Version of Record: https://www.sciencedirect.com/science/article/pii/S0169809523000224
Manuscript_127c2a2c60d4ef06444ef91e70816f12

Variability and trends of the total cloud cover over Italy (1951-2018)

Veronica Manara[1], Michele Brunetti[2], Martin Wild[3], Maurizio Maugeri[1]

[1]Department of Environmental Science and Policy, Università degli Studi di Milano, via Celoria 10, 20133 Milano, Italy

[2]Institute of Atmospheric Sciences and Climate, ISAC-CNR, via Gobetti 101, 40129, Bologna, Italy

[3]Institute for Atmospheric and Climate Sciences, ETH Zurich, Zurich, Switzerland

Corresponding author: Veronica Manara (veronica.manara@unimi.it)

Abstract

A new quality-checked and homogenised dataset of total cloud cover (TCC) series over Italy for the period 1951-2018 is presented and the variability and trends of the obtained regional series is discussed. The diurnal cycle highlights the relevant role of convection induced by solar radiation that, as expected, is more relevant at medium and high elevations and in summer. In parallel, the annual cycle presents a strong minimum in July and a maximum during winter for southern Italy, while it exhibits a more complex behaviour with strong differences between low elevations and high elevations in northern Italy. Moreover, the seasonal and annual TCC series are characterized by a significant negative trend over the whole considered period, which is mainly due to the 1951-1990 period. Even if small differences between northern and southern Italy can be observed, the two regions exhibit a coherent behaviour both for long-term trends and decadal time-scale variability suggesting that the causes of variability and trends of the Italian TCC records are more related to large scale factors rather than to local scale changes. Indeed, the comparison with sea level pressure and 500 hPa geopotential height data highlights that large-scale atmospheric circulation explains a relevant fraction of the signal of the Italian TCC records. Finally, the new TCC dataset shows that the long-term evolution of sunshine duration and surface solar radiation in Italy is only partially influenced by changes in TCC.

Keywords

Total cloud cover, variability, trends, atmospheric circulation, surface solar radiation, Italy

1. Introduction

Considerable changes in the physical climate system have been detected globally and many of them have been attributed with a very high confidence level to anthropogenic influences. Whereas global warming is a phenomenon established with high confidence, the assessment of other climate

© 2023 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

variables and their changes in response to anthropogenic forcing is more uncertain (Sfîcă et al., 2021).

Cloudiness has a direct and profound effect on the Earth's climate. On one side, clouds absorb the longwave radiation emitted by the Earth and emit energy both to space and backward to the surface (e.g., Zelinka and Hartmann, 2010; Zhou et al., 2016). On the other side, they scatter and absorb shortwave radiation originating from the Sun. Together with atmospheric turbidity, cloudiness is the main cause of interannual and decadal variability of shortwave radiation reaching the Earth's surface, contributing to the global energy budget (Ramanathan et al., 1989). Indeed, cloudiness can both heat or cool the Earth, where the cooling effect is rather dominant even if considerable spatial and temporal heterogeneities are present (Henderson-Sellers, 1992). The effect of clouds is strongly dependent on their horizontal coverage (total cloud cover – TCC) but also on their form, height, and optical properties. For example, low layered clouds can cool the surface by reflecting sunlight back into space and so increasing the albedo of the Earth. Differently, high, thin clouds enhance the atmospheric greenhouse effect, as they are almost transparent to shortwave radiation, but they also absorb the outgoing longwave radiation and thus have a warming effect on the Earth. Moreover, clouds play a key role also in the hydrological cycle (Free and Sun, 2014).

Despite the importance of TCC, there are still large uncertainties in the variability and changes of this variable as well as in its response to anthropogenic and natural forcings (e.g., Mendoza et al., 2021). Therefore, clouds are still amongst the largest uncertainties in estimating and interpreting the changes of the Earth's energy budget. In fact, the response of cloudiness to the increase in greenhouse gases and the cloud feedbacks are currently the primary source of uncertainties for climate sensitivity studies and future climate scenarios (e.g., Free and Sun, 2014; Pörtner et al., 2022). These uncertainties are connected to the difficulty in quantifying and attributing the impact of changes in large scale phenomena like atmospheric circulation and warming-induced increase in the atmospheric water vapour content on clouds, but also in quantifying the role of more local phenomena like changes in the aerosol indirect effects (i.e., aerosol induced changes in cloud properties such as lifetime and albedo) (Ramanathan et al., 2001; Rosenfeld et al., 2008) or, for example, changes in land use (e.g., intense urbanization possibly increases the frequency of occurrence of convective clouds due to the heating surface) (Fan et al., 2020).

Observations of cloudiness at weather stations have been performed systematically since the mid-19th century, when national meteorological offices started to be established, reporting not only information about the extension of clouds but sometimes also on the type (Auchmann et al., 2012). However, these observations are available only at selected stations and only over land areas.

Long-term TCC records are mainly based on visual observations, because instrumental measurements have been introduced only in the last decades (e.g. in the early 1990s in the U.S. (Dai et al., 2006)). Despite their subjectivity and local character, these records constitute the only source of information on long time scales. In fact, even if an assessment of the cloudiness variations on global or hemispheric scales including remote areas can only be obtained through the use of remote sensing data, satellite records cover only a short period (few decades) and are often affected by inconsistencies with respect to their trends due to drift in observation time during a satellite mission, instrument changes between missions, and the gradual decay of satellite orbits (Pfeifroth et al., 2018).

In the literature there is a large number of studies analysing the spatial and temporal evolution of TCC and underlying causes based on ground-based observations, satellite data and reanalysis data (see Sanchez-Lorenzo et al., (2012) for a review). Recently, Kambezidis et al., (2021) have presented a mathematical determination of the sky conditions (clear, overcast, intermediate) that refer to the cloud-cover conditions (no clouds, thick clouds, fast moving clouds) via the sunshine-duration parameter basing their methodology on the definition of three ranges for the diffuse fraction. However, the number of studies spanning a long period is limited, due to the lack of complete time series and to the shortage of digitized historical data. Moreover, TCC records often present problems with data homogeneity, particularly concerning biases due to changing observing practices and times of observations (Dai et al., 2006).

The global TCC record over the period 1971-2009 shows a decline during the 1970s and 1980s followed by a steady cloud amount through 2009 with interannual variation on the order of ±1% (Eastman and Warren, 2013). The observed decline is due to decreasing cloudiness at middle and high level with a higher contribution from the first one. Eastman and Warren, (2013) suggest that this signal may be due to the poleward expansion of the Hadley cell and to a subsequent shift in mid-latitude jet streams, while it is not clear if there could be a contribution from aerosol changes. Some studies have analysed TCC observations since the mid-20th century focusing on smaller areas like for example the Mediterranean area (Sanchez-Lorenzo et al., 2017), Russia (Khlebnikova and Sall, 2009), Eurasian Artic (Chernokulsky and Esau, 2019), US (Free and Sun, 2014; Sun et al., 2015), Australia (Jovanovic et al., 2011), China (Xia, 2010; Zong et al., 2013), Tibetan Plateau (You et al., 2014), Italy (Maugeri et al., 2001), Southern India (Biggs et al., 2007), Spain (Sanchez-Lorenzo et al., 2009), Switzerland (Sanchez-Lorenzo and Wild, 2012) or focusing on single stations (Founda et al., 2018). The sign and the intensity of the reported trends depend on the considered region and period.

Italy is a key area due to its location in the heart of the Mediterranean basin, which is considered a "hot spot" for climate change (see e.g. Tuel and Eltahir, (2020)). The Italian territory includes both the Po Plain, one of the most polluted areas in Europe, and the Alpine chain, an area with complex orography (Fugazza et al., 2021). Moreover, in this region, a strong dimming/brightening signal in sunshine duration (SD) (Manara et al., 2015) and surface solar radiation (SSR) records (Manara et al., 2016a) has been observed during about the last seventy years. The analyses suggested that these variations are due to changes in anthropogenic atmospheric aerosols (Manara et al., 2020, 2019a, 2017a), which have shown a significant increase followed by a marked decrease during the period considered (Manara et al., 2019b). An influence of clouds can however not be ruled out. The only study which includes a comprehensive analysis of TCC records for the Italian territory (35 series over the 1951-1996 period), published more than 20 years ago, reports a highly significant negative trend (about 1 okta over 50 years), especially due to a strong decrease after the 1970s (Maugeri et al., 2001). Therefore, in this context, this paper aims to set up a new updated, quality-checked and homogenised dataset of TCC data series (improved both in the covered time period and in the number of series) over the Italian territory for the 1951-2018 period and to study the variability and trends of the obtained regional time series.

After the Introduction, the TCC dataset is presented in Section 2, the data pre-processing is shown in Section 3. The results are presented in Section 4 and discussed in Section 5. Finally, some conclusive remarks are given in Section 6.

2. Data

The data considered in this work comes from the Italian Air Force (Aeronautica Militare Italiana – AM) Surface Synoptic Observations (SYNOPs). They are available from both the AM website (http://clima.meteoam.it/RicercaDati.php) and the National Oceanic and Atmospheric Administration website (NOAA - https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00532). The database includes records from 168 stations over the entire Italian territory with eight observations per day (see Figure 1 for the spatial distribution of the stations selected after the quality controls – Section 3.4). For each station, together with the data, we recovered also the available metadata, which provide useful information during quality/homogeneity checking, allowing to better identify any possible change in the station conditions affecting the quality of the data (Aguilar et al., 2003). For some stations only the coordinates are available while for others, AM provides very detailed metadata reporting an exhaustive history of the station (http://www.meteoam.it/page/elenco-stazioni).

We focus here on the eight daily total cloud cover (TCC) synoptic observations. Specifically, TCC is a human observation for all the considered stations and over the whole considered period, defined as the fraction of the celestial dome covered by clouds. It is measured in oktas (eights) where 0 and 8 signify a completely clear sky and an overcast sky, respectively. The value 9 is used when the sky is obscured by fog or other meteorological phenomena and so the amount of clouds in the sky cannot be correctly estimated (WMO, 2008 and reference therein).

All series were subjected to a preliminary quality control in order to remove negative and higher than 9 values (no such values were found). Moreover, all observations equal to 9 were removed and considered as not available.

3. Data pre-processing

3.1 Gap-filling process for the synoptic observations

The first step of data pre-processing was the development of a methodology to estimate sparse missing observations. The goal of this procedure was to avoid losing the possibility to calculate the monthly values because of few missing days. This methodology has been applied separately for each of the eight daily synoptic observations and for each series (test series). Each daily missing value was calculated as the weighted mean of its estimations from the seven neighbouring series with the highest weights in terms of distance, elevation difference and similarity of cloudiness climatology (reference series).

The weight associated with each candidate reference series was calculated as the product of distance, elevation, climatology (this one to avoid comparison among stations characterized by different cloudiness regimes) and angular weighting functions. The distance, elevation and climatology weights are based on Gaussian functions:

$$w = e^{-\left(\frac{\Delta^2}{c}\right)} \quad (1)$$

with:

$$c = \left(\frac{\Delta_{1/2}{}^2}{ln2}\right) \quad (2)$$

where Δ is the difference between the value of the considered parameter at test and reference station and c is a coefficient that regulates the weight decreasing rate, with $\Delta_{1/2}$ defined as that value of Δ for which the weight is equal to 0.5. For the distance weight, $\Delta_{1/2}$ is set to 100 km (i.e. the distance

168 at which the common variance is about 50%), for the elevation weight it is set to 250 m if the test
169 station has an elevation lower than or equal to 500 m, otherwise it is set to half of the elevation of
170 the test station. For the climatology weight, $\Delta_{1/2}$ is set to 0.5 oktas. The station climatologies were
171 calculated by considering only the common values available for both test and candidate reference
172 series over a window of +/- 30 days centred on the observation to be estimated, and considering the
173 same window also for the previous and the subsequent year. So, for each considered observation,
174 the maximum length of the considered window is equal to 182 days. Finally, an angular term was
175 added in order to take into account the anisotropy in the spatial distribution of the reference stations
176 around the test one (New et al., 2000). A reference series was retained and considered a possible
177 reference series only if the observation corresponding to the missing day in the test series is
178 available and if it has at least 91 values (50%) in common with the test series in the considered
179 window.
180 Once the seven best reference series have been selected, for each of them the percentile associated
181 with the observation corresponding to the missing value of the test series was calculated with
182 respect to the distribution of its values belonging to the window under consideration and was used
183 to extract the simulated value from the corresponding distribution in the test series. This approach is
184 particularly suitable for those variables with a lower and/or upper bounded domain.
185 The different steps of the gap-filling procedure are summarized in Figure 2. The procedure has been
186 tested on the available data, by comparing observed and independently reconstructed values with a
187 leave-one-out approach, by recursively excluding each of them and estimating it by means of the
188 described procedure. The resulting mean bias error (MBE) was almost null in all months. The mean
189 absolute error (MAE) over all synoptic observations was 1.24 oktas, with values ranging from 1.08
190 oktas at 12 UTC to 1.39 oktas at 3 UTC, while the corresponding root mean square error (RMSE)
191 over all synoptic observations was 1.78 oktas, ranging from 1.52 oktas at 12 UTC to 2.03 oktas at 0
192 and 3 UTC. Overall, the mean percentage of daily reconstructed missing data at each of the eight
193 daily synoptic observations was 7.4%.
194
195 **3.2 Calculation of the monthly series**
196 Starting from the gap-filled observations, the monthly series were calculated for each station and
197 each synoptic observation time, estimating the monthly mean only if at least 80% of observations
198 were available. Also monthly records of the half of the interquartile range of the daily observations
199 were calculated to investigate any signal in the TCC day-to-day variability. These records were,
200 however, obtained excluding the reconstructed observations to avoid the gap filling procedure to
201 bias the variance of the records.

