https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Evaluation of reanalysis- and satellite-based estimates of sensible heat flux over Arctic sea ice against airborne eddy-covariance measurements

Joshua J. Müller[1], Ilga Staudinger[2], Florian Gebhardt[3], David J. Simon[4], Laura Köhler[5], Jörg Hartmann[5], André Ehrlich[1], Michael Schäfer[1], Matthew D. Shupe[6,7], and Manfred Wendisch[1]

[1]Leipzig Institute for Meteorology (LIM), Leipzig University, Leipzig, Germany
[2]Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany
[3]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[4]Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany
[5]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
[6]Physical Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA
[7]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

Correspondence: Joshua J. Müller (joshua.mueller@uni-leipzig.de)

Abstract. Turbulent sensible heat flux (SHF) influences the Arctic surface energy budget, but its magnitude and spatial variability over sea ice remain poorly constrained. Direct measurements are scarce, particularly in the shallow atmospheric layers above sea-ice surfaces. We use eddy-covariance (EC) measurements from a sensor package towed below a research aircraft and flown at a mean altitude of 23 m to evaluate bulk SHF estimates driven by in situ, satellite, and reanalysis data, as well as the SHF diagnosed by the global reanalysis ERA5 and the Copernicus Arctic Regional Reanalysis (CARRA). The measurements were collected near Svalbard in April 2024. Bulk estimates driven solely by reanalysis inputs capture only about one third of the amplitude of the observed SHF and lack to represent the spatial variability when confronted with the EC measurements. Replacing the reanalysis surface skin temperature with satellite retrievals improves magnitude and variability of the SHF estimate. The improvement is associated with flight segments where heterogeneous surface temperatures produce variations in SHF that the reanalysis-only estimates fail to reproduce. With collocated in situ inputs, the bulk method reproduces the observed variability best, but overestimates the amplitude of the SHF by about 23 % relative to the EC SHF. A systematic sensitivity analysis shows that, with in situ inputs, the choice of atmospheric stability scheme has little influence under the conditions sampled. These results identify the surface skin temperature, which determines in this setup the variability of the near-surface temperature difference, and the turbulent transfer coefficient as the main controls on the performance of the bulk estimates under the Arctic spring conditions sampled here. Satellite surface skin temperatures improve the representation of spatial SHF variability, while better-constrained transfer coefficients are needed to reproduce its magnitude.

1 Introduction

The Arctic region has been warming nearly four times faster than the globe on average since 1979 (Rantanen et al., 2022; Wendisch et al., 2023). Sea-ice cover and volume have declined, and the remaining sea ice is becoming thinner, younger, and

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

more mobile (Kwok et al., 2009; Maslanik et al., 2011; Spreen et al., 2011). These changes expose a larger fraction of thin sea ice and open ocean to the atmosphere, as simulated for the Beaufort Sea (Rheinlænder et al., 2024), and thereby affect the vertical turbulent heat fluxes over the Arctic Ocean (Taylor et al., 2018). Among these fluxes, the sensible heat flux (SHF) is a key component of the near-surface energy budget of the Arctic climate system, linking atmosphere, sea ice, and ocean (Vihma et al., 2014). It contributes to the thermal structure and stability of the lower atmosphere and influences sea-ice growth and melt (Maykut, 1978). Accurate knowledge of the SHF over sea ice matters for projecting Arctic climate change (Boisvert et al., 2022).

The SHF is defined as the turbulent vertical transport of sensible heat between the surface and the atmosphere, expressed as the covariance $\overline{w' \cdot \theta'}$ of vertical wind speed (w) and potential temperature (θ) fluctuations, multiplied by the density of dry air ρ and the specific heat capacity of dry air c_p,

$$H_\mathrm{S} = \rho \cdot c_\mathrm{p} \cdot \overline{w' \cdot \theta'}. \quad (1)$$

The covariance is typically obtained using the eddy-covariance (EC) method, which is based on high-frequency wind and temperature measurements. Over sea ice, the required measurements are commonly obtained using instruments mounted on towers installed on sea-ice floes (Persson et al., 2002; Grachev et al., 2007; Cox et al., 2023). This approach, however, samples only the immediate surroundings of the instrument and does not capture the variety of surface and atmospheric conditions within a satellite pixel or model grid cell. Airborne EC campaigns offer the spatial coverage that towers lack (Brümmer et al., 1994; Paluch et al., 1997; Schröder et al., 2003; Elvidge et al., 2021), but the atmospheric boundary layer (ABL) over Arctic sea ice is often too shallow to be sampled from an aircraft. During the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) drift, about half of the surface mixing layers observed in winter and spring were less than 150 m deep (Akansu et al., 2023). Stable and very shallow mixed ABLs were frequent, in spring as well as over the whole year (Jozef et al., 2024), and during the Surface Heat Budget of the Arctic Ocean (SHEBA) experiment, the layer of continuous turbulence was at times less than 5 m deep in the very stable regime (Grachev et al., 2008). The surface layer, within which the turbulent fluxes can be treated as height-constant (Stull, 1988), is confined even closer to the sea-ice surface, whereas aircraft typically operate several tens of metres above the surface (Renfrew et al., 2021).

Continuous Arctic-wide estimates of SHF are available from atmospheric reanalyses, most prominently the global reanalysis ERA5 (Hersbach et al., 2020) and the regional, higher-resolution Copernicus Arctic Regional Reanalysis (CARRA; Køltzow et al., 2022; Batrak et al., 2024). Their SHF products are used in assessments of the Arctic surface energy budget and its trends (Taylor et al., 2018) and in process studies of air-mass transformations such as marine cold-air outbreaks (Slättberg et al., 2025), and their meteorological fields serve as atmospheric input for satellite-based flux retrievals (Qu et al., 2019; Zhang et al., 2023, 2025). Evaluations over sea ice, however, reveal substantial uncertainty. The surface skin temperature (T_s) over sea ice in atmospheric reanalyses such as ERA5 contains a warm bias. For ERA5 in winter it is about 2 K on average and 4 K under clear skies (Herrmannsdörfer et al., 2023), and in cold clear-sky events it can reach 5–10 K (Batrak and Müller, 2019). This warm bias is reduced by CARRA (Batrak et al., 2024). A bulk SHF estimate driven by ERA5 deviates from observations at the drifting station of the Norwegian young sea ICE (N-ICE2015) expedition north of Svalbard (January to June 2015) by

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

a root-mean-square error (RMSE) of 18 W m^{-2} (Zhang et al., 2023). Differences in SHF values of this size are common to all major reanalyses (Graham et al., 2019), and the SHF is highly sensitive to the prescribed sea-ice concentration (Uhlíková et al., 2024), whose distribution can be overly smooth (Renfrew et al., 2021).

Where direct turbulence measurements are unavailable, the SHF is instead parameterized with a bulk approach

$$H_{\mathrm{S}} = \rho \cdot c_{\mathrm{p}} \cdot C_{\mathrm{h}} \cdot U \cdot \left[\theta_{\mathrm{s}} - \theta_{\mathrm{air}}\right]. \tag{2}$$

In this parameterization, U is the horizontal wind speed, and θ_{s} and θ_{air} denote the potential temperature at the surface and . The surface value of potential temperature θ_{s} is obtained from the surface skin temperature T_{s} using the corresponding surface pressure. Within the constant-flux surface layer, H_{S} approximates the surface flux, and the height dependence enters only through the measurement height of the atmospheric input. The dimensionless turbulent transfer coefficient of heat C_{h} depends on the scalar roughness length $z_{0,T}$ and the stability of the near-surface atmosphere. It is typically derived from Monin–Obukhov similarity theory (MOST; Monin and Obukhov, 1954). MOST, however, assumes horizontally homogeneous surface conditions and therefore carries additional uncertainty over heterogeneous sea-ice surfaces (Schröder et al., 2003; Pinto et al., 2003). Leads, narrow openings in the sea ice that refreeze quickly, are one prominent example of that heterogeneity. Their fraction rises rapidly in April (Chen et al., 2025), when the calculated surface SHF over open water is still about 500 W m^{-2} (Maykut, 1978). Although they cover only 1–2 % of the central Arctic surface in winter (Marcq and Weiss, 2012) and about 8 % in the peripheral seas (Lindsay and Rothrock, 1995), they account for more than 70 % of the upward heat flux in the central Arctic in winter (Marcq and Weiss, 2012). This high impact arises from buoyancy-driven convection over the warm open water of leads, which enhances the turbulent heat exchange that stable stratification over the surrounding sea ice suppresses (Lüpkes et al., 2008; Chechin et al., 2019). Observational SHF data over leads nevertheless remain sparse (Tetzlaff et al., 2015).

At the pan-Arctic scale, the bulk method (Eq. (2)) is driven in one of two ways. In the reanalysis-only approach, all input values (θ_{air}, T_{s}, U, and the air pressure p) are taken from a reanalysis, which provides coverage also below clouds, at the cost of horizontal resolution (Zhang et al., 2023, 2025). In the combined satellite–reanalysis approach, the atmospheric input values still come from the reanalysis, but T_{s} is retrieved from thermal-infrared satellite radiances (Qu et al., 2019; Yin et al., 2021), which resolves surface heterogeneity at a higher horizontal resolution than the reanalysis but is limited to cloud-free scenes. Reanalyses additionally diagnose the SHF within their own surface energy balance closure, which makes the diagnosed SHF difficult to compare with bulk estimates recomputed afterwards from the model output fields, referred to here as offline estimates. Both the reanalysis-only and the combined approach are limited by their input data. In the reanalysis-only approach, T_{s} is averaged over grid cells of 2.5–30 km (Hersbach et al., 2020; Køltzow et al., 2022). In the combined approach, operational surface temperature products, such as that of the MODerate-resolution Imaging Spectroradiometer (MODIS; Hall and Riggs, 2021), are restricted to 1–4 km horizontal resolution, so that the MODIS surface skin temperature $T_{\mathrm{s,MOD}}$ averages the thermal signal of different surface types into a single mixed-pixel value.

Beyond the input data, the parameterized SHF depends on C_{h} derived from MOST. Because C_{h} depends non-linearly on vertical atmospheric stability, biases in input data propagate into the SHF values derived from either approach. C_{h} is computed either iteratively from the MOST stability functions or with a Louis-type scheme that diagnoses the stability dependence

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

from the mean wind and temperature differences between surface and measurement height (Louis, 1979). In both schemes, C_h is related to the drag coefficient C_d through the roughness lengths for momentum ($z_{0,m}$) and heat ($z_{0,T}$), combined in the scalar roughness ratio $\alpha_T = z_{0,T}/z_{0,m}$. Because form drag of sea-ice floes enhances the transfer of momentum but not that of heat, $z_{0,T}$ is expected to fall below $z_{0,m}$ over rough sea ice. Surface-renewal theory predicts a decrease of α_T with increasing roughness Reynolds number (Andreas, 1987; Andreas et al., 2010). The adopted values of α_T span several orders of magnitude. Schnierstein et al. (2024) use $\alpha_T = 0.1$ in their large-eddy simulations of the Arctic ABL during the MOSAiC drift. The Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) fixes $z_{0,T}$ at 10^{-3} m for sea ice independently of surface conditions, giving $\alpha_T = 1$ over compact sea ice (ECMWF, 2016). Airborne estimates over sea ice vary just as widely (Schröder et al., 2003; Elvidge et al., 2021). This spread translates into poorly constrained systematic uncertainty in satellite- and reanalysis-based SHF estimates over sea ice.

To address unresolved surface heterogeneity, the representation of the stable Arctic ABL in reanalyses, and the poorly constrained α_T, we use measurements carried out by the recently developed Turbulence Bird (T-Bird). The T-Bird, a sensor platform towed beneath the Polar 6 research aircraft, enables EC measurements at low altitudes (15–34 m) and thereby overcomes the altitude constraint of conventional airborne platforms (Jurányi et al., 2025; Simon et al., 2025). In this study, we derive the SHF with the EC method from T-Bird data collected over Arctic sea ice and use it as a reference to evaluate bulk aerodynamic estimates driven by in situ, satellite, and reanalysis data. The evaluation separates the contribution of the meteorological inputs from that of the flux parameterization itself and thereby identifies where satellite- and reanalysis-based SHF over sea ice can be improved.

Section 2 introduces the instrumentation and datasets, Sect. 3 describes the EC and bulk methods and the evaluation metrics, Sect. 4 compares the bulk and EC SHF estimates and presents the sensitivity analysis, Sect. 5 discusses the results, and Sect. 6 concludes. All processing steps are reproducible with the analysis code and documented on a website that is archived with the code (see the Code and data availability section).

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

2 Measurements, instrumentation and reanalyses

2.1 Airborne measurements

We use observations collected during the Boundary Layer and Aerosol and Cloud Study in the Arctic based on Aircraft and T-Bird Measurements II (BACSAM-II) airborne campaign, which was carried out from Longyearbyen (Svalbard) between 5 April and 2 May 2024. Eleven research flights were conducted, accumulating approximately 55 flight hours (Jurányi et al., 2024). Figure 1 shows the campaign region and the flight tracks.

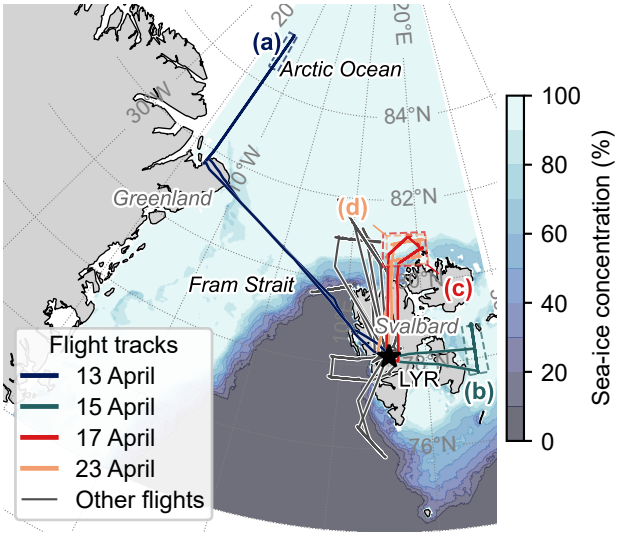


Figure 1. Overview of the BACSAM-II campaign region. Grey lines show the flight tracks of all Polar 6 research flights. The coloured lines highlight the flight paths of the four low-level flights of which data are used in this study. The coloured rectangles (dashed lines) outline the corresponding focus regions shown in Fig. 4a–d. The background shows the mean sea-ice concentration for April 2024 from the merged MODIS–AMSR2 product (Ludwig et al., 2020). The black star marks the operating base Longyearbyen.

The airborne observations combine measurements from the research aircraft Polar 6 (Basler BT-67, operated by the Alfred Wegener Institute; Wesche et al., 2016) with those from the T-Bird, a sensor platform that is towed below the aircraft on a retractable cable (Jurányi et al., 2025). The T-Bird carries a five-hole probe at the nose, which provides the horizontal wind speed U, the vertical wind w, and the wind direction. Two pressure transducers sample the dynamic and static pressure, and a Pt100 resistance thermometer measures the air temperature. All of these quantities are sampled at 100 Hz temporal resolution. An inertial navigation unit (1 Hz GPS, 500 Hz inertial) provides the T-Bird position and altitude. The observational surface skin temperature $T_{\mathrm{s,obs}}$ is obtained from a KT-19 infrared pyrometer (Heitronics KT-19.85 II) mounted on the Polar 6 aircraft, which records nadir-viewing radiances transformed into corresponding thermal-infrared brightness temperatures (9.6–11.5 µm, field

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

of view 2.3°) at 50 Hz. At a flight altitude of 70 m and a true airspeed of 62 m s^{-1}, this corresponds to an instantaneous footprint of approximately 3 m in diameter and an along-track sampling distance of about 1 m.

