Guidance on Operational Practices for Objective Seasonal Forecasting

2020 edition

WEATHER CLIMATE WATER

WORLD
METEOROLOGICAL
ORGANIZATION

WMO-No. 1246

Guidance on Operational Practices for Objective Seasonal Forecasting

2020 edition

WORLD
METEOROLOGICAL
ORGANIZATION

WMO-No. 1246

EDITORIAL NOTE

METEOTERM, the WMO terminology database, may be consulted at https://public.wmo.int/en/meteoterm.

Readers who copy hyperlinks by selecting them in the text should be aware that additional spaces may appear immediately following http://, https://, ftp://, mailto:, and after slashes (/), dashes (-), periods (.) and unbroken sequences of characters (letters and numbers). These spaces should be removed from the pasted URL. The correct URL is displayed when hovering over the link or when clicking on the link and then copying it from the browser.

WMO-No. 1246

Chair, Publications Board
World Meteorological Organization (WMO)
7 bis, avenue de la Paix
P.O. Box 2300
CH-1211 Geneva 2, Switzerland

Tel.: +41 (0) 22 730 84 03
Fax: +41 (0) 22 730 81 17
Email: publications@wmo.int

ISBN 978-92-63-11246-9

PUBLICATION REVISION TRACK RECORD

Date	*Part/ chapter/ section*	*Purpose of amendment*	*Proposed by*	*Approved by*

CONTENTS

ACKNOWLEDGEMENTS

The *Guidance on Operational Practices for Objective Seasonal Forecasting* has been prepared under the auspices of the World Meteorological Organization (WMO) Commission for Climatology (CCl) and Commission for Basic Systems (CBS), through concerted efforts by the following authors:

Mr Arun KUMAR (Climate Prediction Center, United States of America)
Mr Jean-Pierre CERON (former Scientific Deputy Director of Climatology, Météo France)
Mr Caio COELHO (Centre for Weather Forecast and Climate Studies, National Institute for Space Research, Brazil)
Ms Laura FERRANTI (European Centre for Medium-Range Weather Forecasts, United Kingdom of Great Britain and Northern Ireland)
Mr Richard GRAHAM (Met Office, United Kingdom)
Mr David JONES (Bureau of Meteorology, Australia)
Mr William MERRYFIELD (Environment and Climate Change Canada)
Mr Ángel MUÑOZ (International Research Institute for Climate and Society, USA)
Mr Sivananda PAI (India Meteorological Department)
Mr Ernesto RODRIGUEZ (State Agency for Meteorology, Spain)

The document was reviewed by:

Mr Zachary ATHERU (IGAD Climate Predictions and Applications Centre, Kenya)
Ms Soumaya BEN RACHED (National Institute of Meteorology, Tunisia)
Ms Kristina FRÖHLICH (German Meteorological Service)
Ms Clare GOODESS (University of East Anglia, United Kingdom)
Mr Christopher GORDON (Met Office, United Kingdom)
Ms Debbie HUDSON (Bureau of Meteorology, Australia)
Mr Andre KAMGA (African Centre for Meteorological Applications for Development, Niger)
Ms Valentina KHAN (Hydro-meteorological Research Centre of the Russian Federation)
Mr Jeffery KNIGHT (Met Office, United Kingdom)
Mr Rodney MARTINEZ (International Research Centre on El Niño, Ecuador)
Mr Yasushi MOCHIZUKI (Japan Meteorological Agency)
Mr Andrew ROBERTSON (International Research Institute for Climate and Society, USA)
Mr Stefan RÖSNER (German Meteorological Service)
Ms Maria SKANSI (National Meteorological Service, Argentina)
Mr Andrew TAIT (National Institute of Water and Atmospheric Research, New Zealand)
Mr Yuhei TAKAYA (Meteorological Research Institute, Japan)
Mr Adrian TROTMAN (Caribbean Institute for Meteorology and Hydrology, Barbados)
Mr Jin Ho YOO (APEC Climate Centre, South Korea)
Mr Peiqun ZHANG (China Meteorological Administration)

EXECUTIVE SUMMARY

With the formalization of WMO Global Producing Centres for Long-Range Forecasts (GPCs-LRF) in 2006 and the Lead Centre for Long-Range Forecast Multi-model Ensemble (LC-LRFMME) in 2009, the infrastructure for operational seasonal forecasts has reached a mature state. At present, 13 designated GPCs-LRF provide seasonal forecasts on a monthly basis. Data from these forecasts is collected by LC-LRFMME (also a WMO-designated centre), which produces a consolidated seasonal forecast based on a multi-model approach. The combination of GPCs-LRF and LC-LRFMME constitutes a solid foundation for the provision of seasonal forecasts on a global scale and is an authoritative resource for the formulation of specific seasonal forecasts for individual regions, countries and localities. However, in recent years, various WMO forums have indicated that simply having GPCs-LRF and LC-LRFMME as an authoritative resource is not enough. They have recommended the establishment of a standardized strategic approach using global information provided by these entities to develop seasonal forecasts tailored to specific regions or countries.

This document is an attempt to establish such an approach. It provides a set of principles, recommendations, and general technical guidance, all designed to facilitate the development of seasonal forecasts at the regional and national levels based on the seasonal forecasts produced at the global scale.

Background information on the concepts behind the scientific feasibility of seasonal forecasts, interpreting and communicating probabilistic forecasts, bias correction, calibration, downscaling, selecting global models and multi-model ensemble techniques is included in Chapters 1–3. Recommendations on good practices to develop regional and national forecasts, including the use of an *objective* seasonal forecast procedure (defined as a set of steps in a forecast procedure that are traceable, reproducible, and well documented and which allow quantification of forecast quality) are provided in Chapter 4.

Additional chapters provide background information on the WMO infrastructure for seasonal predictions (Chapter 5), other resources for seasonal forecast information (Chapter 6), a discussion on connecting operational aspects of seasonal forecasts to research (Chapter 7), some examples of good practices that are currently being followed (Chapter 8) and future prospects and ongoing developments in long-range forecast practices (Chapter 9). This document is not intended to provide detailed information on methodologies used in seasonal forecasts but rather to recommend good practices that should be used in the development and delivery of seasonal forecasts at the regional and national levels.

The recommendations in this document will facilitate the implementation of the Climate Services Information System (CSIS), the "operational core" of the Global Framework of Climate Services (GFCS). They follow the concept of the cascading forecast process highlighted in the CSIS, whereby seasonal forecast information on the global scale is tailored to provide information at the regional and national levels. The operational implementation of the recommendations will also require the development of some foundational tools, for example, a set of tools for bias correction and calibration; this should be taken up by the relevant WMO Expert and Task Teams. It is also expected that future updates to this guidance will take into account scientific and infrastructure developments relating to seasonal predictions and will incorporate feedback and emerging requirements from Regional Climate Centres (RCCs), Regional Climate Outlook Forums (RCOFs) and National Meteorological and Hydrological Services (NMHSs).

This guidance recommends a set of principles, summarized below, that should be followed in the development of seasonal forecasts:

1. Follow a traceable, reproducible, and well-documented procedure (including model selection, bias correction, calibration and statistical downscaling) that is amenable to assessments of forecast quality (verification);

2. Use dynamical climate models, including multi-model ensembles, as the primary basis for seasonal forecasts;

3. Establish and maintain observational databases (including databases associated with reanalysis and other blended analysis products) of adequate quality, length of record and spatial resolution for verification, bias correction and calibration and to monitor drivers of seasonal predictability;

4. Identify and monitor drivers of predictable climate variability and assess their representation and prediction skill in models;

5. Ensure that forecasts are verified according to established standards, keep archives of past forecasts, and conduct post-season assessments;

6. Provide forecast information together with historical performance (for example, skill and reliability);

7. Use clear and non-technical language to communicate seasonal forecasts, including emphasizing the probabilistic nature and inherent uncertainty of seasonal forecasts;

8. Collaborate across regions influenced by the same climate drivers in forecast production though mechanisms such as RCOFs;

9. Provide seasonal forecasts as well as regular updates on a fixed operational schedule tailored to the applicable decision-making context;

10. Establish user feedback and product upgrade mechanisms and support co-production of tailored products.

Chapter 4 contains a set of recommendations for good practices to put these principles into effect.

At present, for some regions and/or seasons, dynamical models may not be as skilful as statistical (or empirical) models. In such cases, in order to provide the best seasonal forecast information, an objective blend of empirical (statistical) and dynamical models may be used. However, statistical models should be properly trained to avoid overfitting. They should also be cross-validated to adequately assess their past performance (and to ensure that they adhere to other components in the set of principles, for example, the use of a well-documented forecast procedure). The use of statistical models should only be considered an interim measure given the routine advances in dynamical models.

BACKGROUND ON THE REQUIREMENTS FOR THIS DOCUMENT

In order to meet regional and national seasonal forecast requirements, WMO has established RCOFs and RCCs. One of the functions of these entities is to develop user-driven climate products and services and to duly inform users once these products and services have been developed. RCOFs and RCCs currently follow a wide range of practices which at times are not well documented; this is partially due to the chronological evolution of RCOFs, GPCs-LRF, LC-LRFMME and RCCs, but it is also due to geographical and climate differences between regions. RCOFs were established in 1996, followed by GPCs-LRF in 2006 and RCCs and LC-LRFMME in 2009. Because of the timing of the development of the seasonal forecast infrastructure at the global and regional levels, adequate coordination between the global and regional infrastructures was not practical and led to ad hoc and widely differing operational approaches.

Current seasonal forecast practices at RCOFs rely heavily on a consensus-based forecast process. This has several drawbacks: the forecast process is neither traceable nor reproducible (forecasts made in earlier seasons cannot be replicated by a different forecast group); forecasts are seldom available in digitized form and if needed, cannot be used in application models; because of the lack of availability of digitized data, skill assessments regarding forecast quality are difficult to make, and so forth. In light of these drawbacks, the WMO Executive Council at its sixty-ninth session decided to "…consider the adoption of objective sub-seasonal and seasonal forecasts as an overarching technical strategy, particularly at regional and national levels, promoted through RCOFs, by adopting suitable operational practices and capacity development efforts…" (Decision 18 (EC-69)) and further requested that the Commission for Climatology (CCl):

1. Develop a technical guidance on operational predictions from sub-seasonal to longer timescales in collaboration with the Commission for Basic Systems (CBS); and

2. Support the development of operational practices of RCOFs based on objective sub-seasonal and seasonal forecasts.

In accordance with the above-mentioned decision, participants at the WMO International Workshop on Global Review of RCOFs in September 2017 recommended that RCCs "access digital forecast and hindcast data from the WMO LC-LRFMME and produce an objectively consolidated forecast product combining information of various GPCs-LRF to be used as a first estimate for RCOF discussions".

This document follows up on the above points. The initial draft of this document was planned at the Second WMO Workshop on Operational Climate Prediction (OCP-2), held from 30 May to 1 June 2018 in Barcelona, Spain, and prepared by a team of ten subject matter experts under the leadership of the CBS/CCl Inter-Programme Expert Team on Operational Predictions from Sub-seasonal to Longer-time Scales (IPET-OPSLS), with the participation of the World Climate Research Programme (WCRP) Working Group on Sub-seasonal to Inter-decadal Prediction (WGSIP). Reviews of the initial draft document were provided by various stakeholders from GPCs-LRF, RCCs, RCOFs, NMHSs and other institutions with expertise in seasonal predictions.

HOW TO USE THIS DOCUMENT

Chapters 1–3 of this document provide essential background information concerning the scientific foundations for the practice of seasonal predictions: what makes seasonal predictions feasible; the sources of predictability; the infrastructure of dynamical seasonal forecasts; the basic formats of seasonal forecasts, and so forth. The main recommendations for developing and adopting objective seasonal forecast practices are discussed in Chapter 4, and individuals with advanced knowledge of the fundamentals and practice of seasonal predictions can consult this chapter directly. References to the basic concepts of seasonal predictions included in Chapters 1–3 are made throughout the document. Chapters 5–6 refer to WMO and other resources that can used for developing objective seasonal forecast practices. Chapter 8 provides some examples of good practices in the development and delivery of seasonal forecasts that currently exist and can serve as a template for implementing similar practices elsewhere. Chapters 7 and 9 discuss some ancillary activities in the context of seasonal forecasting and its future evolution.

CHAPTER 1. INTRODUCTION TO SEASONAL PREDICTIONS

1.1 Introduction

In the last forty years, facilitated by advances in observing systems and improvements in understanding and modelling of the various components of the Earth system and supported by enhancements in computing capabilities, steady advances in weather and climate prediction have taken place at major operational centres across the world (Bauer et al., 2015).

Complementing these advances in weather and climate prediction, there have been important milestones in advancing the science and operational infrastructure for predictions at longer timescales. The first generations of dynamical seasonal forecast systems were implemented at operational centres in the mid-1990s (Ji et al., 1994; Stockdale et al., 1998). Those initial systems have today grown into a robust infrastructure for operational seasonal predictions that has been formalized within WMO's Global Data-processing and Forecasting System (GDPFS) framework and which provides information on the expected evolution of seasonal climate conditions on a planetary scale. Routine weather and climate forecasts at the global and regional levels now provide information critical for the economic welfare of society and for mitigating losses to life and property.

The development of seasonal forecasts, and the associated infrastructure, fills a user need for predictive information on seasonal timescales which can be used for decision-making purposes. This need spans many sectors, including agriculture and food security, health, energy, water management, resource allocation, and disaster risk reduction; these sectors are also highlighted as priority areas under the GFCS. Advance information about climate conditions for the coming seasons allows users to minimize the risks associated with adverse climate conditions and maximize the benefits of favourable climate conditions.

The current operational infrastructure for seasonal forecasts within WMO, provided by GPCs-LRF, delivers coarse-grained spatial forecast information on a global scale. However, decision-making at the local level requires fine-grained spatial information. To address this discrepancy, WMO initiated the development of RCOFs and RCCs and the concept of the cascading forecast process: GPCs-LRF provide seasonal forecast information on a global scale, and RCCs and RCOFs tailor this information so that it can be used by decision-makers at regional, national and local levels. The purpose of this document is to present a set of principles, recommendations for the implementation of these principles, and good practices that should be followed to ensure the success of the cascading forecast process.

1.1.1 *Societal context of seasonal forecasts*

According to the State of the Climate in 2017 (Blunden, J. et al., 2018), since 1901, the mean annual global (land + ocean) surface air temperature has warmed by 0.7–0.9° Celsius per century, and the rate of warming has nearly doubled since 1975 to 1.5–1.8° Celsius per century. A steady rise in temperature has triggered important changes in the frequency and intensity of extreme weather and climate events such as heat and cold waves, droughts, floods, hurricanes, and so forth over various parts of the globe (Intergovernmental Panel on Climate Change, 2013). These unprecedented long-term climatic changes have also influenced sub-seasonal and seasonal-to-interannual variability and have had a profound impact on the natural environment as well as on the life, health and well-being of human society (Easterling et al., 2000; Meehl and Tebaldi, 2004; Coumou and Rahmstorf, 2012).

As the need has arisen to confront extreme variations in both weather and climate, society has slowly evolved to adapt to such variations and has developed indigenous knowledge acquired through several generations. For example, people in desert regions have changed their lifestyles to adapt to hot and dry conditions, and farmers choose crops suitable for the average local climate and its variability.

In recent years, it has been recognized that in order to sustain human activity in the face of seasonal to interannual variability, seasonal forecasting capabilities must be able to anticipate the development of associated climate anomalies. Seasonal forecasts offer society an opportunity to protect, or even enhance, societal welfare. Depending on the lead time and reliability of seasonal forecasts (see section 1.4), people can take a range of anticipatory actions both to minimize the risk of hazardous climate conditions and to capitalize on climatic opportunities (favourable climatic conditions). For example, in the case of a forecast for a drought, a farmer might decide to plant drought resistant crops or perhaps not to plant at all, and a reservoir water manager can exercise appropriate control on releases of water from the reservoir and plan for the judicious use of water resources.

For seasonal forecasts to be beneficial to society, it is essential to identify the sectors that will potentially use these forecasts and to respond to their needs (Harrison et al., 2008). The GFCS, a United Nations-wide initiative in which WMO Members and inter- and non-governmental, regional, national and local stakeholders work in partnership to develop targeted climate services, has identified five climate-sensitive investment areas supporting both climate adaptation and climate mitigation. Potential uses of seasonal forecasts to minimize risks and maximize benefits in these five sectors are briefly outlined below.

Agriculture and food security: Agriculture and food security in the twenty-first century faces multiple challenges, and climate variability and change is expected to affect all the components that influence food security, including availability, access, stability and utilization. It is now widely recognized that seasonal forecasts can be used to support planning activities and decision-making in agricultural systems. Skilful seasonal forecasts can reduce the uncertainty of future climatic variability, potentially allowing farmers to reduce losses from poor climatic conditions and increase gains from climatic conditions that are beneficial.

Disaster risk management: Disaster risk management (DRM) is the systematic process of using administrative directives, organizations, and operational skills and capacities to implement strategies and policies and improve coping capacities in order to lessen the adverse impacts of climatic hazards and the possibility that they may instigate disasters. Experience shows that there are certain patterns in the occurrence of these events, which provides some early warning indication of the hazards. However, under a changing climate, relying on experience is no longer enough. Seasonal forecast information can be critical for making proactive DRM decisions to cope with the impact of climate variability.

Energy: The impact of climate variability on the energy sector is complex and multidimensional. Climate variability has a profound influence on most forms of energy production (for example, wind and hydropower) and the level of energy consumption. Seasonal changes in temperature, precipitation and the frequency and severity of extreme events affect how much energy needs to be produced, delivered, and consumed. Seasonal climate forecasts can help energy companies to better anticipate spikes in energy demand ahead of time to match the requirements.

Health: Human health is sensitive to several types of climatic variations. Seasonal forecasts can give public health services an unprecedented degree of early warning for the likelihood of epidemics, based on climatic or ecological analysis, before disease organisms appear. Although taking advantage of seasonal forecasts would require sufficient knowledge to link climate parameters to ecological events affecting disease organisms, the availability of skilful seasonal forecasts will provide a catalyst to accelerate the integration of such knowledge.

Water: Water resources are directly dependent on the abundance of precipitation and snowmelt and how we store and use the amount of available water. The potential impact of climate variability with respect to water availability includes changes in precipitation amount, intensity, timing and form (rain or snow), changes in snowmelt timing and changes to evapotranspiration. Seasonal forecasting applications can help assess the influence of such climatic impacts on the decision-making process.

1.1.2 *Historical evolution of seasonal forecasts*

Attempts to provide seasonal forecasting of the summer monsoon rainfall over India started following a devastating famine during the late 1870s. The first operational seasonal forecast of Indian summer monsoon rainfall for the region covering all of India and Burma was issued on 4 June 1886 using an empirical method by Henry Francis Blanford, the first Head of the India Meteorological Department (IMD), established in 1875. This forecast was based on the assumption that the varying extent and thickness of Himalayan snow influences the climate conditions over the plains of northwest India (Blanford, 1884). Blanford used this relationship to prepare experimental seasonal forecasts from 1882 to 1885 before attempting the first ever operational forecast in 1886. After that milestone, the seasonal forecasting of monsoon rainfall became an operational responsibility of IMD.

Sir John Eliot, who succeeded Blanford as the Head of IMD in 1895, applied methods such as analogue year analysis for the seasonal forecasting of Indian Summer Monsoon Rainfall (ISMR). Efforts to improve forecasts for ISMR continued during the period of Sir Gilbert T. Walker (1904–1924), who took over as the Director General of IMD. Realizing the complexities of the forecasting problem, Walker started systematic studies to develop objective techniques for seasonal forecasting (Walker, 1908). Walker (1910, 1914, 1923) also conducted extensive studies of worldwide variations of weather parameters such as rainfall, temperature, pressure, and so forth. The search for candidate predictors led Walker to identify three large-scale see-saw variations in global pressure patterns: the North Atlantic Oscillation (NAO), the North Pacific Oscillation (NPO) and the Southern Oscillation (SO). Walker also introduced the concept of correlation and regression for the first time in seasonal forecasting in order to remove the subjectivity of earlier techniques. The first official objective forecast was issued in 1909 for the seasonal monsoon precipitation over all of India based on the regression technique.

Walker (1910) did pioneering work on the SO and published a prediction technique for India monsoon forecasting containing 22 predictors in six forecast formulae for forecasting precipitation during the whole season (June to September) and for forecasting precipitation over three homogeneous rainfall regions of India during the second half of the season (August–September) (Banerjee, 1950). The SO, a see-saw between sea-level pressure at Tahiti and Darwin, was later linked to the unusual warming of surface waters in the eastern tropical Pacific Ocean, or El Niño, by Jacob Bjerknes in the 1960s (Bjerknes, 1966, 1969). Bjerknes, and subsequently others, defined these linked ocean–atmosphere phenomena as the El Niño-Southern Oscillation (ENSO).

Following the 1972/1973 El Niño and 1973/1974 La Niña events, it was noted that opposite phases of ENSO have significant, and generally opposite impacts on temperature and precipitation patterns across the globe and that these impacts are more pronounced in the tropical regions. However, the 1982/1983 El Niño event and its associated regional climate anomalies (Ropelewski and Halpert, 1987; Rasmusson and Carpenter, 1983) resulted in the recognition that the coupled ocean–atmosphere ENSO phenomenon is the dominant mode of Earth's interannual climate variability (Goddard et al., 2001). These observations were also supported by several theoretical studies (Charney and Shukla, 1981; Palmer and Anderson, 1994; Barnston et al., 1999), which suggested that predictive skill on seasonal timescales is linked with the slowly evolving boundary conditions of the climate system, such as sea-surface temperature, snow cover, soil moisture, sea ice, and so forth. The identification of statistically significant ENSO-associated global climate teleconnections has resulted in ENSO-related parameters, along with other slowly varying climate drivers, being used as predictors in empirical/statistical forecast models for large-scale surface temperature and precipitation anomalies over many countries across the globe (Barnston, 1994; Drosdowsky and Chambers, 2001; Rajeevan et al., 2007; Pai et al., 2017). Statistical models have also provided a benchmark to assess the skill of the now commonly used state-of-the-art dynamical global general circulation models (GCMs) for seasonal climate forecasting.

The first steps towards seasonal climate forecasting based on dynamical models were taken by Norman Phillips in 1956 when he developed a mathematical model to simulate monthly and seasonal tropospheric circulation patterns. Following this, there were targeted efforts by various research groups to further develop general circulation models. The prediction of ENSO variability with a simple coupled ocean–atmosphere dynamical model was first demonstrated by Zebiak and Cane (1987). Subsequent work supported by the establishment of a ten-year (1985–1994) international Tropical Oceans and Global Atmosphere (TOGA) project and the implementation of the ocean observing system in the equatorial Pacific led to the development of today's sophisticated operational ENSO forecast systems (JGR-Oceans, 1998; McPhaden et al., 2010). There has been noticeable improvement in the predictive skill of dynamical models during the last few decades, mainly due to advances in estimating initial ocean and atmospheric conditions as well as advances in the model physics and computing capabilities (Bauer et al., 2015).

The 1997/1998 El Niño event, the strongest El Niño event of the twentieth century, triggered an increased interest in and demand for climate services. The 1997/1998 El Niño provided a dramatic example of the effects of relatively short-term climatic variations on society and the potential value of forecasting them. The period 1997–1998 also coincided with the successful campaign of ocean observations from the Tropical Atmosphere Ocean (TAO) moored array across the equatorial Pacific and the development of several ocean prediction models using TAO data. The positive response of RCOFs, which were first established in 1996 at a meeting held in Victoria Falls, Zimbabwe, to the 1997/1998 El Niño event provided a boost for further developing the concept of RCOFs. The 1997/1998 El Niño, however, did not reproduce the classic patterns of global climate anomalies, such as those seen in the 1982/1983 event (the second strongest event of twentieth century). This discrepancy called into question the notion that different El Niño events may lead to similar impacts on global climate. These observations provided support to the view that the impact of sea-surface temperature (SST) on daily weather is not deterministic as even a small uncertainty in the initial condition can lead to uncertainty in the seasonal forecast beyond a period of one week or so (Lorenz, 1969). The need to communicate the uncertainty in the forecast resulted in the introduction of the concept of probabilistic and ensemble approaches to weather and seasonal forecasting.

At present, operational prediction centres use state-of-the-art ocean–atmosphere coupled GCMs to generate seasonal forecasts. Though atmospheric GCMs (AGCMs) were the first to appear on the scene, coupled GCMs (CGCMs) have gained preference over AGCMs because they are expected to better represent the interactions between the different components of the climate system (atmosphere, ocean, cryosphere, and so forth).

Since 2006, as a part of the development of an infrastructure for its GDPFS, WMO has designated prediction centres with the mandatory responsibility of generating and delivering seasonal forecasts (with global coverage), including associated verification information. There are currently 13 of these so-called Global Producing Centres for Long-Range Forecasts. In 2009, WMO formally endorsed the Lead Centre for Long-Range Forecast Multi-Model Ensemble hosted jointly by the Korea Meteorological Agency (KMA) and the National Centers for Environmental Prediction (NCEP). This has facilitated worldwide access to data from multi-model ensemble-based seasonal forecasts.

To provide seasonal forecast information at the regional domain, WMO established RCCs and RCC Networks covering WMO Regional Associations and the polar regions. In 2009, Beijing and Tokyo were the first two RCCs formally designated by WMO. In addition to the WMO infrastructure for long-range forecasts (LRFs), a consortium of research and operational centres in North America established an experimental multi-model seasonal forecasting system called the North American Multi-model Ensemble (NMME). This consortium consists of coupled models from United States and Canadian modelling centres. In addition, non-governmental centres, such as the Asia-Pacific Economic Cooperation Climate Centre (APCC) in South Korea and the International Research Institute for Climate and Society (IRI) in the USA, also prepare multi-model ensemble-based seasonal forecasts (see Chapters 5 and 6). In Europe, a global multi-model seasonal forecast system called EUROSIP was put in place in April 2005. This system was replaced by the Copernicus Climate Change Service (section 6.2) in October 2019.

1.1.3 *Elements of climate variability*

Variability in the different components of the Earth system, for example, the ocean and the atmosphere, occurs on all timescales and on different spatial scales. Temporal variations occur on scales of seconds (atmospheric turbulence) to centuries and are organized on spatial scales of centimetres to atmospheric planetary waves of over thousands of kilometres (Williams et al., 2017). On weather timescales, cyclones in the tropics and synoptic-scale systems in the extratropics dominate atmospheric variability. The spatial structure of variability on weekly timescales can be discerned as modes such as the Pacific North American (PNA) pattern (Wallace and Guztler, 1981) or the North Atlantic Oscillation/Arctic Oscillation (Barnston and Livezey, 1987; Thompson and Wallace, 1998). The scale interaction between the modes of atmospheric variability, for example, the PNA pattern and synoptic variability, alters the track of extratropical synoptic storms, which in turn helps maintain the former (Pinto et al., 2011). It is generally recognized that atmospheric variability on weather timescales cannot be predicted beyond a lead time of 15 days (Lorenz, 1969; Hoskins, 2013).

In the context of LRF and the potential for making skilful long-range predictions, it is the modes of variability on sub-seasonal and seasonal timescales that play a critical role. Also, as the temporal scale of variability increases, the preferred modes of variability involve interactions across multiple components of the Earth system, for example, ocean, land and atmosphere. On sub-seasonal timescales and global spatial scales, the dominant mode of atmospheric variability is the Madden-Julian Oscillation (MJO), which propagates along the equatorial zone on 30-day to 60-day timescales (Madden and Julian, 1971; Zhang, 2005). The presence of MJO modulates atmospheric variability over different parts of the globe including the frequency of hurricanes in the tropical Atlantic and precipitation over the western United States, South America, Africa, the Maritime Continent and other locations (for example, Zhou et al., 2012).

On seasonal to interannual timescales, the most important mode of variability is ENSO and involves coupled air–sea interaction in the equatorial tropical Pacific. Local influences of ENSO include variations in SSTs in the central and eastern equatorial Pacific. Variations in SSTs, in turn, modulate the location of deep convection over adjacent areas. Changes in the upper level outflow result in changes in tropical upper level circulation which, via meridionally-propagating Rossby waves, alter the atmospheric variability downstream in extratropical regions (Horel and Wallace, 1981; Sardeshmukh and Hoskins, 1988; Trenberth et al., 1998). Remote influences of ENSO, together with ocean thermal inertia, that is, the fact that SSTs vary on longer timescales, provide the potential for skilful seasonal predictions. Indeed, such linkages were the impetus for the initiation of the TOGA programme and the enhancement of the ocean observing system in the equatorial tropical Pacific followed by the initial development of operational coupled seasonal prediction systems in the mid-1990s (JGR-Oceans, 1998; McPhaden et al., 2010).

Beyond sub-seasonal and seasonal timescales, there are also coherent variations on decadal and longer timescales. Examples include the Pacific Decadal Oscillation (PDO) (Newman et al., 2016; Liu, 2018) and the Atlantic Multidecadal Oscillation (AMO) (Grossman and Klotzbach, 2009). Variations on decadal timescales are thought to be due to the natural variability caused by ocean–atmosphere coupled interactions. Variations on timescales longer than decadal can be caused by slow variations in external forcings. These external forcings may be natural in their origin, for example, changes in solar insolation related to the sunspot cycle, or related to anthropogenic causes, for example, an increase in CO_2 and the associated increases in global temperature, or they may be both natural and anthropogenic.

Although weather cannot be predicted deterministically beyond two weeks, the statistics of atmospheric variability, for example, the preferred location of storms in the extratropics and the seasonal frequency of hurricanes and cyclones in the tropics, can be modulated by the variability in the Earth system that occurs on longer timescales. Changes in the statistics of atmospheric variability due to slowly varying components of the Earth system are the underlying basis for skilful seasonal forecasts.

1.2 Scientific basis for seasonal forecasting

In this section, the scientific basis for making skilful seasonal forecasts is discussed further.

The basic elements behind skilful seasonal forecasts over certain parts of the globe come from the local and remote atmospheric and terrestrial influence of longer timescale variations in different components of the climate system (section 1.2.1). The stratosphere and mesosphere provide lagged information on the variability of the troposphere via upward and downward coupling and vertically propagating waves. The land surface provides delayed feedback to the atmosphere due to soil moisture memory. Sea-ice variations change radiation properties and flux exchange between the ocean and the atmosphere, and more generally, ocean–atmosphere fluxes are modulated as a consequence of the large thermal inertia and relatively long dynamical timescales of the oceans.

The remote atmospheric influence of anomalies in boundary conditions is communicated by so-called teleconnections (section 1.2.2). The potential for skilful seasonal predictions also comes from our ability to recognize the slowly evolving trends in surface temperature and precipitation (section 1.2.3) that could be due to either the changes in external forcings or natural causes, for example, decadal variations.

1.2.1 *Role of slowly varying boundary conditions in modulating seasonal atmospheric variability*

Variations in the different components of the Earth system occur according to timescales that are determined by typical wave propagation speeds and thermal inertia. Variations in the atmosphere typically have the shortest timescales, while variations in the ocean surface and land surface (having a larger thermal inertia) occur on a longer timescale. The typical timescale of variations can be characterized by the persistence of anomalies and is generally quantified in terms of their autocorrelation, that is, the timescale over which initial anomalies, if persisted, are a useful forecast for future conditions.

Typical persistence timescales of variations in different components of the Earth system are useful in assessing the predictability and baseline skill attainable for seasonal forecasts. In general, skill assessments based on persistence are improved upon by empirical (statistical) or dynamical prediction methods. It should be noted that for a specific variable, the persistence timescale has a geographical and seasonal dependence. These dependences are associated with regional characteristics of circulation features and interactions across different components of the Earth system. An example of the persistence timescale for monthly mean SST anomalies (quantified in terms of the magnitude of the lag autocorrelation) and its spatial dependence is shown in Figure 1.1.

As indicated earlier, changes in the slowly varying components of the Earth system can modulate the characteristics of variability that occur on shorter timescales, and this modulation could be

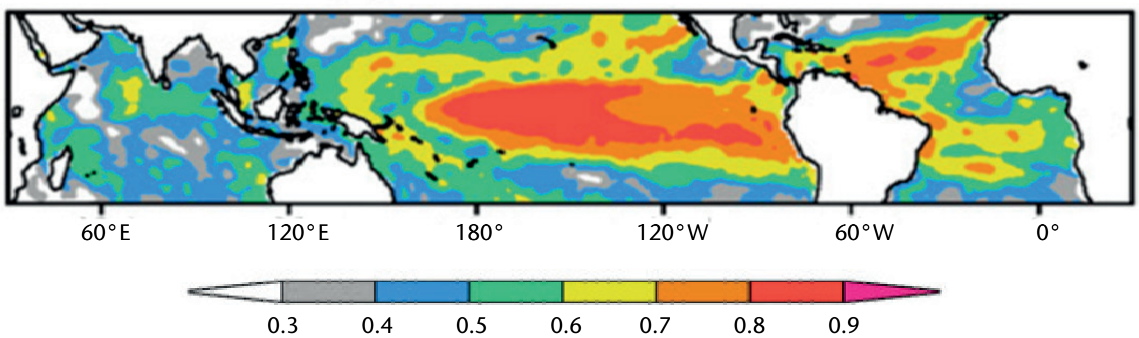

Figure 1.1. Autocorrelation of monthly mean SST variability with a two-month lead. The autocorrelation magnitude relates to the characteristic persistence timescale of observed SST anomalies. Autocorrelation provides a baseline estimate of predictability and skill for long-range prediction and has distinct geographical variations. The autocorrelation timescale for boundary conditions (for example, SST) also provides relevant information about potential seasonal predictability sources for atmospheric and terrestrial variables.

Source: Kumar and Zhu, 2018

due to either local or remote causes. One example of local influence would be the following: drier than normal soil moisture conditions will decrease the proportion of solar heating normally used in evaporation and increase the proportion available for the (sensible) heating of the surface, leading to warmer than normal surface air temperature. Conversely, wetter than normal soil moisture conditions lead to cooler surface temperatures. Snow cover and sea-ice anomalies can also modulate variability in surface temperature in the overlying atmosphere by changing surface albedo, that is, by partitioning the surface insolation into absorbed and reflected components.

Anomalies in SSTs associated with oceanic variability play an important role in modulating shorter timescales for atmospheric variability, and the influence of these anomalies often extends over thousands of kilometres. One important example of this is the SST variations associated with ENSO that have a well-known influence over remote regions (Figure 1.2). In addition, changes in tropical convection associated with ENSO SSTs lead to changes in tropical precipitation. For example, increased SSTs during El Niño in the equatorial eastern and central Pacific, caused by alterations of the Walker circulation, lead to drying over the Maritime Continent. Changes in precipitation brought about by warm SSTs during El Niño also lead to changes in atmospheric and terrestrial variability in other parts of the globe, for example, suppressed hurricane activity in the equatorial Atlantic, above-normal precipitation over the southern United States, and a decrease in Indian summer monsoon precipitation. See Figure 1.2 for other worldwide regions where precipitation tends to be modulated by El Niño.

The modulation of shorter timescale atmospheric and terrestrial variability because of slowly-varying boundary conditions provides the fundamental basis for our ability to make skilful seasonal forecasts. It is also important to note that the local and remote influence of boundary conditions has a strong regional and seasonal dependence; the potential to make skilful seasonal forecasts therefore varies from region to region and according to the time of the year.

1.2.2 *Teleconnections and key drivers of seasonal to interannual climate variability*

Teleconnections are recurrent large-scale anomaly patterns linking climate variability across remote regions. They are the building blocks for our ability to predict climate variability. In the context of seasonal forecasts at the regional level, it is important to understand and quantify which teleconnections influence local climate variability.

The see-saw pattern in surface pressure between Tahiti and Darwin discovered by Sir Gilbert Walker in the nineteenth century, called the Southern Oscillation, is perhaps the earliest discovered example of teleconnection. Today, we know that ENSO, an ocean–atmosphere coupled mode (Philander, 1990) in the tropical Pacific that evolves through positive feedback between SST and trade winds, is the most emblematic example of teleconnection.

ENSO influences the seasonal distribution of temperatures and precipitation over several regions and depends on the season (see, for example, Davey et al., 2014; Trenberth et al., 2018). It is noteworthy that documented influences associated with ENSO represent typical (not guaranteed) responses. For example, it is often noted that no two El Niños are identical in the weather/climate responses they generate around the globe. The stronger the El Niño, the more likely that the influence of teleconnection will manifest. The most reliable effects of El Niño are deficient precipitation over Indonesia and northern South America and excess precipitation in south-eastern South America, eastern equatorial Africa, and the southern United States.

The impact of ENSO on regions outside the tropics can be explained by Rossby wave trains propagating into the extratropics (Hoskins and Karoly, 1981; Trenberth et al., 1998). However, the propagation of these Rossby wave trains is not always identical (due to changes in the mean state), and their diagnosis is crucial for seasonal forecasts in the mid-latitude regions. The influence of ENSO over the extratropics is discussed further at the end of this section.

El Niño and Rainfall

El Niño conditions in the tropical Pacific are known to shift rainfall patterns in many different parts of the world. Although they vary somewhat from one El Niño to the next, the strongest shifts remain fairly consistent in the regions and seasons shown on the map below.

