
Regime-Dependent Predictability and Forecast Error Growth in Kilometer-Scale Models

I-Han Chen



München 2024

Regime-Dependent Predictability and Forecast Error Growth in Kilometer-Scale Models

I-Han Chen

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
I-Han Chen
aus Taipei City, Taiwan

München, den 07.03.2024

Erstgutachter: Prof. Dr. George Craig

Zweitgutachter: Prof. Dr. Hella Garny

Tag der mündlichen Prüfung: 08.04.2024

Parts of this thesis are included in:

Chen, I.-H., J. Berner, C. Keil, Y.-H. Kuo, and G. Craig, 2024: Classification of Warm-Season Precipitation in High-Resolution Rapid Refresh (HRRR) Model Forecasts over the Contiguous United States. *Monthly Weather Review*, **152** (1), 187–201. ©American Meteorological Society. **Used with permission.**

Chen, I.-H., J. Berner, C. Keil, G. Thompson, Y.-H. Kuo, and G. Craig, 2024: To which degree do the details of stochastic perturbation schemes matter for convective-scale and mesoscale error growth? *Submitted to Monthly Weather Review (In review)*.

Kuo, Y.-H., J. Sun, Y. Zhang, **I.-H. Chen**, Y. Ho, and J.-S. Hong, 2024: Impact of Assimilating GNSS Radio Occultation Data on the Prediction of a Squall Line associated with a Mei-Yu front: Part 1. Data Assimilation Study. *To be submitted*.

Chen, I.-H., Y.-H. Kuo, C. Keil, J. Berner, and G. Craig, 2024: Impact of Assimilating GNSS Radio Occultation Data on the Prediction of a Squall Line associated with a Mei-Yu front: Part 2. Predictability Study. *To be submitted*.

Zusammenfassung

Numerische Wettervorhersagemodelle (NWP) haben in den letzten Jahrzehnten erhebliche Fortschritte gemacht. Dank der ständig wachsenden Rechenleistung werden NWP-Systeme, die sich von der synoptischen bis zur konvektiven Skala erstrecken, jetzt routinemäßig in Wettervorhersagezentren weltweit gerechnet. Obwohl sich die Qualität der NWP-Vorhersagen unbestreitbar verbessert hat, stößt jede Vorhersage irgendwann an ihre Grenzen. Insbesondere bei Vorhersagen auf der Kilometerskala (auch bekannt als Vorhersagen auf der konvektiven Skala) wird diese Grenze der Vorhersagbarkeit immer deutlicher, da die Vorhersagefehler um ein Vielfaches schneller wachsen als auf der synoptischen Skala. Das übergeordnete Ziel dieser Dissertation ist es, "das Verständnis der strömungsabhängigen Vorhersagbarkeit und des Wachstums des Vorhersagefehlers in Vorhersagemodellen der konvektiven Skala zu verbessern". In dieser Dissertation werden drei Einzelstudien durchgeführt, um das übergeordnete Ziel zu erreichen.

Die erste Studie nutzt die konvektive Zeitskala, um Wettersituationen objektiv zu identifizieren und eine Klimatologie der konvektiven Wetterlagen im Bereich der kontinentalen USA (CONUS) zu erstellen. Klimatologisch zeigen konvektive Wetterlagen regionale Abhängigkeiten. Im nördlichen CONUS ist der Einfluss der synoptischen Strömung stärker, ähnlich wie in Deutschland, während im südlichen CONUS die lokal angetriebene Konvektion stärker ausgeprägt ist. Diese objektive Klassifizierung spiegelt nachweislich unser grundlegendes physikalisches Verständnis wider, wenn wir die jahreszeitlichen und tageszeitlichen Abhängigkeiten untersuchen. Darüber hinaus zeigt diese Studie, dass Gleichgewichtskonvektion in konvektiven Vorhersagen besser vorhersagbar ist als Nicht-Gleichgewichtskonvektion, was mit den Ergebnissen früherer Studien übereinstimmt, die über viel kleinere Gebiete durchgeführt wurden. Da es sich um die erste klimatologische Studie in diesem Bereich handelt, können die Ansätze und Ergebnisse als wertvolle Orientierungshilfe für zukünftige Studien dienen, die sich auf den CONUS-Bereich konzentrieren.

Unter Verwendung der in der ersten Studie beschriebenen Klassifizierungsmetrik für Wetterlagen werden in der zweiten Studie Fallstudien zu Konvektion ausgewählt, die sich über Winterstürme als auch sommerliche Tage mit starkem und schwachem Antrieb erstrecken. Das strömungsabhängige Wachstum des Vorhersagefehlers wird in einem Wettervorhersagemodell der konvektiven Skala untersucht. Die Ergebnisse deuten darauf hin, dass Winterstürme die höchste Vorhersagbarkeit besitzen, während sommerliche

Konvektion bei schwachem Antrieb, insbesondere hinsichtlich des Niederschlags, die geringste Vorhersagbarkeit aufweist. In dieser Studie wird der Modellfehler bezüglich der Unsicherheit in zwei subgrid-skaligen Prozessen formuliert: Stochastisch gestörte Parameter, angewandt auf die Mikrophysik (SPPMP) und die Unsicherheit in der Grenzschicht, die mit dem physikalisch basierten stochastischen Störungsschema (PSP) beschrieben wird. Der zentrale Gedanke ist, zu verstehen, ob das Wachstum des Vorhersagefehlers durch die Strömungsdynamik oder durch Details der Formulierung des Modellfehlers kontrolliert wird. Die Ergebnisse deuten darauf hin, dass bei längeren Vorlaufzeiten der Vorhersagefehler trotz verbleibender geringer Amplitudenunterschiede hauptsächlich durch feuchte Konvektion und nicht durch Details der Störung beeinflusst wird. Dies erklärt, warum in früheren Studien argumentiert wurde, dass die Auswirkungen der stochastischen Verfahren nicht additiv sind, da ihre Auswirkungen nicht orthogonal sind. Im Gegensatz dazu kann der Unterschied zwischen PSP und SPPMP bei kurzen Vorlaufzeiten ausgeprägter sein, da das Fehlerwachstum hauptsächlich auf der konvektiven Skala stattfindet. Trotz ihrer insgesamt geringeren Wirkung können mikrophysikalische Störungen unter bestimmten Bedingungen wichtig sein, da sie Instabilität früher auslösen können, insbesondere zu Zeiten, in denen das PSP-Schema aufgrund einer stabilen Grenzschicht nicht aktiv ist. Dies ist ein wichtiges Erkenntnis für Forschende, die sich auf die Entwicklung stochastischer Formulierungen des Modellfehlers für NWP-Systeme mit kurzer Vorhersagefrist konzentrieren.

In der dritten Studie wird ein Fall von Starkregen verwendet, um die Auswirkungen der Unsicherheit in den Anfangsbedingungen und des Modellfehlers bei Kurzfrist-Konvektionsvorhersagen zu untersuchen. Im Vergleich zur Unsicherheit in den Anfangsbedingungen ist die Niederschlagsvorhersage nicht empfindlich gegenüber dem Modellfehler, der durch das SPPT-Schema (Stochastically Perturbed Parameterization Tendency) dargestellt wird. Da das SPPT-Schema ein ganzheitlicherer Ansatz ist als die SPPMP- und PSP-Schemata, deutet dieses Ergebnis darauf hin, dass der Unterschied zwischen den beiden Schemata bei kurzen Vorhersagefristen nicht signifikant ist, wenn man kleine Fehler bei den Anfangsbedingungen berücksichtigt. Dieses Ergebnis rechtfertigt weitere Untersuchungen in zukünftigen Studien. Diese Studie schließt an Kuo et al. (2023) an, die gezeigt haben, dass die Integration von Radio-Okkultationsdaten (RO) in ein Modell auf der Kilometerskala die Vorhersagbarkeit von konvektivem Niederschlag verbessern kann. Durch die Darstellung von zufälligen Anfangs- und Modellfehlern untermauert diese Studie die Schlussfolgerung von Kuo et al. (2023) weiter, indem sie zeigt, dass die Verbesserung nicht rein zufällig ist, da Vorhersagen, die von Anfangsbedingungen mit RO-Assimilation ausgehen, im Durchschnitt besser sind als solche ohne.

Zusammenfassend umfasst diese Dissertation eine klimatologische Analyse, eine Untersuchung des Fehlerwachstums bei verschiedenen Modellfehlerformulierungen und eine Erforschung des Fehlerwachstums bei Fehlern in den Anfangsbedingungen. Während jede Studie mehrere Forschungsfragen vertieft, trägt diese Dissertation insgesamt zu einem besseren Verständnis von Vorhersagbarkeit und der Rolle verschiedener Fehlerquellen bei.

Abstract

Numerical Weather Prediction (NWP) models have made significant strides in recent decades. Thanks to the ever-increasing computing power, NWP systems focusing on synoptic- to convective scales are now running routinely at weather forecasting centers worldwide. While there has been undeniable improvement in NWP forecast skills, inevitably, every forecast will ultimately encounter its predictability limit. Especially for kilometer-scale (also known as convection-permitting) forecasts, this predictability limit becomes increasingly evident as their forecast errors grow several times faster than those of synoptic-scale forecasts. In this dissertation, the overarching goal is to “**advance the comprehension of regime-dependent predictability and forecast error growth in kilometer-scale models**”. To achieve the overarching goal, three individual studies are conducted.

The first study utilizes the convective adjustment time scale to identify weather regimes objectively and establish convective regime climatology in the Contiguous United States (CONUS) domain. Climatologically, convective regimes demonstrate regional dependencies. In the northern CONUS, there is a stronger influence from synoptic flow, akin to Germany, whereas in the southern CONUS, locally forced convection is more prevalent. This objective classification is proven to reflect our basic physical understanding when examining seasonal and diurnal dependencies. Furthermore, this study demonstrates that equilibrium convection is more predictable than nonequilibrium convection in convection-permitting forecasts, which aligns with findings from previous studies conducted on much smaller domains. As the first climatological study in this area, the approaches and findings can serve as valuable guidance for future studies focusing on the CONUS domain.

Utilizing the regime classification metric outlined in the first study, the second study selects winter storm, strong forcing, and weak forcing convection cases, and investigates regime-dependent forecast error growth in a convection-permitting model. The findings suggest that winter storm cases possess the highest predictability, while nonequilibrium summer convection, especially regarding precipitation, exhibits the lowest predictability. In this study, the forecast error stems from the uncertainty in the model formulation, as represented by the Stochastically Perturbed Parameter applied to Microphysics (SPPMP) scheme and the Physically based Stochastic Perturbation (PSP) applied to the Planetary

Boundary Layer (PBL) scheme. The central idea is to understand whether forecast error growth is primarily controlled by the flow dynamics or the details of the error representation schemes. The results suggest that for longer lead times, despite remaining small amplitude differences, the forecast error is primarily controlled by moist convection rather than the perturbation details. This explains why previous studies argue that the effects of stochastic schemes are not additive, as their impacts are not orthogonal. In contrast, the distinction between PSP and SPPMP can be more pronounced at short lead times, where the error growth predominantly occurs at the convective scale. Despite its overall smaller effect, microphysics perturbations can be important under certain conditions, as they may trigger instability earlier, particularly during times when the PSP scheme is not active due to a stable PBL. This conveys an important message for researchers focusing on developing stochastic physics suites for short-lead-time NWP systems.

The third study utilizes a heavy rainfall case to explore the impact of initial and model uncertainty in short-range convection-permitting forecasts. Overall, the precipitation forecast is not sensitive to the model uncertainty represented by the Stochastically Perturbed Parameterization Tendency (SPPT) scheme compared to the initial uncertainty. As the SPPT scheme is a more holistic approach than the SPPMP and PSP schemes, this result suggests that the difference between the two schemes at short lead times may not be significant when considering small initial condition errors. This finding warrants further investigation in future studies. This study represents a follow-up to Kuo et al. (2023), which illustrated that assimilating Radio Occultation (RO) data into a kilometer-scale model can improve the predictability of convective precipitation. Through representing random initial and model errors, this study further supports their conclusions by demonstrating that the improvement is not purely coincidental as forecasts initialized from initial conditions with RO assimilation, on average, outperform those without.

In summary, this dissertation encompasses a climatological analysis, an investigation into error growth from model error formulations, and an exploration of error growth from initial condition errors. While each study delves deeper into multiple research questions, collectively, this dissertation advances the understanding of predictability and error growth characteristics across various weather regimes in convection-permitting NWP models.

