Project Overview

This project employs a linear regression technique to assess the relationship between two atmospheric modes of variability (MOVs) that strongly influence the climate of West Antarctica: the Amundsen Sea Low (ASL) and the Southern Annular Mode (SAM). An optimized linear regression (OLR) model is used to visualize the relationship between the two MOVs, and a built-in scikit-learn linear regression function is employed to measure how well different multiple linear regression models predict target values when additional MOVs are included as model features. Our results indicate that linear regression models cannot fully explain the monthly variability observed in our ASL dataset ($R^2 < 0.25$). More importantly, they suggest the need to use distinct machine learning techniques that better capture the nonlinearity of atmospheric MOVs. Lastly, we recognize that this project does not encompass all MOVs known to influence circulation across West Antarctica and propose the inclusion of a wider breadth of MOVs in future studies that employ machine learning techniques to study the ASL.

Introduction

The ASL is a climatological low-pressure center in the Pacific Sector of the Southern Ocean (SO) between 60-180°W. It is the strongest of three low-pressure centers that encircle Antarctica and results from the complex topography and asymmetric geography of the southernmost continent (Baines and Fraedrich 1989; Raphael et al. 2016). As a regional MOV, the ASL strongly influences the near-surface wind field of West Antarctica. In general, northerly flow along the ASL's eastern flank advects relatively warm mid-latitude air to the west of the Antarctic Peninsula while southerly flow along the ASL's western flank advects cold continental air equatorward across the Ross Sea. The combined dynamic and thermodynamic effects of the ASL's cyclonic circulation can result in significant surface temperature (ST) and sea ice concentration (SIC) anomalies that alter the climatic conditions of West Antarctica (Fig. 1). These changes contribute to the mean state of Antarctica, which furthermore regulates global climate patterns.

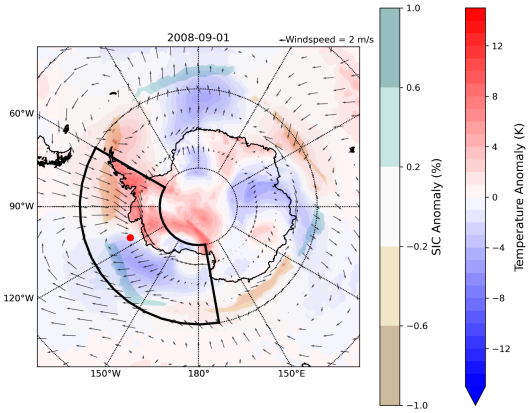

Figure 1: An anomaly plot showing changes in ST and SIC during a strong ASL event in September 2008. Cyclonic circulation surrounds the ASL center (red point).

The depth of the ASL is largely modulated by the SAM, the leading mode of atmospheric variability across the Southern Hemisphere (SH). SAM indices generally relay the phase and strength of the westerly winds that encircle Antarctica by calculating the zonally averaged sea level pressure (SLP) difference between the middle (40°S) and high (65°S) latitudes. In its positive phase (SAM+), the westerly winds strengthen and displace the southern storm track poleward, an action that induces negative (positive) pressure anomalies at high (mid) latitudes

that deepen the ASL (Fogt and Marshall 2020; Thompson and Wallace 2000). Conversely, the negative phase of the SAM (SAM-) is characterized by positive (negative) pressure anomalies near Antarctica (the mid latitudes) that weaken the ASL (Fogt and Marshall 2020). In recent years, increasingly positive SAM conditions have been observed, likely a result of changes in stratospheric ozone levels and greenhouse gas concentrations (Fig. 2) (Arblaster and Meehl 2006).

