

Predictability of Multiscale Waves and Tropical Rainfall and the Impact of Averaging

By
Ying Li

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
(MATHEMATICS)

at the
UNIVERSITY OF WISCONSIN – MADISON
2020

Date of final oral examination: July 24, 2020

The dissertation is approved by the following members of the Final Oral Committee:

Samuel N. Stechmann, Professor, Mathematics

Nan Chen, Assistant Professor, Mathematics

Elizabeth Maroon, Assistant Professor, Atmospheric and Oceanic Sciences

Stephanie Henderson, Assistant Professor, Atmospheric and Oceanic Sciences

Abstract

For tropical rainfall, there are several potential sources of predictability, including synoptic-scale convectively coupled equatorial waves (CCEWs) and intraseasonal oscillations such as the Madden–Julian Oscillation (MJO). In prior work, predictability of these waves and rainfall has mainly been explored using forecast model data. Here the first goal is to estimate the intrinsic predictability using, instead, mainly observational data. To accomplish this, Tropical Rainfall Measuring Mission (TRMM) data is decomposed into different wave types using spectral/Fourier filtering. Predictability of MJO rainfall is estimated to be 22 to 31 days, depending on longitude, as measured by the lead time when pattern correlation skill drops below 0.5. Predictability of rainfall associated with convectively coupled equatorial Rossby waves, Kelvin waves, and a background spectrum or non-wave component are estimated to be 8 to 12 days, 2 to 3 days and 0 to 3 days, respectively. Combining all wave types, the overall predictability of tropical rainfall is estimated to be 3 to 6 days, over the Indian and Pacific Ocean regions, and on equatorial synoptic and planetary length scales. For comparison, outgoing longwave radiation (OLR) was more predictable than rainfall by 5 to 10 days over these regions. Wave-removal tests were also conducted to estimate the contribution of each wave type to the overall predictability of rainfall. In summary, no single wave type dominates the predictability of tropical rainfall; each of the types (MJO, CCEWs, and non-wave component) has an appreciable contribution, due to variance contribution, length of decorrelation time, or a combination of these factors.

How to denoise the data and achieve better forecast skills? Based on conventional wisdom, one would expect a more skillful forecast if predicting weather averaged over one week instead of the weather averaged over one day. Similarly, one would expect a more

skillful forecast when averaging over a large spatial area instead of a smaller spatial area. Meanwhile, in making weather and climate predictions, the goal is often not to predict the instantaneous, local value of temperature, wind speed, or rainfall; instead, the goal is often to predict these quantities after averaging in time and/or space—for example, over one day or one week. What is the impact of spatial and/or temporal averaging on forecasting skill? This question is then investigated using simple stochastic models that can be solved exactly analytically. While the models are idealized, their exact solutions allow clear results that are not affected by errors from numerical simulations or from random sampling. As a model of time series of oscillatory weather fluctuations, the complex Ornstein-Uhlenbeck process is used. To furthermore investigate spatial averaging, the stochastic heat equation is used as an idealized spatiotemporal model for moisture and rainfall. Space averaging and time averaging are shown to have distinctly different impacts on prediction skill. Spatial averaging leads to improved forecast skill, in line with some forms of basic intuition. Time averaging, on the other hand, is more subtle: it may either increase or decrease forecast skill. The subtle effects of time averaging are seen to arise from the relative definitions of the time averaging window and the lead time. These results should help in understanding and comparing forecasts with different temporal and spatial averaging windows.

With the prior knowledge of the theoretical results from the stochastic models, we move forward to systematically investigate the impact of averaging based on data from operational weather forecasts. Data is analyzed for precipitation and surface temperature, for lead times of 3 and 7 days, respectively, and for time- and space-averaging diameters of 1 to 7 days and 100 to 4500 km, respectively. We find that time averaging is actually less effective than spatial averaging at improving forecast skill. A theoretical explanation is provided for the minimal effectiveness of time averaging, based on formulas for the decorrelation time of synthetic weather time series from a stochastic model. In effect, while time averaging creates a time series that is visually smoother, it does not

necessarily cause a substantial increase in the predictability of the time series.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Samuel Stechmann for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research. Meantime, his positive life attitude and efficient, organized methodical working style motivate me throughout the five years. I could not have imagined having a better advisor and mentor for my PhD study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Nan Chen, Prof. Elizabeth Maroon, and Prof. Stephanie Henderson, for their patience and time to serve on my committee.

I enjoyed and went through my graduate education and life with so much help from a lot of people in the friendly working environment within the Department of Mathematics as well as outside the department at UW-Madison. I would like to thank the faculty, staff members and colleagues who helped me and accompanied me throughout the whole program. Among them, I specifically want to mention Prof. David Anderson, Christina Kendzierski, Qin Li and Betsy Stovall, our graduate coordinator Kathie Brohaugh, IT manager Sara Nagreen, my ESL teacher Gail Ibele, CHTC staff Christina Koch, my officemates Geoffrey Bentsen and Polly Yu, my colleagues Hongfei Chen, Yu Feng, Xiao Hou, Yingda Li, my roommates Yating Gong, Chushan Wu for their support and many enjoyable moments.

My sincere thanks also go to all my dear friends and lifetime important people for supporting me in the hard time and bringing me so much happiness and wonderful memories throughout the past fourteen years. Without their support, I can never enjoy

so much about the moments of life and complete the degree. For this, I would like to thank Yuting Bian, Chang He, Sheng Jin, Lim Tau Shean, Lim Tau Yao, Yushan Liu, Yichen Shen, Yuan Tian, Rui Wang, Tianjie Wang, Ziyue Wang, Wen Wu, Qinshu Xue, Yue Yin, Jiye Yu, Yin Zhan, Jingwen Zhu for their countless cares, advice, patience, company and all the fun time that we spent together. I would particularly express my thanks to Qiuyan Xu, who gives me a lot of opportunities to develop myself and helps me to shape my future careers. I indeed appreciate all her selfless guidance from her life experiences and all the enjoyable life moments her family shared with me, from the very beginning when I arrived in the United States for the first time up to now.

My loving family plays important roles not only in my five years of graduate study but also throughout my lifetime. Therefore, last but not the least, I would like to express my deepest appreciation and apology to my parents, my grandmother and all my dear family members for my absence of many family gatherings and life-changing moments. Without their unrequited love and supports, I can never go such a long way on my academic life over these years, come across so many nice lifetime friends, enjoy rich and colorful life experiences and complete my graduate study successfully.

Contents

Abstract	i
Acknowledgements	iv
1 Introduction	1
1.1 Madden-Julian Oscillation	1
1.2 Convectively Coupled Equatorial Waves	5
1.2.1 Equatorial Fluid Dynamics	5
1.2.2 Analysis in the Wavenumber-Frequency Domain	8
1.3 Stochastic Models and Averaging	14
2 Predictability of Tropical Rainfall and Waves: Estimates from Observational Data [40]	15
2.1 Introduction	15
2.2 Data and Methods	18
2.2.1 Data and Setup	18
2.2.2 Wave Decompositions with Fourier Filtering	20
2.2.3 Modeling Wave Signals as Damped Oscillators with Stochastic Forcing	23
2.2.4 Estimating Predictability	26
2.3 Predictability of Individual Wave Modes: MJO, CCEWs, and the Non-wave Component	28
2.3.1 MJO	28
2.3.2 CCEWs	32
2.3.3 Non-wave Component	33

2.4	Predictability of Tropical Rainfall	35
2.4.1	Wave-Exclusion Studies	35
2.4.2	Planetary-Scale Predictions	38
2.5	Comparing predictability of precipitation versus cloudiness (OLR)	40
2.6	Discussion	42
2.7	Conclusion	44
3	Spatial and Temporal Averaging Windows and Their Impact on Fore- casting: Exactly Solvable Examples [39]	48
3.1	Introduction	48
3.2	Models and Methods	50
3.2.1	Mathematical Models	50
3.2.2	Measures of Forecasting Skill: Mean Square Error (MSE) and Pearson Correlation Coefficients (ρ)	55
3.2.3	Temporal and Spatial Averaging: Definitions and Notations . . .	56
3.3	Forecasting the Complex Ornstein-Uhlenbeck Process	58
3.3.1	Forecasting at a Single Time Point	58
3.3.2	Forecasting with Temporal Averaging	62
3.4	Subtle Impacts of Definitions of Averaging Window and Lead Time . . .	65
3.5	Forecasting an Idealized Spatiotemporal Rainfall Model: Stochastic Heat Equation	70
3.5.1	Forecasting at a Single Time Point and a Single Spatial Location	71
3.5.2	Forecasting at a Single Spatial Location with Temporal Averaging	73
3.5.3	Forecasting at a Single Time Point with Spatial Averaging	77
3.5.4	Forecasting with Both Temporal and Spatial Averaging	79
3.6	Conclusions	80
3.7	Appendices	82

3.7.1	Temporal and Spatial Averaging: Model Statistics	82
3.7.2	Formulas of forecasting skill for a both temporal and spatial averaged SPDE	85
3.7.3	Calculations of some important statistics	88
4	Is weather forecast skill improved by time averaging?	91
4.1	Introduction	91
4.2	Methods	92
4.2.1	Data	92
4.2.2	Data Analysis Setup	94
4.2.3	Evaluating Forecast Skills	95
4.2.4	Computational Methods	97
4.3	Impact of averaging on forecast skill	98
4.4	Theoretical explanation	102
4.5	Conclusions	104
4.6	Appendix 1: Additional Sensitivity Tests	105
4.6.1	Analysis on anomaly data	106
4.6.2	Assessing tropical precipitation forecast skills using GPM data . .	107
4.7	Appendix 2: Analytic Formulas for Decorrelation Time	108
5	Conclusions	115
5.1	Research Questions and Main Findings of the Study	115
5.2	Future Work	116
	Bibliography	118

Chapter 1

Introduction

The main research work described in this thesis will be focusing on investigations in the important multiscale waves, such as Madden-Julian Oscillation and convectively coupled equatorial waves, and trying to discover interesting phenomena from data analysis for tropical rainfall, surface temperature, etc.

Madden-Julian Oscillation (MJO) and Convectively Coupled Equatorial Waves (CCEWs) play important roles in tropical climate modelling and weather predictions. In this chapter, we introduce the theory and background about these significant multiscale waves and their influence in the field of atmospheric science, especially in equatorial fluid dynamics.

A brief introduction about the applications of stochastic models and time averaging in the climate and weather modelling appearing in some recent research work will be covered as well in this chapter, followed by a deep dive in the predictability of multiscale waves and impacts of averaging in the next chapters.

1.1 Madden-Julian Oscillation

In the tropical atmosphere, the Madden-Julian Oscillation (MJO) is the largest element of the intraseasonal (30- to 90-day) variability, identified by and then named for Roland Madden and Paul Julian of the American National Center for Atmospheric Research

(NCAR) in 1971. MJO is also known as the 30- to 60-day oscillation, 30- to 60-day wave or intraseasonal oscillation. The MJO is best described as intraseasonal tropical climate variability which varies on a week-to-week basis so that there can be multiple MJO events within a season. As a large-scale coupling between atmospheric circulation and tropical deep atmospheric convection, MJO is revealed by anomalous rainfall most clearly through less intermittent Outgoing Long-wave Radiation (OLR) and it is also closely connected with a lot of other tropical atmospheric phenomena, such as the monsoon, tropical cyclogenesis, El Niño-Southern Oscillation (ENSO), etc. It can also have dramatic impacts in the mid-latitudes, contributing a lot to various extreme events in the United States several times a year, including Arctic air outbreaks during the winter months across the central and eastern areas of the United States [28, 65].

Unlike ENSO, which is stationary, the Madden-Julian Oscillation is a traveling pattern that propagates eastward, at approximately 4 to 8 m/s through the atmosphere above the warm parts of the Indian and Pacific oceans. Consequently, the Madden-Julian Oscillation is characterized by an eastward progression of large regions of both enhanced and suppressed tropical rainfall, observed mainly over the Indian and Pacific Ocean. The MJO is an eastward moving disturbance of rainfall, winds, clouds and pressure that traverses the planet in the tropics and returns to its initial starting point in about 30 to 60 days [46, 47].

Fig. 1.1 shows the activity of outgoing longwave radiation (OLR) throughout the whole year of 2010. The observational data are averaged over the tropical belt. Eastward propagation indicating MJO activities can be clearly observed through the motion of OLR as time changes in this figure.

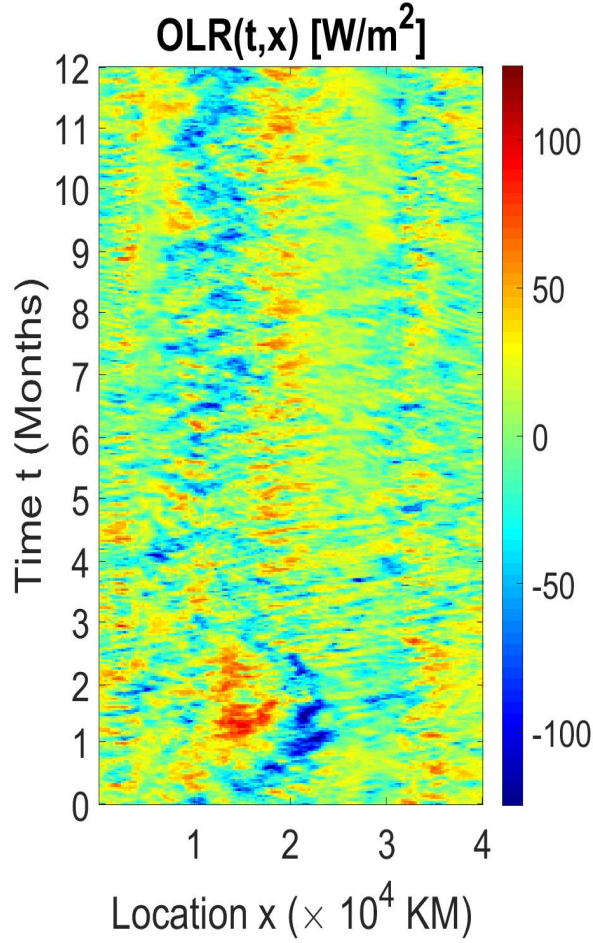


Figure 1.1: A Hovmöller diagram of outgoing longwave radiation (OLR) showing the MJO in the year 2010 after averaging over the tropical belt.

The MJO consists of two phases: the enhanced rainfall phase and the suppressed rainfall phase. The enhanced rainfall phase is also known as the enhanced convective phase [90]. In the enhanced convective phase, winds at the surface converge and air is pushed up throughout the atmosphere. At the top of the atmosphere, the winds reverse. This kind of rising air motion tends to increase condensation and rainfall. On the contrary, in the suppressed convective phase, winds converge at the top of the atmosphere, the air is forced to sink and then to diverge at the surface. The air warms and dries as it sinks from high altitudes, which suppresses rainfall. The planet is often

dissected by the strong MJO into two halves in these two different phases. These two phases produce opposite changes in clouds and rainfall and this entire dipole propagates eastward. More cloudiness, rainfall and storminess in the enhanced convective phase while more sunshine and dryness in the suppressed convective phase are caused by this eastward movement of the dipole structure in the tropics. However, the changes in the rainfall and winds not only impacts the tropics but also the extratropics. The direct influence can be tracked poleward farther than 30 degrees latitude from the equator in both northern and southern hemispheres [29, 71].

Irregular global weather and climate phenomenon will appear according to different statuses of the present MJO's movement. When MJO's movement slow or stall during the Northern Hemisphere summer and early autumn occasionally, it will result in consistently enhanced rainfall for one side of the global and consistently depressed rainfall for the other side. If the MJO goes quite for a period of time, it will lead to non-anomalous storm activity in different regions of the globe. Mechanistically, MJO affects the global climate and weather through its alteration of tropical heating patterns that sets off a Rossby wavetrain/teleconnection pattern.

There exist plenty of MJO-related local effects and downstream effects. For example, the MJO has significant effects on timing and strength of the Indian and West African summer monsoon during the Northern Hemisphere summer season as well as influence on tropical cyclones throughout the boreal warm season in the north Pacific and the north Atlantic basins [6, 37]. The interannual variability in MJO oscillation activities between long periods of strong activities and weak or absent oscillations is also partially connected to the ENSO cycle. Strong links between the MJO events and extreme west coast precipitation events in the North America including storms and floods are suggested by a lot of strong evidence [21], which makes MJO important for weather and climate predictions in an extended range and areas beyond the tropics.

In addition, rapid global warming in ocean temperature changes the residence time of MJO over the tropical oceans, resulting in the changes of rainfall patterns across the globe. From all aspects, MJO is an important component in weather and climate predictions due to its huge impact in different tropical and extratropical climatic events and issues [53].

1.2 Convectively Coupled Equatorial Waves

1.2.1 Equatorial Fluid Dynamics

To describe the deep convection in the fluid dynamics of the atmosphere and ocean, the shallow water system is widely used to model and control these complex processes. The shallow water theory describes motions of incompressible and homogeneous density fluid dynamics in a single thin layer on a rotating sphere that is independent from the vertical direction.

A general form of the shallow equations can be written as

$$\frac{\partial u}{\partial t} - fv = -\frac{\partial \phi}{\partial x} \quad (1.1)$$

$$\frac{\partial v}{\partial t} + fu = -\frac{\partial \phi}{\partial y} \quad (1.2)$$

$$\frac{\partial \phi}{\partial t} + gh_e \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0 \quad (1.3)$$

where u is the zonal velocity, v is the meridional velocity, ϕ is the geopotential, g is the acceleration due to gravity, h_e is the depth of the undisturbed layer of fluid and f is the Coriolis coefficient associated with the Coriolis force. Physically, the Coriolis force is an inertial or fictitious force acting on objects in motion within a frame of reference that rotates with respect to an inertial frame. The Coriolis effect describes the pattern of deflection taken by objects due to the Earth's rotation with different rotation

speed at different latitudes. Both the velocity of the Earth and the velocity of the fluid influence the impact of the Coriolis effect, contributing to a lot of large-scale weather patterns. Specifically, $f = 2\Omega \sin(\theta)$ on earth where Ω is the angular rotation rate ($\pi/12$ radians/hour) of the Earth and θ is the latitude.

In the tropics, an appropriate valid approximation for the Coriolis parameter f is a linear function of y as $f = \beta y$ where y is the distance from the equator. The shallow water equations turn into a tropical version as

$$\frac{\partial u}{\partial t} - \beta y v = -\frac{\partial \phi}{\partial x} \quad (1.4)$$

$$\frac{\partial v}{\partial t} + \beta y u = -\frac{\partial \phi}{\partial y} \quad (1.5)$$

$$\frac{\partial \phi}{\partial t} + gh_e \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0 \quad (1.6)$$

Zonally propagating wave solutions in the form of

$$\begin{pmatrix} u \\ v \\ \phi \end{pmatrix} = \begin{pmatrix} u_0(y) \\ v_0(y) \\ \phi_0(y) \end{pmatrix} \exp [i(kx - \omega t)], \quad (1.7)$$

where k is the zonal wavenumber and ω is the frequency, are sought through substitutions and rearrangements. Then the problem is reduced into a cleaned version of a second-order differential equation of v_0 only:

$$\frac{d^2 v_0}{dy^2} + \left(\frac{\omega^2}{gh_e} - k^2 - \frac{k}{\omega} \beta - \frac{\beta^2 y^2}{gh_e} \right) v_0 = 0 \quad (1.8)$$

With this quantum harmonic oscillator equation, the solutions to this equation are well known which decay away from the equator and the constant part of the coefficient in

parentheses must satisfy

$$\frac{\sqrt{gh_e}}{\beta} \left(\frac{\omega^2}{gh_e} - k^2 - \frac{k}{\omega} \beta \right) = 2n + 1, \quad n = 0, 1, 2, 3, \dots \quad (1.9)$$

The relationship between the frequency ω and the wavenumber k is therefore established through the equation (1.9). The horizontal dispersion relationships for a bunch of important and well-known equatorial waves of the atmosphere and ocean, including eastward inertio-gravity (EIG), westward inertio-gravity (WIG), equatorial Rossby (ER) waves are defined through this relationship. In addition, mixed Rossby-gravity (MRG) waves correspond to the $n = 0$ solution for (1.9) but require special consideration. $v_0 = 0$ for all y is an extra solution to the equations (1.4)-(1.5) but not covered by the relationship (1.9). This extra solution corresponds to the Kelvin wave. More detailed explanations about the equatorial fluid dynamics and convectively coupled equatorial waves can be found at the paper [33] and the book [49].

With the nondimensional frequency $\omega^* \equiv \omega/(\beta\sqrt{gh_e})^{1/2}$ and nondimensional zonal wavenumber $k^* \equiv k(\sqrt{gh_e}/\beta)^{1/2}$, the dispersion curves for these important equatorial waves are plotted in the Fig. 1.2. Kelvin wave solution is labeled as $n = -1$ in Fig. 1.2 for consistency with equation (1.9).

Equatorial waves decay rapidly away from the equator but they can propagate in the longitudinal and vertical directions. As a result of the combination of the rotation of the spherical shaped Earth and rapid increase in the magnitude of the Coriolis force away from the equator, the equatorial waves are trapped close to the equator. These trapped atmospheric and oceanic waves near the equator affect a lot on many climate and weather events such as ENSO. From the perspective of the analytical solutions, full horizontal structures of the wave solutions can be obtained by substitution of v_0 solutions from (1.8) into the original shallow water equations (1.4) - (1.6). Fig. 1.3 shows the full

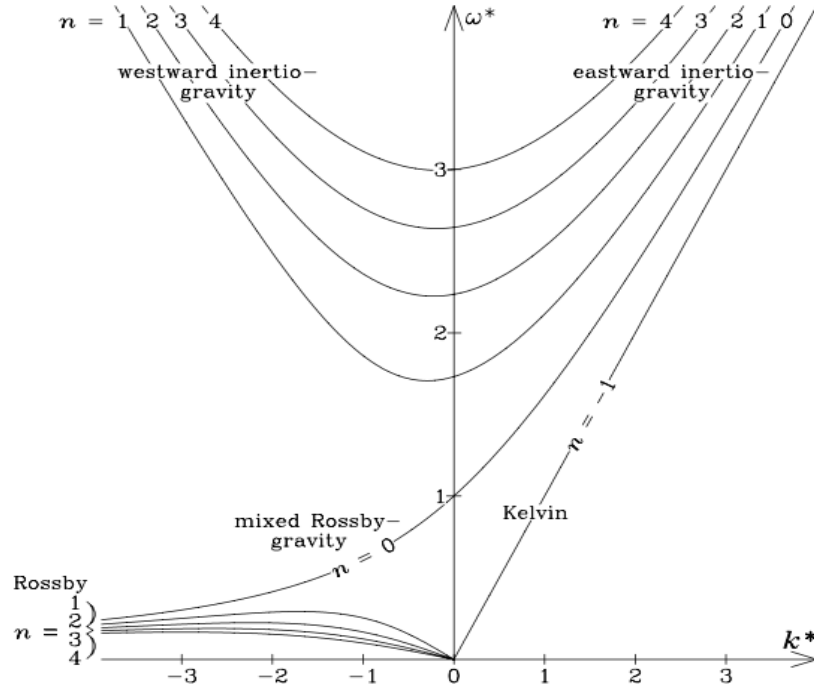


Figure 1.2: Dispersion curves for equatorial waves as a function of the nondimensional frequency ω^* and the nondimensional zonal wavenumber k^* , cited from [33] © Advancing Earth and space science. Used with permission.

structures for some of these equatorial waves. The interesting trapped phenomenon near the equator are clearly and vividly shown in the solutions depicted in Fig. 1.3, playing important roles in not only the tropical but also the global atmospheric physical processes and climate phenomena through their influence on the global circulation patterns.

1.2.2 Analysis in the Wavenumber-Frequency Domain

Satellite data serve as a crucial aid in exploring the location and strength of deep tropical convection because in situ observations over the tropics are very limited. There are some widely and commonly used datasets for research in MJO and CCEWs. The interpolated NOAA-produced outgoing longwave radiation (OLR) product is most utilized and serves as a very good statistical proxy for tropical convection with records starting

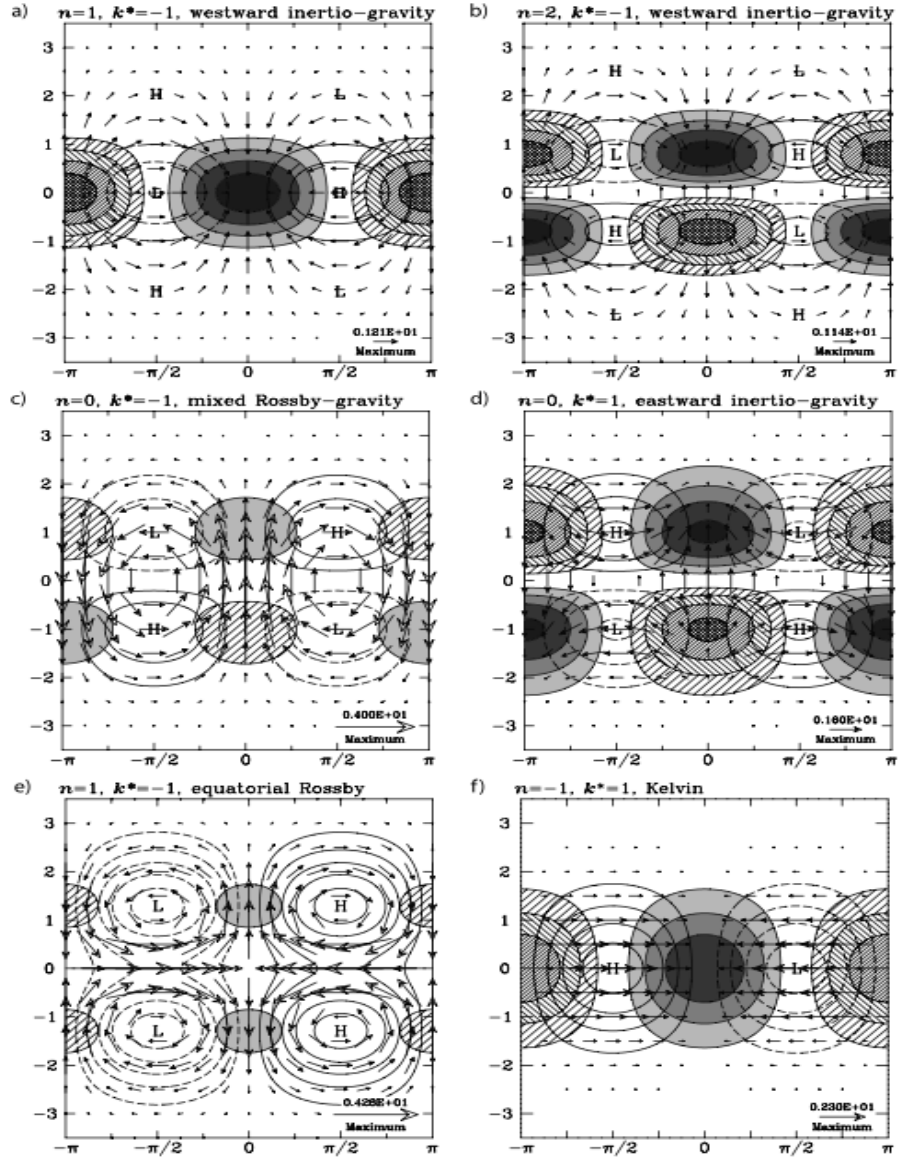


Figure 1.3: Horizontal structures of a subset of the zonally propagating wave solutions to the shallow water equations (equations (1.4)-(1.6)) on an equatorial β plane at nondimensional wavenumber $k^* = 1$ or -1 , cited from [33]
 © Advancing Earth and space science. Used with permission.

from 1974. Measured by infrared-sensing geostationary weather satellites, the degree of outgoing long wave radiation tracks the active phase of MJO. Stronger convection or thunderstorm complexes are expected to appear with lower amount of OLR. The Cloud

Archive User Services (CLAUS) dataset produces a global grid of brightness temperature (T_b) at 0.5° grids and 8 times a day within the tropics through geostationary and polar orbiting images. It is another useful dataset for representing deep convection with an advantage of the higher resolutions compared to the OLR dataset. Reanalysis data are also commonly used for research about the atmospheric dynamics. Reanalysis considers both observational data and first-guess model data to get an estimate about the atmospheric state. Reanalysis is good and efficient at capturing large-scale features of CCEWs but is limited in isolating more subtle details in the wave dynamics due to the use of radiosonde data.

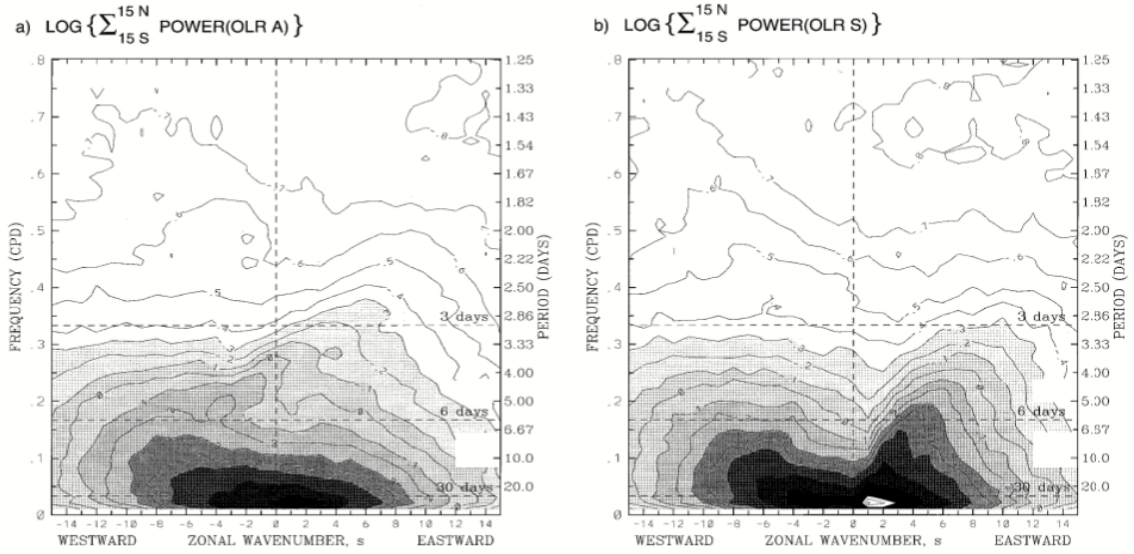


Figure 1.4: Zonal wavenumber-frequency power spectra of the (a) antisymmetric component and (b) symmetric component of OLR, calculated for the entire period of record from 1979 to 1996. For both components, the power has been summed over 15°S – 15°N lat, and the base-10 logarithm taken for plotting. Contour interval is 0.1 arbitrary units. Shading is incremented in steps of 0.2. Certain erroneous spectral peaks from artifacts of the satellite sampling are not plotted. Cited from [84] © American Meteorological Society. Used with permission.

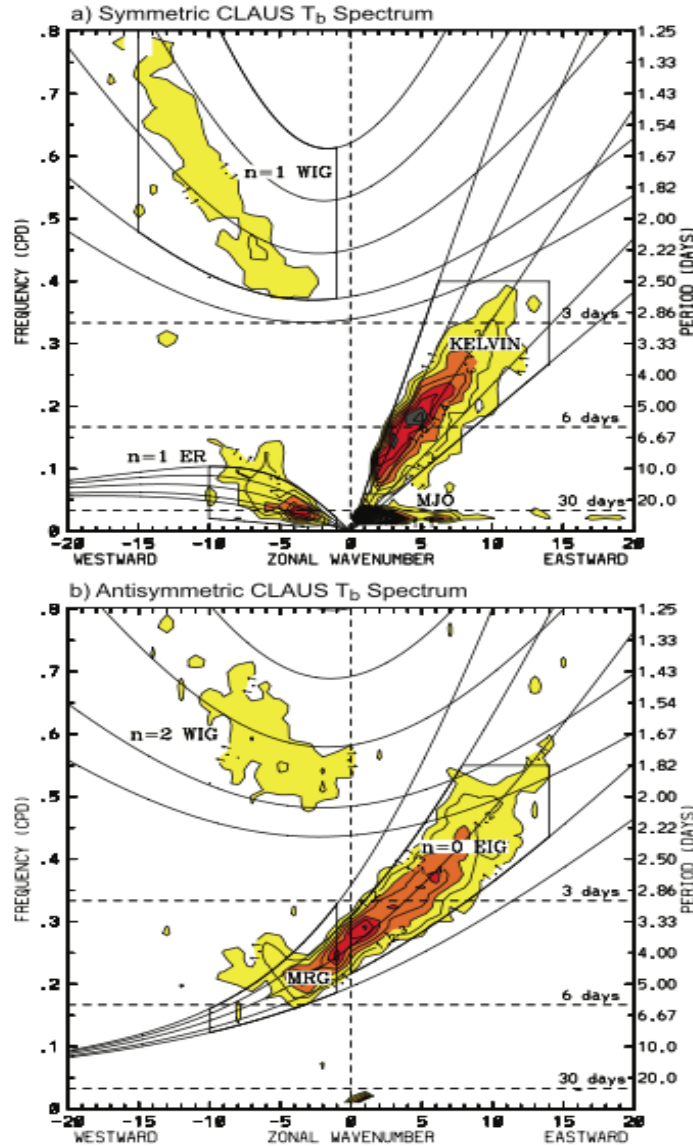


Figure 1.5: Wavenumber–frequency power spectrum of the (a) symmetric and (b) antisymmetric component of Cloud Archive User Services (CLAUS) T_b for July 1983 to June 2005, summed from 15°N to 15°S , plotted as the ratio between raw T_b power and the power in a smoothed red noise background spectrum (see WK99 [84] for details). Contour interval is 0.1, and contours and shading begin at 1.1, where the signal is significant at greater than the 95% level. Dispersion curves for the Kelvin, $n = 1$ equatorial Rossby (ER), $n = 1$ and $n = 2$ westward inertio-gravity (WIG), $n = 0$ eastward inertio-gravity (EIG), and mixed Rossby-gravity (MRG) waves are plotted for equivalent depths of 8, 12, 25, 50, and 90 m. Cited from [33] © Advancing Earth and space science. Used with permission.

As prominent spectral peaks are oriented along the dispersion curves of shallow water modes for a resting basic state from observational data, to isolate the MJO and different equatorial waves, the method of analysis and filtering in wavenumber-frequency domain is widely adopted for researching in these waves [84]. Power spectrum analysis of observational data of CLAUS and OLR along with rainfall provides strong evidence and support for the scientific reasonability and feasibility for analysis in wavenumber-frequency domain analysis. A lot of features in the power spectrum of the observational data are in line with the theory excellently. A spectrum obtained by dividing the raw power in satellite brightness temperature by an estimate of its red noise background for the CLAUS data is shown in Fig. 1.5 ([33]) and the contours of the logarithm of the power in the antisymmetric and symmetric components of OLR are shown in Fig. 1.4 ([84]). One can see the eastward propagating MJO is occurring mostly at eastward wavenumbers 1,2 and 3 and centered at a period of about 48 days in OLRs but not so detectable in OLRA for the raw OLR data, where $OLRS(\phi) = [OLR(\phi) + OLR(-\phi)]/2$ is the symmetric component and $OLRA(\phi) = [OLR(\phi) - OLR(-\phi)]/2$ is the antisymmetric component for OLR as a function of latitude ϕ . At the same time, one can see very clearly that spectral peaks are located very well along the dispersion curves of CCEWs derived from the shallow water equations for the CLAUS data.

