A CMIP5 Model Selection Specific to South Africa's Winter Rainfall Zone

Peter Marsh

Supervisor: Dr Christopher Jack

Abstract

This study undertakes a CMIP5 model selection specific to the Winter Rainfall Zone (WRZ) of South Africa, seeking to reduce the range of future climate projections through identifying a subset of models with increased realism and independence. In order to navigate the subjectivity in identifying relevant circulation metrics to assess models against, the 'Day Zero' drought is used as a characteristic episode. Here initially the extensive literature produced subsequent to the drought has been drawn on to identify and evaluate relevant regional process metrics, before utilising the anomalous conditions during the drought to validate various assessment methods. The dynamics subsequently identified as being most influential to rainfall supply in the WRZ include the South Atlantic subtropical jet stream responsible for steering of mid-latitude storm systems, the South Atlantic subtropical high, and the presence, or preferably absence, of precipitation blocking subsidence, and the prevalence of mid-latitude storm systems, critical for transport and upliftment of moisture to the region. Models were subsequently assessed against these metrics and scored following the technique of McSweeney et al. (2015). Unrealistic models were removed from the ensemble while significantly biased models were also excluded as their absence did not significantly reduce the range of future projections. The same scoring methods were then utilised to create a genealogy of models attaining similar results to that of Knutti, Masson & Gettelman (2013). A subset of 6 CMIP5 models which are more independent and historically more realistic than that of the full ensemble were subsequently identified. While the range of future temperature projections of the final ensemble are somewhat constrained in comparison to the full ensemble, the primary utility is argued to be the reduced number of models where a future researcher may consider each model's projected future climate pathway individually before selecting a model, or models, which best informs their use case, whilst being assured that this model performs suitably well in the region and that the initial ensemble considered adequately represents model uncertainty, while strong similarity between two or more models within the ensemble will not be unduly biasing results.

Acknowledgements

The author is grateful for funding provided for this study by the National Research Foundation and the University of Cape Town. The author would like to further thank CSAG for access to computing facilities as well as staff and students for always being willing and eager to assist. An important thank you is due to my supervisor Dr Christopher Jack for stewardship and engaging with any and all questions. As always, a big thank-you to my parents, or rather officemates during the challenging past year. And finally, thanks to Sonja who is responsible for the majority of full stops.

Contents

Part I

Introduction and Literature Review

Introduction

The WRZ of South Africa is unique to the southern African region as it relies on different regimes for moisture supply and experiences distinct climate modes than that of neighbouring regions. The region can be typified by extreme gradients both spatially and across the meteorological seasons. The spatial extent of the WRZ cannot be easily defined, whilst being the area where winter rainfall dominates is self-explanatory a large transition area exists; north along the length of the Atlantic coast the influence of cold fronts diminishes, before the Namib desert prevails, while to the east summer rainfalls, aided by the warm Agulhas current, become increasingly prevalent contributing to an increasingly all-year round rainfall regime. Inland, to the north-east, the Cape Fold Belt mountains serve to block winter storms from accessing the continental plateau giving way to a convective summer rainfall regime (Reason, 2017).

The WRZ has experienced persistent droughts of varying magnitude over the past decades, often resulting in water restrictions being imposed on the region. The most intense of which occurred over the period 2015-2017 resulting in the most severe water shortage in 113 years and the Western Cape being declared a disaster region (Botai et al., 2017). Thus, reliable and actionable future climate projections are critical for future resource planning in the growing Cape Town metropolis (Ziervogel et al., 2010). Distilling climate information to inform such critical decisions is a complex task which needs to consider multiple lines of evidence across differing scales and disciplines, with the ultimate goal being to provide robust and reliable climate information with low uncertainties, that in an ideal circumstance is tailored to the context of a specific decision maker (Jack et al., 2021).

At present climate models curated through the Coupled Model-Intercomparison Project (CMIP) and utilised in the latest Intergovernmental Panel on Climate Change (IPCC) climate science assessment reports, present the primary source of evidence for what future climate may look like under various emission scenarios (Meehl et al., 2014). Here advances in computation along with improved understanding of physical processes has resulted in increasingly realistic Global Circulation Model's (GCM) as well an increasing number of GCM's within the CMIP ensemble. Various methods of downscaling these GCM's such as through Regional Climate Models (RCM's), curated through projects such as the Coordinated Regional Downscaling Experiment (CORDEX), can provide increased spatial resolution. At very high spatial resolutions this allows for explicit modelling of deep convection (Stratton et al., 2018) and regional moisture transport characteristics and thus potentially more realistic regional climate responses. Alternatively statistical downscaling methods develop statistical relationships between circulation and surface responses such as daily rainfall and can be used to improve the representation of local responses in models. Applied to GCM output or to that of RCM output, can provide

additional sources of evidence for what future climate a region may experience. However, as the models, methods and number of these projections has increased so too has the number of projected future climate scenarios.

Thus, deciding which projections are most likely or robust has become an increasingly difficult task. Traditionally where strong agreement exists within the ensemble that projection is considered to be more likely than those with little agreement. The validity of this approach is however dependant on whether the projections from within the ensemble can be considered independent (Knutti et al., 2010). Alternatives to this approach of equally weighting each model include; model selection studies where a sub-selection of better performing models are selected, or model weighting studies where better performing models are given a higher weighting. However, all model selection studies are constrained in the very nature of assessing a model against an unknown future, thus relying entirely on historical performance. McSweeney et al., (2015) present an approach where instead of considering which models perform best, the problem is considered as which models perform unrealistically or poorly and do not change the range of future projections and can thus be removed. This therefore places emphasis on models which have shown to realistically capture dynamics important to a region and removes those that are unable to capture the relevant dynamics or indeed do capture the relevant dynamics but display significantly more bias in so doing than other models within the ensemble without providing additional information on what future climate may look like. Instead, this approach relies on other models within the ensemble that project a similar future climate pathway and do not display a significant unrealistic bias. This presents an opportunity to reduce uncertainty and range of projections but also provides a platform for selection of ensembles or single models without the risk of selecting an unrealistic or poorly performing model. This would prove particularly useful for future hydrological or downscaling studies where model selection can have a significant impact on results.

In any model selection study, selection of metrics against which to assess models presents a significant challenge - with relating each metric directly to the region in question being scope for a study in itself. This study explores this challenge in the context of the WRZ of South Africa, where a lack of strong regional or global teleconnections, or other clear drivers of regional climate variability are present (Wolski et al., 2020). This study takes a unique approach to metric selection where metrics are considered using an episodic approach, drawing on the storyline approach of Shepherd et al., (2018), where climate change is considered utilising past significant events (episodes) rather than relying on semantic statistical points of reference. The benefits of this are two-fold; firstly, identifying the driving dynamics behind a past period of anomalous climate presents less of a challenge than identifying metrics as a function of long-term climate change induced trends, where separating natural climate variability from significant trends is a challenging task and often hampered by the relatively short timespans of observational or reanalysis data available. Secondly, by considering a known past event the significance of changing climate shifts from a future existential risk to a present and continued threat to present day society, where psychologically these past experienced episodes can be more easily related to by a reader than intangible statistical

quantifications of climate change. By considering climate change to be an already present threat, rather than seeking to identifying what future climate risks a region may experience, the problem is instead framed as what risks the region has experienced in the recent past, or is presently experiencing, and to what extent future climate change may further exacerbate these risks.

In this case Cape Town's 'Day Zero' drought, whose severity and increased likelihood has been attributed to climate change (Otto et al., 2018; Sousa et al., 2018; Burls et al., 2019), is used as the episodic reference point where through leveraging the literature produced subsequent to the 'Day Zero' drought, relevant dynamics are identified by first constructing a storyline describing the anomalous climatic conditions during the 'Day Zero' period and next by considering the ability of subsequently developed metrics to capture these anomalous conditions as a benchmark against which to assess the metrics relevance to the WRZ region. Selected metrics will then be used to score the realism of CMIP5 models, where a model that fails to capture the dynamics that can result in a 'Day Zero' type drought in the WRZ would be considered unfit for utilisation in future climate projections. The aim of this study is thus to first identify a set of dynamics relevant to the WRZ before using these identified metrics to select a reduced ensemble of CMIP5 models, that display increased realism when considered in the context of large-scale climate dynamics relevant to the WRZ, while also ensuring model independence within the ensemble is increased relative to that of the full CMIP5 ensemble.

Literature Review

The Winter Rainfall Zone

The WRZ is distinct from the rest of South Africa as it primarily receives winter rainfall, where the average daily rainfall in winter months (JJA) is 167% greater than that of the non-winter months1. Further-more, rainfall in the region is spatially diverse with steep topography resulting in distinct climatic variation across the Cape Fold Belt Mountains. Rainfall to the region is primarily supplied by eastbound extratropical cyclones. The sudden rise of the Andes mountains over 2000 km to the west disrupts the otherwise continuous 'roaring forties and furious fifties' wind belts, giving genesis to a steady stream of extratropical cyclones (Hoskins & Hodges, 2005). From here, a number of variables including a sufficient supply of moisture across the Atlantic, favourable mid-latitudinal temperature gradients and suitable stratospheric conditions determine whether these sub-tropical systems will traverse the Atlantic and impact the Western Cape peninsula. Invariably, only a minority of extratropical cyclones formed bring sufficient moisture for a rainfall event to occur in the Western Cape. The frequency and intensity of extratropical cyclones impacting the WRZ is influenced by a number of mechanisms.

Influences include shifts in the position and intensity of the South Atlantic High Pressure as well as the presence of transient high pressure ridging activity (Burls et al., 2019), the positioning of the subtropical jet stream, or overall positioning of the westerly wind belt impacting the track and intensity of extratropical cyclones (Mahlalela, Blamey & Reason, 2019; Sousa et al., 2018). Sufficient supply of moisture particularly from tropical South America (Reason, Jagadheesha & Tadross, 2003) and the presence of upstream atmospheric river events (Blamey et al., 2018) further influence rainfall in the region. Increased low level convergence west of the Cape peninsula can also serve to enhance rainfall in the region (Reason, Jagadheesha & Tadross, 2003). Temporally-sporadic cut-off low events can also result in extreme rainfall and flooding in the region (Favre et al., 2013). During the summer when the westerly storm track is located further south, southerly flow, from transient coastal low-pressure systems, bring moisture from the cold Benguela current to the region. On longer seasonal time-scales influences from climate indices such as El Niño – Southern Oscillation (ENSO) (Philippon et al., 2012) and Southern Annular Mode (SAM) (Reason & Rouault, 2005) have been observed, as well as relationships with South Atlantic sea surface temperatures (Reason & Jagadheesha, 2005a) and Antarctic sea-ice anomalies (Reason et al., 2002).

The WRZ is distinct from the rest of Southern Africa in that it does not experience the same strong ENSO teleconnections (Reason & Jagadheesha, 2005b). While some connections have been linked between winter rainfall in the WRZ and South Atlantic sea surface temperature (Reason, Jagadheesha & Tadross, 2003) and Southern Annular Mode (Reason & Rouault, 2005), the teleconnections remain relatively weak resulting in weak statistical correlations. Wolski et al., (2018) utilize a Self-Organizing Map (SOM) based classification for rainfall over Cape Town where the greatest explanatory power is related to atmospheric pressure. This is consistent with the region being dominated by mid-latitude low pressure systems, with the depth of these systems being the primary determinant of rainfall event intensity and duration.

This is contrasted to Johannesburg, located in the northern inland summer rainfall region, where moisture availability has the greatest explanatory power. However, the study finds that even the highest portions of variance explained by SOM node frequencies remains below 60% and can be as low as 8% of variance explained. Daron et al., (2019) utilize a climate process chain technique to better conceptualize climate processes in South Africa, where rather than the traditional reductionist approach of relating climate or weather to a single process, a holistic view is taken where process chains are used to account for multiple cross scale interactions focused on specific climate outcomes of societal importance. This technique is used to relate SAM to Winter Rainfall in the Western Cape, shown in figure 1. Here an example of a specific process chain is given as: "Negative SAM meaning a greater equatorward extent of the circumpolar westerlies →Shifted meridional circulation →Equatorward shift in the subtropical jet location over the South Atlantic →Increased moisture flux from the subtropical Atlantic →Enhanced moisture supply to cold fronts moving towards South Africa →Increased rainfall and a higher risk of wet winters in the Western Cape region with associated heavy rainfall events."

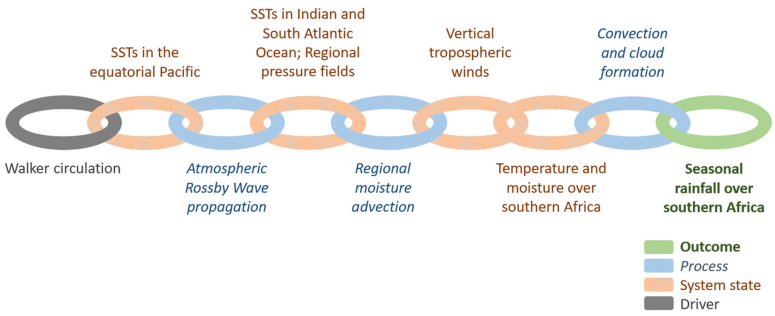


Figure 1: CPC (Climate Process Chain) linking circumpolar westerlies in the Antarctic to heavy rainfall events in southwest southern Africa. Adapted from Daron, J., Burgin, L., Janes, T., Jones, R.G. & Jack, C. 2019. Climate process chains: Examples from southern Africa. International Journal of Climatology. 0(0). DOI: 10.1002/joc.6106.

Future Climate Projection Techniques and future WRZ Climate Projections

Continued global record high temperatures, the proliferation of extreme weather events and climate conscious movements, particularly from the youth (Sabherwal et al., 2021), has once again brought climate change into the global consciousness. While the exponential growth of clean energy and continued policy shifts from leading nations gives cause for hope, the question remains whether the current rate of change is enough to avoid a climate disaster. Locally the 'Day Zero' drought spanning the period 2015-2017, where three consecutively dry winters meant severe water rationing was required to ensure continued water supply to the Cape Town metropolis, has served a stark reminder that the WRZ, like many subtropical regions globally, is particularly vulnerable to changing climate. Thus, increasingly decision makers need reliable and actionable future climate projections to better motivate for mitigation and inform adaptation measures.

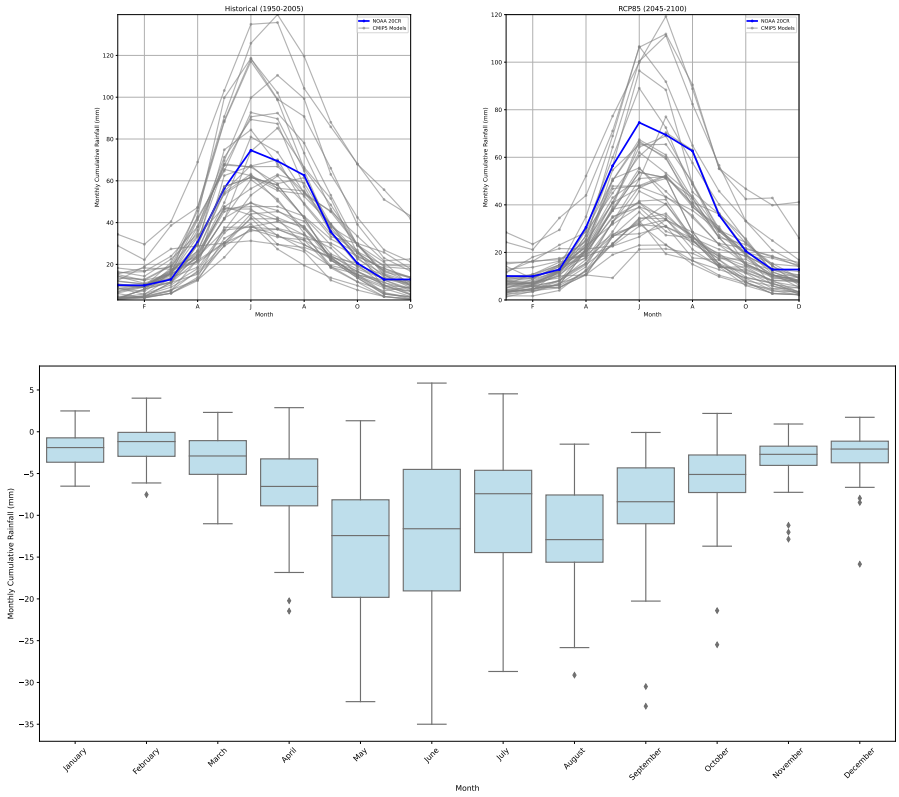


Figure 2: Monthly precipitation (mm) from 27 CMIP5 models (grey) and NOAA 20th Century Reanalysis (blue) utilising the nearest grid point to 34°S 18°E. top left: Historical ensemble considering 1950-2005. Top right: rcp8.5 ensemble considering 2045-2100. Bottom: Rainfall anomaly (rcp8.5 – historical) considering the same periods, 33rd-67th percentile indicated by box and 10th – 90th percentile indicated by whisker extent. Code Link

The Coupled Model Intercomparison Project (CMIP) forms the basis for future climate projections allowing different climate modelling groups to compare results under a common set of anthropogenic greenhouse gas and other radiative forcing scenarios. Climate projections from the 5th CMIP project (CMIP5) indicate a strong probability of continued increasing temperatures under all emission pathways in the WRZ region, while late twenty-first century projections under the high emission rcp8.5 scenario indicate an increased probability of severe drying in the WRZ region (DEA, 2018, pg 110). This combined with the observed widening of the tropical belts resulting in poleward migration of mid-latitude storms (Seidel et al., 2008), consistent with 21st century Global Climate Model (GCM) simulations indicating poleward migration of storm tracks and mid-latitude precipitation (Yin, 2005) highlights the risk to regional water resources and hence the importance of understanding the implications of climate change to the Western Cape.

At present, multi-model coupled Atmosphere-Ocean Global Circulation Model (AOGCM) ensemble projections are considered the forefront in quantifying future climate change. The climate change projections produced by the IPCC in the AR5 special report, which rely

on the ensemble of GCM simulations curated through CMIP5, present the forefront of future climate projections. The CMIP5 ensemble has now been superseded by CMIP6 where outputs are being finalised and will be synthesized in the 2022 IPCC sixth assessment report. The findings from the IPCC reports are however, intended only as a high level summary and do not always provide suitable information at regional scales (Hewitson et al., 2017). As an example the AR5, Adaptation and Vulnerability report does not specifically differentiate between the WRZ and the rest of Southern Africa (IPCC AR5 I, 2014), while the AR6 regional fact sheet differentiates west and east Southern Africa, where the WRZ would fall into the West Southern Africa subregion (IPCC AR6 I, 2021) Instead, decision makers need information at regional and even local scales. Here, given the regions vulnerability to water scarcity, the future climate projections and climate risk resources presently available to a water resource manager in the WRZ are discussed.

Figure 2 considers the projections from the CMIP5 model ensemble for the period 2040-2060 for monthly rainfall over the WRZ region, a metric of particular relevance to water resource planners. Here variations in model projections from wetting to extreme drying can be seen. The multi-model ensemble mean indicates drying across all winter months, however the 10th – 90th percentile ranges from little to no change to extreme drying in winter months. The historical CMIP5 model results all broadly capture the seasonality of the region; however, the model spread is very broad. Further, most models are seen to underestimate the start of the rainy season in April, whilst model spread is largest in June and July when monthly rainfall is at its peak. Some models fail to accurately capture the end of the rainy season in September and two models are seen to overestimate the Autumn and early winter rainfall by a large amount. However, these projections are for a large area, and when the spatial extent of rainfall changes are considered further discrepancies exist, such as the GISS-E2-H-CC model indicating wetting over the Southern tip of the country but overall drying in the south western Domain (Mahlalela, Blamey & Reason, 2019). This may indicate these projections are highly sensitive to the region selected and dynamics may not in all cases be well captured (Mahlalela, Blamey & Reason, 2019).

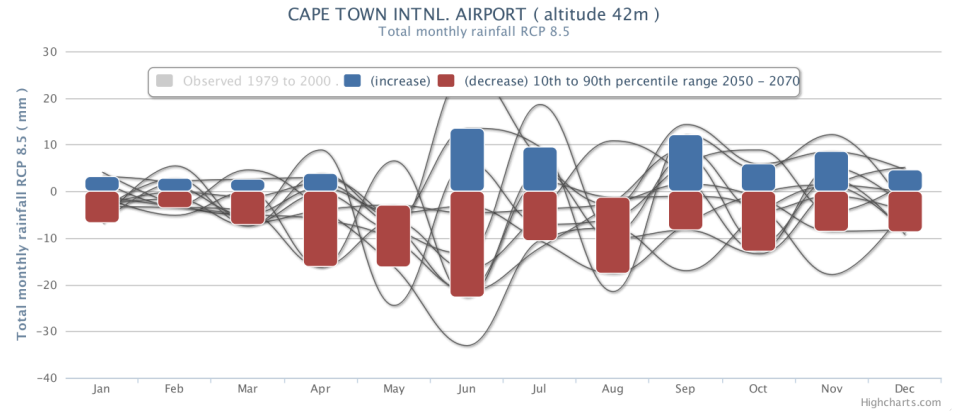


Figure 3: Total monthly rainfall anomaly under the RCP8.5 scenario considering the period 2050-2070 against an 1980-2000 baseline. Figure available from: https:/cip.csag.uct.ac.za/

Utilising this raw model output for decision making contexts without some form of 'expert' guidance is thus not a reasonable approach, nor is this raw data easily available to the public. Nevertheless, making climate data easily available to the public is a key part of climate communication. Climate Information Website's seek to provide a bridge between climate science and society (Hewitson et al., 2017), by providing post processed future climate projections in easy-to-use public facing portals. One such example being the Climate Systems Analysis Group Climate Information Portal (CIP) (https://cip.csag.uct.ac.za/). Figure 3 considers a future projection of total monthly rainfall for a station in the Cape Town Metropolis, available from this particular CIP.

Here an ensemble of 10 different CMIP5 models are statistically downscaled to a weather station located at the Cape Town International Airport. The 10th to 90th percentile of the ensemble projections is used to quantify potential future climate states, while black lines indicate individual models which have purposefully not been named. This visualizations format is intentionally chosen to provide the user with easily attainable and actionable information whilst conveying realistic constraints in the skill and uncertainties of the projection. Here, similar to the projections from the raw CMIP5 ensemble considered in Figure 2, increased likelihood of drying is seen in all seasons and particularly so in the early winter period. However, this projection does indicate somewhat increased likelihoods of wetting or no change than that of the raw CMIP5 output, whether this can be attributed to the influence of statistical downscaling or the reduced ensemble of CMIP models is unclear. Thus, here a balance between enough technical detail to fully capture the data available and simplifying the data enough to be usable for an end user exists (Hewitson et al., 2017).

For decision makers with more resources available an integrated climate service approach provides the most holistic approach to constructing the most likely future climate states. Here through expert analysis climate-information can be synthesised to best suit the needs of a particular decision maker (Hewitt, Mason & Walland, 2012). For instance, the statistical downscaled outputs available from the CIP can be contrasted against dynamically downscaled outputs available through the CORDEX project and the benefits of each method considered. In the case of water resource management coupling these downscaled outputs to a streamflow model can provide a more comprehensive picture, while drawing upon the wealth of literature that already exists around the WRZ can be used to add further sources of evidence to support a particular future scenario. These could include; considering the poleward migration of extratropical cyclones and its implications for subtropical regions such as the WRZ (Sousa et al., 2018), considering the shift in rainfall event intensity in the WRZ linked to Hadley Cell expansion (Burls et al., 2019), considering the prevalence of drought in the early winter period in recent years (Mahlalela, Blamey & Reason, 2019), understanding the role cut-off lows play in extreme rainfall in the region (Abba Omar & Abiodun, 2020; Engelbrecht, Engelbrecht & Dyson, 2013) as well as the role atmospheric river events play in transporting moisture to the region (Blamey et al., 2018). Further, when considering the precise locations relevant to the decision maker different dynamics could be relevant given the extreme spatial rainfall gradients in the region (Wolski et al., 2020). Quantifying the future risk of extreme events, such as drought, is always a challenge when considering future climate. However, through

considering the role climate change played in increasing the likelihood of the 'Day Zero' drought (Otto et al., 2018) as a result of changing rainfall regimes, end users can better infer societal consequences and make better informed decisions.

While the information available through an integrated climate service and a climate information website is significantly different, both fundamentally rely on the CMIP ensemble for providing a baseline quantification of future climate states. Here assessing the realism of these CMIP models in a particular region will add value to both these use cases, through supporting or opposing each model, and the projected scenarios thereof. A climate services group can incorporate these findings into future research and focus downscaling efforts on these better performing models while considering only the most realistic models can provide a more robust outlook in the CIP without adding any further complications to the end user.

Uncertainties in Ensemble Projections

Knutti, (2010) considers the problem of GCM model assessments in a seminal editorial "The End of Model Democracy" where the practice of utilizing ensemble model spreads with a 'one model one vote' approach is questioned. Here the previously held argument that increasing the number of independent models should inherently reduce the uncertainty, as is the case in Bayesian statistics, is questioned. The present status-quo of utilizing the ensemble mean to provide a best guess future climate projection is reliant on considering the error between individual models and observation to be random and independent. Here, when considering a set of independent time series with a random error the mean of the series will have an error lower than that of an individual member such that as the ensemble size approaches infinite the error approaches zero. While in the CMIP ensemble one model may perform better than the ensemble in a single variable, no model performs better than the ensemble across all variables due to the 'cancellation of errors' (Flato, 2011; D'Ercole et al., 2018). This same cancelling out of extremes within the ensemble members also results in the ensemble mean having a lower variance than that of observations, while individual models have a variance more similar to reality and thus can better capture extreme climate events (Bishop & Abramowitz, 2013). While, in most cases, considering the mean of the ensemble does result in reduce error when considering historical runs against observed climatologies, whether this remains true in future projections under differing emissions scenarios and where model spread is significantly larger is an unknown. If models are instead considered to provide 'replicate earths' with discrete trajectories for the Earth, rather than approximating reality, then the ensemble mean would not be expected to converge to reality (Tebaldi & Sansó, 2009; Knutti, 2010; Bishop & Abramowitz, 2013). Further the basis of considering the ensemble mean as a more reliable projection is reliant on models being independent, however many share core components or utilize the same parameterization of unresolved processes such as convection and as such rather than a cancellation of errors an amplification of systemic bias from a particular model group may be present (Knutti, 2010; Knutti, Masson & Gettelman, 2013; Munday & Washington, 2019).

