The Cryosphere, 18, 6027–6059, 2024
https://doi.org/10.5194/tc-18-6027-2024

The Cryosphere Open Access EGU

Temporal stability of a new 40-year daily AVHRR land surface temperature dataset for the pan-Arctic region

Sonia Dupuis[1,2], **Frank-Michael Göttsche**[3], **and Stefan Wunderle**[1,2]

[1]Institute of Geography, University of Bern, Bern, Switzerland
[2]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[3]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence: Sonia Dupuis (sonia.dupuis@unibe.ch)

Received: 24 March 2024 – Discussion started: 10 April 2024
Revised: 11 October 2024 – Accepted: 30 October 2024 – Published: 20 December 2024

Abstract. Land surface temperature (LST) has gained increased attention in cryospheric research. While various global satellite LST products are available, none of them is specially designed for the pan-Arctic region. Based on the recently published EUMETSAT Advanced Very High Resolution Radiometer (AVHRR) fundamental data record (FDR), a new LST product (1981–2021) with daily resolution is developed for the pan-Arctic region. Validation shows good accuracy with an average mean absolute error (MAE) of 1.71 K and a MAE range of 0.62–3.07 K against in situ LST data from the Surface Radiation Budget (SURFRAD) network and Karlsruhe Institute of Technology (KIT) sites. Long-term stability, a strong requirement for trend analysis, is assessed by comparing LST with air temperatures from ERA5-Land (T2M) and air temperature data from the EUSTACE (https://www.eustaceproject.org, last access: 17 December 2024) global station dataset. Long-term stability might not be fulfilled mainly due to the orbit drift of the NOAA satellites. Therefore, the analysis is split into two periods: the arctic winter months, which are unaffected by solar illumination and, therefore, orbital drift, and the summer months. The analysis for the winter months results in correlation values (r) of 0.44–0.83, whereas for the summer months (r) values range between 0.37–0.84. Analysis of anomaly differences revealed instabilities for the summer months at a few stations. The same stability analysis for the winter months revealed only one station with instabilities in comparison to station air temperature. Discrepancies between the temperature anomalies recorded at the stations and ERA5-Land T2M were also found. This highlights the limited influence of orbital drift on the LST product, with the winter months presenting good stability across all stations, which makes these data a valuable source for studying LST changes in the pan-Arctic region over the last 40 years. This study concludes by presenting LST trend maps (1981–2021) for the entire region, revealing distinct warming and cooling patterns.

1 Introduction

In recent decades, the warming in the Arctic has been much faster than in the rest of the world. Studies (Chylek et al., 2022; Rantanen et al., 2022) indicate a warming up to 4 times faster since 1979. This phenomenon, known as Arctic amplification, is visible in both instrumental records and model simulation (Dada et al., 2022). The Arctic Monitoring & Assessment Programme (AMAP) highlights diminishingly long cold spells and increasingly high extreme temperatures, leading to rapid changes in the cryosphere (Arctic Monitoring and Assessment Programme (AMAP), 2021). These episodes affect the sensitive Arctic ecosystem; vegetation dynamics; large-scale circulation patterns; and the distribution of snow, ice and permafrost (Maturilli et al., 2019). Permafrost, a crucial component of arctic ecosystems, is particularly sensitive to increasing air temperatures and changes in the snow regime. Thawing permafrost affects the stability of the bedrock, damages infrastructures, and releases massive quantities of organic carbon (Christensen et al., 2004; Miner et al., 2022). These potential threats highlight the importance of monitoring climate variables such as temperature in the arctic regions (Hachem et al., 2012; Urban et al., 2013). Rapid changes in land surface temperature (LST) pat-

terns have been observed in the Arctic region (Reiners et al., 2021). LST can be used as an indicator of the thermal state of the ground and has, in the last decade, been increasingly used in arctic research and permafrost modelling (Westermann et al., 2009; Obu et al., 2019; Batbaatar et al., 2020; Nielsen-Englyst et al., 2021). LST observations are available from in situ stations or climate models. However, these sources are insufficient to spatially resolve land surface characteristics and their response to climate change at a hemispheric scale (Nitze et al., 2018; Bartsch et al., 2023).

In contrast, satellite data can derive spatially comprehensive information on LST dynamics (Li et al., 2013, 2023b). LST is mainly derived from thermal infrared (TIR) radiation measured by satellites with algorithms based on radiative transfer (RT) equations (Li et al., 2013). Passive microwave (MW) measurements are another source of LST data. MW measurements are less affected by clouds than TIR data, but emissivity is challenging to derive for these wavelengths, especially over snow-covered ground (Jiménez et al., 2017; Ermida et al., 2017). LST is a critical parameter in Earth's surface and water energy balance and is widely exploited across different research fields: cryosphere, geology, vegetation monitoring, hydrology, and urban management (Li et al., 2013; Guillevic et al., 2018). From a climate perspective, LST is needed to evaluate land surface and land–atmosphere exchange processes, constrain surface energy budgets and model parameters, and provide observations of surface temperature change globally and in key regions (Guillevic et al., 2018). LST is defined as an essential climate variable (ECV) by the Global Climate Observing System (GCOS). To retrieve statistically significant changes in ECVs, a time series of at least 30 years is needed (WMO, 2010). Typical LST products include the Pathfinder, GLASS, MODIS, ASTER, and Landsat products (Good et al., 2022; Reiners et al., 2023; Li et al., 2023a), as well as the Spinning Enhanced Visible and Infrared Imager (SEVIRI) LST product produced within the framework of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on Land Surface Analysis (LSA-SAF) (Freitas et al., 2010; Trigo et al., 2011).

Although many satellite LST datasets with different temporal and spatial resolutions exist, only the Advanced Very High Resolution Radiometer (AVHRR) on board the NOAA and MetOp satellites series covers over 4 decades. EUMETSAT published a new AVHRR fundamental data record (FDR) in May 2023 (https://navigator.eumetsat.int/product/EO:EUM:DAT:0862, last access: 17 December 2024). This dataset was homogeneously produced and consists of reflectance and brightness temperatures covering 1978–2021. It is based on reprocessed near-real-time (NRT) observations from 17 AVHRR instruments on board NOAA satellites TIROS-N to NOAA-19, as well as EUMETSAT satellites MetOp-A, MetOp-B, and MetOp-C (EUMETSAT, 2023a). The dataset is provided in the Global Area Coverage (GAC) resolution. Quality control of AVHRR GAC radiances and updates in the retrieval methods offer more accurate results and better uncertainty estimates (Karlsson et al., 2023b). Notably, CLARA-A3, the third edition of the existing cloud albedo and radiation (CLARA) data record, was produced from this FDR (Karlsson et al., 2023a).

Previous LST datasets exist that have a global coverage (Ma et al., 2020; Li et al., 2023b) or were developed for continental (Reiners et al., 2021) or local usage. However, none of them were specifically derived for the pan-Arctic region. The GLASS product, for example, presents data gaps above 45° latitude due to its LST retrieval process that relies on visible channels, which are not available during polar night. Furthermore, the performance of the LST algorithms is strongly tied to the sampling of atmospheric profiles and surface properties used to calibrate the RT models. Different methodologies and data sources have been used in the past, but most LST products rely (entirely or partially) on the SeeBor database, which is built from the Thermodynamic Initial Guess Retrieval (TIGR)-3 database, ERA-40, and radiosonde datasets (Borbas et al., 2005). Taking advantage of the most recent ECMWF version-5 reanalysis (ERA5), Ermida and Trigo (2022) developed a new clear-sky database for the development of LST algorithms. The synthetic database is constructed from ERA5 data chosen with a dissimilarity criterion to ensure a uniform distribution of atmospheric conditions. The clear-sky database shows a significantly wider range of conditions and thus a wider range of brightness temperatures than in the SeeBor database specifically. The pan-Arctic region shows a wide range of temperatures and conditions that are not necessarily common, so it is particularly important to base the RT modelling on a robust and representative database for that region. A new daily LST dataset, presented here for the northern high latitudes ($> 50°$ N), is produced based on the EUMETSAT FDR and the clear-sky database (Ermida and Trigo, 2022).

The new LST dataset represents a valuable source for studying LST dynamics and its impacts on regional climates surface energy balance (Hall et al., 2012; Key et al., 2016), vegetation phenology (Li et al., 2021), and temperature hot spots (Mildrexler et al., 2018) and characterizing land use/land cover dynamics. However, to perform climatological analyses, it is crucial that the satellite LST observations are stable and robust (Waring et al., 2023).

The objectives of this paper are first the description of the LST retrieval and validation methods and, second, the assessment of the stability of the new pan-Arctic LST dataset. Relationships and trends with respect to the following datasets are compared: (i) in situ air temperature (Tair) measurements provided by the EUSTACE database and (ii) 2 m air temperature (T2M) data from ERA5-Land for the overlapping period (1981–2020). The comparisons are made for selected sites in the pan-Arctic region. The presence of trends in T2M is well established and is considered one of the major indicators of anthropogenic climate change (IPCC, 2021). Previous studies (Mildrexler et al., 2011; Hachem et al., 2012; Urban et al.,

2013; Good et al., 2022) obtained a good correlation between LST and T2M/Tair, although these parameters have different physical meanings and are measured or modelled with different procedures. Hachem et al. (2012) found that LST derived from MODIS and daily near-surface air temperatures are comparable. Good (2016) noted that LST and T2M are very similar when solar heating is low or absent. The NOAA satellites do not have a stable orbit (Ignatov et al., 2004; Latifovic et al., 2012), meaning that over the course of their operating years, their Equator crossing time is shifting (Price, 1984). In the case of LST, drifting orbits could lead to artificial trends in long-term records if only one platform is considered. In the present case, the LST from the different platforms is combined (morning and afternoon overpasses), and the final product is generated over multiple satellites. However, this also means that observed trends are more complex to interpret (Lieberherr and Wunderle, 2018). Orbital drift has a more substantial impact in the Southern Hemisphere and on bare soil (Sobrino et al., 2002; Gleason et al., 2002) than in the Northern Hemisphere. For example, this effect has been neglected in previous studies focusing on the Arctic region (Urban et al., 2013) or lakes (Riffler et al., 2015).

In this study, the analysis is carried out for two cases: (a) polar winter, defined here as December and January, and (b) polar summer, defined here as June and July. Incident solar radiation is zero during polar winter (Lund et al., 2017; Wang and Zeng, 2014); therefore, it is expected that a trend analysis for case (a) should not be affected by orbital drift. This paper is structured as follows: the data used to produce the LST are presented in Sect. 2, and the methodology is described in Sect. 3. Validation results for the LST product are presented in Sect. 4.1. Comparisons with air temperature datasets and trend analysis are presented in Sect. 4.2, 4.3, and 4.4. Finally, discussions and conclusions are presented in Sects. 5 and 6.

2 Data

This study uses the newly generated EUMETSAT AVHRR FDR satellite dataset, one reanalysis dataset, and several weather station datasets. Snow cover information is based on the snow water equivalent (SWE) dataset and snow cover fraction (SCF) from the ESA CCI Snow project.

2.1 EUMETSAT AVHRR FDR

The FDR contains AVHRR reflectance and brightness temperatures for each available orbit and channel. The daily AVHRR data from one satellite provide nearly complete coverage of the globe. AVHRR GAC measurements have been processed using pygac – a Python software package to read and transform AVHRR data in GAC format (https://pygac.readthedocs.io/en/latest/#, last access: 17 December 2024) – including the conversion from counts to reflectance or brightness temperature and cross-calibration of the visible channels of the AVHRR sensor. The two thermal channels are calibrated following the platinum resistance thermometer (PRT), space, and internal calibration target (ICT) count procedure (Kidwell, 1995; Walton et al., 1998). Detailed information is available in PyGAC FDR ATBD (EUMETSAT, 2023b). The data are accompanied by additional metadata (such as orbit overlap and Equator crossing time) and basic quality indicators (EUMETSAT, 2023c). Only satellites carrying the newer versions (AVHRR/2 and AVHRR/3) of the AVHRR are considered in this study. The second version (AVHRR/2) has five spectral channels, and the third version (AVHRR/3) has six but transmits only the data from five channels. Brightness temperature channel 4 is centred at 10.8 µm and brightness temperature channel 5 at 12 µm. The IR calibration procedure is satellite-specific, with no cross-calibration between satellites for IR channels (EUMETSAT, 2023d). The PyGAC AVHRR FDR from EUMETSAT (2023) is available in the Network Common Data Form (NetCDF) format and covers the entire globe (−180, 90, 180, −90°). For each satellite, the orbit files are composited by choosing for each pixel only the observation closest to nadir. The composited files have a spatial resolution of $0.05° \times 0.05°$ pixel size and are available for each satellite twice a day (at daytime and nighttime). This study focuses on the pan-Arctic region. Therefore, only data above 50° N have been processed.

2.1.1 Cloud mask

Cloud cover information is obtained from the CM SAF CLARA-A3 dataset (Karlsson et al., 2023b), which is also based on EUMETSAT AVHRR FDR (EUMETSAT, 2023a). The probabilistic cloud mask (CMAPROB), included in the level-2b product, and quality flags are used. Cloud probabilities range from 0 % to 100 % (EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF), 2023). The cloud probabilistic detection is based on the naïve Bayesian theory and offers substantial improvement to previous CLARA-A2 data records (Devasthale and Karlsson, 2023).

2.1.2 Snow cover data

Snow cover fraction (SCF) from optical satellite data and snow water equivalent (SWE) products from passive microwave satellite data from the ESA CCI Snow project (https://climate.esa.int/en/projects/snow/, last access: 17 December 2024) are used to obtain information on snow extent (Luojus et al., 2022; Xiao et al., 2024). The SCF "viewable snow" (SCFV) product is derived from the EUMETSAT AVHRR FDR and applied for this study. A value equal to zero means that the pixel is snow-free, and 100 means that the pixel is fully covered by snow. The SWE variable is indicated in millimetres. Both snow products are combined to

get a snow mask independent of the availability of the visible channels during polar night.

2.2 ERA5-Land 2 m air temperature

ERA5-Land (Muñoz Sabater et al., 2021), provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), is a downscaled version of the land component of the global ERA5 reanalysis. Compared to ERA5, ERA5-Land shows better stability but reduced accuracy (Urraca and Gobron, 2023). However, the accuracy suffices to capture inter-annual variations (Rantanen et al., 2023). Data from the ERA5-Land monthly averaged air temperature 2 m above the surface are compared to the LST dataset.

2.3 In situ 2 m air temperature

In situ air temperature observations from the EU Surface Temperature for All Corners of Earth (EUSTACE) land station dataset (Brugnara et al., 2019; Rantanen et al., 2023) are used for comparison. This database stores daily minimum (T_{min}) and maximum (T_{max}) temperature values recorded at weather stations ~ 2 m above the surface. The station dataset has undergone quality controls, was homogenized, and covers the period from 1850 to 2015. Weather stations from the EUSTACE database were selected according to the following criteria:

- The station lies above 50° latitude.
- The underlying ground is composed of permafrost.
- Different latitudes are represented.
- The surrounding area at a station, corresponding to at least 1 GAC pixel, must be homogeneous.
- The time series should cover at least 30 years.

Based on these criteria, 12 stations have been selected (Fig. 1 and Table 1).

2.4 Auxiliary data

The generation of LST data requires auxiliary datasets:

- Skin temperature (T_{skin}) and total column water vapour (TCWV) from the MERRA-2 reanalysis dataset (M2T1NXSLV, variables are labelled TS and TQV) are used. The data come at hourly temporal resolution with a spatial resolution of $0.5° \times 0.625°$. Nearest neighbour resampling was performed to match the AVHRR spatial resolution and scanline time, i.e. as in the work of Ma et al. (2020). MERRA-2 is preferred over other reanalysis products with finer spatial resolution to allow comparison with the GLASS product (Ma et al., 2020) and to keep the LST retrieval independent of ERA5-Land, which will be used for the stability analysis.
- ESA CCI Land Cover 1992–2015 and Copernicus ICDR Land Cover 2016–2020 datasets are used. Both datasets are consistent with each other. Their spatial resolution was decreased to match AVHRR GAC spatial resolution. The class labels follow the Land Cover Classification System (LCCS) developed by the United Nations (UN) Food and Agriculture Organization (FAO) (Copernicus Climate Change Service, Climate Data Store, 2019).
- Atmospheric profiles from the Clear-Sky Database developed at LSA-SAF (Ermida and Trigo, 2022) are used for the RT modelling (RTM). This database contains atmospheric profiles such as temperature, specific humidity, and ozone on 137 model levels (full vertical resolution), sampled from ERA5 for the 2009–2019 period. The sampling technique follows the method of Chevallier et al. (2000). Surface variables like T2M, surface pressure, T_{skin}, and emissivity are obtained from the combination of ERA5 and satellite data to ensure the best possible representation of the surface conditions. Column variables, such as TCWV and total cloud cover (TCC), are also present in the database. The atmospheric profiles are classified on TCWV varying from 0 to 60 mm and TS ranging from 190 to 340 K.
- In situ LST measurements from the Surface Radiation Budget (SURFRAD) network, the Karlsruhe Institute of Technology (KIT) network (Göttsche et al., 2016; Martin et al., 2019), the Atmospheric Radiation Measurement Climate Research Facility US Department of Energy (ARM) site at the North Slope of Alaska (NSA), and radiation data from the Baseline Surface Radiation Network (BSRN) (Maturilli, 2020; Cox and Halliwell, 2021; Kustov, 2018, 2023) are used for validation. Table 2 lists the stations. The KIT stations, which are part of LSA SAF's validation effort and supported by EUMETSAT, are located in different climate zones (Göttsche et al., 2016). In situ LST is computed from their radiation components as in Martin et al. (2019). The broadband emissivity (BBE) is obtained from channel effective emissivity data provided in the ASTER GED (Hulley et al., 2015) with the linear equation described in Cheng et al. (2013).
- The Copernicus digital elevation model (DEM) GLO-90 upscaled to 0.05° spatial resolution is used (https://doi.org/10.5270/ESA-c5d3d65, Copernicus digital elevation model (DEM) GLO-90, 2024) for the RT modelling. This dataset represents the surface of the Earth and is based on radar satellite data obtained from the TanDem-X Mission.

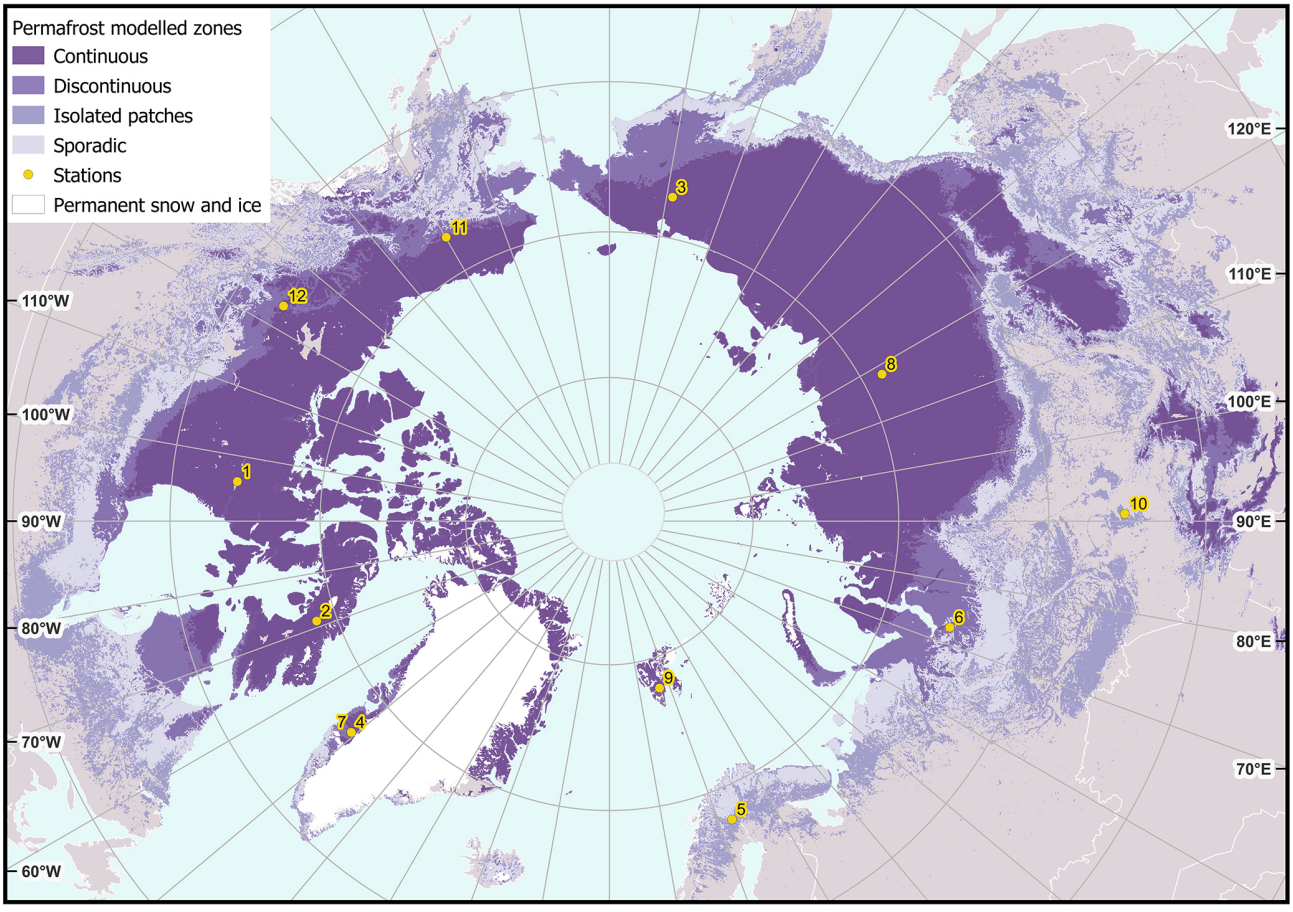

Figure 1. Selected EUSTACE stations with their IDs (Table 1). Permafrost zonation map adapted from Obu et al. (2018).

