Master of Science in Environmental and Geographical Science

University of Cape Town

The Skill Assessment of Seasonal Wind Prediction in South Africa

Trisha Parbhoo
PRBTRI003

Supervisor: Dr Chris Lennard

Acknowledgements

I would like to thank my supervisor, Dr Chris Lennard, for his support, patience and enthusiasm over the past two years. Thank you for sharing your insight and world of knowledge with me – this study would not have been possible without you.

Thank you to the National Research Fund (NRF) for financially supporting me for two years, and to ERM for the financial support in 2020.

I would like to thank my fellow researchers at the University of Cape Town for their constant advice and guidance. Thank you to Phillip Mukwenha, for providing IT support which enabled me to work virtually from the safety of my home.

Lastly, I would like to thank my family – for all you have done, for all you continue to do and for believing in me. I would not be where I am without you.

Abstract

To assess the skill of seasonal wind prediction in South Africa, the period between 1983 and 2018 was studied, unless otherwise stated. A total of sixteen (16) sites were studied across South Africa to ensure that each region across the country was represented in the study. Seasonally, prediction of the wind energy resource is important in order to plan for local plant operations, such as downtime for maintenance during low wind periods. Additionally, if a forecast indicates a large wind energy resource for a particular season, maintenance at other non-wind energy generation plants may be planned. However, currently there is limited information about seasonal prediction of wind in South Africa and no information about the skill of such a forecast. This thesis begins the process of addressing this knowledge gap.

Correlations between the CFS hindcast data and the reanalyses show the windspeed forecast is skilful over the southern regions of South Africa during December – January – February (austral summer), and over northern regions during June – July – August (austral winter), where the Pearson Correlation Coefficient ranges between ~0.5 and ~0.75. This skill is a function of the regional atmospheric stability during the respective seasons which are in turn a function of large-scale circulation features that govern synoptic processes driving the regional wind predictability.

La Niña results in a weakening of the high pressure systems to the west of the country, which could cause lower wind speeds specifically over the north-western regions of South Africa. The deepening of an extended sub-tropical low pressure trough over the interior of South Africa during La Niña potentially causes an increase in the pressure gradient, resulting in higher wind speeds over the eastern, north-eastern and eastern interior regions. During La Niña events literature indicates that the south-easterly wind speeds over regions of the Western Cape increase, and this could be due to the southward movement of the South Atlantic High Pressure.

During a negative SAM phase, the correlation between the SAM Index and the reanalyses indicates that there is an enhanced predictability in the eastern interior regions of the country, specifically during September – October – November (where the *r*-value is < -0.4 over specific regions), which is likely a function of the strong belt of westerly winds that move equatorward and therefore closer to South Africa. A low pressure trough develops and extends across the central interior of country which results in an increased pressure gradient between the interior and the high pressure cells off the coasts of South Africa, resulting in increased wind speeds over the majority of the country. These results indicates that ENSO and SAM are drivers of wind predictability over South Africa at the seasonal scale.

The CFS hindcast captures the ENSO forcing (El Niño, neutral and La Niña) of wind speeds for the majority of the same regions and seasons as found in the reanalysis data. However, the

CFS does not capture the SAM forcing of seasonal wind speeds described above. Therefore the CFS forecast system is a useful system with respect to seasonal wind forecasts given the ENSO forcing is captured, however, this study recommends that model development research should focus on developing forecast systems that capture other large-scale drivers, apart from ENSO, such as the SAM.

The study also demonstrates that it is essential to use multiple reanalyses if assessing the skill of the seasonal forecast using reanalysis products. Results show that there are statistically significant differences between the reanalysis wind datasets. The ERA5 wind speeds were lower in the majority of the sites (14 out of the 16 sites) in comparison to the ERA-Interim and CFSR wind speeds. This indicated that when only one reanalysis was used, it could artificially show higher or lower skill levels.

Based on the results above, this study concludes that there is better predictability in the CFS when the atmosphere is stable, and in certain parts of the country under La Niña and negative SAM during specific times of the year; for example, ENSO has greater effect over South Africa during austral summer months. During these skilful periods, the seasonal prediction system can be used to inform seasonal wind energy potential across the country. Understanding when and where a wind forecast is skilful could assist in efficient energy supply planning, inform plant operations such as maintenance and contribute towards the shift to a low-carbon economy.

Table of Contents

Table of Figures

Table of Equations

Table of Tables

CHAPTER 1

1. Introduction

1.1. Background

1.1.1. Seasonal Wind Prediction

Seasonal wind prediction is a necessary, yet complex process which is relatively new in South Africa. It was only in 1994 that seasonal forecasting was recognised as being an important part of studying the climate, and since then, emphasis has mainly been placed on the seasonal forecasting of temperature and precipitation (Mason, 1998). The use of wind energy is also new in South Africa. Based on the 2019 South African Energy Report, coal was responsible for 69% of the country's "total primary energy supply" in 2016 and renewable energy only accounted for 11%. South Africa has continued to make use of coal due to the easy accessibility and low cost to mine, despite the environmental and social consequences (Department of Energy, 2019: 8). The first major wind farm only became operational in 2014, which is one reason why seasonal wind prediction has not been a notable theme in the country. Currently, there are 33 wind farms in South Africa, 11 of which are still undergoing construction (South African Wind Energy Association, 2019). The use of seasonal wind forecasts combined with historical data, will allow for the optimal running of the wind farms and it will make the maintenance of the wind farms more efficient.

1.1.2. Skill Assessments

The skill assessment of seasonal forecasts in South Africa has mainly been limited to temperature and precipitation, based on the fact that those are the two variables that are being focused on the most. Around the globe, however, the seasonal prediction of wind and the skill assessment of these predictions are taking place. Skill assessments of forecasts began in the United States of America (USA) and Europe and it was driven by those who were interested in knowing the quality of the forecast in order to make use of it. In 1980, there was an increase in verifying forecasts and more attention was paid to various verification processes (Murphy, 1996). These processes were initially for short-term forecasts, however it led to the development of skill assessments for long-range predictions. In 1997 and 1998, an El Niño event resulted in areas around the USA experiencing either extreme droughts or flooding and the same effects were seen years earlier due to the El Niño Southern Oscillation phenomenon. Once the public realised the effects the ENSO phenomenon was causing, more interest was being placed on forecasting, as the value of seasonal prediction became apparent (Hartmann et al., 2002).

1.1.3. Setting the Synoptic Context

South Africa's climate is diverse. It is dominated by various mesoscale systems and there is high variability with regards to the climate across the country, resulting in multiple weather patterns (Blamey and Reason, 2012). The country lies in the domain from 22°S to 35°S and from 17°E to 33°E, and sub-tropical and temperate climate regimes are experienced. The seasonal climate of South Africa is influenced by large-scale global and hemispheric drivers that influence seasonal synoptic-scale drivers. Different regions of the country are influenced by different large-scale and synoptic drivers which raises the question of whether there are regions in South Africa where wind speeds can be skilfully forecast at the seasonal time scale as a function of a particular set of large-scale drivers.

During certain times of the year, there are different prevailing systems which influence the wind characteristics in South Africa. One of the major drivers of climate are the pressure systems, which change in position throughout the year; this is similar to the Westerlies, which also shift and therefore impact the wind patterns in South Africa on a seasonal scale (Section 5.4) (Hahmann and Kruger, 2010). The high pressure zone, which includes the South-Atlantic Anticyclone, the South Indian Anticyclone and the Kalahari High Pressure (Wright et al., 2019), shift over the country depending on the season. During austral summer (winter), the Inter Tropical Convergence Zone (ITCZ) shifts south (north), which causes the high pressure zone to also move south (north). The shifting of the pressure cells impacts the wind speed and direction over South Africa. The wind patterns between the coastal regions and inland regions are different because there are different drivers of wind depending on where the region is located.

The ocean currents also impact the climate over different regions of the country; on the west of the country is the South Atlantic Ocean and the cool Benguela current which results in dry conditions over the western provinces of South Africa. During the austral summer months, the wind speeds over the western provinces increases, due to an increased thermal and pressure gradient. The Indian Ocean and warm Agulhas current impacts the climate of the east of the country. The Agulhas current is responsible for transporting the warm Indian Ocean waters to the South African eastern coastlines (Hutchinson, 2017). During the summer months, the warm currents cause a warm onshore wind which results in rainfall over the eastern and interior regions of the country (Lennard, 2019). Ocean currents regulate and influence the surrounding climate, as they transport moisture and heat around the globe (National Oceanic and Atmospheric Administration, 2018) (South African Government, 2019). The ocean and sea surface temperatures have a long-term impact on the atmosphere. The increasing global temperatures may have an effect on the temperature of surrounding oceans, which in turn will impact the surrounding currents. These changes could result in the pressure systems to be altered and subsequently, wind and rainfall patterns over South Africa may be impacted. For example, the increase in temperatures of the Agulhas current is likely

to affect the rainfall over the eastern provinces in South Africa and seasonal changes may impact the South Atlantic Anticyclone which would not only adjust wind patterns, but also sea surface temperatures and upwelling systems over the Atlantic ocean (Lutjeharms et al., 2001) (Reason et al., 2006). This indicates that modifications in the climate system that occur offshore and internationally do not occur in isolation and that South Africa's weather and climate may experience changes in the future due to these seemingly unrelated modifications.

During austral summer, the South Atlantic High Pressure system moves south by approximately five degrees, and therefore ridges over the southern regions of the country (the Western Cape) (Nchaba et al., 2016), whereas the mesoscale weather patterns that occur over the interior regions of South Africa during austral summer are largely due to the extension of troughs from low pressure systems, which originated from thermal lows. During austral winter, the coastal regions located in the southern parts of the country are impacted by low pressure systems. The westerly wind belt moves north, as the ITCZ moves north, and mid-latitude cyclones pass over the southern parts of South Africa (Wright et al., 2019). The movement of the ITCZ results in the Kalahari High Pressure being situated over the interior of South Africa and the rotation of the Kalahari high pressure is responsible for the direction of wind over the study areas which are located in the interior regions of South Africa (Lennard, 2019). Due to the varying mesoscale systems across the country, the climate of the coastal regions differ to that of the inland regions. This study presents how the wind patterns differ between coastal and inland regions (Section 7.3.).

1.2. Rationale

Seasonal wind prediction is becoming increasingly important, especially since the move towards renewable energy is an essential part of many countries' commitment to the Paris Accord. In order to utilise wind to create wind energy, energy developers require seasonal forecasts to plan plant operations such as downtime for maintenance during low wind periods. A seasonal wind forecast that has a low skill, however, is essentially unhelpful and a forecast can only be used if it is proven skilful enough to use. Wrong forecasts can also result in large expenses. This is one of the reasons why this project has been chosen to study; the only way that seasonal wind forecasts can be useful in South Africa is if their skill is quantified in space and time.

1.2.1. The Global Shift Towards Renewable Energy

There has been a shift towards renewable energy globally. "Energy is essential to our society to ensure our quality of life" (Bull, 2001: 1216), and the development of the economy heavily depends on electricity supply. It is therefore vital that the source of energy is reliable, cost-effective and will not harm the environment, humans or animals. Countries, organizations

and individuals are becoming more aware of the harmful impacts that the burning of fossil fuels has on the environment as well as on human health. This has resulted in an increased use of renewable energy. The benefits of renewable energy have become increasingly apparent; the resources are abundantly available and there is the potential for certain renewable energy sources to be more cost-effective than conventional (fossil fuel-based) energy sources (Bull, 2001; Fulton and Reid, 2018), and the negative consequences on the environment are lower than conventional energy systems. A specific case study in the United States found that the stocks for clean energy outperformed the conventional energy stocks during the Covid-19 pandemic – this indicates that the benefits of clean energy goes beyond being environmentally friendly, and that there is the potential for clean energy to provide financial benefits for investors and shareholders (Hook, 2020). Globally, wind power is becoming increasingly popular (Blaabjerg and Ma, 2017; Willis et al., 2018). The technology associated with wind energy production has improved, which has resulted in the performance of wind turbines to improve and therefore increasing the yield of wind energy (Blaabjerg and Ma, 2017).

1.2.2. The South African Context

In 2002, South Africa signed the Kyoto Protocol (UNFCCC, n.d) and one of the aims of the Kyoto Protocol was to decrease air pollution globally. The Paris Agreement on climate change, which came into force and signed by South Africa in 2016, was seen as an upgrade of the Kyoto Protocol (Department of Environmental Affairs, 2019). There was an international aim to keep the rise of global temperatures below two degrees Celsius from pre-industrial global temperatures and in order to do so, each country that signed the Paris Accord has been encouraged to decrease emissions (UNFCCC, n.d). South Africa has to comply with international commitments, and one way to do so is by moving towards renewable energy. South Africa is within the top twenty countries of highest greenhouse gas emitters, and this is largely due to the burning of coal for energy purposes (McSweeny and Timperley, 2018). The majority of the energy production in the country is coal-based, and coal-fired power plants not only cause acid rain, release harmful fossil fuels into the atmosphere and contribute to global warming, but they also cause health problems for humans and living organisms who come in contact with the fumes (Burt et al., 2013). Additionally, strong winds also impact the built environment, and as wind patterns are better studied and forecasted, the designs for the future can be adapted such that the impacts will be minimised (Kruger et al., 2010).

The combination between South Africa's international commitments and the consequences of fossil fuel-based energy production provides a strong rationale for this study. Wind resources are plentiful, renewable and cheaper than other energy sources and according to South Africa's Integrated Resource Plan developed (Department of Mineral Resources and Energy, 2019), South Africa plans to install 17742 MW of wind energy by 2030, which will be

over 22% of the total installed capacity at the time. This large penetration of wind energy into the power pool means that accurate wind forecasts will be needed to estimate the contribution of wind energy at any particular time to ensure grid stability. This research project aimed at not only assessing the skill of forecasts for regions across the country, but to also apply a methodology that can be used for forecasts in other areas of South Africa, and throughout Africa, thus creating a pathway for the use of renewable energy. This study also aimed to highlight how to improve seasonal wind forecasting in the future by showing the strengths and weaknesses of the forecast, as well as studying the drivers of wind which could potentially enhance seasonal wind predictions.

2. Problem Statement

There has been a lack of attention given to seasonal wind forecasts in South Africa and therefore the skill of the existing forecasts has not been evaluated. The skill assessment of seasonal temperature and precipitation forecasts have been studied, however not wind, and based on the fact that these variables behave differently, the same skill assessments cannot be applied. Wind statistics are rarely considered in defining the South African climate, and only recently have areas of strong wind been studied in detail (Kruger et al., 2010). Wind data in South Africa, in comparison to rainfall and temperature data, is minimal and often there is missing data, which makes it challenging to conduct wind studies in the country.

3. Aim

The aim of this study is to assess the skill of seasonal wind forecasts in different regions across South Africa.

4. Objectives

1. To compare hindcast data to reanalysis data in order to assess the skill of the seasonal forecast, by evaluating tools and metrics being used to assess wind forecast skill.
2. To evaluate three reanalysis datasets for use as a proxy for observed wind data.
3. To understand cross-scale drivers of wind patterns in South Africa including teleconnections, synoptic and regional drivers.

CHAPTER 2

5. Literature Review

A literature review is useful because it allows for current knowledge about a particular topic to be studied and it therefore enhances the level of understanding about the topic at hand. Studying prior knowledge contributions also allows the researcher to find gaps in the existing

knowledge, which can be filled and therefore decreases discontinuities in information. This literature review focuses on seasonal forecasting in South Africa, and continues by examining seasonal wind forecasting around the globe. Types of skill assessment will be studied in order to determine the best methodology for this project. Finally, the large-scale drivers which impact the South African climate will be discussed.

5.1. Seasonal Forecasting in South Africa

The first section of the literature review presents literature on seasonal forecasting in South Africa. Seasonal forecasting, along with gaining a better understanding of the climate, has significant benefits for South Africa. Often, natural hazards cannot be avoided, however the way in which people adapt to and manage a natural hazard can be improved if an accurate forecast is provided (Klopper, 1999). The importance of seasonal forecasts was recognised and in 1994, "the South African Long-lead Forecast Forum" (SALFF) was created by the South African Weather Bureau. The SALFF made use of both upper atmospheric data and surface data to model seasonal forecasts over the South African region. Large scale drivers such as El Niño events, La Niña events and the Quasi-Biennial Oscillation were also studied (Mason, 1998). The El Niño Southern Oscillation (ENSO) has a large impact globally, and therefore there has been considerable focus that has been placed on it (Goddard et al., 2001). In South Africa, the correlation between ENSO and precipitation has been studied, which has allowed for seasonal precipitation forecasts to occur for the country. There are however, some inconsistencies with the correlation between ENSO and rainfall in South Africa (Johnston et al., 2004).

Seasonal forecasting can be complex, especially for regions which have a low predictability through teleconnections and experience high variability (Johnston et al., 2004). Another reason why forecasting can be complex is because the atmosphere is highly dynamic and therefore even slight changes in initial conditions can impact the result of a forecast. In order to increase the accuracy of forecasts, ensemble forecasting is used (Leutbecher and Palmer, 2008). Ensemble forecasting is a numerical weather prediction system that runs multiple forecasts where the initial conditions vary slightly each time, this results in each ensemble member differing to one another depending on the initial conditions. The ensemble members should all have the same likeliness of occurring in the future. One of the reasons why this method of forecasting is useful is because it allows for uncertainties within a forecast to be assessed. An ensemble prediction system (EPS) acknowledges that the weather changes slightly on an hourly and daily basis and that this could impact the output of a seasonal forecast (World Meteorological Organization, 2012).

Seasonal forecasts are created by either using statistical models or dynamical models. Both types of models require some form of input data (Johnston et al., 2004). The South African

Weather Service (SAWS) makes use of the ensemble prediction system in order to produce accurate seasonal forecasts for South Africa and they created an ocean-atmosphere coupled model called the "SAWS Coupled Model" (SCM). The SCM however, did not produce reliable local scale forecasts and therefore the South African Weather Service implemented the "Multi-Model System", which aims at statistically downscaling global forecasts (Landman and Beraki, 2015) (South African Weather Services, 2019). SAWS, along with most scientific and/or academic institutions in South Africa, mainly focus on forecasting precipitation and temperature on a three-monthly time period (Johnston et al., 2004). In order for the seasonal forecasts to be distributed and understood by the South African public, The Southern African Regional Climate Outlook Forum (SARCOF) and the World Meteorological Organisation (WMO) create outlook documents which combine forecasts which are done by university and research institutions. This allows for the user-end to apply the forecast data (Johnston et al., 2004).

There are various types of models which produce seasonal forecasts and many of them are ocean-atmosphere coupled models. Ocean-atmosphere coupled models aim to produce a forecast which is more accurate. In order to increase the skill of forecasts, new models are either being initiated or current models are being adjusted. Models such as the 'Met Office Seasonal Forecasting System (GloSea4)' have been upgraded to produce quality seasonal forecasts (Arribas et al., 2010: 1891). The Intergovernmental Panel on Climate Change (IPCC) wrote a report, and in Chapter 8, they evaluate various climate models. In the report, they state the 'Atmosphere-Ocean General Circulation Models (AOGCMs)' often forecast certain variables better than others, for example, temperature is forecasted more accurately compared to other variables (Randall et al., 2007: 591). The models and systems that are chosen depend on the variable and the site being studied. A research paper written by Bett et al., (2018) aimed at looking at seasonal forecasting data and made use of the European Centre for Medium-Range Weather Forecasts (ECMWF) model data, the Météo-France System data and the GloSea5 data. The study found that each of the models had different levels of skill depending on the region and the variable being studied. This research paper will make use of data from the Climate Forecast System (CFS) (reanalysis and hindcast), the ERA-Interim model and the ERA5 model.

5.2. Seasonal Wind Forecasting

Seasonal wind prediction has been limited in South Africa, therefore this section will assess the literature of seasonal wind forecasts around the globe. Wind prediction is difficult because it is a volatile variable with a chaotic behaviour. Wind is site dependent and the topography, local weather profile and climate can impact seasonal wind predictions (Wu and Hong, 2007). It has been recognised that seasonal predictions are more established in tropical regions rather than temperate and mid-latitudes, which adds to the difficulty of seasonal wind forecasting in South Africa (Palutikof et al., 2002)

In China, wind energy is a fast growing sector due to the fact that the government made a commitment that by 2020, 15% of all electricity will be generated using renewable resources; therefore, seasonal wind forecasts have become vital in order to achieve this goal. China has also been interested in seasonal wind forecasts in order to predict strong winds which could cause damage to infrastructure. There is strong seasonal wind variability in China, where the strongest wind speeds are experienced during winter due to an increase of the pressure gradient which is formed between the Aleutian Low and Siberian-Mongolia High. Changes in wind can also be seen due to the El Niño Southern Oscillation in the south of China as well as the Arctic Oscillation in the north of China (Lockwood et al., 2019) (Thornton et al., 2017). The wind power density (WPD) and the relation between WPD and ENSO has also been studied by Yu et al., (2020), in order to better understand and improve the yield of wind energy.

The use of renewable energy has increased in Europe, and the European Union's future of policy includes decarbonizing the energy sector (Bórawski et al., 2020). The European Commission funded a project in 2012, called the European Provision Of Regional Impacts Assessments on Seasonal and Decadal Timescales (EUPORIAS). EUPORIAS aims at bettering current technology and climate projects such that society will benefit; to increase the resilience of Europe with regards to climate change; and to make climate information more easily available to various parties that may benefit from the information. One of the projects that forms part of EUPORIAS is Project Ukko[1] which focuses on seasonal wind prediction in order to help the European energy sector. Project Ukko has an interactive mode where the user can analyse current and future wind activity. This project also makes use of ensemble predictions, where there are 51 ensemble members which run at the beginning of every month (Project Ukko, n.d).

5.3. Assessing the Skill of Seasonal Wind Forecasts

Seasonal forecasting is becoming a vital part of studying the climate and it is needed to plan for the future. The only way in which seasonal forecasts can be used, is if they have a high level of skill, otherwise the forecast is not helpful (Wu and Hong, 2007). This section of the thesis will discuss the skill analyses of seasonal forecasts that have been used by various researchers. The skill of a specific forecast depends on the variable being studied, the model, the time-frame of the forecast and the study area (Bett et al., 2018). As mentioned before, South Africa's focus has been on the seasonal forecasting of temperature and precipitation; therefore, the skill assessments have been relatively specific to those two variables.

[1] The link to access Project Ukko is: http://project-ukko.net/

Assessing the skill of a forecast can be done by using hindcasts. Hindcast data is when a current forecast model is run at any point in the past, as though it is a forecast. Observed or reanalysis data, that has been collected for the period of time can be compared to the hindcast data, allowing for the quality of the hindcast, and the model data to be established. This method however, is dependent on the quality of the hindcast data and the observed data (Bett et al., 2018).

5.3.1. Correlations

Lockwood et al., (2019) assessed the skill of seasonal wind forecasts in China using two ensemble datasets. Both the ensemble datasets were hindcast datasets, which were created by the same model. The hindcast data was then assessed against reanalysis data, which was used in place of observation data. In order to assess the skill of the forecast, the 'Pearson correlation coefficient (*r*)' was calculated (2019: 3), which is a type of correlation analysis between two variables. If the correlation coefficient is closer to zero, there are weak correlations; whereas if *r* is close or equal to one or negative one, there is a high correlation between the two variables (Benesty et al., 2009). Negative (positive) *r*-values indicate a negative (positive) correlation. Project Ukko, which has been mentioned above, also compares the seasonal wind speed predictions with reanalysis data as a form of testing the quality of the forecast. Project Ukko makes use of the "ranked probability skill score (RPSS)", where a value of one means that the forecast is "perfect" and a value below zero means that there is little skill associated with the forecast (Project Ukko, n.d).

5.3.2. Error

The Root Mean Square Error (RMSE) is a measure-oriented method of analysing the quality of a forecast. This method calculates the level of error between two datasets. Smaller RMSE values indicate that the observed values and the forecasted values are similar and if the two datasets correlate perfectly, the Root Mean Square Error will be equal to zero. The RMSE is a more advanced version of the Mean Error (ME) formula, as the ME does not account for outliers and can therefore be misleading. A negative Mean Error value indicates that a model has under-forecasted, whereas a positive value indicates that a model has over-forecasted. The RMSE accounts for outliers and due to the fact that the values are being squared, the issue of negative values are accounted for. The RMSE is also more effective for large datasets (Nchaba, 2012). Nchaba (2012) provides a clear and detailed description of the other measure-oriented methods. The RMSE formula includes a P value which is the predicted values, O which is the observed values, and n which is the number of timesteps in the dataset (Equation 1).

$$RMSE = \sqrt{\frac{\int_{i=1}^{n}(P_i - O_i)^2}{n}}$$

Equation 1: Root mean square error

5.3.3. Receiver Operating Characteristics and Brier Scores

Landman and Beraki studied the evolution of seasonal forecasting in South Africa and made use of Receiver Operating Characteristics (ROC) scores (2015). ROC Curves are presented as a figure or graph, and are "commonly used to study the capacity of continuous variables (markers) to properly classify subjects into one of two groups" (Pérez-Fernández et al., 2021: 135). Based on the hit rates and false alarm rates, a ROC Curve is plotted (Figure 1). The ROC score system is most effective to measure the skill of variables such as rainfall, because there is a definite outcome of whether there was precipitation in an area or not; whereas with wind, a threshold wind speed needs to be established first, before applying a ROC curve. Based on the fact that wind is not a variable that has a binary outcome, applying ROC scores might not be as effective.

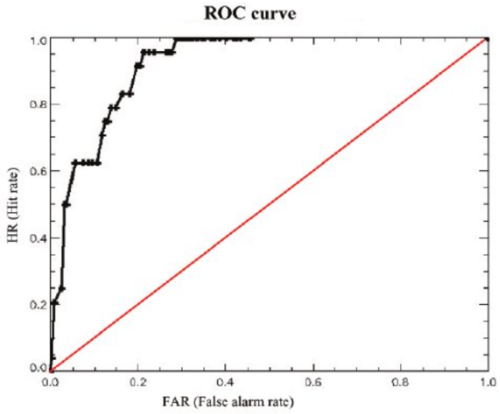


Figure 1: *Example of a ROC Curve for a precipitation forecast (Barrera-Escoda et al., 2007)*

Another skill assessment that has been used in order to test the skill of a forecast is a the Brier Score, which is a distribution-oriented method. The Brier Score was developed in the 1950s and it measures the error in probability forecasts (Mason, 2004; Stefanova and Krishnamurti, 2002). This skill assessment has been used for ensemble forecasts, which have been generated either from a single model or a multi-model forecast. The Brier Score assesses the forecast probability against the climatology probability (Stefanova and Krishnamurti, 2002). The value of the Brier Score will indicate the accuracy of the forecast. Zero means that the forecast is perfect whereas if the Brier Score is equal to one, the forecast has no skill.

A forecast that has the highest skill is graded as a zero, according to the Brier Score (Figure 2). If there are points that lie above the orange diagonal line, the forecast is underconfident, whereas if there are points that lie below the orange diagonal line, the forecast is overconfident (Van Gelder, 2015). The Brier Score is mainly used for binary outcomes, for

example precipitation, where it either rained or it did not. Similar to the ROC score, this skill measure is difficult to apply to evaluate wind forecast skill because it is primarily a binary skill measure.

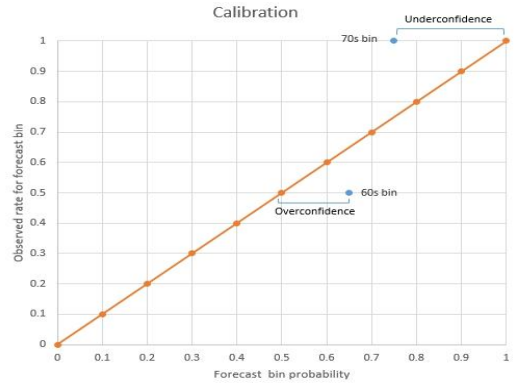


Figure 2: *Example of a Brier Score graph (Van Gelder, 2015)*

5.3.4. Reliability diagrams

Reliability diagrams are used to determine how accurate a "probabilistic forecast" is (Bröcker and Smith, 2007). This is represented on a graph where the forecast probability lies on the x-axis and the observed frequency lies on the y-axis (Figure 3).

A forecast with high reliability, means that the observed frequency and the forecast probability have a high correlation. The Brier Score, mentioned above, can be related to the reliability curve (Figure 3). Reliability diagrams were not used in this study based on the nature of the skill assessment, which suits binary outcomes best.

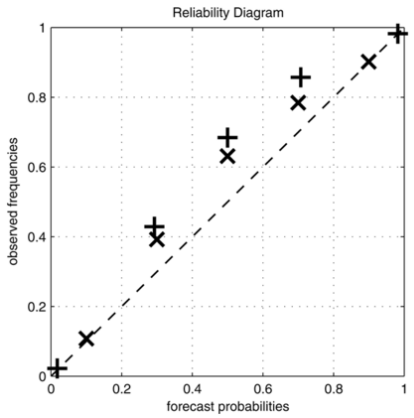


Figure 3: *Example of a Reliability Diagram and the depiction of a forecast that has perfect reliability (Bröcker and Smith, 2007)*

The skill assessments that are mentioned above will help researchers and scientists to improve seasonal forecasting in the future; however, not all skill assessments can apply to all climate variables, especially wind, which behaves differently to temperature or precipitation. It should also be noted that the time of year and the area of study can impact the skill of a forecast because of conditions that cannot be controlled. For example, past studies have

indicated that the skill of a forecast is generally lower during spring and autumn (Landman et al., 2013). Arribas et al., (2010) stated that, based on the research, areas that are impacted by the El Niño Southern Oscillation phenomenon, have higher forecast skill than areas that are not. The skill of forecasts can improve over time as methodologies and technology develop, however it is often a slow and difficult process.

5.4. Synoptic Scale Drivers of Wind in South Africa

5.4.1. Pressure Systems

The pressure systems that influence South Africa are the South Atlantic High Pressure, the Indian Ocean High and a Tropical Low Pressure. The position of the Intertropical Convergence Zone (ITCZ), the pressure systems mentioned above and the westerly wind belt, move depending on the season. The Westerlies impact the south of South Africa, due to the position of the wind belt.

During austral summer there is a ridging high pressure called the South Atlantic High Pressure, on the west of the country (Figure 4). This high pressure causes subsiding air over the Western Cape, and therefore dry conditions in summer. It is also the cause of the 'Cape Doctor', which is a strong south-easterly wind that is known for blowing pollutants away from the Cape Peninsula. The strongest winds are experienced on the outer edges of the anticyclone (Kruger et al., 2010) (Lennard, 2019).

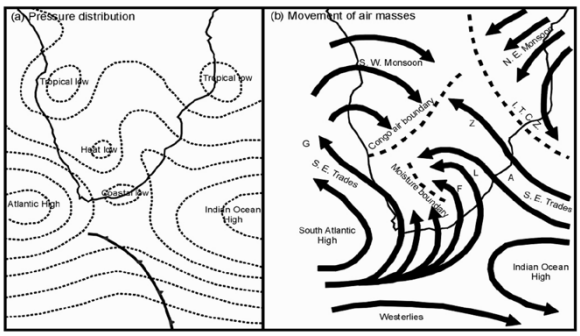


Figure 4: Pressure Systems over South Africa in Summer (Kruger et al., 2010)

During austral winter (Figure 5), the pressure systems move north in relation to the positioning during austral summer (Figure 4), due to the northward movement of the ITCZ. This results in cold fronts passing over the Western Cape and the climate of the interior of the country is controlled by a high pressure system. The South Easter also ceases during winter, due to mid-latitude cyclones which causes the high pressure system to shift positions (Kruger et al., 2010) (Lennard, 2019).

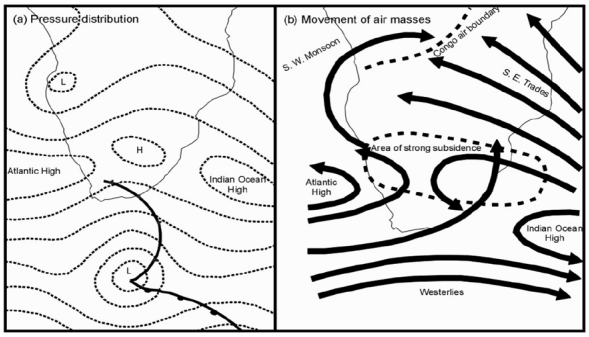


Figure 5: *Pressure Systems over South Africa in Winter (Kruger et al., 2010).*

5.5. Large Scale Drivers of Wind in South Africa

There are certain large-scale drivers of wind in South Africa that need to be understood in order to evaluate and understand a wind forecast. A large-scale driver can be characterised as phenomena that occur on a global or hemispheric scale and impacts smaller scale phenomena, on a synoptic and/or meso-scale. These large-scale drivers impact the amount of wind experienced within South Africa. Macroscale weather systems occur at different timescales and have different impacts on a regional scale (Kruger et al., 2010).

5.5.1. Multivariate ENSO Index

The El Niño Southern Oscillation (ENSO), is one of the most notable large-scale drivers of variability around the globe. Between approximately four to seven years, the trade winds change direction or stop, which results in a change of temperature in the Pacific Ocean (Lakhraj-Govender and Grab, 2018). Changes in equatorial Pacific sea surface temperatures impacts the global climate and specifically precipitation and atmospheric circulation in Southern Africa (Rouault et al., 2010). The pressure and wind speed relationship is also dependent on temperature (Wooten, 2011), therefore due to the relationships between these three variables, an El Niño or a La Niña event impacts the wind speed in a region. Philippon et al., (2012) stated that during an El Niño period, wind speeds decrease, and the opposite occurs during a La Niña over certain regions across South Africa and during specific times of the year (Rouault et al., 2010). The Southern Oscillation Index (SOI) is an indication of the phases of ENSO. The negative phase of the SOI indicates that it is an El Niño phase and the positive phase of SOI indicates a La Niña phenomenon. The predictability of the ENSO phenomenon, although not always accurate, can help to forecast seasonal wind speeds.

5.5.2. Southern Annular Mode

The Southern Annular Mode (SAM), also known as the Antarctic Annular Oscillation (AAO), mainly influences the mid-latitudes and high-latitudes in the Southern Hemisphere. This large-scale driver is caused by a constant change in atmospheric pressure between the south

polar region and the south mid-latitude region. The Westerlies move depending on the phase of the Southern Annular Mode, when SAM is in a positive (negative) phase, the Westerlies move further south (north). It has also been seen that the El Niño Southern Oscillation and the Madden–Julian Oscillation have a teleconnection to the Southern Annular Mode (National Center for Atmospheric Research, n.d).

There are multiple large-scale drivers that impact the wind patterns in South Africa, however based on the scope of this study, only the correlations between El Niño Southern Oscillation and the Southern Annular Mode and wind speed were analysed and discussed.

CHAPTER 3

6. Data and Methodology

This section presents the methodology used in this study to achieve the objectives stated above. The study applied a quantitative analysis, which can be defined as an analysis that involves the use of numerical data and often makes use of statistical analyses (Sheard, 2018).

6.1. Study Area

In order to gain an understanding of wind patterns across South Africa, the full domain as well as 16 sites have been chosen. The sites are Cape Town, Cape Point, Port Elizabeth, Mossel Bay, Port Nolloth, Saldanha, Mabibi, Pietermaritzburg, Johannesburg, Bloemfontein, Polokwane, Upington, Butterworth, Brandvlei, Hanover and Beaufort West (Figure 6) (Table 1). These regions represent coastal regions across the western, southern and eastern coasts along the country and therefore enables the analysis of how the different oceans and ocean currents impact wind patterns over land. The inland regions that were chosen also represent the different climate regimes of South Africa, and facilitates an evaluation of wind forecasts in the different climate and weather regimes.

Figure 6: Map of the study areas (Google Maps, edited by the author)

Table 1: A list of the sites and the corresponding coordinates used in this study (Google Earth)

Site	Coordinates
Cape Town	33.9715° S, 18.6021° E
Cape Point	34.3576° S, 18.4968° E
Port Elizabeth	33.9873° S, 25.6143° E
Mossel Bay	34.1993° S, 22.1165° E
Port Nolloth	29.2415° S, 16.9004° E
Saldanha	33.0277° S, 17.9176° E
Mabibi	27.3851° S, 32.7238° E
Pietermaritzburg	29.6432° S, 30.3965° E
Johannesburg	26.1394° S, 28.2468° E
Bloemfontein	29.0852° S, 26.1596° E
Polokwane	23.8597° S, 29.4533° E
Upington	28.4086° S, 21.2559° E
Butterworth	32.0906° S, 28.1359° E
Brandvlei	30.5162° S, 20.4363° E
Hanover	31.0476° S, 24.4284° E
Beaufort West	32.9667° S, 22.5566° E

6.2. Data Collection

6.2.1. Wind Data

There are four wind speed datasets that has been used for this study. There is one hindcast dataset (CFS) and three reanalysis datasets (ERA5, ERA-Interim and CFSR). The rationale

behind using hindcast data is that it can be seen as forecast data, except that the date of initialization is the past; therefore, the hindcast data has been used in place of the forecasted data. Reanalysis data has been used as a proxy for observed data due to the lack of high quality observed wind data in South Africa. The observed wind datasets often have missing data points, which will result in a weak analysis of the forecasted data. The wind datasets have been discussed in detail below. The four datasets have been re-gridded to cover the area of -35.4° to -20.3° latitude and 14° to 35° longitude.

6.2.1.1. Climate System Forecast (CFS) Hindcast Data

In 2004, the National Center for Environmental Protection (NCEP) "operationally implemented" the Climate Forecast System Version 1 and Version 2 was operationally implemented in 2010 (Saha et al., 2010: 1017) (Saha et al., 2014). The upgrade of the model included a longer time period of data available and the reforecast (hindcast) data improved in skill for specific climate variables (Saha et al., 2014). The Climate Forecast System is a coupled model that captures the interaction between the ocean and land. The CFS hindcast data that has been used in this study was initialized one month in advance. The data is composed of ten ensemble members, from which an ensemble mean has been calculated. Individual ensemble members are usually used when the aim is to overlook smaller scale, erratic features of a forecast and focus on the general pattern of the forecast instead (Atger, 1999). Using the ensemble mean therefore highlighted the larger scale patterns in the CFS wind speed data. There is also generally a smaller error associated with an ensemble mean compared to an individual ensemble member, and therefore often the ensemble member is used to represent the forecasts (Whitaker and Loughe, 1998). The Climate Forecast System hindcast data is available for public access and it is able to be downloaded, which added to the suitability of the dataset for this study.

The spatial resolution of the CFS Hindcast data used in this study is approximately 56 km, and it covers a period from 1982 to 2018. The data is six-hourly; however, it has been manipulated to study the forecast on a daily and seasonal basis.

6.2.1.2. ERA5 Reanalysis Dataset, ERA-Interim Reanalysis Dataset and CFSR

The European Centre for Medium-Range Weather Forecasts (ECMWF) is responsible for the ERA5 and ERA-Interim reanalysis datasets which both originate from the same source. The ECMWF stated that the ERA5 model is an upgrade from ERA-Interim and the upgrade included a higher spatial and temporal resolution. The ERA5 model has data available from 1950, compared to the ERA-Interim model which has data available from 1979. Hersbach et al., (2020: 1999) stated that the ERA5 model "[benefited] from a decade of developments in model physics, core dynamics and data assimilation" and therefore this would likely result in a higher quality forecast. The ECMWF also stated that ERA-Interim will eventually be 'phased out' (Hersbach et al., 2019) (Hersbach et al., 2020). There are still errors that can be seen in

both the ERA datasets compared to the observed data, however that is beyond the scope of this study. The ERA5 and ERA-Interim data used in this study ranges from 1982 to 2017. The third reanalysis model that was used in this study is the CFSR reanalysis data. The CFSR model was made operational by the National Centers for Environmental Prediction (NCEP), and originates from the same model as the CFS hindcast data: the Climate Forecast System.

The spatial resolution of the ERA5 reanalysis data used in this study is 30km, with a hourly temporal resolution. The spatial resolution of the ERA-Interim data is approximately 80km, with a six-hourly temporal resolution. The CFSR wind data's spatial resolution is approximately 38 km and is six-hourly. Unlike the ERA5 and ERA-Interim, the CFSR dataset ranges from 1982 to 2010. The data has been manipulated to study the wind patterns at multiple time scales.

The reanalysis data was used as a proxy for observed data as they provide full spatial coverage for an extended period of time and do not have missing data. Using reanalysis data from multiple models helped to demonstrate which models are most similar to the forecasted data. Three reanalysis datasets were used because it allows for a more robust study as uncertainties between reanalysis data can be quantified in addition to a quantification of skill in the seasonal forecast within this uncertainty.

6.2.2. Multivariate ENSO Index Data

The Multivariate ENSO Index (MEI) demonstrates El Niño phases, La Niña phases and neutral phases. If the MEI is high (low), it is indicative of an El Niño (La Niña) phase. A monthly MEI has been used in this study and it has been correlated to the wind speed data in order to understand whether ENSO impacts wind in South Africa. The MEI was used in this study because it is "considered the most representative since it links six different meteorological parameters measured over the tropical Pacific" (Mazzarella et al., 2010: 23).

6.2.3. Southern Annular Mode Index Data

The Southern Annular Mode (SAM) was studied by using the SAM index, which is available to download on the Climate Data Guide website. It is monthly data, ranging from 1982 to 2017. The SAM index presents the change in pressure and the movement of the atmosphere over 40 degrees south and 60 degrees south. The Southern annular mode impacts the circulation over surrounding regions, and it occurs inter annually. The value of the SAM Index indicates whether the SAM is in a positive phase or a negative phase. The Southern Annular Mode is also known as the Antarctic Oscillation (National Center for Atmospheric Research, n.d). Visbeck (2009), stated that the SAM is "arguably the most significant forcing of intraseasonal to decadal climate variability in the middle- to high-latitude Southern Hemisphere" and that recently, more studies have focused on this phenomenon (Visbeck, 2009: 941).

6.3. Analysis

In order to analyse the quality of the forecasted data and to satisfy the aim of the study, multiple types of analysis were applied to the reanalysis and hindcast wind speed datasets. The three overarching types of analysis were:

1. Comparing the three reanalysis wind speed datasets to one another (Objective 2).
2. Assessing the skill of the hindcast data in comparison to the reanalysis datasets (Objective 1).
3. Understanding how the El Niño Southern Oscillation and the Southern Annular Mode impact wind speed (Objective 3).

In order to analyse the datasets, the methodology described in Section 6.3. was performed.

The time scale of the data varies depending on the type of analysis. The majority of the results either make use of daily, monthly and seasonal wind data with the exception of Section 7.2.2., where four-hourly data was used. The daily, monthly and seasonal data was manipulated in CDO by calculating the daily mean for each dataset. CDO provides the function to calculate the monthly mean for each month of the year for the selected time period. The monthly means were consolidated into one netCDF file, and the correlations were performed on the data. Similarly, the data was grouped into the following seasonal groupings: December – January – February (DJF), March – April – May (MAM), June – July – August (JJA) and September – October – November (SON). This process included grouping all the data for each month according to the season; for example, all the data points from December, January and February were grouped to develop the 'DJF' dataset. In order to complete the daily, monthly and seasonal correlations, the whole timeseries data was used, and the wind speed data was grouped in the relevant timescales prior to the correlation analysis.

6.3.1. Ensemble Members and the Ensemble Mean

In order to determine whether the ensemble mean can be used as being representative of the individual ensemble members, a time series was created. The time series includes the ten ensemble members as well as the ensemble mean for each study area. In addition to the time series' in Section 7.1., an individual box plot was created using each ensemble member for each study area. The box plots for the individual ensemble members for each study site were evaluated by the author, however were not included in the study due to the fact that the results in Section 7.1. indicated the similarity between the individual ensemble members and the ensemble mean adequately. The box plots indicated that the spread of data for each ensemble member was similar, and that the ensemble mean represented the ensemble members, despite the fact that the ensemble mean resulted in a decrease in the spread of data and the number of outliers. The box plots of the combined ensemble members and the ensemble means can be seen in Section 7.3.

6.3.2. Spatial Plots: Seasonal and Four-Hourly

Spatial plots were created to present the change in wind speed over the four seasons: December – January – February (DJF); March – April – May (MAM); June – July – August (JJA); September – October – November (SON); as well as how wind speeds change throughout a twenty-four hour period. Wind speeds at 00:00, 04:00, 08:00, 12:00, 16:00 and 20:00 were plotted over a period of thirty-five years and over the entire study area. These figures were created for the ERA5, ERA-Interim and CFSR wind datasets. A four-hourly interval was chosen to represent the wind ramp that is experienced in many parts of the country (the Eastern Cape and Kwa-Zulu Natal). The wind ramp starts at approximately 04:00 to 05:00 in the morning and in order to capture the wind ramp, both 04:00 and 08:00 needed to be shown. By making use of a three-hourly or six-hourly time interval, the critical times would not have been captured. In addition, it is unlikely that choosing a time interval of three-hourly data would change the diurnal messaging, despite the fact that it is conventional to use a three-hour interval.

6.3.3. Site Analysis

Sixteen (16) study sites across South Africa were chosen for this study (Figure 6). Time series plots, box plots, and wind roses were created for each study area in order to understand and analyse the difference in wind patterns over the study areas, as well as examine the difference between the four datasets (CFS, ERA5, ERA-Interim and CFSR). The time series plots depict the changes in wind between each month over each study area and the box plots present the spread of each dataset. One ensemble member from the CFS hindcast data was used for the wind roses. The ensemble member was chosen for each site, depending on the member that is most similar to the ensemble mean. The wind roses present the change in wind direction and wind speeds over the year for each area. The three types of analysis for each study site resulted in a holistic understanding of the wind patterns across South Africa.

6.3.4. Statistical Significance Tests

The reanalysis datasets (ERA5, ERA-Interim and CFSR) present different wind speed results. In order to analyse the extent of the differences between the reanalyses, two statistical significance tests were computed: the T-Test and the Kolmogorov-Smirnov Test. The tests were performed on a monthly and seasonal timescale. The datasets were first separated into the relevant months and seasons, and the statistical tests were performed on the separate monthly and seasonal datasets. The tests were completed on Python, and the original data manipulation was completed in CDO. The output of the two statistical tests are *p*-values which indicates the level of significant difference between the two datasets. Probability Density Functions (PDFs) and spatial plots accompany the statistical significance tests, in order to visualize the difference between the reanalysis datasets. This section of the results aimed to answer Objective 2, which seeks to evaluate three reanalysis datasets for use as a proxy for observed wind data.

6.3.5. The Pearson Correlation Coefficient

The Pearson Correlation Coefficient was calculated by using the equation shown below:

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}}$$

Equation 2: The Pearson Correlation Coefficient formula

If the value of *r* is close to one, it means that there is a high positive correlation, whereas if the value of *r* is closer to negative one, there is a strong negative correlation. There is little to no correlation if the *r*-value is close to zero. In order to calculate the Pearson Correlation Coefficient, the numpy[2] correlation coefficient function has been used in python.

