First of all, we would like to sincerely thank Torbjörn for the detailed and constructive comments.

We agree that the performance of semi-variogram analysis is sensitive to several methodological choices, including (a) the chosen lag bin widths (and whether they are constant), (b) the careful filtering of data, and (c) the choice of model to fit the empirical variogram. The comments are summarized into points (a)–(c), and we respond to each point below.

(a) Bin width selection:

In practice, bin widths can be defined in two main ways: using a constant value across all datasets, or using an adaptive value derived from the sampling geometry of each transect.

In the manuscript, we adopted constant bin width. Specifically, the bin width was set to 3 m, based on the spacing of the Magnaprobe measurements. Across all test sites, the median nearest-neighbour spacing of each transect varies between approximately 0.5 m and 3 m. The number of lag bins was chosen such that the bin width exceeded the median spacing, thereby ensuring a sufficient number of point pairs per bin while still resolving spatial variability up to the sill. Using a constant bin width also facilitates a consistent comparison of variogram parameters across different sites.

An alternative approach is to compute the median spacing for each transect individually and use this value to define the bin width. We show the results below.

For the selected transect 20191114_PS122-1_7-62_Sloop, we tested the sensitivity of the estimated effective range to the choice of lag-bin width using the unfiltered dataset. As shown in Fig.1(left), when applying constant bin widths of 1, 3, 5, and 10 m, the resulting effective ranges vary between 57 m and 68 m.

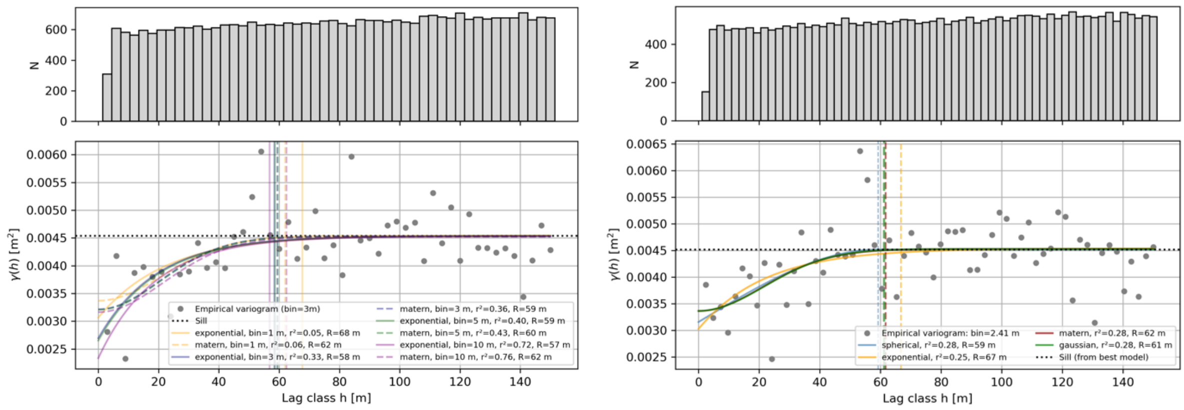

Figure 1. *Left: Empirical semi-variograms calculated using different bin widths. Right: Empirical semi-variograms calculated using an adaptive bin-width approach with different variogram fitting models. Model performance is evaluated using the coefficient of determination (r²). The effective range (R) is interpreted as the correlation length throughout this study.*

We further applied the adaptive bin-width approach, in which the bin width is set to the median nearest-neighbour spacing of the transect. For this transect, the median nearest-neighbour spacing is 2.41 m, yielding a bin width of 2.41 m. Under this configuration, the estimated effective ranges vary between 59 and 67 m across different models, see Fig.1(right).

To further assess the impact of bin choice on our results in the manuscript (Figs. 16 and 17 in the original preprint manuscript), we compared the correlation length between using a constant bin width of 3 m and an adaptive bin-width approach. As shown in Fig. 2, the estimated correlation lengths remain relatively stable across most sites under both cases, with the exception of Weddell Sea (FYI) floe 503. This consistency demonstrates that the choice of a 3 m bin width and the adaptive bin-width approach leads to the similar conclusion: the correlation length is positively related to ice thickness, which is one of the key findings of our study.

