during 1961–2018, a period considerably affected by anthropogenic emissions. Compared with the multi-GCM mean, warming proved to be negligible in June, fairly slow in October and quite rapid in December. All these features were explained rather nicely by dynamical factors. Accordingly, the deviations of the observed regional temperature trends from the multi-GCM mean largely appear to be related to internal variability.

KEYWORDS

atmospheric circulation, climate change, CMIP5 models, internal variability, observed temperature trends, regression
1 | INTRODUCTION

During the industrial era, mean temperatures have increased nearly everywhere in the world (Hartmann et al., 2013, fig. 2.21). There is a good agreement between observations and the mean of climate model simulations for the century-scale global mean warming, but larger differences are often found when studying climatic trends on smaller spatial and temporal scales (IPCC, 2013, fig. SPM.6). Such differences may reflect deficiencies in climate models or the forcing used in the simulations, but differences are also induced by internal variability that either counteracts or amplifies the forced climate changes. For assessing the reliability of future projections, it is essential to study to what extent the model-to-observation differences can be explained by internal variability.

Regional decadal-scale fluctuations in temperature are primarily caused by dynamical variability, that is, temporal variations in the atmospheric circulation (Saffioti et al., 2016, 2017). Correspondingly, dynamical variability is mainly induced by internal variations in the climate system. This was shown by Deser et al. (2016) by studying future trends in multiple realizations simulated by the same global climate model (GCM) with identical forcing but slightly divergent initial conditions. In the individual runs, substantial changes in the sea-level pressure (SLP) occurred, but the ensemble mean response of SLP was close to zero. If the ensemble is large, the ensemble mean change approximately corresponds to the forced response while scatter across the simulations reflects internal variability.

To isolate circulation-related fluctuations in the surface air temperature, several methods have been proposed. For example, Saffioti et al. (2016) resolved the dynamically-induced component of winter temperatures in Europe in 1989–2012 by a regression model in which the coefficients of the five leading Empirical Orthogonal Functions (EOFs) of the Euro-Atlantic SLP were used as predictors. Subtracting this component from the observed temperatures yields a dynamically adjusted temperature in which externally-forced variations are prevalent. Trends calculated from the adjusted temperatures are much closer to the corresponding modelled trends than trends derived from the unadjusted data. Accordingly, dynamical adjustment eliminates a substantial portion of the trend arising from internal variations, thus revealing the actual externally forced trend. In Saffioti et al. (2017), a similar approach was applied to modelled future trends, and inter-model scatter in the temperature trends proved to be somewhat narrower when derived from the dynamically adjusted than unadjusted data. Hu et al. (2019) utilized the two leading EOFs of the surface air temperature to assess the contribution of internal variability to recent summer temperature trends in East Asia.

In the trajectory method employed by Parker (2009) and Räisänen (2019a), one finds, for every time step, the starting point of the trajectory of an air parcel, for example, 2 days before the target time. Thereafter, the sign and magnitude of the temperature anomaly at the target location is assessed for each source region of the trajectories by studying the statistical distribution of the resulting temperatures. Using this information, the temperature trends due to changes in the distribution of source regions can be estimated.

The constructed circulation analogue approach (Deser et al., 2016) is best applicable to analysing GCM output data. One first examines a long preindustrial control integration to find months having a SLP pattern resembling that of the target month. A linear combination of these analogue months is then used to create an approximation for the portion of the target-month temperature anomaly induced by circulation anomalies. To obtain the dynamically adjusted temperature, the resulting circulation-induced temperature anomaly is finally subtracted from the actual anomaly. This approach can also be applied to observational data, but in that case the analogues have to be sought from fairly short time series.

The fourth method for dynamical adjustment, termed the partial least-squares regression (Smoliak et al., 2015; Wang et al., 2020), consists of finding a regressional relationship between the spatial distributions of SLP and the near-surface air temperature. This linkage is utilized to separate the circulation-induced dynamical part of temperature variability.

As an example of a still another approach, Fleig et al. (2015) explained circulation-induced trends in temperature and precipitation in Europe by changes in the frequency and hydrothermal properties of daily Grosswetterlagen.

In the present work, the principal research questions are: (i) to what extent is the observed past temperature increase in Finland consistent with the GCM simulations, considering both the multi-GCM mean and inter-GCM scatter and (ii) is it possible to account for differences between the observed and GCM-simulated evolution of temperature by dynamical factors, that is, inter-decadal variations in the atmospheric circulation? In addition to the annual mean temperatures, the individual calendar months are explored; for example, why warming has virtually stalled in June in the past few decades while the temperature increase has simultaneously been very rapid in December.

In a sense, the present study constitutes an extension to the previous work of Räisänen (2019a), in which
The outline of this article is as follows. In Section 2, the observational datasets, GCM output data and EOF analyses are introduced. The main findings of the study are presented in Section 3: a comparison of the temporal evolution of observed temperatures with the GCM output; explaining the emerging differences by dynamical factors, especially by the spatial distribution of SLP; and the monthly trends of temperature during a sub-period substantially affected by anthropogenic greenhouse gas emissions. The article is finalized by the discussion and conclusions Sections 4 and 5.
both including only those grid squares that fall within the domain of the kriging analysis. The borders of the subregions are portrayed in the upper-left panel of Figure 1. Because of the above-mentioned deficiencies in the analyses, northernmost Lapland (>68°N) is excluded completely. The subregions are larger than those examined by Räisänen (2019a), since the time series are now far longer than in that work and thus more likely exposed to inhomogeneities caused by changes in the station network and the environment of measurement sites. The influence of local inhomogeneities can be minimized by dealing with averages over sufficiently wide areas.