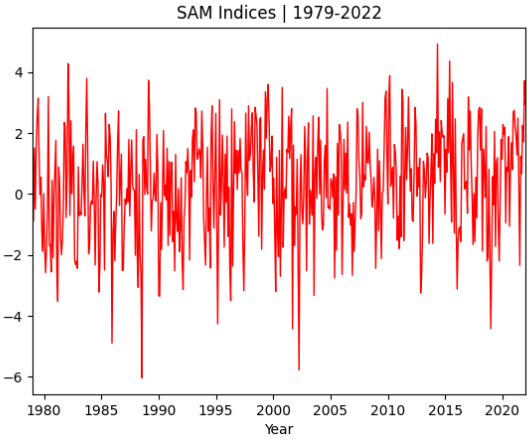

Figure 2: A timeseries of SAM conditions from 1979-2022. SAM polarity has become increasingly positive in recent years.

Our project considers the predictability of ASL actual central pressure (ACP) given the phase and strength of the SAM over the period 1979-2022. We employ a supervised machine learning technique, specifically linear regression, to elucidate the relationship between the two MOVs. After identifying a relatively weak linear relationship between the ASL and the SAM, values representing the phase and strength of additional SH MOVs like El Niño Southern Oscillation (ENSO) and Zonal Wave 3 (ZW3) are included as features in two multiple linear regression models. Statistical metrics vary slightly based on different MOV combinations and will be explored further in the following sections.

Data and Methods

ASL ACP values are obtained from Scott Hosking's ASL Climate Index (https://scotthosking.com/asl_index), which derives variables from fifth-generation European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA5). A stronger ASL is represented by deeper (lower) ACP values and weaker ASLs are denoted by shallow (higher) values.

SAM values are station-based and can be found by visiting Gareth Marshall's SAM Index at: https://climatedataguide.ucar.edu/climate-data/marshall-southern-annular-mode-sam-index-station-based. Positive (negative) index values indicate SAM+ (SAM-) conditions, which in general

coincide with negative (positive) pressure anomalies in the region of West Antarctica that intensify (debilitate) the strength of the ASL.

Although the relationship between the ASL and the SAM is the primary consideration of our project, two secondary atmospheric MOVs that have been found to influence the circulation of West Antarctica are also included for exploratory purposes. The first of these MOVs is ENSO, a climate pattern with origins in the tropical Pacific whose convective behavior at the equator can instigate the propagation of Rossby waves to West Antarctica. Previous studies (Fogt et al. 2011; Wang et al. 2023) have found the cold (warm) phase of ENSO to have a statistically significant relationship with ASL intensity when in-phase with a positive (negative) SAM. While ENSO conditions can be gauged from a number of indices that use a variety of methods to capture ENSO phase and intensity, this project uses values from the National Oceanic and Atmospheric Administration's (NOAA) Southern Oscillation Index (SOI) to represent ENSO-driven changes directly related to the atmosphere. SOI values reflect the standardized difference in SLP between Tahiti and Darwin, Australia and represent ENSO's cold (warm) phase when positive (negative).

Lastly, we consider ASL intensity as it relates to ZW3, an atmospheric MOV characterized by three distinct ridges and troughs that enable the meridional transfer of heat and energy between the middle and high southern latitudes. Raphael (2004) developed a ZW3 index in which monthly mean SLP (MSLP) and 500-hPa geopotential heights are used to determine the strength of the MOV. Index values span the period 1979-2015 and represent the three-month average of three grid points that define the location of each of the three ridges identified by van Loon and Jenne (1972). Positive (negative) values indicate strong (weak) ZW3 conditions when meridional heat flux is large (small).

Prior to modeling, ASL, SAM, and SOI values were compiled into a comma-separated values (CSV) file at a monthly time step for accessibility purposes. ZW3 values were accessed separately from an .nc file provided by Raphael (2004).

Standard regression metrics used to measure the accuracy of our models include root mean square error (RMSE) and coefficients of determination (R^2). RMSE values detail the fit of each model and are generated by square-rooting the average difference between the observed and predicted values. In general, lower RMSE values are preferred as they indicate a smaller difference between observations and predictions. In a similar fashion, R^2 values indicate the goodness of fit of our modeled results; values closer to zero (one) represent a lower (higher) degree of model accuracy. Due to large RMSE values (>5) across all linear and multilinear regression models employed in this analysis, only R^2 values are displayed in our results.

