# Earth and Space Science

# Forecast Skill of Minimum and Maximum Temperatures on Subseasonal-to-Seasonal Timescales Over South Africa

**Steven Phakula[1,2], Willem A. Landman[2], Christien J. Engelbrecht[1,2], and Thabo Makgoale[1]**

[1]South African Weather Service, Pretoria, South Africa, [2]Department of Geography, Geoinfomatics and Meteorology, University of Pretoria, Pretoria, South Africa

**Abstract** Forecast skill of three subseasonal-to-seasonal models and their ensemble mean outputs are evaluated in predicting the surface minimum and maximum temperatures at subseasonal timescales over South Africa. Three skill scores (correlation of anomaly, root-mean-square error, and Taylor diagrams) are used to evaluate the models. It is established that the subseasonal-to-seasonal models considered here have skill in predicting both minimum and maximum temperatures at subseasonal timescales. The correlation of anomaly indicates that the multimodel ensemble outperforms the individual models in predicting both minimum and maximum temperatures for the day 1–14, day 11–30, and full calendar month timescales during December months. The Taylor diagrams suggest that the European Centre for Medium-Range Weather Forecasts model and MM performs better for the day 11–30 timescale for both minimum and maximum temperatures. In general, the models perform better for minimum than maximum temperatures in terms of root-mean-square error. In fact, the skill difference in terms of correlation of anomalies (CORA) is small.

**Plain Language Summary** This study evaluates the forecast skill of three subseasonal-to-seasonal models (European Centre for Medium-Range Weather Forecasts, Centre National de Recherche Meteorologues, and United Kingdom Meteorological Office) in predicting minimum and maximum temperatures during December months at subseasonal timescales over South Africa. All three models have skill to a certain extent in predicting the day 1–14, day 11–30, and the full calendar month for both minimum and maximum temperatures.

## 1. Introduction

For the past few decades National Meteorological and Hydrological Services including the South Africa Weather Service (SAWS) have been routinely issuing weather forecasts (0–7 days) and seasonal climate forecasts (3–9 months) for surface temperatures and rainfall. Forecasting the day-to-day weather is primarily dependent on the atmospheric initial conditions, whereas forecasting at the seasonal to multiannual timescales depends on the slowly evolving components of the Earth system such as sea surface temperatures, soil moisture, and sea ice components (White et al., 2017; Tian et al., 2017). In between weather forecasting and seasonal climate predictions falls the intraseasonal or subseasonal timescale (10–60 days) that has been long neglected because this timescale is difficult to predict (e.g., Luo & Wood, 2006; Hudson et al., 2011; Li & Robertson, 2015; Olaniyan et al., 2018). This timescale is difficult to predict due to the fact that the lead time is sufficiently long that much of the memory of the atmospheric initial conditions is lost and it is too short a time range for the variability of the ocean to have a strong influence on the atmosphere (Black et al., 2017; Vitart, 2013; Vitart et al., 2017; White et al., 2017). Despite the challenges that come with subseasonal predictions, there is an increasing demand from the applications community for skillful forecasts on these timescale (2 weeks to 2 months), recently referred to as the subseasonal-to-seasonal (S2S) timescale.

The need for S2S forecasts triggered government agencies and science community to invest more on resources to improve the skill and to promote the utility of these forecasts in recent years (e.g., Hudson et al., 2013; Li & Robertson, 2015). The accuracy of S2S forecasts relies on the skilful prediction of the large-scale atmospheric circulations, which is closely linked to large-scale teleconnection patterns (Black et al., 2017). These teleconnection patterns reflect large-scale changes in the atmospheric wave and jet stream patterns and thus have strong impacts on temperature, precipitation, and storm tracks over vast geographical

areas (Black et al., 2017). Accurate climate predictions at different timescales including subseasonal and seasonal are very crucial for decision makers in sectors such as agriculture, energy, and health, among others (e.g., Hudson et al., 2011; Tian et al., 2017). In particular, the forecasts for frequency or duration of precipitation and temperature extremes can be directly tailored to different applications need.

S2S forecast information is crucially important since early warning systems for high-impact weather events could be derived from these predictions (e.g., Tian et al., 2017). In particular, many extreme weather events (e.g., floods and droughts, heat, and cold waves) and their corresponding management decisions fall into subseasonal timescales. Furthermore, subseasonal forecast information can be used for developing strategies for proactive natural disaster mitigation (Tian et al., 2017). However, it is worth mentioning that scientific challenges around improving the predictive skill of S2S forecasts and quantifying their limitations and uncertainties remain and are areas of active research. These include design issues around initial conditions, model resolution, and downscaling, to mention a few. Addressing these challenges requires increased inclusion of quantitative information regarding uncertainty and forecast quality.

On the flip side of the coin, skilful S2S predictions are possible due to improvement of numerical prediction models, ensemble prediction techniques, and initialization (e.g., Black et al., 2017; Hudson et al., 2011; Vitart et al., 2008). This improvement created a great opportunity to further improve the skill of S2S forecasts. Improving the skill of S2S forecasts is paramount and could increase their value to society. Enhancing skill begins with understanding sources and limits of S2S predictability within the Earth system. Literature reveals that there are potential sources of predictability for S2S timescale, including the Madden-Julian Oscillation (MJO), the state of El Niño Southern–Oscillation (ENSO), soil moisture, snow cover and sea ice, stratosphere-troposphere interaction, and tropical-extratropical teleconnections (e.g., Ichikawa & Inatsu, 2017; Kim et al., 2014; Vitart et al., 2017; Wang et al., 2016). The objective of this study is to evaluate the skill of the S2S model forecasts in predicting surface temperatures at subseasonal timescales over South Africa. Skilful S2S temperature forecasts at timescales considered in this study could be beneficial to decision makers in sectors such as agriculture in South Africa, as only weather forecasts and seasonal climate outlook are issued by SAWS. The remainder of the paper is organized as follows. Section 2 provides the data sources and the detailed description of methods used to evaluate the skill of S2S models. The results are presented in section 3, and, finally, discussion and conclusions are summarized in section 4.