The Pearson Correlation Coefficient was used to compare the following datasets:

1. The three reanalysis datasets:
    - ERA-Interim and ERA5 daily, monthly and seasonal data
    - CFSR and ERA5 daily, monthly and seasonal data
    - CFSR and ERA-Interim daily, monthly and seasonal data
2. CFS hindcast data and reanalysis datasets:
    - CFS and ERA5 daily, monthly and seasonal data
    - CFS and ERA-Interim daily, monthly and seasonal data
    - CFS and CFSR daily, monthly and seasonal data

6.3.6. Root Mean Square Error (RMSE)

In order to calculate the RMSE, a package called sklearn metrics[3] was used, where there is a built in RMSE function, which makes use of the RMSE formula (Equation 1). The RMSE was calculated between the CFS hindcast data and the three reanalysis datasets. This process helped to quantify the similarities between the hindcast data and the reanalysis data and to test which reanalysis datasets (between the ERA-Interim, ERA5 and CFSR data) are closest to the hindcast data. The RMSE was calculated for each season (DJF, MAM, JJA and SON).

6.3.7. Correlation Analysis of Large-Scale Drivers

A series of analyses were performed to quantify the correlation between the large-scale drivers (ENSO and SAM) with wind speeds over the study region, in order to fulfil the requirements of Objective 3, which aims to understand cross-scale drivers of wind patterns in South Africa including teleconnections, synoptic and regional drivers. These correlations

[2] Numpy is a coding package that is used in Python, in order to analyse data.

[3] Similar to numpy, sklearn metrics is a package in Python which can be used to perform functions on datasets. A comprehensive overview of the package can be found by using the link: https://scikit-learn.org/stable/user_guide.html#user-guide

were conducted over the full available time period for each of the wind speed datasets at a monthly timescale and the correlations at a seasonal timescale (DJF, MAM, JJA, SON) were also evaluated.

6.3.7.1. El Niño Southern Oscillation

Section 7.6.1.1. presents the correlation coefficient values between MEI and wind speed (ERA5, ERA-Interim, CFSR and CFS). The Pearson Correlation Coefficient was calculated by using monthly MEI data, and monthly wind speed data. Section 7.6.1.2. presents the results for seasonal correlations between MEI and wind speed, where the hindcast data and three reanalyses were used. The seasons are defined as: DJF, MAM, JJA and SON and monthly data was used to group the wind speed data into the relevant seasons. The seasonal figures were created by grouping the months into the relevant seasonal bracket in CDO, and this processes was repeated for the four wind datasets and the MEI dataset. The Pearson Correlation Coefficient formula (Equation 2) was performed between the relevant wind datasets and MEI in Python, and the output was presented as spatial plots (Figure 85 to 104).

The results presented in Section 7.6.1.3. were created by extracting all the MEI values (and relevant timesteps), that were greater than and less than 0.5, to represent El Niño and La Niña events respectively, and MEI values between -0.5 and 0.5 were extracted to represent neutral phases of ENSO. The wind speed data points in the CFS, ERA5, ERA-Interim and CFSR wind datasets, were extracted to match the respective timesteps of the El Niño, La Niña and neutral MEI values. The correlation coefficient between the positive (negative) MEI periods and the corresponding wind speed values were calculated and plotted in Python. The El Niño and La Niña events that are defined by the threshold of ± 0.5 will hereafter be referred to as “moderate El Niño” and “moderate La Niña” events. The Pearson Correlation Coefficient was then calculated between each of the four datasets and MEI, and the output was represented on a spatial plot. Additionally, the minimum and maximum *r*-values were calculated for each Pearson Correlation Coefficient calculation, and the results were presented in table form (Table 5 and 6).

The results in Section 7.6.1.4 aimed to present strong El Niño and La Niña events with regards to wind speed. The study made use of a ‘strength index’ for MEI in order to determine strong El Niño phases and strong La Niña phases (Mazzarella and Giuliacci, 2009: 519). Mazzarella and Giuliacci (2009), present the following strength index (SI) scale, where a SI value of one indicates a weak El Niño event, and a SI value of six represents a very strong El Niño event (Table 2).

Table 2: Strength index for an El Niño event

Strength Index (SI)	
1	0.0 < MEI < 0.6
2	0.6 < MEI < 1.2
3	1.2 < MEI < 1.8
4	1.8 < MEI < 2.4
5	2.4 < MEI < 3.0
6	3.0 < MEI

The above strength index was originally created for El Niño phases; however, this study has adapted the index for La Niña events (Table 3).

Table 3: Strength index for a La Niña event

Strength Index (SI)	
1	-0.6 < MEI < -0.0
2	-1.2 < MEI < -0.6
3	-1.8 < MEI < -1.2
4	-2.4 < MEI < -1.8
5	-3.0 < MEI < -2.4
6	MEI < -3.0

The Strength Index was used to create the results in Section 7.6.1.4., where all MEI values greater (less) than 1.2 (-1.2), were extracted from the MEI dataset. The respective timesteps were noted, and the corresponding timesteps were extracted from the CFS, ERA5, ERA-Interim and CFSR wind speed datasets. The correlation coefficient was subsequently calculated for all the El Niño phases (where MEI is greater than 1.2) and wind speed (using the data extracted from the four datasets), and for all the La Niña phases (where MEI is less than -1.2) and wind speed. The results in Section 7.6.1 made use of monthly data. Based on the SI Index (Mazzarella and Giuliacci, 2009), where a threshold of 1.2 was used, months from the following years were extracted to represent a strong El Niño phase: 1982, 1983, 1986, 1987, 1991, 1992, 1993, 1994, 1997, 1998, 2003, 2006, 2010, 2015, 2016 and 2017. Using a threshold of -1.2, months from the following years were extracted to represent a strong La Niña phase: 1988, 1989, 1999, 2000, 2007, 2008, 2010 and 2011.

6.3.7.2. Southern Annular Mode Analysis

The results in Section 7.6.2.1. were created by performing the Pearson Correlation Coefficient calculation between the SAM Index monthly data, and between monthly wind speed data from the CFS (hindcast), ERA5, ERA-Interim and CFSR (reanalyses) datasets. The raw SAM Index data is monthly; however, the wind speed data needed to be averaged from daily data to monthly data. Section 7.6.2.2. presents the results for seasonal correlations between the SAM Index and wind speed, using the whole SAM Index dataset. The seasons were defined as: DJF, MAM, JJA, SON and monthly data was used. Similar to the MEI methodology, the seasonal figures were created by grouping the months into the relevant seasonal bracket in

CDO, and this processes was repeated for the four wind datasets and the SAM Index dataset. The Pearson Correlation Coefficient was applied to the relevant wind datasets and the SAM Index in Python, and the output was presented as spatial plots (Figure 108 to 123).

The final analysis performed using the SAM Index data, was correlating wind speed to positive (0.5 < SAM) and negative (SAM < -0.5) phases of SAM. A positive phase of SAM has been defined as all SAM Index values above 0.5, and a negative phase of SAM has been defined as all SAM Index values below -0.5. The SAM Index values between -0.5 and 0.5 have been evaluated as well, to represent the neutral phase of the Southern Annular Mode. The SAM Index data for the three phases of SAM were extracted in CDO, and the timesteps which correspond to the two phases, were extracted from the four wind datasets. The Pearson Correlation Coefficient was then calculated between the hindcast dataset and the three reanalyses and the SAM Index, for the positive, negative and neutral phases of SAM. Additionally, a seasonal analysis was performed, using the separate positive and negative extractions of SAM, where the output is a spatial plot representing the Pearson Correlation Coefficient between MEI and each of the wind datasets (ERA5, ERA-Interim, CFSR and CFS) for each season (DJF, MAM, JJA and SON).

CHAPTER 4

7. Results

This section presents the results of the CFS ensemble members compared to the ensemble means; a broad analysis of wind speed over a year and over twenty-four hours using the ERA5 reanalysis data; a detailed site analysis using all four datasets; an analysis of the difference between the three reanalysis datasets; the skill assessments between the hindcast data and the reanalysis data; and the correlations between large scale drivers and wind speed over South Africa. The results are accompanied with an analysis of the results. The results follow the same structure as the methodology (Section 6.3.) and is divided into six sub-sections.

The spatial plots presented in Section 7. display the area from -35.4° to -20.3° latitude and 14° to 35° longitude, however, only the area highlighted in green (South Africa), as well as the surrounding South Atlantic ocean and South Indian ocean on either side of the country, will be studied and discussed (Figure 7).

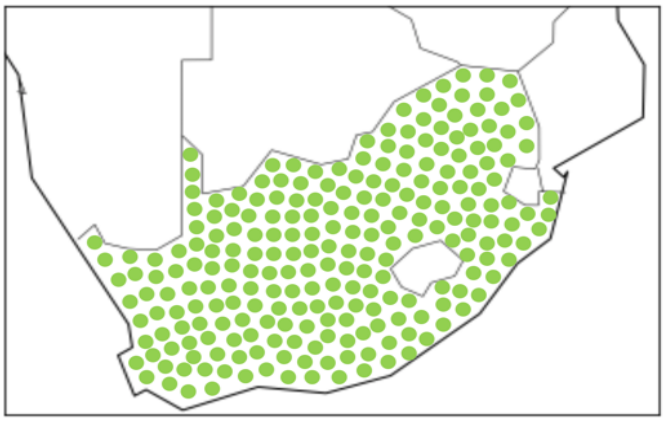
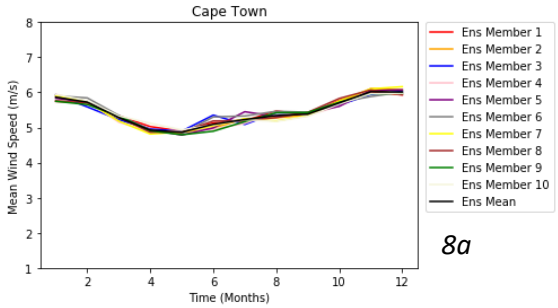


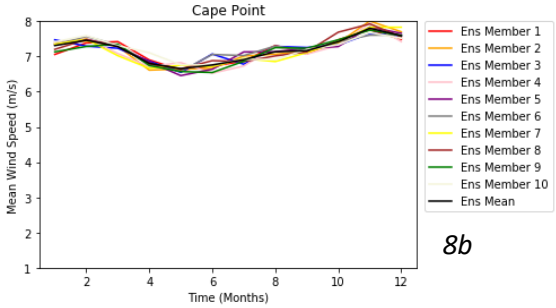
Figure 7: Map of the study area, where South Africa has been highlighted as the country that is studied

7.1. Ensemble Members and Ensemble Means

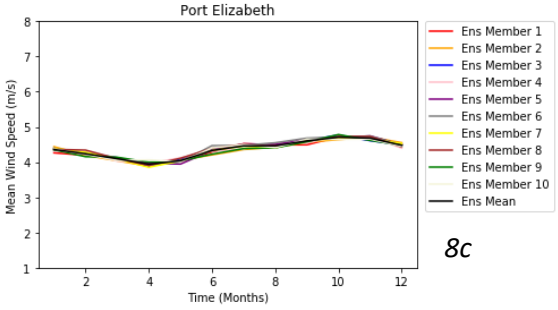
The ensemble members and ensemble means are from the CFS hindcast dataset. There are ten ensemble members, and an ensemble mean. The data has been average from 1983 to 2017, and presents the monthly climatology for each of the 16 sites. The variability between the ten ensemble members and ensemble mean is low, and therefore the ensemble mean will be used for the majority of the results, unless stated otherwise (Figure 8). The highest spread between the ensemble members is at Cape Point, where the spread is between 0.5 m/s to 1.0 m/s (Figure 8b). This could be due to the fact that Cape Point is a topographically complex region. The majority of the other sites have an average spread between 0.1 m/s and 0.2 m/s and therefore the ensemble mean was able to be used for the remainder of the study. These results were the justification for the use of the ensemble mean and subsequently resulted in the CFS hindcast data to be smoother, with lower small-scale variability. This study aimed at analysing wind patterns on a seasonal time scale, and therefore the ensemble mean was suitable.



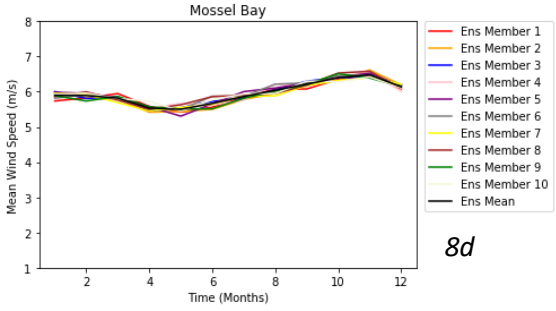
8a



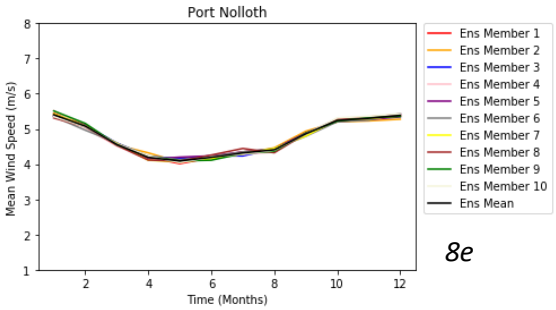
8b



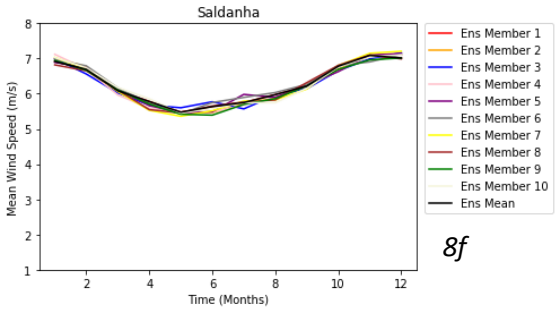
8c



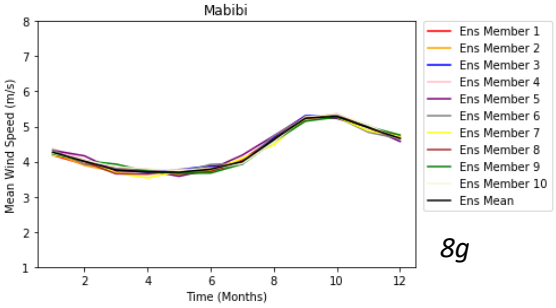
8d



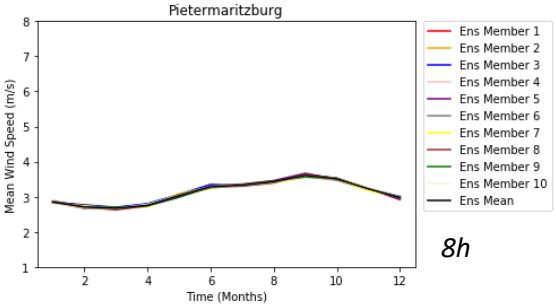
8e



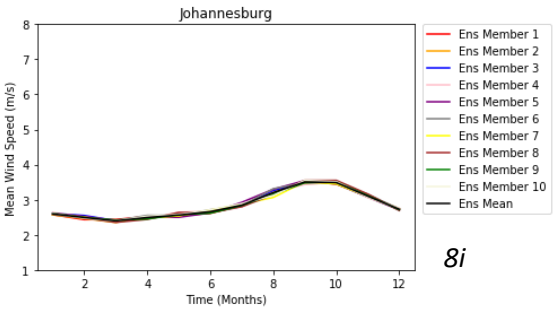
8f



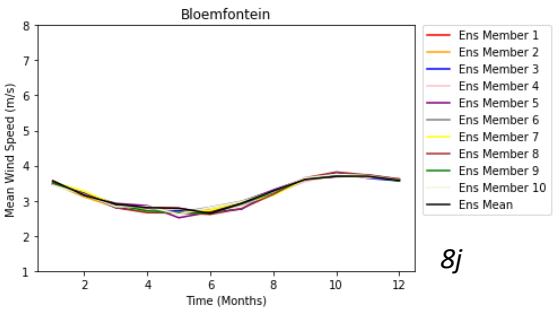
8g



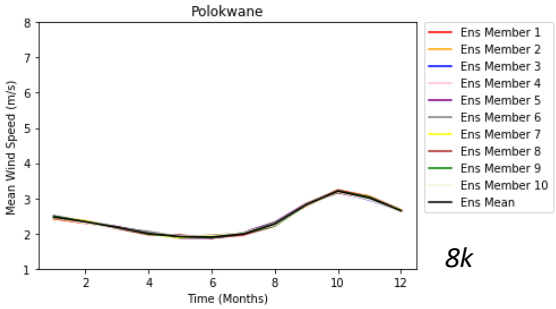
8h



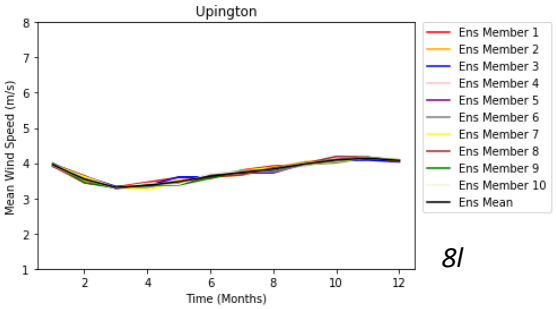
8i



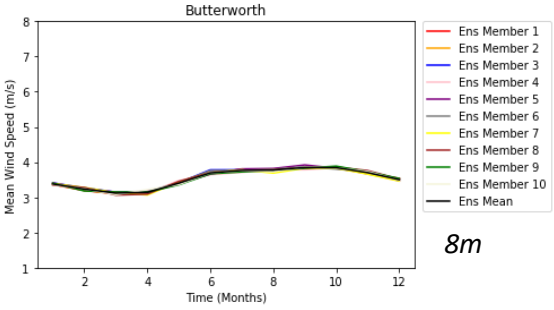
8j



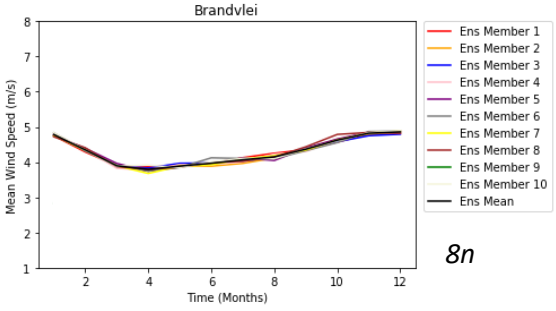
8k



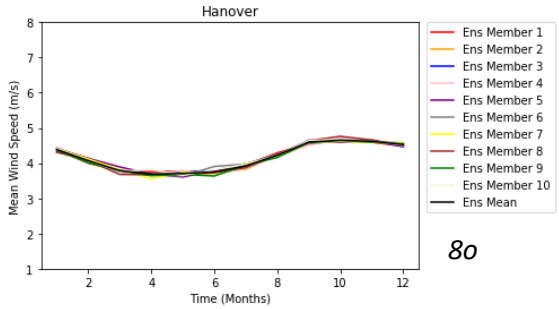
8l



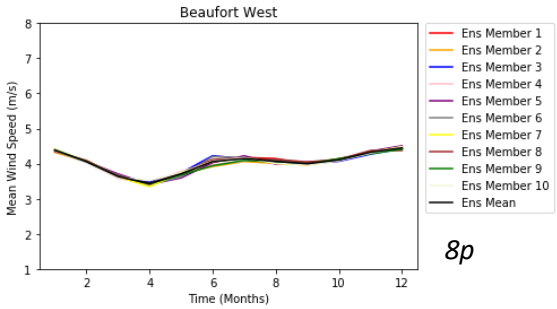
8m



8n



8o



8p

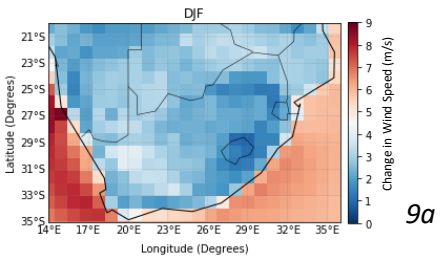
Figure 8: Figures 8a to 8p presents the monthly climatology in the form of a time series of the CFS hindcast data ensemble mean and the ten ensemble members for each of the 16 sites.

7.2. Changes in Wind Speeds over South Africa and Surrounding Oceans

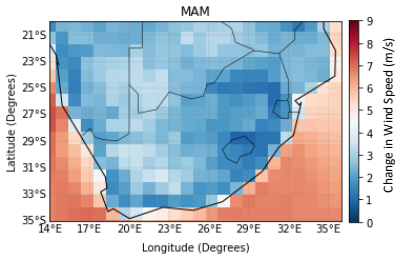
The results below present the way in which the wind speeds over the country change, over the year and over a 24-hour period (South African time). This section aims to give an overview of the wind patterns over South Africa, and to depict the changes in wind speed over the study area depending on the time of year and time of day. The CFSR and ERA-Interim figures were computed and evaluated; however, only the ERA5 figures have been presented in the study due to the fact that the three reanalysis datasets present the same wind speed patterns over the study area.

7.2.1. Seasonal Changes in Wind Speed

Over the eastern and northern interior areas of the country there are consistently lower wind speeds compared to the rest of South Africa. Wind speeds over the southern interior parts of South Africa (which include regions of the Northern Cape, Western Cape and the Eastern Cape), experience higher wind speeds, especially in JJA (Figure 9c) and SON (Figure 9d). The wind speed patterns over the Atlantic Ocean and the Indian Ocean also change between seasons. In DJF (Figure 9a) and SON (Figure 9d), the wind speeds over the Atlantic Ocean increase, whereas in JJA (Figure 9c), the wind speeds increase over the Indian Ocean.



9a



9b

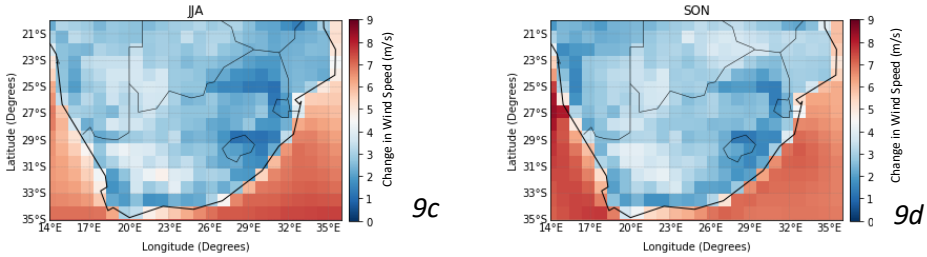


Figure 9: Figures 9a to 9d present the change in seasonal wind speed. The Figures have made use of ERA5 wind speed data over a time period of 1982 to 2017

7.2.2. Four-Hourly Changes in Wind Speed

There are visible changes in wind speed over South Africa, over a period of 24 hours and between the four seasons (Figure 10 to 13). Based on the figures in Section 7.2.2., during all seasons, the highest wind speeds are experienced at 12:00 and 16:00, with maximum wind speeds over the south-western regions of the country. The lowest wind speeds are experienced at 00:00 and 04:00, especially over the eastern and northern regions of South Africa. The majority of the highest wind speeds are experienced during DJF (Figure 10) and SON (Figure 13), during 08:00, 12:00 and 16:00 (Table 4). Over the majority of the regions across South Africa, the weakest wind speeds are experienced during MAM (Figure 11) and JJA (Figure 12). Throughout the 24 hour period, the wind speeds over the Atlantic Ocean and Indian Ocean do not fluctuate as much as over land (Figure 10 to 13), this is due to the fact that the temperature over land fluctuates more than over the ocean, which is caused by the varying topography over land. The temperature gradient over land fluctuates, which results in a greater variance of wind speed throughout a 24-hour period.

December – January – February

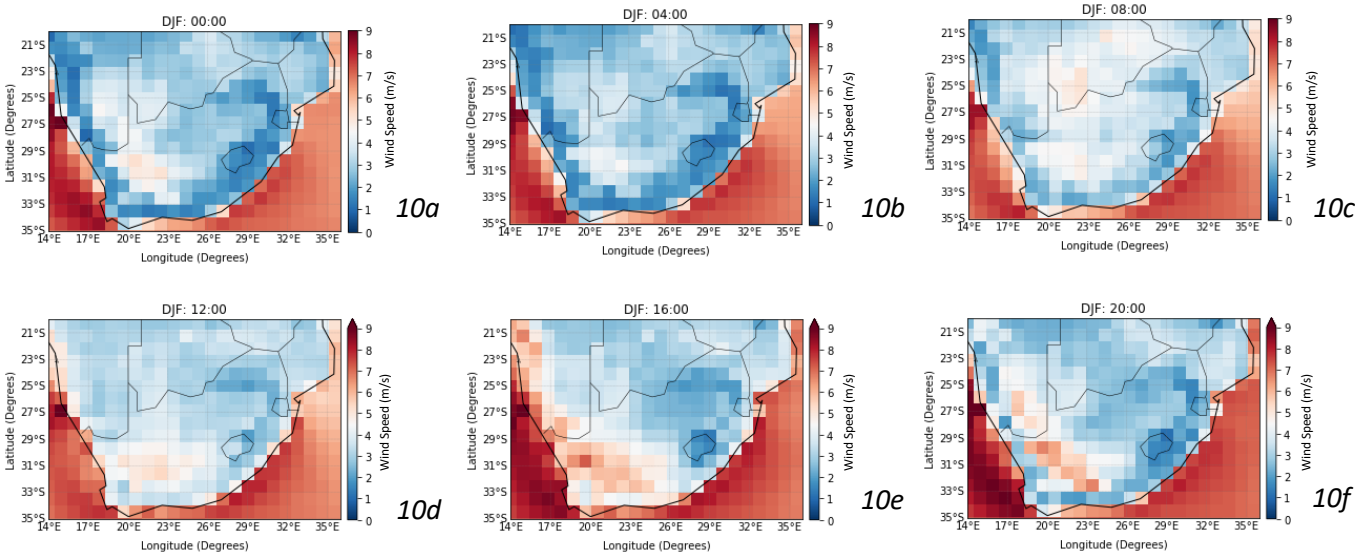


Figure 10: Plot of the 4-hourly change (10a to 10f) in wind speed in DJF, using ERA5 reanalysis wind speed data

March – April – May

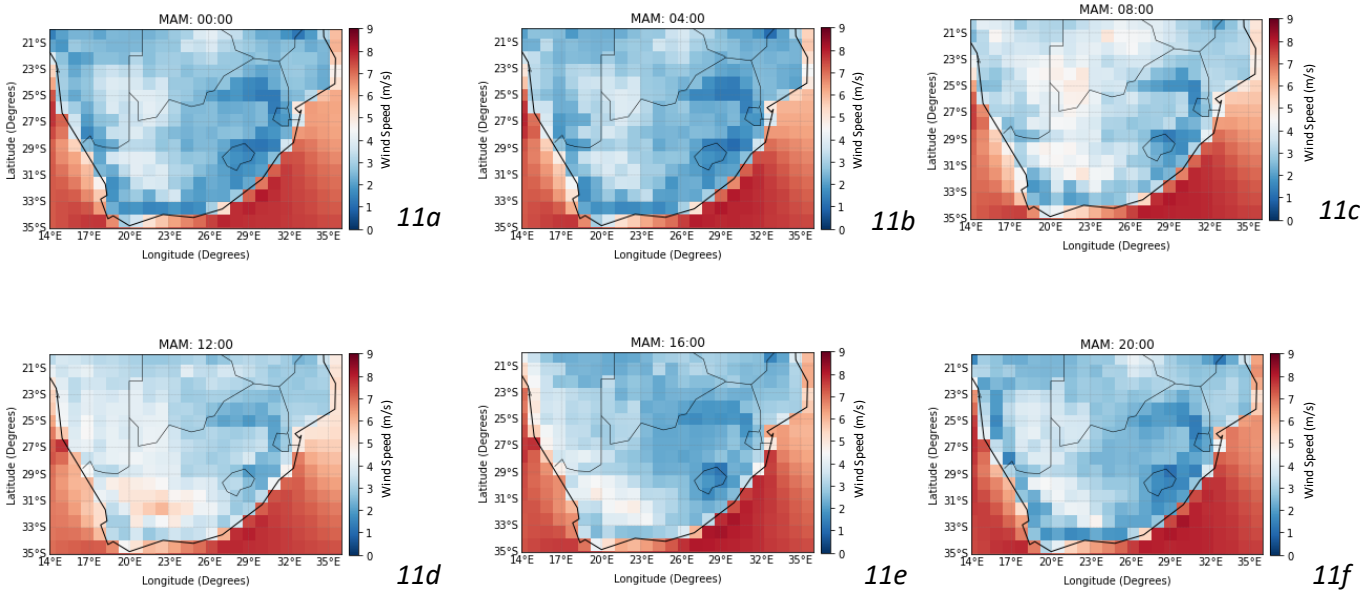


Figure 11: Plot of the 4 hourly change (11a to 11f) in wind speed in MAM, using ERA5 reanalysis wind speed data

June – July – August

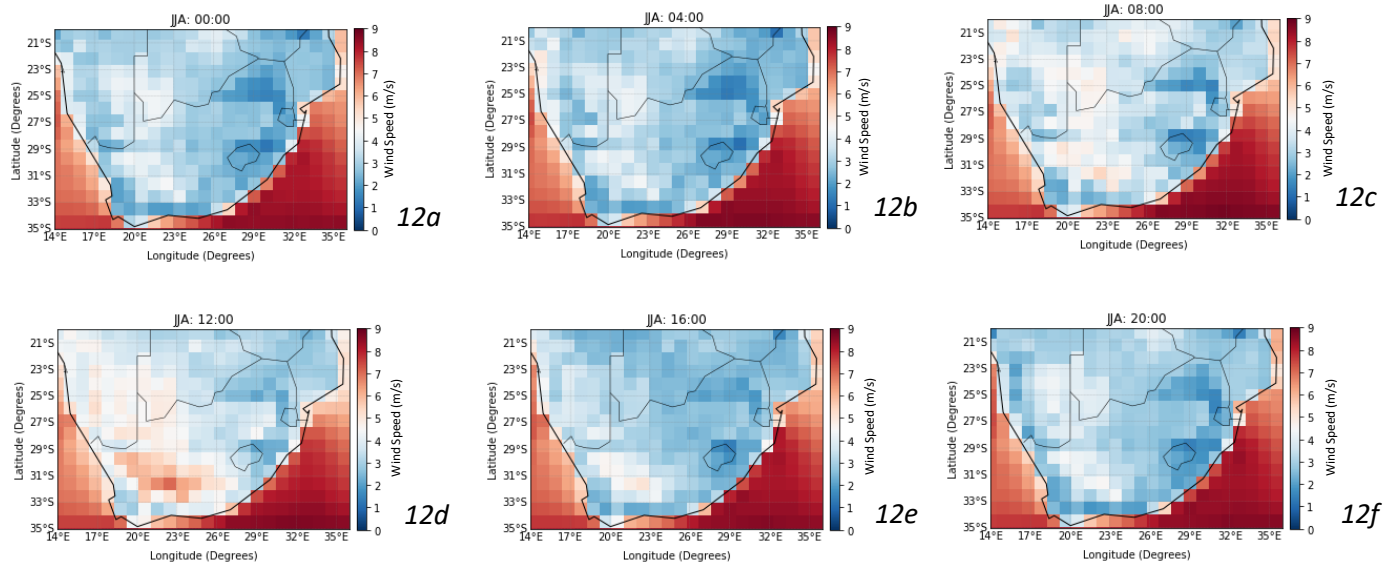
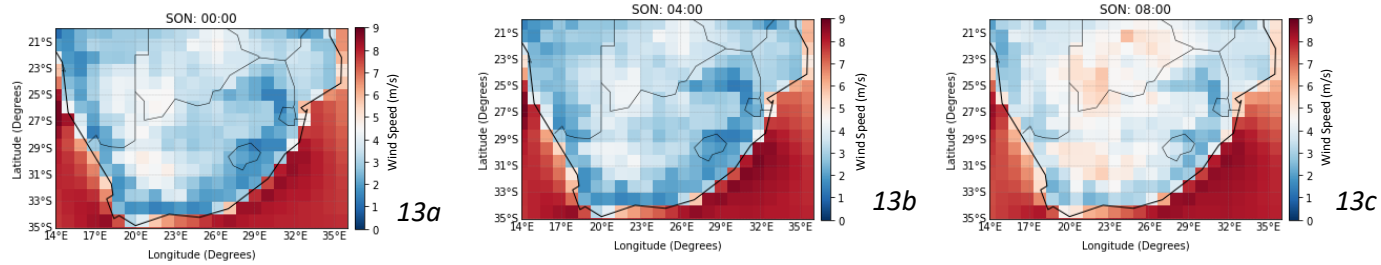


Figure 12: Plot of the 4 hourly change (12a to 12f) in wind speed in JJA, using ERA5 reanalysis wind speed data

September – October – November



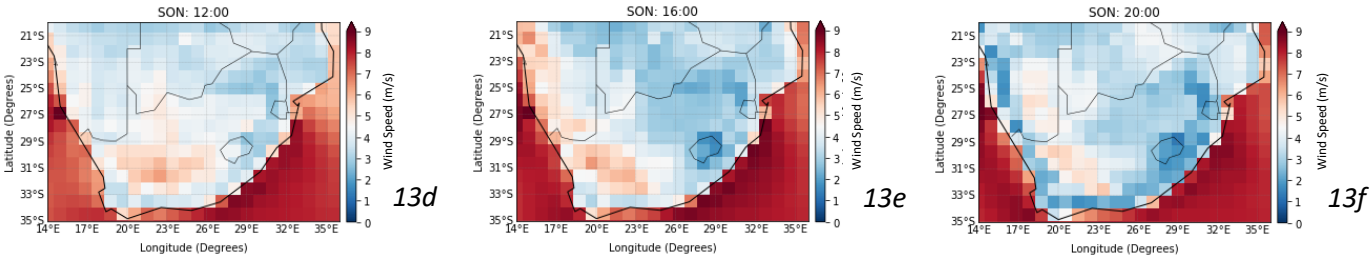


Figure 13: Plot of the 4 hourly change (13a to 13f) in wind speed in SON, using ERA5 reanalysis wind speed data

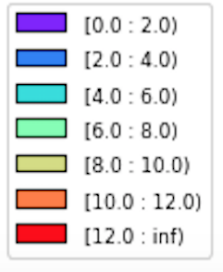
Table 4: Summary table of the results presented in Figure 10, 11, 12 and 13. The wind speed range (minimum and maximum wind speeds) has been given for each season (DJF, MAM, JJA, SON) and for each timestep. Only the wind speeds over land (South Africa) have been considered in the table.

	DJF	MAM	JJA	SON
00:00	1 m/s – 5 m/s	1 m/s – 4 m/s	0.5 m/s – 4 m/s	1 m/s – 5 m/s
04:00	0.5 m/s – 4.5 m/s	1 m/s – 4 m/s	0.5 m/s – 4.5 m/s	1 m/s – 4.5 m/s
08:00	1.5 m/s – 4.5 m/s	2 m/s – 5 m/s	1 m/s – 5 m/s	3 m/s – 5 m/s
12:00	2 m/s – 5 m/s	2 m/s – 6 m/s	2 m/s – 6 m/s	4 m/s – 6 m/s
16:00	2 m/s – 6.5 m/s	1 m/s – 5 m/s	1 m/s – 5 m/s	2 m/s – 6.5 m/s
20:00	1 m/s – 6.5 m/s	1 m/s – 4.5 m/s	1 m/s – 4 m/s	1 m/s – 5.5 m/s

7.3. Site Analysis

This section of the results aims at analysing the wind patterns over South Africa in more detail, by studying 16 sites across the country. For each point and each wind speed dataset, a box plot, a time series and wind roses have been created in order to understand how the wind speed and direction differs between each region. This section also aims to answer Objective 1, by identifying the difference between the four models: CFS hindcast data and three reanalysis datasets: CFSR, ERA-Interim and ERA5; which can be seen in the box plots and the time series graphs. An ensemble member that was most similar to the ensemble mean was used to create the wind roses because the ensemble mean smoothed the hindcast data and was unable to present a clear understanding of the wind directions for each area.

The box plots indicate that the spread of the CFS data is lower than the spread of the three reanalysis datasets. This can be seen on the boxplots where the difference between quartile one and three is smaller for the CFS boxplots, compared to the boxplots of the reanalyses. The three reanalysis wind datasets have a greater amount of outliers, compared to the hindcast data. The lack of outliers seen in the CFS ensemble mean data can be explained by the fact that an ensemble mean has been used, based on the results in Section 7.1. In order to account for the outliers in the CFS data, a combined box plot was created for the ensemble members.



The colour scale is used for the wind roses below. The scale is in m/s, where the purple bin indicates low wind speeds (0 m/s to 2 m/s) and the red bin indicates high wind speeds (12 m/s – infinity). The colour scale is consistent throughout all the wind roses.

7.3.1. Cape Town

Cape Town is in the Western Cape, which is situated towards the south-west of South Africa. Cape Town experiences peak wind speeds between November and January (Figure 14), and minimum wind speeds are experienced in May. The wind speeds decrease over Cape Town due to the equatorward shifting of the South Atlantic high pressure, which results in a decrease in the pressure gradient force over the region (Fawcett et al., 2007). In June and July, the wind speeds increase again. This seasonal pattern can be seen throughout the four wind speed datasets; however, there is a difference in wind speed values between the datasets, where ERA5 has the lowest wind speeds and CFS has the highest wind speeds. The box plots indicates that the CFS ensemble members and the CFSR reanalyses have the most outliers of high wind speeds. The use of an ensemble mean reduces the spread of data; there are fewer outliers and the difference between the first quartile and the third quartile is the smallest compared to the four remaining box plots (Figure 15).

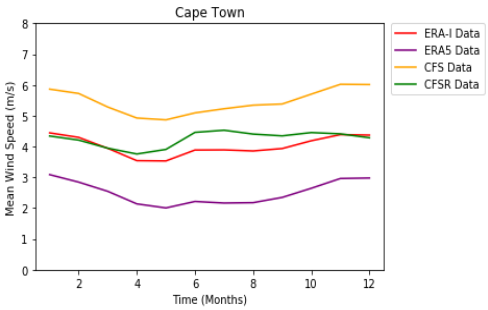


Figure 14: Cape Town time series

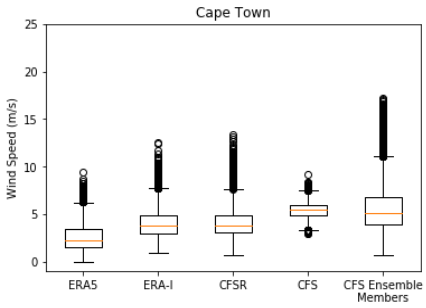


Figure 15: Cape Town box plots

The wind roses for Cape Town present the direction of wind over this point for each month. Between September to April, there is a strong south, south-easterly and south-westerly wind (Figure 16). The ERA-Interim, CFSR and CFS datasets present a stronger south-easterly component, compared to ERA5, during austral summer. During austral winter, the dominant north westerly winds over the region is the result of a low pressure system over the Western Cape (Fawcett et al., 2007). The ERA5 wind roses (Figure 16a) are slightly different to the ERA-Interim, CFSR and CFS wind roses (Figure 16b to 16d). These results could be explained by the high spatial resolution of the ERA5, which is capturing a topographic modification whereas the CFS is a coarser resolution which results in the topography being smoothed, and there is relatively little topographic modification.

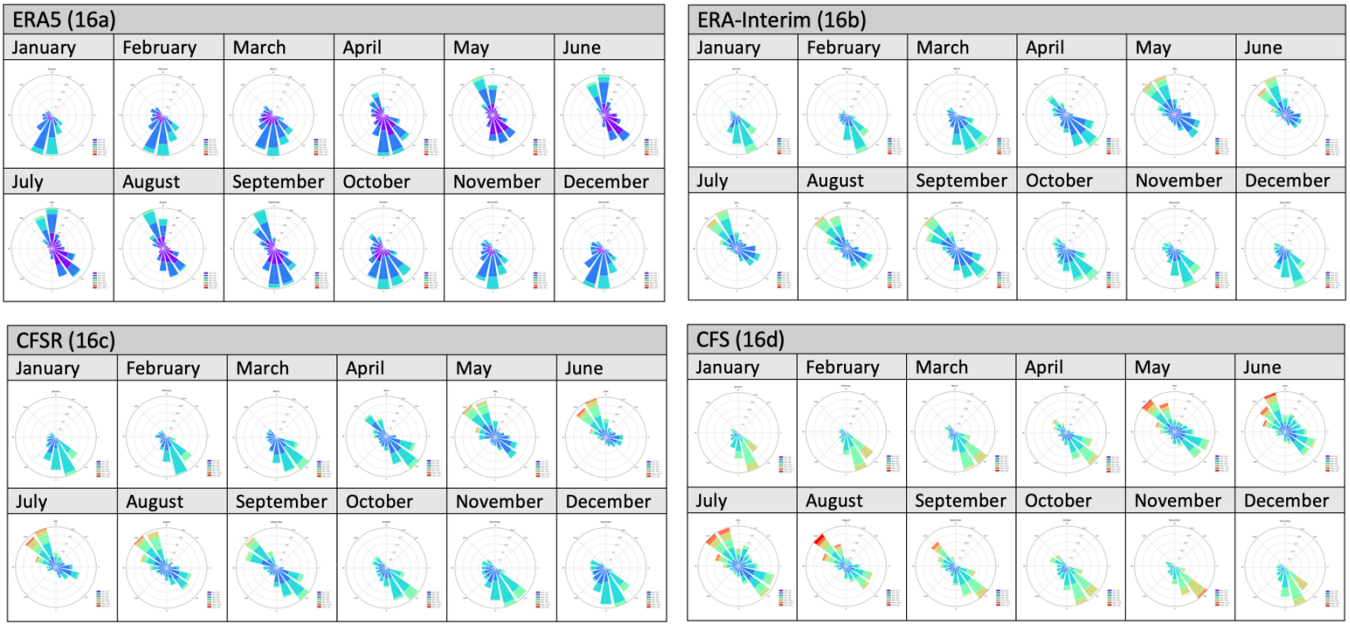


Figure 16: Cape Town wind roses

7.3.2. Cape Point

Cape Point is one of the southern-most areas in South Africa, and it is situated in the Western Cape. The distance between Cape Town (the point used in this study is Cape Town International Airport) and Cape Point is approximately 45 kilometres (based on Google Earth); however, it can be seen that there is a difference in the wind speeds between the two study areas (Figure 14 to 19). The average wind speed over Cape Point is higher than the wind speeds over Cape Town, and this result is consistent throughout the four datasets. The effects of the westerly wind belt can be seen over Cape Point through the high wind speeds (Veitch et al., 2019). Cape Point experiences peak wind speeds in November and February, and a decrease in wind speeds in May. The difference between the four datasets is less noticeable than over Cape Town; however, the CFS mean wind speeds are higher than the reanalysis data (Figure 17).

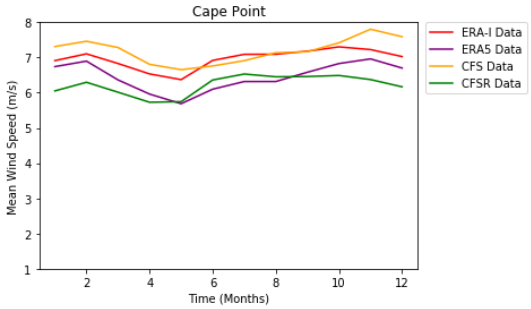


Figure 17: Cape Point time series

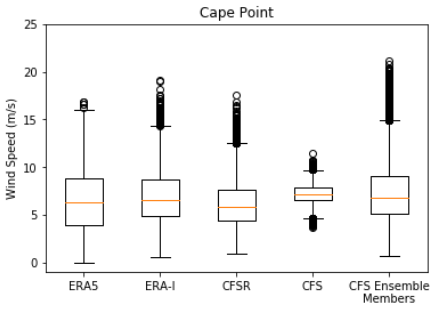


Figure 18: Cape Point box plots

The wind roses show a strong south easterly wind from October to March, and from May to June, north westerly winds are predominant. The wind speeds shown in the wind roses are

consistent with the timeseries' (Figure 17) which show that the wind speeds vary between 6 m/s and 8 m/s. The four datasets present consistent wind direction results (Figure 19).

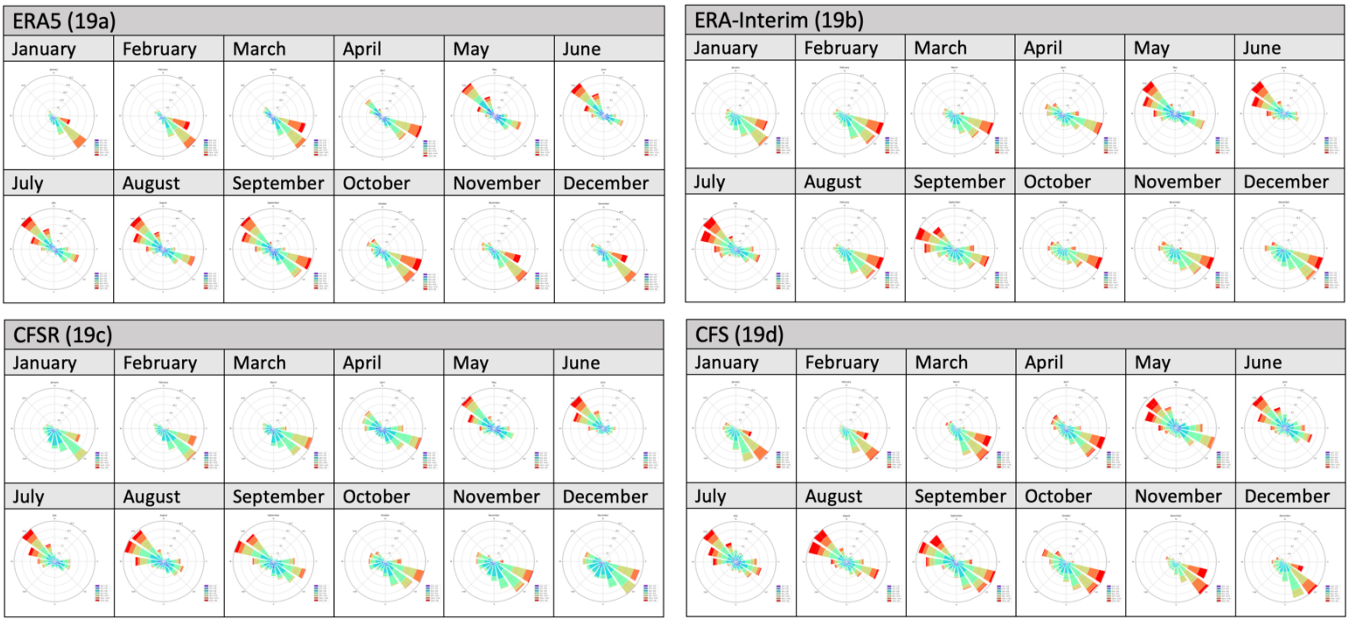


Figure 19: Cape Point wind roses

7.3.3. Port Elizabeth

Port Elizabeth is situated in the Eastern Cape, along the south-east coast of South Africa. Port Elizabeth experiences maximum wind speeds in November to January, and there is a peak in June (Figure 20). The increase in wind speed during June is caused by the strong pressure gradient between the mid-latitude lows (in their most northerly position) and high pressure systems over the interior. The wind speeds decrease in March, and in September. The ERA5 wind speeds are lower than the CFSR, ERA-Interim and CFS wind speeds, by approximately 1.5 m/s (Figure 20 and 21).