Models within the CMIP5 ensemble are often generational which further reduces independence within the ensemble. For example the Australian Research Council Centre of Excellence for

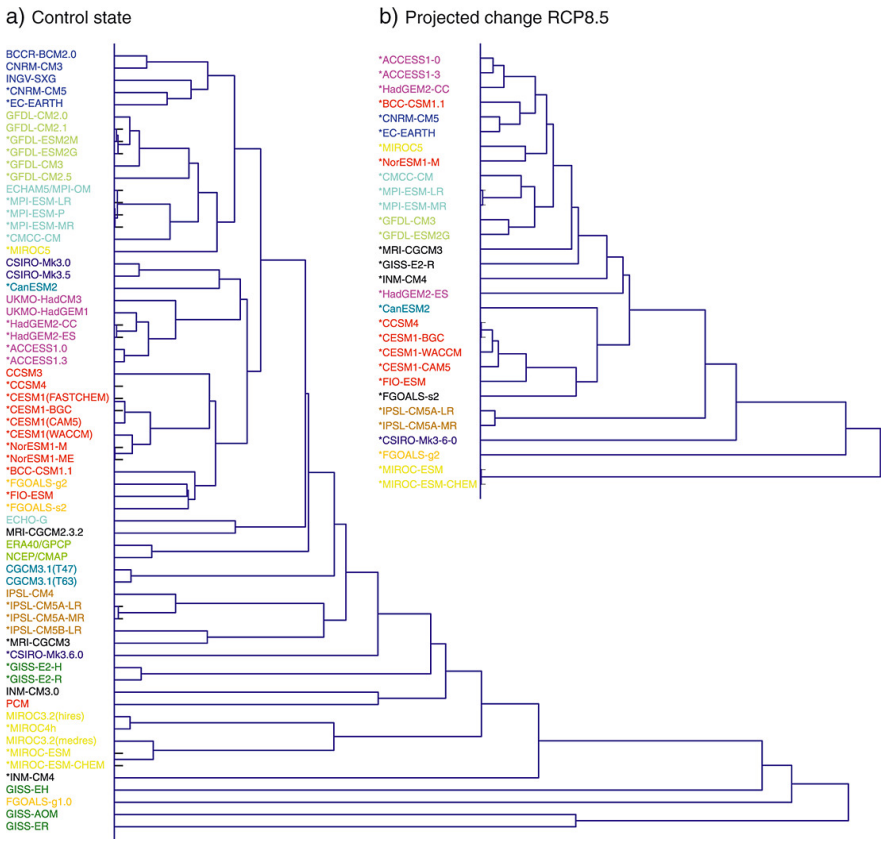


Figure 4: "(a) The model "family tree" from CMIP3 and CMIP5 (marked with asterisks) control climate observations (ERA40/GPCP and NCEP/CMAP), shown as a dendrogram (a hierarchical clustering of the pairwise distance matrix for temperature and precipitation fields) Some of the models with obvious similarities in code or produced by the same institution are marked with the same color. Models appearing in the same branch are close, and similarity is larger the more to the left the branches separate, for a detailed description of the method see Masson & Knutti, (2011)) (b) Same but based on the predicted change in temperature and precipitation fields for the end of the 21st century in the RCP8.5 scenario relative to control." Adapted from Knutti, R., Masson, D. & Gettelman, A. 2013. Climate model genealogy: Generation CMIP5 and how we got there. Geophysical Research Letters. 40(6):1194–1199. DOI: 10.1002/grl.50256

Climate System Science models ACCESS1-0 and ACCESS1-3 are different versions of the same model and hence share significant components, while both ACCESS models are also based on the HadGEM2 atmospheric dynamics core. Knutti, Masson & Gettelman, (2013) capture this relationship between models through considering the linkage distances between models, where model similarity is quantified by a Kullback-Leibler divergence, a distance metric that utilizes the spatial field of monthly values in a control simulation without external forcing and considers the seasonal cycle, the interannual variations, and spatial correlations. These linkage differences are shown in figure 4, where clear groupings within modelling groups are shown. Thus, clear difficulties in considering the mean ensemble projection occur in that groups that have produced more than one model or have shared base code with other groups effectively

have more 'votes' than a modelling group with a single-fully-independently-developed model would have.

The tendency toward similar solutions across all modelling groups is also significant in that while global climate models are primarily based on solving fundamental physical equations, limitations imposed on resolution and radiative transfer computations by todays computing power, mean that complex and multiscale processes such as cloud or ice formation need to be solved through heuristic processes, known as parametrizations. These parameterizations require delicate tuning to ensure that model results and observations match as closely as possible. Hourdin et al., (2016) argue that due to the subjectivity in selecting what parameters to tune the model against and the limited scope for testing different parameterizations, due to the huge complexity of GCM's, the tuning process is as much an art as it is a science. The subjectivity of this tuning process is a primary reason why even models that share no base coding's cannot be considered to be entirely independent and arguably a reason as to why some models are better suited for some applications than others. For example, a model that has had its orographic drag specifically tuned2 may better capture relief induced subsidence than a model that has not and thus would perhaps perform better in capturing the large-scale subsidence across the South Atlantic induced by the Andes Mountain range. However, there is a possibility that tuning to better fit a metric to observations may mean it will no longer be physically consistent with the sensitivity of this particular metric to changing climate, where this tuning is done in an unphysical manor (Eyring et al., 2019). Further tuning against one particular metric may degrade performance against another.

Model Selection

This inherent lack of independence among models indicates that considering each models output equally is suboptimal (Eyring et al., 2019). Evaluation of models against observations to provide a weighting in favour of more realistic models and potentially reduced uncertainty presents a clear way forward. However, here the problem of deciding an appropriate weighting adds yet another layer of uncertainty to the final projection. Selecting a benchmark against which to assess model realism is not a trivial task considering evaluation against the future state they are projecting is an impossibility (Knutti et al., 2010) - the very nature of future climate being uncertain. Instead, models can only be assessed against present day climate, with the sensitivity of models to different future forcing scenarios being a key area where no robust assessment of each model's realism is possible. There is no definitive method of assessing the realism of a model, whereby every assessment is conditional on the particular variable, domains, statistics, and other assumptions made during the assessment.

Sanderson, Knutti & Caldwell, (2015a) analyse both model independence and performance by considering inter-model distances between the principal components of a high dimensional matrix constructed using key model diagnostics, where seasonal mean values from the gridded historical model outputs are normalized and concatenated into a single high dimensional matrix across the CMIP5 models, principal component analysis (PCA) to reduce the data series to a small number of orthogonal components. Next multidimensional scaling is used to construct a 2-dimensional space that exhibits the same interpoint distances as that of the components

from the PCA analysis. Hereby using model error to assess both model independence and skill. Through this analysis the CMIP5 ensemble can be resampled to models that are more similar to observations and are more independent than the full ensemble. However, somewhat paradoxically when the ensemble is weighted to maximise model independence and realism is not accounted for, the error between the ensemble and historical observations is greater than that of the full ensemble. This is as a result of a quasi-natural selection, where through the sharing of code and knowledge across the climate community and some institutions presenting more models than others, the climate community has already upweighted models that historically perform best, therefore demonstrating that the present status quo of considering each model equally is in fact an unconscious model selection. However, when the ensemble is sub-selected to maximise both realism and independence the error between the ensemble and historical observations can be significantly reduced. Nevertheless, whether past performance can be used as an indication for future performance remains a key uncertainty.

McSweeney et al., (2015) outline an alternate approach where models are assessed against present day climate and those which are deemed to be unrealistic or biased and within the range of the remaining ensemble are removed, improving the realism of the ensemble while preserving the maximum range of future climate projections, thus effectively applying a weighting of either 1 or 0 (Overland et al., 2011). This approach is applied over the continental regions; South-East Asia, Africa, and Europe, where the abilities of models to represent crucial dynamics for each region is considered. Models are subsequently classified as: being satisfactory, having biases, having significant biases or being implausible. The models are then considered relative to their influence on the range of projections. Using this solution ensures that satisfactory models are favoured while those with significant biases are only considered if their absence significantly affects the range of outcomes. Those deemed to perform sufficiently poorly as to be implausible are not considered, Overland et al. (2011), refer to this type of approach as "more of a model culling than a model selection". Thus, the goal of reducing the number of poorly performing models within an ensemble whilst still preserving a representative range of future climate projections is achieved. A key difference between this approach to that of Sanderson, Knutti & Caldwell, (2015) is that the final ensemble is selected to maximise discrete future climate pathways, rather than maximise that of independent models, as measured by past performance.

McSweeney et al., (2015) utilize an approach of continental scale model selection under the premise that such an approach reduces the overhead of interfacing GCM's and RCM's for modelling groups and the need for multiple selection studies. However, Overland et al. (2011), highlight that the increased diversity at the regional scale makes elimination of models more feasible and defendable. Further while, this large-scale approach utilized by McSweeney et al. does consider the performance of models in capturing African temperature and precipitation climatologies, an emphasis is placed on the performance of the models in simulating West African monsoon and ENSO teleconnections dynamics, with the former having no directly observed influence and the latter weak teleconnections in the WRZ. However, it is important to remember that the quality of projections needs to be related to more than just mean climate states as it is not clear to what extent models are tuned to represent previous mean climate

states (Overland et al., 2011).

This type of model culling approach presents some clear benefits. In removing models that are unrealistic, the range of model spread for a particular region can potentially be reduced. Gershunov et al., (2019) assess the performance of CMIP5 models in simulating atmospheric river events in western North America and find that by only considering models which accurately capture the dynamics of present-day atmospheric river events the range of future rainfall regime projections in the region can be constrained. When considered in the context of dynamical downscaling projects, presenting a smaller ensemble to be downscaled by removing unrealistic models or removing worse performing models with similar projections to more realistic models computational and human expense can be significantly reduced allowing for resources to be focused on downscaling of models that are most realistic.

Model selection does however not guarantee constrained future projections. Kolusu et al., (2021) considered various common model selection techniques of CMIP5 models over eastern Africa including, considering the ensemble mean or uniform weighting; a binary inclusion or exclusion approach such as Gershunov et al., (2019); the approach of Sanderson, Knutti & Caldwell, (2015) considering both independence and performance; as well as an approach of excluding implausible models that lie on the extremities of predictions, utilised by Rowell, (2019) and not dissimilar to the approach utilized by McSweeney et al. (2015). Here Kolusu et al., (2021) find each method to have little influence on risk profiles in two case studies, whereby sensitivity to model weighting is far exceeded by bias in correction methodologies and only the most aggressive model weighting approach could influence decision making. Here while the range of future climate projections could be somewhat reduced through model selection no impact on the risk profiles in each case study could be made. Thus, in this case the difficulty presented in quantifying the uncertainty of the new distribution outweighs the benefit of narrowing the distribution from the full ensemble in the first place.

Assessment Metrics

While model selections are naturally constrained in having to assess models against past rather than future performance, the metrics through which this is done and how selection of different metrics may affect results is a significant source of uncertainty. Sanderson, Knutti & Caldwell, (2015) consider the variables: precipitation, top of atmosphere long- and short-wave radiation flux, surface temperature, zonal temperature and zonal humidity when constructing their multidimensional matrix to score CMIP5 models. Here intermodal distances calculated using only surface temperature are strongly correlated to that of the multivariate case, while considering any of the metrics alone is seen to have little effect on the final ensemble selection. Whether this remains true at a regional rather than global scale however is uncertain, but it does highlight the interlinked nature of variables within the climate system. Further to what extent model tuning against these parameters may have influenced results is unclear.

Santer et al., (2009) analyse the ability of CMIP3 models to capture water vapour in tropical ocean basins and find no relationship between a model's ability to capture annual mean climatological values and their ability to capture the seasonal cycle or the amplitude or pattern

of variability. Thus, arguing that at the least a model selection needs to consider seasonality as considering only annual climatology is liable to more influence from model tuning techniques, while capturing the seasonal cycle relies on the model accurately representing the underlying physics.

Model selections are often considered on global or continental scales, while climate adaptation decisions are made on regional scales (Hewitson et al., 2017), and thus metric assessments should reflect this. A model which, for instance, captures tropical sea surface temperature and cloud relationships accurately could not necessarily be expected to perform better in a sub-tropical region (Knutti, 2010). In regions with strong relationships with singular climatic drivers, such as sea surface temperature-cloud relationships in East Africa (Rowell, 2019) or atmospheric river event frequency in western North America (Gershunov et al., 2019), identifying metrics with which to assess models is straight forward. However, in regions where climate drivers are diverse selecting metrics presents a considerable task. Quagraine et al., (2020) use co-behaviour to navigate this uncertainty, where automatic feature detection techniques are used to identify various co-behaviour modes across several GCM's and the anomaly in 700 hPa pressure, temperature and precipitation over southern Africa of each model under various co-behaviour modes is assesses, allowing for the assessment of a model's ability to represent the interactions between large-scale process and the impacts on the southern African region.

Thus, given the above constraints to contemporary model selection efforts this work develops a set of model evaluation metrics which are directly relevant to the unique WRZ of South Africa.

Part II

Methodology, Data and Metric Selection

Methodology

The methodology of this study is divided into 4 sections. Initially selection of metrics against which to assess GCMs, this is discussed in the 'Metric Selection' section below while the methods used in each specific metric assessment are discussed in the dedicated chapter 'Metric Evaluation'. After models have been scored against each metric each model is designated as performing either 'realistically', 'biased', 'significantly biased' or as 'implausible'. Next each model will be considered in context of its future projections as well as relative to sibling models to identify models for exclusion. Sibling models will be identified using the same metrics as the models are scored against. Finally, the projections from the final ensemble are contrasted against the full ensemble. These four model selection steps are further discussed below:

1) Selection of metrics to assess models.

Metrics are selected through a literature review, where the ability of the selected metric to capture changes in mean climate during the 'Day Zero' drought period is used as an 'acid test' to ensure the chosen metrics are directly relevant to WRZ climate. This is discussed further in the dedicated 'Metric Selection' section.

2) Evaluate all models and categorize models as 'implausible', 'significantly biased', 'biased' or 'realistic'.

Here each model's performance is considered against each metric and scored as being realistic, biased, significantly biased or implausible. These designations will be made relative to the NOAA 20th Century Reanalysis dataset (NOAA 20CR) (Slivinski et al., 2019). NOAA 20CR is used as the reference reanalysis primarily to the increased timespan available than that the other reanalysis dataset, ERA5 (Hersbach et al., 2020), considered. Models that are the most similar to the reanalysis being scored realistic and the least similar being deemed implausible. Because there are no obvious thresholds for categories each category will be differentiated relative to the worst scoring model. A number of different scoring techniques are considered for each metric to ensure consistent scores are realized before a final scoring technique is selected.

3) Designate 'Model Family' groupings and select only best performing models from these groupings in order to increase model independence within the final ensemble.

Utilising the same techniques used to score models, the similarity between each model is calculated and a pairwise matrices of model similarity is constructed. Ward's minimum variance clustering technique is then used to create a dendrogram of model similarity, where 'sibling'

groupings are designated, and worse performing siblings are eliminated.

4) Refer to each models' future projections complete decision matrix to identify models for further exclusion.

Figure 5 provides a schematic of the final selection process where, after scoring models and designating sibling groupings, implausible models are removed before considering significantly bias models that don't add any further information to the range of projections than that of better performing models are excluded. Finally models with a better performing sibling present are excluded, where sibling models receive the same score the sibling model which serves to increase the range of projections is selected.

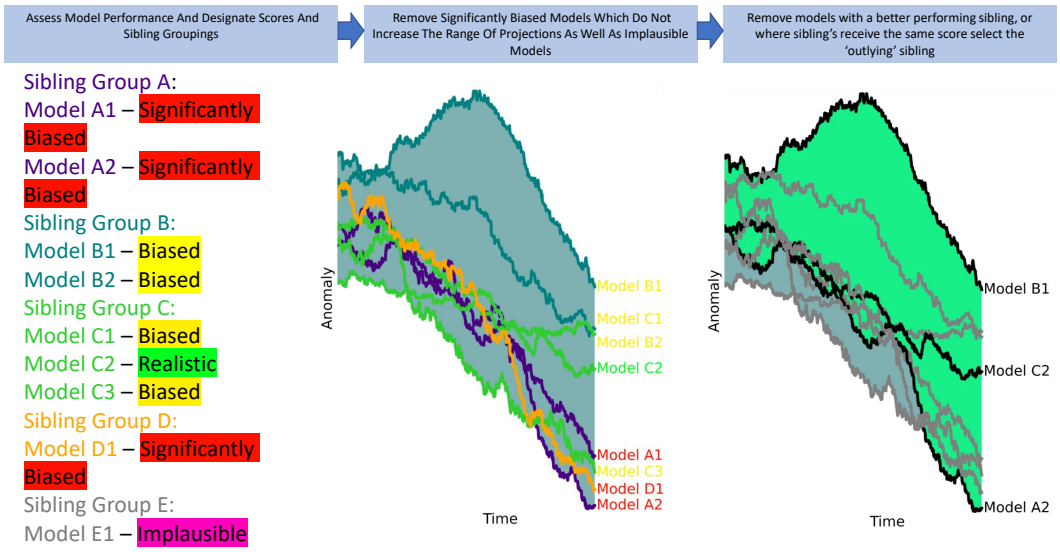


Figure 5: Schematic of the Model Selection Process described above. Where models B1,C2 and A2 are presented as the reduced ensemble, which retains a similar range to that of the initial 9 model ensemble while excluding unrealistic models and increasing ensemble independence by including only a single model from each sibling grouping.

Data

Assessments will be made against the CMIP5 ensemble Historical runs which span the period 1850-2005. This 'historical' run is a core part of the CMIP5 experiment, where models are forced by observed atmospheric composition changes, from both natural and anthropogenic sources, and including land -cover changes. This experiment allows for evaluation of models against observed climate variables over most of the industrial era (Taylor, Stouffer & Meehl, 2012). The intention across this study is to maximises the ensemble wherever possible, thus the ensembles utilised across the three metric evaluations differ due to availability of different variables across models within the CMIP5 ensemble. The largest possible ensemble within each metric is preferred as any persistent bias from the CMIP5 ensemble in a specific metric will be easier to identify in a larger ensemble. While performance of different models can be contrasted against one another. Where performance in a single metric may be binary in that some models capture the dynamic well while others do not or instead a steady decay from better to worse performing models may exist. The final model selection only includes models with the necessary variables available across all three metrics. The primary source for the CMIP5 data is from the Earth System Grid Federation data nodes (https://esgf-node.llnl.gov/projects/cmip5/), made available via the University of Cape Town's Climate Analysis Group (CSAG) storage and computing facilities. Limits on data availability are diverse and can stem from data corruption temporary or otherwise on the ESFG nodes or the local server, and in some cases modelling groups may choose not to preserve certain variables due to storage limitations or model design. In this study models are assessed against the NOAA 20th Century Reanalysis (NOAA 20CR) (Slivinski et al., 2019) over the period 1950-2005. This time span is chosen to maximise the ensemble size as historical CMIP5 runs prior to 1950 are not available across all models, while the longer 1950-2005 span of NOAA 20CR is preferred over the shorter (1980-present) span available from ERA5 where shorter term climate variations may have a larger impact. NOAA 20CR is used as a proxy of 'real world' conditions during this period. NOAA 20CR spans the period 1900-2015 and thus ERA5 reanalysis product (Hersbach et al., 2020) will be used for assessing anomalies during the 2015-2017 'Day Zero' Drought against a 1980-2018 baseline climatology.

Metric Selection

This initial step of selecting metrics presents a considerable task and is the primary source of uncertainty within the study. Relating even a single metric and quantifying its influence over the climatic conditions in the WRZ is scope for a study in itself. The WRZ presents a particularly challenging region due to both its size and the extreme climate gradients within the region. Further rainfall trends within the region are not homogenous. Wolski et al., (2020) studied the spatio-temporal variability of rainfall trends within the Cape Town Region, defined as the area 35°S -31°S and 18°E to 21°E, and highlighted this heterogeneity identifying three subregions with distinct interannual and seasonal variability as well as longer term rainfall trends. In this context it is important to highlight that the purpose of GCM's is not to forecast future climate at any given point (the WRZ is represented as a single grid point in some of the lower resolution GCM's), but rather to capture how global climate dynamics may shift in a warmer future climate. Relating these shifts to the context of a particular region remains an important but separate task.

The metrics against which the models are assessed should reflect this and rather than comparing performance of historical climate simulations in the region directly to that of climatology, a model that performs well in the region is instead defined as one that captures a broad range of dynamics relevant to the region. These dynamics could include those outlined in the climate process chain developed by Daron et al., (2019). Further a model that performs well should capture the relationships between these dynamics and anomalous behaviour in the region.

Shepherd et al., (2018) present storyline-based techniques as a means of navigating the 'cascades of uncertainty' in the chaotic earth system, whereby relating dynamics directly to a real-world event with societal impacts presents a means to reduce uncertainty by simplifying process networks, whilst still maintaining grounding in robust physical aspects. A story line is defined as 'a physically self-consistent unfolding of past events, or of plausible future events or pathways' (Shepherd et al., 2018). In this case the 'Day Zero Drought' provides a convenient reference point as a significant recent event, while the wealth of literature developed subsequent to the drought is used to define which global processes are most relevant to the WRZ.

Among the literature covering the 'Day Zero' drought the key themes across various studies attribute the anomalous condition experienced during the period to: a weaker and poleward displaced jet stream (Sousa et al., 2018; Mahlalela, Blamey & Reason, 2019); increased subsidence and 'blocking' activity induced by the South Atlantic Subtropical High (Mahlalela, Blamey & Reason, 2019; Burls et al., 2019); and reduced rainfall intensity from winter storm systems (Burls et al., 2019). Thus, a storyline could follow as:

> *"A multi-year period of reduced winter rainfall in the Western Cape region due to southward displacement of extratropical cyclones and the South Atlantic Subtropical High resulting in increased blocking frequencies in the region, reducing the intensity and duration of rainfall events."*

Potential exists to expand this storyline however a key emphasis of the storyline approach is to ensure all processes remain physically plausible. Further benefits of this storyline technique are that instead of the traditional semantic approach of considering risk along arbitrary definitions, the risk is directly related to real world experience in an episodic manor, helping people better digest such information in a response known as availability bias (Shepherd et al., 2018; Kahneman, 2011).

It is however important to ensure the metrics selected are important to the region as a whole and not just during a drought period. Thus, the processes against which to assess models are considered foremost from a synoptic point of view, where dynamics that are most critical to rainfall events occurring in the WRZ are considered, before relating these to drought events. The dynamics considered are the subtropical jet stream in the South Atlantic, responsible for steering of mid-latitude storm systems, the South Atlantic subtropical high, and the presence, or preferably absence, of precipitation blocking subsidence, and the prevalence of mid-latitude storm systems, critical for transport and upliftment of moisture to the region. The 'Day Zero' drought is then used as an acid test, such that if the metric fails to capture a meaningful anomaly in reanalysis models during the 'Day Zero' drought period, then the effectiveness of this particular metric to the WRZ is questioned.

In the chapter 'Metric Evaluation' the key dynamics from this story line, namely the jet stream, the high-pressure system, and cold fronts in the South Atlantic, will be discussed before various methods to quantify these systems are considered and their response during 'Day Zero' contrasted. These methods are then used to quantify CMIP5 model performance against that of reanalysis.

Part III

Metric Evaluation

South Atlantic Mid-Latitude Jet Stream

Introduction

The position of the jet stream has a strong influence on climate in the WRZ, where the sharp seasonal temperature and precipitation gradient is a function of the seasonal north-south procession of the jet stream and associated wind belt. The South Atlantic presents a particularly interesting midlatitude region, where a marked double character of the Southern Hemisphere jet exists. Here the upper-level jet stream alternates between a more typical single jet centred around 40°S and a less typical quasi-seasonal double jet structure with the subtropical jet located around 30°S and a polar front jet situated around 60°S, this double jet structure becomes more pronounced as meridional temperature gradients increase through austral winter (Gallego et al., 2005). The influence of the Andes is responsible for this duality where the interruption of the westerly wind belt causes cyclogenesis to occur from relatively low latitudes through to extreme high-latitudes in the Drake passage, resulting in this vacillation between a one-jet and two-jet state (Spensberger & Spengler, 2020) Cold fronts, responsible for the majority of rainfall in the WRZ, are embedded in the jet stream and thus more make landfall at the relatively far north southern tip of Africa during the double structured phase of the South Atlantic jet. Tennant and Reason (2005) find there is no obvious distinction in the seasonal mean kinetic energy of the sub-tropical wind field between wet and dry years in the WRZ, as well as in southwestern Western Australia where the bifurcation of the jet is a more frequent occurrence. This transient nature of jet thus presents a challenge in identifying a climatology in these regions and perhaps explains why automatic feature detection techniques have been favoured in previous studies relating jet stream variability to WRZ climate.

The Southern Annual Mode (SAM) index is often used to describe the non-seasonal meridional movement of the westerly wind belt in the Southern Hemisphere and is the leading mode of variability in the extratropical Southern Hemisphere on weekly and centennial time scales (Lim et al., 2016), whereby a positive (negative) SAM is characterized by a poleward (equatorward) shift of the eddy driven jet and associated extratropical storm tracks (Kidson, 1988). A robust trend in future climate projections is a shift toward a more positive SAM, with clear implications for regions reliant on extratropical systems for moisture supply (Lim et al., 2016). Reason & Rouault (2005) highlight the influence of SAM on the WRZ where 6 out of the past 8 driest winters (JJA) during the period 1948-2004 occurred during a positive SAM phase. Further, using the method outlined by Marshall (2003), defining SAM by utilizing station-based data to consider the pressure difference between 40S and 65S, Sousa et al. (2018) show that indeed the positive SAM trend continued during the 'Day Zero' drought, and for the first time in three decades, a positive SAM occurred in all three parts of the winter rainy season for successive years in a row (Archer et al., 2019). Daron et al., (2019) feature SAM as a primary

link in their WRZ climate process chain, where a negative SAM and associated, equatorward shift in the subtropical jet stream enhances the risk of a wet winter and extreme rainfall over the WRZ, while the converse would be expected during a positive SAM.

However, while a positive SAM is conducive to reduced rainfall in the WRZ, it does not guarantee reduced rainfall, with SAM describing a hemispheric rather than regional process. Instead, a long wave trough and associated anticyclonic anomaly occurring in neighbouring ocean basins, alongside a persistent positive SAM, further exacerbated the drought conditions during the 2015 -2017 period (Archer et al., 2019). While SAM was first defined using station-based observations along the Antarctic coastline and isolated observations from the few island atolls in the Southern Ocean (Marshall, 2003), a now more widely used technique to define SAM within climate simulations is instead to consider the primary principal component of Southern Hemisphere Sea level pressure. However, relating modes of variability to physical process remains a complex task, where Wittman, Charlton & Polvani (2005) observe that across all numerical models, regardless of complexity, the north-south jet wobble is captured as the first principal component, due to the nature of the north-south precession being symmetrical across the jet's maximum. While meridional displacement of the zonal wind flow is strongly tied to the SAM phase, a similar link is not observed between SAM and eddy kinetic energy (Spensberger et al., 2019), where a strengthening of the circumpolar jet is associated with a poleward shift, as defined by the transformed Eulerian momentum equation (Kidston et al., 2015). Further SAM is not homogenous across the Southern Hemisphere during winter, seasonal variations in shape and structure of SAM exist and correlations with SAM are not seasonally consistent. Kidson (1988) attribute this to zonal wavenumber 3 (kz-3) superposition while Ding et al. (2012) attribute this behaviour to a tropical Pacific induced low-frequency Rossby wave influence. These low frequency planetry waves, known as Rossby waves, can be used to quantify such variations in the circumpolar flow. Tennant and Reason (2005) find anomalies in the circumpolar flow during wet winters in the WRZ appear to resemble a wavenumber-1 anomaly, while dry winters are associated with a wavenumber-3 superposition.

