Snowfall and snow accumulation processes during the MOSAiC winter and spring season

David N. Wagner[1,2], Matthew D. Shupe[3,4], Ola G. Persson[3,4], Taneil Uttal[3], Markus M. Frey[5], Amélie Kirchgaessner[5], Martin Schneebeli[1], Matthias Jaggi[1], Amy R. Macfarlane[1], Polona Itkin[6,7], Stefanie Arndt[8], Stefan Hendricks[8], Daniela Krampe[8], Robert Ricker[8], Julia Regnery[8], Nikolai Kolabutin[9], Egor Shimanshuck[9], Marc Oggier[10], Ian Raphael[11], and Michael Lehning[1,2]

[1]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[2]CRYOS, School of Architecture, Civil and Environmental Engineering, EPFL, Lausanne, Switzerland
[3]NOAA Physical Science Laboratory, Boulder, CO, USA
[4]Cooperative Institute for the Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA
[5]British Antarctic Survey - Natural Environment Research Council, Cambridge, UK
[6]University of Tromsø, Tromsø, Norway
[7]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA
[8]Alfred-Wegener Institute, Bremerhaven, Germany
[9]Arctic and Antarctic Research Institute, St. Petersburg, Russia
[10]University of Alaska Fairbanks, Fairbanks, AK, USA
[11]Thayer School of Engineering, Dartmouth College, Hanover, NH, USA

Correspondence: David N. Wagner (david.wagner@slf.ch)

Abstract. Data from the Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition allowed us to investigate the temporal dynamics of snowfall, snow accumulation and erosion in great detail for almost the whole accumulation season (November 2019 to May 2020). We computed cumulative snow water equivalent (SWE) over the sea ice based on snow depth (HS) and density retrievals from a SnowMicroPen (SMP) and approximately weekly-measured snow

5   depths along fixed transect paths. Hence, the computed SWE considers surface heterogeneities over an average path length of 1469 m. We used the SWE from the snow cover to compare with precipitation sensors installed during MOSAiC. The data was compared with ERA5 reanalysis snowfall rates for the drift track. Our study shows that the simple fitted HS-SWE function can well be used to compute SWE along a transect path based on SMP SWE retrievals and snow-depth measurements. We found a accumulated snow mass of 34 mm SWE until 26 April 2020. Further, we found that the Vaisala Present Weather

10 Detector 22 (PWD22), installed on a railing on the top deck of research vessel Polarstern was least affected by blowing snow and showed good agreements with SWE retrievals along the transect, however it also systematically underestimated snowfall. The OTT Pluvio[2] and the OTT Parsivel[2] were largely affected by wind and blowing snow, leading to higher measured precipitation rates, but when eliminating drifting snow periods, especially the OTT Pluvio[2] shows good agreements with ground measurements. A comparison with ERA5 snowfall data reveals a good timing of the snowfall events and good agreement with

15 ground measurements but also a tendency towards overestimation. Retrieved snowfall from the ship-based K_a-band ARM Zenith Radar (KAZR) shows good agreements with SWE of the snow cover and comparable differences as ERA5. Assuming the KAZR derived snowfall as an upper limit and PWD22 as a lower limit of a cumulative snowfall range, we estimate 72 to

107 mm measured between 31 October 2019 and 26 April 2020. For the same period, we estimate the precipitation mass loss
along the transect due to erosion and sublimation as between 53 and 68 %. Until 7 May 2020, we suggest a cumulative snowfall
20   of 98 - 114 mm.

1   Introduction

Snow cover on sea ice has many significant effects on the ice mass balance and general heat exchange processes between the
ocean and the atmosphere. As snow will cover almost all Arctic sea ice by the beginning of the melt season and with albedo
values close to 0.9, a large amount of the incoming solar radiation is reflected rather than absorbed into the snow pack. Due to
25   its potentially very high insulating capacity, snow acts as an inhibitor for heat transfer between ocean, sea ice, and atmosphere
(Holtsmark, 1955; Maykut and Untersteiner, 1971; Sturm et al., 2002). Depending on the season, accumulation, density, and
thermal conductivity of the snow, the sea ice growth and melt varies temporally and spatially. For instance, the underlying sea
ice might undergo faster (slower) growth in autumn when the snow on top is relatively thin (relatively thick). On the other hand,
a thicker (thinner) snow cover might lead to delayed (earlier) sea-ice melt in the melt season. Consequently, the small-scale
30   snow distribution might have a large effect on the ice mass balance as large amounts of snow are accumulated along ridges
or dunes, while large areas of level ice experience little snow accumulation (Lange and Eicken, 1991; Sturm et al., 1998b;
Iacozza and Barber, 1999; Leonard and Maksym, 2011; Trujillo et al., 2016). The snow that has fallen to the ground as fresh
precipitation often gets re-distributed as blowing or drifting snow due to the relatively high average horizontal wind velocities
during the Arctic winter. The high snow transport rates are also a result of the relatively low aerodynamic roughness length of
35   sea ice, where z_0 is typically lower for first-year ice (FYI) than for second- or multi-year ice (SYI/MYI) (Weiss et al., 2011). In
addition, large parts of the snow mass can be expected to get blown into leads and undergo sublimation (Déry and Yau, 2002;
Déry and Tremblay, 2004; Leonard and Maksym, 2011). Besides thermodynamic ice growth at its bottom, snow can directly
contribute to ice formation on top of the sea ice as snow-ice. Snow-ice formation occurs when snow first transforms into slush
due to surface flooding of saltwater or direct brine expulsion through thin ice followed by subsequent refreezing (Ackley et al.,
40   1990; Sturm et al., 1998a; Toyota et al., 2011; Jutras et al., 2016; Sturm and Massom, 2016). The relative mass contribution of
snow-ice towards sea ice by the end of the accumulation season depends strongly on location, with an approximated average
of 6 - 10 % for Arctic sea ice, and with estimated local peaks of up to 80 % (Merkouriadi et al., 2020). As a further term in the
snow mass balance, Webster et al. (2021) mentions sea ice dynamics. However, we can only imagine that the dynamics, such
as ridge formation, can lead to a snow mass decrease when the snow is pushed below the ice or into the water.
45   Considering all effects as snow mass source and mass sink, we can write the mass balance equation of snow over sea ice,
modified from the general mass balance description of snow (e.g. King et al. (2008)) as

$$\frac{dM}{dt} = P \pm E_s \pm E_e - E_D - R - I - \nabla \cdot D - L - S, \tag{1}$$

where $\frac{dM}{dt}$ is the rate of change of the mass of the snow cover over the sea ice at one point in kg m^{-2}, which is equivalent to
SWE per time, P is the snowfall rate, E_s is the sublimation rate and E_e is the evaporation rate of the snow cover, E_D is the

50 drifting and blowing snow particle sublimation rate, R is runoff, I is the snow-ice formation rate, D is the horizontal snow transport rate of blowing and drifting snow, L is the rate of the snow mass blown into leads and S is the mass of snow pushed or dug under the ice due to sea ice dynamics. Considering a larger area, all terms must be considered, while some terms may become zero when considering the equation at one point, e.g. where no open lead is existent at a point, L becomes zero.

55 The first and largest source term in Eq. 1 is, depending on the considered area, P. Quantitative comparisons between datasets are only possible for same compared point locations or same areas and therefore literature values are often not comparable. However, the Central Arctic has a dry climate and depending on location, a yearly average snowfall of approximately 100 to 350 mm can be expected in this area (Serreze and Hurst, 2000; Chung et al., 2011; McIlhattan et al., 2020; Webster et al., 2021). During polar night and its adjacent cold months, the mass decrease of the snow cover by sublimation (E_s) and evaporation

60 (E_e) can probably be assumed as negligible as well as the mass increase due to deposition (re-sublimation) and condensation (Webster et al., 2021). However, sublimation and evaporation terms become larger by beginning of summer in May and stay relatively large until September. Reliable values from literature are hard to determine, but the snow cover decrease as combination of E_s and D (as snow particles that get lifted into suspension) may be up to 50 % (Essery et al., 1999). To estimate the blowing snow sublimation E_D, Chung et al. (2011) applied the often and in various forms used sophisticated PIEKTUK

65 blowing snow model (e.g. Déry et al. (1998); Déry and Yau (1999, 2002); Déry and Tremblay (2004); Leonard et al. (2008); Leonard and Maksym (2011)) for a SHEBA (Surface Heat Budget of the Arctic Ocean) field experiment (Uttal et al., 2002) site, drifting between 74 ° and 81 ° N. They computed 12 mm of SWE blowing snow sublimation over a time period of 324 days between November 1997 to September 1998. As 179 mm of precipitation was found for the same time period, the blowing snow sublimation mass sink was 6 % of the total cumulative snowfall. Within the melt season in summer, R can be expected to

70 be the largest mass sink (Webster et al., 2021). Regarding mass decrease due to snow-ice formation, Merkouriadi et al. (2020) gives an average value of less than 0.05 m snow-ice thickness for the Central Arctic. It is hard to estimate from the 0.05 m a precipitated amount of SWE from recalculation, as the process of snow-ice formation is complex (Jutras et al., 2016). However, when we assume 0.05 m as snow height with an average fresh snow density of $100\,\mathrm{kg\,m^{-3}}$, we expect around 5 mm of SWE decrease, which would mean only about 3 %, relative to the measured 179 mm during SHEBA. The snow-ice formation rate

75 is expected to be highest in the months of September, October and November (Webster et al., 2021). On one hand, the largest sink term in Eq.1 is the erosion outside the melting season, represented as D in the mass balance equation, which may make up to 50 % SWE decrease over sea ice of the total precipitated snow mass (Leonard and Maksym, 2011). On the other hand, locally, the eroded mass may deposit at the windward and leeward side of ridges and fills frozen leads which can locally exceed the precipitated mass. The amount of drifting and blowing snow that is lost and gets melted in open leads L varies strongly

80 depending on location, considered area, ice dynamics and lead properties such as width and orientation relative to the wind. However, the total vanished mass flux from the column of blowing and drifting snow can make locally up to 100 % (Déry and Tremblay, 2004; Leonard and Maksym, 2011). Déry and Tremblay (2004) computed for a 10 km fetch, using the blowing snow model PIEKTUK with a mean lead width of 100 m, an open water fraction of 1 % and a typical lead trap efficiency of 80 %, an annual blowing snow loss of 20 mm SWE. However, in this model setup, saltation mass flux is not considered. Leonard et al.

85   (2008); Leonard and Maksym (2011) were doing computations with the same model-base for Antarctic sea ice, but considering saltation mass flux in addition. They emphasize the relative importance of saltation mass flux in the computation, as they find that all saltated mass flux blown towards an open lead vanishes there, and that although the mass flux within the saltation layer is lower than in the blowing snow column above, the higher frequency of saltation (about 50 % on 23 days in October 2007) compared against blowing snow frequency makes the mass loss due to saltation an important term. However, only a very

90   limited number of studies was carried out that investigate this specific problem and the saltation layer with relative large snow mass flux was not considered in great detail so far. Hence, the existing estimates go along with large uncertainties.

As we will only consider the accumulation time period, we can omit runoff R from Eq.1. Further, snow cover evaporation and sublimation terms are negligable during this time, hence we can neglect the terms E_s and E_e. Then we write the simplified

95   mass balance equation for winter and early spring as:

$$\frac{dM}{dt} = P \pm D - L - E_D - I - S. \tag{2}$$

To investigate all effects of the snow cover over the ice - the insulating effect, the sea ice mass-contribution effect, and the albedo effect - light must be shed into the snow processes that are represented and detailed knowledge of the evolution of total snow mass $\frac{dM}{dt}$, or SWE on top of the ice, is required. However, due to logistical challenges, especially for the winter

100   and spring months, snowfall rate and snow accumulation estimates could only be roughly approximated so far. The past estimates mostly made use of rare point measurements, or rather old time series (Petty et al., 2018) and satellite remote sensing (Petty et al., 2018; Cabaj et al., 2020), leading to high uncertainties in weather-, climate and snow cover models as well as in reanalyses. Batrak and Müller (2019), for instance, could show that a 5 to 10 °C warm bias of the sea ice surface temperature in weather forecasts and reanalyses is due to a missing snow layer modeled on top of the sea ice. For snowfall rates and

105   mass balance estimations, some general problems occur: Limited data about snowfall rates from precipitation gauges currently exist for this region. Snow buoys that measure snow height with acoustic sensors which record long continuous time series along its drift tracks throughout the Central Arctic do exist (Nicolaus et al., 2021). However, uncertainties with point snow measurements arise in those windy regions due to the snow transport processes described above. If using precipitation sensors, the high average horizontal wind velocities make snowfall rate estimates difficult for both weighing gauges (Goodison et al.,

110   1998) and optical sensors (Wong, 2012a). The wind itself may lead to an undercatch for weighing bucket gauges (Goodison et al., 1998), while blowing snow may lead to overestimation for both, weighing principle and optical sensors (Sugiura et al., 2003). Blowing snow typically occurs at heights up to ten meters, while it even can reach several hundreds of meters in altitude (Budd et al., 1966; Scarchilli et al., 2009). Hence, we expect that blowing snow can often falsely be detected as precipitation by snowfall sensors (Sugiura et al., 2003). Some issues caused by the wind can be corrected with scaling factors or transfer

115   functions, but these need to be identified for these specific conditions (Goodison et al., 1998). Another approach is to measure the snow water equivalent (SWE) of the snow cover. From this, one can derive snowfall rates. However, especially during the polar night, the precipitated snow is dry, and as already indicated above is the wind speed often sufficiently high to drift the freshly fallen snow particles away immediately. Hence, single point measurements are not appropriate to estimate snowfall and

horizontal sampling intervals should be kept as short as possible (Trujillo and Lehning, 2015). This becomes more crucial the
120   more windy the location.

During the year-long Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition, during
which the research vessel (RV) Polarstern (Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, 2017)
served as a base moored on two different ice floes, data of snow on the ice as well as of in-situ snowfall was collected in great
125   detail for almost the whole MOSAiC period (October 2019 - October 2020). The dataset includes measurements of the penetra-
tion force into the snowpack with a SnowMicroPen (SMP) (Schneebeli and Johnson, 1998; Schneebeli et al., 1999) from which
snowpack densities can be estimated (Proksch et al., 2015), bulk SWE measurements, weekly repeated transects of snow depth
measurements, and a set of precipitation sensors installed on the ice (Vaisala Present Weather Detector 22 (PWD22) (Vaisala,
2004; Kyrouac and Holdridge, 2019), OTT Pluvio^2 pluviometer (Bartholomew, 2020a; Wang et al., 2019b), OTT Parsivel^2
130   (Bartholomew, 2020b; Shi, 2019), and on board RV Polarstern (Vaisala PWD22, OTT Parsivel^2).

This paper investigates the snow accumulation period from October 2019 to May 2020, where precipitation is solid, and no
significant snowmelt was observed. For this period, the intentions in this paper are as follows:

   - Compute reliable values for SWE evolution along the fixed transect paths that include surface heterogeneities

135   - Use the computed SWE for periods where no drifting snow occurred to compare with snowfall rates from precipitation
     gauges installed during the MOSAiC expedition

   - Evaluate an existing radar reflectivity - snowfall (Z_e-S) relationship (Matrosov, 2007; Matrosov et al., 2008) for the
     ship-based K_a-Band ARM Zenith Radar (KAZR)

   - Evaluate the ERA5 (Hersbach et al., 2020) mean snowfall rates for the MOSAiC drift track

140   - Investigate processes that lead to discrepancies of computed snowfall rates and SWE on the sea ice and shed light into
     the processes described in Eq.2, such as total eroded mass

Section 2 introduces our methods, followed by section 3, where we show the results. In section 4 we discuss our results, and
in section 5 we draw conclusions about our findings and give an outlook about potential future work.

2   Data and methodology

145   All data used for evaluations in the following was collected during the MOSAiC campaign (Krumpen et al., 2020) from the
beginning of leg 1 (24 October 2019) until the end of leg 3 (7 May 2020) (Fig. 1). On 4 October 2019, RV Polarstern moored
along an ice floe that originated in the Siberian shelf (Krumpen et al., 2020).

2.1 Ice conditions and Central Observatory

According to Krumpen et al. (2020), the floe where RV Polarstern moored on had a size of approximately 2.8 km x 3.8 km and
150 was a loose assembly of pack ice, little less than a year old that had survived the 2019 summer melt. Fig. 2 shows a map of
the ice- and snow surface structures and installations by 5 March 2020 of the MOSAiC Central Observatory (CO). Note that
the shown elevation range is only approximate as problems occurred with the inertial navigation system of the laser scanner.
This lead to tilts and the single swaths within the map have staggered heights. To this date, these uncertainties could not be
corrected. However, very bright areas indicate ridges of around 2 m height, with locally 3 m height and more. The CO (all
155 installations in the close vicinity of Polarstern) was distinguished from the Distributed Network, which consisted of remote
autonomous stations at least a few kilometers away from the CO. The core of the floe consisted mostly of deformed second-
year ice (SYI), and the ice surrounding this core mainly consisted of frozen melt ponds (remnant SYI) and partially first-year
ice (FYI). When the ship moored, the heading of RV Polarstern was about 220° in October 2019. Significant changes in ice
conditions occurred the first time around 16 November 2019, when a storm led to strong ice deformations in and around the
160 CO. Another significant ice deformation event occurred around 11/12 March 2020 and the days thereafter. Over time, the floe
rotated counter-clockwise and reached a minimum heading of 75° on 21 March 2020.

We describe the measuring setup and the post-processing for all used data streams in the following.

2.2 Snow cover measurements

2.2.1 SMP force and SWE measurements

165 We measured snow-water equivalent (SWE) with an ETH tube (Haberkorn, 2019), as well as resistance force with the SnowMi-
croPen (SMP) (Schneebeli and Johnson, 1998; Schneebeli et al., 1999) and snow height at different sites (areas shaded in yellow
in Fig. 2). The direct SWE measurements follow the simple principle where the mass of the snow fitting in a tube with a known
cross-sectional area is weighed, which yields the SWE in mm, or $kg\,m^{-2}$. The SMP is a device which measures the penetration
resistance force (N) by means of a rod with a conic tip that is slowly driving vertically into the snowpack. A force sensor is
170 connected to the tip which detects the force that is needed to drive into the snowpack with μm resolution. The output is a
force-snow depth signal. These penetration resistance force signals can be used to estimate snowpack density and detect the
layers in the snowpack (Proksch et al., 2015; King et al., 2020).

The map in Fig. 2 shows the floe state on 5 March 2020, which changed significantly due to ice dynamics that started on
175 11 March 2020. Snow was measured at the different sites as well as along both transect loops. The measurements cover a
large area of the floe, including level, remnant SYI, FYI, and deformed SYI. Details about the used instruments will follow
below. Snow 1 was characterized mainly by a mixture of remnant SYI and deformed SYI. In the beginning, Snow 1 was mostly
flat, but the surface became rougher over the time of the expedition. At Snow 1 we deployed three snow pit sites, which were
maintained until the end of leg 3. The Snow 2 plot was characterized as an open level field, mostly on remnant and deformed
180 SYI with a distinct, high, and long pressure ridge in the center of the plot. On Snow 2, we maintained two snowpit locations
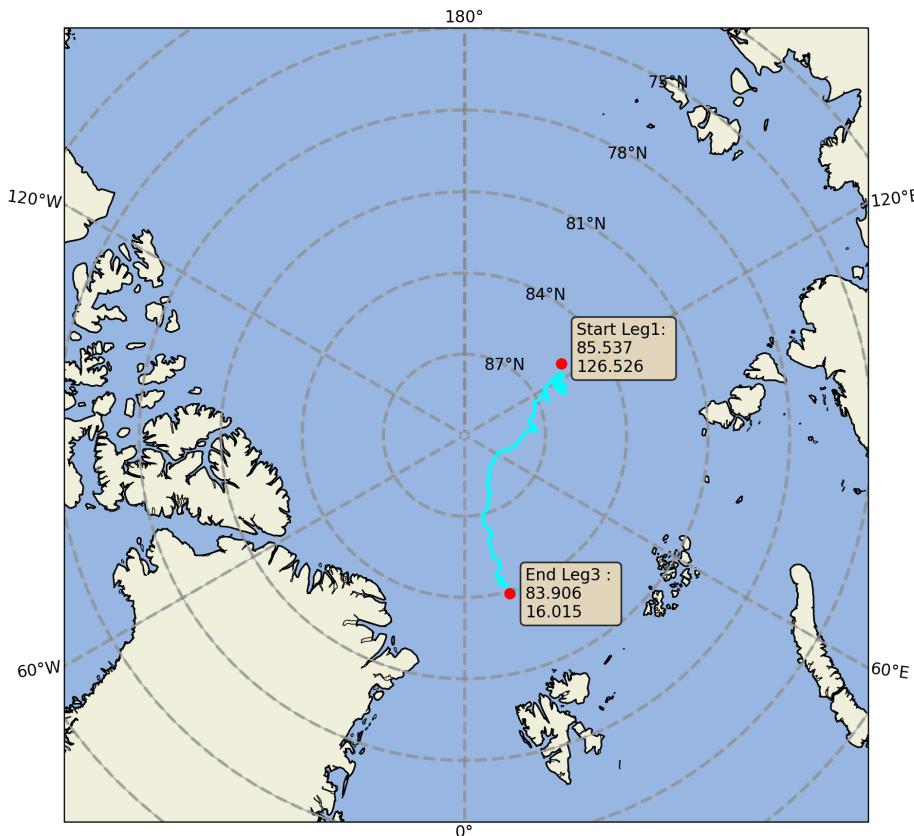
Figure 1. The drift trajectory of RV Polarstern between start of leg 1 (24 October 2019) and end of leg 3 (7 May 2020).

until the end of leg 3. Both sites had very similar underlying ice conditions. Snow 3 was created at a later point, furthest away from the vessel. In the beginning, it was a very flat area with underlying FYI and was maintained during leg 3 but needed to be abandoned due to ice dynamics in mid-March. Further, weekly snow pit measurements were conducted along the south-westerly section of the northern transect loop. Also, transects were conducted infrequently over ridges, and measurements were

185   conducted weekly at the ice coring sites during leg 1 (beyond the map boundaries in Fig. 2, but located north-west of the ship), among other measuring locations. The large variety of locations, their underlying ice types, and snow depths allow us to take the spatial heterogeneities of the snow cover into account. However, since we use a bulk approach with the collected SMP and direct SWE data, detailed information on each measuring site is not needed and will not be provided here.