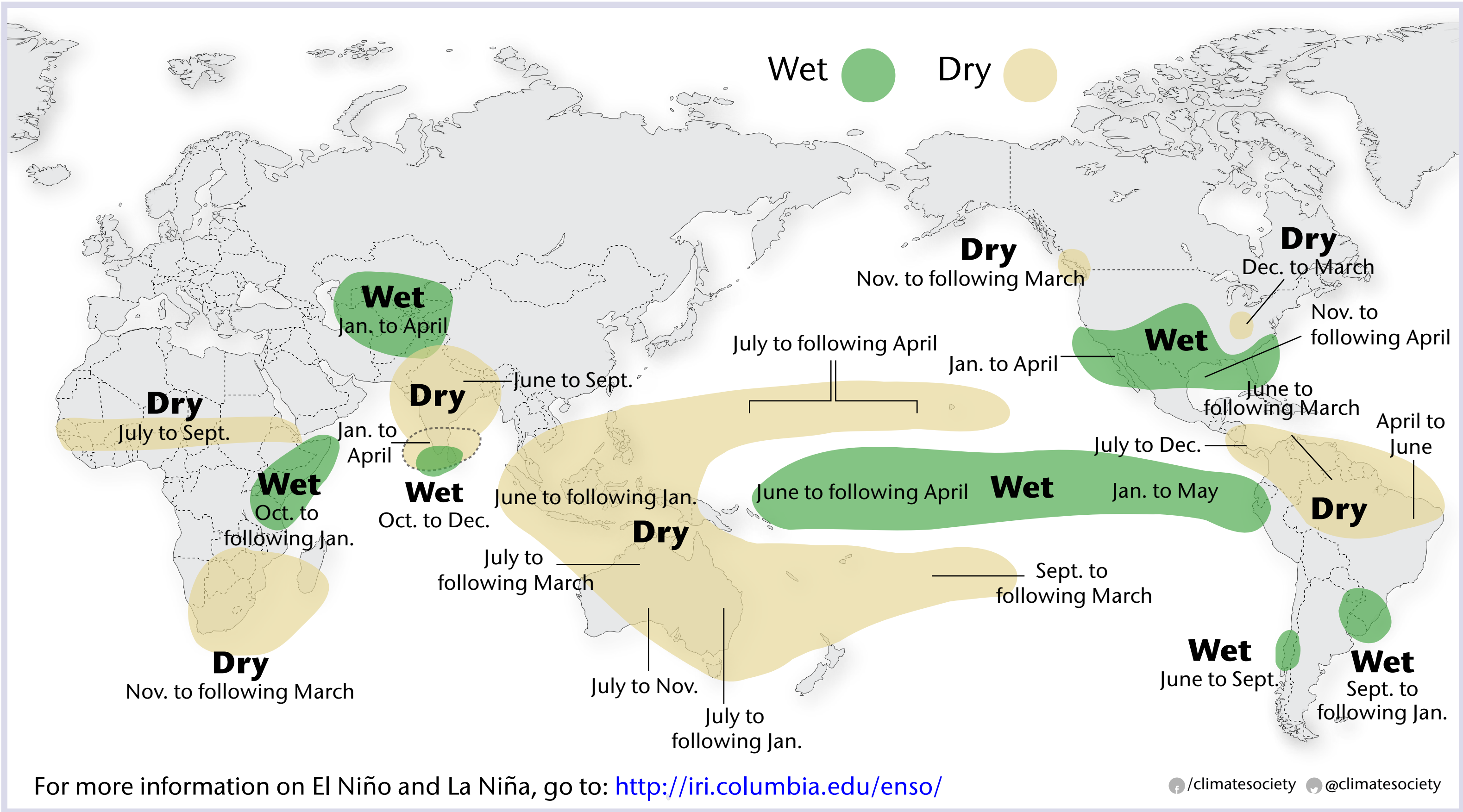

Figure 1.2. One example of teleconnection: the average global influence of El Niño events on precipitation inferred from historical observational data using the compositing technique

Sources: Ropelewski and Halpert, 1987; Mason and Goddard, 2001

ENSO atmospheric teleconnections are generally derived from a relatively large observed sample of El Niño and La Niña events with varying SST characteristics (for example, the eastern Pacific El Niño, when the largest SST warming is in the eastern Pacific Ocean, and the central Pacific El Niño Modoki, when the largest SST anomalies are located in the central Pacific Ocean). "Average" ENSO teleconnections do not always represent in detail the remote impacts in temperature and precipitation associated with specific ENSO events. A number of studies since the mid-2000s (McPhaden, 2004; Kumar et al., 2006; Barnard et al., 2011; Preethi et al., 2015) have focused on eastern Pacific and central Pacific El Niño impacts, but due to the limited number of El Niño events in the record, there are large uncertainties concerning the precise nature of regional ENSO precipitation and temperature teleconnections between different types of events. Since the ocean mean state is also changing due to a combination of natural variability and anthropogenic forcing, we cannot assume that ENSO atmospheric teleconnections in terms of precipitation and temperature over various parts of the globe are stationary (Yeh et al., 2018). The length of the current observational record is not long enough to clearly distinguish which changes in ENSO teleconnections are robust and which are an artefact of sampling variability, although reanalysis datasets can, to some extent, alleviate the problem of an insufficient observational record. Faced with such uncertainties in discerning higher order influences in atmospheric teleconnections related to ENSO, the average ENSO teleconnection information is that which is most widely used.

Another important tropical mode of variability, largely correlated with ENSO, is the Indian Ocean Dipole (IOD). IOD is defined by the SST differences over the western Indian Ocean and the eastern Indian Ocean south of Indonesia. IOD affects the climate of countries that surround the Indian Ocean Basin and is a significant contributor to precipitation variability in this region. Like ENSO, the change in temperature gradients across the Indian Ocean results in changes in the preferred regions of rising and descending air. Understanding these changes is a goal of seasonal prediction efforts (for example, see the forecasts at the UK Met Office).

A positive IOD (positive anomalies in the western Indian Ocean exceed those in the eastern Indian Ocean) is typically associated with easterly low-level wind anomalies across the Indian Ocean, enhanced precipitation over East Africa and reduced precipitation over Indonesia and southern Australia, as well as less cloudiness in northwest Australia. During a negative IOD (cooler SST in the western Indian Ocean), winds become more westerly, typically bringing decreased precipitation to East Africa and increased precipitation to Indonesia, as well as increased cloudiness to northwest Australia and more precipitation over southern Australia. IOD also impacts several regions far from the Indian Ocean. Significant correlation is found throughout Europe, northeast Asia, North and South America, Australia and South Africa concurrent with IOD events (Hameed and Yamagata, 2003; Molteni et al., 2015).

IOD variability is often linked with ENSO phases. Positive IOD events are often associated with warm ENSO (El Niño) and negative events with cold ENSO (La Niña). When IOD and ENSO are in phase, the impacts of El Niño and La Niña events over the regions influenced are generally more extreme; if they are out of phase, the impacts of El Niño and La Niña events can be diminished.

In the northern hemisphere extratropics, NAO and the PNA pattern are the most prominent teleconnection patterns (Wallace and Guztler, 1981) (Figure 1.3). Although NAO and the PNA pattern are associated with internal atmospheric variability, they can be modulated by tropical forcings such as ENSO and tropical organized convection associated with MJO (Madden and Julian, 1971). Long-range predictability of atmospheric anomalies over the extratropics is largely based on the PNA pattern and NAO sensitivity to tropical forcings (Trenberth et al., 1998; Scaife et al., 2014). NAO variability is also linked to stratospheric variability (Baldwin and Dunkerton, 2001, Scaife et al., 2005)

NAO is associated with fluctuations in the strength of the climatological mean jet stream over the western Atlantic Ocean. The positive phase of NAO exhibits below-normal geopotential heights across the high latitudes of the North Atlantic and above-normal heights over the central North Atlantic Ocean, the eastern United States and western Europe. The negative phase has an opposite pattern of height anomalies over these regions. NAO phases modulate temperature and precipitation patterns often extending from eastern North America to western and central Europe. The NAO life cycle is typically around 10 days. In some years, such as 2009/2010, it

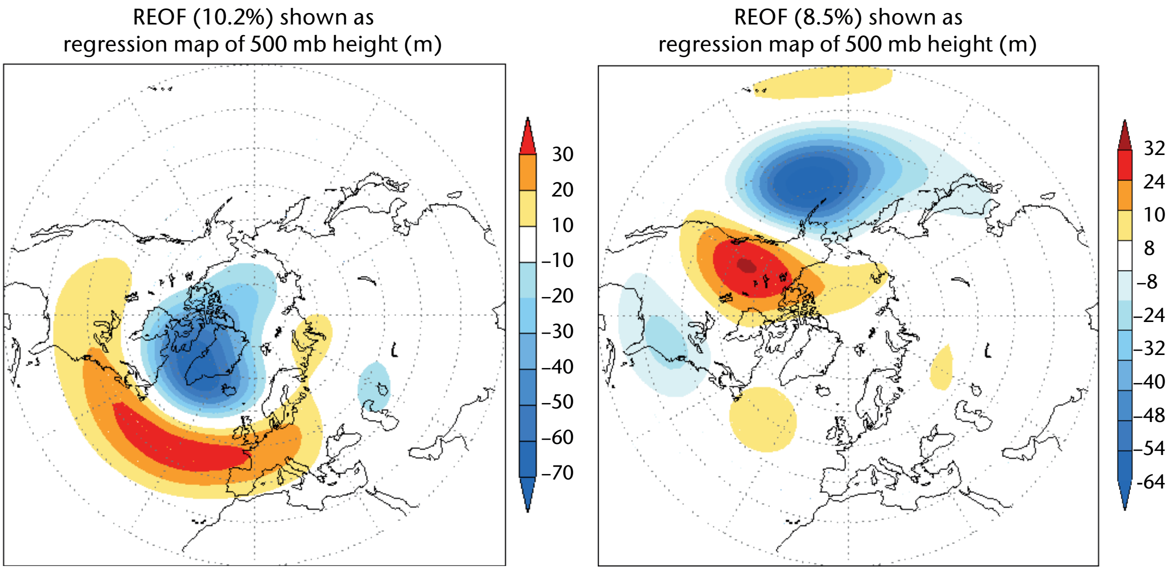

Figure 1.3. NAO and the PNA pattern defined as the first leading mode of Rotated Empirical Orthogonal Function analysis of monthly mean 500 mb height during the 1950–2000 period

Source: Climate Prediction Center

dominated the boreal winter circulation, bringing severe cold conditions over northern Europe and wet mild conditions over southern Europe. Scaife et al. (2014) indicated that the predictive skill for NAO has reached useful levels and suggested that most of the skill in predicting winter European surface temperature anomalies stems from the predictability of NAO.

The PNA pattern is associated with strong fluctuations in the strength and location of the East Asian jet stream. The positive phase of the PNA pattern features above-average heights in the vicinity of Hawaii and over the intermountain region of North America and below-average heights located south of the Aleutian Islands and over the south-eastern United States. The PNA pattern modulates temperature and precipitation patterns over the Pacific sector. PNA pattern variability is strongly influenced by ENSO. The positive phase of the PNA pattern tends to be associated with El Niño, and the negative phase tends to be associated with La Niña. Both the PNA pattern and NAO tend to have their maximum impact on surface temperature variability during the boreal winter; this is related to the seasonality of tropical-extratropical teleconnection patterns. Similarly, in the southern hemisphere, the Southern Annular Mode (SAM) can be a source of predictability (Lim et al., 2013; Hendon et al., 2014).

Although the spatial features of ENSO teleconnections are affected by uncertainties associated with the limitation of observational data, their use in assessing the fidelity of model simulated teleconnections is useful for forecast interpretation and model validation and can serve as a basis for model selection for developing seasonal forecasts over a region.

1.2.3 *Role of low-frequency trends in seasonal forecasts*

Low-frequency climate signals, for example, trends, contribute in different ways to both the variability and the predictive skill of temperature and precipitation, as well as to other climate variables. As forecasts for seasonal anomalies are made relative to a climatological base period, slow trends in variables, for example, warming trends in surface temperature, can contribute to forecast seasonal mean anomalies, and when verified against observations, can add skill. The contribution of trends, for example, those related to anthropogenic global warming, to seasonal variability also depends on the geographical location and the season considered.

A timescale decomposition methodology developed to help assess the contribution of these long-term global warming trends (Greene et al., 2011) shows that the impact on surface temperature (Figure 1.4, upper panel) tends to be more important than the impact on precipitation (Figure 1.4, bottom panel). For example, although most of the Caribbean's surface temperature variability is explained by anthropogenic climate change, the contribution to seasonal precipitation variability is only about 10%.

Volcanic aerosols influence the Earth's energy balance by scattering solar radiation, both by absorbing and by scattering longwave radiation. The overall result is a decrease in global mean surface temperatures and warming of the lower stratosphere (Colose et al., 2016; Pausata et al., 2015). Some trends that are driven by external forcing (such as increases in CO_2) are generally present in dynamical seasonal forecasts and the associated hindcasts (Doblas-Reyes et al., 2006; Cai et al., 2009).

It is common practice to remove any long-term trends from the variables of interest when building statistical prediction models to avoid spurious relationships (Mason and Baddour, 2008). Nonetheless, trends should not be removed (a) if they are present in only the predictor or the predictand, but not in both at the same time, and (b) if there are reasons to expect that trends in the predictands are caused by trends in the predictors. For details, see Mason and Baddour (2008).

Finally, low-frequency trends and low-frequency modes of variability can also modulate predictive skill at the seasonal scale via the impact on different sources of predictability. Thus,

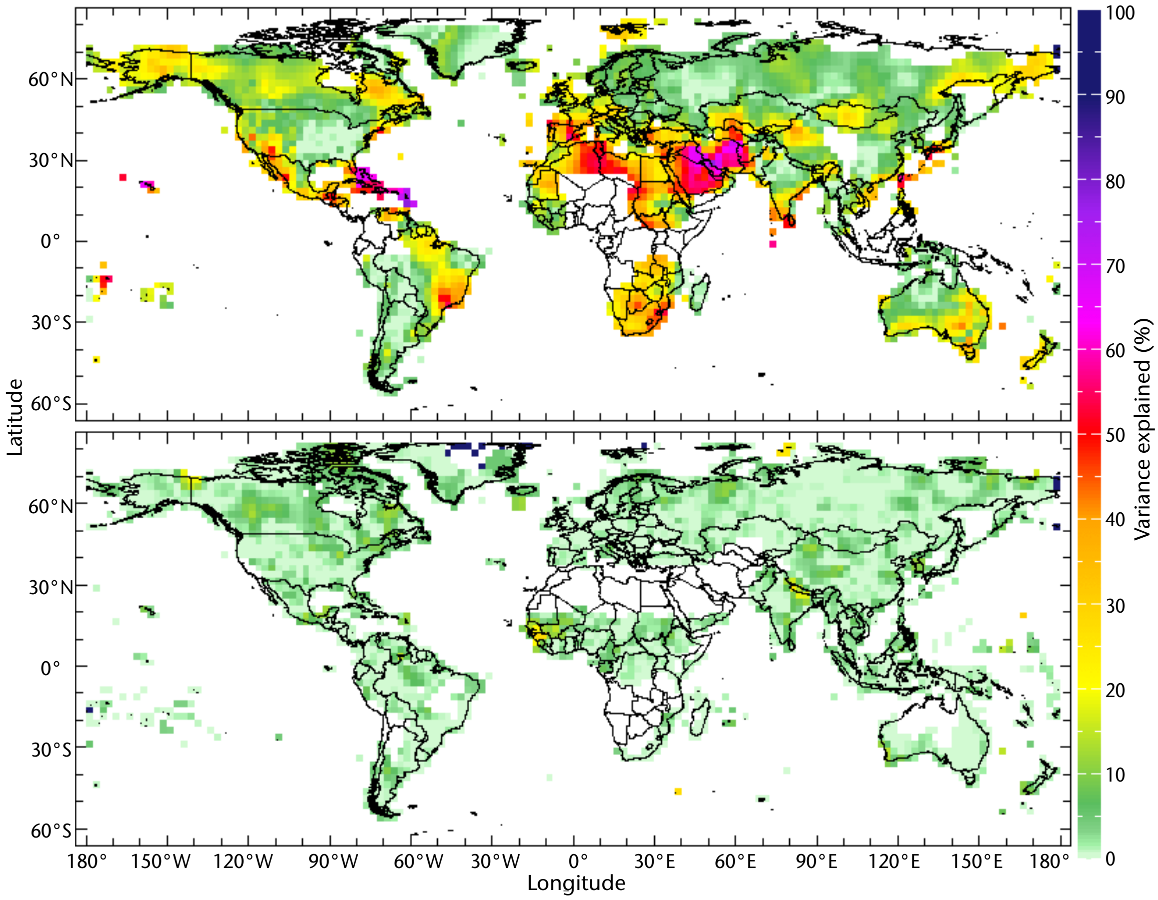

Figure 1.4. Proportion of the total explained variance (in %) associated with anthropogenic long-term trends in surface temperature (upper panel) and precipitation (bottom panel), for the June–August mean. Computed using the IRI Timescale Decomposition Maproom

Source: Greene et al., 2011

understanding these influences is important when assessing seasonal predictive skill for a particular region. Statistical forecast tools such as optimal climate normal (OCN) (Huang et al., 1996) have been used to increase the skill of seasonal forecasts.

1.3 Predictability and prediction skill

Although the influences of slowly varying boundary conditions and external forcings make seasonal forecasts feasible, there are limits on forecast skill, and there is considerable spatial and seasonal dependence with respect to our ability to make seasonal forecasts. The reason for limited prediction skill is due to the inherent limits regarding the predictability of the climate system (due to its non-linear nature) and errors in models that are used to make the forecasts.

The imperfect model formulation and the imperfect knowledge of the initial conditions that are used to initialize dynamical forecasts (see section 1.5.2), together with the non-linear growth of small differences in initial conditions, make predictions of weather and climate intrinsically uncertain. The inherent uncertainty of climate forecasts poses limits on their predictability. To represent and to estimate the uncertainty in the future state of the climate, an ensemble (or set) of seasonal forecasts based on dynamical models is used. Each individual forecast (often called a “realization” or “member”) in the ensemble differs, for example, by small initial perturbations which grow with time due to unpredictable variability or “chaos” (Lorenz, 1969). Other ways to generate the members in the ensemble also exist, for example, by using physical or stochastic perturbations. The importance of an ensemble of forecasts stems from the fact that individual forecasts in the ensemble represent possible outcomes for observations at a future time. Additionally, the forecast spread among the ensemble members can be used to quantify the uncertainty in seasonal prediction.

The question of whether a certain event is predictable depends on how different the ensemble forecast distribution during that event is from the corresponding climatological distribution. Figure 1.5 illustrates the concept of predictability and shows the ensemble forecast distribution and the corresponding climatological distribution for four different forecast lead times (see section 1.4). The shorter the lead time, the more distinguishable is the predicted distribution from the climatological distribution. As the forecast lead time increases and the initial perturbations grow, the distribution of the forecast ensemble slowly evolves towards the climatological distribution and the predictability degrades, and ultimately is lost.

For a well-constructed ensemble prediction system, the ensemble spread is an estimate of the forecast uncertainty, and the rate at which the error grows (spreads) can be used as a predictability measure. When the ensemble spread grows very rapidly, the estimated forecast uncertainty also increases rapidly, and predictability is lower. Conversely, when the ensemble spread grows slowly, the estimated forecast uncertainty increases slowly with the lead time, and the predictability is higher. Predictability is an *a priori* estimate of our ability to make skilful forecasts and quantifies an inherent property of nature; it does not depend on the quality of the forecasting system. Predictability is not a forecast skill measure, but rather the upper limit of predictive skill that a good forecast system can achieve. While predictability is an inherent property of nature, prediction skill (section 2.2.1) is the realization of predictability based on a forecast tool one can use, which could be based on either empirical methods or dynamical models (section 1.5).

Because the rate of error growth depends on the spatial and temporal scale characteristics of the motion involved, the time horizon for predictability varies depending on the spatial-temporal scale of individual weather events. Features with a horizontal scale of approximately 3 000 km, such as a large-scale extratropical cyclone evolving over several days, can be predicted with a useful level of skill out to six days ahead, while smaller-scale features, such as mesoscale convective systems developing within hours, can only be skilfully predicted one day in advance. Some dominant modes of the climate system on global spatial scales with lower frequency variations, such as MJO, ENSO and the Quasi-Biennial Oscillation, are predictable at monthly, seasonal or even longer timescales. These sources of predictable variability, via remote teleconnections, also impart predictability to mid-latitude seasonal mean atmospheric anomalies and to remote regions in tropical latitudes.

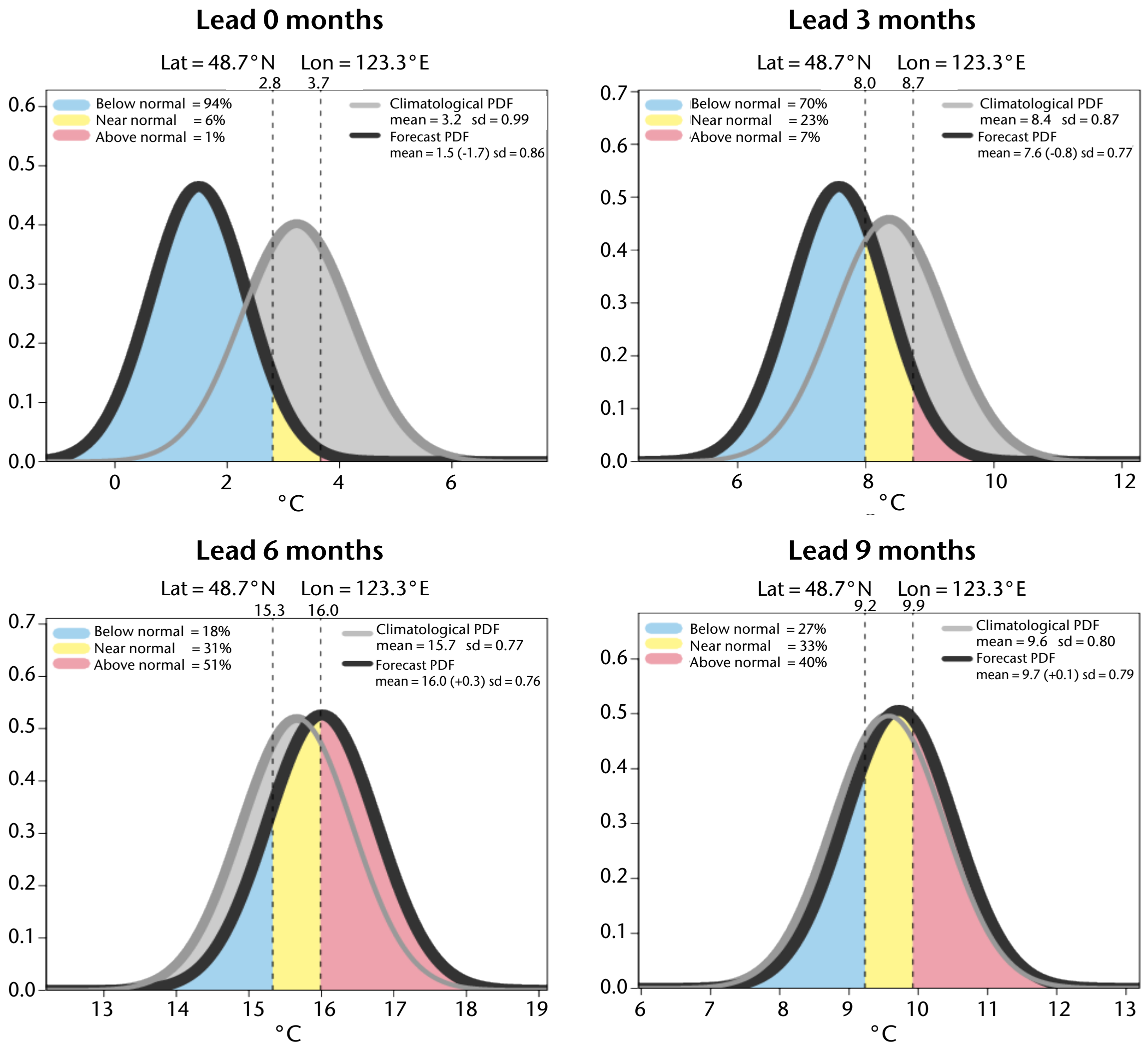

Figure 1.5. Predicted (black) and climatological (grey) probability density functions for a single grid point at different lead times. The difference between the two determines predictability. After George and Sutton (2006)

Observed atmospheric variability can be separated into two components. One is unpredictable: noise resulting from the chaotic nature of the atmosphere; the other is potentially predictable because it is constrained by predictable factors, such as MJO, ENSO, and so forth. Because of the limitations on the observational data, the exact value of predictable components in observations remains unknown and can only be estimated. However, in the context of ensemble predictions using dynamical models, the noise component can be estimated by the difference between an individual ensemble member and its corresponding ensemble mean. Since the noise component is expected to be random, its amplitude diminishes by averaging over the ensemble. Consequently, the ensemble mean provides an estimate of the predictable signal. The relative magnitude of these two components is the signal-to-noise ratio (SNR); the greater the SNR, greater the predictability. Although SNR and predictability estimates can be made based on the ensemble approach, such estimates are influenced by model biases.

Predictability estimates based on model simulation indicate that for seasonal means, predictability is higher in tropical latitudes and lower in extratropical latitudes. In addition, predictability is larger over the regions that have a larger influence due to ENSO variability (Kumar and Hoerling, 1995; Trenberth et al., 1998). Because of regional and seasonal variations in predictability, the potential level of skill also has a strong regional dependence.

1.4 Probabilistic nature of seasonal forecasts

One important aspect of seasonal forecasts is that they are probabilistic in nature. This information, together with its implications with respect to decision-making, needs to be clearly communicated to the user.

Traditional weather forecasts provide information about the likely weather conditions during the next few days. Although it is impossible to skilfully predict daily weather changes beyond about a week in advance (Lorenz, 1965), it is possible to make inferences about likely future conditions averaged over periods of several months. Seasonal forecasts provide information on these long-term time averages, which are usually greater than one month but less than one year. The most common practice in various parts of the world is to use seasonal averages over three-month periods for variables such as precipitation and near-surface temperature.

As day-to-day weather is largely unpredictable on the seasonal timescale, so too is the precise value of the seasonal mean, that is, the average daily weather within the season. For this reason, seasonal forecasts are expressed probabilistically. Probabilities can be estimated from ensembles of predictions obtained from dynamical models, and it is common practice to use ensemble predictions produced by a collection of different models. Ensemble predictions are produced for each individual model so that observational uncertainties may be sampled when initializing the models (see section 1.5.2). A collection of climate models, called a multi-model ensemble, is designed to further sample uncertainties in model formulation. Technically, seasonal forecasts are presented as probability density functions (PDFs) which are derived from the available ensemble of seasonal mean predictions or via statistical methods. Wider PDFs of seasonal forecasts represent more uncertainty in predictions, while narrower PDFs indicate less uncertainty in predictions.

The time difference between the forecast time issuance and the forecast time validity is defined as the forecast lead time. For example, a seasonal forecast issued on 1 January and valid for the forthcoming February-March-April (FMA) season is often referred to as a one-month lead forecast, in other words, a forecast that provides information about the expected climate conditions in FMA one month in advance. The forecast valid for the March-April-May (MAM) season issued on the same date (1 January) is referred to as a two-month lead forecast because it provides information about the expected climate conditions in MAM two months in advance. As the forecast lead time increases, the forecast uncertainty usually increases. This feature is illustrated in Figure 1.6, which shows EUROSIP multi-model El Niño-3.4 (an ENSO index) sea-surface temperature anomaly forecasts produced on 1 May 2007 for the following six months. The spread between individual forecasts in the ensemble provides an estimate about the uncertainty involved in the forecast. The larger (or smaller) the spread, the more (or less) uncertain is the prediction situation. Figure 1.6 illustrates the "cone of uncertainty" in this forecast situation, with less uncertainty (less spread among the individual ensemble predictions) for shorter lead predictions and more uncertainty (more spread among the individual ensemble predictions) for longer lead predictions.

1.5 Seasonal forecast methods

Broadly speaking, there are three kinds of methods one can use to make seasonal forecasts: empirical (or statistical), dynamical, and hybrid. This subsection describes the key concepts of the empirical and dynamical methods. Hybrid approaches make use of a combination of these two methods to try to take advantage of both; they use a physics-based model output to represent the different processes in the climate system and statistical models to bias-correct and calibrate them (see section 2.4).

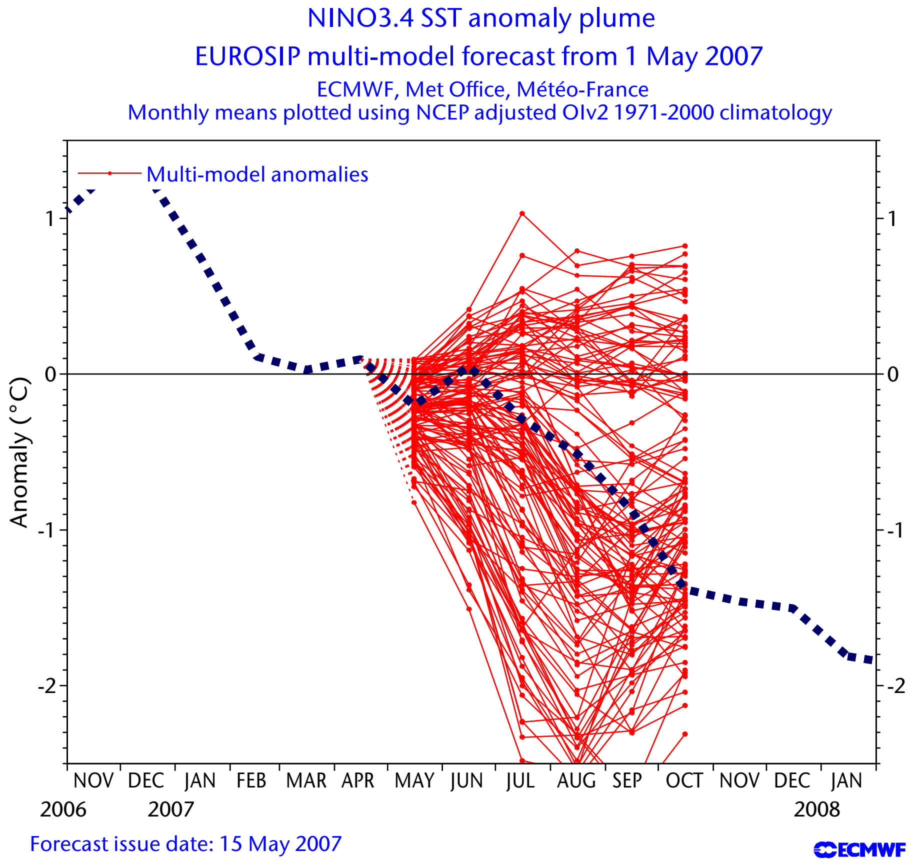

Figure 1.6. EUROSIP multi-model El Niño-3.4 sea-surface temperature anomaly predictions produced on 1 May 2007, issued on 15 May 2007, and valid for the following 6 months (red lines). The dashed blue line shows verified observed Niño-3.4 anomalies.

Source: European Centre for Medium-Range Weather Forecasts (ECMWF)

1.5.1 *Empirical seasonal forecast methods*

Empirical methods use statistical relationships between the predictor(s) – the variable(s) used to make the forecast – and the predictand – the variable to be forecast, usually involving some kind of regression model using observational data. One example of this approach is the use of the Niño-3.4 Index (the predictor) to forecast seasonal mean surface temperature (the predictand).

Predictors are identified through physical analysis of the mechanisms controlling the predictand (see section 1.2). For example, many RCOFs around the world use observed tropical ocean SSTs from the previous month as a predictor field to forecast precipitation and surface land temperatures for the target season.

Although there is a wide variety of statistical methods with different levels of complexity, the most frequently employed models are built using linear regressions. Building a statistical model involves computing the values of the coefficients α and β in the regression equation

$$Y = \alpha X + \beta$$

Models directly using the predictor value are normally called simple linear regression models; when using multiple predictors, they are called multiple linear regression models. To minimize multicollinearity errors (predictors are not completely independent) and multiplicity errors (too many predictors), it is customary to perform a dimensional reduction of the predictor space via calculation of Empirical Orthogonal Functions (EOFs) of the predictor variable. When the regression model uses the predictor EOFs as X, the model is called Principal Component Regression, or PCR.

When dealing with a model for multiple locations, the predictands may be intercorrelated, potentially causing a PCR approach to provide inconsistent predictions due to sampling errors when estimating the model coefficients. In those cases, it is possible to use Canonical Correlation Analysis (CCA) models, which maximize the correlation between linear combinations of selected predictand and predictor EOFs. Other techniques, such as Maximum Covariance Analysis (MCA), are also available. For additional details, see, for example, Mason and Baddour (2008).

Independent of the method used for empirical prediction, the evaluation of the statistical model being built needs to be performed using a cross-validation approach (for example, Mason and Baddour, 2008), as indicated in section 2.2.1.

Some advantages of empirical methods are that they require low computing resources and are easy to implement operationally, they are designed to be consistent with the observations (they are already bias corrected, at least with respect to the mean values), and they offer predictions in terms of both deterministic values and probabilities (the entire PDF, or for particular categories). Some disadvantages are that most common empirical methods assume stationarity in the climate system and therefore cannot adequately represent the presence of trends and other variations, they tend to have problems reproducing the observed variance of the predictands, and since they are generally based on linear relationships, it is difficult for them to represent non-linear interactions within the climate system. Given the continual improvements in dynamical model-based seasonal prediction systems, it is expected that in the future, these systems will be more effective than empirical prediction methods at making accurate seasonal forecasts.

Although empirical methods tend to use observed fields (for example, SST, or atmospheric circulation variables) as predictors, they can also use the output of dynamical models (section 1.5.2). Dynamically predicted precipitation fields, for example, can be used as predictors, while observed precipitation can be used as the predictand. Such a “hybrid” method can also correct for model biases (section 2.4).

IRI’s Climate Predictability Tool (CPT) (Mason and Tippet, 2018) is a freely available tool used worldwide for research and application purposes. The CPT includes several of the empirical prediction methods mentioned in this subsection.

1.5.2 ***Dynamical seasonal forecast methods***

Dynamical seasonal forecast methods, which are carried out using global climate models, are being used with increasing frequency. There are currently two approaches to predicting expected climate conditions for the forthcoming months and seasons:

- **Two-tier system approach:** An atmospheric global general circulation model is forced at its lower boundary with either persisted or predicted SST anomalies. For example, when performing climate prediction runs with persisted SSTs, for a forecast made on 1 January 2018 and valid for the six-month period from January to June 2018, the observed SST anomaly of December 2017 is added to the climatological (long-term mean) SSTs of these six months during the integration of the model (Bengtsson et al., 1993; Coelho et al., 2012). Another approach for specifying future SSTs is to damp the observed SST anomaly by using the so-called damped persistence procedure (Li et al., 2008). When performing climate prediction runs with predicted SST anomalies, SST anomalies predicted by another model are added to the climatological SSTs and used as boundary conditions to force the atmospheric global general circulation model. For atmospheric initialization, different initial conditions are used to sample uncertainties in the initial state to run the atmospheric model and generate an ensemble of predictions (Mason et al., 1999). Two-tier systems have the best utility for short lead time predictions, over which persisting SST anomalies have reasonable skill.

- **One-tier system approach:** A coupled ocean–atmosphere global general circulation model (for example, Saha et al., 2014) is used to generate climate predictions typically up to 6–12 months into the future. In this approach, both the ocean and the atmosphere need to be initialized with small differences in initial conditions to sample uncertainties in the initial states and generate an ensemble of predictions (although there are approaches where stochastic schemes within the model or lagging of ensembles can also generate the ensemble spread, in which case introducing small initial perturbations is not necessary). Other components, such as land, including soil moisture, snow and sea-ice conditions, may also require initialization in some models. Coupled dynamical prediction systems are increasingly being used instead of the two-tier approach. This is because coupled systems can represent two-way air-sea interactions, which are of great importance for seasonal forecasting. Out of 13 seasonal prediction systems used at GPCs-LRF, currently 11 follow the one-tier approach.

Dynamical seasonal prediction is very resource-intensive and requires access to an extensive computing infrastructure both for predictions and for the analysis to generate initial conditions (a process that also requires an infrastructure for data assimilation and ingestion of global in situ and satellite observations in real time) for various components of the Earth system.

One advantage of using dynamical seasonal forecast methods is that predictions of seasonal anomalies are not constrained by assumptions of linearity, which is often an underlying assumption in empirical prediction approaches. Dynamical prediction systems also represent the numerous climate processes that can influence seasonal variability in a region more completely than empirical/statistical approaches can. Dynamical models generate predictions from first physical principles, so they are potentially adept at generating a wide range of possible behaviour, unlike statistical models, which are constrained by observational data. Dynamical seasonal prediction systems are also adept at predicting unprecedented (or rare) climate patterns that have not occurred in the historical data, for example, recent warming trends in temperature or teleconnections associated with extreme ENSO events.

Another important advantage of using dynamical seasonal forecast methods is that the availability of ensembles allows for better quantification of the uncertainty associated with seasonal forecasts and quantifying probabilities of seasonal mean outcomes.

One disadvantage of using dynamical methods is that dynamical models have biases in the representation of the mean and variance of variables such as precipitation, cloudiness and snow, and with respect to the correct reproduction of the spatial patterns (for example, location, extension, shape) of atmospheric and oceanic variables, from SST to circulation fields and associated teleconnections. The need to evaluate these biases through hindcasts of past seasons and to compare them with observed outcomes also adds substantially to the cost and complexity of dynamical systems (section 2.2).

The output from dynamical models can be very extensive and can provide forecasts of numerous variables (temperature, specific humidity, wind, precipitation, soil moisture, surface insolation, and so forth) at daily or higher temporal resolution. Such model outputs can be used to explore predictions of tailored, user-oriented quantities – onset and withdrawal dates of monsoons, seasonal hurricane frequency, and so forth – and to provide forcings/input for application models, for example, streamflow prediction (section 2.5).

It should be noted that as dynamical seasonal forecast systems can simulate non-linear components of teleconnection patterns (for example, Hoerling and Kumar, 1997), their skill, in principle, should be better than linear empirical prediction methods and should increasingly be the preferred approach for making seasonal forecasts.

CHAPTER 2. COMPONENTS OF A SEASONAL FORECAST SYSTEM

Sections 2.1 and 2.2 in this chapter summarize how dynamical seasonal forecasts are produced using global numerical climate models at operational forecasting centres. The remaining sections describe how these forecasts are further processed and combined with other sources of forecast information to create seasonal forecasts serving specific regions and socioeconomic sectors.

2.1 Real-time dynamical forecasts

The ultimate objective of any operational dynamical seasonal forecasting system is to provide reliable real-time predictions to users. Real-time dynamical seasonal forecasts generally require:

1. A global numerical climate model;

2. Comprehensive real-time observations, typically obtained from the WMO Information System (WIS) through its Global Telecommunication System (GTS);

3. Timely model initial conditions from a data assimilation system that constrains model states with these observations;

4. A method for generating ensembles of forecasts, for example, through perturbed initial conditions, which may be provided by the data assimilation system or by other means.

In some respects, these components are similar to those required for numerical weather predictions. Components 1, 2 and 3 are required for both seasonal and weather predictions; however, weather prediction models tend to have higher spatial resolution and are less reliant on forecasting future ocean states, which evolve relatively little on weather timescales. Forecast ensembles initialized in accordance with component 4 are less critical for short-term weather forecasts, for which uncertainties are relatively small. However, for forecasts beyond a few days, ensembles are essential for quantifying forecast uncertainties.

These similarities between weather and longer-range forecasting systems provide a basis for seamless prediction across timescales using essentially the same modelling infrastructure and techniques, which is increasingly becoming a focus of forecast system research and development.

Real-time seasonal forecasts are ordinarily produced by NMHSs that have the capacity to do so, although some research centres, such as those described in Chapter 6, also produce seasonal forecasts in real time which may contribute to multi-model ensembles. NMHSs may also contribute to one or more of these aggregations of forecasts, and as GPCs-LRF, they provide their forecasts to WMO's infrastructure for seasonal prediction (Chapter 5). NMHSs without the computational infrastructure to produce dynamical seasonal predictions tend to use empirical methods (section 1.5.1) or hybrid forecast methods (section 2.4).

Real-time seasonal forecasts are communicated to the public and forecast users through data, forecast maps and sometimes text information on the producing centre's website.

2.2 Hindcasts

Hindcasts, also known as historical forecasts or reforecasts, are produced using the same model and methodologies as a producing centre's real-time forecasts, but "predict" past climate states based on observational information available at the initial time of the hindcast.