2.2 | GCM data

The modelled evolution of mean temperatures in Finland was derived from simulations performed with 28 GCMs participating in Phase 5 of the Coupled Model Intercomparison Project (CMIP5). The GCMs are listed in table 1 of Luomaranta et al. (2014); that article also provides argumentation on the selection of the GCMs. To facilitate the interpretation of inter-GCM differences, only one realization for every GCM has been included. The time series covering the entire period have been compiled by merging the output of the historical
simulations for the period 1901–2005 with the Represen-
tative Concentration Pathway RCP4.5 simulations since
2006. Because the period considered only extends to
2018, the various RCP scenarios do not diverge noticeably
(Ruosteenoja et al., 2016), and therefore the choice of the
scenario is immaterial.

Since we are interested in the trends and temporal
variability of temperature rather than modelling biases,
all GCM-simulated and observational temperatures were
converted to anomalies from their 20th-century
(1901–2000) monthly mean values prior to any further
analysis. The resulting temperature deviations from the
20th century mean are thus commensurable in the GCM
and observational data. Because the focus is on climato-
logical and inter-decadal rather than interannual vari-
tions, 15-year running means ($\bar{T}$) were calculated by
using linearly attenuating weights:

$$
\bar{T}_i = (T_{i-7} + 2T_{i-6} + ... + 7T_{i-1} + 8T_i + 7T_{i+1} + ... + T_{i+7})/64
$$

(1)

where $i$ refers to the ordinal number of the year since
1900. The resulting temporally smoothed modelled tem-
perature deviations were used to calculate the 28-GCM
means and, using the normality assumption, the 5–95%

**FIGURE 2** Composite mean
temperature (upper panels) and SLP
anomalies (lower panels) for the July
months belonging to the 10% subset
with the largest negative (left) and
largest positive (right) temperature
deviation in southern Finland
(60–64°N). Contour interval is 0.5°C
for the temperature and 1 hPa for the
SLP anomaly [Colour figure can be
viewed at wileyonlinelibrary.com]
probability intervals. All the calculations were performed separately for the 12 calendar months and the annual means.

To distinguish inter-decadal variations, in many previous studies (e.g., Räisänen, 2019a) the temperatures have been detrended by subtracting a linear trend from the original time series. In the present work, however, the temporal range is so long (118 years) that a linear trend does not adequately reproduce the long-term evolution; global warming affects the temperatures substantially only during the last few decades of the period. Consequently, a different approach has been adopted: observational temperature deviations from which the corresponding 28-GCM mean deviation has been subtracted are used in place of the detrended temperatures. Figures 1 and 2 display composite averages of the resulting temperature residuals for cold and warm January and July months. The time series of these residuals readily reveal any systematic differences in the long-term evolution of temperature in the observations and the GCM output.

2.3 | ERA-20C and ERA5 analyses; SLP and geostrophic winds

SLP data for the period 1901–2010 were extracted from the ERA-20C reanalyses (Poli et al., 2016). Monthly means were calculated from the original data provided at 3-hr intervals. The ERA5 reanalyses (Hersbach et al., 2020) were used since 2011. For homogeneity, we subtracted the 1995–2010 mean monthly differences between the two analyses from the ERA5 SLP fields of 2011–2018. The resulting merged ERA analyses were then used to calculate the monthly-mean geostrophic wind (\(u_g\), \(v_g\)) and vorticity (\(\zeta_g\)). Temperature data needed in the calculation of the geostrophic winds were taken from the same reanalyses without any adjustments.

The \(\zeta_g\) fields derived from the unmodified SLP data proved to be highly noisy. Therefore, we henceforth use SLP that is smoothed by taking spatial averages over a box consisting of 9 grid points of the 1.125° grid in the zonal and 5 points in the meridional direction. Spatial smoothing does not have any discernible influence on the SLP fields, and \(u_g\) and \(v_g\) are modified only lightly.

The lower panels of Figures 1 and 2 show composite mean SLP anomalies for cold and warm January and July months. In accordance with the previous findings of Tuomenvirta and Heino (1996), in winter mild weather conditions in Finland are accompanied by widespread anomalously strong westerlies and cold conditions with northerly-easterly flow anomalies. This agrees with the well-known relationship between northern European temperature anomalies and the phase of the North Atlantic Oscillation (NAO) (Hurrell, 1995; Iles and Hegerl, 2017). In summer, extreme monthly-mean temperatures are induced by smaller-scale SLP patterns than in winter (Figure 2). Warm summer weather is entailed by anticyclonic southerlies-easterlies in conjunction with a high-pressure area to the east of Finland; a similar SLP pattern under elevated temperatures in Finland has been reported to occur at a daily level by Kim et al. (2018). Conversely, cyclonic north-westerlies tend to cause cool conditions. In anomalous summer months, the composite pressure anomalies in central and western Europe are fairly weak, unlike in winter. Very similar SLP distributions were obtained regardless of whether one studies months having large temperature anomalies in the southern or the northern subregion. Note that, both in January and July, the pressure patterns related to large positive and negative temperature anomalies in Finland are largely mirror images.

Jaagus et al. (2003) found that the lengths of the thermal seasons in northeastern Europe likewise correlate with atmospheric circulation patterns. A zonal flow results in short thermal winters, while an anomalous south-easterly flow acts to make the start of the thermal summer earlier and the termination later. All these findings are consistent with the very different climatological temperature distributions in northern Europe and its adjacent areas in winter and summer. In winter, mean temperatures increase from the east to the west while in summer, the temperature gradient is conversely directed towards the southeast.

Figures 1 and 2 clearly indicate the strong connection between the temperature and circulation anomalies in northern Europe, thus motivating one to develop statistical models that explain temperature anomalies by SLP and other circulation-related quantities.