Spensberger et al. (2019) argue that the Antarctic continent has a disproportionate influence over SAM, where instead SAM is seen to be a measure of the degree of thermal decoupling between the Antarctic continent and that of the Southern Ocean, and as such influence of SAM onsub-tropical jet stream meridional movements has previously been overstated. Thus given the complexity of midlatitude variability in the South Atlantic basin it is preferred instead to opt for a more straight forward method considering the position and shape of the South Atlantic Jet Stream on a daily basis, before developing a

climatology of the distribution of the jet across the South Atlantic. The aim here is to find a relatively straight forward technique to score the performance of CMIP5 models in capturing midlatitude dynamics, where the jet stream is the single most dominant feature in the region. Here methods of comparing jet stream simulations of varying complexity are considered.

Analysis

1-Dimensional Jet Analysis

A common geometric approach is to consider the jet stream as a maximum in the zonal wind flow, where the latitude of maximum wind-speed after fitting a quadratic is used to denote the jet latitude. Barnes and Povlani (2013) use this geometric technique alongside a principal component analysis to study the projected changes in the structure of the jet stream under increased CO2 emissions.

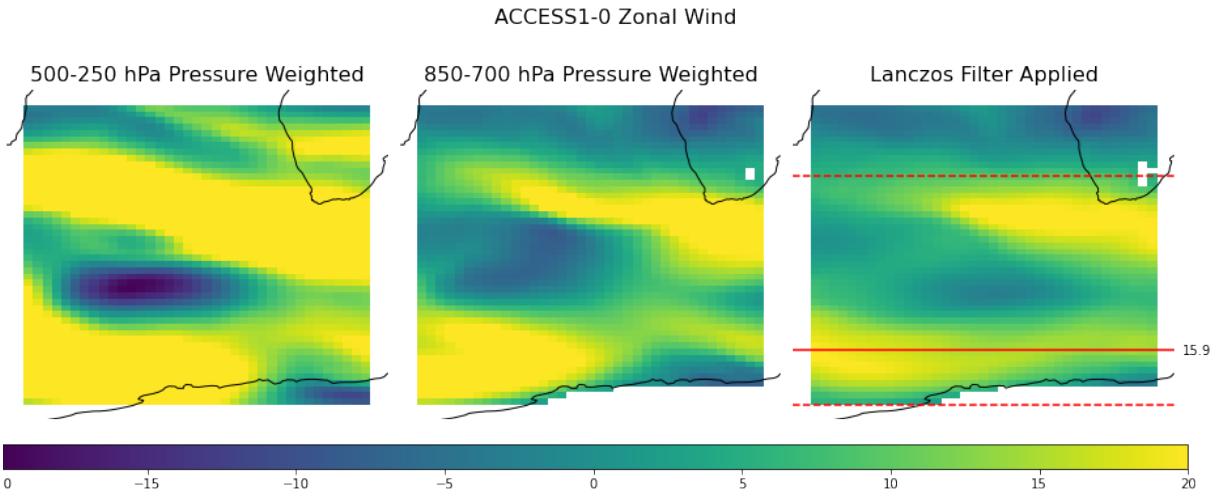


Figure 6: Illustration of the process of calculating the '1-dimensional' jet stream metric using the ACCESS1-0 model as an example. Left – 500-250 hPa pressure weighted average of daily mean zonal wind. Middle – considering only the lower level 850-700 hPa levels. Right – After Lanczos filter has been applied to the 850-700 hPa levels. Jet Latitude denoted by the solid line, FWHM captured as the distance between the dashed lines and Speed as the maximum zonal mean velocity, 15.9 m/s on this particular day. Code Link

The area 40W to 30E is considered to include the entire South Atlantic basin and Indian Ocean sector directly to the east of the WRZ to capture the influence of where the Agulhas current meets the Southern Ocean (Reason, 2017). The pressure weighted average of daily mean zonal wind from 850mb-700 hPa is then calculated. A lower level is preferred as the distinction between the lower eddy-driven jet and the higher thermally driven subtropical jet is strongest. The eddy-driven jet being more easily related to synoptic weather events and where the influence of the rapid ozone depletion and recovery that occurred during the analysis time period is less evident, a particular concern as the CMIP models with interactive chemistry do not necessarily share the same ozone forcing's as non-chemistry models. Further a known limitation of low-top CMIP5 models is a lack of stratospheric variability (Gerber et al., 2012). A 10-day low-pass Lanczos filter is then applied to this using 41 weights. This filtering approach is utilised in order to reduce noise associated with individual synoptic events (Hamming, 1998; Lorenz & Hartmann, 2001; Barnes & Polvani, 2013). The effect of this filtering is highlighted in figure 6. In order to quantify the daily latitude, velocity and

width of the jet stream, the zonal average U-wind value is calculated, where the maximum U-wind value represents the jet velocity. The latitude is then calculated by fitting a quadratic to this maximum value and the two-neighbouring grid-points and considering the vertex of this quadratic as the jet latitude. Jet width is then captured as the 'Full Width Half Mean' (FWHM) where the north and south extent of the jet are classified as the latitudes where jet velocity is half that of maximum jet velocity, the distance between the northern and southern extents then quantifying the width of the jet. The position of the FWHM is estimated through linear interpolation between the nearest 2 grid points.

A summary of the results are included in Table 1. Distributions of the 10-day filtered jet stream metrics are calculated using a Kernel Density Estimation method, where each value is represented by a Gaussian curve centred at its value and at each grid point these values are summed and the area under the resultant curve is normalised to 1. The bandwidth, which determines the number of grid points much like the number of bins in a histogram, is calculated using Scott's rule (Scott, 1979) which is a function of the standard deviation, inter-quartile range and number of independent data points. Function given in equation 1:

$$eq1.bw = 1.059 \times min\left(\sigma, IQR \times 1.34\right) \times 216^{\frac{-1}{5}}$$

There are many different interpretations and implementation of this, however the influence is minor provided the input data is continuous; for instance, considering jet latitude or width without interpolation in lower resolution models yields perturbed results due to grid point intervals being smaller than model resolution.

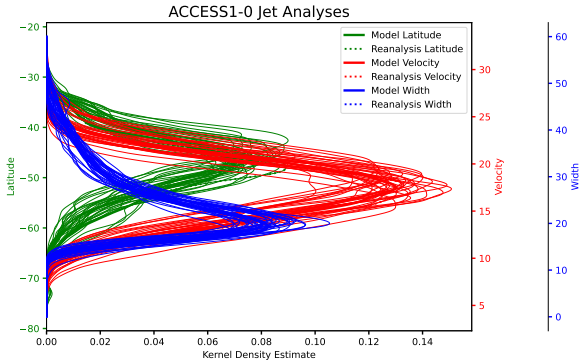


Figure 7: Jet latitude, width and velocity distributions, approximated using the kernel density methods described above. Dark lines indicate the distributions of the ACESS1-0 model in this case, while dashed lines are that of NOAA 20th Century Reanalysis. Lighter lines in the background are that of the CMIP5 models.

Table 1 captures the key descriptors of these metrics. Winter anomaly considers the mean difference between winter (JJA) values and climatology, whilst seasonal precipitation correlation refers to the correlation between the seasonal mean of a metric and that of precipitation anomaly in the WRZ.

Table 1: Results from the Zonal Jet analysis Code Link

results_table_1D

	Mean					Standard Deviation					Winter Anomaly					Seasonal Precipitation Correlation								
Model	Latitude	Velocity	Width	Northern Extent	Southern Extent	Latitude	Velocity	Width	Northern Extent	Southern Extent	Latitude	Velocity	Width	Northern Extent	Southern Extent	Latitude	p-value	Velocity	p-value	Width	p-value	Northern Extent	p-value	Southern Extent	p-value	Model
IPSL-CM5A-LR	-44.240	17.390	22.420	-33.170	-55.590	5.110	2.930	5.610	3.870	5.640	0.040	-0.850	-2.360	-1.150	1.210	0.040	5.79e-01	0.310	3.17e-06	0.460	1.00e-12	0.398	1.31e-09	-0.193	4.51e-03	IPSL-CM5A-LR
MPI-ESM-MR	-46.590	17.170	23.680	-34.620	-58.300	6.040	3.020	6.630	4.880	5.610	-1.180	-0.060	-2.900	-2.430	0.480	0.250	1.73e-04	-0.130	8.22e-02	0.600	2.80e-22	0.609	2.68e-23	0.211	1.78e-03	MPI-ESM-MR
MRI-ESM1	-49.830	18.200	24.790	-36.400	-61.190	6.990	3.490	7.570	5.810	5.990	0.830	-0.760	-2.260	-0.830	1.430	-0.280	2.76e-05	0.250	1.89e-04	0.610	4.84e-23	0.287	1.86e-05	-0.459	1.16e-12	MRI-ESM1
MIROC-ESM-CHEM	-45.860	18.160	23.640	-34.080	-57.720	5.620	3.070	6.920	4.430	5.800	0.360	-0.890	-2.370	-0.990	1.380	-0.150	2.28e-02	0.340	4.40e-07	0.590	1.13e-21	0.415	2.10e-10	0.457	1.44e-12	MIROC-ESM-CHEM
EC-EARTH	-47.610	16.880	23.980	-35.400	-59.380	6.230	3.100	6.720	5.140	5.430	-1.570	-0.950	-2.780	-2.860	-0.080	0.140	4.63e-02	0.310	4.45e-06	0.680	1.40e-30	0.587	2.14e-21	0.301	8.85e-06	EC-EARTH
HadGEM2-CC	-47.750	16.660	24.200	-35.440	-59.640	6.090	2.920	7.560	5.050	6.150	-0.280	0.620	-3.000	-1.770	1.230	-0.050	5.02e-01	-0.470	3.47e-13	0.610	4.85e-23	0.445	6.83e-12	-0.415	2.20e-10	HadGEM2-CC
bcc-csm1-1-m	-49.300	18.990	24.400	-36.520	-60.920	5.500	3.090	5.940	4.790	4.570	-0.800	-0.370	-2.420	-1.820	0.590	0.190	6.28e-03	-0.010	8.97e-01	0.630	1.29e-25	0.582	5.81e-21	-0.338	3.75e-07	bcc-csm1-1-m
FGOALS-g2	-44.530	17.190	23.480	-33.080	-56.560	6.020	2.880	7.450	4.880	6.020	-0.670	-0.610	-3.440	-2.240	1.200	0.030	6.62e-01	0.320	1.19e-06	0.700	1.02e-32	0.592	8.08e-22	0.440	1.24e-11	FGOALS-g2
ACCESS1-0	-48.560	17.150	24.500	-35.900	-60.400	6.080	2.970	7.500	5.250	5.910	-0.020	0.050	-3.020	-1.720	1.290	-0.170	1.01e-02	-0.190	5.07e-03	0.670	4.63e-29	0.385	4.96e-09	-0.509	1.17e-15	ACCESS1-0
MIROC-ESM	-45.720	18.290	23.600	-34.050	-57.650	5.600	3.070	7.040	4.380	6.020	0.250	-0.730	-2.300	-0.940	1.370	-0.120	7.78e-02	0.230	6.60e-04	0.590	1.24e-21	0.384	5.20e-09	-0.448	4.63e-12	MIROC-ESM
CSIRO-Mk3-6-0	-46.820	17.270	24.560	-34.700	-59.260	5.740	2.720	6.880	4.420	5.740	-1.330	0.510	-2.540	-2.260	0.290	0.170	1.49e-02	0.490	9.53e-15	0.640	1.64e-26	0.519	2.91e-16	0.331	8.45e-07	CSIRO-Mk3-6-0
GFDL-CM3	-47.740	17.700	24.490	-35.500	-59.990	6.010	2.830	6.880	4.680	5.820	-1.730	-0.170	-2.610	-2.560	0.050	0.160	1.67e-02	0.020	8.14e-01	0.500	5.25e-15	0.431	3.36e-11	-0.263	9.32e-05	GFDL-CM3
CanESM2	-46.980	18.110	24.610	-34.650	-59.260	5.270	2.720	5.960	4.220	4.880	-0.920	-0.310	-2.570	-1.840	0.730	0.130	6.51e-02	0.040	5.38e-01	0.610	1.00e-23	0.503	3.03e-15	-0.372	1.75e-08	CanESM2
IPSL-CM5B-LR	-44.970	16.080	22.680	-33.600	-56.280	5.940	2.910	6.770	4.710	6.030	-0.210	-0.010	-2.540	-1.480	1.050	-0.210	1.52e-03	-0.150	2.77e-02	0.530	3.15e-17	0.233	5.53e-04	-0.415	2.17e-10	IPSL-CM5B-LR
MPI-ESM-LR	-47.090	17.150	23.680	-35.090	-58.770	6.070	3.060	6.690	5.050	5.530	-1.270	-0.130	-2.810	-2.560	0.250	0.260	1.46e-04	-0.040	5.16e-01	0.550	3.33e-18	0.536	1.77e-17	-0.212	1.71e-03	MPI-ESM-LR
CMCC-CM	-46.810	16.580	24.680	-34.200	-58.880	6.620	3.020	7.010	5.370	5.910	1.390	-0.110	-2.050	-2.400	-0.350	0.260	1.07e-04	-0.090	1.84e-01	0.530	7.77e-17	0.560	3.42e-19	-0.098	1.51e-01	CMCC-CM
GFDL-ESM2G	-47.300	16.920	25.240	-35.050	-60.290	6.190	2.620	8.660	4.850	5.210	-2.790	-0.010	-3.010	-3.410	-0.400	0.410	3.86e-10	-0.180	9.36e-03	0.550	3.40e-18	0.502	3.43e-15	-0.194	4.13e-03	GFDL-ESM2G
MIROC5	-46.660	15.440	24.270	-34.830	-59.100	6.570	2.870	8.260	5.240	7.060	-2.020	0.390	-3.010	-3.290	-0.280	0.450	5.18e-12	-0.370	1.49e-08	0.640	1.62e-26	0.725	1.70e-36	-0.076	2.67e-01	MIROC5
MPI-ESM-P	-47.260	17.070	23.890	-35.080	-58.960	6.150	3.040	6.790	5.040	5.520	-0.840	-0.150	-2.750	-2.090	0.660	0.060	3.48e-01	-0.010	9.03e-01	0.500	9.31e-15	0.409	4.18e-10	-0.291	1.39e-05	MPI-ESM-P
CMCC-CMS	-46.410	17.100	23.850	-34.200	-58.040	6.530	3.150	6.870	5.180	6.130	-0.630	-0.060	-2.650	-2.090	0.570	0.050	4.77e-01	-0.030	8.57e-01	0.560	1.90e-19	0.533	3.10e-17	-0.228	7.39e-04	CMCC-CMS
CNRM-CM5	-46.710	16.640	23.990	-34.730	-58.720	6.060	2.790	6.920	5.160	5.270	-1.700	-0.180	-2.640	-2.790	-0.150	0.460	1.49e-12	-0.030	7.03e-01	0.620	3.12e-24	0.721	5.36e-36	-0.107	1.15e-01	CNRM-CM5
IPSL-CM5A-MR	-45.420	17.480	22.710	-34.140	-56.840	5.110	2.950	5.560	4.080	5.430	-1.000	-0.710	-1.410	-1.660	-0.240	0.360	6.78e-08	0.230	7.77e-04	0.320	1.22e-06	0.583	4.69e-21	0.091	1.81e-01	IPSL-CM5A-MR
inmcm4	-48.830	16.920	23.010	-36.860	-59.870	5.200	2.620	5.480	4.460	4.400	-1.540	-0.420	-1.410	-1.960	-0.540	0.160	1.94e-02	0.140	3.83e-02	0.400	9.94e-10	0.317	1.99e-06	-0.108	1.15e-01	inmcm4
ACCESS1-3	-48.710	17.570	24.500	-36.010	-60.510	6.070	3.040	7.080	5.140	5.720	0.060	0.430	-2.960	-1.550	1.410	-0.160	1.85e-02	0.270	6.23e-05	0.670	1.03e-29	0.352	1.11e-07	0.526	9.34e-17	ACCESS1-3
MRI-CGCM3	-49.780	18.180	25.030	-36.310	-61.340	7.000	3.350	7.650	5.770	5.920	0.540	-0.540	-2.380	-0.810	1.570	-0.280	2.47e-05	0.090	2.03e-01	0.590	2.85e-21	0.218	1.23e-03	-0.519	2.87e-16	MRI-CGCM3
bcc-csm1-1	-47.130	18.060	23.920	-35.060	-58.980	5.550	2.810	6.090	4.600	4.860	-0.440	-0.290	-2.720	-1.530	1.190	-0.050	4.28e-01	0.030	8.90e-01	0.600	7.58e-23	0.375	1.25e-08	0.454	2.11e-12	bcc-csm1-1
CMCC-CESM	-45.670	17.300	22.950	-34.040	-56.990	5.210	2.910	6.020	4.240	5.280	-0.510	-0.220	-2.180	-1.580	0.590	0.100	1.61e-01	0.090	2.09e-01	0.530	2.62e-17	0.551	1.41e-18	0.233	5.72e-04	CMCC-CESM
BNU-ESM	-46.500	18.680	23.750	-34.500	-58.240	5.490	3.070	5.980	4.230	5.200	-0.180	-1.030	-3.390	-1.730	1.660	-0.060	3.63e-01	0.460	5.80e-13	0.630	3.13e-25	0.381	7.20e-09	-0.436	2.07e-11	BNU-ESM
NorESM1-M	-49.040	17.390	24.650	-36.330	-60.980	5.580	2.680	5.610	4.250	4.370	-1.210	-0.040	-2.890	-2.350	0.540	0.200	3.77e-03	-0.030	8.58e-01	0.620	1.06e-24	0.579	1.06e-20	0.323	1.23e-06	NorESM1-M
HadCM3	-47.270	17.480	24.150	-35.330	-59.480	5.510	2.670	6.450	4.770	4.920	-1.750	0.290	-3.300	-3.120	0.180	0.300	5.93e-06	0.210	1.59e-03	0.690	4.08e-32	0.608	3.40e-23	0.255	1.54e-04	HadCM3
GFDL-ESM2M	-46.870	16.940	24.940	-34.640	-59.580	6.090	2.690	6.930	4.790	5.500	-2.760	-0.120	-2.870	-3.140	-0.270	0.440	1.25e-11	-0.070	3.11e-01	0.510	1.85e-15	0.505	2.18e-15	-0.144	3.47e-02	GFDL-ESM2M
MIROC4h	-47.490	18.860	25.130	-34.700	-59.830	5.960	3.080	6.880	4.880	5.710	-1.700	-0.500	-2.460	-2.770	-0.310	0.370	2.44e-08	0.170	1.37e-02	0.660	6.67e-29	0.697	9.17e-33	-0.112	1.01e-01	MIROC4h
NOAA	-47.990	16.820	24.120	-35.540	-59.660	5.140	2.840	5.820	4.280	4.630	-0.630	0.120	-2.430	-1.650	0.780	0.130	5.09e-02	0.140	4.59e-02	0.580	1.30e-20	0.581	6.35e-21	0.301	8.63e-06	NOAA

In order to provide an absolute score to each model the integral of the area difference is calculated by inspection. Figure 8 shows a visualization of this for the HadCM3 model where this technique considers both differences in shape of the distribution as well as displacement. Area differences have been calculated using 1000 inspection points where a limit is seen to be approached after 100 inspection points. This approach was initially chosen to provide both a quantitative score and visual representation of shifts in distribution, as seen in figure 8. In practice a two-sampled Kolmogorov–Smirnov test yields similar results. Results from the area differences have subsequently been normalized assigning each model a score from 1 to 0 where 0 is the most similar to reanalysis and 1 the least. The area difference scores are calculated against the distributions of daily mean jet metrics considered on a season-by-season basis. Figure 9 highlights the need for considering seasons separately where the seasonality of jet width and latitude is captured as a shift in distribution rather than a simple transformation of mean position. The method of disaggregating models into respective climatological seasons to score against reanalysis is preferred as both seasonal and synoptic time-scale variability are considered.

2-Dimensional Jet Analysis

While the double structure observed in the high-level Subtropical jet is less prominent in the low-level eddy-driven jet, this feature does still persist and is subject to influence this method of jet tracking, and perhaps explains why jet width rather than latitude is more strongly correlated to WRZ rainfall anomalies. Thus, a second 2-Dimensional technique is also considered, and the results compared.

Rather than considering zonal wind, here the daily mean wind speed is calculated, before the same Lanczos-filter is applied to the 850-700hPa pressure weighted mean values. Influenced by Sousa et al. (2018), where the jet position is simply considered as the area where seasonal mean wind speed exceeds 20 m/s and the change in area during the 'Day Zero' Drought is contrasted against that of climatology. Sousa et al. (2018) utilize the subtropical jet for this purpose, however a similar shift in the eddy-driven jet is also seen during the 2015-2017 winters and is preferred to remain consistent with the 1-dimensional analysis.

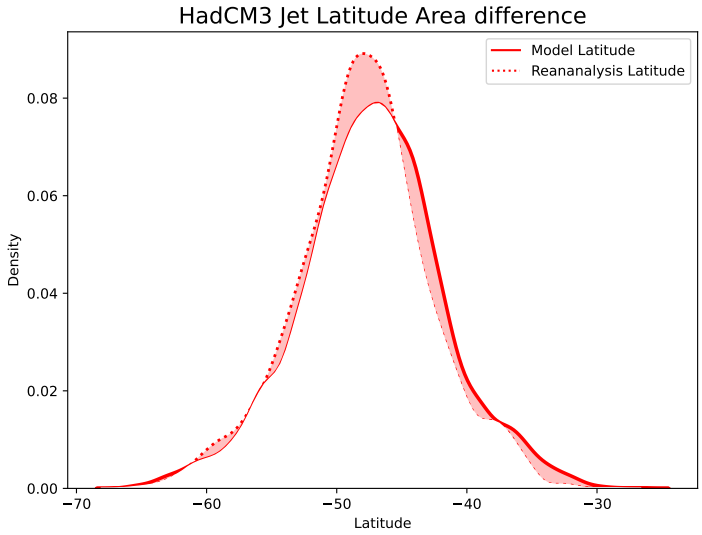
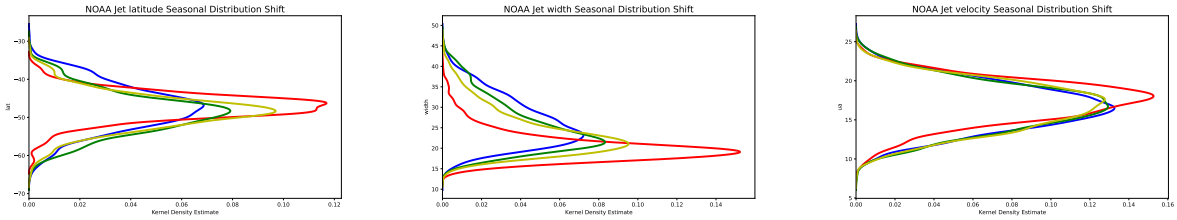


Figure 8: Visualization of the area difference between the latitude distributions of HadCM3 and that of NOAA 20th Century Reanalysis Code Link



(a) *(b)* *(c)*

Figure 9: Jet Latitude (a), Width (b) and Velocity (c) distributions decomposed into different seasons, approximated using the Kernel Density methods described above Code Link

Here the jet is initially defined as the area where seasonal mean wind speed exceeds the 90th percentile of seasonal mean wind speeds, before a 2-dimensional field of the probability of the seasonal mean jet occurring at each grid point is calculated. A further progression of this method is to define the jet on a daily rather than seasonal timescale. This is computed by defining the jet where the daily mean wind speed has exceeded the 90th percentile of daily mean wind speed. Before computing the probability of the jet being defined at each grid point on a season-by-season basis. Models are now ranked as a Root Mean Square Error (RMSE) between the models' probability field and that of Reanalysis. In order to calculate RMSE, the CMIP model and Reanalysis must share a common grid thus each model has been re-gridded to a 2x2 degree grid using a bilinear interpolation technique. The impact of this regridding is believed to be relatively minor given the large scale of the mid-latitude jet structure.

Figure 10 shows the subsequent jet probability field using this metric. Again, due to the large

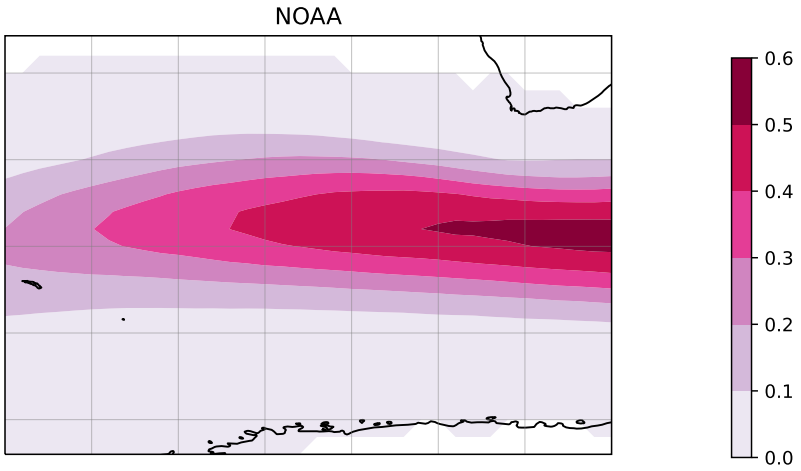
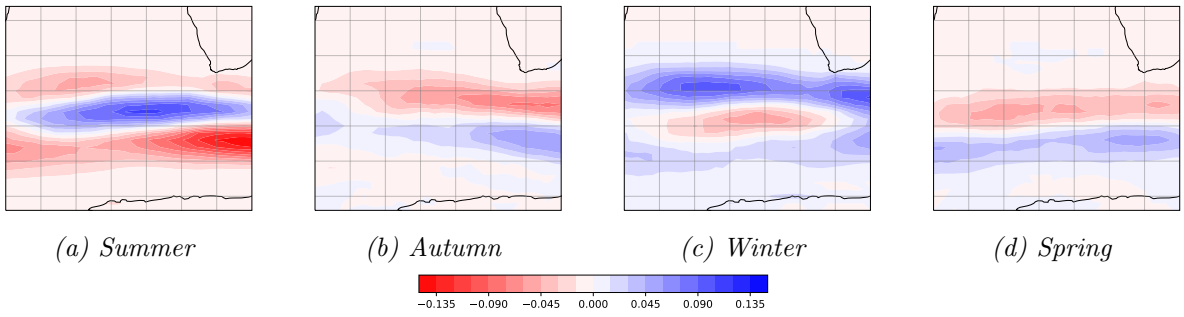


Figure 10: NOAA 20th Century Reanalysis mean jet probability field. Contours indicate probability of jet stream being defined over the grid point in a given day. Correlations between seasonal jet probability and seasonal WRZ precipitation shaded, cross hatching indicates significance at p>0.05. Units reflect probability of a jet stream being defined over a given grid point on a given day. Code Link

scale of the mid-latitude jet structure RMSE is calculated across the full South Atlantic sector as well as the Indian Ocean sector immediately to the east of the Western Cape [26S-60S, 40W-30E]. Similar to the 1-dimensional technique the RMSE of the jet probability field is calculated for each season before combining these scores to form the final score. An area weighted approach is utilised before the mean error across the region is calculated. The area weighting approach is used to account for differences in the area represented by each grid point as longitudinal distances between grid points reduces at higher latitudes.