Contents

Zusammenfassung	iv
Abstract	vii
1 Introduction	1
1.1 Numerical Weather Prediction	2
1.1.1 NWP models of different scales	3
1.1.2 Predictability of atmospheric flows	5
1.1.3 Ensemble forecasting and sources of uncertainty	8
1.2 Atmospheric convection and weather regimes	10
1.2.1 Atmospheric instability and relevant indices	10
1.2.2 Equilibrium and nonequilibrium convection	13
1.2.3 Convective adjustment time scale	14
1.3 Parameterization of subgrid scale processes	16
1.3.1 Planetary boundary layer parameterization	16
1.3.2 Microphysics parameterization	18
1.3.3 Stochastic parameterization	20
1.4 Summary, research goals, and outline	22
1.5 Publications	23
2 Characterization of Convective Regimes and Their Predictability	25
2.1 Background and overview	25
2.2 Methodology	26
2.2.1 Definition of the convective adjustment time scale	26
2.2.2 Description of subdomains, forecast, and observation dataset	27
2.3 Convective regime classification	31
2.3.1 Cases illustrating equilibrium and nonequilibrium convection	31
2.3.2 Convective adjustment time scale climatology	34
2.3.3 Diurnal cycle of precipitation under different convective regimes	36
2.4 Convective adjustment time scale as a predictor of forecast skill	41
2.5 Summary and discussion	43

3	Regime-Dependent Forecast Error Growth in Convection-Permitting Models: The Roles of Stochastic Parameterizations	45
3.1	Background and overview	45
3.2	Methodology	46
3.2.1	Model configuration and experimental design	46
3.2.2	Formulation of the SPPMP scheme	48
3.2.3	The WRF implementation of the PSP scheme	49
3.2.4	Case description	52
3.2.5	Diagnostic metric	55
3.3	Two contrasting cases	57
3.4	Impact of stochastic parameterizations in different weather situations . . .	61
3.5	Role of the diurnal cycle	65
3.6	Summary and discussion	71
4	Predictability of a Heavy Precipitation Event: The Impacts of Initial and Model Uncertainty	75
4.1	Background and overview	75
4.2	Case description and numerical simulations	77
4.3	Methodology	80
4.3.1	Model configuration	80
4.3.2	Experimental design	81
4.3.3	Diagnostic metric	84
4.4	The key analysis correction from RO DA	85
4.5	Impact of random initial and model errors	88
4.5.1	12-hour Quantitative Precipitation Forecasting	88
4.5.2	Hourly Quantitative Precipitation Forecasting	89
4.6	Assessment of practical predictability	90
4.7	Summary and discussion	92
5	Summary and conclusions	95
5.1	Summary	95
5.2	Conclusions	98
6	Appendix	101
6.1	Normalized precipitation spread and NRMDTE	101
6.2	Time series of precipitation spread and intensity for winter storm cases . .	102
6.3	Snapshots of water vapor perturbation generated by the PSP scheme . . .	103
6.4	Time series of RMDTE growth rate for winter storm cases	104
6.5	RMDTE ratios for strong forcing convection, weak forcing convection, and winter storm cases	105
	Acknowledgements	107

CONTENTS

xi

List of Abbreviations	109
List of Figures	116
List of Tables	117
Bibliography	135

Chapter 1

Introduction

Numerical Weather Prediction (NWP) models have evolved rapidly in recent decades and are now running routinely at major operational centers worldwide (Bauer et al., 2015). Their applications span from short-range weather to long-range climate prediction, addressing phenomena ranging from convective storms to synoptic patterns. Despite their undeniable success, every forecast will inevitably encounter its predictability limit (Lorenz, 1963, 1969). This limitation is particularly evident in the prediction of convective storms, where errors amplify several times faster than in the prediction of synoptic weather patterns (Hohenegger and Schar, 2007). In recent decades, efforts to improve convective storm predictability involved the development of kilometer-scale models, observational systems, and data assimilation techniques (Sun, 2005; Gustafsson et al., 2018). In addition, kilometer-scale Ensemble Prediction Systems (EPSs), which generate a range of potential outcomes by representing sources of forecast uncertainty, have been developed in operational and research centers to offer uncertainty quantification in forecasts (Gebhardt et al., 2011; Schwartz et al., 2015; Beck et al., 2016). However, Romine et al. (2014) have reported that kilometer-scale EPSs are often under-dispersive or overly confident, requiring further investigation into the error representations and their associated forecast error growth within these models. One key objective of kilometer-scale NWP models is to provide Quantitative Precipitation Forecasting (QPF) for convective storms, such as their intensity, location, and timing, given their significant economic and societal impacts (Ahmadalipour and Moradkhani, 2019; Brázdil et al., 2021; Turner et al., 2021). According to prior studies (Keil et al., 2014, 2019), the predictability of convective precipitation varies and may depend on their convective regimes, with certain cases demonstrating high predictability while others exhibiting almost no predictability.

This dissertation has an overarching goal to **“improve the comprehension of regime-dependent predictability and forecast error growth within kilometer-scale models”**. In pursuit of the overarching goal, this dissertation identifies three research topics: climatological analysis, forecast error growth from different model error formulations, and forecast error growth from initial condition errors, as further detailed below. As this dissertation primarily focuses on weather systems within the Contiguous

United States (CONUS) domain, the first study systematically identifies convective regimes using the convective adjustment time scale (Done et al., 2006) and establishes convective regime climatology using historical forecast datasets from an operational kilometer-scale NWP system. Furthermore, this study aims to explore whether these historical forecasts over the CONUS domain climatologically exhibit regime-dependent precipitation predictability. Utilizing the regime classification procedure established in the first study, the second study investigates regime-dependent forecast error growth, specifically focusing on forecast errors stemming from different model error representations. This investigation aims to elucidate whether incorporating different sources of model error leads to distinct forecast error growth within a kilometer-scale model. Finally, the third study investigates the predictability of a convective precipitation event occurring in a distinct geographic region over East Asia. Using a kilometer-scale model, the third study compares the impact of initial and model uncertainties on short-range forecasts by examining the evolution of precipitation forecast errors.

In summary, the three studies in this dissertation collectively contribute to understanding predictability and forecast error growth across various weather systems in kilometer-scale NWP models. For example, establishing a convective regime climatology within the CONUS domain and elucidating its connection to regime-dependent predictability can aid future studies in distinguishing convection behavior in forecasts. Moreover, as further exploration of error representations in kilometer-scale EPSs is necessary, investigating forecast error growth originating from initial and model errors can provide valuable insights. In the subsequent sections, I will provide an overview of deterministic and probabilistic NWP systems, subgrid parameterization schemes and their uncertainty representations, deep convection and its classification, and conclude with a summary of the research questions.

1.1 Numerical Weather Prediction

NWP models utilize a set of prognostic equations to predict the temporal evolution of the atmospheric states (e.g., temperature, wind, and water vapor) from the observed current weather. These prognostic equations, such as momentum, continuity, and thermodynamic equations, describe the essential physical processes of atmospheric flow with a certain level of complexity. Its inception can be traced back to the 1920s (Richardson, 1922), with the first operational NWP system established in the United States in the 1950s (Lynch, 2008). Owing to the rapid evolution of NWP models and high-performance computing techniques, NWP systems have become indispensable tools for both operational and research communities. In this section, I will discuss NWP models of different scales, delve into the predictability challenges, and elucidate the emergence of ensemble prediction in recent decades.

1.1.1 NWP models of different scales

Given the challenges in solving partial differential equations analytically, these equations are solved numerically through spatial (Kasahara, 1974) and temporal (Wicker and Skamarock, 2002) discretization. The discretization results in a distinction between resolved and unresolved processes. The trajectory of resolved processes is explicitly predicted by stepping the prognostic equations forward in time (typically called dynamics). On the contrary, unresolved processes occurring at scales smaller than the grid spacing are handled by subgrid parameterization schemes (typically called physics). In a general sense, parameterization schemes use the resolved variables and parameters to estimate sources and sinks in the prognostic equations, playing a crucial role in closing the set of primitive equations (Stensrud, 2007). The processes falling into the physics category can vary among NWP models and may depend on their grid spacing (Yu and Lee, 2010; Pearson et al., 2014).

Before exploring NWP models of different scales, it's important to delineate various types of NWP systems: either deterministic or probabilistic, and either regional or global. Broadly speaking, a comprehensive NWP system comprises two essential components: (1) a numerical model with complete dynamics and physics and (2) a Data Assimilation (DA) system that determines the initial state from observations (Lorenc, 1986). Deterministic forecasting generates a single forecast outcome based on a single realization of the numerical model and initial state, while probabilistic forecasting predicts a range of potential forecast outcomes, as further explained in Section 1.1.3. Global and regional models vary in their spatial coverage, which is configured according to factors such as forecast range and targeted weather phenomena. Since regional models depend on prescribed boundary conditions (Warner et al., 1997; Davies, 2014) and may suffer from incomplete scale interaction, their solutions become increasingly contaminated as forecast length increases (Gustafsson et al., 1998). Therefore, they are not suitable for weather prediction beyond a 4-5 day horizon. Usually, such medium-range forecasts are provided by global NWP models in common practice (Whitaker et al., 2008).

Figure 1.1 illustrates various weather systems with their spatial and temporal scales. To accurately predict a system of interest, models must feature suitable grid spacing to effectively resolve its motion. For example, a grid spacing between 1-4 km is generally required to resolve convective-scale weather systems, such as thunderstorms (Kain et al., 2008; Schwartz et al., 2009). Ideally, one would prefer a high-resolution global model without the need to assign lateral boundary conditions. Yet, due to computational constraints, contemporary operational high-resolution models are typically used regionally and integrate over limited areas of interest (Reinert et al., 2020; Dowell et al., 2022). Another difference between coarse-resolution global and high-resolution regional models lies in the parameterization schemes they employ. An important example is that coarse-resolution models (typical grid sizes ~ 10 km) require convection parameterization to estimate precipitation clouds, modify convective instability, and

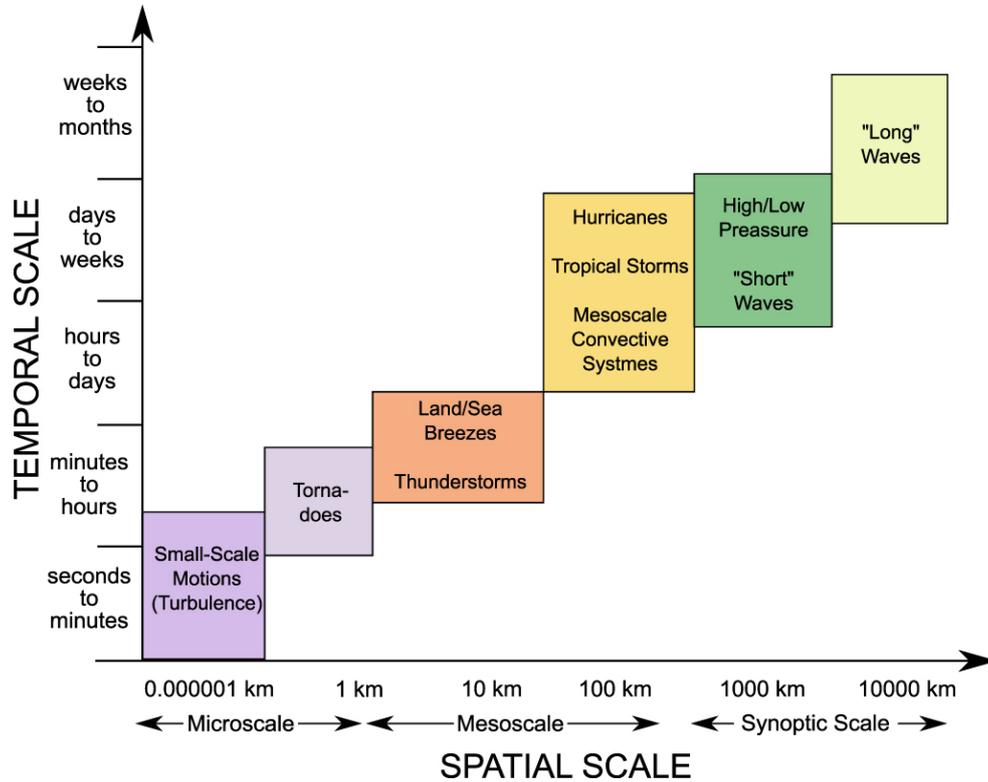


Figure 1.1: A schematic diagram illustrating typical time and length scales of various weather phenomena. The figure is sourced from Brisch and Kantz (2019), and the original publication can be accessed at <https://doi.org/10.1088/1367-2630/ab3b4c>.

redistribute moisture (Kain and Fritsch, 1993; Stensrud, 2007). On the other hand, high-resolution (kilometer-scale) models can realistically resolve convective storms, hence they are referred to as convection-permitting models (Clark et al., 2016). An accurate simulation of precipitating clouds is critical, as it affects the dynamic solution while also influencing other physical processes such as radiation (Boucher et al., 1995).

Bauer et al. (2015) note a significant improvement in NWP forecast skill over the past 40 years, with a gain of about one day per decade in the accuracy of forecasts within the 3 to 10-day range. Although the progress in NWP over the past few decades has been indisputably successful, inevitably, every forecast will eventually encounter its predictability limit. This limit has been reported as two weeks for synoptic-scale (Zhang et al., 2019; Selz et al., 2022) and a matter of hours for convective-scale forecasts (Hohenegger and Schar, 2007), respectively. In the upcoming section, I will discuss the predictability limit and its connection to atmospheric flows.

1.1.2 Predictability of atmospheric flows

According to Lorenz (1996), atmospheric predictability can be further divided into two categories: (1) practical predictability, which refers to the best achievable prediction based on current forecast techniques such as DA systems and NWP models (Melhauser and Zhang, 2012), and (2) intrinsic predictability, which refers to the best possible prediction using nearly flawless initial states and forecast models (Lorenz, 1996; Zhang et al., 2006; Sun and Zhang, 2016). While the practical predictability may be extended by reducing errors in forecast procedures, the intrinsic predictability represents an inherent predictability limit due to the chaotic nature of the flow and cannot be extended. Even with nearly perfect forecast procedures, tiny and inevitable initial condition errors can amplify rapidly and restrict the predictability of all scales (Lorenz, 1963).

