https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

Investigating the drivers of wintertime Southern Ocean sea-ice leads using random forest algorithms

Umesh Dubey, Sascha Willmes, and Günther Heinemann

Department of Environmental Meteorology, Trier University, Trier, Germany

Correspondence: Umesh Dubey (dubey@uni-trier.de) and Sascha Willmes (willmes@uni-trier.de)

Abstract. Sea-ice leads play a crucial role in regulating ocean-atmosphere energy exchange, yet the physical drivers controlling their variability across the Southern Ocean remain unquantified. This study uses a machine learning approach based on random forest regression with permutation importance analysis to identify predictors of Southern Ocean lead frequency during winters (April–September, 2003–2023). The model integrates nine predictors representing atmospheric (wind speed, wind divergence, sea-level pressure, 2 m temperature), oceanic (surface current speed), and sea-ice kinematic variables (ice velocity, divergence, concentration), together with a seasonal descriptor (month). Evaluated on independent test data, the model achieves an evaluation performance correlation of r = 0.70 at the pan-Antarctic scale and r = 0.68–0.78 across regional sectors (MAE = 0.016–0.024). Permutation analysis indicates that 2 m temperature (20 %), wind divergence (13 %), ice divergence (11.9 %), and ocean current speed (11.6 %) collectively explain approximately 57 % of the observed lead frequency variability. Regional analysis reveals sector-specific drivers: The Weddell Sea is controlled by wind and ice divergence; the Ross Sea exhibits contributions from air temperature, wind divergence, and ocean current. The Indian and Pacific Ocean sectors show strong air temperature and ocean current influence, and the Bellingshausen–Amundsen Seas are dominated by seasonal wind forcing. However, the model does not fully resolve fine-scale structures evident in observations, hence a notable portion of the lead frequency variance remains unexplained due to the spatial resolution used in this study. This suggests the need for future work to apply a random forest framework at higher spatial resolution to investigate small-scale regional lead hotspots, including bathymetrically-controlled and coastal lead zones.

1 Introduction

Sea-ice leads, i.e. narrow linear cracks or elongated openings within the consolidated ice pack, play an important role in the climate system by mediating the exchange of heat, moisture, and gas between the ocean and atmosphere (Alam and Curry, 1997; Marcq and Weiss, 2012). When these openings expose relatively warm ocean water to the cold atmosphere, they trigger strong upward fluxes of sensible and latent heat, modify surface albedo, and influence boundary-layer stability (Lüpkes et al., 2008; Heinemann et al., 2022; Tian et al., 2025). As leads refreeze, they promote new ice formation and brine rejection, which in turn alter local salinity, increase water density, and contribute to deep-water formation and regional ocean circulation (Smith et al., 1990; Key et al., 1993; Ohshima et al., 2013). Beyond their thermodynamic significance, leads act as ecological and biogeochemical hotspots, providing habitats for marine fauna and serving as potential sources of methane emissions (Stirling, 1997; Kort et al., 2012; Damm et al., 2010). Through the ice–albedo feedback mechanism, an increased lead fraction enhances solar radiation absorption, which accelerates sea-ice thinning and amplifies regional climate feedbacks (Curry et al., 1995; Nihashi and Cavalieri, 2006). Consequently, lead variability strongly influences air–sea coupling and ocean circulation in both hemispheres (Vihma et al., 2014; von Albedyll et al., 2022).

The Southern Ocean represents a complex environment for studying lead dynamics due to the interplay of strong atmospheric forcing, intricate ocean circulation, and variable bathymetry (Holland and Kwok, 2012). Lead formation and variability are driven by both atmospheric processes, such as wind stress and ice divergence, and oceanic circulation patterns (Reiser et al., 2019; Wang et al., 2023). Recent studies using thermal-infrared remote sensing have enabled the detection of sea-ice leads through surface temperature anomalies derived from the Moderate Resolution Imaging Spectroradiometer (MODIS), which provides daily lead maps at a 1 km^2 spatial resolution (Reiser et al., 2020). However, frequent Antarctic cloud cover limits continuous observation and introduces substantial gaps in daily lead data.

To address these limitations, Dubey et al. (2025a) developed a gap-filled monthly climatology of lead frequency (LF) for winter months (April–September) between 2003 and 2023, offering a consistent long-term dataset for the Southern Ocean. Their analysis revealed ubiquitous and spatially heterogeneous lead occurrence, with pronounced maxima along coastal regions, continental shelf breaks, and key bathymetric features such as Maud Rise and Gunnerus Ridge. Increased LF values (> 0.25) were observed particularly along the Weddell and Ross Seas shelf breaks, where leads frequently align with underlying topographic gradients (Reiser et al., 2019; Dubey et al., 2025a).

While several studies have described lead distributions and temporal variability in both polar regions (Wang et al., 2016; Reiser et al., 2019; Willmes et al., 2023; Dubey et al., 2025a), a comprehensive quantitative framework for identifying and ranking the physical drivers, and quantifying the relative importance of atmospheric, oceanic, and ice-dynamic controls on Southern Ocean leads is still lacking. This knowledge gap limits the representation of lead-related processes in coupled climate models and limits our capacity to assess how future changes in atmospheric or oceanic forcing might influence lead dynamics and their feedback on the polar climate system (Zhang, 2014; Rheinlænder et al., 2022).

Lead formation arises from the complex interplay between multiple coupled non-linear processes, including atmospheric forcing (wind speed, wind divergence, synoptic pressure patterns, and air temperature), sea-ice dynamics (material properties, ice velocity, ice divergence and ice concentration), and oceanic circulation (surface current pattern). Winds generate divergent stress fields that mechanically fracture the ice cover and initiate lead opening (Kimura and Wakatsuchi, 2000). Colder near-surface temperatures enhance turbulent heat fluxes from existing leads and trigger refreezing. In parallel, ocean currents interacting with underlying bathymetry induce shear and strain that both directly open leads through localized stress concentration and indirectly precondition lead formation by mechanically weakening ice through long-term traction and stress accumulation (Holland and Kwok, 2012). Enhanced tidal currents over continental shelf breaks and prominent bathymetric features create localized zones of elevated shear and strain that are particularly conducive to lead opening (Dubey et al., 2025a; Reiser et al., 2019).

The non-linear nature and multi-scale interactions among these drivers are difficult to capture using traditional linear statistical approaches (von Albedyll et al., 2022). Machine-learning methods, particularly random forest (RF) regression models

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

(Breiman, 2001; Biau, 2012), offer a strong framework to capture such non-linear relationships among multiple predictors without requiring predefined functional assumptions. In polar research, RF models have shown strong performance in diverse applications such as sea-ice drift calibration (Palerme and Müller, 2021; Zhang et al., 2024), forecast of sea-ice concentration (Chi and Kim, 2017), and ice-type classification (Shen et al., 2017). Beyond prediction, they also provide interpretable measures of variable importance through permutation-based methods (Strobl et al., 2007), enabling identification of dominant drivers on lead formation.

Earlier work primarily focused on describing the spatial distribution, seasonal cycle, and long-term trends of LF derived from satellite observations (Dubey et al., 2025a; Reiser et al., 2020). However, an explanatory framework that integrates multiple predictors to model LF and identify the relative importance of individual drivers has not yet been developed. This study addresses this gap by developing an RF regression model to reconstruct LF (target variable) across the Southern Ocean using predictors (input variables) from the atmosphere, sea-ice and ocean for the period from 2003 to 2023 (April–September).

Using the gap-filled monthly LF dataset of Dubey et al. (2025a), this study has four main objectives. First, to quantify the contribution of key predictor variables to LF variability across the Southern Ocean. Second, to identify and rank the dominant drivers, and evaluate their relative importance for the pan-Antarctic region and five sub-sectors (Weddell Sea, Indian Ocean, Pacific Ocean, Ross Sea, and Bellingshausen–Amundsen Seas). Third, to assess regional variations in model performance and predictor importance. Fourth, to evaluate the incremental contribution of different predictors to identify the minimal predictor set required for skillful LF reconstruction.

This study is structured as follows. Sect. 2 introduces the LF dataset used as the target variable, outlines the atmospheric, ice-dynamic, and oceanic predictor variables, and describes the RF regression methodology. Sect. 3 presents the model performance using the test dataset for the Southern Ocean and its five regional sectors, evaluates predictor importance and the contribution of incremental predictors, and highlights the monthly model performance for each winter month. We discuss the results in Sect. 4, and finally, Sect. 5 concludes our work.

2 Data and Methods

2.1 Sea-ice lead frequency as target variable

This study uses the monthly LF dataset by Dubey et al. (2025b) for the Southern Ocean covering the months from April to September, 2003–2023. The dataset serves as the target variable in the RF regression framework. The LF represents a temporally integrated quantity indicating the number of days a pixel is detected as a lead during a specified period relative to the number of available clear-sky observations.

For use in the RF framework, the 1 km^2 monthly LF data were spatially aggregated to a 2° latitude × 5° longitude grid. This aggregation is meant to balance spatial resolution with computational efficiency while aligning with the resolution of predictor variables (see Sect. 4.1).

2.2 Auxiliary data as predictor variables

To investigate the drivers of lead formation, this study compiled nine input predictor variables for the same period as the target variable, grouped into four categories: atmospheric forcing, sea-ice kinematics, oceanic currents, and temporal descriptor.

Atmospheric predictors include 10 m wind speed, horizontal wind divergence, sea-level pressure, and 2 m air temperature from the ERA5 reanalysis (Hersbach et al., 2020) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and obtained from the Climate Data Store (CDS) of Copernicus Climate Change Service (C3S). ERA5 provides hourly atmospheric fields at approximately 31 km resolution, which were temporally averaged to monthly means and bilinearly interpolated to the analysis grid. Wind speed and divergence represent the principal mechanical forcing for ice motion and lead opening, capturing both direct effects through wind shear and indirect influences via pressure gradients. Wind divergence is directly used from the ERA5 reanalysis for the 10 m wind field. Sea-level pressure accounts for large-scale synoptic variability that modulates wind stress and convergence–divergence patterns over the ice pack (Tschudi et al., 2019; Hersbach et al., 2020). The 2 m air temperature controls the surface energy balance and thermodynamic processes at the ice–atmosphere interface (Perovich et al., 2002; Elliot, 2014); it influences the magnitude of turbulent heat fluxes over leads and the freezing rate of newly exposed water, both of which are critical to understanding the coupled thermodynamic evolution of lead formation (Mathiot et al., 2012).

Sea-ice kinematic predictors in this study consist of ice velocity, ice divergence, and sea-ice concentration. Ice velocity and divergence are derived from the Polar Pathfinder Daily Sea Ice Motion product (Tschudi et al., 2019) provided by the National Snow and Ice Data Center (NSIDC). This product integrates motion estimates from passive microwave sensors, visible and infrared satellite imagery, together with atmospheric reanalysis winds to generate daily ice drift vectors at 25 km spatial resolution. Ice divergence was calculated from the velocity components using centered finite differences. These variables represent the large-scale dynamic state of the ice cover, capturing mechanical deformation and motion that drive lead formation through divergent (Hutchings et al., 2011). Sea-ice concentration was derived from the ARTIST Sea Ice (ASI) algorithm applied to AMSR-E/AMSR2 satellite observations provided by the University of Bremen (Spreen et al., 2008). Sea-ice concentration is a key indicator of the mechanical state of the ice pack: densely consolidated ice with high concentrations tends to be mechanically stronger and more resistant to deformation, whereas lower concentrations reflect a more fractured, loosely connected ice cover that is more prone to breaking and forming leads (Feltham, 2008; Lüpkes and Gryanik, 2015). For April to June 2012, a gap in AMSR-E/AMSR2 data between October 2011 and June 2012 prevents direct retrieval of sea-ice concentration. For these months, we therefore use the mean sea-ice concentration computed from the periods with available AMSR data. Monthly means of ice velocity, divergence and ice concentration (> 15 %) were therefore calculated and mapped to the analysis grid.

Ocean surface current speed data were obtained from the ECMWF Ocean Reanalysis System 5 (ORAS5) product (Zuo et al., 2019). ORAS5 provides ocean surface current fields at approximately 0.25° spatial resolution, which were also monthly averaged and interpolated to the analysis grid. Ocean currents play a key mechanistic role in lead formation, as mobile sea-ice is continuously subjected to current-induced tractions that generate internal stress, weaken the ice mechanically, and promote divergent ice motion (Leng et al., 2024). The inclusion of ocean current speed as a predictor allows the RF model to account for oceanic control on lead variability that operates alongside atmospheric and ice-dynamic forcing.

Finally, the temporal descriptor, which represents the month, is included to represent the seasonal evolution of ice cover and ice thickness throughout the winter period. This descriptor captures the systematic progression of thermodynamic and ice-dynamic conditions from early winter (April) through late winter (September), allowing the model to account for seasonal shifts in the relative importance of different atmospheric, ice-dynamic and oceanic predictors.

2.3 Random forest regression framework

In this study, we use an RF regression model (Breiman, 2001) to model LF in each grid cell as a function of nine input predictor variables described in Sect. 2.2. RF is an ensemble learning method based on two key principles of randomness: bootstrap aggregating (bagging) and feature selection at each split. This combination allows the RF model to capture complex non-linear relationships and interactions among predictors without requiring predefined assumptions, making it well-suited for representing the non-linear processes driving sea-ice lead formation.