3.3 Homogenisation

All monthly series were then subjected to a homogenisation procedure in order to identify and remove non-climatic signals (Aguilar et al., 2003; Brunetti et al., 2006). Any inhomogeneities were identified by comparing, for each of the eight daily synoptic times, each station series (test series) with a corresponding simulated series (reference series). The simulated series was calculated by removing each entry from the test series (i.e. with a leave-one-out approach) and by reconstructing it as the weighted mean of the reconstructions from the seven reference series with the highest weights and with at least 15 values in common with the test series in the considered month. More specifically, each monthly entry of this series was first estimated from each of the seven best reference series ($\widetilde{TCC}_{test_i}$) by means of the following relation:

$$\widetilde{TCC}_{test_i} = TCC_{ref_i} + \overline{TCC_{test_i}} - \overline{TCC_{ref_i}} \quad (i = 1, \dots, 7) \tag{3}$$

where TCC_{ref_i} is the entry of the corresponding month of the *i-th* reference series and $\overline{TCC_{test_i}}$ and $\overline{TCC_{ref_i}}$ are the average values of the considered month calculated over the common data (i.e. data available both for the test and reference series) of the test and the *i-th* reference series, respectively. Then, it was obtained performing the weighted mean of these seven estimations, with the weights estimated as those of the daily gap-filling procedure.

To better highlight problematic periods, the series of the differences between the simulated and test series were calculated. Moreover, all test series were compared against the corresponding simulated ones, for each synoptic observation, by means of the Craddock test (Craddock, 1979). Specifically, the cumulative normalized differences between the test and simulated series were calculated (Manara et al., 2017b) and the breaks in the test series were identified by a discontinuity in the first derivative of these series. If a discontinuity was evident for all the synoptic observations and from all the comparisons, it was evaluated as an inhomogeneity of the station record. As a rather low number of inhomogeneities were identified, no corrections were performed to adjust the parts of the records that turned out to be inhomogeneous. These parts were simply deleted from the dataset.

The detected inhomogeneities were compared with the available metadata in order to better understand their origin. Some of them turned out to be due, for example, to periods in which measurements had not been taken adequately, others to changes in station location or to permanent or prolonged changes in the area surrounding the station.

Considering the whole dataset and all synoptic observations, the percentage of data eliminated during the homogenisation procedure accounted for 6.3% of the total data.

3.4 Gap-filling procedure for the monthly series

After the homogenisation procedure, only the series with at least 20 years of available data were retained. The number of the selected stations was comprised between 80 stations selected at 0 UTC and 128 stations selected at 12 and 15 UTC. Moreover, the data availability was highest for the 9 and 12 UTC synoptic observations and lowest for the 0 and 21 UTC ones, as expected, considering that some stations make observations only during the day (Figure 1). For all synoptic observations, the data availability increased after the beginning of the investigated period, reaching the maximum during the 1960s and the 1970s (Figure 3) and then it decreased in the subsequent period.

Finally, we filled the gaps in each monthly record using simulated series obtained from the homogenised version of the dataset. Moreover, we applied the same procedure to the monthly half interquartile range records.

To evaluate the performance of the gap-filling procedure, we compared the monthly test and independently simulated (i.e. with a leave-one-out approach) series over the available data for the stations included in the final dataset. Thereby the MBE over all synoptic observations was 0 oktas, the MAE was 0.33 oktas, with values comprised between 0.29 oktas (at 9, 12 and 15 UTC) and 0.37 oktas (at 0 and 3 UTC), while the RMSE was 0.41 oktas, with values comprised between 0.36 oktas (at 9, 12 and 15 UTC) and 0.47 oktas (at 0 UTC).

Finally, for each series and all synoptic observations, we calculated the seasonal (winter: December of the previous year, January, February - spring: March, April, May – summer: June, July, August – autumn: September, October, November) and annual (from December of the previous year to November of the considered year) mean series and then the corresponding seasonal and annual anomaly series. The anomaly series were calculated with an additive approach with respect to the 1961-1990 reference period being the 30-year period with the highest data availability (Figure 3).

3.5 Mean regional anomaly series

To prevent the spatial representativeness of the dataset from being biased by the non-homogeneous distribution of the stations, we generated a 1° resolution gridded version of the dataset of the monthly, seasonal and annual anomaly series and of the corresponding half interquartile range ones. We used an Inverse Distance Weighting interpolation approach (considering both a radial and an angular weight), involving for each node the stations available within about 138.5 km (d) where the d threshold is the mean distance of each grid-point from the subsequent one (increasing both the longitude and the latitude by one degree). Each grid-point series was calculated if there was at least

one available station at a distance lower than *d/2* or two available stations at a distance lower than *d*. The *d* value was also used to define the distance at which the radial weight decreased to 0.5.

Then, the gridded anomaly series for all synoptic observations were clustered into two homogeneous regions by means of a Principal Component Analysis (PCA) (Preisendorfer, 1988; Wilks, 1995). The analysis was performed over the whole considered period. We decided to rotate the first two empirical orthogonal functions (EOFs), which together accounted for a percentage of variance ranging from about 84% at 21 UTC (about 69.6% and 14.7% for the first and the second, respectively) to more than 91% at 12 (about 81.9% and 9.3% for the first and the second, respectively) and 15 UTC (about 82.5% and 9.3% the first and the second, respectively). Consistently with similar analysis performed for other meteorological variables (Manara et al., 2015, 2016a, 2019b) this procedure allowed to divide the Italian territory into two regions: northern and southern Italy (Figure 1).

Finally, for each synoptic observation, we averaged the gridded TCC records (and corresponding half interquartile range series) into northern and southern Italy seasonal and annual mean anomaly series to study the TCC variability and trends over the 1951-2018 period (Section 4.2). The trends were quantified by means of the Sen-Theil method (Sen, 1968; Theil, 1950), while the significance of the trends was evaluated by means of the Mann-Kendall non parametric test (Sneyers, 1992). The trends were investigated both considering the entire investigated period and focusing onto any sub-period at least 20-year long (Manara et al., 2019b).

4. Results

4.1 Analysis of the TCC diurnal and annual cycle

Figure 4 shows the mean seasonal and annual TCC diurnal cycles (over the 1961-1990 reference period), obtained by averaging the single stations' diurnal cycles over northern and southern Italy domains and for different orographic elevation belts. Orographic elevations belts have been defined as in Manara et al., (2019b): stations with elevation lower than or equal to 500 m (low-elevation areas – L), stations with elevation higher than 500 m and lower than or equal to 1400 m (medium-elevation areas – M), and stations with elevation higher than 1400 m (high-elevation areas – H). The last two orographic elevation belts were aggregated into one single layer (mid- and high-elevation areas - MH) for southern Italy, where the elevation range is lower and very few stations are available in mountain areas.

For all regions and seasons, the TCC monthly 1961-1990 averages are lower during the night and higher during the day with highest values for almost all cases at 12 and 15 UTC. Moreover, at annual level for low-elevation areas the TCC values are higher for the northern region than for the

303 southern one, for all synoptic observations (values ranging from about 3.4 oktas to 4.3 oktas in the
304 north and from 2.8 oktas to 4.0 oktas in the south). With increasing elevation, the daily cycle
305 amplitude increases. Specifically, during the night high-elevation sites have lower TCC than low-
306 elevation sites in the north, whereas in the south the values of the different elevation belts are
307 comparable. Differently, during the day high-elevation sites have always higher TCC than low-
308 elevation ones (the maximum values are 4.3 oktas, 4.6 oktas, 4.8 oktas for N-L, N-M and N-H,
309 respectively, and 4.0 oktas and 4.5 oktas for S-L and S-MH respectively).
310 During summer, the values are always higher in the north than in the south and again the values
311 increase with elevation (the maximum value is equal to about 3.5 oktas and to about 5.2 oktas for
312 N-L and N-H, respectively and to about 2.4 oktas and to 3.2 oktas for S-L and S-MH, respectively).
313 The daily cycle is particularly pronounced during this season with a difference between day and
314 night that increases with increasing elevation (from about 1.1 oktas for N-L to about 2.4 oktas for
315 N-H and from about 1.0 okta for S-L to 1.9 oktas for S-MH).
316 Differently during winter in the northern region, the mean TCC decreases with increasing elevation
317 (values at 15 UTC ranging from 4.8 oktas to 4.0 oktas for N-L and N-H respectively), while in the
318 southern region the values are slightly higher for high-elevation sites (values ranging from about 5.1
319 oktas to 5.3 oktas for S-L to S-MH respectively at 15 UTC). Moreover, winter is the only season in
320 which southern Italy has higher TCC values than northern Italy.
321 Finally, spring and autumn present a diurnal cycle more comparable between the different regions
322 even if also in this case high-elevation areas present a more pronounced diurnal cycle than low-
323 elevation ones.
324 Figure 5 shows the mean 1961-1990 annual cycles for each region and for all synoptic observations.
325 In the southern region the monthly 1961-1990 averages show (both for S-L and S-MH regions)
326 higher values during winter and lower values during summer (the minimum value is almost always
327 observed during July) for all synoptic observations, showing a pronounced annual cycle. A similar
328 behaviour is observed for the N-L region even if the difference between the maximum and the
329 minimum values of the annual cycle is less pronounced than for the south (about 1.7 oktas for N-L
330 and about 3.1 oktas for S-L at 15 UTC). An opposite cycle is observed for the high-elevation areas
331 in the north, where the highest values are observed during spring and the lowest ones during winter
332 while the mid-elevation areas show a behaviour more similar to that observed for low-elevation
333 ones.
334 The results presented above are obtained considering all the available observations and so without
335 the implementation of any moonlight criterion. In order to evaluate the impact of those observations
336 that were taken during nights without moon illuminance, we implemented the moonlight criterion

following Hahn et al., (1995). We divided the data into three categories: observations made with sunlight, observations made during nights with moon and observations made during nights without moon. Specifically, observations were classified as made with sunlight when the solar elevation was higher or equal to -9° and a night was classified with moon when the lunar elevation was higher or equal to 0.01° and the lunar illumination was higher than 0.11. At 9, 12 and 15 UTC all observations were performed with sunlight, whereas at 0, 3, 21 UTC sunlight was never present. At 6 and 18 UTC the presence of sunlight depended on the season and also on the location of the considered station. Considering only the observations performed without sunlight, we verified that the mean TCC calculated considering only the nights with moon turned out to be higher of about 0.2 oktas than that calculated considering all the nights. The bias on the average TCC induced by the lack of illuminance in nights without moon can therefore be estimated as about 0.2 oktas. This bias, even though significant, is small in comparison to the daily cycles presented in Figure 4 and it has a negligible effect on the seasonal cycles presented in Figure 5.

4.2 Trend analysis

The regional records of the eight synoptic observations exhibit very similar behaviour both in terms of long-term trends and variability at decadal time scale (figure not shown). Thereby, we present and discuss the temporal evolution of the Italian TCC records focusing only on one observation time. Specifically, we focus onto 12 UTC (Figures 6, 7, 8 and Table 1) which is the synoptic observation with the highest data availability over the considered period (see Figure 3). The annual TCC series are characterized by a comparable significant negative long-term trend for both regions (-0.11±0.02 oktas per decade over the 1951-2018 period – p-value ≤ 0.05 - Table 1). The series are composed of two main periods: the first, running from the beginning of the series to the end of the 1980s, showing a strong significant negative trend (-0.15±0.03 oktas per decade and -0.18±0.03 oktas per decade over the 1951-1990 period, respectively for the north and the south - p-value ≤ 0.05 – Table 1) and the subsequent one where the observed trends are not significant. This behaviour is evident also by considering shorter sub-periods (Figure 8) where the trends are almost always negative and significant if they start before the 1980s, while they are generally not significant if they start after the beginning of the 1990s.

On seasonal level, the behaviour is similar even if some peculiarities can be detected.

During winter for the 1951-2018 period, the trends are significant but slightly stronger for the south than for the north (-0.11±0.04 oktas per decade and -0.12±0.02 oktas per decade for the north and the south, respectively - p-value ≤ 0.05 – Table 1). Also in this case, two sub-periods seem to be detectable, the first (1951-1990) with a negative and significant trend (slightly stronger but less

significant for the north, with about -0.18±0.08 oktas per decade - p-value ≤ 0.10 – Table 1, than for the south, with -0.15±0.05 oktas per decade - p-value ≤ 0.05 – Table 1) and the second (1991-2018) with no significant trend for both regions, apart some shorter sub-periods starting in the 1990s with a significant positive trend (Figure 8).