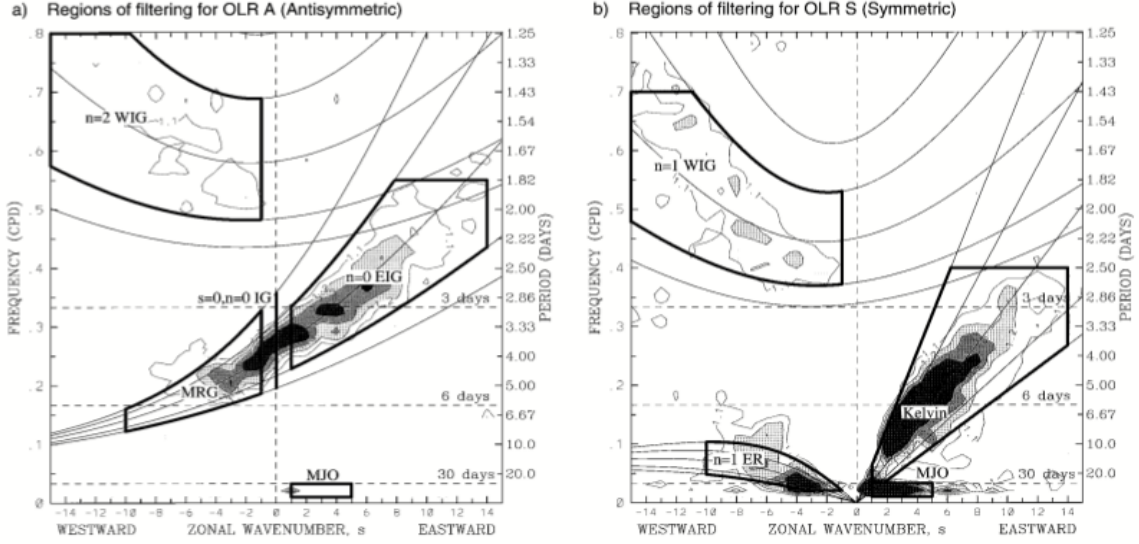


Figure 1.6: Thick boxes indicate the regions of the wavenumber-frequency domain used for filtering of the OLR dataset to retrieve the longitude-time information of the convectively coupled tropical waves for the (a) antisymmetric component and (b) symmetric component of the OLR. The thin lines are the various equatorial wave dispersion curves for the five different equivalent depths of $h = 5, 8, 12, 25, 50$, and 90 m. Cited from [84] © **American Meteorological Society. Used with permission.**

Based on these exploration and evidence, the wavenumber-frequency filtering can be applied to the observational dataset to extract the characteristics of MJO and convectively coupled waves in the time-longitude domain for each latitude. Fig. 1.6 displays the defined regions of filtering for MJO and each equatorial wave.

All these theories, datasets and methodologies serve as scientific cornerstones and provide crucial aid for a lot of tropical climate and weather research and as well as for the further analysis and research in the following chapters.

1.3 Stochastic Models and Averaging

Stochastic models have emerging applications in mathematical modeling in atmospheric science, such as tropical precipitation and water vapor dynamics , tropical waves, etc. Although deterministic models haven played important roles in modeling the fluid dynamics of atmosphere and ocean physically, stochastic models are a trending modeling method to add uncertainty in the models to have a better description of the uncertain quantities in atmospheric science. Through stochastic models, signals can be well modeled and a lot of important characteristics such as the the power spectrum can be captured in the wavenumber-frequency domain. At the same time, time and/or space averaging are widely used techniques in climate and weather predictions to aim at dealing with the limitation of data availability, denoising, predicting quantities over a region or time period. For example, in the research work of [92], time averaging is applied to predict precipitation over a range of time scales from a day to months with short to medium lead times for an operational forecast system, investigating forecast skills from short time predictions to intraseasonal time scale. One would expect improved forecast skills for most of the cases due to the denoising effects embedded in the averaging process based on the conventional wisdom. But systematical analysis in the impacts of time and spatial averaging are still ambiguous and lacking in this field both for theoretical analysis and observational data analysis. Motivated by this, a systematical investigation about this concern has been implemented and addressed in Chapter 3 and Chapter 4.

Chapter 2

Predictability of Tropical Rainfall and Waves: Estimates from Observational Data [40]

2.1 Introduction

Different types of “weather waves” exist in different areas of the globe. While baroclinic eddies generate much of the synoptic-scale weather fluctuations in midlatitudes, convectively coupled equatorial waves (CCEWs) are, in a sense, the “weather waves” of the tropics. CCEWs are a significant component of synoptic-scale variability—i.e., time scales of roughly 2–10 days and length scales of roughly 1,000–10,000 km (e.g., [33, 77, 78, 84]). The Madden–Julian Oscillation (MJO) is another significant contributor to tropical weather and climate, with a larger-scale signature on scales of roughly 30–90 days and 20,000 km (e.g., [36, 83, 90]).

In the present chapter, some of the main questions of interest are: What are the intrinsic limits of predictability of the “weather waves” of the tropics (CCEWs and the MJO)? Furthermore, to what extent do CCEWs and the MJO contribute to predictability of, more generally, tropical rainfall? The goal is to estimate both (i) the predictability of wave signals of each individual type and (ii) the predictability of the

overall precipitation signal, which is comprised of a mixture of the signals of different waves and a “background” signal [22, 84]. If the “background signal” is overwhelmingly strong, then the CCEWs may not contribute much predictability to the overall precipitation signal; on the other hand, it is also possible that the coherence of CCEWs may contribute to an enhancement in the amount of predictability of tropical precipitation, beyond the predictability of the “background signal” alone.

An investigation of these questions serves many purposes. For instance, CCEWs and the MJO have been difficult to simulate in global climate models (e.g., [24, 27, 42]); and for assessing model performance, forecast skill could be a useful metric if upper bounds on predictability are known (e.g., for the MJO, see [57, 83]). Also, if CCEWs are shown to offer a significant source of predictability, then it would provide motivation for the search for improved understanding and simulation of CCEWs.

Some recent work has provided some estimates of the predictability of CCEWs and tropical precipitation, and the approach of the present chapter will differ in many ways. [88] investigated predictability using convection-permitting simulations with the Weather Research and Forecasting (WRF) model. The 9-km horizontal grid spacing provided detailed simulations of multi-scale tropical weather systems, although it was computationally expensive and the investigation was limited to approximately one month. [15] and [25] investigated larger amounts of data from approximately one year and over five years, respectively. These latter two studies used data from several numerical weather predictions models such as the Global Forecast System (GFS) of the National Centers for Environmental Prediction (NCEP), the Integrated Forecast System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), the NCEP Climate Forecast System, version 2 (CFSv2), and the Navy Earth System Model (NESM). In the present chapter, in contrast, predictability will be estimated without a numerical weather prediction model; instead, estimates of predictability will be derived

from primarily observational data. As such, these different approaches provide complementary estimates that together can give a fuller picture of predictability. Some of the advantageous aspects of the present approach are that (i) the use of mainly observational data provides perhaps a more independent estimate, since it is not subject to the particulars of physics assumptions within models, and (ii) the computational expense of the model simulations is eliminated, which allows us to consider relatively long time series with relatively little computational expense.

The concept of predictability can be categorized into intrinsic predictability versus practical predictability. Intrinsic predictability represents the inherent limit of prediction given a nearly perfect forecast model of dynamical system and nearly perfect initial/boundary conditions; in such a setting, the predictability is then an indication of the chaotic nature of the dynamical system [14, 43, 76]. Practical predictability is the ability to make a prediction, given realistic uncertainties in both the forecast model and initial/boundary conditions [44, 45, 91]. The prior work of [15] and [25] was mainly related to practical predictability. In the present chapter, the approach is perhaps more closely aligned with estimating intrinsic predictability, since perfect initial conditions are used, and since the results do not rely on a numerical weather prediction model.

The remainder of the chapter is organized as follows. In section 2.2, the observational data is described, along with the methods for estimating predictability. In section 2.3, estimates of predictability are presented for the rainfall associated with individual wave types (MJO, CCEWs, and the non-wave component). In section 2.4, the different wave types are combined to provide estimates of predictability of the full rainfall signal. In section 2.5, the predictability of rainfall data is compared with the predictability of outgoing longwave radiation (OLR) data. Finally, section 2.7 includes a concluding discussion.

2.2 Data and Methods

2.2.1 Data and Setup

The Tropical Rainfall Measuring Mission (TRMM) data is used here for investigating the intrinsic predictability. The TRMM data mainly used in this chapter has a daily temporal resolution and 0.25° spatial resolution running from January 1st, 1998 to December 31st, 2017. Daily precipitation totals are derived from 3B42 Research Version. The dataset is downloaded from <https://pmm.nasa.gov/data-access/downloads/trmm>.

TRMM data is available at different temporal resolution including the 3 hourly product and the daily product. While the 3-hourly data comes with the advantage of higher temporal resolution, it also has some disadvantages. For instance, the 3-hourly data will include the diurnal cycle, which may a priori need some special treatment, and 3-hourly data is perhaps not necessary for investigating the wave types with the largest spectral peaks (MJO, Kelvin, and Rossby [84]). For this reason, the daily version of TRMM data is mainly used here unless otherwise specified. Some sensitivity tests using 3-hourly data and further discussions are included in section 2.6.

For testing the robustness and sensitivity of the main results, in addition to TRMM precipitation data, gridded daily interpolated OLR data from January 1979 to December 2013 from National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites are also analyzed in this study. OLR data has often been used in the past as a proxy for tropical precipitation and deep tropical convection, so it is interesting to compare the predictability of OLR with the predictability of TRMM data (OLR download link: https://www.esrl.noaa.gov/psd/data/gridded/data.interp_OLR.html). The OLR data initially from NCAR archives has gaps, and the gaps have been filled using temporal and spatial interpolation ([41]) to create the interpolated OLR data used here. The data for each day are archived on a resolution of 2.5° latitude $\times$

2.5° longitude globally. Note that the OLR and TRMM datasets have different native resolutions; therefore, in making comparisons between the two, a spatial filter is used to include only certain wavelengths that are resolved by both datasets; see section 2.5 for the specifications.

For the purpose of assessing model parameters and evaluating the prediction skills, the dataset is split into training data and testing data. The long time data before the year 2011 for TRMM (the year 2005 for OLR) is used as the training period for training parameters in the model and January 2011 to December 2015 (January 2005 to December 2009 for OLR) serves as the prediction period for testing. Data after the year 2015 is not used as testing data since a cosine tapering has been applied for these data for the purpose of Fourier transform. In the preprocessing, a smoothed seasonal cycle of the entire dataset is removed via the annual mean and the first three harmonics for all the data, so the remaining data represent anomalies from the seasonal cycle. An alternative definition of the seasonal cycle was also tested, where the hard cut-off at the third harmonic was replaced by a smoothed cutoff defined by a cosine tapering, and no significant differences in the results were seen.

CCEWs have a meridional structure with equatorial synoptic length scale of $\mathcal{O}(1000)$ km (e.g., [33, 60]). For this reason, instead of taking data at all the spatial and temporal points directly, we are only considering rainfall after averaging over the tropical belt, which provides a rainfall signal $r(x, t)$ for longitude x and time t . In the future it would be interesting to consider the rainfall signal at each latitude y instead of averaging over the tropical belt. The tropical belt average is performed with a Gaussian weight, which can be viewed as a projection onto a parabolic cylinder function. Specifically, the projection to parabolic cylinder mode 0 is used here, namely projecting the data onto the function

$$\phi_0(y) = \frac{1}{\pi^{1/4}} e^{-y^2/2}. \quad (2.1)$$

To define the projection, denote $r(x, y, t)$ as the rainfall data, where x is the longitude, y is the latitude, and t is the time. In (2.1), y is nondimensional, created by scaling with the reference scale 1500km. The discrete version of tropical belt average (the projection) then is

$$r(x, t) = \int_{-\infty}^{\infty} r(x, y, t) \phi_0(y) dy \approx \sum_{90^\circ S}^{90^\circ N} r(x, y_i, t) \phi_0(y_i) \Delta y \quad (2.2)$$

While it is also common to instead average with equal weight over a band of latitudes such as $10^\circ S$ to $10^\circ N$, the Gaussian weight is chosen here because it provides a smoother average, and it provides a connection with the parabolic cylinder functions, which provide a set of meridional basis functions for equatorial waves. Averaging meridionally using a Gaussian weight is also used in some previous papers (e.g., [60, 63, 74, 75]). Note that for some wave types, such as equatorial Rossby waves, an even tighter connection with equatorial wave theory would perhaps use additional parabolic cylinder functions, since the convergence patterns of equatorial Rossby waves include off-equatorial contributions. Also note that, by using a symmetric-in- y average in (2.1)–(2.2), the data is not expected to include the mixed Rossby–gravity (MRG) waves, since their signal appears in the antisymmetric data (i.e., the data from equator-to- $5^\circ N$ -averaged precipitation, minus equator-to- $5^\circ S$ -averaged precipitation; [84]). While the symmetric-in- y data from (2.1)–(2.2) will be the main focus here, the methods have also been repeated for antisymmetric-in- y data by replacing $\phi_0(y)$ in (2.1)–(2.2) with $\phi_1(y) = \pi^{-1/4} \sqrt{2} y \exp(-y^2/2)$. Using $\phi_1(y)$ instead of $\phi_0(y)$ allows the antisymmetric-in- y data to be identified, and it is used in the results below to investigate the predictability of MRG and $n = 0$ eastward inertio-gravity (EIG) waves.

2.2.2 Wave Decompositions with Fourier Filtering

For identifying signals due to different types of waves, the main method used in this study is the space-time spectral analysis (e.g., [84]). A brief overview is as follows.

The method begins with some preprocessing, described above, to remove the seasonal cycle, and to obtain an average over the latitudes near the equator, via a Gaussian weight. The result of the preprocessing is $r(x, t)$, from (2.2). As further preprocessing, a cosine tapering is applied near the beginning and end of the time series, so the values of the time series are zero at the beginning and end, thereby providing a signal that is periodic in time, in preparation for a Fourier Transform.

Next, a spatial Fourier Transform followed by a temporal Fourier Transform is applied on the longitude-time data $r(x, t)$ to convert the signal into the wavenumber–frequency domain, namely,

$$r(x, t) = \sum_k \sum_\omega \hat{r}_{k,\omega} e^{-i\omega t} e^{2\pi i k x / P_e}, \quad (2.3)$$

where k is the zonal wavenumber, ω is the frequency and P_e is the circumference of the Earth at the equator (approximately 40,000 km). The Fourier coefficients $\hat{r}_{k,\omega}$ are then used to identify the different wave types.

In order to isolate the signal for each of the different wave types (e.g., MJO, CCEWs, etc.), we follow the method of Fig. 6 of the research paper [84] (The figure has been shown as Fig. 1.6 in the introduction chapter). In particular, each wavenumber–frequency point (k, ω) is assigned to a different wave type α , where α is an index that indicates the wave type (MJO, Kelvin, Equatorial Rossby, and the non-wave component). The Fourier decomposition in Eqn. 2.3 can then be written as a sum over different wave types α , rather than a sum over different frequencies ω :

$$r(x, t) = \sum_k \sum_\alpha \hat{r}_{k,\alpha}(t) e^{2\pi i k x / P_e}, \quad (2.4)$$

with

$$\hat{r}_{k,\alpha}(t) = \sum_{\omega \in \Omega_{k,\alpha}} \hat{r}_{k,\omega} e^{-i\omega t}, \quad (2.5)$$

where α is an index that indicates the wave type (MJO, Kelvin, Equatorial Rossby, and the non-wave component) and $\Omega_{k,\alpha}$ is the set of frequencies for wave type α at wavenumber k .

For example, for the MJO, $\Omega_{k,\alpha}$ is set to be $\Omega_{k,\alpha} = \{\omega : 1/96 \leq \omega \leq 1/30\}$ for wavenumber $k = 1, 2, 3, 4, 5$ to extract MJO signals $\hat{r}_{k,\alpha}(t)$ via Eqn. 2.5 at wavenumber $k = 1, 2, 3, 4, 5$. In practice, the dataset is filtered to keep the part from 1/96 cpd to 1/30 cpd in the wavenumber-frequency domain for wavenumber $k = 1, 2, 3, 4, 5$ respectively and set all the other part of the data for wavenumber $k = 1, 2, 3, 4, 5$ all zeros. To convert back, an inverse Fourier transform is applied to the filtered data for each wavenumber k from 1 to 5. After that, the first 10% and the last 10% of these data at each wavenumber are cut (due to the cosine tapering applied at the beginning and end of the dataset to facilitate the Fourier transform). Then predictions are made using these filtered data at each wavenumbers independently. The total true MJO signal is considered to be the combination of these final filtered data for wavenumber $k = 1, 2, 3, 4, 5$ (i.e., $\sum_{k=1}^5 \hat{r}_{k,\alpha}(t) e^{2\pi i k x / P_e}$) and the prediction for the total MJO signal is generated by combining predictions at each wavenumber in the same way (i.e., $\sum_{k=1}^5 \hat{r}_{k,\alpha}^{pred}(t) e^{2\pi i k x / P_e}$, where $\hat{r}_{k,\alpha}^{pred}(t)$ is the prediction for $\hat{r}_{k,\alpha}(t)$). Analyses on other CCEWs (e.g., n=1 ER, Kelvin) are performed following the similar procedures by setting a different set to $\Omega_{k,\alpha}$ according to the filtering boxes from Fig. 6 of [84]. The “non-wave component” is defined here as the remaining part of the signal after the MJO, ER, Kelvin waves have been removed from the original spectral data. Westward inertio-gravity (WIG) waves are treated here as a part of the non-wave component, due to the limitations of using daily TRMM observed data, although some explorations of WIG waves with 3-hourly data are described in section 2.6.

2.2.3 Modeling Wave Signals as Damped Oscillators with Stochastic Forcing

As motivation for model choice, we recall that a goal here is to estimate predictability with less reliance on operational forecast models and more reliance on observational data. This goal is made possible here by the wave decomposition method described above, since the signal from one individual wave type can be modeled reasonably well by a simple damped oscillator model, which can be used as a simple forecast model. In this section, the simple model is described along with the method for fitting to observational data.

As a simple model for an individual wave type, a damped oscillator with stochastic forcing will be used. Specifically, the complex Ornstein-Uhlenbeck (cOU) process (see, e.g., [50]) is applied for modeling and predicting filtered signal $\hat{r}_{k,\alpha}(t)$ at each single wavenumber k and each wave type α [recall from (2.4) the details about defining $\hat{r}_{k,\alpha}(t)$]. The traditional Ornstein-Uhlenbeck process is real-valued and does not oscillate ([19]); on the other hand, the cOU process is complex-valued and is a damped oscillator with stochastic forcing. The cOU process is also exactly solvable and meaningful for predicting the complex Fourier coefficient $\hat{r}_{k,\alpha}(t)$ for a single wave. In what follows, we write $r(t)$ in place of $\hat{r}_{k,\alpha}(t)$ to simplify notation. The evolution of the cOU process is a complex linear stochastic differential equation:

$$dr(t) = (-\gamma + i\omega)r(t) + \sigma dW(t) \quad (2.6)$$

where $\gamma, \sigma > 0$ and ω are real numbers and

$$dW(t) \equiv \frac{dW_1(t) + idW_2(t)}{\sqrt{2}} \quad (2.7)$$

is a complex Gaussian white noise where each component satisfies

$$dW_j(t) \equiv \dot{W}_j(t)dt, \quad j = 1, 2, \quad (2.8)$$

that is, white noise $\dot{W}_j(t)$ is intuitively like a “derivative” of the Wiener process $W_j(t)$ and it satisfies the following properties:

$$\mathbb{E}[\dot{W}_j(t)] = 0, \quad (2.9)$$

$$\mathbb{E}[\dot{W}_j(t)\dot{W}_j(s)] = \delta(t - s), \quad (2.10)$$

$$\mathbb{E}[\dot{W}_i(t)\dot{W}_j(s)] = 0 \quad \text{for } i \neq j. \quad (2.11)$$

The exact solution of (2.6) is

$$r(t) = e^{(-\gamma+i\omega)t}r(0) + \sigma \int_0^t e^{(-\gamma+i\omega)(t-s)}dW(s) \quad (2.12)$$

Here, $1/\gamma$ represents the decorrelation time of the signal and $w\omega$ is the oscillation frequency with $2\pi/\omega$ to be the time of one oscillation period. σ is the standard deviation of the white noise. The whole signal $r(t)$ is a periodic decaying signal with random white noise. As $t \rightarrow \infty$, $r(t)$ will converge to a stationary Gaussian distribution with mean 0 and variance $\sigma^2/(2\gamma)$ (e.g., [48, 50]). The autocorrelation function is given by the analytical formula

$$R(\tau) = R(t, t + \tau) = e^{-(\gamma+i\omega)\tau} \quad (2.13)$$

in the stationary state as $t \rightarrow \infty$ and τ is the lag. To summarize, as a forecast model, (2.12) provides an ensemble of forecasts, with a forecast mean of $r(t) = e^{(-\gamma+i\omega)t}r(0)$.

One might wonder why white noise is used here, whereas the tropical rainfall spectrum is known to have a form similar to red noise (e.g., [22, 84]). The formulation

here is, in fact, consistent with a red-noise spectrum of tropical rainfall. The basic feature of spatiotemporal “red noise” is that the variance is decreasing as a function of temporal frequency or spatial wavenumber k . These basic features are actually built into the different values of the parameters γ and σ for different wavenumber k . For instance, for a larger wavenumber k , the fitted cOU process will have a corresponding smaller variance $\sigma^2/2\gamma$ with the fitted values of γ and σ for this particular wavenumber. Different cOU processes are fitted for different wavenumbers independently. With different choices of parameters for different wavenumbers, the basic feature of decreasing variance as a function of wavenumber k is retained.

The model parameters γ, ω are needed to make a forecast, and different values are used for each zonal wavenumber k and wave type α . Here the parameters are chosen by matching the observed autocorrelation of the training data and the analytic autocorrelation function. The model parameters are determined to capture the first maximum/minimum of the real and imaginary parts of the observed autocorrelation for positive lags, except for the non-wave component. For the non-wave component, since it has no propagation direction, its values of ω are nearly zero, so γ, ω are selected by matching the discrete summation of observed autocorrelation function and the integral of the analytic autocorrelation function, a method that provides better model performance. Also, this is consistent with treating the non-wave component like the background spectrum of tropical convection, for which a natural simple model is eddy diffusion ([22]) without any wave oscillations.

An example of the autocorrelation fitting is shown for the MJO in Figure. 2.1. The left panel shows the observed autocorrelation function and the analytical autocorrelation function from (2.13), where the parameters γ and ω were chosen to capture the first maximum/minimum of the real and imaginary parts of the observed autocorrelation function. As seen in the figure, the autocorrelation function of the fitted model has a

quite good fitting up to lags of one month, although there is nonnegligible model error for lags that are larger than one month. It is possible that a nonlinear oscillator model (e.g., [8]) would be able to fit the statistics even more accurately; however, the present chapter is aimed at modeling many different wave types and many different wavenumbers, which involves model parameters for each wave type and each wavenumber; therefore, a simple linear oscillator model is advantageous here for its minimal number of parameters, and it provides reasonable results, as shown in Figure. 2.1.

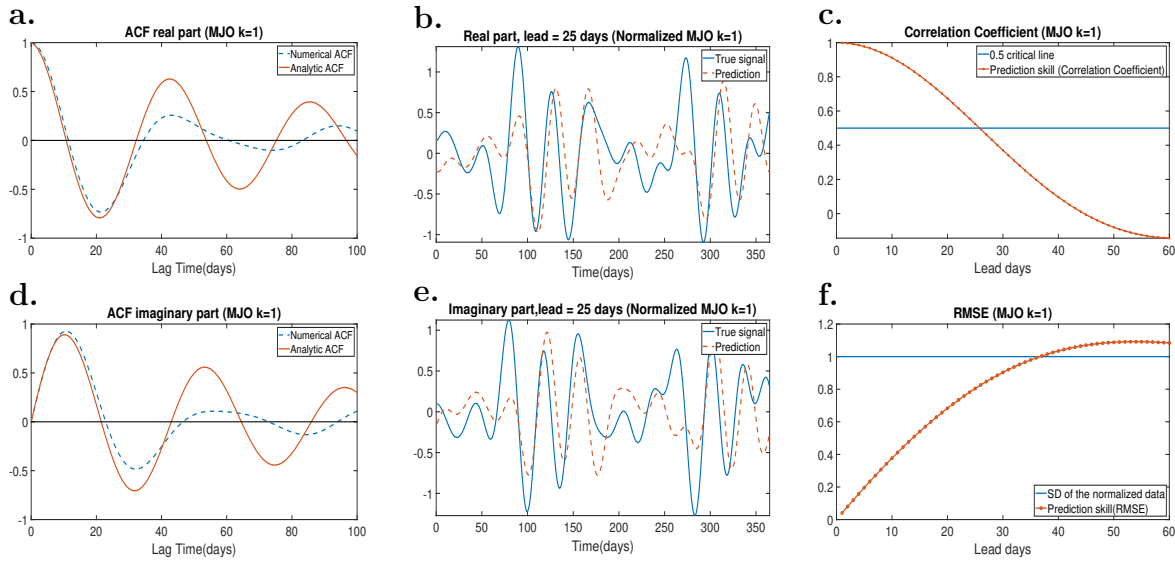


Figure 2.1: Panel **a** (Panel **d**): Real (Imaginary) part of the fitted auto-correlation function (red line) and the numerical auto-correlation function from the observational TRMM data (blue line) for wavenumber $k = 1$ of MJO; Panel **b** (Panel **e**): Real (Imaginary) part of the time series for MJO ($k=1$) observations and forecasts with lead time as 25 days in the year 2011; Panel **c** and **f**: Forecast skill (correlation coefficient and RMSE) for predictions for five-years signals of MJO($k=1$).

2.2.4 Estimating Predictability

To estimate predictability, forecasts are performed using the stochastic, damped oscillator models described in sections 2.2.2–2.2.3. The initial conditions $r(0)$ for the forecast are assumed to be perfect, in which case the formula for the mean prediction is given

by the expected value of (2.12), i.e.,

$$\mathbb{E}[r(t)] = \mathbb{E}[e^{(-\gamma+i\omega)t}r(0) + \sigma \int_0^t e^{(-\gamma+i\omega)(t-s)}dW(s)] = e^{(-\gamma+i\omega)t}r(0). \quad (2.14)$$

Note that such a forecast method would not be applicable to real-time forecasting due to the use of Fourier filtering in time, as described in section 2.2.2; some real-time wave decomposition methods have been proposed (e.g., [32, 58, 59, 60, 74, 75, 85, 86]), although it is not clear that real-time methods are as skillful at wave decompositions as non-real-time methods based on temporal Fourier filtering. In any case, the use of these perfect initial conditions is in line with the main goal here of estimating bounds on intrinsic predictability.

The prediction skill is evaluated by two commonly used criteria, Correlation Coefficient (ρ) and Root Mean Square Error (RMSE). Mathematically, with the true data $\mathbf{X} = (X_1, X_2, \dots, X_N)$ at N points in time and the corresponding predictions $\mathbf{X}^{pred} = (X_1^{pred}, X_2^{pred}, \dots, X_N^{pred})$, the correlation coefficient is calculated by

$$\rho(\mathbf{X}, \mathbf{X}^{pred}) = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_i^{pred} - \bar{X}^{pred})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (X_i^{pred} - \bar{X}^{pred})^2}} \quad (2.15)$$

where $\bar{X}$, $\bar{X}^{pred}$ are the averages of X_i, X_i^{pred} ($i = 1, 2, \dots, N$) respectively and the RMSE is

$$RMSE(\mathbf{X}, \mathbf{X}^{pred}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - X_i^{pred})^2} \quad (2.16)$$

The overall forecasting skill is defined as the lead day where $\rho(\mathbf{X}, \mathbf{X}^{pred}) \geq 0.5$ and $RMSE(\mathbf{X}, \mathbf{X}^{pred}) \leq SD(\mathbf{X})$ (standard deviation of the data of the true signal). In other words, the criterion of a reasonable forecast is a threshold of 0.5 for correlation coefficient and 1 standard deviation for the RMSE. These threshold choices are also

commonly used in other forecast studies (e.g., [7]).

For the purpose of evaluating cOU predictions and wave exclusion tests, two baselines for comparison are used: the zero prediction and persistence prediction. The zero prediction is obtained by predicting that the signal will be identically zero for all future data points. (Note that the time series is centered to mean zero, so this can also be viewed as a climatological prediction, where the predicted value is the climatological mean.) The persistence prediction is obtained by predicting that the future weather condition will be the same as the present condition.

2.3 Predictability of Individual Wave Modes: MJO, CCEWs, and the Non-wave Component

In this section, we investigate the question: What is the intrinsic predictability of MJO-related rainfall, CCEW-related rainfall, and background-spectrum rainfall? Each of the wave types will be considered in isolation to identify the predictability of each individual wave mode.

2.3.1 MJO

MJO predictability is shown for each wavenumber k in Fig. 2.2 and Table 2.1. The predictability is defined as the lead time when the correlation coefficient drops to 0.5. In brief, all MJO wavenumbers have predictability of approximately 25–32 days.

As one example forecast for illustration, time series for $k = 1$ are shown in Fig. 2.1b,e. The figure shows a comparison of the true signal and predicted signal at lead time of 25 days. This lead time was chosen for illustration because it is approximately the predictability limit of the $k = 1$ MJO signal (see Fig. 2.2 and Table 2.1). As seen in Fig. 2.1b,e, the predicted signal catches the overall variability of the oscillations quite

well although it fails to catch the more extreme values with the present simple forecasting framework. In the other panels, in Fig. 2.1c,f, the RMSE and correlation coefficient are shown. The correlation coefficient is seen to decrease as lead time increases, and it decreases to 0.5 at a lead time of 25 days, which is used as the value reported in Fig. 2.2 and Table 2.1.

To move beyond forecasts of individual wavenumbers, we can combine the wavenumbers $k = 1$ to 5 of the MJO signal, using (2.4). In brief, only the MJO signal (from wave $k = 1$ to 5) is kept, and the signals of all other wave types are set to zero. Algorithmically, the space-time data $r_{MJO}(x, t)$ is then obtained using a temporal inverse Fourier transform followed by a spatial inverse Fourier transform. From the space-time data $r_{MJO}(x, t)$, one can observe a forecast of the MJO at each location around the equator. Such a forecast skill at each longitude is shown in Fig. 2.3. In terms of RMSE (Fig. 2.3 Panel **a**), the forecast skill is greatest over the Indian Ocean and western Pacific warm pool (longitudes from roughly 60E to 180), although this is also the region of greatest standard deviation in the MJO signal. In terms of correlation coefficient (Fig. 2.3 Panel **b**), the forecast skill is more nearly equal at each longitude. As a summary of forecast skill, in the Panel **c** in Fig. 2.3, it can be seen that the forecast skill is approximately 25–30 days at each longitude, similar to the forecast skill for individual wavenumbers shown in Fig. 2.2 and Table 2.1. Hence, whether viewed longitude by longitude or wavenumber by wavenumber, the predictability of the MJO is estimated to be 25–30 days.

Wave type	Wavenumber	Forecast skill	γ	ω	$\hat{\omega}$
MJO	k=1	25	0.33	-4.39	-0.023
MJO	k=2	27	0.37	-4.28	-0.023
MJO	k=3	32	0.36	-4.39	-0.023
MJO	k=4	26	0.35	-4.39	-0.023
MJO	k=5	28	0.41	-4.71	-0.025
ER	k=-2	33	0.17	-6.09	-0.032

ER	k=-3	16	0.52	-7.85	-0.042
ER	k=-4	12	1.03	-9.42	-0.050
ER	k=-5	11	1.08	-11.78	-0.063
ER	k=-6	9	1.18	-12.62	-0.067
Kelvin	k=1	17	0.38	-9.00	-0.048
Kelvin	k=2	6	1.79	-15.71	-0.083
Kelvin	k=3	4	2.85	-23.56	-0.125
Kelvin	k=4	4	3.76	-27.49	-0.146
Kelvin	k=5	2	4.81	-39.27	-0.208
Non-wave Component	k=1	2	1.72	0.08	0.001
Non-wave Component	k=2	8	0.76	-0.20	-0.001
Non-wave Component	k=3	5	1.94	-0.26	-0.001
Non-wave Component	k=4	2	3.52	-0.10	-0.001
Non-wave Component	k=5	1	10.04	-0.29	-0.002
MRG	k=-1	5	1.07	-47.12	-0.250
MRG	k=-2	5	1.29	-47.12	-0.250
MRG	k=-3	4	1.93	-47.12	-0.250
MRG	k=-4	3	2.72	-47.12	-0.250
MRG	k=-5	5	2.21	-39.27	-0.208
EIG	k=1	5	1.30	-47.12	0.250
EIG	k=2	3	2.75	-47.12	-0.250
EIG	k=3	2	4.90	-47.12	-0.250
EIG	k=4	2	13.65	-70.69	-0.375
EIG	k=5	3	12.55	-70.69	-0.375

Table 2.1: Summary of forecast skills for the wave types: MJO, ER, Kelvin, Non-wave Component, MRG and $n = 0$ EIG. Note: Forecast skills have the unit days, γ, ω have the unit $2\pi/\text{Month}$ and $\hat{\omega} = \omega/(30 \cdot 2\pi)$ is the frequency (CPD) with the unit $1/\text{day}$.

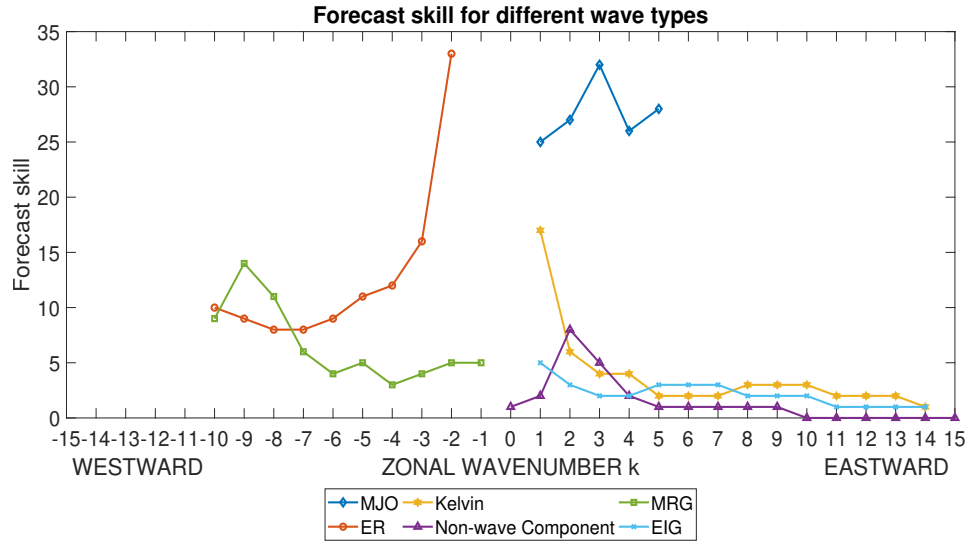


Figure 2.2: Forecast skills for different wavenumbers of MJO, ER wave, Kelvin wave, Non-wave Component, MRG wave and $n = 0$ EIG wave.

