Supplemental material to ”Spatio-temporal variability of small-scale leads based on winter sea ice surface temperatures”

Linda Thielke[1], Gunnar Spreen[1*], Marcus Huntemann[1], Dmitrii Murashkin[2,1]

[1]University of Bremen, Institute of Environmental Physics, Bremen, 28359, Germany

[2]German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Bremen, Germany

[*]Corresponding author: gunnar.spreen@uni-bremen.de

Text S1 - Detailed description of the iterative threshold selection

The surface temperature distribution has its major mode towards the colder part of the distribution and a smaller secondary mode along the tail towards the warmest temperatures which is caused, e.g., by leads. The initial threshold to discriminate leads from the remaining sea ice is the middle value of the temperature distribution, i.e., the average of the minimum and maximum value. Due to the long tail towards the warmer temperatures, it is ensured that the initial threshold is on the warmer side of the major mode so that the iteration can converge towards a minimum between the two modes. Starting from the initial threshold, the threshold is adjusted iteratively based on the new ”lead” mask, defined by the threshold of each iteration, for the temperature array until it reaches the final temperature threshold. The threshold is calculated newly in every loop from the mean of the ”lead” (all values larger than the threshold) and ”sea ice” (all values smaller than the previous threshold) temperatures based on the current threshold. The iteration stop criterion is achieved when the temperature threshold change between one iteration to the next is within the tolerance of 0.02 K, which corresponds to the precision of the IR camera. The main steps are shown in the flowchart in Figure S1, and the Python 3 code is shared below in this section. For three flights (20191224_01, 20191225_01, 20200108_01), the threshold did not converge to a reasonable value; therefore, the tolerance had to be increased to 0.8 K. For the cases with this larger allowed tolerance in the difference, the obtained threshold still results in a reasonable lead classification, i.e., can be

confirmed by the manual classification (see next paragraph).

Evaluation using manual thresholds

We use manual thresholds as a comparison for the reliability of the iterative method. The authors performed the manual threshold selection based on the minima in the distribution and visual approval of the classified map compared to the surface temperature map. The manual selection is a rather arbitrary and conservative choice but it can be used for the evaluation of the automatic, iterative classification method described in the last paragraph. The thresholds are determined for each flight individually because the surface temperature values and their distributions change from flight to flight. The manual selection has, in some cases, warmer thresholds but this does not significantly influence the resulting lead area fractions (see Table S1 below). Although the manual temperature threshold is 1.97 K higher than the one from the automatic and reproducible method, its derived lead area fraction is only 0.1% lower than for the iterative method. The threshold difference does not significantly affect the lead area fraction because it is in the minimum of the surface temperature distribution. The small difference demonstrates that the automatic method aligns well with the manually defined thresholds and the resulting lead classification. We chose results from the automatic method for the discussion in this study because it is reproducible and can also be applied to future flight campaigns.

Code for the iterative threshold selection in Python 3

```
def iterative_classification(Ts_arr,cr=0.02):
    """
    Input:
        - Ts_arr: surface temperatures (2d array)
        - cr: loop limit (float)
    Output:
        - thr: temperature threshold (float)
    """

    thr = 0.5*(np.nanmin(Ts_arr)+np.nanmax(Ts_arr)) # initial threshold
    dT=1

    while dT > cr:
        g = (Ts_arr>thr) # mask with threshold applied
        thr_new = 0.5*(np.nanmean(Ts_arr[g])+np.nanmean(Ts_arr[~g]))
```

```
    dT = np.abs(thr-thr_new)
    thr=thr_new

return thr
```