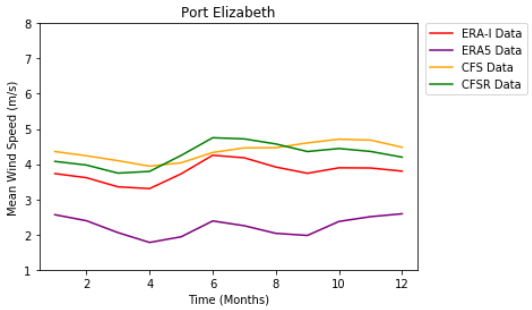


Figure 20: Port Elizabeth time series

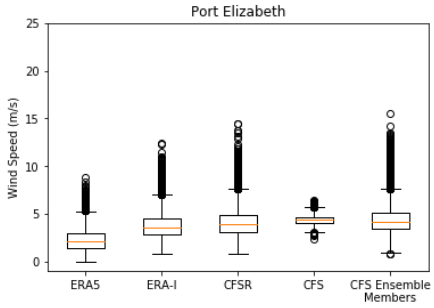


Figure 21: Port Elizabeth box plots

From September to March, the dominant wind directions are south-westerly, southerly and south-easterly winds. The ERA5, ERA-Interim and CFSR wind roses indicate that the wind direction starts to change to a north-westerly wind in April; however, the CFS wind roses indicate that this change in the wind direction only occurs in May. From May to July the dominant wind direction is a north-westerly wind, and this wind direction is consistent

between the four datasets. There is an inconsistency of wind directions in August between the ERA5 (Figure 22a) and the three remaining datasets (Figure 22b to 22d). ERA-Interim, CFSR and CFS indicate that there is a dominant westerly wind, whereas the ERA5 data indicates that the dominant wind direction ranges between a west-south-westerly wind to a north-westerly wind (Figure 22).

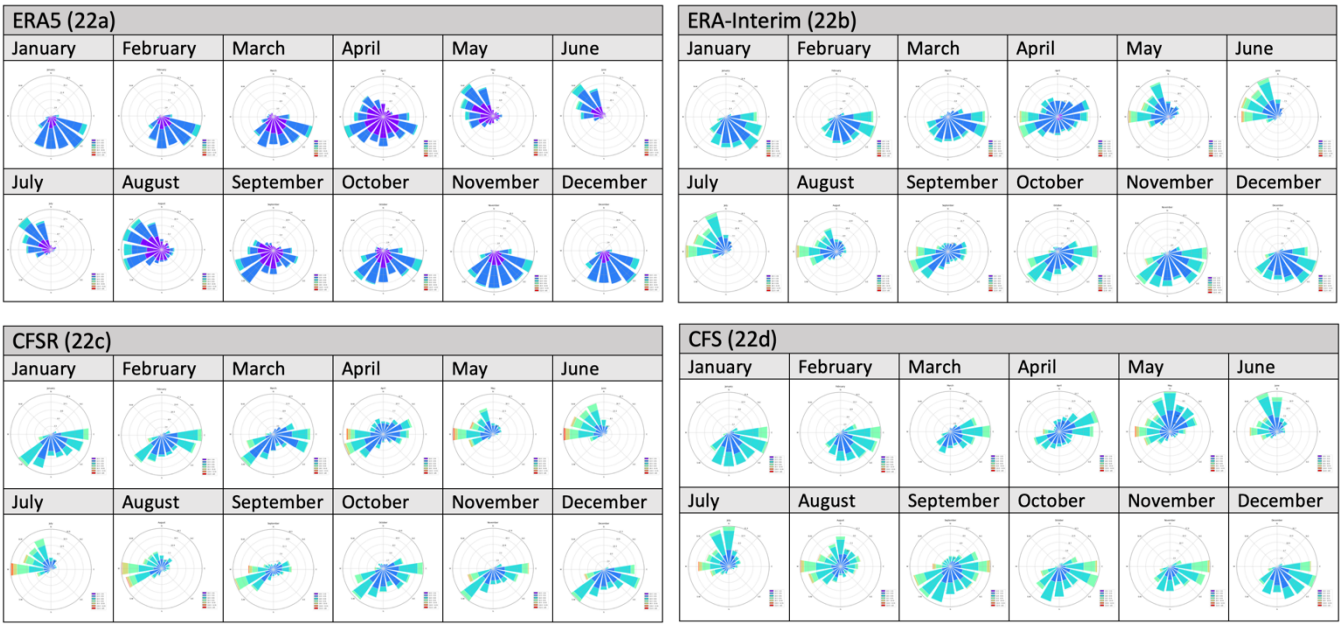


Figure 22: Port Elizabeth wind roses

7.3.4. Mossel Bay

Mossel Bay is in the Western Cape, along the coast of South Africa and it is situated west of Port Elizabeth. The wind speeds over Mossel Bay vary between 4.5 m/s and 6.5 m/s (Figure 23 and 24). The reanalysis datasets show that Mossel Bay experiences a peak in wind speeds in June, however, the CFS hindcast data shows a decrease in wind speed in June. All four datasets indicate that there is a decrease in wind speeds in May. The CFS wind speed data is higher than the reanalysis data for all months except in June (Figure 23).

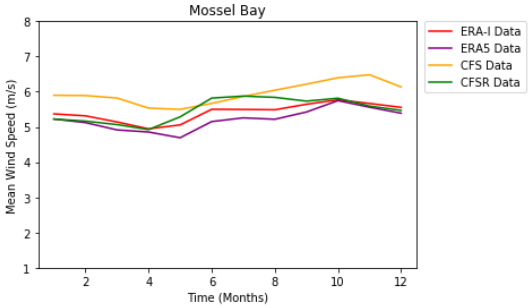


Figure 23: Mossel Bay time series

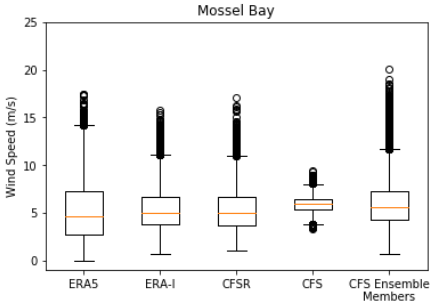


Figure 24: Mossel Bay box plots

There is a strong easterly wind between the months of October to April, which is accompanied by a weaker, yet still notable, west to west-south-westerly wind. Between May to September,

the westerly wind becomes the dominant wind direction. The results are consistent between the ERA5, ERA-Interim, CFSR and CFS datasets (Figure 25).

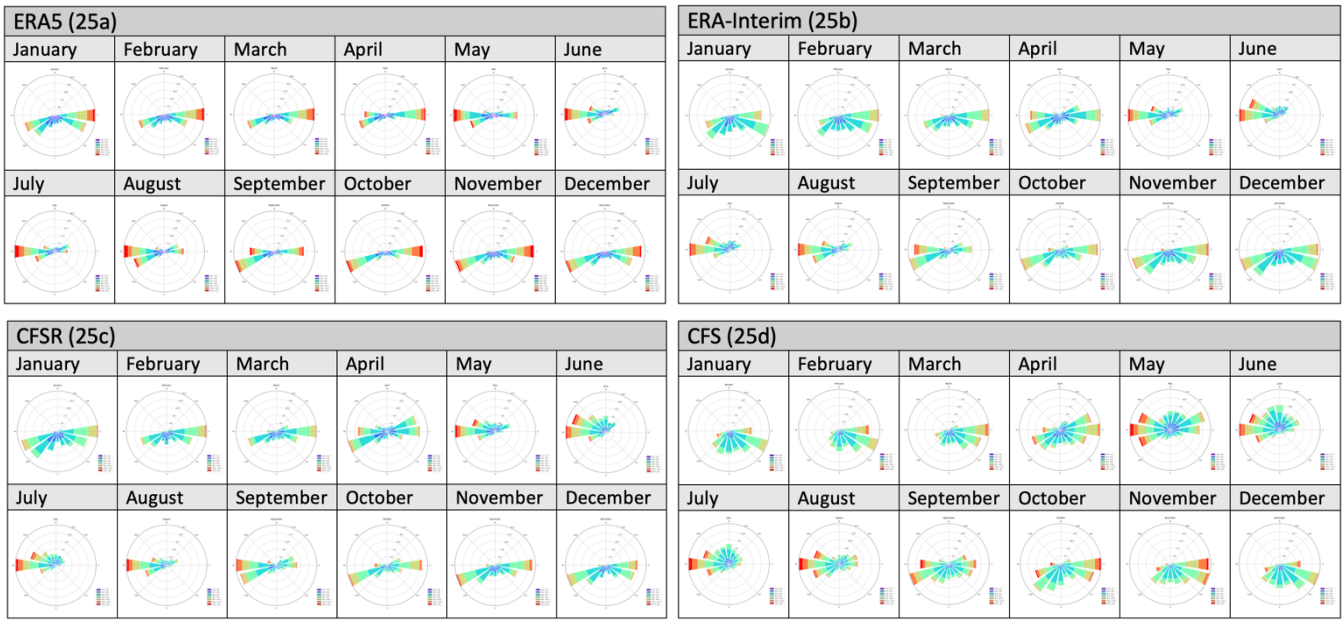


Figure 25: Mossel Bay wind roses

7.3.5. Port Nolloth

Port Nolloth is a coastal town situated along the west coast of South Africa in the Northern Cape. The maximum wind speeds are experienced between October and January, and there is a general decrease in wind speeds during austral winter, with a slight increase in wind speeds during June and July (Figure 26). The ERA5 wind speeds are lower than the three other datasets and that the CFS wind speeds are higher than the reanalyses wind speeds (Figure 26 and 27).

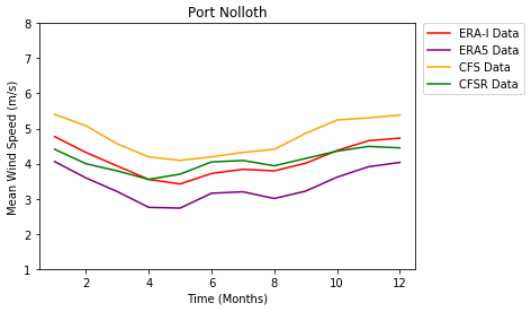


Figure 26: Port Nolloth time series

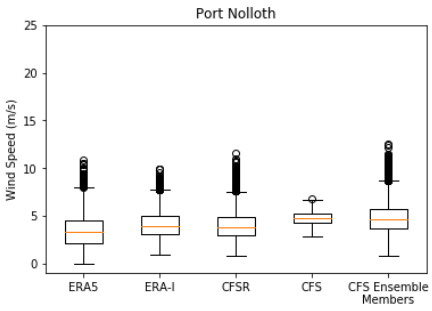


Figure 27: Port Nolloth box plots

There is a dominant southerly wind between September and March over Port Nolloth. In April and May a stronger north-north-easterly wind develops and in June and July, there is a strong north-easterly wind. The South Atlantic High Pressure results in south-easterly and north-easterly winds over Port Nolloth. The wind directions are consistent between the four datasets (Figure 28).

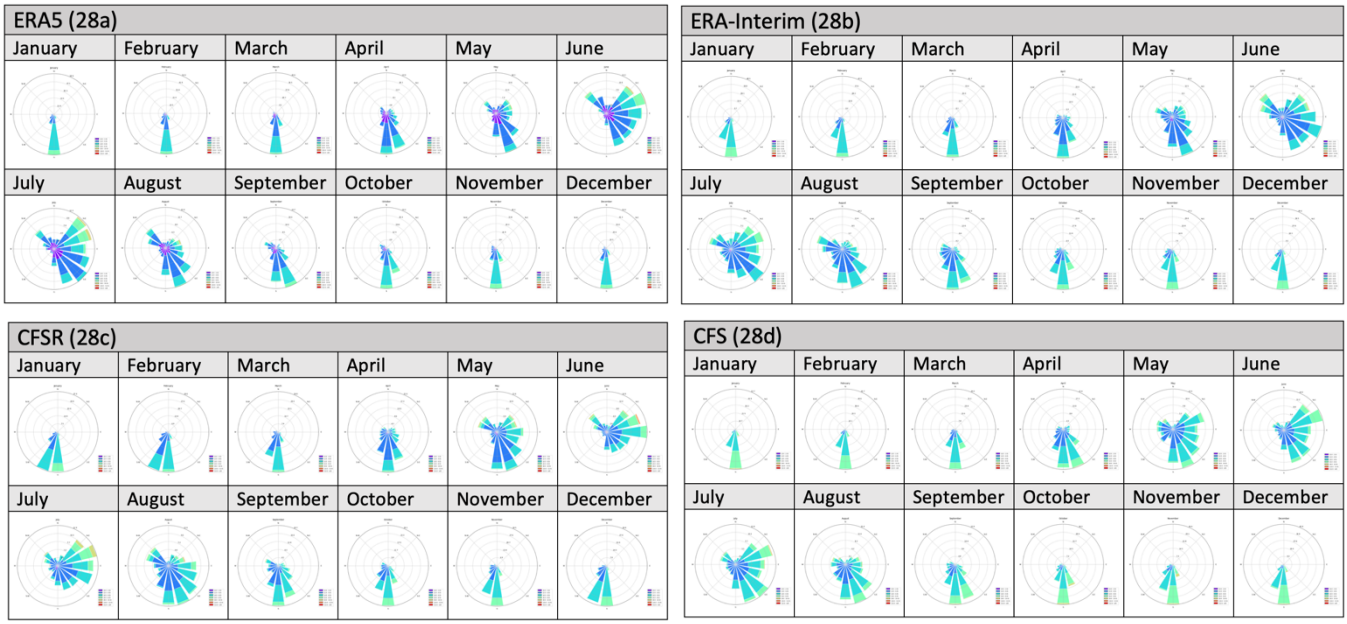


Figure 28: Port Nolloth wind roses

7.3.6. Saldanha

Saldanha is situated along the West Coast of South Africa in the Western Cape. Saldanha experiences peak wind speeds between November and January, and a decrease in wind speeds between April and June. The wind speeds from the CFS hindcast dataset are higher than the wind speeds from the reanalysis datasets; however the three reanalysis datasets present similar wind speed values (Figure 29 and 30).

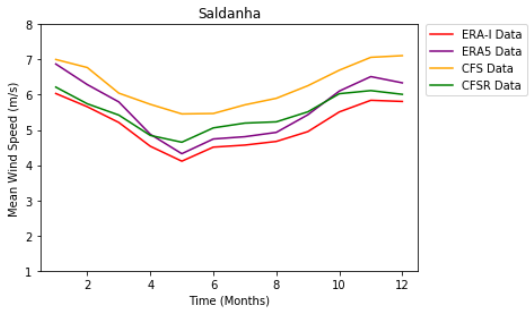


Figure 29: Saldanha time series

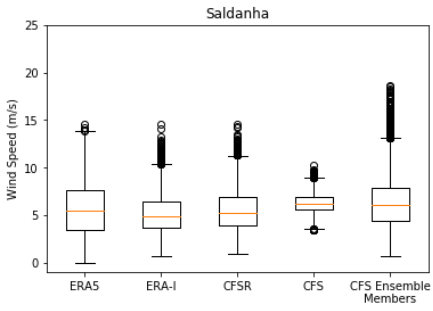


Figure 30: Saldanha box plots

The wind roses for Saldanha indicate that there is a southerly wind during the austral summer months. Between May to August, a north-westerly and north-north-westerly wind develops, along with a weaker southerly to south-south-easterly wind. The results between the four datasets are consistent, except for June, where the ERA-Interim and CFSR data indicates that there is a strong north-west-northerly wind and the ERA5 and CFS datasets show a weaker north-west-northerly wind and a stronger south-easterly wind (Figure 31).

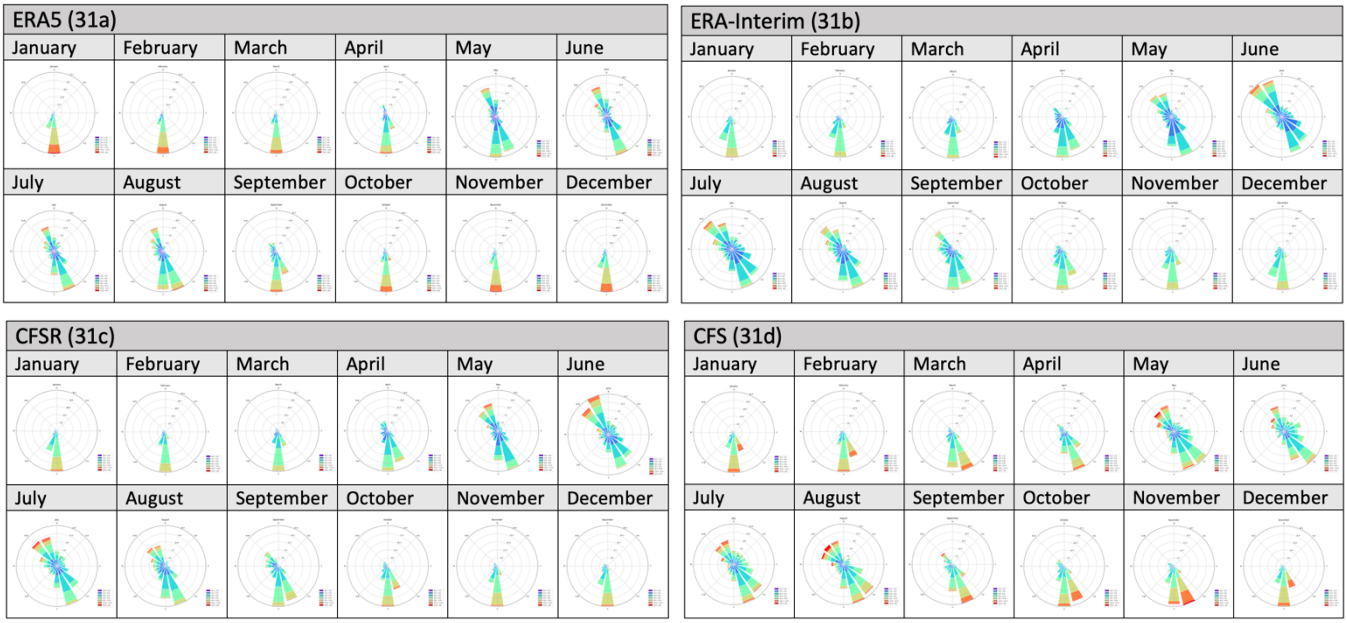


Figure 31: Saldanha wind roses

7.3.7. Mabibi

Mabibi is situated on the east coast of South Africa in northern Kwa-Zulu Natal. Peak winds are experienced in September and October, and wind speeds decrease between March to June. ERA-Interim wind speeds are highest compared to the other datasets and the CFS and ERA5 wind speeds are the same up until April. In April to December, the ERA5 wind speeds are lower than the three other datasets. The wind speeds over this area vary between 3.5 m/s to 6 m/s, depending on the model (Figure 32 and 33).

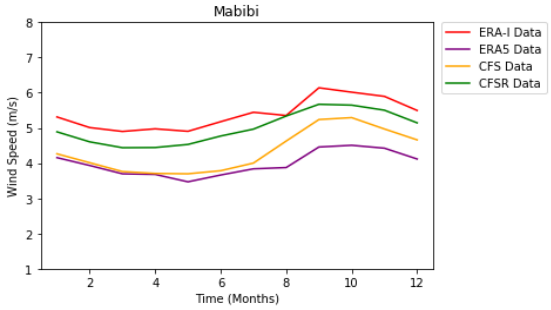


Figure 32: Mabibi boxplots

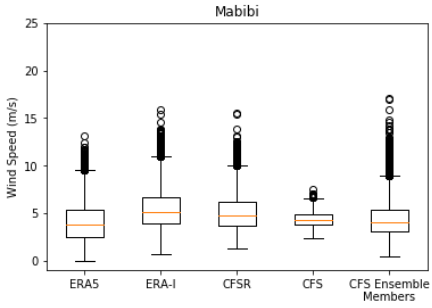


Figure 33: Mabibi timeseries

Between November and April there is a southerly wind and a north-north-easterly wind present; however, in April, the north-north-easterly becomes the predominant wind direction. Between May and October, the southerly wind becomes less dominant, and there is a strong north to north-north-easterly wind. The northern shift of the South Indian High Pressure is responsible for these wind directions during the austral winter months. The ERA5, ERA-Interim, CFSR and CFS datasets present consistent wind roses (Figure 34).

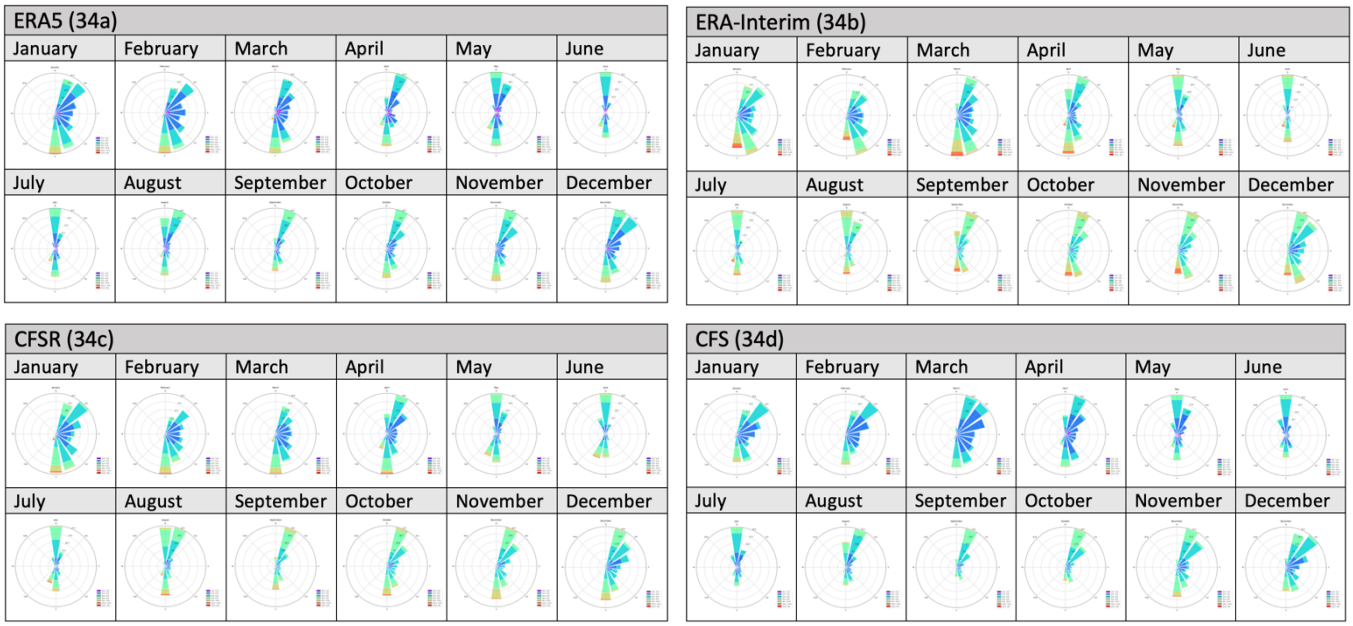


Figure 34: Mabibi wind roses

7.3.8. Pietermaritzburg

Pietermaritzburg is an inland region situated in Kwa-Zulu Natal, and despite the fact that Pietermaritzburg and Mabibi share a province, the wind speed values differ between the two regions. There is a difference in wind speeds between the four datasets over Pietermaritzburg. The ERA5 reanalysis dataset indicates that the wind speed throughout the year fluctuates between 1 m/s and 2 m/s, whereas the CFS hindcast data shows that the annual wind speed ranges between 2.5 m/s to 3.5 m/s. The ERA-Interim and CFSR datasets indicate that the average wind speeds throughout the year are between 2 m/s to 3 m/s. There are also varying wind patterns that emerged. The CFS dataset shows peak wind speeds are experienced between September and October, whereas the CFSR data indicates that there is a peak in wind speed in June and July (Figure 35). The difference between the third quartile and first quartile is small, indicating that throughout the year, the wind speeds over Pietermaritzburg do not fluctuate as much as other areas around South Africa (Figure 36).

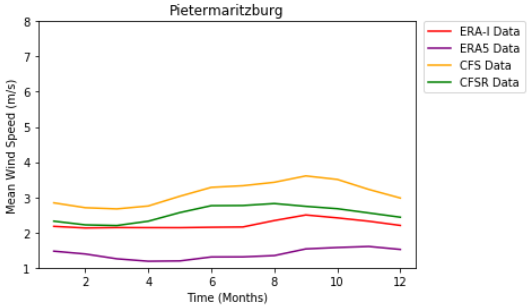


Figure 35: Pietermaritzburg time series

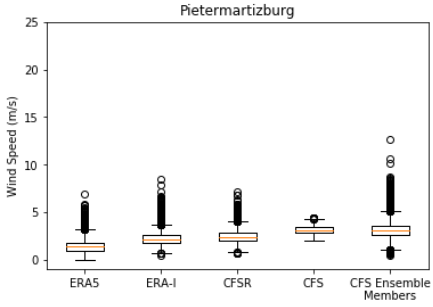


Figure 36: Pietermaritzburg box plots

Multiple wind directions are experienced over Pietermaritzburg during the austral summer months (Figure 37). The wind direction presented between the four datasets differ to one another during certain months (Figure 35 and 37). During April, the ERA5 wind rose and the CFS wind rose are consistent with one another, whereas it differs with the ERA-Interim and CFSR wind roses. During May to July; however, there are consistent wind directions between

the ERA-Interim, CFSR and CFS wind roses and not with the ERA5 wind roses. The low wind speeds presented in the ERA5 dataset (Figure 35), can be seen in the wind roses (Figure 37).

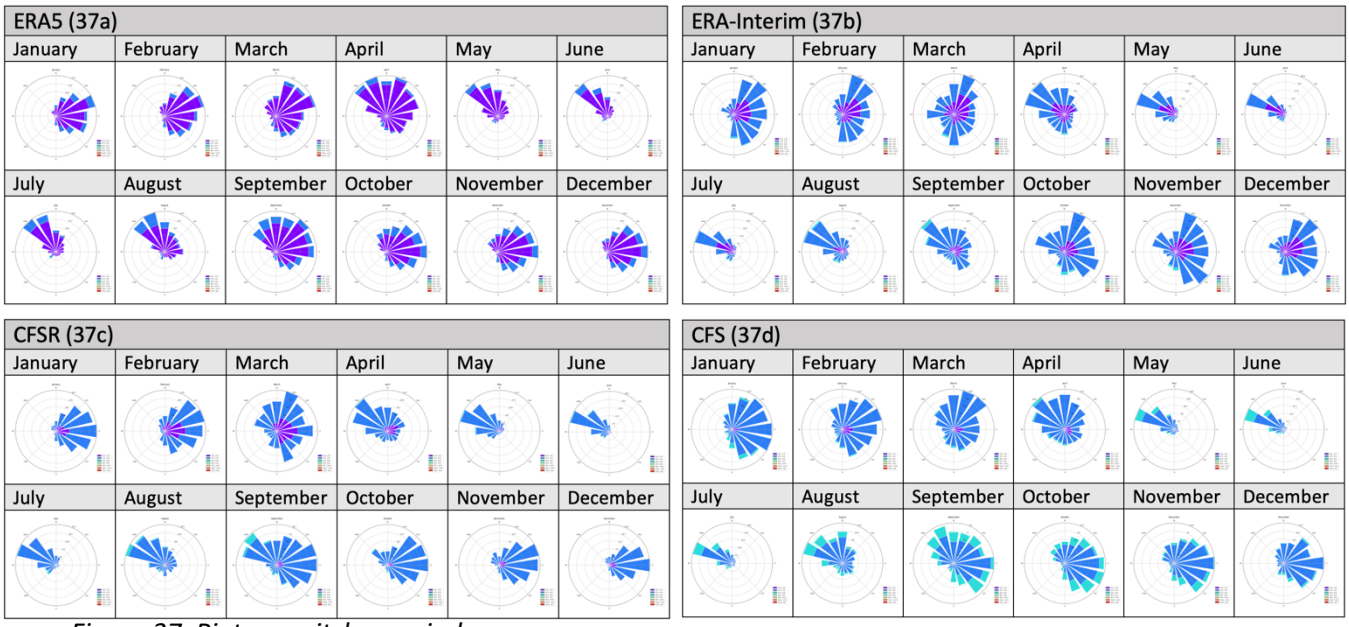


Figure 37: Pietermaritzburg wind roses

7.3.9. Johannesburg

Johannesburg is an inland region situated in Gauteng towards the northern interior of South Africa. The average wind speeds over Johannesburg are lower than the majority of the wind speeds over coastal regions. Wind speeds peak in September and decrease between March and April (Figure 38). Similar to Pietermaritzburg, the difference between the third quartile and first quartile is small, indicating the variability of wind speeds over a year is generally low (Figure 39). The ERA5 wind speeds are lower than the ERA-Interim, CFSR and CFS wind speeds (Figure 38 and 39).

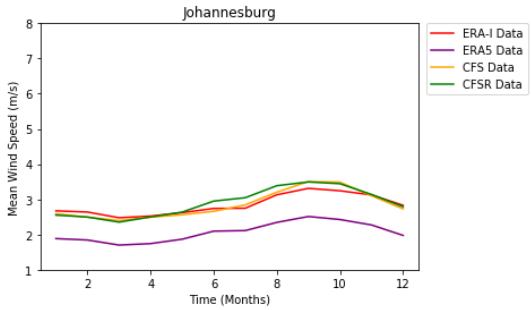


Figure 38: Johannesburg time series

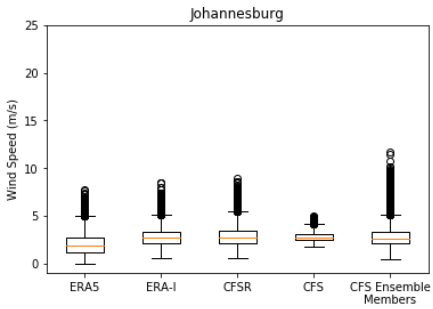


Figure 39: Johannesburg box plots

The four datasets present varying wind directions, which can be seen in May and June specifically. The three reanalysis datasets indicate that during May and June, there is a predominant south-south westerly wind and ERA5 and CFSR indicate that weaker west-north-westerly winds are also present (Figure 40a to 40c). The CFS hindcast data however, presents a different result, where in May there is a predominant north-north-westerly wind and in

June, the wind direction ranges from a south-westerly wind to a easterly wind (Figure 40d). In January and February there is a dominant east-north-easterly wind, seen in all four datasets. There is consistency between the four reanalysis datasets from August to December and the wind speeds are consistent with Johannesburg's time series which shows a peak in wind speed in September, October and November (Figure 38).

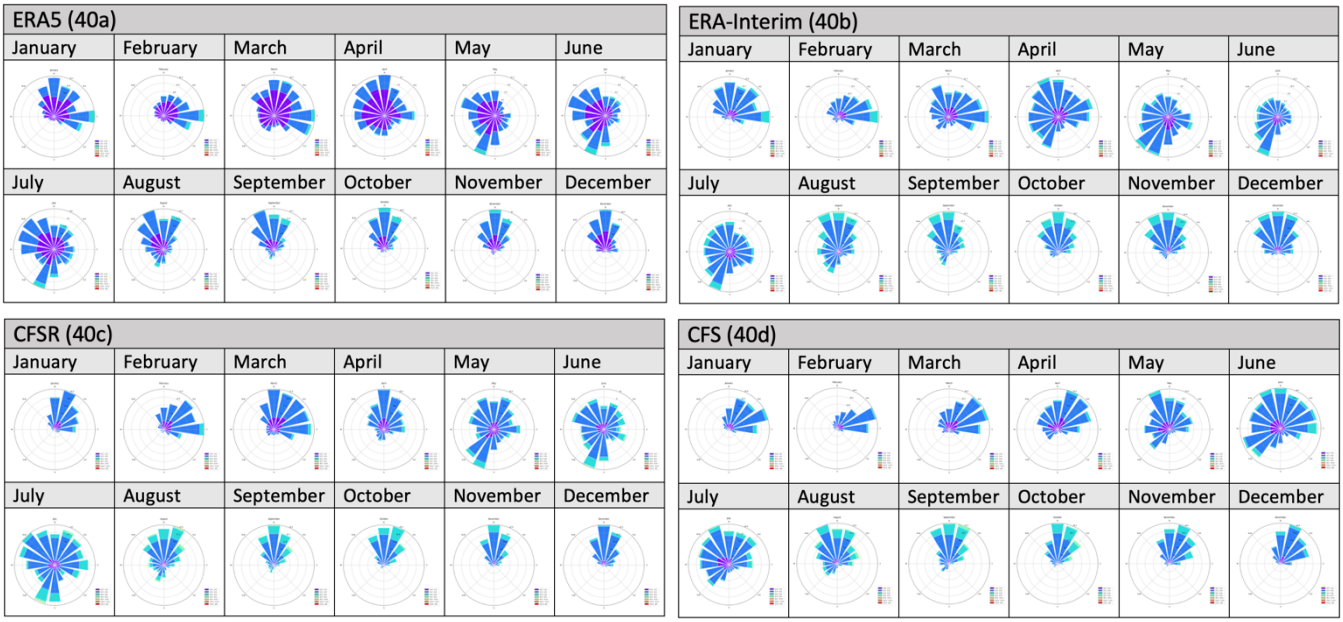


Figure 40: Johannesburg wind roses

7.3.10. Bloemfontein

Bloemfontein is the capital of the Free State, which is situated in the interior of the country, south of Gauteng. The wind speed patterns over Bloemfontein are similar to Johannesburg, where peak wind speeds can be seen in September and October and there is a decrease of wind speeds between April to July. The average wind speeds over Bloemfontein are between 2.5 m/s and 3.5 m/s (Figure 41 and 42). The ERA5 reanalysis wind data has a median wind speed of approximately 2.5 m/s, whereas the median of the ERA-Interim, CFSR and CFS data is approximately 3 m/s (Figure 42).

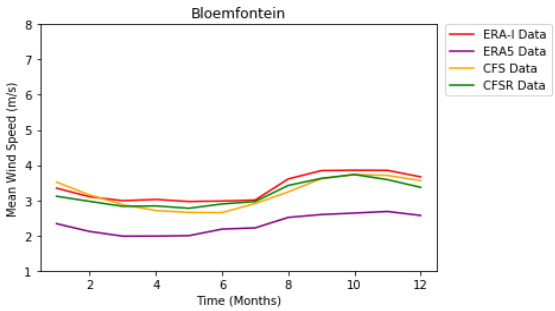


Figure 41: Bloemfontein time series

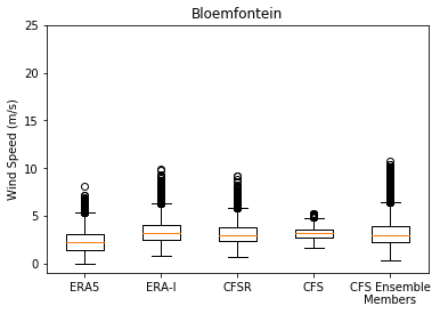


Figure 42: Bloemfontein box plots

The wind direction has little variation over the course of a year in Bloemfontein. There is a dominant north to north-north-easterly wind throughout the year and the ERA5 wind roses

indicate that alongside the northerly winds there are weaker west-south-westerly to north-west-northerly winds during September to November (Figure 43). In September to December, the CFS and ERA-Interim wind roses indicate an increase in wind speed, which is consistent with the time series plot (Figure 41).

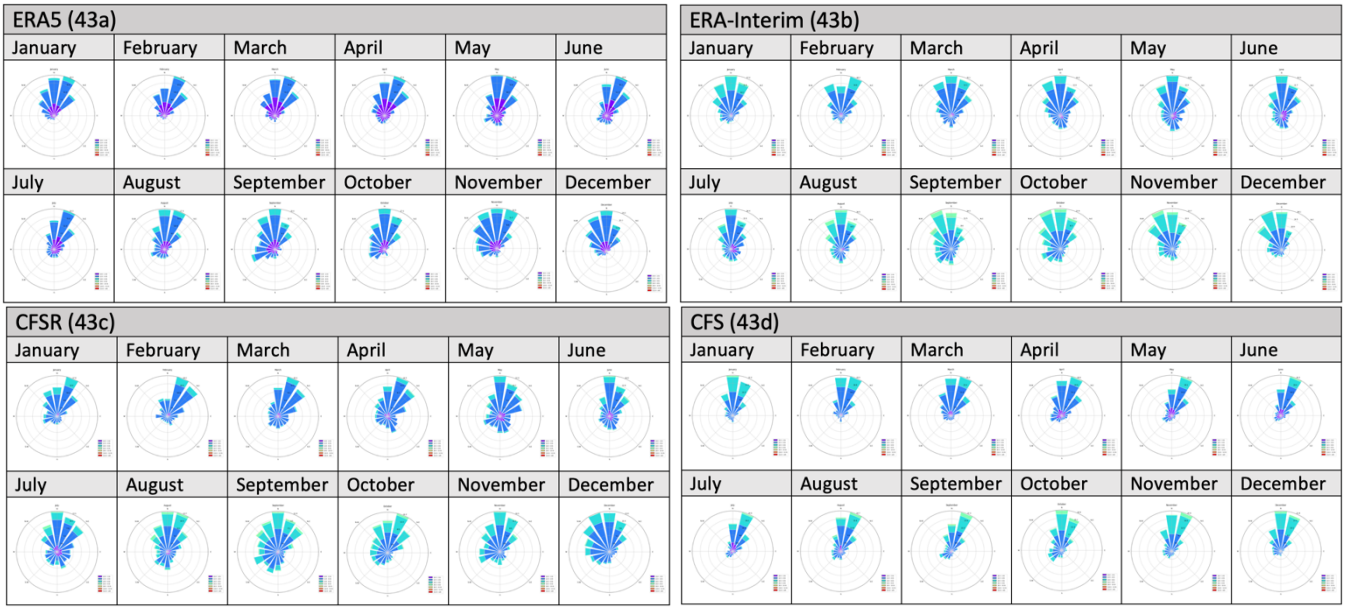


Figure 43: Bloemfontein wind roses

7.3.11. Polokwane

Polokwane is situated in Limpopo which is the northernmost province in South Africa. The wind speeds over Polokwane follow a similar pattern to Johannesburg (Figure 38) and Bloemfontein (Figure 41), with a peak in wind speeds in October and a decrease in wind speeds between April and June (Figure 44). The wind speeds range between 1.5 m/s and 3 m/s, depending on the model. The ERA5 wind speeds are the lowest and the ERA-Interim wind speeds are the highest (Figure 45).

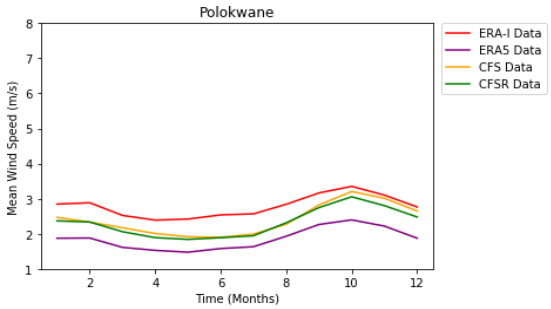


Figure 44: Polokwane time series

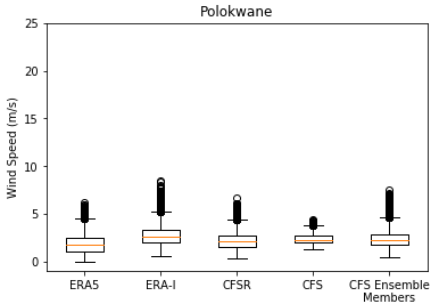


Figure 45: Polokwane box plots

There is a dominant east-north-easterly wind between September and March over Polokwane. The ERA5 wind roses for April to August, indicate that the wind direction becomes more widespread, with a wind direction that ranges from north-easterly winds, northerly

winds and south-easterly winds (Figure 46a); however the ERA-Interim, CFSR and CFS wind roses show a dominant east to east-north-easterly wind during the majority of the austral winter (Figure 46b to 46d). During June, the wind direction varies slightly between the four datasets (Figure 46).

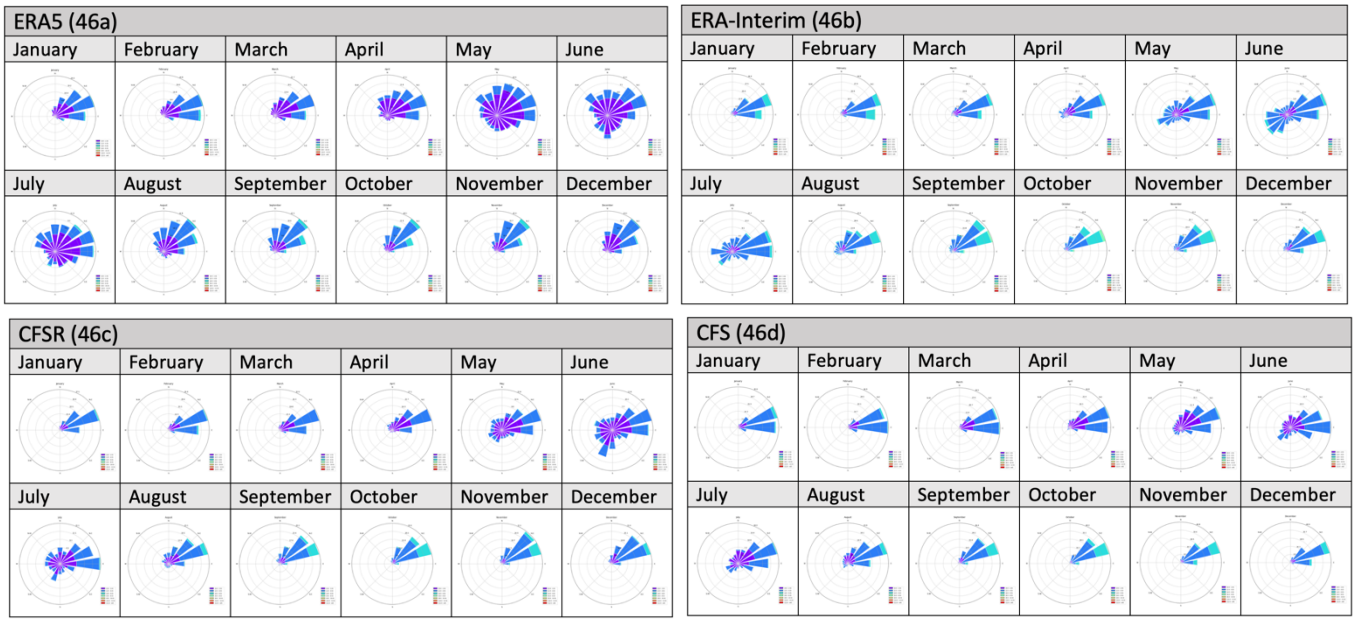


Figure 46: Polokwane wind roses

7.3.12. Upington

Upington is an inland region, situated towards the north-west of the country. The wind speeds experienced over this area ranges from 2.5 m/s to 4.5 m/s, depending on the time of year and dataset, which is higher than the previous inland areas that have been studied (Figure 47 and 48). Minimum wind speeds in Upington are experienced during March, and the wind speeds increase from April. Based on the ERA5 reanalysis dataset, the difference between the maximum and minimum wind speeds is approximately 1.5 m/s to 2 m/s; however, based on the CFS ensemble mean data the difference is approximately 1 m/s (Figure 47).

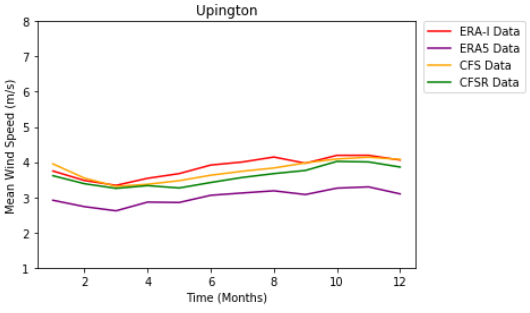


Figure 47: Upington time series

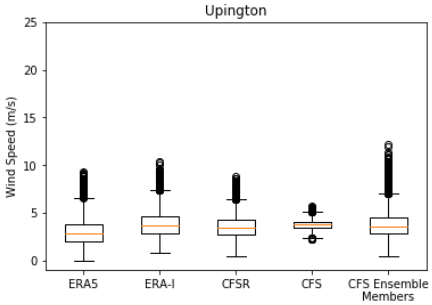


Figure 48: Upington box plots

Between October to February, there is a dominant south-westerly wind presented by the ERA5 and CFS wind roses (Figure 49a and 49d). The ERA5, ERA-Interim and CFSR datasets (Figure 49a to 49c), indicate that in February, a northerly wind develops, whereas the CFS

model (Figure 49d), shows that the northerly winds develop in March. Northerly and north-north-easterly winds are dominant from March to August. The wind roses show that in September, there is a south-westerly and northerly wind present (Figure 49).