2.4 | EOF expansions

The EOF components of monthly-mean SLP were calculated from the temporally unsmoothed 118-year merged ERA dataset by using the climate data operators (CDO) software. For the sensitivity studies discussed below, the EOF analysis was likewise performed for \(u_g\) and \(v_g\). Because data from the edges of the area are lost when calculating \(u_g\) and \(v_g\) by finite differences (Section 2.3), the domain for the EOF analyses is 40°–76°N, 4°W–54°E, which is slightly smaller than the area from which SLP is analysed in Figures 1 and 2. The spatial patterns of EOF components 1–10 for January and July are shown in the Appendix. In components 1–3, the pattern is dipolar and in components 4–5 a quadrupole. Consistently with the
typical behaviour of EOF patterns, higher components exhibit an increasingly fine-scale spatial structure.

EOF components 1–4 account for 94% of the total variance of SLP in January and 87% in July. For 10 components, the corresponding shares are 99.4 and 98.4%. When using 20 EOFs, only 0.1–0.3% of the total variance remains unexplained.

For comparison, the EOF components were likewise inferred from a smaller domain (45°–72°N, 6°–46°E). For most of the components, particularly the leading ones, the patterns proved to be very similar to those derived from the large domain.

In the regression model to be developed in section 3.2, we shall use EOF coefficients 1–10 of SLP to explain temporal variations in the observed temperature (represented as a deviation from the multi-GCM mean). This choice is consistent with the well-known 1 in 10 rule, according to which a regression model should not use more than one predictor per 10 data points to keep the risk of overfitting low (Harrell et al., 1996, see also https://en.wikipedia.org/wiki/One_in_ten_rule). Moreover, as the time series of the EOF coefficients are mutually uncorrelated in time, the risk of obtaining unrealistically large regression coefficients is reduced. The sensitivity of the regression models to the number of EOFs included is briefly discussed in Section 3.3.

3 | RESULTS

3.1 | Observed and GCM-derived temperature changes

The evolution of the observed temporally-smoothed annual mean temperature is shown in Figure 3 (both subregions) and monthly mean temperatures in Figure 4 (southern Finland only). Both figures also display the 28-GCM means along with the 90% uncertainty interval. After 2000, the observed annual means have been substantially higher than in any period during the 20th century. Moreover, when considering the entire time span, observed warming agrees well with the multi-GCM mean. Nevertheless, temperatures have fluctuated considerably on the decadal time scale, the most prominent anomalies being the warmth of the 1930s and the coolness of the 1980s. Even so, observed temperatures have remained within the multi-GCM uncertainty range, with the exception of the 1930s in the north.

Regarding the individual months (Figure 4), in April, May and September recent temperatures have been much higher than during the warmest phases of the preceding century. In June, by contrast, in this century mean temperatures have been nearly the same as in 1940–2000 on average and far lower than in the 1960–1970s. In mid and late winter, recent temperatures have been fairly high but still somewhat lower than in the anomalously warm 1990s. In the other months of the year, temperatures measured after 2000 have been close to those in the

![Figure 3](https://example.com/figure3.png)
FIGURE 4 Temporal evolution of the southern Finland (60–64°N) monthly temperature deviations; 15 year running means. Note that, the scale of the y axis varies among the panels. For further information, see the caption of Figure 3 [Colour figure can be viewed at wileyonlinelibrary.com]
3.2 Regression relationship between SLP and temperature

To uncover dynamical factors behind the differences between the observed and GCM-simulated temperatures, a linear regression model was developed in which monthly temperature deviations from the 28-GCM mean were explained by the 10 leading EOF coefficients of SLP over Europe and the Northeast Atlantic Ocean. The regression model was derived separately for every calendar month using temporally unsmoothed data.

In the majority of the months, the regression equation explained about 60–70% of the total variance of the temperature deviation. In spring and early autumn, the performance was lower. The proportion of explained variance was smallest in April, 41% in the south and 49% in the north. Nevertheless, the F test for 10 predictors and 118 data points indicates that only a 19% proportion of explained variance is needed to confirm statistical significance at the 1% level. Hence, the regression model is statistically highly significant for all the calendar months.

However, in the present work our interest lies on decadal and longer-term variations rather than on yearly anomalies. Therefore, we used Equation (1) to calculate 15-year running means from the temperature deviations given by the regression model. The resulting smoothed deviations were added to the multi-GCM mean temperatures, giving time series of the regression-derived temperatures that are comparable to the observed and GCM-derived temperatures in Figures 3 and 4.

According to Figure 3, the annual mean calculated from the monthly regression fits is able to reproduce most of the observed decadal-scale temperature fluctuations in a qualitative sense, for example, the high annual-mean temperatures in the 1930s and late 1940s and the cold anomalies in the 1960s and 1980s. Nevertheless, the regression model generally tends to underestimate the amplitude of the anomalies, with the exception of the weak positive anomaly in the early 1990s.

For the individual months (Figure 4), the regression model reproduces several anomaly periods qualitatively and even quantitatively correctly. For example, the agreement is good over most of the time in January, June, August and November. Conversely, the regression model fails to find the anomalously cold Aprils in the 1950s and the warm Aprils after 2000, for instance. Further examples of failure: warm Septembers in the 1930s, cool Septembers in the 1970s and warm Octobers in the 1960s. In the 1930s, the regression model predicts cold temperatures for March, even though the observed anomaly is small. In February, the model tends to overestimate the positive anomaly in the 1990s. Most of these finding also hold for the northern subregion.