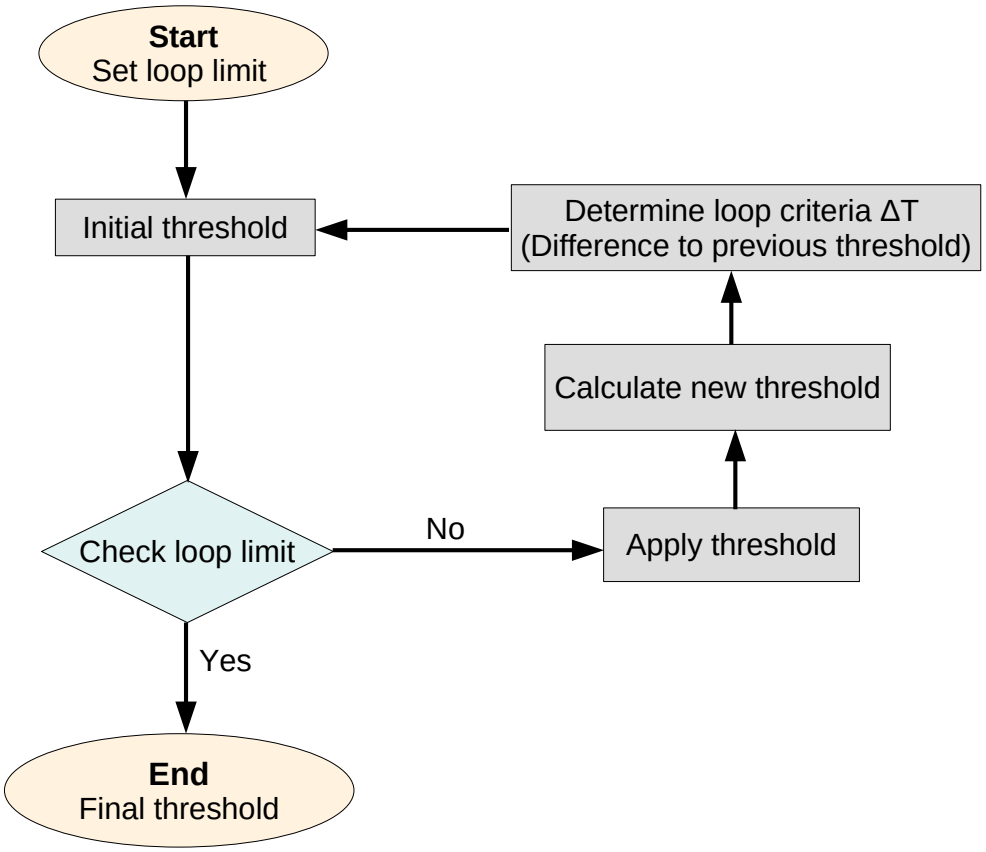

Figure S1. Flowchart for the processing of the iterative threshold selection.

Main processing steps for the automatic iterative threshold selection to determine a temperature threshold for lead classification.

Table S1. All temperature thresholds used for the lead classification.

List for all flights with the respective temperature thresholds used for the lead classification. In column 4 for the iterative method and, for comparison, in column 5 the threshold derived with the manual approach. The flights can be identified by the expedition-specific *Device Operation* and Flight ID that includes the date information (YYYYMMDD_counter).

Device Operation	**Flight ID**	**Flight type**	**Iterative / K**	Manual / K (not used)
PS122-1_2-57	20191002_01	CO	266.83	267.42
PS122-1_2-167	20191020_01	CO	257.75	258.05
PS122-1_5-9	20191029_01	L-Site	259.58	261.98
PS122-1_6-11	20191105_01	CO	247.34	252.15
PS122-1_7-24	20191112_01	L-Site	256.13	260.17
PS122-1_7-25	20191112_02	CO	252.30	250.50
PS122-1_8-23	20191119_01	CO	259.60	258.81
PS122-1_9-98	20191130_01	CO	245.87	242.94
PS122-1_10-78	20191206_01	L-Site	250.82	254.41
PS122-2_17-98	20191224_01	CO	241.47	242.70
PS122-2_17-99	20191225_01	CO	241.63	243.06
PS122-2_17-101	20191228_01	CO	239.93	238.44
PS122-2_18-7	20191230_01	L-Site	251.50	250.64
PS122-2_19-44	20200107_01	CO	238.94	248.30
PS122-2_19-45	20200107_02	L-Site	245.94	251.76
PS122-2_19-46	20200108_01	L-Site grid	242.97	244.41
PS122-2_19-52	20200108_03	L-Site grid	249.28	251.10
PS122-2_19-53	20200108_04	L-Site grid	247.38	249.35
PS122-2_20-52	20200116_01	CO	246.53	247.86
PS122-2_20-53	20200116_02	L-Site	243.12	246.26
PS122-2_21-41	20200121_01	CO	240.15	242.87
PS122-2_21-77	20200123_01	L-Site	244.30	246.58
PS122-2_21-78	20200123_02	Lead-Event	244.50	244.64
PS122-2_21-122	20200125_01	Lead-Event	243.13	248.39
PS122-2_22-16	20200128_01	CO	237.23	239.48
PS122-2_22-97	20200202_01	L-Site (partly)	241.51	244.10
PS122-2_23-14	20200204_01	CO	245.48	247.50
PS122-2_23-109	20200209_01	L-Site	241.28	248.44
PS122-2_24-31	20200212_01	CO	234.55	234.97
PS122-2_25-7	20200217_01	L-Site (partly)	231.66	237.61
PS122-2_25-8	20200217_02	CO	239.11	234.26
PS122-3_29-49	20200227_01	CO	236.17	235.14
PS122-3_32-70	20200321_01	CO	241.78	244.63
PS122-3_32-71	20200321_02	L-Site	242.64	244.42
PS122-3_37-63	20200423_01	CO	256.09	259.52