## 2. Data and Methods

### 2.1. Model Data

The reforecasts (hindcasts) of the European Centre for Medium-Range Weather Forecasts (ECMWF; Vitart, 2014, Vitart et al., 2008), Meteo-France/Centre National de Recherche Meteorologues (CNRM; Voldoire et al., 2013), and the United Kingdom Meteorological Office (UKMO) models from the S2S project database (http://apps.ecmwf.int/datasets/data/s2s) are used. Table 1 provides detailed information of the three models of interest. The ECMWF reforecasts are produced twice a week (Monday and Thursday) "on the fly," meaning every week two new sets of reforecasts are produced to calibrate the Monday and Thursday real-time ensemble forecasts of the following week using the latest version of Integrated Forecast System. Reforecasts of the ECMWF consist of an 11-member ensemble starting the same day and month as a real-time forecast (Monday and Thursday) covering the past 20 years. The UKMO reforecasts are also produced on the fly, on the 1st, 9th, 17th, and 25th each month. The hindcasts consist of a seven-member ensemble per year from 25 March 2017. The CNRM reforecasts data set is a "fixed" data set that means that the reforecasts are produced once from a "frozen" version of the model and are used for a number of years to calibrate the real-time forecasts. The reforecasts consist of a 15-member ensemble starting on the 1st and 15th calendar day of each month for the period 1993–2014. Only the first available December start dates for all the models are considered in this study. Here we have chosen a 1.5° × 1.5° common grid resolution for all the models for the period 1998–2014. The individual model's reforecast for surface minimum temperatures (Tn), maximum temperatures (Tx) over land are considered and the 500-hPa geopotential height fields (Z500) daily data for December months are used to calculate the first fortnight ($d + 1$ to $d + 14$ days) and day 11–30 ($d + 11$ to $d + 30$ days) averages as well as to calculate the full calendar month ($d + 1$ to $d + 31$ days) averages. The $d$ is the start of reforecast dates. Day 1–14 (first fortnight) timescale is considered to check the influence of the

**Table 1**
*The Description of Three S2S Model Forecasts*

| Model | Time range | Resolution | Rfc | Rfc size | Rfc freq | Rfc period |
|-------|-----------|------------|-----|----------|----------|------------|
| ECMWF | 0–46 days | 0.25° × 0.25°, L91 | On the fly | 11 | Twice/week | Past 20 years |
| CNRM | 0–61 days | ~0.7° × 0.7°, L91 | Fixed | 15 | Twice/month | 1993–2014 |
| UKMO | 0–60 days | ~0.5° × 0.5°, L85 | On the fly | 7 | Four/month | 1993–2016 |

*Note.* Time range in forecast lead time (days), resolution is longitude and latitude resolution (°), and the number after L represents the number of vertical levels. Reforecast (rfc) are run using the actual forecast model but for the past several years on the same (or nearby) calendar day as the forecast. The reforecast is used to calibrate the actual forecast. Rfc size is the number members for reforecasts, rfc freq is how often (frequency) the forecast run, rfc period is the number of years the reforecasts are run (source: Vitart et al., 2017).

first 7 days in the models, and this timescale has been explored in Australia (Hudson et al., 2011). They investigated the skill of the Predictive Ocean Atmosphere Model for Australia in predicting Tn and Tx for day 1–14 (first fortnight) and day 15–28 (second fortnight). Day 11–30 is the extended-range weather forecasting timescale and form part of operational product suits at the SAWS, and the monthly calendar falls in the long-range timescale, which is within the definition of S2S timescale (2 weeks up to 2 months) (e.g., World Meteorological Organization, 2016; Viguad et al., 2017; Tian et al., 2017; Ford et al., 2018). The month December is chosen for exploring S2S predictability because it is one of the hottest summer months and most of the heat waves occur during summer months in South Africa (e.g., Herrings et al., 2018; Lyon, 2009). It is worth mentioning that this work is the first to be conducted on S2S predictions in South Africa to our knowledge.

### 2.2. Reanalysis Data

The ECMWF Interim reanalysis (ERA-Interim; Dee et al., 2011) daily surface Tn, Tx, and Z500 data sets are used to calculate the day 1-14 and the day 11-30 averages as well as the full calendar month averages that match the model reforecasts data from 1998 to 2014. The ERA-Interim reanalysis is used here for verification because it is easily accessible from the ECMWF database. The reanalysis data are at the same resolution (1.5° × 1.5°) as the models. It is acknowledged that the ERA-Interim reanalysis might favor ECMWF in the model verification. To address this issue, the ECMWF forecasts were also verified by making use of an independent observational data set (National Oceanic and Atmospheric Administration [NOAA] Climate Prediction Center global temperature data). This revealed no significant difference in the verification scores derived from the independent data set versus that of the ERA-Interim reanalysis data.
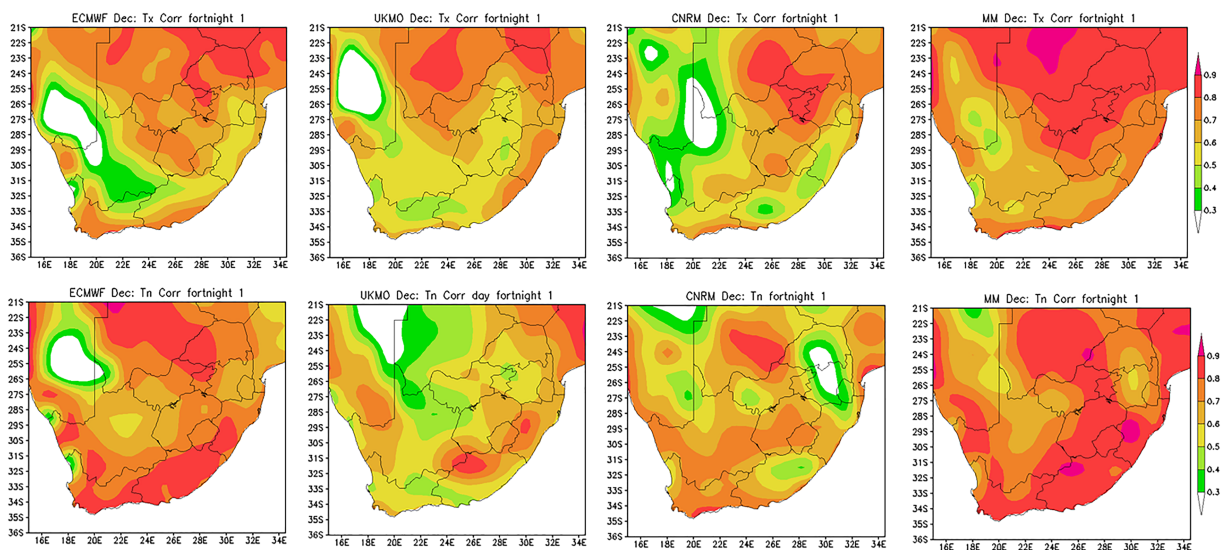