190   At the measuring locations, SWE, snow height, and penetration resistance force measurements with the SMP were done. The SMP measurements at the recurring snowpit locations were conducted as follows: Five SMP measurements were performed at a distance of about 20 cm along a line parallel to the old snowpit wall to account for the spatial heterogeneity of the
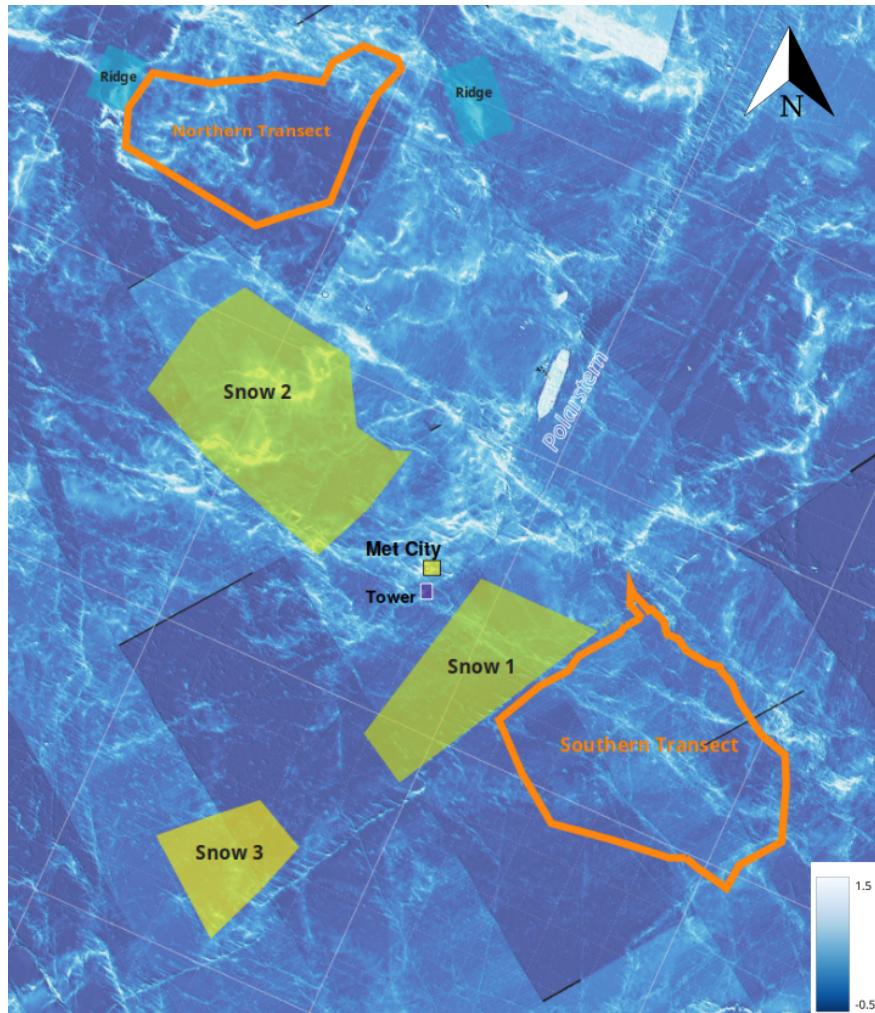
Figure 2. Main snow measuring areas on the MOSAiC floe by 5 March 2020. The bottom layer is a digital elevation map (DEM) from airborne laser scanning (ALS) with the helicopter. The square side length of the underlay grid is 500 m. The transect paths and margins of the shaded measuring areas are based on GPS measurements. The legend for elevation is shown in meters. However, the elevation range is only approximate due to issues with the inertial navigation system which could not be corrected so far.

snowpack. On the ridge sites, for instance, SMP measurements were conducted infrequently as transect over ridges. We used these measurements to estimate SWE along the northern and southern transect loops, which will be explained below.

2.2.2 Transects

Snow depth transects were conducted weekly with a Magnaprobe (Sturm and Holmgren, 2018), if the atmospheric, ice, or overall safety conditions did not prevent it. The transect path was distributed into two loops (Fig. 2):
A northern loop, mostly situated on deformed SYI, and a southern loop, which was mainly situated on FYI and remnant SYI with underlying frozen melt ponds. A transition zone distinguished these two loops, mostly consisting of frozen melt ponds with a very flat surface without significant heterogeneities. The approximate ice conditions and the transect loop locations can be seen in orange on the MOSAiC floe map from 5 March 2020 (Fig. 2). The elevations on the southern transect are mostly below around 1 m height. The elevations are generally higher on the northern transect, although it does not cover ridges of up to 3 m height or more as they have been observed on the Snow 2 plot. The GPS coordinates of the Magnaprobe were transformed into coordinates of a local metric coordinate system, called "FloeNavi" (Fig. 3). Furthermore, the transects were partially corrected for shifts within the ice, which was especially the case for an event with strong ice dynamics on 16 November 2019. Thus, the southern loop transects until (including) 14 November 2019 are marked as a yellow rectangle in Fig. 3. Though the transects may deviate from one another within the FloeNavi coordinate system, the actual transect path was the same. After the coordinate transformation and horizontal correction, for good spatial and temporal comparability, clear margins as shown in the rectangles in Fig. 3 were defined for the "southern" and "northern" transect loop. By the overlays, one can recognize that the transect loops were not significantly impacted by internal differential ice movements.

Ice dynamics affected the transects especially from 11/12 March 2020 on, where leads and cracks opened throughout the paths. Overall, we tried to minimize the influence of these ice deformation events on the transect measurements. However, an impact on the time series cannot be excluded. On the transects, snow height measurements were sampled with the Magnaprobe with an average distance between measuring points of 1.21 m for the northern loop and 1.03 m for the southern loop. Values $z < 0.00\,m$ were discarded as incorrect data. From the measured Magnaprobe snow heights, 0.01 m was subtracted from each measurement to correct for the tip that typically sinks into the ice to this amount. Outliers were detected with a z score, and all outliers with a z score above 3 were discarded.

The northern loop was sampled from 24 October 2019 to 7 May 2020 on 24 days with an average path length of 954 m. The southern loop was sampled from 31 October 2019 to 26 April 2020 14 times with an average transect path length of 974 m.

To take surface roughness and potential snow accumulation at surface irregularities better into account, we looked at the weekly snow height differences of the transects. With the given average horizontal sampling distance (1.03 - 1.21 m), no small-scale patterns are considered for evaluation. However, since the extent of ridges and most types of dunes are in all horizontal
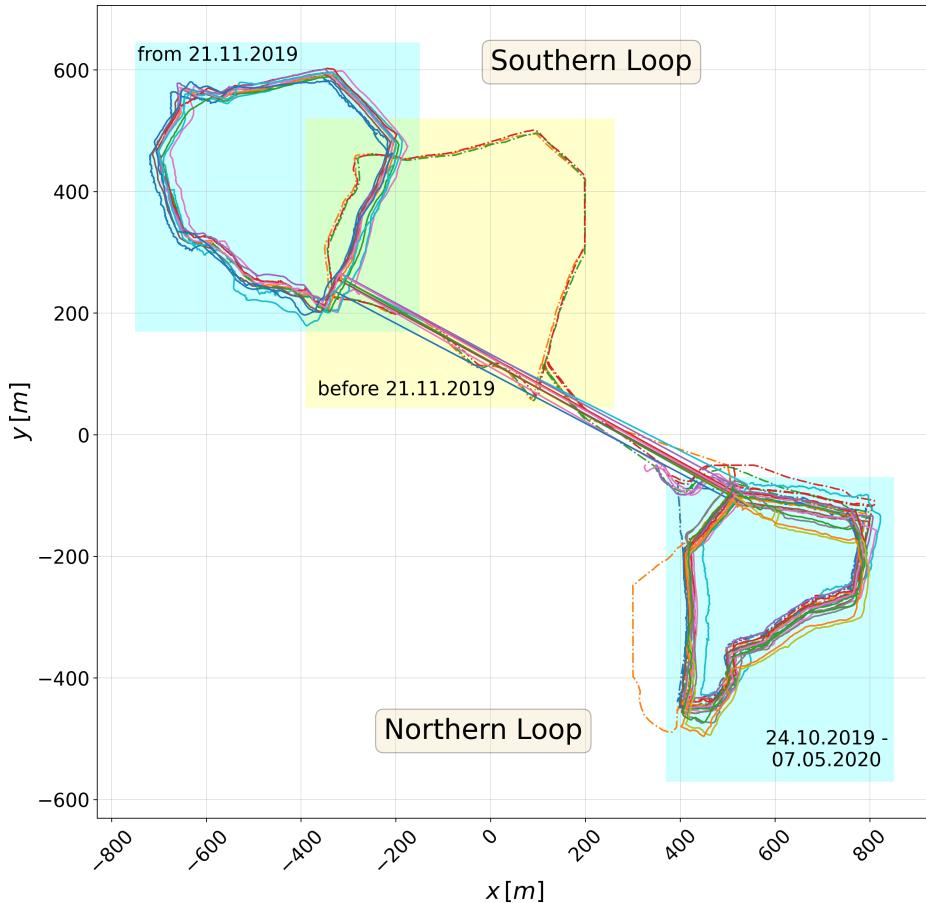
Figure 3. All used Magnaprobe transect paths (leg 1 - leg 3) with coordinates transformed to the FloeNavi grid corrected for ice drift. The rectangles represent the margins that were used as a definition for "Northern transect loop" (upper left) and "Southern transect loop" (bottom right) for good comparability. The shifts between transect paths within a rectangle origin from corrections and coordinate transformation, though the actual transect paths were the same.

directions larger than $1.2\,\mathrm{m}$ (Filhol and Sturm, 2015), we expect our typical horizontal sampling scale to accurately characterize the spatial distribution of accumulation, which we demonstrate in Sec. 2.3.1.

2.3 SMP density retrievals and SWE from transect snow depths

Not many direct SWE measurements or SMP force measurements are available along the transect path. We use the direct SWE measurements for validation but apply a statistical SWE - snow depth (HS) relationship to estimate SWE along the full path (Jonas et al., 2009). Due to its fast and simple usage even during harsh conditions in the polar night, we have considerably more SMP force measurements available ($N = 3007$) than direct SWE bulk measurements ($N = 195$). Furthermore, as we made for each snowpit at least $n = 5$ SMP measurements and the SMP was often used even for ridge transects, these measurements best

characterize the spatial heterogeneity in the snow depth across the sea ice.

235

Snowpack density can be retrieved with a statistical model from SMP snow depth-force signal profiles (Proksch et al., 2015):

$$\rho = a + b\ln(F) + c\ln(F)L + dL \tag{3}$$

where $a\,(kg\,m^{-3})$, $b\,(N^{-1})$, $c\,(N^{-1}\,mm^{-1})$, $d\,(mm^{-1})$ are empirical regression coefficients, F is the penetration force
240 of the SMP (N) and L is the microstructural length scale (Löwe and Herwijnen, 2012). King et al. (2020) calibrated the
corresponding coefficients to snow on Arctic sea ice and found $a = 315.61\,kg\,m^{-3}$, $b = 46.94\,N^{-1}$, $c = 43.94\,N^{-1}\,mm^{-1}$,
$d = 88.15\,mm^{-1}$. The coefficients show a significant improvement in density derivation for snow on sea ice, which is reflected
by the decrease in the root-mean-square error (RMSE) (Proksch et al. (2015): $RMSE = 130\,kg\,m^{-3}$, King et al. (2020):
$RMSE = 41\,kg\,m^{-3}$). Consequently, we used the coefficients from King et al. (2020) for the following SWE computations.
245 From the SMP density estimates we can compute

$$SWE = HS \cdot \overline{\rho} \tag{4}$$

where $HS\,(m)$ is the height of snow over the ice or snow depth, and $\overline{\rho}\,(kg\,m^{-3})$ is the vertically averaged density of the
snowpack. The computed SWE dataset is documented in detail by Wagner et al. (2021). Similar to Jonas et al. (2009), but
applying the function directly to SWE, we fitted the following function to the available bulk SWE measurements as well as
250 SWE retrievals from the SMP:

$$SWE = m \cdot HS^a + b \tag{5}$$

where m is the fitted slope, b is the fitted y-intercept, and a is a fitting coefficient. For the bulk SWE measurements, we found
$m = 333.82$, $a = 1.12$ and $b = +0.92$ (Fig. 4). For the SMP retrievals we found $m = 323.91$, $a = 1.1$ and $b = 2.31$. The func-
tions do not deviate heavily from another, but the y-intercept is obviously shifted downwards in the bulk SWE measurements.
255 Thus, as we have more measurements available for deep snow depths with the SMP but believe that the y-intercept for the
SWE bulk measurements is more reliable, we combined the functions and computed SWE as:

$$SWE = 323.91 \cdot HS^{1.1} + 0.92. \tag{6}$$

From Fig. 4 one can clearly see that the improvement for snow on sea ice of the coefficients found by King et al. (2020) is
valid for MOSAiC leg 1-3 SMP data, too (Fig. 4c) and that the coefficients determined by Proksch et al. (2015) and Calonne
260 et al. (2020) are not appropriate to estimate SWE of snow during MOSAiC leg 1 - 3. Furthermore, the lowest RMSE was found
for the fitted model with the coefficients from King et al. (2020) (7.2 mm SWE) compared against 15.4 mm (Proksch et al.,
2015) and 9.4 mm (Calonne et al., 2020).
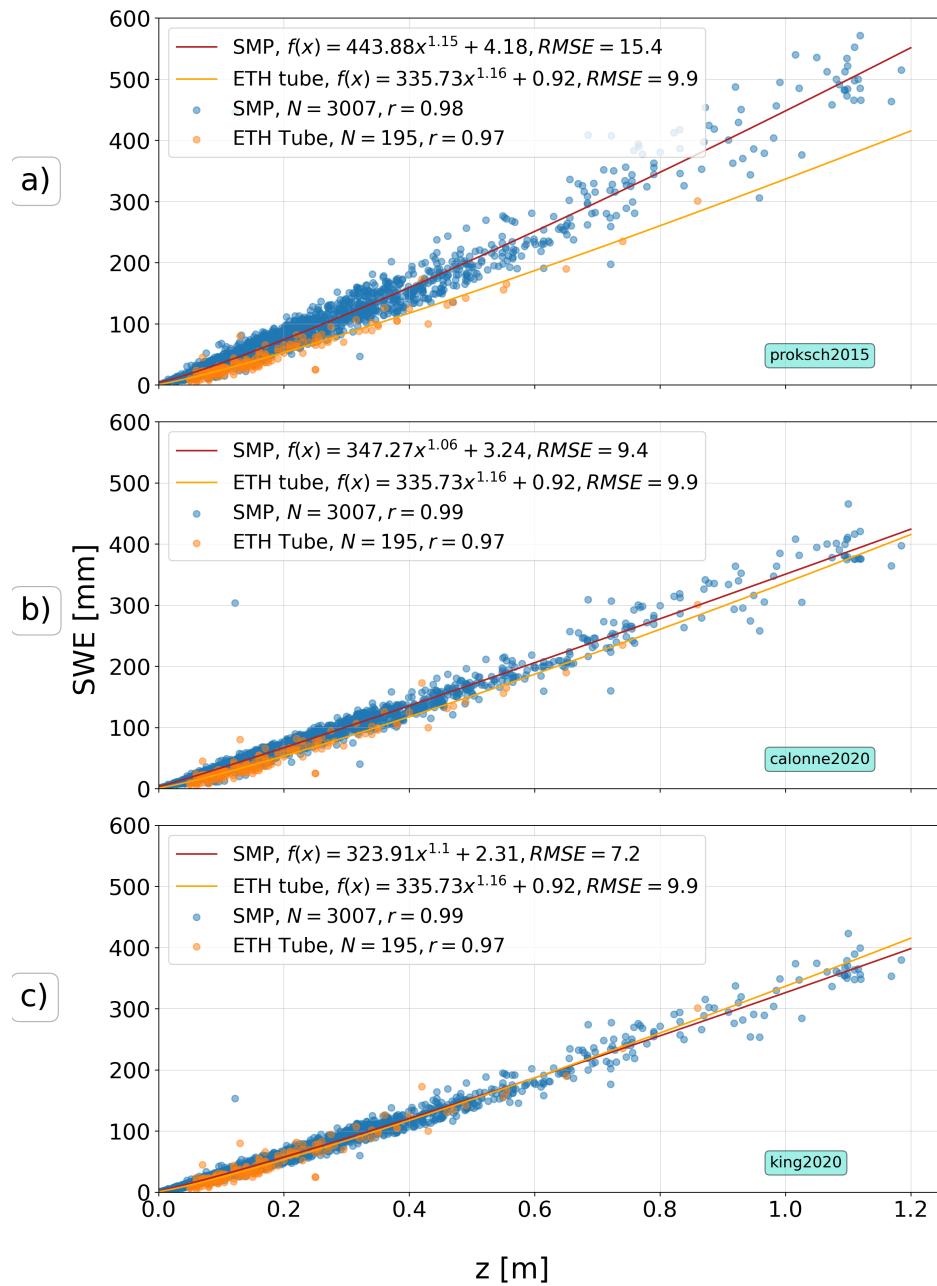
Figure 4. Scatter plots and fitted HS-SWE function of SMP derived SWE and measured SWE with ETH tube for a) The density computation coefficents from Proksch et al. (2015), b) Calonne et al. (2020) and c) King et al. (2020).