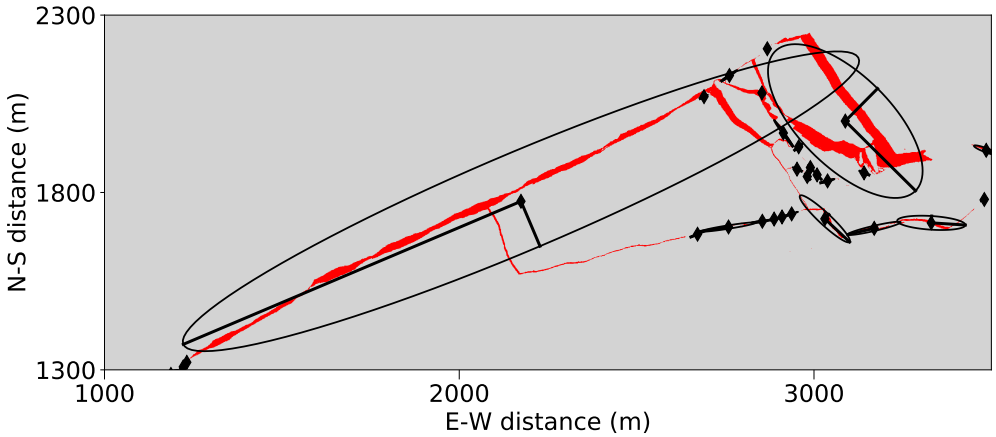

Figure S2. Visualisation of the lead segmentation showing an example for a set of segments.

Selection of the lead classification map for January 23, 2020, showing the ellipse geometry for several segments which are automatically determined and might overlap.