One exception is the Weddell Sea (FYI) floe 503, for which we further examined the impact of bin width on the estimated correlation length (Fig. 3). We find that only when using a bin width of 3 m with the exponential model does the semi-variogram exhibit a physically reasonable shape, with R^2 = 0.17. The pronounced sensitivity of the semi-variogram to bin width in this case is likely due to strong local heterogeneity and undulations caused by local non-stationarities in the snow-depth field over floe 503, where semi-variogram analysis can fail to find change points of slope, necessitating the use of more sophisticated methods (Moon et al., 2019).

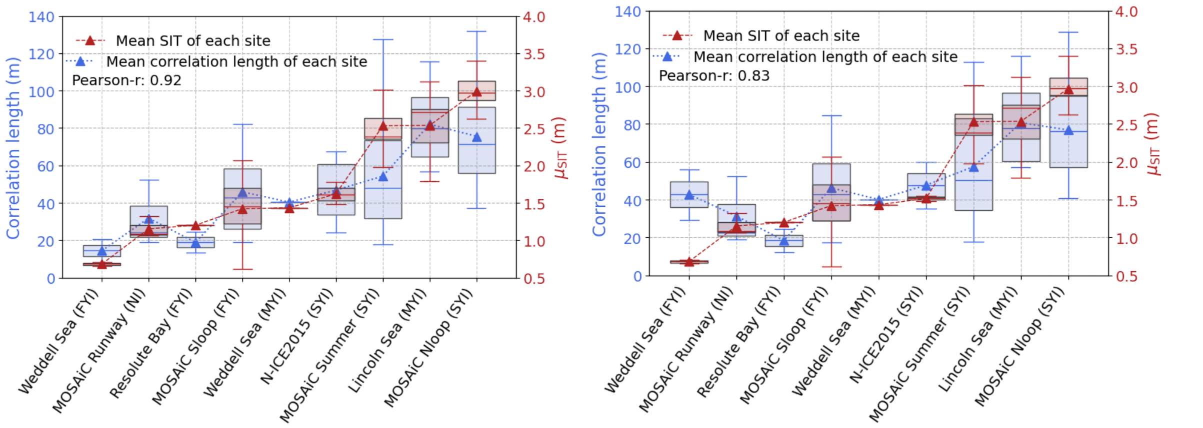

Figure 2. *Relationship between snow depth (SND) correlation length and sea-ice thickness (SIT). The correlation length is estimated using (left) a constant lag-bin width of* ***3 m*** *and (right) an* ***adaptive bin width*** *defined by the median nearest-neighbour spacing along each transect.*

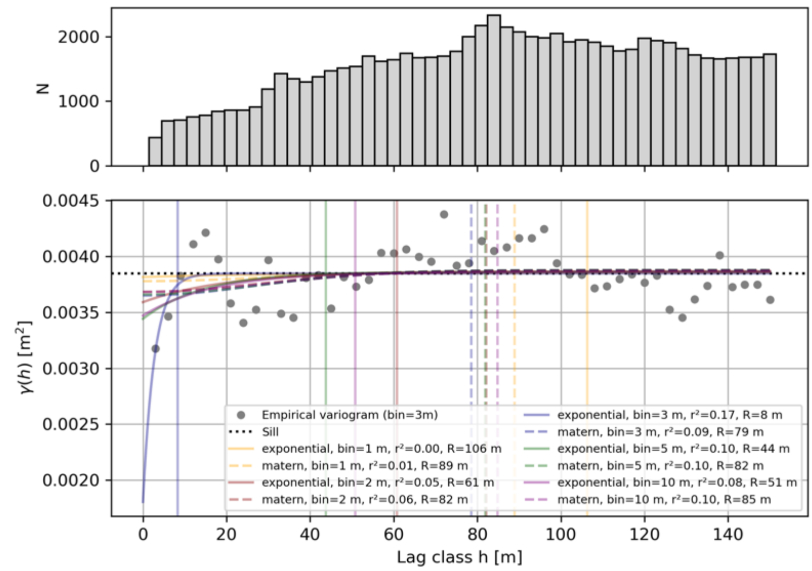

Figure 3. *Weddell Sea first-year ice (FYI) floe 503. Top: Number of data pairs per lag bin for a bin width of 3 m. Bottom: Empirical semi-variograms derived using different lag-bin widths. Model performance is evaluated using the coefficient of determination (r^2). The effective range (R) is interpreted as the correlation length throughout this study.*