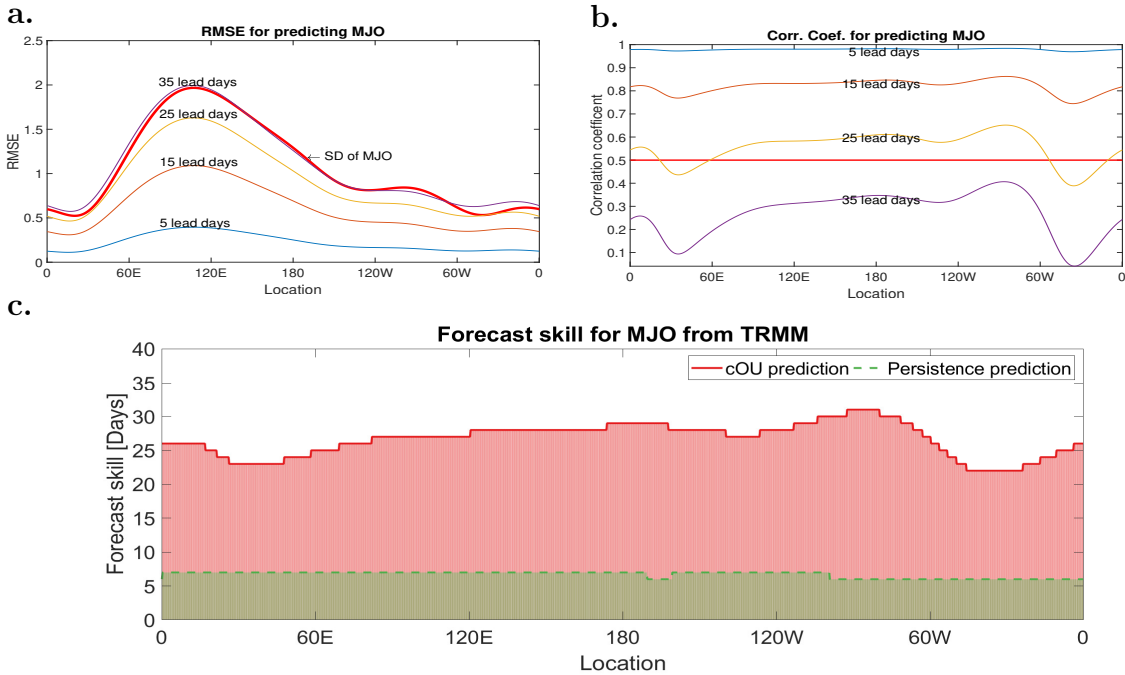


Figure 2.3: Forecast skills (unit: days) for MJO at different locations. Top two panels: RMSE (panel a) and correlation coefficient (panel b) for predicting MJO from TRMM data with different lead days; Bottom panel (panel c): Overall forecast skill for predicting MJO components from TRMM data with different methods cOU (solid pink) and persistence prediction (dashed green).

2.3.2 CCEWs

Predictability is estimated for two CCEWs: the ER and Kelvin waves. The results will be analyzed from two perspectives: wavenumber by wavenumber, and longitude by longitude.

First, predictability is reported for each individual wavenumber, analyzing each wavenumber separately from each other wavenumber, in Table 2.1 and Fig. 2.2. The predictability is seen to depend strongly on wavenumbers. For instance, the ER wave has a predictability of 33 days for wavenumber 1 but a predictability of roughly 8–10 days for wavenumbers 5 to 10; and the Kelvin wave has a predictability of 17 days for wavenumber 1 but a predictability of roughly 2–3 days for wavenumbers 5 to 10.

Overall, one can see a rough general trend in Table 2.1 and Fig. 2.2: predictability tends to increase as wave oscillation period increases. This is consistent with the intuition that waves with longer oscillation periods also tend to have longer decorrelation time γ^{-1} , and longer decorrelation times are associated with longer predictability times. A figure of $\hat{\omega}$ for different wave types versus different wavenumbers along with some additional figures are also provided in the supporting information.

Second, to analyze the predictability at different longitudes, the data from different wavenumbers are combined together to predict, e.g., the ER signal as a function of longitude. The spatial variations of predictability for the ER and Kelvin waves are shown in the first two columns in Fig. 2.4. Both ER and Kelvin have their largest variance over the Indian Ocean and western Pacific warm pool, from about 60°E to 150°W. The predictability of the ER wave varies from 8 days to 12 days over all the locations, while the forecast skill of the Kelvin wave varies from 2 days to 3 days, with little variation from location to location.

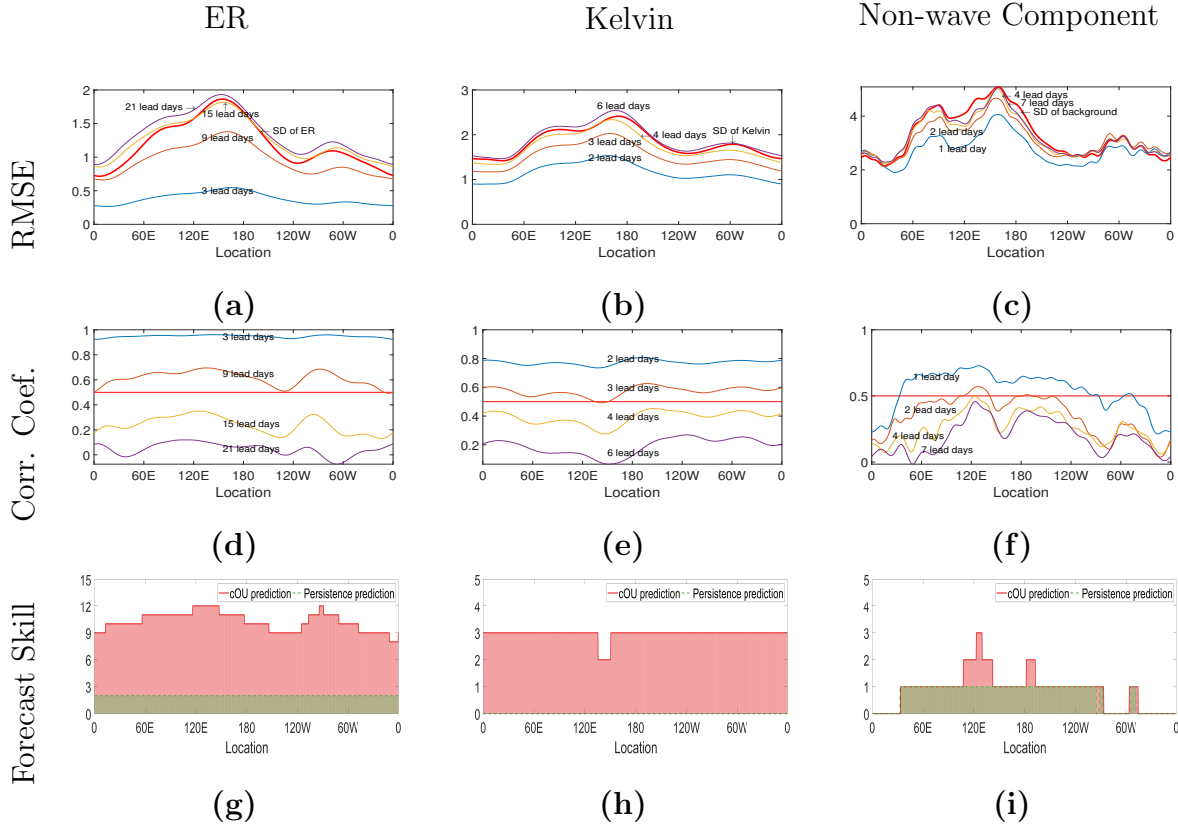


Figure 2.4: Forecast skills for ER wave, Kelvin wave and non-wave component. From the left to right are ER wave (panel **a,d,g**), Kelvin wave (panel **b,e,h**) and non-wave component (panel **c,f,i**), from the top to bottom are RMSE (panel **a,b,c**), correlation coefficient (panel **d,e,f**) and overall forecast skill (panel **g,h,i**).

2.3.3 Non-wave Component

Lastly, consider the non-wave component. The wavenumber-by-wavenumber results are shown in Table 2.1 and Fig. 2.2. Note that the non-wave component is not assigned a particular propagation direction (e.g., eastward or westward), so the $+k$ and $-k$ wavenumbers are analyzed together as a single unit, as wavenumber k . The predictability of the non-wave component is typically low, in the range of 1 to 2 days. An anomaly is seen in wavenumbers 2 and 3, for which the predictability times are 8 and 5 days, respectively; these longer predictability times are likely the result of, e.g., the

MJO signal being partially identified as “non-wave component,” since the MJO signal could potentially influence some frequencies that lie outside the filtering box of Fig. 6 of [84]. Overall, though, when viewed longitude by longitude (see Fig. 2.4 Panel (i)), the non-wave component has low predictability of roughly 1 day.

We note here the possibility of localized regions of enhanced predictability, as illustrated by the case of the non-wave component near 120E longitude; see Fig. 2.4 Panel (i). To ensure that this behavior is not a result of an error in the data analysis, one can trace its source to the plot of correlation coefficients, from which the forecast skill is calculated; see Fig. 2.4 Panel (f). One can see that the curves for two different lead times (e.g., lead times of 2 and 3 days) can sometimes be nearly overlapping when their correlation coefficients are nearly equal; e.g., see 120E longitude. If this occurs for a correlation coefficient of 0.5 (i.e., the cutoff correlation coefficient for defining the “forecast skill”), then the forecast skill can have a sharp change for nearby longitudes, as seen here for the non-wave component near 120E longitude. Such behavior also appears if OLR data is analyzed instead of TRMM data, as shown below in section 2.5.

We speculate that it may be related to the unique geographical features of the Indo-Pacific maritime continent, such as its associated topography and/or land-sea contrast.

This behavior could be eliminated by choosing a different cutoff, such as 0.6 instead of 0.5, since we see this behavior here for the lower correlation coefficient values of 0.55 or lower; but we will retain the cutoff of 0.5 since it is a commonly used definition of forecast skill.

2.4 Predictability of Tropical Rainfall

The previous section assessed predictions of rainfall associated with an individual wave type (MJO, ER, Kelvin, or non-wave component). In this section, in contrast, predictions of the full rainfall signal are analyzed. As a first brief look, see Fig. 2.5. The solid curve indicates the predictability of the full rainfall signal, and it is repeated identically in each panel of the figure. The forecast skill is roughly 3 to 6 days over the Indian and Pacific Ocean regions. This skill represents a substantial improvement over what was seen for the non-wave component alone (see Fig. 2.4), which was predictable for only roughly 1 to 2 days. The improved skill can be attributed to the additional wave types beyond the non-wave component: the CCEWs and the MJO.

In what follows, to provide a more detailed view of the full rainfall signal, wave-exclusion studies are also used in order to assess the contribution of each wave type to overall predictability (Sec. 2.4.1). Also, to assess predictability on different length scales, the planetary length scales (zonal wavenumbers -5 to +5) are investigated in Sec. 2.4.2.

2.4.1 Wave-Exclusion Studies

In this section, we ask: How important is each individual wave type for the predictability of the full rainfall signal? To investigate this question, we exclude the predictions of one wave type in predicting TRMM rainfall data, and evaluate the resulting decrease in predictability. To exclude a wave type, two methods are examined: either (i) setting the prediction of the wave's signal to be zero, or (ii) using a persistence prediction for that wave type. While we are changing the predictions for the component of the one wave type of interest, all the other components of the signal remain being predicted by cOU processes, as in our standard methodology.

The main results of the wave-exclusion studies are shown in Fig. 2.5. The forecast skill is presented as a function of longitude. With all wave types predicted and none excluded, the precipitation is most predictable over the Indian Ocean to Pacific Ocean regions, where the predictability is roughly 3 to 6 days, aside from the longer predictability of 9 days near 120E. When one wave type is excluded, a substantial loss of predictability is typically seen. In particular, if either the MJO, ER, or non-wave component is excluded, then a loss of predictability of several days can be seen over the Indian Ocean and Pacific Ocean regions.

For the Kelvin wave, on the other hand, the results are somewhat mixed. Little predictability is lost if the Kelvin wave is excluded by predicting it to be zero. Over many parts of the Indian and Pacific Oceans, no predictability is lost, and over other parts of the tropics, the loss is 1 to 2 days of predictability. Given that the overall predictability is only 1 to 2 days for many regions outside the Indian and Pacific Oceans, one could possibly view this as a substantial loss. Also, if the Kelvin wave is instead excluded by using a persistence forecast, then a substantial loss in predictability is seen: roughly 1 to 4 days.

Two factors are perhaps sufficient to explain each wave type's importance: decorrelation time and variance. The decorrelation time (γ^{-1}) has a rough correspondence with the predictability, as seen in Table 2.1 (see also Figs. 2.2, 2.3, and 2.4). For instance, the MJO and ER waves have the longest decorrelation times and predictability, whereas the non-wave component and Kelvin waves have shorter decorrelation times and predictability. However, decorrelation time alone is not enough to explain the contribution of each wave type to the predictability of overall rainfall. For instance, the MJO and non-wave component have somewhat similar contributions based on the wave exclusion studies (Fig. 2.5), yet the non-wave component has a very short decorrelation time. Hence, a second factor is needed to explain why the non-wave component has an

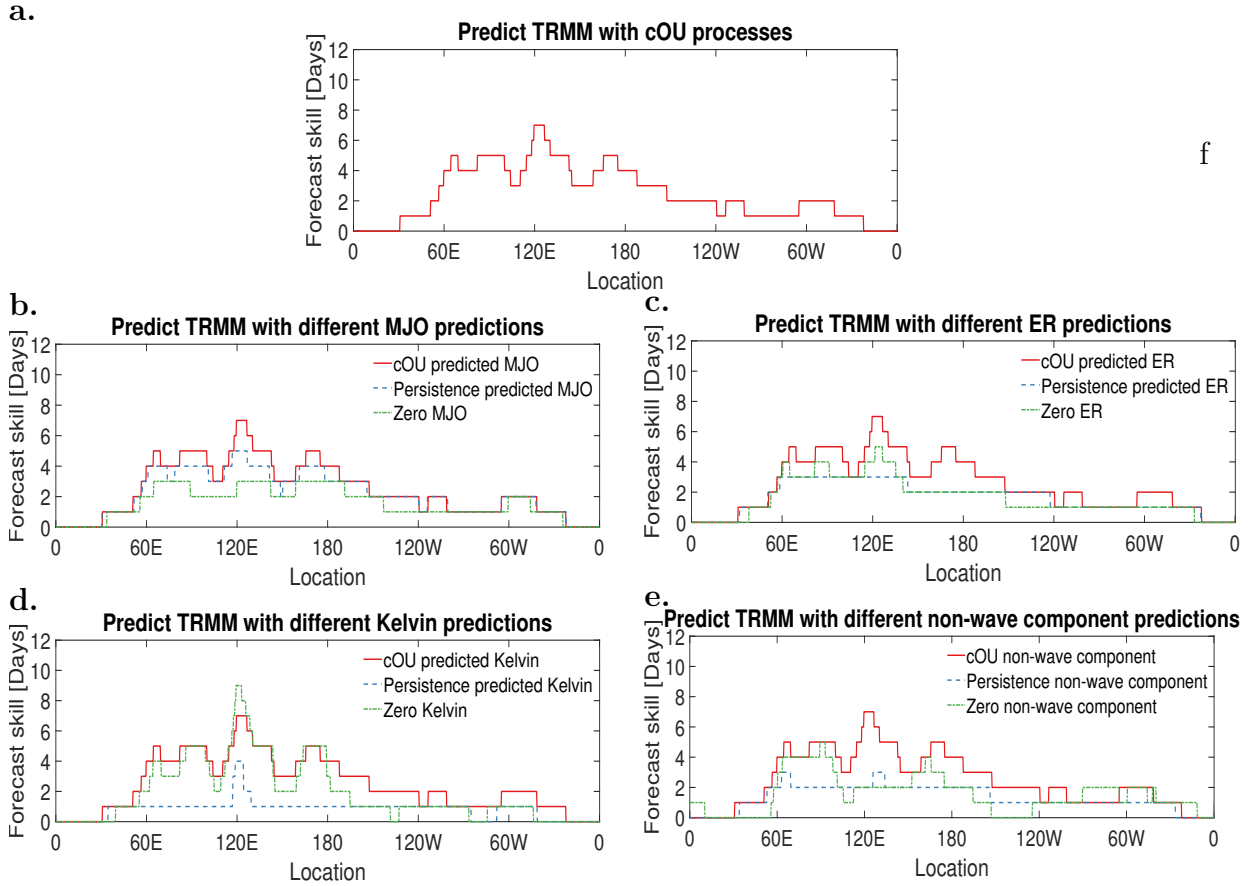


Figure 2.5: Panel **a**: Forecast skill for predicting TRMM data (wavenumber from -15 to 15) with cOU processes for all wave and non-wave components. Panel **b-e**: Forecast skills for predicting five-years TRMM data (wavenumber from -15 to 15) with different predicted MJO (Panel **b**), ER wave (Panel **c**), Kelvin wave (Panel **d**) and non-wave component (Panel **e**). Three forecast methods are used wave-exclusion studies: cOU prediction (pink), persistence prediction (blue) and prediction with all zeros (green).

important contribution: variance. The variance (or, rather, its square root, the standard deviation) is shown for each wave type in Fig. 2.6. The non-wave component has the largest standard deviation of all wave types, so substantial predictability will be lost if a poor forecast is used for such a large share of the total standard deviation. In this way, each of the wave types has an appreciable contribution to the overall predictability of rainfall, due to a long decorrelation time or a large variance or a combination of these

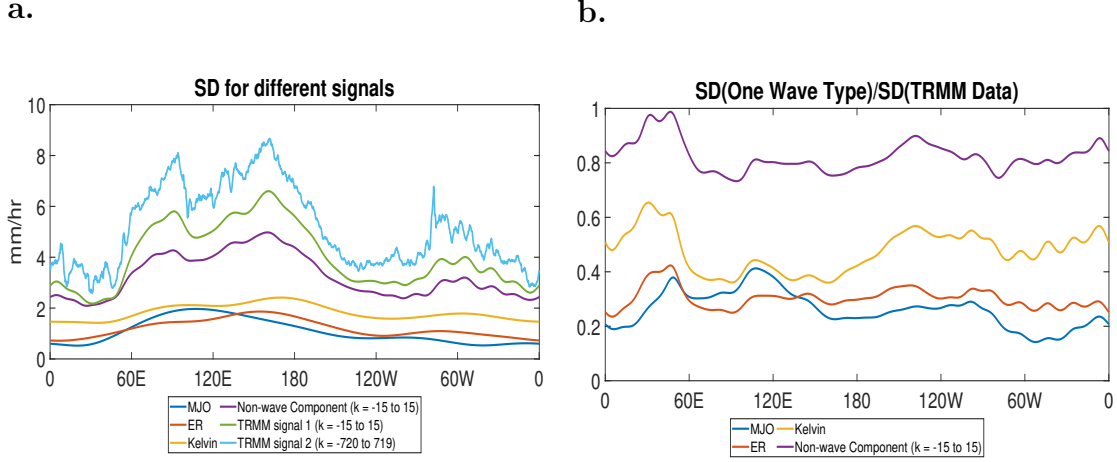


Figure 2.6: Panel **a**. Standard Deviation (SD) for different signals: MJO, ER, Kelvin, Non-wave component ($k=-15$ to 15), true signal from $k=-15$ to 15 of TRMM data and the original TRMM data (all wavenumbers, from $k=-720$ to $k=719$); Panel **b**. Standard deviation of the each wave type signal (MJO, ER, Kelvin, Non-wave component) over the standard deviation of the true signal from $k=-15$ to 15 of TRMM data.

factors.

2.4.2 Planetary-Scale Predictions

In the results above, the equatorial synoptic length scales (zonal wavenumbers -15 to $+15$) were considered, in order to include effects of many CCEWs. Instead, if we are only interested in the large-scale variations of rainfall, is it possible to achieve better prediction skill?

This question is now investigated by considering only zonal wavenumbers $k = -5$ to 5 , in order to represent the planetary-scale zonal variations of rainfall. One might expect enhanced predictability if only the largest scales are considered, consistent with the general idea that spatial averaging will improve forecast skill (e.g., [39]). Physically, we would expect slower phase speeds for some important CCEWs for smaller wavenumber k . For example, it will be true for the Kelvin wave since it has lower frequency for smaller

k defined in the wavenumber-frequency domain. As a result, enhanced predictability will be expected if only signals with small wavenumbers are considered. Exactly how much improvement is seen? The predictability is shown in Fig. 2.7. The forecast skill is quite long: roughly 10 to 20 days over the Indian and Pacific Ocean regions. Comparing Fig. 2.7 to the forecast skills for wavenumbers from -15 to 15 in Fig. 2.5a, the forecast skill for the planetary scale predictions has very significant improvements from 60°E to 150°W. The planetary scale predictions are overall about 5-15 days more predictable in these areas, particularly near 120°E and 180° with about 10-15 days improvements in forecast skills.

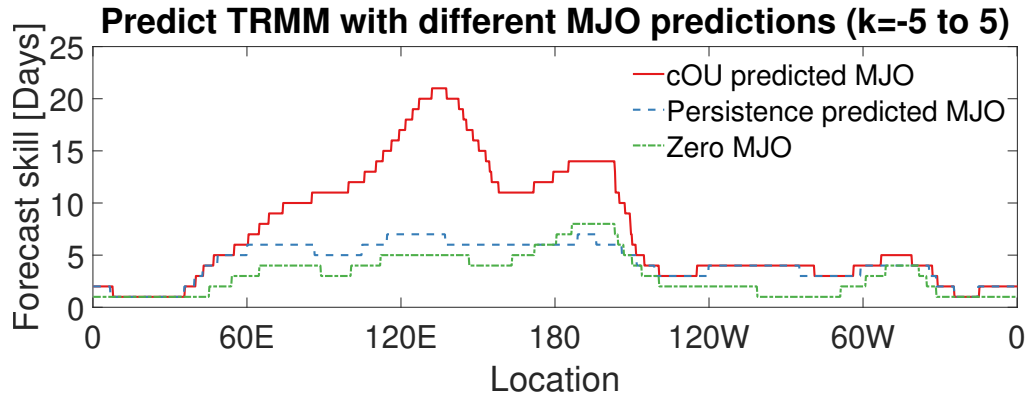


Figure 2.7: Forecast skills for predicting five-years TRMM data on planetary scales (wavenumber from -5 to 5) with different prediction methods for the MJO.

The dominant coherent wave signal on planetary scales is the MJO. How important is the MJO to the predictability of rainfall on planetary scales? To investigate this question, a wave-removal study is implemented where the MJO is removed in one of two different ways: predict the MJO-associated rainfall is zero, or use a persistence prediction for the MJO-associated rainfall. The results are shown in Fig. 2.7. The forecast skill of tropical rainfall is significantly decreased when the MJO is removed from consideration. Specifically, the forecast skill is only 4 to 7 days over the Indian and Pacific Ocean regions if the MJO is removed. Hence, the MJO plays a crucial role in

the predictability of rainfall on planetary scales, and it contributes up to 15 days of additional predictability.

2.5 Comparing predictability of precipitation versus cloudiness (OLR)

Besides precipitation, another quantitative measure of moist convection is OLR. OLR data is commonly used as a proxy for cloudiness, for many purposes, such as identifying CCEWs or the MJO (e.g., [32, 86]). In prior studies on predictions of CCEWs and the MJO, [15] analyzed precipitation whereas [25] analyzed OLR. The predictability could possibly vary significantly depending on which data is used, since, e.g., OLR and precipitation represent distinct physical quantities with potentially different properties. In this section, we investigate the predictability of OLR, and we compare it with the predictability of precipitation.

For comparing OLR and precipitation, results are shown in Fig. 2.8, and can be summarized as follows.

On the one hand, for the broad conclusions regarding comparisons of different wave types, the results are essentially the same, and they are therefore not repeated in detail for the OLR case. For instance, whether OLR or precipitation data is used, each of the wave types is seen to have an appreciable contribution to the predictability of the full rainfall signal. The detailed amounts of each wave's contribution can be different for OLR versus precipitation, as illustrated in Fig. 2.8 for the case of the non-wave component (see section 2.4 for a note about the enhanced predictability of the non-wave component near 120E longitude). Nevertheless, the broad conclusion remains the same: each of the wave types has an appreciable contribution.

On the other hand, one clear difference between OLR and precipitation is that OLR is generally speaking more predictable. Fig. 2.8 quantifies the difference in predictability. The larger predictability of OLR is seen whether viewing individual wave types or the full rainfall signal. The detailed amount can be different for different wave types; for example, for the non-wave component, Fig. 2.8 shows that OLR is more predictable than precipitation by only roughly 1 to 2 days at most locations, whereas for the full rainfall signal, OLR is more predictable than precipitation by roughly 5 to 10 days at many locations over the Indian and Pacific Ocean regions. Other wave types, such as the MJO (not shown), are also noticeably more predictable in terms of OLR versus precipitation.

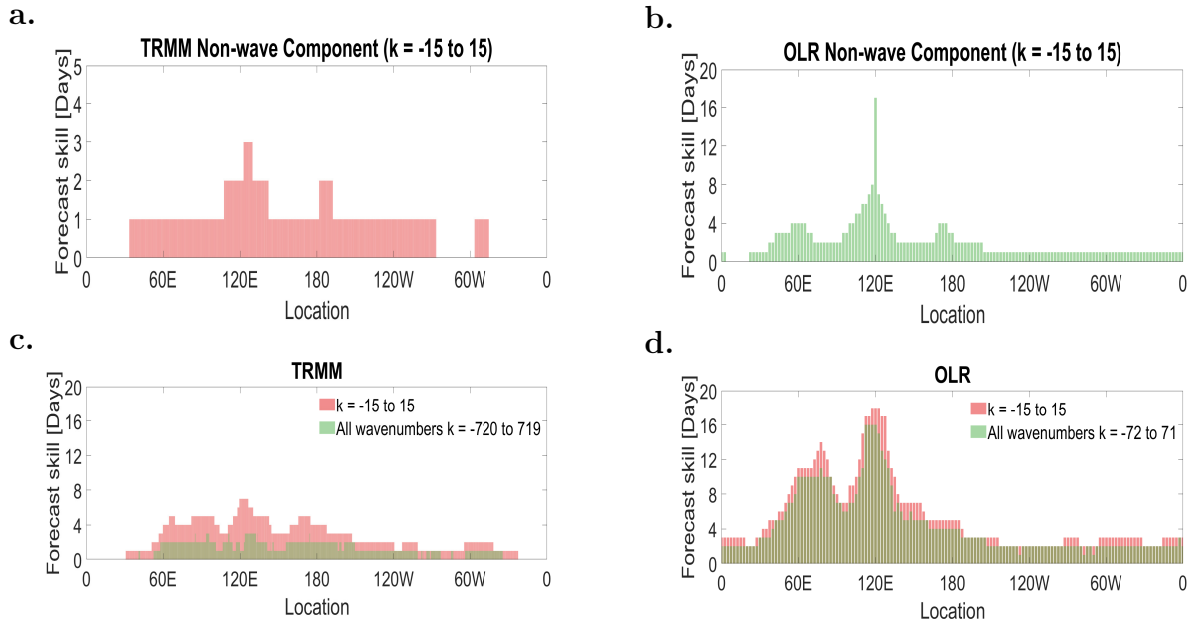


Figure 2.8: Forecast skills for the non-wave component of (a) TRMM and (b) OLR. Forecast skills for the full signal including all wavetypes for (c) TRMM and (d) OLR. Red bar is for the forecast skill from the comparison between the predictions of 31 wavenumbers and the 31 wavenumbers of the true signal, green bar is for the forecast skill from the comparison between the predictions of 31 wavenumbers and the true signal with all wavenumbers.

As some additional sensitivity tests, beyond considering two quantities (precipitation and OLR), we also analyzed the impact of using only zonal wavenumbers -15 to 15 in the predictions. In other words, recall that only the synoptic and planetary scales (zonal wavenumbers -15 to 15) have been considered throughout the present paper. Such a choice was made in part because the predictability of wavenumbers with $10 \leq |k| \leq 15$ is already low (see Table 2.1 and Fig. 2.2) and in part because the focus was on CCEWs and the MJO. The predictability of higher wavenumbers ($|k| > 15$) is low, but they could influence the overall predictability if they account for significant variance. To investigate, we now ask the question: If only zonal wavenumbers -15 to 15 are predicted, and if all higher wavenumbers are predicted to be zero, and if the prediction is assessed via comparison with the full dataset (including $|k| > 15$), then what is the resulting prediction skill? In other words, if only wavenumbers -15 to 15 are modeled, how skillfully can the full dataset (including $|k| > 15$) be predicted? The results are shown in Fig. 2.8c for the full TRMM rainfall dataset, which includes zonal wavenumbers -720 to 720. The predictability drops significantly, consistent with the fact that the higher wavenumbers (smaller scales) contribute an appreciable amount of variance and tend to be less predictable (e.g., see Table 2.1 and Fig. 2.2). A similar drop in predictability is also seen for OLR in Fig. 2.8d, although the higher resolution comparison is not too drastically different in its predictability compared to the standard case. OLR predictability may be less sensitive to changes in resolution because it is generally more smoothly varying than precipitation, or possibly because the higher resolution OLR data includes only wavenumbers -72 to 72 (as opposed to the precipitation case which includes wavenumbers -720 to 720 at its highest resolution).

2.6 Discussion

In this section, several additional tests are discussed, along with some additional discussion and comparisons with other studies.

While the results above were obtained using daily data, another version of TRMM rainfall data is also available as 3-hourly data, which offers additional possibilities, such as the resolution of WIG waves. As one test with the 3-hourly data, we compared two cases: one case where WIG waves were treated as its own wave type and another case where WIG waves were included with the non-wave component. In these tests, for each time available in the time series, a 24-hour running time average was used to average over the effects of the diurnal cycle; note that the result is still a 3-hourly time series, although each data point corresponds to a 24-hour time average. In comparing these two cases, the results were essentially the same, suggesting that, at least when analyzed after a 24-hour running time average, the WIG waves do not contribute a substantial addition to the predictability. This is possibly due to the fact that the spectrum is red, so the variance of high-frequency waves such as the WIG waves is a relatively small contribution to the overall rainfall variance and predictability.

Other sensitivity studies were also carried out to examine different ways of modeling the non-wave components. Since the non-wave component has a wide range of frequencies, whereas the CCEW types were defined over more restrictive ranges of wavenumbers and frequencies, one may want to examine alternative methods where the non-wave component is divided into subcomponents. Two alternative cases were considered. First, the non-wave component was divided into the two subcomponents of oscillation periods less than 5 days and greater than 5 days. As a second case, three components were used: periods less than 5 days, between 5 and 10 days, and greater than 10 days. In these alternative cases, the forecast skill for the non-wave components are nearly the same as in the standard case in Fig. 2.2 and Table 2.1. For the different wavenumbers, the skill is either the same as in the standard case, or sometimes 1 day better skill, and 2 days better skill for a small number of wavenumbers. Hence the precise forecast skill of the non-wave component changes a little if different methods are used, but the overall broad conclusions about different wave types still hold.

It is interesting to make some further comparisons with other prior work. For instance, practical and intrinsic predictability of multiscale weather and CCEWs over Indian Ocean were investigated based on the Weather Research and Forecasting (WRF) Model in [88]. They found that the practical predictability limit decreases rapidly as scale decreases along with many other interesting error analysis and results about the predictability. More specifically, one of the conclusions in [88] is that precipitation has a more limited predictability in comparison to other variables, and its practical predictability limit is only 3 days for large scales and < 12 h for the smaller scales. Both the decreasing predictability for smaller scales and the limited predictability of precipitation are also reflected in the present chapter to some extent. In other work, [57] estimate MJO predictability to be 20-30 days based on single-member hindcast and 35-45 days based on ensemble-mean hindcasts, respectively. Those results are somewhat consistent with the 20-30-day predictability of the MJO estimated in the present chapter, although it is difficult to compare in detail due to the different quantities predicted. [57] focus on predictions of the real-time multivariate (RMM) MJO index of [86], which is based on zonal winds and OLR, variables that are typically seen to be more predictable than precipitation, which was the variable of focus in the present chapter. Beyond these examples of prior work, there are also a number of other interesting papers that examine the practical predictability of current models (e.g., [34, 82, 87] and references therein) .

2.7 Conclusion

In this chapter, the main goal was to analyze the predictability of CCEWs and the MJO, and to assess how important each wave type is to the predictability of the full rainfall signal. The methodology utilized observational data as much as possible, in order to avoid the influence of any particular model's assumptions about detailed physics parameterizations. Also, the methodology allowed the forecasts to be decomposed into the contributions from each wave type (CCEWs, the MJO, and a background spectrum

or non-wave component), and each wave type was treated with an independent forecast model. In this way, one wave type could be considered by itself in isolation, or one wave type could be excluded to measure its influence on the full rainfall signal.

CCEWs and the MJO were seen to provide a significant source of predictability. If the tropics had no CCEWs nor MJO, then it would arguably be a tropics where the entire spectrum looks like the non-wave component, or background spectrum or random scattered thunderstorms (e.g., [22]). Since the non-wave component of rainfall has a predictability of 0 to 3 days, whereas the the overall predictability of tropical rainfall was estimated to be roughly 3 to 6 days, over the Indian and Pacific Ocean regions, one could say that CCEWs and the MJO contribute approximately 3 additional days of predictability. From a slightly different viewpoint, one could say that CCEWs and the MJO double the range of predictability, from 0–3 days to 3–6 days. These results help to quantify the importance of CCEWs and the MJO, in terms of rainfall predictability, as an alternative measure of importance beyond more traditional measures such as climatological variance (e.g., [77, 78, 84]).

Two factors are perhaps sufficient to explain the importance of each wave type: decorrelation time and variance. For the first factor, if a wave type has a long decorrelation time, then it is also likely to be predictable at long lead times. For the second factor, if a wave type has a large variance—i.e., if it contributes a significant fraction of the overall variance—then it is likely to make a significant contribution to the overall predictability as well. In the case of the non-wave component, which has largest contribution to variance of all mode types, the contribution is to limit the overall predictability to several days, in opposition to the other wave types, such as the MJO, which are more predictable but which have lesser contributions to the overall variance. The wave-exclusion studies of section 2.4 helped to quantify the contributions of each wave type, and to illustrate the importance of the two factors of decorrelation time and

variance.