**Figure 1.** CORA of ECMWF, UKMO, CNRM, and multimodel ensemble for average maximum temperature (top panels) and average minimum temperature (bottom panels) anomalies for the December day 1–14 timescale from 1998 to 2014. Any CORA >0.3 is significant at the 5% level based on a permutation method repeated 10,000 times.
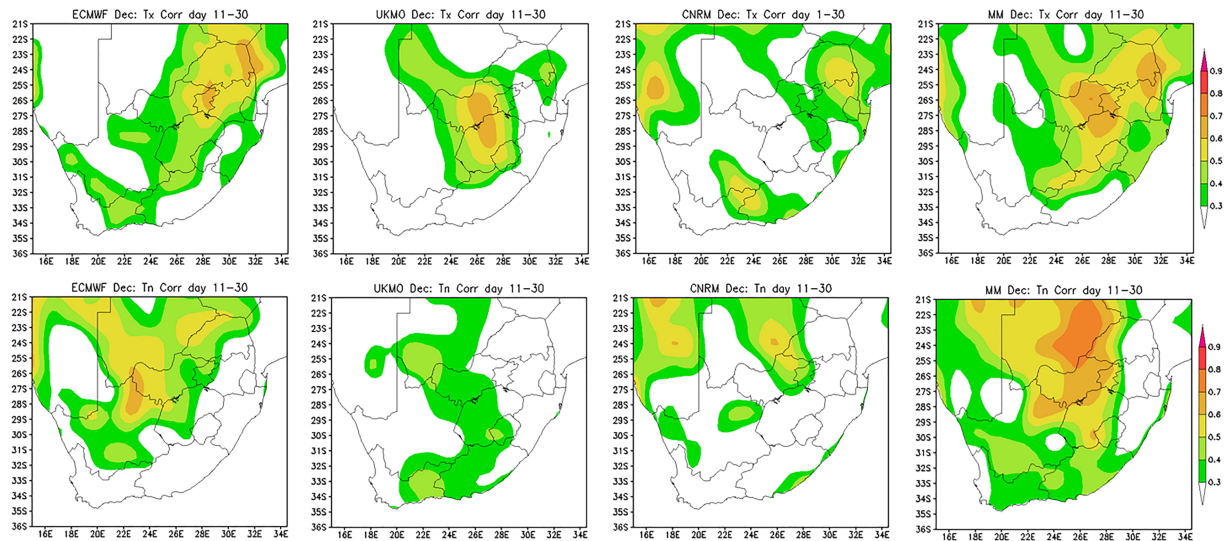
**Figure 2.** As in Figure 1 but for the day 11–30 timescale.

### 2.3. Verification Scores

Forecast verification is the process of determining the quality of a forecast through assessment of the degree of similarity between the forecasts and the observations (Mandal et al., 2007; Wilks, 2006, 2011). Verification of forecasts is mostly performed to check if there is a strong relationship between the forecasts and the observations and if the results provide an accurate indication of how good or bad subsequent forecasts will be (Mason, 2008). Deterministic S2S verification is mostly assessed in a form of correlation referred to as CORA (correlation of anomalies), which measures the linear correspondence between the ensemble mean forecast and the observations (Hudson et al., 2011; Li & Robertson, 2015). The model forecasts and the observed anomalies are computed with respect to their own seasonally varying climatologies, respectively. More importantly, the forecast anomalies are computed with respect to lead-dependent hindcast climatologies. Here the CORA is used to measure a model's performance in predicting Tn and Tx over South Africa. CORA is a correlation over the initialization time dimension of the forecasts. Statistical significance of the CORA is tested at 95% confidence interval using random permutation method repeated 10,000 times.
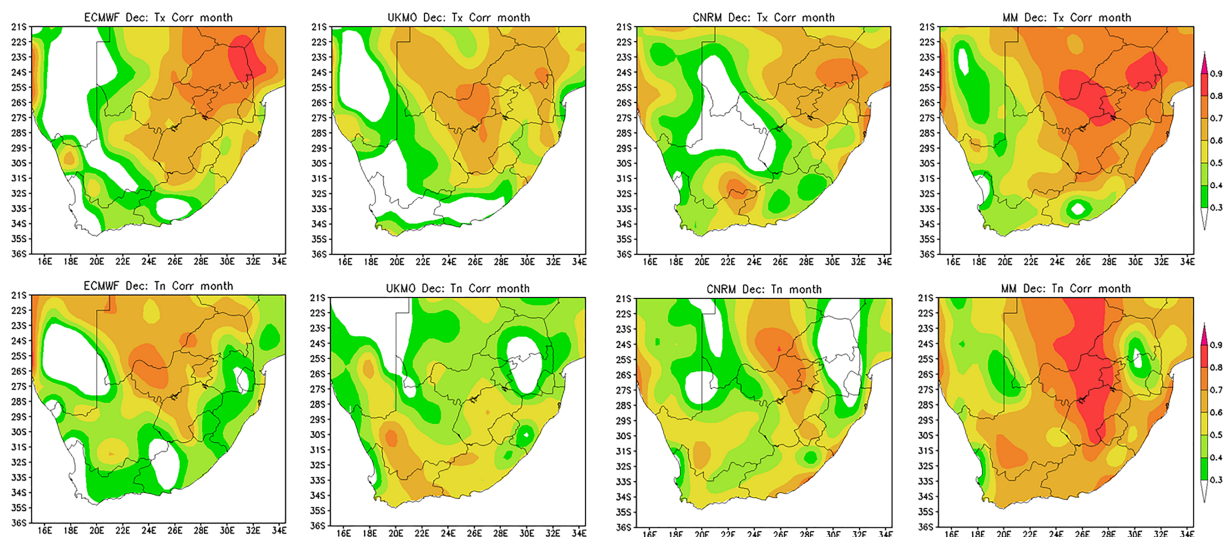