Modeling

```
# Linear regression with consideration for SAM (x) only

SAM = np.array(MOV_data['SAM'])

X = SAM
X = X.reshape(-1,1)
print(X.shape)
y = np.array(MOV_data['ASL'])

print(y.shape) #ensuring dimensional compatibility
print()

model = LinearRegression()

model.fit(X, y)

y_hat_0 = model.predict(X)

rmse_0 = np.sqrt(mean_squared_error(y, y_hat_0))
r2_0 = r2_score(y, y_hat_0)
```

Figure 3: A screenshot showing code related to the implementation of scikit-learn's linear regression class. No model optimization was implemented.

This project utilizes a supervised machine learning technique given that our target variable (ASL ACP) consists of continuous values that can be trained using continuous input features. In particular, we develop an optimized linear regression (OLR) model that yields an expression with two coefficients that quantify the relationship between the ASL and SAM. We then use scikit-learn's linear model module to conduct a series of multiple linear regression experiments that include secondary and tertiary MOVs like ENSO and ZW3. No model optimization is implemented as seen in Figure 3.

Results

Figure 4 shows our baseline analysis of the linear relationship between the ASL and SAM. Although both the training (blue) and test (red) datasets show some level of dispersion, there is a discernible negative correlation between ASL ACP values and SAM indices, thus indicating an increase in ASL intensity (depth) when the polarity of SAM is positive and progressively large. The coefficients contained in our OLR model expression (bottom left) reflect the expected ACP value (976.4 hPa) when SAM is equal to zero and also explain the negative decrease (-1.66) in ACP per change in SAM. A regression line informed by the model expression is shown in green (Fig. 4).

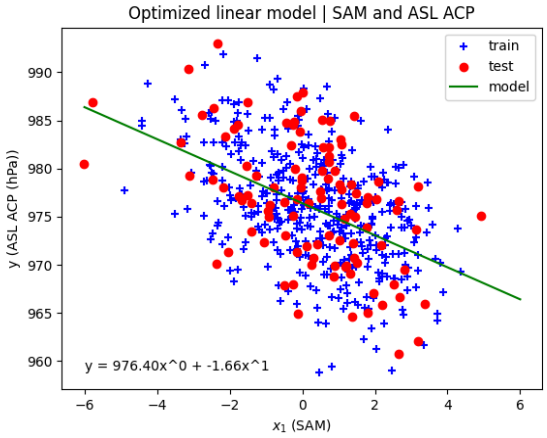

Figure 4: A scatter plot with results from our OLR model relating ASL ACP to SAM indices.

ACP observations (a) and predictions from a linear regression (b) and two multilinear regression models (c, d) are shown in Figures 5a-d. Observed ACP values demonstrate large monthly variability and yield values between 958.79 - 992.94 hPa (Fig. 5a). The mean ACP value during 1979-2022 is 976 hPa, and a visual inspection of the time series indicates a slight negative trend with ASL ACP decreasing over the 43-year study period. Figure 5b displays predictions based on a linear regression model that considers SAM indices exclusively (SAM_only). While the model successfully captures the observational mean, the magnitude of monthly variations is limited to ACP values between 968.27 - 986.62 hPa. With an R^2 value of 0.235 and a lower limit ~10 hPa larger than that of our observations, our SAM_only model results show the model's inability to fully emulate ASL variability, specifically extreme