In a comparison of two datasets, OLR data was seen to be much more predictable than TRMM precipitation data. Such a result is consistent with the well-known general property that OLR data is more smoothly varying in space and time than precipitation data. Here a quantitative comparison was given in terms of predictability. OLR was seen to be more predictable than rainfall, over many locations in the Indian and Pacific Ocean regions, by approximately 5 to 10 days. Given that rainfall predictability was estimated to be roughly 3 to 6 days, the additional predictability of 5 to 10 days for OLR is quite large. One implication is that it could be difficult to compare and contrast different studies, such as [15] who analyzed precipitation and [25] who analyzed OLR.

Finally, results were also provided for another question: What are the decay time scales of CCEWs and the MJO? Observational estimates of the decay time scales were provided here in Table 2.1, and they can be compared with the decay time scales predicted by theoretical models (e.g., [62,73]). As a brief comparison, [73] report theoretical values of MJO decay time scales of roughly 1.5 months in their standard parameter regime. Here the MJO decay time scale was estimated from precipitation observations to be roughly 3 months. Note, however, that the 3-month estimate was based on fitting the autocorrelation function for lags of roughly 0 to 30 days (in order to provide an accurate forecast for these relatively short lag times). Instead, if the autocorrelation function were fit for longer lags, such as 0 to 100 days (see Fig. 2.1a,d), then an estimate of roughly 1.5 months would be more appropriate, based on the decay of the autocorrelation function from lag 0 to a lag of 50 or 100 days. As another estimate, if OLR is used instead of precipitation data, we found decay times of roughly 1.7 months for zonal wavenumbers 1, 2, and 3. In brief, while different methods could lead to different decay time scales from observational estimates, and while different parameter values could lead to different decay time scales from theoretical models, there is some broad agreement

between observations and theory in this preliminary comparison. It would be interesting in the future to make a more detailed comparison between theoretical and observational estimates of CCEW and MJO decay time scales.

Chapter 3

Spatial and Temporal Averaging Windows and Their Impact on Forecasting: Exactly Solvable Examples [39]

3.1 Introduction

Weather predictions are commonly made for *averaged* quantities. For example, the amount of rainfall might be predicted as an average in time over an hour, day, week, month, season, etc. [70, 87, 92]. For longer time averages, the forecasts are often associated with intraseasonal variability such as the Madden–Julian Oscillation (MJO) [1, 8, 31, 36, 46, 47, 51, 90] or El Niño–Southern Oscillation (ENSO) [3, 35, 56, 64, 79]. For shorter time averages, the forecasts are often associated with synoptic variability or individual convective storms. The question of spatial averaging becomes particularly relevant as further details of convective systems are resolved by newer numerical weather prediction models, and many new challenges arise in formulating forecasts of precipitation and in assessing errors and uncertainties [2, 5, 16, 17, 55, 66, 68].

One of the main questions here is: How does prediction skill vary as the space-

and/or time-averaging window is varied? Here this question is investigated using exactly solvable models, which allows precise assessment of forecasts without being tarnished by errors from numerical approximations or pseudo-random Monte Carlo sampling. The first model used here is the complex Ornstein–Uhlenbeck (cOU) process, a stochastic model for time series of weather fluctuations (e.g. [50]). The second model used here is the stochastic heat equation, which has been used as an idealized spatiotemporal model of tropical rainfall [22, 72, 73].

A priori, one would perhaps expect from basic intuition that spatially and/or temporally averaged variables should have longer predictability. Intuition would suggest that averaging should filter out the high-frequency, nearly unpredictable components of the system, and only the low-frequency, more-predictable components of the system would remain. For example, a time average over a week or a month should “average out” the small-scale fluctuations in rainfall associated with individual convective storms, and the remaining, averaged signal should be associated with low-frequency variability such as the MJO and/or ENSO.

In the results shown here, this basic intuition sometimes holds and sometimes does not. Different results will be seen for spatial versus temporal averaging. Spatial averaging leads to improved forecast skill, in line with the basic intuition described above. Time averaging, on the other hand, is more subtle: it may either increase or decrease forecast skill, depending on the relative definition of the lead time.

Our original motivation stemmed from analyzing some preliminary forecasts of precipitation associated with convectively coupled equatorial waves (CCEWs) and the MJO, using the model of [73]. In forecasting the rainfall associated with such multiscale systems, the skill could potentially be measured in a variety of ways, depending on the length scales and time scales and wave types of interest. Some very recent studies have now analyzed the forecast skill in this setting using operational numerical weather

prediction models [15, 26, 88, 89]. In the future it would be interesting to conduct more detailed forecasts with the model of [73], which could offer theoretical perspectives owing to the model's simplicity relative to numerical weather prediction systems.

This chapter is organized as follows. The model equations and prediction setup are defined in section 3.2. Forecasts of time series of the cOU process are presented in section 3.3. Subtleties of time averaging are discussed in further detail in section 3.4, including comparisons of different definitions of lead time relative to the averaging window. Forecasts of the spatiotemporal rainfall model (the stochastic heat equation) are presented in section 3.5, allowing investigation of both time averaging and space averaging, and allowing investigation of time averaging for a more complicated system that has numerous degrees of freedom. Conclusions are described in section 3.6.

3.2 Models and Methods

In this section, the two exactly solvable models are introduced: the complex Ornstein-Uhlenbeck process as a model for a single oscillatory mode, and the stochastic heat equation as a model for spatiotemporal dynamics. In addition, several aspects of the forecasting setup are also described, including measures of forecast skill and definitions of time and space averaging.

3.2.1 Mathematical Models

Complex Ornstein-Uhlenbeck(cOU) Process

A complex Ornstein-Uhlenbeck(cOU) process is applied for the first insight of the forecasting behaviors in predicting signals at single time points or signals averaged over a temporal window. Conducting forecasting analyses on a cOU process is a meaningful attempt for detecting the forecast skill for predicting signals with a single oscillation mode.

We consider a complex linear stochastic differential equation for the Ornstein-Uhlenbeck process

$$du(t) = (-\gamma + i\omega)u(t) + \sigma dW(t), \quad (3.1)$$

where $\gamma, \sigma > 0$ and ω are real numbers and

$$dW(t) \equiv \frac{dW_1(t) + idW_2(t)}{\sqrt{2}} \quad (3.2)$$

is a complex Gaussian white noise where each component satisfies

$$dW_j(t) \equiv \dot{W}_j(t)dt, \quad j = 1, 2 \quad (3.3)$$

that is, white noise is a "derivative" of the Wiener process $W_j(t)$ and it satisfies the following properties

$$\mathbb{E}[\dot{W}_j(t)] = 0 \quad (3.4)$$

$$\mathbb{E}[\dot{W}_j(t)\dot{W}_j(s)] = \delta(t - s) \quad (3.5)$$

$$\mathbb{E}[\dot{W}_i(t)\dot{W}_j(s)] = 0 \quad \text{for } i \neq j. \quad (3.6)$$

The exact solution of (3.1) is

$$u(t) = e^{(-\gamma+i\omega)t}u(0) + \sigma \int_0^t e^{(-\gamma+i\omega)(t-s)}dW(s). \quad (3.7)$$

As $t \rightarrow \infty$, $u(t)$ will converge to a stationary Gaussian distribution with mean 0 and variance $\frac{\sigma^2}{2\gamma}$ (see details in [50]).

In the simulation and plotting in this chapter, we choose the parameters $\gamma = 0.1 \text{ month}^{-1}$, $\omega = \frac{4}{3}\pi \text{ month}^{-1}$, $\sigma = 0.5$ to represent a wave with oscillation period

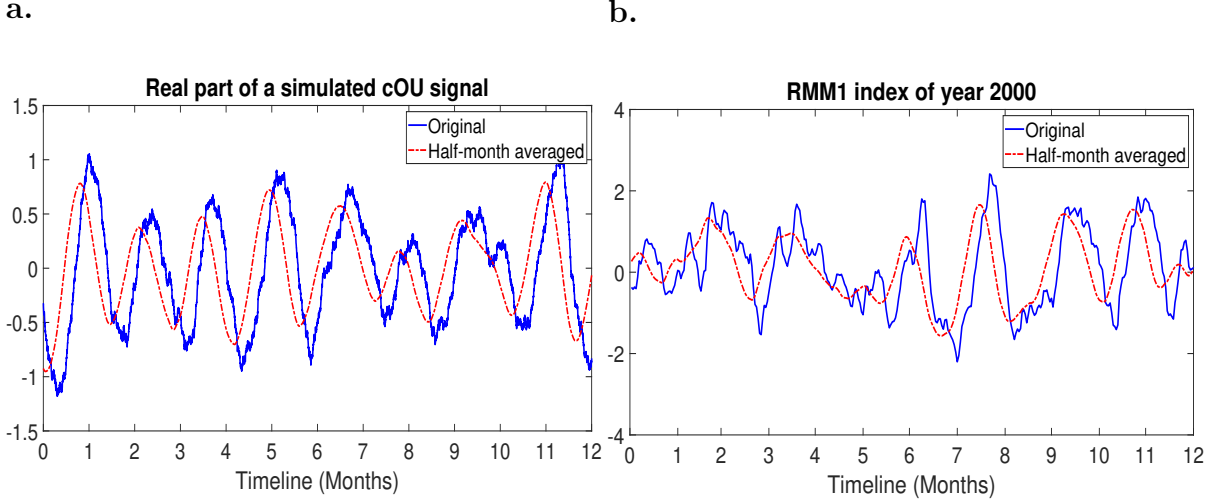


Figure 3.1: Panel **a**: Time series of the real part of a simulated signal from a cOU process over 12 months; Panel **b**: Time series of RMM1 index in the year 2000. In both panels, the blue curve is the original signal while the dashed red curve is the half-month averaged signal (namely $\hat{u}(t, T_w = 0.5 \text{ month})$ from $t = 0$ to $t = 12$ months in (3.12)).

$2\pi/\omega = 1.5$ months like the MJO and a decorrelation time $1/\gamma = 10$ months that is longer than the oscillation period, resulting in the occurrence of sequences of irregular wave oscillations. All the analytic formulas have been checked with the numerical simulations. A simulated time series over 12 months is shown in panel **a** of Fig. 3.1 along with the time series of Real-time Multivariate MJO (RMM) index in the year 2000. The similarities between the two time series show that the cOU signals are good representations for the dynamics to some extent. When a time averaging is applied, small oscillations will be removed from the original signal through the temporal averaging process so that the signal becomes smoother with an expectation for a lower variance.

For investigating more about influences of temporal and spatial averaging in forecasting, a one dimensional model based on the stochastic heat equation is analyzed for some insight. The simple model applied to the water vapor dynamics has been shown that it has behaviors very similar to the observational statistics in the article [22]. The

model is a stochastic PDE

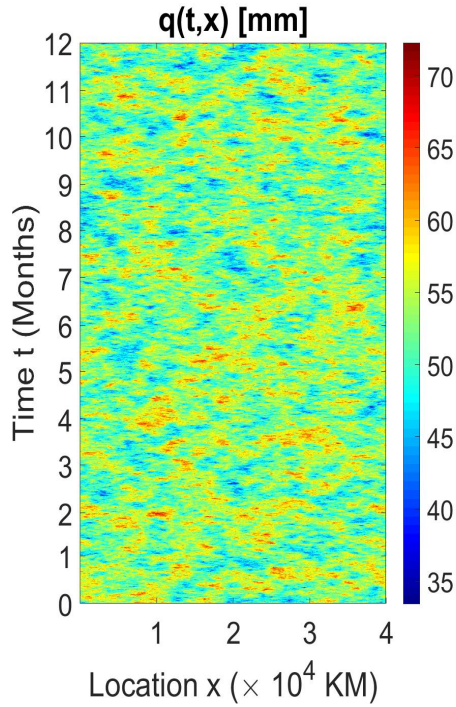
$$\frac{\partial q}{\partial t} = b_0 \nabla^2 q - \frac{1}{\tau}(q - q^*) + F + D_* \dot{W}, \quad (3.8)$$

where $q(t, x)$ is the signal at time t and location x depending on a spatial interaction constant b_0 , the relaxation time τ , the relaxation target q^* , an external force F and the stochastic forcing variance D_*^2 . Now, $\dot{W}$ is a space-time white noise. It can be characterized as a centered Gaussian process such that $\mathbb{E}[\dot{W}(s, x)\dot{W}(t, y)] = \delta(s - t)\delta(x - y)$. We call (3.8) as SHE or SPDE for abbreviations in this chapter. In the one dimensional case, (3.8) can be solved analytically as

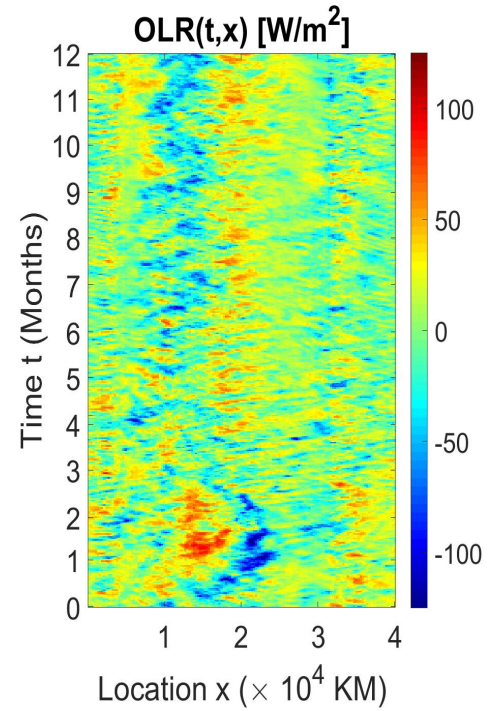
$$\begin{aligned} q(t, x) = & q^* + \tau F + \frac{e^{-t/\tau}}{(4\pi b_0 t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t}} (q(0, y) - q^* - \tau F) dy \\ & + e^{-t/\tau} D_* \int_0^t \frac{1}{(4\pi b_0 |t-s|)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 (t-s)}} e^{s/\tau} \dot{W}(s, y) dy ds. \end{aligned} \quad (3.9)$$

Besides the analytical solution, there are many useful and important statistics that can be computed directly using the analytical solution, such as the temporal covariance $cov[q(t, x)q(s, x)]$ and the spatial covariance $cov[q(t, x)q(t, y)]$ which will help a lot in accessing the forecast skill (See in appendix section 3.7.3). The signal arising from this type of stochastic partial differential equation like (3.8) will enter into a stationary Gaussian distribution as time goes to infinity. The stationary distribution has mean $q^* + \tau F$ and variance $\frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}}$. In addition, in spatial Fourier space, $q(t, x)$ can be decomposed into the integral of a bunch of components satisfying the equation of a cOU process (3.1) independently. All the components have a corresponding decorrelation time $1/\gamma$ in the cOU process that is not greater than τ and $\omega = 0$. From this point of view, $q(t, x)$ can be seen as the combination of numerous different waves with decorrelation time not greater than τ and oscillation frequency $\omega = 0$. A detailed description and derivation for the discrete version of this Fourier transform can be seen in [22].

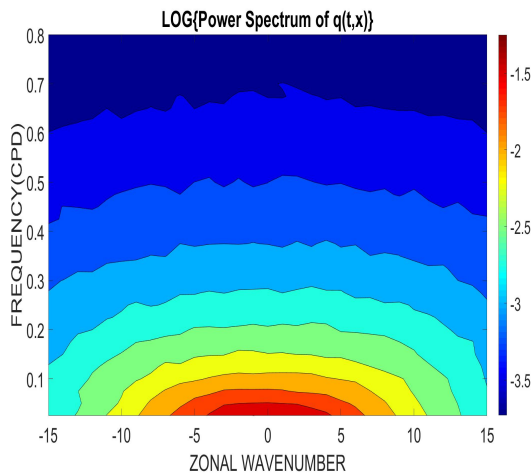
a.



b.



c.



d.

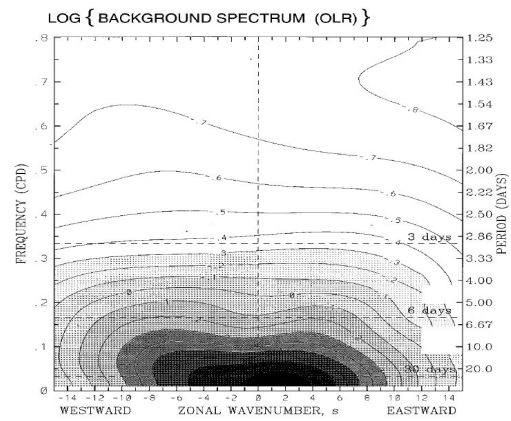


Figure 3.2: Panel **a**. A simulated global image of $q(t,x)$; Panel **b**. A global image of OLR observational data in the year 2010 after averaging over the tropical belt; Panel **c**. The power spectrum density of the simulated data; Panel **d**. Zonal wavenumber-frequency spectrum of the base-10 logarithm of the background power from the OLR observational data cited from Wheeler&Kiladis's paper [84] © American Meteorological Society. Used with permission.

In our simulation and figures, we aim to simulate the rainfall statistics according to the probability distribution in [22] and the power spectrum density plot in [84]. $\tau = 96$ hours, $q^* = 65$ mm, $F = -0.125$ mm $\cdot$ hour $^{-1}$, $b_0 = 10^4$ km 2 $\cdot$ hour $^{-1}$, $D_* = 25$ mm $\cdot$ km $^{1/2}$ $\cdot$ hour $^{-1/2}$ are chosen for this purpose. All the analytic formulas have been checked with the numerical simulation. Both the temporal averaging and the spatial averaging have an influence on eliminating the small oscillations and smoothing. A global dynamic view of the simulated $q(t, x)$ over one year is shown in panel **a** of Fig. 3.2 along with the OLR observational data in the year of 2010 after averaging over the tropical belt, which provides a OLR signal $OLR(t, x)$ for time t and longitude x . The power spectrum density plot of the simulated data in panel **c** of Fig. 3.2 performs a good match with the power density based on the observational OLR data in the paper [84] (shown in panel **d** of Fig. 3.2). The similarities between the observational data and the simulated data provide evidence for that stochastic heat equation is a reasonable and good model for representing the atmospheric dynamic system.

3.2.2 Measures of Forecasting Skill: Mean Square Error (MSE) and Pearson Correlation Coefficients (ρ)

Two types of frequently-used criteria, mean square error(MSE) and the Pearson correlation coefficients, are used in this chapter for evaluating the forecasting skill. The mean square error measures the average of the squares of the errors that is the difference between the predictions and the true signal. When a temporal or spatial averaging is applied, the variance of the averaged signal will also be scaled from the variance of the original signal. Hence, instead of the MSE, the ratio of the MSE over the variance of the averaged signal seems to be a more reasonable and meaningful measure for the assessment. The Pearson correlation coefficient is a measure of the strength and direction of the linear relationship between two variables. Mathematically, if $X(t)$ is the true signal from the cOU process (3.1) or the stochastic heat equation (3.8) and $Y(t)$ is our

prediction, both $X(t)$ and $Y(t)$ are stationary and ergodic processes and $(X(t), Y(t))$ is also a jointly stationary and ergodic process. Stationarity ensures that the first moment $\mathbb{E}[X(t)]$, $\mathbb{E}[Y(t)]$, the second moment $\mathbb{E}[X(t)^2]$, $\mathbb{E}[Y(t)^2]$ and the mixed moment $\mathbb{E}[X(t)Y(t)]$ are all constants that are not depending on the time t . Then the MSE of the estimator $Y(t)$ with respect to the unknown variable $X(t)$ is defined as

$$MSE(X(t), Y(t)) = \mathbb{E}[(Y(t) - X(t))^2], \quad (3.10)$$

which is a constant not depending on the time t . The formula for the Pearson correlation coefficients ρ can be expressed as

$$\rho(X(t), Y(t)) = \frac{cov(X(t), Y(t))}{\sigma_X \sigma_Y}, \quad (3.11)$$

where $cov(X(t), Y(t)) = \mathbb{E}[(X(t) - \mathbb{E}[X(t)])(Y(t) - \mathbb{E}[Y(t)])]$ is the covariance between $X(t), Y(t)$ and σ_X, σ_Y are the standard deviations of X, Y respectively. $\rho(X(t), Y(t))$ is also a constant that doesn't change over time. In a situation of analyzing the real data, we have the true data from the true signal as $X_1, X_2, \dots, X_N$ and our corresponding predictions $Y_1, Y_2, \dots, Y_N$. By ergodicity, as $N \rightarrow \infty$, the sample means $\frac{1}{N} \sum_{i=1}^N X_i$, $\frac{1}{N} \sum_{i=1}^N Y_i$ converge in squared mean to $\mathbb{E}[X(t)]$, $\mathbb{E}[Y(t)]$ and $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_i^2 = \mathbb{E}[X(t)^2]$, $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N Y_i^2 = \mathbb{E}[Y(t)^2]$, $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N X_i Y_i = \mathbb{E}[X(t)Y(t)]$. These connect the sample MSE and sample correlation coefficient of a real dataset with our idealized and theoretical definitions (3.10) and (3.11).

3.2.3 Temporal and Spatial Averaging: Definitions and Notations

When we do the temporal averaging, there are different ways to place the temporal averaging window. Motivated by Fig.1 in the paper [92], the default definitions of the lead time when there is a temporal averaging window in this chapter are shown in

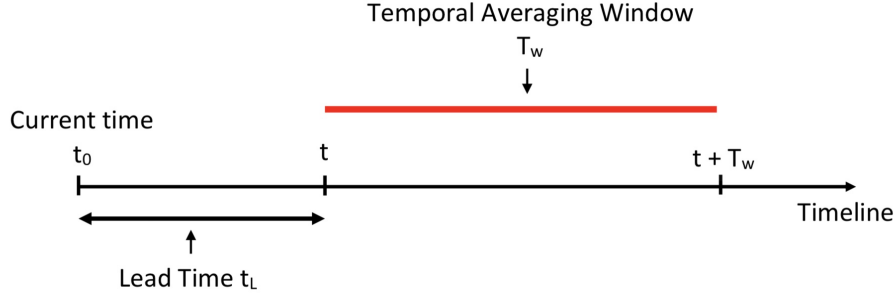


Figure 3.3: Schematic of default lead time and temporal averaging window definitions in the analysis. The horizontal axis represents forecast time from the initial condition at t_0 .

Fig. 3.3. For the temporal averaged cOU signal $\hat{u}(t, T_w)$ at time t with averaging window width T_w , it is defined as the averaged signal from time t to time $t + T_w$, namely

$$\hat{u}(t, T_w) = \frac{1}{T_w} \int_t^{t+T_w} u(s) ds. \quad (3.12)$$

Similarly, the temporal averaged stochastic heat equation signal at time t with averaging window width T_w at a single location x is defined as

$$\hat{q}(t, x, T_w) = \frac{1}{T_w} \int_t^{t+T_w} q(s, x) ds. \quad (3.13)$$

For the spatial averaging at location x , the signals are averaged over a spatial window $[x - L_w/2, x + L_w/2]$ with L_w as the averaging window width and x as the center of the spatial averaging window, namely

$$\tilde{q}(t, x, L_w) = \frac{1}{L_w} \int_{x-L_w/2}^{x+L_w/2} q(t, y) dy. \quad (3.14)$$

The averaged signal with both temporal and spatial averaging is a combination of (3.13) and (3.14) as

$$\bar{q}(t, x, T_w, L_w) = \frac{1}{L_w T_w} \int_{x-L_w/2}^{x+L_w/2} \int_t^{t+T_w} q(s, y) ds dy. \quad (3.15)$$

The changes of variances in the cOU or SHE signals through the temporal and/or spatial averaging are also investigated, which is described in appendix section 3.7.1 in detail.

Note that both the complex Ornstein-Uhlenbeck process and the stochastic heat equation are Markovian processes. With the aids of Markovian properties of the two models, a lot of useful analytical formulas are derived and calculated when we assess the forecasting skills involving temporal and/or spatial averaging in the later sections. It would be interesting in the future to investigate whether similar results are still seen with non-Markovian processes.

3.3 Forecasting the Complex Ornstein-Uhlenbeck Process

3.3.1 Forecasting at a Single Time Point

In an idealized situation, we have the perfect model of a cOU process for predicting, which means we have the exact true values of the parameters γ, ω, σ in (3.1). However, by using the ensemble mean prediction, the value of σ essentially will not affect the prediction skill. The true signal is a known realization from the cOU process, we denote it as

$$u_{truth}(t; \theta) = \text{The true signal} = u(t; \theta) \quad (3.16)$$

with the underlying parameter θ which labels the realization of the stochastic process. By (3.7), treating $t - t_L$ as the current time, one single prediction for time t with lead time

t_L will be another realization starting from $u_{truth}(t - t_L; \theta)$ with underlying parameter ζ as

$$\begin{aligned}
u_{pred}(t, t_L; \zeta, \theta) &= \text{The prediction of the value of } u \text{ at time } t, \\
&\text{given the value of } u \text{ at time } t - t_L, \\
&\text{for ensemble member label } \zeta \\
&= e^{(-\gamma+i\omega)t_L} u_{truth}(t - t_L; \theta) + \sigma \int_{t-t_L}^t e^{(-\gamma+i\omega)(t-s)} dW(s; \zeta). \quad (3.17)
\end{aligned}$$

Note that, technically speaking, $W(s; \zeta)$ here is from another new Wiener process which could be denoted as $\tilde{W}(s; \zeta)$, although the tilde will be left off in order to ease notation. Also note that the prediction u_{pred} can be separated into components from the θ realization and ζ realization according to the Markovian nature of the process, as the future is independent from the past.

At the same time $u_{truth}(t; \theta)$ is also a realization starting from $u_{truth}(t - t_L; \theta)$ with underlying parameter θ as

$$u_{truth}(t; \theta) = e^{(-\gamma+i\omega)t_L} u_{truth}(t - t_L; \theta) + \sigma \int_{t-t_L}^t e^{(-\gamma+i\omega)(t-s)} dW(s; \theta). \quad (3.18)$$

Then the ensemble mean of the predictions in the form of (3.17) serves as the ensemble forecasting for time t with ensemble size ∞ which is

$$\begin{aligned}
u_{pred}(t, t_L; \theta) &= \mathbb{E}_{\zeta}[u_{pred}(t, t_L; \zeta, \theta)] \\
&= e^{(-\gamma+i\omega)t_L} u_{truth}(t - t_L; \theta). \quad (3.19)
\end{aligned}$$

First, we predict the signal at single time points with the perfect cOU model and lead time t_L . The mean square error (MSE_u) and the Pearson correlation coefficient (ρ_u)

between the prediction and the true signal can be calculated as

$$\begin{aligned}
MSE_u(t_L) &= \mathbb{E}_\theta[|u_{truth}(t; \theta) - u_{pred}(t, t_L; \theta)|^2] \\
&= \mathbb{E}_\theta[\sigma^2 \int_{t-t_L}^t e^{(-\gamma+i\omega)(t-s)} dW(s; \theta)^2] \\
&= \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma t_L}),
\end{aligned} \tag{3.20}$$

where $\frac{\sigma^2}{2\gamma}$ is the variance of the true signal.

$$\begin{aligned}
\rho_u(t_L) &= \frac{\mathbb{E}_\theta[(u_{pred}(t, t_L; \theta) - \mathbb{E}_\theta[u_{pred}(t, t_L; \theta)])(u_{truth}(t; \theta) - \mathbb{E}_\theta[u_{truth}(t; \theta)])^*]}{(var[u_{pred}(t, t_L; \theta)])^{1/2} (var[u_{truth}(t; \theta)])^{1/2}} \\
&= \frac{\mathbb{E}_\theta[u_{pred}(t, t_L; \theta) u_{truth}^*(t; \theta)]}{(var[u_{pred}(t, t_L; \theta)])^{1/2} (var[u_{truth}(t; \theta)])^{1/2}} \\
&= \frac{e^{(-\gamma+i\omega)t_L} \mathbb{E}_\theta[u_{truth}(t - t_L; \theta) u_{truth}^*(t; \theta)]}{e^{-\gamma t_L} (var[u_{truth}(t - t_L; \theta)])^{1/2} (var[u_{truth}(t; \theta)])^{1/2}} \\
&= \frac{e^{(-\gamma+i\omega)t_L} e^{(-\gamma-i\omega)t_L} \mathbb{E}_\theta[u_{truth}(t - t_L; \theta) u_{truth}^*(t - t_L; \theta)]}{e^{-\gamma t_L} \frac{\sigma^2}{2\gamma}} \\
&= e^{-\gamma t_L}.
\end{aligned} \tag{3.21}$$

With a perfect model, the mean square error will keep increasing as the lead time increases but will never exceed the variance of the underlying true signal. As lead time grows to infinity, MSE_u is approaching the variance of the true signal. The correlation coefficient $\rho_u(t_L)$ keeps decreasing with respect to the increasing in the lead time and will drop below 0.5 if the lead time t_L is larger than $\ln(2)/\gamma$, approximately 0.69 of the decorrelation time $1/\gamma$. The characteristics of the forecasting skills at a single time point for a cOU process are shown with the blue lines in Fig. 3.4. Since the conclusions from MSE ratios are consistent with those from the correlation coefficients throughout this chapter, we show only MSE formulas but not plots everywhere.

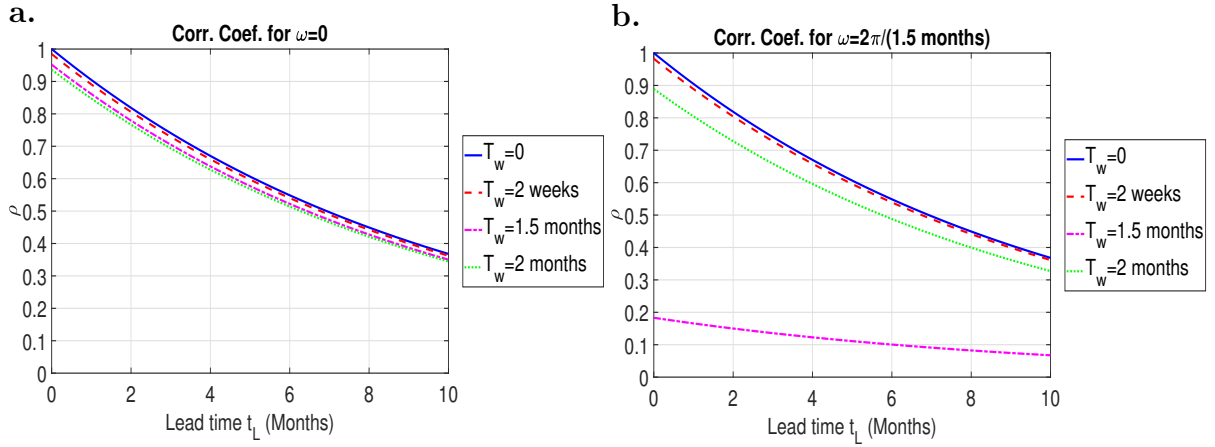


Figure 3.4: Panel **a.** forecasting skill (correlation coefficient) for different temporal averaging windows of a cOU process with $\omega = 0$ under the default definitions clarified in the section 3.2.3; Panel **b.** forecasting skill (correlation coefficient) for different temporal averaging windows of a cOU process with $\omega = 2\pi/(1.5 \text{ months})$ under the default definitions clarified in the section 3.2.3.

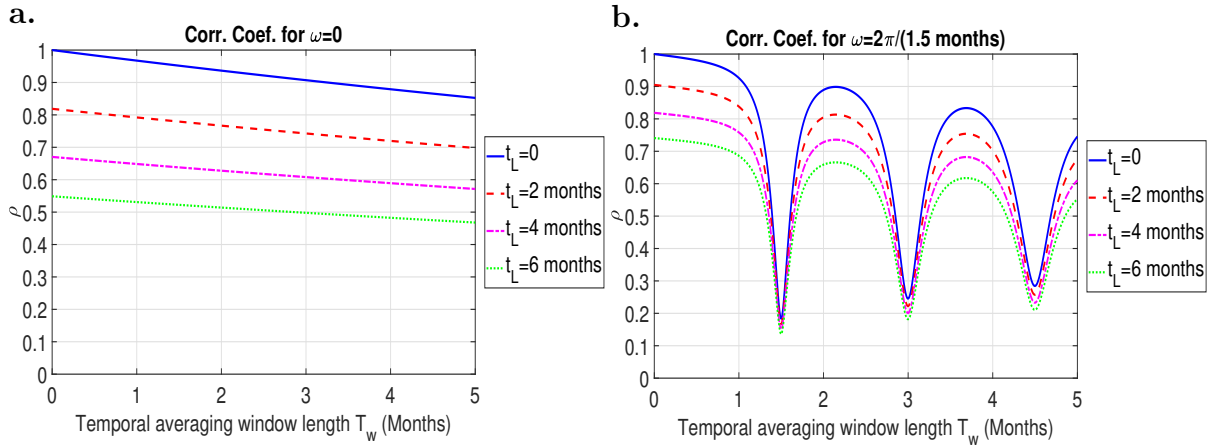


Figure 3.5: Panel **a.** forecasting skill for different lead times of a cOU process with $\omega = 0$ under the default definitions clarified in the section 3.2.3; Panel **b.** forecasting skill for different lead times of a cOU process with $\omega = 2\pi/(1.5 \text{ months})$ under the default definitions clarified in the section 3.2.3.

3.3.2 Forecasting with Temporal Averaging

Now we move on to assess the forecasting skill with a temporal averaging window. The forecasting skill is achieved by comparing the averaged true cOU signal at time t , namely

$$\begin{aligned}\widehat{u}_{truth}(t, T_w; \theta) &= \frac{1}{T_w} \int_t^{t+T_w} u_{truth}(s; \theta) ds \\ &= \frac{1}{T_w} \int_t^{t+T_w} e^{(-\gamma+i\omega)(s-t+t_L)} u_{truth}(t-t_L; \theta) \\ &\quad + \sigma \int_{t-t_L}^s e^{(-\gamma+i\omega)(s-s')} dW(s'; \theta) ds\end{aligned}\tag{3.22}$$

and the ensemble prediction at time t with lead time t_L by the default definition clarified in the section 3.2.3 for this averaged true signal. Seeing $t-t_L$ as the current time, a single prediction with lead time t_L for this averaged signal with underlying ensemble label ζ is specified as

$$\begin{aligned}\widehat{u}_{pred}(t, t_L, T_w; \zeta, \theta) &= \frac{1}{T_w} \int_t^{t+T_w} u_{pred}(s, s-(t-t_L); \zeta, \theta) ds \\ &= \frac{1}{T_w} \int_t^{t+T_w} e^{(-\gamma+i\omega)(s-t+t_L)} u_{truth}(t-t_L; \theta) \\ &\quad + \sigma \int_{t-t_L}^s e^{(-\gamma+i\omega)(s-s')} dW(s'; \zeta) ds.\end{aligned}\tag{3.23}$$

Then the ensemble forecast is the mean of the above forecast with respect to ζ

$$\begin{aligned}\widehat{u}_{pred}(t, t_L, T_w; \theta) &= \mathbb{E}_\zeta[\widehat{u}_{pred}(t, t_L, T_w; \zeta, \theta)] \\ &= \frac{1}{T_w} \int_t^{t+T_w} e^{(-\gamma+i\omega)(s-t+t_L)} u(t-t_L; \theta) ds.\end{aligned}\tag{3.24}$$

Having all these ready, we proceed to compute the forecasting skill in this temporal averaging case.

$$\begin{aligned}
\widehat{MSE}_u(t_L, T_w) &= \mathbb{E}_\theta[|\widehat{u}_{truth}(t, T_w; \theta) - \widehat{u}_{pred}(t, t_L, T_w; \theta)|^2] \\
&= \mathbb{E}_\theta\left[\left|\frac{1}{T_w} \int_t^{t+T_w} \sigma \int_{t-t_L}^s e^{(-\gamma+i\omega)(s-s')} dW(s'; \theta) ds\right|^2\right] \\
&= \text{var}[\widehat{u}_{truth}(t, T_w; \theta)] - \frac{\sigma^2}{2\gamma} \frac{e^{-2\gamma t_L}}{(\gamma^2 + \omega^2) T_w^2} [e^{-2\gamma T_w} - 2e^{-\gamma T_w} \cos(\omega T_w) + 1],
\end{aligned} \tag{3.25}$$

where $\text{var}[\widehat{u}_{truth}(t, T_w; \theta)]$ is given in (3.52).