In summary, we agree that lag-bin width is an important parameter for estimating correlation length. Lag-bin width should be selected in relation to the median spacing of each transect. While the adaptive strategy allows the bin width to respond directly to local sampling density, it results in variable bin widths across transects and test sites, thereby impeding inter-site comparisons. For this reason, and given the relatively narrow range of sampling spacings in our dataset, we adopted a constant bin width in the present analysis.

(b) Filtering of the data

We agree that careful data screening is essential. In this study, measurements associated with distances greater than 10 m were excluded, as these correspond to periods when the Magnaprobe was moved between sampling locations rather than actively measuring snow depth. In addition, measurements with snow depths less than 1 cm or equal to 120 cm which is the operational limits of the Magnaprobe were removed.

For the transect 20191114_PS122-1_7-62_Sloop, we note the presence of a high snow-depth value of 0.98 m. We do not consider this measurement to be spurious solely because it represents an extreme value. Instead, it may correspond to locally thick snow and therefore reflect genuine spatial variability in the snow-depth distribution. Consequently, we retained this sample in the analysis. We appreciate the reviewer raising the data quality. In the revised manuscript, we have added a note acknowledging that such extreme values may influence the estimated correlation length.

(c) Choice of model

We appreciate the reviewer recommending the Matérn model. We evaluated four variogram models: exponential, Matérn, Gaussian, and spherical. For each dataset, we calculated the coefficient r^2 to assess model performance. Overall, both Matérn and exponential models consistently provide the best fits across the datasets. While the Matérn model yields the highest r^2 in many cases, the improvement relative to the exponential model is typically small (on the order of ~0.01), see Fig.4.