During spring, as already observed for the annual series, the trends over the whole considered period are significant, negative and comparable in intensity between the two regions (-0.13±0.03 oktas per decade and -0.13±0.02 oktas per decade for the north and the south, respectively, over the 1951-2018 period - p-value ≤ 0.05 – Table 1). Differently, the trends over the first period (1951-1990) are significant only in the south (-0.15±0.05 oktas per decade - p-value ≤ 0.05 – Table 1), with the trend in the northern region becoming significant only extending the sub-period to mid-1990s (Figure 8). In the second period (1991-2018), the trends are again not significant.

Summer is the only season showing a significant stronger negative trend for the north than for the south over the whole considered period (-0.16±0.02 oktas per decade and -0.09±0.02 oktas per decade for the north and the south respectively - p-value ≤ 0.05 – Table 1). It is mostly due to a stronger decrease over the first period (-0.15±0.04 oktas per decade and -0.11±0.05 oktas per decade in the north and in the south, respectively, over the 1951-1990 period - p-value ≤ 0.05 – Table 1), which persists also in the second period, becoming even stronger (-0.23±0.10 oktas per decade over the 1991-2018 period - p-value ≤ 0.10 – Table 1).

Finally, autumn is the only season for which the trend over the whole considered period is negative and significant only in the southern region (-0.07±0.03 oktas per decade - p-value ≤ 0.05 – Table 1), while in the northern region it is not significant. The trend over the first period is negative and significant for both regions (-0.17±0.08 oktas per decade and -0.27±0.06 oktas per decade for the north and the south, respectively - p-value ≤ 0.05 – Table 1). In the second period the trend is negative and significant only in the north (-0.19±0.11 oktas per decade over the 1991-2018 period – Table 1). Generally, the northern region is characterized by a less persistent tendency than the southern one, as testified by the very few significant trends over shorter sub-periods (Figure 8).

The average summer and annual northern and southern Italian TCC half interquartile range records (Figure 9) show a negative significant trend in both regions while in the other series the trend is not significant. This indicates that, beside a decrease in the TCC, the examined period is characterized at annual level and during summer also by a reduction of the day-to-day TCC variability.

5. Discussion

It is interesting to highlight that the Italian TCC records are not affected by the inhomogeneity present in many other TCC records caused by the shift from human observations to automatic

405 measurements that generally occurred in the last decades (Dai et al., 2006). This is because in Italy
406 the National Meteorological Service is managed by the Air Force, which has continued to perform
407 human observations of clouds (and of other variables such as visibility (Manara et al., 2019b)) also
408 when other Meteorological Services have moved to automatic observations and the Air Force is still
409 continuing to perform them (Vergari and Meli, 2020). The strong interest of the Italian National
410 Meteorological Service in cloud observations is also documented by the high number of available
411 records (Figure 1), the length of the period they cover, and the low fraction of missing data (Figure
412 3). Also the results of the homogenisation procedure give evidence of the high quality of the Italian
413 TCC records: the homogeneity issue could in fact be managed by simply removing some
414 inhomogeneous periods.
415 The mean seasonal and annual TCC diurnal cycles (Figure 4) showed lower values during the night
416 and higher values during the day. In particular, as the elevation increases, the daily cycle becomes
417 more pronounced and high-elevation areas have always higher values during the day than low-
418 elevation ones. This fact can be due to higher convective activities in mountainous regions.
419 Moreover, the mean annual cycles (Figure 5) showed higher values during winter and lower values
420 during summer probably due to more pronounced storm activity in winter and to the northward shift
421 of the subtropical high pressure systems towards the Italian territory during summer. The only
422 exception is represented by the high-elevation areas in the north, where the highest values were
423 observed during spring and the lowest ones during winter, probably due to a stronger convective
424 activity with more insolation in spring than in winter. The highest values were observed in spring
425 and not in summer probably because summertime is frequently affected by anticyclonic circulation,
426 which reduces the generation of convective clouds. Moreover, the regional differences in the annual
427 cycles of TCC are consistent with what has been observed for the annual cycle of precipitation in
428 the different areas of Italy (Crespi et al., 2018).
429 The Italian TCC seasonal and annual anomaly series (Figures 6, 7 and 8) are characterized by a
430 significant negative trend over the whole considered period comparable in intensity for both
431 regions. In parallel, a reduction of the day-to-day TCC variability is observed (Figure 9).
432 Two main sub-periods can be identified: the first, from the beginning of the series to the end of the
433 1980s, showing a significant negative trend and the subsequent one where the observed trends are
434 generally not significant. Even if small differences between northern and southern Italy can be
435 observed, the two regions exhibit a coherent behaviour both for the long-term trends and decadal
436 time-scale variability. Moreover, they exhibit a reasonable agreement in terms of year-to-year
437 variability. A significant fraction of the signal in the TCC records is, therefore, common for the
438 entire investigated area and this explains why the first principal component of the PCA analysis

explains a so high fraction of the variance of the data (82% of the variance of the dataset at 12 UTC). This aspect suggests that the causes of variability and trends of the Italian TCC are related more to large scale factors rather than to local scale changes.

To verify possible connections between the Italian TCC and changes in the atmospheric circulation we downloaded the ERA5 monthly sea-level pressure (SLP) (Bell et al., 2020a; Hersbach et al., 2019a) and the 500-hPa geopotential height (z500) (Bell et al., 2020b; Hersbach et al., 2019b) data from the Copernicus Data Store and we calculated the Italian mean series considering the same domain we used for the Italian TCC series. Moreover, using the same data for the region 40W-50E for longitude and 25N-70N for latitude, we calculated the correlation between each SLP and z500 seasonal and annual grid-point record and the corresponding Italian TCC mean anomaly record.

The Italian average yearly SLP record (Figure 10) gives evidence of a significant positive trend in the 1951-2018 period with an increase of 0.23±0.07 hPa per decade. Moreover, most of this increase is concentrated between 1970 and the beginning of the 1990s, where the TCC decrease is particularly evident (-0.20±0.06 oktas for the 1970-1990 period; p-value $\leq$ 0.05) and where the trend of the annual SLP record is 1.1±0.3 hPa per decade (period 1970-1990; p-value $\leq$ 0.05). This suggests an opposite behaviour of the two variables over Italy. This interesting behaviour is however less clear when the seasonal records are considered. The seasonal SLP records give in fact evidence of a significant (p-value $\leq$ 0.05) increase and a very good agreement only in winter while in the other seasons, even if the sign of the slope is positive, the slopes present strong differences. This is particularly evident in summer that is the season with the lowest increase of the SLP record. Interestingly, all seasonal SLP records exhibit however the strongest positive trend in the 1970s and 1980s. In this period also the extended winter NAO (Jones et al., 1997) gives evidence of a strong increase. This strong increase is also evident in the occurrence of a “High pressure over Europe” according to a weather classification proposed by Schwander et al. (2017); this causes this particular weather pattern to reach the absolute maximum since 1763. In addition, in this case the clearest increase is observed in the winter months.

Considering the 500-hPa geopotential height (Figure 10), the increase in the 1951-2018 period is even more evident as it is also affected by the strong warming, which occurred in this period. Changes in the geopotential height at 500-hPa level are in fact directly related to changes in the temperature of the layer between the surface and the 500-hPa level (Christidis and Stott, 2015), showing a high positive correlation between the 500-hPa geopotential height and global warming (Hafez and Almazroui, 2014). At decadal time-scale the pattern of the 500-hPa geopotential records is, however, coherent with those of the SLP records and also in this case the 1970s and the 1980s

turn out to be a period of a particularly evident increasing tendency, with an increase of 200±46 m^2s^{-2} (p-value ≤ 0.05).

Besides an opposite long-term trend and decadal scale variability, TCC and SLP/500-hPa geopotential records present also an opposite behaviour in terms of year-to-year variability in all seasons. We investigated this issue by subtracting the 11-year window, 3-year standard deviation Gaussian low-pass filters from the corresponding anomaly records and by studying the correlation between the residual records. This avoids correlations to be influenced by long-term trends.

In winter, spring and autumn the correlations with the TCC records are stronger for the SLP records (the values are -0.70, -0.76, -0.67, respectively) than for the 500-hPa geopotential records, whereas in summer the correlation of the TCC record with the 500-hPa geopotential record (-0.72) is much stronger than that with the SLP record (-0.49). Also this analysis seems, therefore, to confirm that the summer TCC signal is only weakly linked to SLP. To further investigate the correlation between the Italian TCC records and the SLP/500-hPa fields, we extended the analysis to a larger area (40W-50E, 25N-70N). The goal was to highlight which are the grid points whose SLP/500-hPa geopotential height mostly influence the Italian year-to-year TCC variability. The results are rather similar for SLP and 500-hPa geopotential height and we, therefore, show them only for the latter variable (Figure 11). As expected, the maximum negative correlation with Italian TCC is a bit displaced to the northwest of the Italian territory: a positive anomaly of the 500-hPa geopotential height over the mid-latitudes of western Europe shields the regions south of the Alps from wet air masses coming from the Atlantic Ocean. Figure 11 gives also evidence of the large spatial extension of the negative correlation between the 500-hPa geopotential height records and the Italian TCC records.

An important aspect that should be considered to understand the causes of changes in TCC is the role of aerosol changes on cloud formation and lifetime (Wild, 2009). Aerosols may affect clouds in different ways depending on the levels of pollution. In pristine areas, additional cloud condensation nuclei strongly increase the formation, lifetime and albedo of clouds. In highly polluted areas, on the other hand, cloud formation is suppressed, since absorbing aerosol layers heat and stabilize the atmosphere, while lower SSR due to the high pollution reduces the available energy for the surface evaporation, which triggers convective clouds. Both stabilization and reduction of surface evaporation lead to a suppression of cloud formation.

The Italian TCC dataset can also help to better understand the strong dimming/brightening signal observed in SD (Manara et al., 2015) and SSR (Manara et al., 2016a) for Italy. The analyses performed in previous papers suggested that these variations are due to changes in anthropogenic atmospheric aerosols (Manara et al., 2020, 2019a, 2017a), which have shown a significant increase

until the end of 1970s/beginning of 1980s, followed by a marked decrease during the next decades (Manara et al., 2019b). The same analyses have, however, also allowed highlighting a role of TCC, which has partially masked the decrease of SD/SSR in the first period, especially for SD that is more influenced by TCC than SSR, SD being almost only driven by the direct radiation component. The comparison of the SD records presented by Manara et al. (2015) with corresponding TCC records over the common period (1951-2013) confirms relevant differences among TCC and SD in the long-term behaviour: considering an average Italian record (Manara et al., 2016b), SD does in fact not increase where TCC shows a very clear decrease (1951-1990), whereas SD increases (p-value ≤ 0.05) in the subsequent period where TCC does not show significant trend. Differently, the influence of TCC on SD is strong if the interannual variability is considered (the correlation between the TCC and SD residuals from a 11-year window, 3-year standard deviation Gaussian low-pass filter is about -0.89 over the whole period). The new Italian TCC dataset allows, therefore, to confirm that the long-term evolution of SD in Italy was only partially influenced by changes in TCC and that changes in anthropogenic atmospheric aerosols turn probably out to be the most relevant factor: if TCC had not changed over the investigated period, probably the dimming until about 1980 would have been stronger, whereas the following brightening would have been weaker. The effect would have been a more balanced strength of the changes in the dimming and in the brightening periods producing a more similar pattern to the one observed by removing the TCC contribution for SSR (Manara et al., 2016a, 2019a).

6. Conclusions

This paper presented a new quality-checked and homogenised dataset of TCC series over the Italian territory for the 1951-2018 period and studied the variability and trends of the obtained regional series. This dataset, with respect to the one presented by Maugeri et al., (2001), was updated to 2018 and thereby extended by more than 20 years of data compared to the previous one. The new dataset includes also a much higher number of stations with respect to the 35 ones considered by Maugeri et al., (2001). Moreover, with respect to the previous dataset, each daily synoptic observation was subjected to a detailed quality check procedure in order to identify and remove non-climatic signals and short missing periods were filled starting from neighbouring series.

The first information provided by the new TCC dataset concerned the diurnal cycle of cloud cover in northern and southern Italy and in selected elevation belts. This cycle highlighted the relevant role of convection induced by solar radiation that, as expected, is more relevant at medium and high-elevation stations and in summer. The second information provided concerned the annual TCC cycle that has a strong minimum in July and a maximum during winter for the southern Italian

stations, whereas the northern Italian stations exhibit a more complex behaviour with strong differences between low-elevation stations (minimum in July and maximum from November to May) and high-elevation stations (minimum in winter and maximum in May). The low amount of TCC in summer is a result of the northward shift of the subtropical high pressure systems towards the Italian territory.