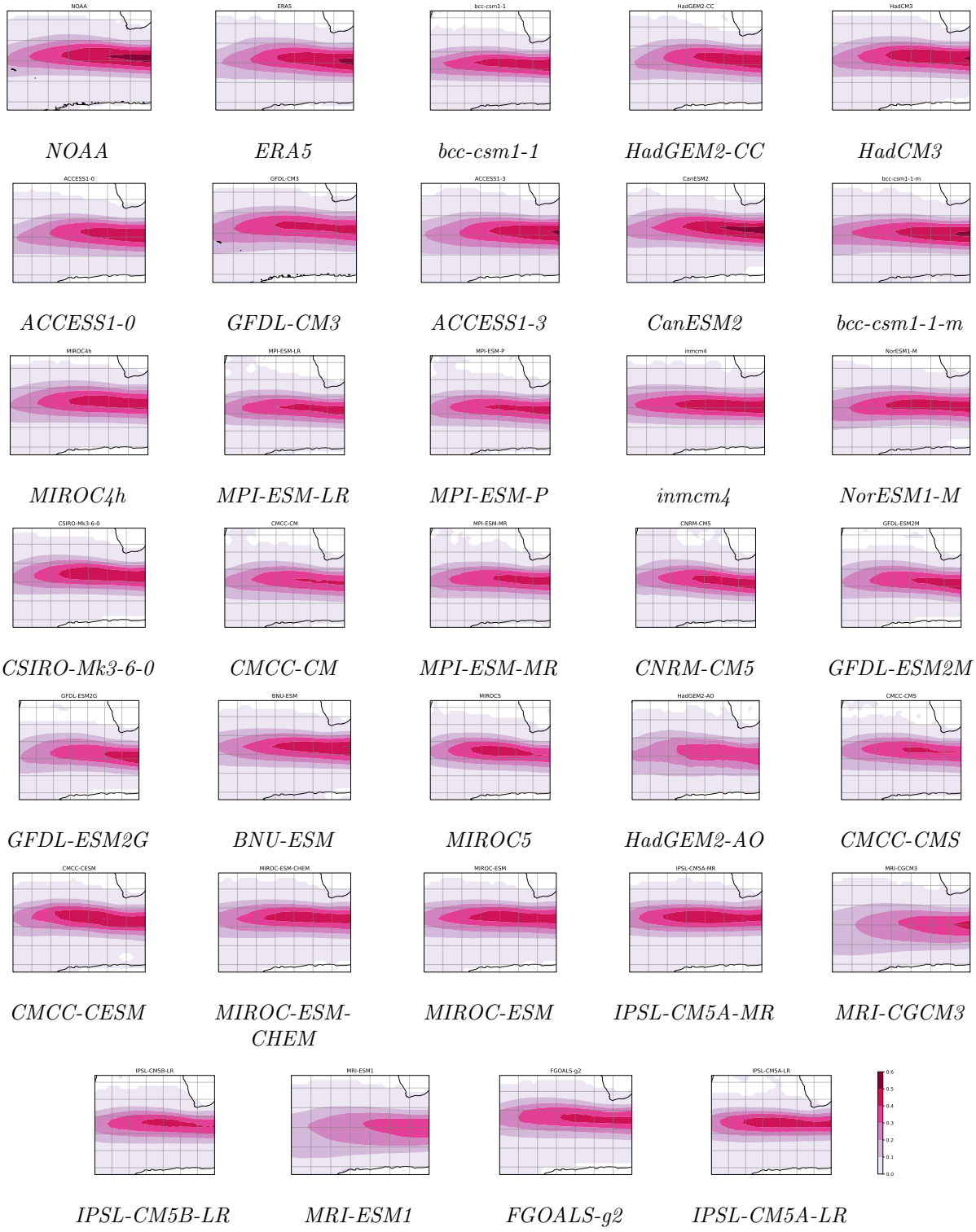


(a) Summer *(b) Autumn* *(c) Winter* *(d) Spring*

Figure 11: NOAA 20th Century Reanalysis, seasonal jet distribution anomaly. Units reflect anomaly of the probability of the jet stream being defined over a given grid point on a given day in a season against the annual probability. Code Link

Figure 11 shows the seasonality of jet density in NOAA 20CR, where a clear increase in jet density at the latitude of the WRZ is observed during winter and to a lesser extent during

shoulder seasons. The double character of jet behaviour in the South Atlantic can also be observed were in peak winter increased jet density is seen at 35 S as well as 65 S while a deficit is observed at 40 S, this same behaviour is seen in Spring and to a lesser extent Autumn (Gallego et al., 2005).



NOAA *ERA5* *bcc-csm1-1* *HadGEM2-CC* *HadCM3*

ACCESS1-0 *GFDL-CM3* *ACCESS1-3* *CanESM2* *bcc-csm1-1-m*

MIROC4h *MPI-ESM-LR* *MPI-ESM-P* *inmcm4* *NorESM1-M*

CSIRO-Mk3-6-0 *CMCC-CM* *MPI-ESM-MR* *CNRM-CM5* *GFDL-ESM2M*

GFDL-ESM2G *BNU-ESM* *MIROC5* *HadGEM2-AO* *CMCC-CMS*

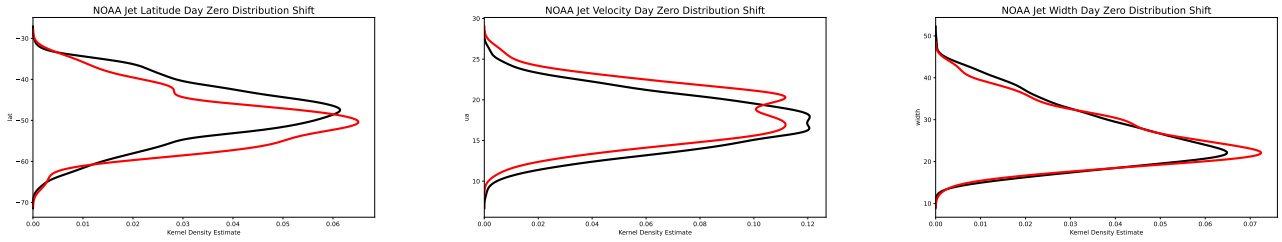
CMCC-CESM *MIROC-ESM-CHEM* *MIROC-ESM* *IPSL-CM5A-MR* *MRI-CGCM3*

IPSL-CM5B-LR *MRI-ESM1* *FGOALS-g2* *IPSL-CM5A-LR*

Figure 12: As in figure 10. Figures are ordered as a function of final RMSE score with the best scoring model top left and worst bottom right Code Link

Day Zero Response

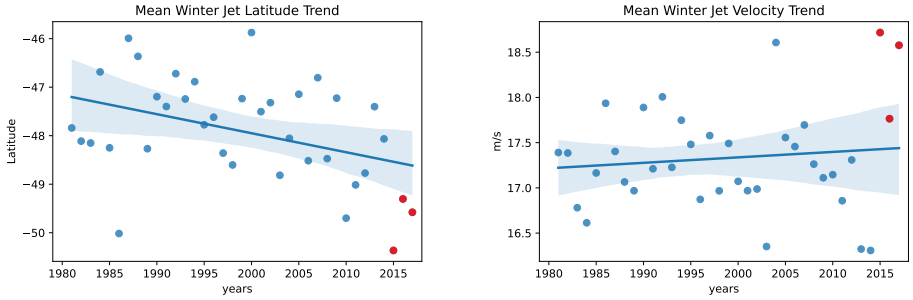
Here as outlined previously the ability of these quantifications to capture the anomalous conditions during the 'Day Zero' drought is considered, where at a minimum, a reliable metric would be expected to capture anomalous behaviour during the 2015-2017 period. Figure 13 shows the anomaly in jet behaviour in the extended winter (AMJJAS) period during 'Day Zero'drought captured by the 1-dimensional jet approximation.



(a) *(b)* *(c)*

Figure 13: Kernel Density Estimates of extended winter (AMJJAS) jet latitude, width and velocity during the period 2015-2017 (red) contrasted against climatology (1980-2017) (black) Code Link

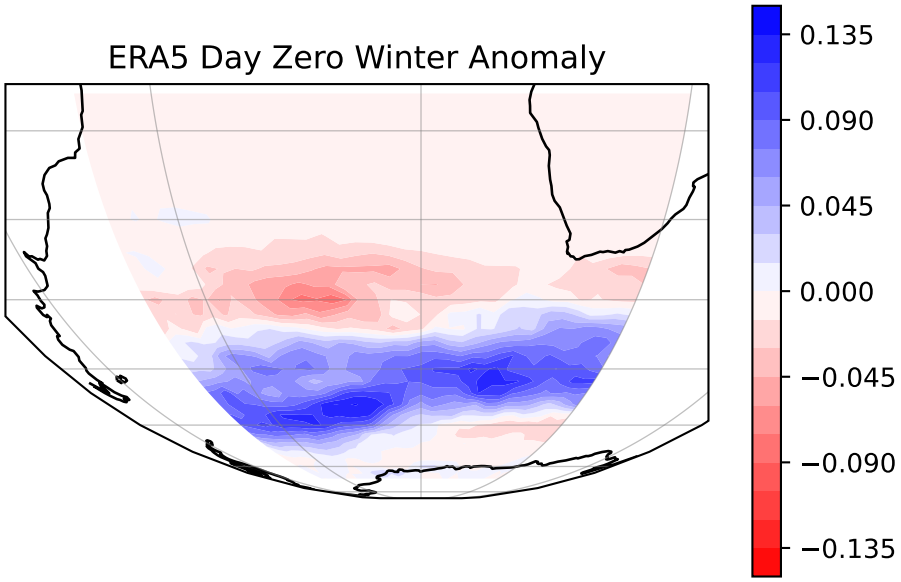
Here a poleward displacement in Jet Latitude is clearly seen during the 'Day Zero' period which is expected and in line with the observed consecutive positive SAM events during these seasons, illustrated by Sousa et al. (2018). A somewhat reduced range of Jet widths are also observed while Jet velocity is seen to be increased. This poleward displacement and intensification of jet streams has been an observed global trend in the late 20th century (Pena-Ortiz et al., 2013) and is thus an expected result if the 'Day Zero' drought was partly contributed to by climate change (Otto et al., 2018). Figure 14 captures this prolonged intensification and poleward displacement of the eddy-driven jet in the South Atlantic, however the 2015-2017 period is still significant in that 3 of the 5 most southerly jet latitude years occur during this period, while mean winter jet velocity was found to be elevated during the 'Day Zero' period.



(a) *(b)*

Figure 14: Mean extended winter Jet latitude and width, blue line represents a linear regression model fitted to the dataset, while light blue shading represents 95% confidence interval of this model computed using a bootstrapping method. Code Link

Figure 15 utilizes the 2-Dimension metric and captures similarly anomalous behaviour during the 'Day Zero' period, with reduced jet activity observed north of 40°S and a marked increase is seen at higher latitudes. Somewhat reduced jet activity is seen at extreme southern latitudes at the longitude of the WRZ, possibly an indicator of a reduced double structured phase.



(a)

Figure 15: ERA5 jet distribution anomaly during the 2015-2017 period. Units reflect anomaly of the probability of the jet stream being defined over a given grid point on a given day during the day zero drought against a 1979-2019 baseline. Code Link

Results

Both metrics are seen to capture the dynamics during the 'Day Zero' drought successfully and are included in the final results in figure 16.

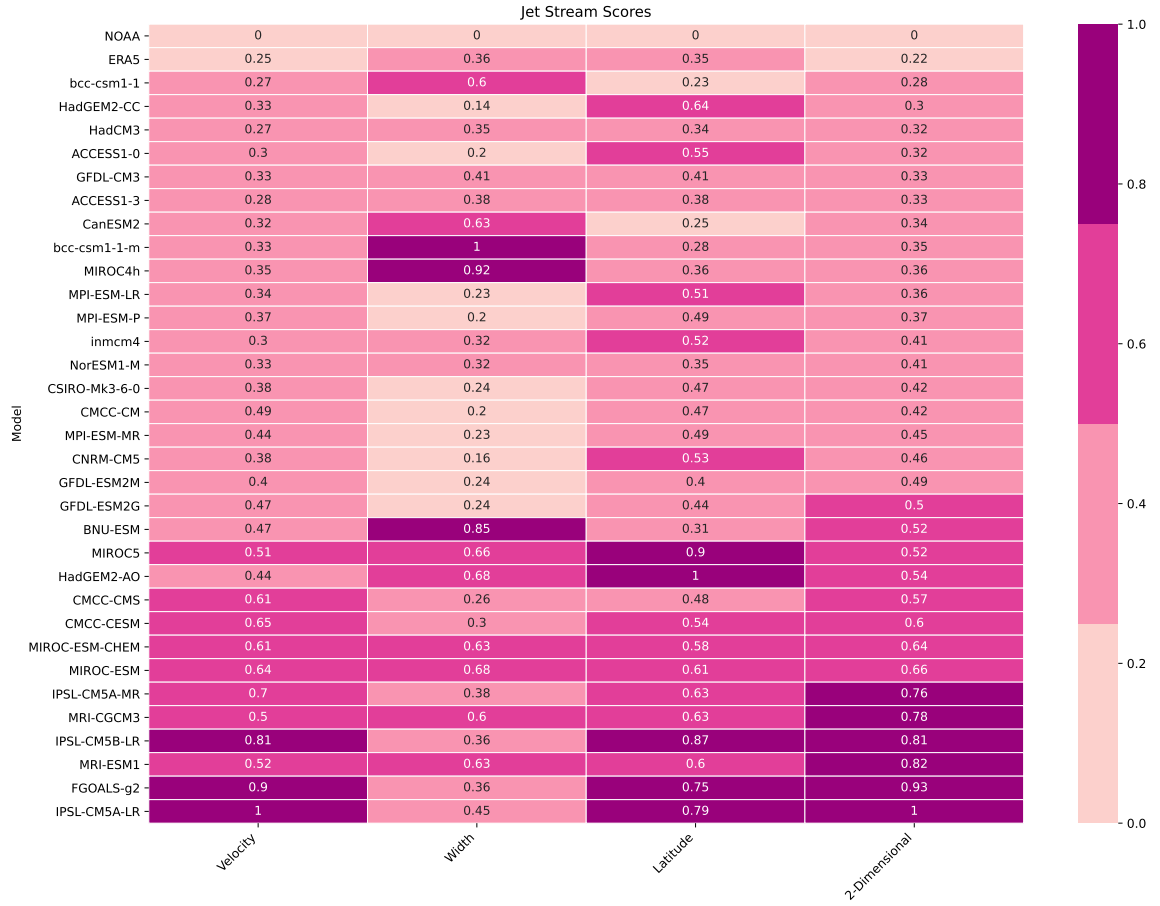


Figure 16: Area difference Scores and Rank across the different metrics, assessed against seasonal distributions(1d) or probabilities(2d) Code Link

Strong agreements between 2-dimensional scores and that of 1-dimensional latitude and to a lesser extent width exist, indicating that despite the known limitation of the 1-dimensional metric defining only a single jet the 1-dimensional technique successfully differentiates between models. 2-Dimensional scores are however preferred for the final results primarily due to this method not requiring the added complexity of interfacing the latitude and width scores and having to decide whether or not to include velocity in the final score. In reality, this choice is of little consequence as scores being are very similar across the various metrics. Thus, the final model section will be made using these '2-dimensional' scores.

South Atlantic Subtropical High

Introduction

The South Atlantic Subtropical High (SASH) is a semipermanent high-pressure system, also known as the Saint Helena High due to its frequent position near the island nation at 25° S 15° W. The system is largely responsible for the low rainfall and persistent SE'ly winds experienced by the WRZ during austral summer months when the SASH is positioned further east and presents a more consistent annular structure. In austral winter a zonally wider system positioned further north and west (Hastenrath, 2012) ultimately allows for more moisture bearing cold fronts to make landfall in the WRZ (Reboita et al., 2019). The non-seasonal component of SASH variability also has a strong influence on WRZ climate. Burls et al., (2019) attribute increased blocking patterns from SASH ridging events as the primary mechanism behind the reduced rainfall during the 'Day Zero' drought period.

Reboita et al., (2019) describe the drivers responsible for the genesis and variability of the SASH as: subsidence from the southern edge of the Hadley Cell, sea air interactions, monsoon driven subsidence from neighbouring continents as well as ENSO and SAM related teleconnections. Richter, Mechoso & Robertson, (2008) use an uncoupled numerical weather model to study the sensitivity of the SASH to perturbations in prescribed orography, soil moisture and sea surface temperatures. Removing South American topography highlights the role the Andes Mountain range plays in blocking westerly flows into the Atlantic and inducing subsidence, where in the absence of South American topography a single zonally homogenous high pressure ridge forms over the South Atlantic and south Pacific oceans. This orographically induced subsidence is the dominant driver of subsidence across the subtropical ocean basin. A possible concern in low resolution CMIP5 models is this disproportionate influence of the relatively narrow mountain chain not being accurately captured by the low-resolution models. While this orographically induced subsidence is responsible for the formation of the SASH, tropical convection and equatorial SST is found to play a key role in modifying SASH position. Richter, Mechoso & Robertson, (2008) find introducinga persistent warm bias in the equatorial southern tropical Atlantic , a common systematic error in CMIP5 and other coupled GCM's (Toniazzo & Woolnough, 2014), results in an unrealistically zonally elongated and westerly displaced SASH. Richter, Mechoso & Robertson (2008) find reducing soil moisture over tropical Africa to desert like conditions resulted in increased zonal pressure gradients across the SASH although the western bias continues to persist. Perturbations in zonal sea surface temperatures are found to have little influence on modifying the SASH in their simulations, however utilizing an uncoupled ocean-atmosphere model may have induced this behaviour (Cabos et al., 2017; Reboita et al., 2019).

The seasonality of the SASH is driven in part by Northern Hemisphere tropical convection. The increased intensity and wider SASH during Austral winter months is a result of increased subsidence from tropical convection during the boreal summer Monsoon's in tropical Africa and India (Rodwell & Hoskins, 2001). While subsidence is the primary driver of the SASH, a number of dynamics contribute to the anticyclonic rotation and position of the system. The northern limb of the SASH is formed as a Kelvin wave response to heating over the

African continent, while the cool Benguela current to the east acts to enhances the eastern limb of the SASH (Peterson & Stramma, 1991). Upwelling of cool water by the northerly flowing Benguela current along the southwestern African coast results in cool sea-surface temperatures and subsidence of dry air. Equatorward flow is then further enhanced by strong temperature gradients between the sea and warm Namib desert, while on the western side of the South Atlantic the warm Brazilian current induces advection and further warming. Increased convection along the Brazilian current during summer results in increased zonal temperature gradients and a more annular SASH (Reboita et al., 2019). This annular structure together with southward displacement, driven by seasonal Hadley cell oscillation, is responsible for increased blocking frequencies in the southwestern Atlantic during summer contributing to the dry and windy summer conditions in the WRZ.

He et al., (2017) find that increased tropospheric stability and changes in diabatic heating under a future warmer climate will result in reduced SASH intensity while continued expansion of the Hadley cell is expected to displace the SASH further south, consistent with more summer like conditions (Sun, Cook & Vizy, 2017; Reboita et al., 2019). Sousa et al., (2018) highlight how a poleward expansion of the Hadley Circulation Cell played a role in driving a poleward shift in moisture corridors in the WRZ as well as how increased high-pressure ridging occurring over the WRZ, contributed to the 'Day Zero' drought. Burls et al., (2019) find that increased prevalence of post-frontal-high-pressure conditions have contributed to the increased prevalence of drought in the WRZ, where ridging events enhanced by Hadley Circulation Cell expansion act to suppress orographically induced rainfall in the region. Drought as a result of increased ridging frequencies are not unique to the WRZ; a similar phenomenon occurred during the similarly severe 2013-2014 drought in California and was dubbed the 'ridiculously resilient ridge' (Swain et al., 2014). Increased GHG emissions have been linked to the increased likelihood of both these events. Expansion of the Hadley Cell across the Southern Hemisphere has been attributed as the primary mechanism behind increased ridging events in the WRZ (Burls et al., 2019).

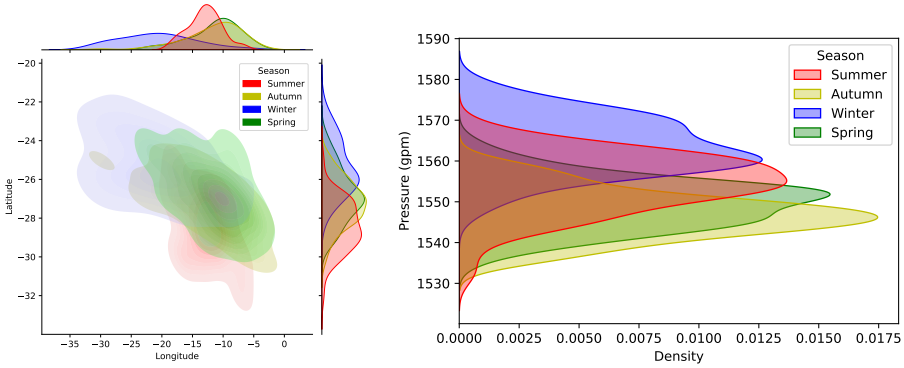
Thus, the dynamics of the SASH are multifaceted and sensitive to a number of perturbations. Here the primary descriptors of SASH, namely central position, intensity, and extent, are first quantified before the CMIP5 models and reanalysis are contrasted and compared with literature. Scoring methods to rank the CMIP5 models are constructed using the SASH descriptors. The ability of these descriptors to capture the 'Day Zero' drought is also considered and the final scoring technique chosen as a function of this.

Analysis

1-Dimensional Analysis

Sun, Cook & Vizy, (2017) developed a climatology of the interannual variability of the South Atlantic Subtropical High (SASH) by considering the latitude, longitude and intensity of maximum 850 hPa geopotential height across the meteorological seasons, utilizing a higher level to lessen the influence of surface topography. Seasons where the position of the SASH is more than 3 std-deviations from the mean position are considered as undefined. These are typically

seasons with increased synoptic variability in the South Atlantic resulting in maximum geopotential occurring over the African continent. The region from the equator to 60°S and 50°W to 20° E is considered here. This method is applied to the 29 available CMIP5 models utilizing the historical experiment and NOAA 20th Century Reanalysis over the period 1950-2005. ERA5 reanalysis over the period 1980-2017 is also considered for reference. Positions of the SASH are considered using seasonal mean geopotential rather than computing climatology as a function of daily SASH centres, as the misinterpretation of transient anticyclones as SASH centres can perturb results (Gilliland and Keim 2018a).



(a) *(b)*

Figure 17: Kernel density estimates of seasonal mean SASH centres (left) and maximum pressure (right) in NOAA 20th Century Reanalysis (1950-2005) Code Link

Of the 220 seasons considered, seasons where the SASH is considered undefined vary from 1-7 seasons across the CMIP5 models while 4 and 3 seasons are excluded from NOAA 20CR and ERA5 respectively. MIROC-ESM is however an exception where no variability in either longitude or latitude of SASH is observed due to maximum geopotential occurring over the African continent in all seasons, the model is thus excluded from further analysis.

Next, using the same technique used in that of the 1-dimensional jet stream analysis, a Kernel Density estimation is constructed from these seasonal values before models are scored against NOAA 20CR. Scores are calculated as a function of mean absolute error between the kernel density estimations, on a season-by-season basis. . Figure 17 provides a visualization of the kernel density estimates constructed for NOAA 20CR across the meteorological seasons. Here the seasonality of the SASH in both the zonal and meridional is captured, where a displacement to the south and east is seen in Summer contrasted to a more north and west position during winter, shoulder seasons are seen to extend the furthest East and have a larger range than that of winter or Summer. Median geopotential is seen to greatest in winter and weakest in shoulder seasons.

Table 2 shows the results of the SASH tracking across the CMIP5 models. Of note is the lack of evidence of a evidence of a systemic southern Bias. Instead, the best performing CMIP5 models accurately capture the mean position of the SASH. ACCESS1-0 is seen to be more similar to NOAA 20CR than that of ERA5, potentially due to the reduced variance of

Table 2: Summary of results from SASH tracks. A positive Latitude Difference indicates a northern displacement while a positive Longitude Difference indicates an eastern displacement Code Link

	Latitude				Longitude				Pressure (gpm)			
Model	Mean	Difference	Std.	Range	Mean	Difference	Std.	Range	Mean	Difference	Std.	Range
IPSL-CM5A-LR	-26.49	0.52	2.2	11.37	-22.08	-8.03	8.35	41.25	1546.62	-6.69	10.97	54.59
MPI-ESM-MR	-26.67	0.33	2.75	11.19	-20.54	-6.49	12.12	46.88	1558.66	5.36	12.09	59.88
MRI-ESM1	-24.14	2.87	2.2	8.97	6.45	20.5	15.09	49.5	1573.27	19.96	13.62	70.92
MIROC-ESM-CHEM	-26.19	0.82	2.38	11.16	-19.14	-5.09	9.71	36.56	1563.06	9.75	9.77	61.15
EC-EARTH	-26.6	0.41	2.59	10.09	-20.46	-6.42	11.73	45.0	1544.92	-8.39	14.09	62.0
HadGEM2-CC	-26.21	0.8	2.23	11.25	-14.93	-0.88	7.75	39.38	1548.54	-4.77	8.81	51.97
bcc-csm1-1-m	-26.04	0.96	2.54	12.34	-19.78	-5.73	11.33	41.62	1571.2	17.89	8.12	45.74
FGOALS-g2	-24.8	2.21	2.16	8.37	-21.9	-7.85	6.51	33.75	1569.64	16.33	9.02	44.8
ACCESS1-0	-26.58	0.43	2.51	12.5	-14.54	-0.49	7.52	39.38	1552.25	-1.06	7.0	38.34
GFDL-CM3	-27.65	-0.65	2.14	10.0	-22.14	-8.09	7.71	37.5	1562.09	8.78	11.87	61.95
CanESM2	-26.06	0.94	2.06	8.37	-22.38	-8.33	8.55	36.56	1559.83	6.53	8.67	45.45
IPSL-CM5B-LR	-25.56	1.44	2.14	9.47	-19.7	-5.65	8.87	41.25	1540.18	-13.13	8.12	40.91
MPI-ESM-LR	-26.61	0.4	2.77	11.19	-22.14	-8.09	10.97	41.25	1556.29	2.98	12.44	62.4
CMCC-CM	-25.68	1.33	2.95	12.72	-19.4	-5.35	10.95	45.75	1550.53	-2.78	9.26	51.65
GFDL-ESM2G	-26.84	0.16	2.12	10.11	-21.32	-7.27	8.25	35.0	1555.78	2.47	11.24	59.29
MIROC5	-27.43	-0.43	2.47	11.21	-17.99	-3.94	7.62	33.75	1564.33	11.02	11.1	54.7
MPI-ESM-P	-26.67	0.33	2.7	11.19	-21.37	-7.32	11.41	41.25	1555.34	2.03	12.69	66.58
CMCC-CMS	-26.51	0.49	2.72	13.06	-19.81	-5.76	12.01	41.25	1557.42	4.11	10.16	61.29
CNRM-CM5	-26.57	0.44	3.09	14.01	-16.14	-2.09	10.7	43.59	1548.61	-4.7	9.91	43.41
IPSL-CM5A-MR	-26.25	0.75	2.84	11.41	-11.93	2.12	17.78	62.5	1561.53	8.22	23.53	78.16
ACCESS1-3	-26.94	0.07	2.61	12.5	-17.73	-3.68	8.3	41.25	1559.05	5.74	7.45	37.1
MRI-CGCM3	-23.93	3.08	2.07	7.85	8.37	22.42	14.01	50.62	1572.27	18.96	14.33	73.14
bcc-csm1-1	-26.1	0.91	2.09	8.37	-18.22	-4.17	7.38	36.56	1565.8	12.49	8.86	45.23
CMCC-CESM	-26.23	0.77	2.09	11.13	-23.6	-9.55	13.11	33.75	1563.2	9.89	10.73	60.77
HadGEM2-AO	-28.51	-1.51	4.17	21.25	-13.41	0.64	15.35	69.38	1562.54	9.23	14.51	77.08
BNU-ESM	-26.44	0.57	2.17	8.37	-18.16	-4.11	8.25	36.56	1570.44	17.14	9.46	51.7
NorESM1-M	-28.55	-1.54	1.92	7.58	-22.34	-8.29	7.38	35.0	1574.38	21.07	10.5	55.15
HadCM3	-27.13	-0.13	2.73	12.5	-19.88	-5.83	7.49	37.5	1532.52	-20.78	10.51	58.26
GFDL-ESM2M	-26.31	0.69	2.5	12.13	-21.33	-7.28	8.98	40.0	1554.93	1.62	10.46	45.34
NOAA	-27.0	-0.0	1.69	10.0	-14.05	-0.0	5.88	26.0	1553.31	0.0	8.99	45.02
ERA5	-26.37	0.63	1.97	9.25	-15.39	-1.34	6.26	30.75	1556.0	2.69	9.24	45.56

NOAA 20CR compared to ERA5 perhaps as a function of lower resolution in NOAA 20CR, as well as ERA5 only considering 1980-2005 and thus not containing 'pre-satellite era' seasons where reduced variance is captured by reanalysis due to fewer input data. Standard deviation of SASH latitude and longitude is seen to be far larger in the CMIP5 models than in both reanalysis models. A consistent western bias is seen across the CMIP5 models, with the notable exception of both Meteorological Research Institute (MRI-ESM1 & MRI-CGCM3) models where an extreme eastern bias is observed. A similar extreme eastern bias occurs in some seasons in IPSL-CM5A-MR.