Hindcasts serve two main purposes: verification, and bias correction/calibration. For verification, the historical performance of the forecast system is assessed through comparison of the hindcasts with verifying observations. This provides measures of forecast quality and skill, where quality describes the association between forecasts and corresponding observations, and skill measures whether the forecast quality exceeds that of reference forecasts such as climatological probabilities (WMO, 2018). For seasonal forecasts to have skill, two necessary conditions must be satisfied: (i) there must be sufficient predictability in the climate system to enable a predictable "signal" to be distinguished from random, unpredictable climate variability ("noise") as described in section 1.3, and (ii) the prediction system must have sufficient fidelity to be able to realize this predictability. The second condition requires that:

- The prediction model adequately represent the processes that give rise to predictability;
- The size of the forecast ensemble be large enough that the predictable signal can be distinguished statistically from random sampling errors and noise; and
- The observations used to initialize and verify the predictions be of sufficient quality to adequately represent the true state of the climate system.

Further aspects of hindcast skill assessment are discussed in section 2.2.1.

For the purposes of bias correction, hindcasts provide a sample of independent predictions that is large enough to enable systematic model errors to be estimated and removed from real-time forecasts. Such errors include:

- Model drift, which occurs because forecast models, when not constrained by observations, tend to simulate conditions that to some extent are different, or biased, compared to the actual climate; thus forecasts, following initialization from an observed climate state, tend to evolve (drift) towards the biased climate that the forecast model prefers; and
- Errors that occur in quantifying the separation between forecast ensemble members, that is, the ensemble spread, relative to the true uncertainty in the climate system; ensemble spreads that are too small lead to overconfident forecasts, whereas ensemble spreads that are too large lead to underconfident forecasts.

The estimation and removal of systematic errors through bias correction and calibration is described in section 2.2.2.

Generally, at least thirty years of hindcasts is usually considered to be an adequate sample for skill estimation and systematic error correction, and the WMO Standardized Verification System for Long-Range Forecasts (SVSLRF) recommends that the hindcast period should be as long as possible (*Manual on the Global Data-processing and Forecasting System* (WMO-No. 485)). A long hindcast period is needed to capture an adequate statistical sample of the influences of climate variability (such as ENSO) that provide predictability, especially for noisier variables such as precipitation.

Among the 13 GPCs-LRF, presently there is some non-uniformity with respect to hindcast periods from different forecasting systems, with some spanning 30 years or more and some spanning shorter periods. Hindcast periods shorter than about 20 years may suffer from inadequate sample size to allow a robust estimation of skill. In addition, a shorter hindcast period impacts the merging of the information coming from different models using different hindcast periods (especially for multi-model ensemble (MME) approaches) because the anomalies and forecast quality are calculated with respect to a hindcast period that is common across all the models.

Another aspect in which hindcasts differ among GPCs-LRF is their production schedule. For some systems, fixed hindcasts are produced before the system becomes operational and remain unchanged throughout its operational lifetime. Other systems run and deliver a new set of hindcasts every time a new forecast is produced, following the so-called "on the fly" approach. Each method has its own advantages. Fixed hindcasts provide stability to users who produce tailored information based on calibrated inputs, providing a long-term basis for system evaluation. On the other hand, on-the-fly hindcasts more readily accommodate frequent model upgrades (motivated by skill improvements) since on-the-fly hindcasts are specific to the initialization date of a given real-time forecast.

One further aspect that can differ among forecasting systems is the hindcast ensemble size relative to that of real-time forecasts. For some systems, hindcast and real-time forecast ensemble sizes are the same, in which case, hindcast skill reflects that of real-time forecasts. For other systems, for reasons of computational economy, hindcasts have fewer ensemble members than real-time forecasts; in such instances, hindcasts provide conservative estimates of real-time skill because skill generally increases with ensemble size (Kumar and Hoerling, 2000; Kharin et al., 2001), whereas bias correction is less sensitive to ensemble size.

Although at present, hindcast and real-time forecast configurations across GPCs-LRF have appreciable differences, which makes using multi-model approaches more difficult, achieving increased homogeneity across seasonal forecast systems is the goal.

Additional potential uses of hindcasts include providing input for application models enabling probabilistic forecasts for crop yields, streamflow, and so forth (Palmer et al., 2005), and serving as a basis for statistical downscaling, as discussed in section 2.7.

2.2.1 ***Establishing the skill of seasonal prediction systems based on hindcasts***

Measures of historical predictive skill are an essential component of seasonal forecasts because they indicate to users how much trust can be placed in the real-time forecast; without this information, the forecast has little value. Access to historical predictive skill measures obtained from hindcasts should therefore be provided together with the forecast itself.

Many predictive skill measures describe the quality of specific forecast attributes. They are estimated by calculating the corresponding properties of the set of hindcasts paired with verifying observations. Some care should be taken in interpreting such measures as they are estimates that are based on the limited sample of hindcasts and are influenced by sampling (Kumar, 2009). Therefore, uncertainties in skill measures associated with the limited sample size should be quantified, for example, by confidence intervals determined through bootstrapping. This should be kept in mind when comparing the quality of different sets of forecasts from various operational systems.

A further important aspect of historical forecast quality evaluation is that it should be calculated in a cross-validated framework. In its simplest form, cross-validation implies that each hindcast should be treated as if only information from the remaining hindcasts is available. This includes bias corrections as outlined in section 2.2.2 as well as the statistics required for the calculation of a given forecast quality measure. In some cases, appropriate variations of this procedure should be applied, for example, if successive forecasts are serially correlated.

If cross-validation is not employed, the resulting values will not be fair estimates of forecast quality because information is used that would not have been available had the individual hindcasts been out of sample in the manner of a real-time forecast. In general, forecast quality measures calculated without cross-validation will tend to be biased towards implying greater forecast quality than exists. The same is true for assessing the skill of empirical forecasts (section 1.5.1).

The process of estimating forecast system quality measures based on a series of paired hindcasts (or real-time forecasts) and observations is referred to as verification (see the *Guidance on Verification of Operational Seasonal Climate Forecasts* (WMO-No. 1220)), although the term “validation” is sometimes incorrectly applied. (Validation in the seasonal forecasting context more correctly refers to the checks that are performed to ensure that a forecasting system is functioning as intended.) It is beyond the scope of this document to provide a detailed description of the utility and calculation of the many forecast quality measures that are employed in seasonal forecasting, except to emphasize, as in the *Guidance on Verification of Operational Seasonal Climate Forecasts*, the particular importance of discrimination, which describes the forecasts’ ability to distinguish between different outcomes, for example, through the relative operating characteristics (ROC) curve, and reliability, which describes how well the forecast probabilities convey the true (or observed) frequency of occurrence of an event. Reliability is essential for probabilistic forecasts to be useful and is discussed further in section 2.9 and elsewhere in this document. In addition to the comprehensive WMO guidance document on verification (*Guidance on Verification of Operational Seasonal Climate Forecasts*), readers are referred to the WMO's Standardized Verification System for Long-range Forecasts (*Manual on the Global Data-processing and Forecasting System*), books by Joliffe and Stephenson (2012) and Wilks (2011), and various online resources listed in the references.

Forecast quality is typically a function of geographical position, predicted season, forecast lead time, and the prediction system itself. Some forecast quality measures such as reliability may require aggregating over some extended spatial region in order to obtain an adequate sample. Also, forecast quality measures are sometimes averaged over spatial region, target season and/or lead time in order to obtain an overall performance assessment.

It should be emphasized that historical forecast quality measures obtained from hindcasts characterize aspects of the expected quality of the forecasts, on average, over many forecasts. In principle, conditional performance pertaining to a subset of the hindcasts may be assessed, for example, to compare forecast quality during El Niño and/or La Niña events with that during non-ENSO conditions. Such evaluations, although highly desirable, should be approached with extreme caution, however, since they require reducing the already limited hindcast sample size, thus magnifying sampling uncertainties which may already be substantial. This, in turn, makes it difficult to distinguish with any confidence forecast quality differences between hindcast subsets or with respect to the hindcasts as a whole.

The quality of a single forecast can be assessed by treating paired forecasts and observations as the sample being analysed. Single-forecast verification is often applied to real-time forecasts and can be useful for quantifying the performance of individual forecasts or tracking any trends in system performance. It is discussed extensively in the *Guidance on Verification of Operational Seasonal Climate Forecasts*. However, the available sample of real-time seasonal forecasts is unlikely to be large enough to provide a clear indication of how the forecast system is likely to perform on average, reinforcing the need for a forecast quality assessment from a much larger set of hindcasts.

A further aspect of the forecast quality assessment process that should be kept in mind is that the results of quality assessments are a function both of the quality of the forecast and that of the verifying observations. A consequence of this is that forecast quality measures can differ depending on the data source used for verification. Forecast quality may give false information if the quality of the verification data is poor. Conversely, if the verification data comes from the same data source used to initialize the forecasts, the forecast quality may be artificially enhanced if errors in the initialization propagate into the forecast and correlate with errors in the verification data. Some of the issues associated with the choice of verification dataset are assessed in Massonnet et al. (2016).

Finally, it should be noted that from the perspective of a forecast user, skill is different from, but complementary to, uncertainty; whereas uncertainty in individual forecasts is conveyed to the user through forecast probabilities of different outcomes, skill indicates whether, on average, the quality of the forecast system exceeds that of some reference forecasts such as climatology or anomaly persistence. In an ideal situation where the forecast probabilities for an event are reliable, they also convey information about the success rate of the forecasts. Skill is also distinct from and does not guarantee forecast value, which measures the economic, social or other benefits that the forecast information potentially can provide (see, for example, Richardson, 2000).

2.2.2 ***Bias correction and calibration and the use of hindcasts***

The global climate models used to produce dynamical seasonal forecasts are highly sophisticated numerical representations of the climate system. They have been developed and improved over the past several decades and continue to improve. Nonetheless, these models, though remarkably realistic, remain imperfect in that their simulated versions of the climate system exhibit systematic differences in relation to observed climate properties in addition to the non-systematic (or random) errors that limit skill.

Corrections for systematic errors fall into two broad but distinct categories. The first adjusts properties of the modelled climate to match those of the observed climate without reference to prediction quality or skill, in other words, without pairing hindcasts and observations as discussed in section 2.2.1 in the context of skill evaluation. This type of adjustment is sometimes referred to as “calibration”, although it is more often called “bias correction” in the scientific literature. This document uses the latter term.

The simplest and perhaps most crucial type of bias correction corrects the biases in the mean climate that develop after initialization as the model drifts from its initial states (which generally are close to observations) towards the biased climate that the model simulates. (It is important to note that biases are a function of location, lead time and starting month of the forecast). For example, if the forecast model is biased cold in a location, the forecasts will systematically be biased cold unless a correction is applied. Illustrations of such biases and drifts are provided in Saha et al. (2006), Molteni et al. (2011) and elsewhere. The procedure usually applied is first to calculate the mean values of a particular variable for each grid location, predicted season and lead time for hindcasts over some standard period (for example, 1981–2010). This value, denoted $< F_{k,l} >$, where k labels the variable, grid location, and predicted season, l labels the lead time, and $< >$ indicates the climatological average over the years in the hindcast period and all ensemble members, is then subtracted from the corresponding value $F_{k,l}$ for a particular forecast, so that

$$F'_{k,l} = F_{k,l} - < F_{k,l} >$$

where $F'_{k,l}$ is the forecast anomaly for variable, grid location, season k and lead time l. If a bias-corrected “full” forecast value $(F_{k,l})_{corr}$ is required, this forecast anomaly can simply be added to the observed climatological average $< O_k >$, where this average is computed for the same years as for the hindcasts:

$$(F_{k,l})_{corr} = F'_{k,l} + < O_k > = F_{k,l} + < O_k > - < F_{k,l} >$$

where the second equality clearly indicates the nature of the bias correction.

Bias corrections can also be applied to other aspects of the forecast values. For example, if the interannual standard deviation of the ensemble mean is biased high or low so that the forecasts have too much or too little year-to-year variability compared to observations, the forecast anomalies can be multiplied by the ratio of the observed to forecast standard deviation. Another approach is to adjust the forecast PDF based on a quantile mapping between the hindcast and observed PDFs (see, for example, Yuan and Wood, 2013).

A second type of systematic error correction involves modifying forecast values to optimize hindcast quality or skill. Although this may also correct certain model biases in the ensemble spread, it requires consideration of paired hindcast and observed values and thus, differs fundamentally from bias correction as defined above, which considers the statistics of the hindcasts and observations independently and adjusts the former towards the latter. This second type of correction is sometimes referred to as “recalibration”, but it is often simply called “calibration” in the scientific literature. The latter term is used here.

A primary purpose of calibration is to improve the properties of probabilistic forecasts, especially their reliability. As introduced in section 2.2, if the spread of the forecast ensemble is too small relative to the true uncertainty in the climate system, forecasts will tend to be overconfident in that the observed frequency of events forecast with high probability is too low and the observed frequency of events forecast with low probability is too high. This corresponds to a slope that is too shallow in the reliability diagram (Figure 2.1, left panel), whereas a slope that is too steep indicates underconfidence (middle panel), and forecast probabilities equalling the observed frequencies indicate perfect reliability (right panel). Uncalibrated forecast probabilities are often overconfident, although in some instances they may be underconfident, in which case, the ensemble spread is too large relative to the true uncertainty. See section 2.9 for further discussion on the importance of reliable probabilistic forecasts.

A typical approach to calibration is first to fit hindcast ensemble values to a parametric probability distribution, such as the normal distribution, which by itself provides an improved estimate of the forecast PDF (Wilks, 2002), and then to adjust the parameters of that distribution to optimize a probabilistic forecast quality measure. In addition, it can be advantageous to smooth the resulting calibration coefficients in time and/or space to reduce sampling uncertainties (Kharin et al., 2017). This chain of post-processing steps, beginning with bias correction of raw forecast values, is illustrated schematically in Figure 2.2, where reliability diagrams are shown in a form known as attributes diagrams, which are explained in Wilks (2011) and elsewhere. Although several distinct approaches to calibrating probabilistic forecasts have been proposed (see, for example, Unger et al., 2009; Kharin et al., 2017; van den Dool et al., 2017), each aims to achieve the same basic objective of improving the reliability of probabilistic forecasts.

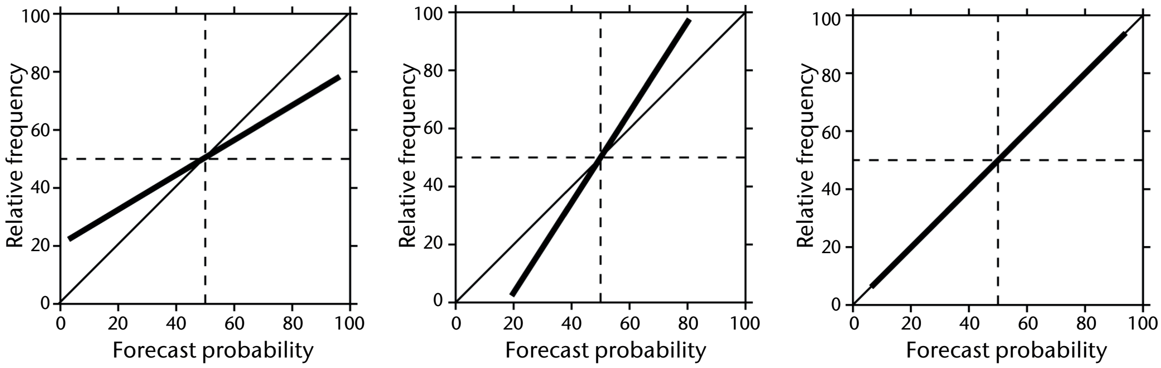

Figure 2.1. Reliability diagrams corresponding to probabilistic forecasts that are overconfident (left), underconfident (middle), and perfectly reliable (right).

Source: Guidance on Verification of Operational Seasonal Climate Forecasts (WMO-No. 1220)

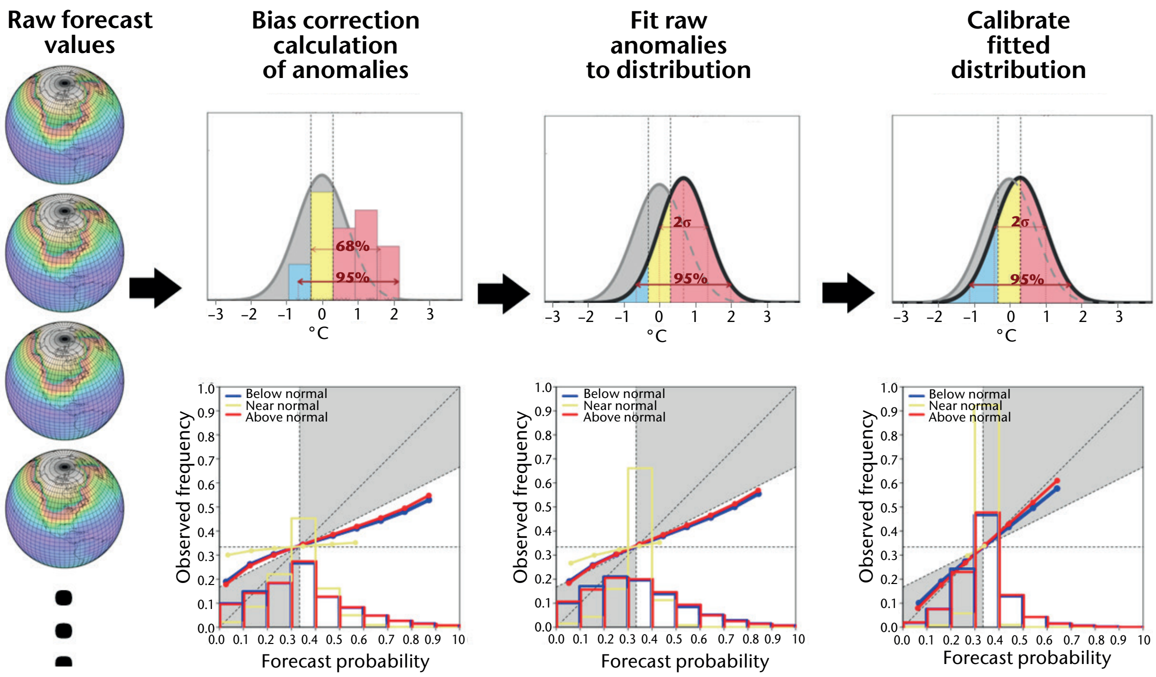

Figure 2.2. Schematic depiction of a chain of post-processing steps leading from raw model output to a calibrated probabilistic forecast. Model output from an ensemble forecast (first column) at a particular location is bias corrected to provide a sample of raw forecast anomalies (second column), where blue, yellow and red denote anomalies in the below- near- and above-normal terciles of the climatological distribution indicated by the grey curve. This sample is fit to a parametric distribution (third column), the parameters of which are adjusted systematically through a calibration procedure to maximize probabilistic performance and improve reliability (fourth column). (After Kharin et al., 2017)

The effects of calibration on probabilistic forecast maps are exemplified in Figure 2.3, which depicts a probabilistic forecast of precipitation over Canada from GPC-LRF Montreal. Precipitation forecast quality in this region is relatively poor, and although relatively high tercile probabilities are predicted, the uncalibrated forecast is overconfident and the near-normal category in particular is grossly overpredicted and has near-zero reliability (left column). In contrast, calibration adjusts the probabilities towards more appropriate confidence, leading to improved reliability (right column).

It should be noted that combining uncalibrated forecasts from multiple models also tends to improve reliability (Hagedorn et al., 2005; Becker and van den Dool, 2016) and may make probabilistic calibration as described above less crucial than for individual models. Such an approach is described in Min et al. (2009) and applied in constructing the multi-model probabilistic forecasts of the WMO LC-LRFMME (section 5.1).

Although its main application is to probabilistic forecasts, calibration can also be applied to deterministic forecasts. For example, ensemble mean anomalies can be rescaled to minimize mean square error (Kharin et al., 2017). This differs from bias-correcting the interannual standard deviation by rescaling the forecast anomalies as described above because under such a

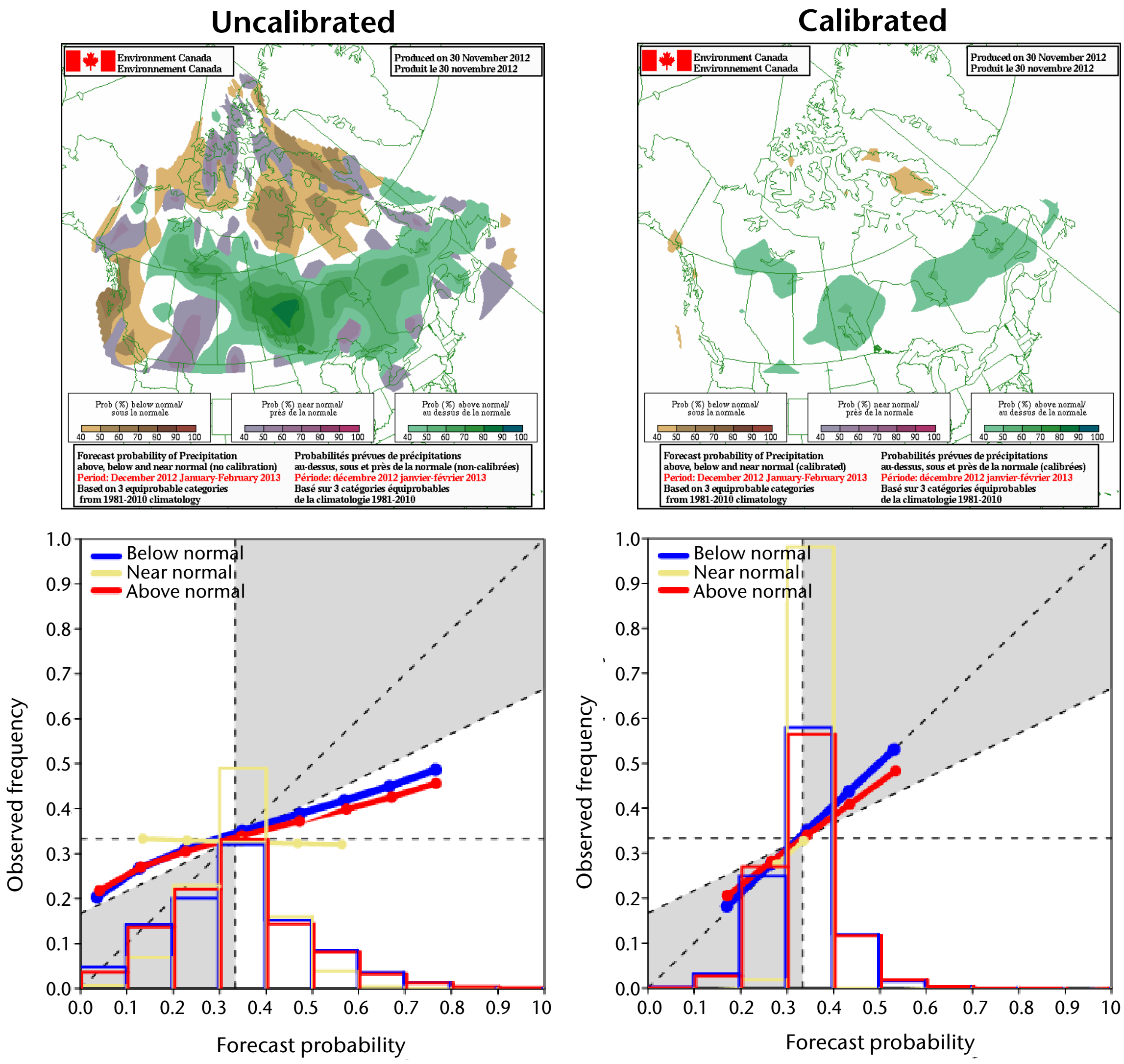

Figure 2.3. Uncalibrated (left) and calibrated (right) precipitation forecasts from GPC-LRF Montreal, with associated reliabilities shown in attributes diagrams

Source: Environment and Climate Change Canada calibration, the rescaling coefficients can be negative; this may occur if the forecasts erroneously represent a teleconnection in such a way that forecast anomalies systematically are opposite in sign to observed anomalies. Under this type of calibration, the mean-square forecast error is never greater than that of a deterministic climatological (zero anomaly) forecast. As for forecast quality assessment, it is important that calibration (and indeed all types of statistical post-processing) be carried out within a cross-validated framework (section 2.2.1).

The bias correction and calibration techniques described above are typical of those applied to produce global seasonal forecasts at individual forecasting centres or centres that aggregate individual forecasts as described in section 5.2 and Chapter 6. A fundamental aspect of these techniques is that they are applied locally at each grid location and do not directly consider errors in representing spatial patterns of variability. Additional methods exist that do apply spatial corrections, for example, using canonical correlation analysis (Barnston and Tippett, 2017) to improve predicted large-scale anomaly patterns. These methods can be useful for improving forecasts in specific regions and are discussed in section 2.4.

2.2.3 *Observed climate data requirements for hindcasts*

From the discussion in section 2.2.2, it should be clear that post-processing hindcasts requires corresponding observations. Further, as bias correction and calibration of seasonal forecasts (see section 2.2.2) are conducted on model outputs, the gridded observed data for these purposes should match the hindcast data in terms of period and spatial resolution. The process of bias correction and calibration is intended to correct model forecasts based on observations. Consequently, any observational errors can negatively affect these post-processing steps. In particular, biases in observations (for example, remotely sensed observations) may introduce additional errors in calibrated seasonal forecasts. The selection of the hindcast period for calibration is also affected by the climatic trends existing in that period. In particular, trends resulting from global warming or decadal variability may be inherited by calibrated seasonal forecasts (section 3.1).

It should be stressed that hindcast periods with predominance of different phases of variability patterns may also influence the estimate of the performance of seasonal forecasts. Recent studies have suggested that considerable success has been achieved in forecasting winter climate anomalies over the Euro-Atlantic area using current-generation dynamical forecast models. This skill is at least partially attributed to the fact that the more recent decades have been dominated by NAO in its positive phase. Using a century-long hindcast, Weisheimer et al. (2017) have shown that there is a complex relationship between the NAO phase/amplitude and forecast skill that can lead to considerable low-frequency variability in skill.

Finally, it should also be stressed that gridded observational data for model calibration are different from station data intended for regional/local calibration and downscaling (section 2.5). The purpose of the latter is mainly to add small-scale detail to the large-scale forecast provided by seasonal forecast models, whereas the objective of the former is to correct forecast fields on the grid of the seasonal models.

2.3 Using multiple forecast tools and multi-model ensembles

Developers of real-time seasonal forecasts have a choice of forecast tools – empirical or dynamical – and a choice of forecasts from various dynamical models, for example, from the 13 GPCs-LRF. Such choices require consideration of which approaches should be followed for the selection of models and to blend forecasts in order to develop consolidated seasonal forecasts over a region.

2.3.1 *Selecting the most appropriate model(s)*

Different models have different biases that influence their ability to predict the observed climate in light of varying climate situations and for various regions. For one climate situation and region, one model may be preferable; however, this same model may not be appropriate for use in other regions and seasons. Moreover, for some regions and seasons, the relatively short historical hindcast period that is generally available may prevent an unequivocal identification of the best model. It is therefore important to be familiar with good practices for selecting the models that should be used.

The first step in the seasonal forecast development process is the selection of a set of candidate seasonal forecasting models. By default, the 13 models included in the WMO LC-LRFMME can be considered for the initial selection process (see section 5.2 and the WMO LC-LRFMME), but the decision of which model to choose can vary based on operational constraints. Although all models included in the WMO LC-LRFMME share a minimum list of characteristics related to spatial/temporal resolution, minimum hindcast period, minimum ensemble size for hindcast and forecast, minimum issuance frequency, and so forth (see LC-LRFMME for a description of each model configuration), additional conditions may guide the choice of model(s), for example:

- Due to operational constraints concerning the provision of regional forecasts on a specific schedule imposed by RCCs or RCOFs, the timely availability of hindcasts in real time is necessary. Because of these constraints, the strategy of some GPCs-LRF to generate hindcasts on the fly, and the consequent difficulty of making these hindcasts available in a timely manner, may put a strain on meeting the operational schedule and may need to be taken into consideration during the selection process.

- There may be a preference for the use of coupled atmosphere–ocean systems over two-tier systems although it should be noted that the skill of the current generation of coupled models may not be unequivocally better than that of two-tier systems.

- When the forecast lead time is on the order of two months (for example, there is a requirement in May to issue forecasts for the July-August-September West African monsoon), the typical range for two-tier systems will be insufficient, and one-tier systems need to be used.

The next set of criteria for selecting models is related to the regional performance of seasonal forecasting systems. A comparison of objective verification scores computed over a common hindcast period provides insight on the quality of different systems. Care should be taken to ensure that the period of verification is the same for all systems considered and the verification dataset is the same. It is also recommended that two observational datasets be used for verification purposes in order to check the dependency of results on the observational dataset selected. In some cases, the required information may be available to assist in the model comparison. For example, the LC-LRFMME website provides global skill maps for all GPCs-LRF using a common hindcast period and verification dataset. More frequently, however, and particularly if statistical methods or calibration methods are employed, it is advisable to conduct the forecast system comparison exercise as part of the seasonal forecast process.

In selecting models to be used at the regional level, it is a good practice to generate for the region of interest tables for different variables and models showing verification scores (both deterministic and probabilistic) month by month. Figure 2.4 shows one verification score (correlation coefficient) for temperature computed for the Mediterranean region using a common hindcast period for six models and for 12 successive one-month lead seasonal forecasts. Based only on this deterministic verification score, it is noteworthy that there is no one model that outperforms the others for all seasons. Moreover, if the maximum hindcast period for each model is used (instead of the common overlapping period), the table changes noticeably (result not shown). Therefore, it is important to underline the strong dependence of verification scores on the specific hindcast period.

One drawback of using skill measures to select models is that the statistical significance of skill scores is often difficult to discern (see also Siegert and Stephenson, 2018; SPECS, 2016). It is therefore recommended that a skill measure approach for selecting models be complemented by attention to the ability of models to simulate large-scale features and processes of the climate system at the global and/or regional level. Considering the large-scale fidelity of simulations as criteria for selection, it is advisable to select models based on an analysis of their ability to simulate climate drivers, climate variability patterns and teleconnections that are relevant at a

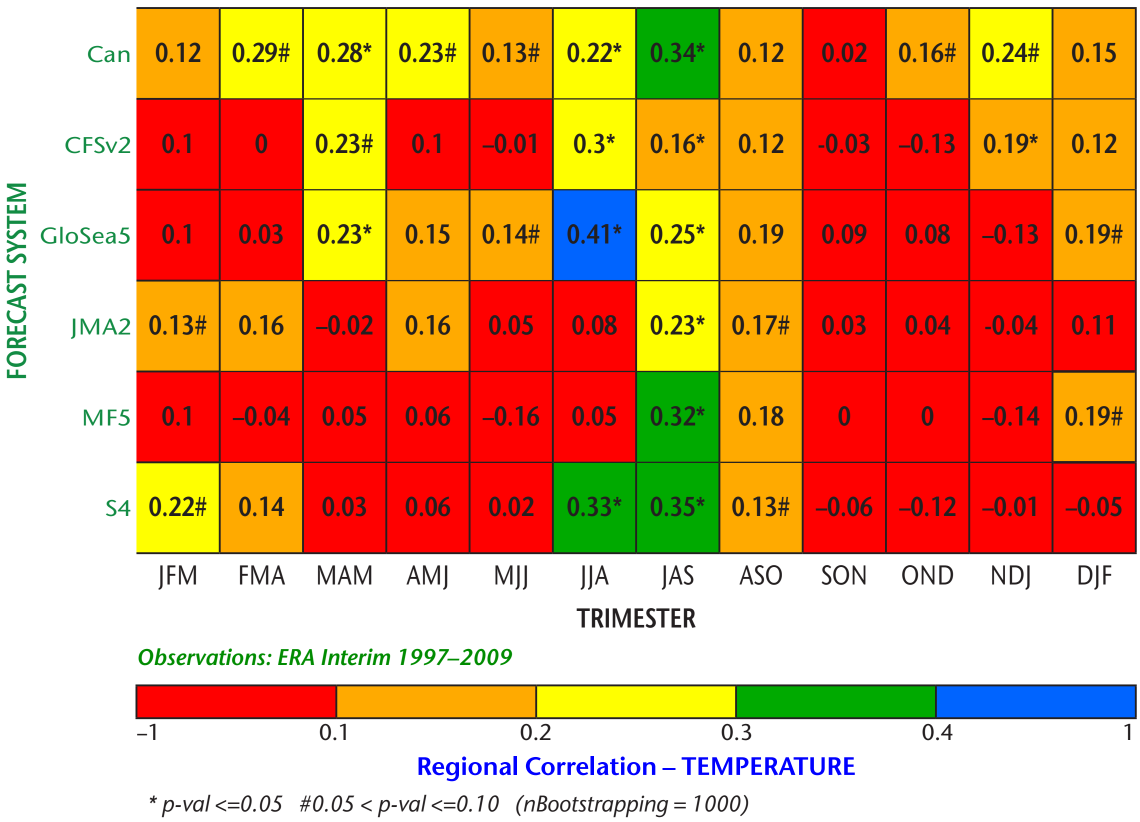

Figure 2.4. Correlation coefficient computed for regional temperature anomalies (without removing trend), for 12 different three-month periods and for a one-month lead forecast over the Mediterranean region (common verification period: 1997–2009). Three-month periods are shown on the X-axis, and systems (GPC Reading S4, GPC Toulouse S5, GPC Tokyo S2, GPC Exeter (GloSea5), GPC Washington v2 and GPC Montreal CanSIPS) are represented on the Y-axis.

Source: Mediterranean Climate Outlook Forum (MedCOF)

seasonal scale over the region of interest. For example, if NAO is the pattern which explains the majority of the variability over the region of interest, the ability and skill of different candidate models to accurately simulate NAO events should be checked. Selected models should also be able to simulate the dynamical linkage between relevant drivers (sea ice, Siberian snow cover, and so forth) and NAO.

In summary, it is recommended that once the relevant drivers and variability patterns have been identified for all seasons, in addition to assessing the skill of the forecast system based on hindcasts, it is a good practice to analyse candidate models for their ability i) to forecast patterns that contribute to the climate variability for each season and ii) to simulate the proper teleconnections linking remote drivers and climate variability patterns. Such a procedure can provide an objective and well-documented methodology for the preselection of models.

2.3.2 *Combining seasonal forecasts from multiple inputs*

Combining predictions from different and complementary models helps improve our predictive ability, and there is documented evidence that an average of forecast inputs (the multi-model ensemble approach) is statistically a better predictor of observed climate than a single model alone and makes combining different climate model predictions advantageous and an advisable approach. (See SPECS (2016) for a review on this topic.)

In general, predictions from several different systems and tools will be available to RCCs and NMHSs to assist them in preparing a consolidated forecast. However, what should these entities do with this information? Should the most skilful system/tool be identified and used as the sole input? Should a blend of several of the more skilful systems be used? Should a more inclusive blend involving all or most of the available inputs be used? These questions are addressed below.

As a good practice, it is usually not recommended to develop forecasts based on a single prediction system for the following reasons:

- For the reasons stated above and in sections 2.2 and 2.3.1, it is not a simple matter to establish with certainty which is the most skilful seasonal prediction system.

- It may not be prudent to be overly dependent on a single system. There may be occasions when the forecast from the selected system is an "outlier", and it may be wise to take into account other perspectives when developing a forecast. In addition, occasionally, technical issues/delays may occur – and may impact timeliness.

- There are several studies in the literature that demonstrate that forecast skill may be enhanced by combining forecasts from different systems (Brown and Murphy, 1996; Graham et al., 2000; Hagedorn et al., 2005; Becker et al., 2014; see also section 2.2.2). This attribute of combined systems is derived largely from the complementarity that often exists – the best performing system may differ according to the nature of the prevailing climate forcing.

While there are no firm guidelines for the number of systems/tools that should be blended, it is recommended that more than a single forecast system should be used. However, depending on resource availability, inputs should be kept to a manageable number.

As discussed earlier, over a single region, different models have different skills. Differences in the levels of model skills may suggest that different models should be weighted differently when combined into a single prediction. However, because of short hindcast periods, different authors have reported that equal weighting generally performs better than the use of unequal weights. The difference is small, but systematic across various investigations (Hagedorn et al., 2005; Weigel et al., 2008; Weigel et al., 2010; DelSole et al., 2013). Recent results show that for hindcast periods shorter than 50 years, unequal weighting schemes are unlikely to improve upon the results of an equal weighting scheme. For larger sample sizes, the weights can be estimated more robustly, and the unequal weighting scheme may produce more skilful forecasts (Siegert and Stephenson, 2018; SPECS, 2016).

The simplest method of blending is to take a simple average of the values predicted by each system (for deterministic forecasts) and an average of predicted probabilities (if forecasts are probabilistic). For probabilistic forecasts, another option is to perform a consistent averaging of all the PDFs, which permits the probability of exceedance of thresholds of interests for the user to be calculated.

Following Min et al. (2009), GPCs-LRF contributions are typically weighted in proportion to the square root of the ensemble size. This method is used to construct probabilities in the WMO LC-LRFMME multi-model system. An approach for weighting individual model probabilities in proportion to the ensemble size is also employed for NMME (Becker and van den Dool, 2016).

Together with dynamical seasonal prediction systems, empirical prediction systems can also be used to improve model combinations. For example, Kilavi and Colman (2012) have explored an objective combined statistical-dynamical prediction methodology for subnational climate zones of Kenya where combining is achieved by using indices of dynamical model output in addition to pre-season SSTs as predictors in linear regression models for each zone. Coelho et al. (2006*a*) proposed the so-called integrated forecast approach for combining and calibrating coupled and empirical model predictions using a Bayesian procedure known as forecast assimilation (Stephenson et al., 2005). Dobrynin et al. (2018) applied a procedure for subsampling ensemble members that reproduce winter NAO states based on a statistical analysis of the initial autumn state of the ocean, sea ice, land and stratosphere.

2.4 Hybrid forecasting methods that combine both dynamical and statistical approaches

In cases where there are known teleconnection responses to the region of interest from large-scale climate modes, such as ENSO and IOD, it may be possible to enhance the skill of real-time predictions available from direct model output. This may be accomplished by using hybrid methods that utilize statistical relationships between predicted model variables (or indices) related to the large-scale modes of variability and the observed variables in the region of interest. Typically, hybrid methods involve comparing information from individual hindcasts (not necessarily for the target region) with simultaneous observations in the target region. As such, they are forms of calibration and deliver outputs that are bias corrected as well as skill-optimized (see also section 2.2.2). In contrast to the calibration approaches discussed in section 2.2.2, in which analysis is performed locally at each grid point using paired hindcast and observed data, hybrid methods typically relate predicted information for remote regions that are far from the target region to observations over the target region. Additionally, the predictor variables may be different from the predictand variable (for example, the model predicted tropical Pacific SST may be used as a statistical predictor for precipitation in a remote, but teleconnected region). Moreover, if the observed predictand is represented by a gridded dataset, the hybrid approach implicitly conducts a downscaling to the resolution of the observed dataset.