The last and most relevant information provided by the new TCC dataset concerned long-term trends and decadal time-scale variability: the seasonal and annual TCC series were characterized by a significant negative trend over the whole considered period comparable in intensity for both regions. This temporal trend was mainly due to strong decrease observed during the 1951-1990 period. Even if small differences between northern and southern Italy can be observed, the two regions exhibited a coherent behaviour with respect to long-term trends and decadal time-scale variability. Moreover, they exhibited a reasonable agreement in terms of year-to-year variability. This aspect suggests that the causes of variability and trends of the Italian TCC are related more to large-scale factors rather than to local-scale changes. The obtained TCC series show an opposite behaviour with respect to SLP and 500-hPa geopotential height showing the same behaviour especially during the 1970s and the 1980s, where the most significant decrease/increase in TCC/SLP and 500-hPa geopotential height is concentrated. Moreover, TCC seems more linked to SLP during winter, spring and autumn while TCC seems more linked to 500-hPa geopotential height during summer, therefore, more linked to higher temperatures.

The role of changes in the atmospheric circulation on the Italian TCC records has to be investigated more in detail in the future. This requires clustering the data according to weather types as e.g. performed by Sfîcă et al., (2021): their analysis supports the idea that the westerly circulation shifted northward in Europe, inducing a decrease of the cloud cover amount. This shift is accompanied by a more frequent occurrence of blocking conditions in Eastern Europe (especially in winter).

An important aspect that should also be considered is the role of aerosol changes on cloud formation and lifetime. Aerosols can in fact influence clouds in different ways depending on the levels of pollution.

Also the relation between temperature and cloud cover trends have to be investigated more in detail: at global scale, the mid tropospheric temperature increase has been found to be connected with the observed decrease in cloud cover due to the Clausius-Clapeyron relation (Mendoza et al., 2021). Moreover, if the shortwave reflection effect of the cloud cover is dominant on a global scale, this relation leads to a predominant positive feedback: if the temperature increases as under the current

warming, the cloud cover decreases and more solar irradiance reaches the surface increasing the temperature even more (Donohoe et al., 2014; Hänsel et al., 2022).

Finally, the new TCC dataset confirmed that the long-term evolution of SD/SSR in Italy was only partially influenced by changes in TCC: if TCC would not have changed over the investigated period, probably the rather low dimming before about 1980 would have been stronger, whereas the following rather strong brightening would have been weaker determining a more similar strength of the changes in the dimming and the brightening periods.

The dataset presented in this paper adds, on one side, new information to the knowledge of the evolution of the TCC over the investigated area, an information not available in literature for the considered period, and, on the other side, it provides a product useful to validate satellite (Pfeifroth et al., 2018) and reanalysis TCC fields. Moreover, it provides new information to better understand the underlying causes of the variability and trends of surface solar radiation (Manara et al., 2016a) and sunshine duration (Manara et al., 2015), which have shown a pronounced dimming and brightening signal over Italy. Finally, the methodology we used to set up the quality-checked dataset of Italian TCC records could be applied in the future to other national TCC data sets.

Acknowledgements

We received the Italian Air Force ("Servizio Meteorologico dell'Areonautica Militare Italiana") SYNOP data for the 1951-2014 period in the frame of an agreement between the Italian Air Force and the Italian National Research Council (for data access refer to https://clima.meteoam.it/istruzioni.php). We thank the National Oceanic and Atmospheric Administration (NOAA) for the SYNOP data for the 2015-2018 period (refer to https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00532 for data access). We thank the Climate Data Store for the ECMWF-ERA5 monthly data of surface level pressure (Bell et al., 2020a; Hersbach et al., 2019a) and 500 hPa Geopotential height (Bell et al., 2020b; Hersbach et al., 2019b) (for data access refer to https://cds.climate.copernicus.eu/cdsapp#!/home).

Veronica Manara was supported by the "Ministero dell'Università e della Ricerca" of Italy [grant FSE – REACT EU, DM 10/08/2021 n. 1062].

References

Aguilar, E., Auer, I., Brunet, M., Peterson, T.C., Wieringa, J., 2003. Guidelines on climate metadata and homogenization. World Clim. Program. Data Monit. WCDMP-No. 53, WMO-TD No. 1186 1186, 50.

Auchmann, R., Brönnimann, S., Breda, L., Bühler, M., Spadin, R., Stickler, A., 2012. Extreme climate, not extreme weather: The summer of 1816 in Geneva, Switzerland. Clim. Past 8, 325–335. https://doi.org/10.5194/cp-8-325-2012

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Thépaut, J.-N., 2020a. ERA5 monthly averaged data on single levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 05-09-2022).

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Thépaut, J.-N., 2020b. ERA5 monthly averaged data on pressure levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 05-09-2022).

Biggs, T.W., Scott, C.A., Rajagopalan, B., Turral, H.N., 2007. Trends in solar radiation due to clouds and aerosols, southern India, 1952-1997. Int. J. Clim. 27, 1505–1518. https://doi.org/10.1002/joc.1487

Brunetti, M., Maugeri, M., Monti, F., Nanni, T., 2006. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. Int. J. Climatol. 26, 345–381. https://doi.org/10.1002/joc.1251

Chernokulsky, A., Esau, I., 2019. Cloud cover and cloud types in the Eurasian Arctic in 1936–2012. Int. J. Climatol. 39, 5771–5790. https://doi.org/10.1002/joc.6187

Christidis, N., Stott, P.A., 2015. Changes in the geopotential height at 500 hPa under the influence of external climatic forcings. Geophys. Res. Lett. 42, 10798–10806. https://doi.org/10.1002/2015GL066669

Craddock, J.M., 1979. Methods of comparing annual rainfall records for climatic purposes. Weather.

Crespi, A., Brunetti, M., Lentini, G., Maugeri, M., 2018. 1961–1990 high-resolution monthly precipitation climatologies for Italy. Int. J. Climatol. 38, 878–895. https://doi.org/10.1002/joc.5217

Dai, A., Karl, T.R., Sun, B., Trenberth, K.E., 2006. Recent trends in cloudiness over the United States: A tale of monitoring inadequacies. Bull. Am. Meteorol. Soc. 87, 597–606. https://doi.org/10.1175/BAMS-87-5-597

638 Donohoe, A., Armour, K.C., Pendergrass, A.G., Battisti, D.S., 2014. Shortwave and longwave
639 radiative contributions to global warming under increasing CO2. Proc. Natl. Acad. Sci. U. S.
640 A. 111, 16700–16705. https://doi.org/10.1073/pnas.1412190111
641 Eastman, R., Warren, S., 2013. A 39-Yr Survey of Cloud Changes from Land Stations Worldwide
642 1971 – 2009 : Long-Term Trends, Relation to Aerosols, and Expansion of the Tropical Belt. J.
643 Clim. 26, 1286–1303. https://doi.org/10.1175/JCLI-D-12-00280.1
644 Fan, J., Zhang, Y., Li, Z., Hu, J., Rosenfeld, D., 2020. Urbanization-induced land and aerosol
645 impacts on sea-breeze circulation and convective precipitation. Atmos. Chem. Phys. 20,
646 14163–14182. https://doi.org/10.5194/acp-20-14163-2020
647 Founda, D., Nastos, P.T., Pierros, F., Kalimeris, A., 2018. Historical observations of cloudiness
648 (1882 – 2012) over a large urban area of the eastern Mediterranean (Athens). Theor. Appl.
649 Climatol. 1–13.
650 Free, M., Sun, B., 2014. Trends in U.S. total cloud cover from a homogeneity-adjusted dataset. J.
651 Clim. 27, 4959–4969. https://doi.org/10.1175/JCLI-D-13-00722.1
652 Fugazza, D., Manara, V., Senese, A., Diolaiuti, G., Maugeri, M., 2021. Snow cover variability in
653 the greater alpine region in the modis era (2000–2019). Remote Sens. 13, 2945.
654 https://doi.org/10.3390/rs13152945
655 Hafez, Y.Y., Almazroui, M., 2014. Recent Study of Anomaly of Global Annual Geopotential
656 Height and Global Warming. Atmos. Clim. Sci. 04, 347–357.
657 https://doi.org/10.4236/acs.2014.43035
658 Hahn, C., Warren, S., London, J., 1995. The effect of moonlight on observation of cloud cover at
659 night, and application to cloud climatology. J. Clim. 8, 1429–1446.
660 Hänsel, S., Hoy, A., Brendel, C., Maugeri, M., 2022. Record summers in Europe: Variations in
661 drought and heavy precipitation during 1901–2018. Int. J. Climatol. 2018, 1–23.
662 https://doi.org/10.1002/joc.7587
663 Henderson-Sellers, A., 1992. Continental cloudiness changes this century. GeoJournal 27, 255–262.
664 https://doi.org/10.1007/BF02482666
665 Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J.,
666 Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N.,
667 2019a. ERA5 monthly averaged data on single levels from 1959 to present. Copernicus
668 Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 05-09-2022).
669 https://doi.org/10.24381/cds.f17050d7
670 Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J.,
671 Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N.,

672 2019b. ERA5 monthly averaged data on pressure levels from 1959 to present. Copernicus
673 Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on 05-09-2022).
674 https://doi.org/10.24381/cds.6860a573
675 Jones, P.D., Jonsson, T., Wheeler, D., 1997. Extension to the North Atlantic Oscillation using early
676 instrumental pressure observations from gibraltar and south-west Iceland. Int. J. Climatol. 17,
677 1433–1450. https://doi.org/10.1002/(sici)1097-0088(19971115)17:13<1433::aid-
678 joc203>3.3.co;2-g
679 Jovanovic, B., Collins, D., Braganza, K., Jakob, D., Jones, D.A., 2011. A high-quality monthly total
680 cloud amount dataset for Australia. Clim. Change 108, 485–517.
681 https://doi.org/10.1007/s10584-010-9992-5
682 Kambezidis, H.D., Kampezidou, S.I., Kampezidou, D., 2021. Mathematical determination of the
683 upper and lower limits of the diffuse fraction at any site. Appl. Sci. 11.
684 https://doi.org/10.3390/app11188654
685 Khlebnikova, E.I., Sall, I.A., 2009. Peculiarities of climatic changes in cloud cover over the Russian
686 Federation. Russ. Meteorol. Hydrol. 34, 411–417. https://doi.org/10.3103/s1068373909070012
687 Manara, V., Bassi, M., Brunetti, M., Cagnazzi, B., Maugeri, M., 2019a. 1990–2016 surface solar
688 radiation variability and trend over the Piedmont region (northwest Italy). Theor. Appl.
689 Climatol. 136, 849–862. https://doi.org/10.1007/s00704-018-2521-6
690 Manara, V., Beltrano, M.C., Brunetti, M., Maugeri, M., Sanchez-Lorenzo, A., Simolo, C., Sorrenti,
691 S., 2015. Sunshine duration variability and trends in Italy from homogenized instrumental time
692 series (1936-2013). J. Geophys. Res. Atmos. 120, 3622–3641.
693 https://doi.org/10.1002/2014JD022560
694 Manara, V., Brunetti, M., Celozzi, A., Maugeri, M., Sanchez-Lorenzo, A., Wild, M., 2016a.
695 Detection of dimming/brightening in Italy from homogenized all-sky and clear-sky surface
696 solar radiation records and underlying causes (1959-2013). Atmos. Chem. Phys. 16, 11145–
697 11161. https://doi.org/10.5194/acp-16-11145-2016
698 Manara, V., Brunetti, M., Gilardoni, S., Landi, T.C., Maugeri, M., 2019b. 1951–2017 changes in
699 the frequency of days with visibility higher than 10 km and 20 km in Italy. Atmos. Environ.
700 214, 116861. https://doi.org/10.1016/j.atmosenv.2019.116861
701 Manara, V., Brunetti, M., Maugeri, M., 2016b. Reconstructing sunshine duration and solar radiation
702 long-term evolution for Italy: a challenge for quality control and homogenization precedures,
703 in: 14th IMEKO TC10 Workshop Technical Diagnostics - New Perspectives in Measurements,
704 Tools and Techniques for System's Reliability, Maintainability and Safety. pp. 13–18.
705 Manara, V., Brunetti, M., Maugeri, M., Sanchez-Lorenzo, A., Wild, M., 2017a. Sunshine duration

and global radiation trends in Italy (1959-2013): To what extent do they agree? J. Geophys. Res. 122, 4312–4331. https://doi.org/10.1002/2016JD026374

Manara, V., Brunetti, M., Maugeri, M., Sanchez-Lorenzo, A., Wild, M., 2017b. Homogenization of a surface solar radiation dataset over Italy, in: Radiation Processes in the Atmosphere And Ocean (IRS 2016) - AIP Conference Proceeding. AIP Publishing: Melville, NY, USA, Auckland, New Zealand, 12-22 April 2016, pp. 090004-1-090004–4. https://doi.org/10.1063/1.4975544

Manara, V., Stocco, E., Brunetti, M., Diolaiuti, G.A., Fugazza, D., Pfeifroth, U., Senese, A., Trentmann, J., Maugeri, M., 2020. Comparison of surface solar irradiance from ground observations and satellite data (1990–2016) over a complex orography region (piedmont—Northwest Italy). Remote Sens. 12, 1–26. https://doi.org/10.3390/rs12233882

Maugeri, M., Bagnati, Z., Brunetti, M., Nanni, T., 2001. Trends in Italian total cloud amount, 1951-1996. Geophys. Res. Lett. 28, 4551–4554. https://doi.org/10.1029/2001GL013754

Mendoza, V., Pazos, M., Garduño, R., Mendoza, B., 2021. Thermodynamics of climate change between cloud cover, atmospheric temperature and humidity. Sci. Rep. 11, 1–11. https://doi.org/10.1038/s41598-021-00555-5

New, M., Hulme, M., Jones, P., 2000. Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. J. Clim. 13, 2217–2238. https://doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2

Pfeifroth, U., Bojanowski, J.S., Clerbaux, N., Manara, V., Sanchez-Lorenzo, A., Trentmann, J., Walawender, J.P., Hollmann, R., 2018. Satellite-based trends of solar radiation and cloud parameters in Europe, in: Advances in Science and Research. pp. 31–37. https://doi.org/10.5194/asr-15-31-2018

Pörtner, H., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B., 2022. IPCC, 2022: Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/9781009325844

Preisendorfer, R.W., 1988. Principal component analysis in meteorology and oceanography. New York. https://doi.org/10.1029/JD093iD09p10815

Ramanathan, V., Cess, R.D., Harrison, E.F., Minnis, P., Barkstrom, B.R., Ahmad, E., Hartmann, D., 1989. Cloud-Radiative Forcing and Climate: Results from the Earth Radiation Budget Experiment. Sciences (New. York). 243, 57–63.