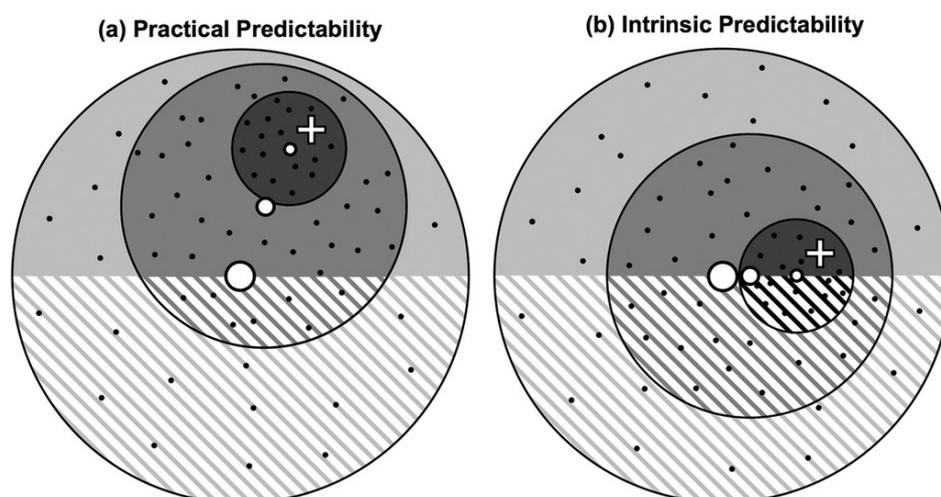


Figure 1.2: An idealized schematic illustrating the distinction between (a) practical and (b) intrinsic predictability. The figure is sourced from Melhauser and Zhang (2012), and the original publication can be accessed at <https://doi.org/10.1175/JAS-D-11-0315.1>. ©American Meteorological Society. Used with permission.

A schematic plot provided by Melhauser and Zhang (2012) delineates the difference between intrinsic and practical predictability. As shown in Fig. 1.2, the black dots represent individual forecasts starting with slightly different initial values, which are referred to as initial uncertainty. The three black circles indicate the gradual reduction of initial uncertainty. The two shaded regions depict two potential outcomes; the white dots denote the mean of forecasts, and the white crosses mark the forecast truth. With the largest initial uncertainty, the forecasts predict that both outcomes are equally likely. Two distinct scenarios occur when initial uncertainty is reduced contingent on whether a system is constrained by practical predictability (Fig. 1.2a) or intrinsic predictability (Fig. 1.2b). Reducing initial uncertainty enhances forecast accuracy in systems constrained by practical predictability, while in systems constrained by intrinsic predictability, reducing

initial errors does not enhance the forecast as the two outcomes remain equally likely.

A clear distinction between “the model predictability of the atmospheric state” and “the predictability of the model state” has been articulated in Bachmann (2019). The former represents the model forecast accuracy when compared to observations, while the latter relies on the perfect model assumption, acknowledging the possibility of other systematic errors in comparison to the truth. To elaborate further, a perfect model experiment assumes that the model processes can realistically describe the atmospheric processes, which holds true to a certain extent (Bachmann, 2019). A perfect model experiment allows for an investigation into how a specific error source influences the evolution of model states, thereby implying its impact on the predictability of actual weather systems. It is important to note that the first study in this dissertation assesses precipitation predictability against true observations. On the other hand, the subsequent two used perfect model experiments to investigate the forecast error growth in convection-permitting EPSs (as introduced later in section 1.1.3). Here, I outline three factors that significantly influence predictability.

Flow-dependence of predictability

Chaos theory explains that nonlinear deterministic equations can lead to entirely random behavior, termed deterministic chaos when these equations are unstable to small initial perturbations (Lorenz, 1963). The atmospheric processes are recognized as nonlinear deterministic systems, and the degree of chaos depends on the characteristic of atmospheric flow (Slingo and Palmer, 2011). For instance, studies have shown that synoptically forced convection exhibits higher predictability with less diverging solutions in the presence of small uncertainties, while weakly forced convection may have large diverging solutions under similar conditions (Keil et al., 2019; Puh et al., 2023). Furthermore, Palmer (1993) suggests that forecasts are expected to encounter a loss of intrinsic predictability during the regime transitions. In this situation, the distribution of forecasts is broad or even multi-modal, and reducing initial uncertainty does not enhance forecast accuracy (like Fig. 1.2b); conversely, the reduction of initial errors enhances forecast accuracy when forecasts possess a dominant solution (like Fig. 1.2a). These examples collectively demonstrate that varying levels of predictability are anticipated within different flow regimes.

Scale interaction and predictability

The predictability of model states is reported to vary as a function of scale, and this scale-dependent predictability is often assessed through the saturation of Kinetic Energy (KE) error in spectral space (Durrán and Gingrich, 2014; Selz and Craig, 2015; Selz, 2019; Selz et al., 2022). The following paragraph explains the concepts illustrated in Fig. 1.3, which is sourced from Durrán and Weyn (2016). The two panels depict the KE errors

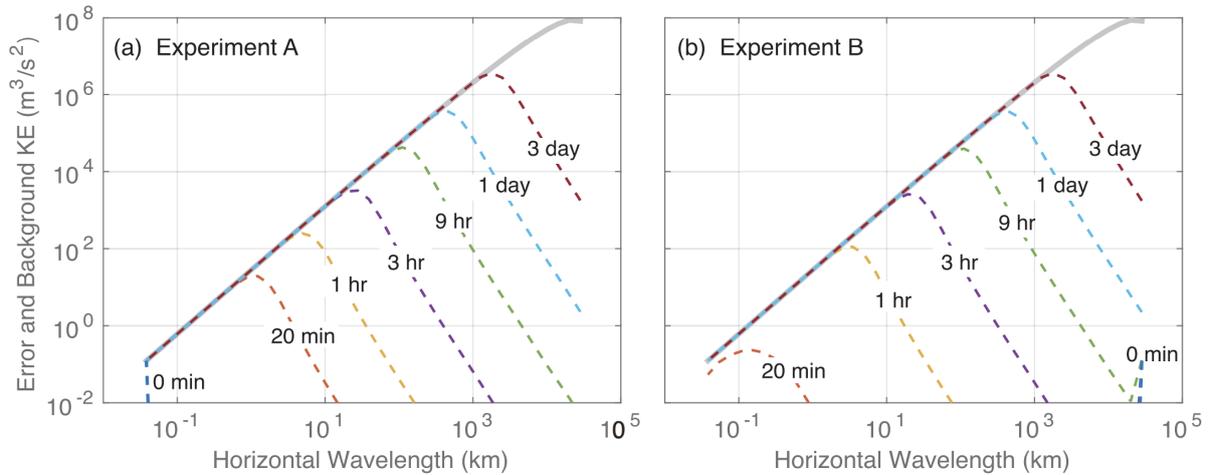


Figure 1.3: Amplitude of background KE (the thick gray line) and KE error (dashed lines) as a function of horizontal scale for (a) Lorenz’s experiment A and (b) Lorenz’s experiment B at various lead times. The figure is sourced from Durran and Weyn (2016), and the original publication can be accessed at <https://doi.org/10.1175/BAMS-D-15-00070.1>. ©American Meteorological Society. Used with permission.

as a function of horizontal scale at various lead times. When the KE error at a specific wavelength matches the background KE (the thick gray line), predictability is lost at that wavelength due to the error saturation. Figure 1.3a illustrates a butterfly effect with upscale error growth originating from small-scale errors, while Figure 1.3b challenges this effect by emphasizing that large-scale errors with equal absolute magnitude can rapidly propagate downscale. These downscaled errors will then cascade upscale, similar to those originating from small-scale errors but holding even more significance at the short lead times. Despite their different initial error propagation, both situations exhibit similar upscale error growth after three hours, suggesting that predictability is inherently lost faster in small scales compared to larger scales regardless of error sources (Durran and Weyn, 2016). For example, the KE errors saturate at horizontal scales smaller than approximately 20 km and 1000 km at 3 hours and 1 day, respectively (Fig. 1.3).

The physical mechanisms of upscale error growth have been explained by a three-stage error growth model proposed by Zhang et al. (2007). During the first stage, moist convection, specifically the latent heat release, governs the error growth processes at the convective scale (Hohenegger et al., 2006). In the second stage, mesoscale error saturation occurs through geostrophic adjustment. Following this, error growth is driven by baroclinic instability on a larger scale. In this dissertation, a similar KE diagnostic is used to estimate the scale-dependent predictability of different weather systems. For convective precipitation with scales of tens of kilometers, the rapid loss of predictability within hours can be anticipated.

Resolution-dependence of predictability

The predictability of the model state is also linked to the horizontal resolution of NWP models. Compared to coarse-resolution forecasts, error growth is about ten times larger at the convection-permitting resolution, primarily attributed to convective instabilities as mentioned earlier (Hohenegger and Schar, 2007). Also, convection-permitting simulations exhibit a higher degree of nonlinearity. In general, the tangent-linear approximation can remain valid for 54 hours in synoptic-scale forecasts; however, it fails much faster in high-resolution forecasts (Ansell and Mass, 2006; Hohenegger and Schar, 2007). This indicates that nonlinear processes occur earlier as model resolution increases, making accurate predictions inherently more challenging in such cases.

The predictability challenges outlined above collectively hinder the accuracy of a single deterministic forecast. For example, accurately predicting convective events within a lead time of a day is already challenging due to their high nonlinearity and small-scale nature. To cope with limited predictability, it is advantageous to employ EPSs that provide a spectrum of potential outcomes by considering sources of forecast uncertainty.

1.1.3 Ensemble forecasting and sources of uncertainty

In response to the limited predictability of atmospheric flow, operational centers have developed EPSs to consider sources of forecast uncertainty and produce probabilistic forecast products (Molteni et al., 1996; Toth and Kalnay, 1997; Winkler et al., 2018). In this context, the term “ensemble” denotes the entire set of forecasts, while each forecast within it is termed an “ensemble member”. To generate ensemble members, it is essential to sample sources of forecast uncertainty as realistically as possible. It’s worth noting that ensemble members are often assumed to have equal probabilities (Unger et al., 2009; Fraley et al., 2010). Figure 1.4 offers a schematic visualization of the ensemble approach sourced from Slingo and Palmer (2011). As illustrated, ensemble members may be derived from initial condition uncertainty or uncertainty in the NWP model. It is important to mention that for regional EPSs, additional lateral boundary condition uncertainty must be considered (Nutter et al., 2004; Kühnlein et al., 2014; Zhang et al., 2023).

Lorenz (1963) employed a Monte Carlo approach to represent errors in the initial conditions. Due to the nonlinear nature of atmospheric motion, these small initial perturbations will amplify and lead to diverging forecasts as the lead time increases. In current practice, the initial uncertainty is generally not randomly sampled but is estimated from ensemble DA systems (Schwartz et al., 2022; Chen et al., 2023) or optimized to ensure sufficient error growth, such as using the singular or breeding vector techniques (Descamps and Talagrand, 2007; Magnusson et al., 2008). Despite efforts to enhance

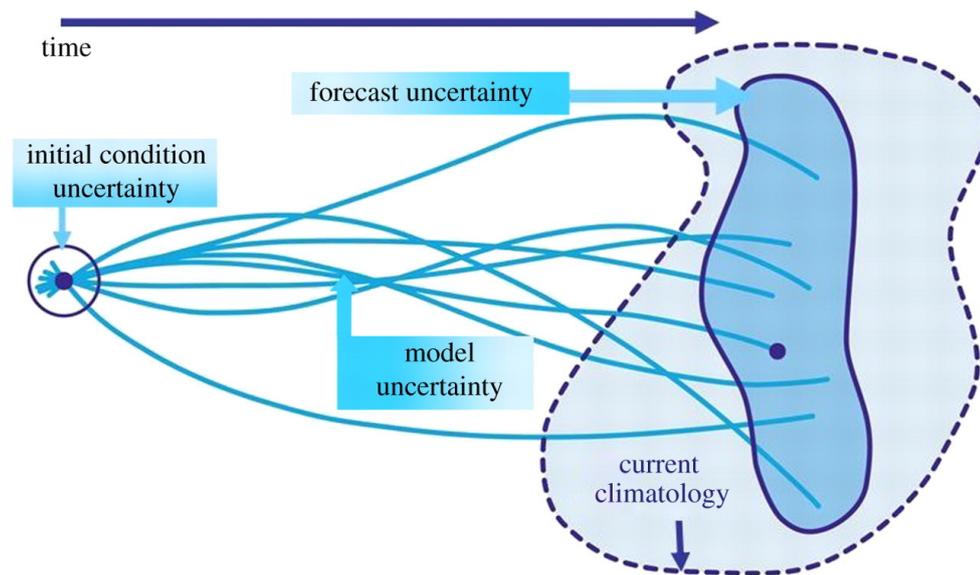


Figure 1.4: A probabilistic forecasting method that considers uncertainties stemming from initial conditions and model formulations. The figure is sourced from Slingo and Palmer (2011), and the original publication can be accessed at <https://doi.org/10.1098/rsta.2011.0161>.

ensemble spread from initial errors, the EPSs are frequently under-dispersive (Buizza et al., 2005; Berner et al., 2009). The under-dispersion, defining a condition where forecast uncertainty is smaller than forecast errors, suggests that forecasts tend to be overly confident, with inadequate error growth.