The potential for skill improvement from such hybrid approaches is illustrated in Figure 2.5 using an example relating large-scale predicted precipitation to regional-scale precipitation over Africa. On large scales, the December–February precipitation response to ENSO is well captured at one-month lead by the GPC-LRF Exeter model (GloSea5 at time of writing) (reference the similarities in predicted and observed spatial characteristics of the anomaly patterns over the tropical Pacific, Indian, and to an extent, Atlantic Oceans). At more granular scales, however, discrepancies emerge in the vicinity of East Africa; the response centre of action in the model (bottom) is located too far east over the Indian Ocean; in southern Africa, the response appears too weak and with too much southerly extent relative to observations. These considerations suggest a potential hybrid approach based on the use of large-scale model predictions, for which performance is typically better (such as over the Indo-Pacific region in Figure 2.5), to infer predictions on regional scales (for example, over eastern and southern Africa) through the intermediary of the observed relationship between these large and regional scales.

This methodology will bring the largest skill improvements to models that have skill in predicting large-scale modes (for example, ENSO), but little skill in predicting the associated teleconnection responses over the region of interest. In cases where models successfully predict both the large-scale mode and the teleconnection response, the direct model output should be skilful, and use of a hybrid approach may not add further value. The quality of representation of the teleconnection responses can vary considerably with the model (for example, Gleixner et al., 2017) because of the challenges models face in accurately representing the complex chain of processes that transfer the "signal" from the source region along the teleconnection pathways (starting, for example, with convective processes transferring heat from the ocean to the atmosphere in the source region). Thus, when investigating the performance of GPCs-LRF for the region/country of interest, it is good practice to be aware of the quality of the teleconnection responses (in the model forecasts) to modes of SST variability that are influential in the region.

There are many examples of hybrid techniques in the literature which could be adapted for use. Two recent examples include Shukla et al. (2014), who used a constructed analogue technique relating predicted precipitation and SST over the tropical Pacific and Indian Oceans to observed March-May precipitation over East Africa; and Gleixner et al. (2017), who found that a linear regression model of GPCs-LRF multi-model predicted Niño-3.4 index on Ethiopian June–September precipitation achieved better skill than the direct precipitation predictions of any single model.

Other methodologies include extended logistic regression (Wilks, 2009), which is used to calibrate GCM output in products issued by the International Research Institute for Climate and Society and Forecast Assimilation (Stephenson et al., 2005; Coelho et al., 2006*a*). Some other methods and tools may be sourced from the Climate Services Toolkit (CST) of the GFCS. CST comprises a wide range of software tools designed to assist data processing and forecasting

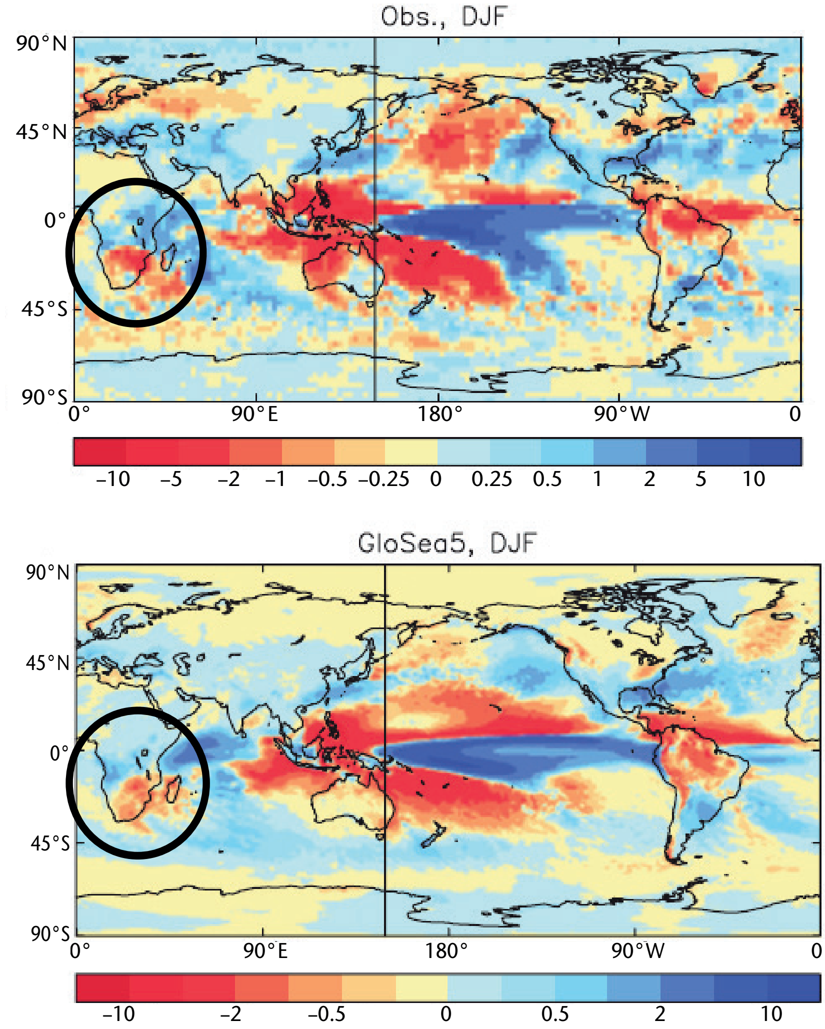

Figure 2.5. December-January-February (DJF) precipitation composites (mm/day) for seasons with El Niño minus composites of seasons with La Niña. Top: observed; bottom: GloSea5 one-month lead hindcasts. (Adapted from MacLachlan et al., 2015)

in support of climate services. Here, we briefly outline one tool, the Climate Predictability Tool (CPT), developed by IRI (see also section 1.5.1), because of its widespread use by many RCCs and NMHSs.

Several statistical methodologies are available within CPT, the most widely used being CCA, which has the advantage of operating on the gridded fields of the predictand to generate spatially varying forecast output. CCA operates by decomposing co-variability in two sets of spatiotemporal data (predictor field X and predictand field Y) into a series of orthogonal paired patterns (referred to as the X and Y CCA modes) whose associated pairs of time series (spatial projections of X on the X mode and Y on the Y mode) have maximum linear correlation. In other words, CCA decomposes two fields into their main associated modes of correlated variability (section 1.5.1). Real-time predictor data can then be projected onto the CCA spatial loading patterns to generate forecasts (Kipkogei et al., 2017). This analysis of covariability between the predictor and predictand fields enables systematic pattern errors in the model predictions to be identified (for example, errors in the spatial positioning of teleconnection responses) and corrections to be applied (Barnston and Tippett, 2017).

Predictor data required as inputs to CCA are the ensemble mean deterministic GPC outputs; the CCA forecast output is also deterministic. CPT includes a post-CCA operation which constructs a probability density function with a mean value equal to the deterministic output and a width equal to the standard error of the CCA forecasts as measured by cross-validated hindcasts. The PDF is then used to obtain probabilistic forecasts for terciles or other user defined categories.

Simply stated, CCA identifies co-variability between the predictor data and observed data such that we can say "when the model predicts a certain spatial pattern X in a given variable (for example, SST or precipitation over a large-scale domain), the dominant observed pattern in the variable/region of interest is typically Y". Thus, with reference to Figure 2.5, the predictor data could be gridded model ensemble mean predictions of precipitation or SST over a tropical Indo-Pacific domain, and the predictand could be gridded observed precipitation data over East or southern Africa. The predictor variable need not be the same as the predictand; model predictions of large-scale SST, precipitation and low-level U and V wind components have been used to predict regional/local precipitation. A large-scale predictor domain is often selected to take advantage of information from remote teleconnections to the region of interest. However, often good results are also achieved from predictor domains that are more or less coincident with the predictand domain. In this configuration, the operation is somewhat akin to the local calibration approaches described in section 2.2.2.

2.5 Tailoring of seasonal forecasts

The conventional three-category tercile probability format (see section 3.2) used to present forecasts by most NMHSs and RCCs as well as at RCOFs is designed to convey the broad depiction of the seasonal outlook. To encourage the forecast to be taken into account in decision-making, some tailoring to customize the information for specific users is often needed. In this section, some examples of tailored products are given with a brief description of the methodology and tools required.

We focus here on tailored forecast products relevant to RCCs and NMHSs that can be generated relatively easily from the one-month and/or three-month mean forecast data supplied as part of the minimum GPCs-LRF mandatory function. Tailoring of this information may be approached using statistical methods similar to those described in section 2.4 and usually requires some simple assumptions, for example, that the variable of user interest is well correlated with a predicted monthly or seasonal mean quantity.

Development of tailored, sector specific, forecasts is a topic of active research collaboration between GPCs-LRF and various sectors (for example, the European energy industry). To fully explore the potential for tailored products, it is recommended that RCCs and NMHSs seek close collaboration with GPCs-LRF to access high-frequency forecast data (for example, daily outputs) which may offer forecast products, such as the number of wet days within a season (rain-day frequency), with similar or greater predictive skill than calibrated seasonal mean precipitation (Moron et al., 2007). Discussed below are some examples of tailored products using GPC-LRF mandatory outputs and statistical facilities available in CPT:

Rain-day frequency: Rain-day frequency is of high interest to the agriculture sector as it is perhaps the simplest proxy to the frequency of wet/dry spells. Probability forecasts of rain-day frequency may be generated using CPT-CCA by retaining the same predictor field used to calibrate forecasts of the seasonal mean (typically predicted SST or precipitation over a large-scale domain) and replacing the predictand with a historical time series of rain-day frequency. Time series of number of days with rain exceeding a threshold (for example, 1 mm) are readily constructed from datasets and online processing available at the IRI Data Library.

User application variables: If a strong correlation can be demonstrated between a predicted seasonal mean quantity and a user application variable, tailored forecasts may be generated that are expressed in terms of the user variable. For example, reservoir inflow may be well correlated with seasonal precipitation totals, particularly when the reservoir's catchment is relatively large. If a sufficiently long time series of the user variable is available, the calibration procedure may be used to relate large-scale predictors to fluctuations in the user variable. An example, which provides forecasts of reservoir inflow in cubic meters/sec/month, generated using the Principal Components Regression facility in CPT (see sections 2.4 and 1.5.1), is provided in Figure 2.6. Further examples may be found in Coelho et al. (2006*b*) and Sahu et al. (2017).

Probabilities of exceeding/not reaching threshold values: Tercile or other quantile thresholds may be readily replaced with absolute values in CPT processing. These thresholds may be selected to reflect user interest. Interest in such tailoring is high in the agriculture sector, where thresholds may be set at the minimum rainfall requirements for certain crops.

Predictions of Standardized Precipitation Index (SPI): The disaster risk reduction sector and users sensitive to drought may welcome a forecast service that each month updates the probability of a nine-month SPI value (for example) falling below a threshold by the close of the period. CPT enables calibrated precipitation forecasts to be appended to the observed precipitation figures to generate a predicted SPI for the nine-month period that can be updated on a monthly basis. As for the other examples, the input predictor may be the GPC-LRF forecast of a variable and spatial domain found to have optimum prediction skill. Refer to WMO's *Standardized Precipitation Index User Guide* (WMO-No. 1090) for details on the background of the index, its introduction and a related description, and see section 8.2 and the link for the Caribbean Climate Outlook Forum (CariCOF).

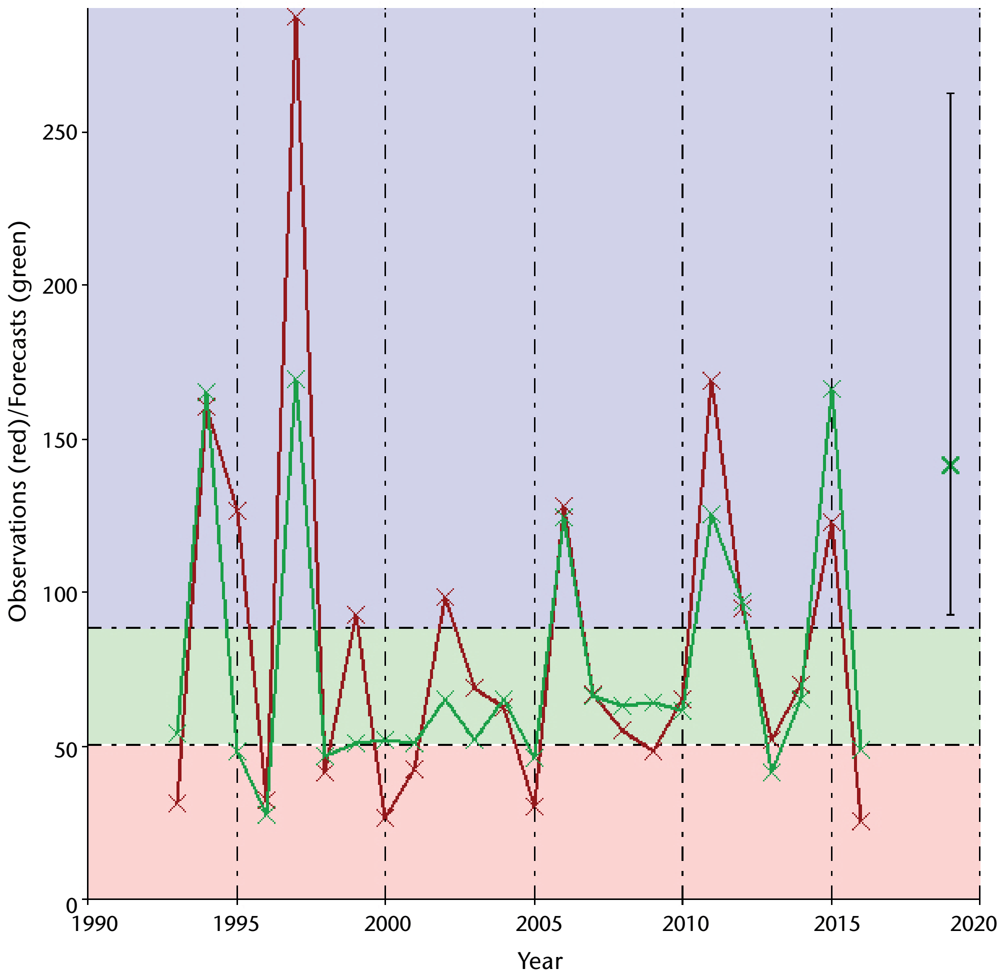

Figure 2.6. October–December inflow (cumecs/month) into the Masinga Reservoir, Kenya, observed (red) and hindcast (green), with a prediction for 2019 including a 70% confidence interval. The prediction was generated with CPT using Principal Components Regression (PCR) to relate predicted Kenya-wide temporal and spatial precipitation variability (from GPC-LRF ECMWF) to observed Masinga inflow 1993–2016. (Courtesy of J. Muhindi and W. Ochieng)

Rainy season onset: A long-standing demand from various application sectors (for example, agriculture and hydropower production) is the availability of rainy season onset predictions. Recent studies have explored the feasibility of probabilistic rainy season onset predictions (Vellinga et al., 2013; Drosdowsky and Wheeler, 2014; MacLeod, 2018; Coelho et al., 2017, Moron et al., 2009). The findings of these studies are encouraging, with (modest) discrimination ability identified in the investigated prediction, and suggest that there is the potential to tailor daily outputs from seasonal forecast models to generate such predictions. Collaboration among GPCs-LRF, RCCs and NMHSs is key for advancing tailored product developments.

2.6 Space-time aggregation and target variables

Very often, spatial aggregation can enhance the skill of the seasonal average, especially in homogeneous climate zones and when the variability of the quantities of interest is driven by regional and/or large-scale phenomena, rather than by meso- and micro-scale phenomena. Spatial aggregation tends to enhance skill because it reduces the noisy character of the local and unpredictable component of variability.

Homogeneous climate zones for spatial aggregation can be identified using different methods, such as principal component analysis involving the empirical orthogonal functions of the variable of interest, or clustering algorithms. Once the homogeneous regions have been identified, averaging over such regions (for example, for temperatures) tends to increase seasonal forecast skill.

In a similar way, temporal aggregation tends to enhance skill at seasonal timescales at the expense of the loss of temporal granularity, but there are locations and cases for which usable skill can be found even at monthly or bimonthly timescales.

Spatial and temporal aggregation can work better for certain variables and even for certain thresholds or percentiles of such variables. As mentioned earlier, the most frequent predictands, or target variables, in seasonal climate forecast systems are accumulated precipitation and average surface temperature, but streamflow and SPI (that is, quantities that integrate information over large spatial and/or long temporal scales) are also good candidates.

It is important to bear in mind that even though temporal and spatial aggregation may increase skill, they may not always be the most desired approach. For example, homogeneous climate zones exhibiting high predictive skill might not fit administrative or governance regions that are relevant to decision-makers; similarly, key climate services for agriculture, hydrology and other sectors require disaggregated temporal information, for example, monthly, weekly or daily time series for the location of interest.

Although by far the most common statistics considered for any predictand are the average and sum (accumulation) over a given period, other alternatives can also be considered. For instance, regarding precipitation, there is evidence (for example, Moron et al., 2007) suggesting that precipitation frequency – normally controlled by large-scale drivers – tends to offer higher predictive skill than intensity, which is usually noisier due to its relation to local phenomena. For decision-makers, providing information on precipitation characteristics (for example, frequency of rainy days or onset/demise/duration of rainy seasons) on top of the more traditional total precipitation is crucial as it offers a better understanding of the expected behaviour of the variable.

Extreme values and other thresholds in both tails of the variable's PDF are also of great interest to decision-makers but tend to be difficult to forecast, especially at the local scale. Especially in the case of extreme events, spatial aggregation tends to increase the associated predictive seasonal skill for the reasons explained above. Some examples for south-eastern South America can be found in Muñoz et al. (2015; 2016).

2.7 Statistical and dynamical downscaling of real-time forecasts

Downscaling is the process of increasing the spatial and/or temporal resolution of the original variable of interest in the forecasts in order to provide more detailed forecast information for use in various applications. Downscaling can be performed using empirical/statistical approaches (basically the approaches described in section 1.5.1), dynamical models, usually known as regional climate models, or hybrid methods involving both methodologies. However, dynamical downscaling is unable to enhance forecast skill over regions for which the global dynamical model used to force the regional model demonstrates poor skill. In contrast, statistical methods involving spatial pattern correction can potentially improve regional skill in cases where skill in representing climate variability patterns is degraded by model errors.

Since present dynamical models cannot yet represent the smallest spatial scales of climate variability, downscaling may be desirable. For example, even if a model could represent the climate averaged over the region defined by a gridbox, in the presence of complex topography, it would probably fail to represent the actual microclimates within the gridbox. However, if high resolution and quality observational data are available, statistical downscaling can be performed to provide seasonal forecasts for those locations not resolved by the coarse resolution model. In general, downscaling may be worthwhile in instances where the forecast being downscaled has skill and there is some reason to expect that downscaling will further add to skill (and will provide further details).

The development of several kinds of climate services, for example, in hydrology or agriculture, also requires that seasonal forecasts be provided at higher temporal resolution, for example, daily timescales, than the mandatory GPC-LRF model output. Although it is not possible to offer useful climate forecasts for any given day within the season, there exist multiple statistical methods to disaggregate seasonal climate forecasts. This temporal statistical downscaling or disaggregation consists of computing sub-seasonal characteristics that are statistically consistent with the seasonal prediction (Mason, 2008) and capturing the spatio-temporal characteristics and correlation structures among the variables of interest (Ailliot et al., 2015). Typical methods include a wide range of “weather generators” (Han and Ines, 2017) that use different approaches to condition sub-seasonal characteristics on seasonal forecasts; some examples include Markovian models (Wilby et al., 2002), hidden Markov models (Robertson et al., 2004) and weather-type-based approaches (Moron et al., 2008; Muñoz et al., 2015).

A possible alternative to statistical downscaling is to downscale seasonal forecasts dynamically, using a relatively high-resolution regional climate model constrained by global model outputs on the boundaries of the regional model domain and sometimes the larger scales of its interior as well. This procedure is used extensively to downscale global climate model projections of future climate under the WCRP’s Coordinated Regional Climate Downscaling Experiment (CORDEX) initiative (Georgi and Gutowski, 2015).

Dynamical downscaling has not been widely applied to seasonal forecasting, although several research studies have examined its possible effectiveness. These studies have found, inter alia, that for seasonal forecasts of precipitation in East Africa, errors in the teleconnection between ENSO and precipitation in the global model also propagate into the regional model, and no clear indication for improved skill is found (Nikulin et al., 2018). Similarly, downscaled summer temperature forecasts for Europe have shown no greater skill than the global model (Manzanas, et al., 2018). A general property of dynamical downscaling is that while the downscaled forecast fields show increased detail, particularly in topographically complex regions, dynamical downscaling does not correct large-scale errors in global model forecasts.

Some potential advantages of dynamical downscaling are that:

- It is objective in that it does not require a set of predictors to be selected, as is the case for statistical downscaling;

- Because it does not require training based on past observations, it may better represent situations for which there is no precedent in the observational record;

- Regional models sharing the same physical formulation and driven directly by the outputs of a host global model could straightforwardly downscale that model's forecasts operationally (Scinocca et al., 2016);

- Even if skill is not enhanced, the addition of small-scale details, particularly in regions of complex topography and in the vicinity of coastlines and large lakes, could make the downscaled forecasts more credible to users. Regional models, moreover, may be better suited to driving distributed hydrologic models, which require gridded inputs and dynamically consistent forcing fields.

On the other hand, disadvantages of dynamical downscaling compared to statistical downscaling include the following:

- The computational expense of dynamical downscaling is far greater and may rival or exceed that of global model-based seasonal forecasts;

- Dynamical downscaling, unless carried out at the operational centre that runs the driving model, requires transferring large volumes of high-frequency model data (typically three-hourly) between centres in order to drive the regional model; depending on the resources required, this may be a prohibitive endeavour;

- The hindcast ensemble, as well as the real-time forecast ensembles, must be downscaled, considerably increasing computational resource requirements;

- There is no reason to suppose that a single downscaled model can exhibit greater skill than can be achieved by combining global forecast models, and the computational and logistical demands are magnified if multi-model forecasts are to be downscaled;

- Regional climate model outputs are also biased, requiring further statistical bias corrections.

Given the great amount of resources required for dynamically downscaled seasonal forecasts, resources which are largely unavailable in various parts of the world, and given that the use of this technique is not established as a credible practice, its use to develop regional forecast information is currently not recommended.

2.8 Verification of real-time forecasts

Another fundamental component of a seasonal prediction service is the quality assessment of real-time forecasts. As important as producing forecasts is, the task of evaluating how good the real-time forecasts have been is of equal importance. It is recommended that a verification assessment of real-time forecasts should be performed in addition to establishing the average predictive skill of the seasonal forecasting system (based on the accompanying hindcasts, section 2.2.1).

With the availability of about 20 years of seasonal forecast outlooks (consensus-based predictions produced at RCOFs and National Climate Outlook Forums (NCOFs), usually in the form of tercile category probability maps) produced in various regions and countries around the world, one can perform a historical forecast quality (verification) assessment. Such an assessment is useful to provide to seasonal forecast users as an evaluation of the service provided over these years (Mason and Chidzambwa, 2009; Coelho, 2013). Note that some caution should be given to these estimates because the forecasts may have progressively improved over the evaluation period, making the long-term skill an underestimate of current capability, thereby potentially discouraging users.

In order to perform such a real-time forecast assessment, the following points should be kept in mind:

- An adequate historical set of observations is required in order to compare seasonal forecast outlooks with the corresponding observations. It is particularly recommended to use observational datasets which are consistent with the outlooks being verified. For example, for the verification of regional outlooks, it is highly recommended that a gridded dataset be used. In addition, the anomalies from the verification dataset should be computed over the same period as the one used for the forecast.

- A digital record of the previously produced seasonal forecast outlooks is required. If this record is not available, it is recommended that the existing outlooks be digitized.

- An adequate interpretation of the produced forecast outlooks is required. For example, are the seasonal outlooks valid for regional (spatial) averages, or for specific locations/stations? Having precise prior knowledge about how the outlooks should be interpreted is fundamental for an appropriate outlook quality assessment.

For procedures and methods for verifying real-time seasonal forecast outlooks, the *Guidance on Verification of Operational Seasonal Climate Forecasts* can be consulted. This guidance document provides recommendations on how to verify (probabilistic) seasonal forecast outlooks produced for previous years, including both past (historical) forecasts and individual (single) forecasts (for example, the outlook produced for the last year). It emphasizes the importance of using various complementary skill measures for a complete examination of outlook forecast quality. The most fundamental attributes recommended to be assessed are discrimination, reliability, resolution, sharpness and skill. The above-mentioned WMO CCl guidance provides detailed information about how to examine these attributes and the corresponding recommended scores. As it is well established that probabilistic verification information can only be adequately interpreted when computed over a large sample of past forecasts, individual outlook forecast verification should therefore be interpreted and communicated with caution.

2.9 Forecast reliability and its implications for seasonal forecasts

The quality of short- and medium-range forecasts, typically spanning a few days, has improved notably in recent years as observations and Numerical Weather Prediction (NWP) models have improved. These forecasts provide significant information and resulting intelligence for users, and their use in decision-making is relatively straightforward because of high levels of skill (that can be attached to forecasts and warnings). In contrast, there remain significant limitations with the low average skill of seasonal forecasts as future climate is not always well conditioned by climate drivers such as ENSO (section 1.2.1), and the resulting uncertainty in outlooks is often high. These limitations mean that the proper characterization of forecast skill and uncertainty is critical for users. In the context of decision-making and making a prior assessment of the benefits of using seasonal forecast information, forecast reliability plays an important role.

The word “reliable”, as applied to seasonal forecasts, has a quantitative meaning beyond generic confidence or trustworthiness. A forecasting system is said to be reliable if the forecast probabilities for an event agree with the observed frequency of occurrence of that event. For example, if a forecast of tercile one precipitation is 60%, reliability means that for these forecasts, the observation will also show 60% of outcomes falling into the tercile one category. It is possible to calculate the reliability of a single forecast over large areas (for example, a hemisphere) since space and time, with sufficient degrees of freedom, are somewhat interchangeable. For smaller regions such as a nation, the reliability of a series of forecasts can be quantified in a meaningful way with the 20- to 30-year hindcast sets typically available from coupled climate models.

Unfortunately, an accurate measurement of reliability requires large samples, and as discussed, it is only viable to measure reliability by pooling the forecasts in time and (usually) across space. The *Guidance on Verification of Operational Seasonal Climate Forecasts* describes in detail the measurement of reliability, including suitable means for its quantification, using several skill scores or their decomposition. SVSLRF describes a range of suitable verification practices which may provide information about reliability.

For more detailed diagnostics of forecast quality, reliability diagrams are recommended. A construction of a reliability diagram involves binning forecasts by probability category (on the x axis) and plotting these values against the observed frequency outcome on the y axis. Information about the spread and frequency of forecasts/outcomes is also shown in the form of accompanying frequency histograms and/or a scaling of dot sizes to reflect the population size in each category (Figures 2.7 and 2.8). The histogram of past forecasts provides information about the sharpness of forecasts (that is, it addresses whether the distribution of seasonal mean outcomes significantly differs from the climatological distribution from one forecast event to another). It is important to keep in mind that reliability diagrams do suffer significant sampling issues as forecasts and outcomes are binned to create matching outcomes. This can create substantial noise in results when the verification period or spatial domain is not large.

Reliability is an important measure for several reasons. The first is the close relationship between forecast reliability and forecast value (Wilks, 2006), noting that reliability measures the extent to which forecast outcomes can be "trusted". Forecast value (and usability) is eroded rapidly if a forecast lacks reliability (Kumar, 2010) and particularly if the forecasts are too emphatic (insufficient spread or dispersion of ensembles). This is because users will tend to attach too much confidence to decisions and overreact. Fortunately, deficiencies in reliability can be addressed through calibration, as discussed in section 2.2.2, and the design of forecast systems (increasing the ensemble size, ensemble spread, and so forth) or by the use of multi-model ensembles, which tend to be more reliable than single models (see, for example, Hagedorn et al., 2005). Statistical calibration methods to improve overconfidence or underconfidence come at the cost of reduced discrimination (less spread in forecasts or a narrower range of forecast probabilities), as depicted in Figure 2.3, for example.

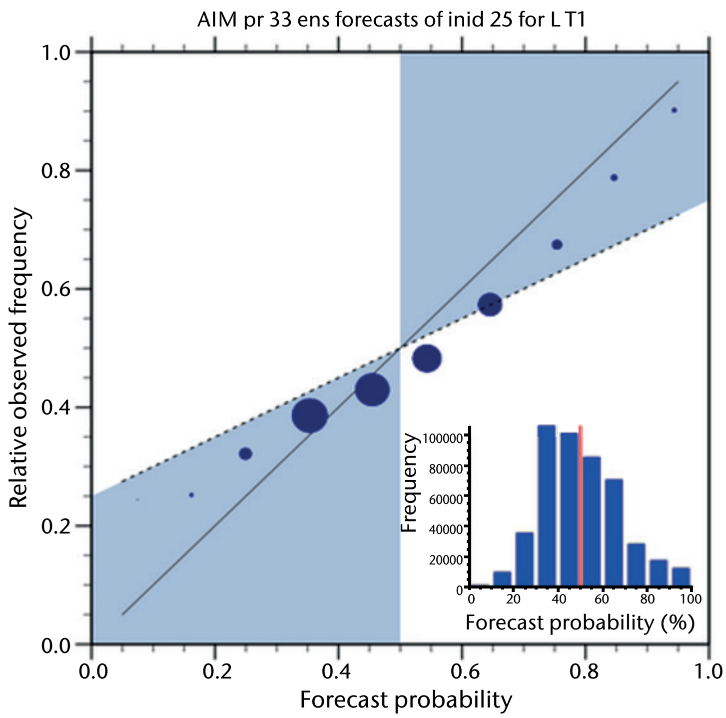

Figure 2.7. Example of a reliability diagram for an operational climate forecast system for above and below median seasonal rainfall calculated for the whole of Australia. Where points lie within the blue shaded region, the corresponding subset of forecasts contributes to a positive Brier Skill Score.

Source: Hudson, 2017

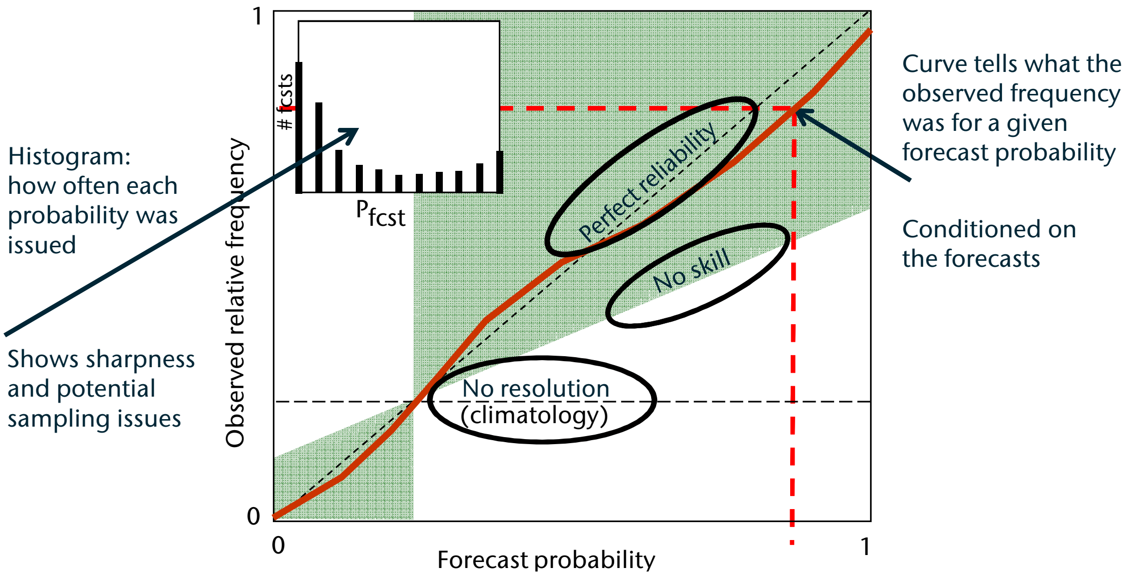

Figure 2.8. The interpretation of an idealized reliability diagram

Source: Hudson, 2017

A reliable forecast can simplify the use of seasonal forecasts as users can take the probabilities of forecast outcomes on face value and act accordingly. A common question asked by users is "How confident can I be in a probability forecast?". Such a question touches on the uncertainty that is already built into the probabilities of forecast outcomes; if those probabilities are reliable, they can be used directly in the decision-making context. A reliable forecast has the added advantage of making it clear when the use of a forecast is likely to be most valuable. For example, large probability shifts in a reliable forecast system indicate significant conditioning of observed outcomes on climate drivers. In a practical sense, this is seen in countries such as Australia, where ENSO is a strong modulator of climate and tends to dominate predictability during winter and spring. In the presence of an El Niño or a La Niña, probabilities for above and below median tend to depart significantly from their climatological value.

CHAPTER 3. SEASONAL FORECAST PRODUCTS

3.1 Reference period and prediction of seasonal anomalies

Real-time forecast anomalies are customarily defined with respect to a standard base period such as 1981–2010 or with respect to the base period spanned by the hindcasts. This allows context for real-time seasonal prediction anomalies to be set relative to a “climate normal” over a selected period. However, forecast anomalies with respect to a shorter and a more recent reference climate might be more relevant in the context of climate change. Temperatures have changed sufficiently over the last 35 years for seasonal mean values to be more frequently warmer than during the early years of the hindcast period. A probabilistic statement about the next season may be correct relative to the 1981–2010 period but of limited use to a typical user wanting to know what to expect relative to a more recent past. Making seasonal forecasts in a period with a strong trend is a challenge, and although the warming trend of the climate is an important effect that should not be discarded, predicting year-to-year variability with respect to a recent normal can also be important. Such considerations may apply to other climate variables also.

3.2 Forecast categories

Based on the availability of historical observations at a location, one can quantify the range of seasonal mean variability of the climate variable one wishes to forecast. The range of historical values observed for the December-January-February (DJF) seasonal mean temperature can be used to estimate the PDF of the variability of DJF seasonal mean. The PDF of the variability can also be used to find the temperature values that divide the temperature range into various categories, for example, terciles. Using a simple procedure, temperatures over a thirty-year period can be arranged in ascending/descending order, and boundaries such that one third of the observed temperatures are below or one third of the observed values are above can be found. Following this procedure, data over thirty years can be partitioned into three groups, and terciles are the two values that divide the thirty-year historical time series into three parts, leaving one third of the observed temperature seasonal means below the “lower tercile” value, one third of the values above the “upper tercile” value and one third of the values between the lower and upper tercile values.

Once the terciles for the below and above normal categories have been established, these terciles also define the boundaries such that in the absence of any forecast information, the probability for the seasonal mean outcome to be in each category is one third; such forecasts are referred to as climatological forecasts or forecasts with equal chances. For a season, a source of predictability (for example, the presence of SST anomalies associated with ENSO) can change the odds of the seasonal mean being in different categories, and seasonal forecasts are cast in terms of probability anomalies. It should be noted that tercile boundaries are a function of season and location, observational data used and the period that is utilized to define the climate normal. Terciles can also be computed for dynamical model-based forecasts. Because of biases in models, the tercile for model forecasts can be different from observations, and as discussed in section 2.2.2, appropriate bias correction and calibration procedures are recommended for error adjustment.

3.3 Deterministic seasonal forecasts

For dynamical seasonal prediction systems, predictions for the forthcoming season are often presented in terms of ensemble mean anomalies. Over each grid point, the three-month mean anomalies for any forecast variable are computed by averaging over the forecast outcome of each member of the ensemble. The ensemble mean anomalies highlight the differences between the forecast ensemble mean and the model "climatological" mean (estimated using the hindcast). Therefore, although the ensemble mean appears as a single value and sometimes is referred to as “deterministic”, it is still the outcome of a probabilistic ensemble prediction system. The ensemble mean, being just an average of many forecast realizations, is not a forecast outcome per se but represents an average of several possible forecast outcomes.

The ensemble mean provides a simple indication of the amplitude of the forecast anomalies. Sometimes, ensemble mean anomalies are presented in combination with statistical tests aimed at identifying regions where the forecast signal is considered not significant due to sampling (for example, because of limited ensemble size or a small signal-to-noise ratio). In such cases, when the signal is not considered statistically significant, the ensemble mean values are “masked out”. In addition, estimates of skill based on hindcast performance are also sometimes used to mask anomalies over regions where typically there is limited skill. Hindcast-based skill masking of seasonal forecasts provides users with guidance about the regions where, over a historical period, the skill of the forecast system has been low (due to either lack of predictability or errors in the forecasting system). Unless forecasts are calibrated and demonstrate good reliability, it is a good practice to show the forecast for a given season, variable and lead time in conjunction with the skill for that season, variable and lead time.

Deterministic categorical forecasts can also be constructed by displaying the tercile category (for example, as defined in section 3.2) that is implied by the ensemble mean anomaly or the category that is predicted with the highest probability. This type of “categorial” forecast provides no information about either the magnitude of the ensemble mean anomaly or the confidence with which the category is predicted, and it is recommended that this format should be avoided.

3.4 Probabilistic seasonal forecasts

Forecast probabilities are typically expressed for three-month averages and are based on forecast probability anomalies for different categories, as in section 3.2. As mentioned in section 1.4, a way to convey the uncertainties inherent in climate predictions is to use probabilities, which can be computed for different parameters and relevant thresholds. Thresholds can be based on predefined values of interest to users or on the statistics of the observed or models’ (hindcast) probability distribution.

In the simplest formulation, probabilities are derived from the number of ensemble members in the real-time forecast displaying an anomaly within the categories defined from the model climatology. As the next step, calibration can attempt to make such a forecast more reliable (section 2.2.2).

Displaying forecast probabilities in conjunction with the climate threshold derived from the observed climate can help with the interpretation of the probabilistic forecast. As for the ensemble mean deterministic forecast, forecast probabilities should be displayed in combination with the relevant probabilistic skill estimates.

3.5 Predicting the probability density function of seasonal mean outcomes and the probability of exceedance

Dividing the climatological forecast distribution into tercile categories and casting seasonal forecasts in terms of probabilities corresponding to each tercile has the disadvantage that information about the full statistical distribution of the variable is lost and the probabilistic forecast information is available only in the context of tercile values. Some users are interested in knowing the change in the probability that the seasonal means will exceed a certain threshold, which may not be the same as the tercile category. For example, in order to decide which seeds to use during the next season, a farmer may be interested in knowing the probability that the mean temperature at his farm will exceed 30°C; if the boundary corresponding to the upper tercile is 25°C, the information required cannot be obtained from seasonal forecasts based on terciles alone. To alleviate this issue and provide seasonal forecast probabilities associated with user-defined threshold values of a variable, the seasonal forecast can be cast to depict the full probability density function. This is achieved following the approach of providing seasonal forecasts in terms of the probability of exceedance (POE).

The concept of POE is related to obtaining the climatological cumulative distribution function (CDF) for the relevant climate parameter. The observed (climatological) and forecast CDFs can then be compared, and the POE for any threshold value can be obtained.