Ramanathan, V., Crutzen, P.J., Kiehl, J.T., Rosenfeld, D., 2001. Aerosols, Climate, and the

740 Hydrological Cycle. Science (80-.). 294, 2119–2124. https://doi.org/10.1126/science.1064034
741 Rosenfeld, D., Lohmann, U., Raga, G.B., O'Dowd, C.D., Kulmala, M., Fuzzi, S., Reissell, A.,
742 Andreae, M.O., 2008. Flood or drought: How do aerosols affect precipitation? Science (80-.).
743 321, 1309–1313. https://doi.org/10.1126/science.1160606
744 Sanchez-Lorenzo, A., Calbó, J., Brunetti, M., Deser, C., 2009. Dimming/brightening over the
745 Iberian Peninsula: Trends in sunshine duration and cloud cover and their relations with
746 atmospheric circulation. J. Geophys. Res. 114, D00D09.
747 https://doi.org/10.1029/2008JD011394
748 Sanchez-Lorenzo, A., Calbó, J., Wild, M., 2012. Increasing cloud cover in the 20th century: review
749 and new findings in Spain. Clim. Past 8, 1199–1212. https://doi.org/10.5194/cp-8-1199-2012
750 Sanchez-Lorenzo, A., Enriquez-Alonso, A., Calbó, J., González, J.A., Wild, M., Folini, D., Norris,
751 J.R., Vicente-Serrano, S.M., 2017. Fewer clouds in the Mediterranean: Consistency of
752 observations and climate simulations. Sci. Rep. 7, 1–10. https://doi.org/10.1038/srep41475
753 Sanchez-Lorenzo, A., Wild, M., 2012. Decadal variations in estimated surface solar radiation over
754 Switzerland since the late 19th century. Atmos. Chem. Phys. 12, 8635–8644.
755 https://doi.org/10.5194/acp-12-8635-2012
756 Schwander, M., Brönnimann, S., Delaygue, G., Rohrer, M., Auchmann, R., Brugnara, Y., 2017.
757 Reconstruction of Central European daily weather types back to 1763. Int. J. Climatol. 37, 30–
758 44. https://doi.org/10.1002/joc.4974
759 Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. J. Am. Stat.
760 Assoc. 63, 1379–1389. https://doi.org/10.1080/01621459.1968.10480934
761 Sfîcă, L., Beck, C., Nita, A.I., Voiculescu, M., Birsan, M.V., Philipp, A., 2021. Cloud cover
762 changes driven by atmospheric circulation in Europe during the last decades. Int. J. Climatol.
763 41, E2211–E2230. https://doi.org/10.1002/joc.6841
764 Sneyers, R., 1992. On the use of statistical analysis for the objective determination of climate
765 change. Meteorol. Zeitschrift 1, 247–256.
766 Sun, B., Free, M., Yoo, H.L., Foster, M.J., Heidinger, A., Karlsson, K.G., 2015. Variability and
767 trends in U.S. cloud cover: ISCCP, PATMOS-x, and CLARA-A1 compared to homogeneity-
768 adjusted weather observations. J. Clim. 28, 4373–4389. https://doi.org/10.1175/JCLI-D-14-
769 00805.1
770 Theil, H., 1950. A rank-invariant method of linear and polynomical regression analysis, in:
771 Proceedings of the Royal Netherlands Academy of Sciences. p. Part I: 386-392.
772 Tuel, A., Eltahir, E.A.B., 2020. Why Is the Mediterranean a Climate Change Hot Spot? J. Clim. 33,
773 5829–5843. https://doi.org/10.1175/JCLI-D-19-0910.1

Vergari, S., Meli, G., 2020. Stima automatica della copertura nuvolosa. Riv. di Meteorol. Areonautica Luglio-Set, 39–48.

Wild, M., 2009. Global dimming and brightening: A review. J. Geophys. Res. 114, D00D16. https://doi.org/10.1029/2008JD011470

Wilks, D.S., 1995. Statistical methods in the atmospheric sciences. Int. Geophys. Ser. 59, 464. https://doi.org/10.1007/s13398-014-0173-7.2

WMO, 2008. Observation of clouds, in: Guide to Meteorological Instruments and Methods of Observation. p. I.15 1-I.15-12.

Xia, X., 2010. Spatiotemporal changes in sunshine duration and cloud amount as well as their relationship in China during 1954-2005. J. Geophys. Res. Atmos. 115, 1–13. https://doi.org/10.1029/2009JD012879

You, Q., Jiao, Y., Lin, H., Min, J., Kang, S., Ren, G., Meng, X., 2014. Comparison of NCEP/NCAR and ERA-40 total cloud cover with surface observations over the Tibetan Plateau. Int. J. Climatol. 34, 2529–2537. https://doi.org/10.1002/joc.3852

Zelinka, M.D., Hartmann, D.L., 2010. Why is longwave cloud feedback positive? J. Geophys. Res. Atmos. 115, 1–16. https://doi.org/10.1029/2010JD013817

Zhou, C., Zelinka, M.D., Klein, S.A., 2016. Impact of decadal cloud variations on the Earth's energy budget. Nat. Geosci. 9, 871–874. https://doi.org/10.1038/ngeo2828

Zong, X.-M., Wang, P.-C., Xia, X.-A., 2013. Variability and Long-Term Trend of Total Cloud Cover in China Derived from ISCCP, ERA-40, CRU3, and Ground Station Datasets. Atmos. Ocean. Sci. Lett. 6, 133–137. https://doi.org/10.1080/16742834.2013.11447069

Figure captions

Figure 1: Spatial distribution of the stations selected after the quality controls (black points) for each synoptic observation (see Section 3.4). The blue and red stars represent the northern and southern Italy grid-points, respectively as clustered by a Principal Component Analysis (see Section 3.5). The panels also show the orography of the region.

Figure 2: Main steps of the gap-filling procedure.

Figure 3: Number of records per year of the final quality-checked and homogenised dataset (series with at least 20 years of available data before the gap-filling procedure). Red line: 0 UTC; black line: 3 UTC; grey line: 6 UTC; orange line: 9 UTC; brown line: 12 UTC; blue line: 15 UTC; green line: 18 UTC; pink line: 21 UTC.

Figure 4: 1961-1990 seasonal and annual mean TCC daily cycles for northern (N) and southern (S) Italy and for the different orographic elevation belts (low (L), medium (M) and high (H) elevations) at the eight synoptic observation times.

Figure 5: 1961-1990 mean TCC annual cycles for the different regions and orographic elevation belts. Each panel is referred to one synoptic observation time.

Figure 6: Average seasonal and annual northern (thin line) Italian TCC records, plotted together with an 11-year window, 3-year standard deviation Gaussian low-pass filter (bold line). The grey area represents the standard deviation of the grid-point records used to obtain the average records. The series are expressed as additive deviations from the 1961-1990 means.

Figure 7: Average seasonal and annual southern (thin line) Italian TCC records, plotted together with an 11-year window, 3-year standard deviation Gaussian low-pass filter (bold line). The grey area represents the standard deviation of the grid-point records used to obtain the average records. The series are expressed as additive deviations from the 1961-1990 means.

Figure 8: Running trend analysis for seasonal and annual northern (left column) and southern (right column) Italian TCC records. The y and x axes represent the length and the first year of the period under analysis, respectively, while the coloured pixels show the trends expressed in oktas per decade with a significance level of $p \leq 0.05$ (large squares) and $p > 0.05$ (small squares).

Figure 9: Average seasonal and annual northern (black thin line) and southern (red thin line) Italian half TCC interquartile range records, plotted together with an 11-year window, 3-year standard deviation Gaussian low-pass filter (bold lines).

Figure 10: Average seasonal and annual Italian TCC (black thin line), SLP (red thin line) and z500/100 (blue thin line) records plotted together with an 11-year window, 3-year standard deviation Gaussian low-pass filter (bold lines). The series are expressed as additive deviations from the 1961-1990 means. The grey lines highlight the period in which the decrease of the series is more pronounced.

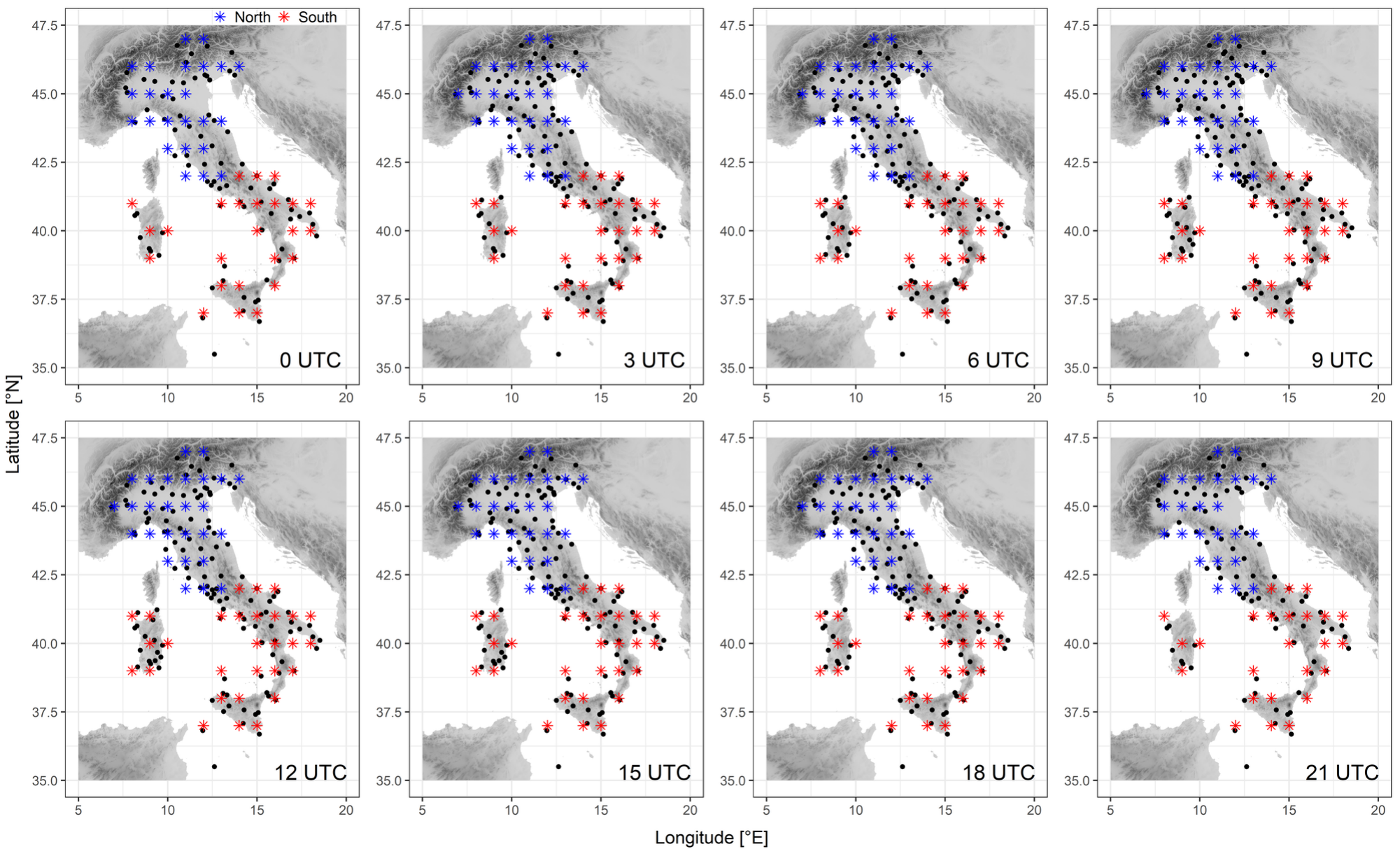
Figure 11: Annual and seasonal correlation maps between the Italian TCC mean anomaly series and the ERA5 500hPa geopotential height records over the 40W-50E and 25N-70N domain.