Table 1. Selected EUSTACE meteorological stations with WMO code, geolocation, elevation. and available period.

Station ID	Station code	Station name	Latitude [°]	Longitude [°]	Elevation [m]	Available period
1	CA002300500	BAKER LAKE A	64.30	−96.08	19	1946–2015
2	CA002401030	DEWAR LAKES	68.65	−71.17	527	1958–2015
3	TX_SOUID148829	ILIRNEJ	67.25	168.97	352	1944–2013
4	TX_SOUID147048	KANGERLUSSUAQ	67.02	−50.70	50	1975–2015
5	TX_SOUID137416	LAINIO	67.76	22.35	315	1965–2015
6	TX_SOUID148484	NADYM	65.47	72.67	14	1959–2013
7	GLW00016504	SONDRESTROM	67.02	−50.80	50	1947–2015
8	TX_SOUID148639	SUHANA	68.62	118.33	78	1938–2013
9	TX_SOUID111376	SVEAGRUVA	77.88	16.72	9	1978–2015
10	TX_SOUID150449	SVETLOLOBOVO	55.10	90.80	326	1958–2013
11	USC00509869	WISEMAN	67.42	−150.11	349	1918–2015
12	CA002204000	WRIGLEY A	63.22	−123.43	150	1943–2013

3 Methods

LST can be retrieved from thermal infrared data with the well-established split-window (SW) method (Ma et al., 2020; Yang et al., 2020; Reiners et al., 2021). Since 1983 (Price, 1984; Prata, 1994) different algorithms have been developed to obtain LST as a function of the satellite-recorded brightness temperature (BT). The split-window approach takes advantage of the different water vapour absorption characteristics of two adjacent channels (Lieberherr et al., 2017; Ma et al., 2020). LST is affected by many factors, which requires additional terms to model the effects of land cover type, viewing angle, and topography (Trigo et al., 2017).

Table 2. Description of the stations used for LST validation. Station name and ID, the network the station belongs to, latitude, longitude, elevation, and the dominant land cover type are listed.

Station name (ID)	Network	Latitude [°]	Longitude [°]	Elevation [m]	LCCS
Bondville, Illinois (BND)	SURFRAD	40.0519	−88.3731	230	Cropland
Desert Rock, Nevada (DRA)	SURFRAD	36.6237	−116.0195	1007	Open shrubland
Fort Peck, Montana (FPK)	SURFRAD	48.3078	−105.1017	634	Grassland
Goodwin Creek, Mississippi (GCM)	SURFRAD	34.2547	−89.8729	98	Wooded grassland
Penn State University, Pennsylvania (PSU)	SURFRAD	40.7201	−77.9309	376	Deciduous broadleaf forest
Sioux Falls, South Dakota (SFA)	SURFRAD	43.73403	−96.62328	1689	Cropland
ARM Southern Great Plains, Oklahoma (SGP)	SURFRAD	36.60406	−97.48525	314	Cropland
Table Mountain, Boulder, Colorado (TBL)	SURFRAD	40.1250	−105.2368	1689	Cropland
Lake Constance, Germany (BOD)	KIT	47.58	9.57	396	Water
Evora, Portugal (EVO)	KIT	38.54	−8.003	300	Mosaic tree and shrubs
North Slope of Alaska, USA (NSA)	ARM	71.323	−156.609	8	Lichens and mosses
Alert, Canada (ALE)	BSRN	82.49	−62.42	127	Bare soil
Ny-Ålesund, Norway (NYA)	BSRN	78.9227	11.9273	11	Bare soil
Tiksi, Russia (TIK)	BSRN	71.5862	128.9188	48	Shrubland

3.1 Generalized split-window algorithm

For this study the generalized split-window (GSW) algorithm, developed by Wan and Dozier (1996) (Eq. 1) and used within the framework of LSA-SAF (Trigo et al., 2008b; Ermida and Trigo, 2022), is selected. The GSW performs well on a global scale and has the highest relative accuracy among a selection of SW algorithms investigated in the work of Yang et al. (2020). The GSW depends on channel-effective emissivity and sensor-specific coefficients that must be determined for the expected atmospheric and surface conditions.

$$T_s = \left(A_1 + A_2\frac{1-\epsilon}{\epsilon} + A_3\frac{\Delta\epsilon}{\epsilon^2}\right)(T_{11}+T_{12}) + \left(B_1 + B_2\frac{1-\epsilon}{\epsilon} + B_3\frac{\Delta\epsilon}{\epsilon^2}\right)(T_{11}-T_{12}) + C, \quad (1)$$

where T_{11} and T_{12} denote BT of channels centred at approximately 11 and 12 µm; $\epsilon = (\epsilon_{11}+\epsilon_{12})/2$; $\Delta\epsilon = (\epsilon_{11}-\epsilon_{12})$; and A_i, B_i, and C are split-window coefficients (SWCs).

The split-window coefficients in Eq. (1) are obtained by applying multi-linear regression on a set of simulated BTs against a calibration database. Simulated BTs are obtained by performing RTM with version 13 of the Radiative Transfer for TOVS (RTTOV) developed at NWC SAF (Saunders et al., 2018). The Python wrapper was used in this study (Hocking et al., 2021). Atmospheric profiles and surface and column variables from the clear-sky profile database developed by Ermida and Trigo (2022), complemented by elevation information and satellite viewing angles, are ingested by RTTOV. The clear-sky profile database contains 97 files, each file containing approximately 1000 atmospheric profiles corresponding to a different class of TCWV and T_{skin}. In total, the database contains 82 793 profiles, and each profile possesses 6 different TS values and 25 pairs of emissivity at 11 and 12 µm. More details can be found in Ermida and Trigo (2022). Convolving the top-of-atmosphere radiances produced by RTTOV with the specific AVHRR response functions yields BTs as seen by the different sensors for different atmospheric and viewing conditions. RT modelling was performed on the calibration dataset for each satellite and 15 different satellite view zenith angles (VZAs) ranging from 0 to 70°. Table 3 summarizes the construction of the simulation dataset. Finally, the calibration was performed independently for each class, and for each class, the samples were split into a training (70 %) and test (30 %) set, and multilinear regression was performed on the resulting BTs. Based on the test sets, look-up tables (LUTs) with coefficients are created for each satellite. The LUTs are organized into classes of TCWV and T_{skin}, allowing the allocation of the right SWC to the encountered atmospheric conditions. Mean absolute error (MAE), the coefficient of determination (R^2), and root mean square error (RMSE) are computed for all coefficients to keep track of the general performance of the RTM.

3.1.1 Land surface emissivity retrieval

Land surface emissivity (LSE) is retrieved by combining the simplified normalized difference vegetation index (NDVI) threshold method ($SNDVI_{THM}$) (Sobrino et al., 2008) based on Sobrino and Raissouni (2000) with channel emissivity data from spectral libraries and static land cover classifications.

First, 10 d NDVI maximum value composites (MVCs) are generated from the AVHRR channels 1 and 2. NDVI thresholds that determine whether a pixel is considered fully vegetated or entirely bare soil are set. In the present case, $NDVI_{soil}$ is set to 0.2, and $NDVI_{veg}$ is set to 0.5 (Sobrino et al., 2001). All pixels that have $0.2 < NDVI < 0.5$ are con-

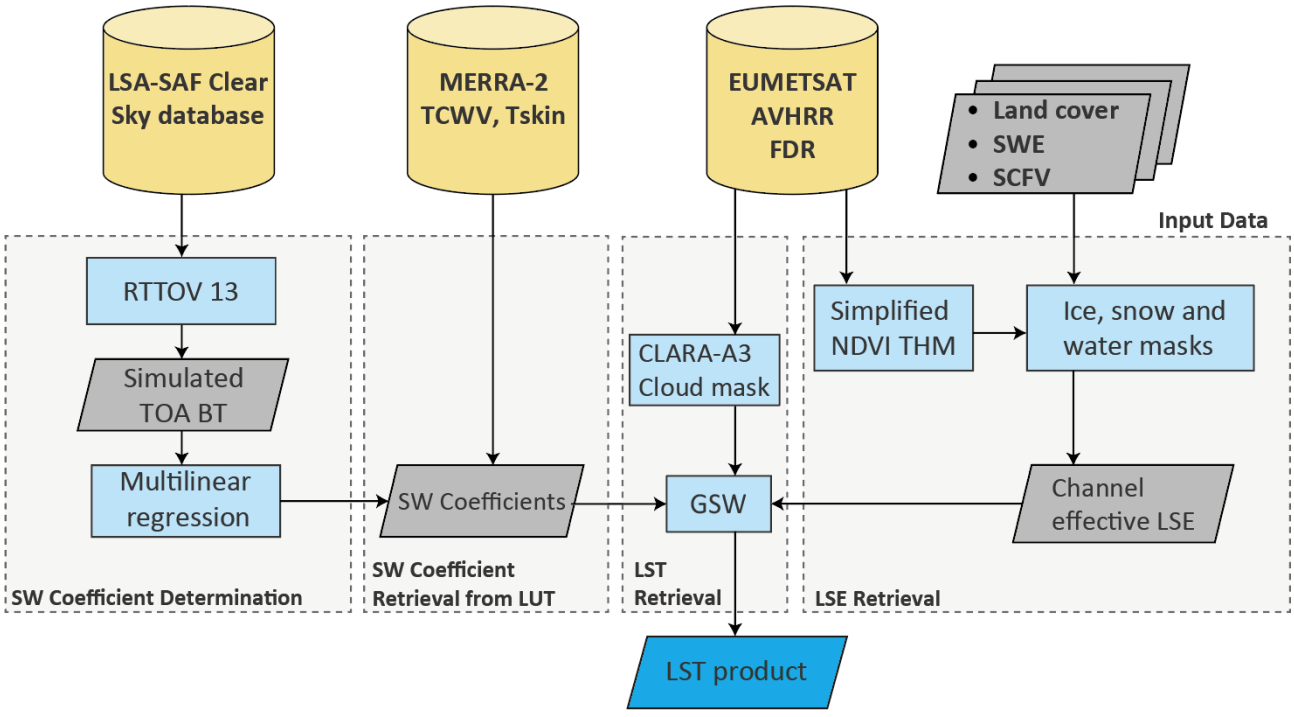

Figure 2. LST retrieval from AVHRR data. Input data are shown at the top. The leftmost box refers to the SW coefficient determination; the next box on the right displays the look-up-table storing coefficients for different atmospheric and surface conditions. The second box from the right presents the cloud mask application to the thermal channels and the rightmost panel the process of generating land surface emissivity (LSE) based on land cover and NDVI information.

Table 3. Summary of the simulation dataset: the number of profiles used and the number of instances of view zenith angles (VZAs), TS, and LSE are shown.

Source	No. profiles	No. VZA	No.TS	No. LSE	Sample size
LSA SAF (Ermida and Trigo, 2022)	82 793	15	6	25	186 284 250

sidered mixed pixels, and the corresponding emissivity is obtained by using the proportion of vegetation (P_v) method (Sobrino et al., 2008) (Eq. 2) that weighs the emissivity of bare soil (ϵ_{si}) and vegetation (ϵ_{vi}) for AVHRR channel i ($i = 4$ or 5).

$$\epsilon_i = \epsilon_{vi} P_v + \epsilon_{si}(1 - P_v) \quad (2)$$

ϵ_{si} and ϵ_{vi} are taken from a LUT based on information from spectral libraries (Trigo et al., 2008a; Peres and DaCamara, 2005). The emissivity of pixels with NDVI < 0.2 or NDVI > 0.5 is set to ϵ_{si} and ϵ_{vi} respectively. Here, the channel emissivities from Trigo et al. (2008b) (Table I) are used. The IGBP classes (Sulla-Menashe and Friedl, 2018) are converted to land cover classes of the ESA CCI project with plant functional type look-up tables (Wang et al., 2023). From 1992 to 2020, the ESA CCI land cover at the 12 selected stations changed very little. Therefore, to reduce emissivity uncertainties due to unknown land cover information before 1992, a static land cover from 2000 is used throughout the project (Freitas et al., 2010). The land cover was previously upscaled to the resolution of the AVHRR dataset. P_v is obtained from NDVI with Eq. (3) (Carlson and Ripley, 1997).

$$P_v = \left(\frac{\mathrm{NDVI} - \mathrm{NDVI}_{soil}}{\mathrm{NDVI}_{veg} - \mathrm{NDVI}_{soil}} \right)^2 \quad (3)$$

Pixels with low NDVI values (NDVI < 0.2) are defined as bare soil. Such NDVI values can also indicate the presence of snow or cloud cover. Snow cover extent information is retrieved from two data products from the ESA CCI Snow project, and cloudy pixels are masked out in the final LST product. A threshold of 70 % of SCFV or 4 mm for SWE is used to categorize the pixel as fully snow-covered. Snow-covered pixels are assigned to laboratory emissivity spectra values of medium snow (Fig. 5 of Hulley et al., 2014). Water pixels and permanent snow and ice areas are retrieved based on land cover information from the ESA land cover CCI for

the year 2000 and are assigned to channel effective emissivity values from Hulley et al. (2014).

3.2 LST retrieval

LST is retrieved as follows: (i) all necessary data (BTs, cloud mask, emissivity, and atmospheric data) are read; (ii) atmospheric data from MERRA-2 (TS and TQV) are downscaled to GAC spatial resolution, and the corresponding timestamp is matched with the scanline time for each pixel; (iii) based on satellite identification number, satellite viewing angle, total column water vapour (TQV), skin temperature (TS), and the corresponding SWC from the LUT are assigned to each pixel; and (iv) LST is computed from channel BTs, emissivities, and the assigned SWC (Eq. 1). Pixels with a satellite view zenith angle greater than 40° and the MAE of the test set in the RTM simulations greater than 0.5 K are masked out.

3.3 Validation procedure

The AVHRR LST dataset is validated against in situ data from different sites (Table 2). In situ LST and AVHRR LST datasets are jointly based on acquisition time and geolocation. The closest pixel to the station is taken, and a time difference of up to 5 min between the satellite overpass and the in situ measurement is considered. For Lake Constance, a time difference of up to 30 min is allowed. Similarly, as in Ma et al. (2020), 3σ filtering was performed to remove outliers (Pearson, 2002). The most accurate surface temperatures can be obtained over large water bodies, such as lakes and reservoirs. Densely vegetated surfaces are also particularly suitable for LST validation (Coll et al., 2009).

3.4 LST AVHRR time series generation

Depending on the heterogeneity of the land cover, between 4 and 9 AVHRR LST GAC pixels are extracted around each station (Table 1). Pixels that have a cloud probability higher than 0.1 are removed, and the average of the remaining pixels is computed. Pixels with a cloud probability of 0.1 or lower are considered cloud-free (Karlsson et al., 2023b). This cloud probability threshold is a compromise between data availability and avoiding cloud contamination. Daytime data from NOAA-7, NOAA-9, NOAA-11, NOAA-14, NOAA-16, NOAA-18, and NOAA-19 (satellites with ascending (northbound) Equator crossing times), as well as the entire MetOp series (satellites with descending (southbound) Equator crossing times), are considered for constructing the time series. The considered period for each satellite is chosen to minimize orbital drift and avoid the outage periods (EUMETSAT, 2023d). The retained periods are listed in Table 4.

Once the relevant periods are extracted, outlier detection is performed based on a 10 d rolling window analysis, and detected outliers are removed. Daily temperature variability is very high (Mildrexler et al., 2011), and AVHRR-derived LST time series are subject to noise. Therefore, monthly means are computed from concatenated daytime time series for further analysis.

Table 4. Considered time period for each satellite and sensor that it carries.

Satellite	Platform	Valid period
NOAA-07	AVHRR-2	24 November 1981–1 February 1985
NOAA-09	AVHRR-2	25 February 1985–7 November 1988
NOAA-11	AVHRR-2	8 November 1988–16 October 1994
NOAA-14	AVHRR-2	20 January 1995–31 December 2000
NOAA-16	AVHRR-3	1 January 2001–30 June 2005
NOAA-18	AVHRR-3	1 July 2005–28 February 2009
NOAA-19	AVHRR-3	1 March 2009–31 December 2015
MetOp-A	AVHRR-3	1 January 2016–31 December 2018
MetOp-B	AVHRR-3	1 January 2016–31 December 2020
MetOp-C	AVHRR-3	3 July 2019–31 December 2020

3.5 Time series analysis

The AVHRR LST monthly means time series is compared with two independent datasets: ERA5-Land and EUSTACE Tair. ERA5-Land pixels collocated to each weather station are selected, and values of T2M for the same position are extracted. The EUSTACE Tair data are filtered to keep only data that passed all quality assurance checks (Menne et al., 2012), and monthly means are computed. This yields time series of LST, ERA5-Land T2M (hereafter referred to as "T2M"), and EUSTACE Tair (hereafter referred to as "Tair") for each station. Before performing the actual stability analysis and starting with trend analysis, the overall relationship between the datasets is assessed by computing statistical parameters such as the Pearson correlation coefficient (r), the MAE, and the RMSE. Previous studies (Mildrexler et al., 2011; Hachem et al., 2012; Westermann et al., 2012; Urban et al., 2013) found variability in the LST–Tair correlation depending on the season and the land cover.

Then, anomalies of the monthly mean time series are computed for the LST, T2M, and Tair time series at all stations to remove the strong seasonal cycle inherent to temperature data (Good et al., 2022). A temperature anomaly describes the difference from a baseline climatology. In the present study, the anomalies are computed by subtracting the mean temperature for the entire time series from the monthly values. The LST, T2M, and Tair anomalies are compared by computing the Pearson correlation coefficient (r). Three different periods are considered for the anomaly analysis: (a) the entire year, (b) polar winter (December, January), and (c) polar summer (June and July). June and July are chosen as summer months to respect the symmetry of the winter period. First, the relationship between the anomalies of the different datasets is evaluated with the Pearson correlation coefficient (r), and the general stability of the LST dataset is assessed

by computing the trends of the anomaly differences. Three sets of differences are computed:

$$\Delta\text{Tanom}_{\text{LST}-\text{T2M}} = \text{LST}_{\text{anom}} - \text{T2M}_{\text{anom}} \quad (4)$$

$$\Delta\text{Tanom}_{\text{LST}-\text{Tair}} = \text{LST}_{\text{anom}} - \text{Tair}_{\text{anom}} \quad (5)$$

$$\Delta\text{Tanom}_{\text{Tair}-\text{T2M}} = \text{Tair}_{\text{anom}} - \text{T2M}_{\text{anom}}, \quad (6)$$

where LST_{anom} denotes the LST anomalies, and T2M_{anom} and $\text{Tair}_{\text{anom}}$ represent the ERA5-Land air temperature anomalies and EUSTACE air temperature anomalies respectively. The anomaly differences are represented by ΔTanom. The stability analysis is performed separately on the summer and winter periods. Non-parametric trend analysis is performed on the anomaly differences using the Theil–Sen slope to quantify the trend and the Mann–Kendall test to determine its significance. The Python implementation of the Mann–Kendall trend test (Hussain and Mahmud, 2019) was used in this work. Finally, trends are computed at each station for the entire year for all three datasets, and summer and winter trends are computed for the whole pan-Arctic region.

4 Results

In the first parts of the Results section (Sect. 4.1.1 and 4.1.2), the performances of the GSW algorithm and the validation results are presented. The remaining subsections present the relationships of the different datasets as well as the stability analysis and trend computation.

4.1 LST validation results

4.1.1 Performance of the split-window algorithm

Multi-linear regression was used to fit Eq. (1) to the RTM results for each class of TCWV and T_{skin} in a training set that consisted of 70 % of the samples, thereby retrieving the corresponding SWCs. The performance of the regression model was evaluated with the remaining independent samples (test set). The performance of the model is assessed by using the MAE and the coefficient of determination (R^2). The values shown in Fig. 3 are mean values across all satellites: the MAEs of the predictions are below 0.5 K for dry conditions (TCWV < 30) and low satellite viewing angle (VZA < 40). In cases with very moist atmospheres and high VZA (VZA > 50), the MAE increases substantially, and R^2 is considerably lower ($R^2 < 0.92$). Higher temperatures and higher VZA also lead to an increased error. The overall MAE is always below 2 K for water vapour below 50 mm TCWV. MAE values above 2 K are reached for high TCWV (TCWV > 50 mm) and high VZA values (VZA $> 50°$). Overall, the amount of water vapour in the Arctic atmosphere is low, which indicates that the model is well suited for the present use case.

4.1.2 Validation with in situ LST

Figures 4 and 5 show the validation results for NOAA-14, NOAA-16, NOAA-17, NOAA-18, and NOAA-19 and MetOp-A, MetOp-B, and MetOp-C against in situ LST from the validation sites in Table 2. The validation is separated for daytime (represented in red) and nighttime (blue). For each validation site, the MAE, RMSE, and mean bias error (MBE) are shown. The match-up with the SURFRAD stations covers the period from 1985 to 2020, the match-up period for the KIT station in EVORA (EVO) starts in 2009 and ends in 2020, the match-up period for Lake Constance starts in 2016 and ends in 2020, and finally the match-up period for the ARM site at the North Slope of Alaska (NSA) starts in 2007 and ends in 2012. Regarding the BSRN data, for the station in Alert (ALE) the match-up period starts in 2004 and ends in 2014, for Ny-Ålesund (NYA) it starts in 2006 and ends in 2020, and for Tiksi (TIK) it starts in 2010 and ends in 2018. EVO is located in an evergreen oak woodland with approximately 33 % tree crown cover, which can affect the satellite-retrieved LST due to directional effects (Rasmussen et al., 2011; Guillevic et al., 2013; Ermida et al., 2014). Due to this anisotropy, the surface in EVO presents high temperature differences between trees and the ground. The nighttime in situ measurements in EVO are therefore more suited than daytime observations. The surface is very heterogeneous at the NSA site, the station being close to lagoons (North Salt Lagoon and Imikpuk Lake) and very close to the coast. This explains why the performances are much worse during summertime than during wintertime when the entire area is snow and ice covered. Similarly, the BSRN stations are all located near the coast. In addition, the terrain around NYA station is mountainous and heterogeneous, leading to worse validation results. The in situ data for Lake Constance (BOD) are collected during the operating hours of the ferry and are thus only available during daytime. Both daytime and nighttime data are considered at the SURFRAD sites. BOD has the fewest available points due to the shorter match-up period. The overall RMSE range for the 10 stations is 0.80–3.43 K. The highest performance is reached at BOD. LSTs are more stable over water bodies, and water emissivity is less prone to inducing significant uncertainties (Masuda et al., 1988; Niclòs et al., 2005). On land, during the daytime, the highest agreement was obtained for the Desert Rock (DRA) and Goodwin Creek (GCM) sites. The lowest agreement was found at Fort Peck (FPK). During nighttime, the highest agreement was obtained at Sioux Falls (SFA), and the worst nighttime agreements were obtained at Desert Rock (DRA) and Fort Peck (FPK). Compared to previous studies on AVHRR LST (Ma et al., 2020; Reiners et al., 2021; Li et al., 2023a), the present dataset shows a similar accuracy and precision.