The RF algorithm builds multiple independent decision trees using bootstrap samples of the training data. Each tree is trained on a randomly selected subset of observations drawn with replacement from the original training data. This bootstrap sampling results in about one-third of the training samples being used in each tree, with the remaining two-thirds (out-of-bag observations) available for internal validation. The aggregation of many diverse trees through averaging significantly reduces variance and overfitting risk while maintaining predictive accuracy.

At each node split during tree construction, the RF algorithm randomly selects a subset of predictor variables rather than evaluating all available features. In this study, we followed standard recommendations for regression problems and set the number of candidate predictors at each split to the square root of the total number of predictors. This feature of randomness further decorrelates the decision trees by encouraging each tree to explore different combinations of predictors, preventing single dominant predictors from dominating all trees. Node splits were optimized using mean squared error as the dissimilarity metric.

The model was implemented in Python 3.10 using the scikit-learn library (Pedregosa et al., 2011; version 1.3). The RF configuration employed 100 decision trees, ensuring sufficient ensemble diversity while maintaining computational efficiency as increasing the number of trees beyond this number provided no measurable improvement in performance. Each tree was allowed a maximum depth, with most leaves containing only a single observation from the training data, and the minimum number of samples required to split an internal node was set to two.

The dataset was randomly split into training and testing subsets in an 80:20 ratio, ensuring balanced representation across months and regional sectors. The RF model was trained on 80 % training subset and subsequently evaluated on the independent 20 % testing subset. Model performance was evaluated using standard regression metrics such as bias (i.e. mean error), mean absolute error (MAE), root mean square error (RMSE), and the Pearson correlation coefficient (r).

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

2.4 Evaluation of predictor importance and incremental contribution

To identify the most influential predictors for sea-ice lead formation, we assessed predictor importance in the RF model using permutation importance (Strobl et al., 2007). This method quantifies the increase in prediction error when the values of a single predictor are randomly shuffled, effectively breaking its association with the target variable while preserving relationships among all remaining predictors. Permutation importance is calculated on held-out test data, avoiding overfitting bias, and accounts for both direct and interaction effects among predictors. It also allows a fair comparison between continuous and categorical (e.g., month) predictors (see Sect. 2.2).

To compare the influence of different predictors, we calculated the relative permutation importance by normalizing the individual importance scores (measured as the increase in mean square error) so that they sum to 100 %. This analysis enabled the identification of the dominant atmospheric, oceanic, and ice-dynamic predictors of lead variability across the Southern Ocean. The resulting relative importance distributions and their regional patterns are presented in Sect. 3.3.

To further evaluate the incremental contribution of predictor categories, we conducted a stepwise predictor addition experiment. The analysis began with a baseline RF model including only atmospheric variables (10 m wind speed, wind divergence, sea-level pressure, and 2 m air temperature) and progressively added sea-ice kinematic variables (ice velocity, divergence, and ice concentration), ocean current speed, and a temporal descriptor (month). Each nested model was trained on 80 % of the data and evaluated on the independent 20 % testing subset. At each stage, model skill was assessed using MAE, RMSE, and r. Differences in these metrics between consecutive models quantified the marginal gain attributable to the newly added variables. This incremental framework isolates the added value of each predictor group, revealing how these variables collectively improve LF reconstruction. All data preprocessing, model training, and evaluation were performed in Python 3.10, and the performance metrics for all nested models are summarized in Table 1. All results presented in Sect. 3 are derived from model predictions and evaluated using an independent test subset (20 % of the dataset) that was withheld during model training.

Table 1. Summary of model performance metrics and top-5 predictors (with percentage importance) for the Southern Ocean and regional sectors during April to September, 2003–2023, based on the test subset. Performance metrics include mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient. (r). Predictor abbreviations: T2m = 2 m air temperature, Wind div = wind divergence, Ice div = ice divergence, Ice conc = sea-ice concentration, SLP = sea-level pressure.

Metric & predictor	Southern Ocean	Weddell Sea	Indian Ocean	Pacific Ocean	Ross Sea	Bellingshausen–Amundsen Seas
MAE	0.022	0.024	0.016	0.016	0.019	0.021
RMSE	0.031	0.034	0.022	0.022	0.026	0.030
r	0.70	0.68	0.78	0.73	0.73	0.77
Top-5 predictors (%)	T2m: 19.6 Wind div: 12.7 Ice div: 11.9 Current speed: 11.6 Wind speed: 11.4	Wind div: 16 Ice div: 14.5 T2m: 14.3 Current speed: 12 Ice velocity: 9.4	T2m: 31 Current speed: 14.1 Ice div: 9.4 Wind speed: 9.3 Wind div: 8.9	T2m: 29.6 Current speed: 13 Wind div: 11.8 Wind speed: 11.5 Ice velocity: 10.1	T2m: 20.8 Wind div: 14 Current speed: 12.1 Ice velocity: 10.8 Ice conc: 9.6	T2m: 25.3 Month: 11.7 Wind div: 10.8 SLP: 10.2 Wind speed: 10

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

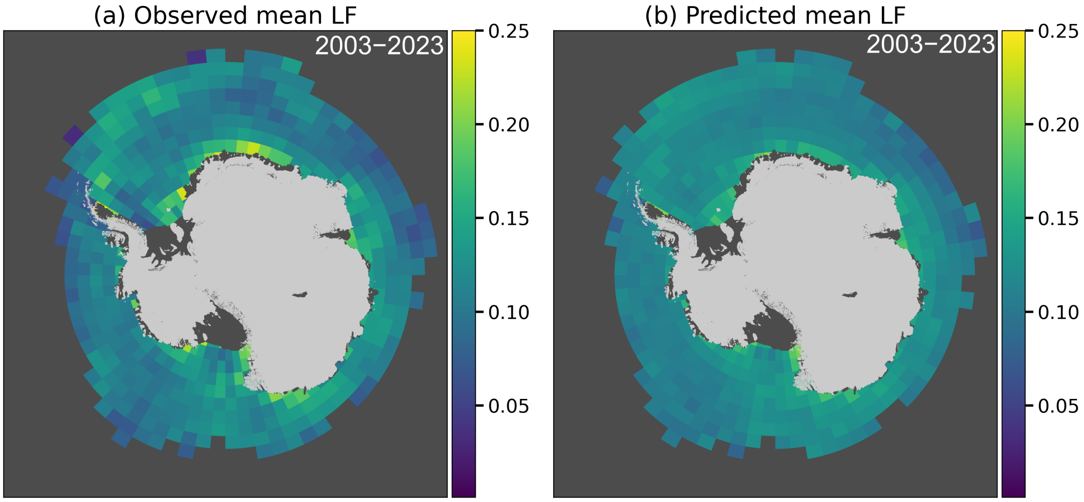

Figure 1. Spatial distribution of (a) observed mean lead frequency and (b) predicted mean lead frequency of the Southern Ocean during the winter period (April–September), 2003–2023, based on the test subset. Data are aggregated to 2° latitude × 5° longitude grid resolution.

3 Results

3.1 Reconstruction of spatial and seasonal lead frequency patterns

The RF regression model demonstrates skill in reproducing observed LF using nine predictor variables based on an independent test data sample across the Southern Ocean during winter months (April–September) for the period 2003–2023.

The model captures the large-scale pan-Antarctic pattern of elevated LF and is broadly consistent with the 21-year observational climatology presented in Dubey et al. (2025a). Reconstructed spatial maxima occur in the same general regions as the main lead-prone zones along the Weddell and Ross Seas continental shelf breaks, where LF values exceed 0.18, indicating that leads were present on more than about 18 % of winter days during the study period shown in Fig. 1. While the model broadly captures the pan-Antarctic regional pattern, it does not reproduce the fine-scale lead hotspots that occur around bathymetric features described in the climatological study (Dubey et al., 2025a; Reiser et al., 2019). The coarse 2° × 5° grid resolution fundamentally limits the model's ability to resolve the small-scale, bathymetrically-controlled divergence activity and shelf-break dynamics that are critical for understanding coastal lead zones. The relationship between bathymetric features and lead formation is smoothed out by spatial aggregation, meaning the model captures only large-scale regional tendencies rather than the explicit role of topographic steering (see Sect. 4.1).

The scatter plot of observed versus predicted LF from the test subset reveals tight clustering around the 1:1 line, particularly for mid-range values between 0.05 and 0.30, suggesting reliable reconstructions for lead frequencies encountered across the Southern Ocean pack ice (Fig. 2a). Density contours in the scatter plots further show that models are more reliable for moderate LF values (0.08–0.20), with a larger spread at both the lowest and highest ends of the distribution. A slight underprediction

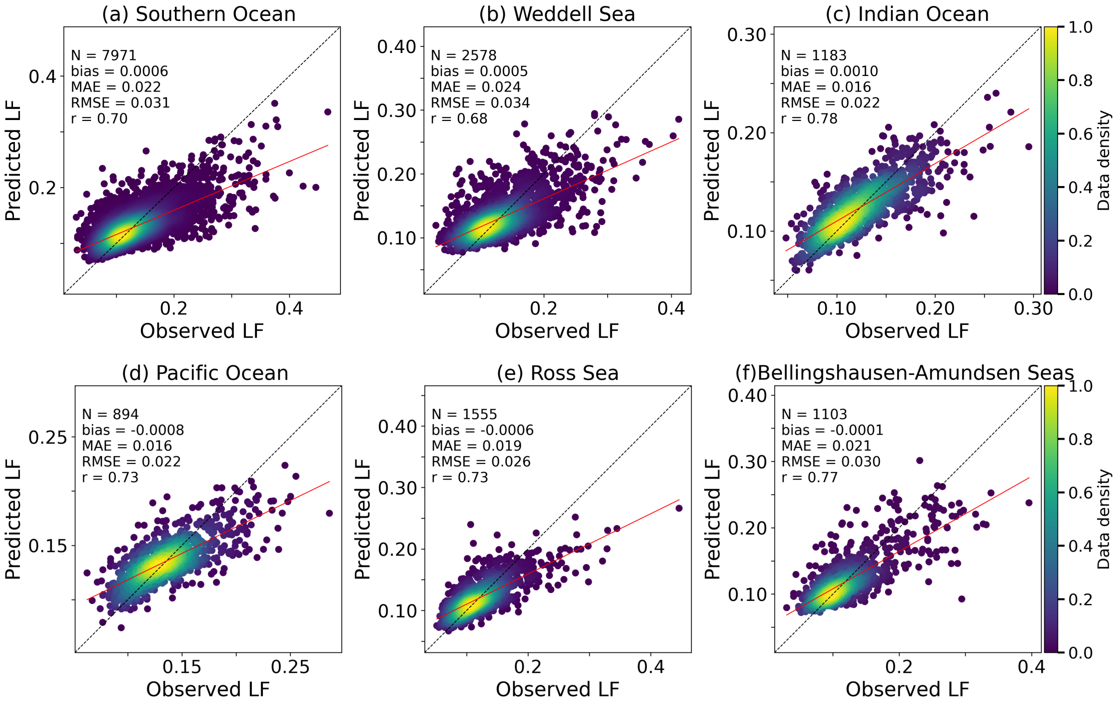

Figure 2. Model performance evaluation comparing observed versus predicted lead frequency based on the test subset for (a) the Southern Ocean and regional sectors: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen–Amundsen Seas during April–September, 2003–2023. The black dashed line represents the 1:1 reference line, and the red line shows the linear regression fit. Color gradients indicate data point density. Performance metrics include sample size (N), bias (mean error), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient (r).

of high lead events is evident (LF > 0.20) from spatial averaging effects that smooth sharp local gradients, particularly near dynamic coastal lead systems. When evaluated on the test dataset (N = 7971, i.e. number of data samples), the model achieved values of r = 0.70, MAE of 0.022, and RMSE of 0.031 for the Southern Ocean in Fig. 2a. Regional validation reveals considerable spatial variability in model performance across the five Southern Ocean sectors (Figs. 2b–f). The Indian Ocean sector shows the highest performance (r = 0.78, MAE = 0.016, RMSE = 0.022), followed by the Bellingshausen–Amundsen Seas (r = 0.77, MAE = 0.021, RMSE = 0.030) and the Ross Sea (r = 0.73, MAE = 0.019, RMSE = 0.026). The Pacific Ocean (r = 0.73, MAE = 0.016, RMSE = 0.022) and Weddell Sea (r = 0.68, MAE = 0.024, RMSE = 0.034) sectors show comparatively lower correlations but still maintain confidence in reconstruction of LF.