From the 11 research flights we use the four flight days (13, 15, 17, and 23 April 2024) on which Polar 6 flew occasionally at altitudes below 80 m over sea ice, which allowed the T-Bird to be winched to altitudes between 15 m and 34 m (mean 23 m). These four low-level days comprise six constant-altitude flight legs (one each on 13 and 15 April and two each on 17 and 23 April, median length 70.2 km), which we cut into 5 km segments for the SHF analysis. We exclude 25 April, when all segments below 50 m were flown over open water. The second leg on 23 April crossed the largely open Whalers Bay polynya and is retained because sea ice was still present (mean sea-ice concentration 28 %). This yields 91 segments, corresponding to 455 km of along-track data. The 100 Hz T-Bird data are used to derive the SHF with the EC method ($H_{\mathrm{S,obs,EC}}$), while data averaged over 5 km along track provide the mean air temperature, pressure, and wind speed for the bulk method ($H_{\mathrm{S,obs,bulk}}$).

Table 1 summarizes the six SHF estimates compared in this study. The EC flux $H_{\mathrm{S,obs,EC}}$ serves as the reference (Sect. 3.1). Three bulk estimates apply Eq. (2) with the same bulk algorithm, including the same segment roughness and reference configuration (Sect. 3.2), and differ only in their input data. The T-Bird bulk estimate $H_{\mathrm{S,obs,bulk}}$ uses in situ inputs only, the ERA5–MODIS bulk estimate $H_{\mathrm{S,RS,bulk}}$ combines the satellite $T_{\mathrm{s,MOD}}$ with the ERA5 atmosphere, and the ERA5-only bulk estimate $H_{\mathrm{S,ERA,bulk}}$ uses ERA5 fields alone. The fluxes diagnosed internally by the two reanalyses, $H_{\mathrm{S,ERA,diag}}$ and $H_{\mathrm{S,CAR,diag}}$, serve as model benchmarks (Sect. 2.2).

Table 1. SHF estimates evaluated in this study. All estimates are collocated with the 5 km T-Bird segments.

Name	Symbol	Input state	Method / role
T-Bird eddy covariance (EC)	$H_{\mathrm{S,obs,EC}}$	T-Bird 100 Hz w, θ	Eddy covariance (Eq. (1))
T-Bird bulk	$H_{\mathrm{S,obs,bulk}}$	KT-19 $T_{\mathrm{s,obs}}$; T-Bird θ, U, z	Bulk estimate (Eq. (2))
ERA5–MODIS bulk	$H_{\mathrm{S,RS,bulk}}$	$T_{\mathrm{s,MOD}}$; ERA5 θ, U, p at model level 136 (about 30 m), closest to the mean T-Bird height	Bulk estimate (Eq. (2))
ERA5-only bulk	$H_{\mathrm{S,ERA,bulk}}$	$T_{\mathrm{s,ERA}}$; ERA5 θ, U, p at model level 136 (about 30 m), closest to the mean T-Bird height	Bulk estimate (Eq. (2))
ERA5 model-diagnosed	$H_{\mathrm{S,ERA,diag}}$	ERA5 surface scheme	Model-diagnosed benchmark
CARRA model-diagnosed	$H_{\mathrm{S,CAR,diag}}$	CARRA surface scheme	Model-diagnosed benchmark

2.2 Reanalysis data and model-diagnosed fluxes

The bulk SHF estimates that do not use in situ atmospheric data (Table 1) take the near-surface input data required for Eq. (2) from the ERA5 reanalysis (Hersbach et al., 2020). We use hourly potential temperature θ, horizontal wind speed U, and pressure p at a horizontal resolution of 0.25°, taken from the ERA5 reanalysis at 30 m above the surface, which is closest to the average T-Bird flight altitude, using the grid cell and hourly time step nearest to each segment. The ERA5 surface skin temperature $T_{\mathrm{s,ERA}}$, the grid-box mean over the open-water and sea-ice tiles of the same 0.25° grid, completes the input of the ERA5-only bulk estimate $H_{\mathrm{S,ERA,bulk}}$, whereas the ERA5–MODIS bulk estimate $H_{\mathrm{S,RS,bulk}}$ replaces it with the satellite

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

$T_{\mathrm{s,MOD}}$ introduced in Sect. 2.3. Both estimates use the same bulk algorithm, roughness, and reference configuration as the T-Bird bulk estimate (Sect. 3.2). Additionally, we evaluate the SHF diagnosed by ERA5 itself, $H_{\mathrm{S,ERA,diag}}$, which results from the ERA5 surface scheme and not from our bulk algorithm. In this scheme, every ocean grid box is split into an open-water and a sea-ice tile. For each tile, SHF is computed from a bulk law analogous to Eq. (2) between the surface and the lowest model level at approximately 10 m, and the grid-box SHF value results from the area-weighted mean of the two (ECMWF, 2016). The scheme obtains C_{h} iteratively from MOST and fixes $z_{0,T}$ over sea ice at 10^{-3} m, which implies $\alpha_T = 1$ over compact sea ice (Sect. 1). The T_{s} of each tile is not prescribed but computed by the model from the energy balance of the tile surface, which includes the radiative, conductive, and turbulent heat fluxes. The diagnosed flux and $T_{\mathrm{s,ERA}}$ are therefore mutually consistent within the model, whereas our offline estimate takes $T_{\mathrm{s,ERA}}$ as given. Consequently, $H_{\mathrm{S,ERA,diag}}$ cannot be reproduced offline by inserting ERA5 output into Eq. (2). It differs from $H_{\mathrm{S,ERA,bulk}}$ in the tile averaging, the input level, the transfer scheme, and the roughness assumptions.

CARRA (Batrak et al., 2024) is based on the HARMONIE-AROME system (Bengtsson et al., 2017) with the SURFEX surface scheme (Masson et al., 2013). Over sea ice it applies the same tile-wise bulk law as ERA5, but obtains the transfer coefficient directly from the bulk Richardson number (Ri_{b}) with a Louis-type scheme, using the scalar roughness ratio $\alpha_T = 0.1$ and no form drag (Batrak et al., 2024; Slättberg et al., 2025). CARRA complements ERA5 at a horizontal resolution of 2.5 km instead of about 30 km. From CARRA we show $T_{\mathrm{s,CAR}}$, height-level data at 30 m and 75 m above the surface, and the model-diagnosed SHF $H_{\mathrm{S,CAR,diag}}$. The two reanalyses represent two types of C_{h} parameterization. CARRA uses a non-iterative, Richardson-number-based scheme of the Louis type, whereas ERA5 uses an iterative scheme based on the Monin–Obukhov similarity functions. Both diagnosed fluxes serve as model benchmarks, for which the reanalyses provide no uncertainty estimates.

2.3 Satellite-derived surface skin temperature and sea-ice concentration

The ERA5–MODIS bulk estimate $H_{\mathrm{S,RS,bulk}}$ requires values of T_{s} that are independent of the in situ measurements. This satellite-derived $T_{\mathrm{s,MOD}}$ is taken from the MODIS/Aqua ice-surface temperature (IST) daily level 3 product at 4 km resolution (version 61; Hall and Riggs, 2021). The underlying retrieval applies a split-window algorithm to the thermal-infrared channels 31 and 32 (10.78–12.27 µm). For each flight day, the $T_{\mathrm{s,MOD}}$ grid of that day is collocated with the 5 km segments in the same way as the reanalysis fields. Grid cells that are cloud-masked or otherwise invalid are filled with the collocated $T_{\mathrm{s,ERA}}$, which applies to 24 of the 91 segments, 23 of them on 23 April. The daily sea-ice concentration is taken from the merged sea-ice concentration product of MODIS and the Advanced Microwave Scanning Radiometer 2 (AMSR2) at 1 km resolution (Ludwig et al., 2020). Because no satellite-derived surface skin temperature is assimilated into ERA5 over sea ice, $T_{\mathrm{s,MOD}}$ and $T_{\mathrm{s,ERA}}$ are independent (Nielsen-Englyst et al., 2023; Tian et al., 2024).

Figure 2 summarizes the thermodynamic inputs used for the SHF estimates. Throughout, the near-surface potential temperature difference is defined as

$$\Delta\theta = \theta_{\mathrm{air}} - \theta_{\mathrm{s}}, \tag{3}$$

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

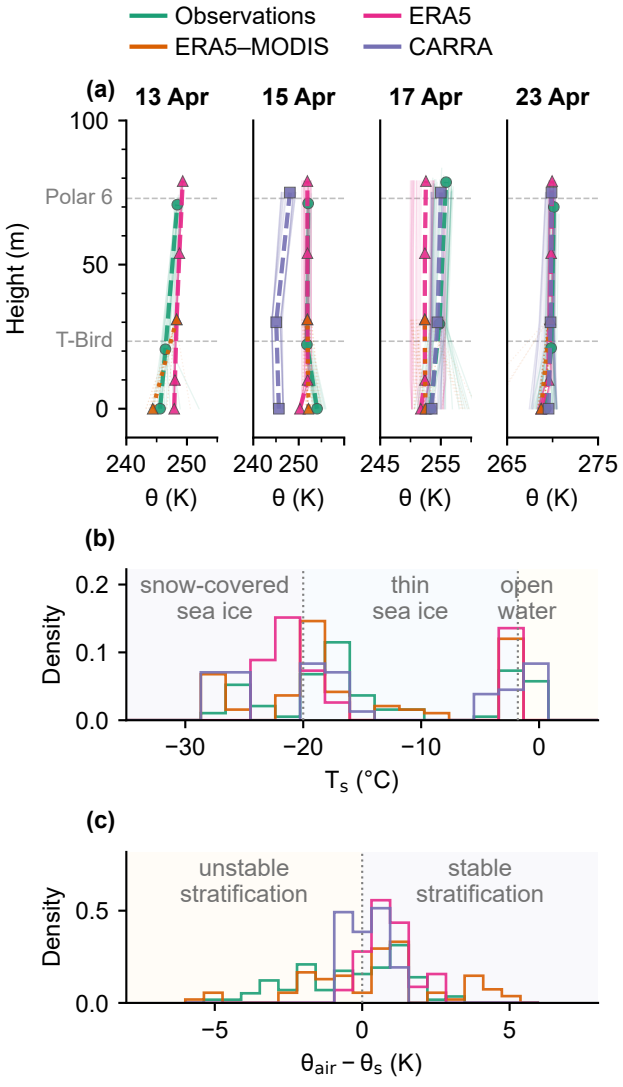


Figure 2. Thermodynamic lower-boundary conditions for the bulk flux estimates. **(a)** Profiles of potential temperature on the four BACSAM-II low-level flight days, constructed from θ_s at the surface and the potential temperature at the T-Bird and Polar 6 flight heights (observations) or at the model and height levels of the reanalyses. Thin lines show individual segments and thick lines the median profile. Dashed horizontal lines indicate the mean flight heights above the surface. Colours distinguish the observations, ERA5–MODIS, ERA5, and CARRA. **(b)** Normalized distributions of T_s, with surface-regime ranges indicated by the shaded background. **(c)** Normalized distributions of the near-surface potential temperature difference.

180 so that $\Delta\theta < 0$ denotes unstable and $\Delta\theta > 0$ stable stratification. The median vertical profiles in Fig. 2a connect θ_s at $z = 0$ with θ_{air} measured at the T-Bird level, on average 23 m above the surface, and at the Polar 6 level, about 70 m above the surface. They indicate predominantly weakly unstable to weakly stable atmospheric stratification. Replacing $T_{s,ERA}$ with $T_{s,MOD}$

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

modifies only the lowest part of the profile, since the atmospheric levels remain unchanged. This improves the near-surface temperature difference on 15 April, makes the median profile too stable on 13 April and too unstable on 17 April, a day on which the observed profiles also differ widely between segments, and leaves the results for 23 April almost unchanged, because cloud cover leaves a valid MODIS pixel on only 3 of the 26 segments. The KT-19 T_s distribution (Fig. 2b) is multimodal, resolving cold snow-covered sea ice, intermediate thin or refrozen ice, and warm sea-ice-free surfaces, whereas the reanalysis distributions also show several modes, but fewer and narrower ones. The observed $\Delta\theta$ frequency distribution is correspondingly broad and spans both weakly unstable and weakly stable regimes (Fig. 2c), in contrast to the narrower $\Delta\theta$ frequency distributions produced by both reanalyses.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

3 Methods

3.1 Eddy-covariance (EC) method

We derive the observational reference SHF $H_{\mathrm{S,obs,EC}}$ from the turbulent fluctuations of w and θ measured by the T-Bird, using the EC method. First, the platform motion is removed from the measured wind vector based on the inertial navigation data (Jurányi et al., 2025). For each 5 km flight segment, corresponding to approximately 81 s of sampling at the mean true airspeed of $62\,\mathrm{m\,s^{-1}}$, we remove a linear trend from the 100 Hz w and θ time series and denote the turbulent fluctuations by w' and θ'. The kinematic heat flux $\overline{w'\theta'}$ is then obtained by integrating the w–θ cospectrum between 0.1 Hz and 20 Hz, and converted to $H_{\mathrm{S,obs,EC}}$ through Eq. (1). The lower limit follows Tetzlaff et al. (2015), who introduced it to remove variability caused by the geometry of a lead. At this true airspeed, the frequency band corresponds to along-track scales of about 3–620 m. The upper limit excludes instrument noise. The lower limit also excludes larger scales, which may carry flux (Mahrt, 1998), as quantified in Appendix A. We use the potential rather than the virtual potential temperature because humidity corrections are negligible under the mostly dry, cold conditions sampled during BACSAM-II. As in Eq. (2), a positive flux denotes upward (surface-to-atmosphere) turbulent transport.

Instead of a quality class, every flux carries its own random error. An integrity check of the resolved turbulence, which requires an inertial-subrange slope of the vertical-wind spectrum within 0.4 of $-5/3$ (Kolmogorov, 1941; Obukhov, 1941), is satisfied by 41 of the 91 segments. It is reported but not used to exclude segments (Appendix A). The random error σ_{rand} of each flux follows Finkelstein and Sims (2001) (Eq. (A1)). Its median is $0.51\,\mathrm{W\,m^{-2}}$, and 75, 58, and 49 of the 91 fluxes exceed 1, 2, and $3\,\sigma_{\mathrm{rand}}$. Fluxes that do not stand out from their random error are valid measurements of a small flux and are retained and reported together with their random error, as recommended by Langford et al. (2015), because eliminating small fluxes can bias the comparison. Appendix A documents these checks and shows that conventional screening changes the results little.

The choice of 5 km is justified by a window-length test on the six constant-altitude legs (Sect. 2.1). We recompute the EC flux and its random error for sub-windows between 1 and 30 km and show the mean flux of the sub-windows as a fraction of the SHF of the stretch they cover in Fig. 3a. This is an internal consistency test rather than an independent accuracy check, because both derive from the same data. For all window lengths between 1 and 30 km this ratio stays close to unity, and at 5 km it is 0.99 ± 0.01 (one standard deviation). The random error of the sub-window fluxes falls from a median of $1.00\,\mathrm{W\,m^{-2}}$ at 1 km to $0.50\,\mathrm{W\,m^{-2}}$ at 5 km (Fig. 3b), and the fraction of sub-windows whose flux exceeds $2\,\sigma_{\mathrm{rand}}$ rises from 45 % to 71 % (Fig. 3c). For longer windows, the random error keeps falling, but this fraction levels off at about 70 %. Longer windows also give fewer independent samples and coarser spatial resolution, so we adopt 5 km as the analysis scale.