844 ***Table captions***

845 *Table 1: TCC trends expressed as oktas per decade over three periods (1951-2018; 1951-1990;*
846 *1991-2018) for northern and southern Italy. The trends are provided with an estimation of their*
847 *standard deviations (half widths of 68% confidence intervals). Values are shown in roman for a*
848 *significance level of* $0.05 < p \leq 0.1$ *and in bold for a significance level of* $p \leq 0.05$*. For non-*
849 *significant trends, only the sign of the slope is given. The trends are estimated with the Theil-Sen*
850 *method, while the significance of the trends is evaluated with the Mann-Kendall non-parametric*
851 *test.*

852

853


North
South
0 UTC
3 UTC
6 UTC
9 UTC
12 UTC
15 UTC
18 UTC
21 UTC
Latitude [°N]
Longitude [°E]
47.5
45.0
42.5
40.0
37.5
35.0
5
10
15
20

The procedure should be executed for each of the eight daily synoptic observations and each missing day in each series (test series).

PART A: select the possible reference series

Consider each record (candidate reference series) with the entry of the day under consideration available.

↓

Select for each couple candidate reference and test series a window of +/- 30 days centered on the observation to be estimated and consider the same window also for the previous and the subsequent year. The maximum length of the considered window is equal to 182 days.

↓

Do the candidate reference and test series have in common at least 50% of available data in that window?

- NO → The considered series is not a possible reference series.
- YES → The considered series is a possible reference series.

↓

Calculate the total weight of each possible reference series as the product of distance, elevation, climatology and angular weight.

PART B: select the reference series to be used

Sort in descending order the weights of the possible reference series identified in PART A

↓

Select the seven series with the highest weight. They are the reference series that will be used in PART C to estimate the missing entry in the test series.

PART C: estimate the missing value in the test series

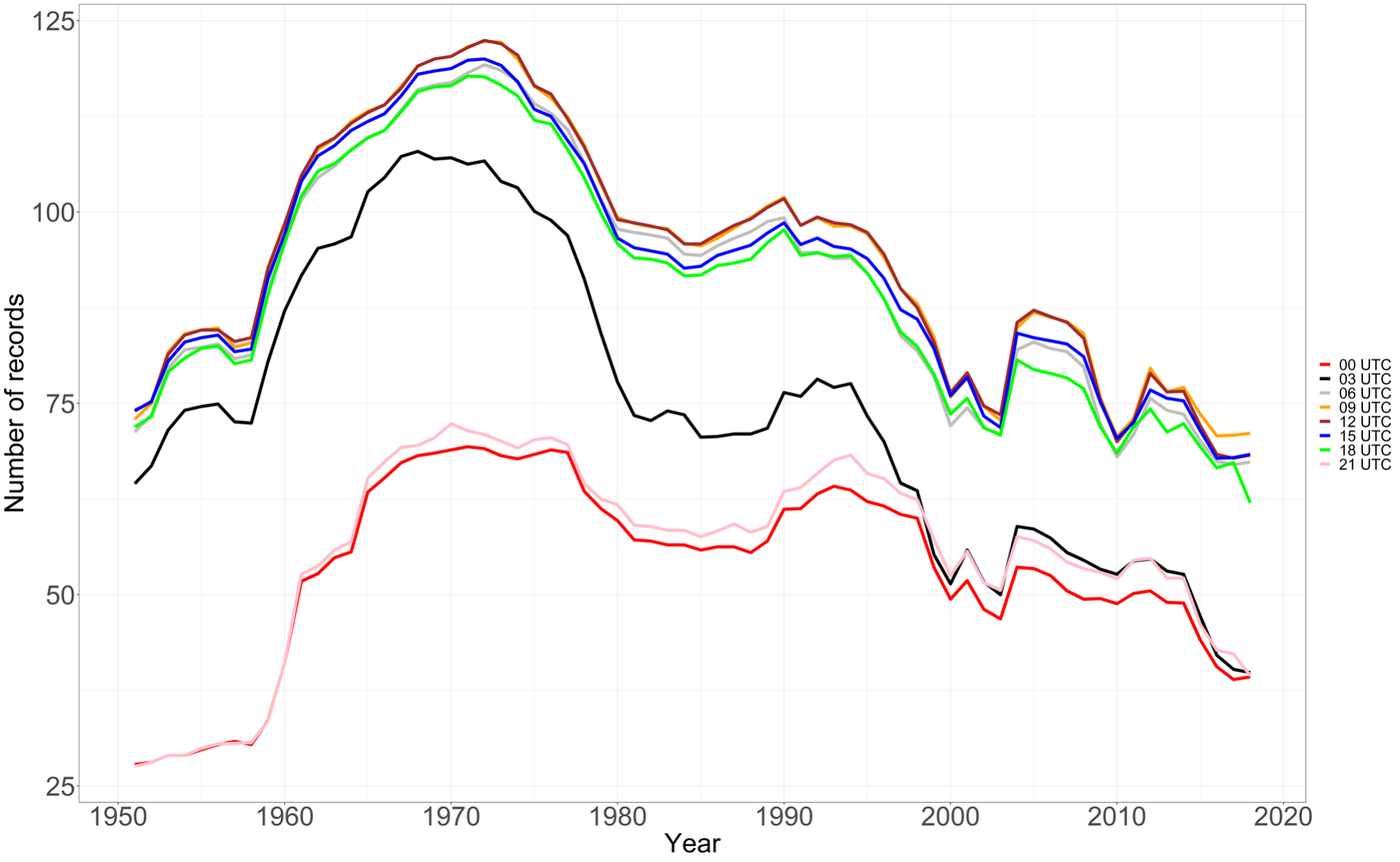
For each reference series selected in PART B calculate the percentile associated to the entry that is missing in the test series. The percentile is extracted by means of the data in the window selected in PART A.

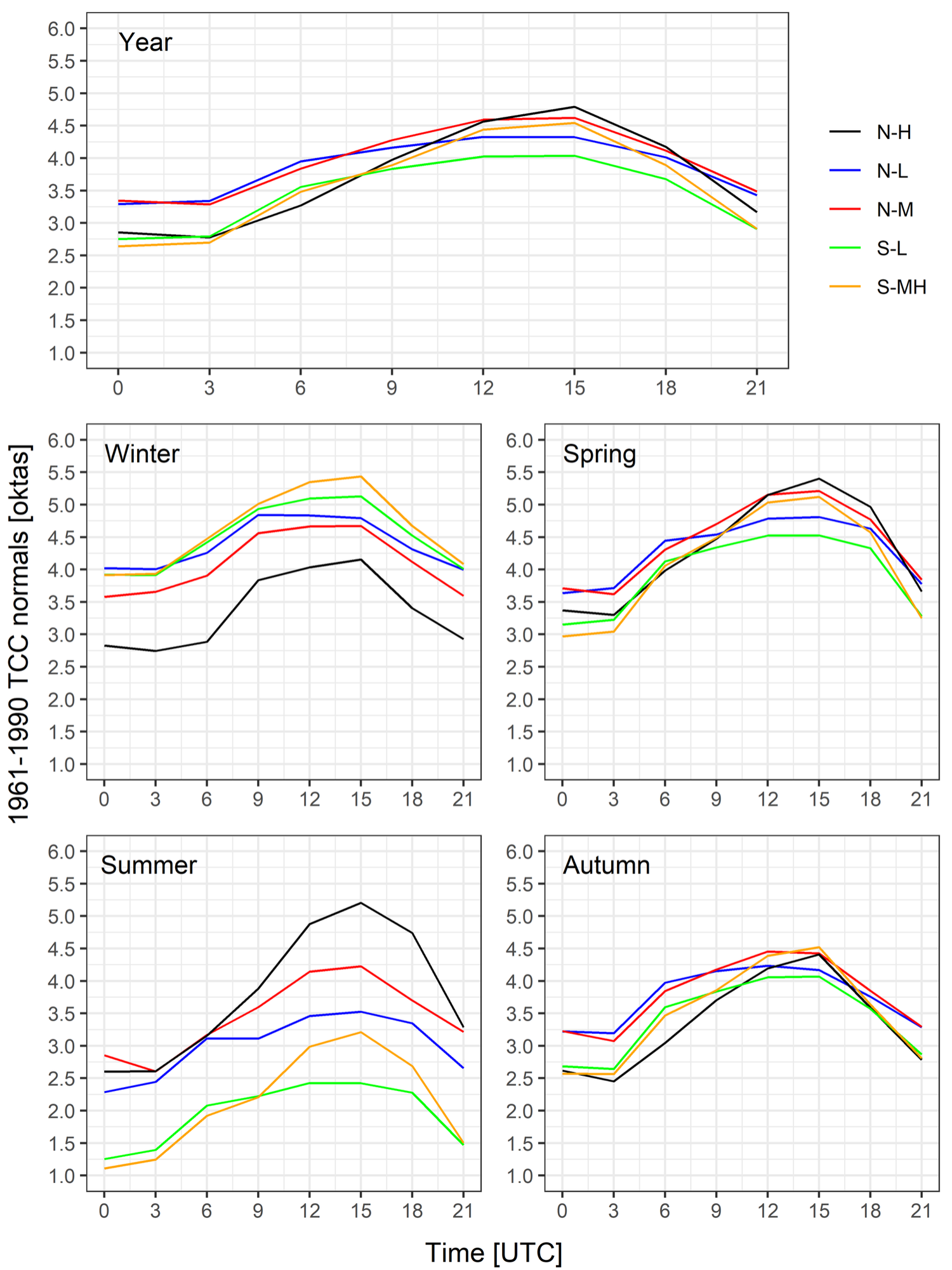
↓

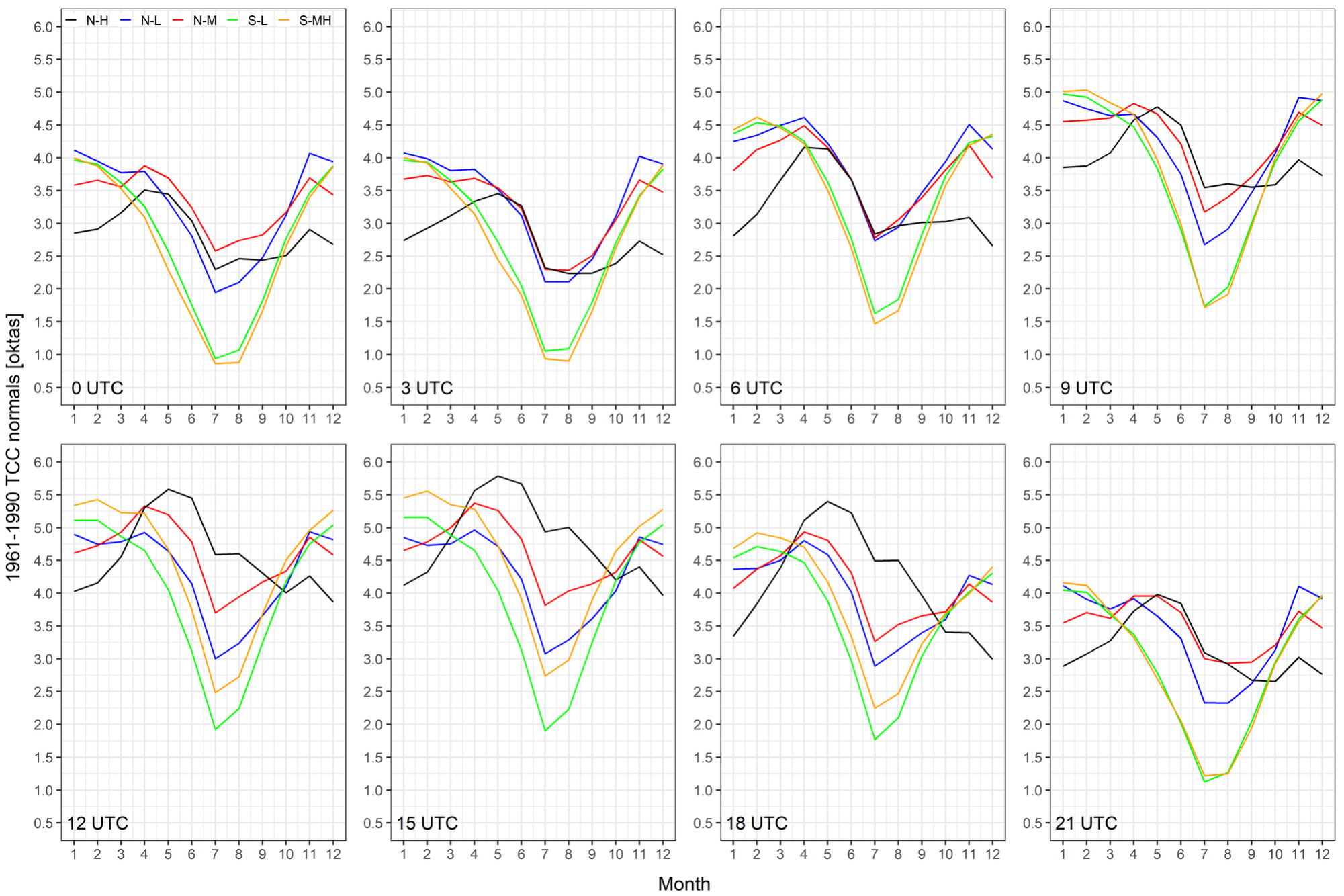
The obtained percentile for each reference series is used to calculate the candidate missing value in the test series. The missing value is calculated using the empirical distribution of the data in the window selected in PART A.