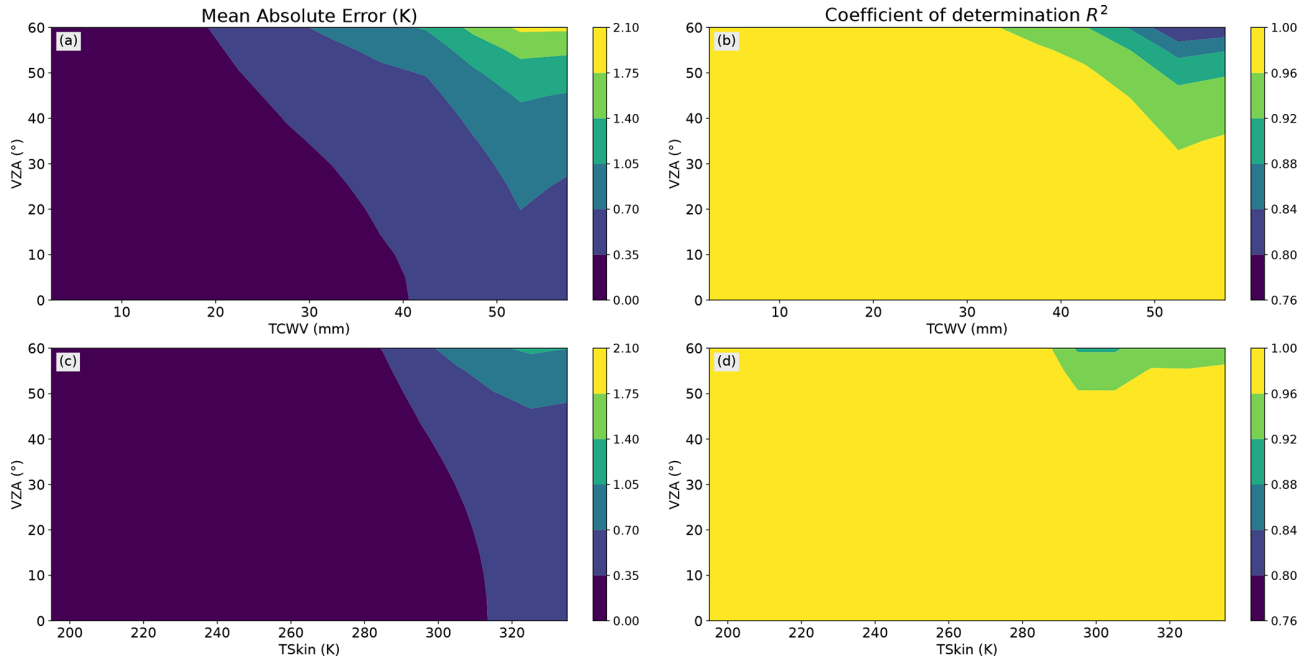

Figure 3. Split-window algorithm (SWA) performances: **(a)** distribution of the mean absolute error (MAE) and **(b)** coefficient of determination as a function of the view zenith angle (VZA) and total column water vapour (TCWV). Panels **(c)** and **(d)** show the corresponding distributions dependent on VZA and T_{skin}.

4.1.3 Comparison with the GLASS dataset

The pan-Arctic AVHRR LST dataset is compared against the well-established GLASS product (Zhou et al., 2019; Ma et al., 2020) that provides twice-daily LST observations for the whole globe for the 1980–2000 period. Figures 6 and 7 present a comparison of monthly means at two stations located in the Arctic (BAKER LAKE A and SVEAGRUVA). The classical GLASS LST, the orbital-drift-corrected (ODC) GLASS LST, and the pan-Arctic AVHRR LST are compared at the pixel closest to the station. In the high northern latitudes, the GLASS product is only available during the summer months. This is particularly visible for the SVEAGRUVA site (located on Svalbard), where very few GLASS observations are available. Also at BAKER LAKE A our product presents considerably more and slightly higher values, which can be explained by the different cloud masking and emissivity computation.

4.2 Relationships between the temperature datasets

The relationship between AVHRR LST and station EUSTACE Tair and the relationship between AVHRR LST and ERA5-Land T2M at each EUSTACE station are first assessed by computing the Pearson correlation coefficient (r), the MAE, and RMSE of the monthly mean temperature data (see Sect. 3.5). Comparisons against $\mathrm{Tair}_{\mathrm{min}}$ show, in general, higher RMSE and MAE values than comparisons against $\mathrm{Tair}_{\mathrm{max}}$ and T2M (Table 5). For the stations in Table 5, the mean MAE values are 8.47, 5.56, and 5.92 K for LST versus $-\mathrm{Tair}_{\mathrm{min}}$, $-\mathrm{Tair}_{\mathrm{max}}$, and $-$ T2M respectively. Stations WRIGLEY A, WISEMAN, and SVEAGRUVA exhibit lower MAE and RMSE values than the other stations. WISEMAN and SVEAGRUVA have slightly fewer comparison samples ($N < 230$). SVEAGRUVA, located in Svalbard, is the northernmost station and experiences persistent cloud cover, leading to fewer usable satellite observations. The EUSTACE time series at WISEMAN is shorter due to missing data. WRIGLEY A exhibits low MAE and RMSE values compared to $\mathrm{Tair}_{\mathrm{max}}$ but presents similar values to other stations compared to $\mathrm{Tair}_{\mathrm{min}}$.

As to be expected, the analysis reveals a high degree of correlation between monthly LST, Tair, and T2M with correlation coefficients all above 0.9 (r) (Table 5). In line with higher RMSE and MAE, $\mathrm{Tair}_{\mathrm{min}}$ has slightly lower (r) values compared to LST. The better performance of $\mathrm{Tair}_{\mathrm{max}}$ can be attributed to the closer daytime overpass of the NOAA–MetOp satellites (Good et al., 2022). Conversely, $\mathrm{Tair}_{\mathrm{min}}$ is generally recorded during night.

The relationship between the monthly AVHRR LST anomalies versus EUSTACE Tair anomalies and ERA5-Land T2M anomalies is assessed with the Pearson correlation coefficient (r) (Table 6). All stations except SUHANA (Siberia) and SVEAGRUVA (Svalbard) display strong positive correlations ($r > 0.5$) between LST and both air temperature datasets (Tair and T2M). Correlations are consistent for both air temperature datasets; only minor differences are visible.

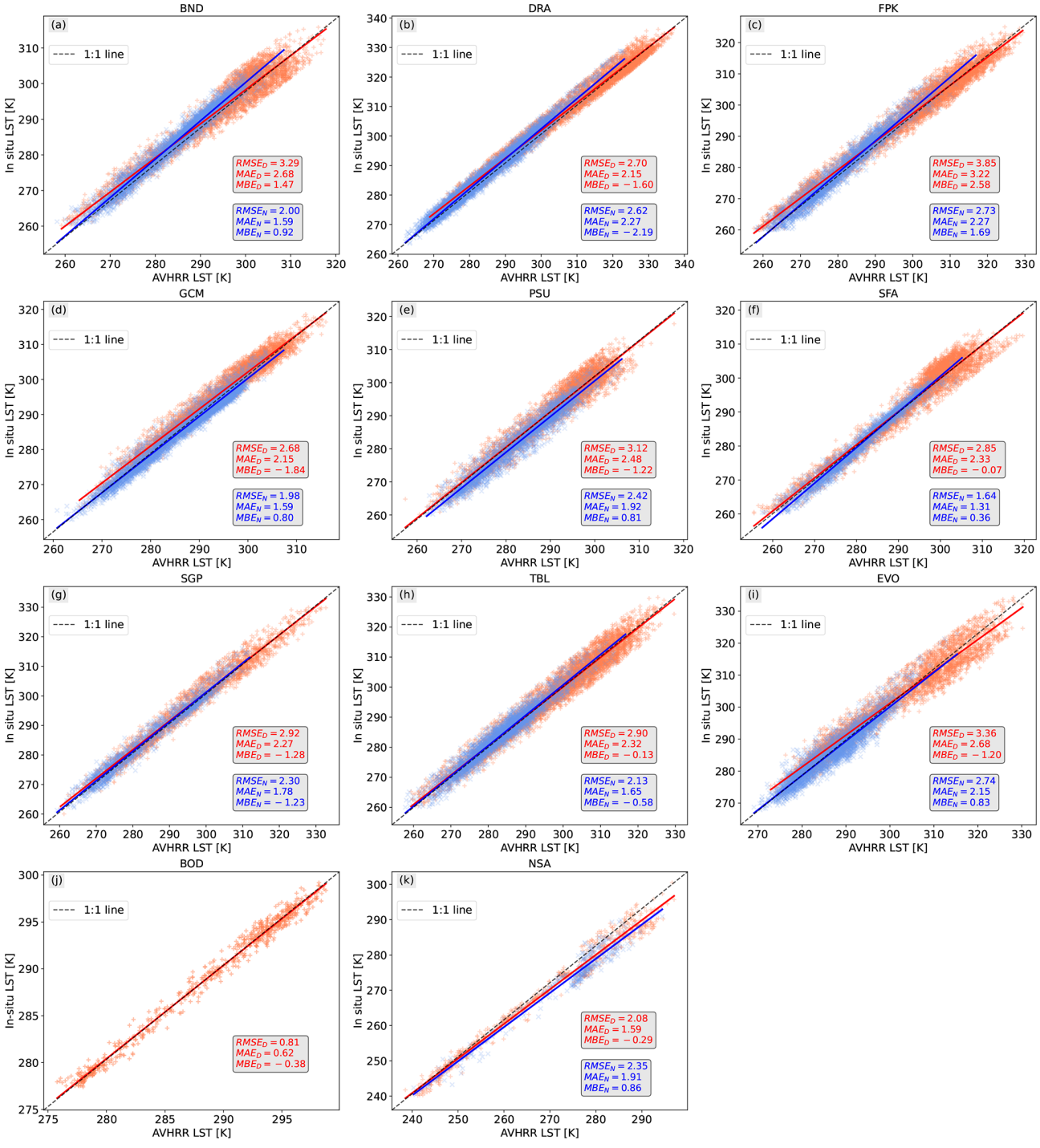

Figure 4. AVHRR LST versus in situ LST at **(a)** Bondville (BND), **(b)** Desert Rock (DRA), **(c)** Fort Peck (FPK), **(d)** Goodwin Creek (GCM), **(e)** Penn State University (PSU), **(f)** Sioux Falls (SFA), **(g)** Southern Great Plains (SGP), **(h)** Table Mountain (TBL), **(i)** Evora (EVO), **(j)** Lake Constance (BOD), and **(k)** North Slope of Alaska (NSA). Red represents daytime measurements, and blue represents nighttime measurements. Match-up periods are provided in the text.

The r values for comparing Tair with T2M are higher than those in the corresponding comparison with LST. Correlations of LST versus T2M anomalies vary between 0.46 (r) and 0.71 (r). For Tair_{max} versus LST, the r values r of are between 0.40 and 0.71, whereas for Tair_{min}, r has a range of 0.35–0.70. In general, slightly higher r values are obtained for the comparison against T2M than Tair. Comparison values of Tair versus T2M have values of r between 0.69 and

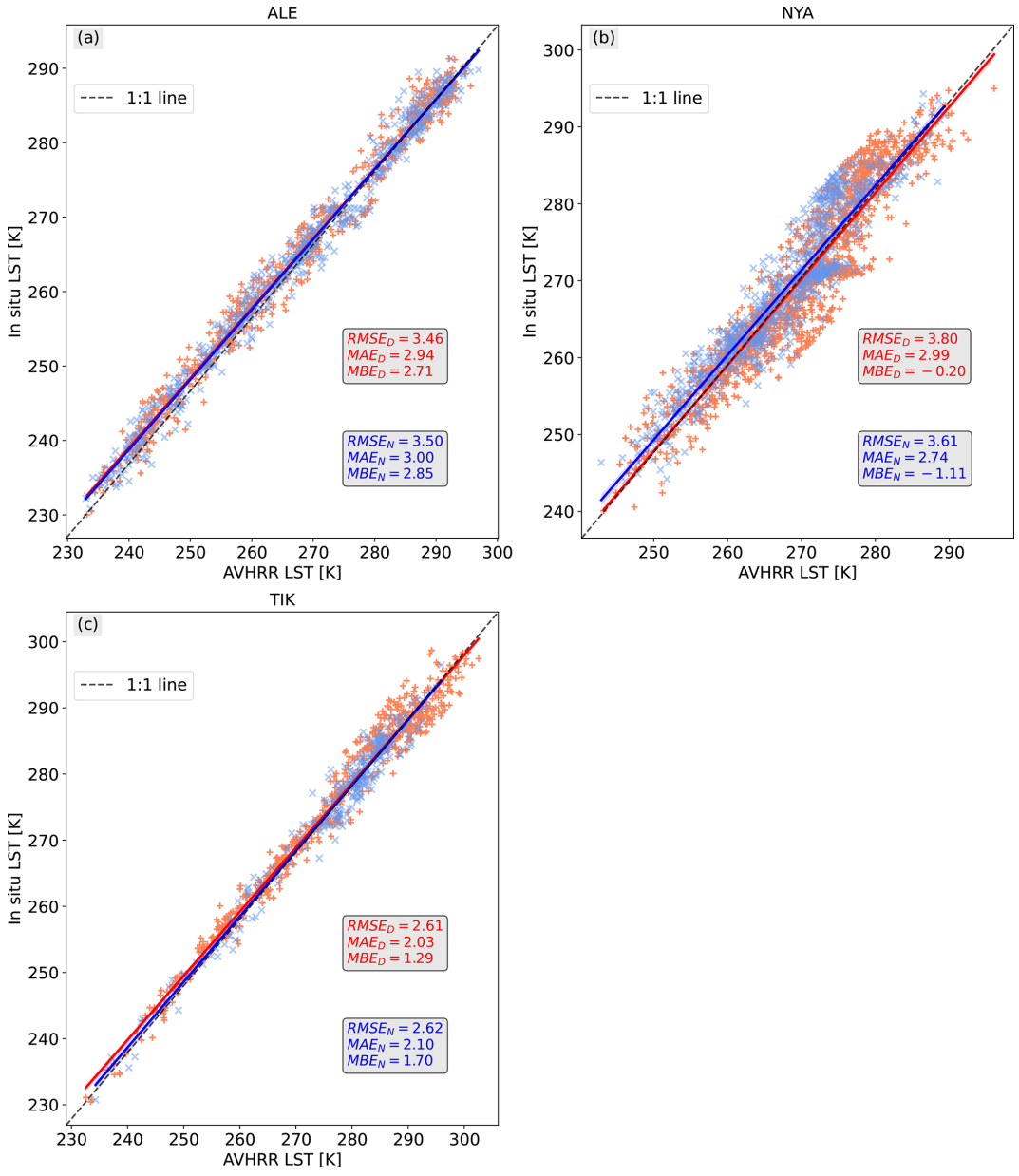

Figure 5. AVHRR LST versus in situ LST at BSRN stations: **(a)** Alert, Lincoln Sea (ALE); **(b)** Ny-Ålesund, Svalbard (NYA); and **(c)** Tiksi, Russia (TIK). Red represents the daytime measurements, and blue represents the nighttime measurements. Match-up periods are provided in the text.

0.97. Similarly to in Table 5, lower r values are found for the comparison with Tair_{min} than for Tair_{max}. The lowest correlation value between T2M and Tair is obtained at WISEMAN (Alaska). NADYM (Russia) shows consistently high correlations ($r \approx 0.7$) across all comparisons and has the highest correlation value for the Tair versus T2M evaluation. Previous studies have also found high correlations between station Tair and LST data for other LST datasets (Urban et al., 2013; Good et al., 2022), e.g. from the ESA CCI project (Ghent et al., 2023). The differences in r between Tables 5 and 6 can be attributed to a phase shift between the anomalies and to a strong seasonal signal present in the time series on monthly means.

To assess the general stability of the LST dataset, the differences between the monthly anomalies of the datasets (Eqs. 4, 5 and 6) and the trends of these differences are calculated. Since high correlation values were obtained for T2M and Tair_{max}, these two datasets are considered for the LST stability analysis. The confidence interval is set to 95 %, meaning that trends with p values below 0.05 present a significant trend. Ideally, the trend of the difference should be zero or very close to zero. The results are shown in Ta-

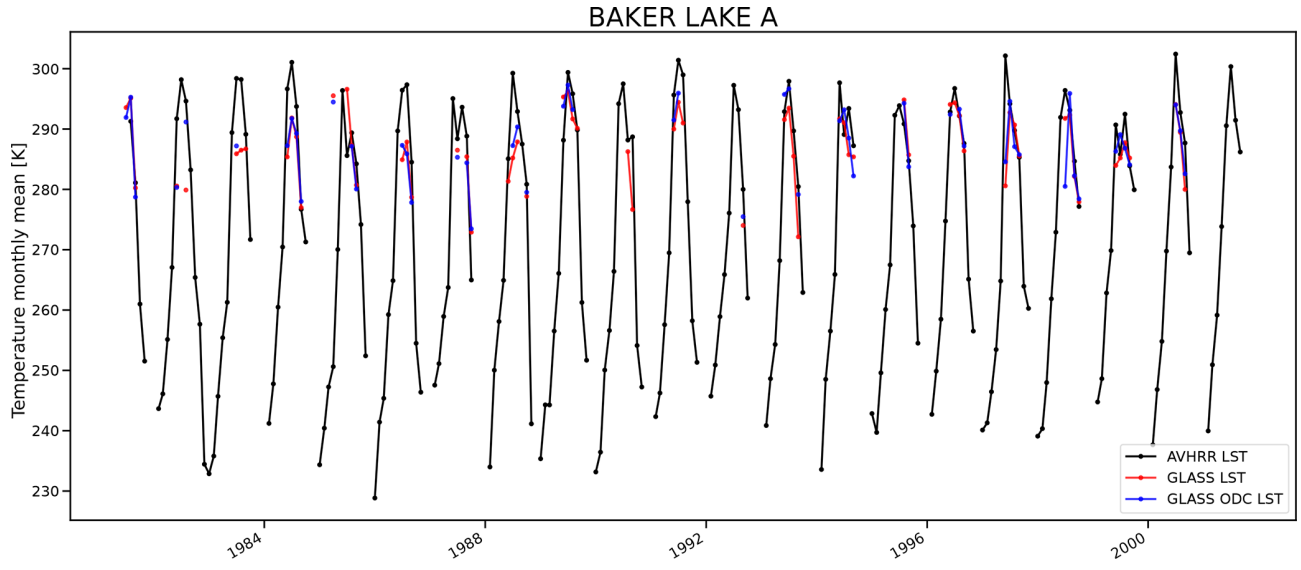

Figure 6. Monthly mean LST product comparisons at BAKER LAKE A.