Further studies addressing the under-dispersion have pinpointed that uncertainty in model formulation can lead to substantial forecast divergence (Buizza et al., 1999; Shutts, 2005; Gebhardt et al., 2011; Bouttier et al., 2012). It is comprehensible that NWP models cannot be perfect because of the necessary approximations in the primitive equations, the numerical discretization, and the physics parameterizations. There are several approaches to incorporate model uncertainty, such as (1) multi-model ensembles (Ziehmann, 2000; Weigel et al., 2008; Roberts et al., 2019), (2) multi-physics ensembles (Berner et al., 2011; Gallo et al., 2017; García-Ortega et al., 2017), and (3) ensembles that sample uncertainty within a single physics scheme (Bouttier et al., 2012; Jankov et al., 2017, 2019; Frogner et al., 2022). Although violating the equal probable rule of the ensemble members, multi-model and multi-physics EPSs have been tested by several operational centers and have demonstrated success (Roberts et al., 2020; Chen et al., 2023). It should be noted that the interdependence of initial and model errors is evident, as model errors are implicitly included in the initial conditions during the DA initialization process (Zeng et al., 2018, 2019).

Convection-permitting EPSs have been reported to be under-dispersive (Romine

et al., 2014; Frogner et al., 2019). One likely cause is the inadequate representation of physical processes (Puh et al., 2023), emphasizing the need for further fundamental research on representing model uncertainty in convection-permitting models. To address this issue, the second study in this dissertation will examine forecast error growth originating from different parameterization schemes. Given the close association of boundary-layer turbulence and cloud microphysics with convective precipitation, the discussion primarily focuses on representing errors and understanding their effects within these processes.

1.2 Atmospheric convection and weather regimes

Severe storms such as tornadoes, thunderstorms, and hailstorms are closely linked to convection (Doswell, 2001, 2014), emphasizing the importance of grasping its dynamics. Convection manifests in diverse forms within the atmosphere, varying between dry and moist (Stevens, 2005), as well as shallow (Atkinson and Wu Zhang, 1996) and deep convection (Stensrud, 2007). One distinction between shallow and deep convection is that the latter is characterized by precipitating clouds that can extend vertically up to the tropopause (Yano and Plant, 2012). This process involves the conversion of water vapor to liquid water, releasing substantial amounts of latent heat that can facilitate the development of intense convective storms (Houze Jr, 2004). This dissertation primarily focuses on deep, moist convection as it frequently evolves into severe convective storms. In this section, I will provide a brief overview of the moist convection dynamics, including relevant measurement indices. Also, this section will introduce types of convection driven by different forcing mechanisms, along with a diagnostic metric for an objective and physically-based regime classification.

1.2.1 Atmospheric instability and relevant indices

The development of convective storms can be explained by atmospheric stability, which hinges on two components: the temperature of the rising/sinking air parcel and the temperature of the stratified environmental air. When an air parcel undergoes vertical motion, its buoyancy is determined by the temperature contrast between these two components. This temperature contrast is closely related to the lapse rate (the rate of temperature change with altitude) in both air parcels and the surrounding environment. In parcel theory, the lapse rates of unsaturated and saturated air parcels conform to the dry and moist adiabatic lapse rates, respectively, and the stability is therefore contingent upon the lapse rate of the stratified environmental air. Both the dry and moist adiabatic processes are reversible, but they differ in whether the air parcel is saturated or not. As illustrated in Fig. 1.5a, the faster the environmental temperature decreases with altitude, the atmosphere tends to become more unstable. In an absolutely stable condition (the blue arrow), air parcels undergoing small vertical displacements will return to their original positions due to negative buoyancy acceleration, making deep convection unlikely as vertical motion faces resistance.

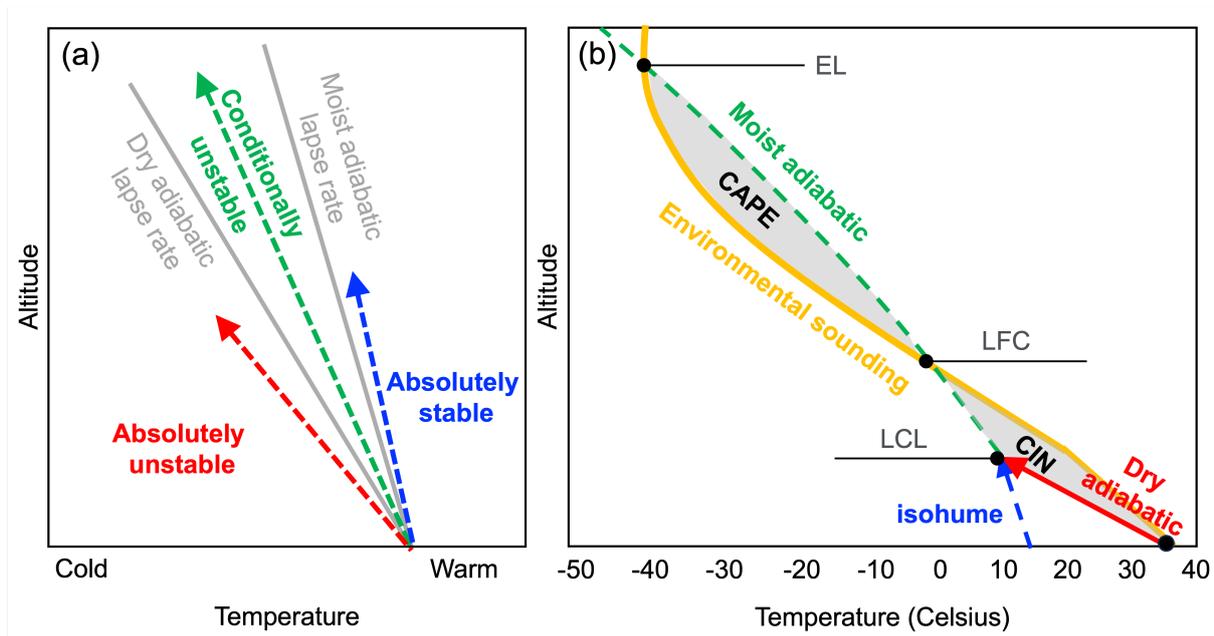


Figure 1.5: (a) A thermodynamic diagram illustrating atmospheric stability. (b) The yellow line illustrates the environmental sounding, with a black circle marking the initial position of an air parcel near the surface. Red solid and blue dashed lines delineate the dry and moist adiabatic lapse rates. Shaded regions denote CIN and CAPE, respectively.

For deep convection to take place, the atmosphere should at least be conditionally unstable, where its lapse rate falls between the moist and dry adiabatic lapse rates (the green arrow in Fig. 1.5a). In this scenario, the buoyancy of an air parcel is contingent upon whether it is saturated or not. As illustrated in Fig. 1.5b, an air parcel rising from the surface will typically undergo the following stages delineated by its position relative to the Lifting Condensation Level (LCL), Level of Free Convection (LFC), and Equilibrium Level (EL):

1. From the surface to the LCL: an unsaturated ascending air parcel cools at the dry adiabatic lapse rate, requiring external forcing to counteract its negative buoyancy. As temperature decreases, the maximum water vapor capacity of air diminishes, ultimately causing the air parcel to reach saturation at the LCL.
2. From the LCL to the LFC: above the LCL, condensation releases latent heat, enabling the parcel to cool at the moist adiabatic lapse rate. It is important to note that the air parcel is typically denser than its surroundings at the LCL and still requires external forcing to counteract negative buoyancy until forced to ascend to the LFC, where its temperature is equal to the environmental temperature.

3. From the LFC to the EL: beyond the LFC, the air parcel cools at the moist adiabatic lapse rate and maintains a higher temperature than its surrounding environment. This positive buoyancy initiates free upward acceleration towards the EL without the need for external forcing.

In this scenario, intense convective storms are more likely to develop. Based on this concept, various indices have been formulated to evaluate the likelihood and intensity of convective storms. Presented here are two key indices: Convective available potential energy (CAPE) and Convective inhibition (CIN) (Markowski and Richardson, 2011; Groenemeijer et al., 2019):

Definition of Convective available potential energy

CAPE quantifies instability in the troposphere by the maximum potential energy that a given parcel could accumulate when ascending from the LFC to the EL. This index is commonly used to estimate the strength of potential convection. As illustrated by the positive area in Fig. 1.5b, CAPE is determined by the vertical integral of positive buoyancy (B). As shown in (1.1), the vertical integration of B is given by the difference between the virtual temperature of the air parcel (T_v) and that of the environment (\bar{T}_v):

$$CAPE = \int_{LFC}^{EL} B dz \approx g \cdot \int_{LFC}^{EL} \frac{T_v - \bar{T}_v}{\bar{T}_v} dz \quad (1.1)$$

As mentioned in Doswell and Rasmussen (1994), neglecting the correction for virtual temperature, which is the temperature a dry air parcel would have if its pressure and density were equivalent to those of a moist air parcel, can be particularly detrimental when the CAPE value is low.

Definition of Convective inhibition

For deep convection to occur, a parcel requires external forcing to counteract the negative buoyancy until it reaches the LFC, at which point the release of CAPE can begin. The energy required to overcome the inhibition can be measured by CIN, as indicated by the negative area in Fig. 1.5b. The larger the area, the higher the CIN value, leading to a lower chance of deep convection. In situations with considerable CAPE and CIN, the large CIN has the potential to suppress convection; however, if external forcing successfully breaks through the inhibition, strong convection can occur. Like CAPE, CIN is defined as the vertical integral of buoyancy but from the surface up to the LFC:

$$CIN = - \int_{surface}^{LFC} B dz \approx -g \cdot \int_{surface}^{LFC} \frac{T_v - \bar{T}_v}{\bar{T}_v} dz \quad (1.2)$$

1.2.2 Equilibrium and nonequilibrium convection

As noted earlier, the initiation of convection requires an external lifting mechanism to overcome the CIN. Numerous forcing mechanisms exist, including convection with synoptic forcing (Doswell and Bosart, 2001; Cannon et al., 2018), orography forcing (Kirshbaum et al., 2018; Schneider et al., 2018), and local thermal forcing (Jucker et al., 2020). Broadly speaking, convection cases can be classified on a continuum with two extremes—equilibrium and nonequilibrium—based on the coupling of convection to external forcing (Done et al., 2006; Keil and Craig, 2011; Zimmer et al., 2011; Flack et al., 2016).

For equilibrium convection, the properties of precipitation (e.g., timing and intensity) are controlled by large-scale destabilization. For this to occur, there should be little collocation between the CAPE and CIN so that convection is free to act immediately in response to CAPE (Done et al., 2006). For example, large-scale uplift can cool air masses and create large amounts of CAPE, which is then rapidly exhausted by convective precipitation and a fast return to equilibrium. Since the CAPE is immediately converted to upward motion, equilibrium convection is typically associated with small CAPE values throughout its life cycle. The predictability of precipitation varies (Done et al., 2012), with equilibrium convection generally being more predictable as it is more constrained by environmental forcing (Keil et al., 2014; Flack et al., 2018). In contrast, nonequilibrium convection is primarily triggered by local processes such as boundary layer turbulence (Hanley et al., 2015; Hirt, 2020) and cold pool lifting (Hirt et al., 2020; Hirt and Craig, 2021). This type of convection occurs mostly in summer, partially due to strong radiative heating. For nonequilibrium convection, CAPE is often co-located with CIN and thus can build up until convection is triggered by overcoming the CIN threshold (Done et al., 2006). Therefore, the properties of nonequilibrium convection largely depend on local processes, such as surface and boundary layer variability, and tend to be less predictable (Keil et al., 2014, 2019; Chen et al., 2024).

Although often used interchangeably in the literature and throughout this dissertation, it's important to note that the distinction between equilibrium and nonequilibrium regimes is not identical to the strong versus weak forcing of convection. There is no standard quantitative definition of strong or weak forcing, but the terms generally reflect the rate at which the atmospheric flow creates instability, for example, through the dynamical forcing of synoptic and mesoscale ascent. Ascent generates CAPE, but also decreases CIN. This makes it easy to trigger convection, which can increase rapidly in response to the creation of CAPE, leading to a rapid adjustment to equilibrium. In contrast, afternoon convection occurring in response to surface heating can be accompanied by a strong temperature inversion at the top of the boundary layer giving a large CIN. Triggering convection is difficult, and the response to increasing CAPE is delayed in a nonequilibrium situation. There is thus a tendency for equilibrium to occur in the presence of strong dynamical forcing, but it is also possible that the convection is, for some time, unable to respond to strong forcing, and the instability rapidly increases. Similarly, there is a ten-

dency for nonequilibrium to prevail when dynamical forcing is weak. However, if surface processes are not dominant, weak equilibrium can occur in response to weak forcing.