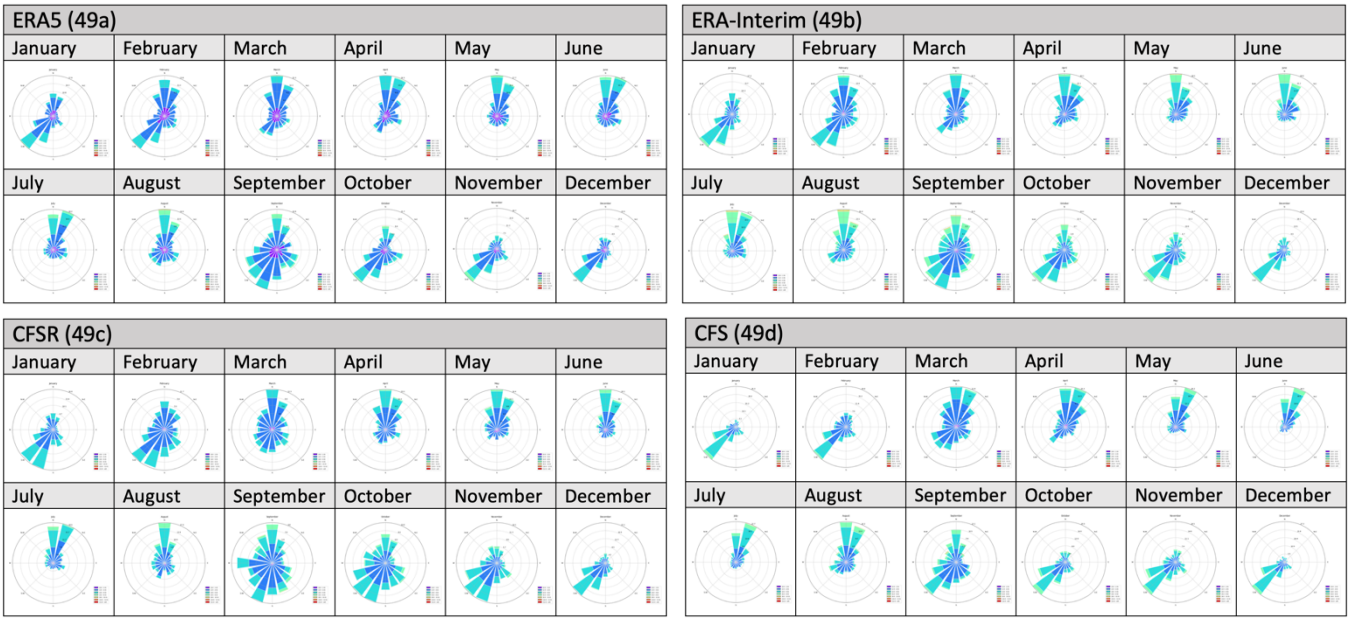


Figure 49: Upington wind roses

7.3.13. Butterworth

Butterworth is an inland region situated near the east coast of South Africa. The three reanalysis datasets (ERA5, ERA-Interim and CFSR) indicate that wind speeds increase in June and decrease between March, April and August. The CFS hindcast data indicates that the wind speeds decrease in April, but increase in May. The ERA5 wind speeds are the lowest, and the CFS wind speeds are the highest (Figure 50). The box plots indicate that the spread of data between the first and third quartile is low in all datasets, including the CFS ensemble members (Figure 51).

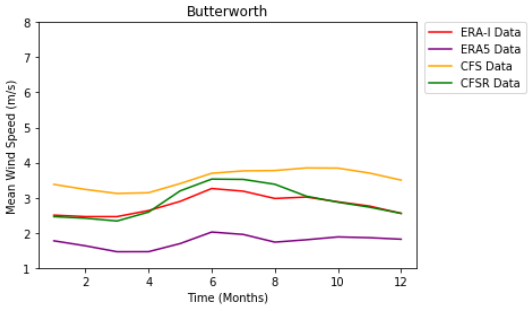


Figure 50: Butterworth time series

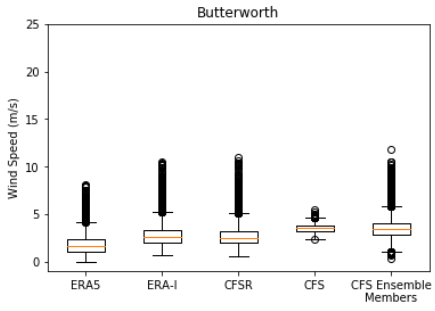


Figure 51: Butterworth box plots

The wind roses below indicate that between April and August there is a dominant north-westerly wind that has been depicted by the four datasets. During September to March, however, the wind rose results differ slightly between the datasets (Figure 52). The wind

directions presented by ERA-Interim (Figure 52b) and CFSR (Figure 52c) during September to March are more widespread, compared to CFS (Figure 52d) and ERA5 (Figure 52a).

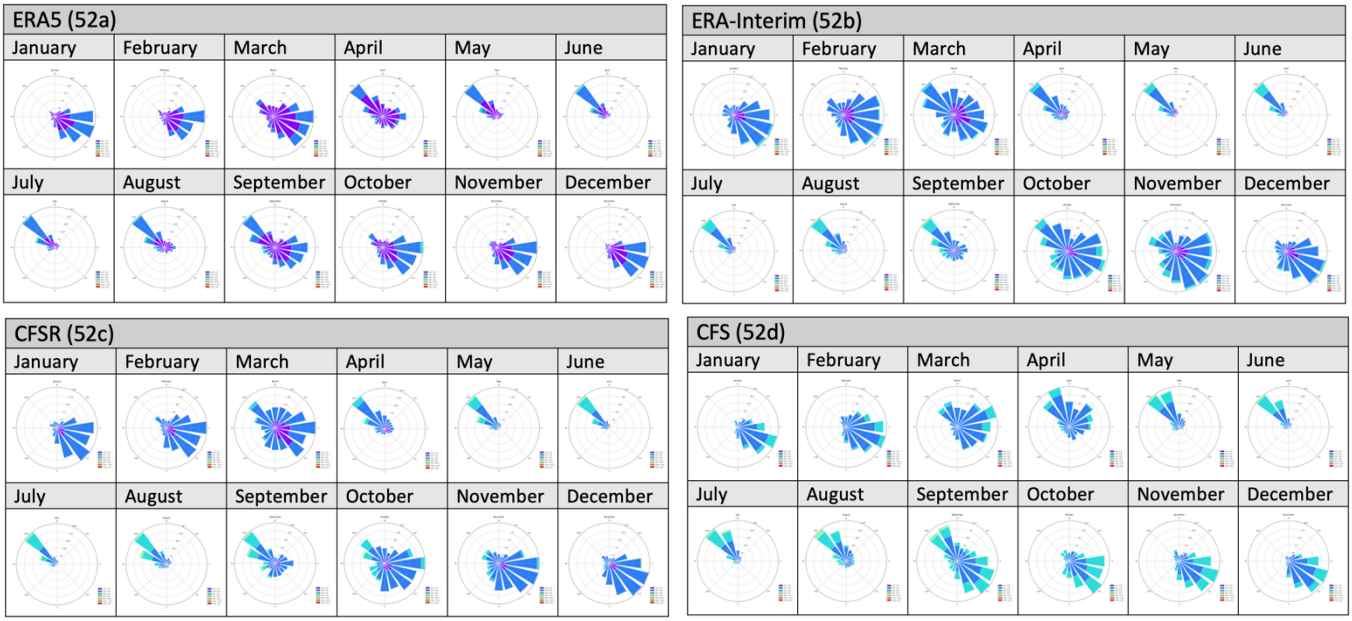


Figure 52: Butterworth wind roses

7.3.14. Brandvlei

Brandvlei is located in the Northern Cape. Peak winds over Brandvlei are experienced from October to January. The ERA5 and ERA-Interim models show that wind speeds increase in June and July, whereas the CFSR and CFS models show that there is a decrease in wind speeds. The four models are consistent in indicating there the minimum wind speeds are experienced in March (Figure 53). The average wind speeds experienced over Brandvlei are between 3 m/s to 5 m/s (Figure 54).

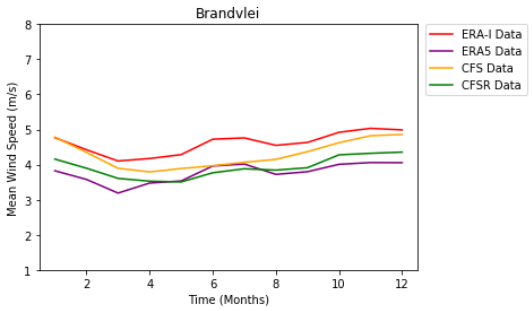


Figure 53: Brandvlei time series

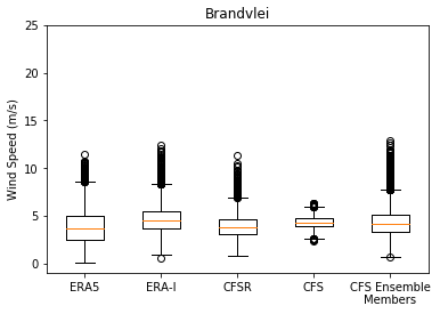


Figure 54: Brandvlei box plots

The wind direction from September to March over Brandvlei is predominantly a south-westerly and west-south-westerly wind, however in April the wind roses indicate that a north-north-easterly wind develops. During May to August the north-north-easterly wind is the dominant wind direction. These results are consistent across the four datasets, except during

May and June, where the wind directions presented by the reanalysis datasets (Figure 55a to 55c) are more widespread than the CFS hindcast data (Figure 55d).

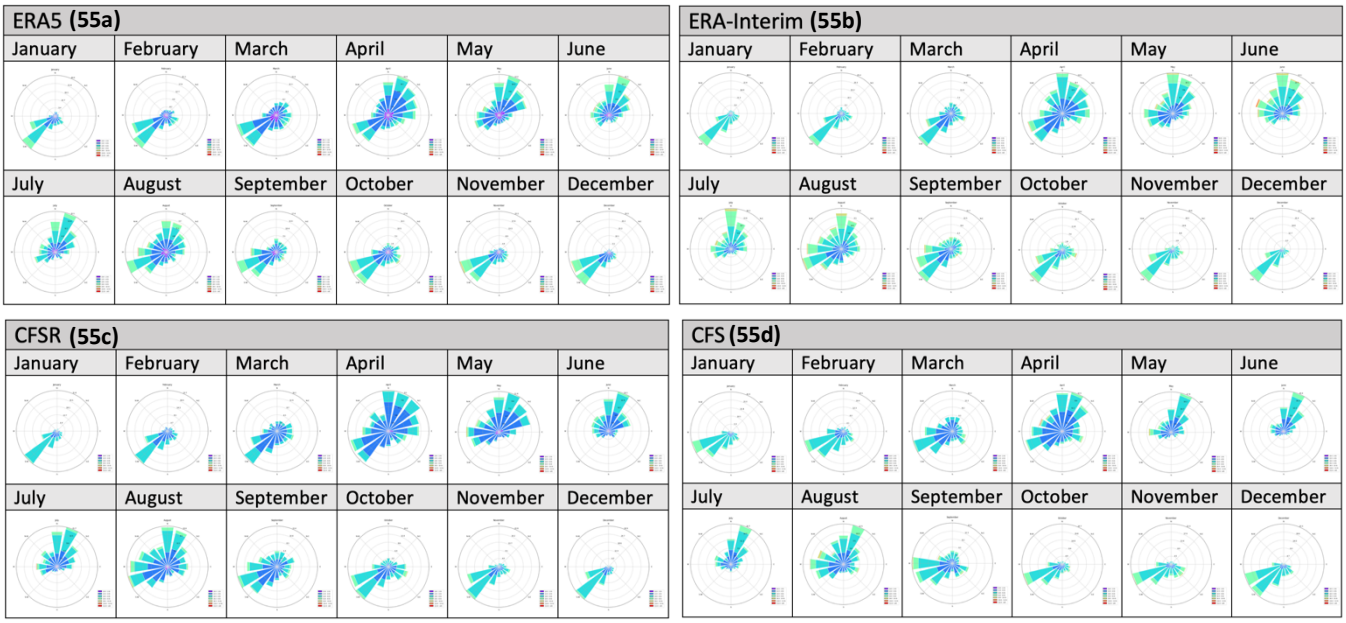


Figure 55: Brandvlei wind roses

7.3.15. Hanover

Hanover is a town in the Little Karoo, situated towards the south of the Northern Cape. The ERA-Interim, CFSR and CFS models show that the wind speeds increase during austral summer months and decrease between May and July. The ERA5 model, however, portrays a different wind pattern. The ERA5 wind data indicates that wind speeds increase (decrease) during austral winter (summer) (Figure 56). The spread of data is consistent between the three reanalysis datasets and the CFS ensemble members and that the spread of data is lower in the CFS ensemble mean dataset (Figure 57).

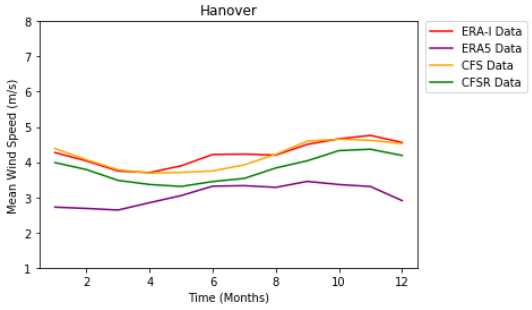


Figure 56: Hanover time series

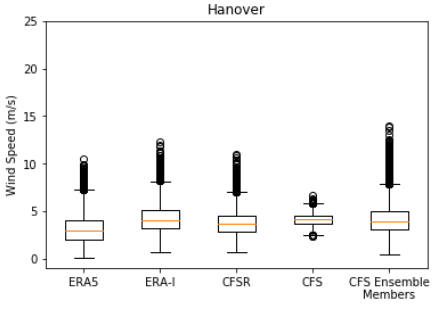


Figure 57: Hanover box plots

The reanalysis datasets (ERA5, ERA-Interim and CFSR) indicate that the wind direction between November and February is predominantly a south-easterly wind accompanied by a weaker north-westerly to north-north-westerly winds (Figure 58a to 59c). The CFS hindcast data however, indicates that that a south-south-easterly wind is the primary wind direction

during the same months (Figure 58d). In March the south-easterly winds become weaker and between April to October, the north-westerly winds dominate the area; these results are consistent between the four datasets (Figure 58).

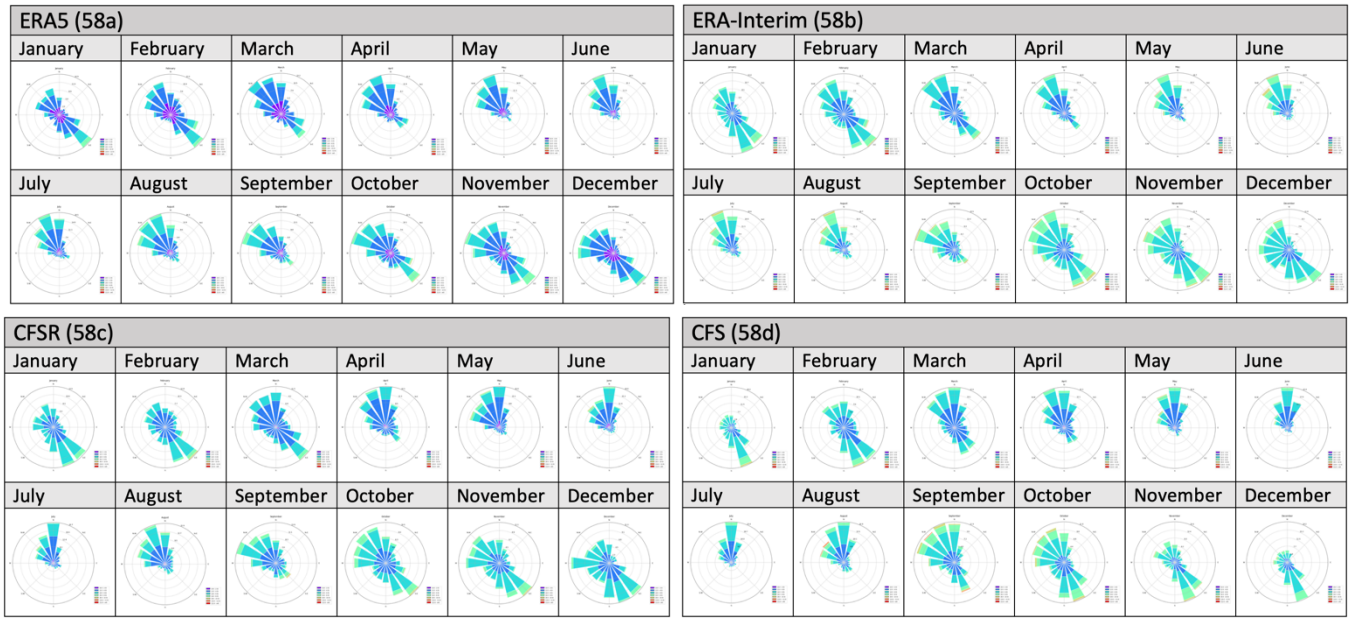


Figure 58: Hanover wind roses

7.3.16. Beaufort West

Beaufort West is an inland region situated in the Western Cape. Wind speeds increase between June and July, and minimum wind speeds are experienced in March and August (Figure 59). The lowest wind speeds are represented by ERA5. The difference between the first and third quartile for the CFS ensemble mean wind speeds is smaller than the difference between the first and third quartile for the reanalysis datasets and the CFS ensemble members and the three reanalyses, which is consistent with the findings that the ensemble mean results in the smoothing of the wind speed data (Figure 59 and 60).

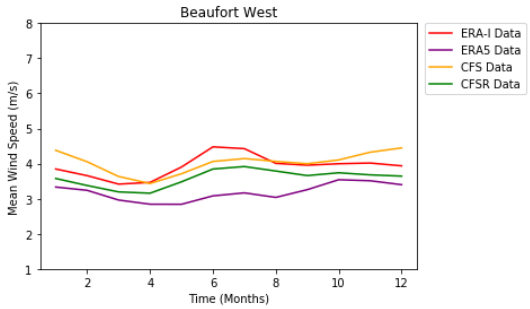


Figure 59: Beaufort West time series

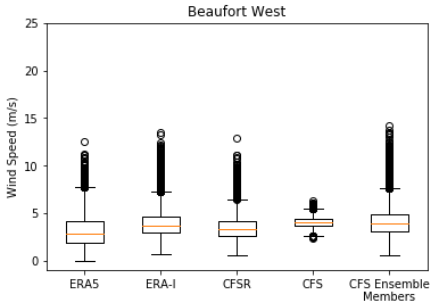


Figure 60: Beaufort West box plots

The wind roses for each dataset show differing wind directions. The ERA-Interim and CFSR wind roses are consistent with one another (Figure 61b and 61c), and the ERA5 wind roses are the outlier (Figure 61a). The strong easterly winds that can be seen throughout the year in the ERA5 wind roses cannot be seen in the remaining three wind rose (Figure 61b to 61d). In September, the CFS wind roses (Figure 61d) differ to the ERA-Interim and CFSR wind roses

(Figure 61b and 61c), as there are predominant south-south-easterly wind directions which do not appear in the other two wind rose diagrams. The majority of the CFS wind roses have a decreased spread of wind directions despite the use of an ensemble member.

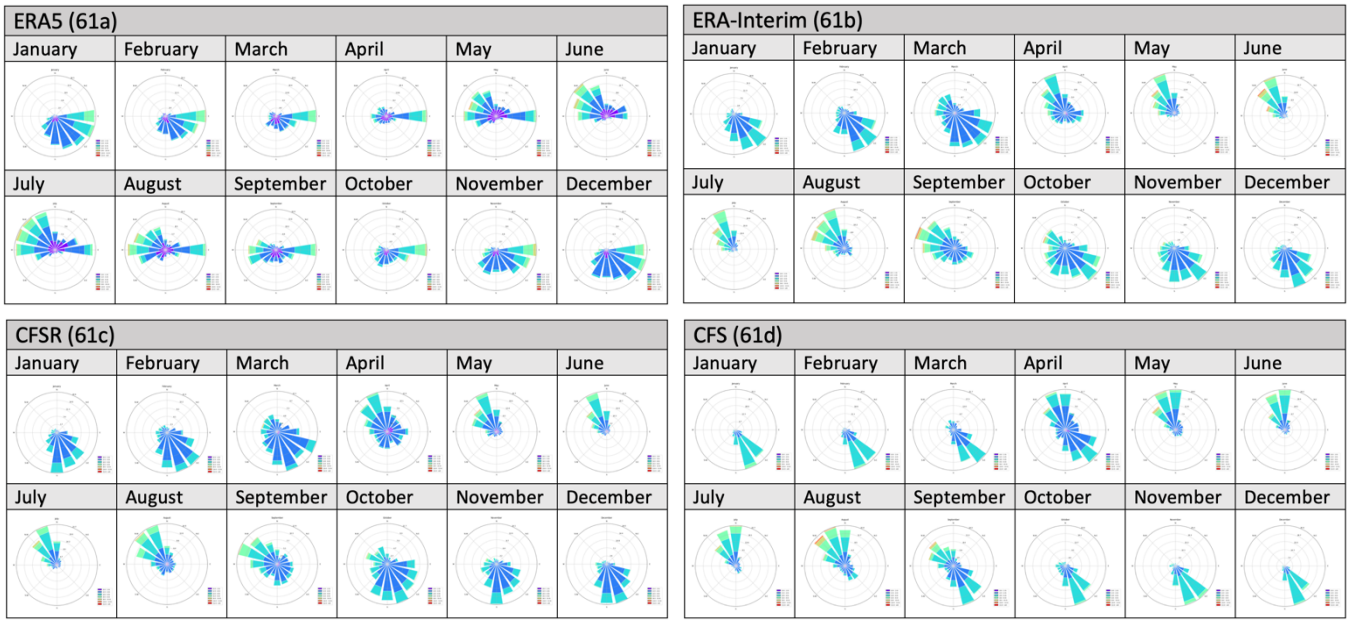


Figure 61: Beaufort West wind roses

The peak wind speeds in August and September over Pietermaritzburg (Figure 35), Johannesburg (Figure 38), Bloemfontein (Figure 41), Polokwane (Figure 44), Upington (Figure 47), Butterworth (Figure 52), and Hanover (Figure 56) can be explained by the mid-latitude cyclones which stretch furthest north during these months. The mid-latitude cyclones result in a high temperature gradient and pressure gradient, which results in higher wind speeds. The timeseries and box plot figures also indicate that the coastal regions have a larger variation in wind speed, whereas the distribution of wind speeds are lower inland.

The results in Section 7.3. presented multiple differences, in wind speed and direction, between the three reanalysis datasets. These differences could have a potential impact on the forecast verification given that observed data has not been used in this study and therefore reanalysis data had to be used as proxies for observations. In order to further analyse the difference between the reanalysis datasets, a statistical analysis has been performed on the datasets in Section 7.4.

7.4. Comparing the difference between reanalysis datasets

Based on Section 7.3. of the results, it can be seen that for the majority (14 out of 16) of the study sites, ERA5 wind data has lower wind speeds than the ERA-Interim wind data and the CFSR wind data over South Africa. There is an average of 1 m/s to 1.5 m/s difference between the ERA-Interim and ERA5 reanalysis datasets over land, and a smaller difference over the ocean (Figure 62a). The ERA5 wind speeds are lower than the CFSR wind speeds, specifically

towards the south and south-western parts of the country (Figure 62b). Over large portions of South Africa, the difference between ERA-Interim wind speeds and CFSR wind speeds are low. Towards the north of the country, over Limpopo and parts of the Northern Cape and the Western Cape, the ERA-Interim wind speeds are higher than the CFSR wind speeds (blue grid points), however along the east coast and over the surrounding oceans, the CFSR wind speeds are higher (Figure 62c).

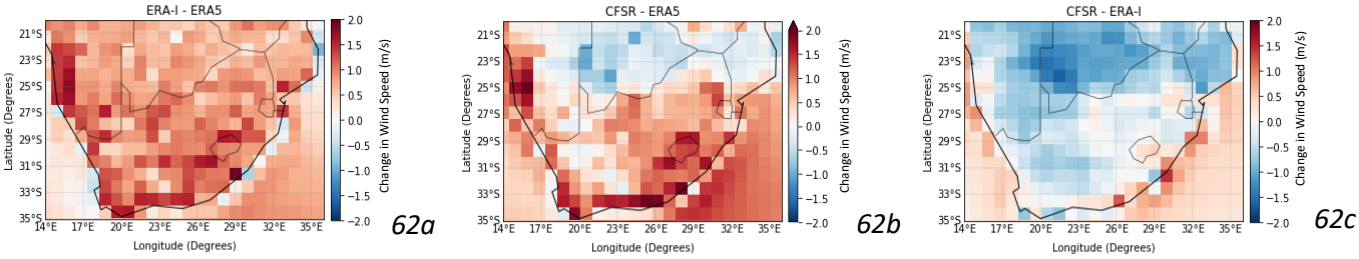


Figure 62: Figure 62 presents the difference in reanalysis datasets, where (a) is between ERA-Interim and ERA5; (b) is between CFSR and ERA5; and (c) is between CFSR and ERA-Interim

7.4.1. Statistical Significance Tests

In order to further analyse the difference between the three reanalysis datasets, statistical significance tests have been performed on the data. This process provides an overview of the spread within the gridded "observed" wind climate and a context in which the forecast data is discussed.

7.4.1.1. T-Test

This section of the results presents two sample T-Tests performed on the reanalysis datasets. The T-Test aims at calculating the difference between the means of two datasets. The output of the T-Test that has been used in this study is the *p*-value which represents the probability of the null hypothesis being true. The null hypothesis for this study is that there is no difference between the means of two reanalysis datasets, and if the *p*-value is smaller than 0.01 (the alpha value), then there is a statistically significant difference between the datasets.

The T-Test has been performed on monthly and seasonal wind speed data. With respect to the ERA-Interim and ERA5 datasets, the *p*-value is low over the interior of the country (Figure 63a), indicating that there is a significant difference between the two datasets. This is consistent with the results in Section 7.3. and 7.4., which shows a visible difference between the two datasets. The *p*-value between ERA5 wind speed data and CFSR monthly wind speed data is higher over the northern parts of the country (regions of Limpopo, the Northern Cape and the Western Cape) (Figure 63b), compared to *p*-value between the ERA datasets (Figure 63a). There are high *p*-values over the interior of the country, and along the west coast between the ERA-Interim and CFSR wind speed data (Figure 63c). There are low *p*-values

towards the north of the study area (Limpopo), where there is a large portion of dark blue grid points. There are low *p*-values over the Indian Ocean and higher *p*-values over the Atlantic ocean (Figure 63a and 63b); however, the T-Test analysis between ERA-Interim and CFSR indicates that both the oceans surrounding South Africa have a high *p*-value. Possible reasons for the differing *p*-values include the fact that the representation of the land surface may be different in each reanalysis dataset and given that 10m wind speeds are being studied, the topography influences the wind speed values. In addition, the 10m wind speeds are a diagnostic variable, rather than a prognostic variable, which may influence the results. Finally, the three reanalysis wind speed datasets have different resolutions, which may explain the differences in the *p*-value (Figure 63).

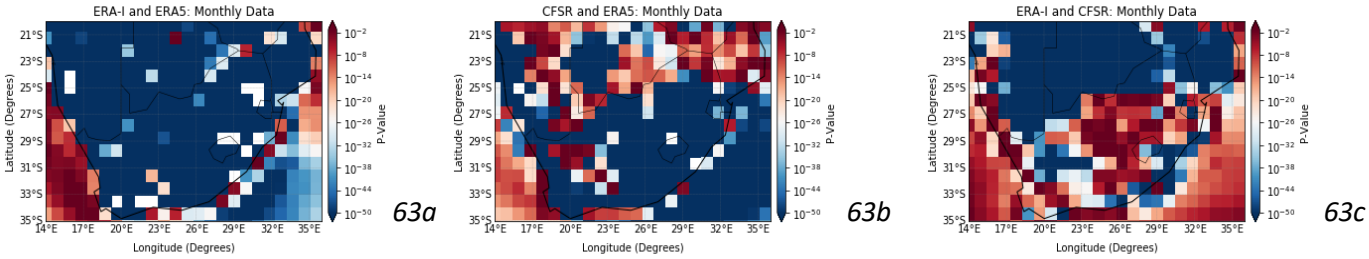


Figure 63: Plot of the t-test p-value between monthly reanalysis wind data where (a) is between ERA-Interim and ERA5; (b) is between CFSR and ERA5; and (c) is between CFSR and ERA-Interim

The change in *p*-values using seasonal data (Figure 64), compared to monthly data (Figure 63) is minimal over South Africa, except for over the northern borders of the country where the p-value increased slightly (Figure 64a). The T-Test between the CFSR and ERA5 datasets (Figure 64b) and between ERA-Interim and CFSR (Figure 64c), using seasonal data compared to monthly data has resulted in higher *p*-values over certain regions of South Africa.

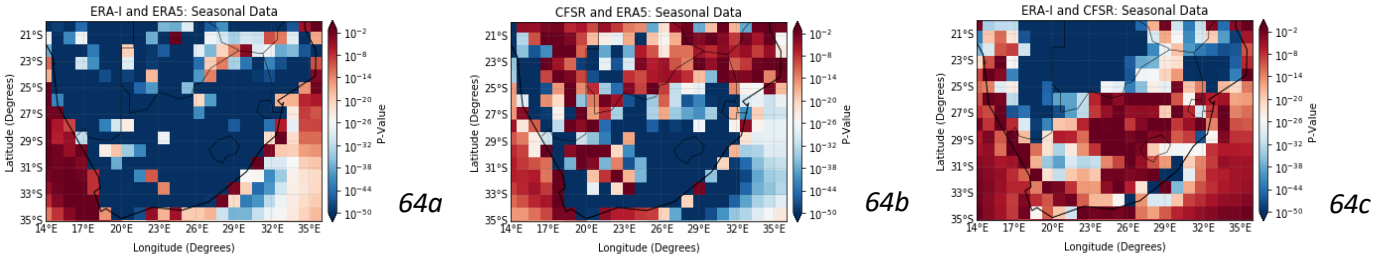
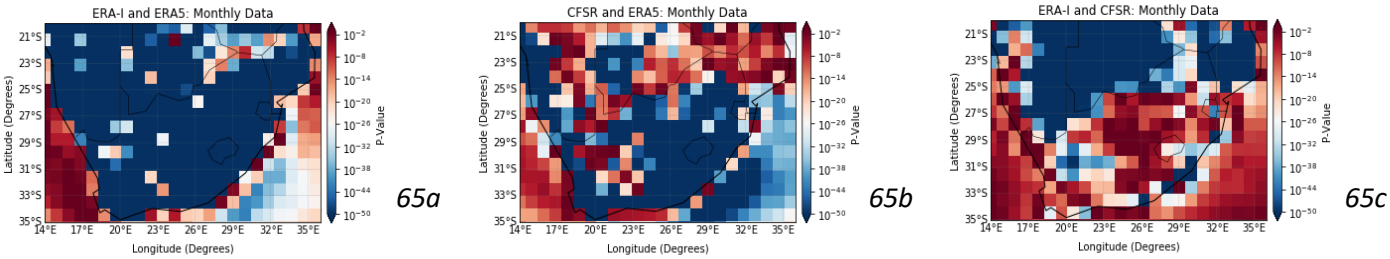


Figure 64: Plot of the t-test p-value between seasonal reanalysis wind data where (a) is between ERA-Interim and ERA5; (b) is between CFSR and ERA5; and (c) is between CFSR and ERA-Interim

7.4.1.2. Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov Test (KS-Test), is another significance test which has been performed on the three reanalysis datasets in order to identify if the difference between the datasets is significant. This section presents the results from the KS-Test, which is also a *p*-value.

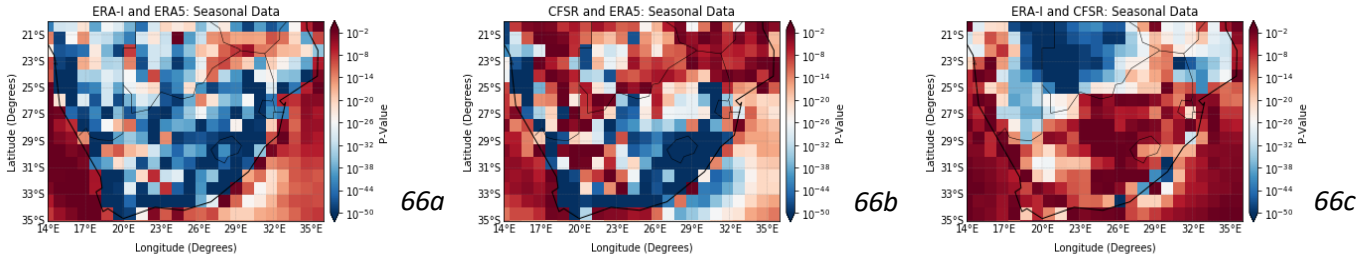
The results presented below (Figure 65) are consistent with the results in Section 1.4.2.1., which indicate that the difference between the ERA5 and ERA-Interim monthly wind speed datasets is significant, as the *p*-value is below 0.01 over the country. The difference between the ERA5 and CFSR dataset decreases over the northern parts of the study area, and over the west interior regions of South Africa, where the colour of the grid points indicate a *p*-value of approximately 0.01, and in some areas greater than 0.01 (Figure 64b and 65b). Over the interior of the country, and over the oceans, where the *p*-value is greater than 0.01, there are no statistically significant differences between the ERA-Interim and CFSR monthly wind speed data. There is a significant difference between the datasets over the north and north-western parts of the study area, which is represented by low *p*-values (Figure 65c). Similar to the *t*-test, the possible reasons as to why the *p*-values differ depending on the reanalysis datasets used, is likely due to the way in which the topography has been represented on each dataset. Given that the 10m wind speeds are a derived variable, rather than a prognostic variable also enhances the differences between the models and likely impacts the results.



65a *65b* *65c*

Figure 65: Plot of the ks-test p-value between monthly reanalysis wind data where (a) is between ERA-Interim and ERA5; (b) is between CFSR and ERA5; and (c) is between CFSR and ERA-Interim

The use of seasonal wind data has resulted in the *p*-value increasing and therefore this indicates that the statistically significant differences between the three reanalyses changes, based on the length of the time periods of the data (Figure 66).



66a *66b* *66c*

Figure 66: Plot of the ks-test p-value between seasonal reanalysis wind data where (a) is between ERA-Interim and ERA5; (b) is between CFSR and ERA5; and (c) is between CFSR and ERA-Interim

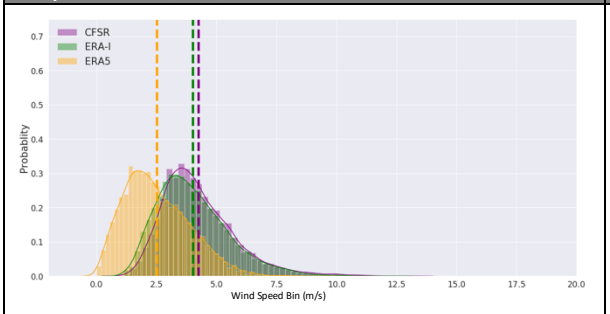
The *p*-values are generally higher when seasonal wind speed values have been used in comparison to monthly wind speed values for both the t-test and the ks-test. This is likely due to the fact that at longer time scales, the variability of the wind speeds decreases. The seasonal wind speed data places more emphasis on general patterns and cancels some of the noise associated with wind speeds at shorter timescales.

7.4.2. Probability Density Functions

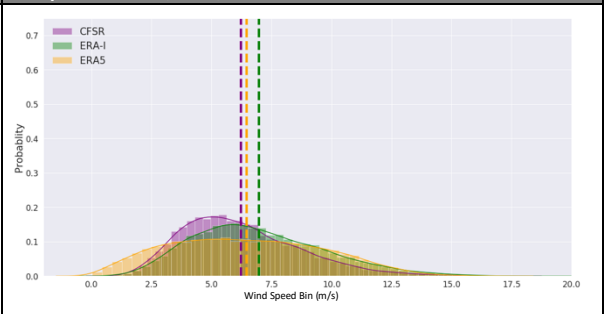
Section 7.4.2. presents Probability Density Functions (PDFs) for the 16 study sites. The PDFs aim to examine the wind distribution at the 16 sites for each reanalysis wind dataset in order to understand the magnitude bias. Wind speed is a continuous variable and therefore specific values will have a probability of zero, therefore the wind speeds are grouped together with a range of wind speed values and sectioned into different bins (Figure 67).

The results from the Probability Density Functions (PDFs) are consistent with the time series figures in Section 7.3. and Section 7.4.1. For the majority of the areas, the ERA5 wind speeds are lower than the CFSR and ERA-Interim wind speeds. The spread of the ERA5 data is also different to the two other reanalysis datasets. The ERA5 wind speed curve is flatter and there is a wider spread of wind speed values, which results in the probability of the mean value occurring being lowered. The most frequent wind speeds are between 3 m/s and 5 m/s over the coastal regions, and 2 m/s and 4 m/s over the inland regions. The majority of the PDFs have longer right tails, than left tails; this indicates that there are a greater amount of high wind speed values that are extreme values, compared to low wind speed values (Figure 67)

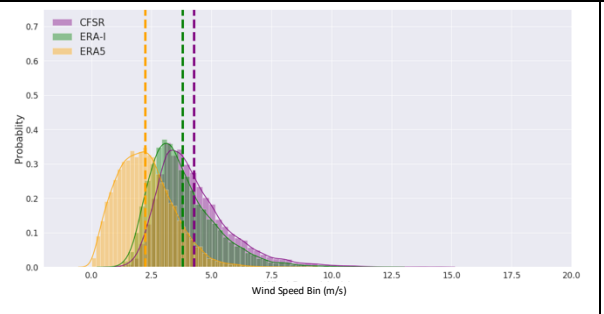
Cape Town *(67a)*



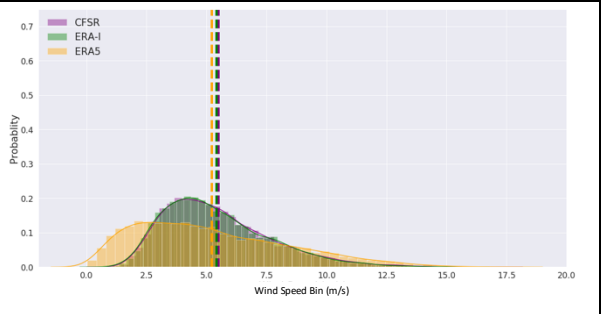
Cape Point *(67b)*



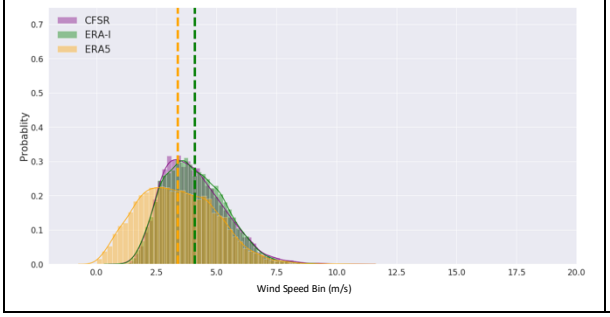
Port Elizabeth *(67c)*



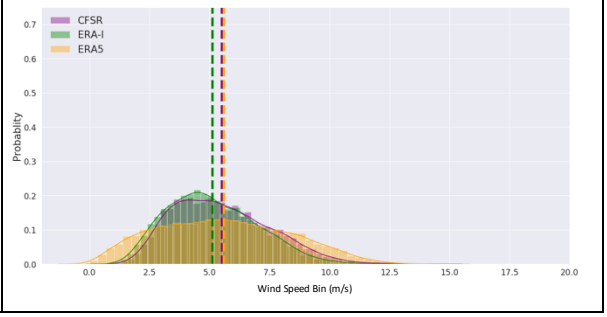
Mossel Bay *(67d)*



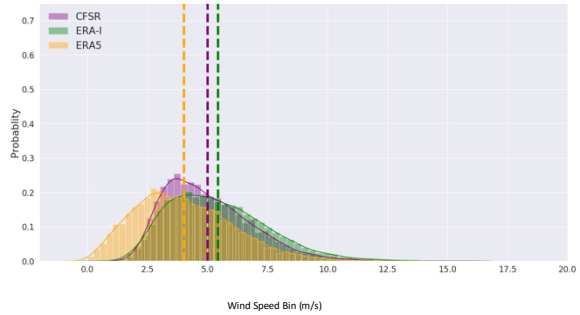
Port Nolloth *(67e)*



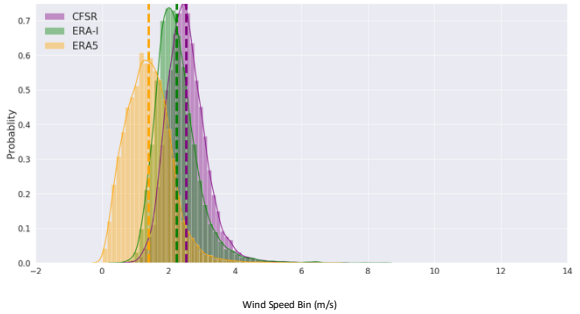
Saldanha *(67f)*



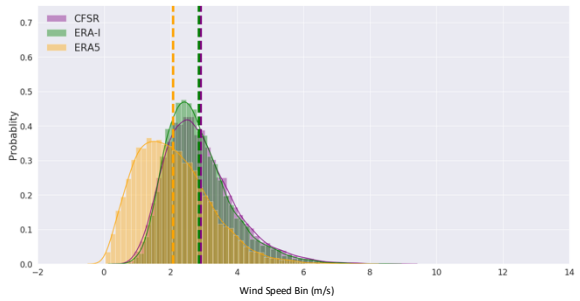
Mabibi *(67g)*



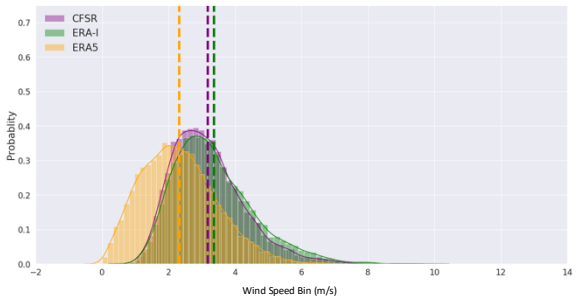
Pietermaritzburg *(67h)*



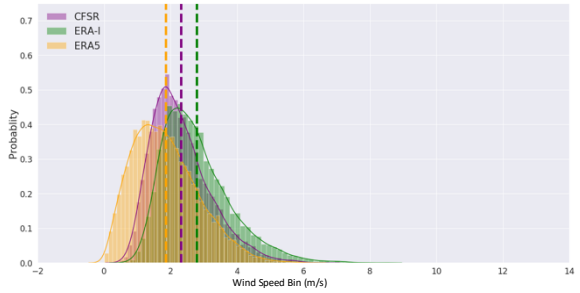
Johannesburg *(67i)*



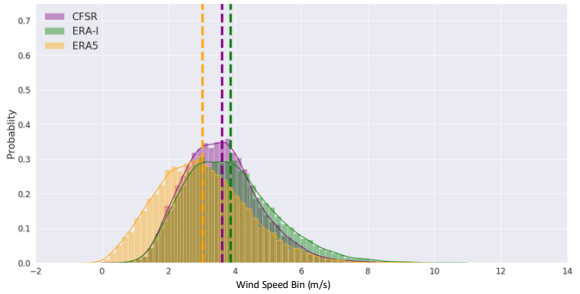
Bloemfontein *(67j)*



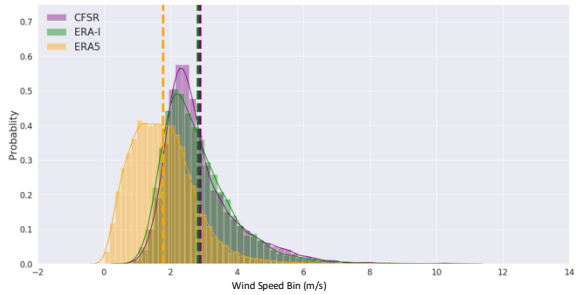
Polokwane *(67k)*



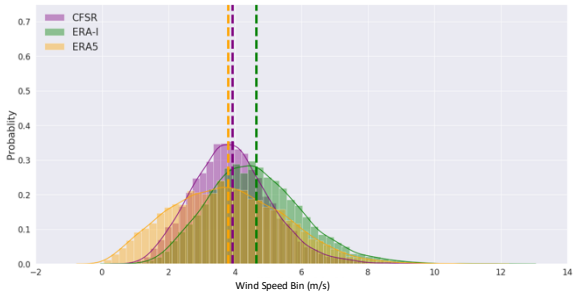
Upington *(67l)*



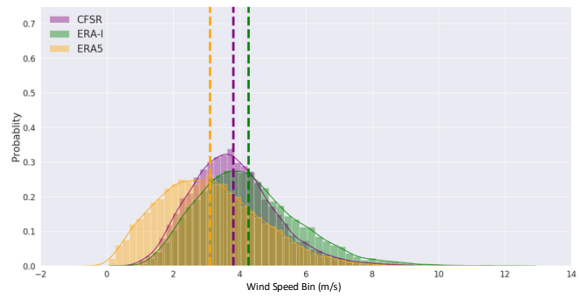
Butterworth *(67m)*



Brandvlei *(67n)*



Hanover *(67o)*



Beaufort West *(67p)*

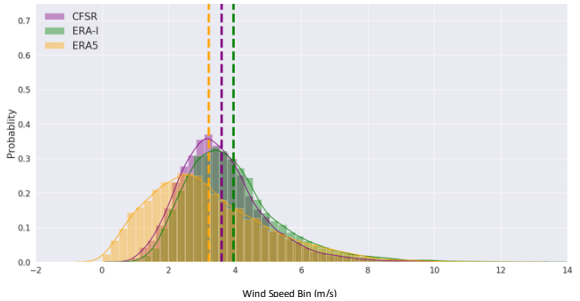


Figure 67: Figure 67a to 67p present PDFs for each of the 16 sites. The coastal regions have an x-axis range from -2 to 20, and the inland regions have an x-axis range from -2 to 14, where the x-axis represents the wind speed bin and the y-axis represents the probability of the wind speeds being experienced at each site

7.4.3. Pearson Correlation Coefficient between the Reanalysis Datasets

The figures below show the Pearson's Correlation Coefficient (*r*) between the three reanalysis datasets. Daily, monthly and seasonal wind data have been used for each dataset. To compute the plots, the whole timeseries was used and re-grouped into the three separate timescales. The analysis period is from 1982 to 2018 (ERA5 and ERA-Interim) and 1982 to 2010 for CFSR. The *r* colour scale ranges from "0" to "1"; however, the Pearson Correlation Coefficient usually ranges from "-1" to "1". The range has been constrained for this part of the study in order to depict the results more clearly. Wind speed correlations are weaker at shorter time scales (daily) compared to longer time periods (monthly and seasonal). This can be explained by wind speeds having a more erratic nature at shorter timescales, making it more difficult to predict. Forecasting wind speeds over a longer time period will allow for patterns to emerge and therefore increase the correlations.

The correlation between ERA-Interim and ERA5 wind data is strongest when seasonal wind data has been used, and weakest using the daily wind data (Figure 68). The areas with the highest *r*-values (between 0.8 and 1), are over the Atlantic Ocean and over the northern parts of the country, and therefore this indicates that there is a strong positive correlation between ERA5 and ERA-Interim reanalysis datasets over these areas. The lowest areas of correlation are along the east coast of South Africa and towards the south-east of the country, over parts of Western Cape and Eastern Cape. Despite the *r*-value decreasing in these areas, there is still a positive correlation between the two datasets.