To obtain an objective measure for the performance of the regression model, we calculated monthly rms differences between the 15-year running means derived from the regression fit and the observed temperature shown in Figure 4. Subsequently, the annual rms difference was calculated as a square root of the mean of the squared monthly rms values. The resulting annual rms error for the regression fit is 0.43°C for the southern (Table 1) and 0.52°C for the northern subregion. These values can be compared with the corresponding annual

| Table 1 | Annual rms differences (in °C) between the estimates given by the various regression models and the actual observed southern Finland (60–64°N) temperature, both deviations from the multi-GCM mean |
| --- | --- |
| EOF truncation | 4 | 6 | 8 | 10 | 20 |
| (i) SLP, large area | 0.492 | 0.479 | 0.444 | 0.427 | 0.396 |
| (ii) SLP, small area | 0.500 | 0.479 | 0.455 | 0.440 | 0.395 |
| (iii) v_g, large area | 0.523 | 0.517 | 0.494 | 0.477 | 0.414 |
| (iv) v_g, large area | 0.563 | 0.545 | 0.559 | 0.554 | 0.446 |

*Note: The differences are given for four categories of regression model, in which the regressors are the 4, 6, 8, 10 or 20 leading EOF coefficients of (i) SLP in the large area, (ii) SLP in the small area and (iii) the zonal and (iv) meridional components of the geostrophic wind over the large area. The rms differences are calculated from 15-year running means. The rms for the baseline regression model is underlined. For comparison, the corresponding rms difference between the observational and multi-GCM mean temperature is 0.657°C.*

warmest decades of the 20th century. Most of these features are also noticeable in the north (not shown).

According to Figure 4, observational monthly temperatures have sporadically resided outside of the GCM-derived 5–95% probability interval in all months apart from February and April. Considering the mean of all 12 months, the proportion of years with the monthly mean temperature outside of the interval is 8.1% for the southern and 9.3% for the northern subregion. If the deviations of the observed temperature from the GCM estimate were purely induced by stochastic internal variability, the expectation for the total length of such externality periods would be 10%. The closeness of the observed proportion to the theoretical expectation supports the notion that inter-decadal fluctuations in temperature largely reflect internal variability. One factor making the proportion slightly smaller than the expectation is that in the last few decades of the span, the inter-GCM scatter is also affected by the different climate sensitivities of the GCMs.
rms difference between the multi-GCM mean and observed temperatures, 0.66°C in the south and 0.80°C in the north. Accordingly, in both subregions, taking into consideration the influence of inter-annual variations in the SLP field through the regression equation reduces the mean squared difference from the observed temperature by 58%.

3.3 Regression models using EOF coefficients—sensitivity assessment

In this section, we compare the performance of the baseline regression model discussed in Section 3.2 with models using other combinations of predictors: a varying number of EOF components of SLP; EOF components of SLP calculated for the smaller domain; and EOF components of $u_g$ or $v_g$. Only southern Finland is considered. Annual rms errors for the various regression models are given in Table 1 and monthly rms errors for a subset of models in Figure 5.

Using an increasing number of EOF coefficients as predictors consistently reduces the annual rms error (Table 1, row (i)). In particular, the lowest truncation ($N_{\text{EOF}} = 4$) yields quite a large rms error for June (Figure 5); in fact, in June it is EOF5 of SLP that has the highest correlation with the temperature deviation. On the other hand, with $N_{\text{EOF}} = 20$ some high EOFs proved to have quite large regression coefficients, which may be an indication of overfitting. Thus, $N_{\text{EOF}} = 10$ used in the baseline regression model would be a reasonable compromise.

Using the EOF functions of the small rather than the large domain has little influence on the performance of the regression analysis. Compared with the models with the large-area EOFs, the rms errors are virtually identical or slightly inferior (Table 1, row (ii)). This is an expected finding since the patterns of the EOF components were very similar for both domains (Section 2.4).

In the two final experiments, the predictors used in the regression model were the EOF coefficients of $u_g$ or $v_g$. In the spatial pattern of $u_g$ and, in particular, $v_g$ small-scale features are much more influential than in SLP, and the proportion of variance explained by a predefined number of EOFs is lower accordingly. Therefore, the rms errors are consistently larger than in a regression model utilizing the same number of EOFs for SLP (Table 1, rows (iii),(iv); Figure 5). Nonetheless, the difference becomes narrower for $N_{\text{EOF}} = 20$. Under a high truncation, all the variables (SLP, $u_g$ and $v_g$) largely incorporate the same information about the structure of the geostrophic circulation field. However, a regression model with an excessive number of predictors may be susceptible to overfitting.

In principle, it would be possible to use the EOF coefficients of different variables as predictors simultaneously. However, since this would make the predictors mutually correlated and thus increase the risk of overfitting, we did not explore such an option.

3.4 A regression model using local circulation variables

As an alternative for regression models based on the EOF expansions, we studied a model in which the predictors represent local circulation conditions in the subregion considered. In winter, the temperature deviations proved to have the highest correlation with $u_g$ ($r > 0$) and SLP ($r < 0$), in summer with SLP ($r > 0$), $v_g$ ($r > 0$) and $\zeta_g$ ($r < 0$). These correlations are in line with the composite maps shown in Figures 1 and 2. For consistency, we chose to use the same predictors throughout the year: SLP, $u_g$, $v_g$ and $\zeta_g$ at a point near the centre of the target...
area (61.875°N, 25.875°E or 66.375°N, 27°E). Unlike in the regression models using the EOF coefficient, the predictors are now mutually correlated. In particular, the anomalies of SLP and $\zeta_g$ tend to be opposite, even though these variables provide information about different spatial scales of the circulation field.