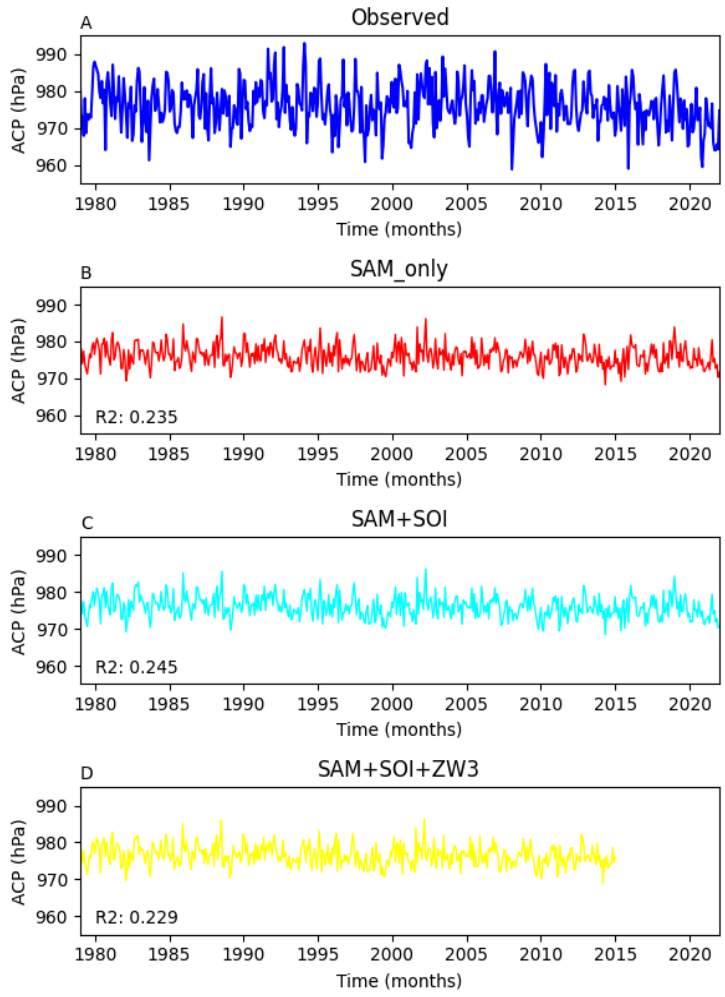

Figure 5: Plots of ASL ACP observations (a) and predictions from our linear regression (b) and multilinear regression models (c, d). In all cases, coefficients of determination show a relatively weak relationship between the ASL and different SH MOVs.

ASL events with lower- or higher-than average ACP values. Figure 5c shows predicted ACP values when a multilinear regression model of SAM and SOI indices is employed. A small improvement in predictability is reflected in the SAM+SOI model's R^2 value of 0.245, but simulated variability is similar to the variability predicted by the SAM_only model (Fig. 5b), meaning that inconsistencies in reproducing the magnitude of ASL extremes persist. A second multilinear regression model is applied to gauge ACP predictions from 1979-2015 when three MOVs - SAM, ENSO, and ZW3 (SAM+SOI+ZW3) - are featured (Fig. 5d). Minimal differences in monthly variability (with range 969 - 986.4 hPa) are identified when compared to previous regression models (Figs. 5b, 5c) and an R^2 value of 0.229 indicates reduced ACP predictability when ZW3 is considered in concert with SAM and ENSO.

Discussion and Conclusion

Our results suggest that ACP predictions based solely on SAM do not fully capture the monthly variability that characterizes the ASL over the period 1979-2022. While SAM is the primary MOV controlling the atmospheric circulation of the SH, it explains only ~23.5% of the fluctuations in the intensity of the ASL, therefore indicating a weak linear relationship between the two MOVs. Given the regional nature of the ASL, we employ a multilinear regression model that tests whether the phase and strength of more localized MOVs can improve ACP predictions. Our model based on the interaction between SAM and ENSO, whose confluence has been found to influence ASL strength (Fogt et al. 2011), shows a slight improvement in ACP predictability of ~1% ($R^2 = 0.245$). Based on these results, we test the inclusion of ZW3 as a tertiary MOV in

the prediction of ACP values. However, the multilinear model based on the convergence of the three MOVs (which span a shorter study period from 1979-2015 due to data availability) does not demonstrate improvements. On the contrary, the SAM+SOI+ZW3 model produces an R^2 value that is lower than its two counterparts ($R^2 = 0.229$), therefore indicating a decline in ACP predictability when ZW3 is included as a model feature.