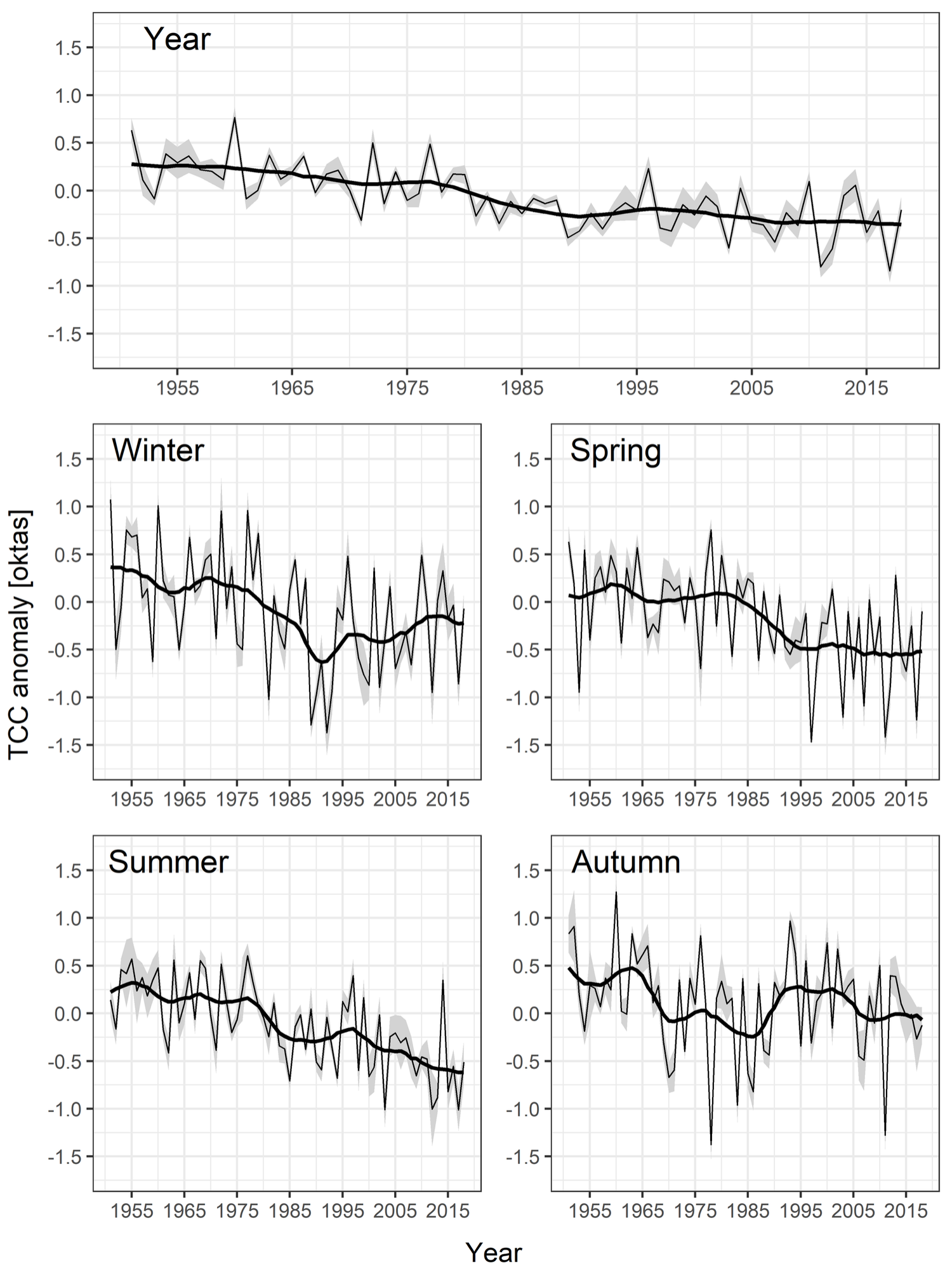
↓

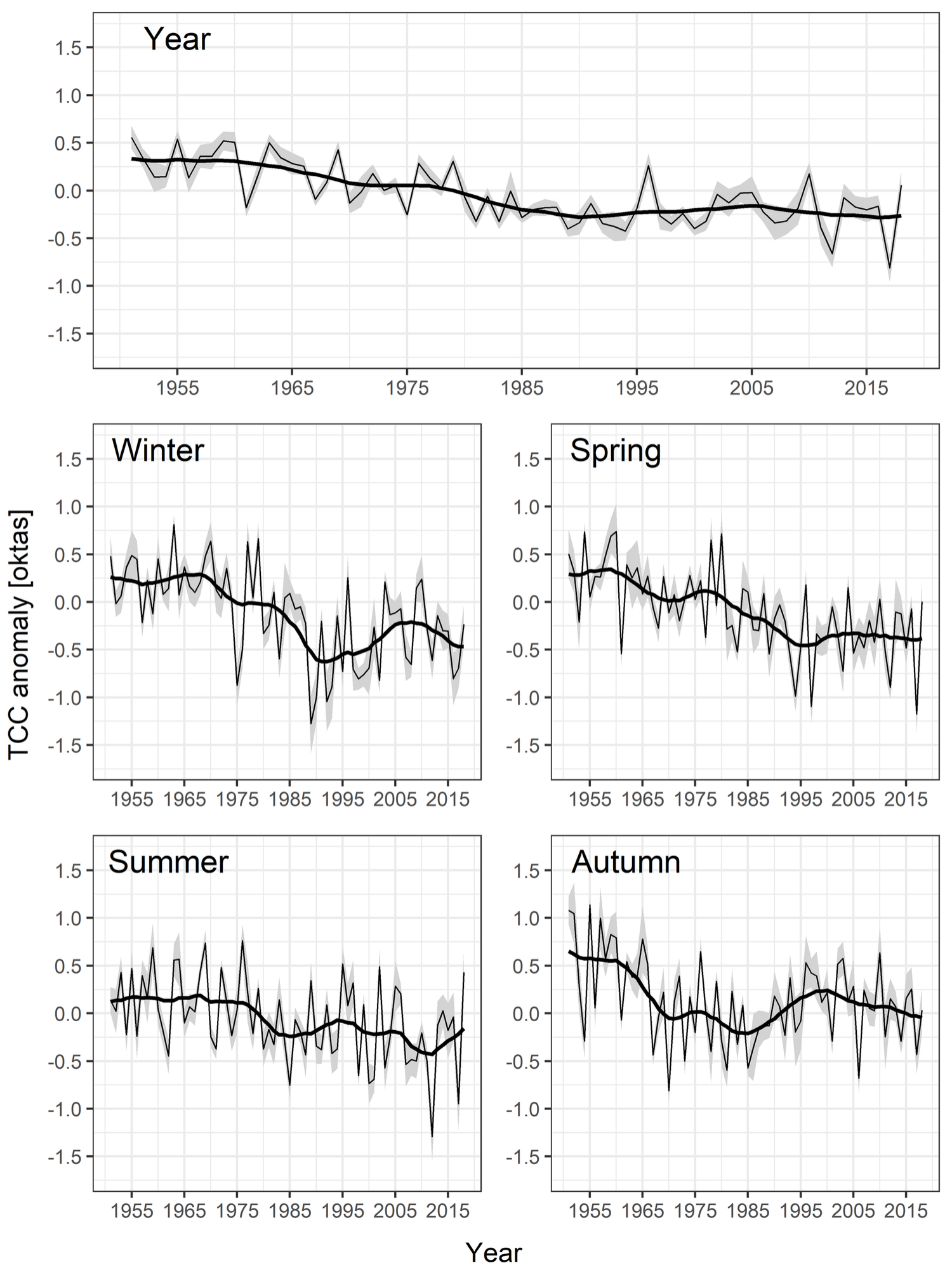
The missing value in the test series is calculated as the weighted mean of the seven estimated values.

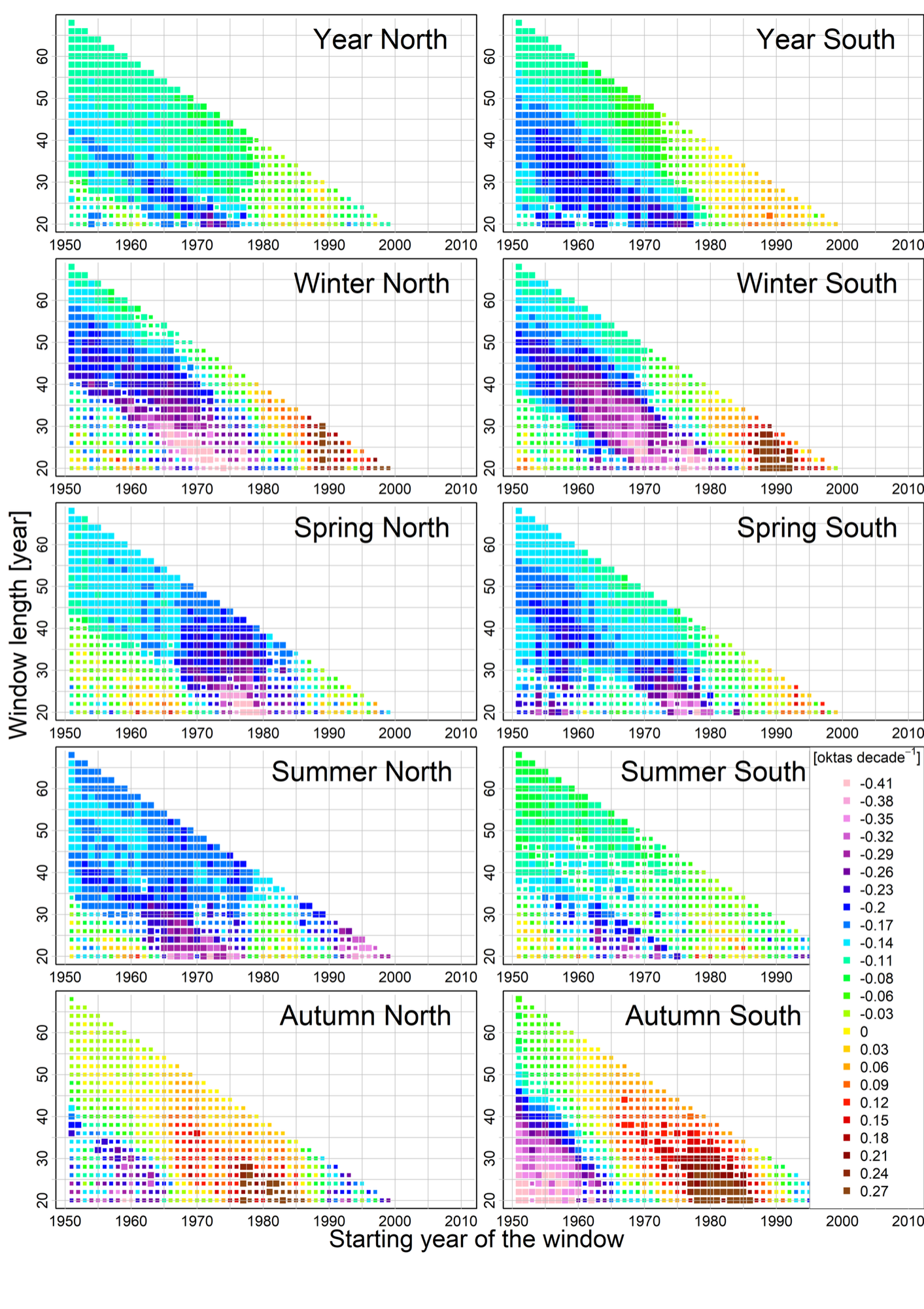

125
100
75
50
25
Number of records
1950
1960
1970
1980
1990
2000
2010
2020
Year
00 UTC
03 UTC
06 UTC
09 UTC
12 UTC
15 UTC
18 UTC
21 UTC