1.2.3 Convective adjustment time scale

To measure the degree of coupling between convection and its environment, the convective adjustment time scale was first proposed by Done et al. (2006). This metric can objectively classify convective weather according to the degree of equilibrium between the convection and the forcing processes. To elaborate further, this metric estimates the timescale at which convective instability is consumed by convective heating, which can be compared to the timescale of the processes that create instability. Processes that create CAPE, such as large-scale uplift, evolve significantly over periods of 6 to 12 hours. A small convective adjustment time scale of a few hours indicates equilibrium convection: convective instability is removed quickly by convective heating, and the atmosphere returns rapidly to equilibrium. A large convective adjustment time scale of a day or more indicates that convection is acting too slowly to follow the evolution of the forcing processes and is indicative of nonequilibrium convection.

Given that the convective adjustment time scale mirrors the strength of the coupling between convection and the environment, it can serve as a valuable tool in NWP to differentiate between forecasts with distinct convective behavior (Keil et al., 2014; Flack et al., 2018). One notable application is to differentiate the predictability of convective precipitation, as equilibrium and nonequilibrium regimes typically exhibit varying degrees of predictability. For example, Bachmann et al. (2020), Keil et al. (2020) and Sobash et al. (2023) demonstrate above-average predictability of total area-averaged precipitation in weather situations controlled by synoptic forcing (equilibrium convection).

Another crucial application is utilizing the convective adjustment time scale to delineate regime-dependent forecast sensitivity. Considering the distinct error growth of initial conditions, boundary conditions, and model uncertainties across different weather regimes, categorizing flow situations can facilitate the interpretation of forecast sensitivity to sources of uncertainty. Studies, primarily focused on Europe, have shown that this metric can effectively indicate the flow-dependent impact of specific model components on forecast performance. Here is a summary of the conclusions from the relevant studies:

- Craig et al. (2012) noted that radar data assimilation has a longer-lasting impact in weak forcing regimes.
- Kühnlein et al. (2014) found that model physics perturbations have a larger impact on precipitation spread in weakly forced situations than in strongly forced situations.
- Keil et al. (2014) noted that lateral boundary perturbations grow faster under strongly forced conditions, while model perturbations have a more pronounced effect under weak forcing conditions.

- Zeng et al. (2018) found that the representation of model error using additive noise is more beneficial during strongly forced weather conditions.
- The relative importance of microphysics uncertainties depends on the significance of synoptic forcing (Surcel et al., 2017; Keil et al., 2019). During weak synoptic forcing, the impact of microphysical uncertainty on precipitation is twice as large as its impact in strong forcing and accounts for nearly one-third of the spread resulting from operational initial and boundary condition uncertainties (Matsunobu et al., 2022).
- Puh et al. (2023) demonstrated a systematic benefit of using a physically based stochastic perturbation scheme for convection-permitting forecasts of convection covering three months. In particular, there is an increased ensemble spread of precipitation during weak synoptic forcing conditions, which is much more difficult to detect in statistics that contain all weather regimes.

So far, the convective adjustment time scale diagnostic has been mostly used as a categorical measure to classify certain weather situations into equilibrium or nonequilibrium flow regimes. Studies over Europe have applied threshold values of 3 hours (Keil et al., 2014, 2020) and 6 hours (Kühnlein et al., 2014; Molini et al., 2011; Grazzini et al., 2020). The choice of these values concurs with the statistical results of Zimmer et al. (2011), who use observations in central Europe and find a continuous distribution of these time scales and conclude that a value somewhere between 3 and 12 h clearly distinguishes between different regimes. However, the climatology of the convective adjustment time scale can also vary across regions. For instance, Flack et al. (2016) established a climatology for the British Isles and observed a twofold increase in equilibrium convection when using the same threshold as Zimmer et al. (2011).

In conclusion, this section has discussed convection-related indices (CAPE and CIN), described equilibrium and nonequilibrium convective regimes, and introduced the convective adjustment time scale for regime classification. These discussions offer fundamental insights into the first research topic in this dissertation, which focuses on identifying convective regimes over the CONUS and studying the regime-dependent predictability in convection-permitting forecasts. As the convective adjustment time scale value varies depending on geographical region, season, and the intricacies of computation (e.g., hourly or three-hourly rainfall accumulations, the exact CAPE calculation, horizontal smoothing of input fields, mean or maximum daily values), studies over Europe may not directly apply to different areas. Therefore, this dissertation represents the first systematic study that utilizes this metric to establish the convective regime climatology over the CONUS and explores whether the classification aligns with fundamental physical understanding and reflects regime-dependent precipitation predictability.

1.3 Parameterization of subgrid scale processes

As mentioned earlier, more fundamental studies regarding model error formulations within convection-permitting EPSs are required to address under-dispersion, especially considering the significant uncertainty stemming from unresolved processes. Despite reaching kilometer and sub-kilometer scales, where NWP models are anticipated to resolve more small-scale features, achieving a complete representation of all such processes remains unattainable (Shin and Hong, 2013; Honnert et al., 2020). However, subgrid processes can alter the evolution of the atmospheric motion, requiring proper representation to maintain the model forecast skills (Stensrud, 2007). Broadly speaking, parameterization schemes are formulated to approximate the influence of subgrid processes by utilizing resolved variables rather than directly simulating them. The approximated values from parameterization schemes then function as tendency forcing terms to advance the prognostic equations forward in time (Stensrud, 2007). It is worth mentioning that certain schemes may be called less frequently due to their high computational cost. For example, the long-wave and short-wave radiation parameterization, which estimates solar and thermal radiative transfer and thus controls energy balance, is typically not updated at every time step (Pincus and Stevens, 2013).

Many other subgrid parameterization schemes, including microphysics, Planetary Boundary Layer (PBL), deep convection, shallow convection, gravity wave drag, orography drag, and surface layer, are developed to represent different subgrid processes (Stensrud, 2007). As discussed in section 1.1, activating all schemes is not essential; their specific usage depends on the spatial grid structure. It's worth clarifying that microphysics parameterization is centered on microscale processes within clouds, while convection parameterizations address the collective effects of cloud within a single column and are often phased out at convection-permitting resolution (Weusthoff et al., 2010; Clark et al., 2016). Since this dissertation focuses on convection, attention is given to two parameterization schemes—microphysics and PBL—due to their high sensitivity as reported (Kober and Craig, 2016; Keil et al., 2019; Hirt et al., 2019; Stanford et al., 2019; Barthlott et al., 2022a,b; Matsunobu et al., 2022). These two subgrid processes can directly influence convective precipitation by modulating latent heat release due to phase change and the thermodynamic structure in the lower troposphere, respectively. In the subsequent sections, a brief introduction to the PBL parameterization and the microphysics parameterization will be provided. Additionally, stochastic parameterization, a widely adopted method for sampling model uncertainty, is introduced.

1.3.1 Planetary boundary layer parameterization

The Planetary Boundary Layer refers to the lower part of the troposphere directly influenced by the properties of Earth's surface (Stensrud, 2007). Its depth typically ranges from several dozen meters to kilometers, exhibiting significant diurnal varia-

tion due to changes in ground temperature caused by solar radiation (Markowski and Richardson, 2011). The PBL responds rapidly to surface forcing, such as momentum, heat, and mass, transmitting their effects between the surface and the atmosphere via transport processes. Turbulence, also known as eddies, dominates vertical transport within the PBL, exerting a significant impact on tropospheric wind, temperature, and humidity (Stull, 1988). Consequently, it can influence the development of various weather systems, including convective thunderstorms (Hirt, 2020). However, as small eddies cannot be explicitly captured within the spatial-temporal scales used in most NWP models, the overall effect of these unresolved eddies on resolved quantities is parameterized.

Reynolds averaging is a fundamental concept used to mathematically link resolved and unresolved processes, but it introduces the turbulence closure problem as illustrated below. The Reynolds decomposition separates flow variables into their mean (resolved) and fluctuating (unresolved) components ($\phi = \bar{\phi} + \phi'$). Another assumption made is that the average of the fluctuating part (ϕ') is zero, while the average of their products does not disappear (correlated). Equation (1.3) exemplifies the application of Reynolds decomposition to the meridional (v) and vertical (w) wind components, while (1.4) illustrates that averaging the product of two Reynolds averaged variables (I) is equal to adding the product of their means (II) with the average of their fluctuation products (III).

$$uw = (\bar{u} + u')(\bar{w} + w') = \bar{u}\bar{w} + \bar{u}w' + \bar{w}u' + u'w' \quad (1.3)$$

$$\underbrace{\overline{uw}}_{\text{(I)}} = \overline{\bar{u}\bar{w}} + \overline{\bar{u}w'} + \overline{\bar{w}u'} + \overline{u'w'} = \underbrace{\overline{\bar{u}\bar{w}}}_{\text{(II)}} + \underbrace{\overline{u'w'}}_{\text{(III)}} \quad (1.4)$$

In numerical solutions of the primitive equations, discretization occurs in both time and space, with only the mean of each grid box and time step being of interest. In line with this concept, resolved-scale model equations can be obtained through Reynolds averaging the primitive equations. Equation (1.5) illustrates a resolved-scale prognostic equation for Reynolds-averaged quantities. It highlights that to predict the grid-scale mean (I), it is necessary to consider the mean contribution of unresolved motions (II) to close the equation. Although models cannot explicitly simulate eddies, their unresolved effect, quantified here by vertical turbulent momentum flux (II), acts as a tendency forcing to step the grid-scale motion forward in time. It's important to note that the derivation of (1.5) yields horizontal flux terms, but they are simplified here. While it's possible to explicitly predict the fluxes terms, it requires dealing with third-order unknown terms. Further predicting higher-level terms results in the issue of closure (Stensrud, 2007).

$$\underbrace{\frac{\partial \bar{u}}{\partial t}}_{\text{(I)}} = -\left[\bar{u}\frac{\partial \bar{u}}{\partial x} + \bar{v}\frac{\partial \bar{u}}{\partial y} + \bar{w}\frac{\partial \bar{u}}{\partial z}\right] + f\bar{v} - \frac{1}{\rho}\frac{\partial \bar{p}}{\partial x} - \underbrace{\frac{\partial \overline{u'w'}}{\partial z}}_{\text{(II)}} \quad (1.5)$$

Since the explicit prediction of all small eddies is unattainable, PBL parameterization works to approximate turbulent fluxes, including kinematic momentum, heat, and moisture fluxes. In most cases, PBL schemes only approximate vertical turbulent fluxes under the assumption of horizontally homogeneous conditions, as shown in (1.5). Based on how they estimate the vertical flux terms, PBL schemes can be classified as either local or non-local and further categorized based on their order. Local schemes estimate flux terms using resolved variables at nearby points, while non-local schemes typically utilize multiple vertical grid points, often covering the entire PBL. The former tend to perform better under stable stratification, while the latter are reported to work better under unstable conditions, as they incorporate the effects of larger eddies (Stensrud, 2007). The order of PBL schemes represents the level at which the cascade of unknowns is truncated. First-order closure entails equations for resolved variables with parameterized covariance terms, while second-order closure involves equations for both resolved variables and covariance terms with parameterized third-order terms. In general, higher-order closure schemes tend to provide more accurate results than lower-order ones, but they require more computational resources (Troen and Mahrt, 1986; Stensrud, 2007).

Following the operational settings, this dissertation includes the use of the Yonsei University (YSU) and Mellor–Yamada–Nakanishi–Niino (MYNN) schemes. The former is a first-order closure scheme that accounts for non-local mixing (Hong et al., 2006), while the latter is a 1.5-order local closure scheme with prognostic second-order turbulent kinetic energy (Nakanishi and Niino, 2006). Given the sensitivity of numerous weather phenomena like deep convection to the PBL evolution, the second study in this dissertation explores forecast error growth arising from parameterized PBL processes, recognizing that eddies can only be partially resolved.

1.3.2 Microphysics parameterization

The accurate representation of cloud microphysical processes, which influence thermodynamic feedback and precipitation, is another crucial yet challenging topic in the NWP (Choudhury and Das, 2017; Bao et al., 2021; Köcher et al., 2023). Broadly speaking, simulating cloud microphysical variables like cloud water, cloud ice, rain, snow, and graupel involves two major components: (1) integrating these variables directly at the grid scale, and (2) parameterizing their subgrid-scale processes. At the grid scale, NWP models predict mass mixing ratios for microphysical variables (x), the exact number of which depends on the scheme employed. As illustrated in Stensrud (2007) and outlined below, the prognostic equations can be simplified as:

$$\frac{\partial q_x}{\partial t} = -M_{advection}(q_x) + M_{turbulence}(q_x) + (S_1 + S_2 + S_3 + \dots), \quad (1.6)$$

where q_x , $M_{advection}$, $M_{turbulence}$ represents the prognostic variable, advection effect, and turbulence effect, respectively. The terms S_1 through S_3 represent sources and sinks

arising from subgrid-scale cloud processes that need to be parameterized, including the formation, growth, and dissipation of hydrometeor (sometimes non-hydrometeors) particles after conversions between categories. Since microphysics processes modulate the latent heat release, they also force the prognostic temperature and can influence large-scale dynamics (Stensrud, 2007).