We decompose the total uncertainty of $H_{\mathrm{S,obs,EC}}$, σ_{EC}, into two independent contributions that are added in quadrature, the random error σ_{rand} and a systematic instrument error σ_{instr}. The latter is dominated by the five-hole-probe flow calibration, which propagates to about 5 % of the flux. Across the 91 segments the median total uncertainty is $0.53\,\mathrm{W\,m^{-2}}$ and consists almost entirely of the random error (median $\sigma_{\mathrm{rand}} = 0.51\,\mathrm{W\,m^{-2}}$, against $\sigma_{\mathrm{instr}} = 0.12\,\mathrm{W\,m^{-2}}$; Table 2). As an independent, purely empirical check of the random error we slide the 5 km window along the full leg in steps of 0.05 km and take for every segment the standard deviation of all window estimates that overlap it, each weighted by its fractional overlap with the segment.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

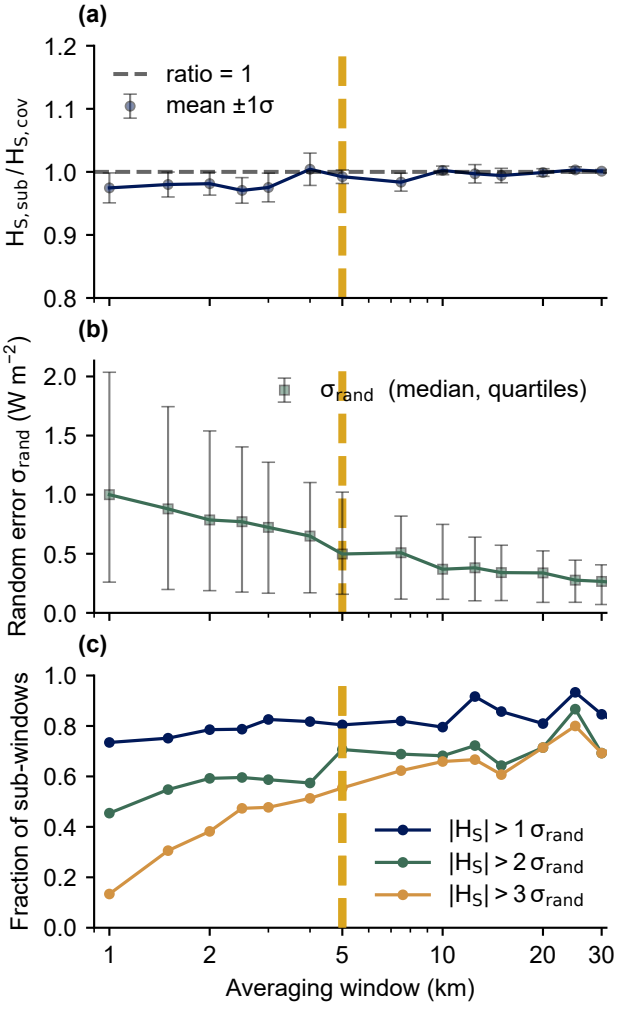


Figure 3. Sensitivity of the EC sensible heat flux to the length of the averaging sub-window for the six T-Bird low-level legs during BACSAM-II (median leg length 70.2 km). **(a)** Mean EC flux of the sub-windows relative to the EC flux of the stretch they cover ($H_{\mathrm{S,sub}}/H_{\mathrm{S,cov}}$), averaged over five positions of the window grid (mean and standard deviation over the legs with a flux of at least $0.5\,\mathrm{W\,m^{-2}}$). **(b)** Random error σ_{rand} of the sub-window fluxes (median and quartiles; Appendix A). **(c)** Fraction of sub-windows whose flux magnitude exceeds 1, 2, and $3\,\sigma_{\mathrm{rand}}$. The dashed vertical line marks the 5 km window used in this study.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

This segment-placement spread has a median of $0.70\,\mathrm{W\,m^{-2}}$ and ranks the segments like σ_{rand} (rank correlation 0.73), which confirms the magnitude of the random error.

3.2 Bulk sensible heat flux method

The bulk estimates are evaluated for the same 5 km along-track segments as the EC reference, so that all SHF estimates are compared over a common analysis domain. Each estimate combines $\Delta\theta$, U, and C_{h} through Eq. (2) and differs from the others only in its input data (Sects. 2.1–2.3 and Table 1).

The transfer coefficient C_{h} in Eq. (2) depends on the roughness lengths for momentum and heat and on the near-surface stability (Appendix B). The momentum roughness length is prescribed for each segment from its surface composition, derived from $T_{\mathrm{s,obs}}$ of the KT-19 as a surface-type proxy. This proxy is suitable because T_{s} separates the surface types of heterogeneous sea ice (Müller et al., 2025), although T_{s} also responds to the atmospheric state (Shupe et al., 2026). Each segment is partitioned into fractions of open water, thin sea ice, and cold snow-covered ice, and the corresponding momentum roughness lengths are combined into an effective value $z_{0,m}^{\mathrm{eff}}$ by a fraction-weighted geometric mean (Table B1 lists the thresholds and $z_{0,m}$ values). These $z_{0,m}$ values are deliberately prescribed rather than derived from the measured momentum fluxes, because the algorithm must not depend on momentum-flux measurements, which a remote approach lacks. Their influence is tested in Sect. 4.2 by replacing them with two alternative sets from the literature (Appendix B), namely the MOSAiC set, based on the roughness length prescribed for the bulk fluxes of the MOSAiC drift (Cox et al., 2023), and the Schnierstein set, taken from large-eddy simulations of a sea-ice lead (Schnierstein and Neggers, 2026). All three bulk estimates use this segment roughness. For the ERA5–MODIS and ERA5-only estimates it is the only in situ information. A spatially uniform $z_{0,m}$, which requires no in situ information, changes the Pearson correlation of the ERA5–MODIS estimate by only 0.03 but its slope from 0.59 to as much as 0.74 (Table C1).

The effective scalar roughness length follows as

$$z_{0,T}^{\mathrm{eff}} = \alpha_T \cdot z_{0,m}^{\mathrm{eff}}, \tag{4}$$

with a reference ratio $\alpha_T = 10^{-1}$, the value used in the large-eddy simulations of Schnierstein et al. (2024) (Sect. 1) and in CARRA (Sect. 2.2), which we refer to as the conventional value. We additionally report an amplitude-matched configuration, $\alpha_T = 10^{-2}$, defined in Sect. 4.2 as the value at which the T-Bird bulk estimate reproduces the observed flux amplitude. Table C1 lists all tested configurations.

Near-surface atmospheric stability is represented by the bulk Richardson number Ri_{b}, which quantifies the ratio of buoyancy to shear from the potential temperature and wind speed differences between the surface and the T-Bird level. The reference configuration uses Richardson-number-based (Louis-type, Louis, 1979) transfer coefficients with the stability functions of Gryanik et al. (2020, GLGS20), which allow a non-iterative parameterization of C_{h} in neutral and stable conditions. The Businger–Dyer stability functions are used in unstable conditions (Businger et al., 1971; Dyer, 1974). In the sensitivity analysis, we additionally test a set of alternative stability functions for stable conditions, implemented following Gryanik and Lüpkes (2023, their Eqs. 11–19b).

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

The bulk estimate is computed entirely from segment-mean input data. Therefore, it carries no random sampling term, and its total 1σ uncertainty follows from first-order Gaussian propagation of the input uncertainties through the bulk formula. We propagate the uncertainties in T_s ($T_{s,obs}$ or $T_{s,MOD}$), the air temperature, and the wind speed, with the assumed 1σ values listed in Table 2. The median total uncertainty is $2.50\,\mathrm{W\,m^{-2}}$ for the T-Bird bulk and $6.95\,\mathrm{W\,m^{-2}}$ for the ERA5–MODIS bulk SHF values. For the ERA5–MODIS approach these inputs represent product (retrieval and model) uncertainties rather than instrument errors. The sensitivity of the bulk flux to the stability scheme, the scalar roughness ratio, and the roughness set is deliberately excluded from this uncertainty estimate and quantified separately in Sect. 4.2.

Table 2. Flux-uncertainty budget for the three estimates. For $H_{S,obs,EC}$ the total 1σ uncertainty σ_{EC} (bold) is the quadrature sum of the random error σ_{rand} and the instrument error σ_{instr}. For the bulk estimates the total uncertainty follows from first-order Gaussian propagation of the input uncertainties stated in the column "Component / input uncertainties". Absolute entries show the median $[p_{25}–p_{75}]$ across the 91 segments in $\mathrm{W\,m^{-2}}$. The relative uncertainty $\sigma/|H_S|$ is computed over segments with $|H_S| > 2\,\mathrm{W\,m^{-2}}$. The input uncertainties for ERA5 and MODIS are assumed values, of the same order as the errors reported by Graham et al. (2019) and Nielsen-Englyst et al. (2023), respectively.

Method	Component / input uncertainties	σ $[p_{25}–p_{75}]$ ($\mathrm{W\,m^{-2}}$)	$\sigma/\lvert H_S\rvert$ $[p_{25}–p_{75}]$ (%)
$H_{S,obs,EC}$	Total σ_{EC}	**0.53 [0.16–0.99]**	**18 [16–21]**
	Random error σ_{rand} (Finkelstein and Sims, 2001)	0.51 [0.16–0.98]	18 [15–21]
	Instrument error σ_{instr}	0.12 [0.01–0.25]	5 [5–5]
$H_{S,obs,bulk}$	$\sigma(T_{s,obs}) = 0.5\,\mathrm{K}$, $\sigma(T_{air}) = 0.3\,\mathrm{K}$, $\sigma(U) = 0.5\,\mathrm{m\,s^{-1}}$	**2.50 [0.13–3.90]**	**53 [38–96]**
$H_{S,RS,bulk}$	$\sigma(T_{s,MOD}) = 1.5\,\mathrm{K}$, $\sigma(T_{air}) = 1.2\,\mathrm{K}$, $\sigma(U) = 1.7\,\mathrm{m\,s^{-1}}$	**6.95 [0.02–11.09]**	**186 [158–237]**

All SHF estimates, bulk and model-diagnosed, are evaluated against the EC reference with the same metrics. The sampled SHF values are small, so the RMSE discriminates poorly in this regime. It rewards estimates that shrink towards zero. Therefore, the estimates are evaluated by three quantities that separate the amplitude of the estimated SHF from their spatial pattern. The amplitude is measured by the dimensionless slope b of the ordinary least-squares regression of the estimate on the EC reference. Since the transfer coefficient scales the flux, an error in C_h appears directly as $b \neq 1$. The ability of the parameterized flux to represent the observed variability is judged by its Pearson correlation r, which is unaffected by any amplitude scaling and thus measures whether the flux maxima and minima fall on the right segments. Furthermore, we calculate the bias as the mean difference between each parameterized SHF estimate and the EC reference, so that a positive bias denotes an overestimate.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

4 Results

Figure 4 shows the four measurement areas, whose locations within the campaign region are marked in Fig. 1. On 13 April the T-Bird sampled consolidated pack ice reaching 85.3° N, where the EC flux is close to zero. The largest fluxes, up to 11.7 W m^{-2}, occur on 15 April east of Svalbard over sea ice with open leads. The two polynya overflights on 17 and 23 April 2024 contrast open-water and ice-covered segments. As Whalers Bay, a polynya north of Svalbard, starts to open (Fig. 4c) and then fully opens (Fig. 4d), the larger positive EC fluxes are confined to the segments crossing open water, whereas the adjacent ice-covered segments show near-zero or negative fluxes, down to −7.3 W m^{-2}.

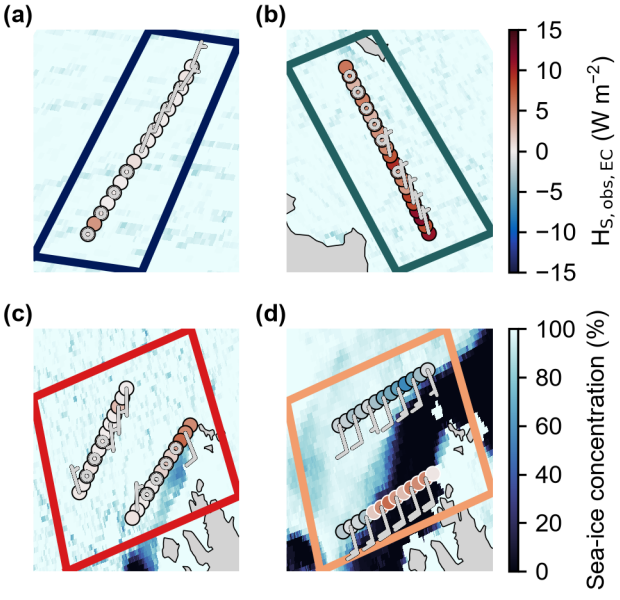


Figure 4. EC fluxes of the 5 km low-level segments flown during BACSAM-II (April 2024). Red/blue colour shading indicates the EC sensible heat flux, white/blue colour shading the sea-ice concentration, and wind barbs denote the mean wind speed and direction. The frame colours correspond to the focus-region rectangles in Fig. 1. **(a)** 13 April 2024, northernmost flight leg over consolidated pack ice. **(b)** 15 April 2024, compact sea ice with leads east of Svalbard. **(c)** 17 April 2024, overflight of the opening Whalers Bay polynya north of Svalbard over mostly thin sea ice. **(d)** 23 April 2024, same area as in (c), with the polynya now exposing open water.

4.1 Comparison of sensible heat fluxes obtained with the different methods

The agreement between the SHF estimates and the EC reference differs markedly between the six flight legs. Over the consolidated pack ice of 13 April (Fig. 5a) the EC time series is close to zero except for a peak near the end of the leg, which $H_{S,RS,bulk}$ also shows. The other estimates are close to zero as well, except $H_{S,obs,bulk}$, which, after a small peak at the start of the leg, is slightly negative before rising to two peaks, the first of which far exceeds the weak EC signal at the same position, where the MODIS/Terra IST imagery shows warm anomalies (Fig. 5g). On 15 April the EC flux varies strongly along the leg (Fig. 5b).

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

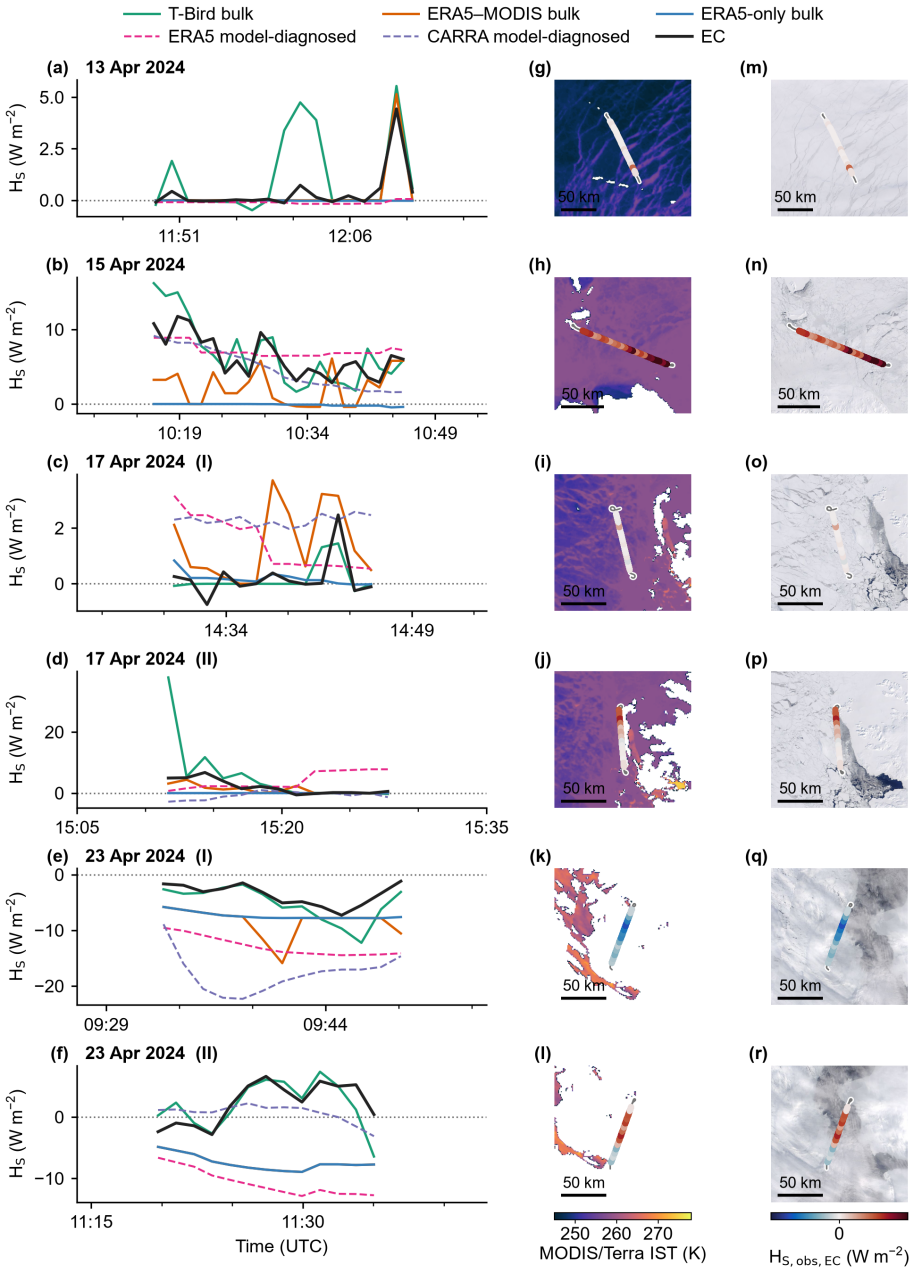


Figure 5. (a)–(f) Along-track SHF time series of the six low-level flight legs during BACSAM-II from the T-Bird bulk, ERA5–MODIS bulk, and ERA5-only bulk estimates, the ERA5 and CARRA model-diagnosed fluxes, and the EC reference (Table 1). The dotted grey line marks zero SHF. Positive values denote upward turbulent heat flux from the surface to the atmosphere. **(g)–(l)** MODIS/Terra IST imagery of the same day (NASA Global Imagery Browse Services) with the analysed flight legs. White areas indicate missing values (cloud-obscured pixels). **(m)–(r)** MODIS/Terra true-colour imagery of the same scenes, with the flight legs coloured by $H_{\mathrm{S,obs,EC}}$.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