Figure 3 shows that the RF model captures the seasonal evolution of monthly mean LF throughout the winter season, 2003–2023. At the pan-Antarctic scale, predicted monthly LF closely follows observations from April to September, capturing both the timing and magnitude of the rise toward mid-winter and the subsequent late-season decline. Predicted LF increases from 0.126 in April to a June peak near 0.140, before decreasing slightly toward September at lowest LF value (Fig. 3a). Regional temporal evolution shows close agreement across sectors (Figs. 3b–f). The Weddell Sea displays a clear mid-winter

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

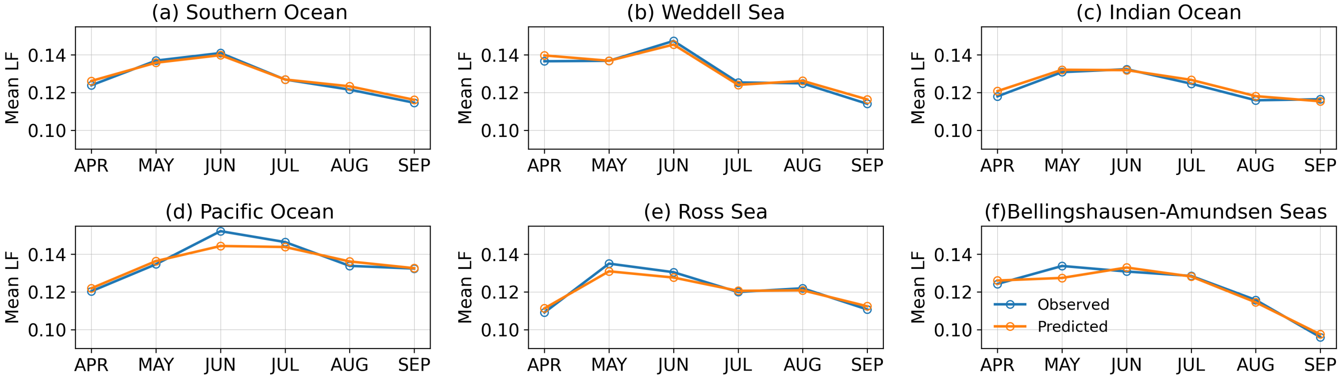

Figure 3. Seasonal evolution of monthly mean lead frequency from April through September, comparing observed (blue) and predicted (orange) values based on the test subset for (a) the Southern Ocean and regional sectors: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen–Amundsen Seas during 2003–2023.

maximum, while the Indian Ocean sector remains comparatively stable with only a modest increase in the May–June (Fig. 3c). Seasonal variations are weaker between observed and predicted LF during mid-winter in the Pacific and Ross Sea sectors (Figs 3d–e) than in other regions. The Bellingshausen–Amundsen Seas show the lowest seasonal values in September, and maintain moderate LF values throughout winter (Fig. 3f). Across all sectors, the generally close match between predicted and observed monthly means indicates that the model captures much of the large-scale temporal variability of LF, although noticeable discrepancies remain in some regions and months.

3.2 Incremental predictor contribution analysis

Figure 4 presents how different variables contribute to overall model performance based on the test subset. For the Southern Ocean, wind speed alone provides limited skill (r ∼0.10) with large errors (MAE = 0.037, RMSE = 0.048), and performance improves steadily as atmospheric variables, including wind divergence, 2 m air temperature, and sea-level pressure, are progressively introduced. Adding wind divergence increases r to ∼0.28 (MAE ∼0.033, RMSE ∼0.045), and the subsequent addition of 2m air temperature brings r to ∼0.40 (MAE ∼0.029, RMSE ∼0.038). Sea-level pressure further increases r to ∼0.45 (MAE ∼0.029, RMSE ∼0.036). The addition of ice-dynamic variables (ice concentration, ice velocity, and ice divergence) progressively improves correlation to r ∼0.60, followed by ocean current speed, raising r to ∼0.62. The final addition of the temporal descriptor (month) yields an enhancement in model performance, bringing the correlation to its full value (r = 0.70) and reducing errors to MAE = 0.022 and RMSE = 0.031 at the Southern Ocean scale shown in Fig. 4a.

Regional patterns show pronounced differences in how predictor additions affect skill across regions. In the Weddell Sea, correlation increases gradually with each predictor addition (r = 0.68), and error metrics correspondingly decrease (MAE = 0.024, RMSE = 0.034), with steep improvements in both r and errors when month is added (Fig. 4b). The Indian and Pacific Ocean sectors demonstrate that 2 m air temperature enhances correlation early ∼0.50 in the sequence, with proportional error reduction (Figs. 4c–d). The Ross Sea (Fig. 4e) shows r ∼0.20 (MAE ∼0.036, RMSE ∼0.047) with wind speed alone, increasing

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

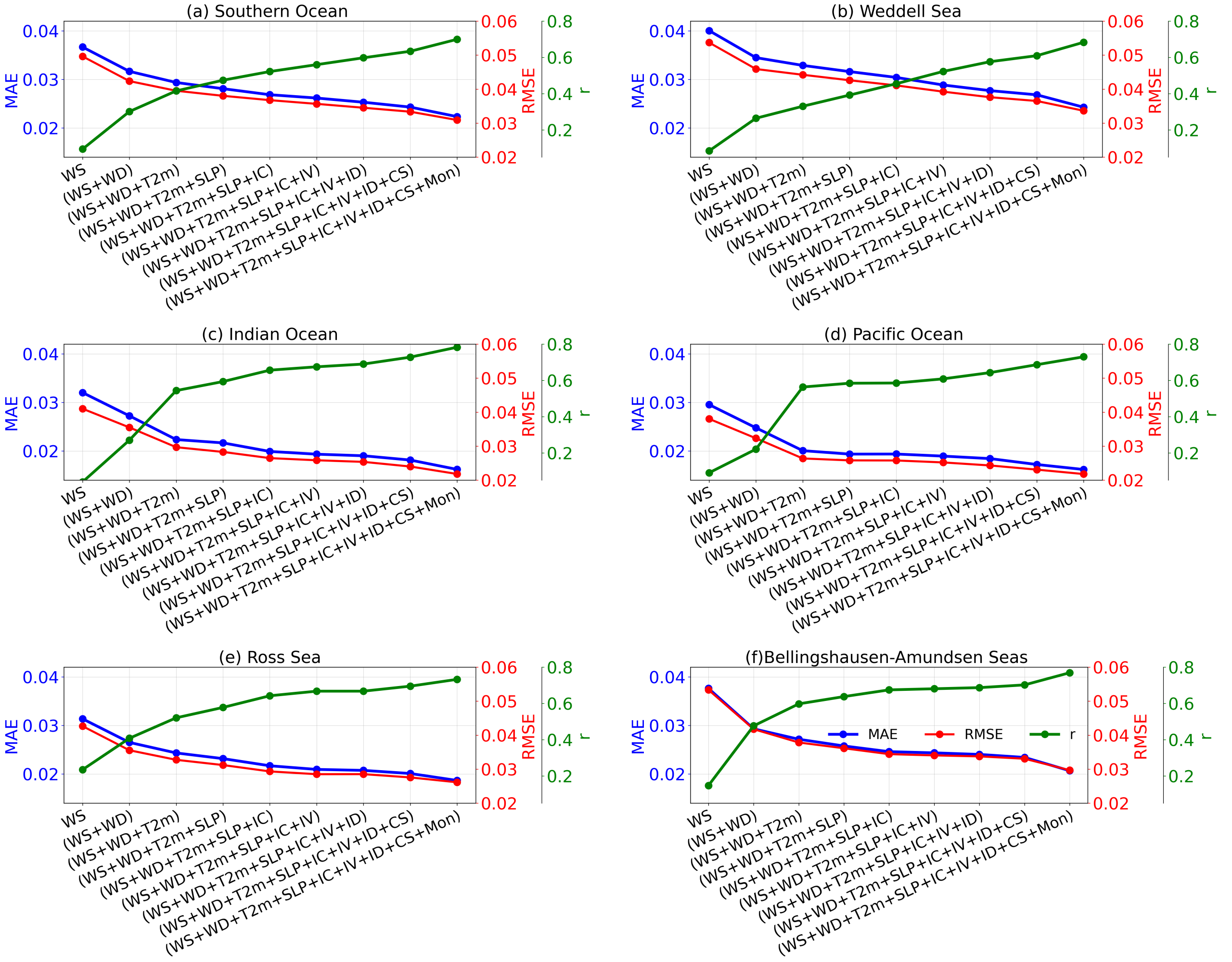

Figure 4. Incremental predictor contribution analysis showing the progressive improvement in model performance as predictor variables are sequentially added for (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen–Amundsen Seas during April–September, 2003–2023. Lines represent the MAE (blue), RMSE (red), and r (green). Predictors are added in the following sequence: wind speed (WS)⟶wind divergence (WD)⟶2 m air temperature (T2m)⟶sea-level pressure (SLP)⟶ice concentration (IC)⟶ice velocity (IV)⟶ice divergence (ID)⟶ocean current speed (CS)⟶month (Mon).

steadily to r ∼0.40 (MAE ∼0.034, RMSE ∼0.044) with wind divergence, and r ∼0.50 with 2m air temperature. Ice-dynamic variables increase r to ∼0.65 (MAE ∼0.028, RMSE ∼0.036), with ocean current speed contributing to r ∼0.68, and month yielding the final r = 0.73 (MAE = 0.019, RMSE = 0.026). The Bellingshausen–Amundsen Seas show consistent, incremental correlation gains throughout the predictor sequence, with coupled error reductions, particularly good improvements in both metrics when 2 m air temperature, wind divergence, and month are added in Fig. 4f.

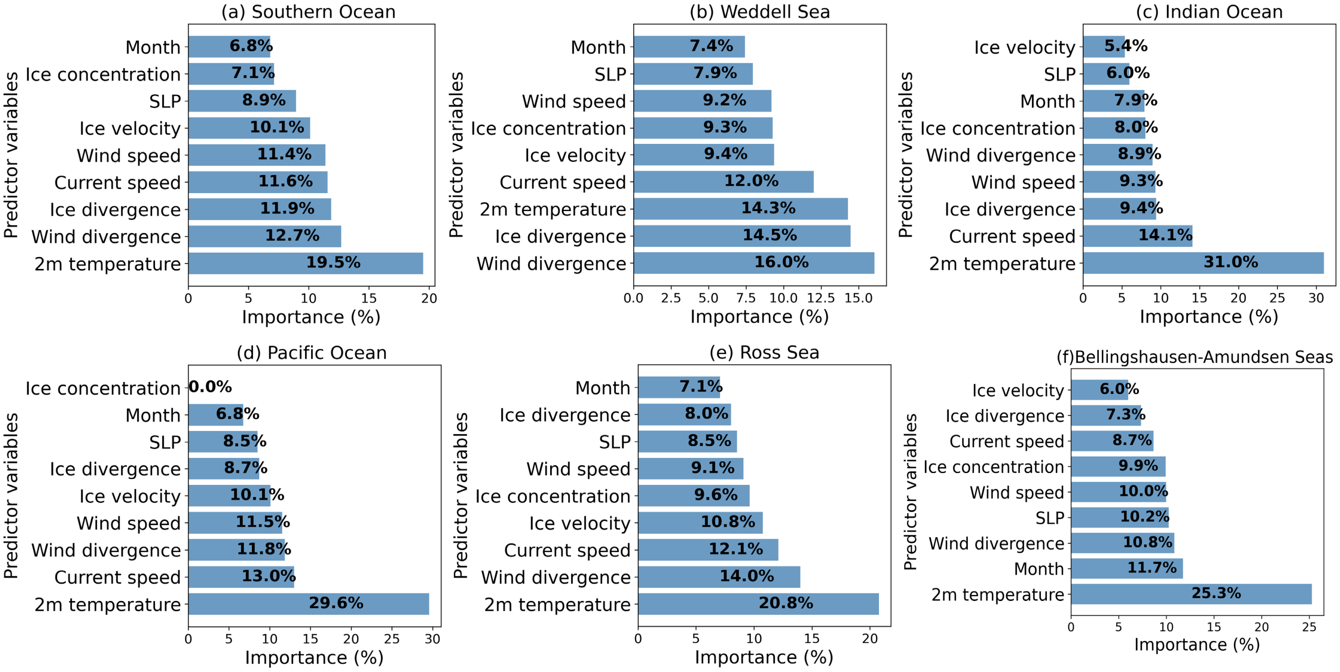

Figure 5. Permutation-based relative importance of predictor variables for lead frequency based on the test subset across (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen–Amundsen Seas during April–September, 2003–2023. Values represent the percentage contribution to the model performance.

The largest improvements in both correlation and error metrics typically occur when 2 m air temperature and wind divergence are introduced, indicating that ice-pack mechanical properties and wind context are critical for capturing lead formation mechanisms.

3.3 Relative importance of predictor variables

Figure 5 shows the permutation-based feature importance analysis, which reveals the dominant predictors of LF across the Southern Ocean and its regions based on the test sample. All feature importance values are reported as whole number percentages, rounded to the nearest integer (based on above and lower than 0.5 decimal), to provide a clearer interpretation of predictor contributions. The same integers but different decimals for a similar region from the original analysis are maintained as they are. At the pan-Antarctic scale, 2 m air temperature emerges as the most influential predictor (20 %), followed by wind divergence (13 %), ice divergence (11.9 %), ocean current speed (11.6 %), and wind speed (11.4 %). Ice velocity, sea-level pressure, ice concentration, and month provide additional contributions.