We do not present the details of the calculations that lead to (3.25), since they are relatively long and they provide limited additional insight. To check the accuracy of the formula in (3.25), we conducted Monte Carlo numerical simulations of the stochastic process in order to find statistical estimates of the forecast skill, and the formula in (3.25) was in agreement with the statistical estimates (not shown). This same scenario will repeat itself numerous times in the remainder of the chapter; namely, many details of calculations will be omitted, but the accuracy of the formulas was verified using numerical statistical estimates.

From the formula in (3.25), one can see the following properties. With $\cos(\omega T_w) \leq 1$, the term $e^{-2\gamma T_w} - 2e^{-\gamma T_w} \cos(\omega T_w) + 1 \geq (e^{-\gamma T_w} - 1)^2 \geq 0$, which indicates that $\widehat{MSE}_u(t_L, T_w) \leq \text{var}[\widehat{u}_{truth}(t, T_w; \theta)]$ always holds. The mean square error is always controlled under the variance of the temporal averaged signal if forecasting with a perfect model. It can be verified that as $T_w \rightarrow \infty$, $\widehat{MSE}_u(t_L, T_w) \rightarrow 0$ and as $t_L \rightarrow \infty$,

$$\widehat{MSE}_u(t_L, T_w) \rightarrow \text{var}[\widehat{u}_{truth}(t, T_w)].$$

For the Pearson correlation coefficient,

$$\begin{aligned} & \widehat{\rho}_u(t_L, T_w) \\ &= \frac{\mathbb{E}_\theta[\widehat{u}_{pred}(t, t_L, T_w; \theta) \widehat{u}_{truth}^*(t; \theta)]}{(\text{var}[\widehat{u}_{pred}(t, t_L, T_w; \theta)])^{1/2} (\text{var}[\widehat{u}_{truth}^*(t; \theta)])^{1/2}} \\ &= \left[\frac{\text{var}[\widehat{u}_{pred}(t, t_L, T_w; \theta)]}{\text{var}[\widehat{u}_{truth}(t; \theta)]} \right]^{1/2} \\ &= \rho_u(t_L) \cdot \left[\frac{(\gamma^2 + \omega^2)(e^{-2\gamma T_w} - 2e^{-\gamma T_w} \cos(\omega T_w) + 1)}{2\gamma T_w(\gamma^2 + \omega^2) - 2(\gamma^2 - \omega^2) + 2e^{-\gamma T_w} \left((\gamma^2 - \omega^2) \cos(\omega T_w) - 2\gamma\omega \sin(\omega T_w) \right)} \right]^{1/2}. \end{aligned} \quad (3.26)$$

The performances of the forecasting skills for different temporal averaging windows and different lead time are demonstrated graphically in Fig. 3.4 and Fig. 3.5. (Note that a special behavior can sometimes arise when the temporal averaging window T_w is an integer multiple of the oscillation period, $2\pi/\omega$. For example, in Fig. 3.5, if T_w equals the oscillation period of 1.5 months, the prediction will have a very bad forecasting skill. This is mainly because the signal is almost completely averaged out when the temporal averaging window equals an integer multiple of the oscillation period; see also appendix 3.7.1.)

From these two figures 3.4 and 3.5, we can see clearly an interesting result: as the temporal averaging window increases, the correlation coefficient actually decreases, which means a worse forecasting for a larger temporal averaging window. This is counter-intuitive. When a temporal averaging window is applied, intuitively we would expect a lot of fast oscillations and unpredictable components would be averaged out, making the resulting signal much easier to predict. But the results from the formulas tell us a

totally counter story that the temporal averaging makes the forecast skill worse!

3.4 Subtle Impacts of Definitions of Averaging Window and Lead Time

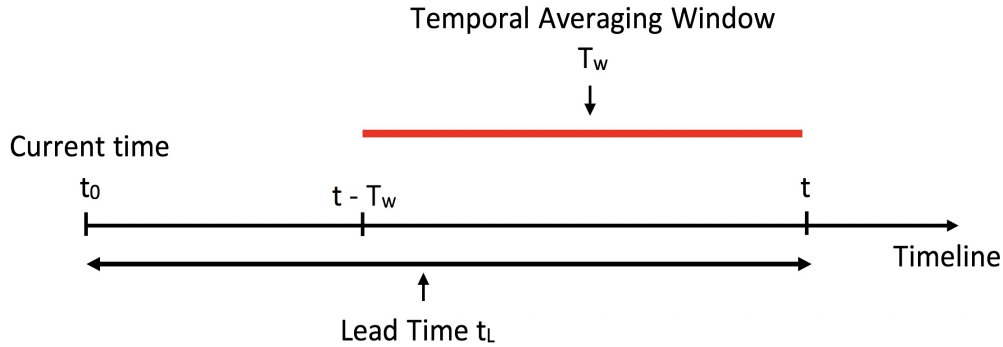


Figure 3.6: Schematic of a **new** lead time and temporal averaging window definition in the analysis. The horizontal axis represents forecast time from the initial condition at t_0 .

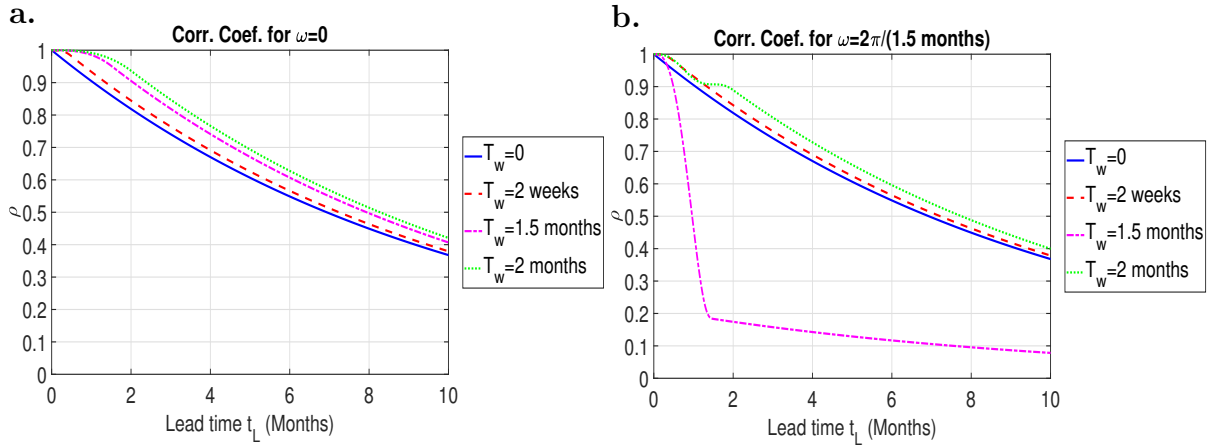


Figure 3.7: Panel **a.** forecasting skill for different temporal averaging windows of a cOU process with $\omega = 0$ under the new definition clarified in the Fig. 3.6 and equation (3.27); Panel **b.** forecasting skill for different temporal averaging windows of a cOU process with $\omega = 2\pi/(1.5 \text{ months})$ under the new definition clarified in the Fig. 3.6 and equation (3.27).

In the previous section, why did a wider temporal averaging window give a worse forecasting skill, counter to what one would expect from basic intuition? With this question in mind, we have tried to review all of the details of the procedure in order to determine the underlying cause. One key factor seems to be the relative definitions of the lead time and the temporal averaging window, and the relative definition seems to cause differences in the forecasting skills under different settings. To illustrate the different ways of making such definitions, here in this section we use an alternative definition: the time of the right endpoint of a temporal averaging interval is taken to be the “base time” of the averaged signal, which means a temporal averaged signal over a temporal averaging window with length T_w is defined as

$$\widehat{u}_{truth}^R(t; \theta) = \frac{1}{T_w} \int_{t-T_w}^t u_{truth}(s; \theta) ds. \quad (3.27)$$

A diagram for this new definition is shown in Fig. 3.6. In this new definition, it needs $T_w < t_L$ to make the entire temporal averaging window lie in the future. If $T_w > t_L$, then part of the averaging window lies in the past which means that only part of the signals in the averaging window that lies in the future needs to be predicted since the past is already known. Comparing the old and new definitions, the forecasting skill can be easily evaluated by replacing the old lead time t_L in (3.25) and (3.26) with $t_L - T_w$ if $t_L \geq T_w$, namely $\widehat{MSE}_u^R(t_L, T_w) = \widehat{MSE}_u(t_L - T_w, T_w)$ and $\widehat{\rho}_u^R(t_L, T_w) = \widehat{\rho}_u(t_L - T_w, T_w)$. If $t_L < T_w$, the prediction in this new definition will be made of a part of already known past values and a part of future forecasting. When $t_L < T_w$, $\widehat{MSE}_u^R(t_L, T_w) = \widehat{MSE}_u(0, t_L) \cdot t_L^2 / T_w^2$ while $\widehat{\rho}_u^R$ needs to be calculated starting from the new definition. Here, we omit the calculation details and directly give out the formulas.

$$\widehat{MSE}_u^R(t_L, T_w)$$

$$= \begin{cases} \text{var}[\widehat{u}_{truth}(t, t_L; \theta)] \frac{t_L^2}{T_w^2} - \frac{\sigma^2}{2\gamma} \frac{1}{(\gamma^2 + \omega^2) T_w^2} [e^{-2\gamma t_L} - 2e^{-\gamma t_L} \cos(\omega t_L) + 1] & \text{if } t_L < T_w \\ \text{var}[\widehat{u}_{truth}(t, T_w; \theta)] - \frac{\sigma^2}{2\gamma} \frac{e^{-2\gamma T_w}}{(\gamma^2 + \omega^2) T_w^2} [e^{2\gamma T_w} - 2e^{\gamma T_w} \cos(\omega T_w) + 1] & \text{if } t_L \geq T_w \end{cases} \quad (3.28)$$

$$\widehat{\rho}_u^R(t_L, T_w) = \begin{cases} \left[\frac{\text{var}[\frac{1}{T_w} (\int_{t+t_L-T_w}^t u_{truth}(s; \theta) ds + \int_t^{t+t_L} u_{pred}(s, s-t; \theta) ds)]}{\text{var}[\widehat{u}_{truth}(t, T_w; \theta)]} \right]^{1/2} & \text{if } t_L < T_w \\ \rho_u(t_L) \left[\frac{(\gamma^2 + \omega^2)(e^{2\gamma T_w} - 2e^{\gamma T_w} \cos(\omega T_w) + 1)}{2\gamma T_w (\gamma^2 + \omega^2) - 2(\gamma^2 - \omega^2) + 2e^{-\gamma T_w} ((\gamma^2 - \omega^2) \cos(\omega T_w) - 2\gamma \omega \sin(\omega T_w))} \right]^{1/2} & \text{if } t_L \geq T_w \end{cases} \quad (3.29)$$

where

$$\begin{aligned} & \text{var}[\frac{1}{T_w} (\int_{t+t_L-T_w}^t u_{truth}(s; \theta) ds + \int_t^{t+t_L} u_{pred}(s, s-t; \theta) ds)] \\ &= \frac{1}{T_w^2} \left[(T_w - t_L)^2 \text{var}[\widehat{u}_{truth}(t, T_w - t_L; \theta)] + \frac{\sigma^2}{2\gamma} \frac{1}{\gamma^2 + \omega^2} (e^{-2\gamma t_L} - 2e^{-\gamma t_L} \cos(\omega t_L) + 1) \right. \\ & \quad \left. + \frac{\sigma^2}{2\gamma} \text{Re}[\frac{2}{(\gamma + i\omega)^2} (1 - e^{(-\gamma - i\omega)t_L})(1 - e^{(-\gamma - i\omega)(T_w - t_L)})] \right] \end{aligned} \quad (3.30)$$

and $\text{var}[\widehat{u}_{truth}(t, T_w; \theta)]$ is given in (3.52).

From Fig. 3.7, we care more about the case when $t_L > T_w$ where the entire temporal averaging window falls in the future. Now, under the new definition, when $\omega = 0$, the forecasting skills are improving when a wider temporal averaging window is applied. For $\omega \neq 0$, things get a little bit more complicated. Forecast is very unskillful when the signal is averaged over a window with length that is multiples of the oscillation period $2\pi/\omega$. But the highest skill in the second cycle for $2\pi/\omega \leq T_w \leq 4\pi/\omega$ is even better than the highest skill in the first cycle when $0 \leq T_w \leq 2\pi/\omega$ for positive lead time.

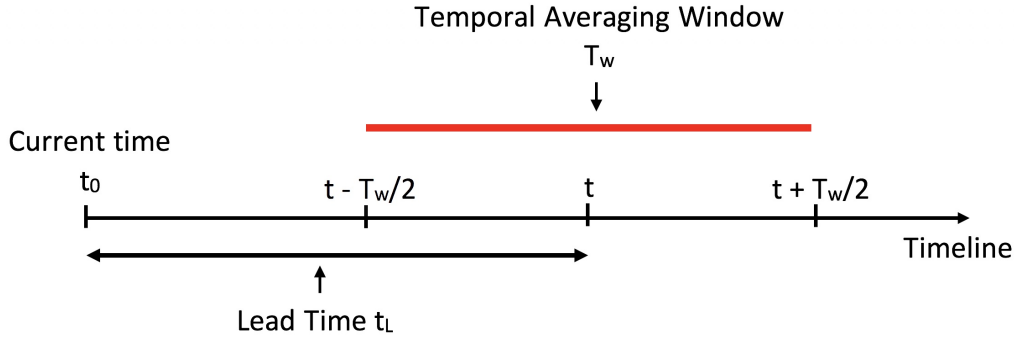


Figure 3.8: Schematic of another **new** lead time and temporal averaging window definition in the analysis. The horizontal axis represents forecast time from the initial condition at t_0 .

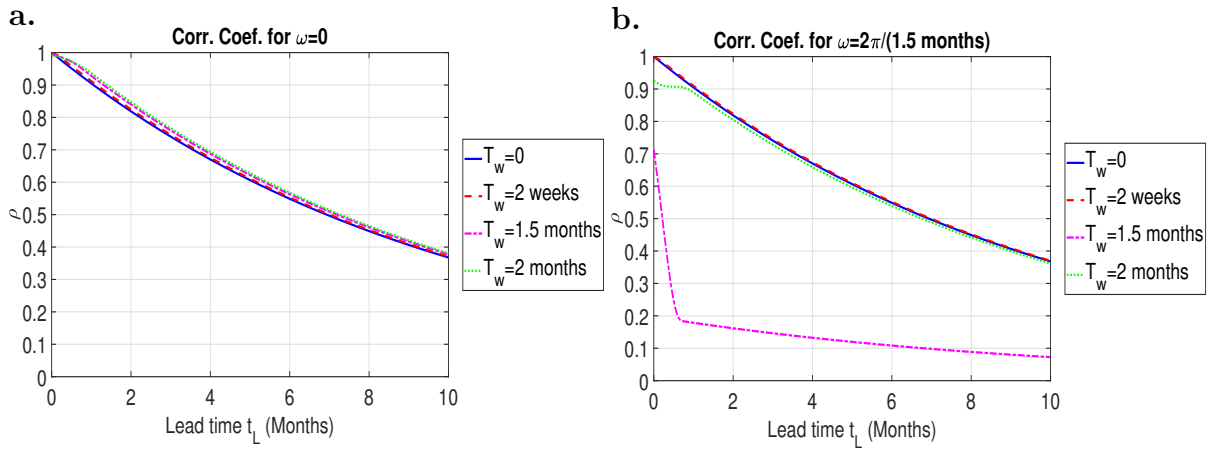


Figure 3.9: Panel **a.** Pearson correlation coefficient of forecasting a cOU process with $\omega = 0$ under the new definition clarified in the Fig. 3.8; Panel **b.** Pearson correlation coefficient of forecasting a cOU process with $\omega = 2\pi/1.5$ months under the new definition clarified in the Fig. 3.8.

These results show in a counter way as those under the default definition we discussed in the section 3.3.2. This is all caused by the different definitions of the averaging window.

Another choice of the temporal averaging window at time t is to put the window center exactly at the time point t as in the Fig. 3.8. With curiosity, we also dig into this

case. The formulas of the new forecasting skills $MSE_u^C(t_L, T_w)$ and $\rho_u^C(t_L, T_w)$ can be derived directly based on (3.28) and (3.29) through the relationship $MSE_u^C(t_L, T_w) = MSE_u^R(t_L + T_w/2, T_w)$ and $\rho_u^C(t_L, T_w) = \rho_u^R(t_L + T_w/2, T_w)$ and their behaviors can be observed in Fig. 3.9. For $\omega = 0$, the forecasting skill is similar to but slightly better than the single time point predictions. For the $\omega = 2\pi/1.5$ months, the very limited forecasting skill stays the same at T_w is multiples of oscillation cycles.

A summary figure, Fig. 3.10, is plotted to give a comprehensive view of the changes in forecast skill with different definitions of the temporal averaging window. As the center of the temporal averaging window over time t moves from the left side to the right side of time t , such as from the definition in Fig. 3.6 to Fig. 3.3, the forecast skill becomes worse. The centered definition in Fig. 3.8 has slightly better forecast skill than the baseline which is the single time point forecasting.

Note that two different γ values are chosen in Fig. 3.10, $\gamma = 0.1$ in panel **a** and $\gamma = 0.5$ in panel **b**. We can see that the differences in forecast skill among different locations of the temporal averaging windows are slight when $\gamma = 0.1$ as the y-axis in panel **a** of Fig. 3.10 ranges from about 0.95 to 1.15 but are much more significant when $\gamma = 0.5$. This is because the correlation coefficient ratio $\hat{\rho}(t_L, T_w)/\rho(t_L)$ under different definitions depends on γT_w , which can be seen as follows. If we think the baseline is the case of single time point forecasting and then define a as

$$a = \frac{\text{The temporal averaging window center minus the lead time } t_L \text{ in baseline}}{T_w}. \quad (3.31)$$

Then a is a variable representing where the temporal averaging window is placed. Denote $\hat{\rho}(t_L, T_w)$ under the definition of the temporal averaging window with a as $\hat{\rho}(t_L, T_w, a)$.

When $\omega = 0$, the correlation coefficient ratio can be calculated as

$$\begin{aligned}\hat{\rho}(t_L, T_w, a)/\rho(t_L) &= \frac{\hat{\rho}(t_L - (T_w/2 - aT_w), T_w)}{\rho(t_L)} \\ &= (\exp(-\gamma T_w))^a \left[\frac{\exp(\gamma T_w) + \exp(-\gamma T_w) - 2}{2\gamma T_w + 2\exp(-\gamma T_w) - 2} \right]^{1/2}.\end{aligned}\quad (3.32)$$

It is depending on γT_w while independent from t_L .

In any case, to summarize, it is clear from Fig. 3.10 that the definition of the time averaging window, relative to the definition of the lead time, has a significant impact on the change in forecast skill. Certain definitions can even lead to the counter-intuitive result that time averaging causes a worse forecast skill. A “centered” time averaging window, which is centered about the target prediction time, appears to offer only a small increase in forecast skill, despite intuitive expectations that time averaging should offer larger increases in forecast skill.

3.5 Forecasting an Idealized Spatiotemporal Rainfall Model: Stochastic Heat Equation

A complex Ornstein-Uhlenbeck process is an interesting test case since it is a good presentation for a single wave. Nevertheless, spatial averaging issues cannot be conducted in a forecasting problem for a cOU signal. Meanwhile, in the real world, we are often trying to predict signals composed of numerous waves or modes. The cOU process is insufficient for getting insights for those situations. Therefore, we proceed to study the forecasting problem for the signal arising from a stochastic heat equation in the form of (3.8). Now we return to the default definition of the time-averaging window from section 3.2.3 and take a look at the forecasting behaviors for this type of signals.

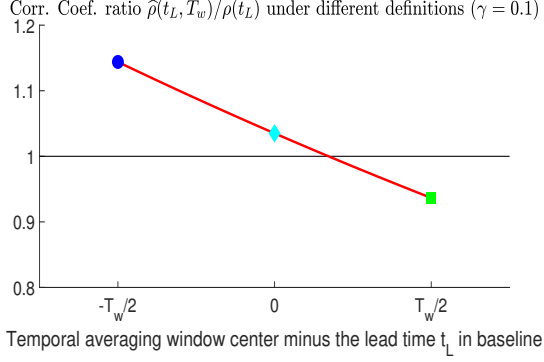
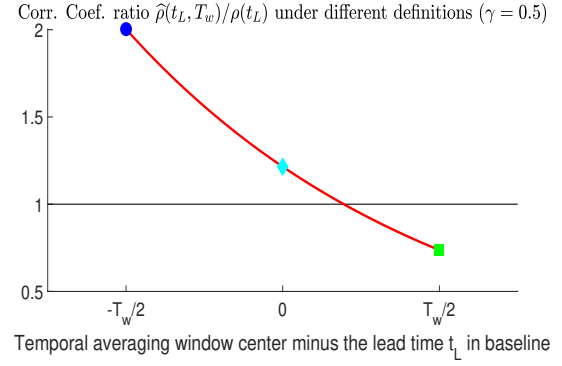
a.**b.**

Figure 3.10: Forecast skills (Pearson correlation coefficients) with temporal averaging under different definitions as in Fig. 3.13 compared to the baseline (single time point forecasting) for predicting cOU processes with $\omega = 0$. $T_w = 2$ months is fixed and the correlation coefficient ratio is not depending on the lead time t_L . Only cases where the entire averaging window for forecasting lies in the future are considered here (in other words, $t_L > T_w$). The blue dot stands for the definition in Fig. 3.6 with better forecast skill, the cyan diamond represents the definition of the similar skill as baseline (slightly better) in Fig. 3.8, the green square is used for the definition of the worse forecast skill in Fig. 3.3. Panel **a** is for the case $\gamma = 0.1$ and panel **b** is for the case $\gamma = 0.5$.

3.5.1 Forecasting at a Single Time Point and a Single Spatial Location

The stochastic PDE (3.8) can be solved analytically. For convenience, let's denote

$$Q(t, x) = q(t, x) - q^* - \tau F, \quad (3.33)$$

then $\mathbb{E}[Q(t, x)] = 0$ for $t \rightarrow \infty$. Similar to the cOU process, the true signal $q_{truth}(t, x; \theta)$ is a realization from (3.8), say it is with underlying parameter θ as

$$q_{truth}(t, x; \theta)$$

$$\begin{aligned}
&= q(t, x; \theta) \\
&= Q(t, x; \theta) + q^* + \tau F \\
&= q^* + \tau F + \frac{e^{-t_L/\tau}}{(4\pi b_0 t_L)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t_L}} Q(t - t_L, y; \theta) dy \\
&+ D_* \int_0^{t_L} \frac{1}{(4\pi b_0 |t_L - s|)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(t-s)}} e^{-(t_L-s)/\tau} \dot{W}(s + t - t_L, y; \theta) dy ds. \quad (3.34)
\end{aligned}$$

A single prediction $q_{pred}(t, x, t_L; \zeta, \theta)$ for predicting $q_{truth}(t, x; \theta)$ with lead time t_L is then a realization with another underlying parameter ζ . In other words, it is

$$\begin{aligned}
&q_{pred}(t, x, t_L; \zeta, \theta) \\
&= q^* + \tau F + \frac{e^{-t_L/\tau}}{(4\pi b_0 t_L)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t_L}} Q(t - t_L, y; \theta) dy \\
&+ D_* \int_0^{t_L} \frac{1}{(4\pi b_0 |t_L - s|)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(t-s)}} e^{-(t_L-s)/\tau} \dot{W}(s + t - t_L, y; \zeta) dy ds. \quad (3.35)
\end{aligned}$$

With lead time t_L , the ensemble forecast for time t is the mean of all the single predictions $q_{pred}(t, x, t_L; \zeta, \theta)$ starting from $q_{truth}(t - t_L, x; \theta) = Q(t - t_L, x; \theta) + q^* + \tau F$, which is

$$\begin{aligned}
q_{pred}(t, x, t_L; \theta) &= \mathbb{E}_{\zeta}[q_{pred}(t, x, t_L; \zeta, \theta)] \\
&= q^* + \tau F + \frac{e^{-t_L/\tau}}{(4\pi b_0 t_L)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t_L}} Q(t - t_L, y; \theta) dy. \quad (3.36)
\end{aligned}$$

Now the forecasting skill can be achieved straightforwardly by comparing the $q_{truth}(t, x; \theta)$ and $q_{pred}(t, x, t_L; \theta)$ and direct calculation.

$$MSE_q(t_L)$$

$$\begin{aligned}
&= \mathbb{E}_\theta[(q_{truth}(t, x; \theta) - q_{pred}(t, x, t_L; \theta))^2] \\
&= \mathbb{E}_\theta[(D_* \int_0^{t_L} \frac{1}{(4\pi b_0 |t_L - s|)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(t-s)}} e^{-(t_L-s)/\tau} \dot{W}(s + t - t_L, y; \theta) dy ds)^2] \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \text{erf}(\sqrt{\frac{2t_L}{\tau}}). \tag{3.37}
\end{aligned}$$

Note $\frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}}$ is the variance of the true signal. Hence the mean square error is again controlled under the variance of the true signal with the perfect model.

$$\begin{aligned}
\rho_q(t_L) &= \frac{\mathbb{E}_\theta[(q_{pred}(t, x, t_L; \theta) - \mathbb{E}_\theta[q_{pred}(t, x, t_L; \theta)])(q_{truth}(t, x; \theta) - \mathbb{E}_\theta[q_{truth}(t, x; \theta)])]}{(var[q_{pred}(t, x, t_L; \theta)])^{1/2} (var[q_{truth}(t, x; \theta)])^{1/2}} \\
&= (\frac{var[q_{pred}(t, x, t_L; \theta)]}{var[q_{truth}(t, x; \theta)]})^{1/2} \\
&= [1 - \text{erf}(\sqrt{\frac{2t_L}{\tau}})]^{1/2}. \tag{3.38}
\end{aligned}$$

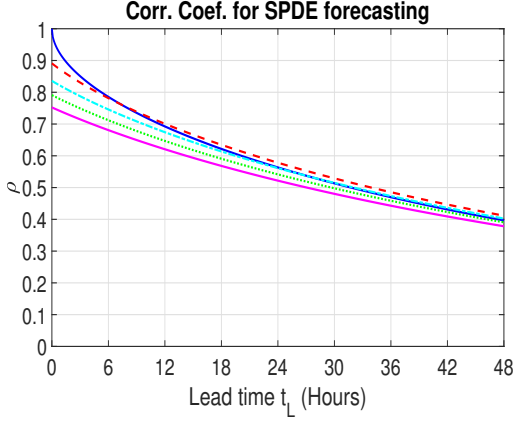
Essentially, the forecasting skill at a single time point and a single spatial location is only related to the ratio of t_L and τ and doesn't involve any other parameters in the model. With $t_L \leq 0.33\tau$, $\rho_q(t_L) \geq 0.5$. In our simulation, for $\tau = 96h = 4 \text{ days}$ in the idealized rainfall model, $t_L \leq 1.32344 \text{ days} = 31.76256 \text{ hours}$ will have a good forecast correlation coefficient $\rho_q(t_L) \geq 0.5$.

3.5.2 Forecasting at a Single Spatial Location with Temporal Averaging

Temporal averaging forecasting skill for a single spatial location for the signal from the stochastic heat equation is studied as well. The temporal averaged signal is

$$\hat{q}_{truth}(t, x, T_w; \theta) = \frac{1}{T_w} \int_t^{t+T_w} q_{truth}(s, x; \theta) ds = \frac{1}{T_w} \int_t^{t+T_w} Q(s, x; \theta) ds + q^* + \tau F, \tag{3.39}$$

a.



b.

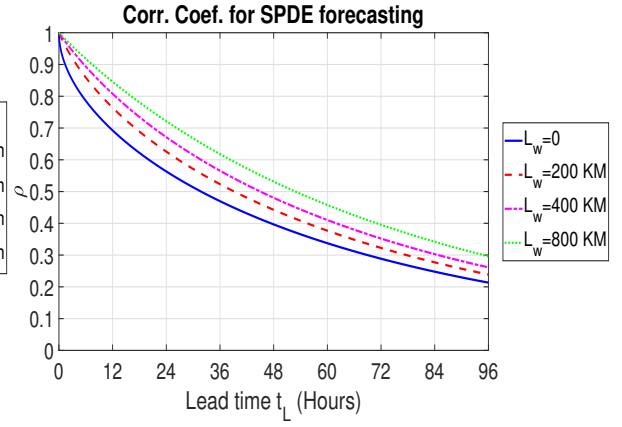


Figure 3.11: Panel **a.** forecasting skill for different temporal averaging windows of a SPDE process under the default definition clarified in the section 3.2.3; Panel **b.** forecasting skill for different spatial averaging windows of a SPDE process under the default definition clarified in the section 3.2.3 .

where

$$\begin{aligned}
 Q(s, x; \theta) &= \frac{e^{-(s-(t-t_L))/\tau}}{[4\pi b_0(s-(t-t_L))]^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(s-(t-t_L))}} Q(t-t_L, y; \theta) dy \\
 &\quad + D_* \int_0^{s-(t-t_L)} \frac{1}{(4\pi b_0|s-(t-t_L)-\mu|)^{1/2}} \cdot \\
 &\quad \left(\int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(s-(t-t_L)-\mu)}} e^{-(s-(t-t_L)-\mu)/\tau} \dot{W}(\mu+t-t_L, y; \theta) dy \right) d\mu \quad (3.40)
 \end{aligned}$$

and the ensemble forecast is achieved by taking the ensemble mean of all the single predictions

$$\begin{aligned}
 \hat{q}_{pred}(t, x, t_L, T_w; \theta) &= \frac{1}{T_w} \int_t^{t+T_w} \frac{e^{-(s-(t-t_L))/\tau}}{[4\pi b_0(s-(t-t_L))]^{1/2}} \cdot \\
 &\quad \left(\int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0(s-(t-t_L))}} Q(t-t_L, y; \theta) dy \right) ds + q^* + \tau F, \quad (3.41)
 \end{aligned}$$

$$\begin{aligned}
\widehat{MSE}_q(t_L, T_w) &= \mathbb{E}_\theta[(\widehat{q}_{truth}(t, x, T_w; \theta) - \widehat{q}_{pred}(t, x, t_L, T_w; \theta))^2] \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \frac{\tau^2}{T_w^2} \left[\frac{2(t_L + T_w)}{\tau} \int_{(2t_L + T_w)/\tau}^{2(t_L + T_w)/\tau} erf(\sqrt{x}) dx \right. \\
&\quad - \frac{2t_L}{\tau} \int_{2t_L/\tau}^{(2t_L + T_w)/\tau} erf(\sqrt{x}) dx \\
&\quad - \frac{2T_w}{\tau} \int_0^{T_w/\tau} erf(\sqrt{x}) dx + \int_{2t_L/\tau}^{(2t_L + T_w)/\tau} x erf(\sqrt{x}) dx \\
&\quad \left. - \int_{(2t_L + T_w)/\tau}^{2(t_L + T_w)/\tau} x erf(\sqrt{x}) dx + 2 \int_0^{T_w/\tau} x erf(\sqrt{x}) dx \right], \quad (3.42)
\end{aligned}$$

where

$$\begin{aligned}
\int_A^B erf(\sqrt{x}) dx &= B erf(\sqrt{B}) - A erf(\sqrt{A}) - \frac{1}{2} erf(\sqrt{B}) + \frac{1}{2} erf(\sqrt{A}) \\
&\quad + \sqrt{\frac{B}{\pi}} exp(-B) - \sqrt{\frac{A}{\pi}} exp(-A) \quad (3.43)
\end{aligned}$$

and

$$\begin{aligned}
\int_A^B x erf(\sqrt{x}) dx &= \frac{1}{2} B^2 erf(\sqrt{B}) - \frac{1}{2} A^2 erf(\sqrt{A}) - \frac{3}{8} erf(\sqrt{B}) + \frac{3}{8} erf(\sqrt{A}) \\
&\quad + \frac{1}{2\sqrt{\pi}} B^{3/2} exp(-B) - \frac{1}{2\sqrt{\pi}} A^{3/2} exp(-A) \\
&\quad + \frac{3}{4\sqrt{\pi}} B^{1/2} exp(-B) - \frac{3}{4\sqrt{\pi}} A^{1/2} exp(-A). \quad (3.44)
\end{aligned}$$

The correlation coefficient can again be proved to be the square root of the ratio between the two variances of the true averaged signal and the prediction as

$$\widehat{\rho}_q(t_L, T_w) = \left(\frac{var[\widehat{q}_{pred}(t, x, t_L, T_w; \theta)]}{var[\widehat{q}_{truth}(t, x, T_w; \theta)]} \right)^{1/2}, \quad (3.45)$$

where

$$var[\widehat{q}_{pred}(t, x, t_L, T_w; \theta)]$$

$$\begin{aligned}
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \left\{ 1 - \frac{\tau^2}{T_w^2} \left[\int_{2t_L/\tau}^{(2t_L+T_w)/\tau} (x \operatorname{erf}(\sqrt{x})) dx - \int_{(2t_L+T_w)/\tau}^{2(t_L+T_w)/\tau} x \operatorname{erf}(\sqrt{x}) dx \right. \right. \\
&\quad \left. \left. - \frac{2t_L}{\tau} \int_{2t_L/\tau}^{(2t_L+T_w)/\tau} \operatorname{erf}(\sqrt{x}) dx + \frac{2(t_L+T_w)}{\tau} \int_{(2t_L+T_w)/\tau}^{2(t_L+T_w)/\tau} \operatorname{erf}(\sqrt{x}) dx \right] \right\} \quad (3.46)
\end{aligned}$$

and $\operatorname{var}[\hat{q}_{truth}(t, x, T_w; \theta)]$ is given in (3.53).