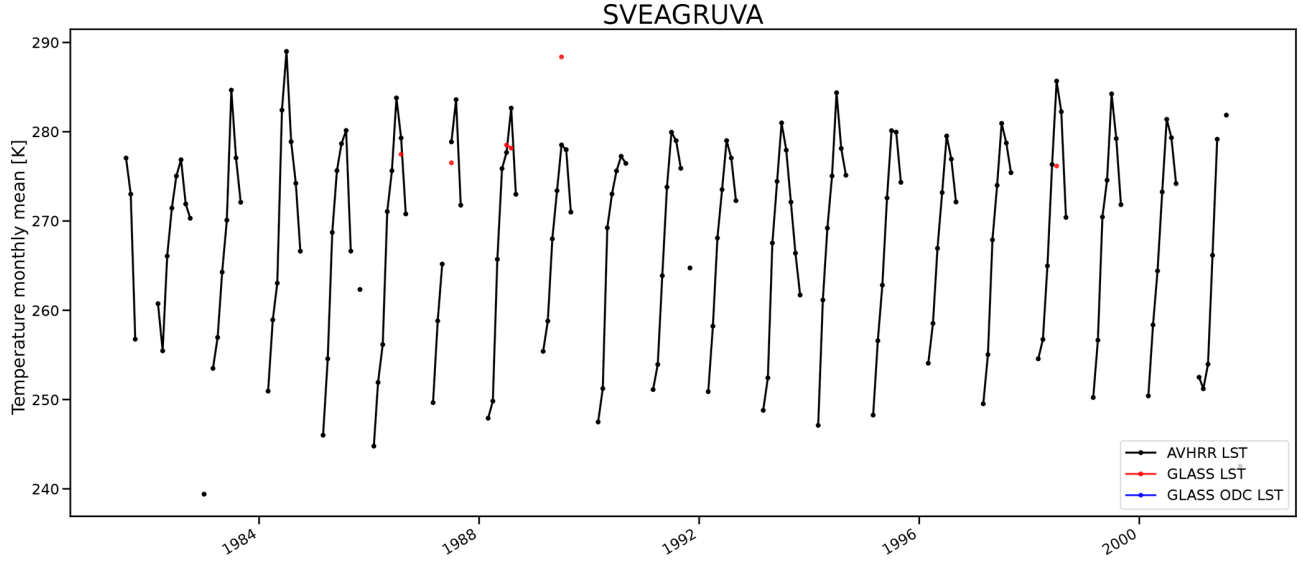

Figure 7. Monthly mean LST product comparisons at SVEAGRUVA.

ble 7: 10 of the 24 trends of the anomaly differences involving LST present statistically significant trends. This indicates that for these datasets, the difference between the anomalies increases or decreases over time and, thus, is not stable. Four stations (DEWAR LAKES, SVEAGRUVA, SVETLOLOBOVO, and WISEMAN) show statistically significant trends when comparing T2M and Tair_{max} anomalies (e.g. see SVEAGRUVA in Fig. 8). However, these stations present stable trends for the comparison with LST anomalies. For example, the comparison of LST versus T2M at SVEAGRUVA is very stable (0.10 K per decade) (see Fig. 8). The same observation can be made at DEWAR LAKES. SVEAGRUVA is a special case as the data are very sparse. Overall, the LST–Tair_{max} trend and LST–T2M trends shown in Table 7 are not consistent across datasets. For example, Fig. 9 shows the trends of the anomalies at KANGERLUSSUAQ. All trends are stable (statistically non-significant), but differences are visible in the trend values for both LST comparisons. Plots of the trends of the other stations are shown in Appendix A1. Several reasons, such as orbit drift, missing data, or corrupt station data, could explain the observed discrepancies. Instabilities are also visible in the Tair versus T2M comparisons. The significant trend for the comparison of Tair versus T2M can be explained by missing station data (Appendix A1). Six of the 12 LST versus T2M experiments are statistically non-significant, suggesting no detectable trends for the anomaly

Table 5. MAE, RMSE, and Pearson coefficient (r) at the selected EUSTACE stations for monthly mean comparisons. N is the number of samples.

Station name	Relationship (MAE, RMSE, and r) of monthly means									N
	LST versus $Tair_{min}$			LST versus $Tair_{max}$			LST versus T2M			
	RMSE	MAE	r	RMSE	MAE	r	RMSE	MAE	r	
BAKER LAKE A	9.56	8.07	0.97	5.27	4.26	0.98	6.40	4.89	0.98	298
DEWAR LAKES	10.62	8.37	0.95	9.92	8.46	0.96	8.17	6.18	0.97	330
ILIRNEJ	10.06	8.58	0.97	5.92	4.98	0.98	7.65	6.2	0.98	316
KANGERLUSSUAQ	10.80	8.65	0.95	8.69	7.28	0.96	8.38	6.91	0.97	357
LAINIO	7.78	6.61	0.96	5.34	4.18	0.97	5.63	4.59	0.97	290
NADYM	9.52	8.05	0.97	6.71	5.33	0.98	7.41	6.06	0.98	282
SONDRESTROM	12.15	9.91	0.95	8.11	6.93	0.96	9.39	7.76	0.97	329
SUHANA	9.82	8.70	0.98	5.76	4.62	0.98	6.37	5.20	0.98	339
SVEAGRUVA	5.03	3.66	0.93	7.17	5.82	0.94	5.65	4.22	0.94	228
SVETLOLOBOVO	15.30	13.08	0.96	7.31	6.06	0.97	10.72	8.94	0.97	343
WISEMAN	9.42	7.98	0.96	6.79	5.54	0.98	6.43	5.20	0.98	176
WRIGLEY A	10.79	9.97	0.97	4.39	3.31	0.98	5.75	4.91	0.98	279

Table 6. Correlation coefficient (r) from all stations for the comparison between the LST monthly mean anomalies versus the T2M and Tair monthly mean anomalies.

Station name	Pearson correlation coefficient (r) of the monthly mean anomalies				
	LST vs. $Tair_{min}$	LST vs. $Tair_{max}$	LST vs. T2M	T2M vs. $Tair_{min}$	T2M vs. $Tair_{max}$
BAKER LAKE A	0.53	0.56	0.57	0.91	0.92
DEWAR LAKES	0.59	0.58	0.63	0.88	0.89
ILIRNEJ	0.56	0.61	0.61	0.86	0.91
KANGERLUSSUAQ	0.66	0.68	0.70	0.93	0.96
LAINIO	0.61	0.65	0.64	0.93	0.94
NADYM	0.70	0.71	0.71	0.97	0.97
SONDRESTROM	0.67	0.70	0.69	0.93	0.96
SUHANA	0.43	0.48	0.51	0.88	0.94
SVEAGRUVA	0.35	0.40	0.46	0.92	0.95
SVETLOLOBOVO	0.59	0.70	0.68	0.91	0.95
WISEMAN	0.47	0.61	0.65	0.69	0.75
WRIGLEY A	0.57	0.57	0.62	0.85	0.87

differences at these stations. Regarding the $Tair_{max}$ anomalies minus LST anomalies trends, 4 out of 12 stations show a significant trend. The stations located at DEWAR LAKES, KANGERLUSSUAQ, LAINIO, SUHANA, SVEAGRUVA, and WRIGLEY A do not show a significant trend in the comparisons to LST (see Table 7).

LST and air temperature trends of stable stations (DEWAR LAKES, KANGERLUSSUAQ, LAINIO, SUHANA, SVEAGRUVA, and WRIGLEY A) have the same order of magnitude (Table 8), and all stations present a warming trend. DEWAR LAKES, located east of Nunavut (Canada), shows a more pronounced trend for LST than for the air temperature (Tair and T2M). Tair captured at the weather station does not present a significant trend. KANGERLUSSUAQ in Greenland shows similar trends in LST as in air temperature. LAINIO, located in a forested area in northern Scandinavia, presents similar trends to LST and Tair, while the corresponding T2M trend is lower. The LST trend in SUHANA (Siberia) does not show a significant trend for LST, but both air temperature datasets show a consistent trend. WRIGLEY A, located in the Northwest Territories in a forested area, shows a higher LST trend than T2M trend.

4.3 Analysis of summer and winter periods

To further investigate whether the orbital drift of the NOAA satellites influences the stability analysis and determined trends, the previous analysis is now performed separately for polar winter and polar summer (see Sect. 3.5). Correlation coefficients (r), trends (slopes), and p values of the differences between LST anomalies minus $Tair_{max}$ anomalies and LST anomalies minus T2M anomalies are calculated.

Table 7. Trend of the anomaly differences for the three pairs of differences. Trends in italic are significant (p value < 0.05), and associated p values are marked in bold.

Station name	LST–T2M		LST–$Tair_{max}$		$Tair_{max}$–T2M	
	Trend [K per decade]	P value	Trend [K per decade]	P value	Trend [K per decade]	P value
BAKER LAKE A	*0.52*	**0.0005**	*0.63*	**0.0003**	−0.07	0.13
DEWAR LAKES	0.04	0.81	0.22	0.22	*−0.11*	**0.03**
ILIRNEJ	*0.35*	**0.005**	0.21	0.21	0.03	0.72
KANGERLUSSUAQ	−0.23	0.13	−0.10	0.62	−0.04	0.42
LAINIO	0.25	0.054	0.10	0.50	0.06	0.08
NADYM	*−0.57*	**0.003**	*−0.59*	**0.01**	−0.00	0.96
SONDRESTROM	*0.37*	**0.04**	*0.55*	**0.008**	−0.05	0.35
SUHANA	−0.16	0.27	−0.33	0.06	−0.01	0.80
SVEAGRUVA	0.10	0.63	0.35	0.10	*−0.13*	**0.000**
SVETLOLOBOVO	*−0.52*	**0.002**	*−0.36*	**0.04**	*−0.17*	**0.005**
WISEMAN	*0.27*	**0.02**	0.05	0.88	*−0.56*	**0.006**
WRIGLEY A	0.11	0.32	0.21	0.31	−0.04	0.65

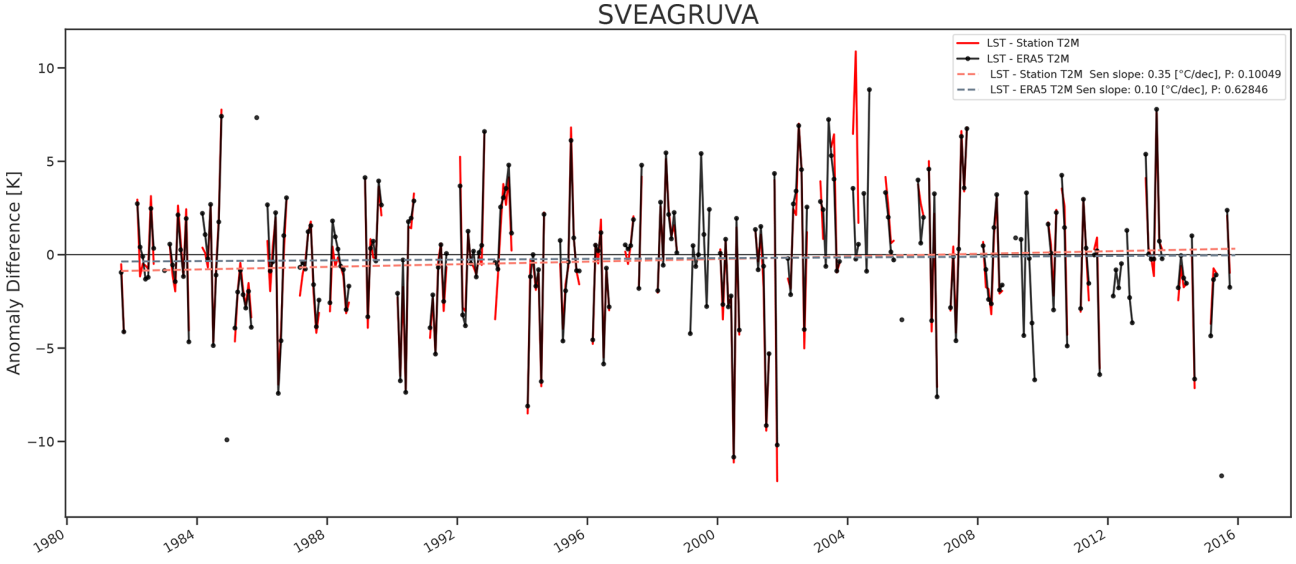

Figure 8. Monthly differences in the anomalies at SVEAGRUVA (Svalbard) between 1981 and 2015.

Correlation coefficients (r) (Table 9) for summer and winter for the analysis against LST lie in the range 0.4–0.8. Correlation results are, on average, slightly higher for LST versus T2M than LST versus $Tair_{max}$, for winter and summer. SVEAGRUVA had insufficient data points in winter, and the trends for summer were not significant. Therefore, this station was removed from the analysis. Winter shows only slightly higher correlation values than summer. At BAKER LAKE A, ILIRNEJ, SONDRESTROM, and WISEMAN, r values in summer are higher than in winter. Correlation values are slightly higher for the separate seasons (Table 9) than for the general analysis (Table 6). Correlation results for the $Tair_{max}$ versus T2M experiment range between 0.57 and 0.96. The lowest correlation values are obtained at WISEMAN and WRIGLEY A. All of the stations except SVETLOLOBOVO exhibit r values in a similar range in summer and winter. The air temperature correlation value at SVETLOLOBOVO is considerably lower in summer than in winter. SVETLOLOBOVO is the station located the furthest south, at 55° latitude. At WRIGLEY A, r values are similar for all significant results.

The results in the previous section show that some instabilities could be detected from the trends of the anomaly differences. The same analysis is now performed separately for summer and winter (Table 10). Five out of 12 stations do not show any significant trend either in summer or in winter. Except for SUHANA, the remaining stations only show a significant trend during summer. For example, KANGERLUSSUAQ (Table 10 and Fig. 10) presents a significant positive trend during summer but no significant trend in win-

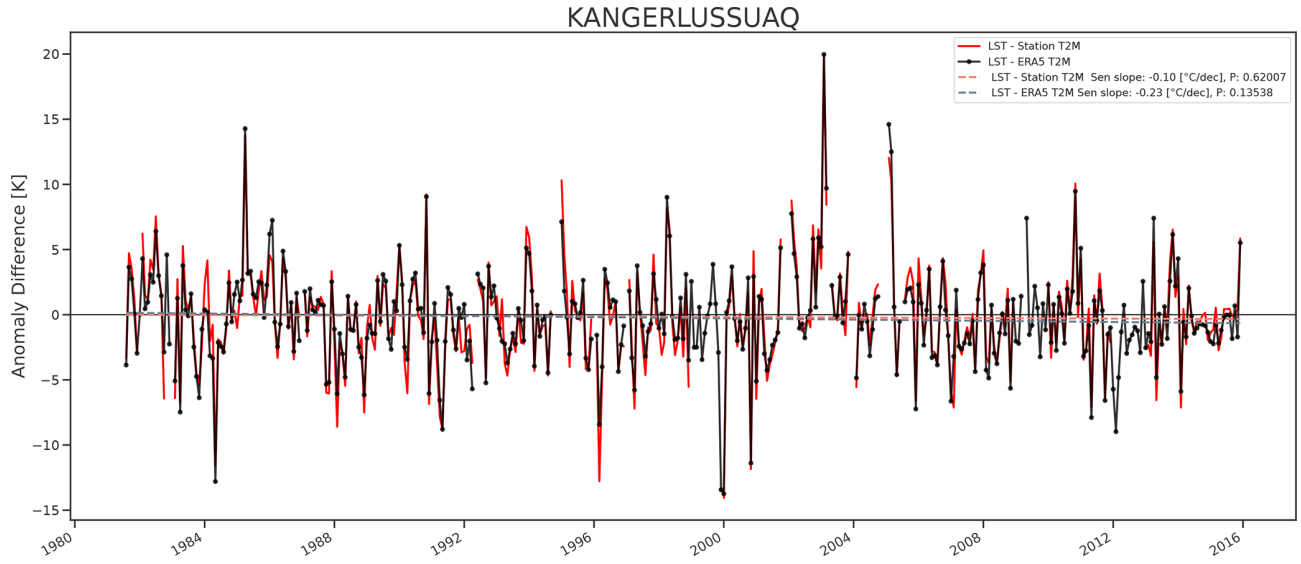

Figure 9. Monthly differences in the anomalies at KANGERLUSSUAQ (Greenland) between 1981 and 2015.

Table 8. Trends in monthly mean anomalies (K per decade) for selected stations. Values in italic are significant (p value < 0.05), and associated p values are marked in bold.

Station name	LST		EUSTACE Tair		ERA5-Land T2M	
	Trend [K per decade]	p value	Trend [K per decade]	p value	Trend [K per decade]	p value
DEWAR LAKES	*0.61*	**0.005**	0.21	0.11	*0.39*	**0.002**
KANGERLUSSUAQ	*0.50*	**0.025**	*0.52*	**0.0008**	*0.55*	**0.0001**
LAINIO	*0.64*	**0.0004**	*0.64*	**0.0001**	*0.50*	**0.0001**
SUHANA	0.26	0.18	*0.47*	**0.007**	*0.54*	**0.0001**
WRIGLEY A	*0.47*	**0.009**	0.34	0.1	*0.32*	**0.005**

ter. SUHANA experiences a significant positive trend in winter for the LST anomaly minus Tair_{max} anomaly difference, but this trend is not observed in comparison to T2M (Table 10 and Fig. 11). From Fig. 11, it can also be noticed that the differences between Tair_{max} and T2M increase over time in winter and summer, which suggests discrepancies between air temperature measured at the station and from ERA5-Land. The magnitude of the LST anomalies is also generally higher than Tair and T2M magnitude (Figs. 10 and 11), which can be explained by the higher amplitude of the LST diurnal cycle pattern compared to air temperature (Good, 2016; Sharifnezhadazizi et al., 2019). SVETLOLOBOVO and KANGERLUSSUAQ present poor correlation results in summer (see Table 9) and also exhibit strong significant summer trends for the anomaly difference, contrary to winter trends that remain stable (Table 10).

4.4 LST analysis for the pan-Arctic region

The previous sections identified a few stability issues in the current LST dataset, particularly during summer months (Tables 7 and 10). However, these limitations are not unique to LST, as analyses of the difference between both air temperature datasets also revealed instabilities (Table 7). While this highlights the stability and accuracy of the LST dataset, the trend analysis should be interpreted with care. The summer months are more prone to instability in the analysis than the winter months. Therefore, the trends of the winter and summer months are computed separately for the entire pan-Arctic region (Fig. 12). This allows us to compare and analyse temperature changes across different seasons. Additionally, mean LST values for summer and winter for the pan-Arctic are calculated (Fig. 13) to further understand temperature distributions during different seasons.

Cold glaciers and mountain zones in western Canada and Alaska are well captured in the summer mean temperatures (Fig. 13). These regions also exhibit a pronounced warming during summer and winter (Fig. 12). Warmer mean winter temperatures are present along the Lena River in Siberia, which is also visible in the mean annual ground temperature (MAGT) map from Obu et al. (2019). Generally, summer and winter mean LST values follow the temperature pattern shown in the MAGT map. For example, valleys in Russia present in the MAGT map are also visible in the mean

Table 9. Pearson correlation coefficient (r) of the monthly anomalies for the summer and winter period.

Station		LST versus $Tair_{max}$ (r)	LST versus T2M (r)	$Tair_{max}$ versus T2M (r)
BAKER LAKE A	Summer	0.65	0.66	0.88
	Winter	0.46	0.47	0.90
DEWAR LAKES	Summer	0.60	0.72	0.91
	Winter	0.62	0.83	0.85
ILIRNEJ	Summer	0.73	0.76	0.91
	Winter	0.44	0.46	0.88
KANGERLUSSUAQ	Summer	0.37	0.38	0.90
	Winter	0.59	0.63	0.96
LAINIO	Summer	0.62	0.75	0.95
	Winter	Not significant	Not significant	0.95
NADYM	Summer	0.46	0.57	0.93
	Winter	0.74	0.67	0.90
SONDRESTROM	Summer	0.70	0.72	0.90
	Winter	0.60	0.63	0.96
SUHANA	Summer	0.84	0.77	0.93
	Winter	Not significant	Not significant	0.83
SVETLOLOBOVO	Summer	0.47	0.37	0.79
	Winter	0.67	0.75	0.92
WISEMAN	Summer	0.60	0.64	0.75
	Winter	0.55	0.59	0.64
WRIGLEY A	Summer	Not significant	0.55	0.57
	Winter	Not significant	Not significant	0.58

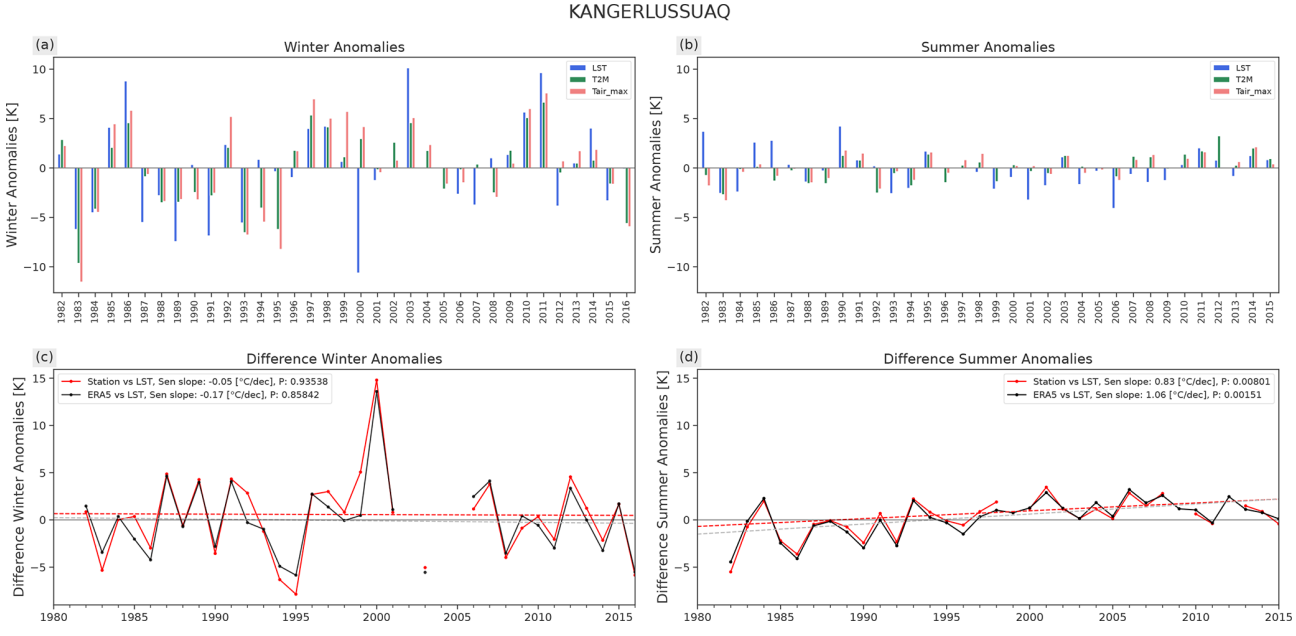

Figure 10. Winter and summer anomalies and the difference between anomalies of LST, T2M, and Tair time series for KANGERLUSSUAQ (Greenland).

Table 10. Trends of the anomaly differences for winter and summer. Trends in bold–italic are statistically significant.