$H_{\mathrm{S,obs,bulk}}$ reproduces this variability, whereas $H_{\mathrm{S,RS,bulk}}$ does so only weakly. As on all legs, the model-diagnosed fluxes vary only smoothly along the track and do not resolve the segment-to-segment variability. On 17 April (Fig. 5c, d), $H_{\mathrm{S,obs,bulk}}$ shows isolated peaks that appear only weakly in the EC flux. The two legs flown on 23 April 2024 (Fig. 5e, f) show the largest spread between the SHF estimates of the different methods, with the model-diagnosed fluxes reaching -14.5 (-22.3) W m^{-2} for ERA5 (CARRA), while the EC flux is negative along the first leg, down to -7.3 W m^{-2}, and mostly positive along the second, and $H_{\mathrm{S,obs,bulk}}$ follows the EC flux on both legs. Cloud cover leaves a valid MODIS pixel on only three segments of the first leg, so $H_{\mathrm{S,RS,bulk}}$ differs from $H_{\mathrm{S,ERA,bulk}}$ on these three segments only. Overall, $H_{\mathrm{S,obs,bulk}}$ captures much of the along-track variability of the EC flux, although individual peaks of $H_{\mathrm{S,obs,bulk}}$ and $H_{\mathrm{S,RS,bulk}}$ lack an EC counterpart (Fig. 5d).

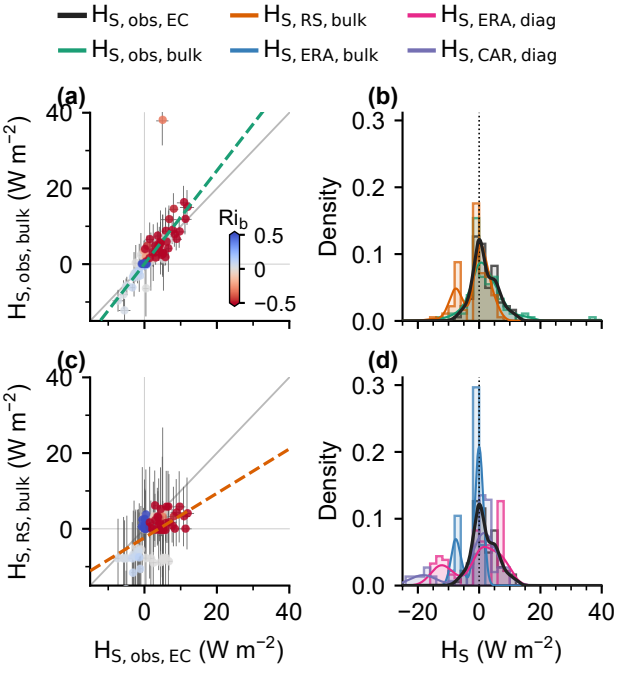


Figure 6. Comparison of the bulk and model-diagnosed sensible heat flux estimates with the EC reference over the 91 segments. **(a)** Scatter comparison of the T-Bird bulk flux $H_{\mathrm{S,obs,bulk}}$ with the EC flux $H_{\mathrm{S,obs,EC}}$, colour-shaded by the in situ bulk Richardson number Ri_{b} of the T-Bird bulk estimate (colour scale saturating beyond ± 0.5). The dashed line shows the ordinary least-squares fit of the estimate on the reference, the grey line the 1:1 relation, and grey bars the 1σ uncertainty of each estimate. **(b)** Histograms of $H_{\mathrm{S,obs,EC}}$, $H_{\mathrm{S,obs,bulk}}$, and $H_{\mathrm{S,RS,bulk}}$. Continuous lines show the smoothed kernel-density estimate. **(c)** As (a), for $H_{\mathrm{S,RS,bulk}}$, with the colours again showing the in situ Ri_{b}. **(d)** As (b), for the reanalysis-based estimates $H_{\mathrm{S,ERA,bulk}}$, $H_{\mathrm{S,ERA,diag}}$, and $H_{\mathrm{S,CAR,diag}}$ ($n = 74$), with the EC reference in black.

A quantitative comparison and scatter plots are shown in Fig. 6 and Table 3. With measured inputs, the T-Bird bulk estimate overestimates the flux amplitude in the reference configuration (GLGS20 stability scheme, conventional $\alpha_T = 10^{-1}$, surface-type-dependent roughness). The regression slope is 1.23, which corresponds to an overestimation of the SHF amplitude by

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

23 %, and the bias is $+0.47\,\mathrm{W\,m^{-2}}$ with respect to the EC reference (Fig. 6a). It nevertheless tracks the observed variability over all 91 segments better than any other estimate, explaining more than half of the variance of the EC flux ($r = 0.77$, $r^2 = 0.59$). The ERA5–MODIS bulk estimate $H_{\mathrm{S,RS,bulk}}$ (Fig. 6c) recovers more than half of the observed amplitude (slope 0.59, $r = 0.48$, $r^2 = 0.24$), whereas the ERA5-only bulk estimate $H_{\mathrm{S,ERA,bulk}}$ recovers only about a third (slope 0.31, $r = 0.36$, $r^2 = 0.13$). Furthermore, the flux distributions differ in shape (Fig. 6b, d). The ERA5-only estimate $H_{\mathrm{S,ERA,bulk}}$ produces mostly zero and negative SHF, while $H_{\mathrm{S,RS,bulk}}$ tracks the positive tail of the observed SHF more closely. Both reanalysis-driven bulk estimates also have a secondary mode of negative SHF, which stems from the two 23 April legs (Fig. 5e, f).

The SHF values that the models diagnose internally behave differently from our offline ERA5-only bulk estimate. Both overestimate the amplitude, ERA5 by 10 % and CARRA by 40 % (slope 1.10 for $H_{\mathrm{S,ERA,diag}}$ and 1.40 for $H_{\mathrm{S,CAR,diag}}$), and their biases are -2.59 and $-3.36\,\mathrm{W\,m^{-2}}$, respectively. Their Pearson correlations with the EC reference ($r = 0.56$ and 0.72, $r^2 = 0.31$ and 0.51) are higher than that of the ERA5-only bulk estimate ($r = 0.36$). CARRA is unavailable on 13 April, so it is evaluated on 74 of the 91 segments. On that common subset, CARRA overestimates the amplitude more than ERA5 but correlates better with the EC reference ($b = 1.17$ and 1.40, $r = 0.58$ and 0.72), while their biases are close (-3.07 and $-3.36\,\mathrm{W\,m^{-2}}$).

The advantage of $H_{\mathrm{S,RS,bulk}}$ over $H_{\mathrm{S,ERA,bulk}}$ is concentrated where T_{s} is heterogeneous. To show this, we characterize the surface heterogeneity of each segment by the standard deviation of the KT-19 $T_{\mathrm{s,obs}}$ within the segment and split the segments at the median of 1.1 K into a heterogeneous and a homogeneous half. In the heterogeneous half, the ERA5-only estimate is anticorrelated with the EC flux ($r = -0.36$, slope -0.02), while replacing its T_{s} with $T_{\mathrm{s,MOD}}$ raises the Pearson correlation to $r = 0.49$ and the slope to 0.26. This gain in Pearson correlation is significant in the heterogeneous half of the segments, whereas in the homogeneous half the Pearson correlations of the two estimates with the EC reference are indistinguishable.

Table 3. Agreement of each SHF estimate with the EC reference over the 91 segments in the reference configuration. The dimensionless slope b of the regression on the EC flux measures the amplitude, r is the Pearson correlation over all segments, r^2 the explained variance, and the bias is the mean of the estimate minus the EC flux. *CARRA is unavailable on 13 April and is evaluated on 74 segments.

Estimate	slope b	r	r^2	bias ($\mathrm{W\,m^{-2}}$)
$H_{\mathrm{S,obs,bulk}}$	1.23	0.77	0.59	+0.47
$H_{\mathrm{S,RS,bulk}}$	0.59	0.48	0.24	−3.05
$H_{\mathrm{S,ERA,bulk}}$	0.31	0.36	0.13	−3.84
$H_{\mathrm{S,ERA,diag}}$	1.10	0.56	0.31	−2.59
$H_{\mathrm{S,CAR,diag}}{}^{*}$	1.40	0.72	0.51	−3.36

4.2 Sensitivity analysis

Beyond the per-segment measurement uncertainty quantified in Sect. 3.2, the retrieved SHF also depends on the methodological choices embedded in the bulk formulation, whose influence we isolate one component at a time for the T-Bird and

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

ERA5–MODIS bulk estimates, with all other settings at the reference configuration. We vary the Monin–Obukhov stability scheme, the scalar roughness ratio $\alpha_T = z_{0,T}/z_{0,m}$, and the momentum roughness set, and we perturb the two governing input variables, the near-surface potential temperature difference $\Delta\theta$ (by ± 0.5 K) and the wind speed U (by $\pm 0.5\,\mathrm{m\,s^{-1}}$, with C_h held at its unperturbed value). Whereas the uncertainty budget (Table 2) measures the random and instrument error of a single configuration, these experiments rank the methodological choices by their effect on the agreement with the EC reference. They are summarized in Table 4 and Fig. 7 and tabulated in full in Table C1 (Appendix C). Since each range depends on the values tested, the comparison below is conditional on those choices.

Table 4. Summary of the bulk-SHF sensitivity experiments for the T-Bird and ERA5–MODIS bulk estimates. For the stability scheme, α_T, and roughness experiments, the listed values are the minimum-to-maximum regression slope b and Pearson correlation r across the tested choices.

	T-Bird bulk		ERA5–MODIS bulk	
Experiment	slope b	r	slope b	r
Stability scheme (8 schemes)	1.20–1.23	0.76–0.77	0.47–0.64	0.43–0.48
α_T (10^{-4}–10^{-1})	0.79–1.23	0.77–0.77	0.45–0.59	0.47–0.48
Momentum roughness set	1.21–1.38	0.76–0.77	0.59–0.66	0.49–0.50
Constant $z_{0,m}$	1.22–1.56	0.77–0.77	0.59–0.74	0.51–0.51
Input perturbations (median [interquartile range] across segments)				
$\partial H_\mathrm{S}/\partial(\Delta\theta)$	-1.80 [-4.56, -0.05] $\mathrm{W\,m^{-2}\,K^{-1}}$		-1.42 [-4.67, 0.00] $\mathrm{W\,m^{-2}\,K^{-1}}$	
$\partial H_\mathrm{S}/\partial U$	0.08 [-0.03, 4.30] $\mathrm{W\,m^{-2}(m\,s^{-1})^{-1}}$		0.00 [-0.86, 2.01] $\mathrm{W\,m^{-2}(m\,s^{-1})^{-1}}$	

The choice of the Monin–Obukhov stability scheme is immaterial here. Across the eight schemes the T-Bird bulk slope is 1.20–1.23 at an almost constant r (0.76–0.77), and ERA5–MODIS behaves similarly (slope 0.47–0.64, with the slopes of all schemes within 0.17 of each other).

The scalar roughness ratio changes the flux amplitude most, yet it does not change the Pearson correlation with the EC reference. Lowering α_T from the conventional 10^{-1} to 10^{-2}, with the momentum roughness held at its reference set so that only $z_{0,T}$ varies, reduces C_h and thereby the T-Bird bulk slope from 1.23 to 1.03 and the bias from $+0.47$ to $-0.01\,\mathrm{W\,m^{-2}}$. The value $\alpha_T = 10^{-2}$ thus reproduces the observed flux amplitude and defines the amplitude-matched configuration used alongside the reference. The Pearson correlation, however, is $r = 0.77$–0.78 across all tested α_T, from 10^{-1} to 10^{-7}, and nearly unchanged for the two reanalysis-driven estimates (0.45–0.48 and 0.36–0.37).

Of the two governing input variables, the near-surface temperature difference has the larger median effect. Per kelvin of $\Delta\theta$ the SHF changes by a median $-1.80\,\mathrm{W\,m^{-2}}$ for the T-Bird bulk and $-1.42\,\mathrm{W\,m^{-2}}$ for ERA5–MODIS, against only 0.08 and $0.00\,\mathrm{W\,m^{-2}}$ per $\mathrm{m\,s^{-1}}$ of wind speed. The temperature derivative is negative because a larger $\Delta\theta$ means a colder surface relative to the air (Eq. (3)) and hence a smaller upward flux. Both medians hide a strong state dependence. The interquartile ranges are wide, and the wind-speed sensitivity even changes sign between segments (Table 4), with the strongest wind-speed response on the segments with large $|\Delta\theta|$.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

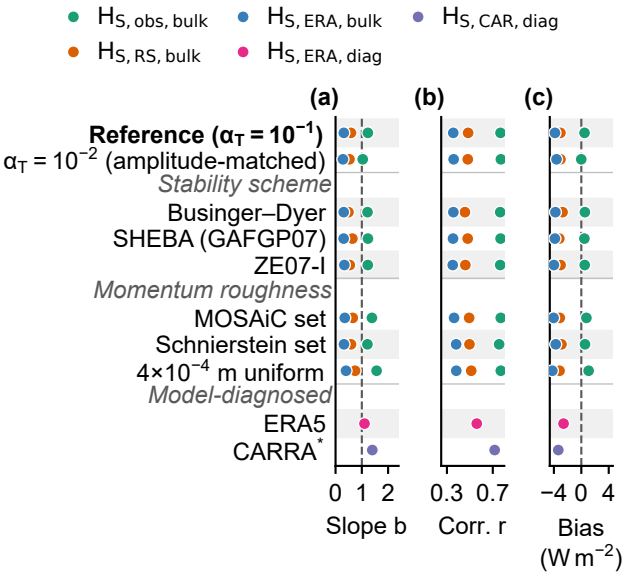


Figure 7. Skill of representative bulk configurations against the EC reference for $H_{\mathrm{S,obs,bulk}}$ (green), $H_{\mathrm{S,RS,bulk}}$ (orange), and $H_{\mathrm{S,ERA,bulk}}$ (blue). **(a)** Amplitude skill, given by the regression slope b on the EC flux. **(b)** Variability skill, given by the Pearson correlation r. **(c)** Bias. The grey dashed lines in (a) and (c) mark the ideal values. The model-diagnosed fluxes $H_{\mathrm{S,ERA,diag}}$ and $H_{\mathrm{S,CAR,diag}}$ * (bottom rows) do not depend on the bulk configuration. The first row is the reference configuration (GLGS20, conventional $\alpha_T = 10^{-1}$, surface-type-dependent roughness). Every other row changes one setting at a time. Only representative members of each experiment are shown. Table C1 lists the slopes and Pearson correlations of the T-Bird and ERA5–MODIS bulk estimates for every tested configuration. *CARRA is unavailable on 13 April and is evaluated on 74 instead of 91 segments.