For positive lead time, the skill has slight improvements for very narrow temporal averaging window T_w and then starts to drop as the averaging window becomes wider and wider. Fig. 3.11 shows this characteristic when lead time t_L is longer than about 6 hours. The SHE signal is composed of a lot of cOU waves with different decorrelation times but no oscillations if we view it from Fourier space. Recalling the discussion from the cOU part, the quickly decaying waves with short decorrelation times are hard to predict and those slowly decaying waves with long decorrelation times are predictable. When a temporal averaging window is used, we average out the quickly decaying waves which improves the forecasting skill while the forecasting skill for each individual predictable wave is dropping which may deteriorate the forecasting. The improvement part can be deduced from both intuition and the correlation formula. When we apply the temporal averaging, we remove those unpredictable waves and diminish the overall variance of the signal, making it easier to predict. The variance contribution of the predictions for those quickly decaying waves is always almost zero in $\operatorname{var}[\hat{q}_{pred}(t, x, t_L, T_w; \theta)]$ while temporal averaging definitely reduces the variance of the true signal $\hat{q}_{truth}(t, x, T_w; \theta)$. The Pearson correlation coefficient appears to be the ratio of these two values indicating that the averaging out of those quickly decaying waves will help the forecast. The forecast on those slowly decaying waves is worse when a temporal averaging window under the default definition in section 3.2.3 is applied. Hence, the overall change in the forecasting skill is a trade off between the improvements brought by averaging out the quickly decaying waves and the decline caused by the worse forecast for those slowly decaying waves.

Whether the forecasting skill improves or degrades for a predictable wave when temporal averaging is used highly depends on how we place the temporal averaging window. If a temporal averaging window is placed as in Fig. 3.6, then the forecasting for the SHE signal increases when a wider temporal averaging window is applied since both the averaging out of the quickly decaying signals and the averaging of the slowly decaying signals are helping the forecast. The same thing will happen if we invoke the averaging definition in Fig. 3.8 with the same reasons.

3.5.3 Forecasting at a Single Time Point with Spatial Averaging

For the spatial averaging, we define the spatial averaged signal at location x with spatial window length L_w by averaging all the signals over the interval $[x - L_w/2, x + L_w/2]$ as in (3.14), namely the averaged truth signal $\tilde{q}_{truth}(t, x, L_w; \theta)$ and its corresponding prediction $\tilde{q}_{pred}(t, x, t_L, L_w; \theta)$ are

$$\tilde{q}_{truth}(t, x, L_w; \theta) = \frac{1}{L_w} \int_{x-L_w/2}^{x+L_w/2} q_{truth}(t, y; \theta) dy \quad (3.47)$$

and

$$\tilde{q}_{pred}(t, x, t_L, L_w; \theta) = \frac{1}{L_w} \int_{x-L_w/2}^{x+L_w/2} q_{pred}(t, y, t_L; \theta) dy. \quad (3.48)$$

We omit more details in this section and directly give the formulas for the forecasting skills as following.

$$\begin{aligned} \widetilde{MSE}_q(t_L, L_w) &= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \left\{ \frac{2\sqrt{b_0\tau}}{L_w} - \frac{2b_0\tau}{L_w^2} \operatorname{erf}(\sqrt{2t_L/\tau}) \right. \\ &\quad \left. + \frac{b_0\tau}{L_w^2} \exp\left(-\frac{L_w}{\sqrt{b_0\tau}}\right) \left(1 - \operatorname{erf}\left(\frac{L_w}{\sqrt{8b_0t_L}} - \sqrt{2t_L/\tau}\right)\right) \right\} \end{aligned}$$

$$\begin{aligned}
& -\frac{b_0\tau}{L_w^2}\exp\left(\frac{L_w}{\sqrt{b_0\tau}}\right)\left(1-\operatorname{erf}\left(\frac{L_w}{\sqrt{8b_0t_L}}+\sqrt{2t_L/\tau}\right)\right) \\
& -2\frac{\sqrt{b_0\tau}}{L_w}\exp(-2t_L/\tau)\left(\operatorname{erf}\left(\frac{L_w}{\sqrt{8b_0t_L}}\right)-\frac{\sqrt{8b_0t_L}}{\sqrt{\pi}L_w}\left(1-\exp\left(-\frac{L_w^2}{8b_0t_L}\right)\right)\right)\Bigg\},
\end{aligned} \tag{3.49}$$

$$\tilde{\rho}_q(t_L, L_w) = \left(\frac{\operatorname{var}[\tilde{q}_{pred}(t, x, t_L, L_w; \theta)]}{\operatorname{var}[\tilde{q}_{truth}(t, x, L_w; \theta)]} \right)^{1/2}, \tag{3.50}$$

where

$$\begin{aligned}
\operatorname{var}[\tilde{q}_{pred}(t, x, t_L, L_w; \theta)] = & \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \left\{ \frac{b_0\tau}{L_w^2} \left[\exp\left(-\frac{L_w}{\sqrt{b_0\tau}}\right) + \exp\left(\frac{L_w}{\sqrt{b_0\tau}}\right) - 2 \right] \right. \\
& \frac{b_0\tau}{L_w^2} \exp\left(-\frac{L_w}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{L_w}{\sqrt{8b_0t_L}} - \sqrt{2t_L/\tau}\right) \\
& + b_0\tau \exp\left(\frac{L_w}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(-\frac{L_w}{\sqrt{8b_0t_L}} - \sqrt{2t_L/\tau}\right) \\
& - \frac{2b_0\tau}{L_w^2} \operatorname{erf}\left(-\sqrt{2t_L/\tau}\right) + \frac{2\sqrt{b_0\tau}}{L_w} \exp(-2t_L/\tau) \operatorname{erf}\left(\frac{L_w}{\sqrt{8b_0t_L}}\right) \\
& \left. - \frac{2\sqrt{b_0\tau}}{\sqrt{\pi}L_w} \frac{\sqrt{8b_0t_L}}{L_w} \exp(-2t_L/\tau) \left(1 - \exp\left(-\frac{L_w^2}{8b_0t_L}\right)\right) \right\} \tag{3.51}
\end{aligned}$$

and $\operatorname{var}[\tilde{q}_{truth}(t, x, L_w; \theta)]$ is given in (3.54).

The spatial averaging improves the forecasting skills unambiguously. A wider spatial averaging window produces a better forecast. Since the spatial averaging window is independent from the lead time, we would expect the improvements in forecasting brought by the spatial averaging no matter whether the averaging window at location x is placed more on the left side of x , centered at x or more on the right side of x .

3.5.4 Forecasting with Both Temporal and Spatial Averaging

a.

b.

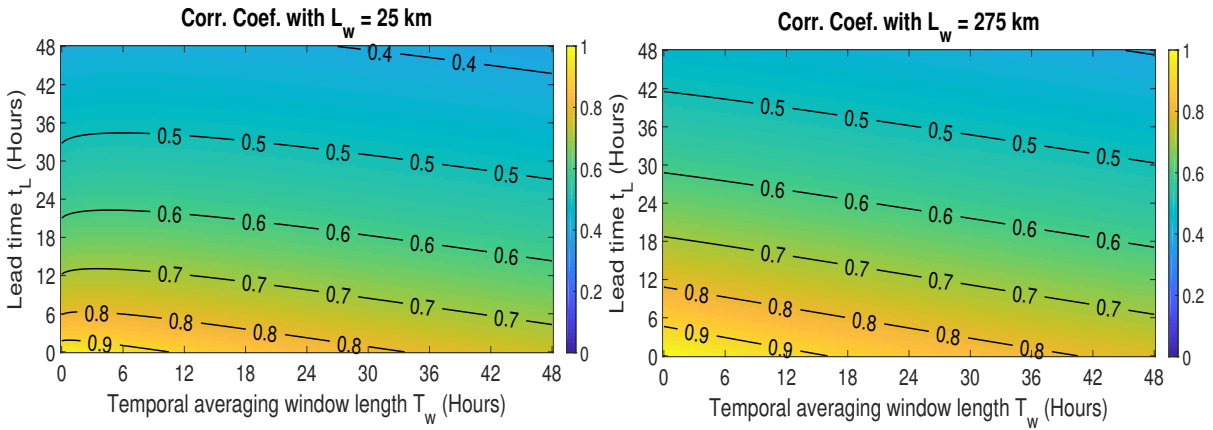


Figure 3.12: Panel **a.** forecasting skill of a temporal averaged SPDE process under the default definition clarified in the section 3.2.3 with a fixed spatial averaging window $L_w = 25$ km; Panel **b.** forecasting skill of a temporal averaged SPDE process under the default definition clarified in the section 3.2.3 with a fixed spatial averaging window $L_w = 275$ km.

Regarding the formulas of forecast skills with both temporal and spatial averaging for predicting a SHE signal, see in appendix section 3.7.2. Instead, we take a look at the patterns through Fig. 3.12 here. The forecasting performance is the combination effects of the temporal and spatial averaging where the spatial averaging always helps to improve the forecasting while the temporal averaging helps the forecasting for very narrow windows and then make the forecasting worse for wider windows. If the definitions shown in Fig. 3.6 or Fig. 3.8 are used, then the temporal averaging is also helping the forecasting so that the skills are always improving with wider spatial and temporal windows in those cases.

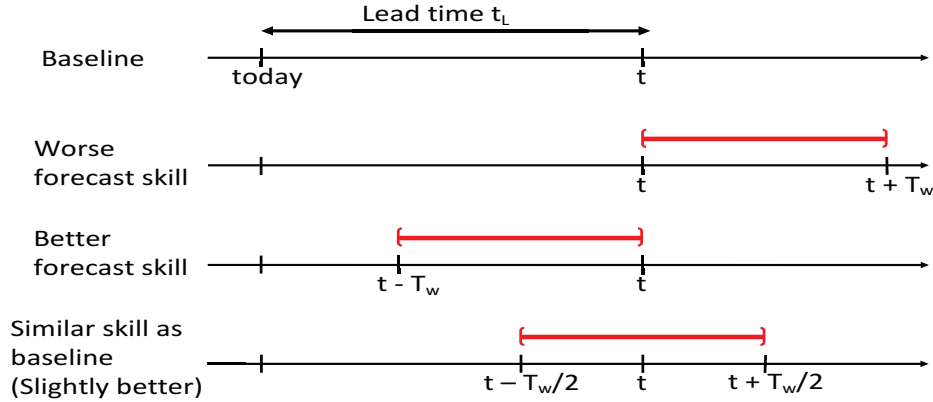


Figure 3.13: Schematics of comparison between different definitions of temporal averaging window and single time point predictions and brief conclusions for predicting cOU processes with $\omega = 0$.

3.6 Conclusions

The effect of time averaging on forecast skill was seen to be significantly different depending on the definitions of the temporal averaging window and the lead time. Two main competing effects contribute to this difference. On the one hand, time averaging should increase forecasting skill by eliminating (“averaging out”) high-frequency, nearly unpredictable components from the signal; such an effect is consistent with some a priori, basic intuition. On the other hand, time averaging can decrease forecasting skill if the time-averaging window $[t, t + T_w]$ lies in the future, beyond the time t of the prediction, since the signal is less predictable for times that are farther in the future. As a result, it is difficult to make a general statement about whether time averaging leads to an increase or decrease in forecasting skill. The answer depends on the relative definitions of the time-averaging window and the lead time; see Fig. 3.13 for a schematic diagram.

Spatial averaging was seen to always improve the forecast skill. It eliminates

(“averages out”) high-frequency, nearly unpredictable components from the signal, and it leaves behind only the low-frequency, more-predictable components of the signal. Such behavior is consistent with a priori, basic intuition. No subtlety exists in the relative definition of lead time and a *spatial* averaging window.

What is the “correct” definition of lead time, relative to the time-averaging window? One might argue that the lead time should be defined with respect to the *end* of an averaging window, $[t - T_w, t]$. Such a definition seems sensible for a number of reasons. For instance, in terms of the initial data for the forecast, the time-averaged state over the past T_w time units would be known at the initial time. Also, this definition leads to enhanced prediction skill, which is in line with some basic intuition for the impact of averaging on prediction skill, and it seems desirable to have agreement with basic intuition. However, it is possible that the enhanced prediction skill is not entirely due to the averaging itself; instead, it is likely enhanced at least partially due to the fact that the time-averaging window $[t - T_w, t]$ contains signal data *before* the target time t , and this “earlier” data should naturally be more predictable than the “later” data at time t . Consequently, it is unclear whether there is a “best” definition of lead time, relative to the time-averaging window, since several competing factors are intermingled.

The models here were chosen to be exactly solvable and to be somewhat representative of atmospheric, oceanic, and climate variability. Therefore, the results here could potentially be expected to hold to some degree even for more complicated dynamical models. However, it is possible that nonlinearities [7, 8, 9, 10, 11, 30, 52, 61, 80] and other factors could potentially impact the conclusions, and it would be interesting to investigate similar questions with more complicated models in the future. As one small step in this direction, we have added advection to the stochastic heat equation by replacing $\partial_t q$ by $\partial_t q + c\partial_x q$ with an advection speed of $c \approx 15$ m/s; numerical results of forecasts with time averaging (not shown) are essentially the same as the main conclusions presented

above. It would be interesting to investigate other models in the future.

3.7 Appendices

3.7.1 Temporal and Spatial Averaging: Model Statistics

Before we evaluate the forecasting skill, we need to investigate a little bit in the change of variance of the signal through the temporal averaging and spatial averaging. Relatively small mean square errors are expected for forecasts with temporal or spatial averaging since averaging will create a smaller variance for the averaged signal. The trends of the variance changes as we do the averaging are shown in Fig. 3.14. For a complex OU process from (3.1), after approaching the stationary distribution arbitrary closely, the averaged mean will be nearly 0 while the variance of the averaged signal over a temporal window with length T_w is close to the stationary value of

$$\begin{aligned} \text{var}[\hat{u}(t, T_w)] &= \frac{\sigma^2}{2\gamma} \cdot \frac{1}{(\gamma^2 + \omega^2)^2 T_w^2} \cdot [2\gamma T_w(\gamma^2 + \omega^2) - 2(\gamma^2 - \omega^2) \\ &\quad + 2e^{-\gamma T_w}((\gamma^2 - \omega^2)\cos(\omega T_w) - 2\gamma\omega\sin(\omega T_w))], \end{aligned} \quad (3.52)$$

where $\sigma^2/2\gamma$ is the variance of the unaveraged cOU signal and this variance has properties

$$\lim_{T_w \rightarrow 0} \text{var}[\hat{u}(t, T_w)] = \frac{\sigma^2}{2\gamma} = \text{var}[u(t)], \quad \lim_{T_w \rightarrow \infty} \text{var}[\hat{u}(t, T_w)] = 0,$$

which agree with one's intuition that the signal will be averaged out if the temporal averaging window is very large. From the panel **a** of Fig. 3.14, we can see that the variance of the cOU signal is monotone decreasing if $\omega = 0$ as we increase the length of the temporal averaging window. When $\omega \neq 0$, the variance of the averaged cOU signal is decreasing at first if we enlarge the temporal averaging window and the signal is almost averaged out after a time period $2\pi/\omega$ according to that its variance is almost below 15%

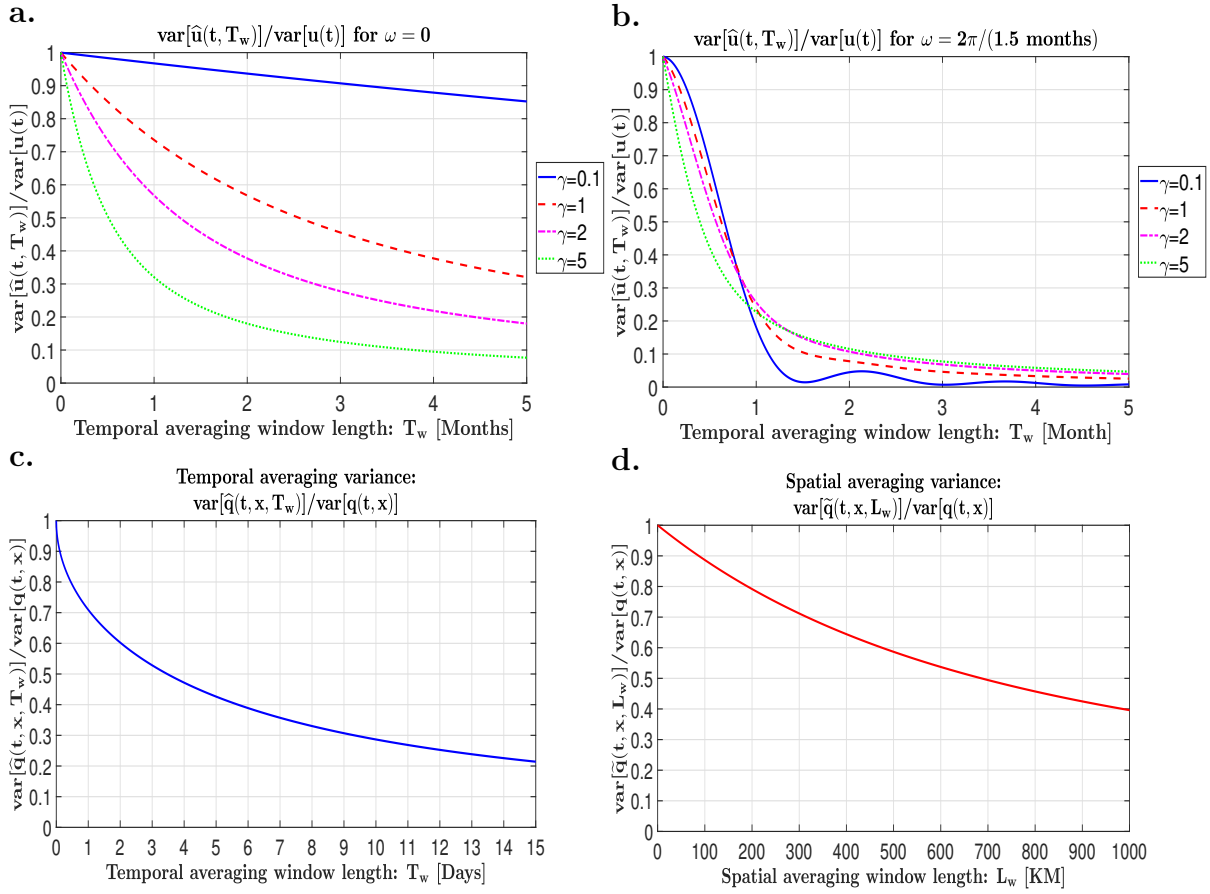


Figure 3.14: Panel **a.** variance change over temporal averaging for $\omega = 0$ and different γ of a cOU process; Panel **b.** variance change over temporal averaging for $\omega = 2\pi/(1.5 \text{ months})$ and different γ of a COU process ; Panel **c.** variance change over temporal averaging for a stochastic PDE process; Panel **d.** variance change over spatial averaging for a stochastic PDE process

of the original variance after a time period in panel **b** of Fig. 3.14. As the decorrelation time $1/\gamma$ decreases (γ increases), the declining rate of the variance is changing from a cosine type decreasing to an exponential decreasing. In the case $\omega = 0$, the variance decline is dominated by γ only. In the case $\omega \neq 0$, the variance always decreases to a very small amount after a time period regardless of the values of γ .

Regarding the stationary process stemming from the stochastic heat equation (3.8), both the impacts of the temporal and spatial averaging on the variance have to

be considered. The averaged signal always has the same mean $q^* + \tau F$. Through the temporal averaging with a temporal window of length T_w , the variance of the temporal averaged signal becomes

$$\begin{aligned} \text{var}[\hat{q}(t, x, T_w)] = & \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \left[1 - \text{erf}\left(\sqrt{\frac{T_w}{\tau}}\right) \left(\frac{3}{4} \left(\frac{\tau}{T_w}\right)^2 - \frac{\tau}{T_w} + 1 \right) \right. \\ & \left. + \exp\left(-\frac{T_w}{\tau}\right) \left(\frac{3}{2\sqrt{\pi}} \left(\frac{\tau}{T_w}\right)^{3/2} - \frac{1}{\sqrt{\pi}} \left(\frac{\tau}{T_w}\right)^{1/2} \right) \right]. \end{aligned} \quad (3.53)$$

Through the spatial averaging with a spatial window of length L_w , the variance of the spatial averaged signal will be

$$\text{var}[\tilde{q}(t, x, L_w)] = \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \cdot 2 \left[\frac{\sqrt{b_0 \tau}}{L_w} - \frac{b_0 \tau}{L_w^2} (1 - \exp(-\frac{L_w}{\sqrt{b_0 \tau}})) \right], \quad (3.54)$$

where $\frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}}$ is the variance of the unaveraged stochastic partial differential equation signal. The variance is decreasing monotonically as the averaging window increases in both the temporal averaging case and the spatial averaging case for the stochastic partial differential equation process by panel **c** and panel **d** of Fig. 3.14. By the analytic formulas (3.53) and (3.54), the variance of the temporal averaged SPDE signal compared to the original unaveraged variance is determined by the ratio of the temporal averaging window T_w and the relaxation time τ only and not related to other parameters in the model (3.8). The variance of the spatial averaged SPDE signal compared to the original unaveraged variance is determined by the ratio of square root of $b_0 \tau$ and the spatial averaging window L_w only.

3.7.2 Formulas of forecasting skill for a both temporal and spatial averaged SPDE

- Mean Square Error $\overline{MSE}_q(t_L, T_w, L_w)$

$$\begin{aligned}
& \overline{MSE}_q(t_L, T_w, L_w) \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \frac{1}{L_w^2 T_w^2} \frac{\tau^2}{2} \left[\frac{2T_w}{\tau} (\mathcal{P}(T_w/\tau, L_w) - \mathcal{P}(0, L_w)) \right. \\
&\quad + \frac{2t_L}{\tau} (\mathcal{P}((2t_L + T_w)/\tau, L_w) - \mathcal{P}(2t_L/\tau, L_w)) \\
&\quad - \frac{2(t_L + T_w)}{\tau} (\mathcal{P}(2(t_L + T_w)/\tau, L_w) - \mathcal{P}((2t_L + T_w)/\tau, L_w)) \\
&\quad - 2(\mathcal{Q}(T_w/\tau, L_w) - \mathcal{Q}(0, L_w)) \\
&\quad - (\mathcal{Q}((2t_L + T_w)/\tau, L_w) - \mathcal{Q}(2t_L/\tau, L_w)) \\
&\quad \left. + (\mathcal{Q}(2(t_L + T_w), L_w)) - \mathcal{Q}((2t_L + T_w)/\tau, L_w)) \right], \tag{3.55}
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{P}(x, y) = & \sqrt{\frac{1}{b_0 \tau}} [(h_2(x, y) + g_2(x, y)) - y(h_1(x, y) + g_1(x, y))] \\
& + (2x - 1)[(h_1(x, y) - g_1(x, y)) - y(h_0(x, y) - g_0(x, y))] \\
& - 4\sqrt{b_0 \tau} \cdot x \cdot \exp(-x) \left(y \cdot \operatorname{erf}\left(\frac{y}{2\sqrt{b_0 \tau x}}\right) - 2\sqrt{\frac{b_0 \tau x}{\pi}} (1 - \exp(-\frac{y^2}{4b_0 \tau x})) \right), \tag{3.56}
\end{aligned}$$

$$\begin{aligned}
\mathcal{Q}(x, y) = & -\frac{1}{4b_0\tau}[(h_3(x, y) - g_3(x, y)) - y(h_2(x, y) - g_2(x, y))] \\
& + \frac{3}{4}\frac{1}{\sqrt{b_0\tau}}[(h_2(x, y) + g_2(x, y)) - y(h_1(x, y) + g_1(x, y))] \\
& + (x^2 - \frac{3}{4})[(h_1(x, y) - g_1(x, y)) - y(h_0(x, y) - g_0(x, y))] \\
& - \sqrt{b_0\tau}(2x + 3)x \cdot \exp(-x) \cdot \\
& \left(y \cdot \operatorname{erf}\left(\frac{y}{2\sqrt{b_0\tau x}}\right) - 2\sqrt{\frac{b_0\tau x}{\pi}}(1 - \exp(-\frac{y^2}{4b_0\tau x})) \right), \tag{3.57}
\end{aligned}$$

$$h_0(x, y) = \int_0^y \exp\left(\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} + \sqrt{x}\right) dz, \tag{3.58}$$

$$h_1(x, y) = \int_0^y z \cdot \exp\left(\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} + \sqrt{x}\right) dz, \tag{3.59}$$

$$h_2(x, y) = \int_0^y z^2 \cdot \exp\left(\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} + \sqrt{x}\right) dz, \tag{3.60}$$

$$h_3(x, y) = \int_0^y z^3 \cdot \exp\left(\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} + \sqrt{x}\right) dz, \tag{3.61}$$

$$g_0(x, y) = \int_0^y \exp\left(-\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} - \sqrt{x}\right) dz, \tag{3.62}$$

$$g_1(x, y) = \int_0^y z \cdot \exp\left(-\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} - \sqrt{x}\right) dz, \tag{3.63}$$

$$g_2(x, y) = \int_0^y z^2 \cdot \exp\left(-\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} - \sqrt{x}\right) dz, \tag{3.64}$$

$$g_3(x, y) = \int_0^y z^3 \cdot \exp\left(-\frac{z}{\sqrt{b_0\tau}}\right) \operatorname{erf}\left(\frac{z}{2\sqrt{b_0\tau x}} - \sqrt{x}\right) dz. \tag{3.65}$$

- Correlation coefficient $\bar{\rho}_q(t_L, T_w, L_w)$

$$\bar{\rho}_q(t_L, T_w, L_w) = \left(\frac{\operatorname{var}[\bar{q}_{pred}(t, x, t_L, T_w, L_w; \theta)]}{\operatorname{var}[\bar{q}_{truth}(t, x, T_w, L_w; \theta)]} \right)^{1/2}, \tag{3.66}$$

$$\begin{aligned}
& \text{var}[\bar{q}_{truth}(t, x, T_w, L_w; \theta)] \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \frac{1}{L_w^2} \left\{ \frac{\tau^2}{T_w^2} \left[\frac{T_w}{\tau} (\mathcal{P}(T_w/\tau, L_w) - \mathcal{P}(0, L_w)) (\mathcal{Q}(T_w/\tau, L_w) - \mathcal{Q}(0, L_w)) \right] \right. \\
&\quad \left. + b_0 \tau \left(\exp\left(\frac{L_w}{\sqrt{b_0 \tau}}\right) + \exp\left(\frac{L_w}{\sqrt{b_0 \tau}}\right) - 2 \right) \right\},
\end{aligned} \tag{3.67}$$

$$\begin{aligned}
& \text{var}[\bar{q}_{pred}(t, x, t_L, T_w, L_w; \theta)] \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \frac{1}{L_w^2} \left\{ b_0 \tau \left[\exp\left(\frac{L_w}{\sqrt{b_0 \tau}}\right) + \exp\left(-\frac{L_w}{\sqrt{b_0 \tau}}\right) - 2 \right] \right. \\
&\quad + \frac{1}{T_w^2} \left[\int_{2t_L}^{2t_L+T_w} f_1(\alpha, L_w) d\alpha - \int_{2t_L+T_w}^{2(t_L+T_w)} f_1(\alpha, L_w) d\alpha \right. \\
&\quad \left. \left. - 2t_L \int_{2t_L}^{2t_L+T_w} f_0(\alpha, L_w) d\alpha + 2(t_L + T_w) \int_{2t_L+T_w}^{2(t_L+T_w)} f_0(\alpha, L_w) d\alpha \right] \right\},
\end{aligned} \tag{3.68}$$

where t_L is the lead time, T_w is the width of the temporal averaging window, L_w is the width of the spatial averaging window and

$$\begin{aligned}
f_0(\alpha, \beta) &= b_0 \tau \exp\left(-\frac{\beta}{\sqrt{b_0 \tau}}\right) \text{erf}\left(\frac{\beta}{2\sqrt{b_0 \alpha}} - \sqrt{\alpha/\tau}\right) \\
&\quad - b_0 \tau \exp\left(\frac{\beta}{\sqrt{b_0 \tau}}\right) \text{erf}\left(\frac{\beta}{2\sqrt{b_0 \alpha}} + \sqrt{\alpha/\tau}\right) \\
&\quad + 2b_0 \tau \text{erf}\left(\sqrt{\alpha/\tau}\right) + 2\sqrt{b_0 \tau} \beta \exp(-\alpha/\tau) \text{erf}\left(\frac{\beta}{2\sqrt{b_0 \alpha}}\right) \\
&\quad - 4b_0 \sqrt{\frac{\tau \alpha}{\pi}} \exp(-\alpha/\tau) \left(1 - \exp\left(-\frac{\beta^2}{4b_0 \alpha}\right)\right),
\end{aligned} \tag{3.69}$$

$$f_1(\alpha, \beta) = \alpha f_0(\alpha, \beta). \tag{3.70}$$

3.7.3 Calculations of some important statistics

By the change of the variable in (3.33), inserting into the SPDE (3.8) gives

$$\frac{\partial Q}{\partial t} = b_0 \triangle Q - \frac{1}{\tau} Q + D_* \dot{W}. \quad (3.71)$$

With one more step of changing the variable by $v(t, x) = e^{t/\tau} Q(t, x)$, the equation (3.71) is simplified to a simple and regular stochastic heat equation

$$\frac{\partial v}{\partial t} = b_0 \triangle v + D_* e^{t/\tau} \dot{W}. \quad (3.72)$$

v then can be solved directly as

$$v(t, x) = \frac{1}{(4\pi b_0 t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t}} v(0, y) dy + D_* \int_0^t \frac{1}{(4\pi b_0 |t-s|)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 (t-s)}} e^{s/\tau} \dot{W}(s, y) dy ds. \quad (3.73)$$

The important statistics of the $q(t, x)$ and $Q(t, x)$ can be calculated through this analytic solution (3.73) as the following.

- **Mean**

For a bounded initial value $|q(0, y)| \leq K$ (K is a finite constant)

$$\begin{aligned} |\mathbb{E}[q(t, x)] - (q^* + \tau F)| &= |\mathbb{E}[Q(t, x)]| \\ &= |\mathbb{E}[e^{-t/\tau} v(t, x)]| \\ &= \left| e^{-t/\tau} \frac{1}{(4\pi b_0 t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t}} v(0, y) dy \right| \\ &= \left| e^{-t/\tau} \frac{1}{(4\pi b_0 t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t}} (q(0, y) - q^* - \tau F) dy \right| \\ &\leq e^{-t/\tau} \frac{1}{(4\pi b_0 t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4b_0 t}} (K + q^* + \tau F) dy \\ &= (K + q^* + \tau F) e^{-t/\tau} \rightarrow 0 \quad \text{as } t \rightarrow \infty \end{aligned}$$

$$\Rightarrow \lim_{t \rightarrow \infty} \mathbb{E}[q(t, x)] = q^* + \tau F. \quad (3.74)$$

• **Covariance and variance**

$$\text{cov}[q(s, x), q(t, y)] \quad (3.75)$$

$$\begin{aligned} &= \text{cov}[Q(s, x), Q(t, y)] \\ &= \text{cov}[e^{-s/\tau} v(s, x), e^{-t/\tau} v(t, y)] \\ &= e^{-(s+t)/\tau} \text{cov}[v(s, 0), v(t, x - y)] \\ &= e^{-(s+t)/\tau} \frac{D_*^2}{4\pi b_0} \cdot \\ &\quad \mathbb{E} \left[\int_0^t \int_0^s \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{|t-r|^{1/2} |s-r'|^{1/2}} e^{-\frac{|x-y-y'|^2}{4b_0(t-r)} - \frac{|y''|^2}{4b_0(s-r')}} \right. \\ &\quad \cdot \dot{W}(r, y') \dot{W}(r', y'') e^{r/\tau} e^{r'/\tau} dy' dy'' dr' dr \left. \right] \\ &= e^{-(s+t)/\tau} \frac{D_*^2}{4\pi b_0} \int_0^{s \wedge t} \int_{\mathbb{R}} \frac{1}{|t-r|^{1/2} |s-r|^{1/2}} e^{-\frac{|x-y-y'|^2}{4b_0(t-r)} - \frac{|y'|^2}{4b_0(s-r)}} e^{2r/\tau} dy' dr \\ &= e^{-(s+t)/\tau} \frac{D_*^2}{(4\pi b_0)^{1/2}} \int_0^{s \wedge t} (s+t-2r)^{-1/2} \exp\left(-\frac{|x-y|^2}{4b_0(s+t-2r)}\right) e^{2r/\tau} dr \\ &= \frac{D_*^2}{4\sqrt{\pi b_0}} \int_{|s-t|}^{s+t} l^{-1/2} \exp\left(-\frac{|x-y|^2}{4b_0 l} - al\right) dl. \end{aligned} \quad (3.77)$$

For the temporal covariance at a single location, take $x = y$ in (3.77),

$$\begin{aligned} \text{cov}[q(s, x), q(t, x)] &= \frac{D_*^2}{4\sqrt{\pi b_0}} \int_{|s-t|}^{s+t} l^{-1/2} \exp(-l/\tau) dl \\ (\text{let } \eta = (l/\tau)^{1/2}) &= \frac{D_*^2}{4\sqrt{\pi b_0}} \int_{\sqrt{|s-t|/\tau}}^{\sqrt{(s+t)/\tau}} 2\sqrt{\tau} \exp(-\eta^2) d\eta \\ &= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \cdot \frac{2}{\sqrt{\pi}} \int_{\sqrt{|s-t|/\tau}}^{\sqrt{(s+t)/\tau}} \exp(-\eta^2) d\eta \\ &= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} (\text{erf}(\sqrt{(s+t)/\tau}) - \text{erf}(\sqrt{|s-t|/\tau})). \end{aligned} \quad (3.78)$$

For the asymptotic spatial covariance at a single time point, take $s = t$ in (3.77),

$$\begin{aligned}
\lim_{t \rightarrow \infty} \text{cov}[q(t, x), q(t, y)] &= \frac{D_*^2}{4\sqrt{\pi b_0}} \int_0^\infty l^{-1/2} \exp\left(-\frac{|x-y|^2}{4b_0 l} - l/\tau\right) dl \\
(\text{Let } \eta = \sqrt{l}) &= \frac{D_*^2}{2\sqrt{\pi b_0}} \exp\left(-\frac{1}{\sqrt{b_0 \tau}}|x-y|\right) \int_0^\infty \exp\left(-\left(\frac{|x-y|}{2\sqrt{b_0 \eta}} - \frac{\eta}{\sqrt{\tau}}\right)^2\right) d\eta \\
&= \frac{D_*^2}{2\sqrt{\pi b_0}} \exp\left(-\sqrt{\frac{1}{b_0 \tau}}|x-y|\right) \cdot \frac{\sqrt{\pi \tau}}{2} \\
&= \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}} \exp\left(-\frac{1}{\sqrt{b_0 \tau}}|x-y|\right). \tag{3.79}
\end{aligned}$$

For the stationary variance for $q(t, x)$, take $x = y$ in (3.79),

$$\lim_{t \rightarrow \infty} \text{var}[q(t, x)] = \frac{D_*^2}{4} \sqrt{\frac{\tau}{b_0}}. \tag{3.80}$$

Chapter 4

Is weather forecast skill improved by time averaging?