Using a parametric approach, the climatological CDF is generally derived by estimating various parameters associated with the PDF of seasonal statistics, for example, the mean and the spread (standard deviation), based on observational data over which the climate normal is defined. The curve is a "fitted" curve because it is defined using a parametric estimate to construct a smooth curve fitted to the observational data. The data may not be so smooth and regular, but the parametric estimate only uses the mean and the typical standard deviations from which the CDF is defined. A similar CDF is obtained for the forecast for the season under consideration. A comparison of the two CDFs can then provide the anomalous probabilities corresponding to any threshold value desired. The use of the POE in seasonal forecasting is commonly known as "predictions in flexible format". For more information see discussions at the Climate Prediction Center (CPC) and IRI sites.

3.6 Interpreting probabilistic seasonal outlooks

A longstanding communication issue with probabilistic seasonal forecasts is their interpretation and how are they used in the decision-making process. Probabilistic forecasts may be perceived to fail whenever a category other than the most probable one is observed. (This is even more of an issue for deterministic forecasts that specify a tercile or other category without stating its probability, which is one reason such forecasts should be avoided.)

It is perhaps human nature to interpret "misses", particularly recent ones, as indicating poor forecast performance, and even sophisticated forecast producers and users sometimes feel such concerns. However, it is crucial to keep in mind that for a reliable forecast system, even categories that are predicted with a low probability of, for example, 10%, should indeed occur 10% of the time, and further, the observed outcome will not always be in the most probable category.

A prominent example of when an unlikely verification spawned concern about the efficacy of seasonal forecasts occurred during the large El Niño of 2015/2016. California (which was in the midst of a multi-year drought) was expected to receive enhanced precipitation during a strong El Niño event (based on a small number of historical events), and so, there was strong anticipation that the 2015/2016 winter rainy season would "break the drought". Furthermore, the consensus of seasonal forecast models was "above normal" as the most likely tercile category, with positive precipitation anomalies over California in the ensemble mean. In reality, much of California, including its southern half, received below normal precipitation that winter, while in the north of the state, precipitation was near to slightly above average. The apparent discrepancy is not obviously attributable to poor forecasting of the El Niño itself since it was predicted reasonably well by the models in accordance with historical skill estimates and average precipitation response during El Niño.

This apparent failure of the seasonal forecasts spawned many theories and scientific studies about "what went wrong". However, one should also examine the null hypothesis that the observed outcome was within the range of possibilities that was forecast. Figure 3.1 shows the forecast from GPC-LRF Montreal on the left, calibrated using the method of Kharin et al. (2017) for the FMA season, which had the lowest observed percentile for precipitation that winter, at a two-month lead time for a grid location near Los Angeles in southern California. The forecast tercile probabilities were 51% for above normal, 31% for near normal, and 18% for below normal. The verification, indicated by the green dot, was near the 25th percentile of the climatological distribution and the 10th to 15th percentile of the forecast distribution. The corresponding reliability diagram, aggregated over North America for the same season and lead time, is shown on the right. It indicates that when the below normal category is forecast with probabilities of 10–20%, the verification falls in the normal category 10–20% of the time. According to the forecast, the probability of the observed outcome was thus comparable to that of a rolled die coming up "1", or of a flipped coin coming up heads three times in a row, which

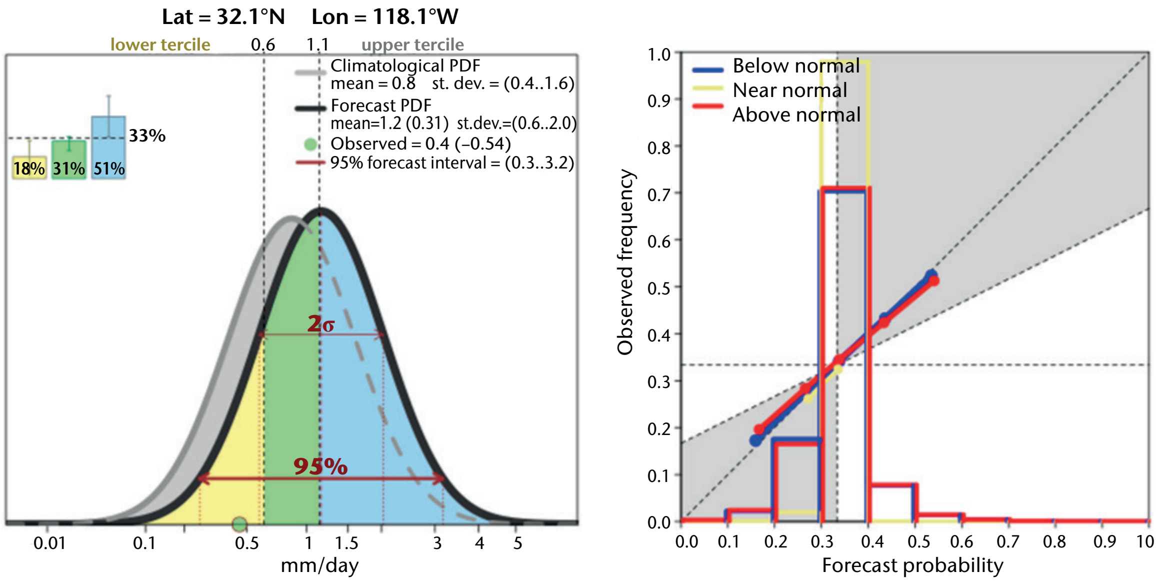

Figure 3.1. Left: Calibrated probabilistic forecast from GPC-LRF Montreal of precipitation in February-March-April 2016 at a lead time of 2 months (initialized near the start of December) for a grid cell in southern California. The grey curve indicates the climatological distribution, the dark curve indicates the forecast distribution, and the colours represent the climatological terciles. The green dot is the verification according to the GPCP2.3 analysis. Right: Reliability diagram, aggregated over North America, for the same season and lead time. (After Kharin et al., 2017)

are not particularly unusual events. A similar perspective that the dry southern California winter in the face of the 2015/2016 El Niño could have been attributable to a random realization of atmospheric internal variability is argued in Chen and Kumar (2018).

The above view should not obscure the fact that there may be instances where the perceived failure of a forecast does indeed point to a problem with the forecast system. Warning signs for failures could be (i) a verification so far outside the range of predicted outcomes even of a calibrated forecast that it should have occurred only with a miniscule probability; (ii) a number of successive forecasts that lead to improbable verifications; or (iii) a system that is consistently in strong disagreement with other systems.

One final point regarding the communication of seasonal forecasts verbally (or in text) is that the probabilistic nature of the forecast should always be emphasized even if the forecast is relatively confident. This could mean, for example, stating that “the odds are tilted towards a wet winter” (as one way of conveying probabilities in plain language), rather than “the winter will be wet”.

CHAPTER 4. GUIDANCE ON GOOD PRACTICES FOR DEVELOPING OBJECTIVE SEASONAL FORECASTS

Following the background on some basic concepts of seasonal predictions discussed in Chapters 1–3, this chapter provides practical guidance on good practices for developing, producing and disseminating seasonal climate forecasts. The recommended good practices draw on the considerations described previously covering product development, verification, contextual information and user feedback.

Recommendations for good practices in this chapter can be thought of as the practical steps in implementing the set of principles outlined in the Executive summary. Sections 4.1–4.7 provide recommendations on the infrastructure required to support the recommended seasonal forecast practice (section 4.8). Sections 4.9 and 4.10 provide recommendations on communicating seasonal forecasts and good practices for establishing credibility for the seasonal forecast providers. In each section, the text with the main recommendations is in ***semibold italic*** font.

While in the recommendations that follow, objective methods and dynamical models are an integral part of the seasonal forecast process and are highly desirable, the need for skilled climatologists and supporting professionals to interpret and to bring the information to users (using mechanisms like RCOFs) remains a critical part of the seasonal forecast process.

4.1 Catalogue and document regional climate variability and its drivers

4.1.1 *Document seasonal climatology*

As part of developing the seasonal forecast process for a region, it is recommended that first, the variability of the regional climate be understood. Understanding the variability of the regional climate, in turn, requires access to climate data, including data regarding the variables for which predictions are to be made (such as rainfall and temperature) as well as predictors (for example, SSTs). Some basic diagnostic analysis, such as analysis of simple statistical features (mean, variance, correlations, and so forth) and composite analysis with different stratifications (relative to anomalies of climate drivers) should be part of this step.

Knowledge of the local climate is important for seasonal forecasts because forecasts are expressed as anomalies and therefore implicitly refer to climatology over the region. The application of seasonal forecast information usually also involves projecting the forecast into real-world applications, and again, this requires knowledge of the regional climate.

In addition to understanding the climate system, one important objective of the diagnostic analysis is to catalogue climate features that are of societal relevance for the region of interest. For instance, documenting the onset/withdrawal dates of the rainy season over a region is crucial in order to develop forecast products that are of high interest in the context of local decision-making. Identifying important regional climate features also helps with the tailoring of forecasts at the regional level (section 2.5).

In quantifying the drivers of regional climate, documenting the influence of ENSO is an essential next step. One should also explore additional important climate drivers, such as the IOD and Sub-tropical Indian Ocean Dipole (SIOD), for the Indian Ocean regions, or the Guinean Gulf SSTs or SSTs of large bodies of water for the African continent. Their importance can be crucial, especially when ENSO, for example during the summer, may have little influence over the target region. The WMO Climate Services Toolkit provides many tools to assist with this analysis process.

Analysis of climate variability for a region can be conducted using reference datasets, such as global reanalysis products, and complemented by the availability of national and regional datasets. Such an approach can provide both large-scale and regional views for different relevant parameters and patterns of climate variability.

It is also important to analyse the simulation and forecast of regional modes of variability in dynamical seasonal forecast systems or other tools that might be used to generate the seasonal forecast. This could help in the selection of the model (see section 2.3.1) for use in seasonal forecasts for the region of interest.

4.1.2 ***Document climatology for the drivers of climate variability***

Depending on the region, the predictable component of climate variability can be small or large and could depend on the season and the characteristics of the drivers of climate variability. If the drivers of regional climate have remote origins, to understand the seasonality for potentially skilful predictions, ***one should also have a basic understanding of the characteristic evolution of regional climate variability drivers***. For example, if the driver of regional climate variability is ENSO, Niño 3.4 is the region where SST variability is the largest (standard deviation in the range of 1°C) (Figure 4.1) and follows a distinct evolutionary cycle – Niño 3.4 SST anomalies generally develop during the boreal summer and reach their peak amplitude during the boreal winter. On the other hand, if the driver of regional climate variability is the Tropical Atlantic, its variability is smaller (about half of that in the Niño 3.4 region).

4.1.3 ***Document recent trends***

Many climate variables show significant trends, mostly in the context of surface temperature but also with respect to annual precipitation (Figure 4.2). ***It is recommended that these trends be documented at the regional level as they are relevant in the context of seasonal forecasts.*** Slowly evolving climate trends may impact the local climate in many ways, for example by disrupting historical teleconnections, shifting the background to a warmer mean state or leading to a change (increase or decrease) in precipitation. While most climate models used for seasonal forecasting can incorporate observed climate change (Doblas-Reyes et al., 2006), understanding these changes is important for interpreting seasonal forecasts.

The magnitude of trends at regional levels is variable dependent and easier to discern for some parameters, such as surface temperature (Figure 4.2) and the number of cold days, than for others. Slowly evolving trends, for example, long-term decreases in snow cover and sea ice or low-frequency variability in ENSO, may also be climate drivers. In a majority of instances, trends in climate drivers are part of the initialization of dynamical forecast systems, and that information is carried over during the forecast.

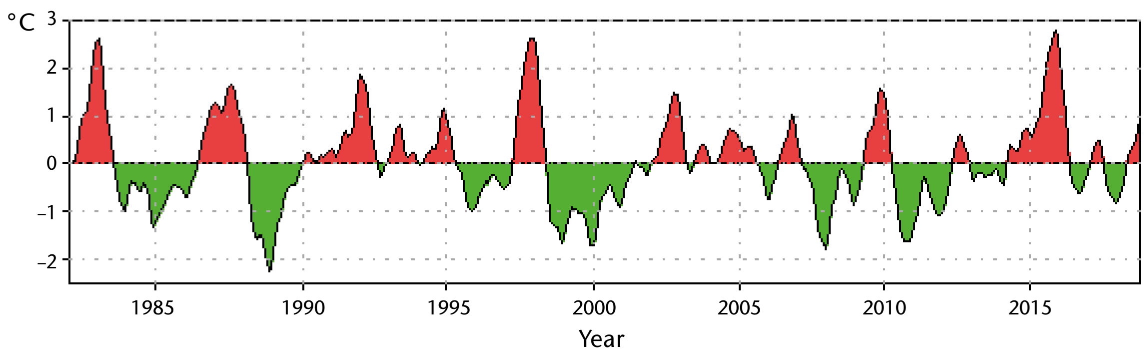

Figure 4.1. An example of the characteristic evolution of the drivers of regional climate variability: the time-series of the Niño 3.4 SST anomaly in the tropical Pacific. The temporal evolution of the Niño 3.4 SST index follows a typical cycle of 3–4 years of warm and cold SST anomalies with maximum amplitude (on occasions) exceeding 2°C.

Source: Climate Prediction Center

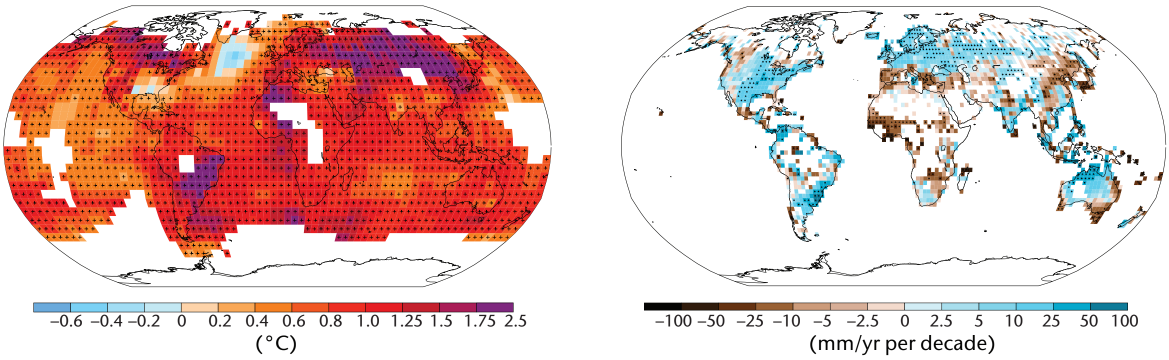

Figure 4.2. Left: map of the observed change in surface temperature from 1901 to 2012; right: map of the observed change in precipitation over land from 1951 to 2010.

Source: Intergovernmental Panel on Climate Change (IPCC), 2014: *Climate Change 2014: Synthesis Report*, page 41

Climate trends clearly have impacts on users and the decision-making process and should be taken into consideration when making seasonal forecasts. However, the inclusion of trends may inflate measures of skill if the trends come to dominate the local climate. For example, over much of the planet, it is becoming increasingly easy to forecast temperature departures from the long-term average because cooler conditions are becoming increasingly unlikely. Depending on the users, it may be relevant to filter these trends when developing seasonal forecast information, for example, by providing anomalies relative to a more recent reference period as discussed in section 3.1.

4.2 Establish a schedule for seasonal forecasts

Seasonal forecasts must be received by users in time to inform their decision-making. A forecast that arrives after key climate-sensitive decisions have been made will not be used. The GDPFS requires GPCs-LRF to adhere to fixed forecast production cycles and to issue forecasts at least once a month. The requirements for RCCs are similar, with the added suggestion that shorter-term forecasts also be issued every 10 days, if possible.

To increase the likelihood of seasonal forecasts being taken into account, NMHSs are encouraged to co-design climate services with users. Some examples of the process of co-production of climate services, for a range of timescales and in the African context, are documented in the WISER-FCFA Manual on Co-production in African weather and climate services. Further, ***it is recommended that the schedule for issuing forecasts should be established in consultation with users, documented and made clearly available***. If the same service or a similar service is provided to a wide range of users, the forecast schedule may be made publicly available (as is the practice at CPC). User and public trust in the service will increase, and users are more likely to use climate products if they are confident about when, where and how the forecast products will become available.

By the same measure, it is recommended that RCCs co-design their regional forecast issuance schedule with the NMHS in their region so that the regional forecast is issued in time for NMHSs to utilize it to prepare their national forecasts, the timing of which should be driven by national users. RCCs and RCOFs can face particular challenges in setting forecast issuance times. It is important to structure the regional domain of RCOFs such that the NMHSs served have the same climatic zone and similar attributes for climate variability. Meeting all the requirements in deciding the forecast schedule may be challenging.

Flexibility in the timing of issuing seasonal forecasts can also be constrained to a degree by the monthly forecast schedule of GPCs-LRF. Currently, there are wide variations in when individual GPCs-LRF issue their forecasts, ranging from the 1st to 23rd of the month. Because RCCs and

NMHSs may be committed to issuing forecasts before all the data from GPCs-LRF is available, it may not be possible to use some models. Such considerations could be part of the model selection process at the regional level (section 2.3.1).

4.3 Review and document the performance of issued forecasts

It is important and recommended that providers of seasonal forecasts review the performance of real-time forecasts. This includes documenting the skill of issued forecasts and highlighting factors which might have influenced the observed outcomes. Documenting the skill of past forecasts can support building trust between forecast providers and users, provide an opportunity to learn about forecast performance, and improve forecast processes in the future (see section 2.8). If the seasonal forecast process is based on dynamical and empirical methods, the skill assessment of issued real-time forecasts complements the long track-record assessments based on hindcasts.

Engagement between users and producers of seasonal forecasts (section 4.7) should include a discussion of the performance of the most recently verified seasonal outlook and user decisions influenced by that outlook. This activity is a standing agenda item at many RCOFs and NCOFs. Its advantages include:

- Emphasizing the producer's "ownership" of the forecast and accountability, which are important in building trust;
- Building an understanding of the strengths and weaknesses of seasonal forecasts and the consequences of their probabilistic nature in decision-making;
- Helping users to refine their strategies in using probability outputs to inform decisions; and
- Increasing producer understanding of the challenges faced in using the probability information and potential improvements in forecast content/presentation that might assist users.

The procedure for the forecast post-mortem will necessarily vary depending on the format of the issued forecast, for example, text only, histograms of quantile probabilities (used in some cases for regions where seasonal variability has high spatial coherency), forecast probability maps, or a mixture of these.

Procedures for approaching single-forecast verification are documented in the *Guidance on Verification of Operational Seasonal Climate Forecasts*. If maps are issued, a good starting point is a simple side by side comparison of the forecast probability map and a map showing the category or percentile which was observed to occur (Figure 4.3). It is recommended to avoid comparisons that do not compare like with like. For example forecast probabilities for tercile categories should be compared with the observed tercile category, not, for example, the observed anomaly relative to the climate mean. Comparisons that do not compare like with like give ambiguous information on the success of the forecast and may be misleading.

4.4 Provide a discussion of the current state of the climate to set the context for the seasonal forecast for the coming season(s)

The impact of future climate anomalies is often dependent on current climate conditions. For example, a wet season occurring after a drought will often be viewed favourably, indicating an improvement in conditions, whereas the same season coming after a flood may make conditions worse. Understanding and documenting prevailing climate conditions, therefore, is an important component of the seasonal forecast development process and should cover both local conditions as well as those affecting the main climate drivers. ***It is thus recommended to start with a brief description of the prevailing oceanic, atmospheric and land surface conditions, paying special attention to features that are relevant at the seasonal timescale*** (see an example in Box 4.1).

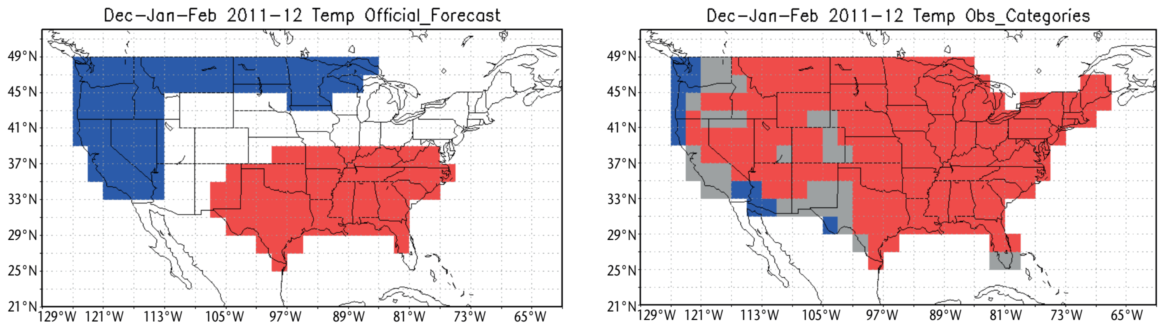

Figure 4.3. Verification of a real-time seasonal forecast at the Climate Prediction Center. At left is the category for the seasonal forecast for DJF 2011/2012 (depicted using the information gathered from the probabilistic seasonal forecast, which was reduced to a categorical forecast in accordance with the category for which the seasonal forecast probability was the highest). At right is the category for the observed anomalies. Blue is for the lower tercile; red is for the upper tercile; grey is for the normal category. For the forecast map (left panel), the white areas are the regions where the forecast was for equal chances, that is, no specific forecast information was available.

Source: Climate Prediction Center

The description of oceanic conditions may, for example, focus on the Pacific tropical regions and include details of the ENSO phase and factors contributing to its probable evolution. Depending on the region of interest, the Atlantic and Indian Ocean states should also be described as they may be an important source of predictability at the regional scale.

Atmospheric conditions related to ENSO are frequently discussed to assess the degree of coupling between ocean and atmosphere that will, to some extent, determine the near-term probable evolution. Inspection of the Hadley-Walker circulation anomalies and teleconnection patterns can shed additional light on the response of the atmosphere to the oceanic conditions. Other atmospheric variability patterns relevant to the region of interest (for example, AO/NAO, PNA pattern, and so forth) may also be monitored and discussed. Real-time reanalysis datasets can be the primary tools to aid this analysis.

Finally, land surface features such as soil moisture, snow cover, sea ice, and so forth carrying memory at the seasonal scale or affecting remote regions through teleconnections should also be discussed. Box 4.1 shows a description of the current state of the climate from the operational seasonal outlook of the United States National Oceanographic and Atmospheric Administration (NOAA) Climate Prediction Center. Note that in this example, land surface features are mainly focused on the North American region.

4.5 Provide seasonal forecasts in probabilistic format

It is recommended that operational seasonal forecasts be in probabilistic format and that accompanying forecast bulletins include some guidance for interpreting maps and other provided products. It is also recommended that the probabilistic nature of seasonal forecasts be emphasized and a description be included of the expression of probabilities used and their meaning (for example, tercile categories, probability of exceedance of some predetermined value, and so forth, see sections 3.4 and 3.5).

Figure 4.4 shows an example of a probabilistic tercile-category forecast, provided by the LC-LRFMME multi-model, which summarizes the forecast for all three categories in a single map. It represents the probabilities for the most likely tercile with different ranges of colours. This type of representation, combining information of the three categories defined by tercile values, has become standard for many centres providing seasonal outlooks.

Box 4.1. Example from the prognostic discussion for long-lead seasonal outlooks at the National Weather Service (NWS) Climate Prediction Center, College Park, Maryland, USA, issued at 8.30 a.m. EDT, Thursday 18 October 2018 (for the text discussion of the latest forecast, see http://www.cpc.ncep.noaa.gov/products/predictions/90day/fxus05.html)

Current atmospheric and oceanic conditions

Starting with ocean conditions, we continue to remain in ENSO-neutral conditions through early October and the El Niño watch continues. In recent weeks, however, observations across the global tropics have trended toward more favorable conditions for development of El Niño in the next couple of months. Sea Surface Temperatures (SSTs) are above-normal across most of the equatorial Pacific basin, generally ranging from +0.5 – +1.0 degrees C. The most recent weekly values of the Niño4 and Niño3.4 indices are now +0.9 and +0.6 degrees C respectively. This somewhat rapid change was considerably aided by the development of a Westerly Wind Burst (WWB) and evolution of an ongoing MJO event and the associated westerly low-level wind anomalies. A strong downwelling oceanic Kelvin wave was initiated as a result of this significant weakening of the Pacific basin trade winds due to these events and continues to shift eastward across the Pacific basin at depth. Accordingly, the equatorial upper-ocean heat content has increased considerably during September and early October making the development of El Niño somewhat more likely. Specific Niño3.4 SST forecasts are reviewed in the section below.

The atmospheric conditions associated with a developing El Niño event remain modest at best. Although westerly low-level wind anomalies are now present, any organized areas of enhanced convection are not yet present. This is most likely, however, in part due to the counter-acting effect from large scale subsidence associated with the suppressed convective phase of the MJO.

Turning to current terrestrial conditions, we note much above-normal soil moisture anomalies from Texas to the Midwest, eastward to the Great Lakes, mid-Atlantic and parts of the northeast, while drought conditions persist across many areas of the west. As we are entering the late autumn and winter months, soil moisture conditions, although near the 99th percentile in many areas, is not expected to contribute any significant impact to the climate conditions over the next several months. Snow cover anomalies (both negative and positive) for September and the first half of October are important to note. Snowfall over areas of Eurasia and Alaska are currently below-normal in many areas, but snowfall has been substantially above-normal for portions of central and eastern Canada.

In addition to terciles, other features of the forecast PDF can be provided, for example, values of a variable associated with key user-defined probabilities (for instance, 10%, 50%, 90%) or probabilities of exceedance associated with specific threshold values for a variable (see also section 3.5).

IRI provides a full estimate of the probability distribution showing interactive maps and point-wise distributions of exceeding (or not exceeding) user selected percentiles (see http://iridl.ldeo.columbia.edu/maproom/Global/Forecasts/NMME_Seasonal_Forecasts/precip_full.html). Figure 4.5 shows the probability of exceedance of a selected value of rainfall for a chosen region and the forecast PDF and probability of exceedance over a selected point in the region (climatological values are also plotted as a reference). A specific threshold (in steps of five percentile points) can then be selected. The user can also specify a quantitative value in physical units (for example, seasonal total rainfall in mm) for the probability of exceeding or not exceeding that threshold.

It is important to underline that using the probability of exceeding some predetermined value, even for a forecast with only climatological PDFs, could also be beneficial to users, as frequently, they may be unaware of the climatological probabilities and values.

As for many users, the interpretation of the maps expressing probabilities is not straightforward, it is also a good practice to include some standard text in the forecast bulletins with a short description of how probabilities have been computed and their meaning (see https://www.climate.gov/news-features/blogs/enso/betting-climate-predictions for a discussion of why probabilistic forecasts may be misinterpreted). A further discussion on including an interpretation of forecast probabilities as part of the seasonal forecast dissemination is included in section 4.9.2.

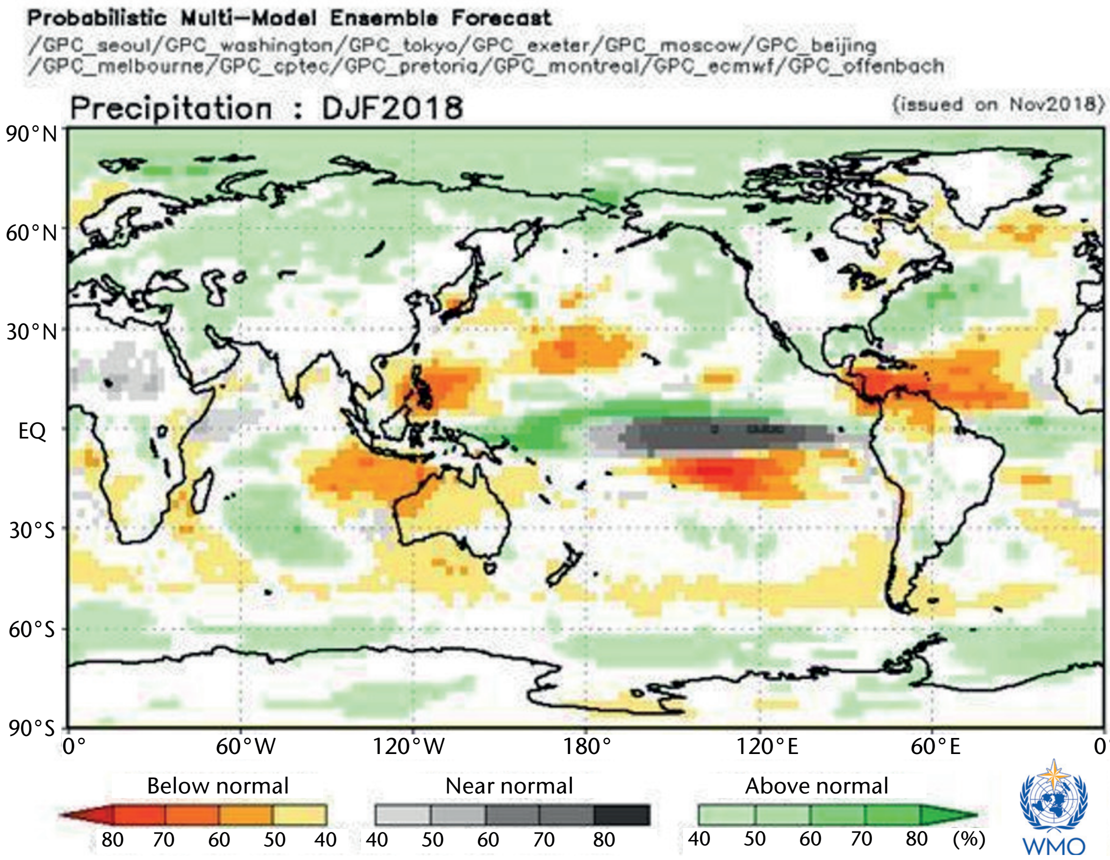

Figure 4.4. Probability of the most probable tercile from the LC-LRFMME multi-model for 2019 winter precipitation (DJF), issued in November 2018, with white indicating "equal chance" and the probabilities of all terciles below 40%

Source: LC-LRFMME, https://www.wmolc.org

4.6 Provide a discussion of the physical basis for the seasonal forecast

Seasonal forecasts are based on a multiplicity of sources, for example, dynamical models, understanding the influence of climate drivers, and so forth, and are expressed in probabilistic terms showing the shift with respect to the climatology. ***It is a commonly accepted good practice to provide a physical explanation (sometimes referred to as a forecast attribution) for this shift.*** Starting from the current state of the climate expressed in terms of the relevant variability patterns over the region of interest, the analysis of known drivers at the seasonal timescale, teleconnections, and model outputs frequently allows the outlook to be described in terms of climate drivers. A "forecast attribution" can be useful in helping users interpret why a forecast shows a particular shift and can increase their confidence in using the probability forecast.

The example of a forecast attribution shown in Box 4.2 corresponds to the outlook for November 2018 to January 2019 issued in October 2018 by Regional Association VI (RA VI) RCC-LRF (Toulouse). As NAO is the pattern which explains most of the winter climate variability over Europe, the text provides a physical explanation of the effect of known drivers on the NAO phase. Other relevant patterns (East Atlantic (EA) and Scandinavian (SCAND) patterns) modulate the effect of NAO. The shifts in temperature and precipitation probabilities are explained in terms of these variability patterns.

Note that the text providing interpretation is an attempt to explain in physical terms a final product that was produced by combining information from different sources. Although the final

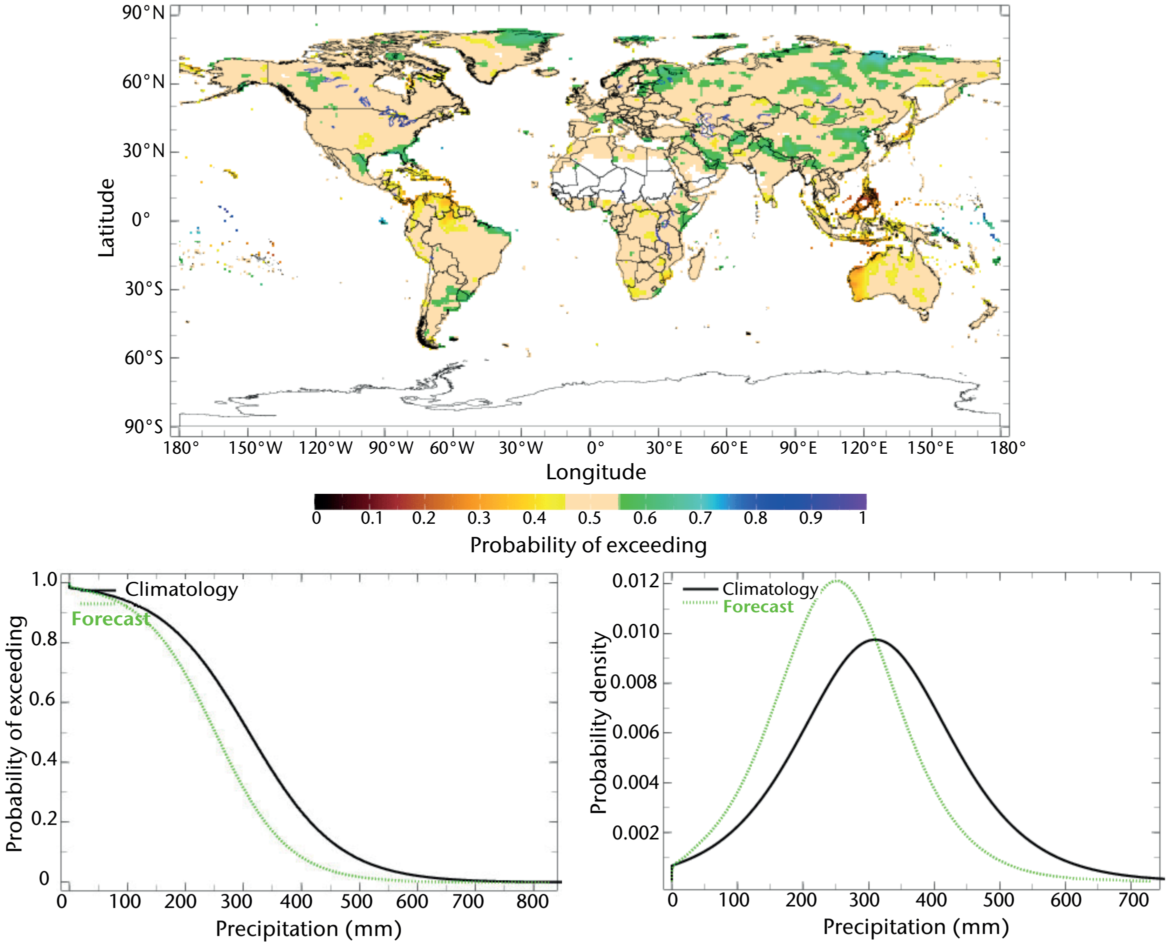

Figure 4.5. Top: winter seasonal precipitation forecast probability (December 2018 – February 2019) of exceeding the 50th percentile of the distribution from historical 1981–2010 climatology, issued in November 2018; (bottom left) probability of exceeding precipitation values for a cursor-selected location showing forecast (green) and climatology (black); (bottom right) same as the bottom left panel but for a probability density function

Source: IRI, http://iridl.ldeo.columbia.edu/maproom/Global/Forecasts/

product may have been produced objectively, the discussion providing an interpretation in terms of known drivers, variability patterns and teleconnections is inherently a subjective process and may have limitations.

4.7 Establish feedback mechanisms and engagement with users

Climate information only delivers benefits for users when it is accessible, understandable and actionable. While climatologists and specialists within NMHSs are well placed to develop climate services, ensuring that these services are useful requires regular engagement with users and feedback mechanisms. ***It is thus recommended that as part of the seasonal forecast process, feedback mechanisms be put in place and users be engaged.***

The *Guidance on Good Practices for Climate Services User Engagement* (WMO-No. 1214), prepared by the WMO Expert Team on User Interface for Climate Services, provides a practical guide to assist with user engagement and feedback. User engagement is also important for ensuring that user desire for products is reflected in service plans and priorities.

Good user feedback mechanisms involve matching user needs with scientific and technical capabilities in a way which is often iterative. Established methods for user engagement can be divided into a number of categories, each with their own strengths and weaknesses, namely:

- Websites and online tools – a relatively passive approach in terms of user engagement but able to reach a large number of users efficiently;
- Interactive activities across a range of groups – a more involved and intensive approach, often involving workshops and seminars, which is targeted to a manageable group of people/organizations;
- Focused relationships between one provider and one user – an intensive approach, specifically targeted and tailored to a specific user group with a strong element of co-design.

The interaction and feedback between service providers and users is illustrated in Figure 4.6, which provides a useful conceptual breakdown of broad approaches.

The way to reach the largest number of users, and the cheapest means of obtaining user input, is through the Internet and online tools such as mobile applications, feedback forms, surveys, rating scales and questionnaires. These methods are extremely useful, but they do suffer from some weaknesses including a focus on existing users, the possibility that the input will not come from as broad a range of users as desired, and the potential that the results will be biased (for example, towards more technical users).

Interactive activities and focused relationships may reflect more traditional good practice when it comes to obtaining user feedback, but they come at a higher cost and often reach a

Box 4.2. Example from the prognostic discussion for the seasonal outlook from RA VI RCC-LRF (Toulouse) issued on 30 October 2018 (see http://seasonal.meteo.fr/)

Discussion
Forecast over Europe

"Predictors" which could influence the weather in Europe for the next three months:

El Niño (weak to moderate): Favours **negative NAO** (and also positive EA and SCAND patterns; see composites at http://ds.data.jma.go.jp/tcc/tcc/products/clisys/enso_statistics/);

Late summer Atlantic SSTs: The observed tripole pattern favours **positive NAO** during the upcoming winter (Cassou, 2004);

Easterly quasi-biennial oscillation (but forecast to switch in early 2019) and very weak solar activity favour sudden stratospheric warming and **negative NAO** circulations.

The Eurasian October snow extent is around normal as of mid-October; **no influence on NAO**.

In conclusion, the odds seem to **favour a negative NAO pattern for the next three months**.

Taking these predictors into account, and after analysing different model outputs, we forecast a scenario resembling the scenario forecast by the European Centre for Medium-Range Weather Forecasts (ECMWF)-S5 (a mix of positive SCAND and EA patterns and negative NAO). It is also noteworthy that the monthly ECMWF forecast for November shows wet (especially for the Mediterranean) and cool conditions.

Temperature: The above conclusions suggest mild conditions for the Mediterranean and cold conditions for eastern Europe and western Russia. Elsewhere, no clear signal emerges, but high variability is expected as a consequence of the predominance of SCAND and EA patterns.

Precipitation: Wet conditions are likely over the Mediterranean, extending toward Portugal, western France and southern England (positive EA pattern). Dry conditions are expected over parts of eastern Europe, Ukraine, and western Russia, consistent with the blocking pattern. Elsewhere, no clear signal emerges.