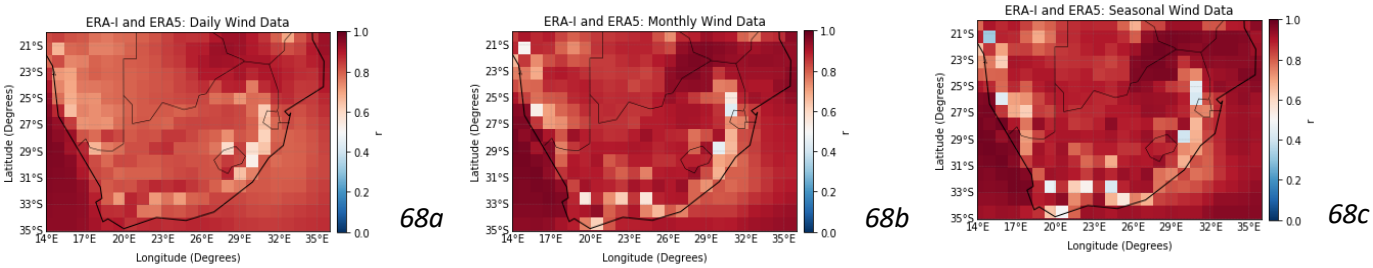


Figure 68: Plot of the Pearson Correlation Coefficient between the ERA-Interim and ERA5 reanalysis wind datasets where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

The correlation between the CFSR reanalysis wind data and the ERA5 reanalysis wind data (Figure 69) is similar to the correlation patterns between the ERA5 and ERA-Interim wind speed datasets (Figure 68). The correlation coefficient increases over the interior of the country, from the daily data to the seasonal data, indicating that the positive correlation is

stronger when longer time periods are used to perform the calculation. There is a weaker correlation between the two reanalysis datasets over the eastern and southern coastal regions, where the *r*-value is between 0.3 and 0.5 (Figure 69c).

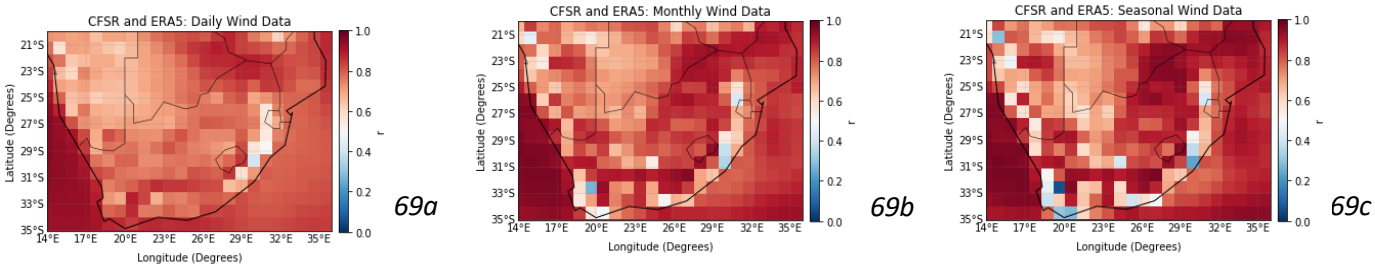


Figure 69: Plot of the Pearson Correlation Coefficient between the CFSR and ERA5 reanalysis wind datasets where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

The Pearson Correlation Coefficient between the CFSR reanalysis data and the ERA-Interim reanalysis data (Figure 70) is higher than when the Pearson Correlation Coefficient calculation was performed using the ERA5 data (Figure 68 and 69). These results are consistent with Section 7.4., which showed the statistical difference between ERA5 and the two other reanalysis datasets (Figure 63 to 66). The *r*-value over the whole study area is between 0.7 and 1, indicating that there is a strong positive correlation between these two datasets. There are lower *r*-values over regions in Kwa-Zulu Natal, along the east coast of South Africa, which is consistent with the results the previous results (Figures 69).

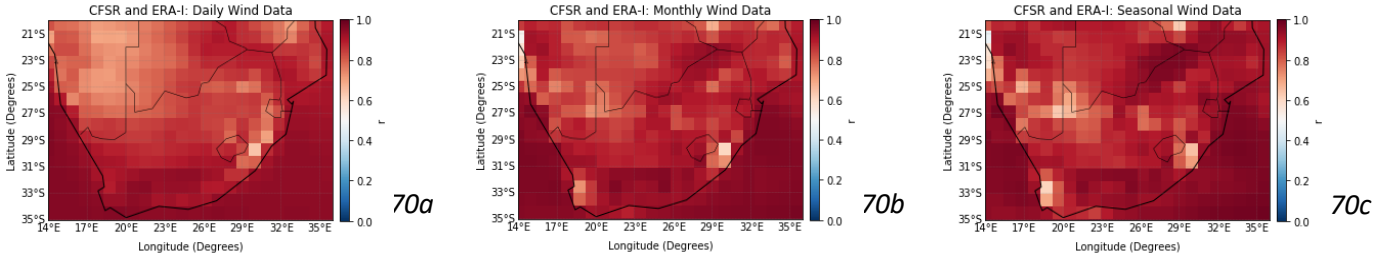


Figure 70: Plot of the Pearson Correlation Coefficient between the CFSR and ERA-Interim reanalysis wind datasets where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

The regions that consistently present weak correlations could be due to the way in which the topography is represented by the models. Given that the weakest correlations occur when ERA5 is being used (Figure 68 and 69), it is likely that the way in which the ERA5 model represents the land surface differs to the way in which ERA-Interim and CFSR represent the land surface, most notably over the eastern and south-eastern regions of South Africa.

7.5. Hindcast Data (CFS) Compared to Reanalysis Data (ERA5, ERA-Interim and CFSR)

7.5.1. Pearson's Correlation Coefficient Results

The Pearson Correlation Coefficient (*r*) between the three reanalysis datasets (CFSR, ERA5 and ERA-Interim) and the hindcast dataset (CFS) is presented in Section 7.5.1. A *r*-value which is greater (less) than zero, indicates that there is a positive (negative) relationship between the two datasets, and a *r*-value which equals to zero indicates there is no correlation between the variables. The Pearson Correlation Coefficient usually ranges from "-1" to "1". The range has been constrained to "0" to "1", for this part of the study in order to depict the results more clearly.

For daily wind speed correlations (Figure 71a), the *r*-values are low (between 0 and 0.2), and therefore this indicates that daily wind speed correlations between the two datasets are weak. Performing the Pearson Correlation Coefficient calculation using monthly wind data results in increased *r*-values over the whole study area; specifically over the northern interior of South Africa and along the west coast and over the Atlantic Ocean (Figure 71b). The correlation coefficient increased when the seasonal wind data was used, however the *r*-value remained low over areas in the Northern Cape, the Western Cape and the Eastern Cape; and along parts of the east coast (Figure 71c). The difference in correlation values between the CFS and ERA-Interim and CFS and ERA5 are likely due to the way in which the land surface is represented by the two reanalysis wind speed datasets. In addition, the planetary boundary scheme is different for the ERA-Interim dataset and the ERA5 dataset, which may impact the correlations with the CFS hindcast wind speed data.

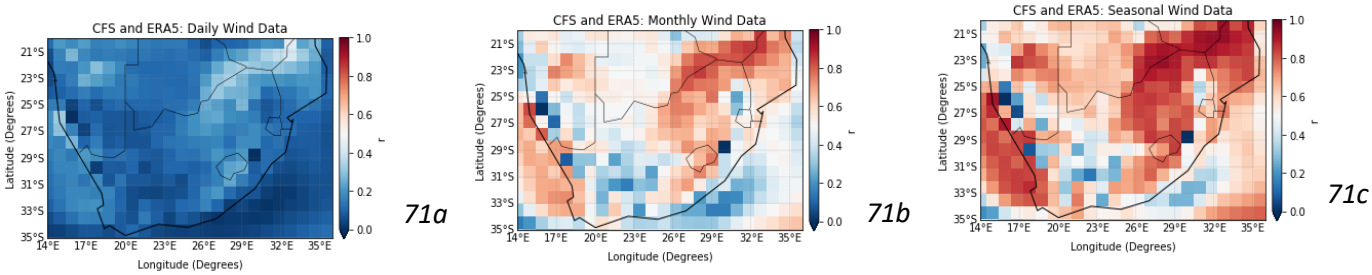


Figure 71: Plot of the Pearson Correlation Coefficient between the CFS hindcast data and ERA5 reanalysis wind data where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

The correlation coefficient values between the CFS and ERA-Interim wind speed data (Figure 72) are higher than the ERA5 and CFS wind speed data; however, the areas of weak correlation are consistent with the correlation figures between CFS and ERA5 (Figure 71). The northern and interior parts of the country have the highest *r*-values and therefore the strongest positive correlation between the two datasets; however, there is still a weak correlation between the forecasted data and the reanalysis data along the west and south coast, and over the regions located towards the south of South Africa (Figure 72b and 72c).

This is consistent with literature that shows higher forecast skill in the northern parts of the country (Landman et al., 2009).

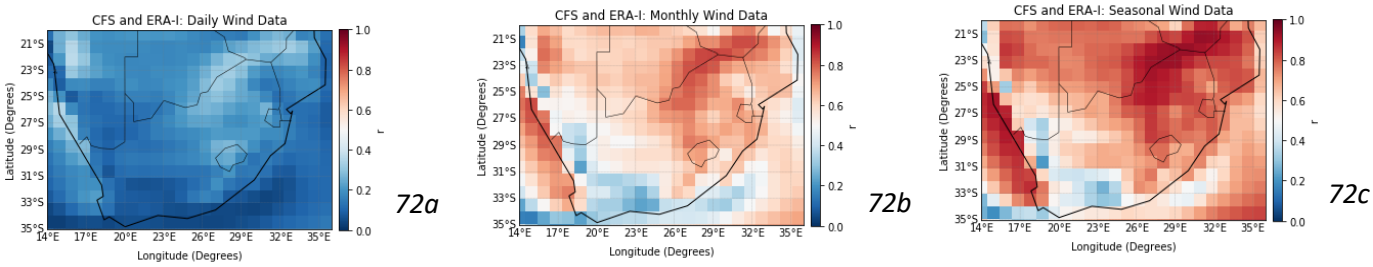


Figure 72: Plot of the Pearson Correlation Coefficient between the CFS hindcast data and ERA-Interim reanalysis wind data where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

The correlation between the CFSR and CFS wind speed data (Figure 73) is the strongest compared to the two ERA reanalyses (Figure 71 and 72). The *r*-value is higher for the daily, monthly and seasonal correlations between CFS and CFSR (Figure 73) compared to the respective results above (Figure 71 and 72). The areas of weak correlation remain consistent throughout all the results in Section 7.5., indicating that there is an interference with the wind speeds on the west to south-west coast and interior of the country.

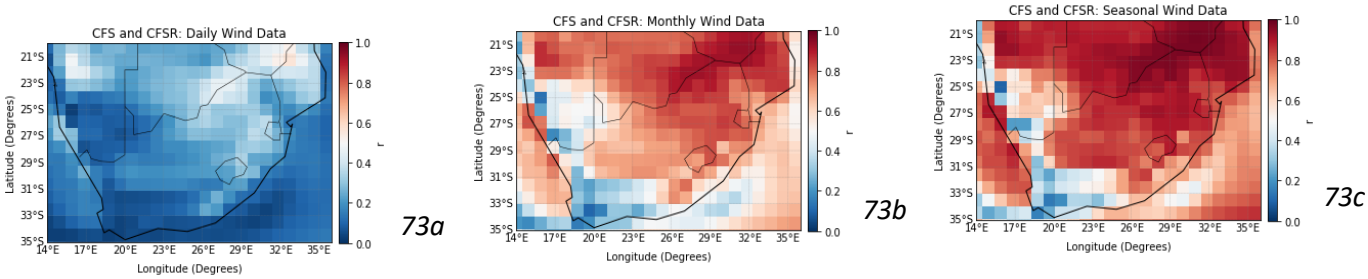


Figure 73: Plot of the Pearson Correlation Coefficient between the CFS hindcast data and CFSR reanalysis wind data where (a) makes use of daily wind speed data; (b) makes use of monthly wind speed data; and (c) makes use of seasonal wind speed data

7.5.1.1. Seasonal Correlations

The results below present the Pearson Correlation Coefficient between the reanalyses wind speed data and CFS wind speed data for the four seasons: DJF, MAM, JJA and SON. The *r* colour scale has been altered in Section 7.5.1.1. to "-1" to "1", which presents the Pearson Correlation Coefficient's range in its entirety.

December – January – February

In DJF, over the interior and southern parts of South Africa, there is a positive correlation between CFS wind speed and the reanalysis wind speed values, with a correlation coefficient of approximately 0.5. Over the northern regions of South Africa (Limpopo and Mpumalanga), there are weak negative and positive correlations in DJF between the reanalyses wind speed

data and the CFS hindcast data, indicating that there is close to no correlation during DJF over the northern parts of the country (Figure 74).

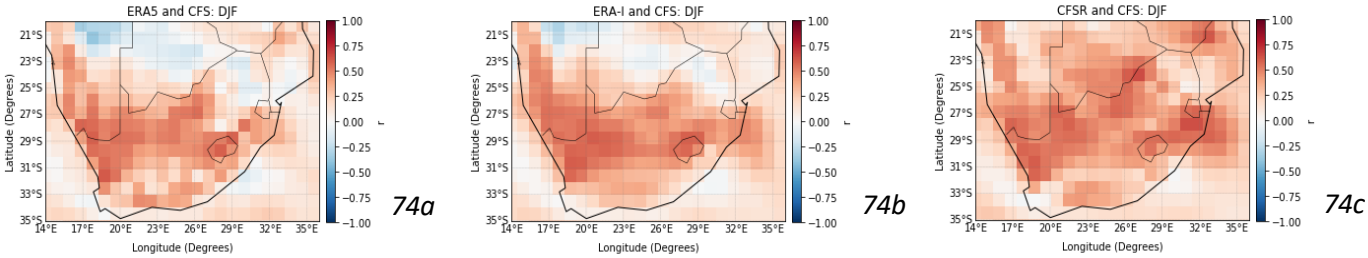


Figure 74: Plot of the Pearson Correlation Coefficient between the hindcast data and the reanalysis datasets for DJF between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

March – April – May

The correlation pattern changes in MAM (Figure 75), compared to DJF (Figure 74), where there is a stronger positive correlation over the northern provinces, and a weak negative correlation over the interior sections of the Northern Cape and the Free State. There are also negative correlations over parts of Kwa-Zulu Natal and Mpumalanga. Along the east and south coast and over the Atlantic Ocean, there are stronger positive correlations; however, the correlations remain weak over the Indian Ocean. The strongest positive correlations can be seen between the CFSR and CFS wind datasets (Figure 75c).

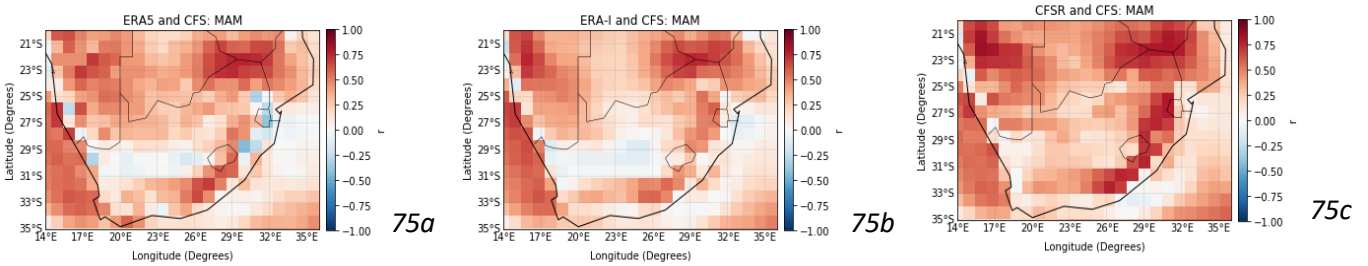


Figure 75: Plot of the Pearson Correlation Coefficient between the hindcast data and the reanalysis datasets for MAM between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

June – July – August

During JJA, over the southern half of South Africa (including the Northern Cape, Western Cape and Eastern Cape), there are weak correlations between the datasets, with an *r*-value between -0.25 and 0.25. Over the majority of the northern parts of the country, the correlation coefficient increases to between 0.25 and 0.5. In comparison to the correlations between the hindcast data and the reanalyses during MAM (Figure 75), the *r*-value decreases over the Atlantic Ocean and over the Indian Ocean the *r*-value increases. The correlation patterns are consistent between the three reanalysis datasets (Figure 76).

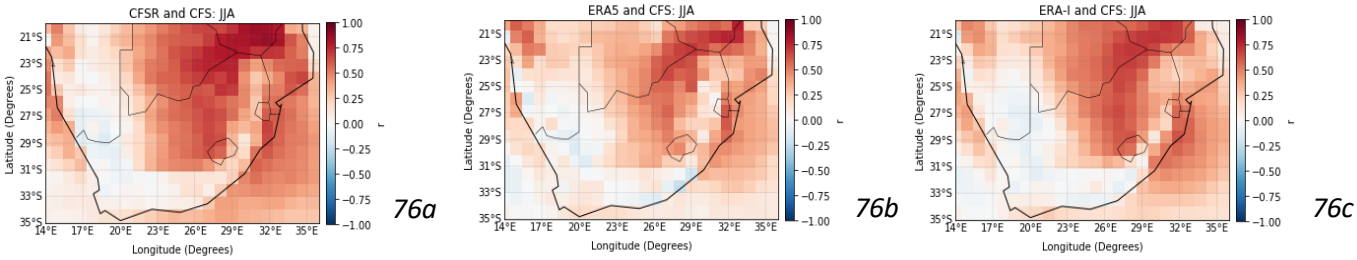


Figure 76: Plot of the Pearson Correlation Coefficient between the hindcast data and the reanalysis datasets for JJA between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

September – October – November

During SON, over the majority of the country, there are weak positive correlations between CFS wind speed data and reanalyses wind speed data. The Pearson Correlation Coefficient increases along the south-west coast. Over the Atlantic Ocean there are strong positive correlations and over the Indian Ocean there are weak correlations (Figure 77).

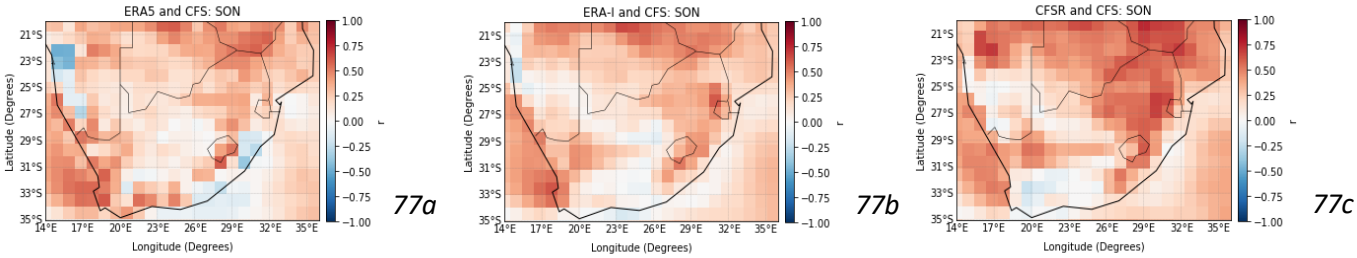


Figure 77: Plot of the Pearson Correlation Coefficient between the hindcast data and the reanalysis datasets for SON between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

7.5.2. Root Mean Square Error

This section of the results presents the Root Mean Square Error (RMSE) between the proxy for observed data and the hindcast data. The results aim at informing the study about the errors in the forecasted data, compared to the observed data. The lower the RMSE value, the higher the correlation between two datasets. A RMSE value of zero, corresponds to a Pearson Correlation Coefficient of one. The RMSE value has no upper limit; the maximum RMSE values depend on the datasets being used. For Section 7.5.2., the RMSE scale is from “0” to “3”.

The RMSE calculation has been performed using monthly (Figure 78) and seasonal data (Figure 79). There are low RMSE values over the northern interior areas of the study area, where the RMSE value is between 0 and 0.5 (Figure 78 and 79). The RMSE value increases further south, and along the coastlines there is a peak in RMSE values. The results are consistent with the results in Section 7.5.1. The areas that have a higher RMSE value are consistent with *r*-values that are closer to zero, indicating that over these areas, there is weaker relationship between the hindcast data and the reanalysis data. Monthly and seasonal error shows similar spatial and magnitude characteristics.

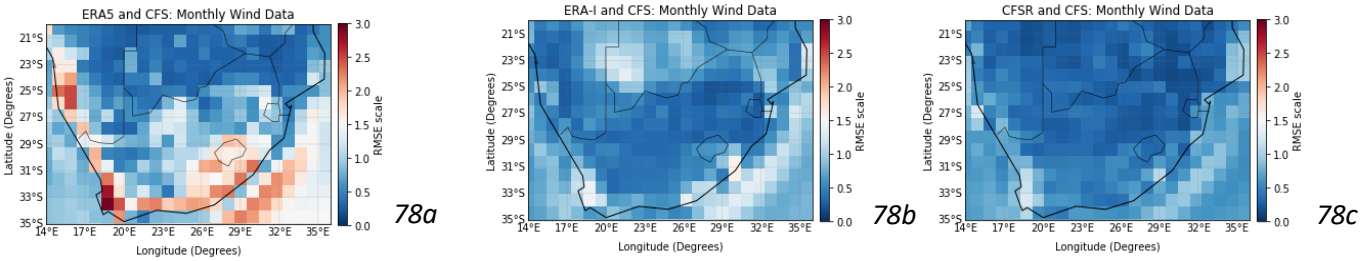


Figure 78: Plot of the Root Mean Square Error between monthly hindcast data and reanalysis data between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

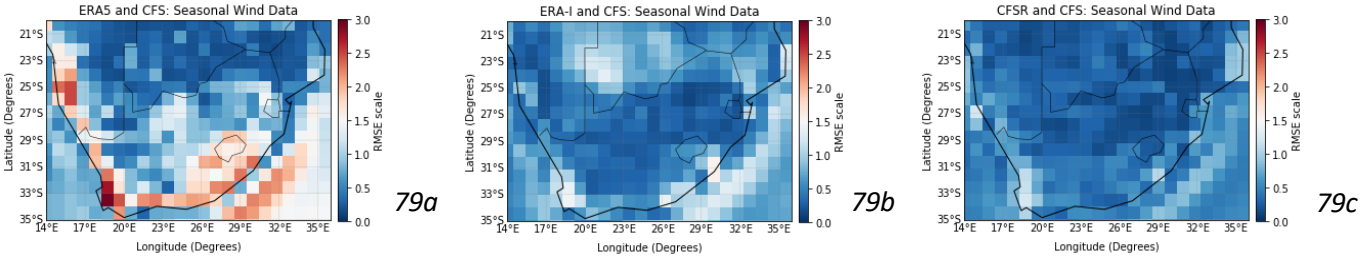


Figure 79: Plot of the Root Mean Square Error between seasonal hindcast data and reanalysis data between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

7.5.2.1. Seasonal Comparisons

The results below present the Root Mean Square Error between the reanalyses wind speed data and CFS wind speed data for the four seasons: DJF, MAM, JJA and SON.

December – January – February

During DJF (Figure 80), the highest RMSE values are over the west and south coast of the country, as well as offshore over the Indian Ocean. The lowest RMSE values are over the north-eastern parts of South Africa. The RMSE results for DJF are not consistent with the Pearson Correlation Coefficient results for DJF (Figure 74). The RMSE calculation indicates that there is a decreased amount of error in the forecasted data compared to the reanalyses over the northern regions of the country and an increased amount of error in the forecasted data over the southern regions of South Africa (Figure 80).

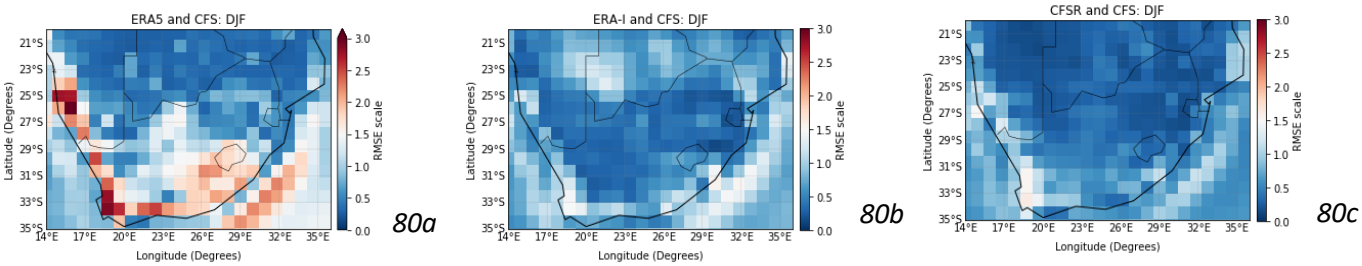


Figure 80: Plot of the RMSE between seasonal hindcast data and reanalysis data for DJF between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

March – April – May

The RMSE values during MAM (Figure 81), are similar to DJF (Figure 80), where the highest values are over the coastlines and the lowest RMSE values are towards the northern interior of the country. The RMSE values between CFSR and CFS decreased along the west coast, compared to DJF. The results that made use of ERA5 and CFSR indicate that the lowest RMSE values are over the northernmost province of South Africa (Limpopo) (Figure 81a and 81c); however, the results created using the ERA-Interim dataset indicate that the RMSE value increases over this area (Figure 81b).

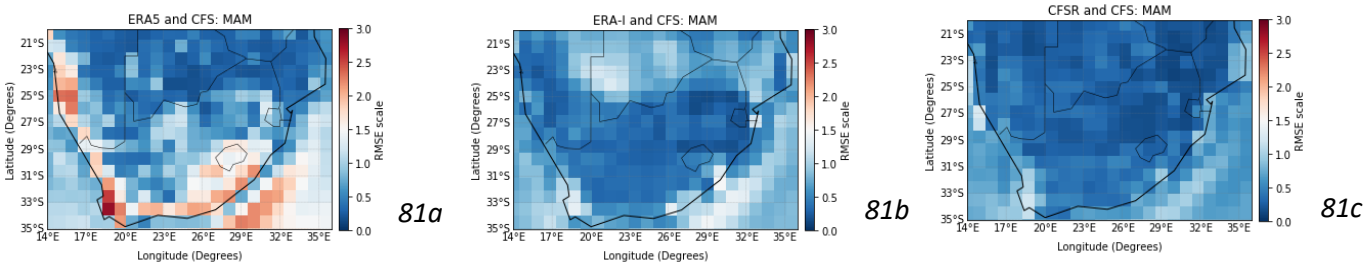


Figure 81: Plot of the RMSE between seasonal hindcast data and reanalysis data for MAM between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

June – July – August

During JJA (Figure 82), the RMSE values between ERA5 and CFS decreased from DJF (Figure 80) over the interior regions of the country. Over the northernmost regions there continues to be the lowest RMSE values between ERA5 and CFS (Figure 82a) and CFSR and CFS (Figure 82c), suggesting that forecast error in this region during JJA is relatively low. The lowest RMSE values between ERA-Interim and CFS can be seen over the interior of South Africa (Gauteng, Free State, Kwa-Zulu Natal, the North West) (Figure 82b).

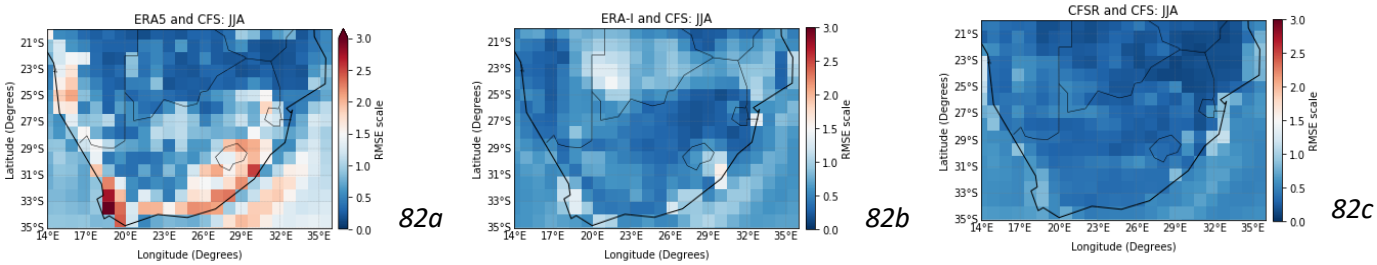


Figure 82: Plot of the RMSE between seasonal hindcast data and reanalysis data for JJA between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

September – October – November

During SON (Figure 83), the RMSE patterns are similar to DJF (Figure 80). The RMSE values increase over the Indian Ocean and along the west coast compared to JJA (Figure 81). The RMSE value between ERA5 and CFS increased over the interior of South Africa (Figure 83a).

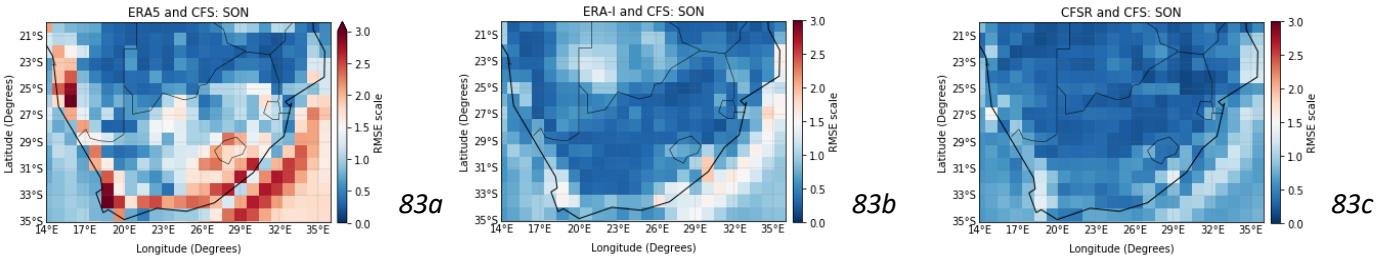


Figure 83: Plot of the RMSE between seasonal hindcast data and reanalysis data for SON between (a) ERA5 and CFS; (b) ERA-Interim and CFS; and (c) CFSR and CFS

The figures in Section 7.5.2. indicate that the RMSE values between the CFSR reanalysis dataset and the CFS hindcast data are the lowest, compared to the ERA5 dataset and the CFS dataset, where the RMSE values are the highest between the three reanalyses. This is expected, given that the CFS and CFSR datasets originated from the same model. The differences highlighted in Section 7.5.2., can also be explained by the fact that the reanalysis models are different in structure regarding how different processes are represented. The differences within the models likely had an impact on the error between each reanalyses and the CFS hindcast data. The results also depict that depending on the reanalysis wind speed dataset used, the level of skill of the forecast will vary which emphasizes the fact that more than one reanalysis wind speed data should be used when assessing the skill of a forecast.

The addition of the RMSE aimed at enhancing the study, as the use of multiple skill measures has the ability to increase the robustness of the results and corroborate key findings. This can be seen when comparing Section 7.5.1. and Section 7.5.2., where there were overlapping results which aided in answering Objective 1 and advanced the understanding of the level of skill of the CFS wind speed forecast. More specifically, the results provide clarity on where the forecast is most skilful over South Africa.

7.6. Large Scale Drivers of Wind

7.6.1. Multivariate ENSO Index

Section 7.6.1. presents the Multivariate ENSO Index (MEI) results, and how the El Niño Southern Oscillation (ENSO) impacts wind characteristics and predictability over South Africa. ENSO fluctuates between El Niño, neutral and La Niña phases and these fluctuations are presented using the Multivariate ENSO Index, where higher MEI values indicate an El Niño and lower MEI values indicate a La Niña. A MEI value of zero indicates that the ENSO phenomenon is not present (Figure 84). The section is divided into three sub-sections, where Section 7.6.1.1. analyses the entire MEI dataset, Section 7.6.1.2. presents the results for El Niño, neutral and La Niña phases and Section 7.6.1.3. presents the results using MEI values which represent strong El Niño and strong La Niña events.

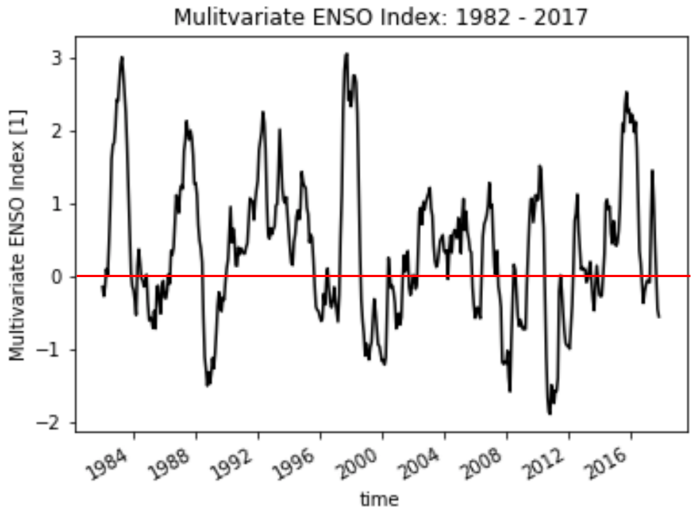


Figure 84: Time series of the Multivariate ENSO Index from 1982 to 2017

The results below show the Pearson Correlation Coefficient (*r*) over time for each grid cell of the study area between the MEI and the three reanalysis wind datasets and MEI and the CFS hindcast dataset. The *r*-value scale has been constrained for the results below to "-0.4" to "0.4", in order to make the results more understandable, however, the Pearson Correlation Coefficient usually ranges from "-1" to "1".

7.6.1.1. MEI and Monthly Wind Speed Correlations

The results below present the Pearson Correlation Coefficient between MEI and monthly wind speed data, from 1982 to 2017 for the ERA5, ERA-Interim and CFS datasets and from 1982 to 2010 for the CFSR dataset. Reanalysis wind speed data and hindcast wind speed data have been used to perform the Pearson Correlation Coefficient calculations with MEI.

Over South Africa, the correlations between the total MEI dataset (1982 to 2017) and the reanalyses wind speeds are weak. The ensemble mean has been used to calculate the Pearson Correlation Coefficient between MEI and CFS wind speeds. The correlation pattern seen is similar to reanalyses figures; however, the areas of positive correlations between the MEI and CFS (Figure 85d) is stronger than between MEI and the reanalysis datasets (Figure 85a to 85c). The CFS figure (Figure 85d) and the ERA-Interim figure (Figure 85b) have the most similarities, with regards to correlation patterns. Over the interior of South Africa, there are positive correlations which increases in strength towards the north of the country (over the North West and Limpopo). There are weak negative correlations between the two variables over the south-west coast of the country, and over the regions in north of South Africa (Mpumalanga) (Figure 85d). There are weak positive correlations over the interior of the country (Northern Cape, Free State, Eastern Cape and the northern areas of the Western Cape), as well as over the north-eastern coast (Kwa-Zulu Natal). There are also areas across the country where there is no correlation between the MEI and wind speed (Figure 85). These

results indicate that the drivers of wind patterns in South Africa depend on the region that is being studied.

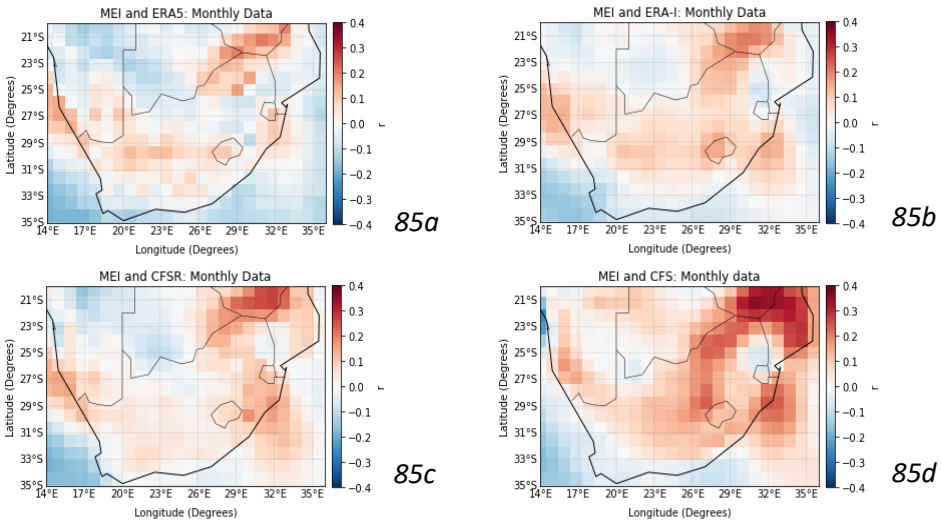


Figure 85: Plot of the Pearson Correlation Coefficient between the total MEI dataset and the four wind datasets between (a) MEI and ERA5; (b) MEI and ERA-Interim; (c) MEI and CFSR; and (d) MEI and CFS

7.6.1.2. Seasonal Correlations

The results in Section 7.6.1.2 present the correlations between MEI and wind speed data from the three reanalysis wind speed datasets and the CFS forecasted data, using the complete MEI dataset. The seasonal forecast during DJF, responds to the ENSO phenomenon, specifically between the CFS hindcast data and the MEI. This however, could be due to the fact that seasonal forecasts are highly likely to respond to an ENSO event based on the way the model has been programmed and therefore it means that the forecast may not be accurate. The results presented in Section 7.6.1.2. indicate that ENSO has a greater effect over South Africa during austral summer months and contributes to answering Objective 3, by highlighting the season in which a large scale driver, such as ENSO, has the most influence over wind speed over the country.

December – January – February

The correlation between MEI and CFS is stronger than between MEI and the reanalysis datasets during DJF (Figure 86). Over the northern and southern most parts of South Africa there is a negative correlation (Figure 86a and 86b). The correlation between MEI and the four wind speed datasets indicate that a portion of the interior of the country and over parts of the east coast, have positive correlations (Free State, Eastern Cape, Kwa-Zulu Natal, Western Cape); the strongest positive correlation between MEI and CFS wind speed is over the interior of South Africa (Figure 86d). There is a consistent negative correlation along the south-west coast, across the four wind datasets (Figure 86).

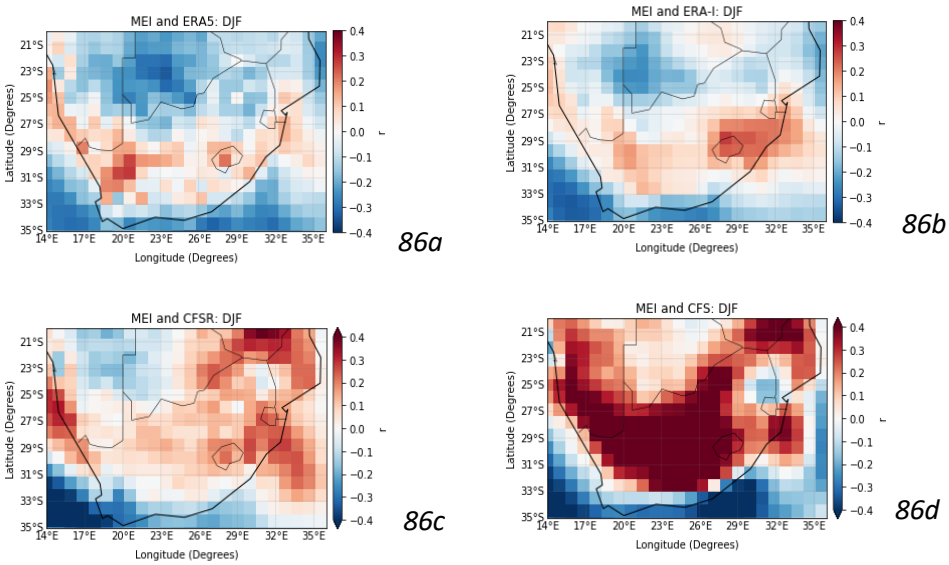


Figure 86: Plot of the Pearson Correlation Coefficient between the MEI and wind speed data for DJF between (a) MEI and ERA5; (b) MEI and ERA-Interim; (c) MEI and CFSR; and (d) MEI and CFS

March – April – May

There are stronger correlations between MEI and CFS (Figure 87d), compared to MEI and the reanalyses during MAM (Figure 87a to 87c). The correlation between MEI and the CFS hindcast data during MAM indicates that there are strong positive correlations ($0.4 < r$) over parts of the east coast and over the northern provinces of South Africa. There are weaker positive correlations between MEI and CFS wind speed over the southern regions of South Africa (Eastern Cape, Western Cape and the Northern Cape) and negative correlations over the western regions of the country (Figure 87d). The results indicate that there are some regions across South Africa where there is no correlation between MEI and wind speed during MAM (Figure 87).

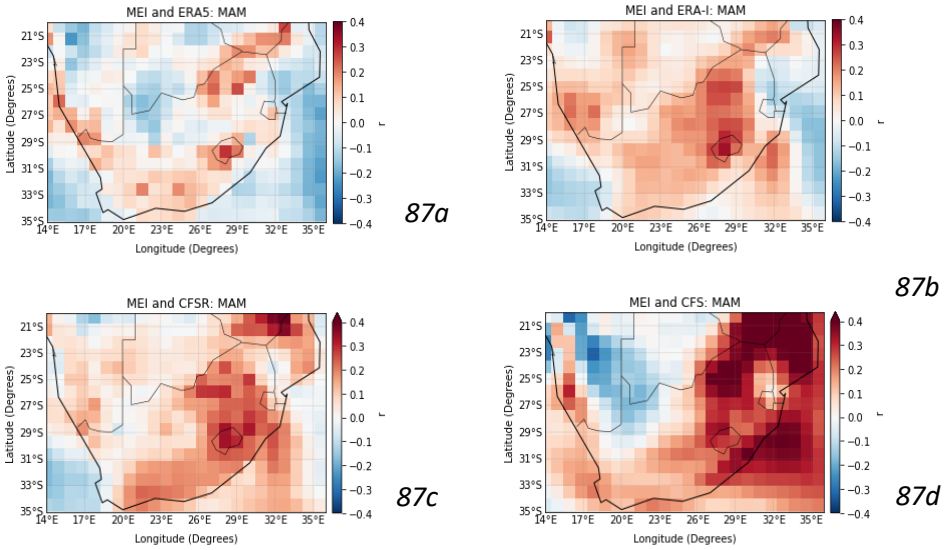


Figure 87: Plot of the Pearson Correlation Coefficient between the MEI and wind speed data for MAM between (a) MEI and ERA5; (b) MEI and ERA-Interim; (c) MEI and CFSR; and (d) MEI and CFS

June – July – August

According to the reanalysis dataset results (Figure 88a to 88c), there are weak correlations between the MEI and wind speed data in JJA over the majority of the country ($-0.2 \leq r \leq 0.2$). There is also a lack of consistency between the correlations between MEI and the reanalyses (Figure 88a to 88c) and the correlations between MEI and the CFS hindcast data (Figure 88d). The correlation between the hindcast data and MEI (Figure 88d), indicates that over the interior of South Africa there is a weak negative correlation between the two variables, whereas the Pearson Correlation Coefficient calculation between MEI and the reanalyses results (Figure 88a to 88c), indicate that there is weak positive correlation over the same area. The regions of consistent results between the four figures are regions towards the north of South Africa (Kwa-Zulu Natal and north-western Limpopo) and regions of the west coast (Figure 88).

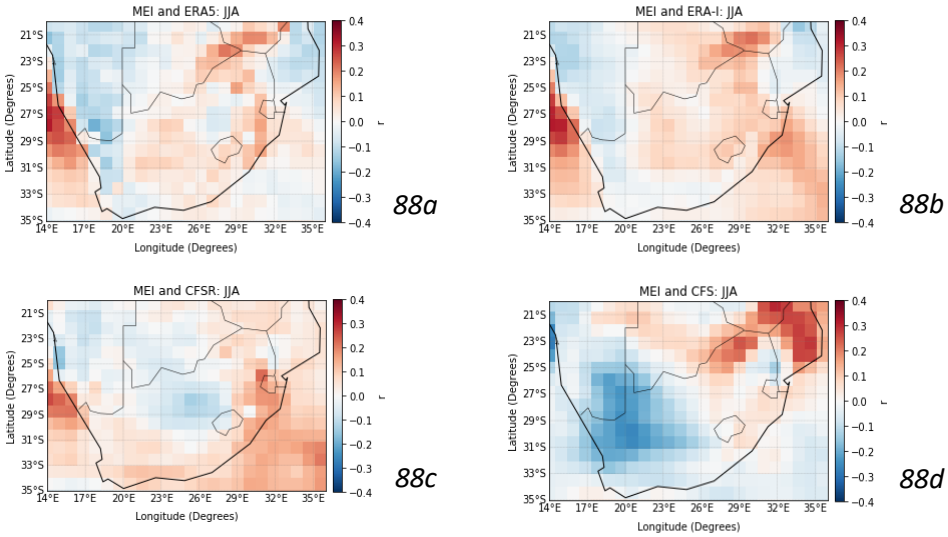


Figure 88: Plot of the Pearson Correlation Coefficient between the MEI and wind speed data for JJA between (a) MEI and ERA5; (b) MEI and ERA-Interim; (c) MEI and CFSR; and (d) MEI and CFS

September – October – November

The ERA5, ERA-Interim and CFSR figures show that over the majority of South Africa during SON, there is a negative correlation between MEI and wind speed, with the exception of a few selected regions towards the north of South Africa (Limpopo and parts of the Northern Cape) (Figure 89a to 89c) and over the interior regions of the country (the Free State) (Figure 89d), where there is a positive correlation. The consistency in the results, especially over the eastern regions of the country, increases the robustness of the results and enhances the understanding the cross-scale drivers of wind patterns in South Africa (Figure 89).

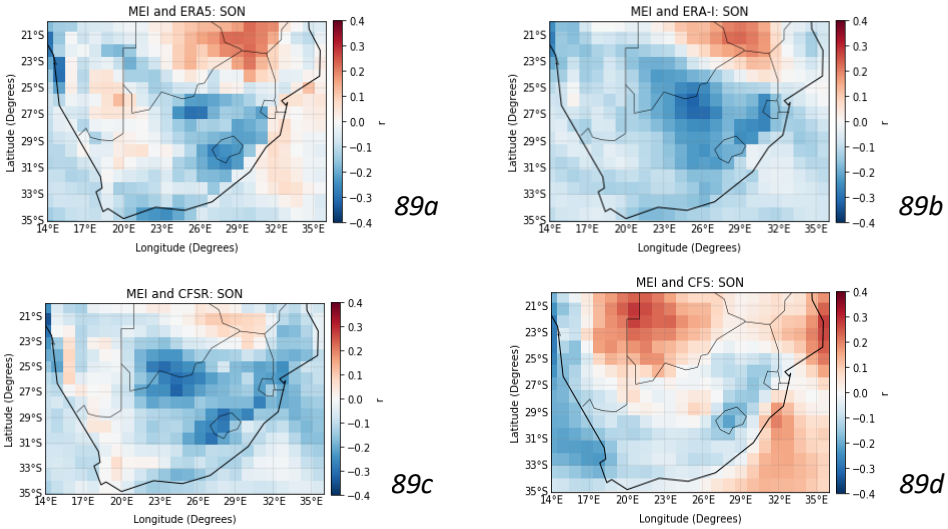


Figure 89: Plot of the Pearson Correlation Coefficient between the MEI and wind speed data for SON between (a) MEI and ERA5; (b) MEI and ERA-Interim; (c) MEI and CFSR; and (d) MEI and CFS

7.6.1.3. El Niño, Neutral and La Niña Phases

The results below present the Pearson Correlation Coefficient between the positive, neutral and negative MEI time periods and monthly ERA5, ERA-Interim, CFSR and CFS wind speed data. Positive MEI values (0.5 < MEI), indicate an El Niño phase, a negative MEI value (MEI < -0.5), indicates a La Niña phase, and MEI values between -0.5 and 0.5 indicate neutral phases. The National Oceanic and Atmospheric Administration uses a threshold of ± 0.5 to indicate the different phases of ENSO, when using the Multivariate ENSO Index (NOAA, 2021).