Regional importance patterns reveal sector-to-sector differences shown in Figs. 5b–f and Table 1. The Weddell Sea is dominated by wind divergence (16 %) and ice divergence (14.5 %), with 2 m air temperature (14.3 %), and ocean current speed (12 %) contributing substantially. The Indian Ocean sector displays the strongest thermodynamic control, with 2 m air temperature

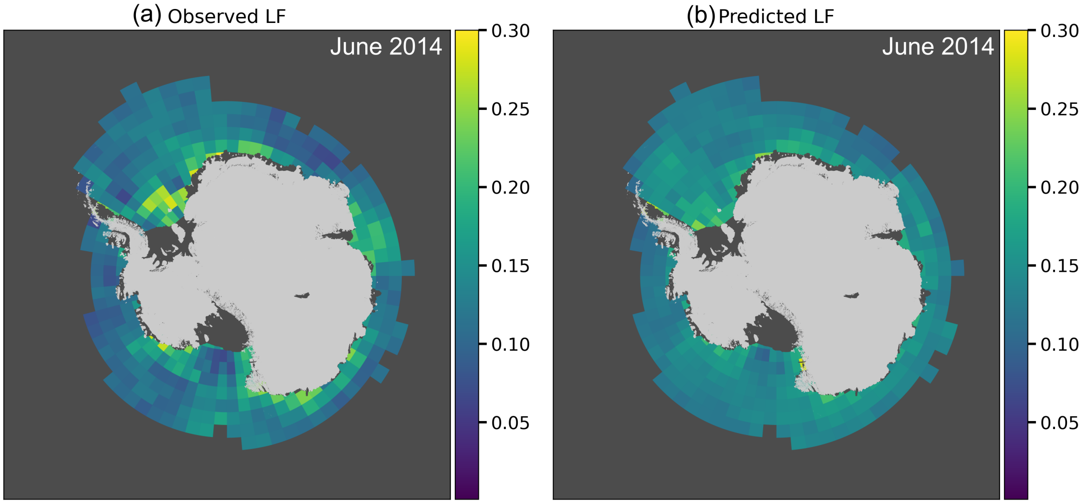

Figure 6. Spatial distribution of (a) observed mean lead frequency and (b) predicted mean lead frequency of the Southern Ocean for June 2014, based on the test subset. Data are aggregated to 2° latitude × 5° longitude grid resolution.

accounting for 31 % of importance, complemented by current speed (14 %), ice divergence (9.4 %), wind speed (9.3 %), and wind divergence (9 %). The Pacific Ocean shows 2 m air temperature influence with 30 %, followed by ocean current speed (13 %), wind divergence (11.8 %), wind speed (11.5 %), and ice velocity (10 %). The Ross Sea exhibits more balanced contributions with 2 m temperature (21 %), wind divergence (14 %), current speed (12 %), ice velocity (11 %), and ice concentration (10 %), reflecting its dynamical environment. The Bellingshausen–Amundsen Seas are dominated by 2 m air temperature (25 %), with substantial contributions from month (12 %), wind divergence (10.8 %), sea-level pressure (10.2 %), and wind speed (10 %).

These results suggest that Southern Ocean lead formation is influenced by regional-scale interactions between atmospheric forcing, ice–pack dynamics, and oceanic processes. Among the predictors at the pan-Antarctic scale, 2 m temperature, wind-ice divergence, and current speed show the strongest influences, underscoring the essential roles of atmospheric forcing, sea-ice deformation and oceanic traction in shaping lead variability, which will be discussed further in Sect. 4. Overall, the feature importance patterns indicate that lead formation arises from coupled non-linear interactions among multiple physical processes rather than from any single dominant predictor.

3.4 Monthly model performance for June

To assess how the dominant predictors and monthly RF model performance evolve during mid-winter using the test subset, we evaluate June as a representative case study. June represents a mid-winter state when overall lead activity peaks across the Southern Ocean (Dubey et al., 2025a). The month is excluded as a predictor in this single-month analysis. To illustrate the spatial distribution of LF during June, we present June 2014 as an example month (Fig. 6), which highlights how the model

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

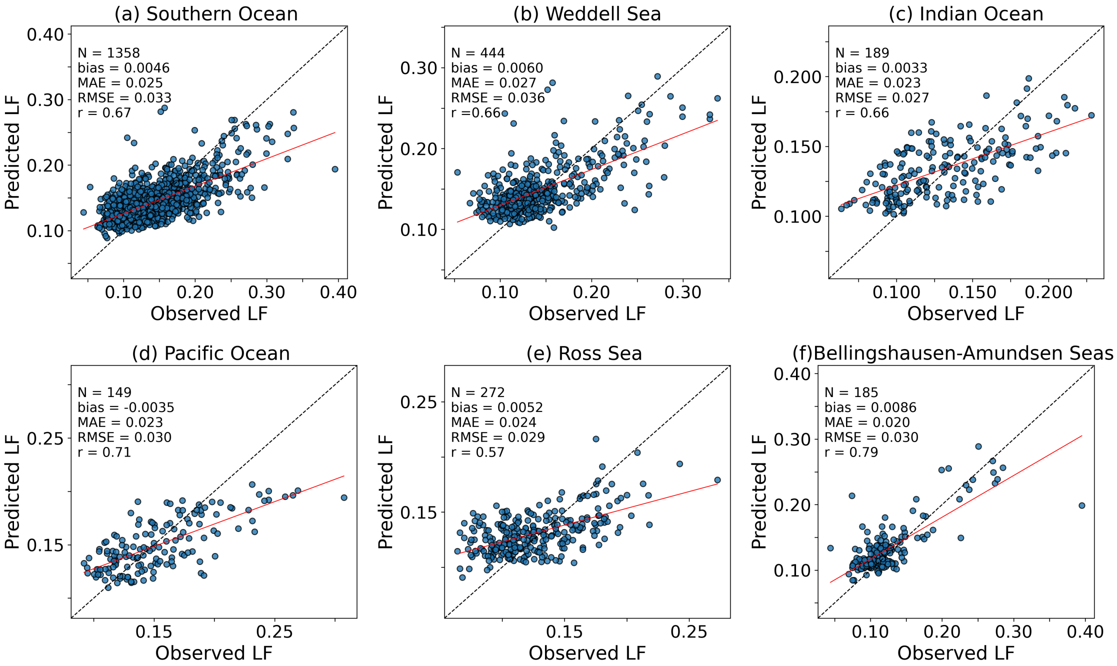

Figure 7. Observed versus predicted lead frequency based on the test subset for June (2003–2023) across (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen-Amundsen Seas. Each point represents an individual 2° × 5° grid cell.

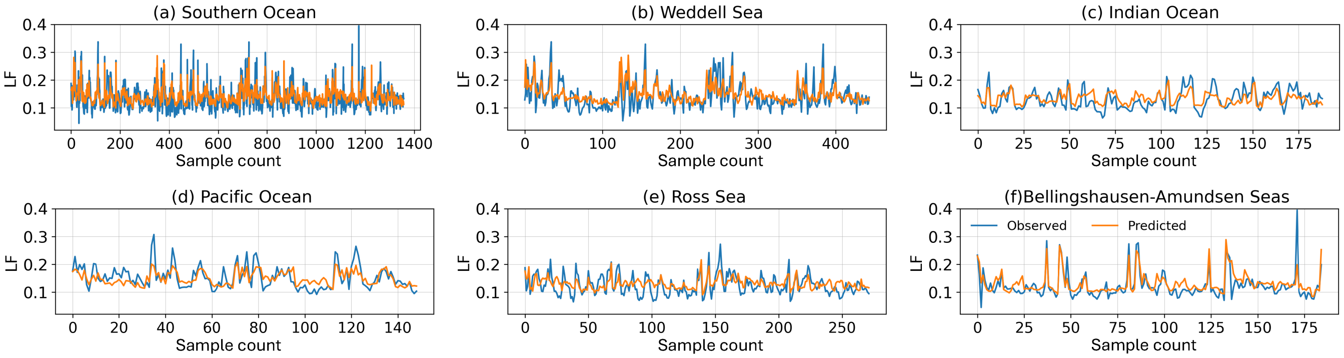

Figure 8. Sample-by-sample comparison of observed (blue) and predicted (orange) lead frequency based on the test subset for June (2003–2023) across (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen-Amundsen Seas. Each point represents an individual 2° × 5° grid cell.

represents large-scale spatial structures. The model shows the broad spatial distribution of elevated LF, with lead activity in the Weddell and Ross Seas as well as in Prydz Bay, where observed and predicted LF values frequently exceed 0.20, but fine-scale coastal features and narrow lead structures are not resolved at the coarse resolution. However, the comprehensive

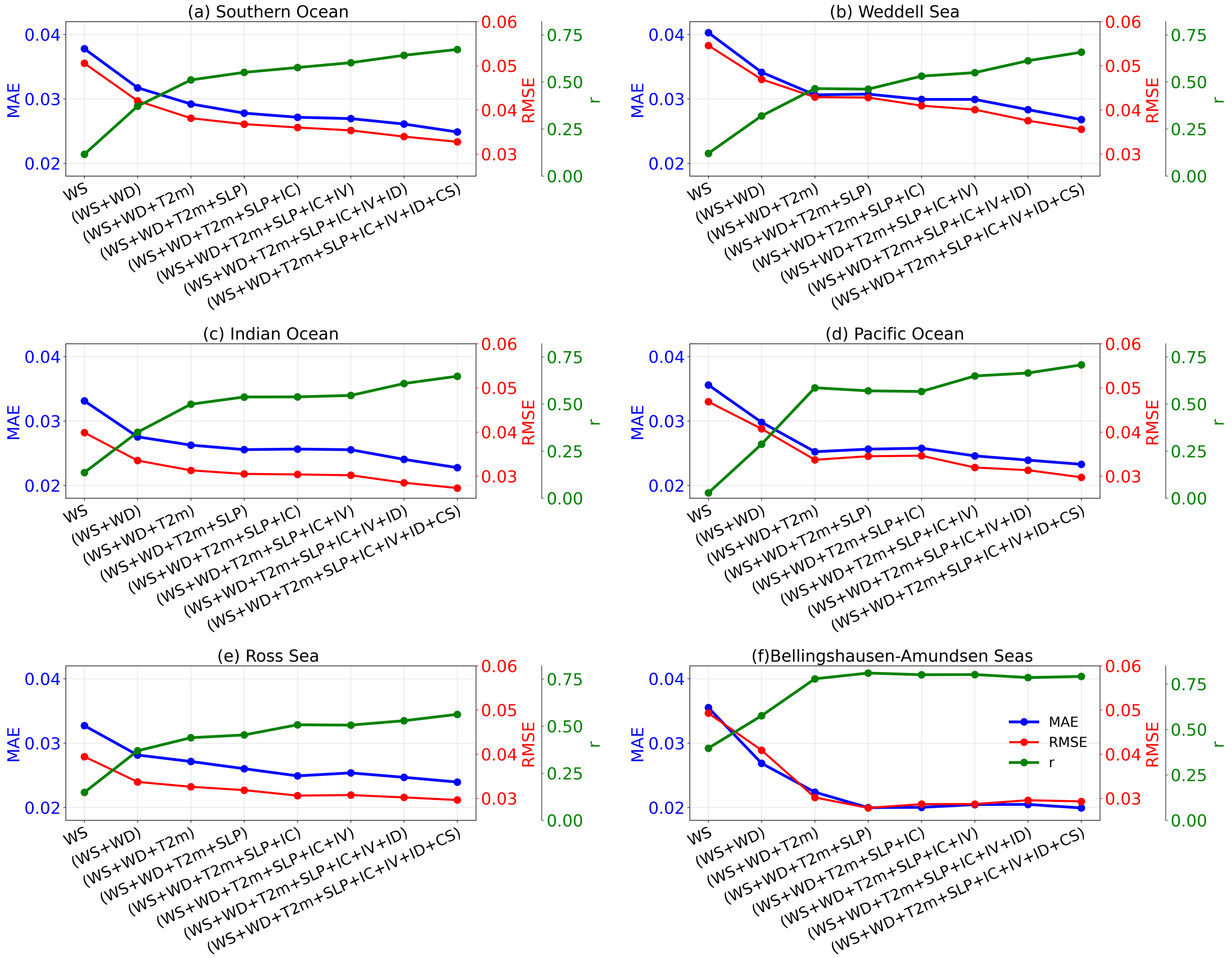

Figure 9. Incremental predictor addition analysis for June (2003–2023) showing MAE (blue), RMSE (red), and r (green) for (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen-Amundsen Seas. The month is excluded as a predictor in this single-month analysis. Predictors are added in the following sequence: wind speed (WS)⟶wind divergence (WD)⟶2 m air temperature (T2m)⟶sea-level pressure (SLP)⟶ice concentration (IC)⟶ice velocity (IV)⟶ice divergence (ID)⟶ocean current speed (CS).

model performance evaluation and predictor importance analysis presented in figures 7–10 are based on the aggregated June data across the entire 2003–2023 period.