Typically, microphysics parameterization schemes describe hydrometeors through their Particle Size Distributions (PSDs), which quantifies the number of particles per unit size interval. Based on how they represent the PSD, microphysics schemes can be categorized into either bulk or bin models. Bulk models predict the total mass mixing ratios and describe their PSDs through predetermined functions, such as Gamma or Marshall–Palmer distributions depending on hydrometeor types. In contrast, bin models partition the PSD into discrete size bins and predict mass mixing ratios and number concentrations for each bin separately (Khain et al., 2015). While bin models are designed to calculate microphysics with high accuracy, their computational cost has hindered their operational use.

Bulk microphysics schemes can be further categorized based on their microphysical moments. Single-moment schemes predict only mass mixing ratios, while double-moment schemes additionally predict their number concentrations. The latter is expected to perform better for vast regions with varying hydrometeor characteristics, but the additional freedom it provides also introduces uncertainty (Jin and Baik, 2023). Here, the formulation of the Thompson double-moment aerosol-aware bulk microphysical scheme (Thompson and Eidhammer, 2014) is used for further illustration. It predicts the mass mixing ratios of cloud ice, cloud water, snow, rain, and graupel, the number concentrations of rain, cloud ice, and cloud water, along with the cloud condensation nuclei and ice nuclei. Then, their PSDs are calculated using gamma functions with slope and intercept parameters derived from the predicted mass mixing ratios and number concentrations (assigned for the one-moment and predicted for the two-moment species), respectively. Except for the snow, all water and ice species are assumed to conform to the generalized gamma distribution given as follows (Thompson et al., 2008):

$$N(D) = N_0 D^\mu e^{-\lambda D} \quad (1.7)$$

where D is the droplet diameter, μ is the shape parameter, N_0 is the intercept parameter, and λ is the slope of the distribution. As shown above, the microphysics parameterization relies heavily on empirical and theoretical hypotheses, leading to considerable uncertainty in the forecast. Moreover, fixed parameters can vary and are not universally applicable (Tao and Li, 2016; Matsunobu et al., 2022). To explore the associated forecast uncertainty, this dissertation also discusses forecast error growth resulting from uncertainty within the microphysics parameterization.

1.3.3 Stochastic parameterization

As mentioned earlier, improving the model error representation is critical to address under-dispersion in convection-permitting EPSs. In earlier times, operational centers often utilized multi-model or multi-physics EPSs to represent model uncertainty and have demonstrated their efficacy. Over the past two decades, there has been a growing trend of incorporating spatial-temporal correlated random noise to represent uncertainty in NWP models. This technique, known as stochastic parameterization, has been demonstrated to alleviate the under-dispersion of EPSs at different centers (Buizza et al., 1999; Bouttier et al., 2012; Jankov et al., 2017). Unlike multi-model and multi-physics EPSs, the stochastic approach upholds the equal probability of ensemble members, an attribute that is essential for equitable evaluation and post-processing techniques. As depicted in Berner et al. (2017), stochastic processes can further reduce systematic model bias through noise-induced rectification as shown in Fig. 1.6. Figure 1.6a depicts that small noise confines the system to the deeper well, yielding an unimodal Probability Density Function (PDF) of states (Fig. 1.6b). Further increasing the noise amplitude can induce a transition to the secondary well (Fig. 1.6c), producing a bimodal PDF (Fig. 1.6d), hence altering both variance and mean. Figure 1.6e-h exemplify how even within a single potential, bimodal PDFs and changes in the mean state can arise due to multiplicative or state-dependent noise.

In general, model uncertainty can be categorized into two groups: (1) uncertainty originates from unresolved processes due to finite grid size, and (2) uncertainty arises from our incomplete understanding of atmospheric phenomena, such as empirically-assigned parameters (Puh et al., 2023). The former is frequently sampled by adding perturbation a posteriori, while the latter is often applied a priori. Examples are provided below.

Early-developed stochastic schemes commonly introduce uncertainty a posteriori. Two widely used schemes are the Stochastically Perturbed Parameterization Tendency (SPPT) and the Stochastic Kinetic Energy Backscatter (SKEB) schemes (Buizza et al., 1999; Shutts, 2005). Both schemes provide a bulk representation for multiple sources of model error, yet they are based on different physics principles. The SPPT scheme perturbs net parameterization tendencies to address the need for sampling subgrid-scale variability as resolution increases (Berner et al., 2017). On the other hand, the SKEB scheme perturbs the stream function and potential temperature tendencies, acting as a means of randomly re-injecting some of the dissipated energy at the grid scale. This approach intends to address model uncertainty arising from unresolved processes and their interactions with resolved flows. In contrast to bulk representations, process-level stochastic schemes that introduce perturbation a posteriori have also been adopted to represent model uncertainty closer to its source (Wastl et al., 2019; Lupo et al., 2022; Zhao and Torn, 2022). In particular, unresolved PBL processes often influence convective initiation, especially when local mechanisms are crucial in overcoming convection inhibition. To tackle this subgrid-scale uncertainty, the Physically based Stochastic Perturbation (PSP) scheme perturbs tendencies from the PBL parameterization scheme, recognizing that

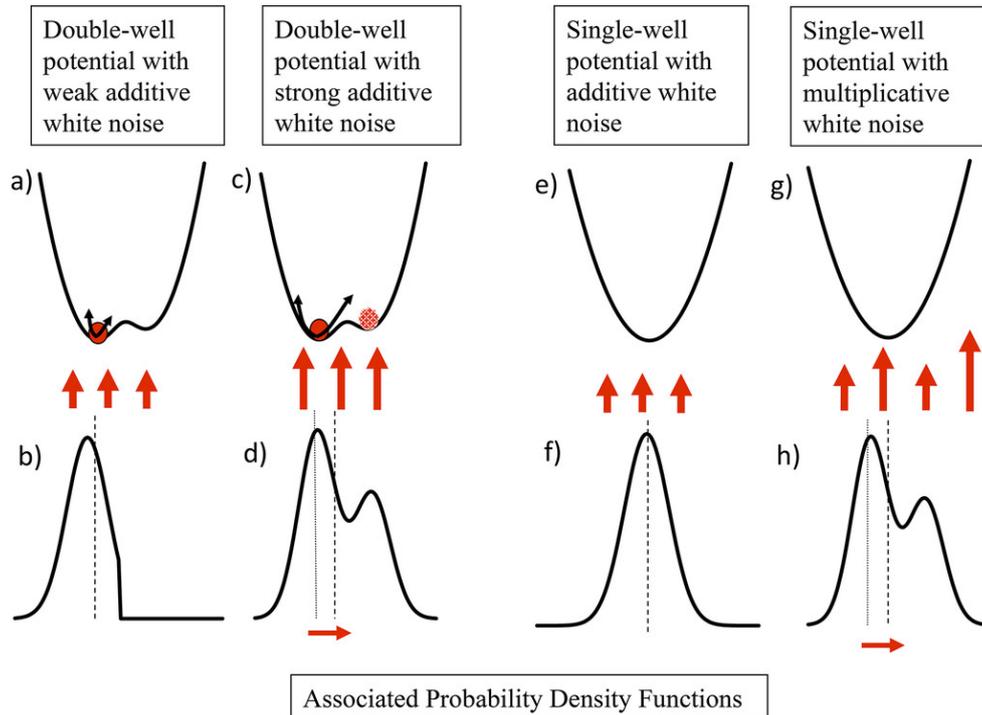


Figure 1.6: System characterized by (a), (c) double-potential or (e),(g) single-potential well. (b), (d), (f), and (h) are their associated PDFs when encountering forcing. The figure is sourced from Berner et al. (2017), and the original publication can be accessed at <https://doi.org/10.1175/BAMS-D-15-00268.1>. ©American Meteorological Society. Used with permission.

eddies can only be partially resolved (Kober and Craig, 2016; Hirt et al., 2019; Puh et al., 2023). While successful in practice, some posteriori schemes have been criticized for violating conservation principles (Leutbecher et al., 2017; Lang et al., 2021).

Recently, there has been an increasing emphasis on stochastic schemes that incorporate uncertainty a priori. For instance, the Stochastic Perturbed Parameter (SPP) approach is formulated to address uncertainty stemming from our incomplete understanding of physical processes, such as uncertain parameters. Specifically, this method introduces variability to uncertain model parameters at the level of individual physical processes, such as the PBL (Jankov et al., 2017, 2019), microphysics (Thompson et al., 2021; Ollinaho et al., 2017), and radiation (Ollinaho et al., 2017). Unlike schemes that apply stochastic perturbations directly to computed physics tendencies, the SPP schemes can maintain the internal consistency and conservation principles of the physical parameterization schemes (Lang et al., 2021). The performance of different SPP schemes has been assessed in various operational systems, showcasing its potential as an option for representing model uncertainty in EPSs (Ollinaho et al., 2017; Lang et al., 2021; Frogner et al., 2022; McTaggart-Cowan et al., 2022).

To sum up, this section presents microphysics and PBL parameterizations because of their notable impact on convective precipitation. Additionally, stochastic parameterization schemes designed with distinct theoretical foundations are introduced. These discussions provide the research background for the second study in this dissertation, which investigates regime-dependent forecast error growth in a convection-permitting model, with errors stemming from unresolved PBL and microphysics processes. The former will be sampled by the process-level posterior PSP scheme, while the latter is represented by a priori SPP scheme. Given the need for more fundamental studies on model error representations in convection-permitting resolution, this study can contribute by elucidating how these two processes impact forecast error growth in different regimes.

1.4 Summary, research goals, and outline

In summary, Chapter 1.1 offers an overview of NWP models, discusses the predictability limits, and introduces the concept of ensemble forecasting. Next, Chapter 1.2 introduces deep convection and its related indices such as CAPE and CIN, discusses equilibrium and nonequilibrium convection, and explains their classification. Lastly, Chapter 1.3 introduces the PBL and microphysics parameterizations, which are crucial for convection prediction. Additionally, stochastic parameterization schemes designed to address different types of model errors are presented. These discussions provide the research background for this dissertation, which has an overarching goal to **“improve the comprehension of regime-dependent predictability and forecast error growth within kilometer-scale models”**. To achieve the overarching goal, this dissertation identifies three main areas of focus as detailed below.

- The first study aims to objectively classify convective regimes across the Contiguous United States (CONUS) domain using the convective adjustment time scale. To systematically evaluate its performance, this study utilizes the operational High-Resolution Rapid Refresh (HRRR) forecast datasets to establish convective adjustment time scale climatology over this region. Through a climatological analysis, this study will investigate whether regime-dependent precipitation predictability is evident in historical convection-permitting forecasts. A central research question addressed in the first study is:

How to effectively classify convective regimes over the vast CONUS domain and can this classification reflect physical understanding and regime-dependent predictability? – discussed in Chapter 2

- Using the objective regime classification outlined in the first study, the second study aims to investigate regime-dependent predictability and forecast error growth within a

convection-permitting EPS. Specifically, it aims to explore how uncertainty in the PBL and microphysics processes influences error growth behavior across various weather regimes. In this study, the central research question is articulated as:

Does forecast error growth vary significantly across different weather regimes and with different stochastic parameterizations in a convection-permitting model? – discussed in Chapter 3

- This third study is a collaborative effort with Kuo et al. (2023), which explores the impact of Data Assimilation (DA) on convection prediction. Utilizing the DA analyses from Kuo et al. (2023), this study investigates the predictability of a convective precipitation event linked with a Meiyu front. Specifically, the study aims to examine how random errors from initial conditions and model processes influence short-range (0 to 12 hours) precipitation forecasts. Although the conclusions may be less conclusive due to the use of a single case, this study contributes to the overarching goal by addressing:

What is the impact of initial and model uncertainty on precipitation predictability and forecast error growth in a convection-permitting model? – discussed in Chapter 4

In summary, this dissertation includes a climatological study, an investigation into error growth from various model errors, and an exploration of error growth from initial condition errors. They all focus on convection-permitting models and collectively cover a variety of weather systems, facilitating the discussion of regime dependency. While individual chapters delve deeper into several research questions, as further elaborated in each chapter, the three studies collectively advance our understanding of how predictability and forecast error growth in kilometer-scale NWP models vary when considering different weather systems and error representations. In addition to the key research findings presented in Chapters 2 to 4, the subsequent structure of this thesis is as follows: Chapter 5 provides the summary and conclusions, and Chapter 6 includes the appendix.

1.5 Publications

This dissertation is composed of chapters that have either been published or are prepared for publication in the following four works:

1. **Chen, I.-H.**, J. Berner, C. Keil, Y.-H. Kuo, and G. Craig, 2024: Classification of Warm-Season Precipitation in High-Resolution Rapid Refresh (HRRR) Model Forecasts

over the Contiguous United States. *Monthly Weather Review*, **152** (1), 187–201.

This publication focuses on regime classification across the CONUS domain. Its content is integrated into Chapter 1, Chapter 2, and Chapter 5 of this dissertation.

2. **Chen, I.-H.**, J. Berner, C. Keil, G. Thompson, Y.-H. Kuo, and G. Craig, 2024: To which degree do the details of stochastic perturbation schemes matter for convective-scale and mesoscale error growth? *Submitted to Monthly Weather Review (In review)*.

This work examines the regime-dependent forecast error growth resulting from different stochastic perturbations. Its content is integrated into Chapter 1, Chapter 3, Chapter 5, and Chapter 6 of this dissertation.