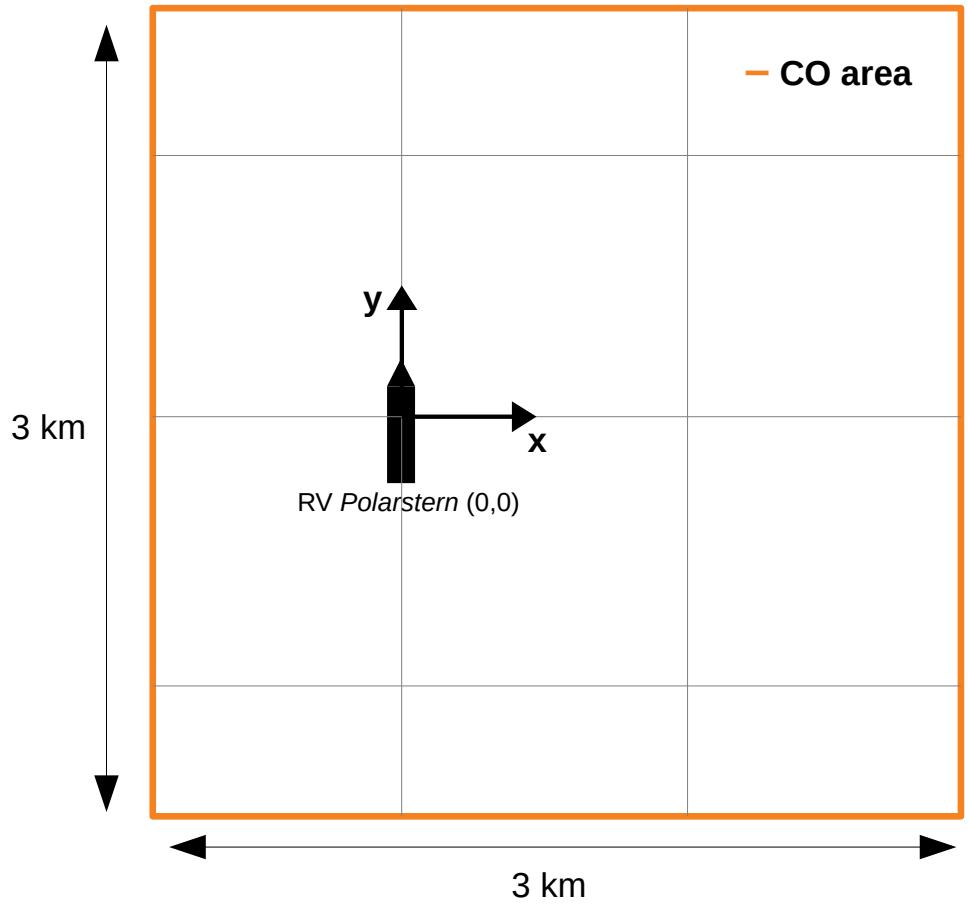

Figure S3. Definition of the Central Observatory (CO) area.

This schematic shows the main CO area with an extent of 3x3 km and the orientation of RV *Polarstern* in the relative coordinates whereas the center of RV *Polarstern* is the point (0,0). In the CO area, most of the measurement sites were located. The edge points are: xmin, xmax, ymin, ymax = −1000 m, 2000 m, −1500 m, 1500 m. The arrows indicate the orientation of the coordinate system for the relative coordinates.

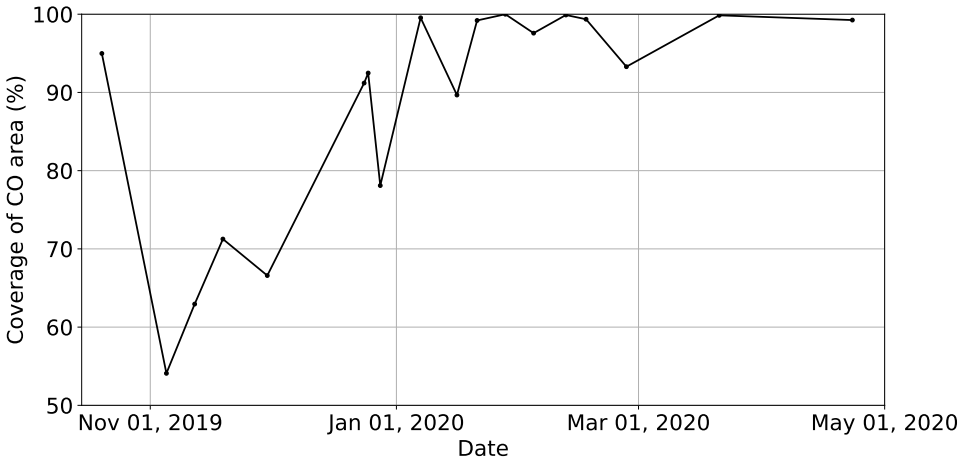

Figure S4. Data coverage of the Central Observatory (CO) area for each flight.

Time series of percentage of data points available in the CO area (as defined in Figure S3) for all local flights.

Text S2 - Additional data on the lead orientations

We find prevailing orientations of −80°, −10°, 30°, or 60° (Table S2). For the description, we focus on three examples from January 7, 2020, January 28, 2020, and March 21, 2020 (Figure 8 in main text). We identify modes of the lead orientation probability distributions of the orientation from −90 to 90°, binned in 5° steps (Figure S5). We constrain our data to elongated ellipse shapes with an axis ratio (major/minor) of at least two. Additionally, we compare the leads of all widths with leads of a width of more than 3 m which is consistent with the valid range of the power law. With the constraint of the axis ratio, the data are reduced to 89% of the full data set. With the minimum width of 3 m the number of leads is reduced to 21% of the complete data set. Starting with the case in March (Figure 8C), there is no major peak in the lead orientation distribution for all lead width. With only wider leads (≥3 m), the distribution of orientation angles is modified to a preferred direction at 35°, but still most orientations are present and not a clear prevailing orientation can be identified. On January 28 (Figure 8B) we have one prominent orientation at −35° (all leads) which is even more emphasised for leads with the minimum width of 3 m (slightly shifted to −30°). There is a second minor peak at 40°.For the case on January 7, 2020 (Figure 8A) we identify one clear main direction of −10°.