The model performance at the pan-Antarctic scale during June is characterized by r = 0.67, MAE = 0.025, and RMSE = 0.033 in Fig. 7a. Regional performance varies considerably. The Bellingshausen-Amundsen Seas (r = 0.79, MAE = 0.020, RMSE = 0.030) and Pacific Ocean (r = 0.71, MAE = 0.023, RMSE = 0.030) exhibit the highest model skill (Figs. 7f and d). The Indian Ocean (r = 0.66, MAE = 0.023, RMSE = 0.027) and Weddell Sea (r = 0.66, MAE = 0.027, RMSE = 0.036) achieve moderate

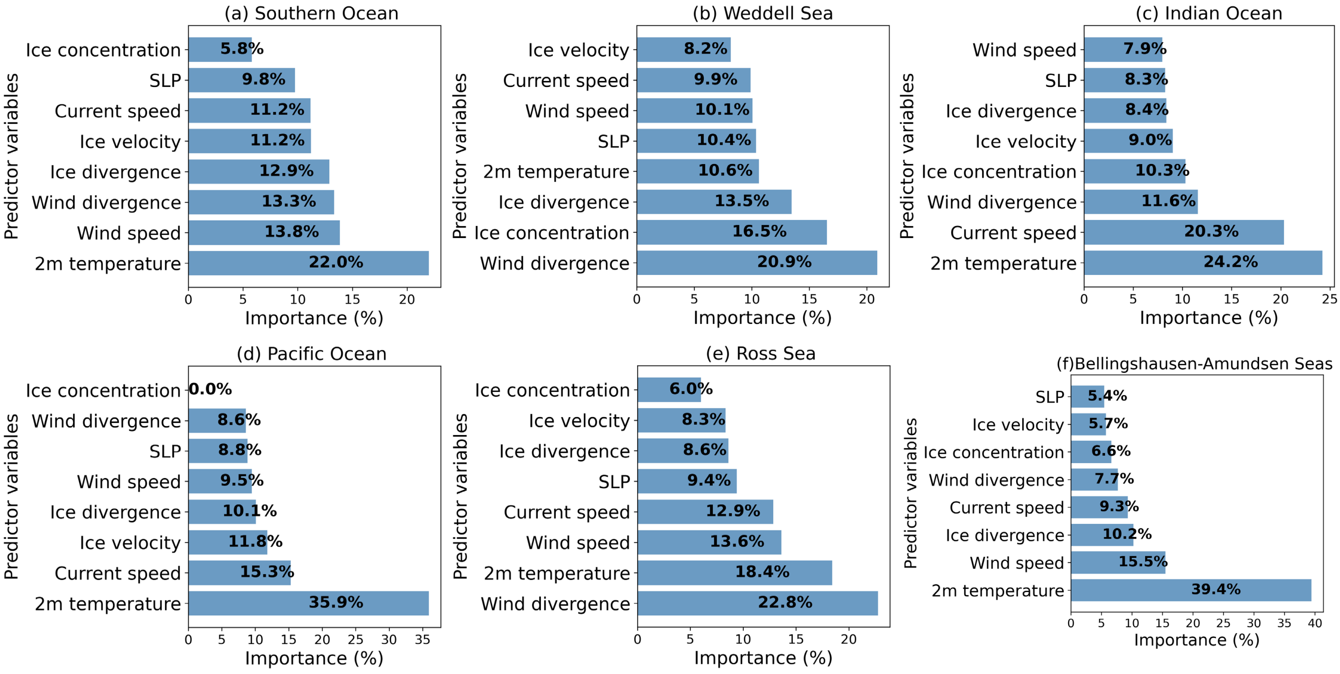

Figure 10. Relative importance of predictor variables for lead frequency based on the test subset for June (2003–2023) across (a) the Southern Ocean and its regions: (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen-Amundsen Seas.

performance, while the Ross Sea maintains lower correlation (r = 0.57, MAE = 0.024, RMSE = 0.029) (Figs. 7c, b, and e). Figure 8 presents sample-by-sample data-point comparisons, which further reveal that the model generally tracks predicted LF variations across individual regions and grid cells, although some discrepancies occur in regions with high variability.

Incremental predictor analysis for June is shown in Fig. 9, which overall indicates that 2 m air temperature and wind forcing contribute more prominently, reflecting stronger atmospheric control during mid-winter. At the Southern Ocean scale, wind speed alone yields limited skill (r ∼0.15, MAE ∼0.035), but adding wind divergence and 2 m air temperature increases correlation to r ∼0.50 and reduces MAE to ∼0.028. The full predictor set achieves r = 0.67 and MAE = 0.025 (Fig. 9a). In the Weddell Sea, wind speed alone provides r ∼0.20, which increases to r ∼0.55 after adding atmospheric variables, reaching r = 0.66 with the complete predictor set (Fig. 9b). The Indian and Pacific Ocean sectors show sharp performance gains from winds and 2 m air temperature addition, with the Indian Ocean improving from r ∼0.10 to r ∼0.60 after adding 2 m temperature, ultimately reaching r = 0.66 (Fig. 9c), and the Pacific Ocean advancing from r ∼0.20 to r ∼0.65 with 2 m temperature, achieving final r = 0.71 (Fig. 9d). The Ross Sea shows gradual improvement from r ∼0.15 with wind speed alone to r = 0.57 with all predictors (Fig. 9e). The Bellingshausen-Amundsen Seas exhibit steady improvements across all predictor groups, progressing from r ∼0.25 to r = 0.79, with the largest gains occurring when wind divergence and 2 m air temperature are added, as shown in Fig. 9f.

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

Figure 10 presents predictor importance rankings for June. All importance values are reported as whole number percentages, rounded to the nearest integer (based on above and lower than 0.5 decimal). The same integers but different decimals for a similar region from the original analysis are maintained as they are. At the pan-Antarctic scale, 2 m temperature (22 %) remains dominant while wind speed (13.8 %) and wind divergence (13.3 %) increase in prominence, reflecting enhanced wind forcing during mid-winter. Ice divergence (13 %), ice velocity (11 %), and ocean current speed (11 %) also contribute meaningfully, with lesser contributions from sea-level pressure (10 %) and ice concentration (6 %) (Fig. 10a). Regional patterns remain distinct: the Weddell Sea is shaped by wind divergence (21 %) followed by ice concentration (17 %) and ice divergence (14 %), with 2 m temperature (10.6 %) and sea-level pressure (10.4 %) providing additional control (Fig. 10b), while the Indian Ocean retains strong 2 m temperature (24 %), current speed (20 %), and wind divergence (12 %) control (Fig. 10c). The Pacific Ocean shows higher contributions from 2 m temperature (36 %), ocean current (15 %), and ice velocity (12 %), followed by ice divergence (10 %) (Fig. 10d), whereas the Ross Sea exhibits particularly high contributions with wind divergence (23 %), 2 m temperature (18 %), wind speed (14 %), and current speed (13 %) influence, along with sea-level pressure (9 %) (Fig. 10e). The Bellingshausen-Amundsen Seas remain primarily 2 m temperature-driven (39 %), followed by wind speed (16 %), ice divergence (10 %), current speed (9 %), and wind divergence (8 %) (Fig. 10f). Overall, the enhanced role of wind fields suggests that atmospheric forcing becomes more influential during mid-winter.

3.5 Monthly model performance analysis for winter months

Figure 11 presents the test subset monthly model performance metrics and top-5 predictors for each winter month (April–September, 2003–2023) across the Southern Ocean and regional sectors. The month is excluded as a predictor in this single-month analysis. This demonstrates the monthly and regional evolution of model skill and progressive shifts in predictor dominance throughout the winter. All feature importance values are reported as whole number percentages, rounded to the nearest integer (based on above and lower than 0.5 decimal). The same integers but different decimals for a similar region from the original analysis are maintained as they are.

At the pan-Antarctic scale (Fig. 11a), 2 m air temperature is the most frequently important driver (17–26 %), while wind-related variables (wind divergence: 12–15 %; wind speed: 12–14 %) and ice divergence also contribute substantially. Model performance varies across months (r: 0.62–0.68), with April and September showing higher correlations (r ∼0.67–0.68), while August shows the lowest (r = 0.62). Early winter (April) shows balanced contributions from thermodynamic and dynamic forcing (2 m temperature: 21 %, wind divergence: 12.3 %, ice divergence: 12 %). Through mid-winter (June), wind forcing increases (wind speed: 13.8 %, wind divergence: 13.3 %), and by late winter (September), ocean currents become more prominent (15 %), indicating a growing role of persistent ice–ocean coupling as the ice pack thickens and consolidates.

The Weddell Sea (Fig. 11b) exhibits seasonal variation, with shifts between thermodynamic and wind-ice dynamic contributions. In April, 2 m temperature (23 %), wind divergence (22 %), and ice divergence (13 %) show comparable importance (r = 0.56). During May, ice divergence (19 %) and ice concentration (15 %) increase. By June, wind divergence (21 %) re-emerges alongside ice concentration (17 %) and ice divergence (14 %) (r = 0.66). Through July and August, no single driver dominates, and lead variability reflects a multi-process regime involving thermodynamic forcing, wind divergence, and ice

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

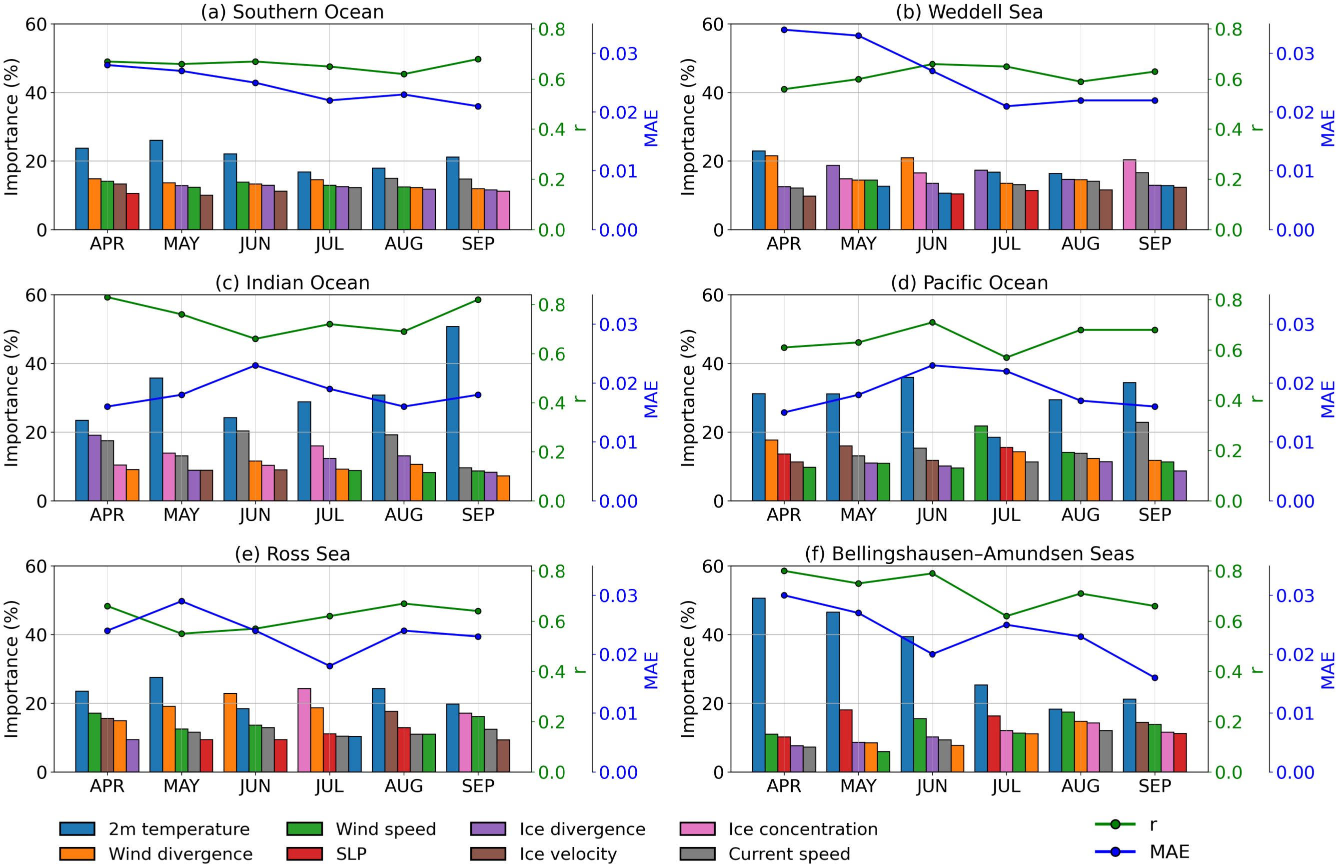

Figure 11. Monthly and regional variation of predictor importance and metrics across Antarctic regions (April–September, 2003–2023). Bar charts show relative importance (%) of top-5 predictors in six Antarctic sectors: (a) Southern Ocean, (b) Weddell Sea, (c) Indian Ocean, (d) Pacific Ocean, (e) Ross Sea, and (f) Bellingshausen–Amundsen Seas. Overlaid lines indicate r (green) and MAE (blue). Predictors include 2m temperature (blue), wind divergence (orange), wind speed (green), SLP (red), ice divergence (purple), ice velocity (brown), ice concentration (pink), and current speed (grey).

deformation. By September, ice concentration (20 %) and ocean currents (17 %) become increasingly important, indicating stronger mechanical control under a thicker late-winter ice pack.

In contrast, the Indian Ocean (Fig. 11c) region is dominated by 2 m temperature throughout winter (23–51 %), with ocean currents playing overall a secondary role (13–20 %). In April, 2 m temperature (23 %) and ice divergence (19 %) with current speed (18 %) are associated with the highest correlation (r = 0.83). During May, temperature dominance increases (36 %), with ice concentration (13.9 %) and current speed (13.1 %) contributing (r = 0.76). By late winter, 2 m temperature remains important: August shows 31 % with ocean current speed (19 %), and September shows 51 % with r = 0.82. The Pacific sector (Fig. 11d) also shows consistent thermodynamic dominance with substantial contributions from ocean currents, with model performance ranging from r = 0.61 to 0.71. In April, 2 m temperature (31 %) and wind divergence (18 %) drive activity (r =

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

0.61). Through late winter, this region maintains temperature-current coupling: August shows 2 m temperature (29 %) with current speed (14 %); and September shows 2 m temperature (34 %) with ocean currents (23 %).

The Ross Sea (Fig. 11e) displays more variable predictor importance (r ∼0.55–0.67). April shows 2 m temperature (24 %), wind speed (17 %), and ice velocity (16 %) as comparable contributors (r = 0.66), while May and June are increasingly wind-driven, with wind divergence peaking in June (23 %). July shows ice concentration becomes important (24 %), while August shows the highest correlation (r = 0.67) with 2 m temperature (24 %) and ice velocity (18 %) as substantial contributors.