7.6.1.3.1. El Niño

There are weak correlations between MEI and wind speed over the majority of South Africa when MEI is greater than 0.5 (El Niño). The regions of no correlation between the MEI and wind speed datasets can be seen across the country (Figure 90). During positive phases of MEI, there are weak correlations between MEI and wind speed over the surrounding oceans. The correlation patterns are consistent between MEI and the three reanalyses (Figure 90a to 90c) and between MEI and the hindcast dataset (Figure 90d), except for the northern-most and southern-most regions of South Africa (Figure 90d). The consistent result over the majority of the country enhances the robustness of the result which indicates that the positive phase of ENSO is weakly correlated to wind speed over South Africa, when not looking at one specific season.

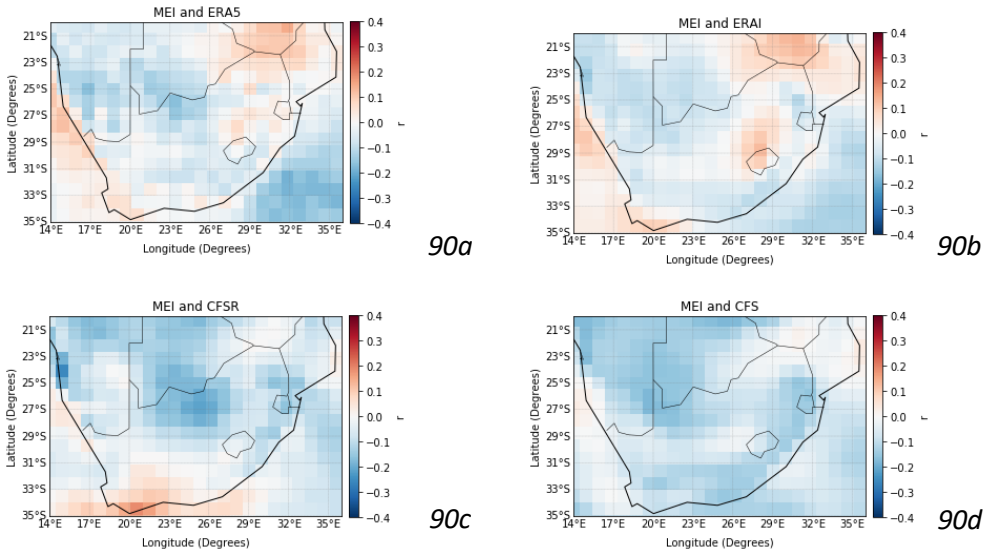


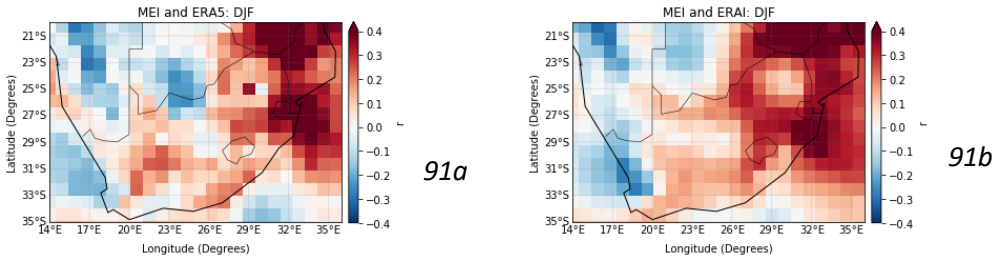
Figure 90: Plot of the Pearson Correlation Coefficient between positive MEI periods (the positive phase of ENSO) and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS

7.6.1.3.1.1. El Niño Seasonal Correlations

This section of the results presents the seasonal correlations (DJF, MAM, JJA and SON), between the MEI and the three reanalysis datasets and MEI and the hindcast dataset during El Niño events (0.5 < MEI).

December – January – February

The correlation between the reanalyses wind speeds and MEI are similar during El Niño phases in DJF. There is a positive correlation along the east coast and eastern-interior of South Africa, as well as over the northern regions of the country (Figure 91a to 91c). There are weaker correlations between MEI and CFS, compared to the reanalyses; however, there are strong negative correlations (MEI < -0.4), over the southern-most region of the country (Figure 91d).



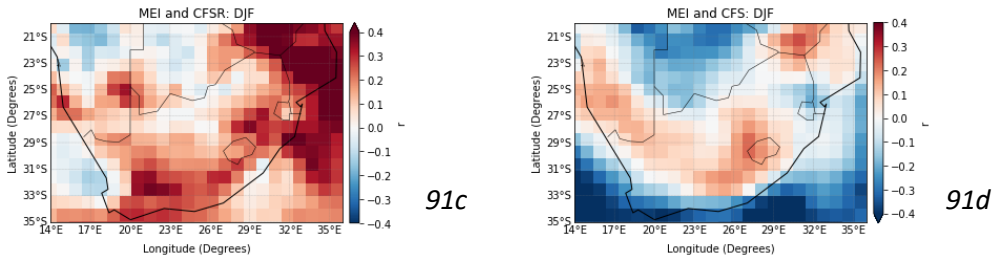


Figure 91: Plot of the Pearson Correlation Coefficient between MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases (0.5 < MEI) in DJF

March – April – May

The correlations between MEI and wind speed during MAM (Figure 92) are weaker compared to DJF (Figure 91); however, there is a similar correlation pattern. There are positive correlations over the northern and interior regions of South Africa and weak positive and negative correlations over the southern regions of the country. These results are consistent across the correlations between the MEI and the four wind speed datasets (Figure 92), which indicates that during an El Niño event in MAM, the below correlation patterns can be expected.

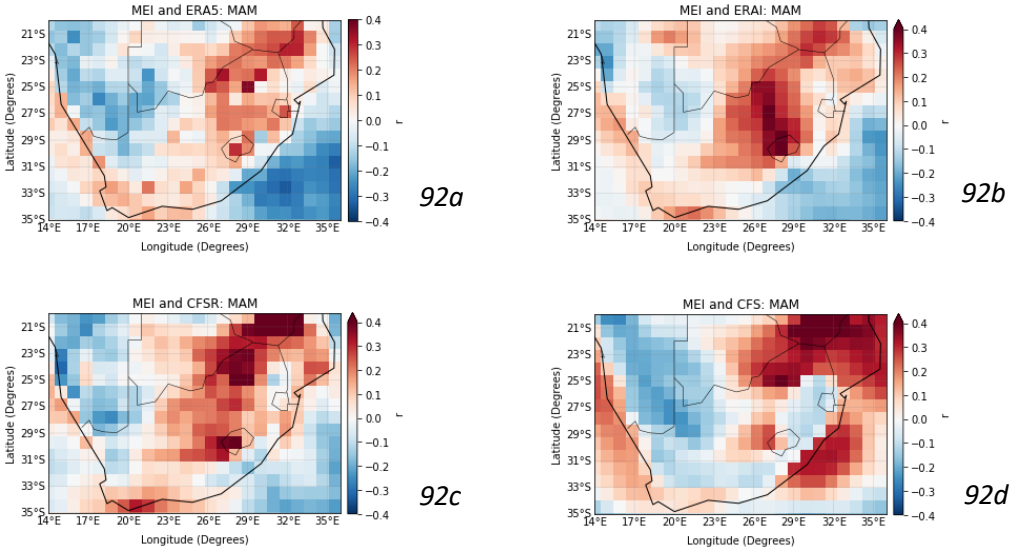


Figure 92: Plot of the Pearson Correlation Coefficient between MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases (0.5 < MEI) in MAM

June – July – August

The correlation pattern during JJA (Figure 93), differs to DJF and MAM (Figure 91 and 92). The MEI correlations with the ERA datasets (Figure 93a and 93b), indicate that there is a negative correlation between wind speed and the MEI during El Niño events. The correlation pattern between the CFSR reanalysis wind speed dataset and MEI indicates that there is a positive correlation between the variables over the southern regions of South Africa (regions in the Western Cape and Eastern Cape), a negative correlation over the interior regions (regions of

the Northern Cape, Free State and North West) and weak correlations over the northern regions of the country (regions of Limpopo, Gauteng, Mpumalanga and the North West) (Figure 93c). The correlation pattern between CFS and MEI, indicates that there is a weak positive correlation over the northern regions of the country and negative correlations over the south-western regions of the country (Figure 93d). There is dataset agreement with regards to a positive correlation between wind speed and MEI during moderate El Niño phases in JJA over the southern-most point of South Africa (Figure 93).

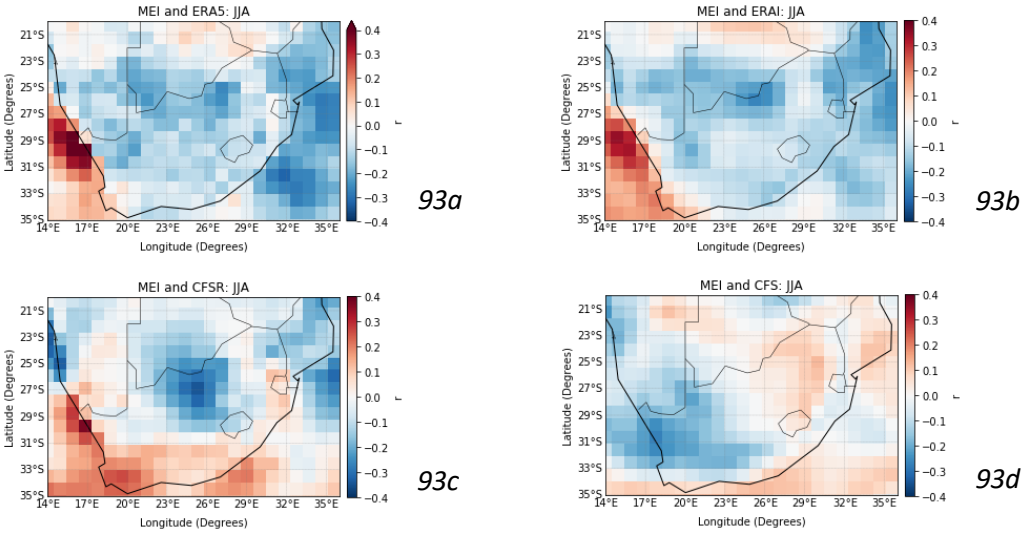
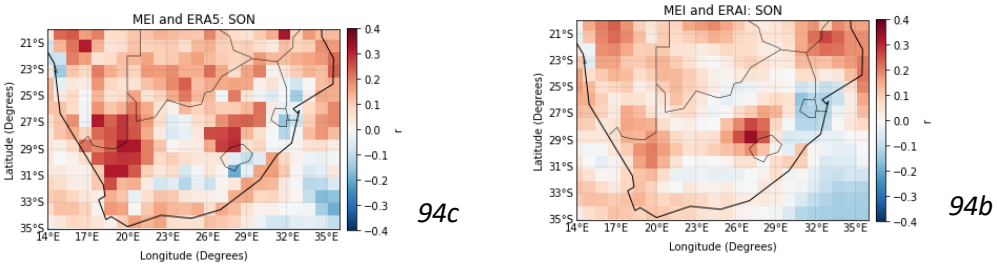


Figure 93: Plot of the Pearson Correlation Coefficient between MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases (0.5 < MEI) in JJA

September – October – November

During moderate El Niño events in SON, the results present a positive correlation between wind speed and MEI over the south-western regions of South Africa (Figure 94); however, the correlation between CFS (Figure 94d) and MEI is weaker than the reanalyses. Over the interior regions of the country, there is a positive correlation between MEI and the ERA reanalyses and CFS (Figure 94a, 94b and 94d); whereas over the same region, the correlation between MEI and CFSR is negative (Figure 94c). The lack of agreement between MEI and the four datasets, indicates that during SON, El Niño as a driver of wind over South Africa is complex and is likely dependent on other factors, which are not assessed in this study.



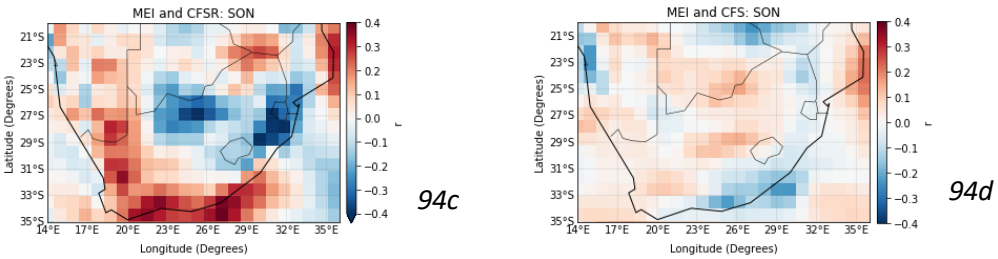


Figure 94: Plot of the Pearson Correlation Coefficient between MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases (0.5 < MEI) in SON

7.6.1.3.2. Neutral Phase

During a neutral phase of ENSO (-0.5 ≤ MEI ≤ 0.5), the *r*-values range between -0.245 and 0.035. There is a is a negative correlation between MEI and wind speed during a neutral phase over the South Africa and the correlation pattern over the study area is uniform over the four datasets; however, the ERA datasets (Figure 95a and 95b) and MEI have weaker correlations compared to the CFS datasets (Figure 95c and 95d) and MEI. The consistency in the results helps answer Objective 3, whereby it can be deduced that during a neutral phase of ENSO, as the MEI decreases (increase), wind speed will likely increase (decrease).

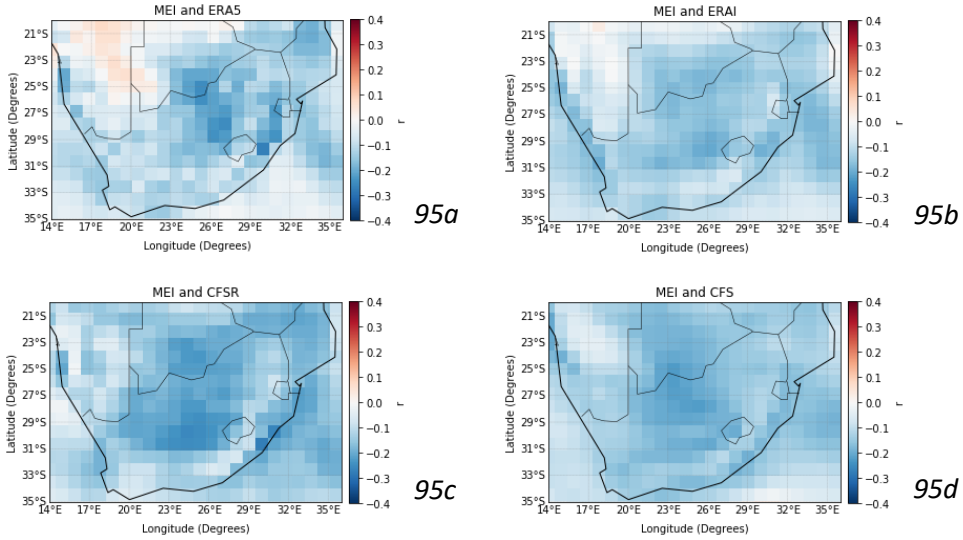


Figure 95: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during neutral ENSO phases where -0.5 ≤ MEI ≤ 0.5

7.6.1.3.2.1. Neutral Phase Seasonal Correlations

Section 7.6.1.3.2.1. of the results presents the seasonal correlations (DJF, MAM, JJA and SON), between the MEI and the three reanalysis datasets and the hindcast dataset during neutral phases of ENSO (-0.5 ≤ MEI ≤ 0.5).

December – January – February

There is a positive correlation between MEI and wind speed over the interior regions of South Africa in DJF during neural phases of ENSO. This results can be seen in the correlation between the ERA reanalyses and MEI (Figure 96a and 96b), as well as between CFSR and MEI, however the positive correlation is weaker (Figure 96c). There is a weak positive correlation between CFS and MEI during DJF over the majority of the country, and weak negative correlations over the southern regions of the country (Northern Cape, Western Cape and North West) (Figure 96d). Over the northern regions of South Africa (Limpopo), there are positive correlations between MEI and the ERA-Interim, CFSR and CFS wind speed datasets (Figure 96b to 96d). Figure 96 alludes to the fact that during DJF and during the neutral phase of ENSO, the correlation depends on the site. In addition, the lack of agreement between the correlation between the reanalyses and MEI and the hindcast data and MEI indicates that the forecast does not accurately capture this driver of wind during DJF.

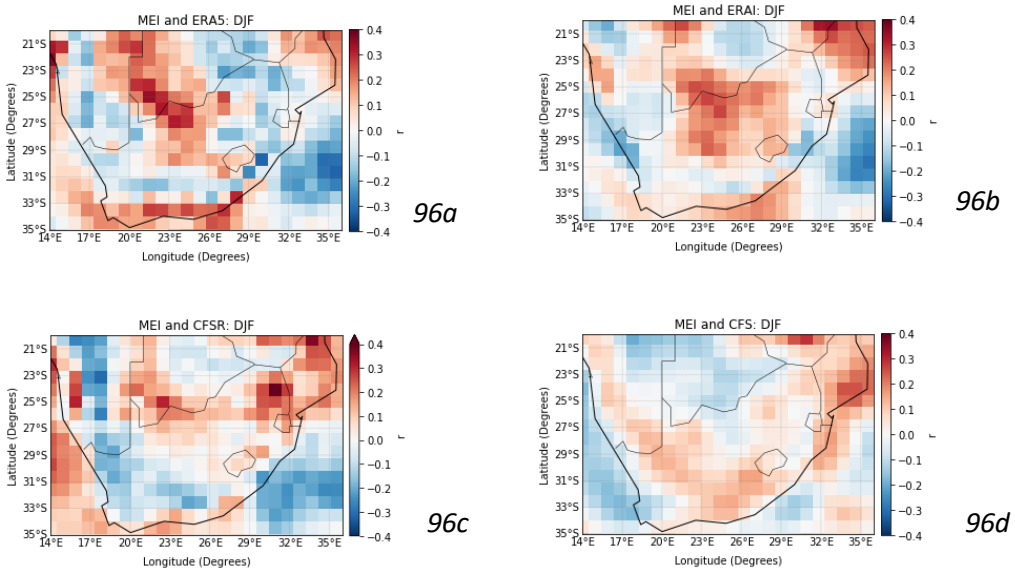


Figure 96: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during neutral ENSO phases in DJF where -0.5 ≤ MEI ≤ 0.5

March – April – May

The correlation between the reanalysis wind speed datasets and MEI indicates that over the interior of South Africa (Northern Cape, North West and the Free State), there are negative correlations with a correlation coefficient value ranging between -0.3 and -0.2 during a neutral phase of ENSO, in MAM (Figure 97a to 97c). Over the same region, there is a weak positive correlation between CFS and MEI (Figure 97d). There is dataset agreement between the four wind speed datasets with regards to the positive correlations over the south-eastern and northern areas of the country (Figure 97). Similar to DJF, the results indicate that during MAM, the wind forecast was unable to capture the large scale driver of wind as accurately as the reanalyses.

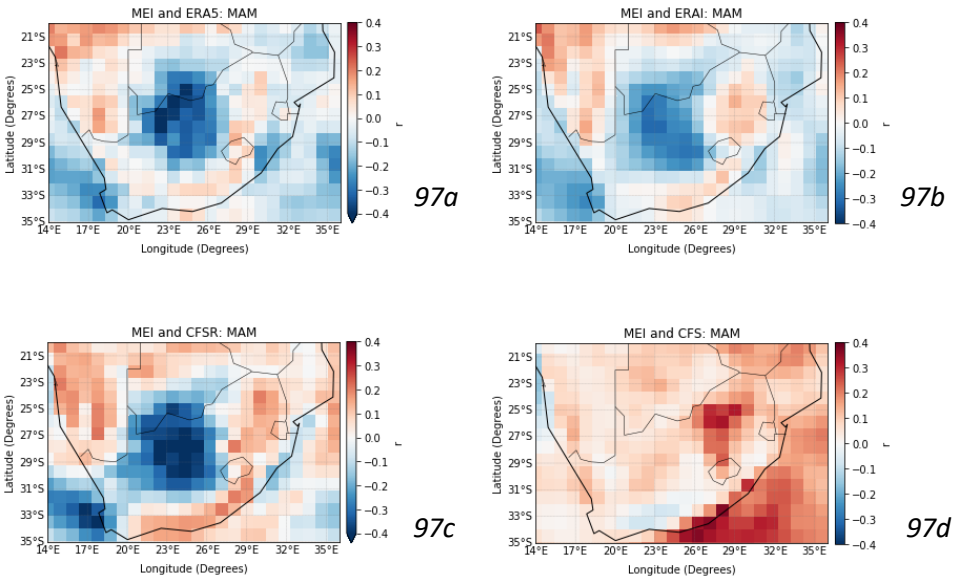


Figure 97: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during neutral ENSO phases in MAM where -0.5 ≤ MEI ≤ 0.5

June – July – August

During a neutral ENSO phase, in JJA, there are negative correlations between wind speed and MEI over the majority of the study area. This result is consistent for the ERA5, ERA-Interim and CFS wind speed datasets (Figure 98a, 98b and 98d). There are weak negative and positive correlations between CFSR and MEI (Figure 98d). Over the southern-most regions of the country, there are positive correlations between wind speed and MEI in JJA. This is demonstrated by the correlations between the four wind speed datasets and the MEI (Figure 98), indicating that there is an increased level of robustness associated with the correlation between wind speed and MEI in JJA over this specific region.

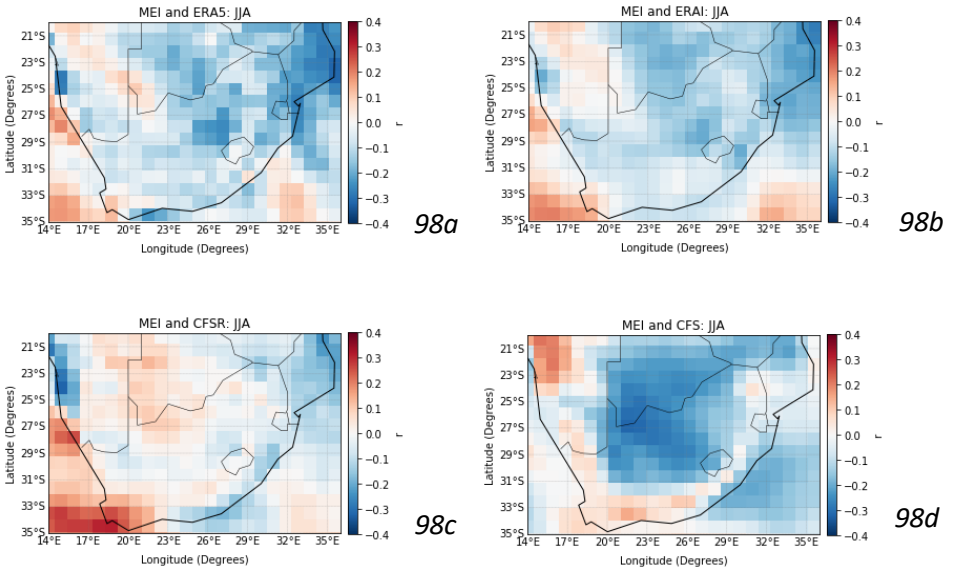


Figure 98: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during neutral ENSO phases in JJA where -0.5 ≤ MEI ≤ 0.5

September – October – November

The correlations between MEI and the ERA5 wind speed dataset, indicate that there are positive correlations with an approximate *r*-value range between 0.1 to 0.2, over the interior and northern regions of the country; whereas the correlations between the ERA-Interim wind speed data and the MEI, indicate weaker correlations over the same area (Figure 99a and 99b). The CFSR and MEI correlations follow a similar correlation pattern, however the positive correlations are weaker and the negative correlations are stronger (Figure 99c). The positive correlations between the CFS wind speed data and the MEI are more isolated to the northern regions of South Africa, and there are negative correlations over the southern regions of South Africa (Figure 99d). There are negative correlations between the MEI and wind speed along the eastern interior and east coast of the country during SON (Figure 99). The lack of consistent results indicate that during SON and during the neutral phase of ENSO, the relationship between ENSO and wind speed cannot be thoroughly understood by only studying the MEI and that additional research is required.

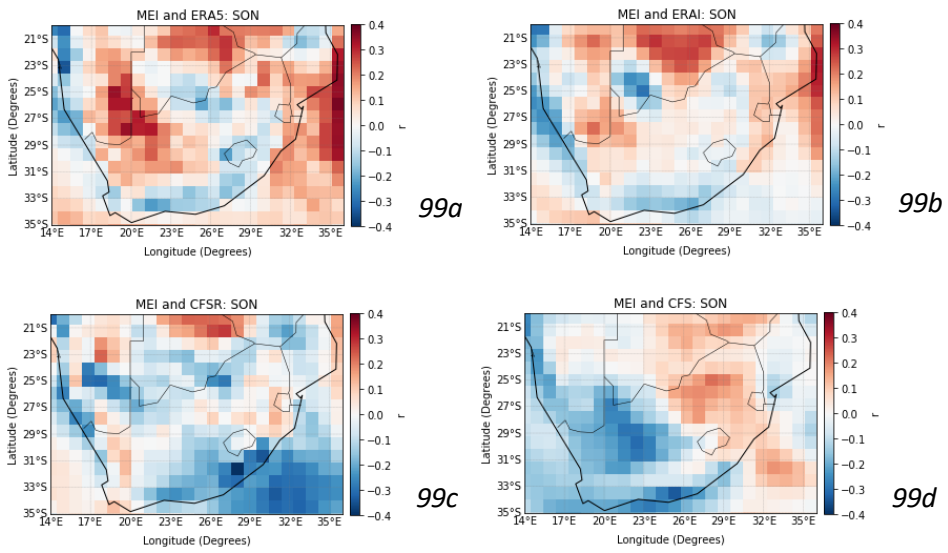


Figure 99: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during neutral ENSO phases in DJF where -0.5 ≤ MEI ≤ 0.5

7.6.1.3.3. La Niña

The three reanalysis wind speed datasets show that there is a positive correlation between negative MEI (moderate La Niña phases) and wind speed over the western regions of the country (Figure 100a to 100c). Both the reanalyses and hindcast correlations with the MEI indicate that there is low skill over the interior regions of South Africa during moderate La Niña events (MEI < 0.5) (Figure 100), the consistency increases the robustness of the result and aids in addressing Objective 3. The strongest negative correlations over land can be seen over the northern-most regions of South Africa (Limpopo) (Figure 100), and the southern-most regions of the country (Western Cape) (Figure 100a to 100c).

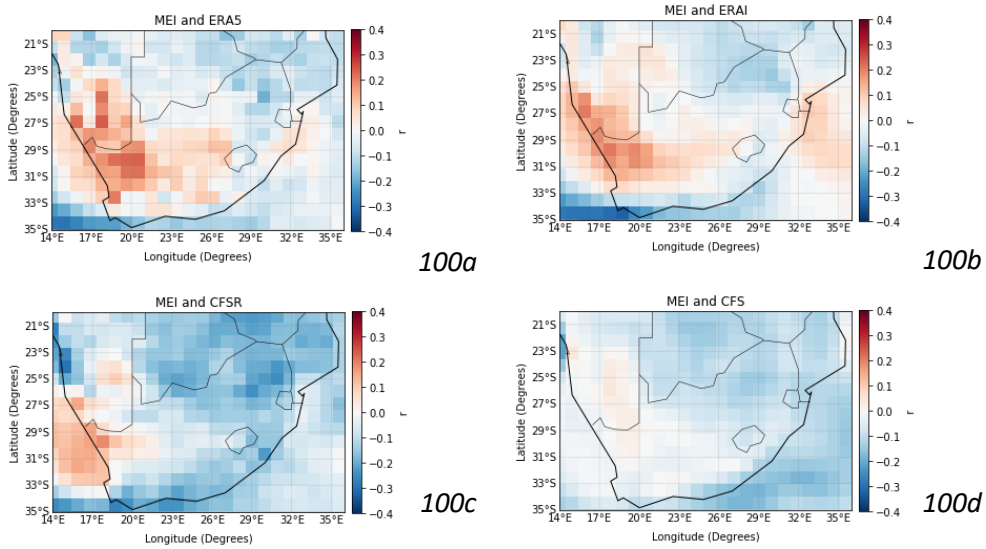


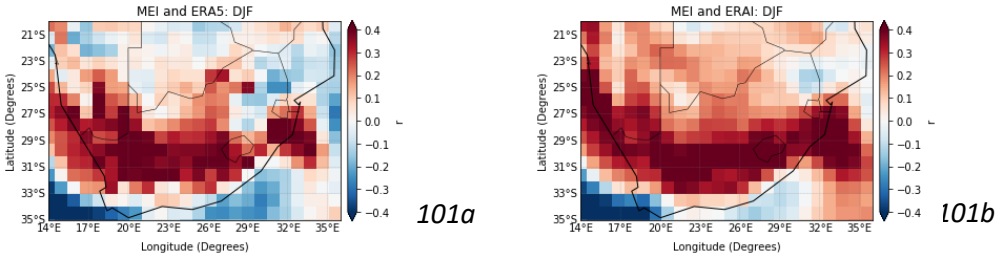
Figure 100: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during La Niña phases where MEI < -0.5

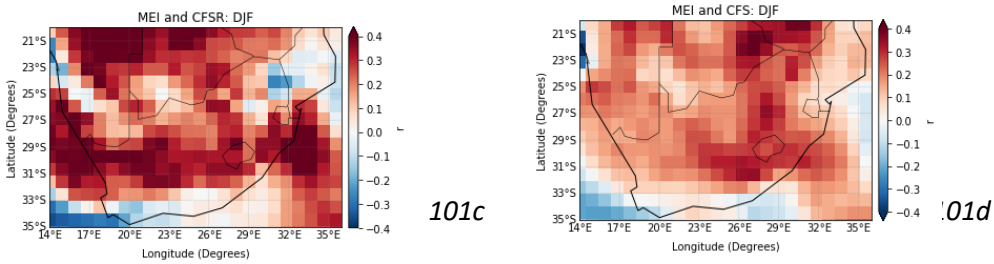
7.6.1.3.3.1. La Niña Seasonal Correlations

The section below presents the seasonal correlations (DJF, MAM, JJA and SON), between the MEI and the three reanalysis datasets and the hindcast dataset during La Niña phases (MEI < -0.5). Due to the fact that there were fewer La Niña events than El Niño and neutral events, the amount timesteps available for each season were limited, and therefore the correlations presented in Section 7.6.1.3.3.1., between wind speed and MEI may be misleading. The limited timesteps has resulted in this study not placing a large emphasis on the seasonal correlations between MEI and wind speed during La Niña events.

December – January – February

During La Niña phases in DJF, there is a strong positive correlation between MEI and wind speed over the interior regions of the country, and a negative correlation over the southern-most regions (Western Cape). The correlation pattern is consistent between the four wind speed datasets (Figure 101).



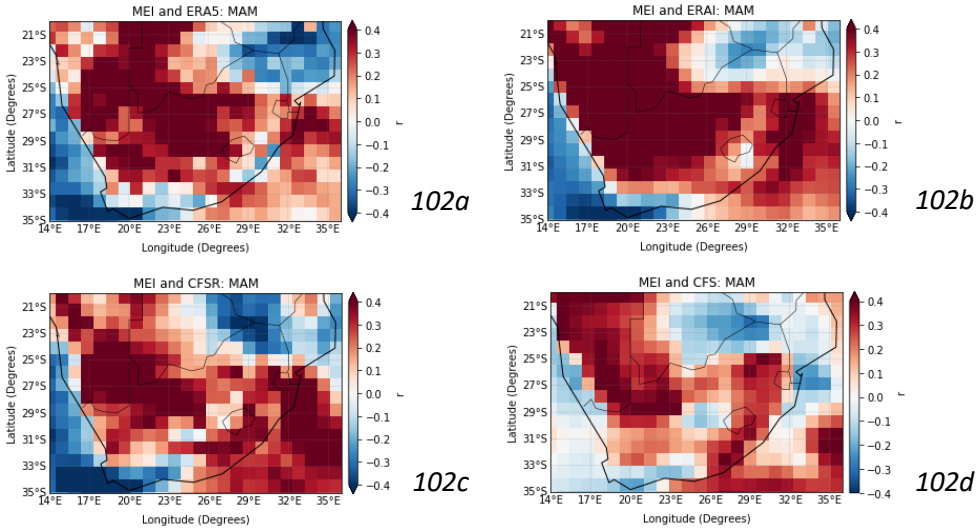


101c 101d

Figure 101: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during La Niña phases in DJF where MEI < -0.5

March – April – May

There are strong positive correlations (0.4 < *r*) between MEI and wind speed over the interior, western and eastern regions of the country and over the east coast over the Indian Ocean. Over the southern and northern-most regions of South Africa, there are negative correlations between wind speed and MEI (Figure 102).

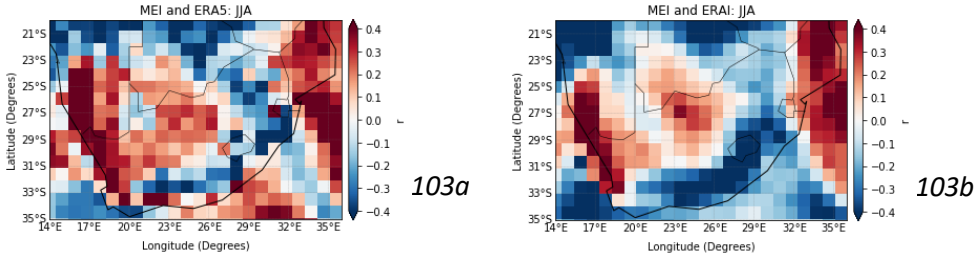


102a 102b 102c 102d

Figure 102: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during La Niña phases in MAM where MEI < -0.5

June – July – August

During JJA, there are negative correlations between MEI and the reanalyses wind speed datasets along the eastern regions of the country and positive correlations over the western and interior regions (Figure 103a to 103c). The correlation pattern between the CFS wind speed data and the MEI differs as there are negative correlations over the majority of the study area and positive correlations along the eastern interior of South Africa (Figure 103d).



103a 103b

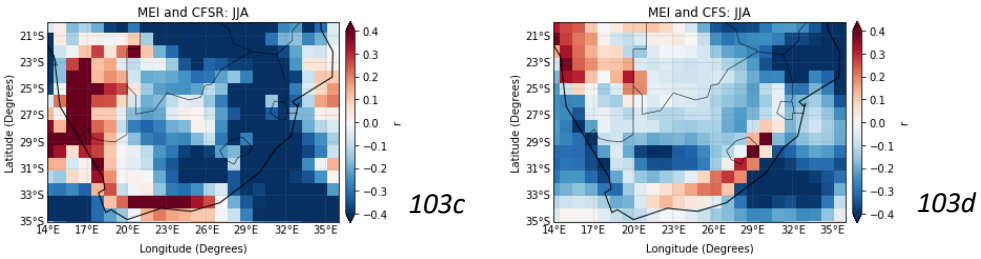


Figure 103: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during La Niña phases in JJA where MEI < -0.5

September – October – November

During La Niña phases in SON, there is limited dataset agreement with regards to correlation patterns; however, over the northern-most regions of the country (Limpopo), there are negative correlations between wind speed and MEI (Figure 104). There are positive correlations between the MEI and the ERA5 and CFSR datasets over the interior regions of the country (Figure 104a and 104c), whereas there is a negative correlation between the MEI and the ERA-Interim and the CFS datasets over the same area (Figure 104b and 104d).

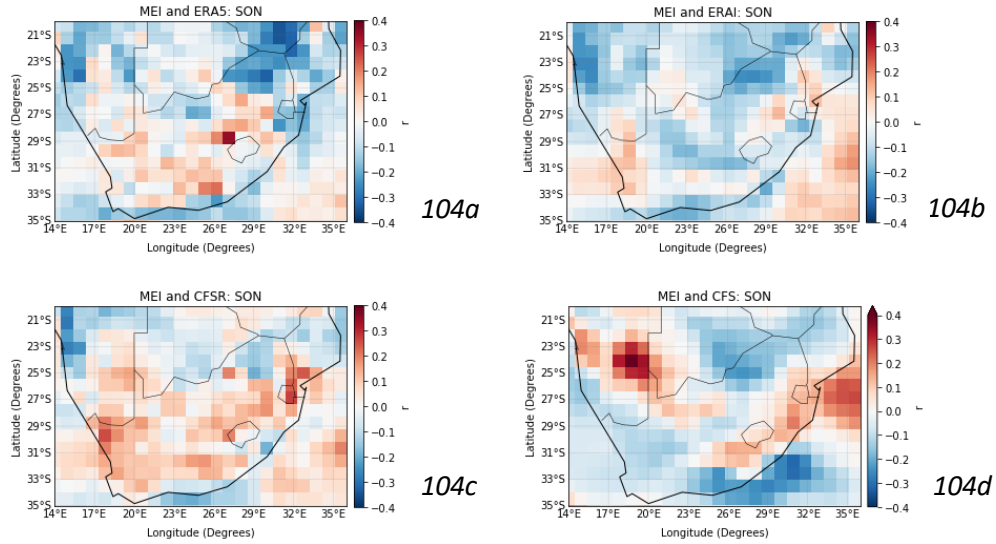


Figure 104: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during La Niña phases in SON where MEI < -0.5

During La Niña phases the correlation range and spatial plots indicate that there are stronger correlations between MEI and wind speed than during an El Niño (Table 5).

Table 5: The maximum and minimum Pearson Correlation Coefficient (r-values) calculated for the correlation between MEI and each wind dataset, based on the data used to create Figure 90, 95 and 100 for El Niño (0.5 < MEI), neutral -0.5 ≤ MEI ≤ 0.5) and La Niña phases (MEI < -0.5).

	CFS	ERA5	ERA-Interim	CFSR
Moderate El Niño Phase				
Maximum *r*-value	0.034	0.149	0.140	0.189
Minimum *r*-value	-0.179	-0.188	-0.149	-0.258
Neutral Phase				
Maximum *r*-value	0.035	0.078	0.0314	-0.004
Minimum *r*-value	-0.227	-0.273	-0.204	-0.277
Moderate La Niña Phase				
Maximum *r*-value	0.045	0.241	0.212	0.191
Minimum *r*-value	-0.213	-0.314	-0.327	-0.301

7.6.1.4 Strong El Niño and La Niña Phases

This section of the results makes use of Mazzarella and Giuliacci's (2009) 'strength index' for MEI in order to determine strong El Niño and strong La Niña phases, which is presented in Section 6.3.7.1. (Table 2 and 3).

7.6.1.4.1 Strong El Niño Phases

During a strong El Niño phase, there are positive correlations over the northern-interior of the country, and negative correlations over the southern regions of South Africa. The correlation patterns across the country are relatively consistent between the wind speed datasets; with a variation in the strength of the correlations between wind speed and the MEI. There are weak correlations between MEI and the four wind datasets over the interior of South Africa (Figure 105).

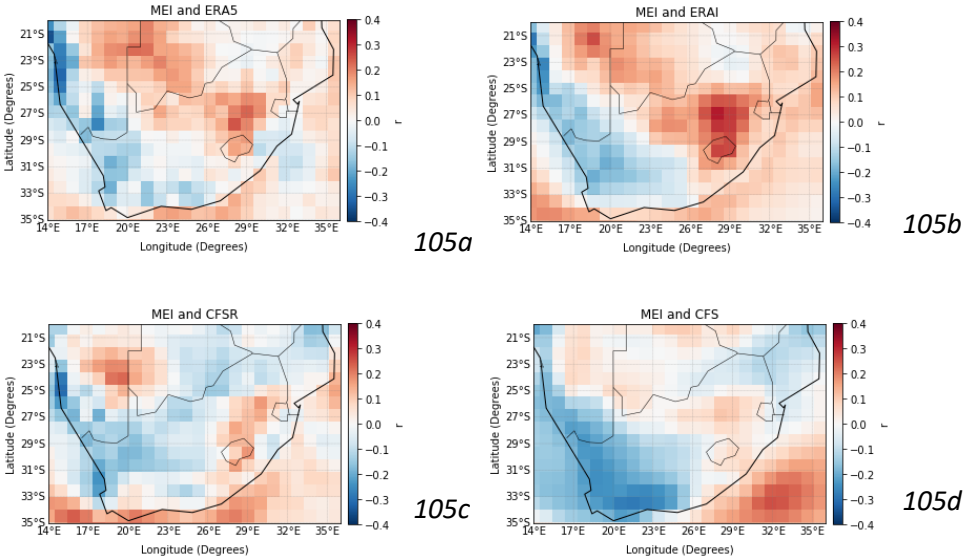


Figure 105: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases where 1.2 ≤ MEI

7.6.1.4.2. Strong La Niña Phases

There are stronger correlation values between wind speed and the MEI during strong La Niña phases (where the MEI is less than -1.2) (Figure 106), compared to strong El Niño phases (Figure 105), suggesting that La Niña events are a good predictor of wind speeds over South Africa, in comparison to El Niño events. The reanalyses results indicate that there are negative correlations over the majority of the regions in South Africa, and positive correlations over parts of the west coast and northern-most areas of the country (Limpopo) (Figure 106a to 106b). There are inconsistencies between the correlations between the MEI and reanalysis wind speed datasets (Figure 106a to 106c), and the hindcast dataset, where the correlation between the CFS data and the MEI indicates that during a strong La Niña event, there are positive correlations over the southern parts of South Africa and negative correlations over the northern and interior areas of the country (Figure 106d). This result indicates that during strong La Niña events, the skill of the CFS to forecast wind speed is low over South Africa. The negative correlations over the Indian Ocean are consistent (Figure 106).

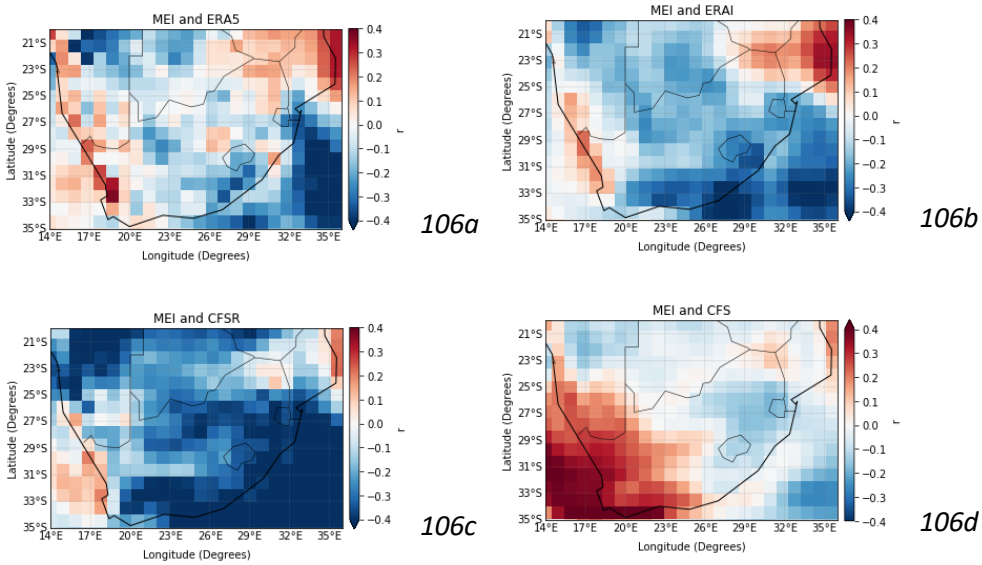


Figure 106: Plot of the Pearson Correlation Coefficient between the MEI and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during El Niño phases where MEI ≤ -1.2

The correlations between wind speed and the MEI during strong La Niña phases are stronger than the correlations between wind speed and the MEI during strong El Niño phases (Table 6).

Table 6: The maximum and minimum Pearson Correlation Coefficient (r-values) calculated for the correlation between MEI and each wind dataset, based on the data used to create Figure 105 and 106 for strong El Niño (1.2 < MEI) and strong La Niña phases (MEI < -1.2)

	CFS	ERA5	ERA-Interim	CFSR
Strong El Niño				
Maximum *r*-value	0.257	0.251	0.317	0.241
Minimum *r*-value	-0.284	-0.352	-0.288	-0.307
Strong La Niña				
Maximum *r*-value	0.467	0.375	0.347	0.227
Minimum *r*-value	-0.261	-0.538	-0.446	-0.611

7.6.2 Southern Annular Mode

The Southern Annular Mode is a climate driver which has three phases: a neutral phase, a positive phase and a negative phase. The SAM index indicates the phases of SAM; when the SAM index is equal to 0, it is indicative of a neutral phase whereas when then SAM index is above (below) 0, it is indicative of a positive (negative) SAM phase (Figure 107).

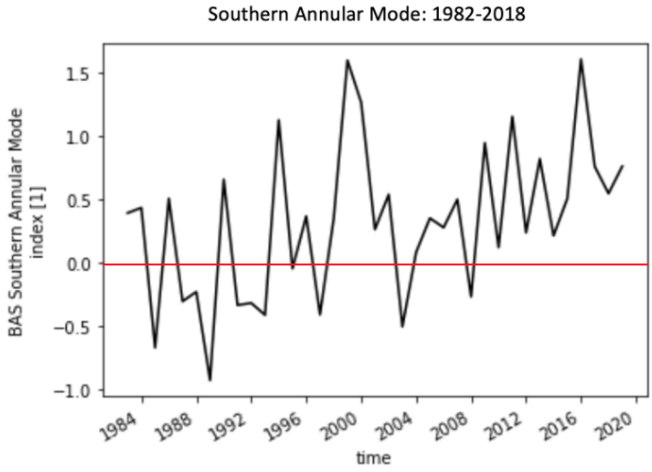


Figure 107: Timeseries of a the Southern Annular Mode from 1982 to 2018

The figures in section 7.6.2.1 present the Pearson Correlation Coefficient (*r*) over time for each grid cell of the study area between the SAM index and the ERA5, ERA-Interim, CFSR and CFS wind speed datasets. The *r* scale range has been constrained for this portion of the study from "-0.4" to "0.4", however, the full r scale extends from "-1" to "1".

7.6.2.1. Monthly Correlations between SAM and Wind Speed

The figures below quantify the relationship between the Southern Annular Mode and wind speed, by comparing the SAM index to the hindcast and reanalysis datasets. Monthly averages have been used.