Station		Trend [K per decade] LST–$Tair_{max}$ (p value)	Trend [K per decade] LST–T2M (p value)
BAKER LAKE A	Summer	−0.57 (0.30)	***−0.80 (0.043)***
	Winter	−2.04 (0.10)	−1.61 (0.28)
DEWAR LAKES	Summer	0.96 (0.23)	0.81 (0.22)
	Winter	−0.00 (1.0)	0.39 (0.45)
ILIRNEJ	Summer	0.06 (0.71)	0.21 (0.44)
	Winter	−0.43 (0.75)	−0.33 (0.54)
KANGERLUSSUAQ	Summer	***0.83 (0.008)***	***1.06 (0.001)***
	Winter	−0.05 (0.93)	0.17 (0.85)
LAINIO	Summer	0.58 (0.08)	0.51 (0.10)
	Winter	–	–
NADYM	Summer	***1.56 (0.002)***	***1.41 (0.001)***
	Winter	−0.97 (0.59)	−0.89 (0.53)
SONDRESTROM	Summer	−0.57 (0.11)	−0.46 (0.21)
	Winter	−1.13 (0.19)	−1.22 (0.16)
SUHANA	Summer	***0.47 (0.028)***	***0.71 (0.0001)***
	Winter	***1.43 (0.01)***	0.80 (0.10)
SVEAGRUVA	Summer	−0.53 (0.45)	−0.42 (0.40)
	Winter	–	–
SVETLOLOBOVO	Summer	***1.22 (0.009)***	***1.14 (0.01)***
	Winter	0.43 (0.41)	0.38 (0.55)
WISEMAN	Summer	0.34 (0.62)	0.34 (0.18)
	Winter	−1.39 (0.44)	−0.71 (0.08)
WRIGLEY A	Summer	0.54 (0.23)	0.17 (0.29)
	Winter	−1.33 (0.39)	−0.81 (0.46)

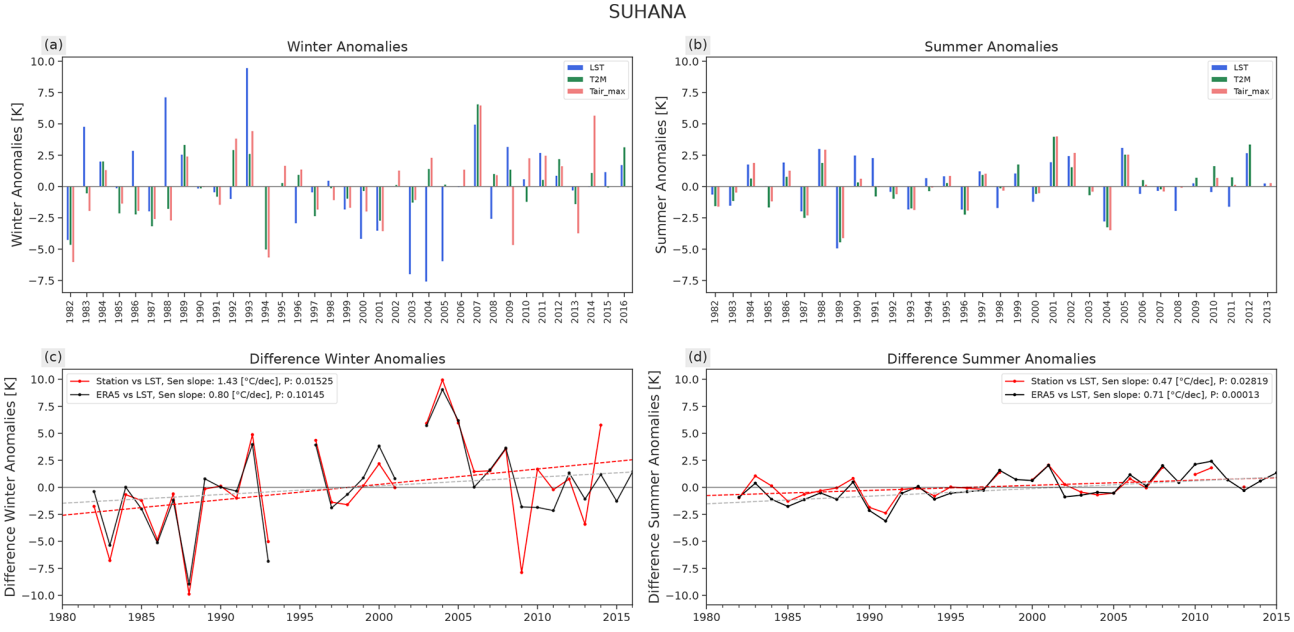

Figure 11. Winter and summer anomalies and the difference between anomalies of LST, T2M, and Tair time series for SUHANA (Siberia).

LST values (Obu et al., 2019). The Verkhoyansk Mountains (Yakutia) present a pronounced winter warming trend. During summer that area does not show a significant temperature trend.

During the winter period, pronounced warming can also be observed in the south of Greenland as well as in eastern Canada (Fig. 12c). Vandecrux et al. (2024) analysed firn and ice temperature at 10 m below the surface (T10m) across the Greenland Ice Sheet and found a general warming trend across the ice sheet. Parts of southern Siberia show negative winter and summer trends. Similar cooling trends for winter are also visible in the AVHRR Polar Pathfinder product over Siberia, as well as warming trends along the Siberian north coast during winter (Key et al., 2016). Large areas at latitudes $> 70°$ suffer from persistent cloud cover and are not covered by satellite LST data. Summer LST trends reveal warming in the northern Siberian lowlands and the Lena Delta area. Northern Canada and most of Greenland also show warming. Cooling trends are visible in the summer period in the south of Siberia and southern Canada.

5 Discussion

5.1 Datasets

5.1.1 EUMETSAT AVHRR FDR

The AVHRR data used in this study are not considered a fundamental climate data record (FCDR) because the calibration for the thermal channels is satellite-specific, and uncertainty quantification is missing. The absence of thermal inter-calibration of the IR sensors on board AVHRR might introduce sensor discontinuities in the dataset (EUMETSAT, 2023c). Therefore, the stability of this dataset needs to be thoroughly assessed before using it for climatological analyses. Solutions to fix this issue might be implemented in the next release of the FDR. However, in the present study, sensor discontinuities were not detected. AVHRR data are affected by orbital drift and, although this effect is negligible for the Northern Hemisphere and monthly means, this still might introduce artificial trends. Different solutions to correct for orbital drift exist (Ma et al., 2020; Julien and Sobrino, 2022). However, for the moment, none of the existing LST datasets applies such a correction to the entire AVHRR time period. The present study revealed artificial trends during summer months for a few stations when compared to ERA5-Land T2M and in situ air temperature.

5.1.2 ERA5-Land T2M

T2M data from reanalyses have the advantage that they are continuous and free of data gaps. However, the coarse resolution of ERA5-Land (9 km) represents a challenge when compared with variables such as LST, which have high spatial variability and are linked to intrinsic properties of the surface such as roughness and moisture (Hulley et al., 2014). Some discrepancies between ERA5-Land T2M and the EUSTACE station Tair data are visible in the anomaly analysis (Figs. 10 and 11). Previous work performed for the Chinese Qilian Mountains (Zhao and He, 2022) found an average RMSE of 2.2 °C between ERA5-Land and air temperature measurements from weather stations, which was mainly attributed to elevation differences between both data sources. However, ERA5-Land performed well in estimating trends. Another study from the northeast of Brazil (de Araújo et al., 2022), which also compares air temperature data from stations to ERA5-Land, indicates that ERA5-Land generally underestimated average air temperature values.

5.1.3 EUSTACE Tair

EUSTACE air temperature data are homogenized and have undergone break detection and quality checks (Brugnara et al., 2019). Large parts of eastern Siberia, northern Canada, and Alaska only have a few EUSTACE weather stations in their territory (Rayner et al., 2020). In general, stations are placed close to settlements and road access, meaning it is difficult to obtain quality weather station data from remote places. Satellite imagery represents, therefore, a valuable source for obtaining temperature data from remote places. In the current study, many stations were discarded due to their proximity to large water bodies or the coast, which would impact the comparison with remote sensing data. In Greenland, two EUSTACE stations are located in close vicinity to each other (KANGERLUSSUAQ and SONDRESTROM). However, they present very different results when compared to ERA5-Land data and satellite data (see Tables 7 and 9). Both stations are assigned to different ERA5-Land and satellite pixels. This area in Greenland is composed of deep fjords with steep hillslopes, which might not be captured well by the spatially coarse model data and satellite data.

5.1.4 Pan-Arctic AVHRR LST

The AVHRR LST dataset developed in this study covers 40 years and represents a valuable data source, complementing data from models and weather stations for obtaining temperature information at a hemispheric scale. A dynamic snow and vegetation cover mask is integrated into the LST algorithm to assign correct emissivities to snow-covered pixels. This is particularly important for cryospheric research at high latitudes. Most of the existing LST datasets do not use a dynamic snow mask and assign snow emissivity based on a static LSE map (Ma et al., 2020). Snow cover onset and snow melt onset are events that can be captured by satellite thermal imagery and are of particular interest for the thermal regime of the ground (Grünberg et al., 2020; Hammar et al., 2023). Westermann et al. (2012) highlighted the importance of using an accurate cloud mask when using LST data for permafrost modelling, as this can lead to high uncertainties and a lack of

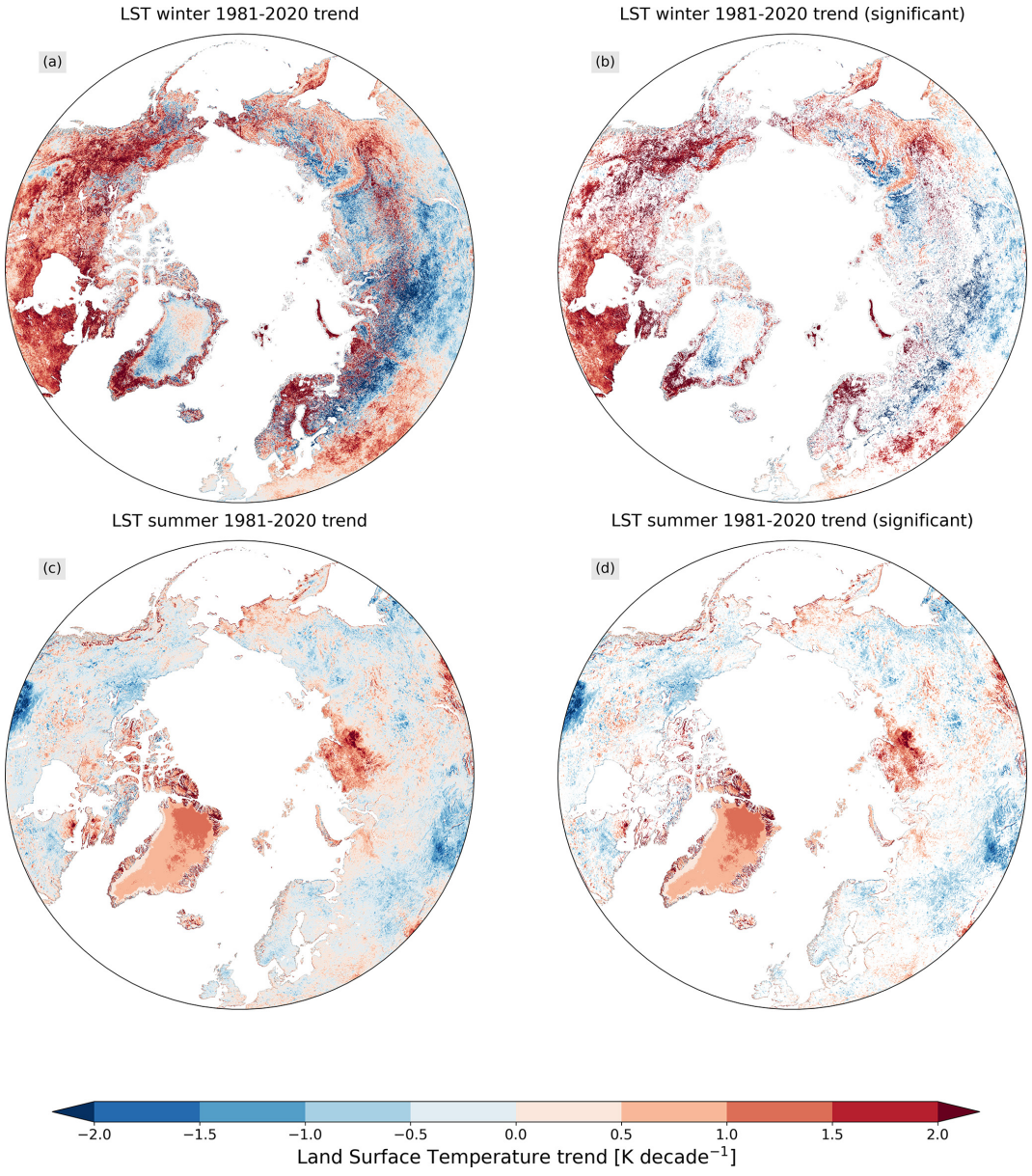

Figure 12. Panels **(a)** and **(b)** show the trends for the winter period and **(c)** and **(d)** the trends for the summer period. The left panels **(a, c)** show the trends independent of their significance. The right panels **(b, d)** show only the significant trends; i.e. areas with statistically insignificant trends ($p > 0.1$) are masked out.

accuracy. The pan-Arctic AVHRR LST dataset incorporates the latest cloud mask from the CLARA-A3 database (Karlsson et al., 2023b), and a low cloud probability threshold (see Sect. 3) was used to avoid cloud contamination.

LST trends computed at different stations (Table 8) are positive and lie in the same ranges as trends computed for Tair and T2M. Furthermore, LST temperature trend maps (Fig. 12) highlight areas that are particularly sensitive to Arctic amplification and present pronounced warming trends. LST trends computed for the summer and winter periods at a hemispheric scale reveal distinct warming areas as well as some regions with cooling trends. Maturilli et al. (2019) determined air temperature trends from weather station data in Ny-Ålesund (Svalbard) from 1993 to 2017 and found that the strongest warming trend occurred during the winter season. They found a summer warming trend of +0.6 K per decade, which corresponds to the LST warming trend for Svalbard shown in Fig. 12. Compared to air temperature trends, satellite-derived LST trends can present a cold bias as only clear-sky days are considered in the LST generation (Westermann et al., 2017). Therefore, the all-sky LST trends might be even higher. All-weather LST datasets have been

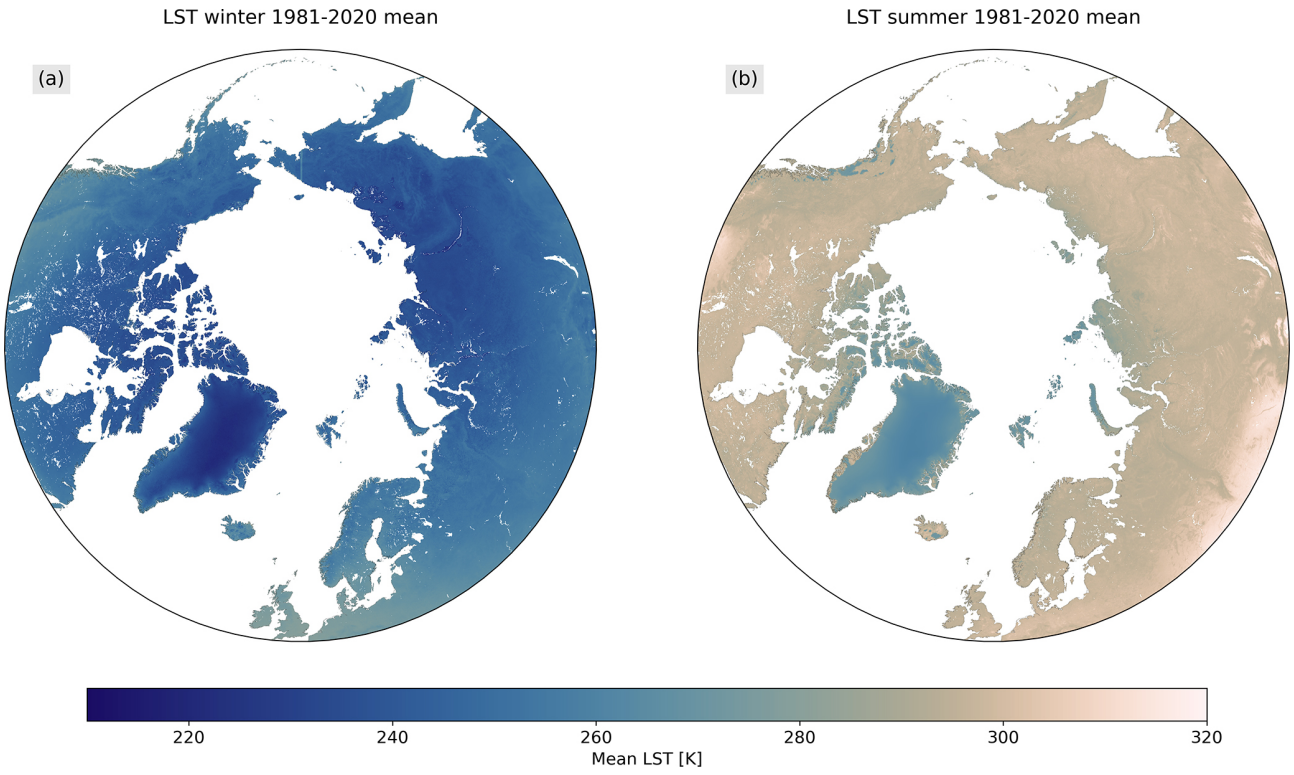

Figure 13. Mean LST (1981–2020) for **(a)** winter and **(b)** summer.

generated in the past by combining energy balance modelling or reanalysis LST data with LST TIR data (Martins et al., 2019; Zhang et al., 2021).

5.2 Point to pixel comparison

The landscape in the pan-Arctic region is very heterogeneous; for example, complex wetland systems are difficult to map from satellite imagery (Olefeldt et al., 2021; Palmtag et al., 2022). Land cover data such as from the CCI project are thus prone to high uncertainties in the high northern latitudes (Bartsch et al., 2024). The nominal spatial resolution of AVHRR GAC pixel is 4 km. Air temperature measured at a nearby location might differ considerably from the corresponding AVHRR LST and depends on vegetation type and water content of the ground.

5.3 Comparison of LST, Tair, and T2M

Day length and, consequently, solar irradiance cover a wide range in the Arctic. During winter, there is constant night, and Tair is in close agreement with LST (Hachem et al., 2012; Urban et al., 2013). Furthermore, during winter, most of the pan-Arctic region is covered by snow. Snow cover variations directly influence LST (Thiebault and Young, 2020). LST and air temperature anomalies exhibit a strong correlation (mean $(r) > 0.6$), and LST anomalies show a similar pattern to air temperature anomalies (Figs. 10 and 11). LST anomalies present a greater magnitude, which can be explained by a greater amplitude in the LST diurnal cycle than in the air temperature diurnal cycle. Differences in winter anomalies are more pronounced than summer anomalies. This can be explained by the clear-sky bias that occurs in satellite LST data (Westermann et al., 2012): cloud cover affects the winter period more than the summer period. Such a cooling bias was also observed over the Greenland Ice Sheet over ice surface temperature (Hall et al., 2012). Regarding monthly means, the air temperature mean is higher than the LST mean value, which only considers clear-sky days. Under cloudy skies, air and surface temperature are in close agreement (Obu et al., 2019). A few missing data occurred in the EUSTACE, which made the ERA5-Land data more reliable for such an analysis. However, EUSTACE records two temperatures per day: T_{max} and T_{min}. Therefore, for analysis at daily frequency, EUSTACE might be more suitable, especially for satellites having an overpass time close to noon. Winter data proved to be stable over all investigated stations. This makes this dataset particularly interesting for studying the winter months in the high latitudes. The winter period is a particularly active time for the ecosystem in the Arctic (Berge et al., 2015), but it tends to be understudied.

Finally, trends for the three different datasets computed at stations with good temporal stability (Table 8) showed good agreement, with LST trends generally exhibiting stronger positive trends than air temperature. In addition, LST winter

trends for central Siberia (Fig. 12) showed similar values as in Waring et al. (2023). Strong summer warming in the north of the Krasnoyarsk region and northern Nunavut (Canada) (visible in Fig. 12d) is also visible in ERA5-Land temperature data for the period 1984–2022 in the work of Larocca et al. (2024). The regions showing positive trends in LST are generally associated with lichens, moss, and herbaceous land cover, whereas regions with negative trends are mostly located in forested areas. Mildrexler et al. (2011) revealed the cooling effect of forests on LST, which might slow the general temperature increase in the Arctic. The presence of thermokarst lakes and wetlands as depicted in Olefeldt et al. (2021) might also slow the warming trends slightly. The Yamal Peninsula, for example, contains many water bodies and thermokarst lakes and shows a slower warming trend than an area with a majority of barren soil. Hemispheric LST data can also be used to highlight fast-warming areas that might lead to abrupt permafrost thaw, which in turn influences carbon fluxes (Treat et al., 2024). LST also has the advantage over air temperature of being more sensitive to changes in vegetation density (Mildrexler et al., 2011), which makes LST a particularly interesting variable for cryospheric research.