3. Kuo, Y.-H., J. Sun, Y. Zhang, **I.-H. Chen**, Y. Ho, and J.-S. Hong, 2024: Impact of Assimilating GNSS Radio Occultation Data on the Prediction of a Squall Line associated with a Mei-Yu front: Part 1. Data Assimilation Study. *To be submitted*.

I am a coauthor involved in the case analysis and visualization of this study. Chapter 4 of this dissertation includes several figures generated from the research data.

4. **Chen, I.-H.**, Y.-H. Kuo, C. Keil, J. Berner, and G. Craig, 2024: Impact of Assimilating GNSS Radio Occultation Data on the Prediction of a Squall Line associated with a Mei-Yu front: Part 2. Predictability Study. *To be submitted*.

This work explores how initial and model perturbations influence the predictability of a heavy rainfall event. Its content is integrated into Chapter 1, Chapter 4 and Chapter 5 of this dissertation.

Chapter 2

Characterization of Convective Regimes and Their Predictability

2.1 Background and overview

As mentioned in Chapter 1.2, the convective adjustment time scale diagnostic has been employed to classify the convective weather regime. Most of these studies concentrated on Continental Europe, and together they suggest that the exact threshold value is contingent upon factors such as geographical region, season, and the complexities of its computation. To the author's knowledge, there is no systematic evaluation of this metric covering the CONUS domain; therefore, this dissertation represents the first comprehensive examination of the convective adjustment time scale climatology over the CONUS, including the regional and seasonal dependencies. To account for geographical variances, the domain is partitioned into multiple subdomains, and the climatology of equilibrium and nonequilibrium regimes are studied regionally. The timing and amount of precipitation are linked with the convective adjustment time scale to examine if the classification is physically meaningful, and its relation to the precipitation forecast skills is also investigated. Given the expansive geographic area of the CONUS domain, such a study could offer valuable insights and serve as a basis for comparison with studies conducted in different parts of Europe.

This chapter is organized as follows. Chapter 2.2 describes the methodology, including the definition of the convective adjustment time scale, details of subdomains, and forecast dataset. Chapter 2.3 and Chapter 2.4 present the convective adjustment time scale climatology and investigate its relation to forecast skill, respectively. The summary and discussion are given in Chapter 2.5.

2.2 Methodology

2.2.1 Definition of the convective adjustment time scale

The convective adjustment time scale (τ_c) measures the timescale at which convective instability is removed by convective heating. Following Done et al. (2006), the τ_c is defined as:

$$\tau_c = \frac{CAPE}{\frac{dCAPE}{dt}}, \quad (2.1)$$

where the change of instability, $dCAPE/dt$, is derived from the column-integrated rate of latent heat release implied by the precipitation rate:

$$\frac{dCAPE}{dt} = \frac{1}{3600} \frac{L_v}{c_p} \frac{g}{\rho_0 T_0} p_r, \quad (2.2)$$

Here, p_r is the hourly precipitation rate ($mm\ h^{-1}$), L_v the latent heat of evaporation ($J\ k_g^{-1}$), c_p the specific heat capacity of air at constant pressure ($J\ k_g^{-1}\ K^{-1}$), g the gravity acceleration, T_0 the reference temperature (K), and ρ_0 the reference density ($k_g\ m^{-3}$). From (2.1) and (2.2), the final formulation of τ_c is derived as

$$\tau_c = \beta \cdot \frac{CAPE}{\frac{dCAPE}{dt}} = \beta \cdot \frac{c_p \rho_0 T_0 CAPE}{L_v g p_r}, \quad (2.3)$$

In (2.3), p_r ($mm\ s^{-1}$) is derived from the hourly accumulated precipitation ($mm\ h^{-1}$). As in previous studies, we use an effective convective adjustment time scale by introducing an extra factor $\beta = 0.5$ in (2.3). The factor is set to account for the reduction of CAPE by other aspects of convection, such as cooling and drying of the boundary layer. As shown in (2.3), τ_c is proportional to CAPE and inversely proportional to the precipitation rate.

There are multiple definitions of CAPE, which include the influence of different depths, such as surface CAPE, mixed-layer CAPE, and most unstable CAPE. Through comprehensive sensitivity tests, we chose to use the instantaneous surface CAPE directly from the model output since it is most closely linked to the 24-hour QPF skill over the CONUS. While τ_c ought to represent local features, it should be smooth enough not to capture variability on scales smaller than the spacing between convective clouds (Craig et al., 2012). In this study, we adopt the same smoothing method as Keil and Craig (2011) and Flack et al. (2016), namely a Gaussian kernel with a half-width of 15 km which is applied to both precipitation and CAPE fields. In general, the regime classification is

insensitive to the details of the smoothing methods, as would be expected as long as the smoothing scale is large enough to average out the effects of individual clouds while not being so large that synoptic forcing features are significantly smoothed. To neglect the impact of very light rainfall and avoid abnormal jumps between convective regimes, we further introduce a precipitation threshold of 1 mm h^{-1} after smoothing as Keil and Craig (2011).

In this study, we utilized hourly τ_c values to establish the τ_c climatology for each respective subdomain. The hourly τ_c values represent domain averages for each subdomain and are reported only if more than a hundred grid points have a valid τ_c value according to the defined criteria. For each case, the daily τ_c values were obtained by averaging the hourly τ_c values, and their correlation with the skill of precipitation forecasts is subsequently investigated.

2.2.2 Description of subdomains, forecast, and observation dataset

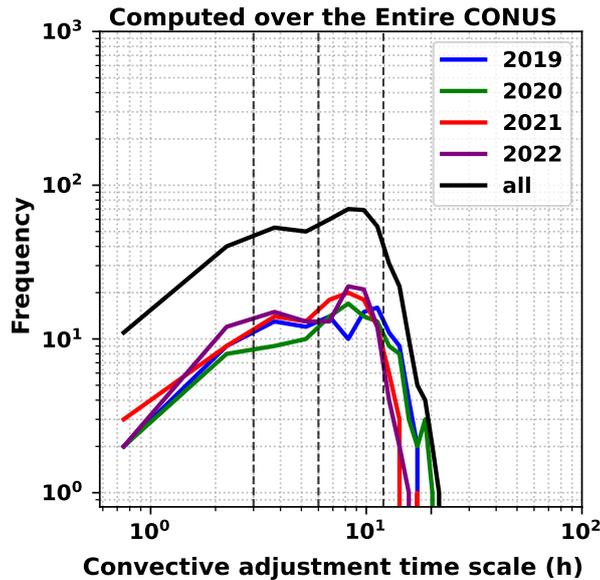


Figure 2.1: Frequency distributions of the convective adjustment time scale shown annually (colors) and 4y-mean (black) for the entire CONUS domain. The dashed vertical lines indicate 3h, 6h, and 12h, respectively.

As shown in Fig. 2.1, the τ_c values computed across the entire CONUS domain generally fall within the range of 1 to 10 hours, with rare occurrences of extreme values. This behavior is primarily caused by averaging over the vast CONUS domain, generally encompassing various weather systems simultaneously. For this reason, this study partition

the CONUS domain into several subdomains to account for regional differences. Guided by the 14 verification regions (<https://www.wpc.ncep.noaa.gov/rgnscr/verify.html>) used at the National Oceanic and Atmospheric Administration (NOAA), several small subdomains with similar climatology are combined. This leaves us with four northern and four southern subdomains roughly divided by the 40°N latitude (Fig. 2.2).

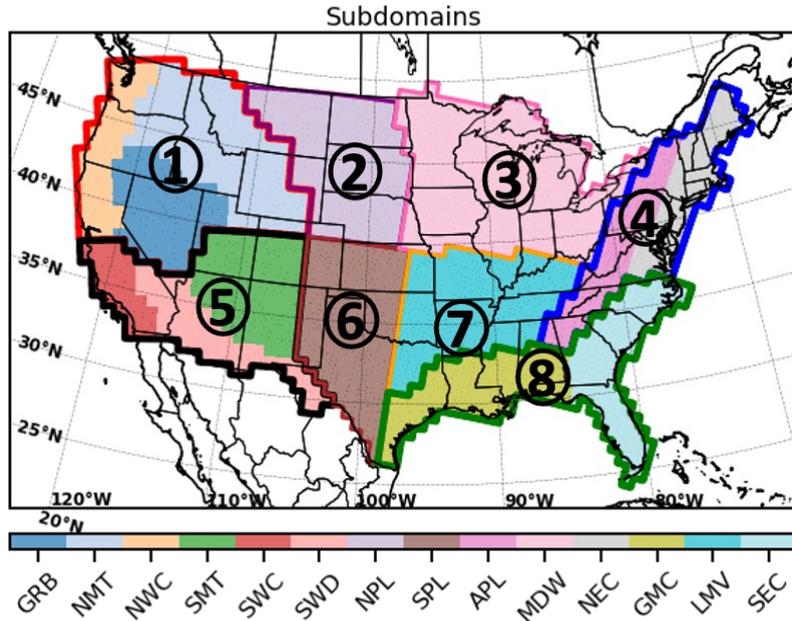


Figure 2.2: Verification and diagnostic subdomains used in the NOAA model QPF verification by regions (shaded) and this study (contour). The black numbers denote the subdomain indexes.

This study uses historical forecast datasets from NOAA’s operational HRRR system (Dowell et al., 2022), which forecast covers the CONUS domain. This convection-permitting HRRR system has been operational since 2014 to provide forecast guidance for rapidly evolving weather. It is built on the non-hydrostatic Weather Research and Forecasting (WRF) model (Skamarock et al., 2008) with a 3 km horizontal grid spacing that explicitly resolves convective storms. It has 51 vertical levels covering from near-surface (roughly from 8 m) up to 15 hPa. It assimilates observations on an hourly basis using the Gridpoint Statistical Interpolation (GSI) Hybrid Three-Dimensional Ensemble-Variational (3DEnVar) DA and with enhancements for radar, land and cloud assimilation, all described in Dowell et al. (2022). Note that the model is running without a deep convective parameterization, which is crucial for this study since such schemes often prescribe a convective adjustment time scale determining how quickly CAPE is consumed.

In this study, the warm season, including May, June, July, and August (MJJA), τ_c climatology is derived from data covering the period 2019 to 2022. The four

years are chosen since HRRRV3 (operational in 2019 and 2020) and HRRRV4 (operational in 2021 and 2022) provide forecasts of up to 36 hours. In this study, the HRRR forecasts were retrieved from the NOAA open data platform (<https://console.cloud.google.com/marketplace/details/noaa-public/hrrr?pli=1>). The 12 UTC run is selected since we aim to cover the whole life cycle of summertime convection. The first three-hour forecasts are excluded from τ_c computation to mitigate the impact of model spin-up. That is, the 4-year MJJA hourly forecast dataset was composed of +3 to +27 h forecast initialized at 12 UTC daily. The three-hour forecast from the 12 UTC run corresponds to 11 am Local Time (LT) at CONUS East (Eastern Daylight Time, UTC-4) to 08 am LT at CONUS West (Pacific Daylight Time, UTC-7).

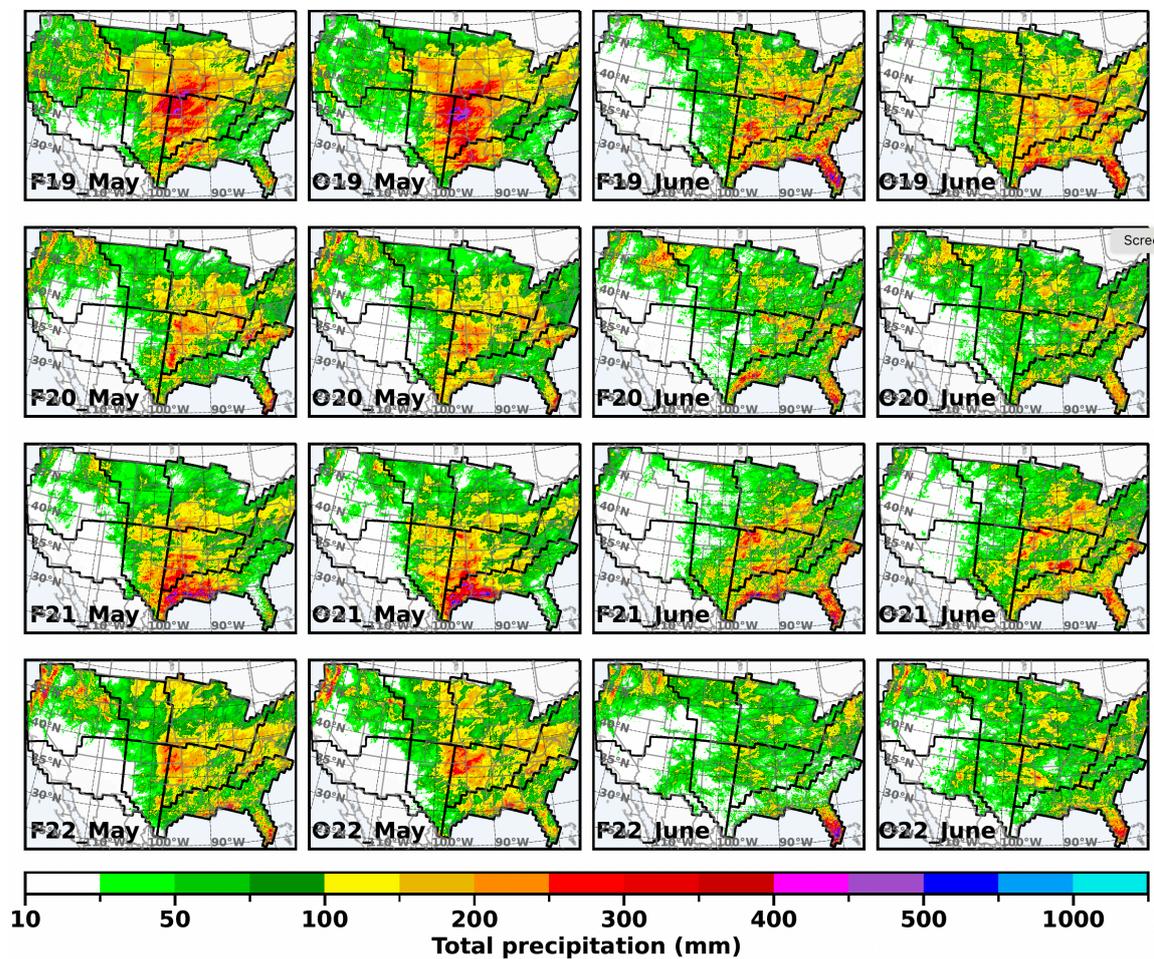


Figure 2.3: Monthly total precipitation (shaded, mm) in May and June from 2019 to 2022 derived from the HRRR 3-27h forecasts (denoted as F) and NCEP StageIV products (denoted as O). Rows are the result of different years, and columns are the result of forecasts (the first and the third columns) and observations (the second and the fourth columns) at different months. Black contours are the subdomain boundaries used in this study.