Figure 18 provides a visualization of CMIP5 SASH positioning vs NOAA 20CR. A somewhat increased spread of SASH centres is seen in the CMIP5 models with HadGEM2-AO being the extreme example. Richter, Mechoso & Robertson, (2008) associate this deficiency in other coupled ocean–atmosphere GCMs with the systematic overprediction of convection over the tropical land surfaces.

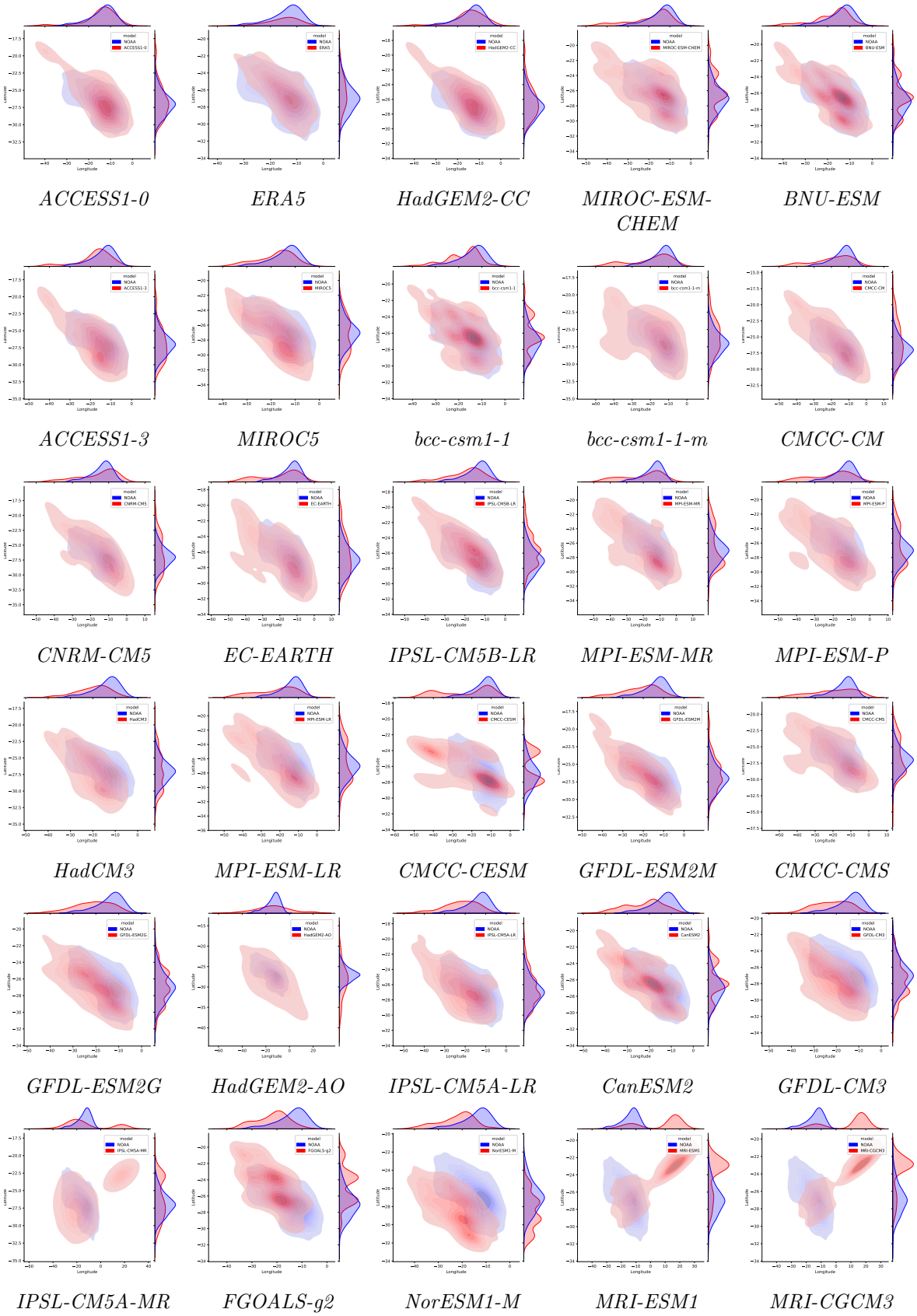


Figure 18: CMIP5 Seasonal SASH centers (red) compared to NOAA 20th Century Reanalysis (blue). Figure is ordered from best to worst performing models as scored using the meridional metric. Best top right to worst bottom left. [Code Link]

2-Dimensional Analysis

To contrast against the results from the 1-dimensional analysis a second 2-dimensional technique is considered here. The same 850 hPa geopotential is considered, but rather than defining SASH as the maximum geopotential, it is defined as the area defined by the contour of 90th percentile geopotential height (gpm), in a similar fashion to that of the 2-Dimensional Jet analyses and again inspired by Sousa et al., (2018) who use the extent of the 1020 hPa mean sea-level pressure contour to visualize shifts in the SASH during the Dday Zero' drought. This 90th percentile extent is defined across each season over the period 1950 – 2005 and a probability field of the SASH being defined over a given point is constructed for each season. Scoring is conducted using the same technique as that of the 2-dimensional jet analysis where first models are regridded to a common 2x2 degree grid, using a bilinear interpolation technique before the SASH extent probability field is constructed. This regridding is believed to have little influence to final results given the large scale structure of the SASH. Models are then scored as a function of RMSE as in the 2-Dimensional Jet Analysis, where an area weighted approach is again utilised before calculating the mean RMSE across the region.

Figure 19 captures the seasonality in the SASH probability field for NOAA 20CR. Here a reduced extent and more annular structure is evident in Summer whilst an increased extent is evident in winter.

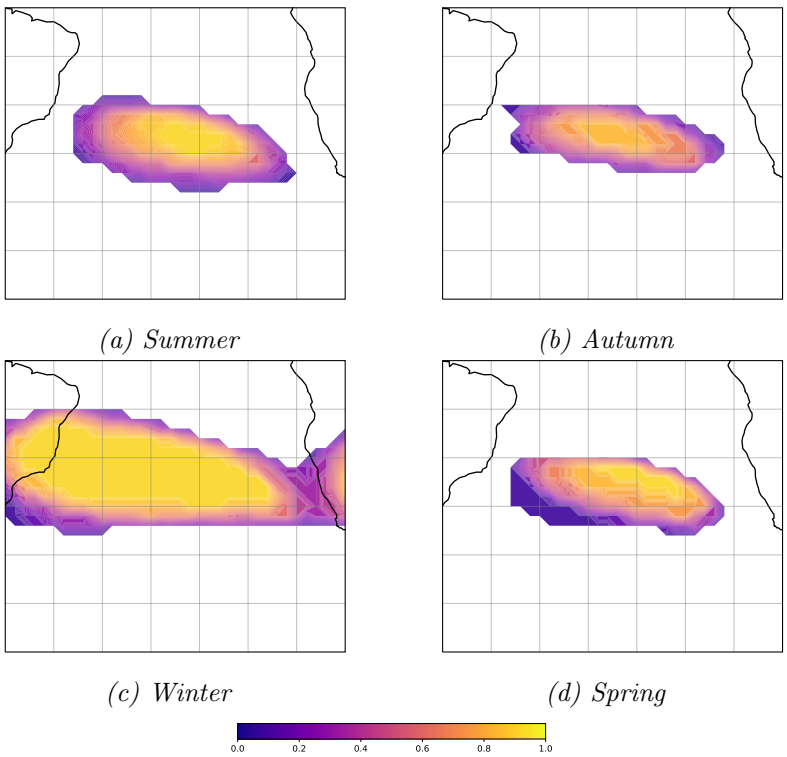


Figure 19: NOAA 20th Century Reanalysis 850 hPa Seasonal Mean Geopotential - 90th percentile contour probability field. Code Link

Figure 20 captures the anomaly between CMIP5 models and NOAA 20CR probability fields, this time considering the probability field across all seasons. Models are ordered as a function of final scores with best scoring top left and worst bottom right. ERA5 over the period 1980-2005, included as a reference, has a somewhat northern bias compared to NOAA 20CR. The best performing CMIP5 models, HadGEM2-CC and ACCESS1-0, display a similar northern bias to ERA5. Notably these same 2 models also perform best in the 1-dimensional analysis. The western bias, observed in the 1-dimensional analysis, is seen to some extent in all CMIP5 models.

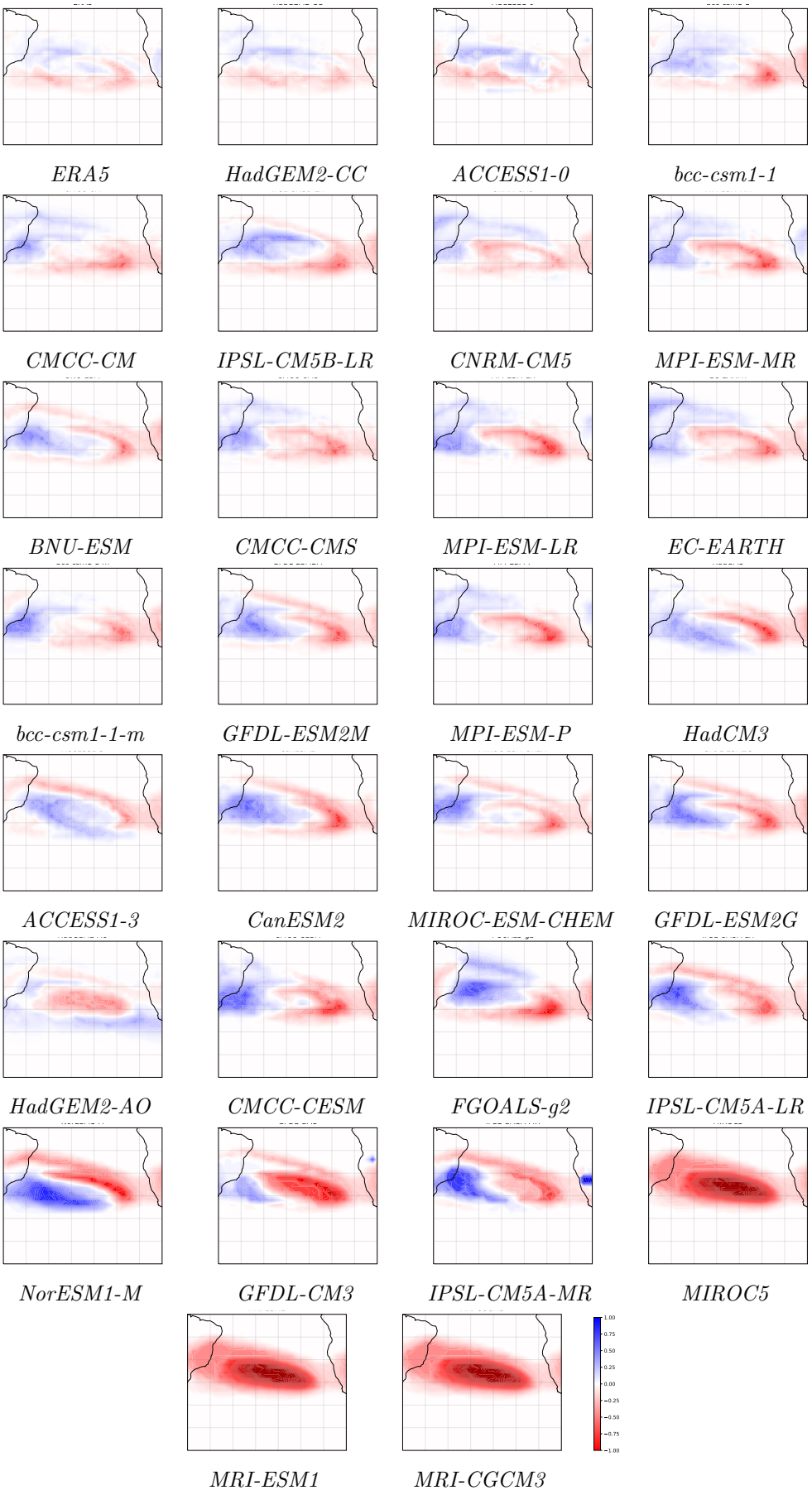
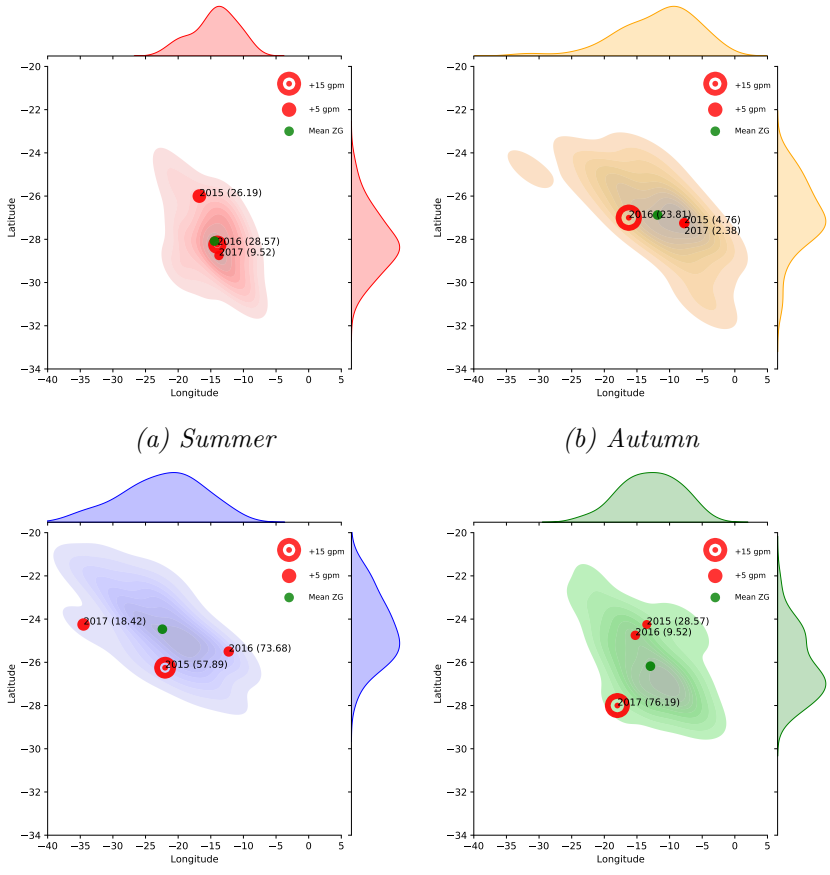


Figure 20: South Atlantic Subtropical High extent probability field anomaly contrasted against NOAA 20th Century Reanalysis. Positive anomaly (blue) indicates increased SASH prevalence while a negative anomaly (red) indicates reduced SASH prevalence. Code Link

Day Zero Response

Here the ability of these SASH metrics to capture the anomalous SASH behaviour during the 'Day Zero' drought is assessed, where CMIP5 models that could best identify future drought risk in the WRZ region should, at a minimum, be able to capture the dynamics behind such events.

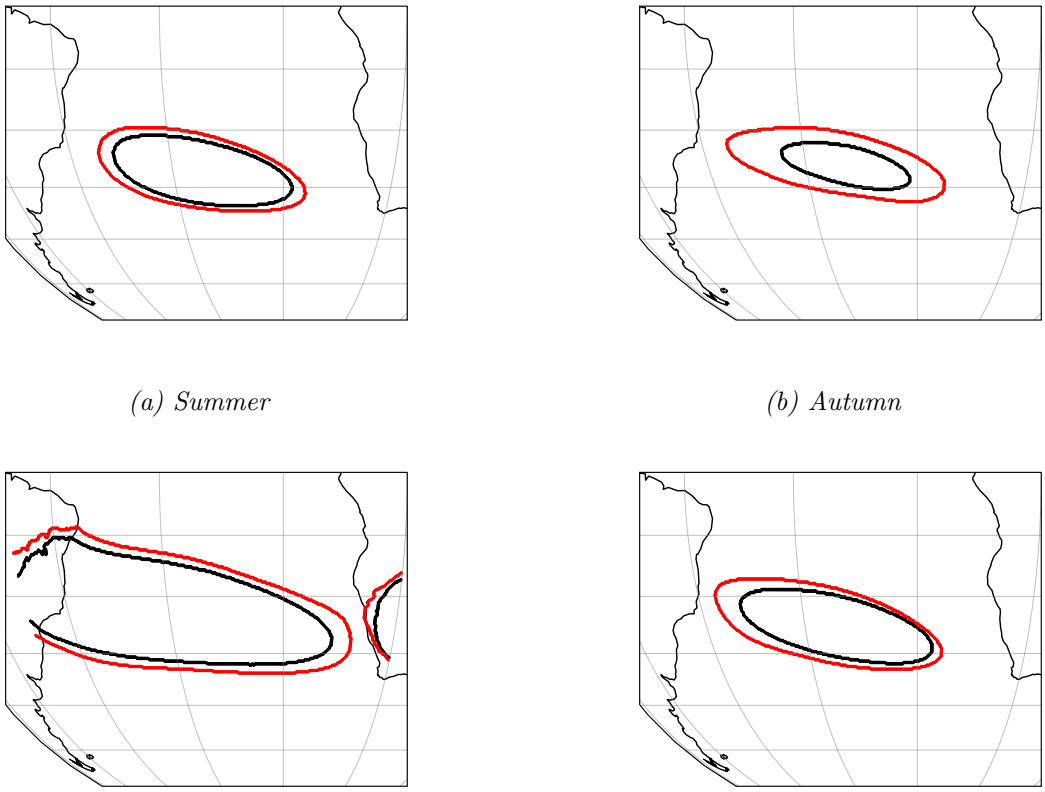
Figure 21 captures the position and intensity of the SASH centres during the 'Day Zero' period, as captured by ERA5 reanalysis, here no clear or consistent anomaly is evident. Mahlalela, Blamey & Reason, (2019) find a failure of early winter rainfall as key driver of drought during the period and indeed the 2015 and 2017 Autumn, where rainfall is markedly reduced, has a somewhat eastern bias. The Spring of 2015 and 2016 where rainfall is also reduced a northern and western bias is seen contrary to what would be expected. The seasons where rainfall is most reduced do appear to have reduced intensity compared to seasons with increased rainfall however these reduced intensities are close to that of mean values.



(a) Summer *(b) Autumn*

(c) Winter *(d) Spring*

Figure 21: Kernel Density Estimation of seasonal SASH central location in ERA5 reanalysis from 1980-2005. Scatter points indicate SASH centers during the 'Day Zero' drought, with years labelled and the corresponding percentile of that season's rainfall in parentheses. Code Link



(a) Summer *(b) Autumn*

(c) Winter *(d) Spring*

Figure 22: Extent of the 90th percentile gpm contour at 850hPa in ERA5 Reanalysis. Climatology (1980-2017) in Black compared to 'Day Zero' (2015-2017) in red. Code Link

Figure 22 instead considers the extent of the 90th percentile gpm contour during the 'Day Zero' period, contrasted against climatology. Here a clearer picture of anomalous behaviour during the period can be seen. Specifically, in Autumn and Winter, an increased extent in the south and eastern sector is evident, the increased extent in this sector is likely driven by the persistent ridging that occurred during the 'Day Zero' period. Thus the 2-dimensional technique is preferred as it has the ability to account for the influence of transient ridges which can be a dominant feature of drought in the WRZ.

Results

The final scores are included in figure 23. Models that perform well in capturing the meridional position of the SASH also perform well in the zonal SASH position, However, pressure does not correlate well to any of the other metrics. The models that perform best in the Meridional and Zonal direction, HadGEM2-CC, ACCESS1-0 and bcc-csm1-1, also perform well in the 2-dimension 'contour' scoring technique. The same is true for the worst performing scores. Thus, the model selection will utilise theses 2-dimensional 'contour' scores.

(a) Summer

Figure 23: Results calculate as Mean Absolute Error between Kernel Density Estimation of SASH centers. Results are normalized such that the worst performing models is scored 1 and a model identical to that of NOAA20R scores zero. Net score is the sum of Meridional and Zonal Scores, again this result is normalized. Code Link

South Atlantic Cold Fronts

Introduction

Atmospheric fronts play a key role in weather and climate systems at mid-latitudes and along with associated mid-latitude cyclones are a primary mechanism for meridional heat and moisture transport. Cold fronts, embedded within mid-latitude cyclones or as a result of atmospheric troughs, are the primary source of moisture for the WRZ. The upliftment of warmer air ahead of the cold front, enhanced by the orography of the Western Cape, results in rainfall over the region. The midlatitude jet stream is largely responsible for the steering of eastbound midlatitude cyclones and associated cold fronts, while the presence of ridging activity in the region may serve to suppress associated upliftment and reduce subsequent rainfall (Odoulami, Wolski & New, 2021).

Lennard & Hegerl, (2015) use a station-based approach to classify synoptic types over the WRZ and note an increasing trend in dry synoptic states during winter in the WRZ over the period 1979-2009, while Odoulami, Wolski & New, (2021) find the 'Day Zero' drought was typified by a reduction in the number of rain-bearing synoptic states as well as a reduction in the amount of rainfall during wet synoptic states. Intuitively this could be associated with a reduced number of fronts impacting the WRZ. Burls et al., (2019), study cold fronts directly during the 'Day Zero' period and instead find a reduction in rainfall associated with cold fronts rather than a reduction in the absolute number of fronts, linking this reduced rainfall to subsidence from increasing ridging activity as a result of Hadley cell expansion. Mechanisms which strongly influence cold fronts have previously been considered in the jet stream and SASH metrics, however here considering these transient fronts directly allows for comparison of a model's ability to capture dynamic systems while also including a metric that does not consider a mean state and thus may be less susceptible to influence from model tuning (Santer et al., 2009).

Scoring models against cold fronts directly rather than considering midlatitude cyclones is preferred as whilst the two are typically associated the influence of a cold front can extend thousands of kilometres from the centre of the cyclone and thus whether a cold front impacts the WRZ is not necessarily predicated by the position of the associated cyclone. Thus, considering cold fronts directly can provide a more robust picture of WRZ dynamics. Automatic identification of cold fronts is not a trivial task where manual identification by a weather forecaster is still the preferred method of many national meteorological institutes. Automatic detection does however provide many advantages particularly so in developing climatologys of frontal passages from reanalysis and identifying future projected shifts from climate models (Simmonds, Keay & Tristram Bye, 2012).

Selection of front tracking algorithms can have a significant influence on results (Schemm, Rudeva & Simmonds, 2015). Broadly frontal detection techniques can be classified by either considering thermal gradients, such as the thermal frontal parameter developed by Renard & Clarke, (1965) which utilizes the first order derivative of wet bulb potential temperature, potential temperature or equivalent potential temperature to define a front, or by considering

shifts in the wind field, such as that developed by Simmonds, Keay & Tristram Bye, (2012) where a front is defined as a wind shift from the northwest to the southwest quadrant over a 6-hour period, along with a change in meridional wind speed that exceeds a given threshold. Schemm, Rudeva & Simmonds, (2015) contrast these two techniques across various synoptic situations and find both perform suitably well. The wind shift method is found to be better suited for the detection of strongly elongated, meridionally orientated moving cold fronts, which are typical of South Atlantic or Southern Ocean conditions and is thus utilised to develop this metric.

Analysis

The method developed by Simmonds, Keay & Tristram Bye, (2012) and utilised by Burls et al., (2019) to study changes in rainfall events associated with the passage of cold fronts during the 'Day Zero' drought is adapted to score the performance of CMIP5 models in capturing cold front dynamics in the South Atlantic region. Here 6-hourly synoptically sampled meridional and zonal wind speeds from 16 CMIP5 models, with sufficiently high temporal resolution data available, as well as ERA5 and NOAA 20CR reanalysis models, are considered and a climatology of cold front activity in the South Atlantic region developed.