6 Conclusions

This study presents the workflow to derive a pan-Arctic LST dataset from the EUMETSAT AVHRR GAC FDR and validates the new LST product against in situ LST. AVHRR LST is derived with the generalized split-window algorithm (Wan and Dozier, 1996), and the corresponding RTM is performed with RTTOV v.13, based on a new calibration database (Ermida and Trigo, 2022). This ensures an optimal representation of atmospheric conditions in the pan-Arctic region. This LST dataset utilizes a recent cloud mask with notable improvements compared to previous cloud products (Karlsson et al., 2023b), which is of particular importance for LST retrieval. The pan-Arctic AVHRR LST product showed good performance, and validation results lie in the range of similar products. The new LST product is assessed for stability in the pan-Arctic region by comparing it to air temperature data from weather stations and T2M data from ERA5-Land. Twelve weather stations belonging to the EUSTACE global station dataset are chosen based on several criteria: latitude $> 50^\circ$ N, a minimum of 30 years of overlapping data, and homogeneous land cover over at least 1 GAC pixel. LST trends and variability are compared to ERA5-Land T2M and EUSTACE Tair maximum data. The correlation coefficients between the datasets indicate good agreement: $r > 0.9$ for the monthly mean correlation and ~ 0.5 to ~ 0.8 for the anomaly analysis. The analysis of the differences in anomalies showed slightly significant trends for the summer months but no artificial trends in the data during the winter months when solar irradiance is absent. LST trends for the winter and summer periods were computed for the entire pan-Arctic region and revealed spatially varying trends. In winter, positive trends in the south of Greenland, Siberia, and eastern Canada were revealed. Summer temperature trends highlight the fast warming of the Greenland Ice Sheet and a large region in the vicinity of the Lena Delta. Its good accuracy and emissivity retrieval based on dynamical vegetation and snow masks make this product suitable for a wide range of research applications in the pan-Arctic region. However, future research investigate a possible spatial upsampling of the dataset (e.g. to 1 km spatial resolution). Orbital drift correction with a robust method would allow one to extend the study to lower latitudes.

Appendix A: Anomaly differences

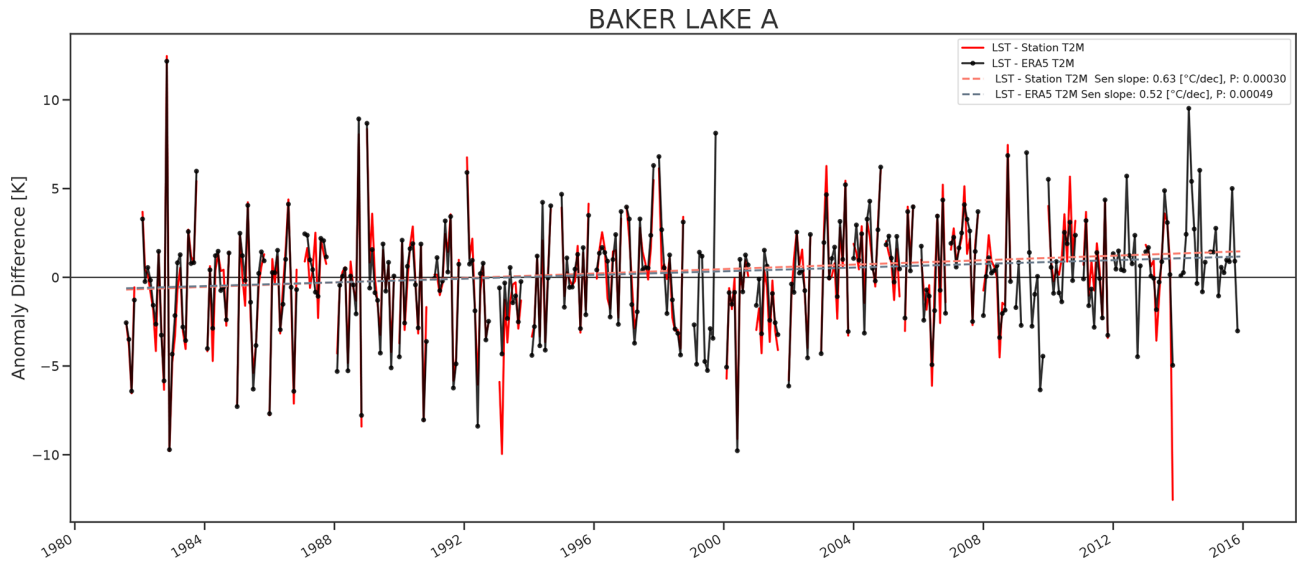

Figure A1. Differences in the anomalies at BAKER LAKE A (Canada) as a time series.

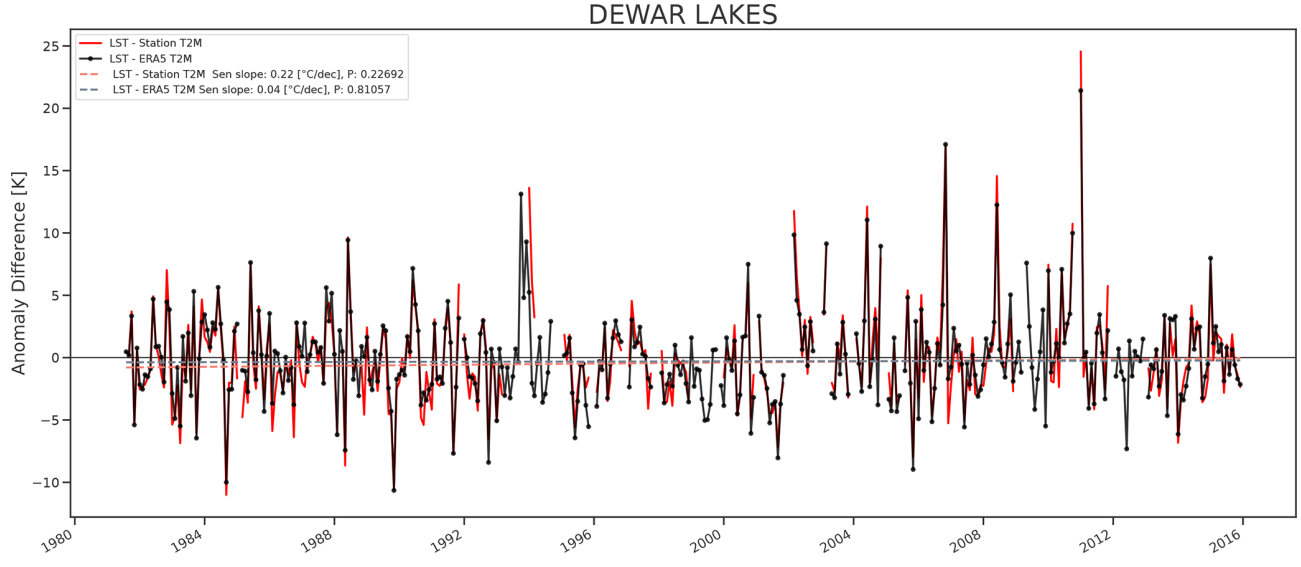

Figure A2. Differences in the anomalies at DEWAR LAKES (Canada) as a time series.

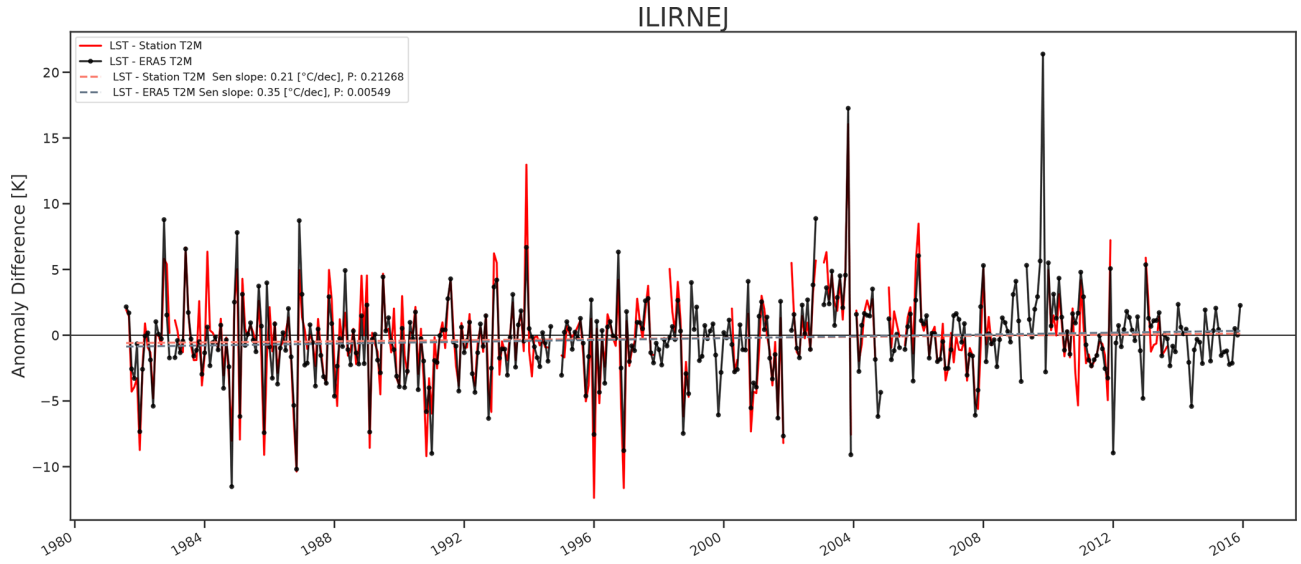

Figure A3. Differences in the anomalies at ILIRNEJ (eastern Siberia) as a time series.

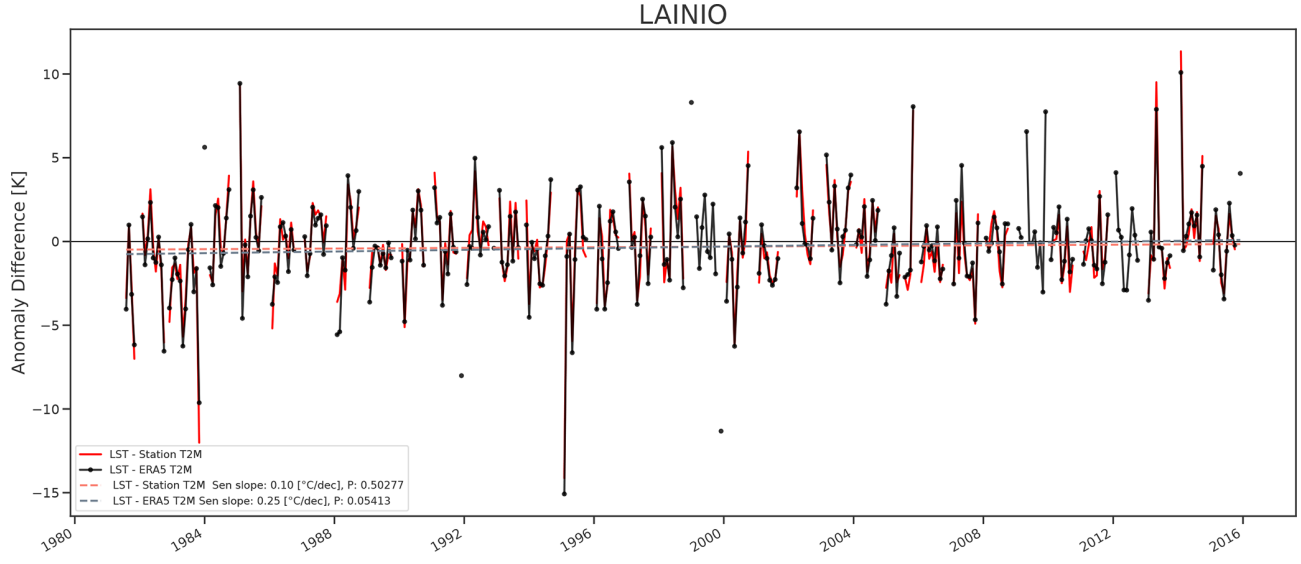

Figure A4. Differences in the anomalies at LAINIO (Norway) as a time series.

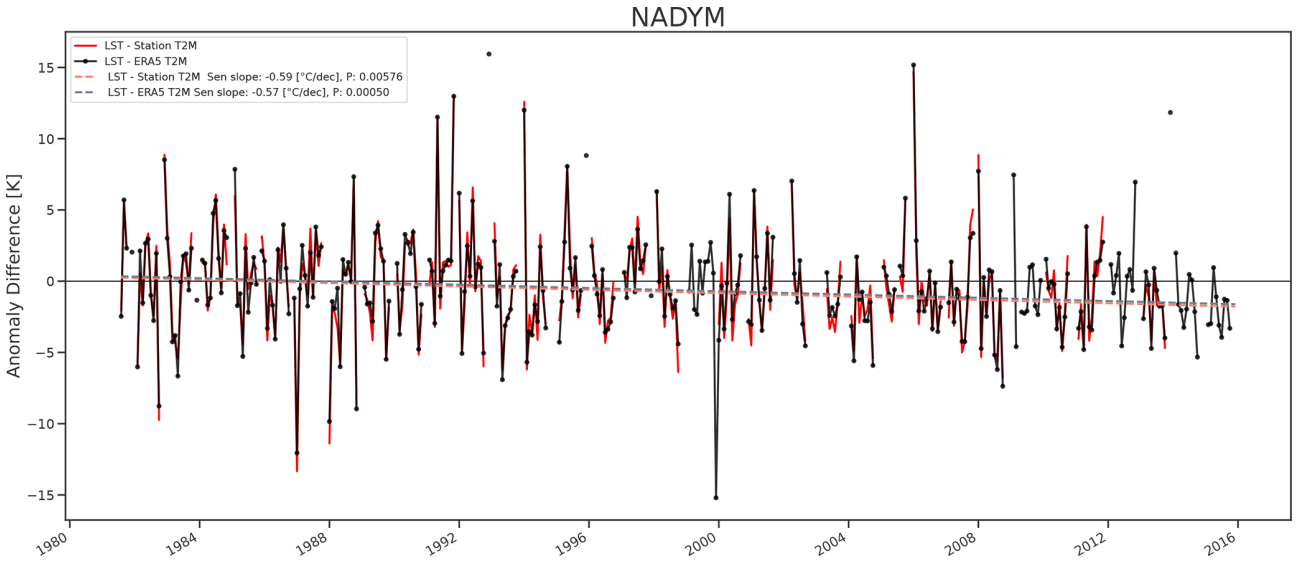

Figure A5. Differences in the anomalies at NADYM (Siberia) as a time series.

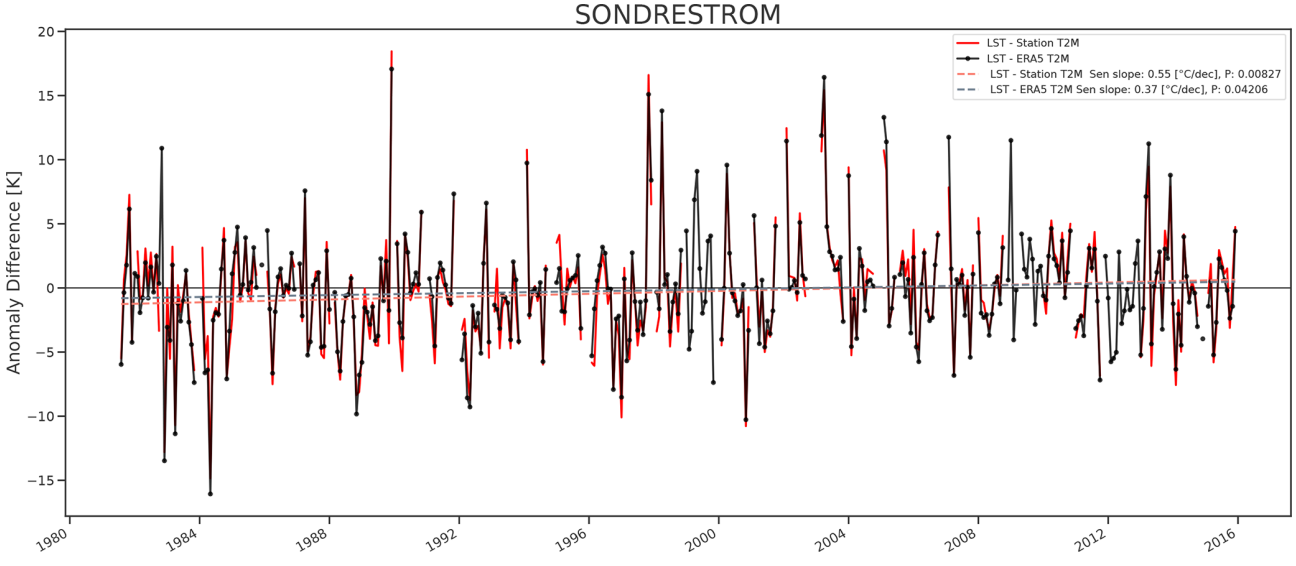

Figure A6. Differences in the anomalies at SONDRESTROM (Greenland) as a time series.

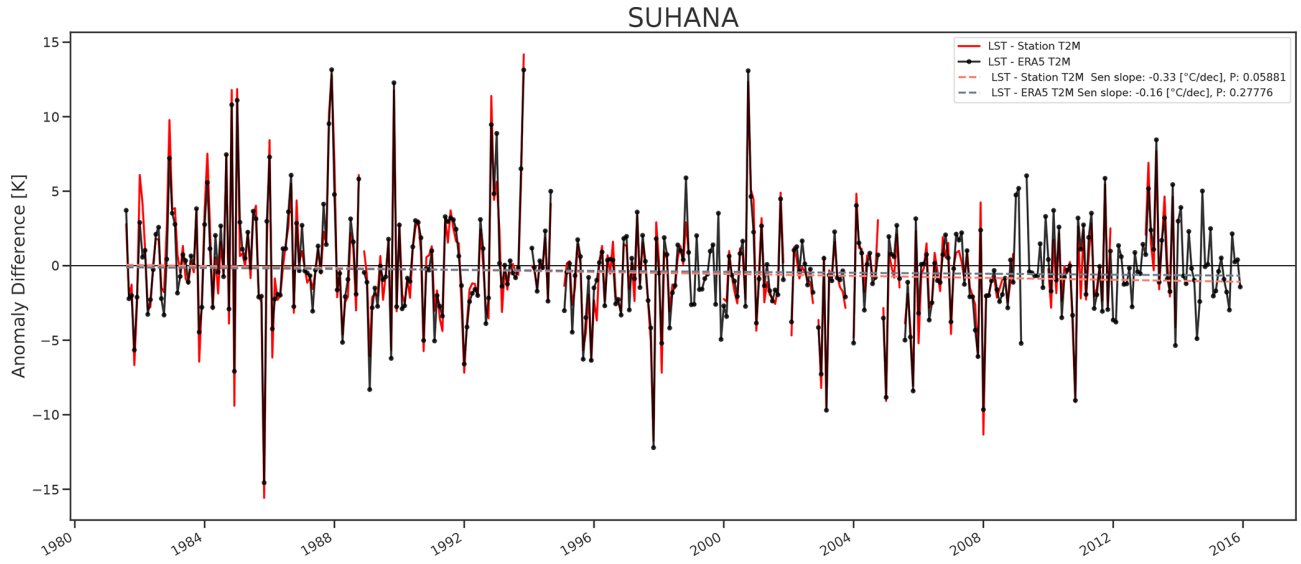

Figure A7. Differences in the anomalies at SUHANA (Siberia) as a time series.

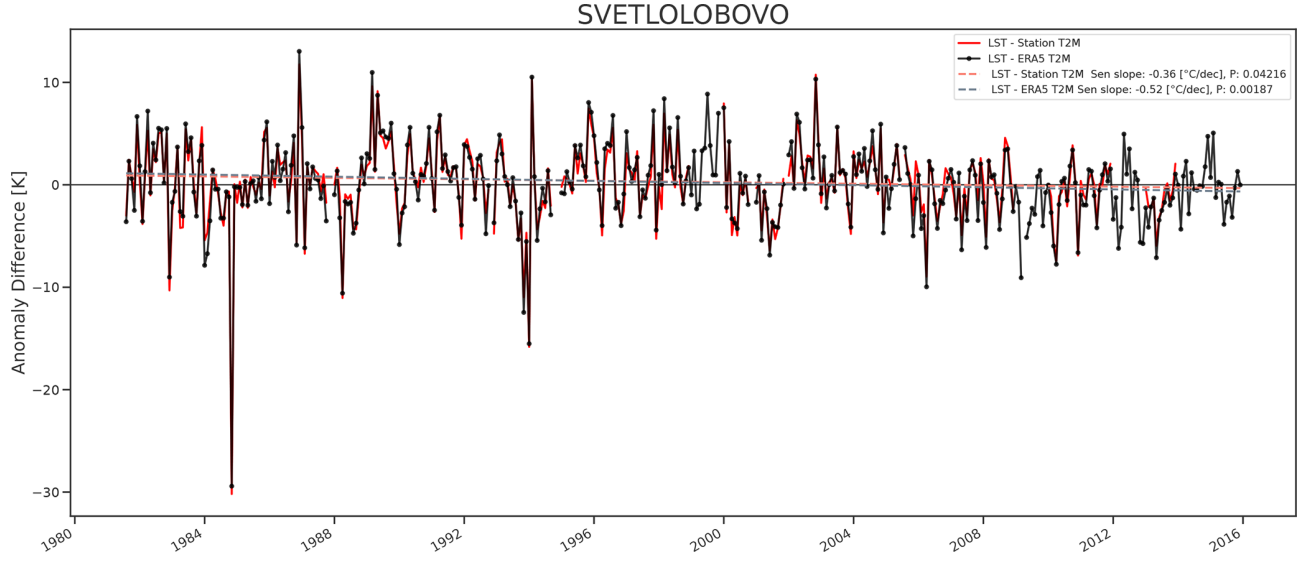

Figure A8. Differences in the anomalies at SVETLOLOBOVO (Siberia) as a time series.