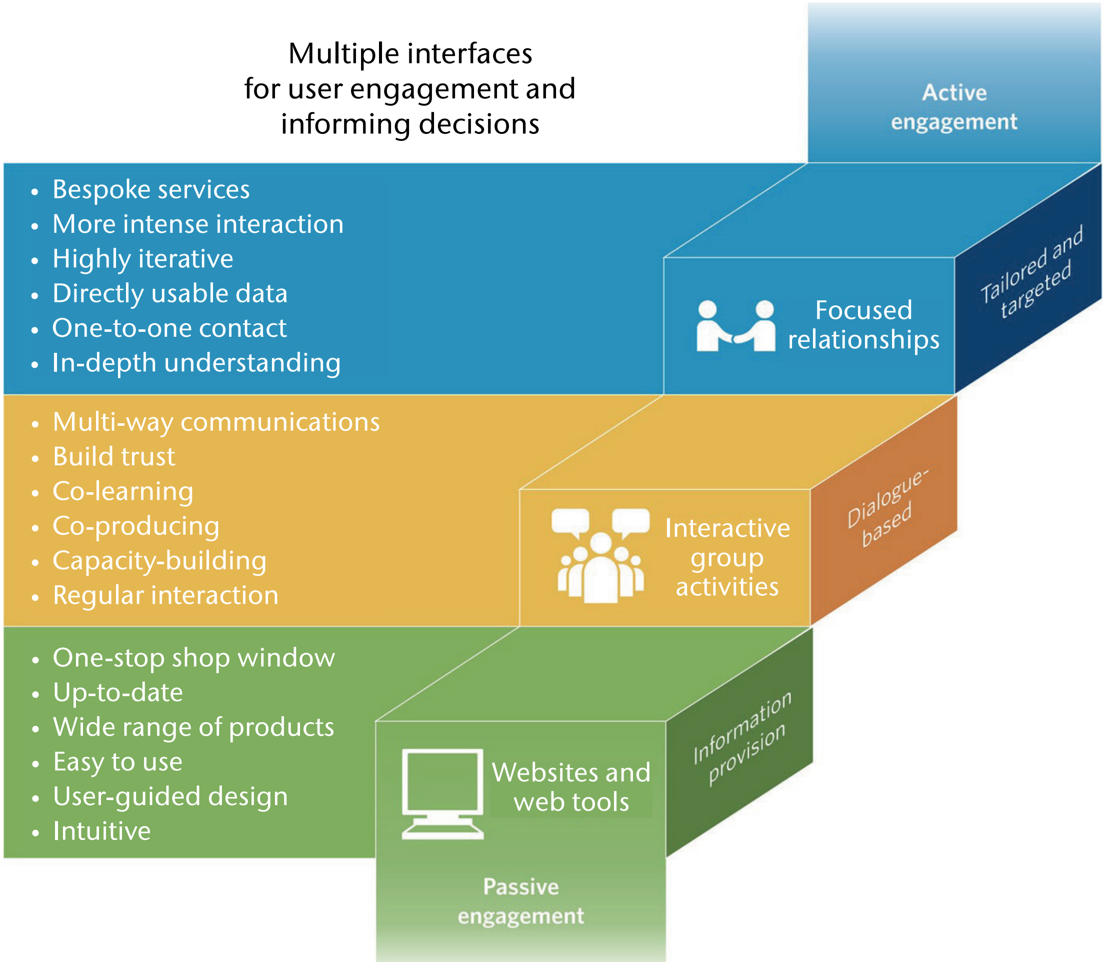

Figure 4.6. Ways in which NMHSs can engage with users

Source: Hewitt et al., 2017

small number of users. Workshop organizers assessing user feedback need to invite the most appropriate people from the industry, sector or community involved (such as sector champions or community leaders), while ensuring a suitable balance of participants (for example, with respect to gender, race, languages); this may require selecting a cross-section of users from within a particular group. Meetings and workshops need to be clearly structured and tend to work best when they are conducted by skilled group facilitators. The outcomes of workshops and meetings are improved if social scientists are involved.

Finally, some users of climate services will require more tailored solutions and will have quite specialized applications for the information they are provided. In these cases, a more focused relationship is often preferred, with an emphasis on co-design principles. These solutions come at significant cost to an NMHS but will often deliver the highest societal and economic impacts.

4.8 Recommend an approach for producing operational objective seasonal forecasts

As discussed previously, ***objective methods for producing seasonal forecasts are preferred and recommended***. Because some components of the current forecasting processes, especially the consensus process in RCOFs, are partly or largely subjective, strong arguments have been put forth for implementing more objective forecasting processes. These arguments were considered by the WMO Executive Council in its sixty-ninth session in May 2017 (Decision 18 (EC-69) – Sub-seasonal and seasonal forecasting systems).

The starting point for producing seasonal forecasts at the regional and national levels is determining the availability of a large-scale forecast. To this end, the WMO LRF infrastructure, namely the GPCs-LRF and LC-LRFMME, are useful resources; currently, forecasts from 13 forecasting models are available from LC-LRFMME and from the individual GPCs-LRF. Given these multiple models, the first question to address should be: which model should be selected to develop a seasonal forecast for a specific region? Because of regional characteristics including climate and operational constraints, there is no universal answer to this question; however, section 2.3.1 provides a recommended approach. We also stress that the model selection process needs the necessary expertise (or support from experts) in the analysis of the models' abilities to capture important climate processes, phenomena and teleconnections as well as the characteristics of observed climate variability in the region of interest. GPCs-LRF and RCCs are expected to play a pivotal role in this regard.

In cases where expertise or support is not available, it is recommended that all models be used. Obviously, operational constraints such as the timely provision of the forecast, or easy access to digital information, may favour a subset of the models. If expertise in interpreting verification scores and diagnosing the state of the climate system, both with respect to observations and with respect to forecast models is available, NMHSs can then move to the model selection process (see section 2.3.1).

When the NMHS is ready to select the appropriate models to be used, the first step is to obtain an estimate (or guess) of the seasonal forecast using a multi-model ensemble, includ**ing, it is important to note, the forecast uncertainty and its expression through probabilities.** On the latter point, we must stress that all the probabilities issued through the MME are expressions of the uncertainty inherent in the climate system (manifested as spread in individual forecasts within the ensemble). ***Additionally, it is important to emphasize that once models are selected for different seasons within a year, the model selection process should not be revisited every time a seasonal forecast is issued.*** Choosing different models for different forecasts is likely to result in jumps between forecasts and inconsistent forecast skill between periods.

Downstream to this MME approach, all necessary and usual post-processing steps should be performed; starting with bias correction and continuing with calibration at the grid point level (see Figure 4.7 and sections 2.2.2 and 2.4). In going through this process, RCCs (and RCOFs) have an important role to play in supporting the NMHS within the region so that climate information at the country level (in terms of both monitoring and forecasting) is consistent across climate forecasts over a broader region.

The next step in the forecast process should be to discuss the first estimate (or guess), and if necessary and justified, to alter the forecast if there are strong reasons for doing so. Different elements should be considered in this step – the current predictability, the state and impact of the regional climate and local drivers, the uncertainty present in both the climate system and the models. ***Note that it is very important to keep a record of all the discussions conducted concerning altering the MME-based forecast in order to build a traceable/documented forecasting process.*** One should keep in mind that MME forecasts may not always conform to the forecast expectations due to the nature of the climate drivers; this discrepancy, however, may be a manifestation of a source of predictability beyond the information contained in (average) teleconnection maps associated with the climate drivers.

The last step in the forecast process is to use local data, either station or gridded, to better understand the local climate and the statistical downscaling process, predominantly via hybrid methods (see sections 2.7 and 2.5). The general flow of developing objective seasonal forecasts is outlined in Figure 4.7.

In the seasonal process recommended above, while it has been a routine practice to take forecast products from GPCs-LRF or LC-LRFMME and ***frequently modify these products manually, this practice is not encouraged without (a) strong justification, (b) relevant expertise and (c) proper documentation***. If the above-mentioned forecast products are modified manually, the modifications should be carried out using methods which have been previously documented and have led to improvements in seasonal forecasts. A manual modification of a forecast should be supported by a detailed and transparent discussion of climate drivers, and all forecasts should

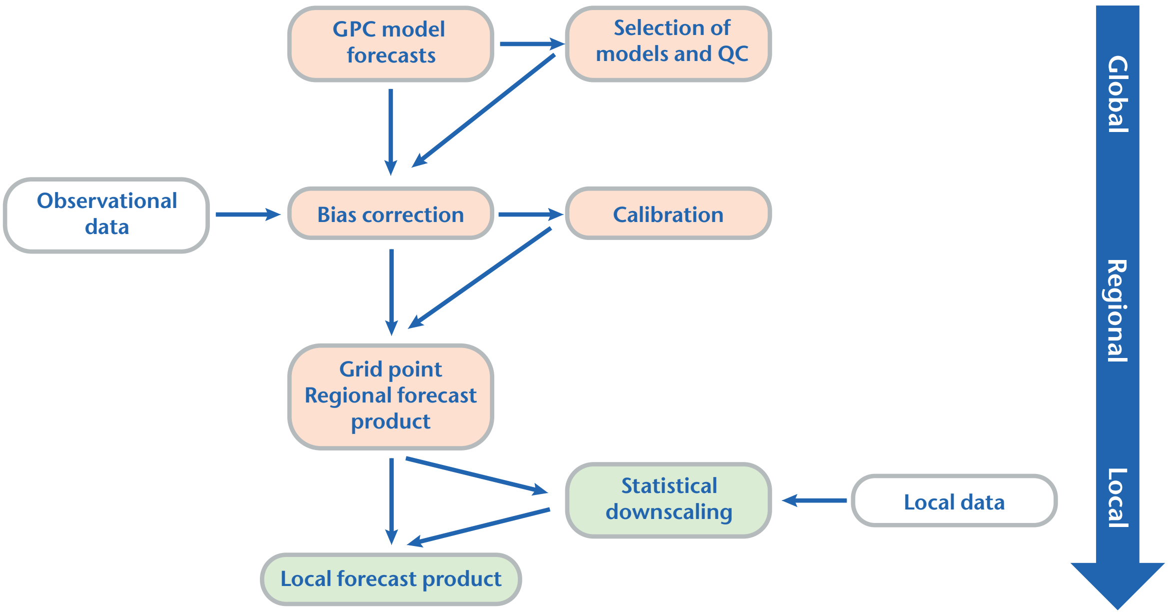

Figure 4.7. An outline of the recommended procedure for developing seasonal forecasts at the regional and national levels starting from the forecasts from GPCs-LRF. The first step in the forecast development process is to choose GPC-LRF models that will be available in a consistent manner, and if desired, to select models to be used over the region. Model forecasts are then bias corrected, calibrated and combined using the multi-model approach, leading to seasonal forecasts at the regional level. If desired, forecasts at the regional level can then be downscaled to the local level. The forecast development process also encapsulates the basic concept of flow of information from the global to the regional to the national level. It should be recognized that, starting from GPC-LRF forecasts as the input, adding various layers of processing (for example, model selection, statistical downscaling) requires additional resources and expertise and should be considered in the establishment of an operational seasonal forecast strategy and in the steps required to develop the necessary capacity.

be verified as per standard practices; this verification should include documenting the skill of the MME and using statistical forecasts as the baseline over which manual intervention shows improvement. ***The process of manual intervention requires significant knowledge and expertise to be able to improve on the MME and statistical forecasts and detailed knowledge of global, regional and local drivers of climate variability and teleconnections.***

At the conclusion of the seasonal forecast process, it is recommended that a technical note supporting the seasonal forecast be prepared and a report of the discussions carried out during each briefing be issued as part of the forecast dissemination process.

4.9 Establish good practices for communicating seasonal forecasts

Once produced, seasonal forecasts need to be communicated to users. Seasonal forecasts, because of limited skill, and because of regional and seasonal dependence in skill, present unique challenges to users; good communication is a key factor in how these forecasts are used and applied. A number of factors frequently make using forecasts difficult. These include: low predictive skill, forecasts which often sit close to the climatological frequency (for example, 50:50 for probabilistic forecasts of above median conditions), the complexity of the forecast information, and general difficulties related to understanding and responding to probabilities. Human cognitive biases tend to compound these issues (Nicholls, 1999) and may affect the value attributed by users because users may apply information incorrectly. To alleviate some of the issues in the usability of seasonal forecasts, additional recommendations are provided below.

4.9.1 *Include information about past forecast quality*

This section adds to the recommendations provided in section 4.3.

SVSLRF provides recommendations for the minimum requirements for GPCs-LRF to follow when assessing and reporting on the quality of seasonal forecasts. Similar practices are expected of RCCs, and these practices provide a suitable base level of accountability for NMHSs when delivering seasonal forecasts. It is important that the forecast quality assessments use established metrics similar to those used in SVSLRF (see the *Manual on the Global Data-processing and Forecasting System*), including measures such as relative operating characteristics and reliability curves.

A challenge for communicating forecast quality is the complexity of the verification measures involved; these measures are rarely familiar to users, and they are often mathematically complex. This issue can be addressed in various ways, including by using simpler measures such as "hit rates" or a "per cent consistent" (Wang et al., 2018). If needed, simplified forecast quality measures can be used to complement scores recommended in SVSLRF to aid user uptake of seasonal forecasts. It should be recognized, however, that these simpler methods have several drawbacks, including that they may appear to condone interpreting probabilistic forecasts too deterministically; they do not take account of how strongly a forecast shifts from the climatological distribution and thereby may be potentially misleading for low probability events.

There has also been some effort to develop user-orientated forecast quality measures; these may be more accessible to users and better match user needs. One example of such measures is the weighted per cent of correct forecasts, which weights forecast quality by the magnitude of observed anomalies meaning that the focus is on those anomalies which are significant for users.

In summary, hindcast and real-time forecast quality should be published, regularly reported and made available alongside operational forecasts so that this information can be used when preparing seasonal forecasts. When possible, both hindcast and real-time forecast quality should be described in the scientific literature, and it should be noted that improving forecast quality is a key reason to invest in climate modelling, science and observations (see, for example, Min et al., 2014, MacLachlan et al., 2015). It may be helpful for operational agencies to report forecast quality as a performance indicator, noting that it will be an approximate estimate of value and impact over time (see Figure 4.8).

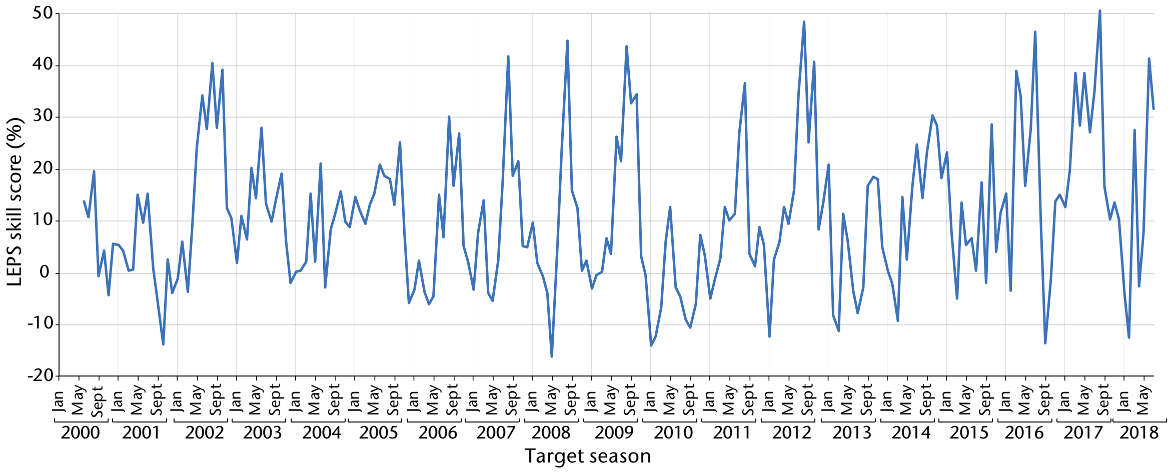

Figure 4.8. Forecast quality of 18 years of maximum temperature seasonal forecasts issued by the Australian Bureau of Meteorology using the Linear Error in Probability Space (LEPS) skill score (Potts et al., 1996). During this period, the Bureau upgraded three forecast models and two statistical models and introduced a 250-kilometre coupled climate model and a 60-kilometre coupled climate model. The forecast skill score shows a clear increase over time, reflecting upgrades made to the forecasting system over this period.

4.9.2 *Include guidance on the interpretation of forecast probabilities*

As discussed in section 4.5, the necessary probabilistic presentation of seasonal forecasts represents one of the most significant barriers to their use. It is important for probabilities to be carefully communicated to users and for the ways in which forecasts are presented to not cause users to be biased with respect to their use. (This is sometimes referred to as the framing bias.)

Several practical steps can be taken to help users utilize and interpret seasonal forecasts. These include:

- Presenting forecast probabilities as numbers, whether in the form of numerical values, contoured maps, graphs or tables. Modern web technology (see Figure 4.9) will often allow these probabilities to be displayed in various ways, so users can access information in the way which best suits them.

- Ensuring that descriptors used to define probabilities are clearly defined, well tested, and used consistently. Terms such as "likely", "very likely", and so forth carry very different meanings for users and are best avoided (Mauboussin and Mauboussin, 2018).

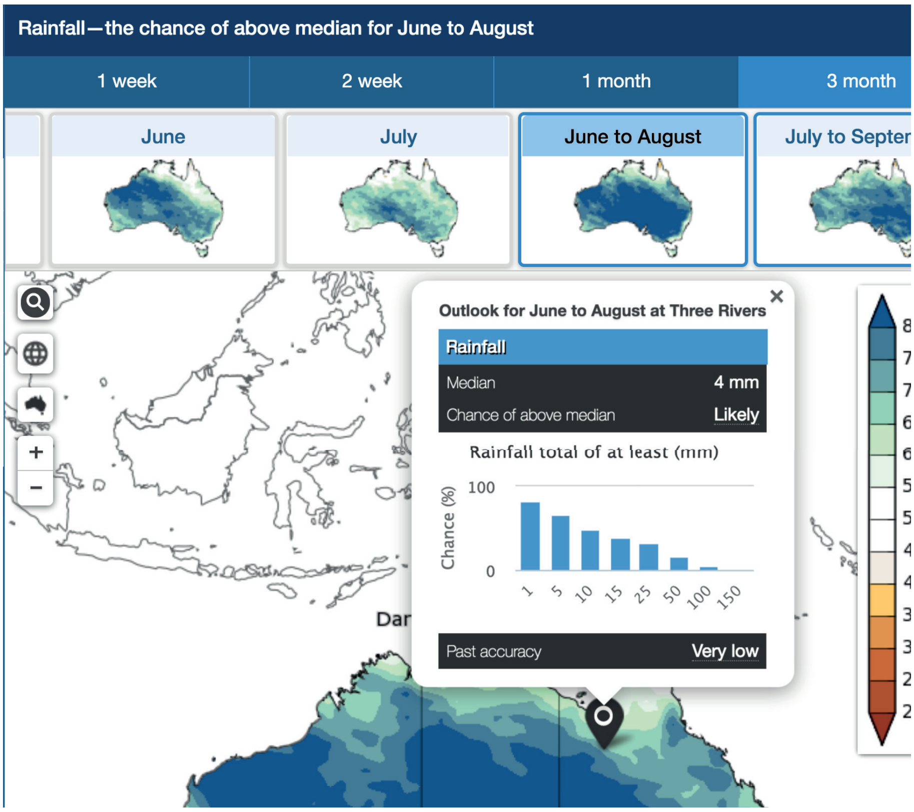

Figure 4.9. An example of presenting a seasonal forecast (for precipitation) concurrently in several ways. The dynamic web page provides probabilities in graphical, tabular and text form, includes information about the base rate, and presents the forecasts as a probability of exceedance. It can be useful to provide forecast information concurrently in several ways and to describe the forecasts in more than one way. For example, a forecast of a 70% chance of above median rainfall can be described as 70%, "7 in 10", "highly likely", and so forth. Across GPCs-LRF, RCCs and NMHSs, there are many examples of good practice with respect to displaying forecast information.

Source: http://www.bom.gov.au/climate/outlooks/ - /rainfall/median/weekly/0

- Carefully choosing colours and types of displays. For example, the pie chart in which probabilities are portrayed as fractional areas of a circle is inefficient and prone to misinterpretation (Kozak et al., 2015). Colours should be intuitive (Teuling et al., 2011) and easily distinguishable.

- Ensuring that base rates are clearly interpreted and described. For example, when describing the likelihood of an El Niño, it is often useful to refer to the climatological base rate, which is near 20% ("one in five years"). In this context, a 60% probability of an El Niño may seem only just over 50:50, but making it clear that this probability is, in fact, three times the long-term base rate will highlight the fact that the risk of this event is indeed greatly elevated.

- Ensuring that the terms used to describe forecasts are defined in advance and do not vary. Humans suffer from a range of cognitive biases which may bias their communication. Forecasts should therefore only contain terms which are consistently defined in advance, and by repeated and consistent use, are internalized by the user community.

- Avoiding anchoring communication to historical analogues, which may influence how people interpret a climate forecast. For example, emphasizing strong El Niño events (and neglecting regional SSTs patterns) may bias user expectations if a forecast mentions that an El Niño is likely. Using analogues based only on ENSO may lead to biases in forecasts and user expectations.

4.9.3 ***Include a discussion on managing user expectations***

Seasonal forecast users are typically familiar with weather forecasts and warnings (for the expected conditions for the forthcoming one to seven days), which are known to be highly accurate and usually presented as a single "deterministic" outcome. Sharper forecasts have improved quality and more value for users, which means users naturally prefer a deterministic forecast over a probabilistic one.

Users typically want the most emphatic forecasts, and it is common for them to convert probabilities into yes/no categorical forecasts to simplify their decision-making process. This creates issues as climate forecasts are often highly uncertain and may have modest skill, and any attempt to reduce a probability to a simple yes or no renders the forecasts unreliable (too emphatic). It is important for the full uncertainty of forecasts to be retained and for forecasts to be provided alongside forecast quality information in order to encourage users to be aware of their limitations.

The fundamental limits with respect to seasonal forecasts have been described previously (see section 1.3), and the consequent need to use probabilities is well established. Ensuring that these concepts are understood by users is important in order to manage expectations regarding the extent to which seasonal forecasts (associated with probabilistic information) will be correct and in order to make sure that users do not feel unduly confident in the use of probability forecasts in their decision-making. There are many excellent articles, audiovisual aids, infographics and online courses which are available to assist with user engagement (see the examples in the next paragraph).

Users of forecasts can be very diverse, and it is important to manage user expectations in ways which are understandable and engaging for them. This will often require translating complex information into simple language. Some innovative examples of how climate information can be simplified are the Pacific Climate Crabs and the Australian Climate Dogs animations and the NOAA ENSO Blog, which use simple concepts and illustrations to explain the role of climate drivers including ENSO. More complex examples include course material produced by WMO (for example, the Theory and Operational Principles series).

In summary, ***managing expectations will require an approach which matches information with users and will usually require more than one approach given the diversity of the user base.***

Customizing forecasts to user needs or incorporating more predictable variables into a forecast can also assist with user expectations. As an example, seasonal forecasts for soil moisture will often have higher skill than forecasts for precipitation, as soil moisture conditions in the upcoming season depend on the current state of various climate variables, including precipitation, evaporation, temperature and soil moisture itself. To illustrate this point, consider the situation of a well-established drought in eastern Australia during the later stages of an El Niño event. Historically, the relationship between El Niño and precipitation is not strong during the summer, but the combination of dry soil, above average temperatures and low seasonal precipitation means that the likelihood of general drought relief is poor. This relationship is seen in objective seasonal forecasts of streamflow and fire danger, which show significantly more predictability than precipitation alone at this time of year. In this case, we might not be too confident about how precipitation will respond over the December to February period, but we can be much more confident that relief from drought is unlikely.

4.10 **Good practices for establishing and maintaining credibility**

Providers of seasonal forecasts should also observe a set of good practices which will further assist in establishing and maintaining credibility. These practices include:

- ***Using objective techniques which are reproducible and traceable;***
- ***Ensuring that forecast models and the methods of post-processing and forecast preparation are clearly documented, peer reviewed, transparent and accessible to users and scientific peers;***
- ***Maintaining records of issued technical notes supporting forecast briefings, reports of the briefings, forecasts; and***
- ***Ensuring that proper verification is undertaken as part of the operational quality assurance process.***

4.10.1 ***Using objective methods which are reproducible and traceable***

In accordance with current scientific knowledge and technical capabilities, objective procedures are used to issue seasonal forecasts based on single models and multi-model ensembles and sometimes in combination with dynamical and empirical (statistical) predictions.

The use of objective methods is recommended when issuing seasonal forecasts because the forecasts then use a common, testable and reproducible approach. In the past, it has been common practice for experts to undertake various forms of forecast intervention, at times changing probabilities or other outputs. It is better to avoid these approaches as they are not reproducible from case to case, cannot avoid subjective judgments or personal preferences, and in the end, may not be able to demonstrate their utility on improvements in forecast skill. The use of objective methods avoids these artifacts, and when applied optimally (that is, to minimize the errors in a forecast), it is to be expected that well-tested automated procedures will perform better. It should be noted that increased reliance on objective methods for seasonal forecasts will result in saving valuable human resources that can be applied to interaction with the user community, for example, to provide interpretation in the use of probabilistic forecasts, to obtain user feedback, to develop tailored forecasts, and so forth.

4.10.2 ***Ensuring that methods are properly documented***

The *Manual on the Global Data-processing and Forecasting System* specifies the mandatory requirement that GPCs-LRF must make available up-to-date information on the characteristics of prediction systems, on methods used to develop forecasts and on contacts for further information through the Internet (Table 4.1). As all GPCs-LRF provide forecasts to LC-LRFMME, the requirements placed on LC-LRFMME extend to GPCs-LRF and cover past forecasts and verification information maintained by LC-LRFMME.

Table 4.1. Adapted from the *Manual on the Global Data-processing and Forecasting System* Appendix 2.2.10: Characteristics of global numerical long-range prediction systems

Date of implementation of the current seasonal forecast system
Whether the system is a coupled ocean–atmosphere forecast system (Y/N)
Whether the system is a Tier-2 forecast system (Y/N)
Atmospheric model resolution
Ocean model and its resolution (if applicable)
Source of atmospheric initial conditions
Source of ocean initial conditions
If Tier-2, the source of SST predictions
Hindcast period
Ensemble size for the hindcasts
Method of configuring the hindcast ensemble
Ensemble size for the forecast
Method of configuring the forecast ensemble
Length of forecasts
Data format
The latest date that predicted anomalies for the next month/season become available
Method of construction of the forecast anomalies
URL where forecasts are displayed
Point of contact

With respect to documenting seasonal forecast procedures at the regional and national levels, up-to-date information (or references) should be kept on forecast methods and observational dates, the methods used to post-process and prepare forecast products and any significant changes made over time.

Peer review in the scientific literature is an important means of assessing the robustness and suitability of seasonal forecasting methods. NMHSs and RCCs are encouraged to document their systems and processes in the scientific literature whenever possible. This may be through established science journals, or in-house technical reports which are subject to peer review (for example, review by external scientists). This documentation can be particularly important on those occasions when forecasts do not match observations well and to support ongoing improvements to services and systems. There are many examples in the literature of operational systems and services; these examples can be used as templates for documentation purposes (see, for example, Graham et al., 2011, Hudson et al., 2017). In addition to scientific documents, as few users will be familiar with the details of climate science, simplified material should be made available to users in the form of websites, brochures and infographics.

4.10.3 *Maintaining archives of past seasonal forecasts*

Alongside up-to-date system information, maintaining a history of past forecasts is good practice for seasonal forecast providers. Ideally, the forecast will sit alongside verification information, allowing users to browse past forecasts and compare them against the observed outcomes. Copies of all issued forecasts should be retained, and ideally, the data which went into the forecast should also be retained, though this may not be practical when dealing with large model datasets. Having a well-documented archive and retention policy is a sound starting point, and some practical guides on this matter are provided by various groups including the International Council on Archives. Many of the practical considerations for climate data management described in the CCl *Guide to Climatological Practices* (WMO-No. 100) are applicable to forecast data. In many countries, there are legal requirements for the retention of official products and services including forecasts, and local authorities may be able to support and assist with process designs.

Establishing Internet archives of digital products should be considered part of website design, including a means for users to navigate through past forecasts. At their most basic, these archives might simply be stores of intuitively named files or part of a climate data management system in which forecasts sit alongside observations and form part of a broader climate record. GPCs-LRF, LC-LRFMME and RCCs all face similar issues around data retention and can be a useful contact for support and suggestions.

4.10.4 *Verification as part of seasonal forecast quality assurance*

A fundamental component of the forecasting process is quantifying forecast quality. As discussed earlier, this is achieved by verifying hindcasts and real-time forecasts and illustrates how accurate (or inaccurate) the previously produced forecasts were. Quantifying forecast quality allows users to become aware of the current limitations in seasonal forecasting over various parts of the world and allows forecast producers to indicate the limits of their products.

CHAPTER 5. WMO INFRASTRUCTURE AND RESOURCES FOR SEASONAL FORECASTS

5.1 Global Producing Centres for Long-Range Forecasts

Within the operational infrastructure of WMO, the seasonal component of GDPFS is represented at the global level by GPCs-LRF. These centres, which are formally designated by WMO, provide seasonal forecasts on an operational basis. In order to be designated as an operational centre, a GPC-LRF must meet the mandatory functions described in sections 2.2.1.5.1 and 2.2.1.5.2 of the *Manual on the Global Data-processing and Forecasting System*. At the time of the writing of this document (December 2018), there were 13 WMO designated GPCs-LRF. Users at the regional and national levels are encouraged to contact GPCs-LRF if a particular seasonal forecast dataset or product is required, for example seasonal forecasts from GPC-LRF Washington.

5.2 Lead Centre for Long-Range Forecast Multi-Model Ensemble

There is a general consensus that because different models have different biases, forecasts based on a multi-model approach benefit from the cancellation of biases and generally provide more skilful and reliable forecast information than forecasts based on a single model (Yoo and Kang, 2005; Min et al., 2014). As indicated in section 5.1, currently there are 13 GPCs-LRF providing seasonal forecast information. However, for users at the regional and national levels, accessing forecasts from each of the GPCs-LRF would present a considerable expense in terms of resources. In light of this, the Lead Centre for Long-Range Forecast Multi-Model Ensemble (LC-LRFMME) was developed to (a) collect seasonal forecast data from individual GPCs-LRF, (b) develop seasonal forecasts based on a multi-model ensemble and (c) provide multi-model-based seasonal forecasts to users at the regional level (Figure 5.1). The concept of a LC-LRFMME was subsequently formalized in the *Manual on the Global Data-processing and Forecasting System* (section 2.2.2.2), where the mandatory and highly recommended functions of the centre are listed. At present, the Korea Meteorological Administration (KMA) and the National Centers for Environmental Prediction (NCEP) are jointly responsible for LC-LRFMME. Users at the regional and national levels can access multi-model seasonal forecast products via the LC-LRFMME website.

5.3 Global Seasonal Climate Update

The Global Seasonal Climate Update (GSCU) is a WMO-coordinated initiative that provides information on current climate conditions and the seasonal forecast for the upcoming season at the global scale. GSCU was originally conceived as an extension of the WMO consensus-based El Niño/La Niña updates. The information provided on large-scale patterns and climate drivers can be used as key inputs by climate information providers at the regional and national levels for their own specific purposes. Operational production of GSCU is carried out by LC-LRFMME under the joint responsibility of KMA and NOAA/CPC. GSCU summarizes information about current climate conditions, including the global surface temperature and precipitation anomalies for the past season and the current state of important climate drivers such as SST anomalies in various El Niño regions, the Indian Ocean Dipole (IOD), and so forth. The current state of climate conditions is complemented by corresponding seasonal forecast information for the upcoming season. Seasonal forecast information is provided by LC-LRFMME and is based on a multi-model ensemble approach utilizing GPC-LRF models. As discussed in section 5.2, the role of LC-LRFMME is to collect seasonal forecast data from the 13 GPCs-LRF and consolidate that information using the multi-model approach.

In GSCU, all graphical products are complemented by text summaries for six regions corresponding to the WMO Regional Associations (RAs).

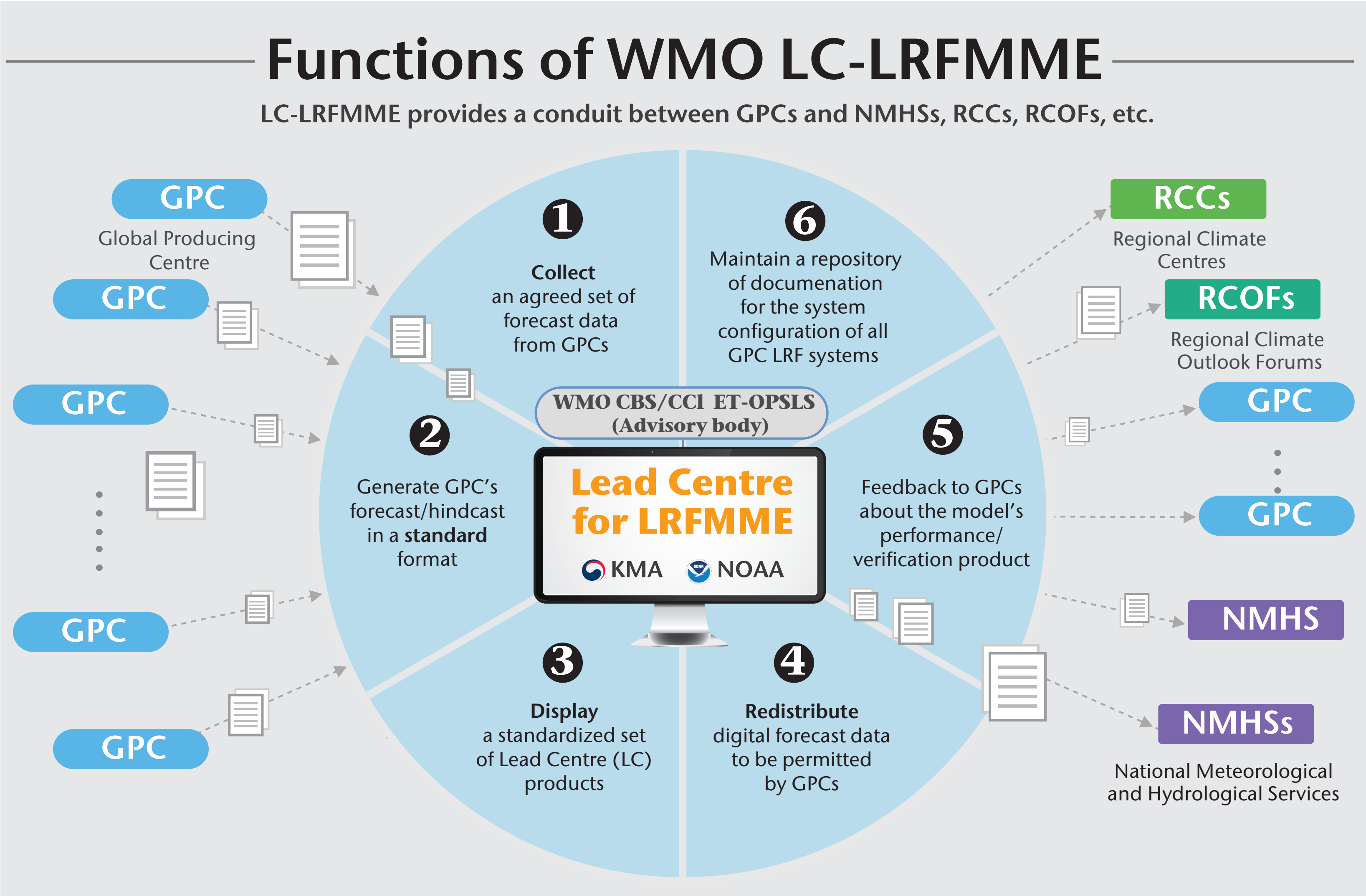

Figure 5.1. Functions of LC-LRFMME. LC-LRFMME collects seasonal forecast data from individual GPCs-LRF and provides multi-model products to users at RCCs, RCOFs, GPCs and NMHSs.

An example of the seasonal prediction graphical display for surface temperature and precipitation is given in Figure 5.2. In the interpretation of the forecast probabilities, it is noted that a particular colour does not signify that the seasonal mean temperature or precipitation is "certain" to be observed in the highlighted (or most likely) forecast category, but rather that its probability of being in that category is the largest. For a further discussion of how forecast probabilities should be interpreted, see the WMO website for GSCU.

The trial phase of GSCU was initiated in 2015 and is now close to being made operational. At present, GSCU is released four times a year around the 25th of the month prior to the season in question (March-April-May (MAM), June-July-August (JJA), September-October-November (SON) and December-January-February (DJF)). Additional and updated information on GSCU is available on the WMO website for GSCU.

5.4 Regional Climate Centres

Regional Climate Centres (RCCs) are Centres of Excellence designated to create regional climate products, including data, and to conduct monitoring and forecasting activities in order to support and strengthen the capacity of WMO Members in a given region to deliver climate services. WMO RCCs have been recognized as key elements of the Climate Services Information System, a foundational pillar of the Global Framework for Climate Services.

Since the designation of the mandatory functions (Appendix 2.2.16 of the WMO *Manual on the Global Data-processing and Forecasting System*) and highly recommended functions (Attachment 2.2.2 of the *Manual*) for RCCs, the number of RCCs has gradually increased. To date, eight RCCs and three RCC Networks have been formally designated and provide climate services and products to NMHSs in the concerned regions on an operational basis. The services provided are in the domains of data and data management, monitoring, long-range forecasting, and training and capacity building. There are also a number of RCCs that are currently in the demonstration phase. In fact, each of the Regional Associations has at least one RCC or RCC Network that has been formally designated or is in the demonstration phase.

An RCC Network is a group of centres performing climate-related activities that collectively fulfils all the required functions of an RCC. Each centre in a designated WMO RCC Network is referred to as a Node, and each Node performs one or several of the mandatory RCC activities for the region or subregion defined by the Regional Association.

A review of RCC operations was conducted in 2018 (Pune, October 2018), and discussions involved possible modifications to some mandatory and highly recommended functions and options for the future evolution of RCCs.

5.5 Regional Climate Outlook Forums

Regional Climate Outlook Forums (RCOFs) are regional platforms that bring together climate experts and sector representatives from countries in a climatologically homogenous region. Currently, RCOFs provide consensus-based climate prediction as a key service to the region; this is generated (typically in a pre-Climate Outlook Forum workshop) with input from GPCs-LRF, RCCs, NMHSs and other climate research centres such as IRI, APEC Climate Centre, and so forth. However, as recommended in this document, a transition from consensus-based to objective seasonal forecast practices should be implemented to take the evolution of RCOFs to the next level.

One of the main goals of RCOFs is to provide information that can be used to achieve socioeconomic benefits in climate-sensitive sectors. At present, there are 20 RCOFs (including the recent Pan-Arctic Climate Outlook Forum – PARCOF) which cover most of the regions across our planet. Based on the needs of specific sectors, specialized, sector-oriented outlook forums, such as the Malaria Outlook Forums (MALOFs) in Africa, are also held in conjunction with Regional Climate Outlook Forums. More information about RCOFs can be found on the WMO website.