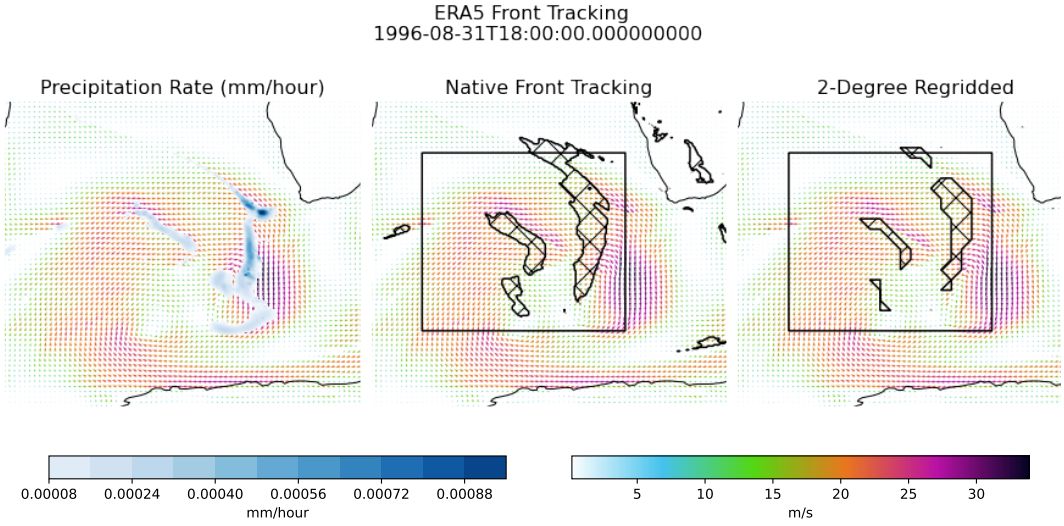


Figure 24: ERA5 Frontal tracking. wind speed and direction with precipitation rate overlayed in blue (left), and the frontal region indicated with cross hatching (middle), finally the frontal region defined after the wind values have been regridded to a 2x2 grid (right) Code Link

A cold front is initially defined as a region were over a 6-hour period the wind has shifted from the northwest to the southwest quadrant (a shift in a westerly wind from a negative to positive meridional component), combined with a change in meridional wind speed exceeding 2 m/s. This is sampled at 00z and 12z. Next, utilizing a segmentation technique grid points defined as a front are linked with neighbouring grid points with an 8-point connectivity, achieved utilizing the watershed segmentation algorithm, common in image processing and available

through the scikit-image python library (Van der Walt et al., 2014). Finally, the length of identified fronts is approximated as an ellipse and only fronts which exceed 500 km in length are considered.

Simmonds, Keay & Tristram Bye, (2012) include further steps to improve visualization and tracking of fronts by considering only easternmost grid points of each front before fitting a resistant smoothing metric to longitude values to create a visually appealing frontal line. Further through defining the 'center of gravity' of the front a cyclone tracking technique adapted from Simmonds, Burke & Keay, (2008) is utilized to track frontal passages, where only fronts that last for longer than a day are included in their final analyses. These final steps however present significant challenges to implement across all the CMIP5 models while tracking of individual fronts is not a priority and as such these steps have been omitted. Not including a minimum time span for a front to be defined does appear to result in increased activity over land areas, especially in higher resolution models, and as such models are only assessed over an oceanic region defined by 60° south to 25° south latitude and 30° west to 10° east longitude, indicated in figure 24.

Cold front identifying algorithms are known to be sensitive to the resolution of the model considered and thus two separate scoring techniques are considered. First frontal zones are defined in the model's native resolution before calculating the average number of fronts defined at each grid point for the meteorological seasons. This seasonal climatology is then regridded to a common 2x2 degree grid to allow for scoring of models against reanalysis, utilising the same RMSE area weighted approach as utilised in the 2-Dimensional SASH and jet stream analyses. The second technique uses the same method however the 6-hourly wind component fields are regridded to a common 2x2 degree grid before the frontal analysis takes place. Figure 24 highlights the influence of interpolating ERA5 reanalysis to a lower resolution before defining frontal zones, where reduced detail of fronts are captured and smaller frontal zones are no longer defined, precipitation rate is included in figure 24 to highlight the role cold fronts play in precipitation at subtropical latitudes.

Across the CMIP5 and reanalysis models the total number of fronts captured per season varies greatly, such that without a normalization factor the ERA5 reanalysis model scores worse than all of the CMIP5 models, when computing fronts from both native and regridded data. ERA5 captures sgnificantly less frontal activity in the scoring region during the period 1981-2005 than that of NOAA 20CR. Thus, the seasonal climatologies are normalized by the maximum number of fronts observed in all seasons. Considering instead the ratio of fronts observed in a season to the maximum observed in all seasons in a model. While this normalization approach no longer considers the ability of a model to capture the absolute number of fronts that occur per season the ability of a model to capture the seasonality of cold front occurrences is still preserved. The ability to accurately capture the absolute number of fronts is an important metric to consider, but the sensitivity of the cold front algorithm to the level of detail captured by the model is such that this is not easily captured.

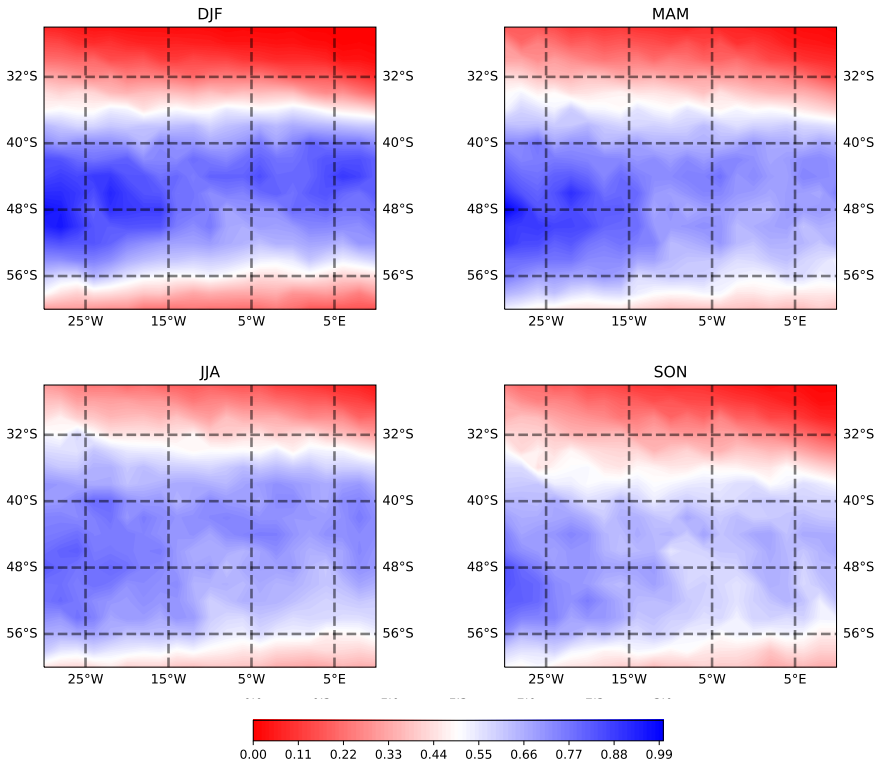
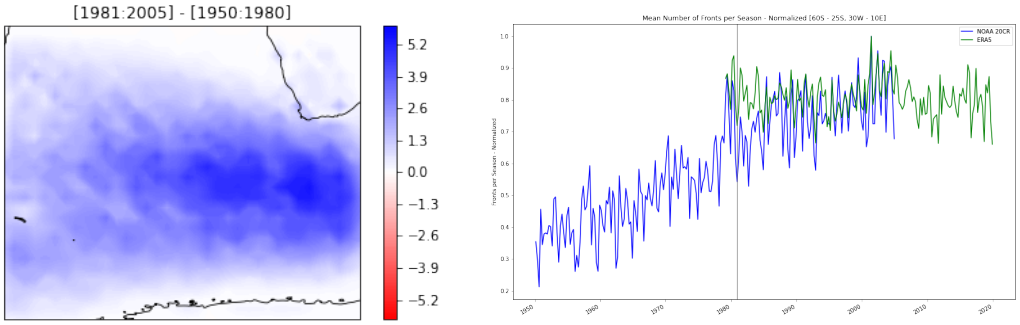


Figure 25: Normalized seasonal frontal climatology in NOAA 20th Century Reanalysis. (top left) Summer, (top right) Autumn, (bottom left) winter, (bottom right) spring. Code Link

Figure 25 highlights the seasonality as captured by the NOAA 20CR model, where a zonal translation of the belt of maximum frontal activity from Summer to Winter occurs, as well as an expansion of the area of maximum activity during winter and shoulder seasons. A final challenge is deciding on an appropriate baseline for scoring the models, previously the maximum 1950-2005 period has been used as default. A complication experienced by all reanalysis products over this time span is that upper air and satellite data only become readily available after 1979 resulting in non-climatic discontinuities within the reanalysis. To avoid this NOAA 20CR assimilates only near surface conventional observations that have been available over the entire time span, these variables include surface pressure and marine winds. The number of assimilated observations is however not continuous through time. An increased number of observations would be expected to result in increased detail within the reanalysis, this lack of observation data is particularly relevant for the remote scoring region considered here. Secondly, NOAA 20CR is given two distinct designations 20CRv3si and 20CRv3mo where the latter has prescribed SST and the former none, This change in designation occurs from 1981 onwards (Slivinski et al., 2019). Figure 26 captures the influence this has on the number of fronts captured by NOAA 20CR over the 1950-2005 period where a significant increasing trend occurs before 1981, presumably associated with increasing number of observations being assimilated, while inclusion of prescribed SST from 1981 onwards results in similar interannual variability and trend as that of ERA5 reanalysis.



(a) *(b)*

Figure 26: (a) NOAA 20CR pre and post 'satellite era' data assimilation anomaly (b) Seasonal mean number of fronts captured in the scoring region in NOAA 20CR (green) and ERA5 reanalysis (red). A normalization factor has been applied. Code Link

The final seasonal climatologies, as depicted in figure 25, are used to score the models using a mean absolute error technique, as utilized in the previous '2-dimensional' metrics. Figures 27 and 28 capture absolute error fields of annual mean frontal activity relative to NOAA 20CR. Where figure 27 utilizes a 1950-2005 baseline, while Figure 28 considers the anomaly's utilizing a 1981-2005 baseline, both methods utilise the 'native' front detection approach, where little difference is seen between native and regridded results. The persistent 'under-reading' bias in the eastern section of the scoring region from NOAA 20CR utilising the longer 1950-2005 baseline is no longer evident when using the shorter 1980-2005 baseline, and as such more varied bias within CMIP5 models are captured.

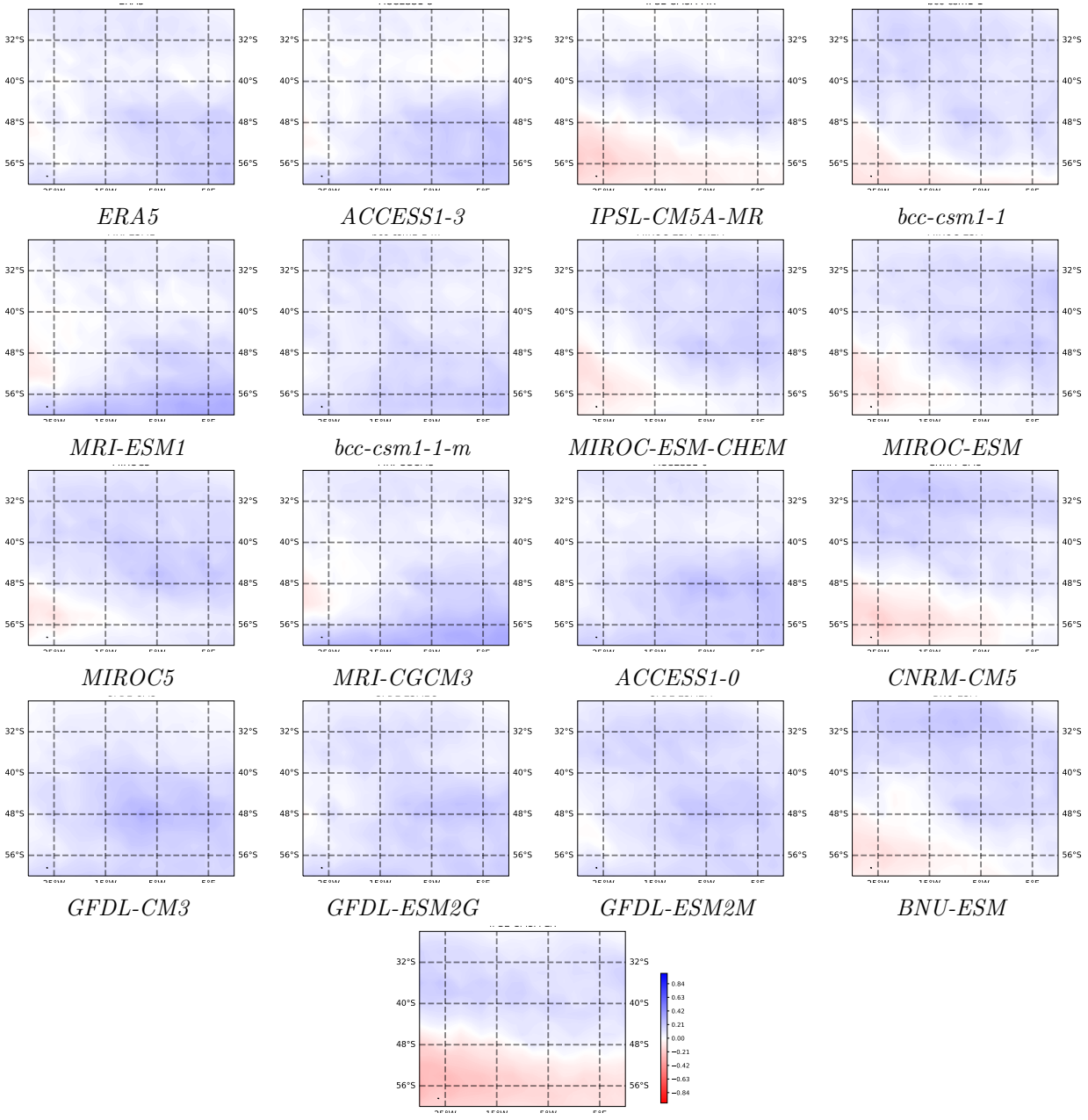


Figure 27: Native annual mean front anomaly relative to NOAA20CR. Models ordered relative to score, best top left worst bottom right Code Link

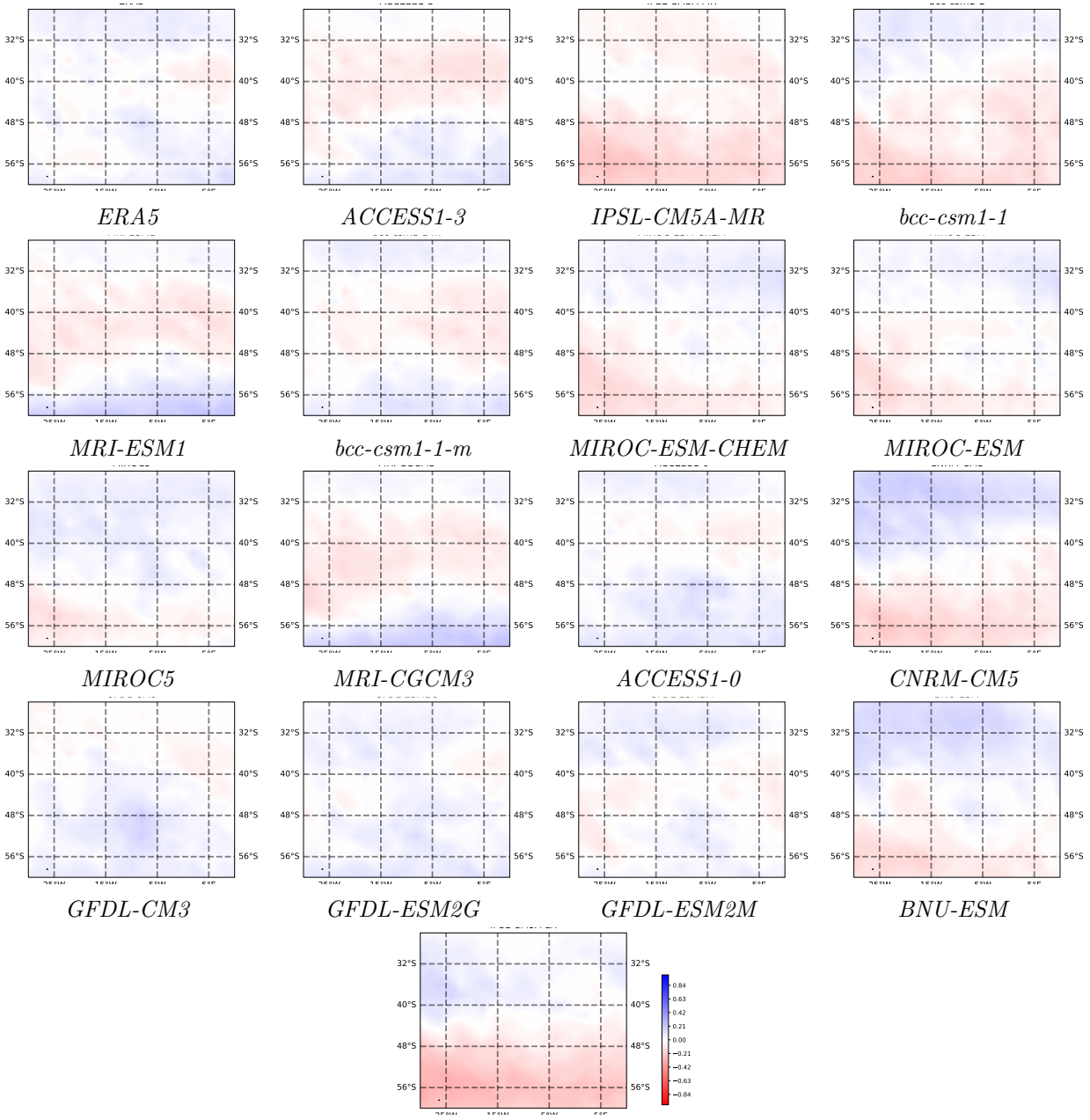
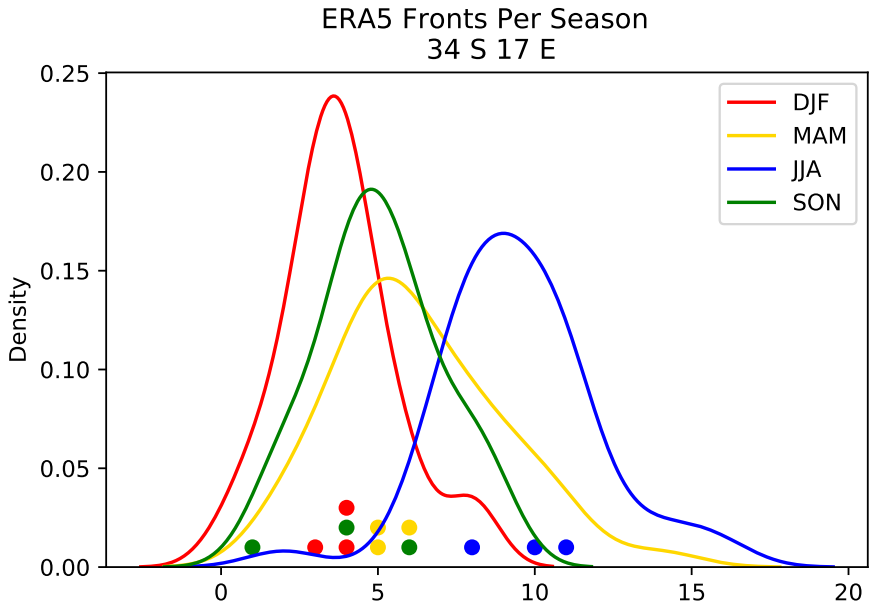


Figure 28: Native annual mean front anomaly relative to NOAA 20CR, utilizing a 1981-2005 baseline. Models ordered relative to score, best top left worst bottom right. [Code Link]

Day Zero Response

Here the response of the cold front climatology developed using this method to the anomalous conditions during the 'Day Zero' period is considered. Figure 29 captures the distribution of cold fronts to the immediate southwest of the WRZ, while scatter points indicate the number of fronts during the 'Day Zero' period. The Spring 2016 is noteworthy as a significantly reduced number of fronts is captured. However, while this was the driest spring during the period the most significant rainfall anomalies were experienced during Autumn where the number of fronts captured is near the median in all years. Remaining seasons are located near the median although no seasons experienced above median number of fronts during the period.



(a)

Figure 29: Kernel Density Estimate of the number of fronts passing 34 S 17 E across the meteorological seasons. Scatter points indicate the number of fronts that occurred during the 2015-2017 meteorological seasons. Code Link

Figure 30 captures the spatial anomaly in the number of fronts captured during the 'Day Zero' period. Here, particularly in Autumn and Winter a positive anomaly south of 40°S and negative anomaly north of 40°S is evident. However, this pattern is not homogenous across space indicating the potential bias considering only a single grid point or small region, such as in figure 29, may introduce. Thus, in the absence of other co-dependent factors, the reduced rainfall experienced during the 'Day Zero' drought cannot be linked to cold fronts incidence alone, however considering the spatial distribution of cold fronts across the South Atlantic can provide a clearer picture than considering the number of cold fronts at a singular region.

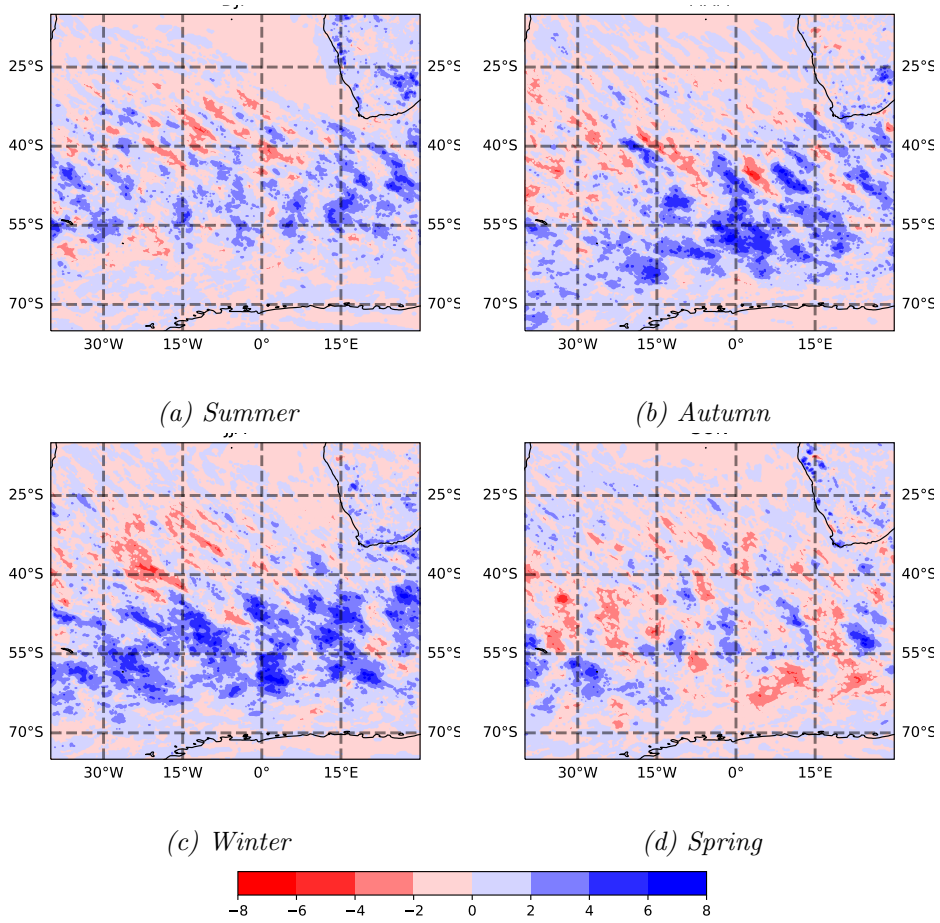


Figure 30: Anomaly in the number of cold fronts captured per season during the 2015–2017 'Day Zero' period against a 1979-2019 baseline. Code Link

Figure 30 captures the spatial anomaly in the number of fronts captured during the 'Day Zero' period, here particularly so in Autumn and Winter a positive anomaly at southern of 40 degrees south and negative anomaly north of 40 degrees south is evident. This pattern is however not homogenous across space indicating the potential bias considering only a single grid point or small region, such as in figure 31a, may introduce. Thus, in the absence of other co-dependent factors, the reduced rainfall experienced during the 'Day Zero' drought cannot be linked to cold fronts incidence alone, however considering the spatial distribution of cold fronts across the South Atlantic can provide a clearer picture than considering the number of cold fronts at a singular region.

Results

The final scores are included in figure 31. Here strong agreement exists between the '2 Degree' and 'native' methods. The native method is preferred as the resolution of a climate model can impose direct limitations on its usefulness to a particular region, thus considering a metric which may be directly influenced by this is preferred. While the same three models, IPSL-

CM5A-LR, CNRM-CM5 and BNU-ESM, perform worst regardless of the method or whether a 1950-2005 or 1981-2005 period is considered, the choice of period does significantly influence the remaining models. Here the shorter 1981-2005 period is preferred in part due to the improved score of ERA5 reanalysis and the increased spread in scores across the CMIP models. Excluding the pre-prescribed SST years from NOAA 20CR is seen to be more beneficial than considering a longer time span. Thus, the final model selection will utilise the scores from the 'Native 1981' technique.