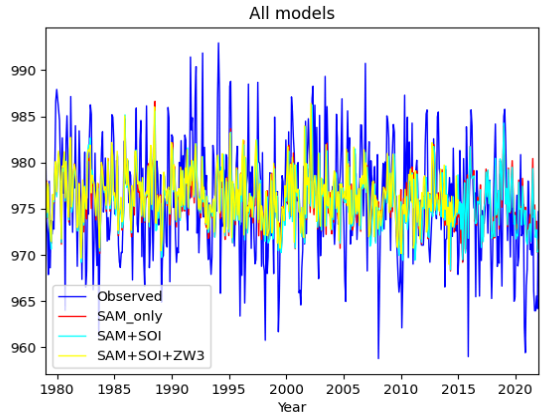

Figure 6: A plot that overlays observational and predicted ACP values based on three distinct regression models. The magnitude of ASL extremes is not faithfully captured as shown by predictions that underestimate the true value of ACP.

We conclude that the strength of the ASL cannot be fully explained by a linear or multilinear relationship to the SAM or other SH MOVs, respectively. This is likely because atmospheric MOVs are non-linear features that result from drivers with large internal and external variability. Although our results suggest that a small portion of the variance in the observational dataset can be explained by one or more MOVs, and though some of our regression models imitate the relative timing of some extreme ASL events, large RMSE values (> 5) across all of our models imply that ACP predictions differ considerably from observations (Fig.6). This is particularly problematic when considering extreme ASL events, which have the largest potential to alter the climate of West Antarctica and whose magnitude is underestimated by our regression models.

ASL intensity may be a function of additional MOVs that produce more localized changes in the circulation of West Antarctica. We thus recommend that future studies of the interaction between the ASL and SH MOVs consider a more comprehensive methodology that includes additional MOVs like the Pacific Decadal Oscillation (PDO) and Indian Ocean Dipole (IOD). An unsupervised machine learning technique like principal component analysis (PCA) could then be used to distill the relative contribution of each to variability in West Antarctica.

References

Arblaster, Julie M., and Gerald A. Meehl. 2016. Contributions of external forcings to Southern Annular Mode trends. *Journal of Climate* 19 (12): https://doi.org/10.1175/JCLI3774.1.

Baines, P. G. and K. Fraedrich. 1989. Topographic effects on the mean tropospheric flow patterns around Antarctica. *Journal of the Atmospheric Sciences* 46 (22): https://journals.ametsoc.org/view/journals/atsc/46/22/1520-0469_1989_046_3401_teotmt_2_0_co_2.xml.

Fogt, R. L., D. H. Bromwich, and K. M. Hines. 2011. Understanding the SAM Influence on the south Pacific ENSO teleconnection. *Climate Dynamics* 36 (7): 1555–76. https://doi.org/10.1007/s00382-010-0905-0.

Fogt, R. L. and G. J. Marshall. 2020. The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere. *WIREs Climate Change* 11 (4): e652. https://doi.org/10.1002/wcc.652.

Raphael, M. N. 2004. A zonal wave 3 index for the Southern Hemisphere. *Geophysical Research Letters* 31 (23). https://doi.org/10.1029/2004GL020365.

Raphael, M. N., G. J. Marshall, J. Turner, R. L. Fogt, D. Schneider, D. A. Dixon, J. S. Hosking, J. M. Jones, and W. R. Hobbs. 2016. The Amundsen Sea low: Variability, change, and impact on Antarctic climate. *Bulletin of the American Meteorological Society* 97 (1): 111–21. https://doi.org/10.1175/BAMS-D-14-00018.1.

Thompson, D. W. J. and J. M. Wallace. 2000. Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate* 13 (5): 1000–1016. https://doi.org/10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2.

van Loon, H. and R. L. Jenne. 1972. The zonal harmonic standing waves in the Southern Hemisphere. *Journal of Geophysical Research (1896-1977)* 77 (6): 992–1003. https://doi.org/10.1029/JC077i006p00992.

Wang, J., H. Luo, L. Yu, X. Li, P. R. Holland, and Q. Yang. 2023. The impacts of combined SAM and ENSO on seasonal Antarctic sea ice changes. *Journal of Climate* 36 (11): 3553–69. https://doi.org/10.1175/JCLI-D-22-0679.1.