**Figure 3.** As in Figure 2 but the for full calendar month timescale.
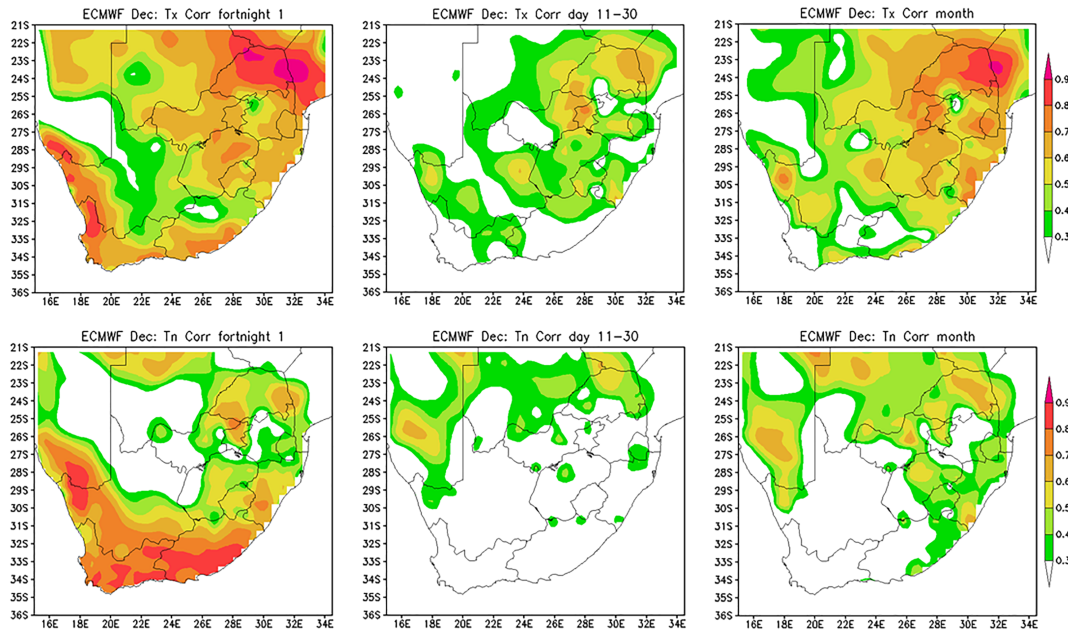
**Figure 4.** CORA of the ECMWF against the CPC global daily temperature data set for average maximum temperature (top row) and average minimum temperature (bottom row) anomalies for the day 1–14, day 11–30, and full calendar month timescales during December from 1998 to 2014. Any CORA >0.3 is significant at the 5% level based on a permutation method repeated 10,000 times.

Another skill score considered is the root-mean-square error (RMSE), which measures the mean square gap between the observed and forecasts data (e.g., Chai & Draxler, 2014; Zhang & Casey, 2000). The RMSE is a popular statistical measure for the performance of numerical models in atmospheric research and is a good criteria to classify the accuracy of a model; a low index indicates higher accuracy (Mugume et al., 2016). In addition to the CORA and RMSE, Taylor diagrams (Taylor, 2001) are also used to evaluate the models. The domain for the Taylor diagrams is 21–36°S and 15–35°E covering southern Africa. These diagrams are widely used to evaluate multiple aspects of complex models and measure the relative skill of many different models against observations (e.g., Intergovernmental Panel on Climate Change, 2013; Kalognomou et al., 2013). The
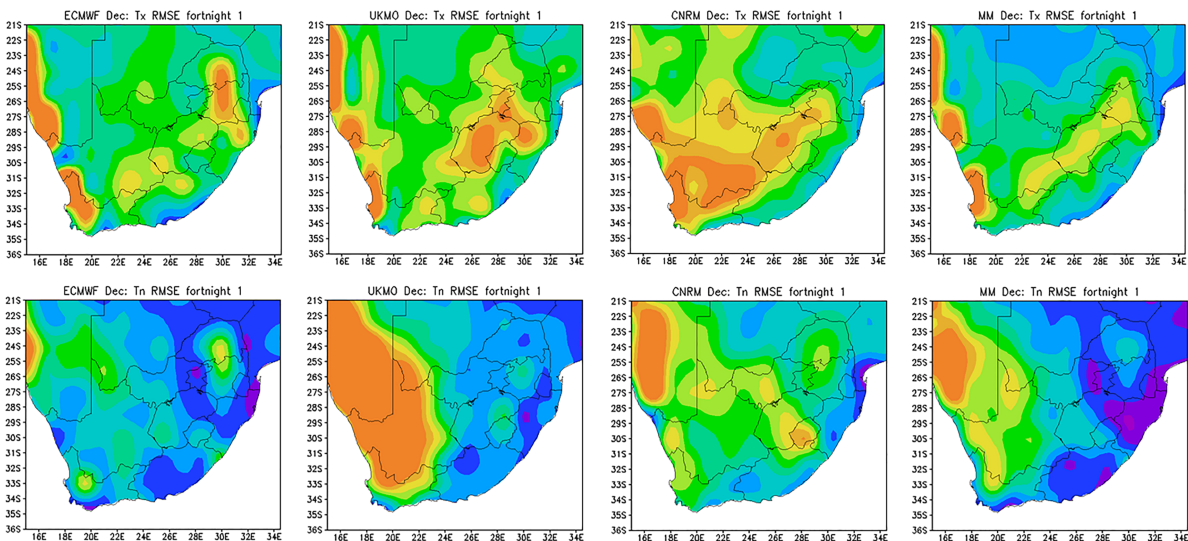