In summary, the amplitude of the retrieved SHF is governed by the transfer coefficient and the near-surface temperature difference. How well an estimate tracks the observed variability is set by its input data alone, namely the surface skin temperature, air temperature, and wind speed that enter Eq. (2), and no parameterization choice tested here improves it.

5 Discussion

Supplied with measured, collocated inputs, the bulk method overestimates the SHF amplitude of the EC reference by 23 %, yet it reproduces the observed variability better than any other estimate. However, individual segments deviate substantially (Fig. 5a). Published evaluations against the N-ICE2015 drift, sampled in far more energetic conditions, report winter RMSE values of 11–41 W m^{-2} for five of six reanalyses and 70 W m^{-2} for the sixth (Graham et al., 2019). With SHF values of at most 11.7 W m^{-2} here, those numbers do not transfer, whereas a fractional amplitude error does.

The benefit of replacing the reanalysis surface skin temperature with a satellite retrieval is largest over heterogeneous sea-ice surfaces. On 15 April 2024, the sea ice was heterogeneous, with leads visible from Polar 6 that also appear in the MODIS/Terra

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

true-colour imagery (Fig. 5n). On this leg, the satellite retrieval reduces the magnitude of the bias and raises the slope, although the correlation within the leg does not improve. Over all segments, the improvement of $H_{\mathrm{S,RS,bulk}}$ occurs mainly over heterogeneous surfaces (Sect. 4.1), indicating that the surface-temperature boundary condition controls the observed spatial variability. This contrasts with Zhang et al. (2023), where satellite-based T_{s} raised the RMSE of the bulk SHF against the N-ICE2015 measurements from 18 to 36–76 W m^{-2}. A possible reason is that these measurements sampled a much more homogeneous sea-ice floe (Walden et al., 2017) than the airborne measurements used here.

The model-diagnosed SHF values disagree in amplitude and in distribution with respect to the EC reference, not merely in the mean, and they explain 31 % (ERA5) and 51 % (CARRA) of the observed variance (Table 3). They nevertheless track the observed variability better than the ERA5-only bulk estimate ($r = 0.56$ and 0.72 versus 0.36), because their internal schemes use their own, mutually consistent skin temperature. The offline estimate approaches the diagnosed SHF only once $T_{\mathrm{s,MOD}}$ replaces $T_{\mathrm{s,ERA}}$, and does not surpass it.

The dominance of $\Delta\theta$ over U (Table 4) is not a strict separation, because U also enters C_{h} through the stratification, which the wind-speed perturbation holds fixed, but over the sampled range it is the spread of the observed $\Delta\theta$ distribution (Fig. 2c) that decides the differences between the estimates. The stability scheme, by contrast, hardly matters here. The stability correction f_{h} itself is not small, but all tested schemes share the Businger–Dyer functions in unstable conditions and differ only in stable conditions (Sect. 3.2). The SHF is small there, and the sampled stratification was predominantly weak (Fig. 2c).

Surface roughness matters more through its magnitude than through its surface-type dependence. Exchanging our class values for the MOSAiC or the Schnierstein set changes the T-Bird bulk slope from 1.23 to 1.38 and 1.21, respectively, at an almost unchanged Pearson correlation ($r = 0.76$–0.77). The effect of the surface-type classification derived from T_{s} is therefore minor. Our blending approach (Eq. (B4)) does not explicitly represent ridges or form drag, but under the near-neutral conditions sampled here a spatially uniform roughness changes the Pearson correlation by at most 0.03 and shifts the regression slope between 1.22 and 1.56, whereas varying α_T over three decades changes it from 0.79 to 1.23 (Table C1).

Reproducing the observed amplitude therefore requires $\alpha_T \approx 10^{-2}$, which reduces the regression slope by 16 %. This value lies one order of magnitude below the conventional value and two orders below the $\alpha_T = 1$ used by ERA5 over compact sea ice. It also lies well below the ratio near unity that surface-renewal theory predicts at the small roughness Reynolds numbers of these segments (Andreas, 1987), so it is an empirical value. Airborne EC estimates over heterogeneous sea ice point to even lower values (Schröder et al., 2003). Published reductions based on a scalar roughness that depends on the roughness Reynolds number concern aerodynamically rough flow (Elvidge et al., 2023; Srivastava et al., 2024), not the small roughness Reynolds numbers sampled here. We do not interpret this value as a direct retrieval of $z_{0,T}$, however, because the inferred amplitude may also reflect residual biases in the input data (T_{s}, θ_{air}, U), scale effects, or the 0.1 Hz lower limit of the EC SHF (Appendix A). For example, on 13 April the T-Bird may have flown above a shallow stable surface layer that did not transport the full surface flux to the measurement height. The KT-19 samples the surface directly below the track, whereas the EC flux integrates over a source area upwind of the T-Bird. Where the wind crossed the legs, as on 17 and 23 April (Fig. 4c, d), warm surfaces such as the polynya can therefore enter one estimate but not the other. The result should thus be interpreted as evidence for an overly large effective transfer coefficient under the sampled spring conditions, rather than as a universal scalar roughness value. Its

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

generality is further limited by the four flight days, the predominantly near-neutral stratification, the simplified surface classes, and the daily rather than instantaneous $T_{\mathrm{s,MOD}}$.

The relevance of these results is relative rather than absolute. In April over consolidated Arctic sea ice, the SHF is a minor term of the surface energy budget, an order of magnitude below the radiative components, so an error of a few W m^{-2} matters little for the energy balance of these particular days, unlike over winter leads and during cold-air outbreaks (Sect. 1). Because the amplitude errors identified here are fractional properties of the bulk estimates, however, they carry into the regimes that dominate the Arctic turbulent heat budget, where the SHF reaches hundreds of W m^{-2} over winter leads (Maykut, 1978; Lüpkes et al., 2008; Marcq and Weiss, 2012; Ólason et al., 2021) and during marine cold-air outbreaks (Brümmer, 1996; Slättberg et al., 2025). Moreover, a campaign characterized by weak SHF and predominantly near-neutral stratification provides a stringent test, because parameterization errors are less easily obscured by large turbulent signals.

6 Summary and conclusions

We present low-level airborne eddy-covariance (EC) measurements of the sensible heat flux (SHF) from the towed Turbulence Bird (T-Bird) platform, collected during the aircraft campaign Boundary Layer and Aerosol and Cloud Study in the Arctic based on Aircraft and T-Bird Measurements II (BACSAM-II, Svalbard, April 2024), and use them to evaluate bulk aerodynamic SHF estimates driven by in situ, satellite, and reanalysis data over Arctic sea ice. Flying the T-Bird at a mean altitude of only 23 m allowed us to measure the near-surface heat exchange over leads, thin ice, and consolidated pack ice. A window-length analysis established 5 km as a suitable averaging length for the comparison.

Three main findings emerge. First, the surface skin temperature governs whether a bulk estimate reproduces the observed flux variability. Driven by ERA5 alone, the bulk method recovers only about a third of the observed flux amplitude and explains 13 % of its variance (slope 0.31, $r = 0.36$). Replacing the reanalysis T_{s} with the MODIS retrieval raises these to more than half of the amplitude and 24 % of the variance (slope 0.59, $r = 0.48$), and most of the gain comes from the half of the segments with the more heterogeneous surface.

Second, in its conventional configuration the bulk method overestimates the flux amplitude. Driven by the KT-19 surface skin temperature and the T-Bird atmospheric measurements, it tracks the observed variability over all 91 segments better than any other estimate ($r = 0.77$), but its regression slope on the EC reference is 1.23 (95 % confidence interval 1.04–1.46). With the EC SHF restricted to 0.1–20 Hz, matching the amplitude requires $\alpha_T \approx 10^{-2}$, an order of magnitude below the conventional value, which reduces the regression slope by 16 %.

Third, the results depend strongly on the scalar roughness ratio and the near-surface temperature difference, and, in the median, more weakly on the wind speed. In the measured stability range, which is close to neutral, the choice of the Monin–Obukhov stability scheme has a minor impact on the T-Bird bulk estimate, although errors in the input data large enough to change the stability regime reintroduce a scheme dependence, as the ERA5–MODIS estimate shows (slopes 0.47–0.64 across the schemes).

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Improving satellite- and reanalysis-based SHF over Arctic sea ice therefore depends on better constraints on the scalar roughness ratio, for which our data, with the EC SHF restricted to 0.1–20 Hz, favour a value below the conventional 10^{-1}, and on representing the sub-grid heterogeneity of the near-surface temperature difference. Models are already moving in this direction by replacing constant ratios with roughness-Reynolds-number and form-drag formulations (Elvidge et al., 2023). Beyond this, the study points towards a pan-Arctic SHF retrieval based not on reanalysis surface skin temperatures alone, but on dedicated satellite surface-temperature products such as the combined sea- and sea-ice surface temperature climate data set of Nielsen-Englyst et al. (2023). Combined with satellite-based surface-roughness information (Zhang et al., 2025) and new high-resolution pan-Arctic reanalyses such as CARRA2 (Pelt et al., 2025), such satellite products could improve our understanding of past SHF trends and help to constrain regional climate models.

Code and data availability. The analysis code is archived at https://doi.org/10.5281/zenodo.23731340 (Müller, 2026a), a companion website to interactively explore the code is available at https://franzflink.github.io/bacsam2-flux-analysis/. The T-Bird measurements (wind, air temperature, pressure, and position), the Polar 6 nose-boom measurements, and the KT-19 surface skin temperatures are available at https://doi.org/10.1594/PANGAEA.995979 (Köhler et al., 2026). The regional subsets of the reanalysis and satellite products listed below and the table of the flight legs, as read by the code, are archived at https://doi.org/10.5281/zenodo.23729591 (Müller, 2026b). The derived sensible heat fluxes are produced by the code from these data. The Polar 6 master tracks are available at https://doi.org/10.1594/PANGAEA.971763 (Jurányi et al., 2024). ERA5 (Hersbach et al., 2020) model-level (Copernicus Climate Change Service, 2023) and single-level data (C3S, 2018), and CARRA data (Copernicus Climate Change Service, 2021b, a), are available from the Copernicus Climate Change Service. The MODIS/Aqua ice-surface temperature product (Hall and Riggs, 2021) is available from the National Snow and Ice Data Center, and the MODIS/Terra ice-surface temperature and true-colour imagery from NASA's Global Imagery Browse Services. The merged MODIS–AMSR2 sea-ice concentration (Ludwig et al., 2020) is available at https://data.seaice.uni-bremen.de/modis_amsr2.

Appendix A: Quality of the EC reference

Most EC SHF values measured during BACSAM-II are close to zero, which makes judging their quality by relative criteria impractical. Both the relative stationarity test of Foken and Wichura (1996) and the coefficient of determination of a linear fit to the cumulative covariance along a segment relate the within-segment variation of the covariance to the magnitude of the SHF itself. In the present data set, these criteria are governed by the SHF magnitude rather than by the measurement quality, with rank correlations of -0.51 and $+0.68$ with $|H_{\mathrm{S,obs,EC}}|$. A near-zero SHF value violates such a criterion irrespective of the quality of the underlying measurement, and a selection based on it would systematically remove small SHF values and bias the comparison towards large ones. The quality assessment therefore consists of an integrity check of the resolved turbulence and the random error of each SHF value. Both are reported for each segment, and neither is applied as a selection criterion. Table A1 shows how conventional selection criteria would change the results.

The integrity check determines whether the 0.1–20 Hz frequency interval used for the EC SHF estimate contains a resolved inertial subrange. Within this interval, the power spectrum of the vertical wind speed is expected to decrease with frequency in

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

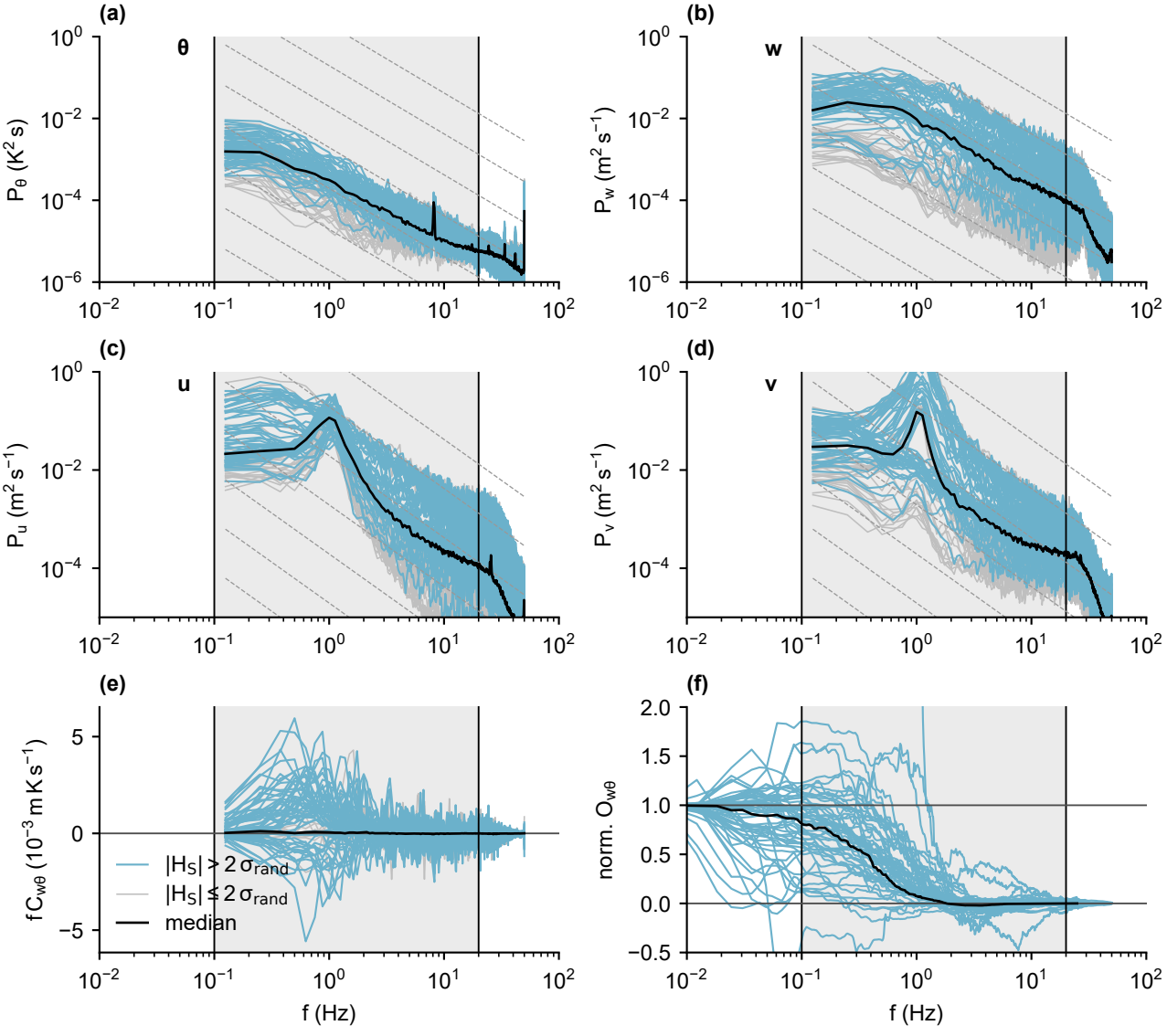


Figure A1. Spectral diagnostics of the 91 T-Bird segments. Blue lines mark segments whose SHF exceeds $2\,\sigma_{\mathrm{rand}}$, grey lines the segments below $2\,\sigma_{\mathrm{rand}}$, and the black line is the median. **(a–d)** One-sided power spectra of potential temperature, vertical wind speed, and the horizontal wind components. Grey shading marks the 0.1–20 Hz frequency interval used for the EC SHF estimate. Diagonal dashed lines show reference slopes proportional to $f^{-5/3}$. **(e)** Frequency-weighted w–θ cospectra. **(f)** Normalized w–θ ogives, the cumulative covariance from high to low frequencies, for the SHF values above $2\,\sigma_{\mathrm{rand}}$. The spectra in (a)–(e) are Welch estimates with 8 s windows and therefore start at 0.125 Hz, whereas the ogives in (f) are computed from the full segment.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

proportion to $f^{-5/3}$, which corresponds to the inertial-subrange law of Kolmogorov (1941) and Obukhov (1941) transformed from wavenumber to frequency with the true airspeed. The spectral slope of each segment is determined by linear regression in log–log space, and a segment passes the check if the slope deviates by at most 0.4 from $-5/3$. Of the 91 segments, 41 fulfil this criterion, and the slopes range from -1.6 to -1.0 with a median of -1.2 (Fig. A1b). The w–θ cospectra peak well inside the interval (Fig. A1e), and the ogives, the cumulative covariance integrated from high to low frequencies, accumulate most of the covariance there (Fig. A1f). The flux outside this interval is not negligible. For the 47 segments with $|H_{\mathrm{S,obs,EC}}| > 2\,\mathrm{W\,m^{-2}}$, the covariance of all resolved frequencies exceeds the band-limited one by a median of $15\,\%$ (quartiles $-3\,\%$ and $+35\,\%$), and against this all-frequency flux the slope of the T-Bird bulk estimate is 1.04 instead of 1.23. The band-limited EC SHF is therefore lower than the total flux on average, though not on every segment, and a slope above unity is conditional on the lower limit. The criterion assesses only the spectral shape of the vertical wind speed and does not constrain the SHF magnitude. Since it is not used for selection, the only conditions for a segment of the four flight days over sea ice (Sect. 2.1) to enter the analysis are a mean T-Bird height below $50\,\mathrm{m}$, complete data, and a T-Bird height variation of less than $30\,\mathrm{m}$ within the segment.