We applied this formula to each snow depth measurement with the Magnaprobe along the transect path to obtain the SWE estimates. A limitation with this approach is that different snow layers are not distinguished by density, even though a wind-

265   packed layer has a higher density than a depth hoar layer. Hence, when high winds lead to drifting snow deposition that is
detected by a snow height increase with the Magnaprobe, the SWE increase is likely to be underestimated, as would be the
eroded mass of a drifting snow layer. It is beyond the scope of this study, and indeed beyond the measurements themselves,
to attempt an approach that distinguishes different snow layers. Instead, for validation, we compared SMP and ETH tube
measurements with transect-computed SWE along a section of the northern transect loop for 14, 21, and 28 November 2019

270   (Fig. 5). The validation measurements were conducted at different positions at each day of measurements. Note that a clear
quantitative comparison is difficult, as the accuracy of GPS measurements (2 m) and the following coordinate transformation
do not allow for cm-scale precision. Thus, for quantitative comparison, SWE computations from direct bulk SWE and SMP
measurements along the transects were plotted over SWE model retrievals. Fig. 5a shows the measuring locations for each
SMP measurement along the northern transect loop (5 measurements at each snowpit location) and Fig. 5b, Fig. 5c and Fig. 5d

275   show the corresponding SWE plotted over the x-axis of the FloeNavi for different days of measurements.
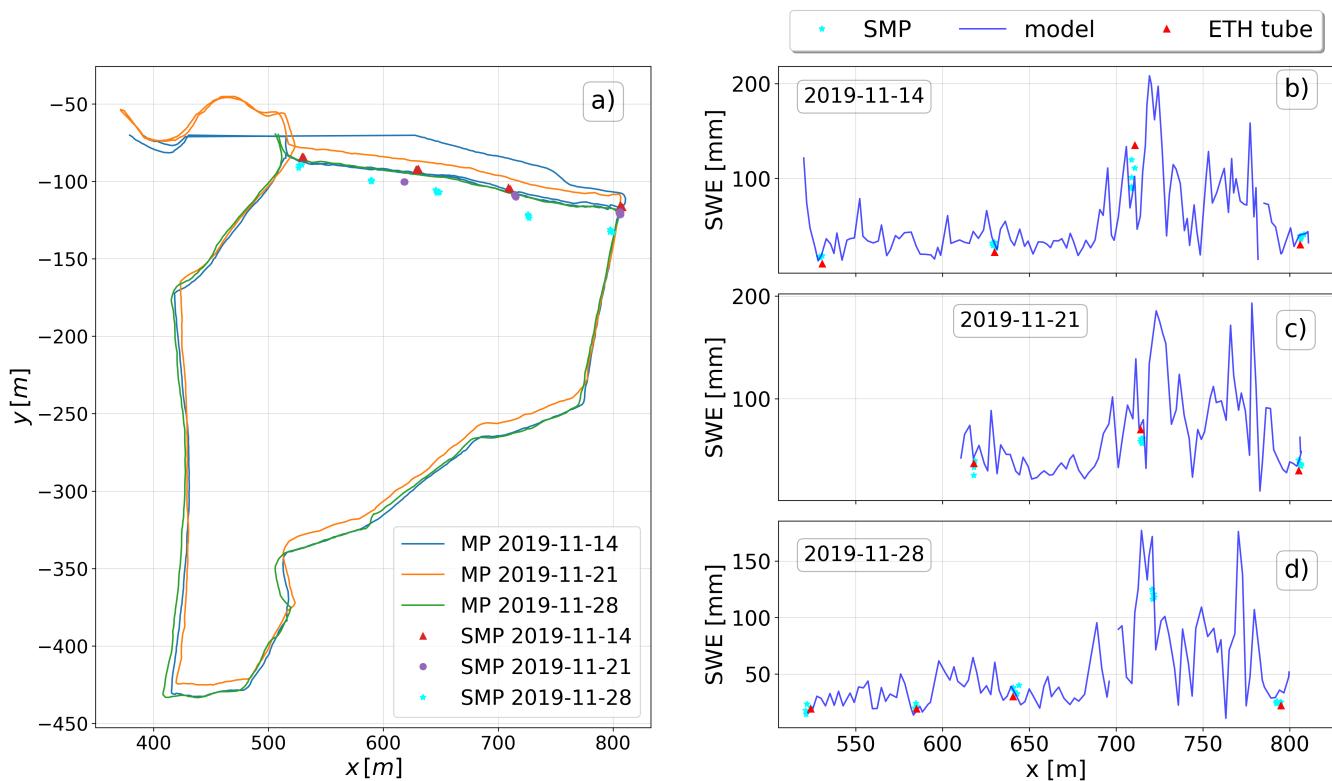
Figure 5. a) SMP measurement locations along the Magnaprobe (MP) transect path on 14, 21, and 28 November 2019. The GPS coordinates
were transformed into local FloeNavi grid coordinates. b), c), and d) show the comparison of SWE estimates from direct SMP measurements,
direct bulk SWE measurements, and SWE derived with the HS-SWE model from the Magnaprobe snow depth measurements along the
northern transect as x-axis location on the FloeNavi grid, for 14, 21 and 28 November 2019.

This comparison shows that the modeled SWE matches the derived SWE from SMP retrievals and bulk SWE measurements quite well during the three chosen time periods, even for higher SWE estimates, where a higher scatter is expected (Fig. 4). The direct SWE measurement with the ETH tube shows a tendency towards lower SWE values. However, we must note that the depth of SWE measurements from ETH tube and SMP have some individual but differing restrictions: Firstly, as the SMP cut-
280 off force signal was set to 40 - 41 N (depending on the device), the snow depth was determined whenever one of those values was reached, which is not necessarily the snow - ice interface. Secondly, during the sampling period, there was no method evolved to distinguish between surface scattering layer (SSL) and snow. Hence, its vertical position was determined visually, which was not always clear. Therefore, a measurement with the ETH tube might or might not include the surface scattering layer which formed during the melt season of 2019. If the SMP was able to penetrate the SSL only partially while it was
285 not measured with the ETH tube, then SWE is overestimated from the SMP measurements. Otherwise, if the SMP could not penetrate the SSL while it was partially measured with the ETH tube, the SMP-based SWE computation overestimates actual SWE. However, as the SMP-SWE retrievals are often close to the direct SWE measurements, one can assume on the whole reliable values. Research to determine exact boundaries between snow and sea ice is ongoing. Furthermore, since the number of measurement points along a transect is large and we do not expect systematic biases, we believe that fluctuations caused
290 by these various sources of uncertainty will largely average out, such that the results from the applied SWE model yield a reasonable estimate along the transect.

2.3.1 Evaluating the sensitivity of the arithmetic mean with respect to horizontal sampling distance

We studied the sensitivity of the horizontal sampling interval for average mass estimates by reducing the sample numbers from using all samples (average sampling distance of 1.1 m) down to considering every 10th sample (about 11 m sample distance)
295 for the average. The process was conducted for each day of sampling and the averages were normalized against the original sampling frequency (Fig. 6).

The results show that for sampling frequencies down to 1/3 of the original frequency (sampling distances ranging from 1.1 to 3.4 m), the average mass estimates vary by less than $\pm$ 1 %. This indicates that a sampling distance up to 3.4 m is mostly robust and that no significant undersampling occurred. This also shows that the impact of variations in sampling interval distance
300 that inevitably occurs with different operators of the Magnaprobe is probably negligible. The larger fluctuations in computed average mass for longer sample interval distances suggests undersampling at those scales and less reliable averages. However, a validation of uncertainties that could accompany varying vertical penetration force leading to different measured snow height, e.g. when a crust within the snow is penetrated or not due to varying operators, is not conducted here. The operators were aware of this issue and tried to apply a similar power for the Magnaprobe sampling.
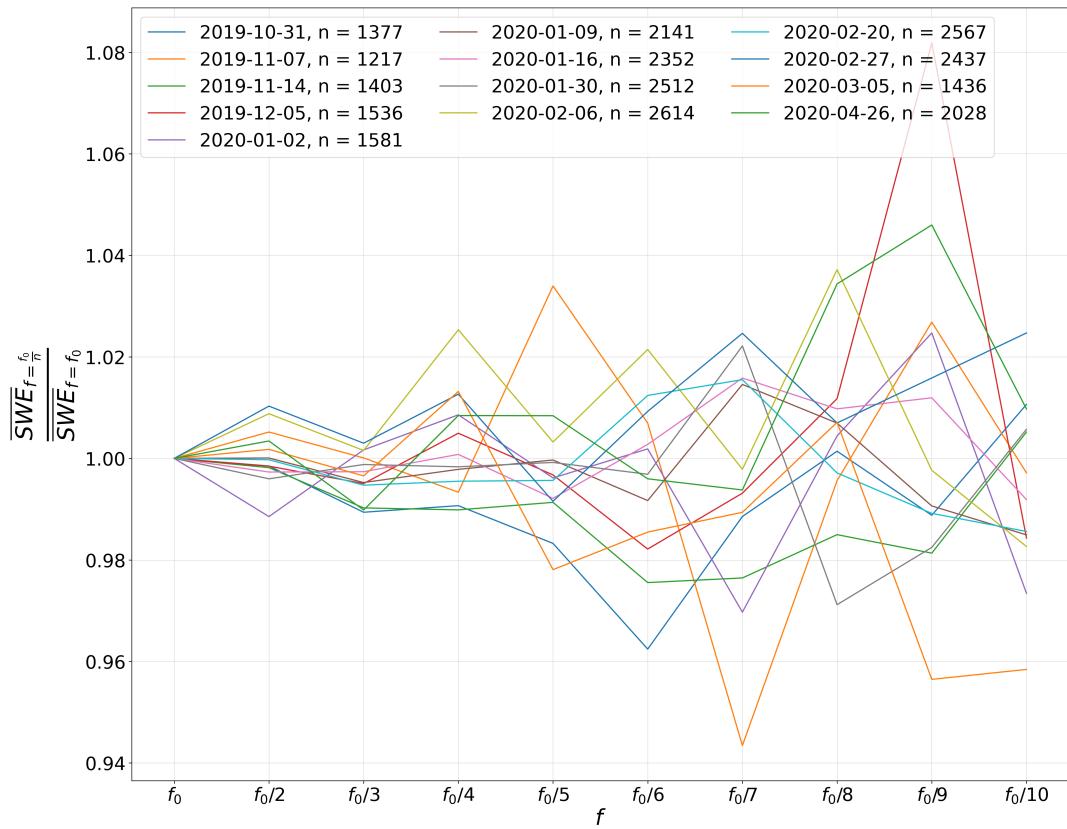
Figure 6. Sensitivity of the transect average in dependence of the Magnaprobe horizontal sampling interval. The x-axis shows the horizontal transect sampling frequency in relation to the original sampling frequency f_0, the y-axis shows the ratio of the average SWE of all tested frequencies to the average SWE of the original sampling frequency for each day of sampling.

2.4 Snowfall rates

2.4.1 Precipitation gauges

Snowfall rates were estimated using standard internal processing software from five distinct precipitation gauges operated by the US Department of Energy Atmospheric Radiation Measurement (ARM) Program. Two sensors investigated here were installed on the railing on the top deck of Polarstern - a Vaisala Present Weather Detector 22 (Vaisala, 2004; Kyrouac and Holdridge, 2019) (in the following referred to as PWD22PS) and an OTT Parsivel[2] laser disdrometer (Shi, 2019; Bartholomew, 2020b) (in the following referred to as P2PS) (Tab. 1). The PWD22PS was installed at 22 m and the P2PS at 24 m above the water line. On the ice, in "Met City" (Fig. 2, in the following, referred to as MC), three precipitation sensors were installed: 1) an OTT Parsivel[2] (P2MC), installed at 1.5 m nominal height above the snow surface, surrounded by a double-alter shield, 2) a PWD22 (PWD22MC), installed at 2 m nominal height, unshielded and 3) an OTT Pluvio[2] L (Wang et al., 2019b; Bartholomew,

2020b), shielded by a double-alter shield and installed at 1 m above the snow (in the following referred to as Pluvio2). Different ARM data levels of the devices are given, where a1 means "calibration factors applied and converted to geophysical units" and b1: "QC checks applied to measurements".

Optical devices evaluated here are the Vaisala PWD22 and the OTT Parsivel[2]. However, the measurement technique and the process of estimating snowfall rates are different. The Parsivel[2] is a laser-disdrometer that processes the voltage signal changes due to light extinction when a hydrometeor falls through the laser-beam. It has an effective measuring area of $54\,\mathrm{cm}^2$ to estimate hydrometeor size and velocity (Löffler-Mang and Joss, 2000). The hydrometeors are classified into size classes which can be used to investigate the particle size distribution. The precipitation type is determined by device-internal spectral signature comparison, where the spectra are determined empirically. Based on particle size, velocity, and estimated precipitation type, device-internal software computes a snowfall estimate. No details are known about the exact formula used by the manufacturer for the snowfall estimate. Its accuracy is given by the manufacturer as $\pm\,20\,\%$ with an intensity range of 0.001 to $1200\,\mathrm{mm\,h}^{-1}$ (Tab. 1). The calibration was conducted in the manufacturer's laboratory and therefore no calibration was needed in the field.

The PWD22 consists of several sensors that are used to compute the snowfall rate: The two core sensors are a transmitter-receiver combination, where the transmitter emits pulses of near-infrared (NIR) light. The receiver on the other side measures the scattered part at $45°$ of the light beam from the emitted signal (sampling volume $100\,\mathrm{cm}^3$). Rapid changes in the scatter signal between transmitter and receiver are used to compute precipitation intensity. The sampling volume allows for the detection of single crystals and aggregates of snow crystals (snowflakes). Furthermore, the PWD22 is equipped with a heated RAINCAP rain sensor, which produces a signal proportional to the amount of water on the sensing element. By means of the ratio from sample volume and water content determined with the RAINCAP sensor, precipitation types are distinguished. In the tube between the transmitter and receiver, another temperature sensor (thermistor) is installed. The detected temperature is used to select the default precipitation type. When frozen precipitation is detected, the PWD22 software multiplies optical intensity with a scaling factor, determined from RAINCAP and optical intensities from the receiver to estimate snowfall intensity as SWE per time unit (Vaisala, 2004). The manufacturer does not provide a value for accuracy, however the intensity measuring range is given as 0.00 to $999\,\mathrm{mm\,h}^{-1}$. There is no calibration principle known for the field but the manufacturer mentions comparisons with closeby reference gauges as calibration method.

The only device that we compare here that uses a weighing principle is the OTT Pluvio[2]. The instrument's core is a sealed load cell that continuously measures the weight of the precipitation falling into the entry of the bucket. The installed variant was an OTT Pluvio[2] L Version 400, with a collecting area of $400\,\mathrm{cm}^2$ and a recording capacity of 750 mm of precipitation. Its accuracy is given by the manufacturer as $\pm\,0.1\,\mathrm{mm\,min}^{-1}$ or $\pm\,6\,\mathrm{mm\,h}^{-1}$, or $\pm\,1\,\%$ and its intensity range as $\pm\,6\,\mathrm{mm\,h}^{-1}$ or 0.1 to $30\,\mathrm{mm\,h}^{-1}$. No calibration for the OTT Pluvio[2] is needed in the field as it was delivered calibrated by the manufacturer. However, calibration weights were used to test for accuracy.

Table 1. Summary of validated installed precipitation sensors and radar during MOSAiC as well details about the ERA5 reanalysis (PS: RV Polarstern, MC: Met City, DA: double alter shield, WL: Water line, DFIR: Double Fence Intercomparison Reference).

Device /Reanalyis	Loc	Abbreviation	Nominal height Snow/WL (m)	Shield	Data reference	ARM data level	Accuracy	Intensity range	Selected existing snowfall validations / Reference setup	Calibration
Vaisala PWD22	PS	PWD22PS	22 (WL)	-	Kyrouac and Holdridge (2019)	b1	-	0.00 - 999 mm/h	-32 % from median / DFIR (Wong, 2012a) +33 % from median / manual (Boudala et al., 2016)	Reference comparison
OTT Parsivel[2]	PS	P2PS	24 (WL)	-	Shi (2019)	b1	± 20 %	0.001 - 1200 mm/h	+46 / +54 % from median / DFIR (Wong, 2012a) +24 / +29 % from median / DFIR (Wong et al., 2012b)	manufacturer
KAZR	PS	KAZR	14 (WL)	-	Lindenmaier et al. (2019)	a1	-	-	23 % relative bias / shielded Nipher gauge (Matrosov et al., 2008)	-
Vaisala PWD22	MC	PWD22MC	2 (Snow)	-	Kyrouac and Holdridge (2019)	b1	-	0.00 - 999 mm/h	see above	Reference comparison
OTT Parsivel[2]	MC	P2MC	1.5 (Snow)	DA	Shi (2019)	b1	± 20 %	0.001 - 1200 mm/h	see above	manufacturer
OTT Pluvio[2]	MC	Pluvio2	1 (Snow)	DA	Wang et al. (2019b)	a1	± 6 mm/h ± 0.1 mm/min ± 1 %	6 - 1800 mm/h	-63 % from median (max) / DFIR -14 % from median (min) / DFIR (Wong, 2012a)	manufacturer / calib. weights
ERA5	-	-	-	-	Hersbach et al. (2020)	-	-	-	+ 62.8 mm cum. SWE / snow buoys (Wang et al., 2019a) slight underestimate in summer / CloudSat (Cabaj et al., 2020) overestimate in other months	-

The data streams were downloaded from the ARM data archive (https://adc.arm.gov/) and scanned for quality control flags.
350 Values with timestamps that correspond to flags indicating maintenance time, suspicious or incorrect values were discarded.

2.4.2 Snowfall retrievals from the Ka-band ARM Zenith Radar.

Snowfall was retrieved from the K_a-Band ARM Zenith Radar (KAZR) (Widener et al., 2012; Lindenmaier et al., 2019) that was installed on a container at the bow of RV Polarstern. Using a radar snowfall retrieval allows investigating snowfall continuously, and eliminates impacts on gauges such as acceleration effects of wind that result in undercatch or overestimation due to blowing
355 snow particles. The KAZR operated at approximately 35 GHz. We computed the snowfall rate S $[mm\,h^{-1}]$ according to the power law

$$Z_e = a\,S^b \qquad (7)$$

where $Z_e\,[mm^6\,m^{-3}]$ is the radar equivalent reflectivity factor and a and b are empirical coefficents. We chose $a = 56$ and $b = 1.2$ as these were found to be good average values for dry snowfall at this radar frequency, and no significant riming was
360 observed (Matrosov, 2007; Matrosov et al., 2008).

Near-field radar measurements can suffer from a variety off issues, such that snowfall retrievals typically must be applied to radar signals that are elevated above the surface. To find an appropriate KAZR range gate to extract snowfall rates, we plotted the cumulative sums of SWE based on KAZR-derived snowfall from reflectivity measurements at different range gates (Fig. 7). The first range gate of 100 m did not yield any measurements, while at 130 m reflectivities were too low. From Fig. 7
365 we see that the differences in the cumulative snowfall from range gates between 220 m, and 280 m are the least. The decrease of computed snowfall with height beyond 280 m is probably due to very low cloud heights in winter (Jun et al., 2016), such that snowfall would get underestimated as these range gates are often at higher elevations within the clouds or even beyond the cloud top. As we found for 280 m the highest snowfall rates, we chose it as range gate from which we extracted the snowfall retrievals. However, based on this simple analysis, the potential differences in snowfall based on this choice of range gate are
370 on the order of about 10 %. With an instrument elevation of 14 m asl, the elevation of the extracted snowfall rates is 294 m asl.
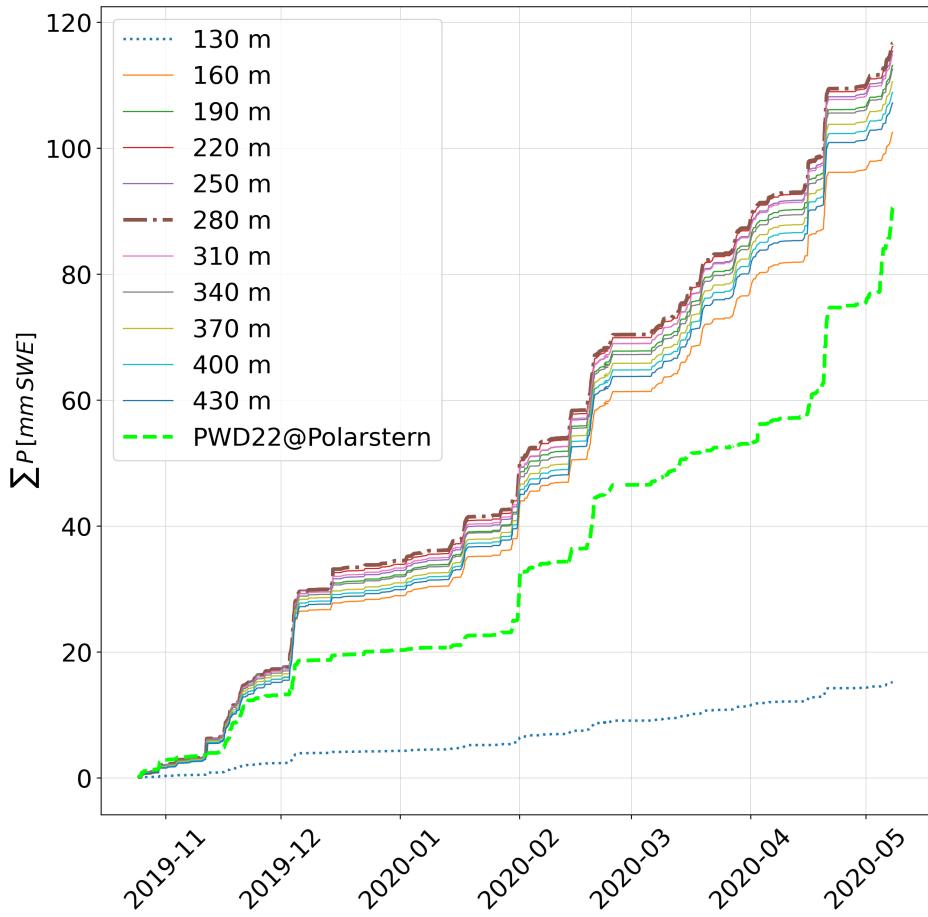
Figure 7. Computed cumulative snowfall for different KAZR range gates and PWD22PS.

2.5 Atmospheric flux station data

A meteorological tower of 10 m height was installed on the ice 558 m away from RV Polarstern at about 60° off the bow of the vessel in the middle of October 2019. However, due to ice dynamics, by the end of leg 3 (beginning of May 2020), the distance was only about 334 m while the direction from the ship stayed approximately the same (Fig. 2). At nominal levels $z = 2\,\text{m}$, 6 m and 10 m above the snow, three-dimensional wind (u, v, w) and temperature were measured at high frequency with METEK uSonic-3 Cage MP anemometers (METEK, 2019), while on the same elevation levels, relative humidity and temperature were measured with Vaisala HUMICAP Humidity and Temperature HMT330 sensors (Vaisala, 2009). The University of Colorado / NOAA surface flux team carried out the post-processing and computed turbulent fluxes, such as momentum flux and turbulent heat fluxes, mixing ratio, or friction velocity. Wind vectors were corrected, i.e., processed wind directions are according to geographic true north. We used wind velocity, wind direction, computed latent heat flux, friction velocity, relative humidity, the temperature at 2 m, and temperature of the snow surface from the described dataset.