The annual rms errors yielded by this regression model were 0.501°C (southern subregion) and 0.608°C (northern subregion). Accordingly, the performance is substantially lower than in the baseline regression model and even slightly inferior to the SLP regression model with $N_{\text{eof}} = 4$ (Table 1). The monthly rms errors are likewise generally larger than in the baseline model (Figure 5).

3.5 | Multi-decadal trends

In addition to inter-decadal variability, we explore the ability of the dynamical variables in explaining multi-decadal trends. In this respect, the period 1961–2018 is of particular interest, since at the beginning of this time span several months were anomalously cold, with the exception of June and October that were fairly warm (Figure 4). This period likewise covers the epoch during which anthropogenically-forced global warming has mainly taken place (fig. 2.20 in Hartmann et al. (2013) and fig. 8.18 in Myhre et al. (2013)).

The monthly linear trends of the observed and multi-GCM mean temperature, along with the estimates inferred from the baseline regression model (Section 3.2) and the local grid-point value model (Section 3.4), are shown in Figure 6. Following the conventions applied in Figures 3 and 4, the regression-model derived trends are given in absolute terms, so that the estimate given by the regression equation is added to the multi-GCM mean trend; this makes all the trends readily comparable.

The observational and multi-GCM annual mean trends are very similar (2.10 vs. 2.06°C in 57 years in the south and 2.25 vs. 2.09°C in the north). Moreover, in nine (south) or six (north) months out of 12, the trends differ at most by 0.5°C/57 a. In both subregions, the largest differences between the observational and GCM-derived trends occur in June (observed $-0.42/+0.14$ and multi-GCM mean $+1.64/+1.66$C per 57 years a in the south/north) and December ($+3.95/+4.27$ vs. $+2.34/+2.68$C). Compared with the trends derived from the individual GCM simulations, the observed southern-subregion trend in June falls near the lower end but not outside of the frequency distribution (Figure 7a); two GCMs out of 28 simulate negative trends for that period, one of these being less and the other more negative than the observed trend. The strong observational warming in December is a less extreme case, since there are five GCMs producing even more positive trends (Figure 7b).

The baseline regression model, utilizing the 10 leading EOF coefficients of SLP, reproduces the seasonal course of the observed trend fairly closely (Figure 6). In particular, the trend derived from the regression fit is far weaker than the multi-GCM mean trend in June and October.
and substantially larger in December, concordantly with observations. Figure 7 reveals that the regressional trends (recall that the 28-GCM mean trends have been added to these trends) lie in the correct portion of the frequency distribution but are less extreme than their observational counterparts, particularly in December. In addition, in the north the regression model tends to underestimate the trends in March and April.

Furthermore, Figure 6 shows the trends inferred from the regression model in which the predictors describe local conditions in southern/northern Finland. The seasonal distribution of the trend is generally reproduced less successfully than by the baseline regression model. For example, the weak trends in June and October are partly misrepresented. Conversely, in December this model even outperforms the baseline model.

Räisänen (2019a) studied temperature trends in Finland in 1979–2018. He likewise found that cooling caused by changes in the atmospheric circulation produced near-zero temperature trends for June and October despite the continuing global warming. As can be seen in Figure 4, the weakness of these monthly trends is to large extent due to negative temperature anomalies after 2000. This explains the similar behaviour of the trends despite the different period examined.

We additionally studied centennial (1919–2018) trends, but for that period, the trends derived from the regression models were consistently closer to the 28-GCM mean trend than the observed trend (not shown). This indicates that in this temporal scale dynamical factors included in the regression model do not succeed well in explaining the deviations of the observed trends from the corresponding GCM-derived trends. Fluctuations in the geostrophic circulation may in general be more capable of explaining inter-decadal temperature anomalies than centennial-scale trends. Dynamically-induced anomalies are indeed most prominent on the decadal time scale (Figure 4), and their influence largely tends to be smoothed out in the long term.

One potential source for the failure of the regression models in explaining the century-scale trend is the warmth of the 1930s. The warm period occurred near the beginning of the time interval, thus engendering a strong torque that acts to reduce the linear trend. Figures 3–4 indicate that the positive temperature anomaly of the 1930s was mainly induced by factors other than circulation anomalies in the Euro-Atlantic area. This topic will be discussed further in the next section.

4 | DISCUSSION

All the regression models examined in section 3 can explain a substantial proportion of inter-decadal variability in the difference between the observed and multi-GCM mean temperature evolution (Table 1). The highest
agreement is achieved by a model in which the predictors describe the spatial distribution of SLP over the European—Northeast Atlantic area. In practice, the SLP pattern can be represented adequately by ten EOF components.

Even so, there is a number of potential reasons why SLP and the other dynamical quantities cannot explain the observation versus multi-GCM differences completely. Evidently, there does not exist any unambiguous linear relationship between the SLP field and the local temperature anomaly. A similar monthly-mean temperature anomaly can occur in conjunction with various SLP patterns. Moreover, even an identical monthly-mean SLP field may be an average of different combinations of daily SLP patterns. For example, a winter month with 15 days of a fresh westerly and 15 days of easterly flow would have a different mean temperature than a month with weak winds throughout the month, even if the monthly mean SLP fields were fairly similar in both cases.

Furthermore, even the instantaneous SLP field does not control the prevailing temperature perfectly. For example, in winter anticyclonic weather can well be either cloudy or clear, the strongly negative net longwave radiation favouring cold temperatures in the latter case. Wintertime temperature conditions are likewise affected by ice conditions in the nearby Baltic Sea and snow cover through their albedo and thermal insulation capacity (Vavrus, 2007). Correspondingly, summer temperatures are influenced by soil moisture through the latent heat flux of evapotranspiration (Seneviratne et al., 2010) and cloudiness (Räisänen, 2019b). These factors partly explain the coincidence of a high SLP and temperature in summer (Figure 2), since anticyclonic weather conditions tend to be clear and dry. When the correlation between the predictor(s) and predictand is <1, a linear regression model by definition tends to underestimate temporal variations. When considering the annual means, the tendency towards underestimation is further enhanced by averaging over the 12 months (Figure 3).