The random error σ_{rand} quantifies the sampling uncertainty of the covariance, namely the standard deviation that repeated realizations of the same $5\,\mathrm{km}$ segment under identical conditions would yield. A segment of this length comprises a limited number of heat-carrying eddies. The longer the integral time scales of w' and θ', the fewer independent samples the segment contains, so the uncertainty of the covariance increases with these time scales and with the cross-correlation of w' and θ'. Following Finkelstein and Sims (2001), σ_{rand} is calculated from the auto- and cross-covariance functions of the band-passed series,

$$\sigma_{\mathrm{rand}}^2 = \frac{(\rho \cdot c_{\mathrm{p}})^2}{N} \sum_{k=-m}^{m} \left[\gamma_{ww}(k)\,\gamma_{\theta\theta}(k) + \gamma_{w\theta}(k)\,\gamma_{\theta w}(k)\right], \tag{A1}$$

with the auto- and cross-covariance functions γ of w' and θ' band-passed to 0.1–$20\,\mathrm{Hz}$, the number of samples N in the segment, and the maximum lag m corresponding to $10\,\mathrm{s}$. The result varies by less than $5\,\%$ for m between 5 and $30\,\mathrm{s}$ (cf. Rannik et al., 2016). Table A1 shows that the selection criteria change the regression slopes of the comparison only slightly. For every screening, the regression slope of the T-Bird bulk estimate remains within the $95\,\%$ bootstrap confidence interval of its reference value (1.04–1.46) and the slopes of the two reanalysis-driven bulk estimates remain below unity. The Pearson correlations depend more on the screening. On the 41 segments that pass the integrity check, the ERA5 model-diagnosed SHF correlates more strongly with the EC reference than the T-Bird bulk estimate ($r = 0.77$ versus 0.65), and with the strictest tolerance (29 segments) the ERA5-only bulk estimate is anticorrelated with it ($r = -0.26$). Consequently, the complete set of 91 segments, each with its random error, is used throughout this study.

Appendix B: Bulk-transfer implementation

This appendix provides the common equations and parameter choices used to derive the bulk sensible heat flux. All quantities are evaluated on the same $5\,\mathrm{km}$ segments as the EC reference.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table A1. Regression slope b and Pearson correlation r of four SHF estimates against the EC reference for alternative screenings of the 91 segments. The first row is the data set used in this study. The relative stationarity test follows Foken and Wichura (1996). σ_{rand} is the random error of the EC flux. s_w is the inertial-subrange slope of the vertical-wind spectrum (Fig. A1b).

		$H_{\mathrm{S,obs,bulk}}$		$H_{\mathrm{S,RS,bulk}}$		$H_{\mathrm{S,ERA,bulk}}$		$H_{\mathrm{S,ERA,diag}}$	
Screening	n	b	r	b	r	b	r	b	r
none (this study)	91	1.23	0.77	0.59	0.48	0.31	0.36	1.10	0.56
$R^2 \geq 0.4$ of the cumulative covariance	77	1.25	0.78	0.62	0.53	0.35	0.41	1.15	0.60
$R^2 \geq 0.6$	71	1.25	0.77	0.65	0.56	0.37	0.45	1.18	0.62
relative stationarity test $< 100\,\%$	83	1.23	0.77	0.61	0.51	0.34	0.40	1.14	0.58
relative stationarity test $< 30\,\%$	53	1.17	0.93	0.69	0.59	0.41	0.51	1.24	0.64
segment halves agree within $3\sqrt{2}\sigma_{\mathrm{rand}}$	68	1.20	0.92	0.54	0.49	0.31	0.38	1.07	0.58
only fluxes above $2\sigma_{\mathrm{rand}}$	58	1.25	0.77	0.74	0.62	0.46	0.56	1.34	0.69
integrity check, $\lvert s_w + 5/3\rvert \leq 0.3$	29	1.12	0.47	0.15	0.21	−0.02	−0.26	0.44	0.55
integrity check, $\lvert s_w + 5/3\rvert \leq 0.4$	41	1.16	0.65	0.46	0.63	0.23	0.54	0.97	0.77
integrity check, $\lvert s_w + 5/3\rvert \leq 0.5$	71	1.25	0.75	0.53	0.48	0.27	0.34	1.03	0.57
integrity check, $\lvert s_w + 5/3\rvert \leq 0.6$	86	1.23	0.76	0.59	0.49	0.32	0.36	1.11	0.57

B1 Surface–air temperature difference and bulk Richardson number

Surface skin temperatures measured by the KT-19, retrieved from MODIS, or taken from ERA5 are converted to the surface potential temperature according to

$$\theta_{\mathrm{s}} = T_{\mathrm{s}} \left(\frac{p_0}{p_{\mathrm{s}}} \right)^{R_{\mathrm{d}}/c_{\mathrm{p}}}, \tag{B1}$$

where $p_0 = 1000\,\mathrm{hPa}$ is the reference pressure, p_{s} is the local surface pressure, $R_{\mathrm{d}} = 287.05\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific gas constant for dry air, and $c_{\mathrm{p}} = 1004\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific heat capacity of dry air. For the reanalysis-driven estimates, p_{s} is the ERA5 surface pressure. For the T-Bird bulk estimate, no surface pressure is measured, and p_{s} is reconstructed from the pressure p_z measured at the T-Bird height z with the hypsometric equation,

$$p_{\mathrm{s}} = p_z \exp\left(\frac{g\,z}{R_{\mathrm{d}}\,\overline{T}} \right), \tag{B2}$$

with $g = 9.80665\,\mathrm{m\,s^{-2}}$ the gravitational acceleration and $\overline{T}$ the mean of the surface and air temperatures. The air potential temperature is computed with the pressure at its own level, the measured T-Bird pressure or the pressure of the ERA5 model level. Using the T-Bird pressure also at the surface would overestimate θ_{s} by about $gz/c_{\mathrm{p}} \approx 0.2\,\mathrm{K}$.

Near-surface stability is quantified by the bulk Richardson number, computed from the two-level scalar inputs at the surface and the T-Bird level:

$$Ri_{\mathrm{b}} = \frac{g/\overline{\theta}\,\Delta\theta\,\Delta z}{\Delta u^2 + \Delta v^2}, \tag{B3}$$

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

where g is the gravitational acceleration, $\overline{\theta}$ the layer-mean potential temperature, $\Delta\theta$, Δu, Δv are finite differences of the potential temperature and of the zonal (u) and meridional (v) wind components between the surface and the T-Bird level, and $\Delta z = z$ is the T-Bird measurement height, or the height of the ERA5 level (about 30 m, Sect. 2.2) for the reanalysis-driven estimates. The horizontal wind at the surface is assumed to vanish (no-slip), so Δu and Δv are the T-Bird-level wind components.

For the reanalysis-driven estimates ($H_{\mathrm{S,RS,bulk}}$ and $H_{\mathrm{S,ERA,bulk}}$), the wind speed entering Eq. (B3) is the effective wind speed of Jordan et al. (1999) for stable and Fairall et al. (1996) for unstable stratification, following Zhang et al. (2025). This accounts for windless exchange and convective gustiness and thereby prevents the divergence of Ri_{b} at vanishing wind speed. No further limitation is applied. For the T-Bird bulk estimate, the measured wind is used directly, and $\varepsilon\,\Delta z^2$ with $\varepsilon = 10^{-4}\,\mathrm{s}^{-2}$ is added to the denominator $\Delta u^2 + \Delta v^2$ of Eq. (B3) to avoid a division by zero at vanishing wind speed. Because the measured wind and $\Delta\theta$ are mutually consistent, $|Ri_{\mathrm{b}}|$ remains below 16.6 on all segments but one, where the wind speed is only $0.4\,\mathrm{m\,s}^{-1}$ ($Ri_{\mathrm{b}} = -18.0$). As the transfer coefficients saturate beyond $|Ri_{\mathrm{b}}| \approx 1$, the flux statistics are insensitive to any additional bound on Ri_{b} (changes of at most $0.08\,\mathrm{W\,m}^{-2}$ in bias and RMSE and 0.02 in slope for bounds of 1, 2.5, and 5, not shown). In the unstable branch, the Monin–Obukhov stability parameter $\zeta = z/L$, with L the Obukhov length, is limited to -20, which holds C_{h} constant in strongly unstable stratification. This limit is reached on 12 of the 91 segments for the T-Bird bulk estimate (19 for ERA5–MODIS and 6 for ERA5-only), and removing it changes the regression slopes by at most 0.01.

B2 Surface-type-dependent roughness

The effective momentum roughness length $z_{0,m}^{\mathrm{eff}}$ is calculated from the fractional coverage f_i of each surface type i within each 5 km segment,

$$\ln z_{0,m}^{\mathrm{eff}} = \sum_{i=1}^{3} f_i \ln z_{0,m,i}, \qquad \sum_{i=1}^{3} f_i = 1, \tag{B4}$$

where $i = 1, 2, 3$ denotes open water, thin ice, and cold snow-covered ice. The scalar roughness length is prescribed as

$$z_{0,T}^{\mathrm{eff}} = \alpha_T\, z_{0,m}^{\mathrm{eff}}, \qquad \alpha_T = 10^{-1}, \tag{B5}$$

in the reference configuration, with the amplitude-matched value $\alpha_T = 10^{-2}$ reported alongside (Sect. 4.2).

The surface classes are derived from all KT-19 $T_{\mathrm{s,obs}}$ samples within the segment. The cold-ice threshold is deliberately selected conservatively. It defines a roughness end-member for cold snow-covered ice rather than a direct retrieval of physical ice thickness. The two alternative sets tested in Sect. 4.2 assign either $z_{0,m} = 3.27 \times 10^{-4}$ m, the value prescribed for the MOSAiC bulk fluxes (Cox et al., 2023), to open water and cold snow-covered ice and 2×10^{-4} m to thin ice (MOSAiC set), or the lead value of Schnierstein and Neggers (2026), 1×10^{-4} m, to open water and thin ice and their sea-ice value, 5×10^{-3} m, to cold snow-covered ice (Schnierstein set).

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table B1. Surface-temperature classes and prescribed surface roughness used to derive the effective roughness length of each 5 km segment.

Surface class	KT-19 surface temperature	$z_{0,m}$
Open water	$T_\mathrm{s} > -1.8\,^\circ\mathrm{C}$	2.0×10^{-4} m
Thin ice	$-20\,^\circ\mathrm{C} < T_\mathrm{s} \leq -1.8\,^\circ\mathrm{C}$	1.0×10^{-4} m
Cold snow-covered ice	$T_\mathrm{s} \leq -20\,^\circ\mathrm{C}$	4.0×10^{-4} m

B3 Transfer coefficient and stability correction

The neutral transfer coefficient is

$$C_{\mathrm{h,n}} = \frac{\kappa_{\mathrm{vK}}^2}{\mathrm{Pr}_{t,0}\,\ln\!\left(z/z_{0,m}^{\mathrm{eff}}\right)\ln\!\left(z/z_{0,T}^{\mathrm{eff}}\right)}, \tag{B6}$$

where $\kappa_{\mathrm{vK}} = 0.40$ is the von Kármán constant and $\mathrm{Pr}_{t,0}$ the neutral turbulent Prandtl number of the respective stability formulation. The stability-corrected coefficient is written generally as

$$C_\mathrm{h} = C_{\mathrm{h,n}}\, f_\mathrm{h}\!\left(Ri_\mathrm{b},\, z/z_{0,m}^{\mathrm{eff}},\, z/z_{0,T}^{\mathrm{eff}}\right), \tag{B7}$$

where f_h is the scheme-specific stability correction (GLGS20 in the reference configuration), specified in Gryanik and Lüpkes (2023, their Eqs. 11–19b).

Appendix C: Full sensitivity matrix

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table C1. Sensitivity of the T-Bird and ERA5–MODIS bulk estimates to every tested configuration (Fig. 7). Listed are the regression slope b on the EC reference and the Pearson correlation r. "Ref." marks a parameter held at its reference value.