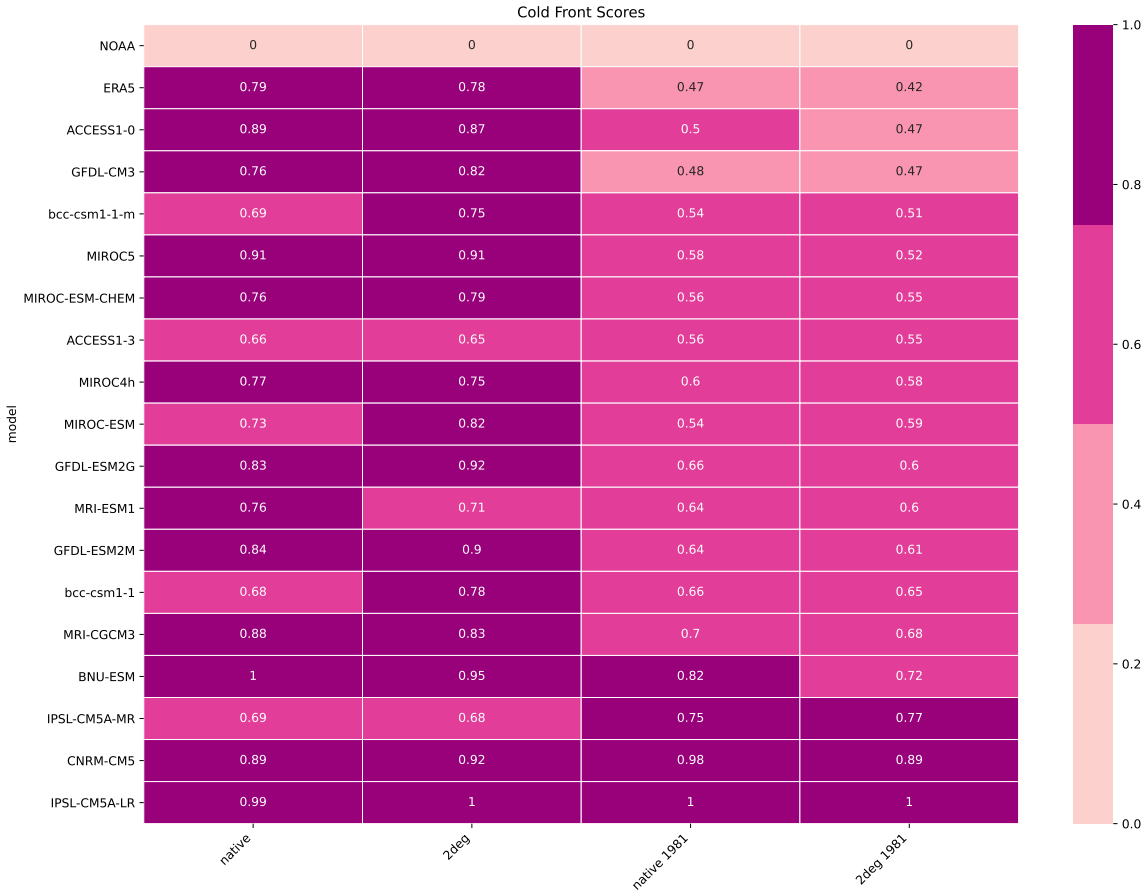


Figure 31: Results from the scoring technique described above. Results are normalized such that the worst performing models is scored 1 and a model identical to that of NOAA20R scores zero. Code Link

Part IV

Model Independence, Selection and Projection Implications

Model Independence

In order to ensure the final ensemble increases both model realism and model independence it is necessary to first quantify the similarity between models to identify 'sibling' model groupings. Projects such as the Earth system Documentation Project (ES-DOC) provide a platform to better understand the particular composition of various climate models, such as the institute leading the project as well as the underlying atmospheric, ocean and sea-ice models that make up the model and the various physics scheme used therein (Greenslade et al., 2014), however comparing various models remains difficult as the fundamental architecture of the models varies greatly and as such comparison on a like-for-like basis of physics schemes is not possible. Further while naming conventions tend to outline model groupings this is not a given and as such it is preferred to have a more objective method to group the models.

Table 3: Pairwise Distances Between CMIP5 Models, as calculated utilising the final Jet, SASH and Cold Front assessment techniques described above. Code Link

Model	ACCESS1-0	ACCESS1-3	BNU-ESM	CNRM-CM5	ERA5	GFDL-CM3	GFDL-ESM2G	GFDL-ESM2M	IPSL-CM5A-LR	IPSL-CM5A-MR	MIROC-ESM-CHEM	MIROC5	MRI-CGCM3	MRI-ESM1	NOAA	bcc-csm1-1	bcc-csm1-1-m
ACCESS1-0	0.0	0.93	2.023	1.898	0.922	1.51	1.426	1.455	2.403	2.568	2.091	2.396	1.671	1.706	1.077	1.512	1.392
ACCESS1-3	0.962	0.0	2.213	2.088	1.146	1.605	1.303	1.339	2.32	2.381	2.324	2.533	1.778	1.811	1.241	1.603	1.456
BNU-ESM	1.664	1.758	0.0	1.532	1.6	1.872	1.865	1.762	1.44	1.893	1.323	2.469	2.358	2.408	1.7	1.551	1.421
CNRM-CM5	1.682	1.812	1.602	0.0	1.598	1.908	1.644	1.508	1.713	2.077	1.714	2.03	2.426	2.458	1.668	1.413	1.617
ERA5	0.884	1.063	1.874	1.72	0.0	1.605	1.387	1.416	2.332	2.494	1.968	2.33	1.802	1.849	0.818	1.237	1.303
GFDL-CM3	1.364	1.406	2.009	1.897	1.501	0.0	1.372	1.507	2.131	2.055	2.017	1.811	1.627	1.623	1.518	1.942	1.451
GFDL-ESM2G	1.318	1.162	2.065	1.667	1.33	1.417	0.0	0.605	2.149	2.285	2.226	2.062	2.01	2.055	1.504	1.699	1.595
GFDL-ESM2M	1.274	1.127	1.847	1.454	1.282	1.484	0.573	0.0	2.0	2.18	2.028	1.964	2.006	2.046	1.417	1.561	1.452
IPSL-CM5A-LR	2.563	2.403	1.906	2.14	2.586	2.549	2.518	2.487	0.0	1.186	1.644	2.916	2.949	2.949	2.564	2.289	2.389
IPSL-CM5A-MR	2.392	2.147	2.056	2.208	2.409	2.12	2.312	2.322	1.02	0.0	1.684	2.678	2.617	2.606	2.293	2.247	2.217
MIROC-ESM-CHEM	1.687	1.803	1.264	1.574	1.655	1.816	1.952	1.882	1.208	1.499	0.0	2.353	2.303	2.309	1.658	1.579	1.389
MIROC5	2.068	2.104	2.484	1.931	2.076	1.696	1.929	1.913	2.339	2.426	2.379	0.0	1.332	1.37	2.044	2.298	2.217
MRI-CGCM3	1.771	1.831	2.963	2.888	1.98	1.941	2.297	2.393	3.0	3.0	2.945	1.957	0.0	0.336	2.158	2.33	2.287
MRI-ESM1	1.798	1.854	2.998	2.904	2.015	1.92	2.331	2.424	2.968	2.954	2.924	1.981	0.33	0.0	2.187	2.333	2.289
NOAA	0.976	1.102	1.851	1.668	0.772	1.572	1.494	1.48	2.19	2.269	1.862	2.221	1.92	1.961	0.0	1.288	1.288
bcc-csm1-1	1.285	1.309	1.676	1.468	1.106	1.843	1.602	1.575	1.949	2.191	1.774	2.463	1.955	1.981	1.236	0.0	1.374
bcc-csm1-1-m	1.204	1.215	1.488	1.585	1.164	1.431	1.496	1.441	1.904	2.102	1.532	2.298	1.895	1.907	1.232	1.378	0.0

Knutti, Masson & Gettelman, (2013) create a genealogy of CMIP5 and CMIP3 models by considering root mean square error of monthly mean and annual climatologies of precipitation and temperature fields from the unperturbed pre-industrial control runs of the CMIP5 and CMIP3 models. Before utilising a hierarchical clustering to create model family trees. Here a similar approach is utilised to create sibling groupings of the models considered, but instead of utilising temperature and precipitation fields to define the pairwise distance between models the previously defined metrics are utilised. Where rather than scoring models against only NOAA 20CR each model is scored against the next and the results of the 'JET2D', 'SASH Contour' and 'Front 1981 native' assessments are normalized and summed to create the pairwise matrix seen in table 3.

Finally, the dendrogram shown in figure 32, is constructed from these pairwise distances utilising the Ward's minimum variance clustering technique. To further validate these results the colour coding's from the results of Knutti, Masson & Gettelman, (2013) are transposed

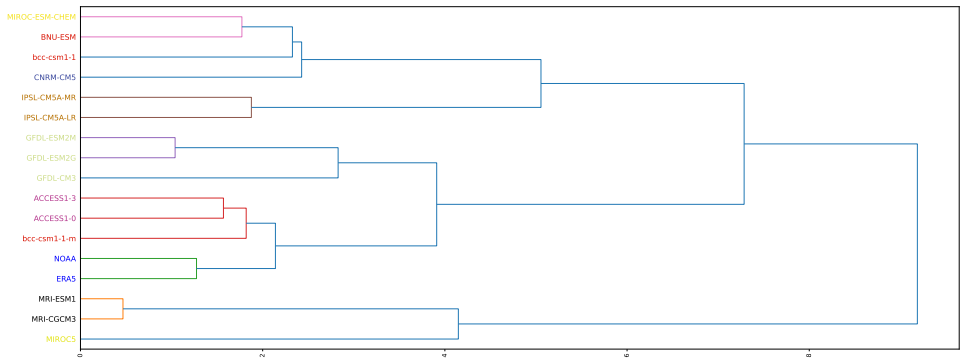


Figure 32: Dendrogram showing hierarchical clustering of the pairwise distances shown in Figure 32. Color coding of labels have been transposed from Knutti, Masson & Gettelman, (2013): Climate model genealogy: Generation CMIP5 and how we got there - replicated in this paper in Figure 1a. *Code Link*

on to the labels of figure 32. Here model clustering's are seen to be largely confined to models from the same institution, with the exception of MIROC5 and MIROC-ESM-CHEM, where MIROC5 is considered distinct model from its MIROC-ESM counterpart despite coming from the same institution. Largely similar groupings to those identified by Knutti, Masson & Gettelman, (2013) are identified, including that of the MIROC exception. The final sibling groupings are indicated on figure 33.

Final Ensemble Selection

The results from the chosen Jet Stream, Subtropical High and Cold Front assessment metrics have been transposed on to the final results in Figure 33, while dark lines indicate divisions between sibling groupings. Here the results have been normalized such that ERA5 reanalysis represents a 'perfect' score of 0 whilst the worst performing model in this ensemble receives a score of 1. This differs subtly to the scores in the previous metric sections where NOAA 20th Century was used as both the 'reference' model and 'perfect' model where now NOAA 20th Century remains the 'reference' model but instead 'ERA5' is used as the 'perfect' model against which to normalize scores. In practice this helps increase the range of scores within the CMIP5 models.

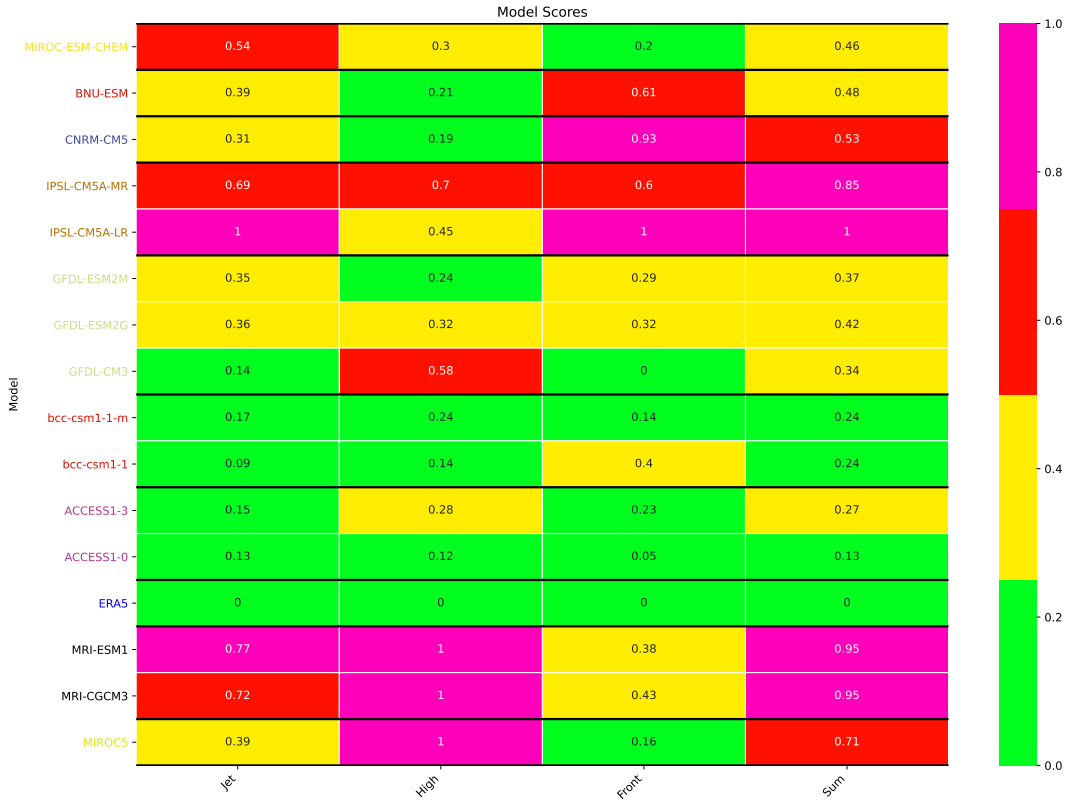
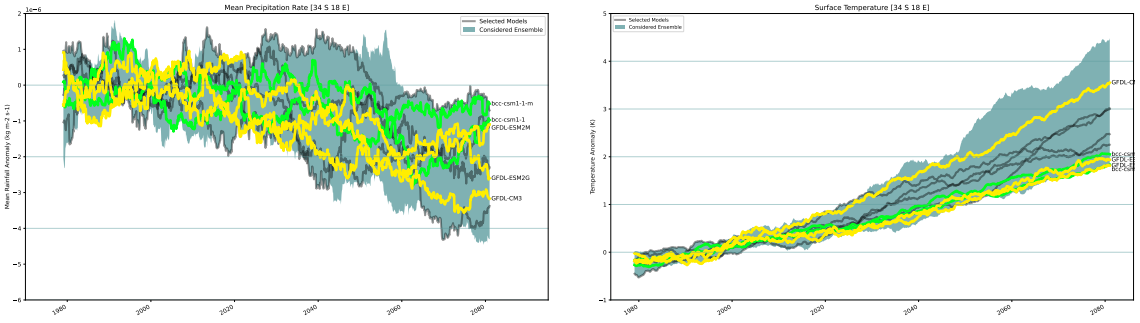


Figure 33: Final Model Scores from the Metric Evaluations. Here the results have been normalized so the model least similar to NOAA 20th Century reanalysis scores 1 while the most similar model, which in all cases is ERA5 reanalysis scores 0. Thus utilising ERA5 as reference for a perfectly realistic model. Results are colour coded whereby: green corresponds to realistic, yellow - bias, red - significantly bias and purple - implausible. Code Link

Here the intention is to select only a single model from each sibling grouping to increase model independence. Divisions between scoring categories are entirely subjective and such it was decided to simply proceed with four categories of equal size such that a score of: 1-0.75 corresponds to implausible; 0.75-0.5 corresponds to significantly bias; 0.5-0.25 corresponds to bias and 0.25-0 corresponds to realistic. Where a score of 1 indicates the worst performing CMIP5 model and 0 a suitably realistic model. Models that score implausible are not considered further, while models scored significantly bias are considered in the context of their future projections and only included in the final ensemble if the model's absence would significantly reduce the range of future possible climates. Models that score either bias or realistic are included in the final ensemble unless a sibling model performs better. Where sibling models receive the same scores, they are considered in context of future projections and the model that would increase the range of future projections is selected, in a similar fashion to the selection matrix of a significantly bias model.

Following this decision matrix an initial ensemble of MIROC-ESM-CHEM, BNU-ESM, CNRM-CM5 and ACCESS1-0 can be selected as these models score either realistic or bias and receive a better score than their sibling models, whilst the sibling groupings of 'bcc-csm1-1' and 'bccc-csm1-1-m' and 'GFDL-ESM2M', 'GFDL-ESM2G' and 'GFDL-CM3' are considered in context of their future projections as each model receives the same score as its sibling model. Figure 34 considers the projections of these two sibling groupings against that of the initially selected ensemble. Here 'GFDL-CM3' is seen to project significantly more severe temperature increases than that of the initial ensemble and as such is included, and its sibling models excluded. 'bcc-csm1-1-m' projects a similar temperature trend to that of its sibling however it projects a less severe drying trend than that of its sibling model as well as the initial ensemble and is thus selected and its sibling excluded.

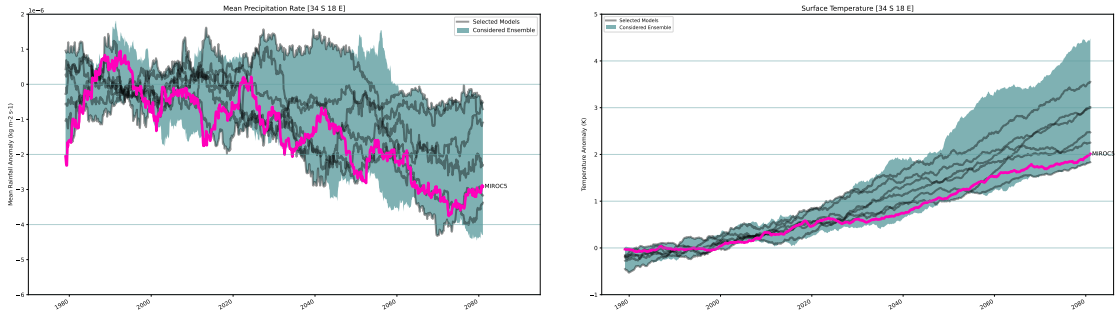


(a) pr *(b) tas*

Figure 34: Projected rainfall (a) and temperature (b) at [34 S 18 E] under the rcp85 emissions scenario of the initial ensemble of [MIROC-ESM-CHEM, BNU-ESM, CNRM-CM5 and ACCESS1-0] and that of the sibling groupings of 'bcc-csm1-1' and 'bccc-csm1-1-m'] – coloured in green corresponding to the realistic scores of these models - and ['GFDL-ESM2M', 'GFDL-ESM2G' and 'GFDL-CM3] – coloured in yellow corresponding to the bias scores of these models. (a) - Code Link (b) - Code Link

Figure 35 considers MIROC5 which is scored as significantly bias. Here the projections of

MIROC5 are within the range of the already selected ensemble and as such it is not considered in the final ensemble.



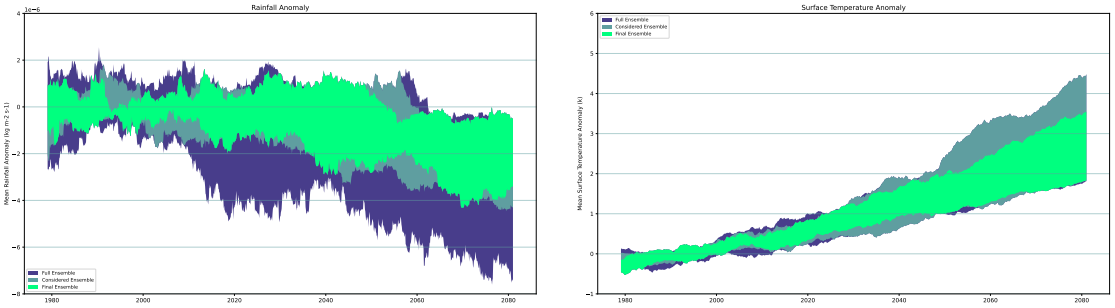
(a) pr *(b) tas*

Figure 35: Projected rainfall (a) and temperature (b) at [34 S 18 E] under the rcp85 emissions scenario of the initial ensemble of [MIROC-ESM-CHEM, BNU-ESM, CNRM-CM5, ACCESS1-0, 'bcc-csm1-1-m' and 'GFDL-CM3'] and MIROC5 – coloured in red corresponding to the significantly bias score of this model. MIROC5 is the only model to score as significantly bias without a better performing sibling and as such is only included in the final ensemble is its absence would significantly reduce the range of future projections. (a) - Code Link (b) - Code Link

Thus, the final ensemble consists of MIROC-ESM-CHEM, BNU-ESM, CNRM-CM5, ACCESS1-0, 'bcc-csm1-1-m' and 'GFDL-CM3', hereby reducing the ensemble to six models from the 15 models considered in all three metric evaluations.

Future Climate Projection Implications

Figure 36 shows the implications this selection has on future temperature and precipitation projections in the WRZ. Here the range of future surface temperature projections in the WRZ is seen to be significantly reduced by the model selection with the upper limit of projections reduced. Precipitation projections remain similar to that of the considered ensemble with data availability being the primary driver of reducing the range of precipitation projections from that of the full ensemble.



(a) pr *(b) tas*

Figure 36: Projected rainfall (a) and temperature (b) at [34 S 18 E] under the rcp85 emissions scenario of the final ensemble of [MIROC-ESM-CHEM, BNU-ESM, CNRM-CM5, ACCESS1-0, 'bcc-csm1-1-m' and 'GFDL-CM3'] in green against the ensemble of 15 models considered in light blue and the maximum ensemble of 40 models with precipitation and surface temperature variable available in purple. Red highlights the result of excluding reliance on sub-daily resolution data from the model selection. (a) - Code Link (b) - Code Link

The sub-daily temporal resolution data required for the cold front analysis is the primary driver behind the reduced considered ensemble size where excluding this metric can increase the considered ensemble size from 15 to 26 models and help increase the range of projections, as indicated in figure 36. Excluding the cold front metric from the selection process does however degrade the genealogy distinctions somewhat making distinction between model families less clear, presumably due to the reduced number of metrics. Thus, finding a third metric which could serve as a proxy for mid-latitude dynamics without reliance on sub-daily resolution data could improve results. Further while the ensemble size does increase when excluding the 'Cold Front' scores and thus reliance on sub-daily temporal resolution data the range of precipitation projections of the considered ensemble remains lower than that of the full ensemble.

Figure 36 highlights the drying trends projected by the final ensemble in the WRZ region, consistent with recent drying trends in Southern Hemisphere subtropical regions (Burls et al., 2019). Figure 37 captures how this projected decrease in precipitation and temperature increase is accompanied by an increased probability of prolonged drought. This is calculated by considering 12-month Specific Precipitation Index (SPI-12), calculated utilising a 1950-2005 baseline, following a similar method to that described by Kam et al., (2021) where a drought period is defined as starting when SPI drops below -0.8 and ending when SPI-12 is

greater than 0.2. A long-term drought is then defined as a drought period exceeding 2 years. Finally, the probability of long term drought occurring during the periods 1985-2005 and 2040-2060 is calculated for each model. Here the reduced final ensemble projects increased likelihood of long-term drought under the rcp8.5 emissions scenario, compared to that of the full ensemble.

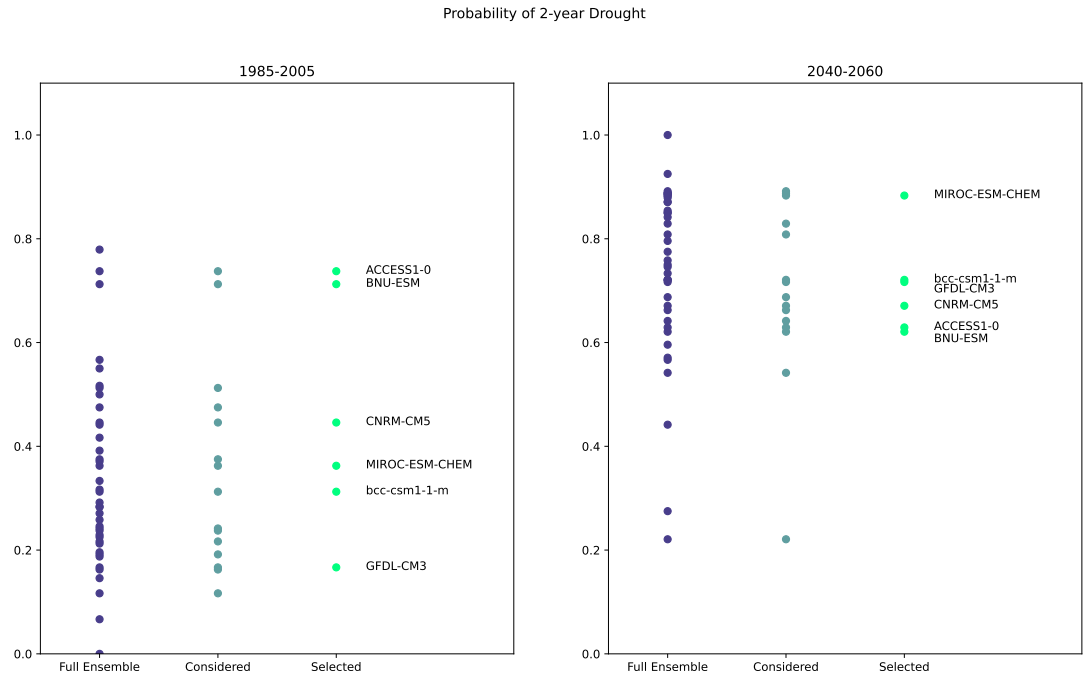


Figure 37: Probability of a 2-year drought occurring in a given year in the WRZ. Code Link

Figure 38 captures projected change in jet stream and SASH positioning in the final selected ensemble as capture by the previously described metrics developed for model scoring. Sub daily resolution rcp85 data was not available for a number of the models and as such a Cold Front anomaly has not been included. Here similar regime changes can be seen across all models with a distinct southward migration of the mid-latitude jet stream. The SASH is seen to be defined less frequently in the central South Atlantic suggesting a shift to a more transient nature, MIROC-ESM-CHEM indicates increased SASH prevalence in the western South Atlantic while GFDL-CM3 indicates a notable southward migration of the SASH. MIROC-ESM-CHEM and GFDL-CM3 which indicate the most distinct shift in SASH capture the largest reduction in precipitation, 103 mm/year and 83 mm/year respectively

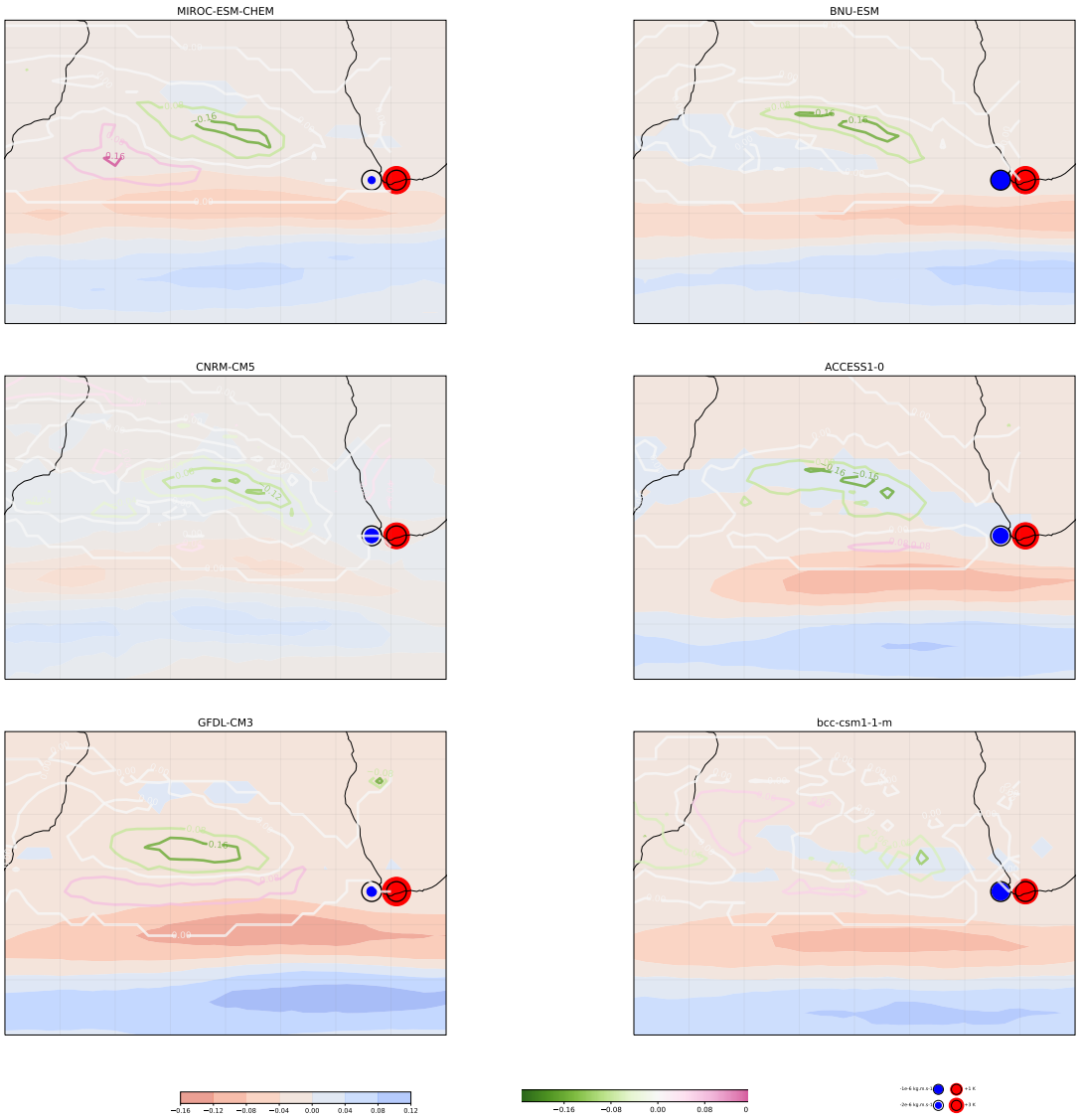


Figure 38: Future ([2045-2100] - [1950-2005]) Jet Stream (shaded) and SASH (contours) anomaly under the rcp85 emission pathway. Code Link

Part V

Conclusion

The 2015-2017 'Day Zero' drought in South Africa's WRZ highlighted the need for robust and reliable future climate projections for future water supply and agricultural planning in the region. The large uncertainties within these future climate projections are however a significant impediment to decision making. This study has investigated the potential to sub-sample a multi-model ensemble of climate projections with the aim to reduce uncertainty as well as identify a realistic subset of plausible future climate pathways in the region. A model selection of this nature requires evaluation of model realism with respect to relevant regional climate dynamics, as well as consideration of model independence within the ensemble, such that the final reduced ensemble is subject to less influence from unrealistic models whilst ensuring model-based uncertainties are sufficiently sampled.