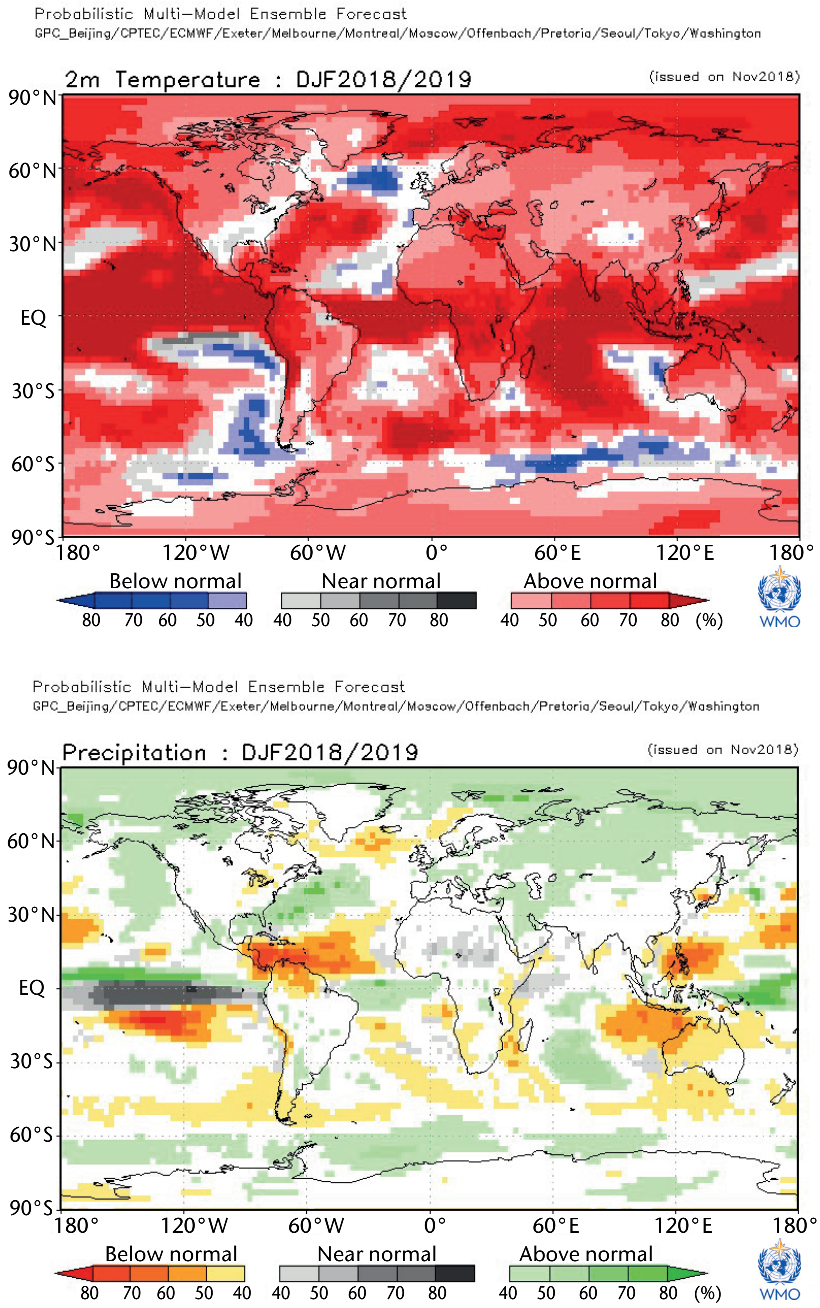

Figure 5.2. An example of seasonal forecasts for surface temperature (top) and precipitation (bottom) from the Global Seasonal Climate Update. Seasonal forecasts are issued by LC-LRFMME and are based on multi-model ensembles. Colour is indicated only for the tercile category that has the highest probability of occurrence. For example, for regions highlighted in red, the seasonal mean surface air temperature forecast indicates that the most likely forecast category to occur is warmer than normal. Similarly, blue indicates regions where the colder than normal category is the most likely to occur, and grey indicates regions where the near normal category is the most likely to occur. Deeper shades of colours highlight increasing probability for the seasonal mean temperature to be in the indicated category. White areas indicate that all categories are equally likely.

Source: WMO Global Seasonal Climate Update

The RCOF process, pioneered in Africa, typically includes the following:

- A training workshop on seasonal climate prediction to strengthen the capacity of national and regional climate scientists and to help NMHSs prepare their national forecasts;

- Meetings of national, regional and international climate experts to develop a consensus for the regional climate outlook, typically in a probabilistic form;

- The Forum itself, in which climate scientists and representatives from user sectors identify the impacts and implications of the predicted seasonal climate on the region and formulate response strategies; and

- Special outreach sessions involving media experts to develop effective communications strategies.

On the twentieth anniversary of the successful implementation of the RCOF concept, WMO organized the Global RCOF Review 2017 (a comprehensive review of the RCOF process) to review the activities of individual RCOFs, to analyse the different aspects of current operational practices at RCOFs, including reaching consensus, verification, capacity development and user engagement, and to discuss opportunities for an improved and sustained RCOF process. The participants unanimously recognized the progress achieved, particularly with respect to the contribution of RCOFs in promoting the wider use and better interpretation of seasonal forecasts at the national level, and agreed on recommendations for the way forward towards the new generation of RCOFs (RCOF v2.0).

CHAPTER 6. OTHER SOURCES OF SEASONAL PREDICTION PRODUCTS

6.1 North American Multi-Model Ensemble

The North American Multi-model Ensemble (NMME) is an initiative of NOAA in the United States that combines forecasts from the operational models of GPC-LRF Washington and GPC-LRF Montreal and models developed at the National Aeronautics and Space Administration, the Geophysical Fluid Dynamics Laboratory and the National Center for Atmospheric Research (Kirtman et al., 2014). Although the specific combination of models has evolved since NMME was established in 2011, there have consistently been six to eight contributing models with approximately 100 ensemble members in total.

NMME forecasts are posted online by the 9th of each month at http://www.cpc.ncep.noaa.gov/products/NMME/. Information on this site includes:

- Maps of seasonal temperature, precipitation and SST anomalies at one- to five-month leads;
- Corresponding maps for monthly-mean anomalies;
- Uncalibrated and calibrated probabilistic forecasts;
- Deterministic (anomaly correlation) and probabilistic skills;
- Verifications of past forecasts; and
- Niño3.4 index plumes.

Deterministic spatial anomaly forecasts include ensemble means for each contributing model plus a multi-model forecast that weights each model equally. The probabilistic forecasts pool all of the ensemble members, weight each ensemble member equally and employ the calibration method described in van den Dool et al. (2017).

NMME provides free access to its numerical data, which includes:

- Real-time and hindcast raw monthly-mean data through the IRI Data Library;
- Real-time anomaly and probabilistic forecasts (current and archived);
- Model hindcast climatologies; and
- Monthly and daily hindcast data for a wide range of variables (NMME Phase-II data).

Links to each of these resources are provided at http://www.cpc.ncep.noaa.gov/products/NMME/data.html.

NOAA also provides regionalized NMME resources including forecast maps and data at http://www.cpc.ncep.noaa.gov/products/international/nmme/nmme.shtml, and NMME currently serves as the basis for the calibrated multi-model probabilistic seasonal forecasts provided by IRI (section 6.4).

An important benefit of free access to NMME hindcast and real-time forecast data is that, in addition to serving forecast users, these data have stimulated a wide range of research activities relating to seasonal forecasting. A partial list of the scientific publications that have resulted from such studies is provided at http://www.nws.noaa.gov/ost/CTB/nmme_pub.htm.

6.2 Copernicus Climate Change Service

The Copernicus Climate Change Service (C3S) is one of the six thematic services provided by the European Union's Copernicus Programme. The Copernicus Programme is managed by the European Commission, and C3S is implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF). The objective of the Copernicus Climate Change Service is to build a European Union knowledge base in support of mitigation and adaptation policies for Climate Change and Global Warming. The ultimate goal of C3S is to provide reliable information about the current and past state of the climate, forecasts on a seasonal timescale, and the most likely projections in the coming decades for various scenarios of greenhouse gas emissions and other climate change contributors.

C3S regularly publishes seasonal forecast products at https://climate.copernicus.eu/seasonal-forecasts. All C3S products are free and available to the public. Seasonal forecasts, digital data and graphical products are updated every month, currently on the 13th day at 1200 UTC, and cover a time period of six months. The graphical products consist of maps for several forecast variables (air and sea-surface temperature, atmospheric circulation and precipitation) for single models and multi-model combinations.

Digital data for a larger set of variables are available from each of the contributors to the multi-model system in the original time resolution and as monthly means (including monthly-mean anomalies); these data can be accessed through the C3S Climate Data Store (CDS). The centres currently providing forecasts to C3S are ECMWF, the Met Office, Météo-France, the German Weather Service (Deutscher Wetterdienst, DWD) and the Euro-Mediterranean Centre for Climate Change (Centro Euro-Mediterraneo sui Cambiamenti Climatici, CMCC). Both real-time forecast data and hindcast data are available.

C3S has developed a good technical infrastructure to support users in accessing and processing data. From CDS, users can access observational and reanalysis data useful for validation and calibration of seasonal forecasts. C3S is currently developing forecast performance (verification) estimates to accompany forecast products wherever possible.

6.3 APEC Climate Centre

The Asia-Pacific Economic Cooperation (APEC) Climate Centre (APCC) collects climate forecast information from 14 seasonal climate forecasting centres and institutes, including GPCs-LRF and major research groups, in 10 countries. These forecasts are combined using a multi-model ensemble scheme (Min et al., 2009) and are disseminated through the APCC website: https://www.apcc21.org/ser/outlook.do?lang=en and email list.

The following information is available on the APCC website:

- Deterministic/probabilistic forecasts for the upcoming six months from the multi-model ensemble and from individual models for six variables (precipitation, surface temperature, 850 hPa temperature, 500 hPa geopotential height, sea-surface temperature, and 200 hPa wind field);

- Forecasts for major oceanic climate drivers (ENSO, El Niño Modoki, and IOD) and probabilistic forecasts of ENSO state and strength;

- Verifications of seasonal forecasts for hindcasts and recent real-time forecasts;

- Information on current climate conditions, including major climate indices and drought/flood monitoring;

- Monitoring and multi-model forecasts of the Boreal Summer Intraseasonal Oscillation (BSISO).

APCC has developed a web-based tool (CLimate Information toolKit, CLIK) which allows users to construct their own multi-model ensemble forecast comparing hindcast performances when selecting different model combinations. It also provides a statistical downscaling function to convert large-scale forecast information into pointwise information (http://clik.apcc21.org). Raw forecast data is available via FTP and OpenDAP (http://adss.apcc21.org).

6.4 International Research Institute for Climate and Society

The Columbia University International Research Institute for Climate and Society (IRI) provides probabilistic seasonal climate forecast products based on a recalibration of model output from the United States National Oceanographic and Atmospheric Administration's NMME Project (section 6.1). The output from each NMME model is recalibrated prior to combining in a multi-model ensemble to optimize the reliability of the probability forecasts. The forecasts are presented on a one-degree latitude-longitude grid and are made available at https://iri.columbia.edu/our-expertise/climate/forecasts/seasonal-climate-forecasts/. IRI also has extensive seasonal hindcast data available on its data library portal http://iridl.ldeo.columbia.edu/index.html?Set-Language=en.

6.5 EURO-BRazilian Initiative for improving South American seasonal forecasts (EUROBRISA)

The Brazilian Centre for Weather Forecast and Climate Studies at the National Institute for Space Research provides precipitation seasonal climate forecast and verification products for South America based on a combination of three European dynamical coupled ocean–atmosphere models (ECMWF System 5, UK Met Office GloSea 5 GC2 and Météo-France System 6) and an empirical model that uses Pacific and Atlantic sea-surface temperatures as predictor variables (Coelho et al., 2006a). These products are made available at http://eurobrisa.cptec.inpe.br/.

CHAPTER 7. OTHER ASPECTS OF SEASONAL PREDICTIONS AND VARIABILITY

7.1 Attribution and forecast post-mortem

Skill assessments for real-time forecasts establish the accuracy of forecast seasonal anomalies but do not provide reasons for the predicted anomalies or why predictions are accurate in some cases and inaccurate in others (Barnston et al., 2005). The reasons for inaccurate seasonal forecasts may include an excessive contribution of the noise component to the observed seasonal mean or errors in forecasts of sources of predictability themselves.

In cases in which GPC-LRF forecasts form a key component in the consolidated forecast, WMO recommends a post-mortem analysis of the performance of the key model fields used and any related circulation patterns (see the *Guidance on Verification of Operational Seasonal Climate Forecasts*). There may be cases in which inaccurate predictions have a knock-on detrimental impact on the consolidated forecast probabilities for the region of interest, for example, shortcomings in model predictions of the timing of El Niño development, or the location of the main warming in the tropical Pacific (in the case of El Niño). An analysis of such cases can lead to an improved understanding and a strengthened forecast process.

Several research institutions are now focusing on understanding climate variability by evaluating the performance of previous seasons' predictions. This post-mortem analysis, in combination with additional model experiments, aims to identify sources of predictability for a specific event. For example, the effect of the initial surface conditions (ocean, dry soil, excess snow cover, and so forth) can be separated from the effect of the atmospheric variability (noise) by comparing climate model simulations with initial predictions. In addition, the common features among the ensemble of model runs can be identified as the predictable signal, and the spread among the ensemble members can be viewed as the unpredictable component. Separating predictable and unpredictable components and relating predictable components to external factors in the initialized seasonal predictions is generally referred to as an "attribution" analysis. At the NOAA Climate Prediction Center, the attribution of seasonal mean climate anomalies is available in real time at http://www.cpc.ncep.noaa.gov/products/people/mchen/AttributionAnalysis/.

The attribution analysis based on forcing atmospheric model simulations with observed boundary conditions also provides an assessment of the extent to which the predictable signal is compromised due to limitations and uncertainties in predicting boundary conditions themselves and aids in the post-mortem of the forecast performance. It is also expected that in the long run, attributing observed climate anomalies to physical causes (the sources of predictability) by building the knowledge base for sources of predictability will help improve future seasonal forecasts (Barnston et al., 2005) (section 4.3). With enhanced GPC/RCC/NMHS interaction, such attribution studies may add considerable depth to the forecast post-mortem process.

7.2 Connections of operations with research

As detailed in Chapter 5, extensive networks and an extensive infrastructure for the production of real-time seasonal forecasts operate at the global, regional and national levels between individual GPCs-LRF and the WMO through LC-LRFMME, RCCs and RCOFs. At the same time, vigorous research activities relating to climate predictability and prediction on seasonal and other timescales are occurring under WMO's WCRP and World Weather Research Programme (WWRP) and at research centres and universities worldwide. It is a key priority of WMO to promote two-way information exchange to inform forecast producers about research developments that may lead to new or improved forecast capabilities and products and to inform researchers about needs and gaps within the operational and user communities. Feedback from the operational community regarding modelling groups is also important in order to point out specific behaviour concerning the models (for example, biases in summer forecasts, and so forth).

Coordinating bodies under WCRP and WWRP with terms of reference and research programmes relevant to seasonal prediction include:

- The WCRP Working Group on Sub-seasonal to Interdecadal Prediction (WGSIP), which coordinates a programme of numerical experimentation for sub-seasonal to interdecadal variability and predictability, with an emphasis on assessing and improving predictions. WGSIP's programmes include the Climate-system Historical Forecast Project (CHFP), which archives hindcast datasets from many seasonal prediction systems as a resource for predictability and prediction research, as well as targeted research projects examining key topics such as the impact of land initialization, teleconnections, and model systematic errors;

- The WCRP Working Group on Coupled Modelling (WGCM), which fosters the development and evaluation of coupled climate and Earth system models, some of which are applied to seasonal prediction; and

- The WWRP/WCRP Working Group on Numerical Experimentation (WGNE), which fosters atmospheric model development, including refining numerical techniques and formulating atmospheric physics processes in models, for application to predictions on all timescales.

Advances in seasonal forecasting are also being supported by major research and service initiatives around the globe. These include projects supported by the European Commission's Horizon2020 and Copernicus programmes and by NOAA's Modeling, Analysis, Prediction, and Projections (MAPP) programme.

Mechanisms through which an exchange of information between the seasonal forecasting operational and research communities is being facilitated include:

- WMO Workshops on Operational Climate Prediction, the first two of which were held in Pune, India in 2015 and Barcelona, Spain in 2018;

- Major research conferences with themes linking research to operations, such as the International Conferences on Sub-seasonal to Decadal Prediction held in Boulder, USA in 2018; and

- Cross-membership and meeting attendance between WGSIP and the WMO Inter-Programme Expert Team on Operational Predictions from Sub-seasonal to Longer Timescales, which guides the coordination and development of WMO's seasonal forecasting infrastructure.

An overall trend in climate prediction research is a broadening from seasonal forecasting of basic meteorological variables to climate forecasting across sub-seasonal, seasonal and annual to decadal timescales, while developing cross-cutting approaches to fundamental aspects of climate prediction such as initialization, ensemble generation, and verification. For example, the joint WWRP/WCRP Subseasonal-to-Seasonal Prediction Project is developing a scientific foundation for sub-seasonal forecasting drawing upon hindcasts and real-time forecasts provided by participating operational centres, while the science of annual to decadal prediction is being advanced by the WCRP's Decadal Climate Prediction Project. Both of these developments feed into emerging WMO operational capabilities as discussed in Chapter 9.

An additional trend is that climate forecasts increasingly employ comprehensive climate models that simulate coupled interactions of the atmosphere, land, oceans, and sea ice and, in the case of Earth system models, ecosystems and biogeochemical cycles. This poses challenges to the observing and data assimilation systems that provide forecast initial conditions but also creates opportunities for forecasting elements of the climate system beyond basic meteorological variables. A further challenge having clear societal importance is predicting shifting risks of climate and weather extremes, both averaged over and within a verification period. All of these developments, along with advancements in modelling and methodology that improve forecast quality, will lead to the greater use of and benefits from climate predictions across timescales (National Academies of Sciences, 2016).

7.3 Exploring historical data

The sources of climate data can be classified as in situ, remote sensing, models (for example, reanalysis) and combinations of the above.

In situ data involves stations, weather balloons, buoys, ships and any other source of climate records acquired in a given location, including paleoclimate proxies. Although some station data go back approximately to the year 1600 (Overpeck et al., 2011), the most frequently available observational datasets start around 1860–1900. Paleoclimate data involves tree rings, ice cores, fauna and other sources to help understand climate variability at centennial to millennial timescales.

Remote sensing data is mostly acquired by satellites, but it is also acquired by devices on aircraft and, more recently, drones. Satellite data became available in the late 1960s, and although imperfect, it is considered very useful for monitoring the evolution of several environmental variables at high spatial and temporal resolution.

Climate models are used to reconstruct the evolution of the climate in the past. When systematically combined with in situ and satellite information, models produce "reanalyses" that are extremely useful representations of the climate system on Earth. Although they have some key biases, model-based reanalysis products are a major source of readily available data to improve our understanding of climate mechanisms and to make decisions. The current suite of real-time reanalysis efforts includes:

- NCEP/NCAR Reanalysis
- JRA-55
- ERA-5
- CFSR (http://cfs.ncep.noaa.gov/)
- MERRA (https://gmao.gsfc.nasa.gov/reanalysis/MERRA/)

Many reanalysis products are available in one place via the NASA CREATE project and the IRI Data Library: https://cds-cv.nccs.nasa.gov/CREATE-V/.

Given the problems related to data availability, quality, and spatial and temporal heterogeneity in multiple locations around the world, different types of combinations of data sources are used to create hybrid data products. Sophisticated interpolation techniques can use station and satellite information to create high resolution gridded datasets, and although these techniques have their own uncertainties, they can be used to estimate the behaviour of the variable of interest in places where no stations exist. Initiatives like Enhancing National Climate Services – ENACTS – (Dinku et al., 2016) have successfully merged station data, model output and satellite data in several countries in Africa, helping quality control and helping to homogenize and complete the observational record.

Although paleoclimate datasets bring their own share of uncertainties, they can be used in conjunction with local knowledge of the region under study to try to extend the historical record. These analyses require carefully constructed spatio-temporal models and a good understanding of the mechanisms involved in governing climate variability at multiple timescales.

To extend the record of observational data, international efforts are involved in attempting to recover historical observational data (often referred to as data mining) regarding surface terrestrial and marine global weather observations, for example, Atmospheric Circulation Reconstructions over the Earth (ACRE). Enhancements in the observational database as a result of data mining efforts subsequently feed into extending the record of reanalysis efforts to support an improved understanding of climate variability and validation of climate models over a longer period. As the quality of the models and techniques for data assimilation improve, reanalysis efforts are also periodically repeated over regular intervals.

CHAPTER 8. EXAMPLES OF GOOD PRACTICES CURRENTLY FOLLOWED AT NMHSS, RCCS AND RCOFS

This document provides recommendations on good practices for the development of forecasts at the regional/national level based on the concept of the cascading forecast process (Chapter 4). The recommendations stress the use of an objective process that is reproducible, traceable and verifiable. A set of overarching principles was also developed. In this chapter, we describe some examples of good current seasonal forecast practices that are in place and align with the principles and associated recommendations. The examples given below are by no means exhaustive.

8.1 Presenting and communicating seasonal forecasts

- The Met Office three-month outlook product for contingency planners provides a novel presentation of forecast uncertainties based on a display of ensemble members. The product also includes substantial contextual information on the historical climate, including an illustration of the observed month-to-month variability within the three-month season.
https://www.metoffice.gov.uk/services/public-sector/contingency-planners

- The Bureau of Meteorology, Australia provides a video presentation and explanation of the seasonal forecast, including the context of current conditions.
http://www.bom.gov.au/climate/outlooks/#/overview/video

- Climate Prediction Center, NOAA: The final seasonal forecast is associated with a text statement that describes (a) the current state of climate conditions (including those associated with sources of predictability) and (b) possible reasons (attribution) for forecast anomalies. Seasonal forecasts are presented in several formats – probability of tercile categories, entire probability density function in the form of probability of exceedance, and so forth. All seasonal forecasts and digital data are archived, and the history of verification for real-time seasonal forecasts is maintained and provided publicly.
http://www.cpc.ncep.noaa.gov/products/predictions/long_range/
http://www.cpc.ncep.noaa.gov/products/predictions/long_range/tools.html

- Environment and Climate Change Canada (ECCC) provides GPC-LRF Montreal seasonal forecasts up to 12 months in advance in an interactive format that simultaneously displays all three tercile probabilities plus a simple per cent correct skill measure at any location indicated by a mouse-positioned cursor or tap and allows the associated numerical data to be downloaded.
http://climate-scenarios.canada.ca/?page=cansips-global (Global)
http://climate-scenarios.canada.ca/?page=cansips-prob (Canada)

- The Mediterranean Climate Outlook Forum (MedCOF) produces consensus-based seasonal forecasts for the Mediterranean region including (a) a description of the current state of the climate, (b) a verification of the previous forecast and (c) seasonal forecasts expressed as a probability of tercile categories with a text statement. Further, in order to assist in selecting models to be used at the regional level, tables are provided for different variables, models and predetermined sub-regional domains showing verification scores (both deterministic and probabilistic) computed over a common hindcast period.
http://medcof.aemet.es/

- The Regional Climate Outlook Forum for Northern Africa, known by its French acronym PRESANORD, covers five countries in North Africa: Algeria, Egypt, Libya, Morocco and Tunisia. The forum generates the seasonal precipitation and temperature outlook for northern Africa using the probability and tercile categories.
http://rccnara1.marocmeteo.ma/moroccovlov.php

- The Tokyo Climate Centre’s Monthly Discussion on Seasonal Climate Outlook is mainly intended to assist NMHSs in the Asia-Pacific region in interpreting and assessing GPC-LRF Tokyo’s products for a three-month prediction and in understanding the current conditions of the climate system. Its sections include Summary and Discussion, Latest State of the Climate System, Three-month Predictions, Warm Season Predictions and Explanatory Notes.
https://ds.data.jma.go.jp/tcc/tcc/products/model/monthly_discussion/latest.pdf

8.2 Tailoring of seasonal forecasts

- The Caribbean Regional Climate Centre (RCC Caribbean), hosted by the Caribbean Institute for Meteorology and Hydrology (CIMH), generates Standardized Precipitation Index (SPI) forecasts with CPT for the purposes of drought watch and warning at the CariCOF.
ftp://ftp.cimh.edu.bb/CariCOF/CariCOF_outlooks/2018/ASO_outlook/technical/CARICOF_ASO-NDJ_2018_technical_outlook.pdf

- The International Research Institute for Climate and Society (Earth Institute/Columbia University, IRI) generates flexible seasonal forecasts based on the full estimate of the probability distribution. IRI routinely produces interactive maps and pointwise distributions of exceeding (or not exceeding) user selected percentiles for probabilistic temperature and precipitation forecasts.
http://iridl.ldeo.columbia.edu/maproom/Global/Forecasts/index.html

- The North EurAsia Climate Centre (NEACC), along with multimodel seasonal forecasts based on SL-AV (Tolstykh et al., 2015) and the Voeikov Main Geophysical Observatory’s dynamical models, generates seasonal forecasts of climate indices (Kiktev et al., 2015) such as the Eurasian Pattern Oscillation, West Atlantic Oscillation, West Pacific Oscillation, Pacific-North American Pattern, North Atlantic Oscillation, and Arctic Oscillation in support of seasonal outlook discussions.
(Access to the forecasts online requires a password; see http://neacc.meteoinfo.ru/neacc/links for contact information.)

- The National Meteorological Services of Colombia and Guatemala (IDEAM and INSIVUMEH, respectively) have recently implemented their NextGen system of seasonal climate forecasts, producing objective, calibrated and flexible predictions with CPT, using the recommendations of this Guidance Document.

8.3 User feedback

- User engagement and user-centred design requires specialized skills, including design methods, awareness of cognitive biases (Nicholls, 1999), accessibility (for example, providing access to the young, the visually or hearing impaired and the disadvantaged) and familiarity with the context in which forecasts might be used. User engagement, therefore, should include language consideration as many countries have large populations that are linguistically diverse. Further, engagement on product development should start early, and where possible, look to co-design methods in which NMHSs and their users work together to develop services.

- Some limitations in obtaining user feedback can be overcome by categorizing feedback by user attributes (age, location, and so forth), and large sample sizes do offset some concerns around quality. Social media such as Facebook and Twitter can provide additional means for feedback, allowing users to directly question, comment on, or promote services. Social media allows real-time feedback on services, for example by indicating high levels of user engagement or user frustration. As an example, an analysis of Twitter by the Australian Bureau of Meteorology has revealed that adding contextual (climate) information around extreme events (for example, when the last similar El Niño and related climate drivers occurred) greatly increases user engagement, as evidenced by metrics such as retweets and likes.

- Modern web technologies can allow sophisticated information to be gathered about how users interact with and consume climate services. This is a rapidly developing field of user metrics which can provide insights which are not obvious through surveys or other forms of feedback. For example, a high "bounce rate" (defined as the percentage of users who visit the website landing page then leave without browsing any further) might indicate that a product is not particularly user friendly, a dominance of referrals through Google may indicate that a site is difficult to navigate or not very accessible, and so forth. There are many specialized resources for using and interpreting the web for feedback and many businesses which can provide assistance with Google Analytics, a popular means for exploring web traffic.

- When resources allow, a tiered method for feedback is usually the best practice for bringing together the strengths of the various approaches. As an example, the Australian Bureau of Meteorology recently carried out extensive rounds of user engagement and feedback involving online surveys, interactive sessions and focused relationships with sophisticated users. This process formed the basis for a major improvement in seasonal forecast services, with an emphasis on the agriculture sector. Among the key findings were the following:

    - The lowest levels of satisfaction were expressed by those with the poorest understanding of seasonal forecasts, which suggested that many users were misapplying or misunderstanding the service;

    - The highest priority among users was higher accuracy, followed by local information, an ability to tailor information and new forecast products such as forecasts for extremes;

    - Users were often uncomfortable with technical terms; for example, users did not like "tercile forecasts" because the term was complex and unfamiliar;

    - For most users, the seasonal forecast was just a piece in a complex decision chain, meaning that it needed to be simple to use and accessible; and

    - Specialized users wanted access to raw gridded data to support decision-making tools which they were familiar with, including crop, pasture and streamflow models.

8.4 Managing change and continuity of operations

- Quality (or change) management is a key part of climate service delivery, and NMHSs are encouraged to follow the *Guidelines on Quality Management in Climate Services* (WMO-No. 1221). Quality management is a process for ensuring that all the activities necessary to design, develop, and deliver a product or service are conducted effectively and efficiently. It focuses not only on product and service quality but also on the means to achieve it. By utilizing quality assurance and control of functions and products, it is possible to achieve more consistent outputs for users.

- To support robustness in the generation and delivery of objective seasonal forecasts and services, the roles and responsibilities of individuals involved in preparing, delivering, receiving and acting on a climate service should be documented in the Standard Operating Procedures (SOPs). SOPs are used in many contexts and are not unique to meteorology or climatology. SOPs are step-by-step instructions compiled by an organization to help workers carry out complex routine operations, of which preparing a seasonal forecast is an example. SOPs aim to achieve efficiency, quality output and uniformity of performance while reducing miscommunication and failure. Accompanying SOPs should be an assessment of risk which documents the ways in which a service may fail or have an adverse outcome and strategies to manage those risks to a suitable level. It is not unusual for individual model runs to fail or data transfers to be affected by information technology issues, and SOPs would typically provide information about how these situations are to be managed.

8.5 Examples of evolving towards an objective seasonal forecast process

- South-West Indian Ocean Climate Outlook Forum (SWIOCOF) example: The proposed methodology for evolving the seasonal forecast process is based on the availability of seasonal forecasts provided by Global Seasonal Prediction Centres, the latest generation of climate models and statistical methods of downscaling. Dedicated software (coded under "R") and access to the necessary dataset through a web platform developed at the regional level by the Indian Ocean Commission (with technical support from Météo-France) are the core elements of the operational part. A guidance document is provided to NMHSs to support the application of the methodology to their islands (available in French and in English).

- Use of GCMs and calibration: The Intergovernmental Authority on Development (IGAD) Climate Prediction & Applications Centre (ICPAC) has integrated GPC-LRF data into seasonal outlooks and has evaluated the performance of GPC-LRF prediction systems and bias correction methods for the Greater Horn of Africa.
  https://rcc.icpac.net/index.php/long-range-forecast/regional-tailored-products/gcm

  ICPAC issued its first fully objective forecast based on GPC outputs at GHACOF-52 for the June–September 2019 season. An article with more information and a technical summary of the process can be found at https://medium.com/@icpac/improved-seasonal-forecast-for-eastern-africa-57872645f449.

CHAPTER 9. FUTURE PROSPECTS FOR SEASONAL AND OTHER LONG-RANGE FORECASTS

The discussion and recommendations provided in this document related to seasonal forecasts have kept in mind the current state of the operational infrastructure for seasonal predictions within the purview of WMO (as included in the *Manual on the Global Data-processing and Forecasting System*). This infrastructure is used to provide the monthly mean forecast fields (at a coarse horizontal resolution) for a limited number of meteorological variables from GPCs-LRF. The gridded data for these forecasts can be obtained either from individual GPCs-LRF (at least in some cases) or from LC-LRFMME. It is anticipated that in the future, advances in the science of seasonal predictions, increased data exchange capabilities (for example, provisions for the exchange of seasonal forecasts at daily temporal resolution), improved prediction systems (for example, higher resolution) and emerging forecast requirements at the global, regional and national levels – predictions relating to sea ice, the onset and withdrawal dates for monsoons, the number of days of extreme heat, and so forth – will lead to modifications in the recommendations provided in the current version of this guidance document.

Operational forecast capabilities are also emerging on sub-seasonal and decadal timescales. The *Manual on the Global Data-processing and Forecasting System* already contains designations for Global Producing Centres for Annual to Decadal Climate Predictions (GPCs-ADCP) and a corresponding Lead Centre for Annual to Decadal Climate Predictions (LC-ADCP). Including Global Producing Centres and Lead Centres for sub-seasonal forecasts in future versions of the *Manual* is also under discussion.

With these advances for long-range forecasts to complement the current infrastructure for seasonal forecasts, similar guidance for the cascading of forecast information from the global to the regional scale will need to be developed for other timescales. Although the many concepts provided and recommendations proposed in this document are also applicable to other LRF timescales, there are some aspects unique to specific timescales that will have to be considered. For example, the causes for underlying predictability are different for different timescales, and latency and update frequency also differ. Taking sub-seasonal forecasts as an example, one of the primary sources of predictability is MJO, the latency for forecast information is much shorter, and annual to decadal predictions are issued less frequently than either seasonal or sub-seasonal predictions.

Another development that will impact future updates to this guidance document is the emerging concept of seamless predictions and the evolution of the current GDPFS towards a seamless GDPFS. The concept of a seamless GDPFS addresses emerging requirements from services-oriented programmes, such as hydrology, aeronautics, marine services and agriculture, as well as requirements for implementing disaster mitigation strategies. To complement such advances towards a seamless GDPFS, appropriate guidance for objective tools will also need to be provided.

APPENDIX 1. GLOSSARY

Anomaly. The difference between an observed value of a meteorological variable (for example, seasonally averaged temperature) for a single period (for example, January–March (JFM) 2000) and its long-term average (for example, JFM 1961–1990). In the case of seasonally averaged temperature, a positive anomaly occurs when the temperature for the season in question is higher than average, and a negative anomaly occurs when the season is colder than average. Anomalies can also be calculated for model-based seasonal forecasts. [Text adapted from the *Guidance on Verification of Operational Seasonal Climate Forecasts*]

Attribution. An exercise to link observed seasonal mean atmospheric and terrestrial anomalies to their possible causes, for example, anomalies in sea-surface temperature. Given that atmospheric variability alone can also contribute to seasonal mean anomalies, not all observed anomalies can be attributed. The exercise of attribution helps understand limits and sources of seasonal predictability.

Bias correction. Altering seasonal forecasts based on systematic differences between a history of forecasts and corresponding observations. Systematic differences can be quantified based on a large set of retrospective forecasts (also referred to as hindcasts). Bias correction adjusts properties of the modelled climate to match those of the observed climate without reference to prediction quality or skill, in other words without pairing hindcasts and observations.

Bootstrapping. A means of estimating sampling errors in the value of a parameter (for example, a verification score) by resampling with replacement from the original dataset. Bootstrapping is recommended for estimating the uncertainty in each verification score given that the sample size of seasonal forecasts is generally very small. The procedure involves recalculating a verification score a large number of times and then examining the distribution of these values. Typically, the distribution is summarized by identifying one of the lowest and one of the highest score values (but not the absolute lowest and highest), and thus defining a range or "interval" (or percentiles, for example, the 2.5 and 97.5 percentiles to characterize 95% confidence) between which the true score is thought to lie. [Text adapted from the *Guidance on Verification of Operational Seasonal Climate Forecasts*]

Calibration. A systematic error correction of forecast anomalies that involves modifying forecast values to optimize hindcast quality or skill. Although this process may also correct certain model biases as in ensemble spread, calibration requires consideration of paired hindcast and observed values and thus differs fundamentally from bias correction, which only considers the statistics of the hindcasts and observations independently and adjusts the former towards the latter.

Cross-validation. A model validation or a forecast verification technique in which the sample being validated or verified is not included in computing population statistics. For example, for correcting forecast bias, the forecast for a particular season being corrected is not included in the estimating model's climatology based on hindcasts.

Dynamical Prediction Systems. A method for making seasonal predictions based on atmospheric or coupled general circulation models. Dynamical seasonal prediction systems require extensive computational resources and an extensive supporting infrastructure.

Downscaling. The process of increasing the spatial and/or temporal resolution of the original variable of interest in the forecasts in order to provide more detailed forecast information for use in various applications. Downscaling can be performed using empirical/statistical approaches or dynamical models, usually known as regional climate models, or hybrid methods involving both methodologies.

El Niño-Southern Oscillation. The El Niño-Southern Oscillation (ENSO) is the strongest mode of coupled variability on an interannual timescale in the tropical Pacific. ENSO is responsible for warming and cooling of sea-surface temperatures in the equatorial tropical Pacific and influences weather and climate variability over remote regions of the globe. ENSO is the primary source of our ability to make skilful seasonal predictions.

Empirical prediction systems. A method for making seasonal predictions based on quantifying statistical relationships between two variables in observational data. One variable in the statistical relationship is used to make the forecast (referred to as the predictor) and the other variable is the forecast (referred as the predictand).

Ensemble forecast. A collection of forecasts, starting from small differences in the initial condition, designed to sample uncertainty in future outcomes. The ensemble forecast technique can be used to assign probabilities to a range of possible future outcomes.

Forecast quality. A measure of the association between forecasts and corresponding observations. [Text adapted from the *Guidance on Verification of Operational Seasonal Climate Forecasts*]

Forecast lead time. The time difference between the forecast time issuance and the forecast time validity is defined as the forecast lead time. For example, a seasonal forecast issued on 1 January and valid for the forthcoming February-March-April (FMA) season is often referred to as a one-month lead forecast. In other words, the forecast provides information about the expected climate conditions in FMA one month in advance. The forecast valid for the March-April-May (MAM) season issued on the same date (1 January) is referred to as a two-month lead forecast (that is, it provides information about the expected climate conditions in MAM two months in advance).

Forecast skill. An attribute of forecast quality, specifically, a comparative measure of forecast quality in which a set of forecasts has positive skill if it scores better on one or more forecast attributes than another set, known as the reference set. Forecast skill is usually measured against a naïve forecasting strategy, such as random guessing, perpetual forecasts of one category, or climatological probabilities of all categories, but it can be calculated using any reference set. [Text adapted from the *Guidance on Verification of Operational Seasonal Climate Forecasts*]

Hindcast. Hindcasts, also known as historical forecasts or reforecasts, are produced using the same model and methodologies as a producing centre's real-time forecasts, but "predict" past climate states based on observational information available at the initial time of the hindcast. Hindcasts are used for (a) assessing the overall skill of the forecast system and (b) bias correction and calibration of real-time forecasts.

Hybrid prediction systems. A forecast method that uses a combination of dynamical and empirical prediction systems. For example, dynamical forecasts of an index (for instance, the amplitude of the ENSO sea-surface temperature anomaly) can be used as a predictor in an empirical forecast system.

Indian Ocean Dipole (IOD). An irregular oscillation characterized by anomalous cooling of sea-surface temperature in the south-eastern equatorial Indian Ocean and anomalous warming of the sea-surface temperature in the western equatorial Indian Ocean. Similar to ENSO, IOD also has some influence on the climate of the regions in the vicinity of the Indian Ocean.

Modes of climate variability. Seasonal climate variability can be characterized by preferred spatial patterns of variability that are referred to as "modes of climate variability". Such modes are often also involved in communicating the influence of boundary conditions over remote regions of the globe.