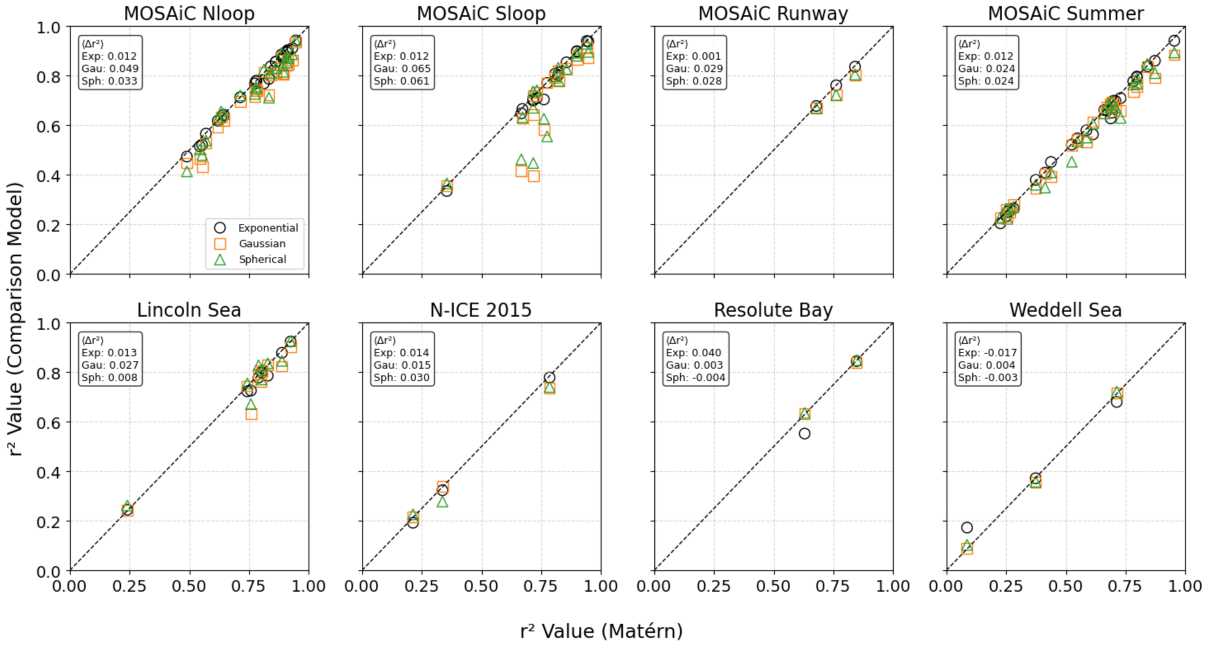

Figure 4. *Comparison of variogram model performance across all datasets. The Matérn and exponential models provide the best fits overall. r^2 values from the exponential, Gaussian, and spherical models are plotted against those from the Matérn model to illustrate relative performance.*

Moreover, in Fig 5, we observe that in some cases the Matérn model fails to produce a reasonable estimate of the effective range, e.g., MOSAiC Nloop 6 Feb, 20 Feb and 27 Feb, Sloop 26 Dec. Semi-variogram analysis can fail to detect slope change points at large length scales because local non-stationarities introduce undulations in the semi-variogram (Moon et al. 2019). Considering the marginal performance gain and stability, we adopted the exponential model for the main analysis, as it provides a good balance between simplicity, robustness, and accuracy.

In Fig. 6, the resulting correlation lengths derived from the exponential model do not differ significantly from those obtained using the Matérn model. The only exception is again the Weddell Sea (FYI) floe 503, see reasons in previous text**.**

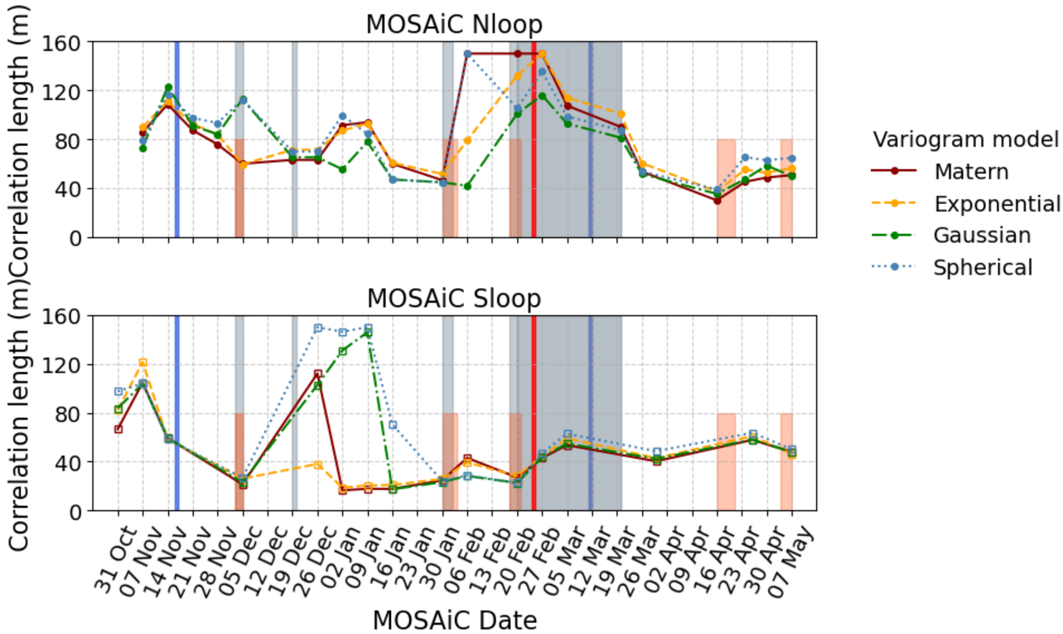

Figure 6. *The time-series correlation lengths over MOSAiC Nloop and Sloop. The correlation length is estimated using different models.*

