Chapter 1

Data and Data Processing

This will be a summary of the data we use in this project and some data which we don't (but is used by other researchers) How is it collected? What format is it in? How reliable is it? This chapter will also contain a brief overview of some of the preprocessing methods carried out on the variety of datasets used for this project.

1.1 Antarctic Ice

Sea Ice

Land Ice

Standardising Ice values

We can try to standardise the different ice datasets so we can use them as one variable without comparing them separately.

The NSIDC data comes in a concentration % with an associated area data file. We can therefore calculate the total area of ice which changes from year to year, or leave the data as a %.

The GRACE data

1.2   Temperature and other environmental variables

1.3   Climate Indices

SAM

ENSO

DMI

IPO

Non-Pacific Climate Modes

1.4   Pre-processing

Temporal Averaging

Spatial Regridding

Because we use a variety of datasets which come in a variety of structures, it is important that standardise the spatial dimensions of each data source. One way we do this is by interpolating each dataset to have a consistent spatial arrangement. This allows for better quality results and makes it easier to calculate measures such as the correlation between 2m temperature and sea ice concentration.

We do the interpolation using the python package Scipy, which makes use of a piecewise cubic, continuously differentiable (C1), and approximately curvature-minimising polynomial surface to determine the value of our given variable at a chosen location.

We converted the temperature data to the projection the sea ice data is provided in; a south polar stereographic projection with regular grid cells of 25km×25km. We found this resolution to have a good balance between rea-

sonable run-times and good quality results.

Temporal Anomalies

Erroneous Values