4.1 Introduction

It is natural to expect more accurate forecasts of time-averaged quantities. For instance, one would expect that the rainfall averaged over one week can be forecast with greater skill than the rainfall averaged over one day. Likewise, one would expect that the rainfall averaged over a large region (such as the area of a state or province) could be predicted with greater skill than the rainfall over a smaller region (such as a city or town).

Is this conventional wisdom true? If so, how much additional forecast skill is gained by spatial and/or temporal averaging?

Some past work has examined forecast skills for different spatial or temporal averaging [67, 92], although the studies were not set up in a way that allows direct assessment of the conventional wisdom. To most clearly assess the conventional wisdom, it is necessary to compare forecasts in a way that holds all parameters fixed (lead time, etc.) except for changes in the averaging window. Such an assessment is the goal of the present chapter.

In the previous chapter, the conventional wisdom was assessed, but on simple

dynamical systems rather than operational weather forecasts. For simple (stochastic) dynamical systems, it was seen that time averaging was ineffective at improving forecast skill, while spatial averaging did improve forecast skill.

Will the weather system show a similar behavior? Is time averaging ineffective at improving weather forecasts? Or, on the other hand, does conventional wisdom hold for weather forecasts, due to some fundamental difference between weather and simple stochastic dynamical systems?

The rest of this chapter is organized as follows. Methods and data are described in section 4.2. The impact of averaging on forecast skill is assessed in section 4.3, and in section 4.4 a theoretical explanation is provided for the limited effect of time averaging on forecast skill. Section 4.5 is a concluding section. Additional sensitivity studies are presented in section 4.6. Theoretical calculations are included in appendix section 4.7.

4.2 Methods

4.2.1 Data

The operational weather forecast data used here is from the Global Forecast System (GFS) [18] and for two variables: precipitation and surface temperature. The 3 hourly and 1° gridded global GFS dataset is used here for allowing assessment of longer available lead times. The open source data can be downloaded at <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>. GFS analysis data at lead time of 0 days (also called the day-0 analysis) is used to serve as the truth for both precipitation rate and surface temperature for assessing the forecast skill.

For some sensitivity tests, Global Precipitation Measurement (GPM) [23] data are also used as a baseline truth, for comparison with the use of GFS day-0 analysis data

as truth for precipitation. The 30-min research/final run version of GPM data with spatial resolution 0.1° is used here for the variable “precipitationCal” (Multi-satellite precipitation estimate with gauge calibration). The GPM data are fully available from 60°N - 60°S while much data is missing outside this range. Hence, only GPM data within 60°N - 60°S are used in the analysis here to ensure stable and reliable results. GPM data can be downloaded from <https://pmm.nasa.gov/data-access/downloads/gpm>.

As another type of sensitivity test, data has also been used for anomalies from the seasonal cycle. To create the anomaly data, a smoothed seasonal cycle has been subtracted from the GFS forecast data before moving forward with further analysis. For the seasonal cycle of precipitation rate, the enhanced monthly long term mean CPC Merged Analysis of Precipitation (CMAP) [12] with 2.5° latitude x 2.5° longitude global grid derived from years 1981 to 2010 is used (and it can be downloaded from <https://psl.noaa.gov/data/gridded/data.cmap.html>). For creating the seasonal cycle for surface temperature, GFS analysis data with lead 0 from January 1, 2005, to December 31, 2017 are used. In the preprocessing, the 6 hourly smoothed seasonal cycle is calculated via the annual mean and the first three harmonics at every location globally through discrete Fourier transform and inverse discrete Fourier transform. Particularly, a bilinear interpolation follows for the smoothed CMAP precipitation seasonal cycle to interpolate the data into 1° grids to match the grids of the GFS dataset. After removing the smoothed seasonal cycle from the GFS forecast data, the remaining data represent anomalies from the seasonal cycle.

Four main assessments are conducted here, and they cover a variety of quantities of interest (precipitation vs. surface temperature), datasets, time ranges (in line with the availability of the datasets), etc.: (1) 3.5-months GFS precipitation forecast data from June 17, 2019 to October 08, 2019 are compared with GFS day-0 analysis data for precipitation. (2) 11-months GFS surface temperature forecast data from March 01, 2018

to January 28, 2019 are compared to GFS day-0 analysis data for surface temperature. (3) 11-months GFS precipitation forecast data from November 01, 2018 to August 31, 2019 are compared with GPM dataset. (4) The assessments 1 and 2 were also repeated with GFS precipitation/temperature data as anomalies from the seasonal cycle. GFS forecast data for precipitation rate and surface temperature (Assessments 1 and 2) are mainly used in this main manuscript. Analysis involving GPM dataset and the same analysis on the anomaly data of GFS (Assessments 3 and 4) are also investigated as sensitivity tests. More details and results can be found in the next section.

4.2.2 Data Analysis Setup

To calculate the forecast skills, 7 days lead time is used for the surface temperature data. Since precipitation usually is less predictable than the surface temperature, 3 days lead time is used for precipitation rate forecasts. Time averaging window ranges from 1 day to 6 days with an increment of 0.5 day, and the diameter of a spatial averaging window ranges from 100 km, 450 km to 4500 km with an increment of 450km.

We start with 1-day time averaging and 100-km space averaging. Then, for varying the time averaging window, the 100-km space averaging diameter is held fixed, and the time averaging window is increased. Similarly, for varying the spatial averaging diameter, the 1-day time averaging window is held fixed, and the spatial averaging diameter is increased.

To achieve the comparability between the forecast dataset and the truth data and a well-defined setup for the time averaging window, some data processing is implemented before calculating the forecast skill with different averaging windows.

First, for comparisons involving spatial averaging windows larger than or equal to 900km, 30 min GPM data are averaged at each latitude and longitude to be coarsened

from 0.1° to 1° at the very beginning. The GFS forecast/analysis data and GPM data are then averaged over a circular region with diameter of 450km (or 900 km, ... , 4500 km) followed by the comparison.

Second, for comparisons involving time averaging, the baseline is with spatial averaging with a diameter of 100km. For high resolution GPM data with 0.1° , the data is directly averaged over a circular region with diameter of 100km to create a preprocessed GPM data with 100km spatial averaging with the guarantee of enough sample points within 100km for each integer latitude and longitude. For GFS data with 1° , an underlying problem arising for those tropical locations is that they don't have enough sample points within 100km spatial averaging due to the coarse grid. To achieve a better estimate, a bilinear interpolation has been applied to the GFS dataset to create new data with 0.2° grids, which is then spatial averaged over a circular region with diameter of 100km as a new preprocessed GFS dataset. Using the preprocessed datasets, we can move forward to do time averaging with 1 day, 1.5 days, ..., 6 days.

4.2.3 Evaluating Forecast Skills

The Pearson correlation coefficient ρ is used here as the main metric for evaluating the forecast skill. The coefficient ρ can be expressed as

$$\rho(X(t), Y(t)) = \frac{\text{cov}(X(t), Y(t))}{\sigma_X \sigma_Y}, \quad (4.1)$$

where $\text{cov}(X(t), Y(t)) = \mathbb{E}[(X(t) - \mathbb{E}[X(t)])(Y(t) - \mathbb{E}[Y(t)])]$ is the covariance between $X(t), Y(t)$ and σ_X, σ_Y are the standard deviations of X, Y respectively. In practice, the expected value $\mathbb{E}$ is calculated as a time average, so $\rho(X(t), Y(t))$ is a constant that doesn't change over time. In particular, in a situation of analyzing the real data, we have the true data from the true signal as $X_1, X_2, \dots, X_N$ and our corresponding predictions $Y_1, Y_2, \dots, Y_N$, each given at N points in time. The correlation coefficient is calculated

by

$$\rho = \frac{\frac{1}{N} \sum_{i=1}^N (X_i - \frac{1}{N} \sum_i X_i)(Y_i - \frac{1}{N} \sum_i Y_i)}{(\frac{1}{N} \sum_{i=1}^N (X_i - \frac{1}{N} \sum_i X_i)^2)^{1/2} (\frac{1}{N} \sum_{i=1}^N (Y_i - \frac{1}{N} \sum_i Y_i)^2)^{1/2}} \quad (4.2)$$

For time averaging, a centered averaging window is used in this chapter. Namely, as Fig. 4.1 shows, for a time averaging window having length T_w and centered at time t , the averaged signal at time t is defined as the average of the signals over the window $[t - T_w/2, t + T_w/2]$. If t_0 is the current time and the lead time is t_L , the averaged forecast with the lead time t_L is the averaged forecast over $[t - T_w/2, t + T_w/2]$ where $t = t_0 + t_L$. The true signal will be the averaged true signals averaged over $[t - T_w/2, t + T_w/2]$. Then the averaged forecast and the averaged true signal at time t will be compared for further forecast skill evaluation.

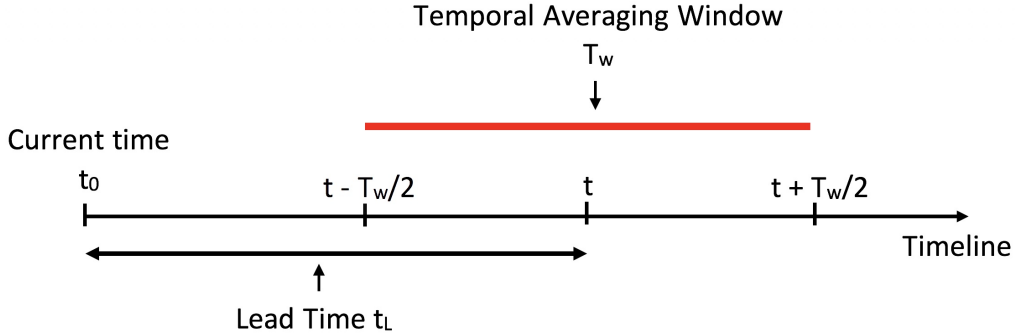


Figure 4.1: Schematic of lead time and temporal averaging window definition in the analysis. The horizontal axis represents forecast time from the initial condition at t_0 .

Another measure used here is the ratio r , which measures the impact on forecast skill brought by space averaging versus by time averaging. Some details of its definition are as follows. Recall that the baseline case is the forecast skill when one day and 100km averaging are applied. Denote $\rho(L_w, T_w)$ as the correlation coefficient for the forecast skill with spatial averaging diameter L_w and time averaging window duration T_w . Then

for every location on the earth, r is defined as

$$r = \frac{\rho(900 \text{ km}, 1 \text{ day}) - \rho(100 \text{ km}, 1 \text{ day})}{\rho(100 \text{ km}, 2 \text{ day}) - \rho(100 \text{ km}, 1 \text{ day})} \quad (4.3)$$

For creating a global map of the ratio r , some manual settings are made to better present the comparison between time averaging and spatial averaging. Areas are marked with white if a negative ratio r occurs; such a situation corresponds to a forecast skill increase by spatial averaging but the forecast skill decrease by time averaging. In the opposite scenario, when a forecast increase is seen for time averaging and a forecast skill decrease for spatial averaging, the ratio r is manually set to be 0.001 for visualizations with blue colors to represent better performance in forecast skills through time averaging. Grey areas are locations where either missing values are present or where the forecast skill decreases for both spatial and time averaging, since this latter scenario should be distinguished from positive r values due to forecast skill increases for both spatial and time averaging. To summarize, in the global map presented in this chapter, only blue color represents a better forecast skill for the time averaging, grey color represents locations where a comparison is invalid due to missing values or both forecast skills decrease after time averaging or spatial averaging, and all the other colors (yellow, forest green, medium sea green, white colors) are locations where spatial averaging has better performance in increasing the forecast skill compared to the time averaging.

4.2.4 Computational Methods

A substantial computational cost is required in order to calculate forecast skills with various spatial and time averaging windows and at all global locations. For example, consider the assessment of the forecast skill for GFS temperature data with different time averaging windows (1 day to 14 days) and different spatial averaging, using as input a global dataset with approximately one year of data at 6 hourly time intervals.

Furthermore, the GFS temperature forecast dataset includes lead times from lead 0 to lead 16 days. To analyze such a dataset, memory about 70-75 GB is needed for each latitude to keep the running time of the calculation within about 8 hours for one latitude. To manage the calculations, high-throughput computing (HTC) is used to evaluate forecast skills at all latitudes simultaneously, to reduce the overall calculation time.

4.3 Impact of averaging on forecast skill

Weather forecast skill is analyzed for the Global Forecast System (GFS). GFS is the operational numerical weather forecast model of the US National Centers for Environmental Prediction (NCEP). To analyze the impact of time averaging, the forecast data is averaged over a particular window duration, such as 1 day, for assessing the forecast skill. Additional durations are also investigated, such as 2 days, 3 days, etc., and the skills for different durations are compared. Similarly, for analyzing the impact of spatial averaging, the forecast data is averaged over a circular region with diameter of 100 km, and, for comparison, additional diameters are also considered, such as 900 km, 1800 km, 2700 km, etc.

Fig. 4.2 shows the forecast skill for different time and space averaging windows. The forecast skill is reported as a correlation coefficient, although it is an average of correlation coefficients. Further details are described in the Methods section. In brief, the skill (correlation coefficient) was first calculated for each location on the Earth, and then the skill was averaged over all locations globally, leading to the globally averaged forecast skill that is plotted in Fig. 4.2. Results are shown for both precipitation (panel a) and surface temperature (panel b).

Fig. 4.2 demonstrates one important basic result: spatial averaging is much more

effective than time averaging at improving forecast skill at the global level.

Specifically, as the averaging window increases in size, the forecast skill with spatial averaging (blue squares) is increasing much faster than the forecast skill with time averaging (red circles).

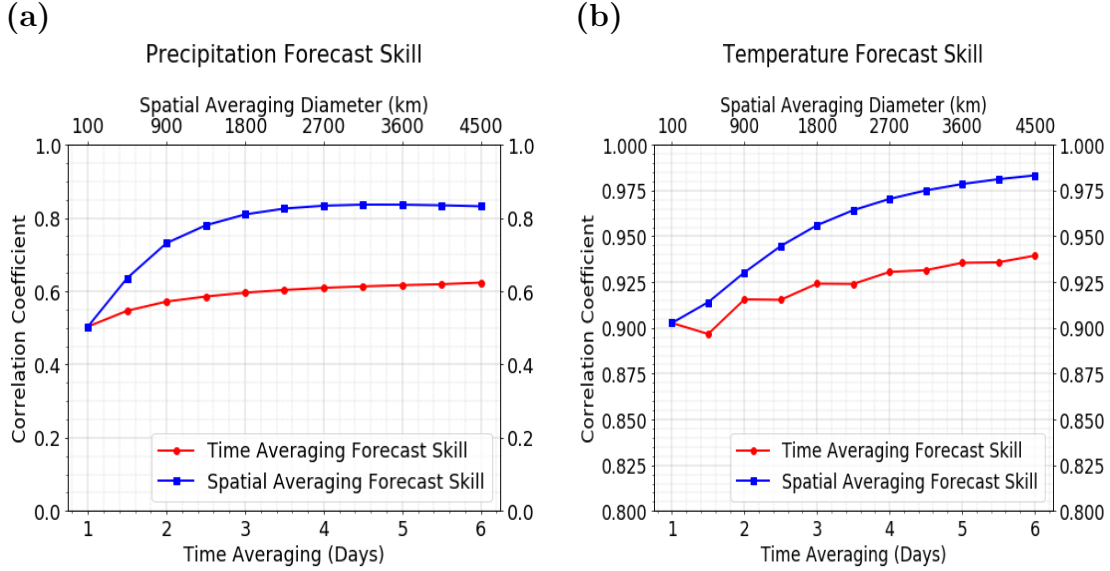


Figure 4.2: Effect of time and space averaging on forecast skill, averaged globally. (a) Precipitation, at a lead time of 3 days. (b) Surface temperature, at a lead time of 7 days. Bottom x-axis: Time averaging window length; Left y-axis: Forecast skill (correlation coefficient) when time averaging is applied; Top y-axis: Spatial averaging window length; Right y-axis: Forecast skill (correlation coefficient) when spatial averaging is applied.

To take a closer look, beyond a global average, Fig. 4.3 shows a metric of forecast skill increase and displays it at each location around the globe. The metric is a ratio r that is defined as

$$r = \frac{\Delta_x \rho}{\Delta_t \rho} \quad (4.4)$$

where ρ is the forecast skill (correlation coefficient), and $\Delta_t \rho$ is the change in forecast skill due to an increase in time-averaging window duration (from 1 day to 2 days), and $\Delta_x \rho$

is the change in forecast skill due to an increase in spatial-averaging window diameter (from 100 km to 900 km). Hence, if $r > 1$, then spatial averaging is more effective than time averaging at improving forecast skill, whereas if $0 < r < 1$, then time averaging is more effective than spatial averaging at improving forecast skill. Fig. 4.2 shows that, at nearly all locations worldwide, $r > 1$ and therefore spatial averaging is more effective than time averaging at increasing forecast skill.

Some additional interesting features also can be seen in Fig. 4.3. For instance, for precipitation, in Fig. 4.3a, over a large fraction of the global locations, spatial averaging is three times more effective ($r > 3$, yellow areas) than time averaging at improving the forecast skill. It is only in very few small regions that time averaging performs more efficiently (blue areas). For surface temperature, in Fig. 4.3b, an interesting characteristic is an apparent land–sea contrast: time averaging is least effective compared to spatial averaging ($r > 3$, yellow color) over ocean regions, less effective ($1 < r < 3$, green colors) over land regions, and more effective ($0 < r < 1$) over coastal regions. These geographic features are so apparent in the spatial map of r values that the outlines of the continents arise naturally in Fig. 4.3b. Time averaging is more effective than spatial averaging at only a very small number of locations, along the coastlines of the continents; this is possibly due to the unique conditions associated with coastal microclimates.

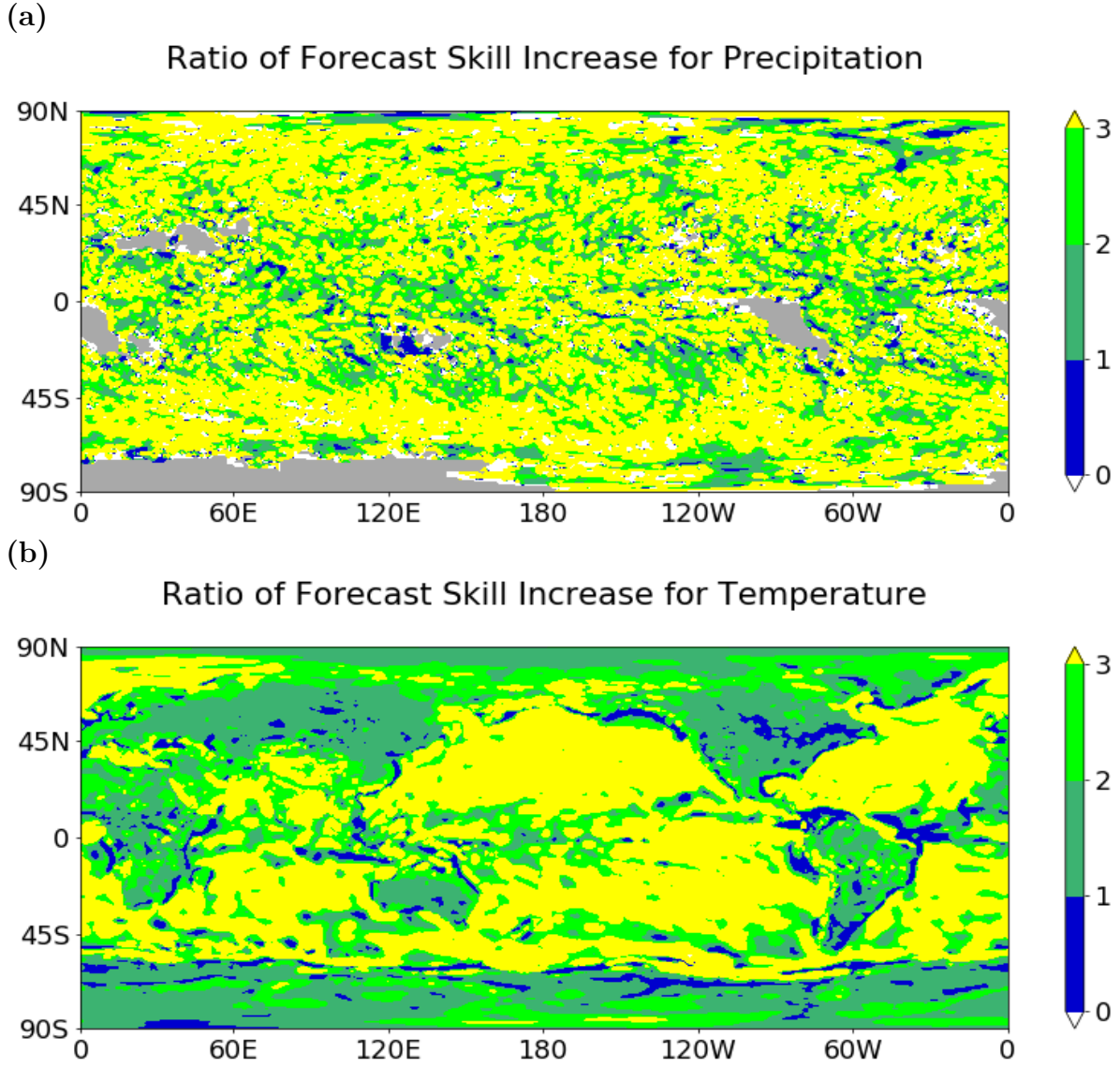


Figure 4.3: Global map of the ratio of forecast skill increase, defined in (4.4) as $r = \Delta_x \rho / \Delta_t \rho$, where $\Delta_x \rho$ is the change in forecast skill due to increased spatial averaging, and $\Delta_t \rho$ is the change in forecast skill for increased time averaging. **(a)** Precipitation. **(b)** Surface temperature. Detailed explanations for the white and grey colors can be found in the Methods section.

4.4 Theoretical explanation

Why does time averaging have limited impact on forecast skill? Conventional wisdom would say that time averaging should provide a smoother time series, and a smoother time series should be more predictable.

Here we show that, in fact, if a time series is smoother, it is not necessarily much more predictable. A theoretical illustration of this can be seen using an auto-regressive (AR) model (also called red noise, correlated noise, colored noise, or the Ornstein–Uhlenbeck process in the continuous-time version). The AR(1) model takes the form

$$u^{n+1} = u^n + (\Delta t/T_d)u^n + \xi^n, \quad (4.5)$$

where u^n is the quantity of interest (e.g., precipitation or surface temperature) at time step n , Δt is the time step, T_d is the decorrelation time, and ξ^n is a white noise process. Such a model is often used to represent atmospheric and oceanic phenomena [13, 20, 40]. An example time series is shown in Fig. 4.4a. The example was created to be similar to a surface temperature signal, and it uses a decorrelation time of $T_d = 3.8$ days. Also shown in Fig. 4.4a is a smoothed version of the AR(1) time series which was created by time averaging over a window of 2 days. The time-averaged signal is much smoother than the original signal, and one might therefore expect the time-averaged signal to be much more predictable.

In Fig. 4.4b, an experiment is implemented for testing the impacts of time averaging on the simulated time series from the AR process using 3 days as the lead time. One can see that, as the time averaging window increases, the forecast skill has only slight improvements, similar to what had been seen in the operational weather forecast data in Fig. 4.2.

Analytical formulas can be derived to better understand the effects of time averaging on the AR model. As one quantity of interest, the decorrelation time provides a measure of predictability, and its value $\bar{T}_d$ for the time-averaged process can be calculated by

$$\bar{T}_d = \int_0^\infty ACF(\tau) d\tau \quad (4.6)$$

where $ACF(\tau)$ is the autocorrelation function with time lag τ . For the signal $u(t)$ from the original AR(1) model (in its continuous-time form as the Ornstein–Uhlenbeck process), the autocorrelation function is $ACF(\tau) = \exp(-\tau/T_d)$, and the decorrelation time is T_d . For a time-averaged signal ($\bar{u}(t) = \frac{1}{T_w} \int_{t-T_w/2}^{t+T_w/2} u(s) ds$) averaged over a time window with length T_w , we find the decorrelation time to be

$$\bar{T}_d = T_d \cdot \frac{(T_w/T_d)^2}{2[(T_w/T_d) - 1 + e^{-T_w/T_d}]} \approx T_d + \frac{1}{3}T_w \quad (4.7)$$

Details of the calculation of $\bar{T}_d$ are shown in Sec. 4.7.

The analytical formula in (4.7) provides insight into how the time averaging window length, T_w , affects the decorrelation time, $\bar{T}_d$. When time averaging is performed on the signal, how does the decorrelation time $\bar{T}_d$ of the smoothed signal compare with the original signal's decorrelation time, T_d ? As seen in (4.7), $\bar{T}_d$ is proportional to T_d , and the proportionality factor depends on the key quantity T_w/T_d . If T_w/T_d is small—i.e., if the time-averaging window length is smaller than the original signal's decorrelation time—then time averaging has only a small impact on the decorrelation time and forecast skill. For example, in Fig. 4.4, where $T_d = 3.8$ days, a time averaging window of $T_w = 2$ days leads to a $\bar{T}_d$ that is only 18% larger than T_d . This is a somewhat small change given what might be expected from how much smoother the time-averaged signal looks in Fig. 4.4a. Even for larger values of T_w , which would create even smoother time series, the corresponding increase in $\bar{T}_d$ is somewhat small compared with what one might expect. For example, when $T_w = T_d$, so that the time-averaged signal is greatly

smoothed, the corresponding decorrelation time is $\bar{T}_d \approx 1.36T_d$, for a 36% increase even in this largely smoothed case. Both of these examples are in line with the approximation, shown in (4.7), of $\bar{T}_d \approx T_d + (1/3)T_w$, which suggests that the additional decorrelation time brought about by time-averaging is only one-third as large as the time averaging window duration T_w . These slight changes in decorrelation time will then translate into only slight changes in forecast skill, as can be seen in either analytical formulas [39] or in the plot in Fig. 4.4b.

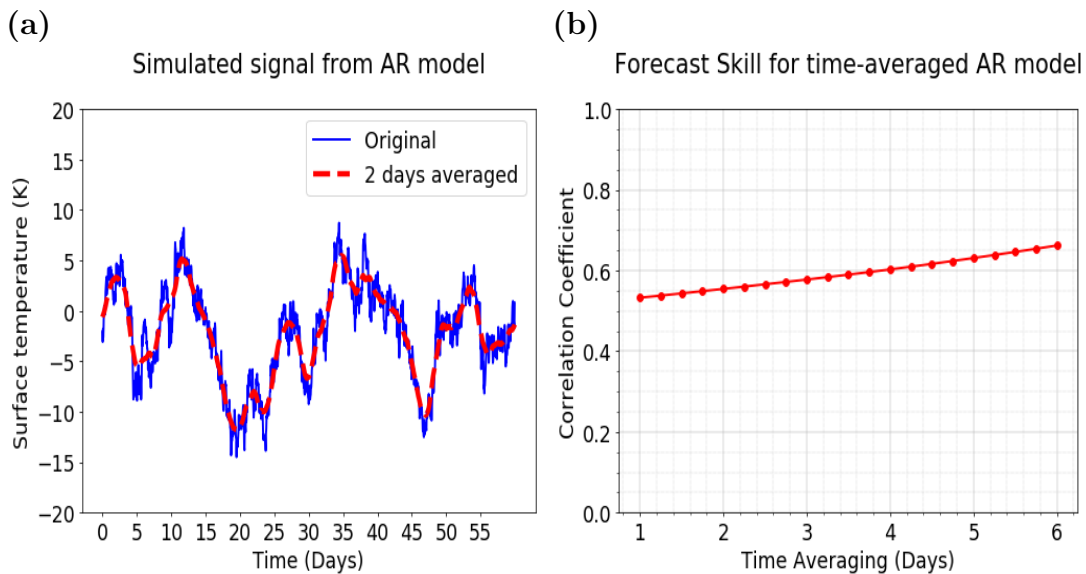


Figure 4.4: Effects of time averaging on an AR(1) model. **(a)** An example time series of the simulated AR(1) process (blue solid line) and a smoothed version of it (red dashed line) obtained by time averaging with a 2-day window. **(b)** Forecast skills with different time averaging windows for a fixed lead time of 3 days for the AR(1) model.

4.5 Conclusions

An investigation was conducted to assess the conventional wisdom that time averaging and space averaging will provide increases in forecast skill. The results show spatial averaging can enhance skill more than time averaging. Consistent evidence is seen in

forecast data from both operational weather forecasts and simple stochastic models (either time series models or spatiotemporal stochastic models) that can be studied analytically. Taken together, this is evidence from a variety of models that time averaging may create a smoother time series that is not necessarily much more predictable.

One could speculate on the possible origin of the conventional wisdom (that time averaging should improve forecast skill significantly). It is possible that the conventional wisdom arose in an early era of forecasting, prior to the advent of weather forecasting with ensembles. It was only in the 1990s that operational weather forecasts began to use ensembles [2,4,38]. If only a single ensemble member is used, then one would indeed expect improvement in a weather forecast from time averaging. If an ensemble is used for the forecast, then an average over the ensemble members will likely create a similar type of benefit as a time average of an individual ensemble member's forecast.

It is also possible that the effects of time averaging depend on the time scales of the averaging. Here, lead times of 3 and 7 days were investigated, along with time averaging windows of duration up to 6 days. It would be interesting to investigate other time scales, such as subseasonal to seasonal predictions or short-term climate predictions [54,69,81]. Nevertheless, one might expect the principles of the theoretical AR model (Fig. 4.4) to hold no matter the time scales.

4.6 Appendix 1: Additional Sensitivity Tests

In this section, to test the robustness of the results in the main manuscript, additional sensitivity tests are implemented. One test involves using anomaly data (using anomalies from the seasonal cycle), and another test involves using another dataset, the GPM (Global Precipitation Measurement) dataset, as the truth signal for precipitation.

4.6.1 Analysis on anomaly data

Forecast skill changes are now investigated in the GFS precipitation rate/surface temperature anomaly data. The anomaly data is defined by subtracting a smoothed seasonal cycle before assessing the forecast skills.

For the precipitation rate, the enhanced monthly long term mean CPC Merged Analysis of Precipitation (CMAP) is used for calculating the seasonal cycle. For the surface temperature, GFS analysis data with lead 0 are used. More details about the seasonal cycles are explained in the Method section.

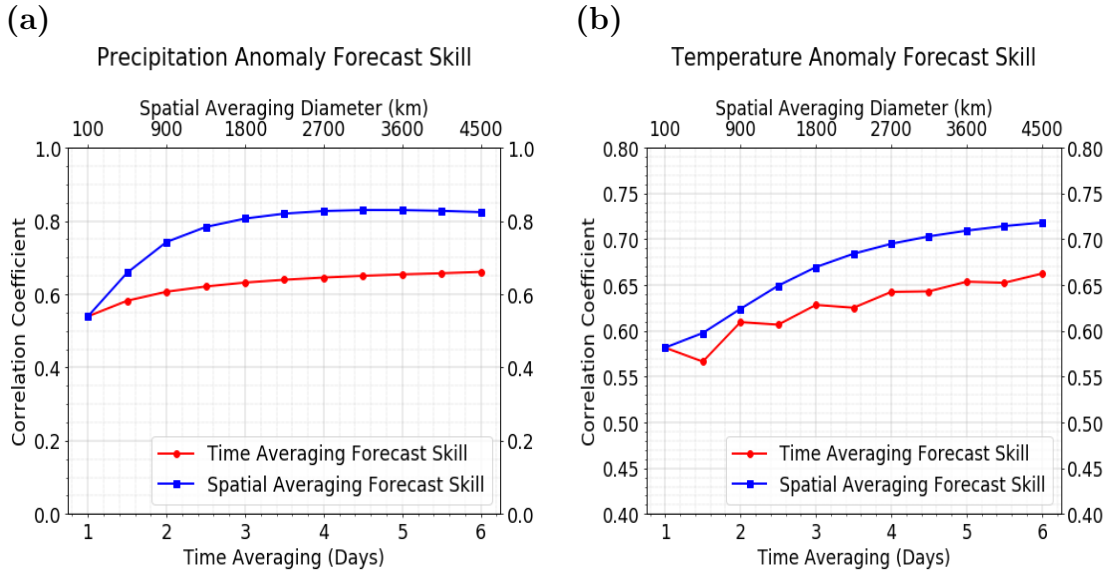


Figure 4.5: Effect of time and space averaging on forecast skill, averaged globally. **(a)** Precipitation anomaly, at a lead time of 3 days. **(b)** Surface temperature anomaly, at a lead time of 7 days.

The main results are shown in Fig. 4.5. In the comparison between forecast performance changes via time averaging vs. spatial averaging, essentially the same result is obtained: spatial averaging is more efficient than time averaging in improving the forecast skills, both for the precipitation and the temperature. One difference seen in Fig. 4.5b is that the temperature forecast skill is lower for the anomalies from the seasonal

cycle, in comparison to the case of the main text, where the temperature forecast skills were higher. This difference indicates that variations from the seasonal cycle contribute a substantial amount of the forecast skill. Nevertheless, as seen in this additional test, the effects of time averaging vs. space averaging are essentially the same whether or not the seasonal cycle is retained or removed.

4.6.2 Assessing tropical precipitation forecast skills using GPM data

In the main text, the truth signal was taken to be the day-0 analysis from the GFS model. Here, as an additional test, an observational data product is used as the truth signal instead of the model data product. In particular, GPM data is used to serve as the true precipitation data for testing the robustness as well. We compare the GFS precipitation forecast data, at a lead time of 3 days, to the GPM data to assess the forecast skill.

Due to the GPM only having full observations from $60^{\circ}S$ to $60^{\circ}N$ while a maximum of 4500 km spatial averaging diameter is required in our analysis, the results of this additional test are limited to tropical locations from $12^{\circ}S$ to $12^{\circ}N$, so that the entire spatial averaging disc is contained in the GPM data coverage from $60^{\circ}S$ to $60^{\circ}N$.

As seen in Fig. 4.6, it is again seen that spatial averaging outperforms time averaging in increasing forecast skill. One difference here is that, for short spatial/time averaging, the tropical averaged precipitation rate forecast skills are lower than the global averaged ones shown in the main text. This is within expectation based on earlier work since tropical precipitation has less predictability than extratropical precipitation. This additional test can also be taken as a special case of analyzing one region (i.e., the tropics), and again the main conclusion is robust: spatial averaging improves forecast skills more than time averaging.