2.6 Drifting and blowing snow mass flux

On the meteorological tower described under Sec. 2.5, two snow particle counters (SPCs) (Sato et al., 1993) were installed. The devices continously detect number and sizes of snow particles which are transported through a laser beam. The devices rotate
385 with very low friction on a vertical axis and mounted wind vanes at the back of the sensor keep the laser beam 90°C towards the wind. One SPC was installed at 0.5 m (SPC1106) and one at 10 m (SPC1206) above the snow. The lower SPC1106 ran with only a few interruptions from 2 December 2019 until 7 May 2020 (data availability for this period 96.6 %). The upper SPC1206 ran with only a few interruptions from 14 October 2019 until 7 May 2020 (data availability for this period 94.9 %). One bigger data gap for the SPC1206 was between 17 November 2019 03:05:00 UTC and 18 November 11:15:00 UTC because there was
390 a power interruption due to sea ice dynamics resulting in broken power lines.

To determine periods, where snow transport and erosion has occurred, horizontal mass flux $\left(kg\,m^{-2}\,s^{-1}\right)$ for both SPCs were computed as (Sugiura et al., 2009):

$$Q_{SPC} = \frac{\pi\,\rho_p}{6} \sum_{n=1}^{64} S_n\,N_n\,D_n^3 \tag{8}$$

where ρ_p is the density of a drifting snow particle, which we assumed here as the density of ice $\rho_p = 917\,kg\,m^{-3}$, S_n is the
395 shape factor of snow particles of the n-th class, which we assumed as 1 here, N_n is the particle flux of the n-th class $\left(m^{-2}\,s^{-1}\right)$, which is the number of particles per class passing the SPC sensor area A_s in a second, and D_n is the diameter of a drifting snow particle of the n-th class (m). We consider the density assumption for the lower SPC at 0.5 m above the snow as appropriate, as the fragmentation of snow crystals due to saltation leads to spherical particles (Comola et al., 2017). However, for the SPC 10 m above the snow we expect a mixture of dendritic precipitation particles and spherical blowing snow particles that were
400 lifted from the ground and passed into suspension. Hence, for the upper SPC, an overestimation of the mass flux is expected. We are mostly interested in the timing of drifting and blowing snow events, an distinguishing between dendritic and spheric particles is not possible with the available setup, hence, the distinction is not made here.

To compare drifting snow periods with periods determined from measured mass flux from the SPCs, the critical friction
405 velocity for snow particles was calculated as (Bagnold, 1941):

$$u_{*t} = A \cdot \sqrt{\frac{\rho_{ice} - \rho_{air}}{\rho_{air}}\,g\overline{d}} \tag{9}$$

where A is a threshold parameter and is here assumed as 0.18 as found by Clifton et al. (2006) for drifting snow initiation, $\rho_{ice} = 917\,\text{kg}\,\text{m}^{-3}$ is the density of ice, ρ_{air} is the density of air, $g = 9.81\text{m}\,\text{s}^{-1}$ is gravity acceleration on earth and $\overline{d}$ is an average particle diameter, which we assumed as $260\,\mu m$ which was found as lowest particle diameter on the surface where snow transport was observed by Clifton et al. (2006). ρ_{air} could be retrieved from the meteorological tower data. The computed
410 thresholds were applied to computed u_* from the tower. If $u_* > u_{*t}$, particles begin to get lifted from the ground and drifting

snow flux is initiated.

Although the formula of Bagnold (1941) considers cohesion of the particles with the constant threshold parameter A, it is independent of temperature in the formula. In the 15 experiments Clifton et al. (2006) made, the temperature span was -16 to
415   0 °C. Li and Pomeroy (1997) included, besides wind speed, air temperature in the parameterization of their formula for wind transport threshold to take account for the inhibiting effect on snow transport from sintering. The measurements from the dataset that were used to fit the equation were entirely collected over the main land of Canada. Clifton et al. (2006) compared snow transport threshold formulas from Bagnold (1941), Li and Pomeroy (1997) (but tested without varying ambient temperature) and Lehning et al. (2000), which is reformulated after Schmidt (1980), with experiment results from a wind tunnel. They
420   found good agreements for the Bagnold (1941) formula with $A = 0.18$. Clifton et al. (2006) found better agreements with experimental results for Bagnold (1941) over Li and Pomeroy (1997) and they state that the predictive power of a statistical analysis is limited compared against the approaches from Bagnold (1941) and Schmidt (1980). The formula by Lehning et al. (2000) would require detailed snowpack properties which is not feasible at this point. Hence, we decided to compute u_{*t} after Bagnold (1941) here.

425   ## 2.7 ERA5 mean snowfall rates

For the drift track coordinates of RV Polarstern, shown in Fig. 1, we extracted ERA5 (Hersbach et al., 2020) mean snowfall rates, which is the sum of the convective and large-scale snowfall in ERA5. While the large-scale snowfall is generated from the cloud scheme in the ECMWF Integrated Forecasting System (IFS) (IFS , 2020), the convective snowfall is generated from the IFS convection scheme. The resolution for ERA5 over the sea is 0.28125° x 0.28125°, which is about 31 x 31 km. Hence,
430   the extracted snowfall rate from ERA5 for the drift track does not refer to points but represents an averaged value over these grid cells closest to the drift track coordinates. The purpose here is to compare the ERA5 mean snowfall against snow cover SWE and sensors in this study.

2.8 Sensor and reanalysis comparison method

We computed the average SWE for the northern- and southern loop as we expect this combination of deformed SYI, remnant
435   SYI, and FYI is more representative for an overall snow accumulation estimate. This average computed snow cover SWE serves then as our reference for the precipitation sensors and ERA5 snowfall. Note that for the averaging process, data were discarded when only the northern- or southern transect loop was measured on one day, except when the temporal distance between the measurement of northern and southern loop was short and there was no snowfall in between. This is only the case for 24 (northern) and 26 (southern) April. Hence, the transect averaged SWE starts on 31 October 2019 and ends on 24/26 April
440   2020, with a significant reduction of days of sampling, compared to the loop-differentiated time series that start on 24 October 2019 and end on 5 May 2020. RMSE between snow cover SWE and sensor- and reanalysis estimated SWE was computed for the time period where no SWE decrease in between the days of the transect sampling has been detected by means of the computed snow cover SWE (until and including 20 February 2020). The RMSE is computed as mm and always refers to the

precipitation sum between days of transect sampling. To discuss in more detail the erosion influence on potential discrepancies

445   of snow cover SWE and sensor estimated snowfall, RMSE was also computed for days after time periods where no significant

amounts of horizontal mass flux were detected with the SPCs, i.e. where no erosion between two days of transect sampling

was expected. The detailed comparison follows in the results section.

3   Results

3.1   Snow mass accumulation and decrease

450   To estimate average accumulated snow mass along the transects, the SWE was computed for both transect loops over time as

described in section 2.3. The SWE was rounded rounded to integers for the following description in the text, except when two

values that are compared are very similar. No rounding was conducted before any computations. The average computed SWE

will also serve as reference comparison with snowfall sensors and ERA5 as described in section 2.8. Fig. 8a shows the derived

SWE evolution as box- and whisker plots for the northern and the southern transect. The initial average SWE values for the

455   northern loop (64 mm on 31 October) are naturally higher than for the southern loop (28 mm on 31 October), as the northern

transect was situated mostly on deformed SYI. After 31 October 2019, the value for the northern loop decreased step by step

to around 54 mm until 5 December, while average SWE for the southern loop increased, especially between 14 November to

5 December by 24 mm to a very similar value of 56 mm. From 5 December 2019 on, SWE on both transect loops increased

until 20 February 2020. From then on, there was first a slight decrease and then a slight increase observed for both loops. It is

460   noteworthy that even though the initial SWE over the remnant SYI and FYI (southern loop) was approximately only half of

the value on the northern loop, it reached 89 % of the snow mass of the northern loop by the end of the accumulation period

(northern: 81 mm on 24 April, southern: 72 mm on 26 April). This indicates a "saturation" snow accumulation, which we will

discuss later. It can also be observed that the standard deviation increased for both loops over time (Fig. 9), though it was

initially higher for the rougher deformed ice (35 mm on 31 October 2019) than for remnant, or FYI (13 mm on 31 October

465   2019). Most notable, standard deviations for both loops converged almost completely, where we find 46.8 mm for the northern

loop on 24 April 2020 and 46.2 mm for the southern loop on 26 April 2020. For both loops, the standard deviation did not

increase considerably after 20 February, where 43.8 mm was reached on the southern loop and 43.1 mm was reached on the

northern loop. This was coincident with the average SWE values that did not increase considerably from 20 February until the

end of the accumulation season.

470

We present results about snow mass decrease in the following. The mass decrease computed with the HS-SWE function

from the transect is temporally compared against computed cumulative snow mass flux from the snow particle counters. Note

that the cumulative horizontal mass flux is only an indicator for the strength of the erosion, but cannot be translated into actual

eroded mass. To distinguish easier in the text between computed SWE decrease of the snow cover and cumulative mass flux

475   and to avoid confusions, we keep the designation SWE for the snow cover but use $kg\,m^{-2}$ for cumulative mass flux in the

following, although SWE has the same units as $kg\,m^{-2}$. Notably, we find a net mass decrease, computed with the HS-SWE
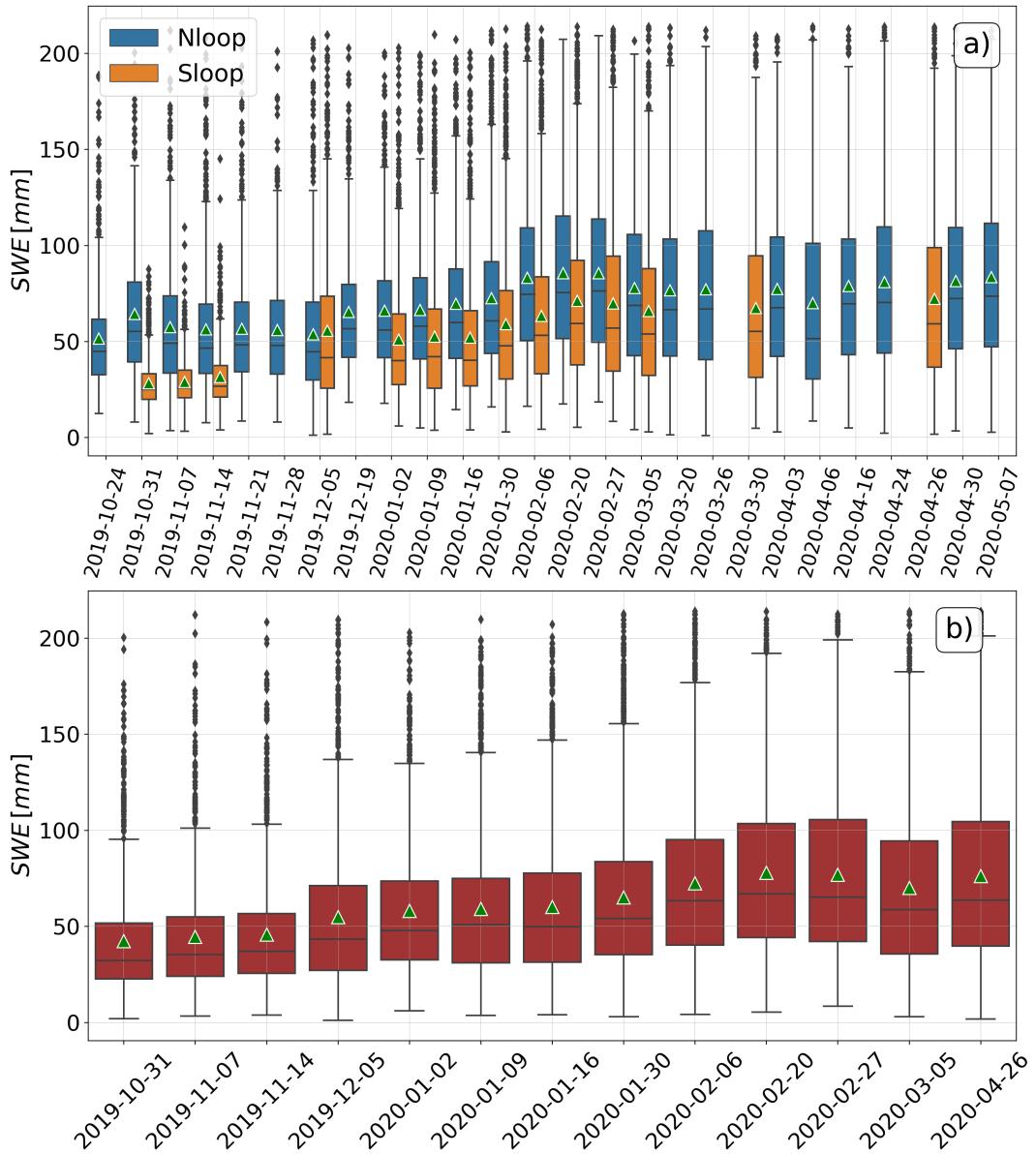
Figure 8. a) Box- and whisker plots for SWE estimates along the northern ("Nloop") and southern ("Sloop") transect, respectively. b) Shows box- and whisker plots for averages of northern- and southern transect loop. Horizontal lines show the median, green triangles the average, the boxes show the interquartile ranges (IQR) (25 - 75 %), the whiskers represent 1.5 times the upper- and lower values of the IQR. The dots represent outliers that are beyond 1.5 times the IQR. Note the different dates between a) and b) as data where only data for one loop was available was discarded for computation of b). The exception is the 24, and 26 April 2020 as the temporal distance was so close that these were averaged, too.
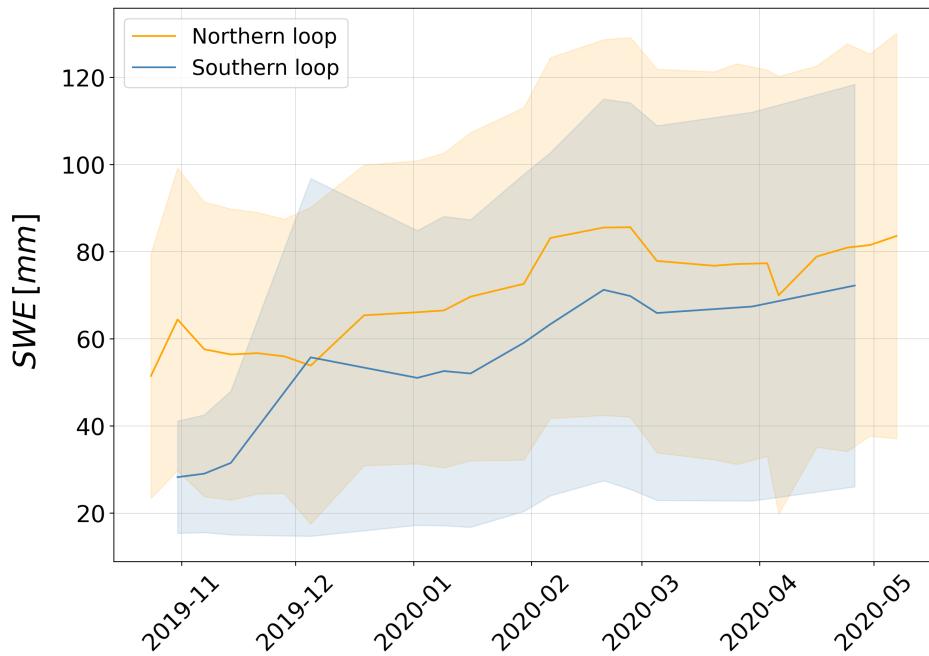
Figure 9. Average cumulative SWE of the snowpack (solid lines) and standard deviation (shaded areas) from this average for the Northern-and Southern transect loop.

function, of 9 mm between 27 February and 20 March for the northern loop. From 20 March on, the mass increased again. For the southern loop, we find a mass decrease from 20 February on, from 71 mm down to 66 mm on 5 March. After 5 March, the mass increased again. The time from 20 February to 20 March exactly falls into the period, where 1) the discrepancy between

480   cumulative SWE from the precipitation sensors and SWE from the transect becomes large (Fig. 11b, Fig. 12a) and 2) where about 45 % (640 kg m^{-2}) of the total cumulative horizontal snow mass flux at 0.5 m above the surface over the whole measuring period of the SPC1104 has occurred (Fig. 12f). That means, over 45 % of drifted snow-mass appeared on only 19 % (30/158) of the days from the whole measuring time of the lower SPC. The period is marked as green shaded areas in Fig. 12. Most distinct in this period was 24 - 25 February (marked as red shaded areas in Fig. 12), in which 323 kg m^{-2} of cumulative mass

485   flux was detected with the lower SPC - which is 23 % of the total detected cumulative mass flux on 1.3 % of the days the device was running. During this storm, a maximum peak of around 11 m s^{-1} was detected in the 1-h averaged wind speed data, which means that the measured peak at shorter time intervals was probably higher.

We computed the sum of the mass decrease that occurred between days of sampling and that is driven by erosion (but does not reflect the total eroded mass), for the northern loop as $\Delta SWE = -27$ mm (between 31 October 2019 - 7 May 2020) and

490   for the southern loop as $\Delta SWE = -11$ mm (between 31 October 2019 - 24 April 2020). However, as erosion occurred on different days for each loop, we find a much lower average of both transects as $\Delta SWE = -17$ mm, between 31 October 2019

and 26 April 2020.

Fig. 10 shows SWE for the same section of the northern loop transect as in Fig. 5, for 20 February and 5 March 2020, which
495 are before and after a drifting snow event. As the section on 5 March had twice the average horizontal Magnaprobe sampling
distance compared to 20 February, and for better illustration, a simple moving average with a window of $n = 4$ was applied to
the section on 20 February while a moving average with a window $n = 2$ was applied to the data from 5 March. A significant
re-distribution due to wind is recognizable. The decrease for this section was $12\,\mathrm{mm}$ SWE, while for the whole northern loop
the decrease was $8\,\mathrm{mm}$ during the same period.

500
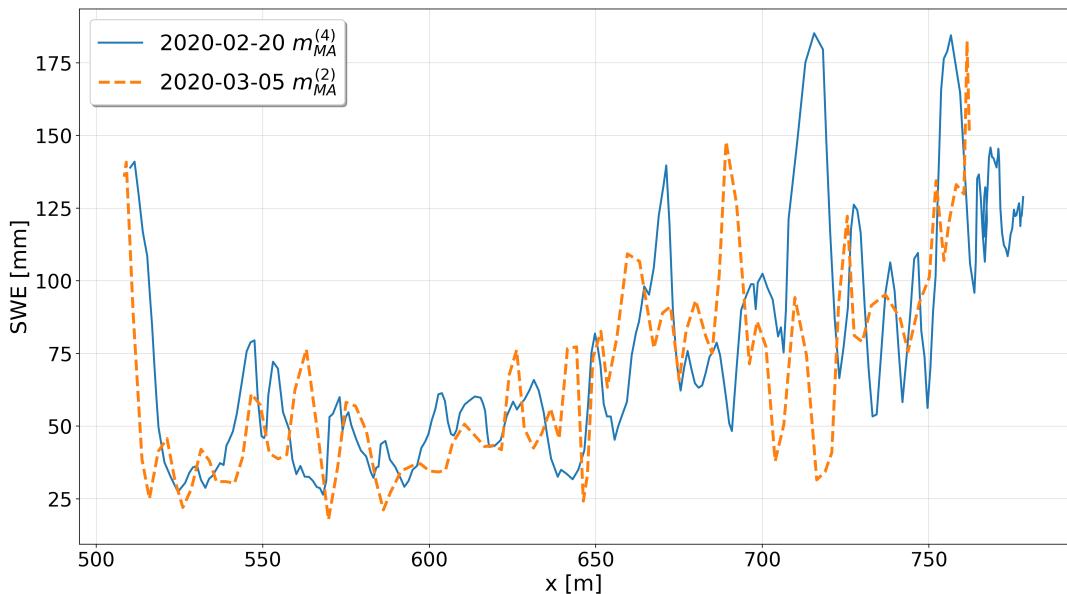
Figure 10. Retrieved SWE from same Magnaprobe section as shown in Fig. 5 on 20 Februrary before a strong drifting snow event that
occurred on 24 - 25 February and mass distribution on 5 March 2020 after the drifting snow event.