There is a weak negative correlation between the SAM and wind speed over the interior of South Africa, as well as over the surrounding oceans, when ERA5, CFSR and CFS wind speed data was used (Figure 108a, 108c and 108d). There is a weak positive correlation with ERA-Interim wind speed data and the SAM index, indicating that there are inconsistencies between the reanalysis datasets (Figure 108b). This result aligns with Objective 2, whereby the results indicate that either observed data should be used when assessing the skill of a wind forecast or multiple reanalysis wind speed datasets need to be used in order to ensure that the forecast and emerging patterns are thoroughly verified.

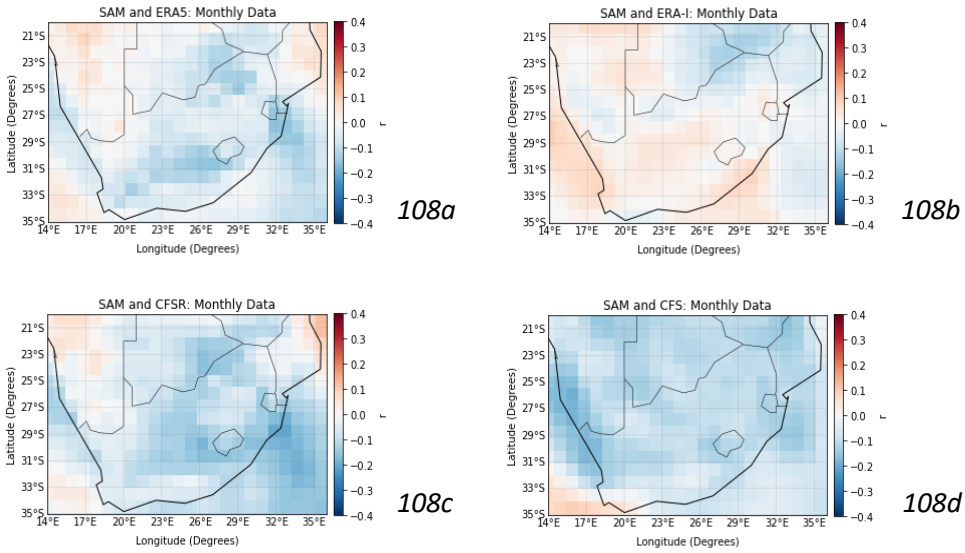


Figure 108: Plot of the Pearson Correlation Coefficient between the total SAM Index dataset and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS

7.6.2.2. Seasonal Comparisons

Section 7.6.2.2. presents the Pearson Correlation Coefficient between the SAM Index and wind speed data at a seasonal timescale (DJF, MAM, JJA and SON).

December – January – February

Correlation patterns between the SAM Index and ERA5 wind speed differs compared to the correlation patterns between the SAM index and ERA-Interim, CFSR and CFS wind speed data during DJF (Figure 109). The ERA-Interim, CFSR and CFS wind speed datasets show negative correlations between the SAM and wind speed (Figure 109b to 109d), however, the correlation between the ERA5 wind speed data and the SAM Index indicates that there are weak positive correlations over the interior of the country, over the same region (Figure 109a). The correlation patterns over the surrounding oceans are also inconsistent; ERA-Interim and CFSR indicate that there is a strong negative correlation over the Indian Ocean, and the ERA5 figure indicates that there is a positive correlation over the same area.

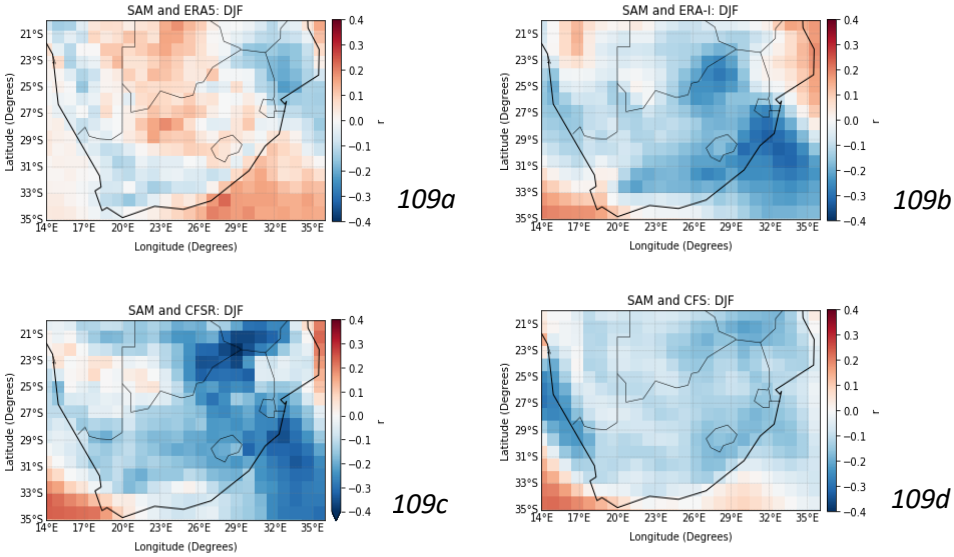


Figure 109*: Plot of the Pearson Correlation Coefficient between the SAM Index and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during DJF*

March – April – May

The areas of positive and negative correlation between the SAM index and wind speed are consistent between the four figures above during MAM (Figure 110), which indicates that the CFS hindcast data has accurately captured the impact of SAM on wind speeds during MAM over South Africa. The figures show that over the eastern areas of South Africa, there are positive correlations between the two variables and over parts of the northern and interior regions of the country (Northern Cape, the Free State and Limpopo), there is a negative correlation.

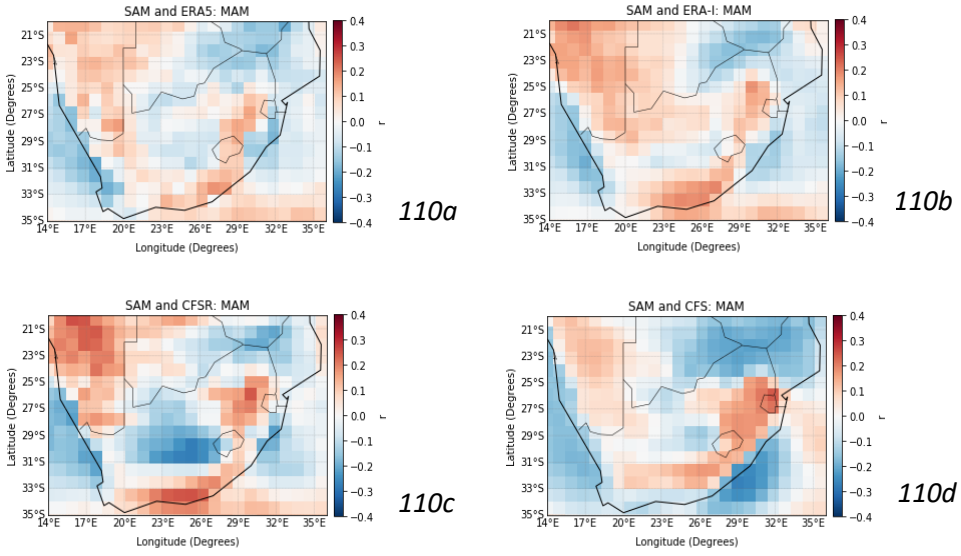


Figure 110: Plot of the Pearson Correlation Coefficient between the SAM Index and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during MAM

June – July – August

During JJA there is a negative correlation between SAM and wind speed over the southern regions of South Africa and weak positive correlations over the northern regions (Figure 111). The correlation between the CFSR data and the SAM Index shows negative correlations across the country, with the exception of two small regions in the northern and western regions of South Africa (Figure 111c).

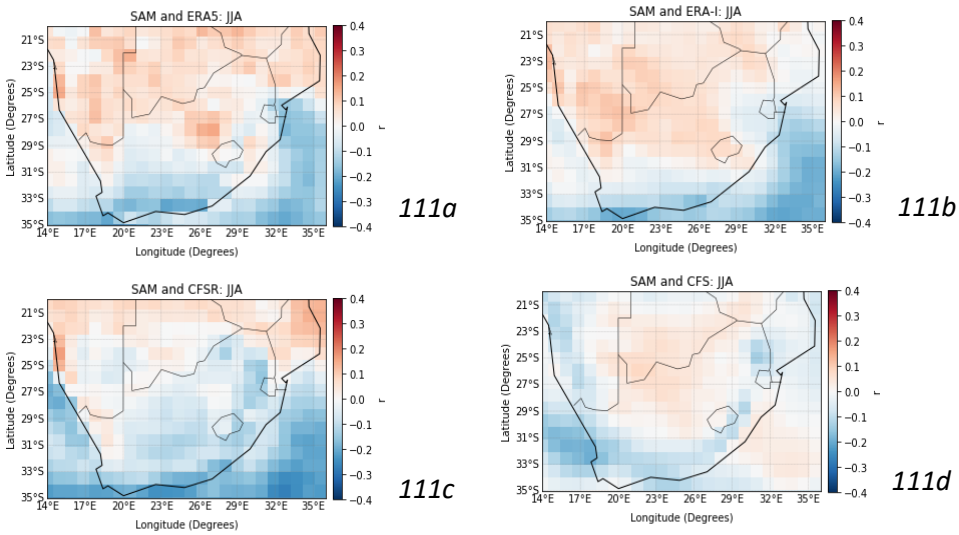
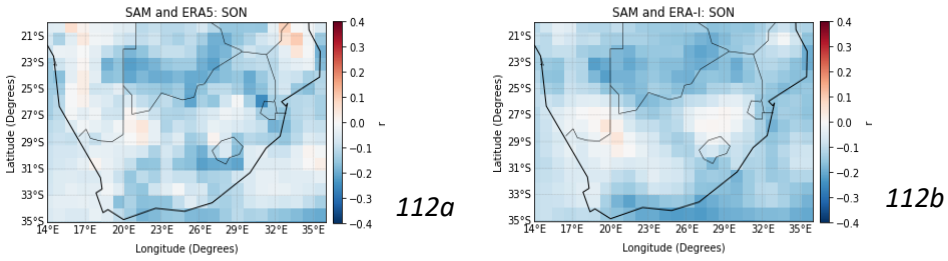


Figure 111: Plot of the Pearson Correlation Coefficient between the SAM Index and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during JJA

September – October – November

The reanalyses results show that there is a weak negative correlation over the whole country during SON (Figure 112a to 112c). The correlation figure between the SAM and the CFS wind speed data indicates that there are weak positive correlations over the southern-most and south-easterly regions of South Africa, however the remaining areas also present negative correlations (Figure 112d)



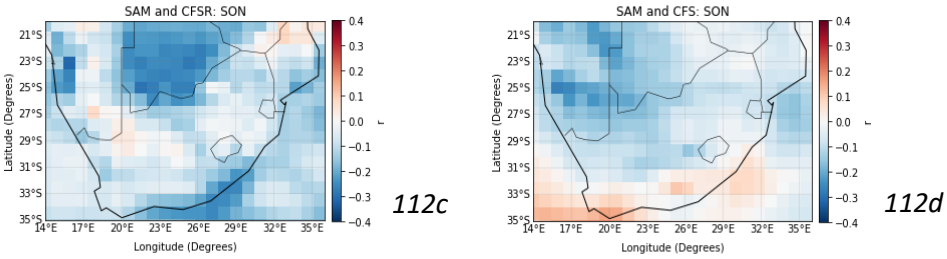


Figure 112: Plot of the Pearson Correlation Coefficient between the SAM Index and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during SON

7.6.2.3. Positive, Neutral and Negative SAM Phases

This section presents the three phases of SAM, where a threshold of -0.5 and 0.5 was used to define a negative SAM phase and a positive SAM phase respectively, and the values between -0.5 and 0.5 were used to define a neutral SAM phase. Seasonal figures (DJF, MAM, JJA and SON) have been presented for each phase.

7.6.2.3.1. Positive SAM Phase

Along the south and west coast, there is a weak positive correlation between the positive phases of SAM and wind speed (Figure 112). The results using the four datasets consistently show that there is a weak negative correlation over the interior of South Africa and weak positive correlations over the south and south western parts of the country. Given the weak correlations and inconsistent results between SAM and the wind speed data, this study was unable to determine, with certainty, how the positive phase of SAM impacts wind speeds over South Africa. This result however, does aid in answering Objective 3, by depicting that further research is required to understand the other factors that may impact the relationship between positive SAM phases and wind speed over the country.

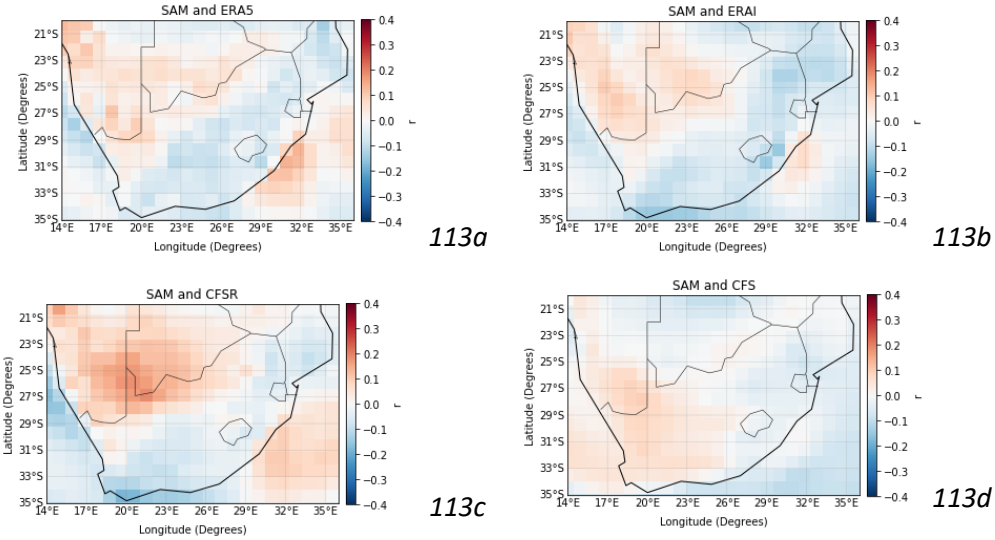


Figure 113: Plot of the Pearson Correlation Coefficient between positive SAM periods and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS

7.6.2.3.1.1. Seasonal Correlations

Section 7.6.2.3.1.1. of the results presents the seasonal correlations (DJF, MAM, JJA and SON), between the SAM Index and the three reanalysis datasets and the hindcast dataset during positive phases of the SAM.

December – January – February

The correlations between the SAM Index and the reanalysis wind datasets indicate that there is a negative correlation over the interior regions of the country, and positive correlations along the western and eastern regions of South Africa (Figure 114a to 114c). There are weak positive correlations between the SAM Index and the CFS wind speed data over the majority of the country and a negative correlation over a region towards the north-east of South Africa (Mpumalanga and Limpopo) (Figure 114d).

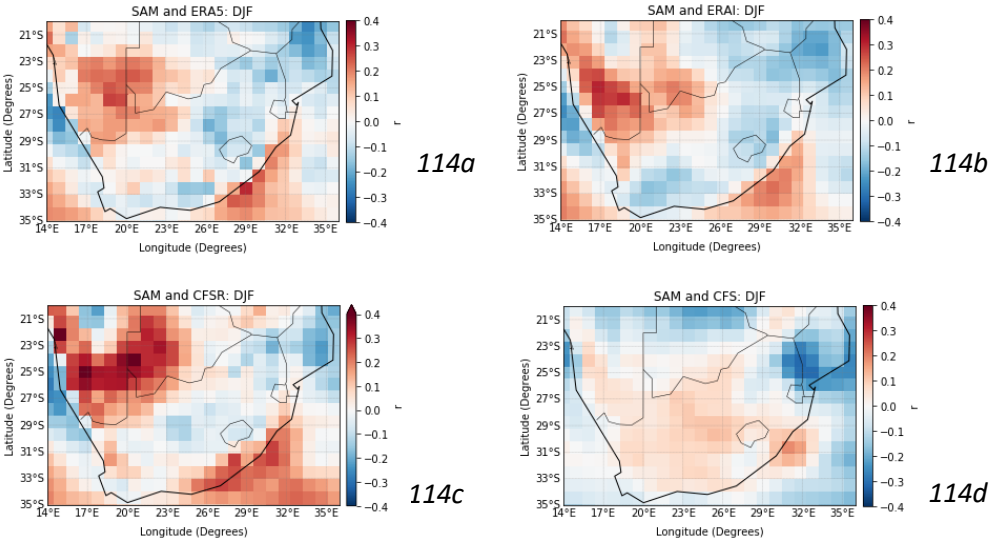
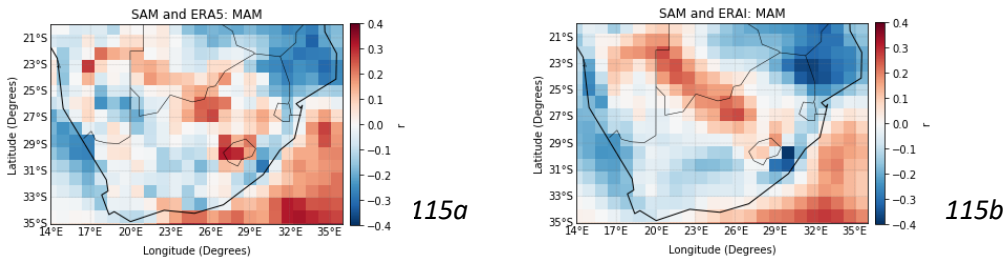


Figure 114*: Plot of the Pearson Correlation Coefficient between positive SAM periods and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during DJF*

March – April – May

There is a negative correlation between wind speed and the SAM over the northern and southern parts of the country, and positive correlations over the northern interior regions. This correlation pattern is consistent over the four wind speed datasets (Figure 115). The reanalyses correlation figures indicate that there is a positive correlation off the east coast, and a negative correlation off the west coast of South Africa (Figure 115a to 115b).



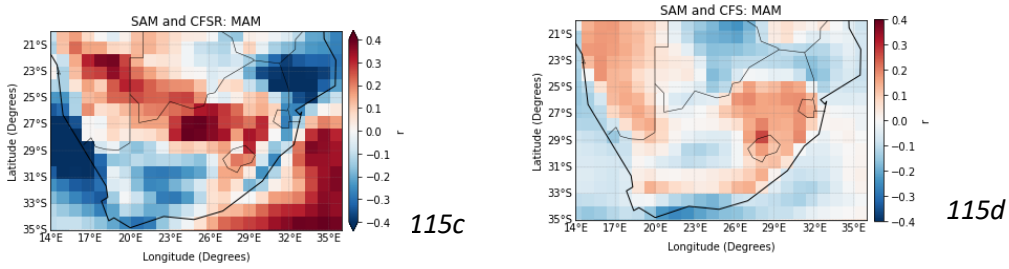


Figure 115: Plot of the Pearson Correlation Coefficient between positive SAM periods and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during MAM

June – July – August

The correlation patterns during positive SAM phases in JJA are consistent between the four wind speed datasets, indicating that the hindcast data was accurately able to detect the correlation between wind speed and positive SAM during JJA in South Africa. Over the south-western region (Northern Cape and the Western Cape), there are positive correlations and over the remaining area of the country there are negative correlations. There are strong negative correlations (-0.3 < *r* < -0.4), over the southern-most regions of South Africa (Figure 116).

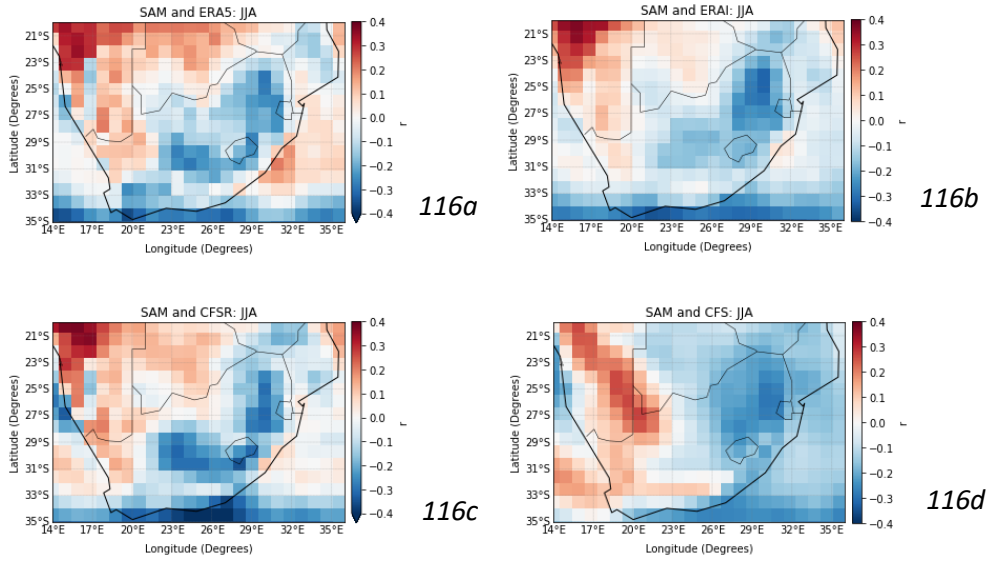


Figure 116: Plot of the Pearson Correlation Coefficient between positive SAM periods and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during JJA

September – October – November

During SON, there are inconsistent correlation patterns between the SAM Index and the four wind speed datasets. The areas of consistency between the SAM Index and the reanalyses are over the north-western regions of the country, and a north-eastern area (Kwa-Zulu Natal) of

the country (Figure 117a to 117b). The correlation between the SAM Index and the CFS wind speed data indicates that there are weak negative correlations over the northern half of the country and positive correlations over the southern half of the country (Figure 117d).

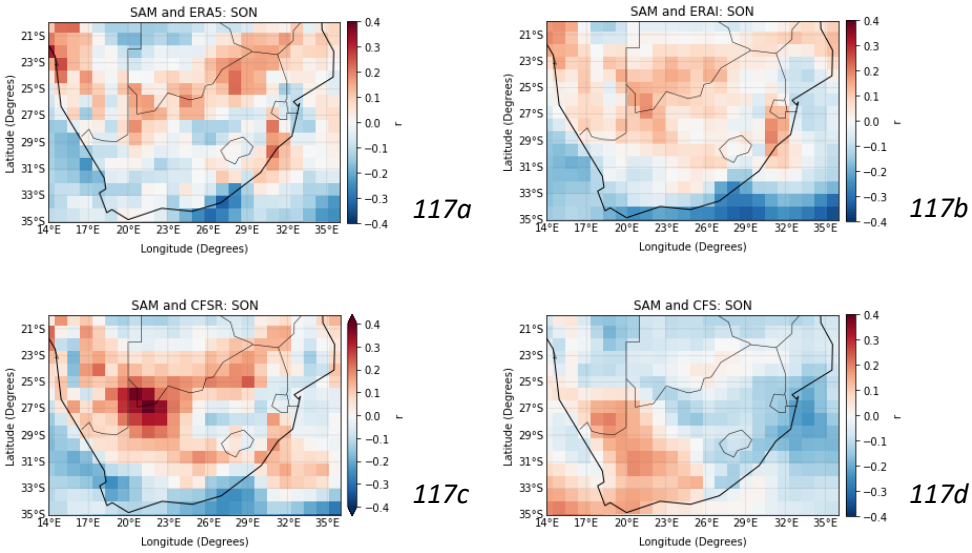


Figure 117: Plot of the Pearson Correlation Coefficient between positive SAM periods and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS during SON

7.6.2.3.2. Neutral SAM Phase

During the neutral phase of SAM (-0.5 ≤ SAM ≤ 0.5), there are positive correlations between the SAM Index and wind speed over the west coast and over the western interior. The areas of negative correlations between the SAM index and wind speed are inconsistent between the four wind speed datasets, indicating that during a neutral phase of SAM, the large scale driver might not be a good predictor of wind speed over South Africa (Figure 118).

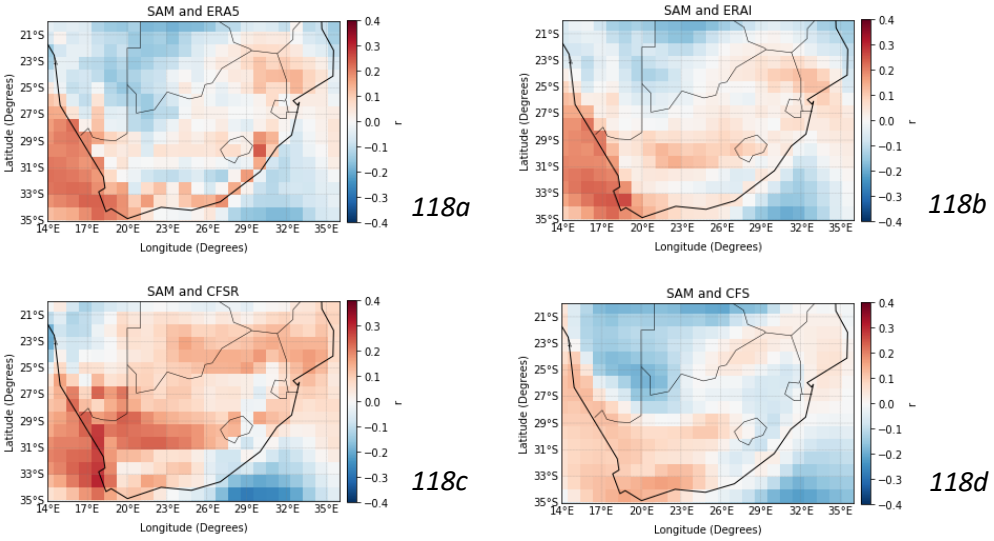


Figure 118: Plot of the Pearson Correlation Coefficient between neutral SAM phases (-0.5 ≤ SAM ≤ 0.5) and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS

The correlations between SAM and wind speed on a seasonal scale during the neutral phase SAM have not been included in this study as the results did not enhance the study nor provide any additional information. This study focuses on the positive and negative phases of SAM.

7.6.2.3.3. Negative SAM Phase

The results indicate that there are stronger correlations between the reanalyses and the SAM index during negative SAM phases compared to positive SAM phases (Figure 119a to 119c). The areas of consistent stronger negative correlations between the reanalysis wind speed datasets and the SAM Index are along the south-east coast of South Africa and over the eastern interior regions. The correlation between SAM and the hindcast wind speeds is low, where the average *r*-value is between -0.05 and 0 (Figure 119d). A negative correlation between the wind speed datasets and the SAM Index are consistent during negative phases of SAM.

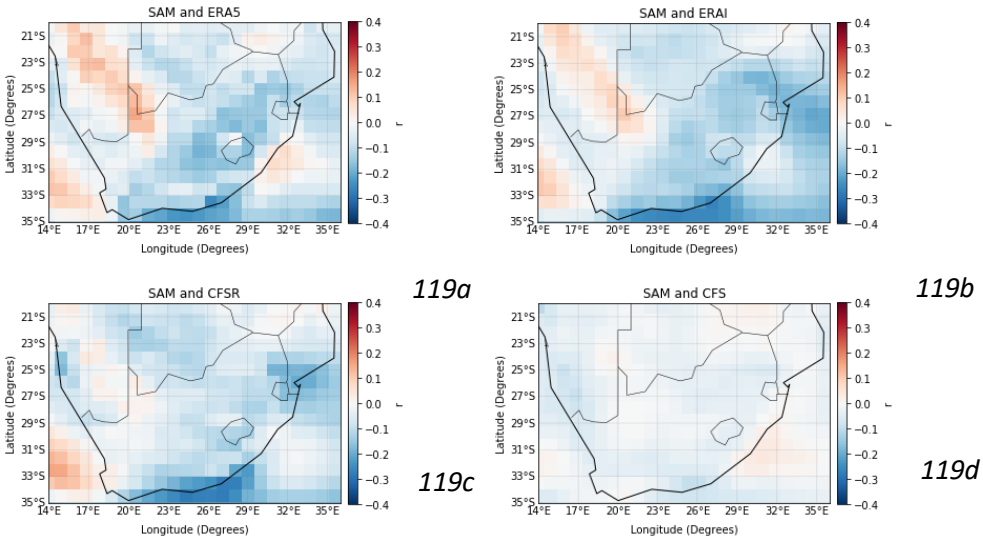


Figure 119: Plot of the Pearson Correlation Coefficient between negative SAM phases and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS

7.6.2.3.3.1. Seasonal Correlations

The results presented in this section are the seasonal correlations (DJF, MAM, JJA and SON), between the SAM Index and the three reanalysis datasets and the hindcast dataset during negative phases of SAM (SAM < -0.5).

December – January – February

The correlation pattern is similar between the SAM Index and the three reanalysis wind speed datasets (ERA5, ERA-Interim and CFSR); there are negative correlations over the interior of the country and positive correlations along the north-eastern region of South Africa

(Limpopo, Mpumalanga and Kwa-Zulu Natal), as well as over southern parts of the country (Figure 120a to 120c). There are positive correlations between the SAM Index and the CFS hindcast data over the majority of South Africa, which is inconsistent with the correlation patterns of the reanalyses (Figure 120d).

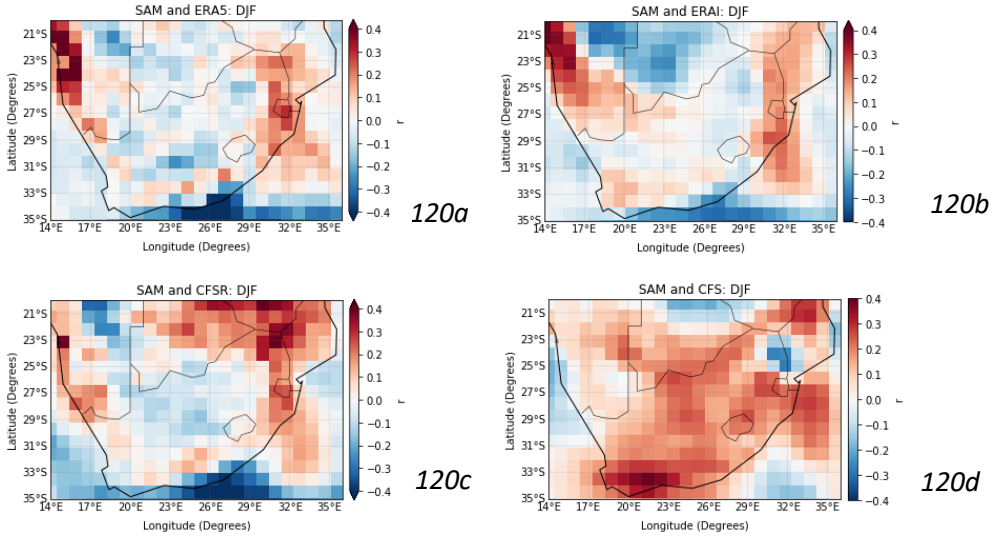


Figure 120: Plot of the Pearson Correlation Coefficient between negative SAM phases and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS in DJF

March – April – May

The Pearson Correlation Coefficient calculation between the SAM Index and the reanalyses indicates that there is a positive correlation over the interior and western regions of the country, and negative correlations over the northern and south-eastern regions. There are negative correlations offshore over the Indian Ocean and positive correlations over the Atlantic Ocean, west of South Africa (Figure 121a to 121c). The correlation differs between the SAM Index and the CFS wind speed data, compared to the reanalyses; there are negative correlations over land and positive correlations along the north-east coast (Figure 121d).

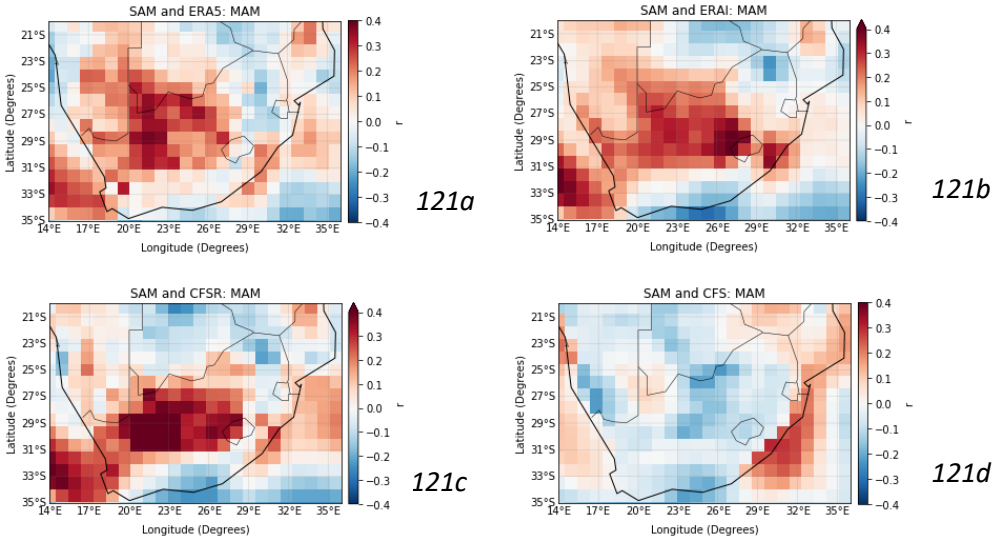


Figure 121: Plot of the Pearson Correlation Coefficient between negative SAM phases and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS in MAM

June – July – August

During JJA, there are positive correlations over the south-western regions and negative correlations over the north and north-eastern regions of South Africa between the SAM Index and the reanalysis wind speed datasets (Figure 122a to 122c). The correlation figure between the SAM Index and the CFS hindcast data, however, presents a different correlation pattern to the reanalyses figures. There is a strong negative correlation ($r \leq -4$), over the southern parts of the country and weak correlations over the northern regions (Figure 122d).

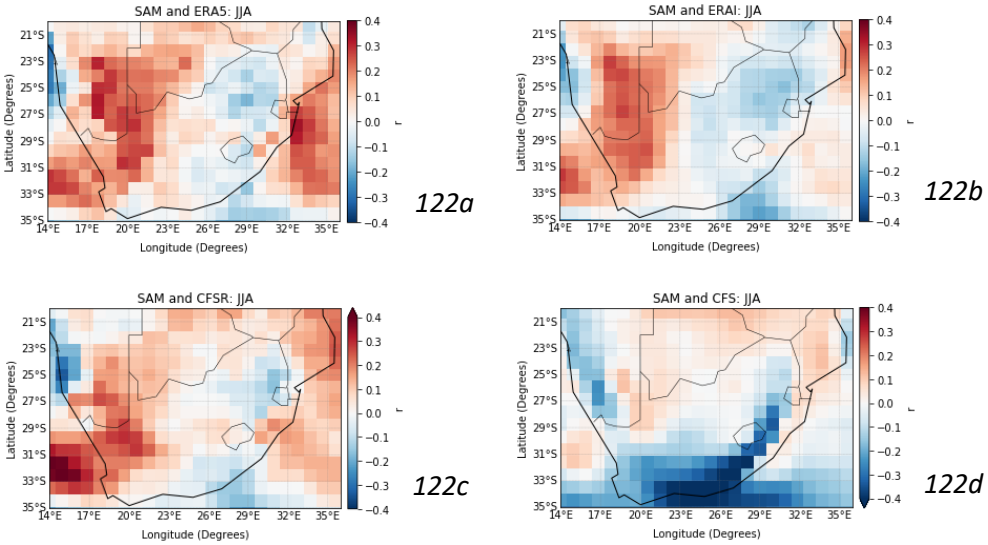


Figure 122: Plot of the Pearson Correlation Coefficient between negative SAM phases and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS in JJA

September – October – November

There are strong negative correlations ($r \leq -4$), over regions of the study area, between SAM and ERA5, ERA-Interim and CFSR. The strong negative correlations are most notable over eastern interior and east coast of South Africa (Figure 123b and 123c), and over the interior regions of the country (Figure 123a). The strong negative correlations between the SAM Index and wind speed indicate that during negative SAM phases, in SON, wind speed increases over the parts of South Africa mentioned above. The correlation between the CFS hindcast data and the SAM Index, however, does not follow this pattern, as there are weak negative and positive correlations over the study area (Figure 123d). This result indicates the forecast does not accurately capture this large-scale driver of wind over the country during SON. The Southern Annular Mode can however be used as a predictor of wind during SON, given the strong correlation values seen between the reanalyses and negative SAM events.

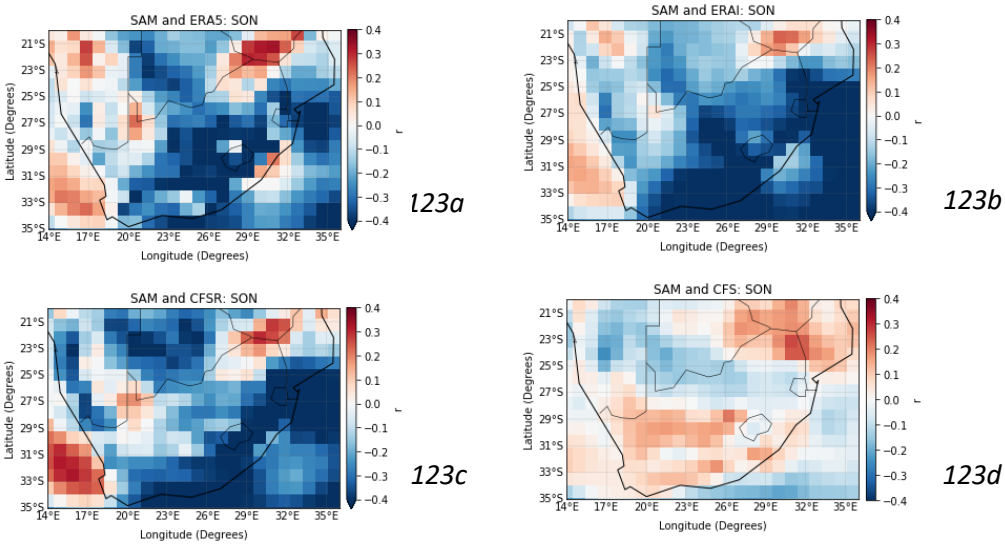


Figure 123: Plot of the Pearson Correlation Coefficient between negative SAM phases and (a) ERA5; (b) ERA-Interim; (c) CFSR; and (d) CFS in SON

During negative SAM phases, the negative correlations are generally stronger than the negative correlations during positive SAM phases and neutral SAM phases. The strongest positive correlations between SAM and wind speed can be seen during neutral phases of SAM (Table 7). This results enhances the response to Objective 3, by indicating which phase of SAM drives the predictability of wind speed.

Table 7: The maximum and minimum r-values calculated for each wind dataset, based on the data used for Figure 113, Figure 118 and Figure 119, for positive (0.5 < SAM), neutral SAM phases (-0.5 ≤ SAM ≤ 0.5) and negative SAM phases (SAM < -0.5)

	CFS	ERA5	ERA-Interim	CFSR
Positive SAM Phase				
Maximum *r*-value	0.056	0.138	0.106	0.189
Minimum *r*-value	-0.079	-0.137	-0.161	-0.182
Negative SAM Phase				
Maximum *r*-value	0.0418	0.155	0.109	0.158
Minimum *r*-value	-0.106	-0.254	-0.268	-0.282
Neutral SAM Phase				
Maximum *r*-value	0.164	0.245	0.263	0.303
Minimum *r*-value	-0.219	-0.210	-0.214	-0.307

CHAPTER 5

8. Discussion

This section provides a discussion for the evidence obtained through the analysis and report in Section 7. It addresses the following objectives:

1. To compare hindcast data to reanalysis data in order to assess the skill of the seasonal forecast, by evaluating tools and metrics being used to assess wind forecast skill.
2. To evaluate several reanalysis datasets for use as a proxy for observed wind data.
3. To understand cross-scale drivers of wind patterns in South Africa including teleconnections, synoptic and regional drivers.

Section 8 is divided into three components. The first discusses the difference between the three reanalysis datasets and the importance of using multiple reanalysis datasets (Objective 2); the second assesses the skill of the hindcast data (CFS), in comparison to the three reanalysis datasets (ERA5, ERA-Interim and CFS), based on the results (Objective 1); and, the third sub-section continues by discussing the large-scale drivers of wind over the region and the implications for seasonal wind prediction over South Africa (Objective 3).

8.1. The importance of using multiple reference datasets

Various studies have described the benefits of using ERA5 reanalysis data, compared to other reanalysis datasets (Hersbach et al., 2019) (Hersbach et al., 2020) (Belmonte Rivas and Stoffelen, 2019) (Jourdier, 2020). The ERA5 is an upgrade compared to ERA-Interim, due to the higher spatial and temporal resolution and an increased number of daily observations (Hersbach et al., 2019) (Hersbach et al., 2020). The ERA5 and ERA-Interim surface wind speeds and wind biases were assessed by comparing the ERA datasets to the Advanced Scatterometer (ASCAT) observations, and based on the study, it was seen that "ERA5 winds show a 20% improvement relative to ERA-Interim", specifically over oceans (Belmonte Rivas and Stoffelen, 2019: 831). Ramon et al., (2019) completed a study which compared surface and near-surface winds from various reanalysis models, including ERA5 and ERA-Interim, to *in situ* observations and concluded 'that the ERA5 surface winds [offered] the best agreement' to the observed data "in 35.1% of the tall tower sites on a daily time-scale" (2019: 3236). The study was conducted for wind speeds around the globe, therefore these results are not specific to South Africa. Similarly, Jourdier (2020), conducted a study over France using the MERRA-2 and ERA5 reanalysis wind datasets, and found that ERA5 "shows great improvements over previous reanalyses"; however, there were underestimations with the ERA5 wind speeds, specifically over mountainous areas (Jourdier, 2020: 65). The underestimation of wind speeds could be due to representativeness errors, where reanalysis datasets cannot capture what a wind mast is able to capture, such as the effect that topography and orography has on wind patterns over an area.

This study made use of three reanalysis datasets and demonstrates that it is necessary to use more than one reanalysis dataset when studying wind despite the sentiment in the literature cited above. Differences exist between the three reanalysis datasets (Section 7.3. and Section 7.4.) and over the majority of the study areas, the ERA5 wind speeds were lower than the ERA-Interim and CFSR wind speeds, by approximately 2 m/s at the majority of the sites. The T-Test, Kolmogorov-Smirnov Test (KS-test) and Probability Density Functions indicated that there is a statistical significant difference between the ERA5 and ERA-Interim datasets, especially over South Africa. The *p*-value between ERA5 and CFSR is higher than the *p*-value between the ERA-Interim and ERA5 datasets, specifically over the northern and southern-interior regions of South Africa. The p-value was as low as 10^{-50} in some regions between ERA5 and ERA-Interim. The *p*-value is highest between ERA-Interim and CFSR (where the p-value is greater than 10^{-2}), indicating that the statistically significant difference is the least between these two reanalysis datasets. Therefore, in order to assess the skill of a wind forecast, hindcast data must be compared to multiple reanalyses as the differences in the reanalysis datasets makes it necessary to use more than one reanalysis dataset when studying wind. A single reanalysis comparison, even if this is the ERA5, is likely to lead to incorrect skill assessments and conclusions.

This study did not make use of observed data due to the lack of available and reliable station data. Additionally, station data was not used to verify the skill of the forecasted data, in order to avoid the comparison between point data and grid data. Point-source wind data from a measurement station will reflect the effect local features such as topography and vegetation have on wind speeds that a forecast product with a grid size of 30km smooths out and cannot capture.

8.2. Assessing the skill of CFS hindcast data

Synoptic scale features can be predictors of forecast wind speed values at the seasonal scale, for example, the semi-permanent high over the interior during winter results in higher skill levels in the forecast. However, quantifying the skill of a particular wind forecast depends on the region, the season and the “observed” data. In this study, the highest Pearson Correlation Coefficient values can be seen between the CFSR and CFS wind speed datasets; this is because both datasets are the product of the same model: the Climate Forecast System model, which was developed by the National Center for Environmental Prediction (NCEP) (NOAA, n.d). Similarly, the lowest RMSE values occur between the CFSR and CFS data, and the highest values between the ERA5 and CFS data. These results reiterate the necessity to use multiple reanalysis wind datasets, as the skill assessment outcomes depend on which reanalyses are used to a large measure. Within the seasonal forecast, wind speed correlations are weaker at shorter time scales (daily) compared to longer time periods (monthly and seasonal) (Section 7.5.1.). This is because wind speeds have a higher variation and are more erratic at shorter

timescales, making it more difficult to predict. Forecasting wind speeds over a longer time period will allow for patterns to emerge within the data.

Synoptic scale circulations such as the South Atlantic High Pressure (SAHP) is responsible for the south and south-easterly winds that are seen over Cape Town, Cape Point and Saldanha during summer. The strongest winds are experienced on the outer edges of the anticyclone (Section 5.4.1.), which is why the time series figures show that the regions in the Western Cape experience peak wind speeds during austral summer when the system is at its most poleward extent (Kruger et al., 2010) (Lennard, 2019) (Nchaba et al., 2016). The seasonal movement of the SAHP over particularly the Western Cape results in 'prevailing winds' that are highly predictable (Fawcett et al., 2007: 21).

Seasonally, there is a distinct change in correlation patterns between the hindcast data and the reanalysis data between December – January – February and June – July – August (Section 7.5.1.1.). In DJF, over the northern areas of South Africa the correlation values indicate a weak correlation between the hindcast data and the reanalyses datasets, specifically between CFS and the ERA datasets (Figure 74). The northern parts of the country (Limpopo, regions of Mpumalanga and regions of the North West) experiences summer convection and in this unstable environment it is difficult for a forecast model to accurately predict the wind speeds. During austral summer, low-pressure troughs extend over the interior of South Africa and is responsible for the wind directions over Johannesburg, Bloemfontein and Polokwane. The trough is responsible for an unstable atmosphere, over the interior of the country and therefore there are weaker correlation values. Increased cloud cover results in "large uncertainty in forecast since numerical weather models have different opinions and accuracies when predicting cloud formation and movement" (Wright et al., 2019: 108).

During JJA, the weaker correlation values over the southern parts of South Africa result from the low pressure systems (cold fronts) whose transient nature result in low forecast skill generally. Past studies have shown low skill in rainfall forecasts over the Western Cape region as a result of these transient systems (Landman et al., 2005). These mid-latitude cyclones pass over the southern parts of the country which introduce vertical instability and an increase in wind variability, and can therefore affect the ability of a model to forecast wind speeds accurately (Wright et al., 2019). Over the interior of the country a high pressure system produces a stable environment resulting in higher forecast skill during JJA.

Overall, during the four seasons, the lowest RMSE values are over the northern areas (Limpopo, North West, Mpumalanga), and the interior of South Africa (Figure 78 and 79); this is consistent with the Pearson Correlation Coefficient results (Figure 71 to 73), which indicates that the predictability of wind over the northernmost areas, and specific interior regions of the country are relatively high. Previous studies have found that maximum skill associated with of a rainfall forecast is over "northeastern and central-western interior" regions of South

Africa (Landman et al., 2012: 500). The majority of the coastal regions experience higher wind speeds than the inland regions. The increase in wind speeds is due to the difference in pressure between the land and the ocean, this results in a strong pressure gradient force and subsequently higher wind speeds (Nhesvure, 2020), and can result in a more accurate and skilful seasonal forecasts over coastal regions.