Models are assessed against synoptic scale circulation features and associated statistics, rather than simulated rainfall in the region, in order to assess the physical realism of the model rather than the realism of regional rainfall which will be subject to model parameterization. To navigate the subjectivity in selecting metrics to assess models against, the 'Day Zero' drought has been used as an episodic reference where relevant metrics should be able to capture the anomalous conditions during this event, consistent with the amplification of this event as a result of climate change. The extensive literature produced subsequent to the 'Day Zero' drought has been leveraged to identify metrics which capture relevant regional climate features.

Three regional climate features were selected: the South Atlantic Jet Stream, the South Atlantic Subtropical High and South Atlantic Cold Fronts. These are the primary dynamics behind moisture supply to the WRZ region, and anomalies therein are crucial to the 'Day Zero' drought. Various methods of quantifying these features have been developed and the ability of each method to capture the anomalous conditions during the 'Day Zero' period were compared and contrasted before selecting a final method for scoring of CMIP5 models against NOAA20CR and ERA5 reanalysis. Multiple methods were tested in order to ensure that the metrics are robust and not overly sensitive to the method formulation. Importantly, in most cases, the different methods used to evaluate each metric produced very similar results. This means that, while the optimal method was chosen in each case, the metrics are not very sensitive to particular details of each method and thus are likely to be robust.

Following the method described by McSweeney et al., (2015) each model was assigned a score of 'realistic', 'biased', 'significant-biased' or 'unrealistic' as a function of the performance across the three metrics. Unrealistic models were subsequently removed from the ensemble while significantly biased models were also excluded as their absence did not significantly reduce the range of future projections. These same metrics were then used to create a genealogy of models, demonstrating that even utilising only three simple metrics it is possible to identify models developed by the same institutions, here attaining the same grouping as

in Knutti, Masson & Gettelman (2013). This process highlights the importance of model selection in the first case, whereby considering each model within the ensemble equally is an unconscious weighting towards institutions that have developed more models (Sanderson, Knutti & Caldwell, 2015a). The model groupings identified were then utilised to select only the best performing model from each grouping, further reducing ensemble size whilst increasing independence within the ensemble. An important result is that performance with respect to the different metrics was consistent across models, with models designated as unrealistic or significantly biased performing poorly across all metrics, while the best performing models are consistent across all metrics. This suggests that the evaluation metrics are not spurious and are indeed evaluating the fundamental realism of the models effectively and can therefore be considered robust.

Thus, after considering 16 CMIP5 models across all metrics, a set of 6 CMIP5 models, namely: 'MIROC-ESM-CHEM', 'BNU-ESM', 'CNRM-CM5', 'ACCESS1-0', 'GFDL-CM3' and 'bcc-csm1-1-m' are selected. These models are shown to all perform suitably well in capturing the dynamics that may result in prolonged drought in the WRZ, while all performing better than their sibling models, therefore presenting an ensemble of more independent and historically more realistic models than that of the full ensemble. Further, this final ensemble has been selected to ensure the range of temperature and precipitation projections is similar to that of a larger ensemble of suitably well performing models. Despite the emphasis on ensuring a wide range of future climate scenarios are preserved, eliminating poorly performing models has significantly reduced the range of projected future climate outcomes. The most extreme temperature projection within the final ensemble projects temperature anomalies, under the RCP8.5 scenario, of a ± 3.5 °C increase by 2080 from a 1980-2005 baseline compared to ± 4.5 °C from the full ensemble.

A key challenge in implementing the South Atlantic cold front method was the availability of high temporal resolution fields. Not all models in the ensemble have archived data for all fields and this resulted in an artificial subsampling of the ensemble to models with sub-daily temporal resolution wind fields available. As a result, projections of absolute precipitation are somewhat reduced in range with the worst-case anomaly half that of the full ensemble – however data availability has contributed significantly to this reduction. Conversely, probability of a 2-year drought is seen to be increased in the final ensemble compared to that of the full ensemble. While these constrained future projections may be a welcome consequence of eliminating unrealistic models, the primary utility of the final ensemble is instead in the reduced size of the ensemble, where with only 6 models presented, a future researcher may consider each model's projected future climate pathway individually before selecting a model, or models, which best informs their use case, whilst being assured that this model performs suitably well in the region. This is of particular value for impacts modellers looking to consider multiple models who can issue a subset of future climate scenarios that are sufficiently independent and yet still represents model uncertainty, while strong similarity between two or more models within the ensemble will not be unduly biasing results.

This study has thus demonstrated the use of relevant, robust, regional model realism metrics in successfully sub-sampling the CMIP5 ensemble in such a way that unrealistic models can

be removed while model independence can still be maximised. The resulting sub-sampling has particular value for further downstream analysis, while to some extent reducing uncertainties of future projections. The use of model realism metrics to sub-sample multi-model ensemble will always have some degree of subjectivity. Here the selection of realism metrics, though still subjective, is strongly guided through literature. Thresholds for discriminating between model scoring categories are also somewhat arbitrary and subjective and could be chosen differently which would ultimately affect the final ensemble sub-selection. However, if these subjective choices are made transparently and multiple measures and metrics are used and compared, the results may be openly evaluated by others.

Further work could consider different regional circulation features, or different approaches to evaluate realism; perhaps considering historical trend or the prevalence of extreme events. For example, considering the ability of models to reproduce historical trends in jet stream or SASH statistics could add further rigour to the assessment. Assessing to what extent different assessment metrics would impact the final sub-selection would be of interest, assessing whether a similar ensemble and range of projections is preserved or not. stream or SASH statistics, could add further rigour to the assessment. Considering to what extent different assessment metrics would impact the final sub-selection would be of interest, assessing whether a similar ensemble and range of projections is preserved or not.

Part VI

References

Abba Omar, S. & Abiodun, B.J. 2020. Characteristics of cut-off lows during the 2015–2017 drought in the Western Cape, South Africa. Atmospheric Research. 235:104772. DOI: https://doi.org/10.1016/j.atmosres.2019.104772.

Archer, E., Landman, W., Malherbe, J., Tadross, M. & Pretorius, S. 2019. South Africa's winter rainfall region drought: A region in transition? Climate Risk Management. 25:100188.

Bishop, C.H. & Abramowitz, G. 2013. Climate model dependence and the replicate Earth paradigm. Climate dynamics. 41(3–4):885–900.

Blamey, R.C., Ramos, A.M., Trigo, R.M., Tomé, R. & Reason, C.J.C. 2018. The Influence of Atmospheric Rivers over the South Atlantic on Winter Rainfall in South Africa. Journal of Hydrometeorology. 19(1):127–142.

Botai, M.C., Botai, O.J., de Wit, P.J., Ncongwane, P.K. & Adeola, M.A. 2017. DOI: 10.3390/w9110876.

Burls, N.J., Blamey, R.C., Cash, B.A., Swenson, E.T., Fahad, A. al, Bopape, M.-J.M., Straus, D.M. & Reason, C.J.C. 2019. The Cape Town "Day Zero" drought and Hadley cell expansion. npj Climate and Atmospheric Science. 2(1):27. DOI: 10.1038/s41612-019-0084-6.

Cabos, W., Sein, D. V, Pinto, J.G., Fink, A.H., Koldunov, N. V, Alvarez, F., Izquierdo, A., Keenlyside, N., et al. 2017. The South Atlantic Anticyclone as a key player for the representation of the tropical Atlantic climate in coupled climate models. Climate Dynamics. 48(11):4051–4069. DOI: 10.1007/s00382-016-3319-9.

D'Ercole, J., Mahetaji, K., Speedie, J. & Raufdeen, F. 2018. Climate Model Diversity: Future Climate Predictions from the CMIP5 Multi-Model Ensemble.

Daron, J., Burgin, L., Janes, T., Jones, R.G. & Jack, C. 2019. Climate process chains: Examples from southern Africa. International Journal of Climatology. 0(0). DOI: 10.1002/joc.6106.

DEA. 2018. South Africa's Third National Communication under the United Nations Framework Convention on Climate Change. Department of Environmental Affairs, Republic of South Africa, Pretoria.

Ding, Q., Steig, E.J., Battisti, D.S. & Wallace, J.M. 2012. Influence of the Tropics on the Southern Annular Mode. Journal of Climate. 25(18):6330–6348. DOI: 10.1175/JCLI-D-11-00523.1.

Engelbrecht, C.J., Engelbrecht, F.A. & Dyson, L.L. 2013. High-resolution model-projected changes in mid-tropospheric closed-lows and extreme rainfall events over southern Africa. International Journal of Climatology. 33(1):173–187. DOI: 10.1002/joc.3420.

Eyring, V., Cox, P.M., Flato, G.M., Gleckler, P.J., Abramowitz, G., Caldwell, P., Collins, W.D., Gier, B.K., et al. 2019. Taking climate model evaluation to the next level. Nature Climate Change. 9(2):102–110.

Favre, A., Hewitson, B., Lennard, C., Cerezo-Mota, R. & Tadross, M. 2013. Cut-off lows in the South Africa region and their contribution to precipitation. Climate dynamics. 41(9–10):2331–2351.

Flato, G.M. 2011. Earth system models: an overview. Wiley Interdisciplinary Reviews: Climate Change. 2(6):783–800.

Gallego, D., Ribera, P., Garcia-Herrera, R., Hernandez, E. & Gimeno, L. 2005. A new look for the Southern Hemisphere jet stream. Climate Dynamics. 24(6):607–621. DOI: 10.1007/s00382-005-0006-7.

Gerber, E.P., Butler, A., Calvo, N., Charlton-Perez, A., Giorgetta, M., Manzini, E., Perlwitz, J., Polvani, L.M., et al. 2012. Assessing and understanding the impact of stratospheric dynamics and variability on the Earth system. Bulletin of the American Meteorological Society. 93(6):845–859.

Gershunov, A., Shulgina, T., Clemesha, R.E.S., Guirguis, K., Pierce, D.W., Dettinger, M.D., Lavers, D.A., Cayan, D.R., et al. 2019. Precipitation regime change in Western North America: The role of Atmospheric Rivers. Scientific Reports. 9(1):9944. DOI: 10.1038/s41598-019-46169-w.

Greenslade, M., Murphy, S., Treshansky, A., DeLuca, C., Guilyardi, E. & Denvil, S. 2014. The Earth System (ES-DOC) Project. In EGU General Assembly Conference Abstracts. 12988.

Hamming, R.W. 1998. Digital filters. Courier Corporation.

Hastenrath, S. 2012. Climate dynamics of the tropics. V. 8. Springer Science & Business Media.

He, C., Wu, B., Zou, L. & Zhou, T. 2017. Responses of the Summertime Subtropical Anticyclones to Global Warming. Journal of Climate. 30(16):6465–6479. DOI: 10.1175/JCLI-D-16-0529.1.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., et al. 2020. The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society. 146(730):1999–2049.

Hewitson, B., Waagsaether, K., Wohland, J., Kloppers, K. & Kara, T. 2017. Climate information websites: an evolving landscape. Wiley Interdisciplinary Reviews: Climate Change. 8(5):e470.

Hewitt, C., Mason, S. & Walland, D. 2012. The global framework for climate services. Nature Climate Change. 2(12):831–832.

Hoskins, B.J. & Hodges, K.I. 2005. A new perspective on Southern Hemisphere storm tracks. Journal of Climate. 18(20):4108–4129.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., et al. 2016. The Art and Science of Climate Model Tuning. Bulletin of the American Meteorological Society. 98(3):589–602. DOI: 10.1175/BAMS-D-15-00135.1.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jack, C.D., Marsham, J., Rowell, D.P. & Jones, R.G. 2021. Climate Information: Towards Transparent Distillation. In Climate Risk in Africa. Palgrave Macmillan, Cham. 17–35.

Kahneman, D. 2011. Thinking, fast and slow. Macmillan.

Kam, J., Min, S.-K., Wolski, P. & Kug, J.-S. 2021. CMIP6 Model-Based Assessment of Anthropogenic Influence on the Long Sustained Western Cape Drought over 2015–19. Bulletin of the American Meteorological Society. 102(1):S45–S50.

Kidson, J.W. 1988. Interannual Variations in the Southern Hemisphere Circulation. Journal of Climate. 1(12):1177–1198. DOI: 10.1175/1520-0442(1988)001¡1177:IVITSH¿2.0.CO;2.

Kidston, J., Scaife, A.A., Hardiman, S.C., Mitchell, D.M., Butchart, N., Baldwin, M.P. & Gray, L.J. 2015. Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. Nature Geoscience. 8(6):433–440.

Knutti, R. 2010. The end of model democracy? Climatic Change. 102: 395. DOI: https://doi.org/10.1007/s
010-9800-2.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. & Meehl, G.A. 2010. Challenges in combining projections from multiple climate models. Journal of Climate. 23(10):2739–2758.

Knutti, R., Masson, D. & Gettelman, A. 2013. Climate model genealogy: Generation CMIP5 and how we got there. Geophysical Research Letters. 40(6):1194–1199. DOI: 10.1002/grl.50256.

Kolusu, S.R., Siderius, C., Todd, M.C., Bhave, A., Conway, D., James, R., Washington, R., Geressu, R., et al. 2021. Sensitivity of projected climate impacts to climate model weighting: multi-sector analysis in eastern Africa. Climatic Change. 164(3):1–20.

Lennard, C. & Hegerl, G. 2015. Relating changes in synoptic circulation to the surface rainfall response using self-organising maps. Climate Dynamics. 44(3–4):861–879.

Lim, E.-P., Hendon, H.H., Arblaster, J.M., Delage, F., Nguyen, H., Min, S.-K. & Wheeler, M.C. 2016. The impact of the Southern Annular Mode on future changes in Southern Hemisphere rainfall. Geophysical Research Letters. 43(13):7160–7167. DOI: 10.1002/2016GL069453.

Lorenz, D.J. & Hartmann, D.L. 2001. Eddy–Zonal Flow Feedback in the Southern Hemisphere. Journal of the Atmospheric Sciences. 58(21):3312–3327. DOI: 10.1175/1520-0469(2001)058¡3312:EZ

Mahlalela, P.T., Blamey, R.C. & Reason, C.J.C. 2019. Mechanisms behind early winter rainfall variability in the southwestern Cape, South Africa. Climate Dynamics. 53(1):21–39.

DOI: 10.1007/s00382-018-4571-y.

Marshall, G.J. 2003. Trends in the Southern Annular Mode from Observations and Reanalyses. Journal of Climate. 16(24):4134–4143. DOI: 10.1175/1520-0442(2003)016¡4134:TITSAM¿2.0.CO;2.

Masson, D. & Knutti, R. 2011. Climate model genealogy. Geophysical Research Letters. 38(8). DOI: https://doi.org/10.1029/2011GL046864.

McSweeney, C.F., Jones, R.G., Lee, R.W. & Rowell, D.P. 2015. Selecting CMIP5 GCMs for downscaling over multiple regions. Climate Dynamics. 44(11–12):3237–3260.

Meehl, G.A., Moss, R., Taylor, K.E., Eyring, V., Stouffer, R.J., Bony, S. & Stevens, B. 2014. Climate model intercomparisons: Preparing for the next phase. Eos, Transactions American Geophysical Union. 95(9):77–78.

Munday, C. & Washington, R. 2019. Controls on the diversity in climate model projections of early summer drying over southern Africa. Journal of Climate. 32(12):3707–3725.

Odoulami, R.C., Wolski, P. & New, M. 2021. A SOM-based analysis of the drivers of the 2015–2017 Western Cape drought in South Africa. International Journal of Climatology. 41:E1518–E1530.

Otto, F., Wolski, P., Lehner, F., Tebaldi, C., van Oldenborgh, G.J., Hogesteger, S., Singh, R., Holden, P., et al. 2018. Available: https://www.worldweatherattribution.org/the-role-of-climate-change-in-the-2015-2017-drought-in-the-western-cape-of-south-africa/.

Overland, J.E., Wang, M., Bond, N.A., Walsh, J.E., Kattsov, V.M. & Chapman, W.L. 2011. Considerations in the selection of global climate models for regional climate projections: The Arctic as a case study. Journal of Climate. 24(6):1583–1597.

Pena-Ortiz, C., Gallego, D., Ribera, P., Ordonez, P. & Alvarez-Castro, M.D.C. 2013. Observed trends in the global jet stream characteristics during the second half of the 20th century. Journal of Geophysical Research: Atmospheres. 118(7):2702–2713. DOI: https://doi.org/10.1002/jgrd.50305.

Peterson, R.G. & Stramma, L. 1991. Upper-level circulation in the South Atlantic Ocean. Progress in Oceanography. 26(1):1–73. DOI: https://doi.org/10.1016/0079-6611(91)90006-8.

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. DOI: 10.1002/joc.3403.

Quagraine, K.A., Hewitson, B., Jack, C., Wolski, P., Pinto, I. & Lennard, C. 2020. Using Co-Behavior Analysis to Interrogate the Performance of CMIP5 GCMs over Southern Africa. Journal of Climate. 33(7):2891–2905.

Reason, C.J.C. 2017. DOI: 10.1093/acrefore/9780190228620.013.513.

Reason, C.J.C. & Jagadheesha, D. 2005a. Relationships between South Atlantic SST Variability and Atmospheric Circulation over the South African Region during Austral Winter.

Journal of Climate. 18(16):3339–3355. DOI: 10.1175/JCLI3474.1.

Reason, C.J.C. & Jagadheesha, D. 2005b. A model investigation of recent ENSO impacts over southern Africa. Meteorology and Atmospheric Physics. 89(1):181–205. DOI: 10.1007/s00703-005-0128-9.

Reason, C.J.C. & Rouault, M. 2005. Links between the Antarctic Oscillation and winter rainfall over western South Africa. Geophysical Research Letters. 32(7). DOI: 10.1029/2005GL022419.

Reason, C.J.C., Rouault, M., Melice, J.-L. & Jagadheesha, D. 2002. Interannual winter rainfall variability in SW South Africa and large scale ocean–atmosphere interactions. Meteorology and Atmospheric Physics. 80(1):19–29. DOI: 10.1007/s007030200011.

Reason, C.J.C., Jagadheesha, D. & Tadross, M. 2003. A model investigation of inter-annual winter rainfall variability over southwestern South Africa and associated ocean-atmosphere interaction. South African journal of science. 99(1–2):75–80.

Reboita, M.S., Ambrizzi, T., Silva, B.A., Pinheiro, R.F. & da Rocha, R.P. 2019. Available: https://www.frontiersin.org/article/10.3389/feart.2019.00008.

Renard, R.J. & Clarke, L.C. 1965. Experiments in numerical objective frontal analysis. Monthly Weather Review. 93(9):547–556.

Richter, I., Mechoso, C.R. & Robertson, A.W. 2008. What determines the position and intensity of the South Atlantic anticyclone in austral winter?—An AGCM study. Journal of climate. 21(2):214–229.

Rodwell, M.J. & Hoskins, B.J. 2001. Subtropical anticyclones and summer monsoons. Journal of Climate. 14(15):3192–3211.

Rowell, D.P. 2019. An observational constraint on CMIP5 projections of the East African long rains and southern Indian Ocean warming. Geophysical Research Letters. 46(11):6050–6058.

Sabherwal, A., Ballew, M.T., van der Linden, S., Gustafson, A., Goldberg, M.H., Maibach, E.W., Kotcher, J.E., Swim, J.K., et al. 2021. The Greta Thunberg Effect: Familiarity with Greta Thunberg predicts intentions to engage in climate activism in the United States. Journal of Applied Social Psychology. n/a(n/a). DOI: https://doi.org/10.1111/jasp.12737.

Sanderson, B.M., Knutti, R. & Caldwell, P. 2015a. A representative democracy to reduce interdependency in a multimodel ensemble. Journal of Climate. 28(13):5171–5194.

Sanderson, B.M., Knutti, R. & Caldwell, P. 2015b. Addressing interdependency in a multimodel ensemble by interpolation of model properties. Journal of Climate. 28(13):5150–5170.

Santer, B.D., Taylor, K.E., Gleckler, P.J., Bonfils, C., Barnett, T.P., Pierce, D.W., Wigley, T.M.L., Mears, C., et al. 2009. Incorporating model quality information in climate change detection and attribution studies. Proceedings of the National Academy of Sciences. 106(35):14778–14783.

Schemm, S., Rudeva, I. & Simmonds, I. 2015. Extratropical fronts in the lower troposphere–global perspectives obtained from two automated methods. Quarterly Journal of the

Royal Meteorological Society. 141(690):1686–1698.

Scott, D.W. 1979. On Optimal and Data-Based Histograms. Biometrika. 66(3):605–610. DOI: 10.2307/2335182.

Seidel, D.J., Fu, Q., Randel, W.J. & Reichler, T.J. 2008. Widening of the tropical belt in a changing climate. Nature geoscience. 1(1):21.

Shepherd, T.G., Boyd, E., Calel, R.A., Chapman, S.C., Dessai, S., Dima-West, I.M., Fowler, H.J., James, R., et al. 2018. Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. Climatic Change. 151(3):555–571. DOI: 10.1007/s10584-018-2317-9.

Simmonds, I., Burke, C. & Keay, K. 2008. Arctic climate change as manifest in cyclone behavior. Journal of Climate. 21(22):5777–5796.

Simmonds, I., Keay, K. & Tristram Bye, J.A. 2012. Identification and climatology of Southern Hemisphere mobile fronts in a modern reanalysis. Journal of Climate. 25(6):1945–1962.

Slivinski, L.C., Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C., Allan, R., Yin, X., et al. 2019. Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. Quarterly Journal of the Royal Meteorological Society. 145(724):2876–2908.

Sousa, P.M., Blamey, R.C., Reason, C.J.C., Ramos, A.M. & Trigo, R.M. 2018. The 'Day Zero'Cape Town drought and the poleward migration of moisture corridors. Environmental Research Letters. 13(12):124025.

Spensberger, C. & Spengler, T. 2020. Feature-based Jet Variability in the Upper Troposphere. Journal of Climate. 33(16):6849–6871.

Spensberger, C., Reeder, M.J., Spengler, T. & Patterson, M. 2019. The Connection between the Southern Annular Mode and a Feature-Based Perspective on Southern Hemisphere Midlatitude Winter Variability. Journal of Climate. 33(1):115–129. DOI: 10.1175/JCLI-D-19-0224.1.

Stratton, R.A., Senior, C.A., Vosper, S.B., Folwell, S.S., Boutle, I.A., Earnshaw, P.D., Kendon, E., Lock, A.P., et al. 2018. A pan-African convection-permitting regional climate simulation with the Met Office Unified Model: CP4-Africa. Journal of Climate. 31(9):3485–3508.

Sun, X., Cook, K.H. & Vizy, E.K. 2017. The South Atlantic Subtropical High: Climatology and Interannual Variability. Journal of Climate. 30(9):3279–3296. DOI: 10.1175/JCLI-D-16-0705.1.

Swain, D.L., Tsiang, M., Haugen, M., Singh, D., Charland, A., Rajaratnam, B. & Diffenbaugh, N.S. 2014. The extraordinary California drought of 2013/2014: Character, context, and the role of climate change. Bull. Am. Meteorol. Soc. 95(9):S3–S7.

Taylor, K.E., Stouffer, R.J. & Meehl, G.A. 2012. An overview of CMIP5 and the experiment design. Bulletin of the American meteorological Society. 93(4):485–498.

Tebaldi, C. & Sansó, B. 2009. Joint projections of temperature and precipitation change from multiple climate models: a hierarchical Bayesian approach. Journal of the Royal Statistical Society: Series A (Statistics in Society). 172(1):83–106.

Tennant, W.J.& Reason, C.J.C. 2005. Associations between the global energy cycle and regional rainfall in South Africa and Southwest Australia. Journal of climate. 18(15):3032–3047.

Toniazzo, T. & Woolnough, S. 2014. Development of warm SST errors in the southern tropical Atlantic in CMIP5 decadal hindcasts. Climate dynamics. 43(11):2889–2913.

Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E. & Yu, T. 2014. scikit-image: image processing in Python. PeerJ. 2:e453.

Wittman, M.A.H., Charlton, A.J. & Polvani, L.M. 2005. On the Meridional Structure of Annular Modes. Journal of Climate. 18(12):2119–2122. DOI: 10.1175/JCLI3394.1.

Wolski, P., Jack, C., Tadross, M., van Aardenne, L. & Lennard, C. 2018. Interannual rainfall variability and SOM-based circulation classification. Climate Dynamics. 50(1):479–492. DOI: 10.1007/s00382-017-3621-1.

Wolski, P., Conradie, S., Jack, C. & Tadross, M. 2020. Spatio-temporal patterns of rainfall trends and the 2015–2017 drought over the winter rainfall region of South Africa. International Journal of Climatology. n/a(n/a). DOI: 10.1002/joc.6768.

Yin, J.H. 2005. A consistent poleward shift of the storm tracks in simulations of 21st century climate. Geophysical Research Letters. 32(18).

Ziervogel, G., Johnston, P., Matthew, M. & Mukheibir, P. 2010. Using climate information for supporting climate change adaptation in water resource management in South Africa. Climatic Change. 103(3):537–554. DOI: 10.1007/s10584-009-9771-3.