Finally, we present results of the average SWE of the northern- and southern loop (Fig. 8b). The average initial SWE value
on 31 October 2019 was $42\,\mathrm{mm}$ and increased to $76\,\mathrm{mm}$ on 26 April 2020. Hence, we estimate the total mass increase over
time as about $34\,\mathrm{mm}$. Over the transect average, the SWE decrease during the snow transport event on 24 - 25 February 2020
was $5.5\,\mathrm{mm}$.

505 **3.2 Precipitation sensor and radar snowfall retrieval comparisons**

To compare with the different estimates of snowfall, the cumulative SWE values were examined for each approach. The plot
of cumulative snowfall between 31 October 2019 and 7 May 2020 without any corrections applied can be seen in Fig. 11a. The

snowfall rates deviate heavily from another between sensors and locations. The P2PS shows the highest cumulative snowfall, while the P2MC shows the lowest, although with limited data availability (Bartholomew, 2020b). As a result of the limited availability and wide spread between these two identical systems operated at different locations, we no longer consider these instruments. The PWD22PS shows the lowest cumulative snowfall, with 97.6 mm while the highest is estimated by the P2PS (290.3 mm). It also stands out that the cumulative sum of ERA5 (110 mm) by 7 May 2020 is very similar to that of the snowfall retrieved from the KAZR (114 mm).

Fig. 11b compares the northern, southern, and average transect loop SWE values with the uncorrected cumulative snowfall from a subset of the sensors and the ERA5 mean snowfall. Especially the PWD22PS is well in line with SWE from the snow cover until mid-end of February 2020. Afterwards, more snow over the ice was eroded, which is indicated by high horizontal drifting and blowing snow mass flux measured with both SPCs (Fig. 12f). After mid-February, the snow cover SWE did not increase significantly and instead stagnates, while the sensors indicate periodic snowfall. Thus, the discrepancy between sensor snowfall rates and snowpack SWE became larger.

We computed the RMSE for the snowfall sensors, KAZR retrieved snowfall and ERA5 relative to the SWE of the snow cover for the time period 31 October 2019 - 20 February 2020, before the first decrease of the SWE from the transect average was observed (Fig. 13, Tab. 2). In this time period, a SWE increase of the snow cover of about 36 mm was detected. In addition to the whole time period (Fig. 13a, cumulated within the intervals in between n = 9 days), RMSE was computed only for days when no drifting or low drifting snow occurred (detected with the SPCs) in between the last and the current day of sampling (Fig. 13b, n = 6 days). The detected drifting snow periods until 20 February were 3 - 5 December, 19 December 2019, 30 Jan - 2 February and 18-20 February 2020 (Fig. 12f). Hence, we discarded 5 December 2019, 6 February and 20 February 2020 of the transect SWE time series. For this time period, an increase of 13.7 mm SWE was detected for the transect. For the whole evaluated time period (n = 9 days), all sensors appear to overestimate. However, this indicates, that erosion occurred in the time periods between the days the transects were sampled, which leads to a systematic positive bias of the sensors. For this period - assuming that no erosion occurred - the PWD22PS is most similar to the SWE (RMSE = 2.5 mm), followed by ERA5 (4.2 mm), the KAZR (5.5 mm), Pluvio2 (7.2 mm), PWD22MC (9.8 mm) and P2PS (26.7 mm). However, when eliminating the days before which drifting snow has been detected with the SPCs, the differences are reduced significantly for all devices and ERA5, too. In this case, PWD22PS still shows the closest comparison to the SWE (RMSE = 1.1 mm) but with a systematic average bias of -1.1 mm. It reveals that also ERA5 performs well (RMSE = 1.2 mm), but with an overestimation tendency, with only one negative biased value found on 7 November of -1.4 mm, while all other values were positively biased with an average of 1.0 mm. The KAZR performs similarly, with a bias of -1.2 mm on 7 November while all other values were positively biased with an average of 1.4 mm. The Pluvio2 performs well without wind influence (RMSE = 1.4 mm), with 4 negatively biased values (-1.5 mm avg) and 2 positively biased values (0.5 mm). The largest difference relative to the SWE is found for P2PS (RMSE = 2.9 mm). However, it must be noted that all sensors except of the P2PS underestimate the snowfall by 7 November 2019, hence we can question the validity of the comparison for this day. In this case, we can speak of an unconditional overestimation of ERA5 and the KAZR. Note that due to the strong cumulative aspect (i.e. we compare snowfall
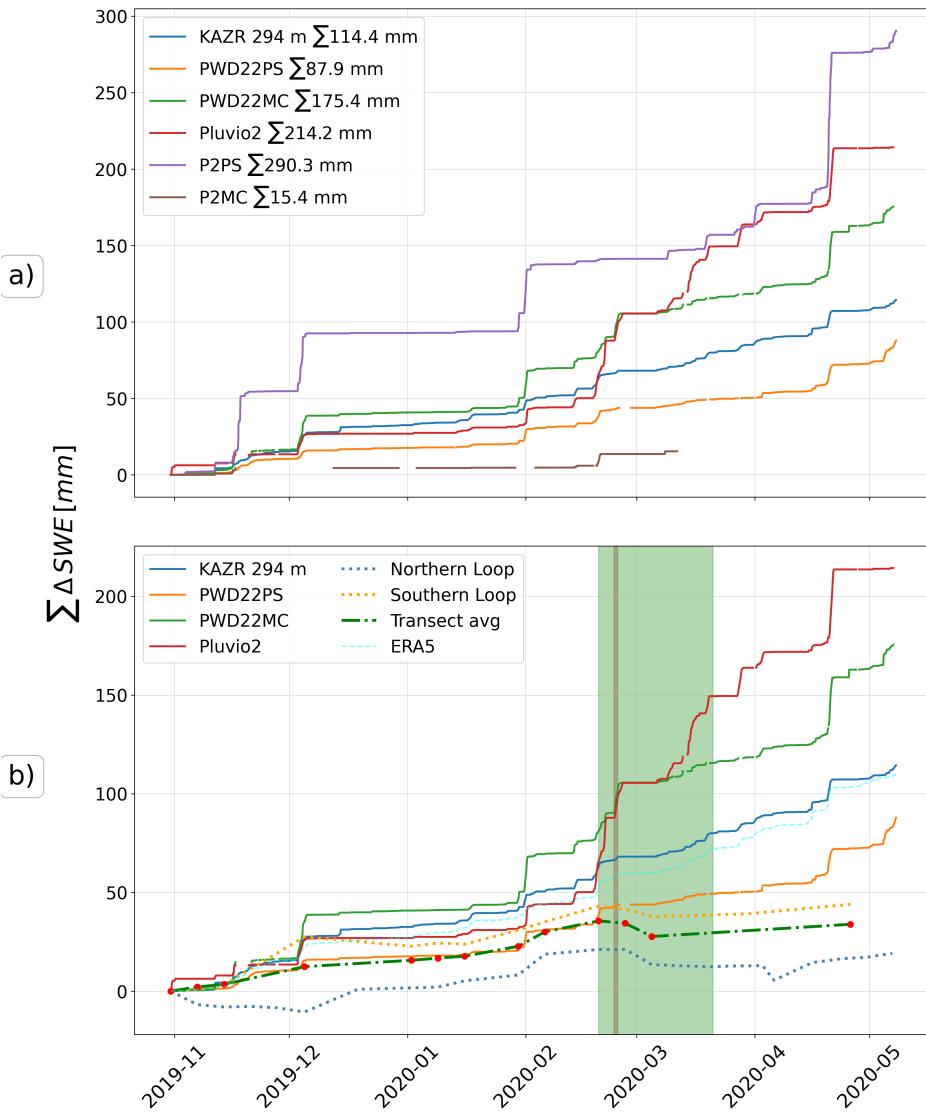
Figure 11. a) Cumulative snowfall for different installed precipitation sensors during MOSAiC from 31 October 2019 to 7 May 2020. b) Sensors, ERA5 estimates and SWE of the snow cover. The red dots show the days on which both transect loops were sampled. The red shading shows the time period of the strong drifting snow event on 24 - 25 February 2020. The green shaded areas mark the strong drifting snow period where 45 % of all cumulative horizontal mass flux was detected (Fig. 12).

that is accumulated always between the days of transect measurements), the difference is naturally reduced when reducing the sample number. Nonetheless, the fact that there is an overall tendency towards a decrease in the apparent overestimation of the

545 sensors relative to the SWE, indicates that erosion likely did occur before the days eliminated in Fig. 13b. For the PWD22PS, the difference compared to SWE is only reduced by about 50 %, while for the PWD22MC, for instance, the difference was

Table 2. RMSE of different devices and ERA5 relative to the computed SWE from the snow cover. The second column shows RMSE before and including 20 February 2020, excluding time periods where drifting snow was detected with the SPCs ($\sum_{transect}$ =13.7 mm). The third column shows RMSE that does not neglect drifting snow days but for the same time span ($\sum_{transect}$ =35.6 mm).

Device/Reanalyis	RMSE [mm] excluding detected snowdrift	RMSE [mm] including detected snowdrift
PWD22PS	1.1	2.5
ERA5	1.2	4.3
Pluvio2	1.4	7.2
KAZR 294 m (Matrosov, 2007)	1.6	5.5
PWD22MC	1.7	9.8
P2PS	2.9	26.7

reduced to about 30 % of its original value. For the Pluvio2, the difference was reduced to 26 % of its initial value. While the apparent overestimation of the sensors in Fig. 13a is likely due to erosion (and hence strongly biased), the different magnitudes of RMSE reduction (Fig. 13b, Tab. 2) suggest that PWD22PS is less affected by overestimation due to high wind speeds that

550 accompany blowing snow, compared to PWD22MC or Pluvio2, both of which were installed near the surface. This is mainly due to a a much lower bias for PWD22PS on 5 December 2019 and 6 February 2020, before eliminating drifting snow periods, compared with Pluvio2 and PWD22MC.

In summary, if we only consider time periods without drifting snow, PWD22PS and ERA5 compare most favorably with

555 SWE, with reasonable results for Pluvio2, KAZR and PWD22MC, as well. The P2PS shows the largest differences. When also considering high wind speeds and blowing snow, the PWD22PS still appears to compare most favorably with SWE, while especially P2PS, PWD22MC and Pluvio2 appear to be most negatively affected by high wind speeds.

We find in total of five significant snowfall events. If we use PWD22MC as reference, we find for 3 - 5 December 2019: ≈ 5.5 mm, 30 January - 3 February 2020: ≈ 10 mm, 18 - 21 February 2020: ≈ 8.5 mm, 16 - 21 April 2020: ≈ 16.5 mm, 4 -

560 7 May 2020: ≈ 14 mm. Hence, about 54 mm of snow fell during events, while the other 33 mm fell in between, e.g. as trace precipitation or diamond dust.

To better illustrate the blowing snow influence on sensors that is already suggested by Fig. 13, we made scatter plots of snowfall rates from different sensors with respect to the PWD22PS. Fig. 14 shows a scatter plot for the short time of two

565 days between 24 and 25 February, where high drifting snow mass fluxes were detected. We can clearly see that the Pluvio2 (Fig. 14b) and the PWD22MC (Fig. 14d) strongly overestimate snowfall relative to PWD22PS, while P2PS (Fig. 14a) and KAZR (Fig. 14d) stay largely unaffected and only measure trace precipitation of 0.1 to 0.2 mm h^{-1}. This becomes even clearer when we look at snowfall rates of Pluvio2 (Fig. 15a) and PWD22MC (Fig. 15b) versus the horizontal mass flux detected by the SPCs. Scatter plots for the whole period (31 October 2019 - 7 May 2020) for different sensors versus horizontal mass flux show
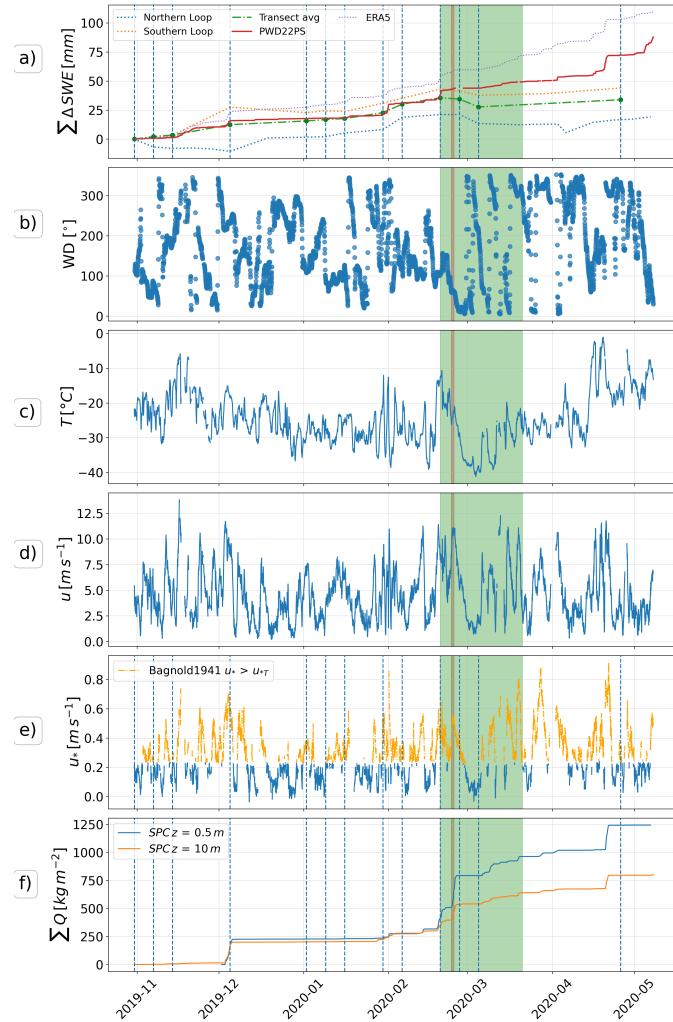
Figure 12. Time series from 31 October 2019 to 7 May 2020 for a) estimated SWE from the transect and HS-SWE model as well as cumulative snowfall from ERA5, cumulative snowfall from PWD22 on Polarstern, KAZR derived snowfall rates. b) wind direction at 2 m, c) air temperature at 2 m above the snow d) wind speed at 2 m height, e) computed friction velocity threshold for snow transport after (Bagnold, 1941), f) cumulative horizontal mass flux with the snow particle counter at 0.5 m above the snow and at 10 m height, respectively. The green shaded areas mark the strong drifting snow period where 45 % of all cumulative horizontal mass flux was detected. The vertical blue dashed lines mark the days where both transect loops were sampled.

570   the influence of drifting and blowing snow, too (Fig. 16). Pearson correlation coefficients show medium positive correlations for mass flux and snowfall from Pluvio2 (Fig. 16b, r = 0.55 - 0.58) and PWD22MC (Fig. 16d, r = 0.54 - 0.55) while a weak negative correlation is observed for P2PS (Fig. 16a, r = -0.19) and for PWD22PS (Fig. 16c, r = 0.26). These results indicate
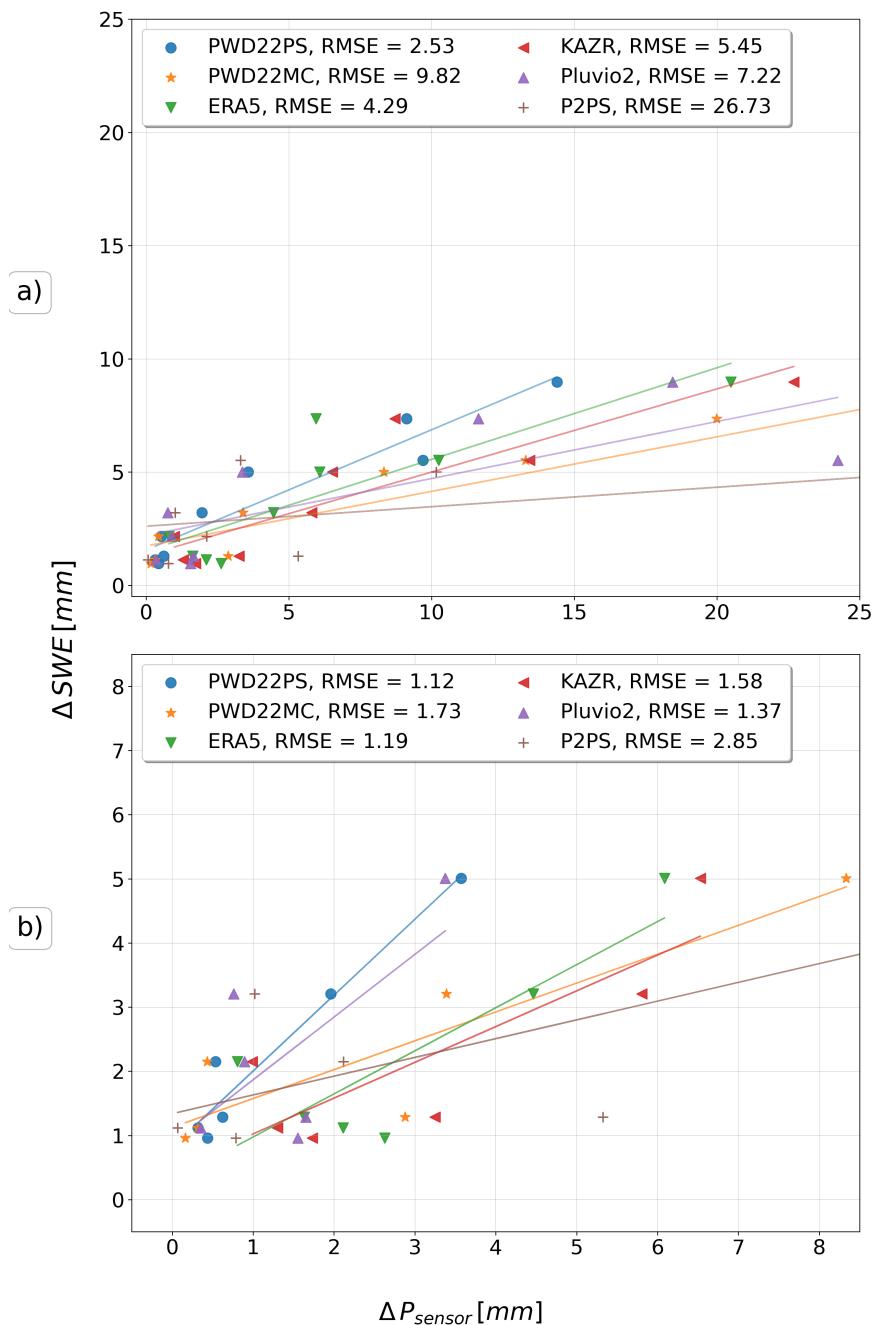
Figure 13. a) RMSE [mm] of the sensors and ERA5 with respect to snow cover SWE, for the time period before 20 February 2020 with n = 9 days of sampling, including days when drifting snow was detected. b) as a) but without days when drifting snow was detected before (n = 6 days).

that instruments collocated on the ice at about 1 - 1.5 m height are much more affected by drifting and blowing snow than instruments installed on Polarstern at 22 m height.
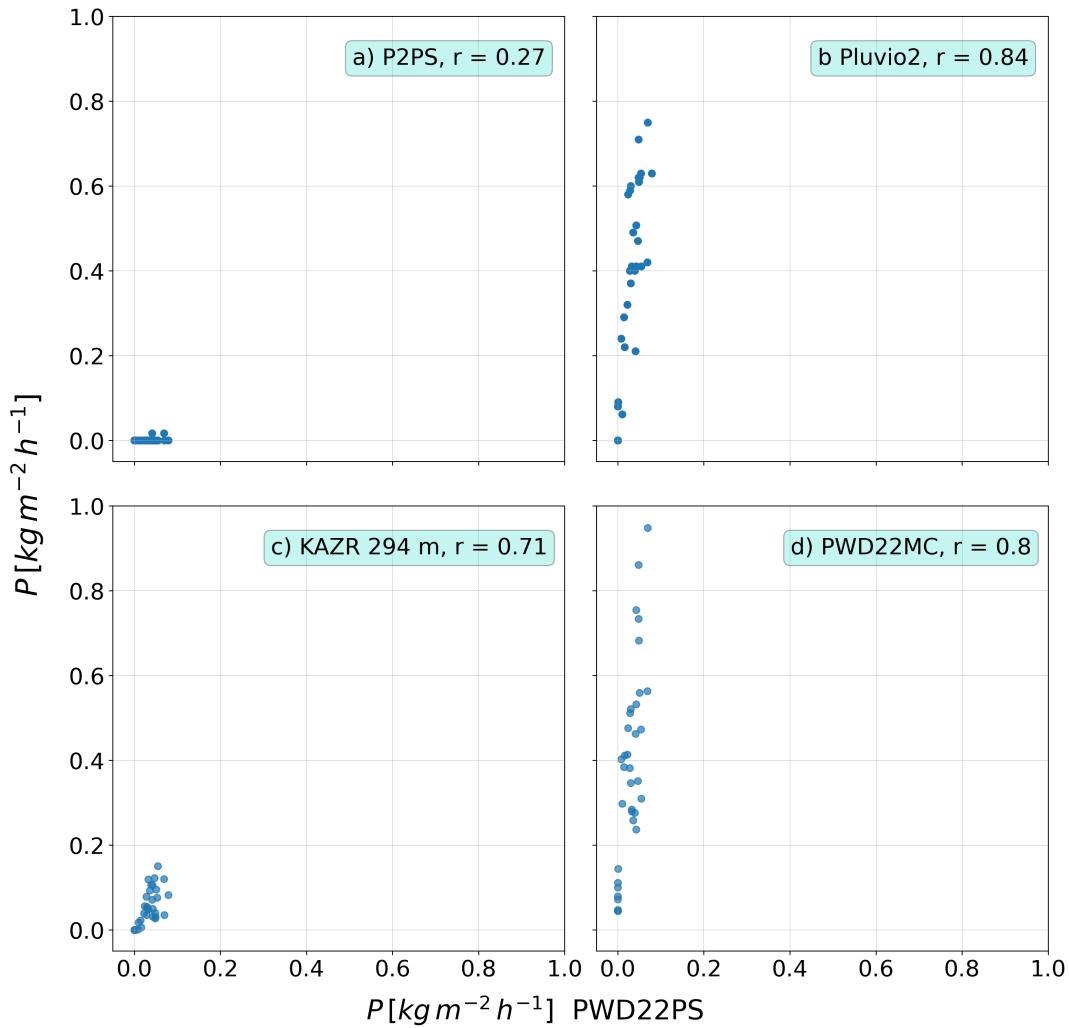
Figure 14. Scatter plots of PWD22PS snowfall rates vs. different sensor snowfall rates for the drifting snow event on 24 - 25 February 2020 for a) P2PS, b) Pluvio2, c) KAZR and d) PWD22MC.