The performance of the regression models proved to be lowest in April. According to fig. 5 of Räisänen (2019b), in southern Finland April is the month when the contribution of both the surface albedo and ground thermodynamical processes (above all, melting of snow and ice) to the temperature is larger than in any other month. In some years, there is still heavily of snow and sea-ice in April, the high albedo and latent heat of melting reducing the temperatures; in other years, snow and ice have largely disappeared. Observations indicate that in the long run the spring-time disappearance of snow cover in Finland (Luomaranta et al., 2019) and ice in the northern Baltic Sea (Jevrejeva et al., 2004; Jaagus, 2006; Haapala et al., 2015) have shifted significantly earlier as a consequence of warming. Secondly, April constitutes a transition period with a more winter-like state of the atmospheric circulation at the beginning than end of the month. In the cold half-year, anomalously warm weather conditions are mainly induced by westerly, in the warm half-year by southerly winds (Figures 1 and 2). In the cold climate of the early 20th century, April has been a more wintry month than around 2000 and afterwards, and thus the circulation patterns producing cold/warm anomalies evidently have not remained the same during the 118 year period examined. Both the above-discussed phenomena act to weaken the regressional relationship between monthly mean circulation patterns and the temperature anomalies.

In the time series of the observed temperature deviation, the most prominent anomaly is the warmth of the 1930s; this feature is visible both in the annual means (Figure 3) and multiple individual months in summer, autumn and early winter (Figure 4). This temperature anomaly is reproduced deficiently by the present regression models. According to Hegerl et al. (2018), the anomaly was indeed largely induced by factors other than atmospheric circulation patterns over Europe and the Northeast Atlantic Ocean. In the 1930s, the NAO index on average manifested a weak negative rather than positive anomaly (see also fig. 1a in Hurrell (1995)). Conversely, the Arctic ice extent was low and the Atlantic Multidecadal Oscillation (AMO) index in its positive phase, accompanied by anomalously high sea surface temperatures in the North Atlantic Ocean. Globally, the warm period was contributed by a quiet phase in volcanic activity, but in Europe, the main cause of the warmth was evidently internal variability in the climate system (Hegerl et al., 2018). According to Hartmann et al. (2013) (p. 193), anomalously high temperatures in the 1930–1940s mainly occurred in the middle and high latitudes of the Northern Hemisphere.

We studied SLP patterns occurred in the 1930s for selected months (not shown). In that decade December, for instance, was characterized by a large positive temperature anomaly (Figure 4). In the average SLP pattern, by contrast, the dominant feature was a high-pressure anomaly over northern Europe; this is in disagreement with the typical conditions of mild winters (Figure 1, bottom-right panel). Correspondingly, in March, April and September the circulation patterns resembled those producing cold anomalies for the respective season, whereas the actual temperature anomalies were neutral or slightly positive (Figure 4).

Parker (2009) studied the influence of circulation fluctuations on the Central England temperature anomalies.
The impact of dynamical factors was clearly discernible in winter and to some extent in spring, while in summer and autumn the signal was weak. Parker (2009) suggest that one factor reducing the contribution of atmospheric circulation anomalies is the large influence of sea surface temperatures, since England is an island surrounded by ocean. In Finland, climate is more continental than in England; this may explain the evident year-round contribution of dynamical factors (Figure 4).

Deser et al. (2016) and Wang et al. (2020) scrutinized large ensembles of parallel runs performed with a single GCM by varying the initial conditions. The resulting trends in temperature and, in particular, SLP diverged widely across the simulations. This confirms the previous conclusion of Deser et al. (2012) that variations in SLP are mainly induced by internal variability in the climate system, the contribution of external forcing being negligible. In the temperature response, the ensemble-mean change (forced component) was comparable to the scatter among the simulations (the contribution due to internal variability). Wang et al. (2020) state that internal variability can well either cancel or double the forced temperature change signal. In the temperature trends of Finland in 1961–2018, June approximately represents the former and December the latter alternative (Figure 6).

5 | CONCLUSIONS

When studying the annual means, the observed temperature increase in Finland since the early 20th century is in good accordance with the average response of the 28 GCMs (Figure 3). Because of such a high agreement, there is no material reason to suspect that multi-GCM derived future projections of temperature (e.g., Ruosteenoja et al., 2016) would be systematically under- or overestimated. Of course, the degree of agreement may partly be fortuitous, in particular, considering the large uncertainty in anthropogenic aerosol forcing (Knutti, 2008) and the potential contribution of internal variability in observed temperatures (Deser et al., 2016). Observed temperature changes have already been used to constrain the future global-mean temperature response to increasing greenhouse gas concentrations (Liang et al., 2020), but the benefit of this approach on regional scales is still unresolved.

We examined in more detail temperature trends in the period 1961–2018 that is significantly influenced by human-induced warming. During this period, the observed and GCM-derived annual-mean temperature trends are very similar. This finding widely holds for the individual months as well, the differences between the observed and GCM mean trends usually being smaller than 0.5°C/57 a (Figure 6). The observational and GCM-derived trends diverge most seriously in October, December and especially in June. Nevertheless, even in these months, the observed trend does not fall outside of the frequency distribution of the GCM-derived trends (Figure 7). Accordingly, the differences do not indicate any fundamental contradiction between the observational and GCM-derived temperature trends, neither on the annual nor monthly level. Moreover, as was discussed in Section 3.1, the proportion of time that the observed temperature deviations lie outside of the GCM-derived 5–95% probability interval is close to its theoretical expectation of 10%. Consequently, it appears plausible that the differences of the observed evolution of temperature from the multi-GCM mean largely reflect internal variability.