**Figure 5.** RMSE of the ECMWF, UKMO, CNRM, and multimodel ensemble for average maximum temperatures (top row) and average minimum temperatures (bottom row) for the December days 1–14 timescale from 1998 to 2014. The
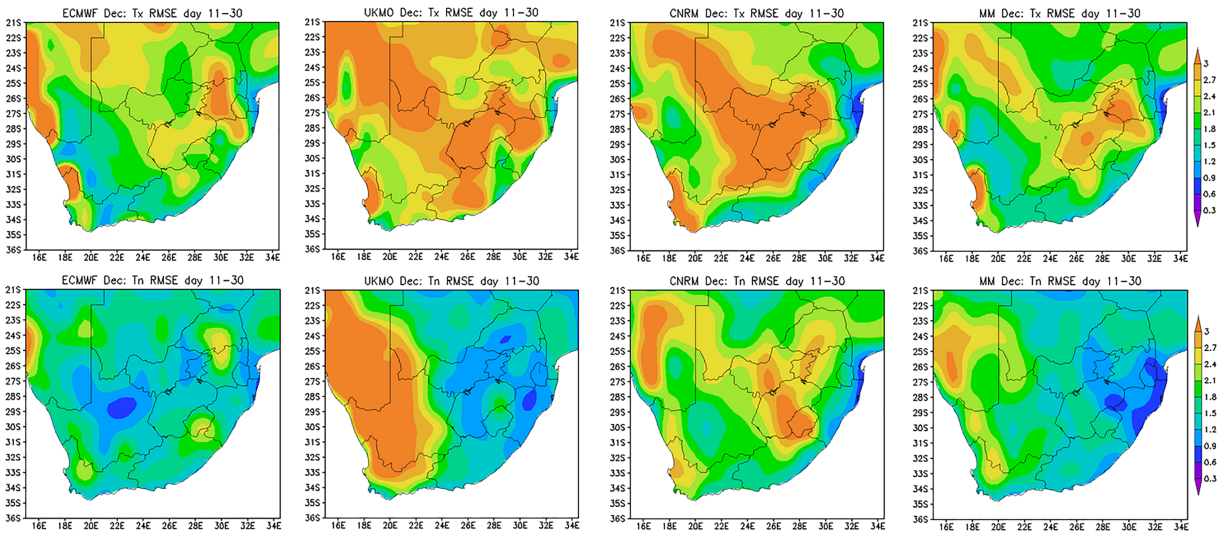
**Figure 6.** As in Figure 5 but for the day 11–30 timescale.

diagrams provide a statistical way of graphically summarizing how well the modeled pattern (or a set of patterns) matches observations in terms of their correlation, their centered RMSE, and the amplitude of their variations (represented by their standard deviations) (Heo et al., 2013). The diagram is usually visualized as a series of points on a polar plot. The azimuth angle pertaining to each point is such that the cosine of the angle is equal to the correlation coefficient between the modeled and observed data. Radial distance from the origin in the Taylor diagram represents the ratio of the standard deviation of the simulation to that of the observation, and the correlation between two fields is given by the azimuthal position. The distance from the reference point (observations) is a measure of the centered RMSE.

## 3. Results

### 3.1. CORA

The ECMWF, UKMO, and CNRM model outputs as well as the multimodel (the average ensemble mean of the three models), here referred to as MM, are evaluated against the Era-Interim reanalysis in predicting Tn and Tx at subseasonal timescales during December months over South Africa. In fact, the MM is considered
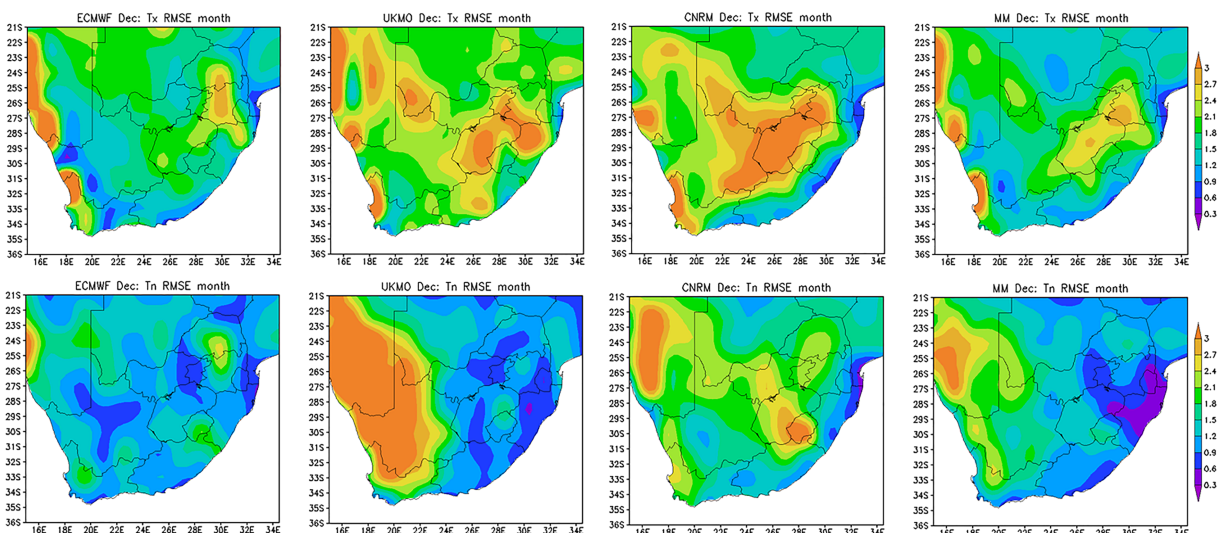


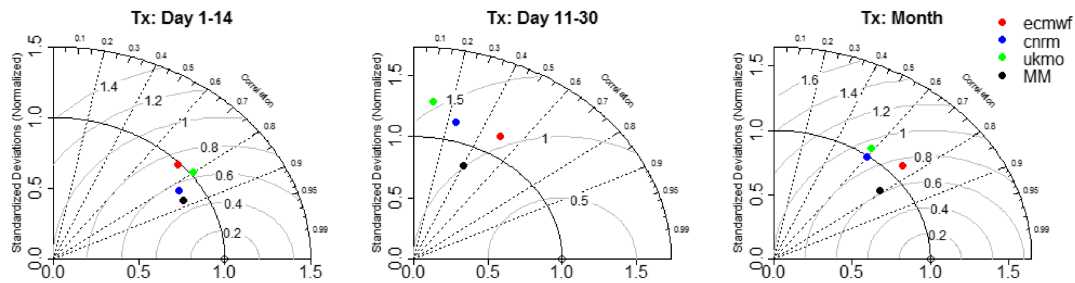**Figure 7.** As in Figure 6 but for the full calendar month timescale.