Finally, the Bellingshausen–Amundsen Seas (Fig. 11f) are persistently dominated by thermodynamic forcing. Across most months, 2 m air temperature accounts for 40–50 % of predictor importance, with particularly high skill in April, May, and June (r ∼0.78–0.80). Secondary contributions from wind speed, sea-level pressure, and ice velocity emerge intermittently, especially in late winter, but temperature remains the overarching control. This strong thermodynamic imprint is consistent with the region's sensitivity to offshore winds and seasonal ice-thickness evolution. Overall, these patterns suggest that lead formation results from distinct combinations of atmospheric, oceanic, and ice-dynamic physical processes that vary across sectors and evolve throughout the season.

4 Discussion

4.1 Model performance and approach

The random forests are predictive algorithms in a statistical sense, they are used here as an explanatory tool to reconstruct observed LF and to evaluate the relative importance of the drivers used within the period 2003–2023, rather than to provide operational forecasts of future lead occurrence. By using the monthly LF dataset with the RF framework, we show that wintertime LF arises from a coupled interaction of dynamic atmospheric processes, ice-pack dynamics, oceanic circulation, and seasonal context (i.e., month). The agreement between observed and predicted large-scale LF patterns (Fig. 1) and the seasonal evolution across all sectors (Fig. 3) is consistent with a physically meaningful influence of the identified LF predictors, while the remaining discrepancies point to processes that are not fully represented in the current framework.

The 2° latitude × 5° longitude grid used in this study represents a spatial aggregation from the native 1 km^2 resolution of LF and the < 31 km resolution of predictor variables. This coarsening was necessary for computational tractability. In particular, coastal divergence activity and shelf-break dynamics features that drive lead formation along the continental regions are substantially smoothed by this aggregation. Nevertheless, the current resolution is sufficient to identify the dominant large-scale physical drivers and their regional variations, which was the primary objective of this study.

Our results show that leads are widespread throughout the Southern Ocean pack ice during winter (2003–2023), with model performance (r = 0.70) indicating that the used predictors capture a substantial fraction of the variability in lead formation. However, a notable portion of the LF variance remains unexplained with the used predictors, reflecting both unresolved processes and limitations of the coarse spatial and temporal resolution. This highlights the need for more detailed observations of sea-ice dynamics, particularly high-resolution lead-resolving modelling for small-scale regional features such as

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

bathymetrically-controlled and coastal lead zones, to better understand ocean–ice–atmosphere interactions and to improve future climate predictions.

A prominent outcome of this study is the clear ranking of drivers governing Southern Ocean lead formation, with 2 m air temperature emerging as the single most important predictor (20 %), followed by wind divergence (13 %), ice divergence (11.9 %), and ocean current speed (11.6 %) (Fig. 5a). Collectively, the nine predictors explain the majority of the model's skill in reconstructing observed LF, with sectoral correlation ranging from lowest 0.68 in the Weddell Sea to highest 0.78 in the Indian Ocean (Table 1).

This finding is consistent with earlier work, which shows that Antarctic leads tend to recur in areas influenced by bathymetric steering, oceanic fronts, and coastal divergence zones (Reiser et al., 2019; Dubey et al., 2025a). The highest lead frequencies occur along the coastline, especially near the Antarctic Slope Currents, continental shelf breaks, and major ridges. The continental slope plays a key role in shaping spatial lead patterns, with ice divergence in these regions driven by the Antarctic Slope Currents and tidal flows (Stewart et al., 2019; Heil et al., 2009; Hutchings et al., 2012). These regions coincide with well-known flaw-lead zones where wind-driven divergence, tidal flows, and bathymetric guidance promote new ice formation and dense shelf water production (Silvano et al., 2018; Goosse and Fichefet, 1999; Jacobs and Weiss, 1998). Likewise, the enhanced activity around features such as Maud Rise and Gunnerus Ridge reflects the influence of tidal variability, mesoscale eddies, and upwelling processes that increase sea-ice motion and divergence (Beckmann et al., 2001; Guelk et al., 2023). While the model captures the large-scale patterns of enhanced LF, it does not fully reproduce the fine-scale structures visible in the observations due to its coarse resolution. Nevertheless, the agreement at the basin scale supports the interpretation that lead formation is not governed by a single forcing. Instead, it is anchored in bathymetrically-controlled deformation regimes that provide the structural template on which atmospheric and oceanic processes act (Holland and Kwok, 2012; Kurtz et al., 2013).

When leads open during winter, they initiate intense air–sea energy exchanges that can significantly modify the upper-ocean stratification (Venables and Meredith, 2014; Cohanim et al., 2021) as well as the structure of the atmospheric boundary layer (Heinemann et al., 2022; Marcq and Weiss, 2012). In this context, the role of leads in shaping the feedback between sea-ice, atmosphere and ocean must not be neglected and requires a robust identification of forcings, for which we think the presented lead climatology will yield a useful basis.

4.2 Atmosphere and leads

Atmospheric forcing variables collectively represent the largest contribution to model skill. 2 m air temperature accounts for 20 % of the explained variance in LF across the Southern Ocean, reaching as high as 31 % in the Indian Ocean sector (Table 1). While lower air temperatures thermodynamically promote rapid refreezing, our data indicate that high lead frequencies often coincide with lower 2 m air temperatures. This inverse relationship likely reflects the influence of strong, cold offshore winds, which mechanically open coastal leads and polynyas while simultaneously subjecting them to intense cooling and refreezing (Mathiot et al., 2012; Morales Maqueda et al., 2004). Thus, we hypothesize that 2 m air temperature serves as an indirect effect and a proxy for the thermodynamic processes that the offshore wind forcing drives the lead formation.

Wind divergence and speed also emerge as dominant physically interpretable predictors once the influence of thermodynamic preconditioning is accounted for. Together, these wind variables contribute roughly 24 % of the total permutation importance across the Southern Ocean (Fig. 5a), with particularly high impact during mid-winter June, reaching about 27 % combined in the Southern Ocean (wind speed: 14 %, wind divergence: 13 %), higher in the Weddell Sea (21 % from wind divergence alone), and Ross Sea (37 %, from wind divergence: 23 % and wind speed: 14 %) (Figs. 10–11). This agrees with some studies showing that divergent wind stress promotes tensile failure in the ice cover by weakening internal ice strength and enhancing shear deformation (Simmonds et al., 2008; Kimura and Wakatsuchi, 2000).

The seasonal change of wind importance can be related to the changing mechanical state of the ice pack throughout winter. The monthly analysis (Fig. 11) reveals that atmospheric control is higher during early to mid-winter. In April, wind divergence and wind speed together contribute ∼29 % at the Southern Ocean scale, with similar patterns in May (∼26 %). The pronounced atmospheric influence in early winter reflects the vulnerability of newly formed ice, which responds more readily to synoptic-scale pressure patterns and wind-induced divergence (Hutchings et al., 2011; Heil and Allison, 1999). The June case study further illustrates that wind forcing remains influential after air temperature control, particularly in the Weddell and Ross Seas regions, where offshore outflow and topographic channelling near the coast strengthen wind–ice coupling.

At the same time, wind divergence and speed are most effective in regions already preconditioned by the mechanical state of the ice pack and long-term deformation patterns. This underscores the interplay between dynamic atmospheric forcing and the underlying ice mechanical properties, while winds initiate lead opening, their impact depends on the existing ice strength, consolidation state, and accumulated internal stress.

4.3 Sea-ice motion and leads

Ice velocity, ice divergence, and ice concentration provide additional model skill, contributing about 29 % to overall model performance (Fig. 5a). Ice divergence alone contributes 12 %, while ice velocity and ice concentration add 10 % and 7 %, respectively. Despite this performance, the importance of ice predictors, along with their regional contrasts (Fig. 5), is consistent with the well-established link between deformation processes and lead formation.

The Weddell Sea shows relatively high ice kinematic importance (15 % for ice divergence and 9 % for ice velocity), in line with its persistent Weddell Gyre and shear zones and characteristic cyclonic drift patterns (Haumann et al., 2016). In contrast, the higher importance in the Ross Sea (11 % for ice velocity after wind and ocean forcing, Table 1) likely reflects the Ross Gyre and hence higher LF variability captured within monthly-mean gridded products in this region. Ice concentration plays a notable role in the Bellingshausen-Amundsen Seas (10 %), acting as a proxy for the mechanical weakness of the pack ice.

Ice concentration specifically emerges as a critical indicator of the ice pack's mechanical state, representing the degree of consolidation and ice cover. Lower ice concentrations indicate more fractured, loosely connected ice that is mechanically weaker and more susceptible to lead formation under divergent stress (Wilchinsky and Feltham, 2011; Lüpkes and Gryanik, 2015; Frew et al., 2025). The importance of ice concentration varies by month and region. While it contributes only 7 % to overall model skill (Table 1), it becomes more influential in specific months and regions, reaching 17 % in the Weddell Sea during June and 20 % in September and rising to 24 % for July in the Ross Sea (Fig. 11). Overall, this pattern aligns with the importance of ice preconditioning; regions with persistently thin ice concentration are more prone to lead opening, while thick consolidated pack ice resists deformation.

Although the presented dataset does not explicitly include polynya openings, it remains highly complementary to the long-term pan-Antarctic polynya record described in Lin et al. (2024). The occurrence and variability of Southern Ocean polynyas, and the associated ice production and brine release, have been the focus of numerous studies (e.g., Golledge et al., 2025; Macdonald et al., 2023; Paul et al., 2015; Tamura et al., 2008; Tamura et al., 2016). However, most of this work has not considered the role of small-scale leads within the pack ice, largely because the satellite data used were unable to resolve surface features at km-scale resolution or below.

4.4 Ocean and leads

Ocean currents show an increase in importance on lead formation as winter progresses, with their contribution rising from around 7–13 % in April–May to roughly 14–15 % in August–September, and reaching up to 19–23 % in specific regions during June, August and September (Fig. 11). While this increase in regional points is notable, it suggests that as the ice thickens and internal stress builds, lead formation becomes more controlled by persistent ocean-driven shear and ice–ocean momentum exchange. The June analysis reveals that ocean current speed contributes 20 % in the Indian Ocean and 15 % in the Pacific Ocean, demonstrating particularly higher oceanic control during mid-winter peak lead activity (Fig. 10). This regional elevation in ocean current importance suggests that the thermodynamically-controlled Indian and Pacific sectors, with enhanced coastal currents and eddy activity, experience enhanced ice–ocean coupling relative to wind-dominated sectors. Ocean currents interact with underlying bathymetry to induce both direct stress intensification (which opens leads locally) and indirect mechanical weakening through long-term traction and cumulative stress. The influence of ocean currents has been highlighted in several Antarctic studies, which show that current shear and bathymetrically guided flows modulate ice-drift pathways and contribute to deformation fields that shape lead networks over long timescales (Reiser et al., 2019; Holland and Kwok, 2012; Dubey et al., 2025a).

The late-winter strengthening of oceanic influence reflects the increasing role of sustained bathymetrically guided flows and intensified ice–ocean coupling beneath thickened ice. The Antarctic Coastal Current, Weddell Gyre circulation, and mesoscale eddy fields all contribute to current-induced deformation, which becomes increasingly important as winter progresses. In the Ross, Pacific and Indian sectors, where interactions among the Antarctic Coastal Current, Ross Gyre circulation, and variability in the Amundsen Sea Low are especially pronounced (Meehl et al., 2019; Turner et al., 2017), the elevated oceanic importance (Figs. 5b-e) suggests that dynamic preconditioning combined with large-scale circulation variability plays a key role in maintaining lead-prone zones. The Weddell and Ross Seas' higher ocean current influence ($\sim$12–13 %; Table 1) reflects the circulation structure and the importance of bathymetry, Antarctic Slope Current and gyre dynamics in these regions.

4.5 Regional predictors and leads

The regional contrasts in predictor contributions are shown in Fig. 5 and Table 1. The Weddell Sea shows a good mechanical imprint (16 % wind divergence, 15 % ice divergence), consistent with its wind divergence and quasi-stationary gyre circulation and recurrent lead activity (Wang et al., 2023; Cheon et al., 2014). During June, the Weddell Sea's mechanical character strengthens further, with wind divergence reaching 21 % and ice concentration and divergence contributing 17 and 14%, respectively (Fig. 10b), reflecting the consolidated mid-winter ice pack's mechanical response to wind forcing. The Indian Ocean sector exhibits the highest 2 m air temperature importance (31 %), reflecting strong north–south gradients imposed by zonal storm tracks and meridional bathymetric contrasts (Shi et al., 2023). During June, the Indian Ocean maintains strong thermodynamic control (24 %) while ocean current speed becomes particularly influential (20 %), suggesting the coupled atmosphere-ocean control on lead formation in this sector (Fig. 10c). The Pacific sector displays a distribution of predictors (2 m air temperature 30 %, ocean current speed 13 %), shaped by the influence of the Amundsen Sea Low (Turner et al., 2017), complex seafloor topography, and variable ice-drift pathways (Meehl et al., 2019). The Ross Sea shows comparatively higher model skill (r = 0.73, Fig. 2e), which likely stems from its intense episodic offshore wind events (Heil et al., 2009) and polynya activity (Silvano et al., 2018) that operate at temporal scales finer than monthly resolution and are therefore excluded from monthly averaged products. In June, the Ross Sea exhibits particularly high wind divergence influence (23 %), reflecting its dynamic coastal and shelf-break environment (Fig. 10e). The Bellingshausen–Amundsen Seas remain primarily thermodynamically controlled (2 m air temperature 25 %), with secondary contributions from seasonal evolution (month at 12 %), indicating that seasonal changes in ice thickness and thermodynamic wind-controlled regimes are particularly important for this region (Elliot, 2014).