Furthermore, we elaborated a regression model to ascertain to what extent it is possible to explain differences between the observed and GCM-derived evolution of temperature by dynamical factors, in particular, the distribution of SLP in the European—Northeast-Atlantic area. Dynamical factors explain, for example, the negligibility of warming in June, weakness in October and the rapid rate in December after 1961, compared to the average GCM trend. In quantitative terms, the regression model still tends to underestimate differences in the trends to some extent.

In addition to the baseline regression model explaining temperature anomalies by the spatial distribution of SLP, we evaluated the performance of several alternative regression models using varying combinations of regressors. For example, a satisfactory albeit somewhat inferior agreement is obtained by a model in which the local values of SLP, \( u_g \), \( v_g \) and \( \zeta_g \) are used as predictors.

One key issue is whether the recent anomalies in the observed monthly temperature trends persist in the near future. The present findings suggest that a substantial portion of differences between the evolution of the observed and multi-GCM mean monthly temperatures is caused by dynamical factors. Since dynamically-induced fluctuations are primarily related to internal variability in the climate system (Deser et al., 2016), their contribution to future annual and monthly temperature anomalies is likely to be largely unpredictable. Hence, there is no particular reason to anticipate that the recent seasonal distribution of the temperature trends, a hiatus phase in June and rapid warming in December, for instance, would continue in a similar manner in the future. More likely, the evolution of the temperatures in such anomalous months will return closer to that in the adjacent months.

Nevertheless, it is likewise evident that substantial dynamically-induced anomalies continue to occur in the coming decades, and anomalously strong and weak, maybe even negative, multi-decadal temperature trends
will be experienced in some months. Even so, it is virtually impossible to foresee what are the particular months that might exhibit such divergent trends.

ACKNOWLEDGEMENTS

Financial support for this research was provided by the Academy of Finland through the Heat and health in the changing climate project (HEATCLIM, decision number: 329307). The CMIP5 GCM data were downloaded from the Earth System Grid Federation data archive (http://pcmdi9.llnl.gov). Observational temperature analyses for Finland were supplied by Pentti Pirinen. The two anonymous reviewers are acknowledged for their positive feedback.

ORCID

Kimmo Ruosteenoja https://orcid.org/0000-0002-4370-0782
Jouni Räisänen https://orcid.org/0000-0003-3657-1588

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.

Deser, C., Phillips, A., Bourdette, V. and Teng, H. (2012) Uncertainty in climate change projections: the role of internal variability. Climate Dynamics, 38, 527–546. https://doi.org/10.1007/s00382-010-0977-x.

Deser, C., Terray, L. and Phillips, A.S. (2016) Forced and internal components of winter air temperature trends over North America during the past 50 years: mechanisms and implications. Journal of Climate, 29, 2237–2258. https://doi.org/10.1175/JCLI-D-15-0304.1.

Fleig, A.K., Tallaksen, L.M., James, P., Hisdal, H. and Stahl, K. (2015) Attribution of European precipitation and temperature trends to changes in synoptic circulation. Hydrology and Earth System Sciences, 19, 3093–3107. https://doi.org/10.5194/hess-19-3093-2015.

Haapala, J.J., Ronkainen, I., Schmelzer, N. and Sztobryn, M. (2015) Recent change – sea ice. In: Second Assessment of Climate Change for the Baltic Sea Basin. The BACC II Author Team, Ed., chapter 8. Heidelberg, Germany: Springer, pp. 145–153.

Harrell, F.E., Lee, K.L. and Mark, D.B. (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine, 15, 361–387.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M. and Zhai, P. (2013) Observations: atmosphere and surface. 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.) chapter 2Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York, USA: Cambridge University Press, pp. 159–254.

Hegerl, G.C., Brönnimann, S., Schurer, A. and Cowan, T. (2018) The early 20th century warming: anomalies, causes, and consequences. WIREs Climate Change, 9, e522. https://doi.org/10.1002/wcc.522.

Herschbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hölm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J.N. (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Hu, K., Huang, G. and Xie, S.P. (2019) Assessing the internal variability in multi-decadal trends of summer surface air temperature over East Asia with a large ensemble of GCM simulations. Climate Dynamics, 52, 6229–6242. https://doi.org/10.1007/s00382-018-4503-x.

Hurrell, J.W. (1995) Decadal trends in the North Atlantic oscillation: regional temperatures and precipitation. Science, 269, 676–679. https://doi.org/10.1126/science.269.5224.676.

Iles, C. and Hegerl, G. (2017) Role of the North Atlantic oscillation in decadal temperature trends. Environmental Research Letters, 12, 114010. https://doi.org/10.1088/1748-9326/aa9152.

IPCC. (2013) Summary for policymakers. 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, United Kingdom and New York, USA: Cambridge University Press.

Jaagus, J. (2006) Trends in sea ice conditions in the Baltic Sea near the Estonian coast during the period 1949/1950–2003/2004 and their relationships to large-scale atmospheric circulation. Boreal Environment Research, 11, 169–183.

Jaagus, J., Truu, J., Ahas, R. and Aasa, A. (2003) Spatial and temporal variability of climatic seasons on the east European plain in relation to large-scale atmospheric circulation. Climate Research, 23, 111–129.