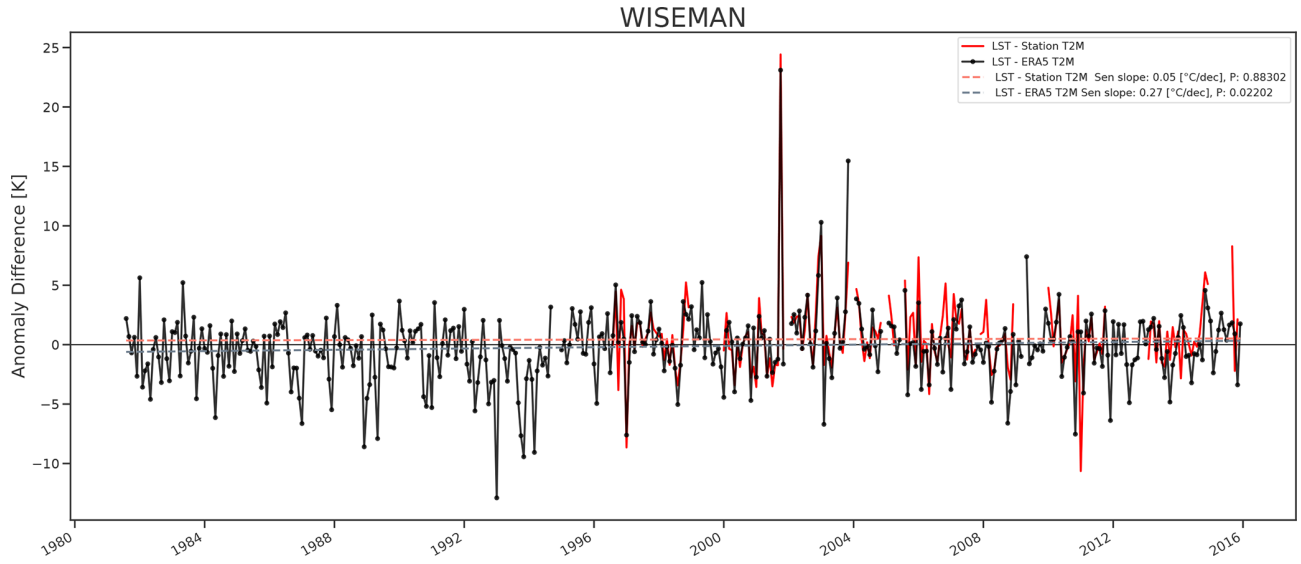

Figure A9. Differences in the anomalies at WISEMAN (Canada) as a time series.

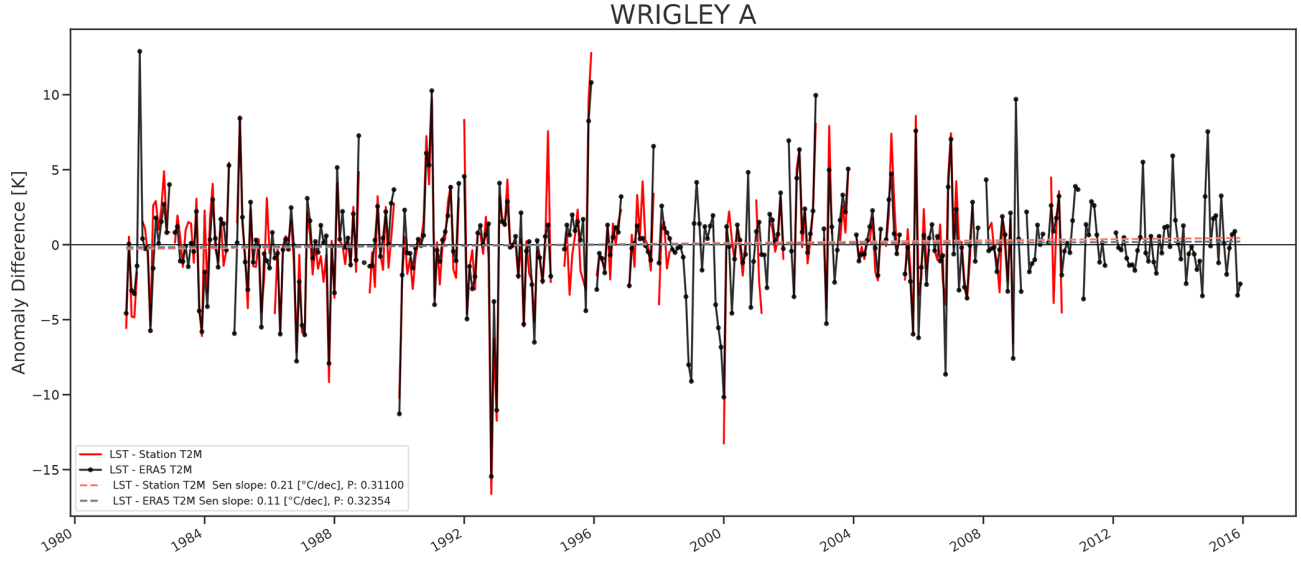

Figure A10. Differences in the anomalies at WRIGLEY A (Alaska) as a time series.

Data availability. The EUMETSAT AVHRR FDR, the basis for this study, is available through the EUMETSAT Data Portal: https://doi.org/10.15770/EUM_SEC_CLM_0060 (EUMETSAT, 2023a). The pan-Arctic LST monthly means used in this study are available on Zenodo: https://doi.org/10.5281/zenodo.13361744 (Dupuis et al., 2024). The probabilistic cloud masks from CLARA-A3 can be found here: https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003 (EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF), 2023). The land cover data, which were used for the emissivity maps and the water masks, can be obtained from the Copernicus Climate Data Store: https://doi.org/10.24381/cds.006f2c9a (Copernicus Climate Change Service, Climate Data Store, 2019). In addition, the CCI Snow project datasets are available here: https://doi.org/10.5285/4647cc9ad3c044439d6c643208d3c494 (Luojus et al., 2022) for the SWE dataset, and https://doi.org/10.5285/7491427f8c3442ce825ba5472c224322 (Xiao et al., 2024) for the fractional snow cover dataset. The clear-sky database used for the radiative transfer modelling is accessible through Zenodo (https://doi.org/10.5281/zenodo.5779543, Ermida and Trigo, 2021).

Author contributions. The idea was conceptualized by SD and refined by SW. The software development, data processing, data analysis, and writing were mainly conducted by SD. FMG and SW supervised the manuscript writing and provided supervision of the research.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We thank the Dr. Alfred Bretscher-Fonds für Klima- und Luftverschmutzungsforschung from the University of Bern for funding this project. We also acknowledge the team at EUMETSAT that compiled the AVHRR FDR data record, Sofia Ermida and Isabel Trigo for compiling and providing the clear-sky database, and the in situ data (SURFRAD, EUSTACE, and BSRN) providers. We acknowledge the use of the scientific colour maps developed by Fabio Crameri (https://doi.org/10.5281/zenodo.1243862, Crameri, 2023). We also want to mention the open-source Python library Xarray, which has been a great resource for this project. We acknowledge the two anonymous reviewers, who helped to improve the manuscript.

Financial support. This work was supported by the Dr. Alfred Bretscher-Fonds für Klima- und Luftverschmutzungsforschung from the University of Bern.

Review statement. This paper was edited by Marie Dumont and reviewed by two anonymous referees.

References

Arctic Monitoring and Assessment Programme (AMAP): Arctic Climate Change Update 2021: Key Trends and Impacts. Summary for Policy-makers, AMAP, 2021.

Bartsch, A., Strozzi, T., and Nitze, I.: Permafrost Monitoring from Space, Surv. Geophys., 44, 1579–1613, https://doi.org/10.1007/s10712-023-09770-3, 2023.

Bartsch, A., Efimova, A., Widhalm, B., Muri, X., von Baeckmann, C., Bergstedt, H., Ermokhina, K., Hugelius, G., Heim, B., and Leibman, M.: Circumarctic land cover diversity considering wetness gradients, Hydrol. Earth Syst. Sci., 28, 2421–2481, https://doi.org/10.5194/hess-28-2421-2024, 2024.

Batbaatar, J., Gillespie, A. R., Sletten, R. S., Mushkin, A., Amit, R., Trombotto Liaudat, D., Liu, L., and Petrie, G.: Toward the Detection of Permafrost Using Land-Surface Temperature Mapping, Remote Sens., 12, 695, https://doi.org/10.3390/rs12040695, 2020.

Berge, J., Daase, M., Renaud, P. E., Ambrose, W. G., Darnis, G., Last, K. S., Leu, E., Cohen, J. H., Johnsen, G., Moline, M. A., Cottier, F., Varpe, O., Shunatova, N., Bałazy, P., Morata, N., Massabuau, J. C., Falk-Petersen, S., Kosobokova, K., Hoppe, C. J., Węsławski, J. M., Kukliński, P., Legeżyńska, J., Nikishina, D., Cusa, M., Kędra, M., Włodarska-Kowalczuk, M., Vogedes, D., Camus, L., Tran, D., Michaud, E., Gabrielsen, T. M., Granovitch, A., Gonchar, A., Krapp, R., and Callesen, T. A.: Unexpected levels of biological activity during the polar night offer new perspectives on a warming arctic, Curr. Biol., 25, 2555–2561, https://doi.org/10.1016/j.cub.2015.08.024, 2015.

Borbas, E., Seemann, S. W., Huang, H.-L., Li, J. Y., and Menzel, W. P.: Global profile training database for satellite regression retrievals with estimates of skin temperature and emissivity, in: Proceedings of the International TOVS Study Conference-XIV, Beijing, China, 25–31 May 2005, https://cimss.ssec.wisc.edu/training_data/data/borbas_trainingData_poster_ITSC14.pdf (last access: 17 December 2024), 2005.

Brugnara, Y., Good, E., Squintu, A. A., van der Schrier, G., and Brönnimann, S.: The EUSTACE global land station daily air temperature dataset, Geosci. Data J., 6, 189–204, https://doi.org/10.1002/gdj3.81, 2019.

Carlson, T. N. and Ripley, D. A.: On the relation between NDVI, fractional vegetation cover, and leaf area index, Remote Sens. Environ., 62, 241–252, https://doi.org/10.1016/S0034-4257(97)00104-1, 1997.

Cheng, J., Liang, S., Yao, Y., and Zhang, X.: Estimating the Optimal Broadband Emissivity Spectral Range for Calculating Surface Longwave Net Radiation, IEEE Geosci. Remote Sens. Lett., 10, 401–405, https://doi.org/10.1109/LGRS.2012.2206367, 2013.

Chevallier, F., Chédin, A., Cheruy, F., and Morcrette, J.-J.: TIGR-like atmospheric-profile databases for accurate radiative-

flux computation, Q. J. R. Meteorol. Soc., 126, 777–785, https://doi.org/10.1002/qj.49712656319, 2000.

Christensen, T. R., Johansson, T., Åkerman, H. J., Mastepanov, M., Malmer, N., Friborg, T., Crill, P., and Svensson, B. H.: Thawing sub-arctic permafrost: Effects on vegetation and methane emissions, Geophys. Res. Lett., 31, L04501, https://doi.org/10.1029/2003GL018680, 2004.

Chylek, P., Folland, C., Klett, J. D., Wang, M., Hengartner, N., Lesins, G., and Dubey, M. K.: Annual Mean Arctic Amplification 1970–2020: Observed and Simulated by CMIP6 Climate Models, Geophys. Res. Lett., 49, e2022GL099371, https://doi.org/10.1029/2022GL099371, 2022.

Coll, C., Hook, S. J., and Galve, J. M.: Land surface temperature from the advanced along-track scanning radiometer: Validation over inland waters and vegetated surfaces, IEEE T. Geosci. Remote, 47, 350–360, https://doi.org/10.1109/TGRS.2008.2002912, 2009.

Copernicus Climate Change Service, Climate Data Store: Land cover classification gridded maps from 1992 to present derived from satellite observation, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.006f2c9a (last access: 1 March 2024), 2019.

Copernicus digital elevation model (DEM) GLO-90: https://doi.org/10.5270/ESA-c5d3d65, last access: 1 March 2024.

Cox, C. J. and Halliwell, D.: Basic measurements of radiation at station Alert (2004-08–2014-06), PANGAEA [data set], https://doi.org/10.1594/PANGAEA.932867, 2021.

Crameri, F.: Scientific colour maps (8.0.1), Zenodo [code], https://doi.org/10.5281/zenodo.8409685, 2023.

Dada, L., Angot, H., Beck, I., Baccarini, A., Quéléver, L. L. J., Boyer, M., Laurila, T., Brasseur, Z., Jozef, G., de Boer, G., Shupe, M. D., Henning, S., Bucci, S., Dütsch, M., Stohl, A., Petäjä, T., Daellenbach, K. R., Jokinen, T., and Schmale, J.: A central arctic extreme aerosol event triggered by a warm air-mass intrusion, Nat. Commun., 13, 5290, https://doi.org/10.1038/s41467-022-32872-2, 2022.

de Araújo, C. S. P., Campos e Silva, I. A., Ippolito, M., and Guedes Cabral de Almeida, C. D.: Evaluation of air temperature estimated by ERA5-Land reanalysis using surface data in Pernambuco, Brazil, Environ. Monit. Assess., 194, 381, https://doi.org/10.1007/s10661-022-10047-2, 2022.

Devasthale, A. and Karlsson, K.-G.: Decadal Stability and Trends in the Global Cloud Amount and Cloud Top Temperature in the Satellite-Based Climate Data Records, Remote Sens., 15, 3819, https://doi.org/10.3390/rs15153819, 2023.

Dupuis, S., Wunderle, S., and Göttsche, F.: 40-year monthly mean AVHRR GAC Land Surface Temperature data for the Pan-Arctic region (Pan-Arctic AVHRR LST), Zenodo [data set], https://doi.org/10.5281/zenodo.13361744, 2024.

Ermida, S. L. and Trigo, I. F.: Clear-sky profile database for the development of Land Surface Temperature algorithms (v0.0.0) [data set], Zenodo, https://doi.org/10.5281/zenodo.5779543, 2021.

Ermida, S. and Trigo, I.: A Comprehensive Clear-Sky Database for the Development of Land Surface Temperature Algorithms, Remote Sens., 14, 2329, https://doi.org/10.3390/rs14102329, 2022.

Ermida, S. L., Trigo, I. F., DaCamara, C. C., Göttsche, F. M., Olesen, F. S., and Hulley, G.: Validation of remotely sensed surface temperature over an oak woodland landscape – The problem of viewing and illumination geometries, Remote Sens. Environ., 148, 16–27, https://doi.org/10.1016/j.rse.2014.03.016, 2014.

Ermida, S. L., Jiménez, C., Prigent, C., Trigo, I. F., and DaCamara, C. C.: Inversion of AMSR-E observations for land surface temperature estimation: 2. Global comparison with infrared satellite temperature, J. Geophys. Res., 122, 3348–3360, https://doi.org/10.1002/2016JD026148, 2017.

EUMETSAT: AVHRR Fundamental Data Record – Release 1 – Multimission, EUMETSAT [data set], https://doi.org/10.15770/EUM_SEC_CLM_0060, 2023a.

EUMETSAT: PyGAC FDR ATBD, Tech. rep., EUMETSAT, ISBN 4961518077, https://doi.org/10.15770/EUM_SEC_CLM_0060, 2023b.

EUMETSAT: PyGAC AVHRR FDR Release 1 Product Users Guide, Tech. rep., EUMETSAT, https://doi.org/10.15770/EUM_SEC_CLM_0060, 2023c.

EUMETSAT: PyGAC AVHRR FDR Release 1 Validation Report, Tech. rep., EUMETSAT, https://doi.org/10.15770/EUM_SEC_CLM_0060, 2023d.

EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF): Product User Manual CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3) Cloud Products, Tech. rep., EUMETSAT CM SAF, https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003, 2023.

Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., and Göttsche, F. M.: Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat, IEEE T. Geosci. Remote, 48, 523–534, https://doi.org/10.1109/TGRS.2009.2027697, 2010.

Ghent, D., Dodd, E., Veal, K., Perry, M., Jimenez, C., and Ermida, S.: Algorithm Theoretical Basis Document: WP2.1 – DEL-LST-CCI-D2.2-ATBD, Tech. rep., LST_CCI, https://admin.climate.esa.int/media/documents/LST-CCI-D2.2-ATBD_-_i4r0_-_Algorithm_Theoretical_Basis_Document.pdf (last access: 17 December 2024), 2023.

Gleason, A. C., Prince, S. D., Goetz, S. J., and Small, J.: Effects of orbital drift on land surface temperature measured by AVHRR thermal sensors, Remote Sens. Environ., 79, 147–165, https://doi.org/10.1016/S0034-4257(01)00269-3, 2002.

Good, E. J.: An in situ-based analysis of the relationship between land surface "skin" and screen-level air temperatures, J. Geophys. Res., 121, 8801–8819, https://doi.org/10.1002/2016JD025318, 2016.

Good, E. J., Aldred, F. M., Ghent, D. J., Veal, K. L., and Jimenez, C.: An analysis of the stability and trends in the LST_cci Land Surface Temperature datasets over Europe, Earth Space Sci., 9, e2022EA002317, https://doi.org/10.1029/2022EA002317, 2022.

Grünberg, I., Wilcox, E. J., Zwieback, S., Marsh, P., and Boike, J.: Linking tundra vegetation, snow, soil temperature, and permafrost, Biogeosciences, 17, 4261–4279, https://doi.org/10.5194/bg-17-4261-2020, 2020.

Guillevic, P., Bork-Unkelbach, A., Gottsche, F. M., Hulley, G., Gastellu-Etchegorry, J. P., Olesen, F. S., and Privette, J. L.: Directional viewing effects on satellite land surface temperature products over sparse vegetation canopies-a multi-

sensor analysis, IEEE Geosci. Remote S., 10, 1464–1468, https://doi.org/10.1109/LGRS.2013.2260319, 2013.

Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J. A., Remedios, J., Román, M., and Camacho, F.: Land Surface Temperature Product Validation Best Practice Protocol. Version 1.1, Tech. rep., Committee on Earth Observation Satellites Working Group on Calibration and Validation Land Product Validation Subgroup, https://doi.org/10.5067/doc/ceoswgcv/lpv/lst.001, 2018.

Göttsche, F., Olesen, F., Trigo, I., Bork-Unkelbach, A., and Martin, M.: Long Term Validation of Land Surface Temperature Retrieved from MSG/SEVIRI with Continuous in-Situ Measurements in Africa, Remote Sens., 8, 410, https://doi.org/10.3390/rs8050410, 2016.

Hachem, S., Duguay, C. R., and Allard, M.: Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain, The Cryosphere, 6, 51–69, https://doi.org/10.5194/tc-6-51-2012, 2012.

Hall, D. K., Comiso, J. C., Digirolamo, N. E., Shuman, C. A., Key, J. R., and Koenig, L. S.: A satellite-derived climate-quality data record of the clear-sky surface temperature of the greenland ice sheet, J. Climate, 25, 4785–4798, https://doi.org/10.1175/JCLI-D-11-00365.1, 2012.

Hammar, J., Grünberg, I., Kokelj, S. V., van der Sluijs, J., and Boike, J.: Snow accumulation, albedo and melt patterns following road construction on permafrost, Inuvik–Tuktoyaktuk Highway, Canada, The Cryosphere, 17, 5357–5372, https://doi.org/10.5194/tc-17-5357-2023, 2023.

Hocking, J., Roquet, P., and Brunel, P.: Python/C/C++ wrapper for RTTOV v13, NWP SAF (EUMETSAT), https://nwp-saf.eumetsat.int/site/download/documentation/rtm/docs_rttov13/rttov-wrapper.pdf (last access: 17 December 2024), 2021.

Hulley, G., Veraverbeke, S., and Hook, S.: Thermal-based techniques for land cover change detection using a new dynamic MODIS multispectral emissivity product (MOD21), Remote Sens. Environ., 140, 755–765, https://doi.org/10.1016/j.rse.2013.10.014, 2014.

Hulley, G. C., Hook, S. J., Abbott, E., Malakar, N., Islam, T., and Abrams, M.: The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale, Geophys. Res. Lett., 42, 7966–7976, https://doi.org/10.1002/2015GL065564, 2015.

Hussain, M. and Mahmud, I.: pyMannKendall: a python package for non parametric Mann Kendall family of trend tests, J. Open Source Softw., 4, 1556, https://doi.org/10.21105/joss.01556, 2019.

Ignatov, A., Laszlo, I., Harrod, E. D., Kidwell, K. B., and Goodrum, G. P.: Equator crossing times for NOAA, ERS and EOS sun-synchronous satellites, Int. J. Remote Sens., 25, 5255–5266, https://doi.org/10.1080/01431160410001712981, 2004.

IPCC: 2021, Summary for Policymakers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., 3–32, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896.001, 2021.

Jiménez, C., Prigent, C., Ermida, S. L., and Moncet, J. L.: Inversion of AMSR-E observations for land surface temperature estimation: 1. Methodology and evaluation with station temperature, J. Geophys. Res., 122, 3330–3347, https://doi.org/10.1002/2016JD026144, 2017.

Julien, Y. and Sobrino, J. A.: Toward a Reliable Correction of NOAA AVHRR Orbital Drift, Front. Remote Sens., 3, 851933, https://doi.org/10.3389/frsen.2022.851933, 2022.

Karlsson, K.-G., Riihelä, AkuandTrentmann, J., Stengel, Martin andSolodovnik, I., Meirink, J. F., Devasthale, A., Jääskeläinen, E., Kallio-Myers, V., Eliasson, S., Benas, N., Johansson, E., Stein, D., Finkensieper, S., Håkansson, N., Akkermans, T., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: CM SAF cLoud, Albedo and surface RAdiation dataset from AVHRR data – Edition 3, Satellite Application Facility on Climate Monitoring (CM SAF), [data set], https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003, 2023a.

Karlsson, K.-G., Stengel, M., Meirink, J. F., Riihelä, A., Trentmann, J., Akkermans, T., Stein, D., Devasthale, A., Eliasson, S., Johansson, E., Håkansson, N., Solodovnik, I., Benas, N., Clerbaux, N., Selbach, N., Schröder, M., and Hollmann, R.: CLARA-A3: The third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023, Earth Syst. Sci. Data, 15, 4901–4926, https://doi.org/10.5194/essd-15-4901-2023, 2023b.

Key, J., Wang, X., Liu, Y., Dworak, R., and Letterly, A.: The AVHRR Polar Pathfinder Climate Data Records, Remote Sens., 8, 167, https://doi.org/10.3390/rs8030167, 2016.