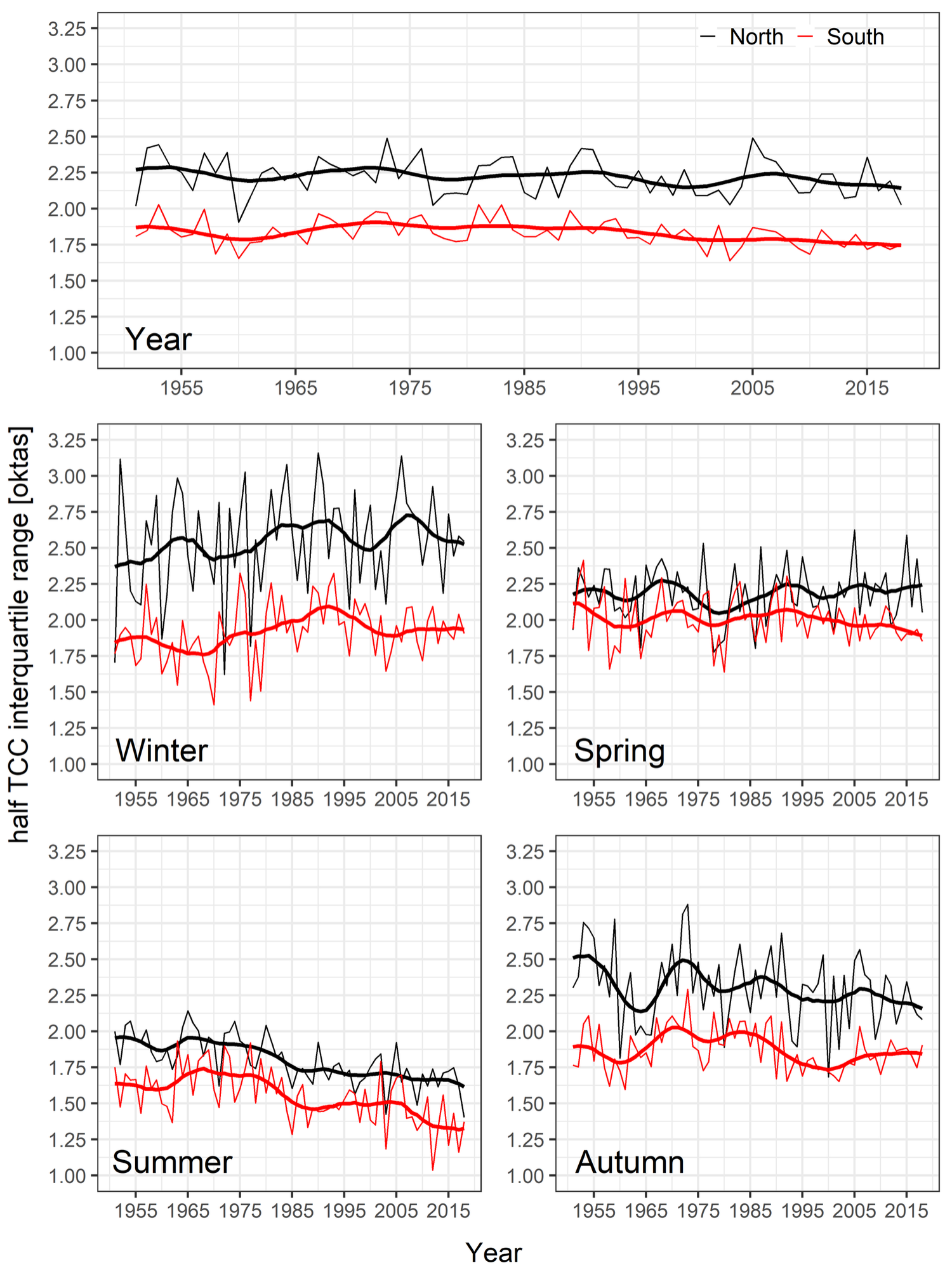

Year
Winter
Spring
Summer
Autumn
N-H
N-L
N-M
S-L
S-MH
1961-1990 TCC normals [oktas]
Time [UTC]
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
3
6
9
12
15
18
21

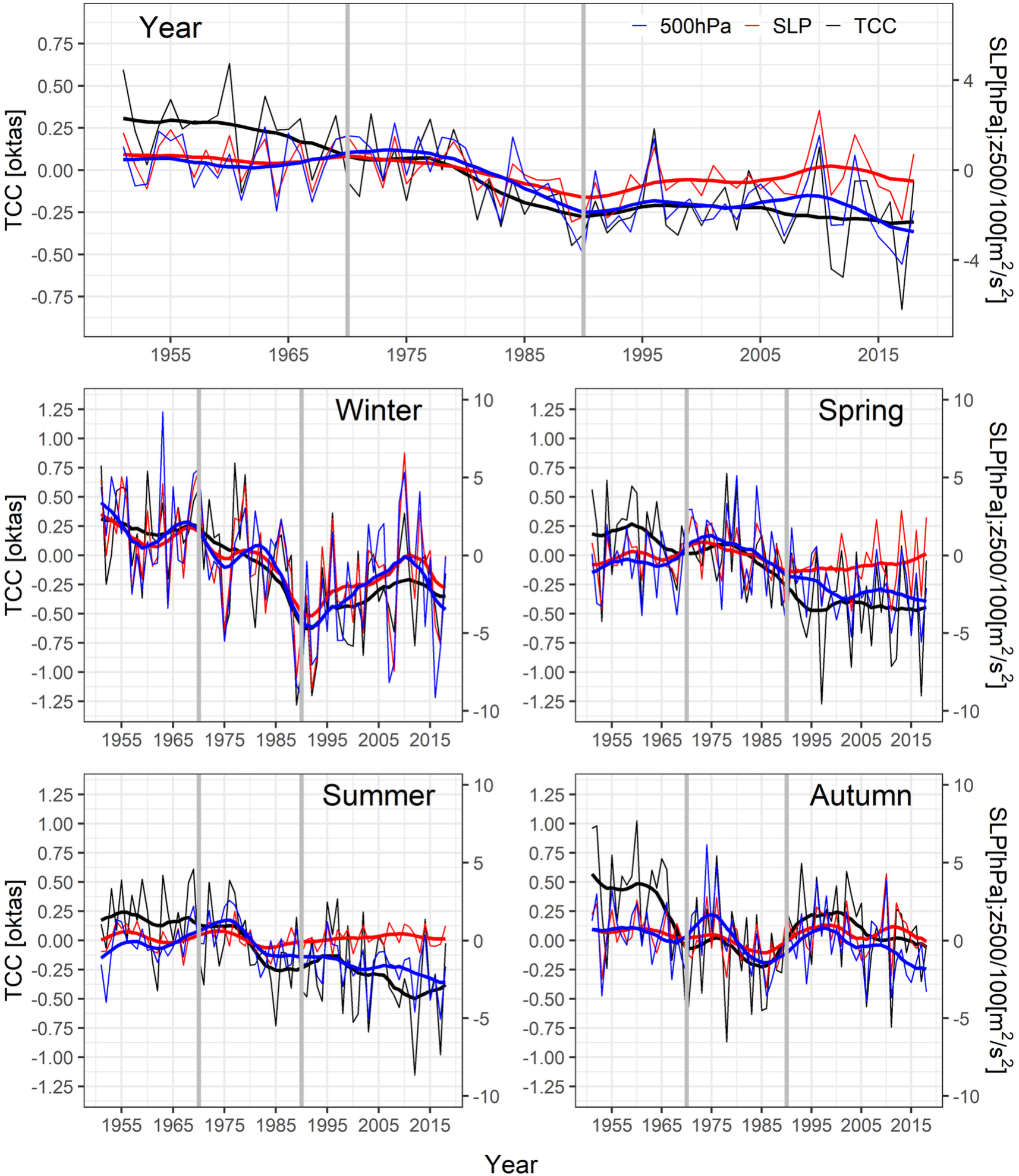

N-H
N-L
N-M
S-L
S-MH
0 UTC
3 UTC
6 UTC
9 UTC
12 UTC
15 UTC
18 UTC
21 UTC
1961-1990 TCC normals [oktas]
Month

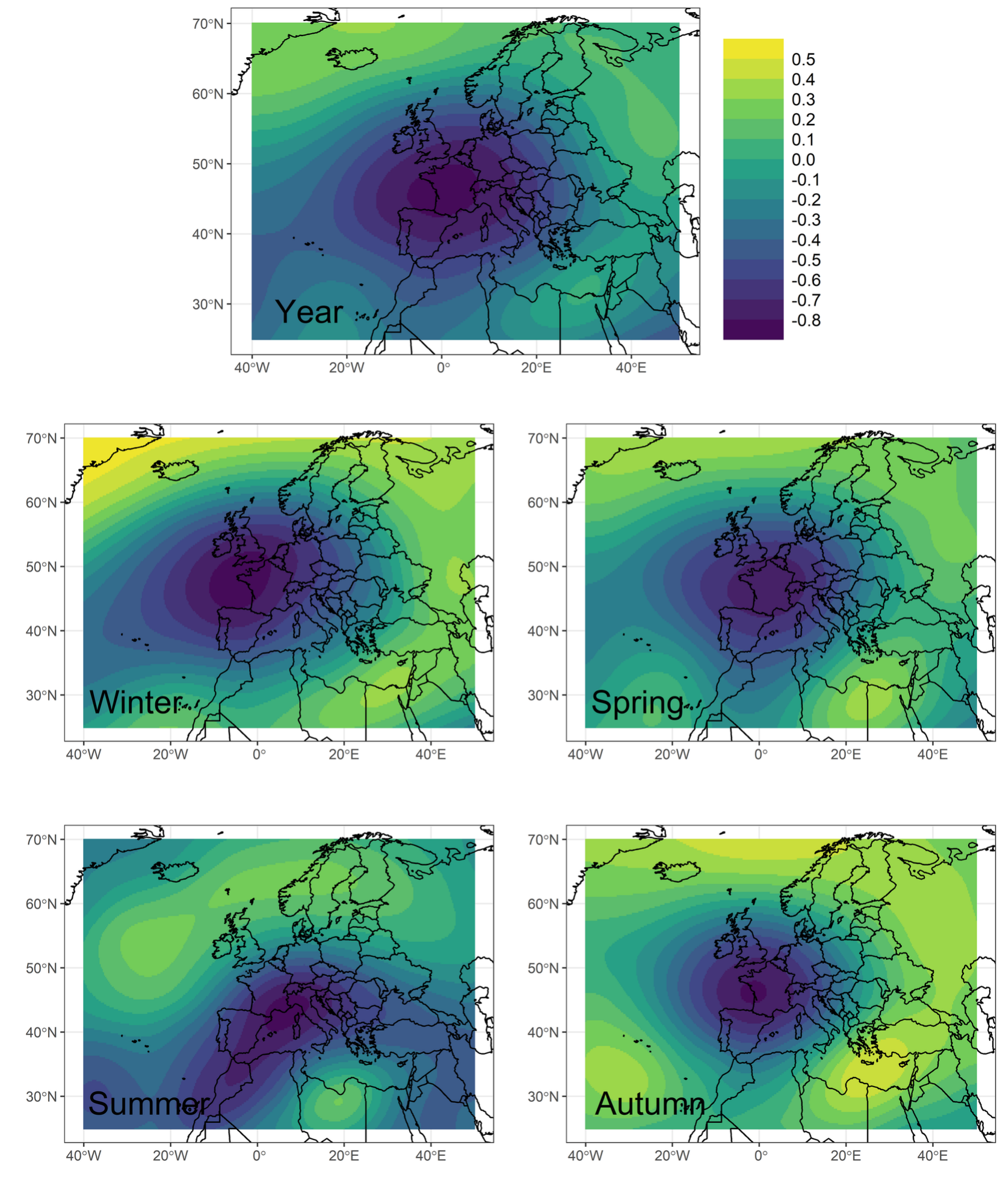

Year
Winter
Spring
Summer
Autumn
TCC anomaly [oktas]
Year
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1955
1965
1975
1985
1995
2005
2015


Year
Winter
Spring
Summer
Autumn
TCC anomaly [oktas]
Year
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1955
1965
1975
1985
1995
2005
2015


Year North
Year South
Winter North
Winter South
Spring North
Spring South
Summer North
Summer South
Autumn North
Autumn South
Window length [year]
Starting year of the window
[oktas decade^{-1}]
-0.41
-0.38
-0.35
-0.32
-0.29
-0.26
-0.23
-0.2
-0.17
-0.14
-0.11
-0.08
-0.06
-0.03
0
0.03
0.06
0.09
0.12
0.15
0.18
0.21
0.24
0.27


North
South
Year
Winter
Spring
Summer
Autumn
half TCC interquartile range [oktas]
Year
3.25
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
1955
1965
1975
1985
1995
2005
2015


Year
500hPa
SLP
TCC
TCC [oktas]
SLP[hPa];z500/100[m²/s²]
Winter
Spring
Summer
Autumn
Year


Year
Winter
Spring
Summer
Autumn
70°N
60°N
50°N
40°N
30°N
40°W
20°W
0°
20°E
40°E
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8

		Winter	Spring	Summer	Autumn	Year
1951-2018	North	**-0.11±0.04**	**-0.13±0.03**	**-0.16±0.02**	-	**-0.11±0.02**
	South	**-0.12±0.02**	**-0.13±0.02**	**-0.09±0.02**	**-0.07±0.03**	**-0.11±0.02**
1951-1990	North	*-0.18±0.08*	-	**-0.15±0.04**	**-0.17±0.08**	**-0.15±0.03**
	South	**-0.15±0.05**	**-0.15±0.05**	**-0.11±0.05**	**-0.27±0.06**	**-0.18±0.03**
1991-2018	North	+	-	*-0.23±0.10*	*-0.19±0.11*	-
	South	+	+	-	-	+