575

3.3   ERA5

We compared ERA5 mean snowfall rates against the snow cover SWE and against PWD22PS. As described above, when we consider the comparisons illustrated in Fig. 13b and Tab. 2, ERA5 shows reasonable results with a relatively low RMSE of 1.2 mm, an overestimation tendency and even an unconditional overestimation when neglecting 7 November 2019 as evaluation
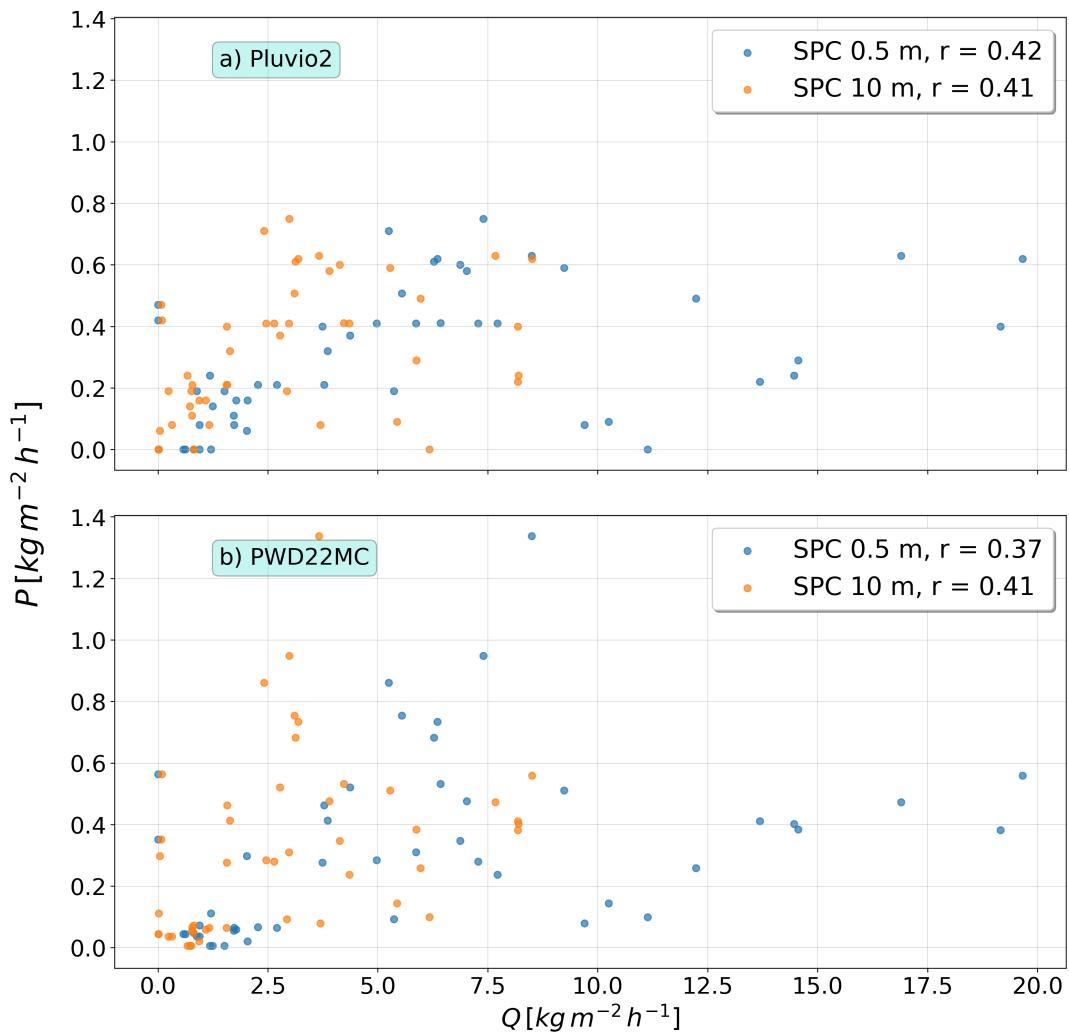
Figure 15. Scatter plots of sensor snowfall rates vs. SPC mass flux for the drifting snow event on 24 - 25 February 2020 for a) Pluvio2 and b) PWD22S3.

value. Assuming PWD22PS as reference, ERA5 shows an overall good timing of the snowfall events (Fig. 11b, Fig. 12a). As for the transect SWE validation, it overestimates snowfall relative to PWD22PS, too, in this case systematically and with an acceleration of the positive bias from the end of February on. This leads to an overestimation (relative to PWD22PS) of the total accumulation of almost 22 mm (+25 %) by the end of the investigation period. We computed the RMSE for the snowfall rate relative to PWD22PS as 0.06 mm h^{-1} for the whole time period from 31 October to 7 May 2020.
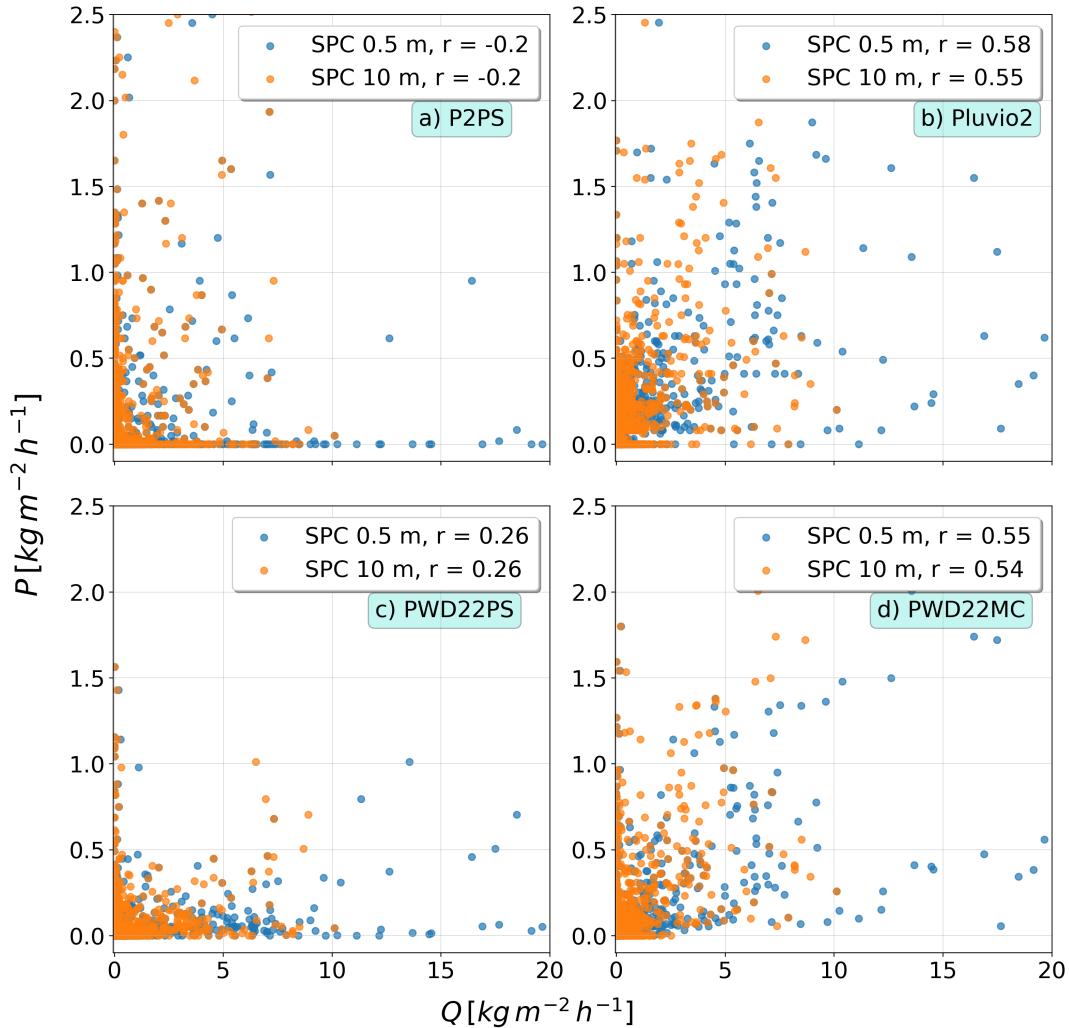
Figure 16. Scatter plots of sensor computed snowfall rates vs. SPC mass flux for the whole time period for a) P2PS, b) Pluvio2, c) PWD22PS and d) PWD22MC.

4 Discussion

4.1 Snow mass balance

With the fitted HS - SWE function, we were able to retrieve the SWE of snow cover over the ice for the transect loops. We found a SWE increase from the beginning of the measurements until around 20 February 2020, while from then on, the average SWE did not increase anymore, although snowfall was occurring. Following three significant snowfall events (3 - 5 December 2019: ≈ 5.4 mm, 30 January - 3 February 2020: ≈ 10 mm, 18 - 21 February 2020: ≈ 8.5 mm), a strong drifting and blowing snow event occurred on 24 - 25 February. Consequently, the total SWE decrease on the southern transect loop was about 5 mm

until 5 March 2020, while the decrease was almost 9 mm on the northern transect loop. The net decrease generally includes eroded mass in addition to incoming precipitated mass, hence the eroded mass was often larger than the precipitated mass. Followed by the drifting snow event on 24 -25 February, we find two significant snowfall events. When considering PWD22PS as reference, the snowfall sums for these events were: 16 - 21 April 2020: ≈ 16.5 mm, 4 - 7 May 2020: ≈ 14 mm. Hence, about 54 mm of snow fell during five events in total, while the other 33 mm fell in between, e.g., as trace precipitation or diamond dust.

We observed an average snow mass increase of 36 mm until 20 February 2020, and from then on a slight decrease was observed to 34 mm until end of leg 3. Within this time period, the decreased was strongest between 20 February 2020 and 5 March 2020, with almost 7 mm down to 28 mm. From 20 February 2020, also the standard deviation stabilized for both transect loops, on deformed SYI and remnant SYI/FYI at around 44 to 45 mm (Fig. 9). This raises the question of a saturation mass of snow that can accumulate in this windy environment. As the time of the maximum mass goes along with the stabilized standard deviation, the roughness of the surface might be a limiting factor. It remains to be investigated why most drifting snow and erosion occurred from 20 February 2020 on, as indicated by transect SWE decrease and horizontal mass flux. One might consider that lower horizontal sampling distance ($\overline{D}_{5Mar} = 1.6$ m) could have led to an underestimation of the accumulated snow. However, from Fig. 6 one can see that reducing the sampling frequency by half has no significant effect on the average. One can only see a significant fluctuation from reducing the frequency to 1/3 and below. Hence, a reduced sampling frequency does not explain the mass decrease. Ice dynamics that affected the transect occurred from 11/12 March 2020 and thereafter. Hence, no ice dynamics influenced the sampling on 5 March 2020. Nevertheless, an impact on the time series cannot be excluded from 12 March 2020 on. Looking at horizontal mass flux and decreased computed snow mass together, we suggest that erosion was largely responsible for the mass decrease. Other than observed, the critical friction velocity after Bagnold (1941) (Fig. 12e) does not indicate more drifting snow compared to other time periods. However, the formula after Bagnold (1941) is rather simple and neglects varying temperatures and therefore varying bond strenghts of snow particles at the top of the snowpack. The bond strength depends strongly on weather history. The preceding long period of air and snow surface temperatures often below -30 °C (Fig. 12c) could have inhibited sintering, which is strongly temperature-dependent and develops more slowly at lower snow temperatures (Colbeck et al., 1997; Colbeck, 1998; Blackford, 2007). The Bagnold formula does not consider splash entrainment (Comola et al., 2017; Comola and Lehning, 2017) or surface roughness and atmospheric stability, which largely affects the near-surface wind field. Furthermore, the used bond strength parameter of $A = 0.18$ that we used in the computation was found by (Clifton et al., 2006) in a wind tunnel, with a temperature range of -16 to 0 °C which was undershot most of the time during our investigation period of MOSAiC (Fig. 12c).

In any case, we expect the transect time series until and including 5 March 2020 to be mostly valid. However, the usage of the transect SWE for sensor- and reanalyis comparisons is limited, as erosion has very likely occurred over some periods between the transect sampling days. Hence, the amount of erosion is not quantifiable based on the available transect data. Nonetheless, days where it is likely that no erosion occurred since the previous transect sampling prior to 20 February 2020 could be detected by means of snow particle counters. The cumulative transect SWE of these time periods was then compared

against cumulative SWE from the sensors and ERA5 for the same time periods, which will be discussed in the next section. The validity of the SWE time series decreased with increasing ice dynamics from 11/12 March 2020 on and with decreasing temporal sampling frequency on both northern and southern transect loop. The largest gap in sampling occurred between 26 630 March and 24/26 April.

If we assume, based on the findings, that PWD22PS is least affected by blowing snow, and provides a reasonable estimate of snowfall, then the snowfall between 31 October 2019 and 26 April 2020 was about 72 mm. We know that the PWD22PS showed the lowest cumulative snowfall and a systematic negative bias compared against ground truth (Fig. 13). Further, it has 635 been suggested that PWD22PS tends to underestimate snowfall, where values found of 30 % less compared to a Geonor gauge in a DFIR (Wong (2012a), Tab. 1). Until 7 May 2020 as end of MOSAiC leg 3, the total snowfall would then be at least about 98 mm in 189 days. If we assume the retrieved snowfall from the KAZR as upper limit for cumulative snowfall (as it was demonstrated that Pluvio2 and PWD22MC were affected by wind and blowing snow towards overestimation), we find about 114 mm. This value is comparable with the cumulative snowfall from ERA5 (110 mm). Hence, during 189 days, was assume 640 that the total snowfall was between 98 and 114 mm. With the total mass increase of 34 mm for the transect SWE during the full observation time, we can approximately compute the minimum total eroded mass as 38 mm until 26 April 2020, which is more than the value of the mass that has accumulated along the transect during the same period. With more than 50 % of eroded snow mass, we find comparable magnitudes as Leonard and Maksym (2011), although they investigated snow over Antarctic sea ice and their time period for investigation was only about 1 month. However, our findings also compare well to results from 645 Essery et al. (1999). Further sensor assessments discussions are made in the next section.

Sublimation of snow crystals during trace precipitation or diamond dust could have led to snowfall detection in the optical sensors but no SWE increase of the snow cover. However, the study design does currently not allow to investigate that but it is recommendable to be investigated in detail. Earlier studies suggest that blowing snow sublimation may be responsible for about 650 6 % for the mass sink (Chung et al., 2011). The relatively low sublimation rates also arise from the high relative humidities found over sea ice, inhibiting further sublimation due to a quick saturation of the air. Snow cover sublimation can generally be expected to be negligible during polar night (Webster et al., 2021).

Studies of Déry and Tremblay (2004), Leonard et al. (2008) and Leonard and Maksym (2011) indicate that, besides subli- 655 mation, large parts of drifting and blowing snow will drift into the open water of leads. However, for the drifting snow event on 24 - 25 February, there was not significant open water in the vicinity. Thus, without such a local sink, we would expect that even though snow was eroded, more snow-mass would be delivered from the upwind side at approximately the same amount. However, the given mass of drifting snow from the upwind side depends on the low-pressure system's extent and trajectory associated with the high wind speeds. In Fig. 12 we see that after the event, some atmospheric parameters have 660 changed significantly and rapidly (a temperature drop, a decrease in wind speed). The rapid change of atmospheric conditions could explain why no "new" snow has been transported from the windward side. Another important factor could be, that the

area that the transect covers, has a relatively low surface roughness, with ridges that were generally shorter than in the surrounding region. Erosion might be large over these relatively smooth areas that the transect covers, and the delivered drifting snow from the upwind side could have been deposited at higher ridges upwind. But even a just frozen lead will be filled by drifting snow after a while, that would lead to a mass sink on the upwind side, inhibiting further snow transport downwind. Considering the ALS-based digital elevation map of the floe (Fig. 2) and the wind direction around 24 - 25 February (between North and North-East) (Fig. 12b), it is evident that the northern transect was likely to be partially wind-shadowed by higher surface structures on the upwind-side. This wind-shadowing can lead to less wind transport downwind but maybe also to less erosion due to lower surface friction. This is suggested by our findings, too, as we found more substantial erosion over the northern transect loop. During the drifting snow event, the southern transect was often in the wind-shadow of the vessel and other installations (Fig. 2), leading to decelerations and less surface friction that could have led to snow transport or erosion. In the case of more deposition upwind, the mass over the sampled area would decrease even if the friction velocity behind the obstacles is decreased, which leads to less erosion. The computed SWE would not be representative for the whole floe in this case. While potentially interesting, the method used in our study does not allow to investigate these balances. Another reason for a computed SWE decrease during the later part of the observations might be wind-induced compaction of the upper snowpack. With our approach, we assume almost a linear relationship between snow depth and SWE. However, snowdrifts consist of dense packed, rounded grains (Fierz et al., 2008). Hence, in this case, we probably under-estimate the increase (decrease) of the SWE in case of drifting snow deposition (erosion). A thorough investigation is recommended, but beyond the scope of this study. Another reason for underestimating snow depth with spatio-temporal sampling approaches might be that parts of the snow are caught in the porous parts of ridges, depending on wind direction and on the age of the ridges. This is something that is not covered by this study. Furthermore, for the erosion, the upper snowpack's microstructural composition plays a role, especially the strength of the bonds associated with sintering and temperature.

4.2 Snowfall sensor estimates

Snowfall rates from precipitation gauges show large differences among each other. Battaglia et al. (2010) already showed that OTT Parsivel instruments overestimate the number of large particles. Further, the uncertainty they found regarding fall velocity was high. Wong (2012a) also found a large overestimation of snowfall for the OTT Parsivel of about 50% (Tab. 1), while the overestimation became larger during high wind speeds. The relatively low accuracy of $\pm 20\%$ for snowfall as given by the manufacturer compared to other instruments does make strong overestimations not surprising. Observations examined here appear to confirm this overestimation. Although the installation on the top deck of RV Polarstern at 24 m height appears to be well protected against blowing snow (Fig. 16a) the overestimation of P2PS rather appears to be due to wind itself as also found in earlier studies.

Regarding Vaisala PWD22 and snowfall, Boudala et al. (2016) found that in comparison with manual measurements, it overestimated snowfall by about 33 % due to detected snow crystals not observed by the human observer. Wong et al. (2012b) instead, found good agreements with two Vaisala VRG 101 (double Alter shielded) and two OTT Pluvio (Tretyakov shielded)

gauges. Compared to a Geonor gauge in a Double Fence Intercomparison Reference (DFIR), Wong (2012a) found for the PWD22 an underestimation for solid precipitation of 32 %. Hence, we expect and also can confirm an underestimation of snowfall during MOSAiC. However, the RMSE was with 1.1 mm the lowest compared against all other validated sensors. Wong (2012a) found little wind influence on the PWD22 measurements, which we can confirm, as the RMSE relative to SWE

700 estimates was reduced least (as relative value) compared against all other sensors after drifting snow periods were eliminated from the comparison. Compared against SPC mass flux, we demonstrated that the installation on the railing on the top deck of RV Polarstern has protected the sensor from blowing snow. On the contrary, the same device installed in Met City on the ice showed a relative overestimate, suggesting an influence from blowing snow. Comparisons with the KAZR-derived snowfall, which is also unaffected by blowing snow, further support the notion that the PWD22PS was not significantly impacted by

705 blowing snow.