Configuration label	Stability function	α_T	$z_{0,m}$ [open/thin/thick] (m)	T-Bird bulk		ERA5–MODIS	
				slope b	r	slope b	r
Reference configuration (conventional α_T)							
GLGS20_ref	Gryanik et al. (2020)	10^{-1}	2×10^{-4} / 1×10^{-4} / 4×10^{-4}	**1.23**	**0.77**	**0.59**	**0.48**
Amplitude-matched configuration							
GLGS20_ztr1e-02	GLGS20	10^{-2}	Ref.	**1.03**	**0.77**	**0.53**	**0.48**
Stability scheme (all other settings at reference)							
BD_ref	Businger et al. (1971); Dyer (1974)	10^{-1}	Ref.	1.22	0.76	0.49	0.46
HB88_ref	Holtslag and de Bruin (1988)	10^{-1}	Ref.	1.22	0.77	0.64	0.48
BH91_ref	Beljaars and Holtslag (1991)	10^{-1}	Ref.	1.22	0.76	0.64	0.48
CB05_ref	Chenge and Brutsaert (2005)	10^{-1}	Ref.	1.20	0.76	0.61	0.48
GAFGP07_ref	Grachev et al. (2007)	10^{-1}	Ref.	1.23	0.77	0.64	0.48
ZE07-I_ref	Zilitinkevich and Esau (2007)	10^{-1}	Ref.	1.23	0.76	0.54	0.46
ZE07-II_ref	Zilitinkevich and Esau (2007)	10^{-1}	Ref.	1.22	0.76	0.47	0.43
Scalar roughness ratio α_T (GLGS20)							
GLGS20_ztr1e-03	GLGS20	10^{-3}	Ref.	0.90	0.77	0.49	0.48
GLGS20_ztr1e-04	GLGS20	10^{-4}	Ref.	0.79	0.77	0.45	0.47
GLGS20_ztr1e-05	GLGS20	10^{-5}	Ref.	0.71	0.78	0.43	0.46
GLGS20_ztr1e-06	GLGS20	10^{-6}	Ref.	0.64	0.78	0.40	0.46
GLGS20_ztr1e-07	GLGS20	10^{-7}	Ref.	0.59	0.78	0.38	0.45
Momentum roughness set (GLGS20)							
GLGS20_mosaic	GLGS20	10^{-1}	3.3×10^{-4} / 2×10^{-4} / 3.3×10^{-4}	1.38	0.77	0.66	0.49
GLGS20_nschnier	GLGS20	10^{-1}	1×10^{-4} / 1×10^{-4} / 5×10^{-3}	1.21	0.76	0.59	0.50
Constant (surface-type-independent) $z_{0,m}$ (GLGS20)							
GLGS20_z0c1e-04	GLGS20	10^{-1}	$z_{0,m} = 1\times10^{-4}$ m	1.22	0.77	0.59	0.51
GLGS20_z0c2e-04	GLGS20	10^{-1}	$z_{0,m} = 2\times10^{-4}$ m	1.38	0.77	0.66	0.51
GLGS20_z0c4e-04	GLGS20	10^{-1}	$z_{0,m} = 4\times10^{-4}$ m	1.56	0.77	0.74	0.51

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions. JJM conceptualized the study, performed the analysis, produced the figures, and wrote the manuscript. IS contributed to the EC retrieval and processing. DJS contributed to the flight segmentation and participated in the field campaign. LK and JH performed the preparation and correction of the 100 Hz aircraft data and participated in the field campaign. FG, AE, MDS, MS, and MW contributed to the conceptualization and development of the study. All authors contributed to the interpretation of the scientific results, discussed the findings, revised the manuscript, and contributed to the formulation of the final text.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques.

Acknowledgements. We thank Christof Lüpkes for the fruitful discussions concerning the manuscript. We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project Number 268020496, TRR 172, within the framework of the Transregional Collaborative Research Center "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms $(\mathcal{AC})^3$". MDS was supported by a Mercator fellowship with $(\mathcal{AC})^3$ and by the NOAA Global Ocean Monitoring and Observing program https://doi.org/10.13039/100018302 via cooperative agreement NA22OAR4320151. The authors are thankful to the Alfred Wegener Institute (AWI) for providing and operating the Polar 6 aircraft and T-Bird. We thank the crews and the technicians of the research aircraft for excellent technical and logistical support. The generous funding of the flight hours for the Polar 6 aircraft by AWI is greatly appreciated. The authors acknowledge the support of the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, through grant number AWI_PA_02172. This publication was supported by the Open Access Publishing Fund of Leipzig University. We thank the Institute of Environmental Physics, University of Bremen for the provision of the merged MODIS–AMSR2 sea-ice concentration data at https://data.seaice.uni-bremen.de/modis_amsr2 (last access: 16 March 2026). This study contains modified Copernicus Climate Change Service information (2026), namely the ERA5 reanalysis (Hersbach et al., 2020) and the Copernicus Arctic Regional Reanalysis (CARRA; Copernicus Climate Change Service, 2021b, a). Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. We acknowledge the use of imagery provided by services from NASA's Global Imagery Browse Services (GIBS), part of NASA's Earth Science Data and Information System (ESDIS). This manuscript benefited from language improvements and structural suggestions generated with the assistance of large language models (ChatGPT with GPT-4o and GPT-5; Claude Sonnet 4.6, Opus 4.6, Fable 5.1, and Opus 5.5). All code was written in collaboration with agentic AI tools (Claude Code with Claude Sonnet 4.6, Opus 4.6, Fable 5.1, and Opus 5.5). All scientific reasoning, data analysis, interpretation, and final wording were conceived, verified, and approved by the authors. The authors remain fully responsible for the content.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Akansu, E. F., Dahlke, S., Siebert, H., and Wendisch, M.: Evaluation of methods to determine the surface mixing layer height of the atmospheric boundary layer in the central Arctic during polar night and transition to polar day in cloudless and cloudy conditions, Atmospheric Chemistry and Physics, 23, 15 473–15 489, https://doi.org/10.5194/acp-23-15473-2023, 2023.

Andreas, E., Persson, O., Grachev, A., Jordan, R., Horst, T., Guest, P., and Fairall, C.: Parameterizing Turbulent Exchange over Sea Ice in Winter, Journal of Hydrometeorology, 11, 87–104, https://doi.org/10.1175/2009JHM1102.1, 2010.

Andreas, E. L.: A theory for the scalar roughness and the scalar transfer coefficients over snow and sea ice, Boundary-Layer Meteorology, 38, 159–184, https://doi.org/10.1007/BF00121562, 1987.

Batrak, Y. and Müller, M.: On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice, Nature Communications, 10, 4170, https://doi.org/10.1038/s41467-019-11975-3, 2019.

Batrak, Y., Cheng, B., and Kallio-Myers, V.: Sea ice cover in the Copernicus Arctic Regional Reanalysis, The Cryosphere, 18, 1157–1183, https://doi.org/10.5194/tc-18-1157-2024, 2024.

Beljaars, A. C. M. and Holtslag, A. a. M.: Flux Parameterization over Land Surfaces for Atmospheric Models, Journal of Applied Meteorology, 30, 327–341, https://doi.org/10.1175/1520-0450(1991)030<0327:FPOLSF>2.0.CO;2, 1991.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, X., and Køltzow, M. Ø.: The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System, Monthly Weather Review, 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2017.

Boisvert, L. N., Boeke, R. C., Taylor, P. C., and Parker, C. L.: Constraining Arctic Climate Projections of Wintertime Warming With Surface Turbulent Flux Observations and Representation of Surface-Atmosphere Coupling, Frontiers in Earth Science, 10, 765 304, https://doi.org/10.3389/feart.2022.765304, 2022.

Brümmer, B.: Boundary-layer modification in wintertime cold-air outbreaks from the Arctic sea ice, Boundary-Layer Meteorology, 80, 109–125, https://doi.org/10.1007/BF00119014, 1996.

Brümmer, B., Busack, B., Hoeber, H., and Kruspe, G.: Boundary-layer observations over water and arctic sea-ice during on-ice air flow, Boundary-Layer Meteorology, 68, 75–108, https://doi.org/10.1007/BF00712665, 1994.

Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F.: Flux-Profile Relationships in the Atmospheric Surface Layer, Journal of the Atmospheric Sciences, 28, 181–189, https://doi.org/10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2, 1971.

C3S: ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.adbb2d47, 2018.

Chechin, D. G., Makhotina, I. A., Lüpkes, C., and Makshtas, A. P.: Effect of Wind Speed and Leads on Clear-Sky Cooling over Arctic Sea Ice during Polar Night, Journal of the Atmospheric Sciences, 76, 2481–2503, https://doi.org/10.1175/JAS-D-18-0277.1, 2019.

Chen, S.-Y., Spreen, G., Hui, F.-M., Cheng, X., and Zhang, Z.-L.: High-resolution Arctic sea ice lead variations during wintertime (2016–2023) from SAR observations, Advances in Climate Change Research, 16, 922–934, https://doi.org/10.1016/j.accre.2025.08.006, 2025.

Chenge, Y. and Brutsaert, W.: Flux-profile Relationships for Wind Speed and Temperature in the Stable Atmospheric Boundary Layer, Boundary-Layer Meteorology, 114, 519–538, https://doi.org/10.1007/s10546-004-1425-4, 2005.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Copernicus Climate Change Service: Arctic Regional Reanalysis on Height Levels from 1991 to Present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.8679900d, accessed: 2026-06-15, 2021a.

Copernicus Climate Change Service: Arctic Regional Reanalysis on Single Levels from 1991 to Present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.713858f6, accessed: 2026-06-15, 2021b.

Copernicus Climate Change Service: Complete ERA5 global atmospheric reanalysis, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.143582cf, 2023.

Cox, C. J., Gallagher, M. R., Shupe, M. D., Persson, P. O. G., Solomon, A., Fairall, C. W., Ayers, T., Blomquist, B., Brooks, I. M., Costa, D., Grachev, A., Gottas, D., Hutchings, J. K., Kutchenreiter, M., Leach, J., Morris, S. M., Morris, V., Osborn, J., Pezoa, S., Preußer, A., Riihimaki, L. D., and Uttal, T.: Continuous observations of the surface energy budget and meteorology over the Arctic sea ice during MOSAiC, Scientific Data, 10, 519, https://doi.org/10.1038/s41597-023-02415-5, 2023.

Dyer, A. J.: A review of flux-profile relationships, Boundary-Layer Meteorology, 7, 363–372, https://doi.org/10.1007/BF00240838, 1974.

ECMWF: IFS Documentation CY41R2 – Part IV: Physical Processes, IFS Documentation 4, ECMWF, Reading, UK, https://doi.org/10.21957/tr5rv27xu, 2016.

Elvidge, A. D., Renfrew, I. A., Brooks, I. M., Srivastava, P., Yelland, M. J., and Prytherch, J.: Surface Heat and Moisture Exchange in the Marginal Ice Zone: Observations and a New Parameterization Scheme for Weather and Climate Models, Journal of Geophysical Research: Atmospheres, 126, e2021JD034 827, https://doi.org/10.1029/2021JD034827, 2021.

Elvidge, A. D., Renfrew, I. A., Edwards, J. M., Brooks, I. M., Srivastava, P., and Weiss, A. I.: Improved Simulation of the Polar Atmospheric Boundary Layer by Accounting for Aerodynamic Roughness in the Parameterization of Surface Scalar Exchange Over Sea Ice, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 305, https://doi.org/10.1029/2022MS003305, 2023.

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., and Young, G. S.: Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment, Journal of Geophysical Research: Oceans, 101, 3747–3764, https://doi.org/10.1029/95JC03205, 1996.

Finkelstein, P. L. and Sims, P. F.: Sampling error in eddy correlation flux measurements, Journal of Geophysical Research: Atmospheres, 106, 3503–3509, https://doi.org/10.1029/2000JD900731, 2001.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agricultural and Forest Meteorology, 78, 83–105, https://doi.org/10.1016/0168-1923(95)02248-1, 1996.

Grachev, A., Andreas, E., Fairall, C., Guest, P., and Persson, O.: Turbulent measurements in the stable atmospheric boundary layer during SHEBA: Ten years after, Acta Geophysica, 56, 142–166, https://doi.org/10.2478/s11600-007-0048-9, 2008.

Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., and Persson, P. O. G.: SHEBA flux–profile relationships in the stable atmospheric boundary layer, Boundary-Layer Meteorology, 124, 315–333, https://doi.org/10.1007/s10546-007-9177-6, 2007.

Graham, R. M., Cohen, L., Ritzhaupt, N., Segger, B., Graversen, R. G., Rinke, A., Walden, V. P., Granskog, M. A., and Hudson, S. R.: Evaluation of Six Atmospheric Reanalyses over Arctic Sea Ice from Winter to Early Summer, Journal of Climate, 32, 4121–4143, https://doi.org/10.1175/JCLI-D-18-0643.1, 2019.

Gryanik, V. M. and Lüpkes, C.: A Package of Momentum and Heat Transfer Coefficients for the Stable Surface Layer Extended by New Coefficients over Sea Ice, Boundary-Layer Meteorology, 187, 41–72, https://doi.org/10.1007/s10546-022-00730-9, 2023.

Gryanik, V. M., Lüpkes, C., Grachev, A., and Sidorenko, D.: New Modified and Extended Stability Functions for the Stable Boundary Layer based on SHEBA and Parametrizations of Bulk Transfer Coefficients for Climate Models, Journal of the Atmospheric Sciences, 77, 2687–2716, https://doi.org/10.1175/JAS-D-19-0255.1, 2020.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Hall, D. K. and Riggs, G. A.: MODIS/Aqua Sea Ice Extent and IST Daily L3 Global 4km EASE-Grid Day (MYD29E1D), Version 61, NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], https://doi.org/10.5067/MODIS/MYD29E1D.061, 2021.

Herrmannsdörfer, L., Müller, M., Shupe, M. D., and Rostosky, P.: Surface temperature comparison of the Arctic winter MOSAiC observations, ERA5 reanalysis, and MODIS satellite retrieval, Elementa: Science of the Anthropocene, 11, 00085, https://doi.org/10.1525/elementa.2022.00085, 2023.

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

Holtslag, A. A. M. and de Bruin, H. A. R.: Applied Modeling of the Nighttime Surface Energy Balance over Land., Journal of Applied Meteorology, 27, 689–704, https://doi.org/10.1175/1520-0450(1988)027<0689:AMOTNS>2.0.CO;2, 1988.

Jordan, R. E., Andreas, E. L., and Makshtas, A. P.: Heat budget of snow-covered sea ice at North Pole 4, Journal of Geophysical Research: Oceans, 104, 7785–7806, https://doi.org/10.1029/1999JC900011, 1999.

Jozef, G. C., Cassano, J. J., Dahlke, S., Dice, M., Cox, C. J., and de Boer, G.: An overview of the vertical structure of the atmospheric boundary layer in the central Arctic during MOSAiC, Atmospheric Chemistry and Physics, 24, 1429–1450, https://doi.org/10.5194/acp-24-1429-2024, 2024.

Jurányi, Z., Stratmann, F., Herber, A., Wendisch, M., and Borrmann, S.: Master tracks in different resolutions during POLAR 6 campaign P6-247_BACSAM2_2024, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.971763, 2024.

Jurányi, Z., Lüpkes, C., Stratmann, F., Hartmann, J., Schaefer, J., Jörss, A.-M., Schulz, A., Wetzel, B., Simon, D., Gebhard, E., Stöhr, M., Hofmann, P., Kalmbach, D., Grawe, S., Wendisch, M., and Herber, A.: The T-Bird – a new aircraft-towed instrument platform to measure aerosol properties and turbulence close to the surface: introduction to the aerosol measurement system, Atmospheric Measurement Techniques, 18, 3477–3494, https://doi.org/10.5194/amt-18-3477-2025, 2025.

Köhler, L., Borrmann, S., Hartmann, J., Hartmann, M., Henning, S., Jörss, A.-M., Jurányi, Z., Müller, J., Pithan, F., Schaefer, J., Schulz, A., Simon, D. J., Stratmann, F., Wendisch, M., Wetzel, B., and Herber, A.: Observations from the BACSAM II aircraft campaign with Polar 6, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.995979, 2026.

Kolmogorov, A. N.: The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers, Doklady Akademii Nauk SSSR, 30, 301–305, english translation: Proceedings of the Royal Society of London. Series A, 434, 9–13, 1991, doi:10.1098/rspa.1991.0075, 1941.

Kwok, R., Cunningham, G. F., Wensnahan, M., Rigor, I., Zwally, H. J., and Yi, D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, Journal of Geophysical Research: Oceans, 114, C07005, https://doi.org/10.1029/2009JC005312, 2009.

Køltzow, M., Schyberg, H., Støylen, E., and Yang, X.: Value of the Copernicus Arctic Regional Reanalysis (CARRA) in representing near-surface temperature and wind speed in the north-east European Arctic, Polar Research, 41, 8002, https://doi.org/10.33265/polar.v41.8002, 2022.

Langford, B., Acton, W., Ammann, C., Valach, A., and Nemitz, E.: Eddy-covariance data with low signal-to-noise ratio: time-lag determination, uncertainties and limit of detection, Atmospheric Measurement Techniques, 8, 4197–4213, https://doi.org/10.5194/amt-8-4197-2015, 2015.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



Lindsay, R. W. and Rothrock, D. A.: Arctic sea ice leads from advanced very high resolution radiometer images, Journal of Geophysical Research: Oceans, 100, 4533–4544, https://doi.org/10.1029/94JC02393, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/94JC02393, 1995.