**Figure 8.** Taylor diagrams for area-averaged maximum temperatures for the day 1–14, 11–30, and monthly maximum temperature for South Africa. The open circle along the horizontal axis represents the observation (ERA-Interim) data, and the individual S2S models are shown by solid colored circles. The correlation between models and observation is shown on the azimuthal axis.

here because it is known to improve forecast skill in climate models (e.g., Viguad et al., 2017). The CORA maps show that all three models have skill in predicting both the Tn and Tx for the day 1–14 timescale over South Africa (Figure 1). Moreover, the correlation is highest over the northeastern parts of the country for Tx and central to southern parts for Tn. It is evident that the MM improves the forecast skill for both the Tn and Tx. The skill over the eastern parts of South Africa could be linked to ENSO, as ENSO plays an important role on the S2S timescale, as a source of predictability. Furthermore, the skill over the same areas is found in seasonal prediction studies (e.g., Lazenby et al., 2014). However, it is clear that all the three S2S models exhibit relatively low skill in predicting both the Tn and Tx for the day 11–30 timescale (Figure 2), with forecast skill diminishing significantly as compared to predicting the day 1–14 timescale. Moreover, the ECMWF model outperforms the UKMO and the CNRM models for predicting the day 11–30 timescale. For the calendar month timescale the CORA indicates that all the individual models perform better compared to predicting the day 11–30 (Figure 3). In general, the forecast skill for all models is concentrated over the eastern half of South Africa for the calendar month timescale. It is worth noting that the MM outperforms the individual models in predicting both the Tn and Tx for all the timescales. To assess if there is bias between the ECMWF model and the ERA-Interim, we used NOAA Climate Prediction Center global temperature data set (https://www.esrl.noaa.gov/psd/data/gridded/index.html) for verification and comparable results were found for Tx (Figure 4).

### 3.2. RMSE

The RMSE maps show that all three individual models are more skilful in predicting the Tn as compared to Tx for all the timescales (Figures 5–7). The errors for the day 1–14 and calendar months timescales are smaller as compared to that of the day 11–30 timescale. In general, the ECMWF and UKMO models have a small error as compared to the CNRM model for all the timescales considered. This result is in agreement with the CORA results in the previous paragraph. The errors for predicting Tx are much higher as compared to those for Tn for all the timescales. Moreover, this result suggests that Tn is more predictable than Tx for all the timescales considered here. Once again the MM of the three models improves the forecast skill in predicting both the Tn and Tx for all timescales.
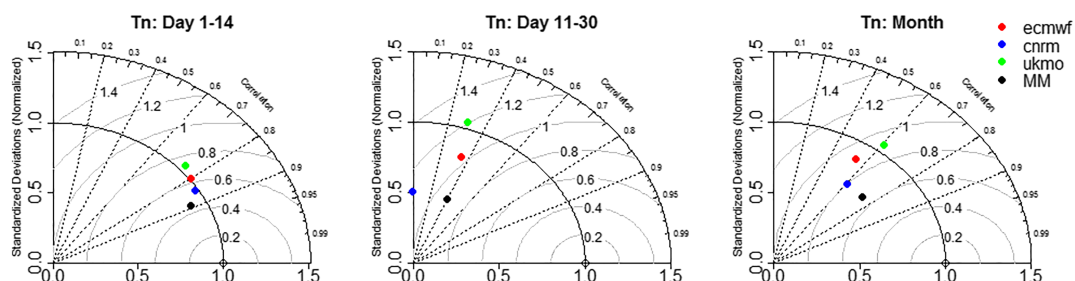


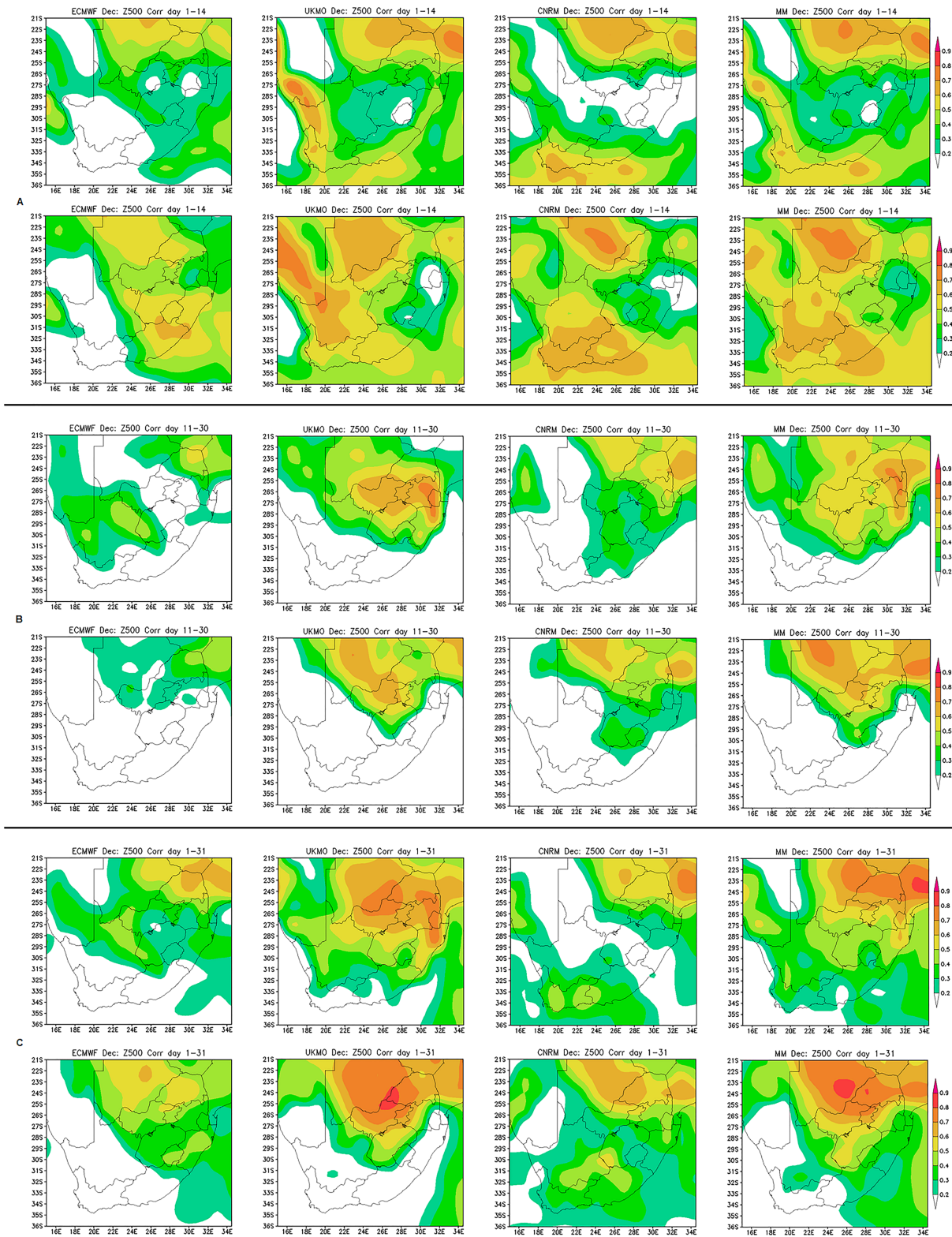**Figure 9.** As in Figure 8 but for the minimum temperatures.

