

Data Files Accompanying: Machine learning for daily forecasts of Arctic sea-ice motion: an attribution assessment of model predictive skill.

Authors: Hoffman, L; Mazloff, M. R.; Gille, S. T.; Giglio, D.; Bitz, C. M.; Heimbach, P.; Matsuyoshi, K.

Contact: Lauren Hoffman. lahoffma@eng.ucsd.edu. <https://orcid.org/0000-0002-9563-8925>. Scripps Institution of Oceanography and Department of Chemical Engineering, University of California, San Diego, La Jolla, CA 92037.

Cite This Data: <https://doi.org/10.6075/J0X06774>

This data package includes the data and analysis code associated with the manuscript listed above. The directory contains (i) links to raw data and (ii) processed data from five sources:

- (1) Polar Pathfinder Version 4 Daily Sea Ice Motion Vectors;
- (2) Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data;
- (3) Japanese 55-year Reanalysis derived for ocean-ice models (JRA55-do);
- (4) International Bathymetric Chart of the Arctic Ocean (IBCAO);

This data package also includes data generated from the following source:

- (5) Modeled outputs from a set of statistical models using MATLAB and Tensorflow/Keras;

Below are lists of the data and MATLAB R2021b and Tensorflow/Keras codes used for processing and plotting (Tables 1-3). Descriptions of processing steps are listed below. Figures can be generated by running the codes listed in Table 3 after the processing listed below has been completed. Raw data files are not included due to their size. Instead, all necessary steps to download and reproduce results are provided. Processed data are provided so that figures can be reproduced.

Required software packages: (i) MATLAB R2021b and (ii) Python 3.8-3.11 (iii) TensorFlow 2

**Note: Make sure to change the descriptor 'loadpath' in each file that is run to match the filepath you're accessing on your machine.*

Label	File Name	Source Data
0.	Raw Data	
R1	Polar Pathfinder Version 4 Daily Sea Ice Motion Vectors	
R2	Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave	
R3	Japanese 55-year Reanalysis derived for ocean-ice models (JRA55-do)	

R4	International Bathymetric Chart of the Arctic Ocean (IBCAO)	
<p>I. Download & Process Raw Data [P1_process_raw_data.m]: This code runs files PX-PX to download the raw data. Make sure to open and read the file, as the various datasets have different download methods (i.e. direct link vs. ftp, etc.).</p>		
P1	D_download_seaicemotion_loop	R1
P2	D_download_jra55do	R3
P3	D_nsidc-download_NSIDC-0051.001_2021-02-04.py	R2
P4	https://www.gebco.net/data_and_products/gridded_bathymetry_data/arctic_ocean/	R4
P5	D_seaicemotion_8921.m	R1
P6	D_read_JRA55_do_wind.m	R3
P7	D_hourly_arctic_JRA55_do_wind_8921.m	R3
P8	D_daily_arctic_JRA55_do_wind_8921.m	R3
P9	D_regrid2EASE_wind.m	R3
P10	D_read_iceconc_8921.m	R2
P11	D_read_sic_grid.m	R2
P12	D_regrid2EASE_sic_8921.m	R2
P13	D_read_IBCAO.m	R4
P14	D_regrid2EASE_IBCAO.m	R4
P15	D_seaicemotion_uncertainty_7821.m	R1
P16	D_uncertainty_counts_7921.m	R1
<p>II. Process data for input into statistical models [P2_data4model.m]: This file runs code PX-PX, which take the downloaded raw data and prepare it for input into the statistical models</p>		
P17	S1_satellite_normalize_NaN_TVT_Mn.m	R1-3
P18	S1_satellite_normalize_NaN_Te_properties_Mn.m	R1-3
P19	S1_input_anomaly_maps.m	R1-3
P20	D_sic_threshold.m	R1-3
P21	D_coast_distance.m	R4
<p>III. Run statistical models [P3_process_runmodels.m]: This file runs code PX-PX to run the various statistical models. Make sure to open and read the file, as the various models are run in different coding environments (i.e. MATLAB vs. python).</p>		
P22	M_PS_loop_v1.m	R1-3
P23	M_LR_global_loop_complex_ridge_NaN_v1.m	R1-3
P24	M_LR_gridwise_loop_complex_ridge_NaN_v1.m	R1-3
P25	M_CNN_TVT_loop_reg_NaN_vX.py	R1-3

IV. Model Evaluation [P4_process_modevaluation.m]: This file runs code PX-PX to evaluate the model outputs and generate the data for plotting that are included in this DOI (listed in Table 2).		
P26	S2_ensemble_mean_overall_skill_corr_PS_LR_CNN_zm.m	R1-3
P27	S2_ensemble_mean_overall_skill_corr_LR_global.m	R1-3
P28	S2_PS_test_pred_zm.m	R1-3
P29	S2_LR_test_pred_zm.m	R1-3
P30	S2_CNN_test_pred_zm.m	R1-3
P31	S2_LR_parameters.m	R1-3
P32	S3_M3_map_significant.m	R1-3
P33	S3_metric_diff_divisions.m	R1-3
P34	S3_ensemble_mean_monthlyline_division_prop.m	R1-3
P35	S3_ANOVA_monthlyline.m	R1-3
P36	S3_metrics_input_percentiles05.m	R1-3
P37	S3_interdivisional_metrics.m	R1-3
P38	S3_interdivisional_props.m	R1-4
P39	S3_ensemble_mean_daily_division_prop.m	R1-4
P40	S3_corr_metrics_props_ensemble.m	R1-4

Table 1. Processing files.

Label	File Name
D1	siconc_threshold_20_0.mat
D2	grid_EASE.mat
D3	coastdistance.mat
D4	depth_avgEASE.mat
D5	ensemblemean_overall_statistics_zm_NaN_PS_LR_CNN_8921_C.mat
D6	ensemblemean_overall_meanstd_LRglobal_zzm_zersi_8921.mat
D7	M3_significance_CNN_LR_PS_zm_NaN_C.mat
D8	inputs_anomalyC_map.mat
D9	ANOVA_monthlyline.mat
D10	yearly_monthly_line_property_division_absdsic_NaN_zm_C.mat
D11	input_percetiles05_prop_metrics_zm_NaN_CNN_LR_PS_C.mat
D12	LR_complex_ensemble_mean_noparam_R5.mat
D13	interdivisional_props_C.mat
D14	interdivisional_metrics_C.mat
D15	ensemble_spatial_zm_corr_metric_prop_C.mat
D16	ensemble_temporal_corr_metric_prop_C.mat
D17	PP_uncertainty_counts_removeNaN_7921.mat

D18	LR_complex_ensemble_mean_noparam_icewind_R5.mat
-----	---

Table 2. Processed data files.

Figure	File Name	Source Data
T1	T1_metrics_overall.m	D5, D6
F1	F1_map_intro.m	D1-D4, D7
F3	F3_map_metrics_props.m	D1-3, D7, D8
F4	F4_pdf_metrics_depth_dist.m	D1-3, D7
F5	F5_temporal_metrics_props.m	D9, D10
F6	F6_input_percentiles.m	D11
F7	F7_map_LRparams.m	D12
F8	F8_pdf_metrics_LRparams.m	D1, D7, D12
F9	F9_interdivisional.m	D13, D14
F10	F10_corr_metrics_props.m	D15, D16
S1	S1_uncertainty_PP.m	D17
S2	S2_LRparam_icewind	D18

Table 3. Figure generation files.