Jevrejeva, S., Drabkin, V.V., Kostjukov, J., Lebedev, A.A., Leppäranta, M., Mironov, Y.U., Schmelzer, N. and Sztobryn, M. (2004) Baltic Sea ice seasons in the twentieth century. Climate Research, 25, 217–227. https://doi.org/10.3354/cr025217.

Kim, S., Sinclair, V.A., Räisänen, J. and Ruuhela, R. (2018) Heat waves in Finland: present and projected summertime extreme temperatures and their associated circulation patterns. International Journal of Climatology, 38, 1393–1408. https://doi.org/10.1002/joc.5253.

Knutti, R. (2008) Why are climate models reproducing the observed global surface warming so well? Geophysical Research Letters, 35, L18704. https://doi.org/10.1029/2008GL034932.

Liang, Y., Gillett, N.P. and Monahan, A.H. (2020) Climate model projections of 21st century global warming constrained using
the observed warming trend. *Geophysical Research Letters*, 47, e2019GL086757. https://doi.org/10.1029/2019GL086757.

Luomaranta, A., Aalto, J. and Jylhä, K. (2019) Snow cover trends in Finland over 1961–2014 based on gridded snow depth observations. *International Journal of Climatology*, 39, 3147–3159. https://doi.org/10.1002/joc.6007.

Luomaranta, A., Ruosteenoja, K., Jylhä, K., Gregow, H., Haapala, J. and Laaksonen, A. (2014) Multimodel estimates of the changes in the Baltic Sea ice cover during the present century. *Tellus A*, 66, 22617. https://doi.org/10.3402/tellusa.v66.22617.

Myhre, G., Shindell, D., Bréon, F.M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T. and Zhang, H. (2013) Anthropogenic and natural radiative forcing. In: Stocker, T.F., Qin, D., Plattner, G.K., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.) chapter 8. *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, United Kingdom and New York, USA: Cambridge University Press, pp. 659–740.

Parker, D.E. (2009) Anomalies of Central England temperature classified by air source. *Journal of Climate*, 22, 1069–1081. https://doi.org/10.1175/2008JCLI2250.1.

Poli, P., Hersbach, H., Dee, D.P., Berrisford, P., Simmons, A.J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.N., Trémolet, Y., Hólm, E.V., Bonavita, M., Isaksen, L. and Fisher, M. (2016) ERA-20C: an atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29, 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1.

Räisänen, J. (2019a) Effect of atmospheric circulation on recent temperature changes in Finland. *Climate Dynamics*, 53, 5675–5687. https://doi.org/10.1007/s00382-019-04890-2.

Räisänen, J. (2019b) Energetics of interannual temperature variability. *Climate Dynamics*, 52, 3139–3156. https://doi.org/10.1007/s00382-018-4306-0.

Ruosteenoja, K. (1988) Factors affecting the occurrence and lifetime of 500 mb height analogues: a study based on a large amount of data. *Monthly Weather Review*, 116, 368–376. https://doi.org/10.1175/1520-0493(1988)116<0368:FATOAL>2.0.CO;2.

Ruosteenoja, K., Jylhä, K. and Kämäräinen, M. (2016) Climate projections for Finland under the RCP forcing scenarios. *Geophysical Research Letters*, 51, 17–50 http://www.geophysica.fi/pdf/geophysica_2016_51_1-2_017_ruosteenoja.pdf.

Saffioti, C., Fischer, E.M. and Knutti, R. (2017) Improved consistency of climate projections over Europe after accounting for atmospheric circulation variability. *Journal of Climate*, 30, 7271–7291. https://doi.org/10.1175/JCLI-D-16-0695.1.

Saffioti, C., Fischer, E.M., Scherrer, S.C. and Knutti, R. (2016) reconciling observed and modeled temperature and precipitation trends over Europe by adjusting for circulation variability. *Geophysical Research Letters*, 43, 8189–8198. https://doi.org/10.1002/2016GL069802.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B. and Teuling, A.J. (2010) Investigating soil moisture-climate interactions in a changing climate: a review. *Earth-Science Reviews*, 99, 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004.

Smoliak, B.V., Wallace, J.M., Lin, P. and Fu, Q. (2015) Dynamical adjustment of the Northern Hemisphere surface air temperature field: methodology and application to observations. *Journal of Climate*, 28, 1613–1629. https://doi.org/10.1175/JCLI-D-14-00111.1.

Tietäväinen, H., Tuomenvirta, H. and Venäläinen, A. (2010) Annual and seasonal mean temperatures in Finland during the last 160 years based on gridded temperature data. *International Journal of Climatology*, 30, 2247–2256. https://doi.org/10.1002/joc.2046.

Tuomenvirta, H. and Heino, R. (1996) Climatic changes in Finland – recent findings. *Geophysica*, 32, 61–75.

Vavrus, S. (2007) The role of terrestrial snow cover in the climate system. *Climate Dynamics*, 29, 73–88. https://doi.org/10.1007/s00382-007-0226-0.

Wang, Y., Huang, G. and Hu, K. (2020) Internal variability in multidecadal trends of surface air temperature over Antarctica in austral winter in model simulations. *Climate Dynamics*, 55, 2835–2847. https://doi.org/10.1007/s00382-020-05412-1.

How to cite this article: Ruosteenoja K, Räisänen J. Evolution of observed and modelled temperatures in Finland in 1901–2018 and potential dynamical reasons for the differences. *Int J Climatol.*, 2021;41:3374–3390. https://doi.org/10.1002/joc.7024.
APPENDIX: The EOF Functions

The patterns of the EOF components 1–10 of SLP, separately for January and July, are displayed in Figure A1.

FIGURE A1  Spatial patterns of EOF components 1–10 of SLP for January (left panels) and July (right panels) [Colour figure can be viewed at wileyonlinelibrary.com]