Louis, J.-F.: A parametric model of vertical eddy fluxes in the atmosphere, Boundary-Layer Meteorology, 17, 187–202, https://doi.org/10.1007/BF00117978, 1979.

Ludwig, V., Spreen, G., and Pedersen, L. T.: Evaluation of a New Merged Sea-Ice Concentration Dataset at 1 km Resolution from Thermal Infrared and Passive Microwave Satellite Data in the Arctic, Remote Sensing, 12, 3183, https://doi.org/10.3390/rs12193183, 2020.

Lüpkes, C., Gryanik, V. M., Witha, B., Gryschka, M., Raasch, S., and Gollnik, T.: Modeling convection over arctic leads with LES and a non-eddy-resolving microscale model, Journal of Geophysical Research: Oceans, 113, C09 028, https://doi.org/10.1029/2007JC004099, 2008.

Mahrt, L.: Flux Sampling Errors for Aircraft and Towers, Journal of Atmospheric and Oceanic Technology, 15, 416–429, https://doi.org/10.1175/1520-0426(1998)015<0416:FSEFAA>2.0.CO;2, 1998.

Marcq, S. and Weiss, J.: Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere, The Cryosphere, 6, 143–156, https://doi.org/10.5194/tc-6-143-2012, 2012.

Maslanik, J., Stroeve, J., Fowler, C., and Emery, W.: Distribution and trends in Arctic sea ice age through spring 2011, Geophysical Research Letters, 38, L13 502, https://doi.org/10.1029/2011GL047735, 2011.

Masson, V., Le Moigne, P., Martin, E., et al.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes, Geoscientific Model Development, 6, 929–960, https://doi.org/10.5194/gmd-6-929-2013, 2013.

Maykut, G. A.: Energy exchange over young sea ice in the central Arctic, Journal of Geophysical Research: Oceans, 83, 3646–3658, https://doi.org/10.1029/JC083iC07p03646, 1978.

Monin, A. S. and Obukhov, A. M.: Basic Laws of Turbulent Mixing in the Surface Layer of the Atmosphere, Trudy Geofizicheskogo Instituta AN SSSR, 24, 163–187, originally published in Russian. English translation by John Miller for the Geophysics Research Directorate, Air Force Cambridge Research Center, Translation No. 1234, January 1959; later retyped in LaTeX and edited by Keith McNaughton (2008)., 1954.

Müller, J. J.: BACSAM-II sensible-heat-flux analysis (code), version 1.0.1, Zenodo [code], https://doi.org/10.5281/zenodo.23731340, 2026a.

Müller, J. J.: ERA5, CARRA and MODIS subsets and flight-leg table for evaluating sensible heat flux over Arctic sea ice (BACSAM-II, April 2024), Zenodo [data set], https://doi.org/10.5281/zenodo.23729591, 2026b.

Müller, J. J., Schäfer, M., Rosenburg, S., Ehrlich, A., and Wendisch, M.: High-resolution maps of Arctic surface skin temperature and type retrieved from airborne thermal infrared imagery collected during the HALO–$(\mathcal{AC})^3$ campaign, Atmospheric Measurement Techniques, 18, 4695–4708, https://doi.org/10.5194/amt-18-4695-2025, 2025.

Nielsen-Englyst, P., Høyer, J. L., Kolbe, W. M., Dybkjær, G., Lavergne, T., Tonboe, R. T., Skarpalezos, S., and Karagali, I.: A combined sea and sea-ice surface temperature climate dataset of the Arctic, 1982–2021, Remote Sensing of Environment, 284, 113 331, https://doi.org/10.1016/j.rse.2022.113331, 2023.

Obukhov, A. M.: On the Energy Distribution in the Spectrum of a Turbulent Flow, C. R. Acad. Sci. URSS, 32, 19–21, 1941.

Paluch, I. R., Lenschow, D. H., and Wang, Q.: Arctic boundary layer in the fall season over open and frozen sea, Journal of Geophysical Research: Atmospheres, 102, 25 955–25 971, https://doi.org/10.1029/97JD01563, 1997.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Pelt, S., Støylen, E., Schyberg, H., Yang, X., Dahlgren, P., Ødegård Køltzow, M. A., Bojarova, J., Nielsen, K. P., and Peralta, C.: Pan-Arctic Regional Reanalysis Sub-Daily Data from 1986 to Present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/f5effe24, accessed: 2026-06-24, 2025.

Persson, P. O. G., Fairall, C. W., Andreas, E. L., Guest, P. S., and Perovich, D. K.: Measurements near the Atmospheric Surface Flux Group tower at SHEBA: Near-surface conditions and surface energy budget, Journal of Geophysical Research: Oceans, 107, 8045, https://doi.org/10.1029/2000JC000705, 2002.

Pinto, J. O., Alam, A., Maslanik, J. A., Curry, J. A., and Stone, R. S.: Surface characteristics and atmospheric footprint of springtime Arctic leads at SHEBA, Journal of Geophysical Research: Oceans, 108, 8051, https://doi.org/10.1029/2000JC000473, 2003.

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal images, The Cryosphere, 13, 1565–1582, https://doi.org/10.5194/tc-13-1565-2019, 2019.

Rannik, Ü., Peltola, O., and Mammarella, I.: Random uncertainties of flux measurements by the eddy covariance technique, Atmospheric Measurement Techniques, 9, 5163–5181, https://doi.org/10.5194/amt-9-5163-2016, 2016.

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, Communications Earth & Environment, 3, 168, https://doi.org/10.1038/s43247-022-00498-3, 2022.

Renfrew, I. A., Barrell, C., Elvidge, A. D., Brooke, J. K., Duscha, C., King, J. C., Kristiansen, J., Lachlan-Cope, T., Moore, G. W. K., Pickart, R. S., Reuder, J., Sandu, I., Sergeev, D., Terpstra, A., Våge, K., and Weiss, A.: An evaluation of surface meteorology and fluxes over the Iceland and Greenland Seas in ERA5 reanalysis: The impact of sea ice distribution, Quarterly Journal of the Royal Meteorological Society, 147, 691–712, https://doi.org/10.1002/qj.3941, 2021.

Rheinlænder, J. W., Regan, H., Rampal, P., Boutin, G., Ólason, E., and Davy, R.: Breaking the Ice: Exploring the Changing Dynamics of Winter Breakup Events in the Beaufort Sea, Journal of Geophysical Research: Oceans, 129, e2023JC020 395, https://doi.org/10.1029/2023JC020395, 2024.

Schnierstein, N. and Neggers, R. A. J.: Lead Impacts on the Moist Static Energy Budget of the Low-Level Arctic Atmosphere in Large-Eddy Simulations Based on MOSAiC Data, Journal of Advances in Modeling Earth Systems, 18, e2025MS005 560, https://doi.org/10.1029/2025MS005560, 2026.

Schnierstein, N., Chylik, J., Shupe, M. D., and Neggers, R. a. J.: Standardized Daily High-Resolution Large-Eddy Simulations of the Arctic Boundary Layer and Clouds During the Complete MOSAiC Drift, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 296, https://doi.org/10.1029/2024MS004296, 2024.

Schröder, D., Vihma, T., Kerber, A., and Brümmer, B.: On the parameterization of turbulent surface fluxes over heterogeneous sea ice surfaces, Journal of Geophysical Research: Oceans, 108, 3195, https://doi.org/10.1029/2002JC001385, 2003.

Shupe, M. D., Persson, P. O. G., Cox, C. J., Gallagher, M. R., Solomon, A., Sledd, A., Blomquist, B., Brooks, I. M., Costa, D., Osborn, J., Perovich, D., Riihimaki, L. D., and Uttal, T.: The two radiative states of the Arctic atmosphere and their impacts on the surface energy budget of sea ice, Elementa: Science of the Anthropocene, 14, 00 100, https://doi.org/10.1525/elementa.2025.00100, 2026.

Simon, D. J., Hartmann, J., Schaefer, J., Zeppenfeld, S., Lüpkes, C., Hartmann, M., Wetzel, B., Heinold, B., Jurányi, Z., Schulz, A., Köhler, L., Jörss, A.-M., Herber, A., Henning, S., Pöhlker, M. L., Roberts, G. C., and Stratmann, F.: Turbulent Aerosol Fluxes From Airborne Measurements Over the Arctic Ocean, Geophysical Research Letters, 52, e2025GL117 094, https://doi.org/10.1029/2025GL117094, 2025.

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Slättberg, N., Dahlke, S., and Maturilli, M.: Fram Strait Marine Cold Air Outbreaks in CARRA and ERA5: Effects on Surface Turbulent Heat Fluxes and the Vertical Structure of the Troposphere, Journal of Geophysical Research: Atmospheres, 130, e2024JD042 908, https://doi.org/10.1029/2024JD042908, 2025.

Spreen, G., Kwok, R., and Menemenlis, D.: Trends in Arctic sea ice drift and role of wind forcing: 1992–2009, Geophysical Research Letters, 38, L19 501, https://doi.org/10.1029/2011GL048970, 2011.

Srivastava, P., Brooks, I. M., Prytherch, J., Salisbury, D. J., Renfrew, I. A., Elvidge, A. D., and Yelland, M. J.: Turbulent Heat Fluxes over Arctic Sea Ice: Measurements and Evaluation of Recent Parameterizations, Boundary-Layer Meteorology, 190, 46, https://doi.org/10.1007/s10546-024-00887-5, 2024.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, Dordrecht, https://doi.org/10.1007/978-94-009-3027-8, 1988.

Taylor, P. C., Hegyi, B. M., Boeke, R. C., and Boisvert, L. N.: On the Increasing Importance of Air-Sea Exchanges in a Thawing Arctic: A Review, Atmosphere, 9, 41, https://doi.org/10.3390/atmos9020041, 2018.

Tetzlaff, A., Lüpkes, C., and Hartmann, J.: Aircraft-based observations of atmospheric boundary-layer modification over Arctic leads, Quarterly Journal of the Royal Meteorological Society, 141, 2839–2856, https://doi.org/10.1002/qj.2568, 2015.

Tian, T., Yang, S., Høyer, J. L., Nielsen-Englyst, P., and Singha, S.: Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis, Communications Earth & Environment, 5, 111, https://doi.org/10.1038/s43247-024-01276-z, 2024.

Uhlíková, T., Vihma, T., Karpechko, A. Y., and Uotila, P.: Effects of Arctic sea-ice concentration on turbulent surface fluxes in four atmospheric reanalyses, The Cryosphere, 18, 957–976, https://doi.org/10.5194/tc-18-957-2024, 2024.

Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjernström, M., Lüpkes, C., Nygård, T., Notz, D., Weiss, J., Marsan, D., Cheng, B., Birnbaum, G., Gerland, S., Chechin, D., and Gascard, J. C.: Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: a review, Atmospheric Chemistry and Physics, 14, 9403–9450, https://doi.org/10.5194/acp-14-9403-2014, 2014.

Walden, V. P., Hudson, S. R., Cohen, L., Murphy, S. Y., and Granskog, M. A.: Atmospheric components of the surface energy budget over young sea ice: Results from the N-ICE2015 campaign, Journal of Geophysical Research: Atmospheres, 122, 8427–8446, https://doi.org/10.1002/2016JD026091, 2017.

Wendisch, M., Brückner, M., Crewell, S., Ehrlich, A., Notholt, J., Lüpkes, C., Macke, A., Burrows, J. P., Rinke, A., Quaas, J., Maturilli, M., Schemann, V., Shupe, M. D., Akansu, E. F., Barrientos-Velasco, C., Bärfuss, K., Blechschmidt, A.-M., Block, K., Bougoudis, I., Bozem, H., Böckmann, C., Bracher, A., Bresson, H., Bretschneider, L., Buschmann, M., Chechin, D. G., Chylik, J., Dahlke, S., Deneke, H., Dethloff, K., Donth, T., Dorn, W., Dupuy, R., Ebell, K., Egerer, U., Engelmann, R., Eppers, O., Gerdes, R., Gierens, R., Gorodetskaya, I. V., Gottschalk, M., Griesche, H., Gryanik, V. M., Handorf, D., Harm-Altstädter, B., Hartmann, J., Hartmann, M., Heinold, B., Herber, A., Herrmann, H., Heygster, G., Höschel, I., Hofmann, Z., Hölemann, J., Hünerbein, A., Jafariserajehlou, S., Jäkel, E., Jacobi, C., Janout, M., Jansen, F., Jourdan, O., Jurányi, Z., Kalesse-Los, H., Kanzow, T., Käthner, R., Kliesch, L. L., Klingebiel, M., Knudsen, E. M., Kovács, T., Körtke, W., Krampe, D., Kretzschmar, J., Kreyling, D., Kulla, B., Kunkel, D., Lampert, A., Lauer, M., Lelli, L., Lerber, A. v., Linke, O., Löhnert, U., Lonardi, M., Losa, S. N., Losch, M., Maahn, M., Mech, M., Mei, L., Mertes, S., Metzner, E., Mewes, D., Michaelis, J., Mioche, G., Moser, M., Nakoudi, K., Neggers, R., Neuber, R., Nomokonova, T., Oelker, J., Papakonstantinou-Presvelou, I., Pätzold, F., Pefanis, V., Pohl, C., Pinxteren, M. v., Radovan, A., Rhein, M., Rex, M., Richter, A., Risse, N., Ritter, C., Rostosky, P., Rozanov, V. V., Donoso, E. R., Garfias, P. S., Salzmann, M., Schacht, J., Schäfer, M., Schneider, J., Schnierstein, N., Seifert, P., Seo, S., Siebert, H.,

https://doi.org/10.5194/egusphere-2026-6889
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



Soppa, M. A., Spreen, G., Stachlewska, I. S., Stapf, J., Stratmann, F., Tegen, I., Viceto, C., Voigt, C., Vountas, M., Walbröl, A., Walter, M., Wehner, B., Wex, H., Willmes, S., Zanatta, M., and Zeppenfeld, S.: Atmospheric and Surface Processes, and Feedback Mechanisms Determining Arctic Amplification: A Review of First Results and Prospects of the (AC)3 Project, Bulletin of the American Meteorological Society, 104, E208–E242, https://doi.org/10.1175/BAMS-D-21-0218.1, 2023.

Wesche, C., Steinhage, D., and Nixdorf, U.: Polar aircraft Polar5 and Polar6 operated by the Alfred Wegener Institute, Journal of large-scale research facilities JLSRF, 2, A87, https://doi.org/10.17815/jlsrf-2-153, 2016.

Yin, Z., Li, X., Ge, Y., Shang, C., Li, X., Du, Y., and Ling, F.: Estimating subpixel turbulent heat flux over leads from MODIS thermal infrared imagery with deep learning, The Cryosphere, 15, 2835–2856, https://doi.org/10.5194/tc-15-2835-2021, 2021.

Zhang, Z.-L., Hui, F.-M., Vihma, T., Granskog, M. A., Cheng, B., Chen, Z.-Q., and Cheng, X.: On the turbulent heat fluxes: A comparison among satellite-based estimates, atmospheric reanalyses, and *in-situ* observations during the winter climate over Arctic sea ice, Advances in Climate Change Research, 14, 347–362, https://doi.org/10.1016/j.accre.2023.04.004, 2023.

Zhang, Z.-L., Hui, F.-M., and Cheng, X.: Satellite observations reveal the impact of changing sea ice on wintertime surface turbulent fluxes over the Arctic Ocean, Advances in Climate Change Research, 16, 935–947, https://doi.org/10.1016/j.accre.2025.04.017, 2025.

Zilitinkevich, S. S. and Esau, I. N.: Similarity theory and calculation of turbulent fluxes at the surface for the stably stratified atmospheric boundary layer, Boundary-Layer Meteorology, 125, 193–205, https://doi.org/10.1007/s10546-007-9187-4, 2007.

Ólason, E., Rampal, P., and Dansereau, V.: On the statistical properties of sea-ice lead fraction and heat fluxes in the Arctic, The Cryosphere, 15, 1053–1064, https://doi.org/10.5194/tc-15-1053-2021, 2021.