Table S2. Main angles of lead orientation per flight.

List of nine selected flights with the main lead orientation angle (0° = north-south line) for all data (column 2) and the data restricted to the lead width ≥ 3 m (column 3). The flights are identified with their Flight ID (column 1).

Flight ID	Peak (all)	Peak (≥ 3 m)
20191020_01	−80°/85°	−80°/90°
20191119_01	−70°	−75°
20191224_01	60°	65°
20191225_01	60°	60°
20200107_01	−10°	−10°
20200128_01	−35°	−30°
20200204_01	35°	30°
20200321_01	None	(35°)
20200423_01	None	−5°

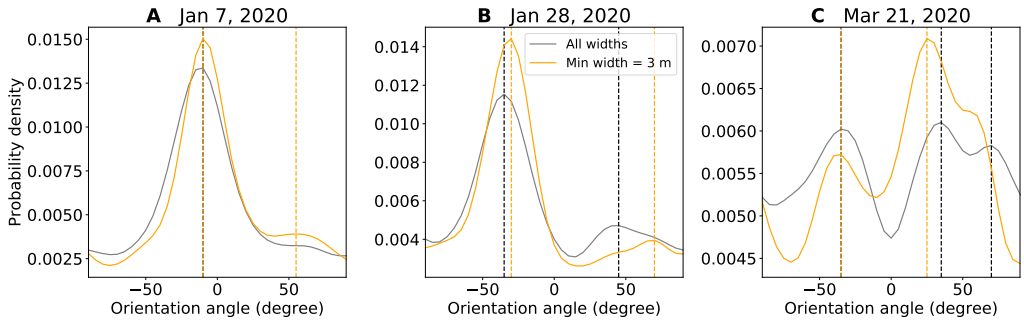

Figure S5. Distribution of lead orientation angles for three example flights.

Probability density functions and peaks of their modes (dashed lines) for the 3 example flights on January 8, January 28, and March 21, 2020, which are also shown in Figure 8 in the main text.

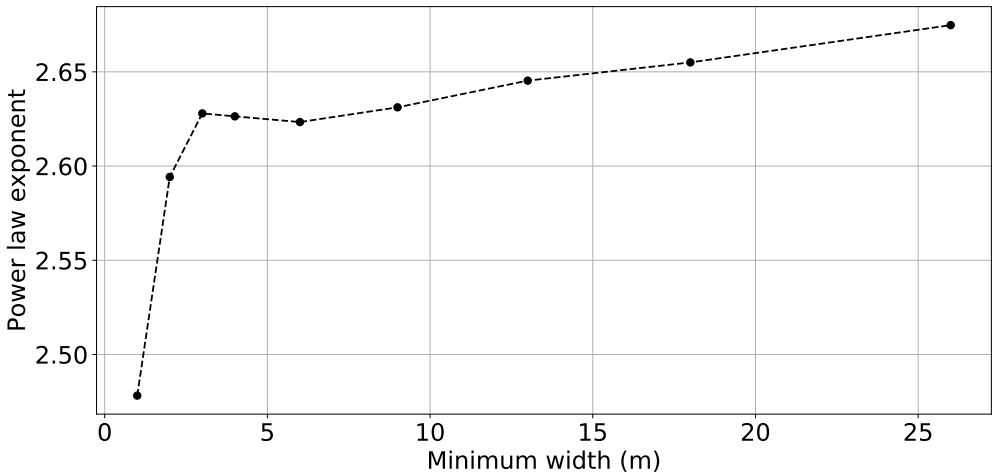

Figure S6. Dependence of the power law exponent on the minimum lead width.

Relation between power law exponent and applied minimum width for the linear fit in the log-log room of lead width against probability.