The consistency of 2 m air temperature importance across all regions (14–31 %), combined with marked regional variations in secondary predictors, indicates that thermodynamic control is fundamental while mechanical ice and oceanic forcings provide region-specific modulations.

Large-scale climate modes such as the Southern Annular Mode (SAM), El Niño–Southern Oscillation (ENSO), and the Indian Ocean Dipole (IOD) are well known to shape Antarctic sea-ice extent and the position of the marginal ice zone by modulating atmospheric circulation, wind fields, and ocean currents (Hall and Visbeck, 2002; Yuan, 2004; Blanchard-Wrigglesworth et al., 2021). These same processes are likely to influence the sea-ice leads by altering patterns of divergence and ice drift. When considering LF trends, the atmosphere is usually seen as the dominant driver, while the ocean largely determines how those trends vary from season to season over the years (Holland and Kwok, 2012; Hobbs et al., 2016).

Understanding the connection between lead variability, short-lived atmospheric events, and slow, pan-Antarctic scale climate shifts remains an important avenue for future work, particularly in the context of the variability observed in Antarctic sea-ice since 2016.

5 Conclusions

This study provides a quantitative assessment of the physical drivers governing wintertime Southern Ocean LF using an RF regression framework. By combining a 21-year gap-filled satellite-derived LF dataset along with atmospheric, sea-ice, and oceanic predictor variables in the model, we identify and rank the processes controlling lead formation across the Southern Ocean and five regional sectors. The overall winter model reconstructs the observed LF over the April–September period from

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

2003 to 2023, achieving a basin-wide correlation of 0.70 on an independent test dataset, with regional correlation ranging from 0.68 to 0.78. The model captures the large-scale spatial pattern of elevations in LF across the Southern Ocean, although the coarse 2° × 5° grid smooths out the realistic representation of fine-scale lead hotspots, including bathymetrically- controlled and coastal lead regions. This RF performance suggest that the selected predictors capture the dominant physical mechanisms shaping Southern Ocean lead variability. Permutation-based importance analysis from the overall model reveals a clear order of LF drivers. 2 m air temperature emerges as the most influential predictor (20 %), highlighting the fundamental role of thermodynamic control and its coupling with atmospheric forcing, probably offshore winds. This likely reflects an indirect dynamical effect, where sharp temperature drops serve as a proxy for the cold, offshore air masses that mechanically drive coastal leads. Wind and ice divergence, and ocean current speed each contribute ∼11–13 %, suggesting that lead formation is governed by coupled atmosphere-ice-ocean interactions rather than any single process. Regional analyses further underscore the heterogeneous nature of Southern Ocean lead dynamics. The Indian and Pacific Ocean sectors are controlled by strong air temperature with ocean currents, and the Bellingshausen–Amundsen Seas are mainly seasonally thermodynamically wind-influenced, whereas the Weddell Sea exhibits a dominant mechanical influence from wind and ice divergence. The Ross Sea reflects a balanced contribution from air temperature, wind divergence, and ocean currents, consistent with its dynamic coastal and shelf-break activity.

Monthly analyses reveal seasonal shifts in driver dominance. April–June, lead variability is most sensitive to atmospheric forcing, including air temperature and wind fields, as well as ice divergence and velocity. In contrast, July–September shows an increasing influence from persistent thermodynamic offshore wind forcing, ocean currents, and consolidated ice-pack mechanics. The rising importance of ocean current (∼15 %) toward late winter shows the growing role of ice–ocean momentum exchange and bathymetrically guided circulation in sustaining leads within the thick winter ice-pack.

The RF model provides a quantitative ranking of LF predictors that was previously unavailable for the Southern Ocean. Our results show that Southern Ocean lead formation arises from coupled, non-linear interactions among the atmosphere, ocean, and sea-ice, rather than from episodic wind forcing alone. The presented framework offers a new insight for future work, linking lead variability to large-scale climate modes and improving lead-related fluxes under changing Antarctic sea-ice conditions. In addition, an analysis using the RF model for small-scale regional features such as bathymetrically-controlled and coastal lead zones (Dubey et al., 2025a) could also be part of future studies.

Appendix A

A1 Southern Ocean regional sectors for LF analysis

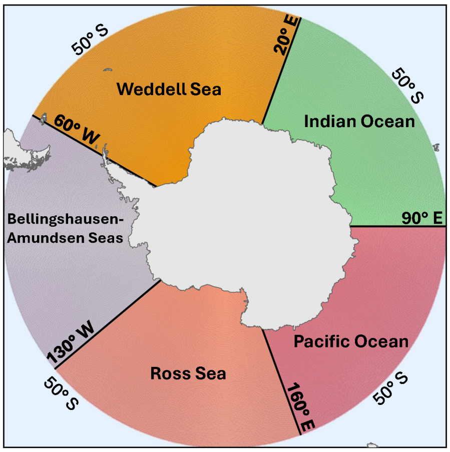

Figure A1. Map of the Southern Ocean showing the five regional sectors used in this study: Weddell Sea (orange), Indian Ocean (green), Pacific Ocean (pink), Ross Sea (coral), and Bellingshausen–Amundsen Seas (grey). The sectors are defined by meridional longitude boundaries, and major latitude circles (50° S) are shown for geographic reference.

Data availability. The monthly lead frequency dataset for the Southern Ocean (April–September, 2003–2023) is available on PANGAEA (AntLeads: Monthly wintertime sea-ice lead maps for the Antarctic, April–September, 2003–2023, https://doi.org/10.1594/PANGAEA.977634) (Dubey et al., 2025b). Atmospheric data were downloaded from the ERA5 reanalysis (Hersbach et al., 2020) produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) and obtained from the Copernicus Climate Data Store (CDS), available at https://cds.climate.copernicus.eu/, last access: 5 January 2026. Sea-ice motion data were obtained from the Polar Pathfinder Daily Sea Ice Motion product (Tschudi et al., 2019) provided by the National Snow and Ice Data Center (NSIDC) at https://nsidc.org/data/explore-data, last access: 5 January 2026. Sea-ice concentration data were provided by the University of Bremen (Spreen et al., 2008), available at https://www.meereisportal.de/, last access: 5 January 2026, and ORAS5 data (Zuo et al., 2019) were acquired from the CDS, https://cds.climate.copernicus.eu/, last access: 5 January 2026).

Author contributions. UD analyzed the data and wrote the main script. SW contributed to the research design. SW and GH contributed to the interpretation of results, manuscript preparation and revisions. The final version was prepared with contributions from all co-authors. All authors have read and agreed to the submitted version of the manuscript.

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

Acknowledgements. The authors would like to thank the ECMWF, NSIDC and the University of Bremen for providing the data used in this study.

Financial support. The research was funded by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority programme "Antarctic Research with comparative investigations in Arctic ice areas" (SPP1158) under grant WI 3314/6-1. The publication was funded/-supported by the Open Access Fund of Trier University and the German Research Foundation (DFG).

References

Alam, A. and Curry, J. A.: Determination of surface turbulent fluxes over leads in Arctic sea ice, Journal of Geophysical Research, 102, 3851–3860, https://doi.org/10.1029/96JC03606, 1997.

Beckmann, A., Timmermann, R., Pereira, A. F., and Mohn, C.: The effect of flow at Maud Rise on the sea-ice cover numerical experiments, Ocean Dynamics, 52, 0011–0025, https://doi.org/10.1007/s10236-001-8173-5, 2001.

Biau, G.: Analysis of a random forests model, J. Mach. Learn. Res., 13, 1063–1095, 2012.

Blanchard-Wrigglesworth, E., Roach, L. A., Donohoe, A., and Ding, Q.: Impact of winds and Southern Ocean SSTs on Antarctic sea ice trends and variability, Journal of Climate, 34, 949–965, https://doi.org/10.1175/JCLI-D-20-0386.1, 2021.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/a:1010933404324, 2001.

Cheon, W. G., Park, Y.-G., Toggweiler, J. R., and Lee, S.-K.: The relationship of Weddell Polynya and open-ocean deep convection to the Southern Hemisphere westerlies, Journal of Physical Oceanography, 44, 694–713, https://doi.org/10.1175/JPO-D-13-0112.1, 2014.

Chi, J. and Kim, H.-c.: Prediction of Arctic Sea Ice Concentration Using a Fully Data Driven Deep Neural Network, Remote Sensing, 9, 1305, https://doi.org/10.3390/rs9121305, 2017.

Cohanim, K., Zhao, K. X., and Stewart, A. L.: Dynamics of eddies generated by sea ice leads, Journal of Physical Oceanography, 51, 3071–3092, https://doi.org/10.1175/JPO-D-20-0169.1, 2021.

Curry, J. A., Schramm, J. L., and Ebert, E. E.: Sea Ice-Albedo Climate Feedback Mechanism, Journal of Climate, 8, 240–247, https://doi.org/10.1175/1520-0442(1995)008<0240:siacfm>2.0.co;2, 1995.

Damm, E., Helmke, E., Thoms, S., Schauer, U., Nöthig, E., Bakker, K., and Kiene, R. P.: Methane production in aerobic oligotrophic surface water in the central Arctic Ocean, Biogeosciences, 7, 1099–1108, https://doi.org/10.5194/bg-7-1099-2010, 2010.

Dubey, U., Willmes, S., and Heinemann, G.: Southern Ocean sea-ice leads: first insights into regional lead patterns, seasonality, and trends, 2003–2023, The Cryosphere, 19, 3535–3552, https://doi.org/10.5194/tc-19-3535-2025, 2025a.

Dubey, U., Willmes, S., and Heinemann, G.: AntLeads: Monthly wintertime sea-ice lead maps for the Antarctic, April-September, 2003–2023, https://doi.org/10.1594/PANGAEA.977634, 2025b.

Elliot, D. H.: Antarctica: Global Science from a Frozen Continent. Edited by David H. W. Walton: Cambridge: Cambridge University Press, 2013. 342 pp. $55.00 (hardcover). ISBN 978-1-107-00392-7, Arctic, Antarctic, and Alpine Research, 46, 286-$ $-287,$ https://doi.org/10.1657/1938 $-4246-46.1.286,2014.$

Feltham, D. L.: Sea Ice Rheology, Annual Review of Fluid Mechanics, 40, 91–112, https://doi.org/10.1146/annurev.fluid.40.111406.102151, 2008.

Frew, R. C., Bateson, A. W., Feltham, D. L., and Schröder, D.: Toward a marginal Arctic sea ice cover: changes to freezing, melting and dynamics, The Cryosphere, 19, 2115–2132, https://doi.org/10.5194/tc-19-2115-2025, 2025.

Golledge, N. R., Keller, E. D., Gossart, A., Malyarenko, A., Bahamondes-Dominguez, A., Krapp, M., Jendersie, S., Lowry, D. P., Alevropoulos-Borrill, A., and Notz, D.: Antarctic coastal polynyas in the global climate system, Nature Reviews Earth amp; Environment, 6, 126–139, https://doi.org/10.1038/s43017-024-00634-x, 2025.

Goosse, H. and Fichefet, T.: Importance of ice-ocean interactions for the global ocean circulation: A model study, Journal of Geophysical Research, 104, 23 337–23 355, https://doi.org/10.1029/1999JC900215, 1999.

Guelk, B., Roquet, F., Naveira Garabato, A. C., Narayanan, A., Rousset, C., and Madec, G.: Variability and remote controls of the warm-water halo and Taylor Cap at Maud Rise, Journal of Geophysical Research: Oceans, 128, e2022JC019 517, https://doi.org/10.1029/2022JC019517, 2023.

Hall, A. and Visbeck, M.: Synchronous variability in the Southern Hemisphere atmosphere, sea ice, and ocean resulting from the annular mode, Journal of Climate, 15, 3043–3057, https://doi.org/10.1175/1520-0442(2002)015<3043:SVITSH>2.0.CO;2, 2002.

Haumann, F. A., Moorman, R., Riser, S. C., Sallée, J.-B., and Sarmiento, J. L.: Sea-ice transport driving Southern Ocean salinity and its recent trends, Nature, 537, 89–92, https://doi.org/10.1038/nature19101, 2016.

Heil, P. and Allison, I.: The pattern and variability of Antarctic sea-ice drift in the Indian Ocean and western Pacific sectors, Journal of Geophysical Research, 104, 15 789–15 802, https://doi.org/10.1029/1999JC900076, 1999.

Heil, P., Massom, R. A., Allison, I., Worby, A. P., and Lytle, V. I.: Role of off-shelf to on-shelf transitions for East Antarctic sea ice dynamics during spring 2003, Journal of Geophysical Research: Oceans, 114, https://doi.org/10.1029/2008jc004873, 2009.

Heinemann, G., Schefczyk, L., Willmes, S., and Shupe, M. D.: Evaluation of simulations of near-surface variables using the regional climate model CCLM for the MOSAiC winter period, Elementa: Science of the Anthropocene, 10, https://doi.org/10.1525/elementa.2022.00033, 2022.