Kidwell, K. B.: NOAA Polar Orbiter Data Users Guide, Tech. rep., National Climatic Data Center (U.S.). National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Climatic Data Center, Satellite Data Services Division, https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.star.nesdis.noaa.gov/mirs/documents/0.0_NOAA_KLM_Users_Guide.pdf (last access: 17 December 2024), 1995.

Kustov, V.: Basic and other measurements of radiation at station Cape Baranova (2016-09), PANGAEA [data set], https://doi.org/10.1594/PANGAEA.892295, 2018.

Kustov, V.: Basic and other measurements of radiation at station Tiksi (2010-06 et seq), PANGAEA [data set], https://doi.org/10.1594/PANGAEA.959213, 2023.

Larocca, L. J., Lea, J. M., Erb, M. P., McKay, N. P., Phillips, M., Lamantia, K. A., and Kaufman, D. S.: Arctic glacier snowline altitudes rise 150 m over the last 4 decades, The Cryosphere, 18, 3591–3611, https://doi.org/10.5194/tc-18-3591-2024, 2024.

Latifovic, R., Pouliot, D., and Dillabaugh, C.: Identification and correction of systematic error in NOAA AVHRR long-term satellite data record, Remote Sens. Environ., 127, 84–97, https://doi.org/10.1016/j.rse.2012.08.032, 2012.

Li, J.-H., Li, Z.-L., Liu, X., and Duan, S.-B.: A global historical twice-daily (daytime and nighttime) land surface temperature dataset produced by Advanced Very High Resolution Radiome-

ter observations from 1981 to 2021, Earth Syst. Sci. Data, 15, 2189–2212, https://doi.org/10.5194/essd-15-2189-2023, 2023a.

Li, K., Guan, K., Jiang, C., Wang, S., Peng, B., and Cai, Y.: Evaluation of Four New Land Surface Temperature (LST) Products in the U.S. Corn Belt: ECOSTRESS, GOES-R, Landsat, and Sentinel-3, IEEE J. Sel. Top. Appl., 14, 9931–9945, https://doi.org/10.1109/JSTARS.2021.3114613, 2021.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sens. Environ., 131, 14–37, https://doi.org/10.1016/j.rse.2012.12.008, 2013.

Li, Z.-L., Wu, H., Duan, S.-B., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B.-H., Yan, G., and Zhou, C.: Satellite remote sensing of global land surface temperature: Definition, methods, products, and applications, Rev. Geophys., 61, e2022RG000777, https://doi.org/10.1029/2022RG000777, 2023b.

Lieberherr, G. and Wunderle, S.: Lake Surface Water Temperature Derived from 35 Years of AVHRR Sensor Data for European Lakes, Remote Sens., 10, 990, https://doi.org/10.3390/rs10070990, 2018.

Lieberherr, G., Riffler, M., and Wunderle, S.: Performance Assessment of Tailored Split-Window Coefficients for the Retrieval of Lake Surface Water Temperature from AVHRR Satellite Data, Remote Sens., 9, 1334, https://doi.org/10.3390/rs9121334, 2017.

Luojus, K., Moisander, M., Pulliainen, J., Takala, M., Lemmetyinen, J., Derksen, C., Mortimer, C., Schwaizer, G., Nagler, T., and Venäläinen, P.: ESA Snow Climate Change Initiative (Snow_cci): Snow Water Equivalent (SWE) level 3C daily global climate research data package (CRDP) (1979–2020), version 2.0, NERC EDS Centre for Environmental Data Analysis [data set], https://doi.org/10.5285/4647cc9ad3c044439d6c643208d3c494, 2022.

Lund, M., Stiegler, C., Abermann, J., Citterio, M., Hansen, B. U., and van As, D.: Spatiotemporal variability in surface energy balance across tundra, snow and ice in Greenland, Ambio, 46, 81–93, https://doi.org/10.1007/s13280-016-0867-5, 2017.

Ma, J., Zhou, J., Göttsche, F.-M., Liang, S., Wang, S., and Li, M.: A global long-term (1981–2000) land surface temperature product for NOAA AVHRR, Earth Syst. Sci. Data, 12, 3247–3268, https://doi.org/10.5194/essd-12-3247-2020, 2020.

Martin, M. A., Ghent, D., Pires, A. C., Göttsche, F.-M., Cermak, J., and Remedios, J. J.: Comprehensive In Situ Validation of Five Satellite Land Surface Temperature Data Sets over Multiple Stations and Years, Remote Sens., 11, 479, https://doi.org/10.3390/rs11050479, 2019.

Martins, J. P. A., Trigo, I. F., Ghilain, N., Jimenez, C., Göttsche, F.-M., Ermida, S. L., Olesen, F.-S., Gellens-Meulenberghs, F., and Arboleda, A.: An All-Weather Land Surface Temperature Product Based on MSG/SEVIRI Observations, Remote Sens., 11, 3044, https://doi.org/10.3390/rs11243044, 2019.

Masuda, K., Takashima, T., and Takayama, Y.: Emissivity of pure and sea waters for the model sea surface in the infrared window regions, Remote Sens. Environ., 24, 313–329, https://doi.org/10.1016/0034-4257(88)90032-6, 1988.

Maturilli, M.: Basic and other measurements of radiation at station Ny-Ålesund (2006-05 et seq), PANGAEA [data set], https://doi.org/10.1594/PANGAEA.914927, 2020.

Maturilli, M., Hanssen-Bauer, I., Neuber, R., Rex, M., and Edvardsen, K.: The Atmosphere Above Ny-Ålesund: Climate and Global Warming, Ozone and Surface UV Radiation, pp. 23–46, Springer International Publishing, https://doi.org/10.1007/978-3-319-46425-1_2, 2019.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An overview of the global historical climatology network-daily database, J. Atmos. Ocean. Tech., 29, 897–910, https://doi.org/10.1175/JTECH-D-11-00103.1, 2012.

Mildrexler, D. J., Zhao, M., and Running, S. W.: A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests, J. Geophys. Res., 116, G03025, https://doi.org/10.1029/2010JG001486, 2011.

Mildrexler, D. J., Zhao, M., Cohen, W. B., Running, S. W., Song, X. P., and Jones, M. O.: Thermal anomalies detect critical global land surface changes, J Appl. Meteorol. Clim., 57, 391–411, https://doi.org/10.1175/JAMC-D-17-0093.1, 2018.

Miner, K. R., Turetsky, M. R., Malina, E., Bartsch, A., Tamminen, J., McGuire, A. D., Fix, A., Sweeney, C., Elder, C. D., and Miller, C. E.: Permafrost carbon emissions in a changing Arctic, Nat. Rev. Earth Environ., 3, 55–67, https://doi.org/10.1038/s43017-021-00230-3, 2022.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Niclòs, R., Valor, E., Caselles, V., Coll, C., and Sánchez, J. M.: In situ angular measurements of thermal infrared sea surface emissivity – Validation of models, Remote Sens. Environ., 94, 83–93, https://doi.org/10.1016/j.rse.2004.09.002, 2005.

Nielsen-Englyst, P., Høyer, J. L., Madsen, K. S., Tonboe, R. T., Dybkjær, G., and Skarpalezos, S.: Deriving Arctic 2 m air temperatures over snow and ice from satellite surface temperature measurements, The Cryosphere, 15, 3035–3057, https://doi.org/10.5194/tc-15-3035-2021, 2021.

Nitze, I., Grosse, G., Jones, B. M., Romanovsky, V. E., and Boike, J.: Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic, Nature Commun., 9, 5423, https://doi.org/10.1038/s41467-018-07663-3, 2018.

Obu, J., Westermann, S., Kääb, A., and Bartsch, A.: Ground Temperature Map, 2000-2016, Northern Hemisphere Permafrost, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.888600, 2018.

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., Khomutov, A., Kääb, A., Leibman, M. O., Lewkowicz, A. G., Panda, S. K., Romanovsky, V., Way, R. G., Westergaard-Nielsen, A., Wu, T., Yamkhin, J., and Zou, D.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km^2 scale, Earth-Sci. Rev., 193, 299–316, https://doi.org/10.1016/j.earscirev.2019.04.023, 2019.

Olefeldt, D., Hovemyr, M., Kuhn, M. A., Bastviken, D., Bohn, T. J., Connolly, J., Crill, P., Euskirchen, E. S., Finkelstein, S. A., Genet, H., Grosse, G., Harris, L. I., Heffernan, L., Helbig, M., Hugelius, G., Hutchins, R., Juutinen, S., Lara, M. J., Malhotra,

A., Manies, K., McGuire, A. D., Natali, S. M., O'Donnell, J. A., Parmentier, F.-J. W., Räsänen, A., Schädel, C., Sonnentag, O., Strack, M., Tank, S. E., Treat, C., Varner, R. K., Virtanen, T., Warren, R. K., and Watts, J. D.: The Boreal–Arctic Wetland and Lake Dataset (BAWLD), Earth Syst. Sci. Data, 13, 5127–5149, https://doi.org/10.5194/essd-13-5127-2021, 2021.

Palmtag, J., Obu, J., Kuhry, P., Richter, A., Siewert, M. B., Weiss, N., Westermann, S., and Hugelius, G.: A high spatial resolution soil carbon and nitrogen dataset for the northern permafrost region based on circumpolar land cover upscaling, Earth Syst. Sci. Data, 14, 4095–4110, https://doi.org/10.5194/essd-14-4095-2022, 2022.

Pearson, R. K.: Outliers in process modeling and identification, IEEE T. Contr. Syst. T., 10, 55–63, https://doi.org/10.1109/87.974338, 2002.

Peres, L. F. and DaCamara, C. C.: Emissivity maps to retrieve land-surface temperature from MSG/-SEVIRI, IEEE T. Geosci. Remote, 43, 1834–1844, https://doi.org/10.1109/TGRS.2005.851172, 2005.

Prata, A. J.: Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 2. Experimental results and validation of AVHRR algorithms, J. Geophys. Res., 99, 13025–13058, https://doi.org/10.1029/94JD00409, 1994.

Price, J. C.: Land surface temperature measurements from the split window channels of the NOAA 7 Advanced Very High Resolution Radiometer, J. Geophys. Res.-Atmos., 89, 7231–7237, https://doi.org/10.1029/JD089iD05p07231, 1984.

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, Nat. Commun. Earth Environ., 3, 168, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Rantanen, M., Kämäräinen, M., Niittynen, P., Phoenix, G. K., Lenoir, J., Maclean, I., Luoto, M., and Aalto, J.: Bioclimatic atlas of the terrestrial Arctic, Nat. Sci. Data, 10, 40, https://doi.org/10.1038/s41597-023-01959-w, 2023.

Rasmussen, M. O., Gottsche, F.-M., Olesen, F.-S., and Sandholt, I.: Directional Effects on Land Surface Temperature Estimation From Meteosat Second Generation for Savanna Landscapes, IEEE T. Geosci. Remote, 49, 4458–4468, https://doi.org/10.1109/TGRS.2011.2144604, 2011.

Rayner, N. A., Auchmann, R., Bessembinder, J., Brönnimann, S., Brugnara, Y., Capponi, F., Carrea, L., Dodd, E. M., Ghent, D., Good, E., Høyer, J. L., Kennedy, J. J., Kent, E. C., Killick, R. E., van der Linden, P., Lindgren, F., Madsen, K. S., Merchant, C. J., Mitchelson, J. R., Morice, C. P., Nielsen-Englyst, P., Ortiz, P. F., Remedios, J. J., van der Schrier, G., Squintu, A. A., Stephens, A., Thorne, P. W., Tonboe, R. T., Trent, T., Veal, K. L., Waterfall, A. M., Winfield, K., Winn, J., and Woolway, R. I.: The EUSTACE project: Delivering global, daily information on surface air temperature, B. Am. Meteorol. Soc., 101, E1924–E1947, https://doi.org/10.1175/BAMS-D-19-0095.1, 2020.

Reiners, P., Asam, S., Frey, C., Holzwarth, S., Bachmann, M., Sobrino, J., Göttsche, F. M., Bendix, J., and Kuenzer, C.: Validation of avhrr land surface temperature with modis and in situ lst – a timeline thematic processor, Remote Sens., 13, 3473, https://doi.org/10.3390/rs13173473, 2021.

Reiners, P., Sobrino, J., and Kuenzer, C.: Satellite-Derived Land Surface Temperature Dynamics in the Context of Global Change – A Review, Remote Sens., 15, 1857, https://doi.org/10.3390/rs15071857, 2023.

Riffler, M., Lieberherr, G., and Wunderle, S.: Lake surface water temperatures of European Alpine lakes (1989–2013) based on the Advanced Very High Resolution Radiometer (AVHRR) 1 km data set, Earth Syst. Sci. Data, 7, 1–17, https://doi.org/10.5194/essd-7-1-2015, 2015.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., Bormann, N., and Lupu, C.: An update on the RTTOV fast radiative transfer model (currently at version 12), Geosci. Model Dev., 11, 2717–2737, https://doi.org/10.5194/gmd-11-2717-2018, 2018.

Sharifnezhadazizi, Z., Norouzi, H., Prakash, S., Beale, C., and Khanbilvardi, R.: A Global Analysis of Land Surface Temperature Diurnal Cycle Using MODIS Observations, J. Appl. Meteor. Climatol., 58, 1279–1291, https://doi.org/10.1175/JAMC-D-18-0256.1, 2019.

Sobrino, J., Raissouni, N., and Li, Z.-L.: A Comparative Study of Land Surface Emissivity Retrieval from NOAA Data, Remote Sens. Environ., 75, 256–266, https://doi.org/10.1016/S0034-4257(00)00171-1, 2001.

Sobrino, J. A. and Raissouni, N.: Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco, Int. J. Remote Sens., 21, 353–366, https://doi.org/10.1080/014311600210876, 2000.

Sobrino, J. A., Jiménez-Muñoz, J. C., Labed-Nachbrand, J., and Nerry, F.: Surface emissivity retrieval from Digital Airborne Imaging Spectrometer data, J. Geophys. Res., 107, ACL 24–1–ACL 24–13, https://doi.org/10.1029/2002JD002197, 2002.

Sobrino, J. A., Jiménez-Muñoz, J. C., Sòria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., and Martínez, P.: Land surface emissivity retrieval from different VNIR and TIR sensors, IEEE T. Geosci. Remote, 46, 316–327, https://doi.org/10.1109/TGRS.2007.904834, 2008.

Sulla-Menashe, D. and Friedl, M. A.: User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product, Tech. rep., USGS, https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf (last access: 17 December 2024), 2018.

Thiebault, K. and Young, S.: Snow cover change and its relationship with land surface temperature and vegetation in northeastern North America from 2000 to 2017, Int. J. Remote Sens., 41, 8453–8474, https://doi.org/10.1080/01431161.2020.1779379, 2020.

Treat, C. C., Virkkala, A., Burke, E., Bruhwiler, L., Chatterjee, A., Fisher, J. B., Hashemi, J., Parmentier, F. W., Rogers, B. M., Westermann, S., Watts, J. D., Blanc-Betes, E., Fuchs, M., Kruse, S., Malhotra, A., Miner, K., Strauss, J., Armstrong, A., Epstein, H. E., Gay, B., Goeckede, M., Kalhori, A., Kou, D., Miller, C. E., Natali, S. M., Oh, Y., Shakil, S., Sonnentag, O., Varner, R. K., Zolkos, S., Schuur, E. A., and Hugelius, G.: Permafrost Carbon: Progress on Understanding Stocks and Fluxes Across Northern Terrestrial Ecosystems, J. Geophys. Res.-Biogeo., 129, e2023JG007638, https://doi.org/10.1029/2023JG007638, 2024.

Trigo, I., Monteiro, I. T., Olesen, F., and Kabsch, E.: An assessment of remotely sensed land surface temperature, J. Geophys. Res., 113, D17108, https://doi.org/10.1029/2008JD010035, 2008a.

Trigo, I., Peres, L. F., DaCamara, C. C., and Freitas, S. C.: Thermal land surface emissivity retrieved from SEVIRI/Meteosat, IEEE T. Geosci. Remote, 46, 307–315, https://doi.org/10.1109/TGRS.2007.905197, 2008b.

Trigo, I., Dacamara, C. C., Viterbo, P., Roujean, J. L., Olesen, F., Barroso, C., Camacho-De-Coca, F., Carrer, D., Freitas, S. C., García-Haroj, J., Geiger, B., Gellens-Meulenberghs, F., Ghilain, N., Meliá, J., Pessanha, L., Siljamo, N., and Arboleda, A.: The satellite application facility for land surface analysis, Int. J. Remote Sens., 32, 2725–2744, https://doi.org/10.1080/01431161003743199, 2011.

Trigo, I., Freitas, S., Bioucas-Dias, J., Barroso, C., Monteiro, I., Viterbo, P., and Martins, J. P.: Algorithm Theoretical Basis Document for Land Surface Temperature (LST), PRODUCTS: LSA-002 (ELST), Tech. rep., EUMETSAT LSA SAF, https://nextcloud.lsasvcs.ipma.pt/s/NsW275gpDAfekzc?dir=undefined&path=%2FATBD-Algorithm_Theoretial_Basis_Document&openfile=27090 (last access: 17 December 2024), 2017.

Urban, M., Eberle, J., Hüttich, C., Schmullius, C., and Herold, M.: Comparison of satellite-derived land surface temperature and air temperature from meteorological stations on the pan-arctic scale, Remote Sens., 5, 2348–2367, https://doi.org/10.3390/rs5052348, 2013.

Urraca, R. and Gobron, N.: Temporal stability of long-term satellite and reanalysis products to monitor snow cover trends, The Cryosphere, 17, 1023–1052, https://doi.org/10.5194/tc-17-1023-2023, 2023.

Vandecrux, B., Fausto, R. S., Box, J. E., Covi, F., Hock, R., Rennermalm, Å. K., Heilig, A., Abermann, J., van As, D., Bjerre, E., Fettweis, X., Smeets, P. C. J. P., Kuipers Munneke, P., van den Broeke, M. R., Brils, M., Langen, P. L., Mottram, R., and Ahlstrøm, A. P.: Recent warming trends of the Greenland ice sheet documented by historical firn and ice temperature observations and machine learning, The Cryosphere, 18, 609–631, https://doi.org/10.5194/tc-18-609-2024, 2024.

Walton, C. C., Sullivan, J. T., Rao, C. R. N., and Weinreb, M. P.: Corrections for detector nonlinearities and calibration inconsistencies of the infrared channels of the advanced very high resolution radiometer, J. Geophys. Res., 103, 3323–3337, https://doi.org/10.1029/97JC02018, 1998.

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, IEEE T. Geosci. Remote, 34, 892–905, https://doi.org/10.1109/36.508406, 1996.

Wang, A. and Zeng, X.: Range of monthly mean hourly land surface air temperature diurnal cycle over high northern latitudes, J. Geophys. Res., 119, 5836–5844, https://doi.org/10.1002/2014JD021602, 2014.

Wang, L., Arora, V. K., Bartlett, P., Chan, E., and Curasi, S. R.: Mapping of ESA's Climate Change Initiative land cover data to plant functional types for use in the CLASSIC land model, Biogeosciences, 20, 2265–2282, https://doi.org/10.5194/bg-20-2265-2023, 2023.

Waring, A. M., Ghent, D., Perry, M., Anand, J. S., Veal, K. L., and Remedios, J.: Regional climate trend analyses for Aqua MODIS land surface temperatures, Int. J. Remote Sens., 44, 4989–5032, https://doi.org/10.1080/01431161.2023.2240522, 2023.

Westermann, S., Lüers, J., Langer, M., Piel, K., and Boike, J.: The annual surface energy budget of a high-arctic permafrost site on Svalbard, Norway, The Cryosphere, 3, 245–263, https://doi.org/10.5194/tc-3-245-2009, 2009.

Westermann, S., Langer, M., and Boike, J.: Systematic bias of average winter-time land surface temperatures inferred from MODIS at a site on Svalbard, Norway, Remote Sens. Environ., 118, 162–167, https://doi.org/10.1016/j.rse.2011.10.025, 2012.

Westermann, S., Peter, M., Langer, M., Schwamborn, G., Schirrmeister, L., Etzelmüller, B., and Boike, J.: Transient modeling of the ground thermal conditions using satellite data in the Lena River delta, Siberia, The Cryosphere, 11, 1441–1463, https://doi.org/10.5194/tc-11-1441-2017, 2017.

World Meteorological Organization (WMO) : Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, Tech. Rep. GOOS-184, GTOS-76, WMO-TD/No. 1523, WMO, Geneva, Switzerland, 2010.

Xiao, X., Naegeli, K., Neuhaus, C., Salberg, A.-B., Schwaizer, G., Wiesmann, A., Wunderle, S., and Nagler, T.: ESA Snow Climate Change Initiative (Snow_cci): Daily global Snow Cover Fraction - viewable (SCFV) from AVHRR (1979–2022), version 3.0, NERC EDS Centre for Environmental Data Analysis [data set], https://doi.org/10.5285/7491427f8c3442ce825ba5472c224322, 2024.

Yang, J., Zhou, J., Göttsche, F. M., Long, Z., Ma, J., and Luo, R.: Investigation and validation of algorithms for estimating land surface temperature from Sentinel-3 SLSTR data, Int. J. Appl. Earth Obs., 91, 102136, https://doi.org/10.1016/j.jag.2020.102136, 2020.

Zhang, X., Zhou, J., Liang, S., and Wang, D.: A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature, Remote Sens. Environ., 260, 112437, https://doi.org/10.1016/j.rse.2021.112437, 2021.

Zhao, P. and He, Z.: A First Evaluation of ERA5-Land Reanalysis Temperature Product Over the Chinese Qilian Mountains, Front. Earth Sci., 10, 907730, https://doi.org/10.3389/feart.2022.90773, 2022.

Zhou, J., Liang, S., Cheng, J., Wang, Y., and Ma, J.: The GLASS Land Surface Temperature Product, IEEE J. Sel. Top. Appl. Earth Obs., 12, 493–507, https://doi.org/10.1109/JSTARS.2018.2870130, 2019.