**Figure 10.** CORA for the 500Z against maximum (top row) and minimum (bottom row) temperatures for the day 1–14 (a), day 11–30 (b), and day 1–31 (c) timescales. Any CORA >0.2 is significant at the 5% level based on a permutation method repeated 10,000 times.

### 3.3. Taylor Diagrams

Figures 8 and 9 show the Taylor diagrams of Tx and Tn, respectively. The models are more skilful in predicting the day 1–14 and calendar month timescales as compared to the day 11–30, with correlation coefficient greater than 0.5 for Tx. The correlation is relatively low (less than 0.5) for the day 11–30 timescale, except for the ECMWF Tx forecasts. The MM outperforms all the individual models with a higher correlation coefficient of about 0.9 for the day 1–14 and calendar month timescales. The RMSE of MM is lower than all individual models in all forecasts; however, the correlation coefficient is relatively low for the day 11–30 timescale. With regard to evaluation of Tn, similar results are found as that of evaluating Tx forecasts (Figure 9). The models exhibit a general good performance for the day 1–14 and calendar months simulation and poor performance for the day 11–30 forecasts. Again the MM performs relatively well as compared to individual models with higher correlation coefficient and lower RMSE.

### 3.4. CORA for Large-Scale Atmospheric Circulations

The Z500 large-scale atmospheric circulations of the ECMWF, UKMO, and CNRM S2S models are used to evaluate the potential source of prediction skill of Tx and Tn for the day 1–14, day 11–30, and day 1–31 (full calendar month) timescales during December over South Africa. It is worth mentioning that the 850-hPa geopotential height fields (Z850) and Z500 have been used to predict both Tx and Tn at seasonal timescale over southern Africa (e.g., Lazenby et al., 2014). In fact, the SAWS is using the Z850 to predict rainfall and temperatures (Tx and Tn) on an operational basis. The Z500 CORA exhibit relatively low to no significant values over the central eastern parts of South Africa for both Tx and Tn for the day 1–14 timescale (Figure 10). Furthermore, the ECMWF model CORA values are not significant over the western parts of the country. For the day 11–30 timescale, high CORA values for both Tx and Tn are generally concentrated over the northeastern parts of South Africa, with no significant values over the western parts of the country for all the models. All the models exhibit relatively high CORA values for both Tx and Tn over north eastern parts of South Africa for full calendar month timescale. It is worth noting that for all the timescales the MM exhibits improved CORA values as compared to all three individual models suggesting a correlation between the Z500 and temperature anomalies, thus the atmospheric circulation seems to be a potential source of predictability.

## 4. Discussion and Conclusions

National Meteorological and Hydrological Services around the world, including SAWS, routinely issue weather forecasts and seasonal climate forecasts for surface minimum and maximum temperatures and rainfall. In between weather forecasting and seasonal climate predictions falls the intraseasonal or subseasonal timescale that has been long neglected because this timescale is difficult to predict. Despite the challenges that come with subseasonal predictions, there is an increasing demand from the applications community for skillful forecasts on these timescales. Subseasonal predictions are important since they are the ones that early warning systems for high-impact weather events could be derived from. In particular, since many extreme weather events (e.g., floods and droughts, heat, and cold waves) and their corresponding management decisions fall into subseasonal timescales. Furthermore, subseasonal forecast information can be used for developing strategies for proactive natural disaster mitigation.

In this study the skill of the ECMWF, UKMO, and CNRM S2S model outputs, and their average ensemble mean are evaluated in predicting surface maximum and minimum temperatures at subseasonal timescales over South Africa. The RMSE indicates that all three models perform better predicting minimum temperatures as compared to maximum temperatures. The error is much higher for maximum temperatures for all the timescales during the December months. The CORA and the Taylor diagrams show that all the three models perform better in predicting both average minimum and maximum temperatures for day 1–14 and calendar month timescales as compared to the day 11–30 timescale during the December months. The reason for this could be the fact that the first 7 days (weather forecasting time scale) are included in both the day 1–14 and calendar month timescales. In fact, all the models considered here exhibit relatively low skill in predicting the day 11–30 timescale for both the minimum and maximum temperatures. It is evident from the CORA and the Taylor diagrams that the multimodel outperforms all the models in predicting both minimum and maximum temperatures for all the timescales. The Z500 large-scale atmospheric circulations seem to have an influence in the predictability of Tx and Tn, particularly over the eastern half of South Africa.

We can conclude from the results that the MM outperforms the individual S2S models considered here in predicting both minimum and maximum temperatures at subseasonal timescales considered here over South Africa. In fact, it was established that the MM for subseasonal forecasts generally improves on the individual model skill in most cases (e.g., Moron et al., 2018). For this reason, multimodel forecasting systems for minimum and maximum extreme temperatures at subseasonal timescales can be developed for South Africa.

# References

Black, J., Johnson, N. C., Baxter, S., Feldstein, S. B., Harnos, D. S., & L'Heureux, M. L. (2017). The predictors and forecast skill of Northern Hemisphere teleconnection patterns for lead times of 3-4 weeks. *Monthly Weather Review*, *145*, 2855–2877. https://doi.org/10.1175/MWR-D-16-0394.1

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Ford, T. W., Dirmeyer, P. A., & Benson, D. O. (2018). Evaluation of heat wave forecasts seamlessly across subseasonal timescales, *npj Climate and Atmosphere Science*, 2018. https://doi.org/10.1038/s4161-018-0027-7.

Heo, K.-Y., Ha, K.-J., Yun, K.-S., Lee, S.-S., Kim, Y.-J., & Wang, B. (2013). Methods for uncertainty assessment of climate models and model predictions over East Asia. *International Journal of Climatology*, *34*(2), 377–390. https://doi.org/10.1002/joc.3692

Herrings, S. C., Christidis, N., Hoell, A., Kossin, J. P., Schreck, C. J. III, & Stott, P. A. (Eds) (2018). Explaining extreme events of 2016 from a climate perspective. *Bulletin of the American Meteorological Society*, *99*(1), S1–S157.