The undercatch of snowfall due to the wind for weighing bucket gauges is well known (Goodison et al., 1998; Boudala et al., 2016; Kochendorfer et al., 2017). However, we observed a strong positive bias in the cumulative snow mass for the Pluvio2 gauge when comparing with PWD22PS. This bias was largely due to a strong drifting snow event at the end of February, prob-

710 ably leading to blowing snow being lifted from the ground and landing in the bucket (Fig. 15,16). Available transfer functions (e.g., Goodison et al. (1998); Boudala et al. (2016); Kochendorfer et al. (2017, 2018)) cannot correct for this type of blowing snow event, as they correct the underestimation with increasing wind speed. However, we suggest that these events can be detected using snow particle counters and removed before applying other corrections.

715 The snowfall retrieval from the KAZR using a Z_e-S relationship with coefficients determined by Matrosov (2007) shows a slight overestimation relative to the snow cover SWE but performs comparably well as ERA5. The difference to ERA5 in cumulative snowfall was only about 5 mm by end of the investigation period. Considering the known low bias in PWD22 measurements, and the fact that the radar snowfall retrievals are not affected by wind or blowing snow, it is likely that this radar-based estimate provides reasonable results. However, such radar-based retrievals are dependent on the inherent proper-

720 ties of the snowfall observed in the datasets from which they are derived. If the snowfall observed at MOSAiC had distinct properties, such as aspect ratio distributions or mass-size relations, then the relationship used here might have an increased uncertainty. More effort could also be taken to find the best range gate for the retrievals.

4.3 ERA5 performance

725 ERA5 performed reasonably well, with a slight overestimation when compared with snow cover SWE. Overall, the timing of the snowfall events is represented well by ERA5. Relative to PWD22PS, the cumulative ERA5 snowfall would have been overestimated by about 25 % by 7 May 2020. However, as already discussed, due to the light underestimation of the PWD22PS, the true snowfall lies probably between PWD22PS and ERA5 or the KAZR. Relative to PWD22PS, ERA5 appears to perform better before about March 2020, although the available data does not allow to prove that it generally performs worse before. As

730 the ERA5 performance depends on input of measurements and numerical weather data, we can at least point out that there was
a substantial decrease of the air- (minus 50 - 75 % between March and May 2020, Chen (2020)) and ship based observations
which may lead to a worsening of the performance. Chen (2020) found globally a worsening of temperature forecasts of up to
2 °C in the time period March - May 2020, compared against February 2020. Nonetheless, a comparison with the cumulative
sum of the KAZR (Fig. 11b) rather speaks against this theory as the ERA5 cumulative snowfall is always below the KAZR
735 snowfall. Exact reasons with respect to ERA5 should be thoroughly investigated. Cabaj et al. (2020) also found a general
positive bias (in daily snowfall rates) comparing ERA5 with CloudSat data. However, CloudSat is only available up to 82 °N,
while MOSAiC leg 1 - 3 was most of the time north of 84 °N and for large parts even above 87 °N (Fig. 1). Further, in
Cabaj et al. (2020), they found a decrease in the positive bias towards the spring months compared with the winter months of
December, January, and February while we find an increased apparent bias during this time when comparing with PWD22PS.
740 Wang et al. (2019a) compared ERA5 snowfall with data from several snow buoys for a comparable drift track in the Arctic
ocean to MOSAiC and a comparable time spans as our validation, from autumn to spring in the following year. Although they
do not give details about the method, they mention a positive bias of ERA5 cumulative snowfall of about 63 mm compared
with snow buyos by end of the investigation period. They found partially a negative bias and very varying results but on
average a positive bias. However, as already pointed out, it cannot be concluded from a few point measurements on the overall
745 accumulated snow on a larger area. Further, from our study we see that we cannot conclude from the computed SWE of the
snow cover alone on the precipitated sum as the erosion appears to be large, even when measured over larger areas. Nonetheless,
the results from Wang et al. (2019a) also indicate that a lot snow mass gets eroded over time, which is even more evident as
we as well as Cabaj et al. (2020) find an overestimation tendency for ERA5 snowfall. We must further notice that ERA5 does
not consider blowing snow sublimation in their computations. Although Orsolini et al. (2019) do not find a mass effect of
750 including blowing snow sublimation in ERA5, Chung et al. (2011) for instance, computed a larger effect of blowing snow
sublimation of 12 mm y^{-1} over sea ice over 324 days. The non-existent blowing snow particle sublimation may be a reson for
the overestimation tendency found in our study.

5 Conclusions

We fitted a HS-SWE function to computed SWE values from SMP force signals and applied the model to snow depths from
755 Magnaprobes for the northern- and southern transect loops of MOSAiC, leg 1 to leg 3. We could show that the SWE recon-
structions compared well against direct SWE measurements. Besides transect paths, other snow depth sources, such as snow
height differences from terrestrial laser scans, could be used with this function to compute SWE. A validation of the horizontal
sampling frequency of the Magnaprobe along the transect path shows that the transect averages for all days of sampling are
mostly robust against sampling frequency changes down to about about 3.4 m. Although we cannot account for uncertain-
760 ties due to sampling differences by different operators regarding the applied force, our validation indicates that no significant
under-sampling occurred and the observed average is likely a good estimate of the actual SWE.

We could show that SWE differences between snow on deformed SYI and snow on remnant SYI and FYI decrease until
the end of the snow accumulation season at the beginning of May. SWE on remnant SYI and FYI, while starting out at only
about 50 % in late October 2019, reached almost 90 % of the value for snow on deformed SYI by early May. As the SWE
differences decreased and the total SWE increased, the standard deviation for both loops equalized at around 44 to 45 mm from
20 February 2020 on, suggesting that the snow surface roughness becomes larger over time until 20 February but stagnates
from then on. Although erosion and deposition occurred at different times for both underlying ice types before 20 February
2020, it is noteworthy that the first decrease of the average SWE occurred at the same time as the average maximum standard
deviation on both loops. Since SWE also did not increase much after this time, these observations raise the question as to
whether there is a saturation point for snow mass accumulation and snow surface roughness over sea ice. We suggest a range
of 38 mm (53 %) to 69 mm (68 %) of precipitation that has been eroded over time, with the PWD22PS lower and the KAZR
as upper limit of cumulative snowfall. The fate of the eroded snow is unclear, but it is likely that a significant amount was
deposited around higher ridges or filled in the gaps of frozen leads, limiting its transport to the areas covered by the transects.

We detected a remarkable snow erosion event between 24 to 25 February 2020. However, as transects were conducted
approximately weekly, processes in the snow that have occurred in the meantime are not detected. More snow erosion accompanies the temporal adjustment of the surface roughness between remnant SYI and deformed SYI. Besides temporal sampling
frequency of the transect, another limitation of the study is that layering of the snowpack was not considered for estimating deposited and eroded snow mass. However, validation measurements at different points in time suggest that the impact
of this effect might be small. A thorough investigation of layer density is recommended. Further research, connecting snow
microstructure with snow transport rates, for instance, investigating sintering and bond strengths of snow grains that depend
on temperature, could help to elucidate high Arctic snow processes. This is also important as the eroded snow has influenced
several sensors' measurements.

Although the unquantified eroded snow mass limits the potential for intercomparisons, we found that the PWD22PS showed
the smallest differences with the estimated SWE during periods without indication of drifting snow. We demonstrated that the
PWD22 on RV Polarstern was mostly protected from blowing snow influences while this blowing snow led to high correlations between horizontal mass flux and precipitation rates of Pluvio2 and PWD22 at Met City. Similar to PWD22PS, the KAZR
snowfall retrievals are likely not affected by wind or blowing snow. It shows a low bias relative to snow cover SWE although
it appears to overestimate snowfall. This overestimation however can be reduced when doing more research on selecting range
gates and fitting of the coefficients. The OTT Parsivel[2] showed a strong tendency towards overestimation. We assume that the
OTT Pluvio[2] originates to significant parts from blowing snow that was blown into the sensor opening. For PWD22PS, over
189 days, we find a total cumulative snowfall of 98 mm. We suggest the KAZR snowfall as possible upper cumulative limit
with 114 mm. ERA5 showed a reasonable performance, with a good timing of snowfall events but a relatively safe tendency
towards overestimation. They slight overestimation may arise from non-existing blowing snow sublimation. However, we made
the first validation of ERA5 snowfall for the high Arctic based on relatively reliable measurement data to our knowledge.

This study sets the base for future snow mass balance research for MOSAiC and further general snow research for Arctic sea ice. We are better aware of snow mass quantities that accumulate and erode over the central Arctic sea ice surface for almost the whole accumulation season, including the polar night. The data can be used to improve numerical snow cover-, sea-ice, weather- and climate models and for more detailed process understanding research across disciplines. Initial sensor validations were conducted, which allow for more specified, more thorough research.

Data availability. Derived SWE from the SnowMicroPen is available at https://doi.pangaea.de/10.1594/PANGAEA.927460. All ARM-related data (KAZR, all used precipitation gauges in the study) are available in the ARM data archive https://adc.arm.gov/discovery/). Computed SWE along the transects and SWE ETH tube measurements are soon available on PANGAEA. Met tower data will soon be available on the Arctic Data Center (https://arcticdata.io/). SPC data will soon be available on the UK Polar data centre (https://www.bas.ac.uk/data/uk-pdc/).

Appendix A: Central Observatory map by 30 March 2020

The map shows the Central Observatory (CO) on 30 March 2020 at a different condition compared to the map of 5 March 2020.
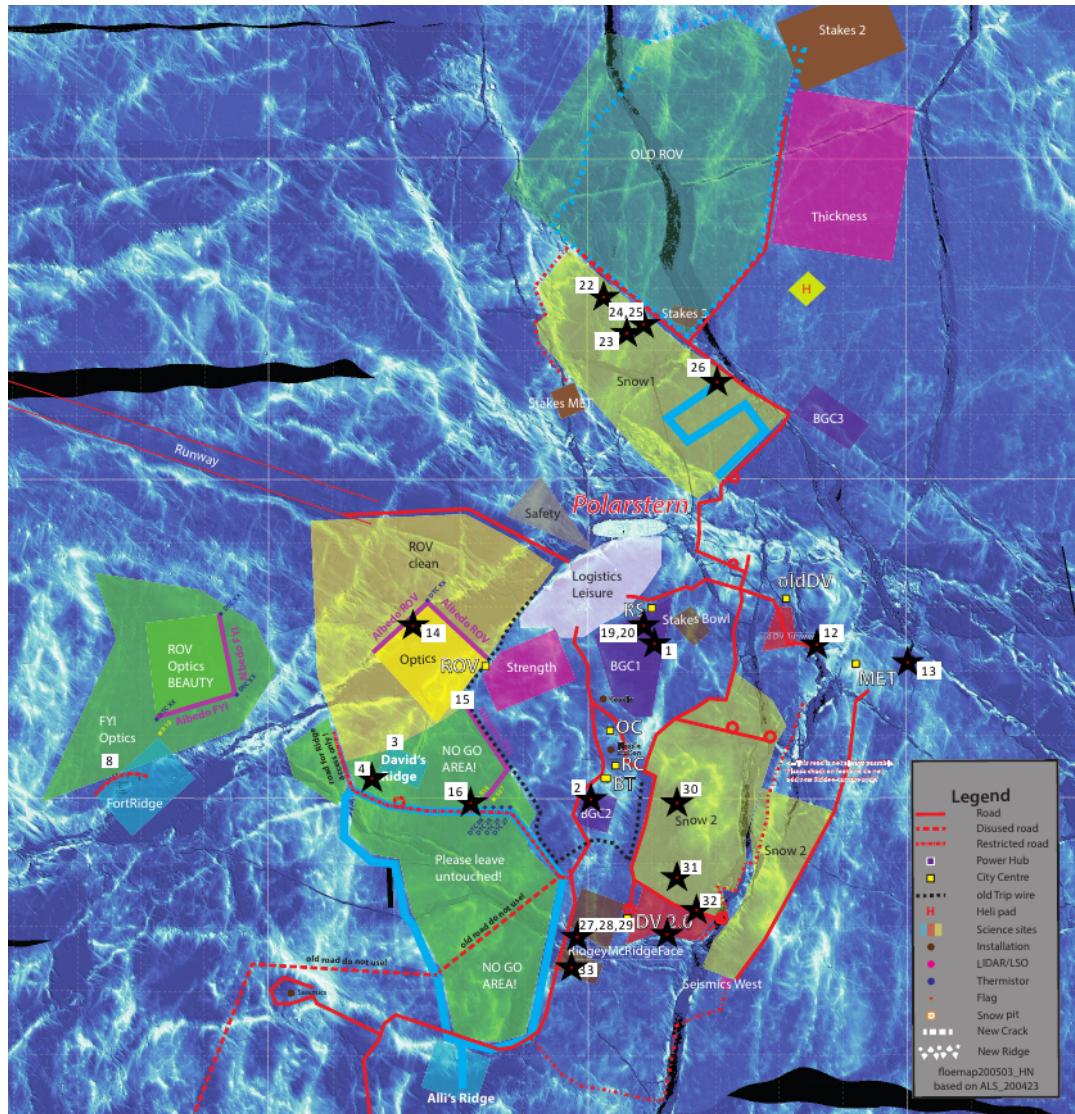
Figure A1. CO map by 30 March 2020. The map shows the Central Observatory (CO) at a different condition and with all installations on the floe compared to the map of 5 March 2020. Snowpit locations are marked here by black stars and the floe is rotated counter-clockwise compared against the map of 5 March, with a ship's heading of about 90°.
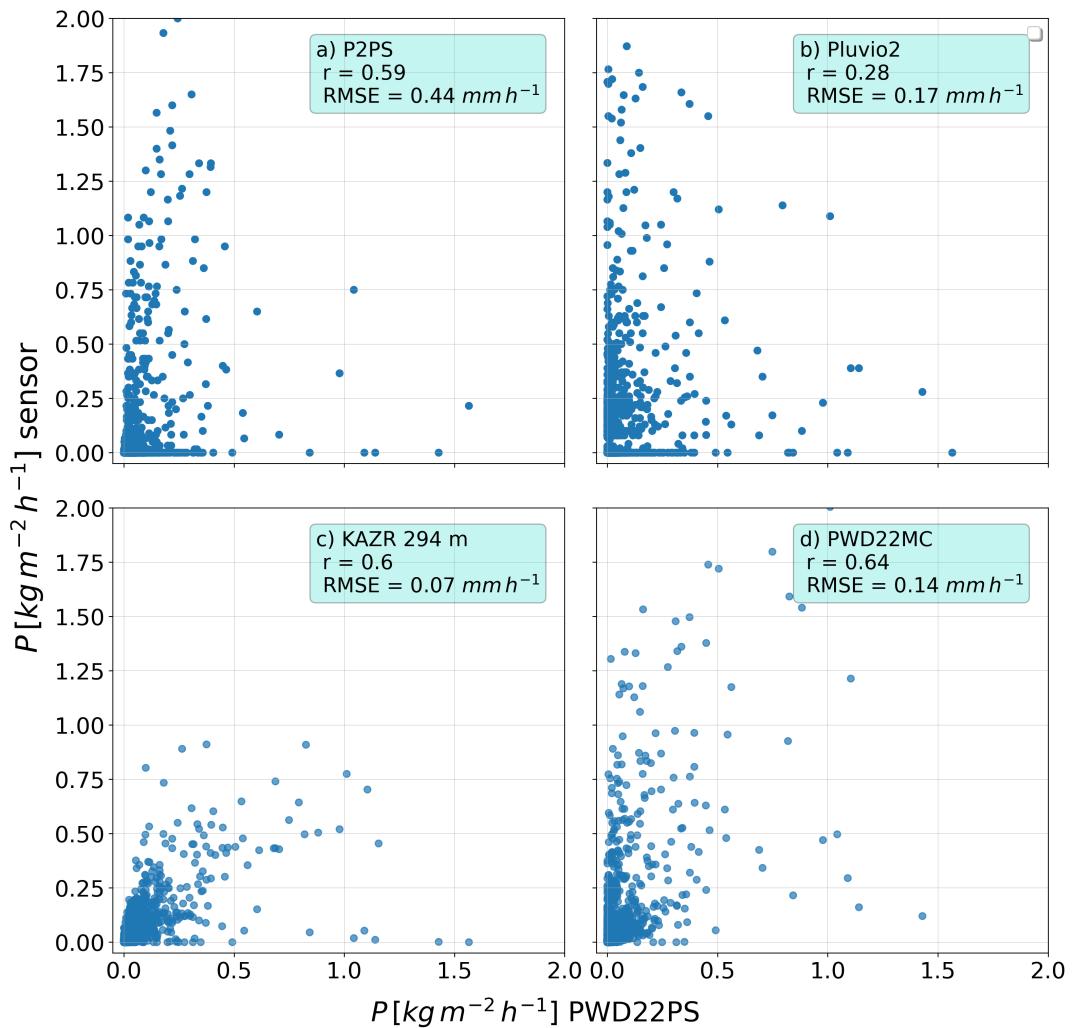
Figure B1. Scatter plots of PWD22PS snowfall rates vs. different sensor snowfall rates for the whole time period for a) P2PS, b) Pluvio2, c) KAZR and d) PWD22MC.

Author contributions. DNW took the lead in writing the manuscript and implemented comments and improvements from and findings from discussions with co-authors. DNW, JR, NK, ES, MO, IR, MS, MJ, DK, ARM, and SA conducted the snowpit measurements. DNW and ARM processed the snowpit data. The principal investigator for snow pits is MS. SH, MO, PI, MJ, SA, and RR conducted transect measurements. The principal investigator for the transect is PI. Transect data was post-processed by DNW. MDS, OGP, TU, MMF, and AK were responsible

815   for the met tower measurements, including SPC. The principal investigator for the tower is MDS. The University of Colorado / NOAA flux team processed the met tower data. The principal investigator for the SPC is MMF. DNW post-processed the SPC data. Team ICE created floe maps during the cruise. SH provided the code to correct the coordinate transformation from global to FloeNavi coordinates. The KAZR and all precipitation sensors were set up, maintained, and processed by the ARM team, and MDS was the principal investigator for the ARM involvement in MOSAiC. DNW conducted further KAZR and gauges processing. ML was involved in MOSAiC planning and supervised

820   the work during the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Data used in this manuscript was produced as part of the international Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAIC20192020 and the project ID: AWI_PS122_00. KAZR reflectivities and precipitation data were obtained from the Atmospheric Radiation Measurement (ARM) user facility, a U.S. Department of Energy (DOE)

825   Office of Science user facility managed by the Biological and Environmental Research Program. We thank Sergey Matrosov for valuable discussions about snowfall and for helping with his radar expertise. The authors thank all MOSAiC participants for contributing to the paper in one way or another: the MOSAiC logistics team, the ship's personnel, all the different team members (ICE, ATMOS, BGC, ECO, OCEAN), the ARM crew, the project coordination, and the MOSAiC project board. Met tower data was supported by the National Science Foundation OPP-1724551. MDS was supported by the DOE Atmospheric System Research Program (DE-SC0019251, DE-SC0021341).

References

IFS Documentation CY47R1 - Part IV: Physical Processes, no. 4 in IFS Documentation, ECMWF, https://doi.org/10.21957/cpmkqvhja, 2020.

Ackley, S., Lange, M., and Wadhams, P.: Snow cover effects on Antarctic sea ice thickness., 1990.

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung: Polar Research and Supply Vessel POLARSTERN Operated by the Alfred-Wegener-Institute, Journal of large-scale research facilities, 2, https://doi.org/10.17815/jlsrf-3-163, 2017.

Bagnold, R. A.: The physics of blown sand and desert dunes / by R.A. Bagnold, William Morrow New York, 1941.

Bartholomew, M.: Weighing Bucket Rain Gauge Instrument Handbook., Tech. rep., ed. by Robert Stafford, ARM user facility. DOE/SC-ARM-TR-232., 2020a.

Bartholomew, M.: Laser Disdrometer Instrument Handbook., Tech. rep., ed. by Robert Stafford, ARM user facility. DOE/SC-ARM-TR-137., 2020b.

Batrak, Y. and Müller, M.: On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice, Nature Communications, 10, 1–8, https://doi.org/10.1038/s41467-019-11975-3, 2019.

Battaglia, A., Rustemeier, E., Tokay, A., Blahak, U., and Simmer, C.: PARSIVEL Snow Observations: A Critical Assessment, Journal of Atmospheric and Oceanic Technology, 27, 333–344, https://doi.org/10.1175/2009JTECHA1332.1, 2010.

Blackford, J. R.: Sintering and microstructure of ice: a review, Journal of Physics D: Applied Physics, 40, R355–R385, https://doi.org/10.1088/0022-3727/40/21/r02, 2007.