Spatially there are areas of weak correlation that are consistent between the three reanalysis datasets and the hindcast data. This indicates that there are circulation characteristics that are associated with low correlation values, and therefore hinders the ability of a model to forecast wind accurately. One of the areas which has consistently weak correlation values is over the western interior of South Africa. It can be seen that this is where south-easterly winds converge with north-westerly winds (Figure 124). The converging winds have resulted in weak correlations because the wind speeds over this area are variable and therefore they are difficult to predict accurately (Wright et al., 2019).

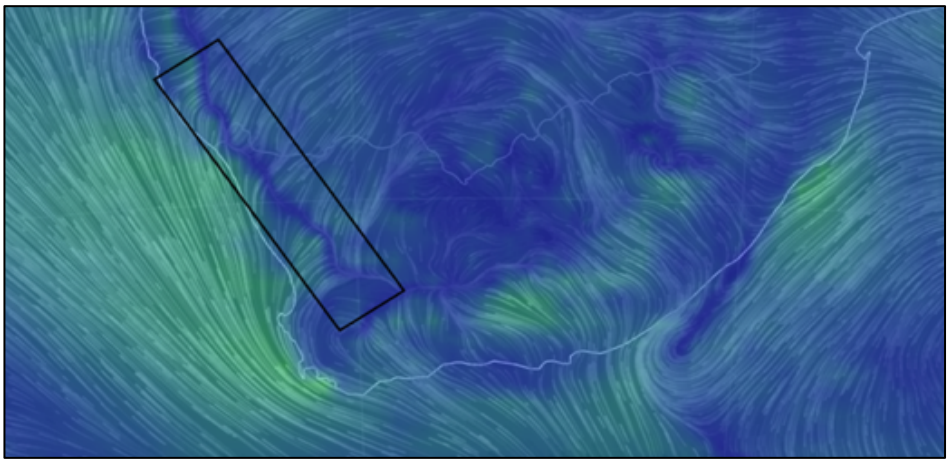


Figure 124: Wind patterns over South Africa and the surrounding oceans. The black box indicates the region of generally weak correlation between the hindcast and reanalysis data (WindFinder, 2020)

8.3. Large scale drivers of wind

This section discusses the El Niño Southern Oscillation (ENSO) and the Southern Annular Mode (SAM) controls of wind over South Africa and how this affects predictability at the seasonal scale. These are two of the dominant large-scale drivers of variability in the South African climate, including wind, and understanding the link between them and wind circulation over South Africa could result in a better prediction of wind and more skilful wind forecasts on a seasonal scale (Nchaba et al., 2016).

8.3.1. El Niño Southern Oscillation

During a neutral ENSO phase, the synoptic circulation features over South Africa during the austral summer months include a low pressure system over the interior of the country, a

ridging South Atlantic High Pressure system over the south-west coast and the South Indian High Pressure system towards the south-east of South Africa. ENSO has the strongest impacts over South Africa during austral summer and during La Niña and El Niño events, these pressure systems are perturbed through modifications of the Walker and Hadley circulation (Rouault et al., 2019).

8.3.1.1. El Niño

During an El Niño event, the sea surface temperature of the Pacific Ocean and the Indian Ocean increases which results in the poleward motion of air. Over Southern Africa, the air decreases in temperature and subsides and therefore a high pressure over South Africa develops. The increased surface pressure inhibits convection (Rouault et al., 2019) and is often associated with decreased summer rainfall over the interior regions of the country. During stronger El Niño events, stronger high pressure systems are present over South Africa (Kruger, 1999). The low pressure systems which are usually extended over regions of South Africa are weakened, and the pressure over the interior of South Africa increases resulting in a summer high pressure anomaly over Southern Africa (Hoell et al., 2017) (Kruger, 1999). The weakening of the low pressure systems, and the increased pressure over South Africa results in a lower pressure gradient force, which could potentially cause a decrease in wind speeds (Table 7). This statement correlates to the results presenting the correlations between strong El Niño (1.2 < MEI) and wind speed over the northern regions of the country (parts of the Free State, Gauteng, Mpumalanga, the North West, Kwa-Zulu Natal and the Northern Cape) (Section 7.6.1.4.1.). There is a negative correlation between El Niño and wind speed over the southern regions of South Africa (Western Cape) (Section 7.6.1.3.1.), specifically, in DJF (Figure 101) and SON (Figure 104) which is a result consistent between the reanalysis products. This finding is consistent with the results stated in the literature by Nhesvure (2020), Rouault et al., (2010) and Philippon et al., (2012); specifically when Rouault (2010) stated that during austral summer along the coast by the Western Cape, "weaker (stronger) upwelling favourable wind is present during El Niño (La Niña)" (2010: 243). Despite the evidence that has been presented, Rouault et al., (2019), stated that "El Niño [events] have different signatures is southern Africa and at regional scale", and therefore the impacts that ENSO has on South Africa is not always constant nor can it be solely relied on.

8.3.1.2. La Nina

During La Niña events, Nhesvure (2020) and Rouault et al., (2010) show that the south-easterly wind speeds over the Western Cape increase, and this could be due to the southward movement of the South Atlantic High Pressure. This can be seen in the results where moderate La Niña events have been presented (MEI < -0.5) (Section 7.6.1.3.3.), however the "strong La Niña" results (MEI < -1.2) (Section 7.6.1.4.2.), do not reflect this result.

La Niña phases result in a weakening of the high pressure systems to the west of South Africa and a deepening of troughs over interior regions of South Africa (Kruger, 1999). The weakening of the high pressure system to the west of the country, would result in a lower land-sea pressure gradients over the western regions of South Africa and can therefore explain the positive correlations between La Niña and wind speed (indicating lower wind speeds) over regions of South Africa (Figure 101 to 104) (Figure 106). Over the northern regions of South Africa, there is a negative correlation between MEI and wind speed (Section 7.6.1.3.3.), as well as over the eastern regions (Section 7.6.1.4.2.). This could be due to the deepening of an extended sub-tropical low pressure trough, which causes an increase in the pressure gradient and results in higher wind speeds over the region. Southern Africa's climate is also impacted by tropical and sub-tropical low pressure cells during the austral summer months. During a La Niña, tropical lows are shifted further south, and this can result in increased convection and precipitation over the country (Howard et al., 2019), and could potentially increase the pressure gradient over the northern and eastern regions of South Africa. The strong negative correlation over the eastern regions seen during strong La Niña phases, indicates higher wind speeds which could be a result of the land-sea pressure gradient which is formed due to the extended low pressure system over land and the South Indian High Pressure system (Figure 106a to 106c).

In terms of wind speed, the impact of La Niña is heterogeneous and results in higher or lower wind speeds over different parts of South Africa. Over the majority of areas across South Africa, the results show that the correlations between MEI and wind speed are stronger during strong La Niña events. The results suggest that La Niña events are a good predictor of wind speed, and that there is agreement between the three reanalysis datasets over certain regions of the country, which suggests further that a La Niña event can be used to predict wind speeds with a certain level of robustness (Table 8). Similarly, the strength of the correlation between wind speed and La Niña increased as the threshold for defining a La Niña event increased, this indicates that the stronger the La Niña event, the more impact La Niña will have on wind speed over South Africa. This output contributes to answering Objective 3 which aims at understanding the cross-scale drivers of wind patterns in South Africa. In addition, the results can enhance decision-making by monitoring ENSO events and being able to roughly predict the wind speeds associated with the El Niño or La Niña events. This method has historically been used by farmers, who use El Niño events to predict rainfall patterns. Similarly, there is a level of uncertainty associated with the forecast abilities.

8.3.1.3. Summary

The weakest correlations between MEI and wind speed is found during El Niño events in all reanalysis products and even weaker correlations between MEI and the CFS (Section 7.6.1.4). This indicates that El Niño is a poor predictor of wind speeds over South Africa generally.

There is a stronger link between wind speeds and ENSO during a La Niña phase, over South Africa, compared to an El Niño phase (Section 7.6.1.4).

The results indicate a lack of agreement between the results presented by the CFS hindcast and the reanalyses. Based on Section 7.6.1.3., the CFS hindcast performs with a certain level of skill during December – January – February, however the performance of the forecast is not consistent. This indicates that the skill of the seasonal wind forecast depends on the season and the region (Table 8).

Table 8: A summary table of how the different phases of ENSO impact wind speed over specific regions. The wind speed for each dataset, is the average wind speed during the moderate El Niño (0.5 < MEI), La Niña (MEI < -0.5) and Neutral (-0.5 ≤ MEI ≤ 0.5) phases for specific grid points. The correlation range has been determined based on the results in Section 7.6.1.3. The dataset agreement are the wind speed datasets that show the same correlation pattern.

Phase	Wind speeds (m/s)				Correlation Range	Dataset Agreement
	CFS	ERA5	ERA-I	CFSR		
Cape Town						
El Niño Phase	5.413	2.446	3.944	4.233	$0.05 \leq r \leq 0.1$	ERA5, ERA-I, CFSR
La Niña Phase	5.628	2.700	4.215	4.292	$-0.15 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	5.407	2.475	3.998	4.254	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Cape Point						
El Niño Phase	7.144	6.325	6.852	6.129	$0.05 \leq r \leq 0.1$	ERA5, ERA-I, CFSR
La Niña Phase	7.368	6.685	7.142	6.335	$-0.2 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	7.147	6.449	6.974	6.257	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Port Elizabeth						
El Niño Phase	4.359	2.232	3.798	4.289	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.398	2.337	3.755	4.247	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	4.348	2.186	3.774	4.256	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Mossel Bay						
El Niño Phase	5.927	5.141	5.365	5.452	$-0.05 \leq r \leq 0.1$	ERA5, ERA-I, CFSR, CFS
La Niña Phase	6.044	5.412	5.505	5.521	$-0.15 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	5.927	5.197	5.415	5.493	$-0.15 \leq r \leq -0.01$	ERA5, ERA-I, CFSR, CFS
Port Nolloth						
El Niño Phase	4.706	3.376	4.072	4.132	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.990	3.541	4.306	4.161	$0 \leq r \leq 0.2$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	4.679	3.289	4.002	3.988	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Saldanha						
El Niño Phase	6.199	5.435	5.006	5.459	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
La Niña Phase	6.564	6.072	5.499	5.786	$-0.05 \leq r \leq 0.1$	ERA5, ERA-I, CFSR, CFS
Neutral Phase	6.192	5.496	5.037	5.405	$-0.2 \leq r \leq -0.1$	ERA5, ERA-I, CFSR, CFS
Mabibi						
El Niño Phase	4.331	4.007	5.412	5.027	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.437	4.026	5.399	5.005	$-0.1 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS

Neutral Phase	4.265	3.986	5.426	4.952	-0.2 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Pietermaritzburg						
El Niño Phase	3.138	1.401	2.258	2.585	-0.1 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
La Niña Phase	3.029	1.437	2.219	2.432	-0.2 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
Neutral Phase	3.125	1.380	2.237	2.526	-0.25 ≤ r ≤ -0.1	ERA5, ERA-I, CFSR, CFS
Johannesburg						
El Niño Phase	2.849	2.059	2.852	2.929	0 ≤ r ≤ 0.1	ERA5, ERA-I
					-0.1 ≤ r ≤ 0	CFSR, CFS
La Niña Phase	2.810	2.069	2.814	2.838	-0.2 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Neutral Phase	2.831	2.067	2.823	2.898	-0.15 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Bloemfontein						
El Niño Phase	3.186	2.314	3.357	3.188	-0.05 ≤ r ≤ 0.05	ERA5, ERA-I, CFSR, CFS
La Niña Phase	3.281	2.366	3.426	3.212	-0.05 ≤ r ≤ 0.1	ERA5, ERA-I, CFSR, CFS
Neutral Phase	3.169	2.292	3.301	3.148	-0.25 ≤ r ≤ 0-0.1	ERA5, ERA-I, CFSR, CFS
Polokwane						
El Niño Phase	2.381	1.851	2.797	2.315	0 ≤ r ≤ 0.1	ERA5, ERA-I
					-0.05 ≤ r ≤ 0	CFSR, CFS
La Niña Phase	2.506	1.872	2.795	2.404	-0.2 ≤ r ≤ -0.1	ERA5, ERA-I, CFSR, CFS
Neutral Phase	2.339	1.836	2.746	2.252	-0.15 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Upington						
El Niño Phase	3.757	3.027	3.899	3.597	-0.1 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
La Niña Phase	3.777	2.987	3.819	3.649	-0.05 ≤ r ≤ 0.05	ERA5, ERA-I, CFSR, CFS
Neutral Phase	3.742	2.994	3.817	3.569	-0.2 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Butterworth						
El Niño Phase	3.555	1.801	2.857	2.938	-0.1 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
La Niña Phase	3.472	1.751	2.689	2.715	-0.05 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
Neutral Phase	3.533	1.750	2.833	2.911	-0.1 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
Brandvlei						
El Niño Phase	4.267	3.824	4.657	3.923	-0.1 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.407	3.772	4.664	4.033	0 ≤ r ≤ 0.2	ERA5, ERA-I, CFSR, CFS
Neutral Phase	4.255	3.739	4.578	3.864	-0.15 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Hanover						
El Niño Phase	4.135	3.126	4.288	3.809	-0.1 ≤ r ≤ 0	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.240	3.012	4.230	3.947	0.05 ≤ r ≤ 0.1	ERA5, ERA-I
					-0.1 ≤ r ≤ 0	CFSR, CFS
Neutral Phase	4.126	3.098	4.195	3.730	-0.2 ≤ r ≤ -0.05	ERA5, ERA-I, CFSR, CFS
Beaufort West						
El Niño Phase	4.024	3.165	3.983	3.611	-0.05 ≤ r ≤ 0.05	ERA5, ERA-I, CFSR, CFS
La Niña Phase	4.068	3.298	3.863	3.566	0 ≤ r ≤ 0.1	ERA5, ERA-I
					-0.15 ≤ r ≤ 0	CFSR, CFS
Neutral Phase	4.006	3.180	3.942	3.569	-0.15 ≤ r ≤ 0.05	ERA5, ERA-I, CFSR, CFS

The seasonal correlations using the whole MEI dataset indicate that the CFS correlations with MEI are larger in magnitude than those of the reanalysis data with MEI (Figure 86). This

suggests that the forecast model seems to respond more to ENSO, compared to the reanalyses models. This is likely because seasonal forecast models are "tuned" to ENSO; for example Landman and Beraki (2012) indicated that rainfall forecast models were more skilful during ENSO phases, compared to neutral phases and this is likely due to the tuning of the forecast model.

8.3.2. The Southern Annular Mode

There is a stronger negative correlation between negative SAM phases and wind speed compared to positive SAM phases and wind speed (Table 7). During a negative SAM phase the strong belt of westerly winds moves equatorward and therefore closer to South Africa, which could result in stronger winds, low pressure systems and unstable conditions to occur over the country. During a negative SAM phase, troughs extend across the country which results in an increased pressure gradient between the low pressures over the interior and the high pressure cells off the coasts of South Africa, resulting in increased wind speeds over the majority of the country. The increase in wind speeds, however, are approximately only 0.1 m/s greater than wind speeds during a positive SAM phase (Table 9).

A positive SAM phase indicates that the westerly winds have moved south and this should result in westerly winds over South Africa being weaker than normal. During a positive SAM phase, there are high pressure systems over South Africa, which restricts low pressure systems from forming over the country and therefore a weaker pressure gradient should be present, resulting in wind speeds decreasing. The results presented in Section 7.6.2.3.1. indicate that the correlations between the SAM Index during positive SAM phases and wind speed are weak, therefore this study was unable to determine whether wind speeds decrease over South Africa during positive SAM phases.

Table 9: A summary table of how the different phases of SAM impact wind speed over specific regions. The wind speed for each dataset, is the average wind speed during positive and negative phases of SAM for a specific grid point. The correlation range has been determined by Figures 113 and 119. The dataset agreement indicates the wind speed datasets that show the same correlation pattern.

	Wind speeds (m/s)				Correlation Range	Dataset Agreement
	CFS	ERA5	ERA-I	CFSR		
Cape Town						
Positive SAM	5.478	2.519	4.048	4.258	$-0.1 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	5.466	2.532	4.026	4.277	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Cape Point						
Positive SAM	7.207	6.531	7.005	6.237	$-0.1 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	7.209	6.399	6.936	6.255	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Port Elizabeth						

Positive SAM	4.359	2.248	3.772	4.243	$-0.05 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.393	2.254	3.811	4.309	$-0.2 \leq r \leq -0.1$	ERA5, ERA-I, CFSR
Mossel Bay						
Positive SAM	5.952	5.243	5.401	5.460	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Negative SAM	5.995	5.221	5.456	5.530	$-0.2 \leq r \leq -0.1$	ERA5, ERA-I, CFSR, CFS
Port Nolloth						
Positive SAM	4.781	3.333	4.077	4.065	$0 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.770	3.417	4.127	4.096	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Saldanha						
Positive SAM	6.306	5.601	5.150	5.511	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	6.286	5.574	5.124	5.485	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Mabibi						
Positive SAM	4.335	3.999	5.386	4.977	$-0.05 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.381	4.033	5.483	5.025	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Pietermaritzburg						
Positive SAM	3.093	1.404	2.223	2.505	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Negative SAM	3.145	1.414	2.277	2.572	$-0.05 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Johannesburg						
Positive SAM	2.824	2.046	2.810	2.875	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	2.878	2.099	2.878	2.937	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Bloemfontein						
Positive SAM	3.216	2.306	3.341	3.154	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	3.223	2.346	3.395	3.231	$-0.2 \leq r \leq -0.1$	ERA5, ERA-I, CFSR
Polokwane						
Positive SAM	2.398	1.845	2.756	2.314	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	2.423	1.874	2.810	2.312	$-0.15 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Upington						
Positive SAM	3.769	3.007	3.856	3.592	$0 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	3.773	3.034	3.873	3.623	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Butterworth						
Positive SAM	3.517	1.773	2.793	2.836	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Negative SAM	3.563	1.804	2.850	2.948	$-0.2 \leq r \leq -0.1$	ERA5, ERA-I, CFSR, CFS
Brandvlei						
Positive SAM	4.318	3.770	4.620	3.915	$0 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.303	3.828	4.669	3.948	$-0.1 \leq r \leq 0$	ERA5, ERA-I, CFSR, CFS
Hanover						
Positive SAM	4.162	3.064	4.237	3.796	$-0.05 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.188	3.140	4.310	3.860	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS
Beaufort West						
Positive SAM	4.037	3.214	3.921	3.572	$-0.1 \leq r \leq 0.05$	ERA5, ERA-I, CFSR, CFS
Negative SAM	4.039	3.218	4.007	3.626	$-0.1 \leq r \leq -0.05$	ERA5, ERA-I, CFSR, CFS

Over the past decades, the Southern Annular has displayed a trend of an increased frequency of positive phases, which has the potential to impact the climate systems in the Southern

Hemisphere (Fogt and Marshall, 2020) (Figure 107). Highest correlations are found during negative SAM so the trend towards positive SAM would result in a decreased control of wind speeds over South Africa (Figure 119). Correlation strengths vary between seasons and would impact the skill of the seasonal forecast. Section 7.6.2.3 indicates that during September – October – November (SON), the correlations between wind speed and SAM are stronger compared to DJF, MAM and JJA, and that the correlation patterns are consistent between the three reanalysis datasets, however not the hindcast dataset. Over the eastern and interior regions of South Africa, during negative SAM phases, there is a negative correlation between the SAM index and wind speed in all three reanalysis datasets.

These results show that the SAM index can be used as a predictor of wind speeds during SON, however, the CFS does not forecast the same pattern as the reanalyses. The CFS hindcast data is a constant outlier with regards to the wind speed correlations and the SAM Index. During negative SAM phases (Section 7.6.2.3.3.), the CFS correlation figure presents weaker correlations than the corresponding reanalyses figures, and the conclusion can be made that during SAM events, the CFS hindcast does not perform with a high level of skill. The fact that CFS does not capture the SAM forcing of seasonal wind speeds indicates that more research should be focused on developing forecast systems that capture other large-scale drivers like SAM, apart from ENSO.

The results in Section 7.6. indicate that during a negative SAM phase and a La Niña event, wind speeds are on average higher over the east, north-eastern and interior regions of South Africa. A forecast system that predicts higher wind speeds over these areas during a negative SAM event and/or a La Niña event, suggests that the forecast is more reliable compared to forecasts when these drivers are absent. More confidence can be attached to the forecast if the forecast indicates that wind speed increases during both events over specific regions across South Africa.

CHAPTER 6

9. Conclusion

This thesis has been a first step in assessing the predictability of wind at the seasonal scale over South Africa. Unlike temperature and rainfall, seasonal wind predictability has not been well studied so the aim of this study has been to assess the skill of seasonal wind forecasts in different regions across South Africa and the source of the skill when present. Regionally, South Africa experiences different regional wind characteristics that are governed by synoptic scale circulations which are in turn a function of global and hemispheric scale drivers. Therefore the ability of a seasonal forecast model to predict regional wind speeds skilfully depends on its ability to capture these cross scale relationships. Through addressing the

objectives of this study some conclusions about the regional predictability of wind at the seasonal scale across South Africa in terms of location, time of year, skill metric and atmospheric drivers can be drawn.

There are no uniformly recognised measures of skill for wind forecasts. This study used the Pearson Correlation Coefficient to calculate the Correlation Coefficient (*r*) of wind speeds between the CFS and three reanalysis products for each grid cell over the study period and provided an understanding of where the forecast has skill, as *r* highlights skill patterns in the data, rather than the actual wind speed values only. Additionally, the Root Mean Square Error calculations supported the Pearson Correlation Coefficient results, and enhanced the robustness of the results by presenting similar results as the Pearson Correlation Coefficient calculation. This study can recommend the use of Pearson Correlation Coefficient calculation, along with an additional skill metric to help corroborate the results, in order to identify similar patterns between the forecast and reanalysis datasets and to measure the skill of a wind forecast.

The Climate Forecast System hindcast has skill, in particular regions of South Africa at particular times of the year. According to the Pearson Correlation Coefficient calculation (high *r*-values) and the Root Mean Square Error calculation (low RMSE values), over the northern regions of South Africa during March-April-May and June-July-August (the r-values range between 0.5 and 0.7 and the RMSE values range between 0 and 0.5, where the maximum value is 3), the CFS hindcast performs well. Based on the stronger correlations, where the *r*-values equal ~0.5, between the three reanalysis datasets and the hindcast data, during the austral summer months (DJF), the CFS hindcast has higher skill over the interior and south-western regions of the country (parts of the Northern Cape, Western Cape, Free State and North West). There is dataset agreement between the three reanalyses indicating that there is a positive correlation between the reanalyses and hindcast data. This indicates that the skill of the CFS model is partly a function of the stability of the atmosphere.

Multiple reanalysis datasets should be used for the evaluation of wind forecasts and quantification of wind forecast skill, but also in larger wind-orientated studies (Nchaba et al., 2018). Reanalysis data was used as a proxy for observed data for this study; however there were statistically significant differences between the reanalysis datasets. The statistical analysis (t-test and ks-test), indicated that over certain regions in South Africa, the *p*-value was as low as 10^{-50} which highlighted the difference between the reanalyses wind speed datasets. Using only one reanalysis product in such a study would likely lead to erroneous skill levels being recorded. The Pearson Correlation Coefficient between the CFS hindcast wind speed dataset and the three reanalysis wind speed datasets highlighted the importance of using more than one reanalyses, where the *r*-value between ERA5 and CFS averaged at approximately 0.4, and the *r*-value between CFSR and CFS averaged at approximately 0.8. Using an ensemble of reanalyses as proxy for observational data is not only specific to wind

and it is becoming more common as seen in skill assessments of precipitation and temperature forecasts (Bosilovich et al., 2008).

Large-scale teleconnections (ENSO and SAM) are drivers of predictability of wind speeds in different regions of the country, specifically, over the parts of the eastern, north-eastern and interior regions of South Africa. The study concludes that La Niña is a useful predictor for seasonal wind speeds over South Africa and if a forecast predicts higher than normal wind speeds over the above-mentioned regions during a La Niña, there is a high probability that the forecast would be correct. The La Niña phase of ENSO drives predictability in the eastern, north-eastern and interior regions of the country, where the *r*-value over certain regions was < -0.4, and predictability depends on the strength of the La Niña event. La Niña results in a deepened trough over the country that likely creates a stronger pressure gradient between the interior low and oceanic highs and results in higher wind speeds.

The SAM facilitates wind speed predictability during September – October – November when the SAM is in its negative phase (*r*-values decrease to < -0.4). However, this may change if the observed positive trend in SAM continues which could result in fewer negative SAM phases and lower wind speed predictability in South Africa. This may, however, be a moot point given that currently the CFS does not capture the SAM forcing of seasonal wind speeds even though the hindcast data does capture the ENSO relationship; this highlights an area for improvement in the CFS hindcast wind speed dataset.

In conclusion, seasonal predictability of wind as described in this thesis suggests that under certain large-scale forcing conditions, seasonal wind forecasts can contribute to the operational planning of wind energy producers across the country as well as other seasonal planning systems such as fire and coastal defence. Further studies of wind predictability in South Africa will help promote the use of renewable energy in the country, decrease the reliance on fossil fuels for power and help the country move towards sustainable development in the energy sector and the emergence of a green, low-carbon economy.

10. Areas for Further Study

This study aimed to assess the skill of seasonal wind forecasts over South Africa and to highlight the importance and benefits of having a skilful forecast. There are assumptions that has had to be made when studying wind. An assumption that was made in this study was the reanalysis data is a usable proxy for observational data. The reanalysis data and the hindcast data results in the spatial resolution is low. This results in a general understanding of the wind patterns over South Africa, rather than assessing the skill of a forecast at a specific point or site. Wind farms, for example, measure the wind speed at a specific point, which accounts for topography; model data, however is unable to replicate the high resolution of observational data.

There are four fundamental areas that would require further study in order to best utilise seasonal wind forecasts. This section will provide the following recommendations for future studies as the recommendations were out of the scope of this study.

10.1. Skill Measures

The use of appropriate skill measures for seasonal wind forecasts should be given consideration because classical skill measures are often not appropriate. Based on this study, the use of the Pearson Correlation Coefficient calculation and the Root Mean Square Error calculation serve as adequate tools to assess the skill of a wind forecast. Alternatively, skill measures that do not require a binary-outcome based variable, could be beneficial in the assessing the skill of wind forecasts. The use of multiple skill measures have the ability to increase the robustness of the results and corroborate key findings, as the RMSE results corroborated the Pearson Correlation Coefficient results.

10.2. The Link between Large-Scale Drivers to Synoptic-Scale Driver Variability

This study has indicated that there are large-scale drivers of wind in South Africa that are not understood thoroughly nor have been studied extensively. A better understanding of the link between ENSO and SAM to synoptic scale driver variability will enhance the ability to forecast wind over South Africa. This study has demonstrated predictability in seasonal wind forecasts through the large-scale drivers, however the link from the large-scale drivers to synoptic-scale drivers is less clear. The CFS does not capture the SAM forcing of seasonal wind speeds and therefore this study can recommend that more research should be focused on developing forecast systems that capture other large-scale drivers like SAM, rather than directing all focus on ENSO. To advance seasonal wind predictability over South Africa an integrated approach should be developed that addresses circulation patterns and drivers at all scales. Additionally, other forecast systems, like the ECMWF should be assessed for their ability to capture the processes examined in this study.

10.3. Seasonal Wind Energy Forecasts

In order to apply and use seasonal wind forecasts in relation to wind energy, there needs to be a development of seasonal wind energy forecasts derived from seasonal wind forecasts. This requires a skilful seasonal wind forecast to be translated into a skilful wind energy forecast, and the skill of energy forecasts can be evaluated using data from wind energy suppliers. This recommendation can be linked to Section 1.2., which indicates that there is a global shift towards the use of renewable energy over fossil fuel-based energy generations, and seasonal wind energy forecast along with skilful wind forecasts can contribute to this movement.

10.4. Site-Specific Seasonal Wind Forecasts

The development of site-specific seasonal wind forecasts for evaluation against site observations will help achieve a more robust assessment of seasonal wind forecasts over specific regions and the assumptions made will also decrease. Site-specific seasonal wind forecasts will also increase the usability of the forecast for wind energy purposes. Based on the results presented in this study, the process of quantifying added value of downscaling forecasts over regions during La Niña phases and negative SAM events could result in an increased amount of predictability of wind over regions of South Africa.

11. References

Arribas, A., Glover, M., Maidens, A., Peterson, K., Gordon, M., Maclachlan, C., Graham, R., Fereday, D., Camp, J., Scaife, A. A., Xavier, P., Mclean, P., Colman, A., Cusack, S. 2011. The GloSea4 Ensemble Prediction System for Seasonal Forecasting. *Monthly Weather Review*. 139(6): 1891-1910.

Atger, F. 1999. The Skill of Ensemble Prediction Systems. Monthly Weather Review. 127 (9): 1941-1953.

Barrera-Escoda, A., Altava-Ortiz, V., Llasat, M. C., M. Barnolas, M. 2007. Heavy rain prediction using deterministic and probabilistic models - The flash flood cases of 11-13 October 2005 in Catalonia (NE Spain). *Advances in Geoscience*. 12: 121-126.

Belmonte Rivas, M. and Stoffelen, A. 2019. Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT. *Ocean Science*. 15(3), 831–852.

Benesty, J., Chen, J., Huang, Y., Cohen, I. 2009. Pearson Correlation Coefficient. *Noise Reduction in Speech Processing. Springer Topics in Signal Processing*. 2: 1-4.

Bett, P, E., Thornton, H., Troccoli, A. 2018. Skill assessment of energy-relevant climate variables in a selection of seasonal forecast models. *ECMWF Copernicus Report: Zenodo*. ECEM Deliverable D2.2.1 version 2: 6-57.

Blaabjerg, F., Ma, K. 2017. Wind Energy Systems. *Proceedings of the IEEE*. 105(11): 2116-2131.

Blamey, R. C., Reason, C. J. C. 2012. Mesoscale Convective Complexes over Southern Africa. *Journal of Climate*. 25(2): 753-766.

Bórawski, P., Bełdycka-Bórawska, A., Jankowskia, J., Dubisa, B., Dunn, J. W. 2020. Development of wind energy market in the European Union. *Renewable Energy*. 161: 691-700.

Bosilovich, M. G., Chen, J. Robertson, F. R., Adler, R. F. 2008. Evaluation of Global Precipitation in Reanalyses. Journal of Applied Meteorology and Climatology. 47(9): 2279-2299.

Bröcker, J., Smith, L. A. 2007. Increasing the Reliability of Reliability Diagrams. *Weather and Forecasting*. 22(3): 651-661.

Bull, S. R. 2001. Renewable Energy Today and Tomorrow. *Proceedings of the IEEE*. 89(8): 1216-1226.

Burt, E., Orris, P., Buchanan, S. 2013. Scientific Evidence of Health Effects from Coal Use in Energy Generation. *Chicago and Washington: School of Public Health, University of Illinois and Health Care Without Harm*. Available at: https://noharm-uscanada.org/sites/default/files/documents-files/828/Health_Effects_Coal_Use_Energy_Generation.pdf

[Accessed on 21 June 2020]

Constitution of the Republic of South Africa [South Africa]. 1996. Available at: https://www.gov.za/documents/constitution-republic-south-africa-1996
[Accessed on 20 July 2020]

Department of Energy. 2019. The South African Energy Sector Report. Available at: http://www.energy.gov.za/files/media/explained/2019-South-African-Energy-Sector-Report.pdf
[Accessed on 19 October 2020]

Department of Energy. No Date. Renewable Energy. Available at: http://www.energy.gov.za/files/renewables_frame.html
[21 September 2020]

Department of Environmental Affairs. 2019. South Africa signs Paris Agreement on Climate Change in New York. Available at: https://www.environment.gov.za/mediarelease/southafricasignsparisagreementonclimate
[18 September 2020]
Department of Mineral Resources and Energy. 2019. Integrated Resource Plan 2019. Available at: http://www.energy.gov.za/IRP/2019/IRP-2019.pdf
[Accessed on 8 March 2021]

Fawcett, A., Pitcher, G. C., Bernard, S., Cambella, A. D., Kudela, R. M. 2007. Contrasting wind patterns and toxigenic phytoplankton in the southern Benguela upwelling system. *Marine Ecology Progress Series*. 348: 19 – 31.

Fogt, R. L., and Marshall, G. J. 2020. The Southern Annular Mode: variability, trends, and climate impacts across the Southern Hemisphere. *Wiley Interdisciplinary Reviews: Climate Change.* 11(4): 1-24.

Fulton, M., Reid, S. 2018. In Sight of the Clean Trillion: Updates on an expanding landscape of investor opportunities. *Ceres Inc*. Available at: https://edx-files.s3.amazonaws.com/climate/Ceres_In_Sight_Clean_Trillion_May10_2018.pdf
[Accessed on 02 February 2021]

Goddard, L., Mason, S, J., Zebaik, S. E., Ropelewski, C. F., Basher, R., Cane, M. A. 2001. Current Approaches to Seasonal-to-Interannual Climate Predictions. *International Journal of Climatology*. 21(9): 1111-1152.

Hahmann, A. Kruger, A. 2010. What is the large-scale wind regime in South Africa?. *Wind Atlas for South Africa (WASA) Project Workshop*. Available at: http://www.wasaproject.info/docs/2_windclimatesafrica2010.pdf
[Accessed on 21 May 2019]

Hartmann, H.C., Pagano, T.C., Sorooshiam, S. Bales, R. 2002. Confidence Builders: Evaluating Seasonal Climate Forecasts from User Perspectives. *American Meteorology Society*. 84: 683-698.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J. 2020. The ERA5 Global Reanalysis. *Quarterly Journal of the Royal Meteorological Society*. 146: 1999 – 2049.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., Dee, D. 2019. Global reanalysis: goodbye ERA-Interim, hello ERA5. Available at: https://www.ecmwf.int/en/newsletter/159/meteorology/global-reanalysis-goodbye-era-interim-hello-era5
[Accessed on 04 September 2020]

Hoell, A., Funk, C., Zinke, J., Harrison, L. 2017. Modulation of the Southern Africa precipitation response to the El Niño Southern Oscillation by the subtropical Indian Ocean Dipole. *Climate Dynamics*. 48(7): 2529–2540

Hook, L. 2020. Clean power stocks outperform fossil fuel peers during pandemic. Available at: https://www.ft.com/content/08675019-1386-49ac-a718-031d6ab85051
[Accessed on 6 February 2021]

Howard, E., Washington, R., Hodges, K. I. 2019. Tropical Lows in Southern Africa: Tracks, Rainfall Contributions, and the Role of ENSO. *JGR Atmospheres*. 124(21): 11009-11032.

Hutchinson, K., 2017. Why deeper insights into the Agulhas Current can shed light on climate patterns. Available at: https://www.news.uct.ac.za/article/-2017-07-19-why-deeper-insights-into-the-agulhas-current-can-shed-light-on-climate-patterns#:~:text=The%20Agulhas%20Current%20transports%20warm,moisture%20and%20carry%20it%20inland
[Accessed on 29 November 2020]

Johnston, P. A., Archer, E. R. M., Vogel, C. H., Bezuidenhout, C. N., Tennant, W. J., Kuschke, R. 2004. Review of seasonal forecasting in South Africa: producer to end-user. *Climate Research*. 28(1): 67-82.

Jourdier, B. 2020. Evaluation of ERA5, MERRA-2, COSMO-REA6, NEWA and AROME to simulate wind power production over France. *Advances in Science and Research*. 17: 63 – 77.

Klopper, E. 1999. The use of seasonal forecasts in South Africa during the 1997/98 rainfall season. *Water SA*. 25(3): 311-315.

Kruger, A. C. 1999. The influence of the decadal-scale variability of summer rainfall on the impact of El Nino and La Nina events in South Africa. *International Journal of Climatology*. 19: 59-68.

Kruger, A. C. 1999. The Relationship between ENSO, Seasonal Rainfall, and Circulation Patterns in South Africa. Master of Science in Atmospheric Science. University of Cape Town.

Kruger, A. C., Goliger, A. M., Retief, J, V., Sekele, S. 2010. Strong Wind Climatic Zones in South Africa. *Wind and Structures*. 13(1): 1-19.

Lakhraj-Govender, R., Grab, S, W. 2018. Assessing the impact of El Niño–Southern Oscillation on South African temperatures during austral summer. *International Journal of Climatology*. 39(1): 143-156.

Landman, W. A., Beraki, A. 2012. Multi-model forecast skill for mid-summer rainfall over. *International Journal of Climatology*. 32: 303-314.

Landman, W. A., Beraki, A. 2015. ENSO Forecasts in South Africa. Council for Scientific and Industrial Research and the South African Weather Service. Available at: https://researchspace.csir.co.za/dspace/bitstream/handle/10204/8547/Landman2_2015.pdf?sequence=1&isAllowed=y
[2 May 2019]

Landman, W. A., Beraki, A. 2015. Operational Seasonal Forecast System Development in South Africa. *CSIR and South African Weather Services*. Available at: https://www.wcrp-climate.org/images/modelling/WGSIP/14th_session/Landman_summerschool.pdf
[Accessed on 19 October 2019]

Landman, W. A., Botes, S., Goddard, L., Shongwe M. 2005. Assessing the predictability of extreme rainfall seasons over southern Africa. *Geophysical Research Letters*. 32(23): 1-4.

Landman, W. A., Dewitt, D., Dong-Eun, L., Beraki, A., Lotter, D. 2012. Seasonal Rainfall Prediction Skill over South Africa: One-versus Two-Tiered Forecasting Systems. *Weather and Forecasting*. 27(2): 489-501.

Landman, W. A., Engelbrecht, F., McGregor, J. L., van der Merwe, J. H. 2013. Atmospheric Modelling for Seasonal Prediction at the CSIR. *South African Society for Atmospheric Sciences.* Available at: https://www.researchgate.net/publication/277198034_Atmospheric_Modelling_for_Seasonal_Prediction_at_the_CSIR
[Accessed on 15 May 2019]

Landman, W.A., Kgatuke, M.J., Mbedzi, M., Beraki, A., Bartman, A., Piesanie, A.D., 2009. Performance comparison of some dynamical and empirical downscaling methods for South Africa from a seasonal climate modelling perspective. *International Journal of Climatology: A Journal of the Royal Meteorological Society*. 29(11): 1535-1549.

Lennard, C. 2019. Multi-Scale Drivers of the South African Weather and Climate. The Geography of South Africa: Contemporary Changes and New Directions. *World Regional Geography Book Series*. 9: 81-90.

Leutbecher, M., Palmer, T. N. 2008. Ensemble Forecasting. *Journal of Computational Physics*. 227(7): 3515-3539.

Lockwood, J. F., Thornton, H, E., Dunstone, N., Scaife, A, A., Bett, P, E., Li, C., Ren, H. 2019. Skilful seasonal prediction of winter wind speeds in China. *Climate Dynamics*. 52: 1-19.

Lutjeharms, J. R. E., Monterio., P. M. S., Tyson, P. D., Obura, D. 2001. Oceans around Southern Africa and regional effects of global change. *South African Journal of Science*. 97(3): 119-120.

Mason, S. J. 1998. Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *International Journal of Climatology*. 18(2): 47-164.

Mason, S. J. 2004. On Using "Climatology" as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Monthly Weather Review*. 132(7): 1891-1895.

Mazzarella, A., Giuliacci, A. 2009. The El Niño events: their classification and scale-invariance laws. *Annals of Geophysics*. 52(5): 517-522.

Mazzarella, A., Giuliacci, A., Liritzis, L. 2010. On the 60-month cycle of multivariate ENSO index. *Theoretical and Applied Climatology*. 100(1): 23-27.

Mazzarella, A., Giuliacci, A., Scafetta, N. 2012. Quantifying the Multivariate ENSO Index (MEI) coupling to CO2 concentration and to the length of day variations. *Theoretical and Applied Climatology*. 111: 601-607.

McSweeny, R., Timperley, J. 2018. The Carbon Brief Profile: South Africa. Available on: https://www.carbonbrief.org/the-carbon-brief-profile-south-africa#:~:text=South%20Africa%20is%20the%20world's,fuel%2C%20towards%20gas%20and%20renewables.
[20 September 2020]

Murphy, A. L. 1996. The Finley Affair: A Signal Event in the History of Forecast Verification. *Weather Forecasting*. 11(1): 3–20.

National Center for Atmospheric Research. No date. Available at: https://climatedataguide.ucar.edu/climate-data/marshall-southern-annular-mode-sam-index-station-based
[Accessed on 22 February 2020]

National Oceanic and Atmospheric Administration. 2018. Ocean Explorer. Available at: https://oceanexplorer.noaa.gov/facts/climate.html
[Accessed on 15 May 2019]

National Oceanic and Atmospheric Administration. 2021. Multivariate ENSO Index Version 2 (MEI.v2). Available at: https://psl.noaa.gov/enso/mei/
[Accessed on 05 March 2021]

Nchaba, T. 2012. Verification of gridded seasonal wind speed forecasts over South Africa. *University of Cape Town.* Available at: https://open.uct.ac.za/handle/11427/4970
[Accessed on 8 February 2020]

Nchaba, T., Mpholo, M., Lennard, C. 2016. Long-term austral summer wind speed trends over southern Africa. *International Journal of Climatology*. 37(6): 2850-2862.

Nhesvure, B. 2020. Impacts of ENSO on coastal South African Sea Surface Temperatures. Master of Science in Physical Oceanography. University of Cape Town.

Palutikof, J. P., Holt, T., and T. J. Osborn. 2002. Seasonal forecasting of strong winds over Europe. *American Meteorology Society*. 3(12): 1-4.

Pérez-Fernández, S., Martínez-Camblor, P., Filzmoser, P., Corral, N. 2021. Visualizing the decision rules behind the ROC curves: understanding the classification process. *Advances in Statistical Analysis*. 105: 135-161.

Philippon, N., Rouault, M., Richard, Y., Favre, A. 2012. The influence of ENSO on winter rainfall in South Africa. *International Journal of Climatology*. 32(15): 2333-2347.

Project Ukko. No Date. Seasonal Wind Predictions for the Energy Sector. Available at: http://project-ukko.net/
[Accessed on 3 May 2019]

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor. 2007. Climate Models and Their Evaluation. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press*. 8: 591-648.

Ramon, J., Lledó, L., Torralba, V., Soret, A., Doblas-Reyes, F. J. 2019. What global reanalysis best represents near-surface winds?. *Quarterly Journal of the Royal Meteorological Society*. 145(724): 3236-3251.

Reason, C. J. C., Landman, W., Tennant, W. 2006. Seasonal to Decadal Prediction of Southern African Climate and its Links with Variability of the Atlantic Ocean. *American Meteorological Society*. 87(7): 941–956.

Rouault, M., Monyela, B., Anicet, R., Kounge, I., Stella, A., Njouodo, N., Bastien, D. 2019. Ocean Impact on Southern African Climate Variability and Water Resources. *Water Research Commission*. WRC Report No. 2425/1/18: 1-134.

Rouault, M., Pohl, B., Penven, P. 2010. Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. *African Journal of Marine Science*. 32(2): 237–246.

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y., Chuang, H., Juang, H. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber,J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C., Liu, Q., Chen, Y., Han, Y., Lidia Cucurull, L., Reynolds, R. W., Rutledge, G., Goldberg, M. 2010. The NCEP Climate Forecast System Reanalysis. *Bulletin of the American Meteorological Society*. 91(8): 1015-1057.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E. 2014. The NCEP Climate Forecast System Version 2. *Journal of Climate*. 27(6): 2185 – 2208.

Sheard, J. 2018. Quantitative data analysis. *Research Methods (second edition).* Chandos Publishing. 429-452.

South African Government. 2019. Geography and climate. Available at: https://www.gov.za/about-sa/geography-and-climate
[15 May 2019]

South African Weather Service. 2019. Seasonal Climate Watch. Available at: http://www.weathersa.co.za/images/data/longrange/gfcsa/scw.pdf
[Accessed on 29 April 2019]

South African Wind Energy Association. 2019. Available at: https://sawea.org.za/
[Accessed on 29 April 2019]

Stefanova, L. Krishnamurti, T. N. 2002. Interpretation of Seasonal Climate Forecast Using Brier Skill Score, The Florida State University Superensemble, and the AMIP-I Dataset. *Journal of Climate.* 15(5):537-544.

Thornton, H., Bett, P., Clark, R., Scaife, A., Hoskins, B., Brayshaw, D. 2017. Skilful seasonal predictions for the European Energy Industry. Met Office. Available on: https://www.wcrp-climate.org/images/modelling/WGSIP/19th_session/6-1-WGSIP_Oct2017_HT.pdf [02 May 2019]

United Nations Framework Convention on Climate Change (UNFCCC). No date. The Paris Agreement. Available at: https://cop23.unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement
[Accessed on 20 September 2020]

United Nations Framework Convention on Climate Change (UNFCCC). No date. What is the Kyoto Protocol?. Available at: https://unfccc.int/kyoto_protocol

[Accessed on 18 September 2020]

Van Gelder, T. 2015. Brier Score Composition. Available at: https://timvangelder.com/2015/05/18/brier-score-composition-a-mini-tutorial/
[Accessed on 13 June 2019]

Veitch, J., Rautenbach, C., Hermes, J., Reason, C. 2019. The Cape Point wave record, extreme events and the role of large-scale modes of climate variability. *Journal of Marine Systems*. 198(103185).

Visbeck, M. 2009. A Station-Based Southern Annular Mode Index from 1884 to 2005. *Journal of Climate*. 22(4): 940-950.

Whitaker, J. S., Loughe, A. F. 1998. The relationship between ensemble spread and ensemble mean skill. *Monthly Weather Review*. 126(12): 3292– 3302.

Willis, D, J., Niezrecki, C., Kuchma, D., Hines, E., Arwade, S. R., Barthelmie, R. J., DiPaola, M., Drane, P. J., Hansen, C. J., Inalpolat, M., Mack, J. H., Myers, A. T., Rotea, M. 2018. Wind energy research: State-of-the-art and future research directions. Renewable Energy. 125: 133-154.

WindFinder. Available at: https://www.windfinder.com/about/
[Accessed on 5 March 2020]

Wooten, R. D. 2011. Statistical Analysis of the Relationship Between Wind Speed, Pressure and Temperature. *Journal of Applied Sciences*. (11): 2712-2722.

World Meteorological Organization. 2012. Guidelines on Ensemble Prediction Systems and Forecasting. 1091: 7-32. Available at: https://library.wmo.int/doc_num.php?explnum_id=7773
[Accessed on 2 June 2019]

Wright, J., Landwehr, G., Chartan, E. 2019. Assessing the value of improved variable renewable energy forecasting accuracy in the South African power system. *Journal of Energy in South Africa*. 30(2): 2413 – 3051.

Wu, Y., Hong, J. 2007. A literature review of wind forecasting technology in the world. *IEEE Power Tech*. (246): 504-509.

Yu., J., Zhou, T., Jiang, Z. 2020. Interannual variability of the summer wind energy over China: A comparison of multiple datasets. *Wind Energy*. 23(11): 1-13.