Hudson, D., Alves, O., Hendon, H. H., & Marshall, A. G. (2011). Bridging the gap between weather and seasonal forecasting: Intraseasonal forecasting for Australia. *Quarterly Journal of the Royal Meteorological Society*, *137*, 673–689. https://doi.org/10.1002/qj.769

Hudson, D., Marshall, A. G., Yin, Y., Alves, O., & Hendon, H. H. (2013). Improving intraseasonal prediction with a new ensemble general strategy. *Monthly Weather Review*, *141*(12), 4429–4449. https://doi.org/10.1175/MWR-D-13-00059.1

Ichikawa, Y., & Inatsu, M. (2017). An alternative estimate of potential predictability on the Madden-Julian oscillation phase space using S2S models. *Atmosphere*, *8*(12), 150. https://doi.org/10.3390/atmos8080150

Intergovernmental Panel on Climate Change (2013). In T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 1535). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/CBO9781107415324

Kalognomou, E.-A., Lennard, C., Shongwe, M., Pinto, I., Favre, A., Kent, M., et al. (2013). A diagnostic evaluation of precipitation in CORDEX models over southern Africa. *Journal of Climate*, *26*(23), 9477–9506. https://doi.org/10.1175/JCLI-D-12-00703.1

Kim, H.-M., Webster, P. J., Toma, V. E., & Kim, D. (2014). Predictability and prediction of the MJO in two operational forecasting systems. *Journal of Climate*, *27*(14), 5364–5378. https://doi.org/10.1175/JCLI-D-13-00480.1

Lazenby, M. J., Landman, W. A., Garland, R. M., & DeWitt, D. G. (2014). Seasonal temperature prediction skill over Southern Africa. *Meteorological Applications*, *21*(4), 963–974. https://doi.org/10.1002/met.1449

Li, S., & Robertson, A. W. (2015). Evaluation of submonthly precipitation forecast skill from global ensemble prediction system. *Monthly Weather Review*, *143*(7), 2871–2889. https://doi.org/10.1175/MWR-D-14-00277.1

Luo, L., & Wood, E. F. (2006). Assessing the idealized predictability of precipitation and temperature in the NCEP climate forecast system. *Geophysical Research Letters*, *33*, L04708. https://doi.org/10.1029/2005GL025292

Lyon, B. (2009). Southern Africa summer drought and heat waves: Observation and coupled model behavior. *Journal of Climate*, *22*, 60–6046. https://doi.org/10.1175/2009JCLI3101.1

Mandal, V., Utpal, K. D., & Biplab, K. B. (2007). Precipitation forecast verification of the Indian summer monsoon with intercomparison of three diverse regions. *Weather and Forecasting*, *22*, 428–443. https://doi.org/10.1175/WAF1010.1

Mason, S. J. (2008). Understanding forecast verification statistics. *Meteorological Applications*, *15*, 31–40. https://doi.org/10.1002/met.52

Moron, V., Robertson, A. W., & Vitart, F. (2018). Editorial: Sub-seasonal to seasonal predictability and prediction of monsoon climates. *Frontiers in Environmental Science*, *6*, 83. https://doi.org/10.3389/fenvs.2018.00083

Mugume, I., Basalirwa, C., Waiswa, D., Reuder, J., Mesquita, M. S., Tao, S., & Ngailo, T. J. (2016). Comparison of parametric and non-parametric methods for analysing the bias of a numerical model. *Modelling and simulation in Engineering*, Volume 2016, Article ID 7530759, 7 pages. https://doi.org/10.1155/2016/7530759

Olaniyan, E., Adefisan, E. A., Oni, F., Afiesimama, E., Balogun, A. A., & Lawal, K. A. (2018). Evaluation of the ECMWF sub-seasonal to seasonal precipitation forecasts during the peak of West Africa monsoon in Nigeria. *Frontiers in Environmental Science*, *6*, 4. https://doi.org/10.3389/fenvs.2018.00004

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Tian, D., Wood, E. F., & Yuan, X. (2017). CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States. *Hydrology and Earth System Sciences*, *21*(3), 1477–1490. https://doi.org/10.5194/hess-21-1477-2017

Viguad, N., Robertson, A. W., & Kippett, M. K. (2017). Multimodel ensembling of subseasonal precipitation forecasts over North America. *Monthly Weather Review*, *145*(10), 3913–3928. https://doi.org/10.1175/MWR-D-17-0092.1

Vitart, F. (2013), Evolution of ECMWF sub-seasonal forecast skill scores over the past 10 years. *ECMWF Technical Report*. http://www.ecmwf.int/sites/default/files/elibrary/2013/12932-evolution-ecmwf-sub-seasonal-forecast-skill-scores-over-past-10-years.pdf

Vitart, F. (2014). Evolution of ECMWF sub-seasonal forecast skill scores. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1889–1899. https://doi.org/10.1002/qj.2256

Vitart, F., Ardilouzo, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2017). The subseasonal to seasonal (S2S) prediction project database. *American Meteorological Society*, *98*(1), 163–173. https://doi.org/10.1175/BAMS-D-16-0017.1

Vitart, F., Buizza, R., Balmased, M. A., Balsamo, G., Bidlot, J.-R., Bonet, A., et al. (2008). The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Quarterly Journal of the Royal Meteorological Society*, *134*(636), 1789–1799. https://doi.org/10.1002/qj.322

Voldoire, A., Sanchez-Gomez, E., Melia, D. S., Decharme, B., Cassou, C., & Chauvin, F. (2013). The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dynamics*, *40*(9-10), 2091–2121. https://doi.org/10.1007/s00382-011-1259-y

Wang, K., Lin, Z.-H., Ling, J., Yu, Y., & Wu, C.-L. (2016). MJO potential predictability and predictive skill in IAP AGCM4.1. *Atmospheric and Oceanic Science Letters*, *9*(5), 388–393. https://doi.org/10.1080/16742834.2016.1211469

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J. T., Lazo, J. K., Kumar, A., et al. (2017). Review potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, *24*(3), 315–325. https://doi.org/10.1002/met.1654

Wilks, D. (2006). *Statistical methods in atmospheric science* (2nd ed.). Burlington, MA: Academic Press.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed., p. 676). Amsterdam: Academic Press.

World Meteorological Organization (2016). Use of climate predictions to manage risks, WMO-No. 1174, ISBN 978-92-63-11174-7.

Zhang, H., & Casey, T. (2000). Verification of probability forecasts. *Weather and Forecasting*, *15*(1), 80–89. https://doi.org/10.1175/1520-0434(2000)015<0080:VOCPF>2.0.CO;2