Boudala, F. S., Isaac, G. A., Filman, P., Crawford, R., Hudak, D., and Anderson, M.: Performance of emerging technologies for Measuring Solid and Liquid Precipitation in Cold Climate as compared to the Traditional Manual Gauges, Journal of Atmospheric and Oceanic Technology, 34, 167–185, https://doi.org/10.1175/jtech-d-16-0088.1, 2016.

Budd, W. F., Dingle, W. R. J., and Radok, U.: The Byrd Snow Drift Project: Outline and Basic Results, pp. 71–134, American Geophysical Union (AGU), https://doi.org/https://doi.org/10.1029/AR009p0071, 1966.

Cabaj, A., Kushner, P. J., Fletcher, C. G., Howell, S., and Petty, A. A.: Constraining Reanalysis Snowfall Over the Arctic Ocean Using CloudSat Observations, Geophysical Research Letters, 47, https://doi.org/10.1029/2019gl086426, 2020.

Calonne, N., Richter, B., Löwe, H., Cetti, C., Schure, J. t., Herwijnen, A. V., Fierz, C., Jaggi, M., and Schneebeli, M.: The RHOSSA campaign: multi-resolution monitoring of the seasonal evolution of the structure and mechanical stability of an alpine snowpack, The Cryosphere, 14, 1829–1848, https://doi.org/10.5194/tc-14-1829-2020, 2020.

Chen, Y.: COVID-19 Pandemic Imperils Weather Forecast, Geophysical Research Letters, 47, e2020GL088 613, https://doi.org/10.1029/2020gl088613, 2020.

Chung, Y.-C., Bélair, S., and Mailhot, J.: Blowing Snow on Arctic Sea Ice: Results from an Improved Sea Ice–Snow–Blowing Snow Coupled System, Journal of Hydrometeorology, 12, 678–689, https://doi.org/10.1175/2011JHM1293.1, 2011.

Clifton, A., Rüedi, J.-D., and Lehning, M.: Snow saltation threshold measurements in a drifting-snow wind tunnel, Journal of Glaciology, 52, 585–596, https://doi.org/10.3189/172756506781828430, 2006.

Colbeck, S.: Sintering in a dry snow cover, Journal of Applied Physics, 84, 4585–4589, https://doi.org/10.1063/1.368684, 1998.

Colbeck, S., of Engineers, U. S. A. C., Research, C. R., and (U.S.), E. L.: A Review of Sintering in Seasonal Snow, AD-a335 556, U.S. Army Cold Regions Research and Engineering Laboratory, https://books.google.ch/books?id=D_CHKAAACAAJ, 1997.

Comola, F. and Lehning, M.: Energy- and momentum-conserving model of splash entrainment in sand and snow saltation, Geophysical Research Letters, 44, 1601–1609, https://doi.org/10.1002/2016GL071822, 2017.

Comola, F., Kok, J. F., Gaume, J., Paterna, E., and Lehning, M.: Fragmentation of wind-blown snow crystals, Geophysical Research Letters, 44, 4195–4203, https://doi.org/10.1002/2017GL073039, 2017.

870 Déry, S. J. and Tremblay, L.-B.: Modeling the Effects of Wind Redistribution on the Snow Mass Budget of Polar Sea Ice, Journal of Physical Oceanography, 34, 258–271, https://doi.org/10.1175/1520-0485(2004)034<0258:MTEOWR>2.0.CO;2, 2004.

Déry, S. J. and Yau, M. K.: A climatology of adverse winter-type weather events, Journal of Geophysical Research: Atmospheres, 104, 16 657–16 672, https://doi.org/10.1029/1999jd900158, 1999.

Déry, S. J. and Yau, M. K.: Large-scale mass balance effects of blowing snow and surface sublimation, Journal of Geophysical Research: 875 Atmospheres (1984–2012), 107, ACL 8–1–ACL 8–17, https://doi.org/10.1029/2001jd001251, 2002.

Déry, S. J., Taylor, P. A., and Xiao, J.: The Thermodynamic Effects of Sublimating, Blowing Snow in the Atmospheric Boundary Layer, Boundary-Layer Meteorology, 89, 251–283, https://doi.org/10.1023/a:1001712111718, 1998.

Essery, R., Li, L., and Pomeroy, J.: A distributed model of blowing snow over complex terrain, Hydrological Processes, 13, 2423–2438, https://doi.org/10.1002/(sici)1099-1085(199910)13:14/15<2423::aid-hyp853>3.0.co;2-u, 1999.

880 Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., Mcclung, D., Nishimura, K., Satyawali, P., and Sokratov, S.: The 2008 International Classification of Seasonal Snow on the Ground, p. 579–580, 2008.

Filhol, S. and Sturm, M.: Snow bedforms: A review, new data, and a formation model, Journal of Geophysical Research: Earth Surface, 120, 1645–1669, https://doi.org/https://doi.org/10.1002/2015JF003529, 2015.

Goodison, B., Louie, P., and Yang, D.: WMO Solid Precipitation Measurement Intercomparison, World Meteorological Organization-885 Publications-WMO TD, 67, 1998.

Haberkorn, A.: European Snow Booklet – an Inventory of Snow Measurements in Europe, https://doi.org/http://dx.doi.org/10.16904/envidat.59, 2019.

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, 890 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.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Holtsmark, B. E.: Insulating Effect of a Snow Cover on the Growth of Young Sea Ice, Arctic, 8, 60–65, http://www.jstor.org/stable/40506652, 895 1955.

Iacozza, J. and Barber, D. G.: An examination of the distribution of snow on sea-ice, Atmosphere-Ocean, 37, 21–51, https://doi.org/10.1080/07055900.1999.9649620, 1999.

Jonas, T., Marty, C., and Magnusson, J.: Estimating the snow water equivalent from snow depth measurements in the Swiss Alps, Journal of Hydrology, 378, 161–167, https://doi.org/10.1016/j.jhydrol.2009.09.021, 2009.

900 Jun, S.-Y., Ho, C.-H., Jeong, J.-H., Choi, Y.-S., and Kim, B.-M.: Recent changes in winter Arctic clouds and their relationships with sea ice and atmospheric conditions, Tellus A, 68, 29 130, https://doi.org/10.3402/tellusa.v68.29130, 2016.

Jutras, M., Vancoppenolle, M., Lourenço, A., Vivier, F., Carnat, G., Madec, G., Rousset, C., and Tison, J.-L.: Thermodynamics of slush and snow–ice formation in the Antarctic sea-ice zone, Deep Sea Research Part II: Topical Studies in Oceanography, 131, 75–83, https://doi.org/10.1016/j.dsr2.2016.03.008, 2016.

905 King, J., Howell, S., Brady, M., Toose, P., Derksen, C., Haas, C., and Beckers, J.: Local-scale variability of snow density on Arctic sea ice, The Cryosphere, 14, 4323–4339, https://doi.org/10.5194/tc-14-4323-2020, 2020.

King, J. C., Pomeroy, J. W., Gray, D. M., and Fierz, C.: Snow accumulation, in: Snow and Climate, edited by Armstrong, R. L. and Brun, E., chap. 3.4, pp. 83–92, Cambridge University Press, 2008.

Kochendorfer, J., Rasmussen, R., Wolff, M., Baker, B., Hall, M. E., Meyers, T., Landolt, S., Jachcik, A., Isaksen, K., Brækkan, R., and
910 Leeper, R.: The quantification and correction of wind-induced precipitation measurement errors, Hydrology and Earth System Sciences, 21, 1973–1989, https://doi.org/10.5194/hess-21-1973-2017, 2017.

Kochendorfer, J., Nitu, R., Wolff, M., Mekis, E., Rasmussen, R., Baker, B., Earle, M. E., Reverdin, A., Wong, K., Smith, C. D., Yang, D., Roulet, Y.-A., Meyers, T., Buisan, S., Isaksen, K., Brækkan, R., Landolt, S., and Jachcik, A.: Testing and development of transfer functions for weighing precipitation gauges in WMO-SPICE, Hydrology and Earth System Sciences, 22, 1437–1452, https://doi.org/10.5194/hess-
915 22-1437-2018, 2018.

Krumpen, T., Birrien, F., Kauker, F., Rackow, T., Albedyll, L. v., Angelopoulos, M., Belter, H. J., Bessonov, V., Damm, E., Dethloff, K., Haapala, J., Haas, C., Harris, C., Hendricks, S., Hoelemann, J., Hoppmann, M., Kaleschke, L., Karcher, M., Kolabutin, N., Lei, R., Lenz, J., Morgenstern, A., Nicolaus, M., Nixdorf, U., Petrovsky, T., Rabe, B., Rabenstein, L., Rex, M., Ricker, R., Rohde, J., Shimanchuk, E., Singha, S., Smolyanitsky, V., Sokolov, V., Stanton, T., Timofeeva, A., Tsamados, M., and Watkins, D.: The MOSAiC ice floe: sediment-
920 laden survivor from the Siberian shelf, The Cryosphere, 14, 2173–2187, https://doi.org/10.5194/tc-14-2173-2020, 2020.

Kyrouac, J. and Holdridge, D.: Surface Meteorological Instrumentation (PWD)., atmospheric Radiation Measurement (ARM) user facility., 2019.

Lange, M. A. and Eicken, H.: The sea ice thickness distribution in the northwestern Weddell Sea, Journal of Geophysical Research: Oceans, 96, 4821–4837, https://doi.org/https://doi.org/10.1029/90JC02441, 1991.

925 Lehning, M., Doorschot, J., and Bartelt, P.: A snowdrift index based on SNOWPACK model calculations, Annals of Glaciology, 31, 382–386, https://doi.org/10.3189/172756400781819770, 2000.

Leonard, K. C. and Maksym, T.: The importance of wind-blown snow redistribution to snow accumulation on Bellingshausen Sea ice, Annals of Glaciology, 52, 271–278, https://doi.org/10.3189/172756411795931651, 2011.

Leonard, K. C., Tremblay, L., MacAyeal, D. R., and Jacobs, S. S.: Interactions of wind-transported snow with a rift in the Ross Ice Shelf,
930 Antarctica, Geophysical Research Letters, 35, https://doi.org/10.1029/2007gl033005, 2008.

Li, L. and Pomeroy, J. W.: Estimates of Threshold Wind Speeds for Snow Transport Using Meteorological Data, Journal of Applied Meteorology, 36, 205–213, https://doi.org/10.1175/1520-0450(1997)036<0205:eotwsf>2.0.co;2, 1997.

Lindenmaier, I., Nelson, D., Isom, B., Hardin, J., Matthews, A., Wendler, T., and Castro, V.: Ka ARM Zenith Radar (KAZRCFRGE), https://doi.org/10.5439/1498936, atmospheric Radiation Measurement (ARM) user facility., 2019.

935 Löffler-Mang, M. and Joss, J.: An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors, Journal of Atmospheric and Oceanic Technology, 17, 130 – 139, https://doi.org/10.1175/1520-0426(2000)017<0130:AODFMS>2.0.CO;2, 2000.

Löwe, H. and Herwijnen, A. v.: A Poisson shot noise model for micro-penetration of snow, Cold Regions Science and Technology, 70, 62–70, https://doi.org/10.1016/j.coldregions.2011.09.001, 2012.

Matrosov, S. Y.: Modeling Backscatter Properties of Snowfall at Millimeter Wavelengths, Journal of the Atmospheric Sciences, 64, 1727–
1736, https://doi.org/10.1175/jas3904.1, 2007.

Matrosov, S. Y., Shupe, M. D., and Djalalova, I. V.: Snowfall Retrievals Using Millimeter-Wavelength Cloud Radars, Journal of Applied
Meteorology and Climatology, 47, 769–777, https://doi.org/10.1175/2007jamc1768.1, 2008.

Maykut, G. A. and Untersteiner, N.: Some results from a time-dependent thermodynamic model of sea ice, Journal of Geophysical Research
(1896-1977), 76, 1550–1575, https://doi.org/10.1029/JC076i006p01550, 1971.

McIlhattan, E. A., Kay, J. E., and L'Ecuyer, T. S.: Arctic Clouds and Precipitation in the Community Earth System Model Version 2, Journal
of Geophysical Research: Atmospheres, 125, e2020JD032 521, https://doi.org/10.1029/2020jd032521, 2020.

Merkouriadi, I., Liston, G. E., Graham, R. M., and Granskog, M. A.: Quantifying the Potential for Snow-Ice Formation in the Arc-
tic Ocean, Geophysical Research Letters, 47, e2019GL085 020, https://doi.org/https://doi.org/10.1029/2019GL085020, e2019GL085020
2019GL085020, 2020.

METEK: 3D Ultrasonic Anemometer uSonic-3 Cage MP Manual, Tech. rep., 2019.

Nicolaus, M., Hoppmann, M., Arndt, S., Hendricks, S., Katlein, C., Nicolaus, A., Rossmann, L., Schiller, M., and Schwegmann, S.: Snow
Depth and Air Temperature Seasonality on Sea Ice Derived From Snow Buoy Measurements, Frontiers in Marine Science, 8, 377,
https://doi.org/10.3389/fmars.2021.655446, 2021.

Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., de Rosnay, P., Zhu, C., Wang, W., Senan, R., and Arduini, G.: Evalu-
ation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations,
The Cryosphere, 13, 2221–2239, https://doi.org/10.5194/tc-13-2221-2019, 2019.

Petty, A. A., Webster, M., Boisvert, L., and Markus, T.: The NASA Eulerian Snow on Sea Ice Model (NESOSIM) v1.0: initial model
development and analysis, Geoscientific Model Development, 11, 4577–4602, https://doi.org/10.5194/gmd-11-4577-2018, 2018.

Proksch, M., Löwe, H., and Schneebeli, M.: Density, specific surface area, and correlation length of snow measured by high-resolution
penetrometry, Journal of Geophysical Research: Earth Surface, 120, 346–362, https://doi.org/10.1002/2014JF003266, 2015.

Sato, T., Kimura, T., Ishimaru, T., and Maruyama, T.: Field test of a new snow-particle counter (SPC) system, Annals of Glaciology, 18,
149–154, https://doi.org/10.3189/S0260305500011411, 1993.

Scarchilli, C., Frezzotti, M., Grigioni, P., Silvestri, L. D., Agnoletto, L., and Dolci, S.: Extraordinary blowing snow transport events in East
Antarctica, Climate Dynamics, 34, 1195–1206, https://doi.org/10.1007/s00382-009-0601-0, 2009.

Schmidt, R. A.: Threshold Wind-Speeds and Elastic Impact in Snow Transport, Journal of Glaciology, 26, 453–467,
https://doi.org/10.3189/S0022143000010972, 1980.

Schneebeli, M. and Johnson, J. B.: A constant-speed penetrometer for high-resolution snow stratigraphy, Annals of Glaciology, 26, 107–111,
https://doi.org/10.3189/1998AoG26-1-107-111, 1998.

Schneebeli, M., Pielmeier, C., and Johnson, J. B.: Measuring snow microstructure and hardness using a high resolution penetrometer, Cold
Regions Science and Technology, 30, 101 – 114, https://doi.org/https://doi.org/10.1016/S0165-232X(99)00030-0, 1999.

Serreze, M. C. and Hurst, C. M.: Representation of Mean Arctic Precipitation from NCEP–NCAR and ERA Reanalyses, Journal of Climate,
13, 182–201, https://doi.org/10.1175/1520-0442(2000)013<0182:romapf>2.0.co;2, 2000.

Shi, Y.: Laser Disdrometer (LD).), https://doi.org/10.5439/1498731, atmospheric Radiation Measurement (ARM) user facility., 2019.

Sturm, M. and Holmgren, J.: An Automatic Snow Depth Probe for Field Validation Campaigns, Water Resources Research, 54, 9695–9701,
https://doi.org/10.1029/2018wr023559, 2018.

Sturm, M. and Massom, R. A.: Snow in the sea ice system: friend or foe?, chap. 3, pp. 65–109, John Wiley and Sons, Ltd, https://doi.org/10.1002/9781118778371.ch3, 2016.

Sturm, M., Morris, K., and Massom, R.: Antarctic Sea Ice: Physical Processes, Interactions and Variability, pp. 1–18, https://doi.org/10.1029/ar074p0001, 1998a.

980 Sturm, M., Morris, K., and Massom, R.: The Winter Snow Cover of the West Antarctic Pack Ice: Its Spatial and Temporal Variability, pp. 1–18, American Geophysical Union (AGU), https://doi.org/https://doi.org/10.1029/AR074p0001, 1998b.

Sturm, M., Perovich, D. K., and Holmgren, J.: Thermal conductivity and heat transfer through the snow on the ice of the Beaufort Sea, Journal of Geophysical Research: Oceans, 107, SHE 19–1–SHE 19–17, https://doi.org/10.1029/2000JC000409, 2002.

Sugiura, K., Yang, D., and Ohata, T.: Systematic error aspects of gauge-measured solid precipitation in the Arctic, Barrow, Alaska, Geophys-
985 ical Research Letters, 30, https://doi.org/10.1029/2002gl015547, 2003.

Sugiura, K., Ohata, T., Yang, D., Sato, T., and Sato, A.: Application of a snow particle counter to solid precipitation measurements under Arctic conditions, Cold Regions Science and Technology, 58, 77–83, https://doi.org/10.1016/j.coldregions.2009.03.010, 2009.

Toyota, T., Massom, R., Tateyama, K., Tamura, T., and Fraser, A.: Properties of snow overlying the sea ice off East Antarctica in late winter, 2007, Deep Sea Research Part II: Topical Studies in Oceanography, 58, 1137 – 1148,
990 https://doi.org/https://doi.org/10.1016/j.dsr2.2010.12.002, antarctic Sea Ice Research during the International Polar Year 2007-2009, 2011.

Trujillo, E. and Lehning, M.: Theoretical analysis of errors when estimating snow distribution through point measurements, The Cryosphere, 9, 1249–1264, https://doi.org/10.5194/tc-9-1249-2015, 2015.

Trujillo, E., Leonard, K., Maksym, T., and Lehning, M.: Changes in snow distribution and surface topography following a snowstorm on Antarctic sea ice, Journal of Geophysical Research: Earth Surface, 121, 2172–2191, https://doi.org/10.1002/2016JF003893, 2016.

995 Uttal, T., Curry, J. A., McPhee, M. G., Perovich, D. K., Moritz, R. E., Maslanik, J. A., Guest, P. S., Stern, H. L., Moore, J. A., Turenne, R., Heiberg, A., Serreze, M. C., Wylie, D. P., Persson, O. G., Paulson, C. A., Halle, C., Morison, J. H., Wheeler, P. A., Makshtas, A., Welch, H., Shupe, M. D., Intrieri, J. M., Stamnes, K., Lindsey, R. W., Pinkel, R., Pegau, W. S., Stanton, T. P., and Grenfeld, T. C.: Surface Heat Budget of the Arctic Ocean, Bulletin of the American Meteorological Society, 83, 255–276, https://doi.org/10.1175/1520-0477(2002)083<0255:SHBOTA>2.3.CO;2, 2002.

1000 Vaisala: Present Weather Detector PWD22 User's Guide, Tech. rep., 2004.

Vaisala: HUMICAP®Humidity and Temperature Transmitter Series HMT330, Tech. rep., 2009.

Wagner, D. N., Jaggi, M., Macfarlane, A. R., Arndt, S., Krampe, D., Regnery, J., Oggier, M., Kolabutin, N., Raphael, I., Shimanchuck, E., and Schneebeli, M.: Snow water equivalent retrievals from SnowMicroPen data from MOSAiC Leg 1 - Leg 3, https://doi.org/10.1594/PANGAEA.927460, pANGAEA, 2021.

1005 Wang, C., Graham, R. M., Wang, K., Gerland, S., and Granskog, M. A.: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution, The Cryosphere, 13, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019, 2019a.

Wang, D., Bartholomew, M., and Cromwell, E.: Weighing Bucket Precipitation Gauge (WBPLUVIO2)., atmospheric Radiation Measurement (ARM) user facility., 2019b.

1010 Webster, M. A., DuVivier, A. K., Holland, M. M., and Bailey, D. A.: Snow on Arctic Sea Ice in a Warming Climate as Simulated in CESM, Journal of Geophysical Research: Oceans, 126, e2020JC016 308, https://doi.org/10.1029/2020jc016308, 2021.

Weiss, A. I., King, J., Lachlan-Cope, T., and Ladkin, R.: On the effective aerodynamic and scalar roughness length of Weddell Sea ice, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD015949, 2011.

Widener, K., Bharadwaj, N., and Johnson, K.: Ka-Band ARM Zenith Radar (KAZR) Handbook., Tech. rep., aRM user facility. DOE/SC-
    ARM/TR-106., 2012.

Wong, K., Fischer, A., and Nitu, R.: Measurement of Precipitation at AWS in Canada: Configuration, Challenges and Alternative Approaches,
    2012b.

Wong, K. C.: Performance of Several Present Weather Sensors as Precipitation Gauges, in: WMO Technical Conference on Meteorological
    and Environmental Instruments and Methods of Observation (Teco 2012), 2012a.