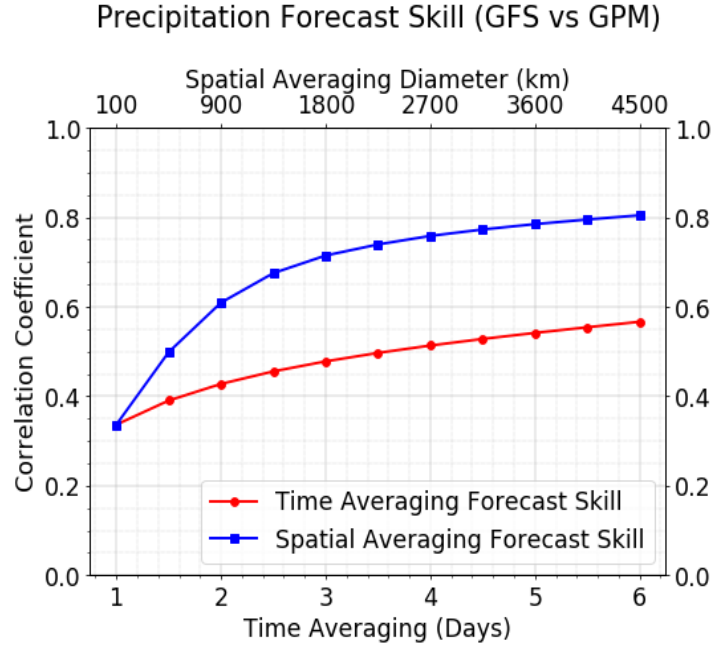


Figure 4.6: Effect of time and space averaging on forecast skill, averaged over tropical areas from 12°S to 12°N for comparing GFS precipitation rate forecast with GPM data, at a lead time of 3 days

4.7 Appendix 2: Analytic Formulas for Decorrelation Time

In this section, we present the derivation of the analytic formula for the decorrelation time, including the decorrelation time of the time-averaged process. In the calculation, we use a continuous-time version of the AR(1) model called the Ornstein–Uhlenbeck (OU) process.

The Ornstein-Uhlenbeck process is defined as

$$du(t) = -u(t)/T_d + \sigma dW(t) \quad (4.8)$$

with positive parameters $T_d, \sigma > 0$ and $dW(t)$ is related to a Gaussian white noise $\dot{W}(t)$

via

$$dW(t) = \dot{W}(t)dt, \quad (4.9)$$

that is, white noise is a “derivative” of the Wiener process $W(t)$ and it satisfies the following properties

$$\mathbb{E}[\dot{W}(t)] = 0, \quad (4.10)$$

$$\mathbb{E}[\dot{W}(t)\dot{W}(s)] = \delta(t-s). \quad (4.11)$$

The exact solution of (4.7) is

$$u(t) = e^{-t/T_d}u(0) + \sigma \int_0^t e^{-(t-s)/T_d} dW(s) \quad (4.12)$$

As $t \rightarrow \infty$, or as the initial time tends to $-\infty$, $u(t)$ will converge to a stationary Gaussian distribution with mean 0 and variance $\sigma^2 T_d/2$ [19, 50]. In what follows, we consider the stochastic process in its stationary state.

For the definition of the time averaged signal $\bar{u}(t)$, we use the values of $u(t)$ averaged over a centered time-averaging window with length T_w , so the averaged signal is defined as

$$\bar{u}(t) = \frac{1}{T_w} \int_{t-T_w/2}^{t+T_w/2} u(s) ds \quad (4.13)$$

To define the decorrelation time $\bar{T}_d$, we use the auto-correlation function, $ACF(\tau) = \mathbb{E}[\bar{u}(t)\bar{u}(t+\tau)]/\text{var}[\bar{u}(t)] = \bar{C}(\tau)/\bar{C}(0)$ for time lag τ where $\bar{C}(\tau)$ is the auto-covariance function for lag τ . Then the analytic formula for the de-correlation time of $\bar{u}(t)$ is defined

as

$$\begin{aligned}
\bar{T}_d &= \int_0^\infty ACF(\tau) d\tau \\
&= \int_0^\infty \frac{\bar{C}(\tau)}{\bar{C}(0)} d\tau \\
&= \frac{\int_0^\infty \bar{C}(\tau) d\tau}{\bar{C}(0)}.
\end{aligned} \tag{4.14}$$

The next goal is to find the formula for $\bar{C}(\tau)$. To do this, it is convenient to re-write the definition of $\bar{u}(t)$ as

$$\begin{aligned}
\bar{u}(t) &= \frac{1}{T_w} \int_{t-T_w/2}^{t+T_w/2} u(s) ds \\
&= \frac{1}{T_w} \int_{-\infty}^{\infty} u(s) \mathbb{1}_{[t-T_w/2, t+T_w/2]}(s) ds \\
&= \int_{-\infty}^{\infty} u(\alpha) g(t - \alpha) d\alpha \\
&= (g * u)(t)
\end{aligned} \tag{4.15}$$

where $\mathbb{1}(t)$ is the indicator function and $g(t) = \frac{1}{T_w} \mathbb{1}_{[-T_w/2, T_w/2]}(t)$. Hence, this expresses $\bar{u}(t)$ as a convolution between the original signal $u(t)$ and a rectangular function $g(t)$. By now transforming to Fourier space, the convolution will become a multiplication and allow several useful formulas to be obtained:

$$\begin{aligned}
\widehat{\bar{u}}_\nu &= \widehat{g}_\nu \widehat{u}_\nu, \\
|\widehat{\bar{u}}_\nu|^2 &= |\widehat{g}_\nu|^2 |\widehat{u}_\nu|^2, \\
E|\widehat{\bar{u}}_\nu|^2 &= |\widehat{g}_\nu|^2 \cdot \mathbb{E}|\widehat{u}_\nu|^2, \\
\widehat{\bar{C}}(\nu) &= \widehat{G}(\nu) \cdot \widehat{C}(\nu)
\end{aligned} \tag{4.16}$$

$$\begin{aligned}
\widehat{G}(\nu) &= |\widehat{g}(\nu)|^2 = \left| \int_{-\infty}^{\infty} \frac{1}{T_w} \mathbb{1}_{[-T_w/2, T_w/2]}(x) e^{-2\pi x \nu i} dx \right|^2 \\
&= \left| \frac{1}{T_w} \int_{-T_w/2}^{T_w/2} e^{-2\pi x \nu i} dx \right|^2 \\
&= \left| \frac{1}{-2\pi T_w \nu i} (e^{-\pi \nu T_w i} - e^{\pi \nu T_w i}) \right|^2 \\
&= \left| \frac{1}{2\pi T_w \nu} i (\cos(\pi \nu T_w) - i \sin(\pi \nu T_w) - \cos(\pi \nu T_w) - i \sin(\pi \nu T_w)) \right|^2 \\
&= \left| \frac{1}{\pi T_w \nu} \sin(\pi \nu T_w) \right|^2 \\
&= \frac{\sin^2(\pi T_w \nu)}{\pi^2 T_w^2 \nu^2}
\end{aligned} \tag{4.17}$$

From these expressions, one can see that the problem of finding $\overline{C}(\tau)$ is equivalent to the problem of finding $\widehat{\widehat{C}}(\nu)$, since the two quantities are related by a Fourier transform.

To find an analytic expression for $\widehat{\widehat{C}}(\nu)$, we start from the definition of the OU process to find $\hat{u}_\nu$:

$$\begin{aligned}
\frac{du(t)}{dt} &= -u_t/T_d + \sigma \dot{W}_t \\
\int_{-\infty}^{\infty} \frac{du(t)}{dt} e^{-2\pi i \nu t} dt &= -\hat{u}_\nu/T_d + \sigma \widehat{\dot{W}}_\nu \\
u(t) e^{-2\pi i \nu t} \Big|_{-\infty}^{\infty} + 2\pi i \nu \int_{-\infty}^{\infty} u(t) e^{-2\pi i \nu t} dt &= -\hat{u}_\nu/T_d + \sigma \widehat{\dot{W}}_\nu \\
2\pi i \nu \hat{u}_\nu &= -\hat{u}_\nu/T_d + \sigma \widehat{\dot{W}}_\nu \\
\hat{u}_\nu &= \frac{\sigma \widehat{\dot{W}}_\nu}{1/T_d + 2\pi \nu i}
\end{aligned} \tag{4.18}$$

$$\begin{aligned}
\mathbb{E}[\widehat{u}_\nu \widehat{u}_{\nu'}^*] &= \frac{\sigma}{1/T_d + 2\pi \nu i} \cdot \frac{\sigma}{1/T_d - 2\pi \nu' i} \mathbb{E}[\widehat{\dot{W}}_\nu \widehat{\dot{W}}_{\nu'}^*] \\
&= \frac{\sigma^2}{1/T_d^2 + 4\pi^2 \nu^2} \delta(\nu - \nu')
\end{aligned} \tag{4.19}$$

Hence

$$\widehat{C}(\nu) = \frac{\sigma^2}{1/T_d^2 + 4\pi^2\nu^2} \quad (4.20)$$

By the equality in (4.16),

$$\widehat{\overline{C}}(\nu) = \widehat{G}(\nu) \cdot \widehat{C}(\nu) = \frac{\sin^2(\pi T_w \nu)}{\pi^2 T_w^2 \nu^2} \frac{\sigma^2}{1/T_d^2 + 4\pi^2 \nu^2} \quad (4.21)$$

Returning now to $\overline{C}(\tau)$, the numerator $\int_0^\infty \overline{C}(\tau) d\tau$ in the definition of the decorrelation time (4.14) can be derived as

$$\begin{aligned} \int_0^\infty \overline{C}(\tau) d\tau &= \frac{1}{2} \int_{-\infty}^\infty \overline{C}(\tau) d\tau \\ &= \frac{1}{2} \int_{-\infty}^\infty \overline{C}(\tau) e^{-2\pi i \nu \tau} d\tau \Big|_{\nu=0} \\ &= \frac{1}{2} \widehat{\overline{C}}(\nu=0) \\ &= \frac{1}{2} \lim_{\nu \rightarrow 0} \widehat{\overline{C}}(\nu) \\ &= \frac{1}{2} \lim_{\nu \rightarrow 0} \frac{\sin^2(\pi T_w \nu)}{\pi^2 T_w^2 \nu^2} \frac{\sigma^2}{1/T_d^2 + 4\pi^2 \nu^2} \\ &= \frac{\sigma^2 T_d^2}{2} \end{aligned} \quad (4.22)$$

Then what remains needed from (4.14) is the denominator, $\overline{C}(0)$. From the equation (52) in the paper [39], for a complex-valued OU process, the variance of the time averaged signal has been calculated. By taking the $\gamma = 1/T_d$ and $\omega = 0$, the auto-covariance function for the averaged real-valued OU process can be found as

$$\overline{C}(0) = \frac{\sigma^2 T_d^3}{T_w^2} (T_w/T_d - 1 + e^{-T_w/T_d}) \quad (4.23)$$

Finally, from the definition of the decorrelation time in (4.14), we find

$$\begin{aligned}
\overline{T}_d &= \frac{\int_0^\infty \overline{C}(\tau) d\tau}{\overline{C}(0)} \\
&= \frac{\sigma^2 T_d^2 / 2}{\frac{\sigma^2 T_d^3}{T_w^2} (T_w/T_d - 1 + e^{-T_w/T_d})} \\
&= T_d \cdot \frac{(T_w/T_d)^2}{2(T_w/T_d - 1 + e^{-T_w/T_d})},
\end{aligned} \tag{4.24}$$

which is the desired result of an analytic expression for the decorrelation time of the time-averaged process $\overline{u}(t)$.

As a consistency check, it can be verified that, for the original unaveraged signal $u(t)$, the decorrelation time is just T_d ; to see this, take the limit $T_w \rightarrow 0$ in (4.24) to see that

$$\begin{aligned}
\lim_{T_w \rightarrow 0} \overline{T}_d &= \lim_{T_w \rightarrow 0} T_d \cdot \frac{(T_w/T_d)^2}{2(T_w/T_d - 1 + 1 - T_w/T_d + \frac{(T_w/T_d)^2}{2} + O(T_w^3))} \\
&= T_d \lim_{T_w \rightarrow 0} \frac{(T_w/T_d)^2}{(T_w/T_d)^2 + O(T_w^3)} \\
&= T_d.
\end{aligned}$$

Furthermore, a simpler, approximate version of (4.24) can be derived as follows,

assuming that T_w/T_d is small:

$$\begin{aligned}
\bar{T}_d &= T_d \cdot \frac{(T_w/T_d)^2}{2(T_w/T_d - 1 + e^{-T_w/T_d})} \\
&= T_d \cdot \frac{(T_w/T_d)^2}{2(T_w/T_d - 1 + (1 - T_w/T_d + \frac{(T_w/T_d)^2}{2!} - \frac{(T_w/T_d)^3}{3!} + O((T_w/T_d)^4))} \\
&= T_d \cdot \frac{(T_w/T_d)^2}{(T_w/T_d)^2 - \frac{1}{3}(T_w/T_d)^3 + O((T_w/T_d)^4)} \\
&= T_d \cdot \frac{1}{1 - \frac{1}{3}(T_w/T_d) + O((T_w/T_d)^2)} \\
&= T_d(1 + \frac{1}{3}(T_w/T_d) + O((T_w/T_d)^2)) \\
&= T_d + \frac{1}{3}T_w + T_d \cdot O(T_w^2/T_d^2)
\end{aligned} \tag{4.25}$$

For small T_w/T_d , equation (4.25) leads to

$$\bar{T}_d \approx T_d + \frac{1}{3}T_w. \tag{4.26}$$

Chapter 5

Conclusions

5.1 Research Questions and Main Findings of the Study

In this thesis, we start with an introduction about Madden-Julian oscillation (MJO) and Convectively Coupled Equatorial Waves (CCEWs) which play important roles in weather forecasting in both tropical and extratropical regions. We then use the observational datasets to estimate the predictability of tropical rainfall and waves. Through modeling the signal components at each wavenumber and within each type of the equatorial waves, we find out the predictability of MJO and CCEWs as well as their importance in predicting the rainfall.

Carrying the question about how to improve the forecast after researching in the predictability of the different waves and tropical rainfall, we move on to consider about the time and spatial averaging impact on weather forecast since intuitively they are expected to remove high-frequency unpredictable components and improve the forecast skill.

With the help of the simple complex Ornstein-Uhlenbeck (cOU) process and stochastic heat equation, we research in the impacts of time and spatial averaging to get some insights first. It shows the effect of time averaging depends on the definitions of

the temporal averaging window and the lead time while spatial averaging always helps in improving the forecast skill. We surprisingly find out that time averaging is very limited in improving forecast when a centered or right defined time averaging window is applied, particularly compared to the effect of spatial averaging.

After the analytical investigation in the simple idealized stochastic models, we start to wonder whether the same conclusion about the limited improvement in forecast by time averaging still holds in modern operational forecast data or not. The GFS precipitation and surface temperature are analyzed to explore the effects of time and spatial averaging on forecast on the real forecast datasets, which validates that time averaging has much more limited improvement in forecast skill compared to spatial averaging again.

5.2 Future Work

Throughout the thesis, we systematically investigate in the predictability of the tropical rainfall and the effects of time and spatial averaging on weather forecast. We find interesting results and raise new questions through this process, which is definitely worth further research in the future.

As next steps, higher resolution dataset can be used to see how much more predictability will be brought for MJO, CCEWs and the tropical rainfall. At the same time, the effect of finer spatial averaging such as spatial averaging with a diameter smaller than 100km can be analyzed without the concern of applying bilinear interpolation for sampling and smaller increments in the diameter of the spatial averaging instead of about a 450km increment as what we did in Chapter 4 will be allowed with the availability of higher resolution data.

It would also be interesting to investigate the effects of time averaging on other

time scales, such as subseasonal to seasonal predictions or short-term climate predictions.

In Chapter 4, when we compare the spatial and time averaging, the coastline shows a very big difference from the land and ocean. Investigations about the global behaviors of time and spatial averaging by excluding the coastline or separating the land and ocean will also be a good topic to be researched.

Furthermore, how can we integrate the information we find out to pick one optimal averaging window for improving the forecast is worth exploring as well. Another idea is how to improve the forecast skill at finer spatial locations with the help of spatial averaging. Since larger spatial averaging helps in improving forecast skill more, perhaps Bayesian method can be applied to help in forecasting quantities on finer grids from the better forecast on coarser grids.

There are indeed a lot of ideas and research can be implemented on related questions and topics in the future beyond the work shown here!

Bibliography

- [1] R. ALEXANDER, Z. ZHAO, E. SZÉKELY, AND D. GIANNAKIS, *Kernel analog forecasting of tropical intraseasonal oscillations*, J. Atmos. Sci., 74 (2017), pp. 1321–1342.
- [2] F. ATGER, *Verification of intense precipitation forecasts from single models and ensemble prediction systems*, Nonlinear Processes in Geophysics, 8 (2001), pp. 401–417.
- [3] A. G. BARNSTON, M. H. GLANTZ, AND Y. HE, *Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Niño episode and the 1998 La Niña onset*, Bull. Amer. Meteorol. Soc., 80 (1999), pp. 217–244.
- [4] P. BAUER, A. THORPE, AND G. BRUNET, *The quiet revolution of numerical weather prediction*, Nature, 525 (2015), pp. 47–55.
- [5] B. CASATI, G. ROSS, AND D. B. STEPHENSON, *A new intensity-scale approach for the verification of spatial precipitation forecasts*, Meteorological Applications, 11 (2004), pp. 141–154.
- [6] C. CASSOU, *Intraseasonal interaction between the Madden-Julian oscillation and the North Atlantic oscillation*, Nature, 455 (2008), pp. 523–527.
- [7] N. CHEN AND A. J. MAJDA, *Predicting the real-time multivariate Madden–Julian oscillation index through a low-order nonlinear stochastic model*, Mon. Wea. Rev., 143 (2015), pp. 2148–2169.

- [8] N. CHEN, A. J. MAJDA, AND D. GIANNAKIS, *Predicting the cloud patterns of the Madden–Julian oscillation through a low-order nonlinear stochastic model*, Geophys. Res. Lett., 41 (2014), pp. 5612–5619.
- [9] S. CHEN, A. J. MAJDA, AND S. N. STECHMANN, *Multiscale asymptotics for the skeleton of the Madden–Julian oscillation and tropical–extratropical interactions*, Math. Clim. Weather Forecast., 1 (2015), pp. 43–69.
- [10] ———, *Tropical–extratropical interactions with the MJO skeleton and climatological mean flow*, J. Atmos. Sci., 73 (2016), pp. 4101–4116.
- [11] S. CHEN AND S. N. STECHMANN, *Nonlinear traveling waves for the skeleton of the Madden–Julian oscillation*, Comm. Math. Sci., 14 (2016), pp. 571–592.
- [12] CLIMATE PREDICTION CENTER, NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION, NATIONAL WEATHER SERVICE, NOAA, U.S. DEPARTMENT OF COMMERCE, *CPC Merged Analysis of Precipitation (CMAP)*, 1995. <https://doi.org/10.5065/4QKP-PF57>.
- [13] T. DELSOLE, *Stochastic models of quasigeostrophic turbulence*, Surv. Geophys., 25 (2004), pp. 107–149.
- [14] A. DHOOGHE, W. GOVAERTS, Y. A. KUZNETSOV, W. MESTROM, A. M. RIET, AND B. SAUTOIS, *Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics*, Journal of the Atmospheric Sciences, 64 (2007), p. 3579–3594.
- [15] J. DIAS, M. GEHNE, G. N. KILADIS, N. SAKAEDA, P. BECHTOLD, AND T. HAIDEN, *Equatorial waves and the skill of NCEP and ECMWF numerical weather prediction systems*, Mon. Wea. Rev., 146 (2018), pp. 1763–1784.

- [16] E. E. EBERT, *Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework*, Meteorological applications, 15 (2008), pp. 51–64.
- [17] ———, *Neighborhood verification: A strategy for rewarding close forecasts*, Weather and Forecasting, 24 (2009), pp. 1498–1510.
- [18] N. N. C. FOR ENVIRONMENTAL PREDICTION (NCEP), *NOAA/NCEP Global Forecast System (GFS) atmospheric model*, 2011. Online Resource Linkage: <https://www.ncep.noaa.gov>.
- [19] C. W. GARDINER, *Handbook of stochastic methods: for physics, chemistry & the natural sciences*, vol. 13 of Springer Series in Synergetics, Springer–Verlag, Berlin, 2004.
- [20] J. HARLIM AND A. J. MAJDA, *Filtering turbulent sparsely observed geophysical flows*, Monthly Weather Review, 138 (2010), pp. 1050–1083.
- [21] W. HIGGINS, J. SCHEMM, W. SHI, AND A. LEETMAA, *Extreme precipitation events in the western united states related to tropical forcing*, J. Climate, 17 (2000), pp. 793–820.
- [22] S. HOTTOVY AND S. N. STECHMANN, *A spatiotemporal stochastic model for tropical precipitation and water vapor dynamics*, J. Atmos. Sci., 72 (2015), pp. 4721–4738.
- [23] G. J. HUFFMAN, D. T. BOLVIN, D. BRAITHWAITE, K. HSU, R. JOYCE, C. KIDD, E. NELKIN, S. SOROOSHIAN, J. TAN, AND P. XIE, *NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), Version 06*, 2019. <ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/>.

- [24] M.-P. HUNG, J.-L. LIN, W. WANG, D. KIM, T. SHINODA, AND S. J. WEAVER, *MJO and convectively coupled equatorial waves simulated by CMIP5 climate models*, J. Climate, 26 (2013), pp. 6185–6214.
- [25] M. JANIGA, C. SCHRECK, R. JA, M. FLATAU, N. BARTON, E. METZGER, AND C. REYNOLDS, *Subseasonal forecasts of convectively coupled equatorial waves and the mjo: activity and predictive skill*, Monthly Weather Review, 146 (2018), pp. 2337—2360.
- [26] M. A. JANIGA, C. SCHRECK, J. A. RIDOUT, M. FLATAU, N. BARTON, E. J. METZGER, AND C. REYNOLDS, *Subseasonal forecasts of convectively coupled equatorial waves and the MJO: Activity and predictive skill*, Mon. Wea. Rev., (2018), p. in press.
- [27] X. JIANG, D. E. WALISER, P. K. XAVIER, J. PETCH, N. P. KLINGAMAN, S. J. WOOLNOUGH, B. GUAN, G. BELLON, T. CRUEGER, C. DEMOTT, ET AL., *Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations*, Journal of Geophysical Research: Atmospheres, 120 (2015), pp. 4718–4748.
- [28] N. C. JOHNSON, D. C. COLLINS, S. B. FELDSTEIN, M. L. L’HEUREUX, AND E. E. RIDDLE, *Skillful wintertime north american temperature forecasts out to 4 weeks based on the state of ENSO and the MJO*, Wea. Forecasting, 29 (2014), p. 23–38.
- [29] C. JONES AND L. CARVALHO, *Active and break phases in the south american monsoon system*, J. Climate, 15 (2002), pp. 905–914.
- [30] B. KHOUIDER AND A. J. MAJDA, *A simple multicloud parameterization for convectively coupled tropical waves. Part II: Nonlinear simulations*, J. Atmos. Sci., 64

(2007), pp. 381–400.

- [31] B. KHOUIDER, A. J. MAJDA, AND S. N. STECHMANN, *Climate science in the tropics: waves, vortices and PDEs*, Nonlinearity, 26 (2013), pp. R1–R68.
- [32] G. N. KILADIS, J. DIAS, K. H. STRAUB, M. C. WHEELER, S. N. TULICH, K. KIKUCHI, K. M. WEICKMANN, AND M. J. VENTRICE, *A comparison of OLR and circulation-based indices for tracking the MJO*, Mon. Wea. Rev., 142 (2014), pp. 1697–1715.
- [33] G. N. KILADIS, M. C. WHEELER, P. T. HAERTEL, K. H. STRAUB, AND P. E. ROUNDY, *Convectively coupled equatorial waves*, Rev. Geophys., 47 (2009), p. RG2003.
- [34] H.-M. KIM, F. VITART, AND D. E. WALISER, *Prediction of the Madden–Julian oscillation: A review*, J. Climate, 31 (2018), p. 9425–9443.
- [35] M. LATIF, D. ANDERSON, T. BARNETT, M. CANE, R. KLEEMAN, A. LEETMAA, J. O’BRIEN, A. ROSATI, AND E. SCHNEIDER, *A review of the predictability and prediction of ENSO*, J. Geophys. Res.: Oceans, 103 (1998), pp. 14375–14393.
- [36] W. K. M. LAU AND D. E. WALISER, eds., *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, Springer, Berlin, 2nd ed., 2012.
- [37] S. LAVENDER AND A. MATTHEWS, *Response of the west african monsoon to the Madden-Julian oscillation*, J. Climate, 22 (2009), pp. 4097–4116.
- [38] J. M. LEWIS, *Roots of ensemble forecasting*, Monthly Weather Review, 133 (2005), pp. 1865–1885.
- [39] Y. LI AND S. N. STECHMANN, *Spatial and temporal averaging windows and their*

- impact on forecasting: exactly solvable examples*, Math. Clim. Weather Forecast., 4 (2018), pp. 23–49.
- [40] ———, *Predictability of tropical rainfall and waves: Estimates from observational data*, Q. J. Roy. Met. Soc., 146 (2020), pp. 1668–1684.
- [41] B. LIEBMANN AND C. A. SMITH, *Description of a complete (interpolated) outgoing longwave radiation dataset*, Bull. Amer. Meteor. Soc., 77 (1996), pp. 1275–1277.
- [42] J.-L. LIN, G. N. KILADIS, B. E. MAPES, K. M. WEICKMANN, K. R. SPERBER, W. LIN, M. WHEELER, S. D. SCHUBERT, A. DEL GENIO, L. J. DONNER, S. EMORI, J.-F. GUEREMY, F. HOURDIN, P. J. RASCH, E. ROECKNER, AND J. F. SCINOCCHA, *Tropical intraseasonal variability in 14 IPCC AR4 climate models Part I: Convective signals*, J. Climate, 19 (2006), pp. 2665–2690.
- [43] E. N. LORENZ, *Atmospheric predictability as revealed by naturally occurring analogues*, Journal of the Atmospheric Sciences, 26 (1969), pp. 636–646.
- [44] ———, *Atmospheric predictability experiments with a large numerical mode*, Tellus, 34 (1982), p. 505–513.
- [45] ———, *Predictability—a problem partly solved*, Proc. Seminar on Predictability, Reading, United Kingdom, ECMWF, (1996), p. 1–18.
- [46] R. A. MADDEN AND P. R. JULIAN, *Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific*, J. Atmos. Sci., 28 (1971), pp. 702–708.
- [47] ———, *Description of global-scale circulation cells in the tropics with a 40–50 day period.*, J. Atmos. Sci., 29 (1972), pp. 1109–1123.
- [48] A. MAJDA, I. TIMOFEYEV, AND E. VANDEN EIJNDEN, *A mathematical framework*

- for stochastic climate models*, Communications on Pure and Applied Mathematics, 54 (2001), pp. 891–974.
- [49] A. J. MAJDA, *Introduction to PDEs and Waves for the Atmosphere and Ocean*, vol. 9 of Courant Lecture Notes in Mathematics, American Mathematical Society, Providence, 2003.
- [50] A. J. MAJDA AND J. HARLIM, *Filtering Turbulent Complex Systems*, Cambridge University Press, 2012.
- [51] A. J. MAJDA AND S. N. STECHMANN, *The skeleton of tropical intraseasonal oscillations*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 8417–8422.
- [52] —, *Nonlinear dynamics and regional variations in the MJO skeleton*, J. Atmos. Sci., 68 (2011), pp. 3053–3071.
- [53] E. D. MALONEY AND D. L. HARTMANN, *Modulation of hurricane activity in the Gulf of Mexico by the Madden–Julian oscillation*, Science, 287 (2000), pp. 2002–2004.
- [54] A. MARIOTTI, P. M. RUTI, AND M. RIXEN, *Progress in subseasonal to seasonal prediction through a joint weather and climate community effort*, npj Climate and Atmospheric Science, 1 (2018), pp. 1–4.
- [55] C. MARSIGLI, A. MONTANI, AND T. PACCANGNELLA, *A spatial verification method applied to the evaluation of high-resolution ensemble forecasts*, Meteorological Applications, 15 (2008), pp. 125–143.
- [56] J. D. NEELIN, D. S. BATTISTI, A. C. HIRST, F.-F. JIN, Y. WAKATA, T. YAMAGATA, AND S. E. ZEBIAK, *ENSO theory*, J. Geophys. Res.: Oceans, 103 (1998),

pp. 14261–14290.

- [57] J. M. NEENA, J. Y. LEE, D. WALISER, B. WANG, AND X. JIANG, *Predictability of the Madden–Julian oscillation in the intraseasonal variability hindcast experiment (ISVHE)*, J. Climate, 27 (2014), pp. 4531–4543.
- [58] H. R. OGROSKY AND S. N. STECHMANN, *Assessing the equatorial long-wave approximation: asymptotics and observational data analysis*, J. Atmos. Sci., 72 (2015), pp. 4821–4843.
- [59] ———, *The MJO skeleton model with observation-based background state and forcing*, Q. J. Roy. Meteor. Soc., 141 (2015), pp. 2654–2669.
- [60] ———, *Identifying convectively coupled equatorial waves using theoretical wave eigenvectors*, Mon. Wea. Rev., 144 (2016), pp. 2235–2264.
- [61] H. R. OGROSKY, S. N. STECHMANN, AND S. HOTTOVY, *Instability and nonlinear evolution of the MJO in a model with vertically-varying convective adjustment*, J. Adv. Model. Earth Syst., (2018), p. submitted.
- [62] ———, *Instability and nonlinear evolution of the MJO in a tropical channel model with vertically-varying convective adjustment*, Theor. Comp. Fluid Dyn., (2019), p. in press.
- [63] H. R. OGROSKY, S. N. STECHMANN, AND A. J. MAJDA, *Boreal summer intraseasonal oscillations in the MJO skeleton model with observation-based forcing*, Dyn. Atmos. Oceans, 78 (2017), pp. 38–56.
- [64] C. PENLAND AND T. MAGORIAN, *Prediction of Niño 3 sea surface temperatures using linear inverse modeling*, J. Climate, 6 (1993), pp. 1067–1076.

- [65] E. E. RIDDLE, M. B. STONER, N. C. JOHNSON, L. M. L., D. C. COLLINS, AND S. B. FELDSTEIN, *The impact of the mjo on clusters of wintertime circulation anomalies over the north american region*, *Climate Dyn.*, 40 (2013), pp. 1749–1766.
- [66] N. ROBERTS, *Assessing the spatial and temporal variation in the skill of precipitation forecasts from an nwp model*, *Meteorological Applications*, 15 (2008), pp. 163–169.
- [67] ———, *Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model*, *Meteorological Applications*, 15 (2008), pp. 163–169.
- [68] N. M. ROBERTS AND H. W. LEAN, *Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events*, *Monthly Weather Review*, 136 (2008), pp. 78–97.
- [69] A. W. ROBERTSON, A. KUMAR, M. PEÑA, AND F. VITART, *Improving and promoting subseasonal to seasonal prediction*, *Bulletin of the American Meteorological Society*, 96 (2015), pp. ES49–ES53.
- [70] M. J. RODWELL AND F. J. DOBLAS-REYES, *Medium-range, monthly, and seasonal prediction for Europe and the use of forecast information*, *J. Climate*, 19 (2006), pp. 6025–6046.
- [71] Z. S., M. L’HEUREUX, S. WEAVER, AND A. KUMAR, *A composite study of mjo influence on the surface air temperature and precipitation over the continental united states*, *Climate Dyn.*, 38 (2012), pp. 1459–1471.
- [72] S. N. STECHMANN AND S. HOTTOVY, *Cloud regimes as phase transitions*, *Geophys. Res. Lett.*, 43 (2016), pp. 6579–6587.

- [73] —, *Unified spectrum of tropical rainfall and waves in a simple stochastic model*, Geophys. Res. Lett., 44 (2017), pp. 10,713–10,724.
- [74] S. N. STECHMANN AND A. J. MAJDA, *Identifying the skeleton of the Madden-Julian oscillation in observational data*, Mon. Wea. Rev., 143 (2015), pp. 395–416.
- [75] S. N. STECHMANN AND H. R. OGROSKY, *The Walker circulation, diabatic heating, and outgoing longwave radiation*, Geophys. Res. Lett., 41 (2014), pp. 9097–9105.
- [76] Y. Q. SUN AND F. ZHANG, *Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect*, Journal of the Atmospheric Sciences, 73 (2016), p. 1419–1438.
- [77] Y. N. TAKAYABU, *Large-scale cloud disturbances associated with equatorial waves. I: Spectral features of the cloud disturbances*, J. Meteor. Soc. Japan, 72 (1994), pp. 433–449.
- [78] —, *Large-scale cloud disturbances associated with equatorial waves. II: Westward-propagating inertio-gravity waves*, J. Meteor. Soc. Japan, 72 (1994), pp. 451–465.
- [79] S. THUAL, A. J. MAJDA, N. CHEN, AND S. N. STECHMANN, *Simple stochastic model for El Niño with westerly wind bursts*, Proc. Natl. Acad. Sci., 113 (2016), pp. 10245–10250.
- [80] S. THUAL, A. J. MAJDA, AND S. N. STECHMANN, *A stochastic skeleton model for the MJO*, J. Atmos. Sci., 71 (2014), pp. 697–715.
- [81] H. VAN DEN DOOL, *Empirical Methods in Short-Term Climate Prediction*, Oxford University Press, 2007.
- [82] F. VITART AND A. W. ROBERTSON, *The sub-seasonal to seasonal prediction*

- project (s2s) and the prediction of extreme events*, npj Climate and Atmospheric Science, 1 (2018), p. 3549.
- [83] D. WALISER, *Predictability and forecasting*, in Intraseasonal Variability in the Atmosphere–Ocean Climate System, W. K. M. Lau and D. E. Waliser, eds., Springer, Berlin, 2012.
- [84] M. WHEELER AND G. N. KILADIS, *Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber–frequency domain*, J. Atmos. Sci., 56 (1999), pp. 374–399.
- [85] M. WHEELER AND K. M. WEICKMANN, *Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability*, Mon. Wea. Rev., 129 (2001), pp. 2677–2694.
- [86] M. C. WHEELER AND H. H. HENDON, *An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction*, Mon. Wea. Rev., 132 (2004), pp. 1917–1932.
- [87] M. C. WHEELER, H. ZHU, A. H. SOBEL, D. HUDSON, AND F. VITART, *Seamless precipitation prediction skill comparison between two global models.*, Q.J.R. Meteorol. Soc., 143 (2017), pp. 374–383.
- [88] Y. YING AND F. ZHANG, *Practical and intrinsic predictability of multiscale weather and convectively coupled equatorial waves during the active phase of an MJO*, J. Atmos. Sci., 74 (2017), p. 3771–3785.
- [89] —, *Potentials in improving predictability of multiscale tropical weather systems evaluated through ensemble assimilation of simulated satellite-based observations*, J. Atmos. Sci., 75 (2018), pp. 1675–1698.

- [90] C. ZHANG, *Madden–Julian oscillation*, Rev. Geophys., 43 (2005), p. RG2003.
- [91] F. ZHANG, A. M. ODINS, AND J. W. NIELSEN-GAMMON, *Mesoscale predictability of an extreme warm-season precipitation event*, Wea. Forecasting, 21 (2006), p. 149–166.
- [92] H. ZHU, M. C. WHEELER, A. H. SOBEL, AND D. HUDSON, *Seamless precipitation prediction skill in the tropics and extratropics from a global model*, Mon. Wea. Rev., 142 (2014), pp. 1556–1569.