Predictability. An a priori estimate of our ability to make skilful forecasts. Predictability quantifies an inherent property of nature. It is not a forecast skill measure but rather an upper limit of predictive skill that a good forecast system can achieve. A good forecast tool helps realize inherent predictability as prediction skill.

Probability of exceedance. Probability that the observed value of the forecast variable will exceed a certain threshold. Seasonal forecasts provided in the probability of exceedance form allow a user to select the threshold and determine the forecast probability that the observed value will exceed that threshold.

Reliability. An attribute of the quality of probabilistic forecasts, specifically, the correspondence between the forecast probabilities and the conditional observed relative frequencies of events. Forecasts are reliable if, for all forecast probabilities, the observed relative frequency is equal to the forecast probability (that is, an event occurs on 40% of the occasions for which the forecast probability is 40%, 50% of the occasions for which the forecast probability is 50%, and so on). Forecast probabilities that deviate from the observed frequency of occurrence are either over- or under-confident.

Reliability curve. A plot of the conditional observed relative frequencies of events (on the y-axis) against forecast probability (on the x-axis).

Teleconnection. Teleconnections are recurrent large-scale anomaly patterns linking climate variability across remote regions. They are the building blocks of our ability to predict climate variability. In the context of seasonal forecasts at the regional level, it is important to understand and quantify which teleconnections influence local climate variability.

Tercile. One of two values that divide the distribution of data into three equal parts. The upper tercile is the higher of the two terciles and is frequently used to define the lower limit of the above-normal category. The lower tercile is frequently used to define the upper limit of the below-normal category. The normal category is bounded by the two terciles.

Verification. The measurement of the quality of a forecast or of a series of forecasts. [Text adapted from the *Guidance on Verification of Operational Seasonal Climate Forecasts*]

Validation. Checks necessary to ensure that a forecasting system or model is functioning as intended.

APPENDIX 2. ACRONYMS

AGCM	Atmospheric General Circulation Model
APEC	Asia-Pacific Economic Cooperation
APCC	APEC Climate Centre
AMO	Atlantic Multi-decadal Oscillation
C3S	Copernicus Climate Change Service
CBS	Commission for Basic Systems
CCA	Canonical Correlation Analysis
CCl	Commission for Climatology
CDF	Cumulative Distribution Function
CDS	Climate Data Store
CGCM	Coupled General Circulation Model
CHFP	Climate-system Historical Forecast Project
CIMH	Caribbean Institute for Meteorology and Hydrology
CORDEX	Coordinated Regional Climate Downscaling Experiment
CPC	Climate Prediction Centre
CPT	Climate Predictability Tool
CSIS	Climate Services Information System
CST	Climate Services Toolkit
DRM	Disaster Risk Management
EC-69	Sixty-ninth session of the WMO Executive Council
EOF	Empirical Orthogonal Function
ENACTS	Enhancing National Climate Services
ENSO	El Niño-Southern Oscillation
EUROSIP	European Seasonal to Inter-annual Prediction
GCM	General Circulation Model
GDPFS	Global Data-processing and Forecasting System
GFCS	Global Framework for Climate Services
GPCs-LRF	Global Producing Centres for Long-Range Forecasts

GSCU	Global Seasonal Climate Update
GTS	Global Telecommunication System
IMD	India Meteorological Department
IPET-OPSLS	CBS/CCl Inter-Programme Expert Team on Operational Predictions from Sub-seasonal to Longer-time Scales
IRI	International Research Institute for Climate and Society
IOD	Indian Ocean Dipole
ISMR	Indian Summer Monsoon Rainfall
KMA	Korea Meteorological Administration
LC-LRFMME	Lead Centre for Long-Range Forecast Multi-Model Ensemble
LRFs	Long-Range Forecasts
MAPP	NOAA's Modeling, Analysis, Prediction, and Projections programme
MCA	Maximum Covariance Analysis
MJO	Madden Julian Oscillation
MME	Multi-Model Ensemble
NAO	North Atlantic Oscillation
NCEP	National Centers for Environmental Prediction
NCOF	National Climate Outlook Forum
NEACC	North EurAsia Climate Centre
NMHS	National Meteorological and Hydrological Service
NMME	North American Multi-Model Ensemble
NOAA	National Oceanographic and Atmospheric Administration
NPO	North Pacific Oscillation
NWP	Numerical Weather Prediction
PCR	Principal Component Regression
PDF	Probability Density Function
PDO	Pacific Decadal Oscillation
PNA	Pacific North American
POE	Probability of Exceedance
RAs	WMO Regional Associations

RCC	Regional Climate Centre
RCOF	Regional Climate Outlook Forum
ROC	Relative Operating Characteristic
SAM	Southern Annular Mode
SIOD	Sub-tropical Indian Ocean Dipole
SNR	Signal-to-Noise Ratio
SO	Southern Oscillation
SPI	Standardized Precipitation Index
SST	Sea-surface Temperature
SVSLRF	Standardized Verification System for Long-Range Forecasts
TAO	Tropical Atmosphere Ocean
TOGA	Tropical Ocean Global Atmosphere programme
WCRP	World Climate Research Programme
WGCM	Working Group on Coupled Modelling
WGNE	Working Group on Numerical Experimentation
WGSIP	Working Group on Sub-seasonal to Interdecadal Prediction
WIS	WMO Information System
WMO	World Meteorological Organization
WWRP	World Weather Research Programme

APPENDIX 3. RESOURCES

Books

Simon J. Mason – Guidance on Verification of Operational Seasonal Climate Forecasts
https://library.wmo.int/index.php?lvl=notice_display&id=20618#.XukJA25uI2y

Daniel S. Wilks: Statistical Methods in the Atmospheric Sciences
https://www.amazon.com/Statistical-Atmospheric-Sciences-International-Geophysics/dp/0123850223

Hans von Storch and Francis W. Zwiers: Statistical Analysis in Climate Research
https://www.amazon.com/Statistical-Analysis-Climate-Research-Storch-ebook/dp/B00CF0JKZE

Ian T. Jolliffe and David B. Stephenson: Forecast Verification. A Practitioner's Guide to Atmospheric Science
https://www.amazon.com/Forecast-Verification-Practitioners-Atmospheric-Science/dp/0471497592

Huug van den Dool: Empirical Methods in Short-Term Climate Prediction https://www.amazon.com/Empirical-Methods-Short-Term-Climate-Prediction/dp/0199202788

Training modules

EUMeTrain Training Module on Forecast Verification: http://www.eumetrain.org/data/4/451/english/courses/msgcrs/index.htm

Seasonal Forecast Course Package T.O.P.: https://sites.google.com/view/top-seasonal-forecast/home

Statistical Tools

Ensemble Verification Metrics: https://www.ecmwf.int/sites/default/files/elibrary/2017/17626-ensemble-verification-metrics.pdf

WWRP/WGNE Joint Working Group on Forecast Verification Research: https://www.wmo.int/pages/prog/arep/wwrp/new/jwgfvr.html

CPT: https://iri.columbia.edu/our-expertise/climate/tools/cpt/

Climate Services Toolkit: https://www.wmo.int/pages/prog/wcp/ccl/meetings/ICT-CSIS/documents/2016/presentations/CST_Presentation_Dec_4_2016.pdf

General

S2S Prediction Project: http://s2sprediction.net/

WMO GSCU page: https://public.wmo.int/en/our-mandate/climate/global-seasonal-climate-update

Manual on the Global Data-processing and Forecasting System: https://library.wmo.int/index.php?lvl=notice_display&id=12793 - .XujDU2ozZMA

Standardized Verification System for Long-Range Forecasts: https://www.wmo.int/pages/prog/www/DPS/LRF/ATTACHII-8SVSfrom%20WMO_485_Vol_I.pdf

CPC's attribution page: https://www.cpc.ncep.noaa.gov/products/people/mchen/AttributionAnalysis/

GFCS implementation plan: https://www.wmo.int/gfcs/implementation-plan

Datasets

NMME: https://www.cpc.ncep.noaa.gov/products/NMME/

C3S: https://climate.copernicus.eu/seasonal-forecasts

LC-LRFMME: https://wmolc.org/

Climate Explorer: https://climexp.knmi.nl/start.cgi

ESRL PSL: https://www.esrl.noaa.gov/psd/data/

NCAR Climate Data Guide: https://climatedataguide.ucar.edu/

IRI Data Library: https://iridl.ldeo.columbia.edu

NOAA/NCEI: https://www.ncdc.noaa.gov/cdo-web/

User interaction

Pacific Climate Crab: https://www.pacificclimatechangescience.org/pacific-adventures-of-the-climate-crab

Australian Climate Dogs: http://agriculture.vic.gov.au/agriculture/weather-and-climate/understanding-weather-and-climate/climatedogs

NOAA ENSO Blog: https://www.climate.gov/news-features/department/enso-blog

APPENDIX 4. REFERENCES

Ailliot P., D. Allard, V. Monbet., P. Naveau, 2015: Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistique*, 156:101–113.

Baldwin, M.P. and T.J. Dunkerton, 2001: Stratospheric Harbingers of Anomalous Weather Regimes. *Science*, 294:581–584.

Barnard, P.L., J. Allan, J.E. Hansen, G.M. Kaminsky, P. Ruggiero, and A. Doria, 2011: The impact of the 2009–10 El Niño Modoki on U.S. West Coast beaches, *Geophys*. Res. Lett., 38, doi:10.1029/2011GL047707.

Barnston, A.G., A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal predictions practices through attribution of climate variability. *Bull. Ame. Meteor. Soc.*, 85:59–72.

Barnston, A.G., 1994: Linear short-term climate predictive skill in the northern hemisphere. *J. Climate*, 7:1513–1564.

Barnston, A.G. and R.E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, 11:1083–1126.

Barnston, A. and Co-authors, 1999: NCEP forecasts of the El Niño of 1997-98 and its U.S. impacts. *Bull. Amer. Meteo. Soc.*, 80:1829–1852.

Barnston, A.G. and M.K. Tippett, 2017: Do Statistical Pattern Corrections Improve Seasonal Climate Predictions in the North American Multimodel Ensemble Models? *Journal of Climate*, 30:8335–8355.

Banerjee, S.K., 1950, Methods of foreshadowing monsoon and winter rainfall in India, *Indian J. Met. Geophysics*, 1:4–14.

Bauer P., A. Thopre and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, 525:47–55.

Becker, E., H. Van den Dool and Q. Zhang, 2014: Predictability and forecast skill in NMME. *J. Climate*, 27:5891–5906, https://doi.org/10.1175/JCLI-D-13-00597.1

Becker, E. and H. Van den Dool, 2016: Probabilistic seasonal forecasts in the North American Multimodel Ensemble: A baseline skill assessment. J. Climate, 29:3015–3026, doi:https://doi.org/10.1175/JCLI-D-14-00862.1

Bengtsson, L., U. Schlese, E. Roeckner, M. Latif, T. P. Barnett and N. Graham, 1993: A Two-Tiered Approach to Long-Range Climate Forecasting. *Science*. 261:1026–1029.

Bjerknes, J., 1966: A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, 18:820–829.

Bejerkes, J., 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.*, 97:163–172.

Blanford, H.F., 1884, On the connexion of the Himalayan snow with dry winds and seasons of droughts in India. *Proceedings of the Royal Society of London*, 37:3–22.

Blunden, J., D.S. Arndt, and G. Hartfield (eds.), 2018: State of the Climate in 2017. *Bull. Amer. Meteor. Soc.*, 99:8, Si–S310, doi:10.1175/2018BAMSStateoftheClimate.1.

Brown, B.G.; A. H. Murphy, 1996: Improving forecasting performance by combining forecasts: The example of road-surface temperature forecasts. *Meteorological Applications*, 3:257–265.

Cai, M., and Co-authors, 2009: The Role of Long-Term Trends in Seasonal Predictions: Implication of Global Warming in the NCEP CFS. *Weather and Forecasting*, 24:965–973.

Charney, J. G. and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics*, Editors: Sir James Lighthill and R.P. Pearce, Cambridge University Press, pp. 99–109.

Chen and Kumar, A., 2018: Winter 2015/16 atmospheric and precipitation anomalies over North America: El Niño response and the role of noise. *Mon. Wea. Rev.*, 146:909–927,

Coelho C.A.S., D. B. Stephenson, M. Balmaseda, F.J. Doblas-Reyes and G.J. van Oldenborgh, 2006*a*: Towards an integrated seasonal forecasting system for South America. *J. Climate*, 19:3704–3721.

Coelho C.A.S., D. B. Stephenson, F.J. Doblas-Reyes, M. Balmaseda, A. Guetter and G.J. van Oldenborgh, 2006*b*: A Bayesian approach for multi-model downscaling: Seasonal forecasting of regional rainfall and river flows in South America. *Meteorological Applications*, 13:73–82.

Coelho, C.A.S., 2013: Comparative skill assessment of consensus and physically based tercile probability seasonal precipitation forecasts for Brazil. *Meteorological Applications*, 20:236–245.

Coelho C.A.S., M.A.F. Firpo, A.H.N. Maia, C. MacLachlan, 2017: Exploring the feasibility of empirical, dynamical and combined probabilistic rainy season onset forecasts for São Paulo, Brazil. *International Journal of Climatology*, 37:398–411.

Coelho, C.A.S., I.A.F. Cavalcanti, S. M.S. Costa, S.R. Freitas, E.R. Ito, G. Luz, A.F. Santos, C.A. Nobre, J.A. Marengo and A.B. Pezza, 2012: Climate diagnostics of three major drought events in the Amazon and illustrations of their seasonal precipitation predictions. *Meteorological Applications*, 19:237–255.

Colose, C.M., A.N. LeGrande and M. Vuille, 2016: Hemispherically asymmetric volcanic forcing of tropical hydroclimate during the last millennium. *Earth System Dynamics*, 7: 681–696.

Coumou, D. and S. Rahmstorf, 2012: A decade of weather extremes. *Nature Climate Change*, 2:491–496.

Davey, M.K., A. Brookshaw and S. Ineson, S. 2014: The probability of the impact of ENSO on precipitation and near-surface temperature. *Climate Risk Management*, 1:5–24.

DelSole, T., L. Jia and M.K. Tippett, 2013: Scale-selective ridge regression for multimodel forecasting, *J. Climate*: 26, 7957–7965.

Doblas-Reyes, F.J., R. Hagedorn, T.N. Palmer, and J-J. Morcrette, 2006: Impact of increasing greenhouse gas concentration in seasonal ensemble forecasts. *Geophys. Res. Lett.*, 33, L07708. doi:10.1029/2005GL025061.

Dobrynin M., D. I. V. Domeisen, W.A. Müller, L. Bell, S. Brune, F. Bunzel, A. Düsterhus, K. Fröhlich, H. Pohlmann, J. Baehr, 2018: Improved teleconnection-based dynamical seasonal predictions of boreal winter. *Geophysical Research Letters*, 43:3605–3614.

Dinku, T. et al., 2016: The ENACTS Approach: Transforming climate services in Africa one country at a time', World Policy Papers, pp. 1–24. doi:10.1037/t07885-000.

Drosdowksy, W. and L.E. Chambers 2001: Near-Global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *J. Climate*, 14:1677–1687.

Drosdowksy, W. and M. Wheeler, 2014: Predicting the onset of the North Australian wet season with the POAMA dynamical prediction system. *Weather Forecasting*, 29:150–161.

Easterling, D.R. and Co-authors, 2000: Climate Extremes: Observations, modeling and impacts. *Science*, 289:2068-2074.

George, S.E. and R.T. Sutton, 2006: Predictability and skill of boreal winter forecasts made with the ECMWF Seasonal Forecasting System II. *Quart. J. Roy. Meteorol. Soc.*, 132:2031–2053. doi:10.1256/qj.04.180

Giorgi, F. and W.J. Gutowski Jr., 2015: Regional Dynamical Downscaling and the CORDEX Initiative, *Annu. Rev. Env. Resour.*, 40:467–490, doi:10.1146/annurev-environ-102014-021217.

Gleixner S., N.S. Keenlyside, T.D. Demissie, F. Counillon, Y. Wang and E. Viste, 2017: Seasonal predictability of Kiremt rainfall in coupled general circulation models. *Environ. Res. Lett.*, 12:114016, https://doi.org/10.1088/1748-9326/aa8cfa.

Goddard, L., S.J. Mason, S.E. Zebiak, C.F. Ropelewski, R. Basher and M.A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climato.*, 9:1111–1152.

Graham, R.J., A.D.L. Evans, K.R. Mylne, M.S.J. Harrison and K.B. Robertson, 2000: An assessment of seasonal predictability using atmospheric general circulation models. *Quarterly Journal of the Royal Meteorological Society*, 126:2211–2240.

Graham, R., and Co-authors, 2011: Long-range forecasting and Global Framework for Climate Services. *Climate Research*, 47:47-55, doi:10.3354/cr00963.

Greene, A.M, L. Goddard and R. Cousin, 2011. Web tool deconstructs variability in twentieth-century climate, *Eos Trans.* 397, doi:10.1029/2011EO450001

Grossmann, I. and P.J. Klotzbach, 2009: A review of North Atlantic modes of natural variability and their driving mechanisms, *J. Geophys. Rea.*, 114, https://doi.org/10.1029/2009JD012728

Hagedorn, R., F.J. Doblas-Reyes and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept. *Tellus A, 57*:219–233.

Hameed, S. and T. Yamagata, 2003: Possible impacts of Indian Ocean Dipole Mode events on global climate. *Climate Research*, 25:151–169.

Han, E. and A.V.M. Ines, 2017: Downscaling probabilistic seasonal climate forecasts for decision support in agriculture: A comparison of parametric and non-parametric approach. *Climate Risk Management*, 18:51–65. doi:10.1016/j.crm.2017.09.003

Harrison, M., A. Troccoli, M. Coughlan and J.B. Williams, 2008: Seasonal Forecasts in Decision Making. In: *Seasonal Climate: Forecasting and Managing Risk* (A. Troccoli, M. Harrison, D.L.T. Anderson and S.J. Mason, eds.), Dordrecht, Springer.

Hendon, H., E.-P. Lim and H. Nguyen, 2014: Seasonal Variations of Subtropical Precipitation Associated with the Southern Annular Mode. *J. Climate*, 27:3446–3460.

Hewitt, C.D., R.C. Stone and A.B. Tait, 2017: Improving the use of climate information in decision-making, *Nature Climate Change*, 7:614–617, https://www.nature.com/articles/nclimate3378.

Hoerling, M.P., A. Kumar and M. Zhong, 1997: El Niño, La Niña and the nonlinearity of their teleconnections. *J. Climate*, 10:1769–1786.

Horel, J.D., and J.M. Wallace 1981: Planetary-Scale Atmospheric Phenomena Associated with the Southern Oscillation. Mon. Wea. Rev., 109, 813–829. *J. Atmos. Sci.*, 45:1228–1251.

Hoskins, B., 2013: The potential for skill across the range of the seamless weather-climate prediction problem: a stimulus for our science. *Q. J. R. Meteorol. Soc.*, 139:573–584.

Hoskins, B.J. and D.J. Karoly, 1981: The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing, *J. Atmos. Sci.*, 38:1179–1196.

Huang, J., H.M. vanden Dool and A.G. Barnston, 1996: Long-lead seasonal temperature predictions using optimal climate normal. *J. Climate*, 9:809–817.

Hudson, D., 2017: Ensemble Verification Metrics, https://www.ecmwf.int/sites/default/files/elibrary/2017/17626-ensemble-verification-metrics.pdf.

Hudson, D., Shi, L., Alves, O., Zhao, M., Hendon, H.H. and Young, G. 2017. Performance of ACCESS-S1 for key horticultural regions. Bureau Research Report, No. 20. Bureau of Meteorology Australia, http://www.bom.gov.au/research/research-reports.shtml

Intergovernmental Panel on Climate Change, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Ji, M., A. Kumar and A. Leetmaa 1994: A multiseason climate forecast system at the National Meteorological Center. *Bull. Amer. Meteo. Soc.*, 75:569–577.

JGR-Oceans, 1998: The TOGA Decade: Reviewing the Progress of El Niño Research and Prediction. 103:14167–14510.

Jolliffe, I.T. and D. B. Stephenson, 2012: Forecast Verification: A Practitioner's Guide in Atmospheric Science. 2nd Edition, Wiley-Blackwell, Oxford.

Kharin V.V., F.W. Zwiers and N. Gagnon, 2001: Skill of seasonal hindcasts as a function of the ensemble size. *Climate Dynamics*, 17:835–843.

Kharin, V.V., W.J. Merryfield, G. J. Boer and W.-S. Lee, 2017: A post-processing method for seasonal forecasts using temporally and spatially smoothed statistics. *Mon. Wea. Rev.*

Kilavi, M. and Colman, A.W. 2012: Integration of statistical and dynamical methods in seasonal forecasting: an example for Kenya. WCRP Africa Newsletter, November 2012.

Kiktev D.B., E.N. Kruglova and I.A. Kulikova, 2015: Large-scale modes of atmospheric variability. Part 1. Statistical analysis and hydrodynamic modeling, *Meteorology and Hydrology*, 40:5–22

Kipkogei, O., A.M. Mwanthi, J.B. Mwesigwa, Z.K.K. Atheru, M.A. Wanzala and G. Artan, 2017: Improved Seasonal Prediction of Rainfall over East Africa for Application in Agriculture: Statistical Downscaling of CFSv2 and GFDL-FLOR. *Journal of Applied Meteorology and Climatology.* 56:3229–3243.

Kirtman, B., et al., 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction., *Bull. Amer. Meteo. Soc.*, 95:585–601.

Kozak, M., J. Hartley, J. Wnuk and M. Tartanus, 2015: Multiple Pie Charts: Unreadable, Inefficient, and Over-Used. Journal of Scholarly Publishing. 46:282–289. DOI:10.3138/jsp.46.3.05.

Kumar, A. and J. Zhu, 2018: Spatial Variability in Seasonal Prediction Skill of SSTs: Inherent Predictability or Forecast Errors? *J. Climate* 31, 613–621, https://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0279.1

Kumar, A. and M.P. Hoelring, 1995: Prospects and Limitations of Seasonal Atmospheric GCM Predictions. *Bull. Ame. Meteor. Soc.*, 76:335–345.

Kumar, A., and M.P. Hoerling, 2000: Analysis of a conceptual model of seasonal climate variability and implications for seasonal predictions. *Bull. Amer. Meteor. Soc.*, 81 255–264.

Kumar, A., 2009: Finite samples and uncertainty estimates for skill measures for seasonal predictions. *Mon. Wea. Rev.*, 137:2622–2631.

Kumar, K.K., B. Rajagopalan, M. Hoerling, G. Bates and M. Cane, 2006: Unraveling the mystery of Indian monsoon failure during El Niño. Science, 314:115–119.

Lorenz, E.N., 1965: A study of atmospheric predictability of a 28-variable atmospheric model. *Tellus A*, 17:321–333.

Lorenz, E.N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus A*, 21:289–307.

Li, S., L. Goddard, and D.G. DeWitt, 2008. Predictive skill of AGCM seasonal climate forecasts subject to different SST prediction methodologies. *J. Climate*, 21:2169–2186.

Lim, E.-P., H. Hendon and H. Rashid, 2013: Seasonal predictability of the southern annular mode due to Its association with ENSO. *J. Climate*, 26:8037–8034.

Liu, Z. and E. Di Lorenzo, 2018: Mechanisms and predictability of Pacific Decadal Oscillation, *Current Climate Change Reports*, 4:128–144.

MacLachlan C., A. Arribas, K.A. Peterson, A. Maidens, D. Fereday, A.A. Scaife, M. Gordon, M. Vellinga, A. Williams, R. E. Comer, J. Camp and P. Xavier, 2015. Description of GloSea5: the Met Office high resolution seasonal forecast system. *Q. J. R. Met. Soc.*, DOI:10.1002/qj.2396.

MacLeod, D., 2018: Seasonal predictability of onset and cessation of the east African rains. *Weather and Climate Extremes*, 21:27–35.

Madden, R.A. and P.R. Julian, 1971: Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific. *J. Atmos. Sci.*, 28:702–708.

Manzanas, R., J.M. Gutiérrez, J. Fernández, E. van Meijgaard, S. Calmanti, M.E. Magariño, A.S. Cofiño and S. Herrera, 2018: Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications. *Climate Services*, 9:44–56.

Mason, S.J. and L. Goddard, 2001: Probabilistic precipitation anomalies associated with ENSO. *Bull. Am. Meteorol. Soc.* 82, 619-638.

Mason, S.J., L. Goddard, N.E. Graham, E. Yulaeva, L. Sun, and P.A. Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Nino event. *Bull. Amer. Meteo. Soc.*, 80:1853–1873.

Mason, S.J., 2008: From dynamical predictions to seasonal forecasts. In Troccoli, A., M.S.J. Harrison, D.L.T. Anderson, and S.J. Mason (eds.), *Seasonal Climate Variability: Forecasting and Managing Risk*, Springer Academic Publishers, Dordrecht, 205–234.

Mason, S.J. and O. Baddour, 2008: Statistical Modelling. in Troccoli, A. et al., (eds.) *Seasonal Climate: Forecasting and Managing Risk*. Dordrecht: Springer Netherlands (Nato Science Series: IV: Earth and Environmental Sciences). doi:10.1007/978-1-4020-6992-5.

Mason, S.J., and S. Chidzambwa, 2009: Verification of African RCOF Forecasts. Position Paper, IRI Technical Report 09-02, International Research Institute for Climate and Society. Palisades, NY

Mason, S.J and M.K. Tippet, 2018: Climate Predictability Tool version 15.7.6. https://doi.org/10.7916/D8475T49

Massonnet et al., 2016: Using climate models to estimate the quality of global observational data sets. *Science*, 354:452-455, https://doi.org/10.1126/science.aaf6369

Mauboussin A. and M. J. Mauboussin, 2018 If You Say Something Is "Likely," How Likely Do People Think It Is? Harvard Business Review, https://hbr.org/2018/07/if-you-say-something-is-likely-how-likely-do-people-think-it-is

McPhaden, M., 2004: Evolution of the 2002/03 El Niño. *Bull. Amer. Meteo. Soc.*, 85:667-695.

McPhaden, M., A.J. Busalacchi and D.L.T. Anderson, 2010: A TOGA perspective, *Oceanography*, 23:85–103.

Meehl, G.A. and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in 21st century. *Science*, 305:994–997.

Min, Y.M., V.N. Kryjov, and C.K. Park, 2009: A probabilistic multimodel ensemble approach to seasonal prediction, *Weather Forecast*, 24:812–828.

Min, Y.M., V.N. Kryjov and S.M. Oh, 2014: Assessment of APCC multi-model ensemble prediction in seasonal climate forecasting: ret-rospective (1983–2003) and real-time forecasts (2008–2013). *J. Geophys Res.*, 119:12132-12150.

Molteni F., T. Stockdale, M. Balmaseda, G. Balsamo, R. Buizza, L. Ferranti, L. Magnusson, K. Mogensen, T. N. Palmer and F. Vitart F. 2011. The New ECMWF Seasonal Forecast System (System 4), Technical Memorandum 656. ECMWF: Reading, UK.

Molteni, F., T.N. Stockdale and F. Vitart, 2015: Understanding and modelling extra-tropical teleconnections with the Indo-Pacific region during the northern winter. *Climate Dynamics*, 45:3119–3140.

Moron V., A.W. Robertson, M.N. Ward and P. Camberlin, 2007: Spatial coherence of tropical rainfall at the regional scale. *J. Climate*, 20:5244–5263.

Moron, V., A.W. Robertson, M.N. Ward, and O. Ndiaye, 2008: Weather types and rainfall over Senegal. Part I: observational analysis. *J. Climate*, 21, 266–287.

Moron, V., A. Lucero, F. Hilario, B. Lyon, A. W. Robertson and E.D. Dewitt, 2009: Spatio-temporal variability and predictability of summer monsoon onset over the Philippines. *Climate Dyn.*, 33:1159–1177.

Muñoz, Á.G., L. Goddard, A.W. Robertson, Y. Kushnir and W. Baethgen, 2015: Cross–time scale interactions and rainfall extreme events in southeastern South America for the austral summer. Part I: Potential predictors. *J. Climate*, 28:7894–7913.

Muñoz, Á.G., L. Goddard, S.J. Mason and A.W. Robertson, 2016: Cross-time scale interactions and rainfall extreme events in southeastern South America for the austral summer. Part II: Predictive skill. *J. Climate*, 29:5915–5934.

Newman, M., and Co-authors, 2016: The Pacific Decadal Oscillation, Revisited. *J. Climate*, 29:4399–4427.

Nicholls, N., 1999: Cognitive illusions, heuristics, and climate prediction. *Bull. Amer. Meteorol. Soc.*, 80:1385–1398.

Nikulin G. and Co-authors, 2018: The effects of 1.5 and 2 degrees of global warming on Africa in the CORDEX ensemble. *Environ. Res. Lett.*, 13:Number 6.

National Academies of Sciences, Engineering, and Medicine. 2016: Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts. Washington, DC: The National Academies Press. doi:10.17226/21873.

Overpeck, J., G. Meehl, S. Bony, and D. Easterling, 2011: Climate data challenges in the 21st century. *Science*, 331:700–702.

Pai, D.S., A. Suryachandra Rao, S. Senroy, M. Pradhan, P. A. Pillai, and M. Rajeevan, 2017, Performance of the operational and experimental long-range forecasts for the 2015 southwest monsoon rainfall, *Current Science*, 112:68–75.

Palmer, T.N. and D. L. T. Anderson, 1994: The prospects for seasonal forecasting – A review paper. *Quart. J. Roy. Meteo. Soc.*, 518:755-793.

Palmer M.A. and Co-authors, 2005: Standards for Ecologically Successful River Restoration. *Journal of Applied Ecology*, 42:208–217.

Pausata, F.S., L. Chafik, R. Caballero and D.S. Battisti, 2015: Impacts of high-latitude volcanic eruptions on ENSO and AMOC. Proceedings of the National Academy of Sciences of the United States of America, 112:13784–13788. doi:10.1073/pnas.1509153112

Pinto, J.G., M. Reyers and U. Ulbrich, 2011: The variable link between PNA and NAO in observations and in multi-century CGCM simulations. *Climate Dynamics*, 36:337–354.

Philander, S. G. H. 1990: *El Niño, La Niña, and the Southern Oscillation*. Academic Press.

Potts J. M., C. K. Folland, I. T. Jolliffe and D. Sexton, 1996: Revised 'LEPS' scores for assessing climate model simulations and long-range forecasts. *J. Climate*, 9:34–53.

Preethi B., T. P. Sabin, J. A. Adedoyin and K. Ashok, 2015: Impacts of the ENSO Modoki and other Tropical Indo-Pacific Climate-Drivers on African Rainfall. Science Reports, 5:doi:10.1038/srep16653.

Rasmussen, E. M. and T. H. Carpenter, 1983: The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka. *Mon. Wea. Rev.*, 111:517–528.

Rajeevan M., D. S. Pai, R. Anil Kumar, and B. Lal, 2007, New statistical models for long-range forecasting of southwest monsoon rainfall over India, *Climate Dynamics*, 28:813–828.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, 126:649–667.

Robertson, A.W., S. Kirshner and P. Smyth, 2004: Downscaling of daily rainfall occurrence over Northeast Brazil using a hidden Markov model. *J. Climate*, 17:4407–4424.

Ropelewski, C. F. and M. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Nino/Southern Oscillation. *Mon. Wea. Rev.*, 110:1606–1626.

Saha, S. and Co-authors, 2006: The NCEP Climate Forecast System. *J. Climate*, 19:3483–3517.

Saha, S. and Co-authors, 2014: The NCEP Climate Forecast System Version 2. *J. Climate*, 27:2185–2208.

Sahu, N. and Co-authors, 2016: Probabilistic seasonal streamflow forecasts of the Citarum River, Indonesia, based on general circulation models, *Stoch. Env. Res. Risk A.*, 31:1747–1758, https://doi.org/10.1007/s00477-016-1297-4.

Sardeshmukh, P. D. and B. Hoskins, 1988: The Generation of Global Rotational Flow by Steady Idealized Tropical Divergence. *J. Atmos. Sci.*, 45: 1228–1521.

Scaife, A.A., J.F. Knight, G.K. Vallis and C.K. Folland, 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophys. Rea. Lett.*, https://doi.org/10.1029/2005GL023226.

Scaife, A.A. and Co-authors, 2014: Skillful long-range prediction of European and North American winters, *Geophys. Res. Lett.*, 41:2514–2519.

Scinocca, J.F. and Co-authors, 2016: Coordinated global and regional climate modeling. *J. Climate*, 29:17–35.

Shukla S., C. Funk and A. Hoell, 2014: Using constructed analogs to improve the skill of National Multi-Model Ensemble March–April–May precipitation forecasts in equatorial East Africa. *Environmental Research Letters*, Volume 9, Number 9.

Siegert S. and D. Stephenson, 2018: Forecast recalibration and multi-model combination, In: Sub-seasonal to Seasonal Prediction: The Gap Between Weather and Climate Forecasting, A. Robertson and F. Vitart (Eds), Elsevier, ISBN: 978-0-12-811714-9.

SPECS, 2016: "Methodologies for calibration and combination of global and downscaled s2d predictions", SPECS project report, http://www.specs-fp7.eu/sites/default/files/u1/SPECS_D53.1.pdf.

Stephenson, D.B., C.A.S. Coelho, F.J. Doblas-Reyes, and M. Balmaseda, 2005: Forecast assimilation: A unified framework for the combination of multi-model weather and climate predictions. *Tellus*, 57:253–264.

Stockdale, T.N., D.L.T. Anderson, J.O.S. Alves, and M.A. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature*, 392:371–373.

Teuling, A.J., R. Stöckli, and S.I. Seneviratne, 2011. "Bivariate Colour Maps for Visualizing Climate Data." *International Journal of Climatology*, 31:1408–1412.

Thompson, D.W., and J.M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, 25:1297–1300.

Tolstykh, M. and Co-authors, 2015: Development of the Multiscale Version of the SL-AV Global Atmosphere Model. *Russian Meteorology and Hydrology*, 40:374–382.

Trenberth, E.K., G.W. Branstrator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.-Oceans*, 107:14291–14324.

van den Dool, H.M., E. Becker, L.-C. Chen, and Q. Zhang, 2017: The probability anomaly correlation and calibration of probabilistic forecasts. *Wea. Forecasting*, 32:199–206

Unger, D., H. van den Dool, E. O'Lenic, and D. Collins, 2009: Ensemble regression, *Mon. Weather Rev.*, 137:2365–2379.

Vellinga M., A. Arribas and R. Graham, 2013: Seasonal forecasts for regional onset of the West African monsoon, *Climate Dynamics*, 40:3047–3070.

Walker, G.T., 1908: Correlation in seasonal variation of climate (Introduction). Mem. India Meterol, Dept., 20:117–124.

Walker, G.T., 1910: Correlation in seasonal variations of weather, II, Mem. India Meteorol. Dept., 21:22–30.

Walker, G.T., 1914: A further study of relationships with Indian monsoon rainfall, II, Mem. India Meteorol. Dept., 23:123–129.

Walker, G.T., 1923, Correlation in seasonal variations of weather, VIII, A preliminary study of world weather, Mem. India Meteorol. Dept., 24:75–131.

Wallace, J.M. and D.S. Gutzler, 1981: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. *Mon. Wea. Rev.*, 109:784–812.

Wang, W.X.D., A. Watkins, D. Jones, 2019: A user-oriented forecast verification metric: Weighted Percent Correct. *Meteorol. Z. (Contrib. Atm. Sci.)*, PrePub DOI:10.1127/metz/2019/0882.

Weigel, A.P., M. A. Liniger and C. Appenzeller, 2008: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Q J R Meteorol Soc.*, 134:241–260.

Weigel, A.P., R. Knutti, M. A. Liniger, and C. Appenzeller, 2010: Risks of model weighting in multimodel climate projections, *J. Climate*, 23:4175–4191.

Weisheimer, A., N. Schaller, C. O'Reilly, D. A. MacLeod and T. N. Palmer, 2017: Atmospheric seasonal forecasts of the twentieth century: Multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution. *Quarterly Journal of the Royal Meteorological Society*, 143:917–926.

Wilby, R.L., C.W. Dawson and E.M. Barrow, 2002: SDSM - a decision support tool for the assessment of regional climate change impacts, *Environ. Modell. Software*, 17:147–159.

Williams, P.D., and Co-authors, 2017: A census of atmospheric variability from seconds to decades. *Geophys. Res. Lett.*, 44:11201–11211.

Wilks, D.S., 2011: Statistical Methods in the Atmospheric Sciences, 3rd Edition, *Academic Press*

Wilks, D.S., 2009: Extending logistic regression to provide full-probability distribution MOS forecasts. *Meteorol. Appl.*, 16:361–368.

Wilks, D.S., 2006: On "field significance" and the false discovery rate. *J. Appl. Meteor. Climatol.*, 45:1181–1189.

Wilks, D.S. 2002. Smoothing forecast ensembles with fitted probability distributions. *Q. J. R. Meteorol. Soc.*, 128:2821–2136.

World Meteorological Organization, 2012: *Standardized Precipitation Index User Guide* (WMO-No. 1090). Geneva.

World Meteorological Organization, 2018: *Guidance on Verification of Operational Seasonal Climate Forecasts* (WMO-No. 1220). Geneva.

World Meteorological Organization, 2018: *Guidance on Good Practices for Climate Services User Engagement* (WMO-No. 1214). Geneva.

World Meteorological Organization, 2018: *Guidelines on Quality Management in Climate Services* (WMO-No. 1221). Geneva.

World Meteorological Organization, 2018: *Seasonal Climate Forecast – COURSE PACKAGE T.O.P.: Theory and Operational Principles, Project Report* (ETR-No. 22). Geneva.

World Meteorological Organization, 2019: *Manual on the Global Data-processing and Forecasting System* (WMO-No. 485). Geneva.

Yeh, S.-W. and Co-authors, 2018: ENSO atmospheric teleconnections and their response to greenhouse gas forcing. Reviews of Geophysics, 56:185–206.

Yoo, J.H. and I.-S. Kang, 2005): Theoretical examination of a multi model composite for seasonal prediction, *Geophys. Res. Lett.,* 32, doi:10.1029/2005GL023513

Yuan, X., and E. Wood, 2013: Multimodel seasonal forecasting of global drought onset. *Geophysical Research Letters*, 40, doi:10.1002/grl.50949.
Zebiak, S.E. and M.A. Cane, 1987: A model El Nino=Southern Oscillation. Mon. Wea. Rev., 115:2262–2278.
Zhang, C., 2005: Madden-Julian Oscillation. *Reviews of Geophysics*, 43:1–36.
Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar, 2012: A composite study of the MJO influence on the surface air temperature and precipitation over the continental United States. *Climate Dynamics*, 38:1459–1471.

For more information, please contact:

World Meteorological Organization

7 bis, avenue de la Paix – P.O. Box 2300 – CH 1211 Geneva 2 – Switzerland

Strategic Communications Office

Tel.: +41 (0) 22 730 87 40/83 14 – Fax: +41 (0) 22 730 80 27

Email: communications@wmo.int

public.wmo.int

JN 2079