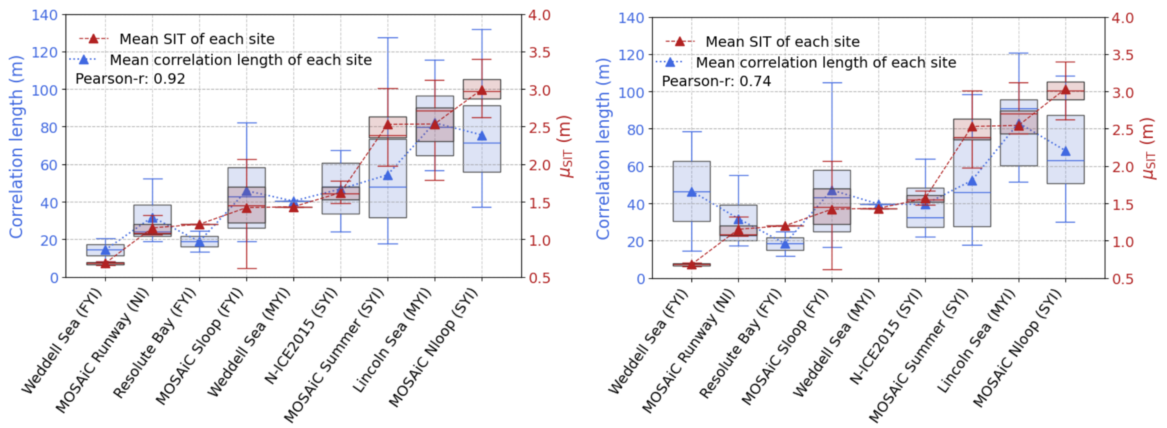

Figure 5. *Relationship between snow depth (SND) correlation length and sea-ice thickness (SIT). The correlation length is estimated using (left) the exponential model and (right) the Matérn model.*

In summary, the semi-variogram method is a simple and widely used approach within the sea-ice community for quantifying the spatial variability of snow depth (Iacozza and Barber, 1999, 2010; Sturm, 1998, 2002; Petrich 2012, Willatt, 2023). In the revision, we retain the results obtained using constant bin = 3m and exponential model in order to ensure stability and comparability across sites. We have added text on selecting bin width selection, variogram model choice, and data quality in Section 3.2.

However, semi-variogram analysis can fail to detect slope change points at large length scales because local non-stationarities introduce undulations in the semi-variogram (Moon et al. 2019). In such cases, the effective range cannot be robustly estimated from the semi-variogram, and more advanced approaches such as multifractal temporally weighted detrended fluctuation analysis may be required (Moon et al. 2019). Nevertheless, in this study the semi-variogram method performs robustly and provides new insights into snow spatial variability: the correlation length is positively related to ice thickness, which is one of the key findings of this study. This relationship is consistent with the chosen bin width and fitting model. We have added the above text in Section 4.3.

Reference

Iacozza, J., and Barber, D. G. (1999): An examination of the distribution of snow on sea ice. *Atmosphere–Ocean*, **37**, 21–51.

Iacozza, J., and Barber, D. G. (2010): An examination of snow redistribution over smooth land-fast sea ice. *Hydrological Processes*, **24**, 850–865.

Moon, W., *et al.* (2019): Physical length scales of wind-blown snow redistribution and accumulation on relatively smooth Arctic first-year sea ice. *Environmental Research Letters*, **14**, 104003.

Petrich, Chris, et al. "Snow dunes: A controlling factor of melt pond distribution on Arctic sea ice." Journal of Geophysical Research: Oceans 117.C9 (2012).

Sturm, M., Morris, K., and Massom, R. (1998): The winter snow cover of the West Antarctic pack ice: Its spatial and temporal variability. In *Antarctic Sea Ice: Physical Processes, Interactions and Variability*, **74**, 1–18.

Sturm, M., Holmgren, J., and Perovich, D. K. (2002): Winter snow cover on the sea ice of the Arctic Ocean at the Surface Heat Budget of the Arctic Ocean (SHEBA): Temporal evolution and spatial variability. *Journal of Geophysical Research: Oceans*, **107**, 8047.

Willatt, R., *et al.* (2023): Retrieval of snow depth on Arctic sea ice from surface-based, polarimetric, dual-frequency radar altimetry. *Geophysical Research Letters*, **50**, e2023GL104461.