Hersbach, H., Bell, B., Berrisford, P., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hobbs, W. R., Massom, R., Stammerjohn, S., Reid, P., Williams, G., and Meier, W.: A review of recent changes in Southern Ocean sea ice, their drivers and forcings, Global and Planetary Change, 143, 228–250, https://doi.org/10.1016/j.gloplacha.2016.06.008, 2016.

Holland, P. R. and Kwok, R.: Wind-driven trends in Antarctic sea-ice drift, Nature Geoscience, 5, 872–875, https://doi.org/10.1038/ngeo1627, 2012.

Hutchings, J. K., Roberts, A., Geiger, C. A., and Richter-Menge, J.: Spatial and temporal characterization of sea-ice deformation, Annals of Glaciology, 52, 360–368, https://doi.org/10.3189/172756411795931769, 2011.

Hutchings, J. K., Heil, P., Steer, A., and Hibler, W. D.: Subsynoptic scale spatial variability of sea ice deformation in the western Weddell Sea during early summer, Journal of Geophysical Research, 117, C01 002, https://doi.org/10.1029/2011JC006961, 2012.

Jacobs, S. S. and Weiss, R. F., eds.: Ocean, Ice, and Atmosphere: Interactions at the Antarctic Continental Margin, vol. 75 of *Antarctic Research Series*, American Geophysical Union, Washington, D.C., https://doi.org/10.1029/AR075, 1998.

Key, J., Stone, R., Maslanik, J., and Ellefsen, E.: The detectability of sea-ice leads in satellite data as a function of atmospheric conditions and measurement scale, Annals of Glaciology, 17, 227–232, https://doi.org/10.3189/s026030550001288x, 1993.

Kimura, N. and Wakatsuchi, M.: Relationship between sea-ice motion and geostrophic wind in the northern hemisphere, Geophysical Research Letters, 27, 3735–3738, https://doi.org/10.1029/2000gl011495, 2000.

Kort, E. A., Wofsy, S. C., Daube, B. C., Diao, M., Elkins, J. W., Gao, R. S., Hintsa, E. J., Hurst, D. F., Jimenez, R., Moore, F. L., Spackman, J. R., and Zondlo, M. A.: Atmospheric observations of Arctic Ocean methane emissions up to 82° north, Nature Geoscience, 5, 318–321, https://doi.org/10.1038/ngeo1452, 2012.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, The Cryosphere, 7, 1035–1056, https://doi.org/10.5194/tc-7-1035-2013, 2013.

Leng, H., He, H., Chen, D., Lin, P., Yang, Y., and Wang, Z.: Bathymetry-constrained ocean geostrophic currents play a key role in shaping the sea ice circulation in the Canada Basin, Arctic Ocean, Environmental Research Letters, 19, 094 040, https://doi.org/10.1088/1748-9326/ad6baa, 2024.

Lin, Y., Nakayama, Y., Liang, K., Huang, Y., Chen, D., and Yang, Q.: A dataset of the daily edge of each polynya in the Antarctic, Scientific Data, 11, https://doi.org/10.1038/s41597-024-03848-2, 2024.

Lüpkes, C. and Gryanik, V. M.: A stability-dependent parametrization of transfer coefficients for momentum and heat over polar sea ice to be used in climate models, Journal of Geophysical Research: Atmospheres, 120, 552–581, https://doi.org/10.1002/2014jd022418, 2015.

Lüpkes, C., Vihma, T., Birnbaum, G., and Wacker, U.: Influence of leads in sea ice on the temperature of the atmospheric boundary layer during polar lows: observational and model results, Journal of Geophysical Research, 113, D03 104, https://doi.org/10.1029/2007GL032461, 2008.

Macdonald, G. J., Ackley, S. F., Mestas-Nuñez, A. M., and Blanco-Cabanillas, A.: Evolution of the dynamics, area, and ice production of the Amundsen Sea Polynya, Antarctica, 2016–2021, The Cryosphere, 17, 457–476, https://doi.org/10.5194/tc-17-457-2023, 2023.

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

Mathiot, P., Jourdain, N. C., Barnier, B., Gallée, H., Molines, J. M., Le Sommer, J., and Penduff, T.: Sensitivity of coastal polynyas and high-salinity shelf water production in the Ross Sea, Antarctica, to the atmospheric forcing, Ocean Dynamics, 62, 701–723, https://doi.org/10.1007/s10236-012-0531-y, 2012.

Meehl, G. A., Arblaster, J. M., Chung, C. T. Y., Holland, M. M., DuVivier, A., Thompson, L., Yang, D., and Bitz, C. M.: Sustained ocean changes contributed to sudden Antarctic sea ice retreat in late 2016, Nature Communications, 10, 14, https://doi.org/10.1038/s41467-018-07865-9, 2019.

Morales Maqueda, M. A., Willmott, A. J., and Biggs, N. R. T.: Polynya dynamics: a review of observations and modeling, Reviews of Geophysics, 42, RG1004, https://doi.org/10.1029/2002RG000116, 2004.

Nihashi, S. and Cavalieri, D. J.: Observational evidence of a hemispheric-wide ice–ocean albedo feedback effect on Antarctic sea-ice decay, Journal of Geophysical Research: Oceans, 111, https://doi.org/10.1029/2005jc003447, 2006.

Ohshima, K. I., Fukamachi, Y., Williams, G. D., Nihashi, S., Roquet, F., Kitade, Y., Tamura, T., Hirano, D., Herraiz-Borreguero, L., Field, I., Hindell, M., Aoki, S., and Wakatsuchi, M.: Antarctic Bottom Water production by intense sea-ice formation in the Cape Darnley polynya, Nature Geoscience, 6, 235–240, https://doi.org/10.1038/ngeo1738, 2013.

Palerme, C. and Müller, M.: Calibration of sea ice drift forecasts using random forest algorithms, The Cryosphere, 15, 3989–4004, https://doi.org/10.5194/tc-15-3989-2021, 2021.

Paul, S., Willmes, S., and Heinemann, G.: Long-term coastal-polynya dynamics in the southern Weddell Sea from MODIS thermal-infrared imagery, The Cryosphere, 9, 2027–2041, https://doi.org/10.5194/tc-9-2027-2015, 2015.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Perovich, D. K., Grenfell, T. C., Light, B., and Hobbs, P. V.: Seasonal evolution of the albedo of multiyear Arctic sea ice, Journal of Geophysical Research: Oceans, 107, https://doi.org/10.1029/2000jc000438, 2002.

Reiser, F., Willmes, S., Hausmann, U., and Heinemann, G.: Predominant Sea Ice Fracture Zones Around Antarctica and Their Relation to Bathymetric Features, Geophysical Research Letters, 46, 12 117–12 124, https://doi.org/10.1029/2019gl084624, 2019.

https://doi.org/10.5194/egusphere-2026-514
Preprint. Discussion started: 10 February 2026

Reiser, F., Willmes, S., and Heinemann, G.: A new algorithm for daily sea ice lead identification in the Arctic and Antarctic winter from thermal-infrared satellite imagery, Remote Sensing, 12, 1957, https://doi.org/10.3390/rs12121957, 2020.

Rheinlænder, J. W., Davy, R., Ólason, E., Rampal, P., Spensberger, C., Williams, T. D., Korosov, A., and Spengler, T.: Driving Mechanisms of an Extreme Winter Sea Ice Breakup Event in the Beaufort Sea, Geophysical Research Letters, 49, https://doi.org/10.1029/2022gl099024, 2022.

Shen, X.-y., Zhang, J., Meng, J.-m., Zhang, J., and Ke, C.-q.: Sea ice type classification based on random forest machine learning with Cryosat-2 altimeter data, in: 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), pp. 1–5, IEEE, https://doi.org/10.1109/rsip.2017.7958792, 2017.

Silvano, A., Rintoul, S. R., Peña-Molino, B., Hobbs, W. R., van Wijk, E., Aoki, S., Tamura, T., and Williams, G. D.: Freshening by glacial meltwater enhances melting of ice shelves and reduces formation of Antarctic Bottom Water, Science Advances, 4, https://doi.org/10.1126/sciadv.aap9467, 2018.

Simmonds, I., Burke, C., and Keay, K.: Arctic climate change as manifest in cyclone behavior, Journal of Climate, 21, 5777–5796, https://doi.org/10.1175/2008JCLI2366.1, 2008.

Smith, S. D., Muench, R. D., and Pease, C. H.: Polynyas and leads: An overview of physical processes and environment, Journal of Geophysical Research: Oceans, 95, 9461–9479, https://doi.org/10.1029/jc095ic06p09461, 1990.

Spreen, G., Kaleschke, L., and Heygster, G.: Sea ice remote sensing using AMSR-E 89-GHz channels, Journal of Geophysical Research: Oceans, 113, https://doi.org/10.1029/2005jc003384, 2008.

Stewart, A. L., Klocker, A., and Menemenlis, D.: Acceleration and Overturning of the Antarctic Slope Current by Winds, Eddies, and Tides, Journal of Physical Oceanography, 49, 2043–2074, https://doi.org/10.1175/jpo-d-18-0221.1, 2019.

Stirling, I.: The importance of polynyas, ice edges, and leads to marine mammals and birds, Journal of Marine Systems, 10, 9–21, https://doi.org/10.1016/s0924-7963(96)00054-1, 1997.

Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T.: Bias in random forest variable importance measures: Illustrations, sources and a solution, BMC Bioinformatics, 8, https://doi.org/10.1186/1471-2105-8-25, 2007.

Tamura, T., Ohshima, K. I., and Nihashi, S.: Mapping of sea ice production for Antarctic coastal polynyas, Geophysical Research Letters, 35, https://doi.org/10.1029/2007gl032903, 2008.

Tamura, T., Ohshima, K. I., Fraser, A. D., and Williams, G. D.: Sea ice production variability in Antarctic coastal polynyas, Journal of Geophysical Research: Oceans, 121, 2967–2979, https://doi.org/10.1002/2015JC011537, 2016.

Tian, T., Davy, R., Ponsoni, L., and Yang, S.: Impact of modulating surface heat flux through sea ice leads on Arctic sea ice in EC-Earth3 in different climates, The Cryosphere, 19, 2751–2768, https://doi.org/10.5194/tc-19-2751-2025, 2025.

Tschudi, M., Meier, W. N., Stewart, J. S., Fowler, C., and Maslanik, J.: Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 4, NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, https://doi.org/10.5067/INAWUWO7QH7B, 2019.

Turner, J., Phillips, T., Marshall, G. J., Hosking, J. S., Pope, J. O., Bracegirdle, T. J., and Deb, P.: Unprecedented springtime retreat of Antarctic sea ice in 2016, Geophysical Research Letters, 44, 6868–6875, https://doi.org/10.1002/2017GL073656, 2017.

Venables, H. J. and Meredith, M. P.: Feedbacks between ice cover, ocean stratification, and heat content in Ryder Bay, western Antarctic Peninsula, Journal of Geophysical Research: Oceans, 119, 5323–5336, https://doi.org/10.1002/2013JC009669, 2014.

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

von Albedyll, L., Hendricks, S., Grodofzig, R., Krumpen, T., Arndt, S., Belter, H. J., Birnbaum, G., Cheng, B., Hoppmann, M., Hutchings, J., Itkin, P., J"akel, E., Jutila, A., Katlein, C., Kolabutin, N., Lei, R., Nicolaus, M., Ricker, R., Rohde, J., Suhrhoff, M., Timofeeva, A., Watkins, D., Webster, M., and Haas, C.: Link between multi-year sea ice and atmospheric boundary layer feedbacks in the Arctic, Journal of Geophysical Research: Oceans, 127, e2022JC018 774, https://doi.org/10.1029/2022JC018774, 2022.

Wang, Q., Danilov, S., Jung, T., Kaleschke, L., and Wernecke, A.: Sea ice leads in the Arctic Ocean: Model assessment, interannual variability and trends, Geophysical Research Letters, 43, 7019–7027, https://doi.org/10.1002/2016gl068696, 2016.

Wang, Y., Ji, Q., Pang, X., Qu, M., Cha, M., Zhang, F., Yan, Z., and He, B.: Distribution Characteristics and Influencing Factors of Sea Ice Leads in the Weddell Sea, Antarctica, Remote Sensing, 15, 5568, https://doi.org/10.3390/rs15235568, 2023.

Wilchinsky, A. V. and Feltham, D. L.: Modeling Coulombic failure of sea ice with leads, Journal of Geophysical Research, 116, https://doi.org/10.1029/2011jc007071, 2011.

Willmes, S., Heinemann, G., and Schnaase, F.: Patterns of wintertime Arctic sea-ice leads and their relation to winds and ocean currents, The Cryosphere, 17, 3291–3308, https://doi.org/10.5194/tc-17-3291-2023, 2023.

Yuan, X.: ENSO-related impacts on Antarctic sea ice: a synthesis of phenomenon and mechanisms, Antarctic Science, 16, 415–425, https://doi.org/10.1017/S0954102004002238, 2004.

Zhang, J.: Modeling the Impact of Wind Intensification on Antarctic Sea Ice Volume, Journal of Climate, 27, 202–214, https://doi.org/10.1175/jcli-d-12-00139.1, 2014.

Zhang, L., Shi, Q., Lepp"aranta, M., Liu, J., and Yang, Q.: Estimating winter Arctic sea ice motion based on random forest models, Remote Sensing, 16, 581, https://doi.org/10.3390/rs16030581, 2024.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.