To ensure that the HRRR system produces a realistic rainfall climatology, we first compare the 3-27 h precipitation forecast to the National Centers for Environmental Prediction (NCEP) stage IV Quantitative Precipitation Estimation (QPE) (Nelson et al., 2016). As shown in Fig. 2.3 and Fig. 2.4, the composed monthly forecast dataset captures the precipitation amount and structure realistically (compare also to Liu et al. 2017). Heavy precipitation (> 200 mm) covering much of the Great Plains is apparent in May on average over these 4 years in both observations and in the HRRR forecast precipitation. In June, the intensity of precipitation systems decreases and shifts toward the east of CONUS, particularly toward the southeast coast. Here, the rainfall associated with hurricanes is excluded by masking precipitation within 350 km from the hurricane center (defined by best track) partly since hurricanes can bring long-lasting heavy rainfall that is not related to the convective rainfall studied here. In this study, we have identified 18 tropical systems and excluded associated precipitation, primarily over eastern CONUS. Note that this method assumes that the model forecasts have a realistic hurricane track compared with the best track.

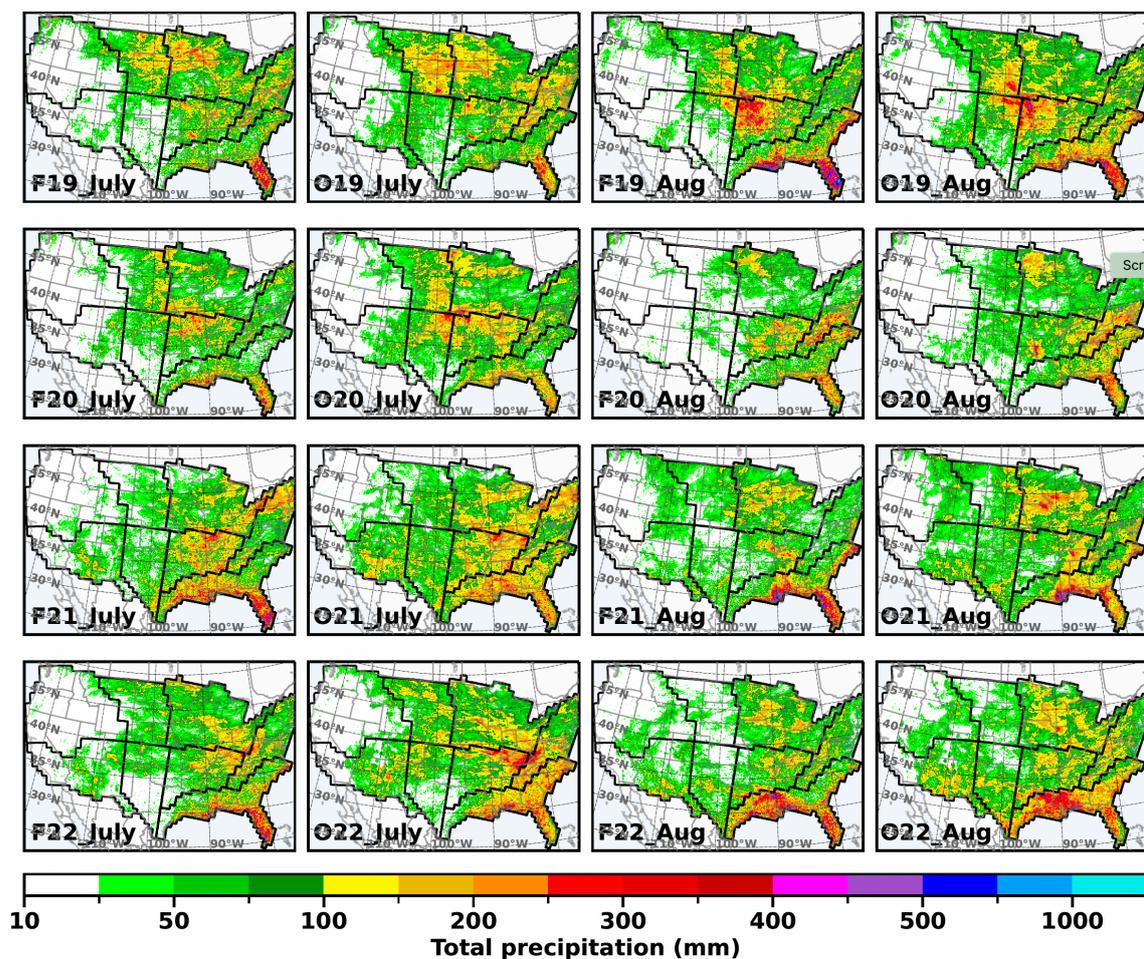


Figure 2.4: Same as Fig.2 but for July and August.

2.3 Convective regime classification

Two cases are selected to illustrate the link between convective precipitation, synoptic forcing, and the resulting convective adjustment time scale for different weather regimes.

2.3.1 Cases illustrating equilibrium and nonequilibrium convection

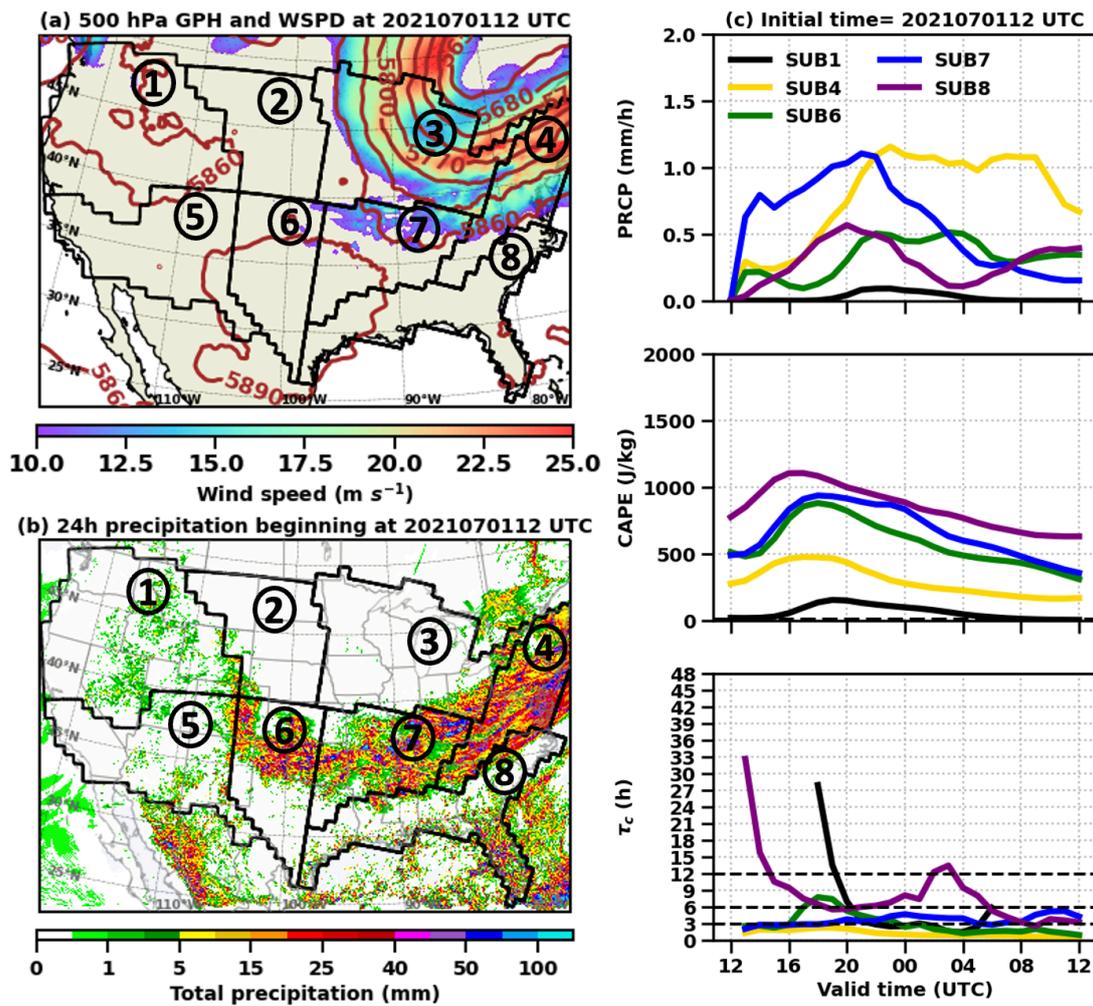


Figure 2.5: (a) 500 hPa geopotential height (contour, m) and wind speed (shaded, $m s^{-1}$) valid at 1200 UTC 01 July 2021. (b) 24h accumulated precipitation (shaded, mm) beginning at 1200 UTC 01 July 2021. Black contours are subdomain boundaries. (c) Time series of hourly precipitation (upper, mm), CAPE (middle, $J kg^{-1}$), and convective adjustment time scale (lower, hours) averaged over different subdomains. The dashed line in the convective adjustment time scale (τ_c) plot marks the 6h threshold between equilibrium and nonequilibrium convection.

The first case initialized at 1200 UTC 01 July 2021, features precipitation events with clear synoptic forcing, where equilibrium behavior is expected (Fig. 2.5). The 24h accumulated precipitation shows that there is an elongated precipitation system located across subdomains 4, 6, and 7 (Fig. 2.5b). This precipitation system is located downstream of a 500 hPa trough (Fig. 2.5a), associated with temperature, moisture, and vorticity advection providing a favorable environment for precipitation since it is conducive to synoptic scale ascent (Doswell, 2001). Precipitation within subdomains 1 and 8 is more scattered and rather weakly forced compared with the major rainband.

The precipitation time series (Fig. 2.5c) shows that precipitation continues for 24 h in subdomains 4, 6, 7, and 8. In contrast, precipitation in subdomain 1 occurs at around 16 LT, exhibiting the lowest intensity among all the subdomains considered. The major rainband initially develops within subdomain 7 and subsequently moves into subdomain 4. While being part of the major rainband, the precipitation is less intense in subdomain 6 than in subdomains 4 and 7. In contrast, precipitation in subdomain 8 exhibits a clear diurnal pattern, implying a thermally-forced trigger mechanism. Since low-level moisture is generally higher in southern subdomains, these regions typically attain larger CAPE values than the northern subdomains. The instability in this case is modest, with $\text{CAPE} \leq 1000 \text{ J Kg}^{-1}$. The CAPE values do not vary strongly throughout the forecast period, suggesting that CAPE does not accumulate and is converted immediately to convective heating. For subdomains 4, 6, and 7, the τ_c values remain low, consistent with equilibrium convection in a strong forcing environment. In contrast, subdomains 1 and 8 have initially larger τ_c values suggesting they are situated within weaker synoptic flow, which agrees with the general flow situation depicted in Fig. 2.5a and the precipitation texture shown in Fig. 2.5b. The rapid decrease of τ_c in these cases would indicate an adjustment to equilibrium. Note that there are no τ_c values at the beginning of the forecast due to low precipitation rates (see chapter 2.2). In conclusion, this case illustrates an equilibrium rainband that is tightly coupled to the governing flow, showing low τ_c values within affected subdomains. Also, this case gives an example of the simultaneous presence of various weather situations across the entire CONUS domain that are only distinguishable by their τ_c values when splitting CONUS into subdomains.

The second case, initialized at 1200 UTC on 27 July 2021, features convective precipitation under weak synoptic forcing (Fig. 2.6). In contrast to the trough pattern in the previous example, this day is dominated by a high-pressure system, which corresponds to sinking drier air and, thus, a more stable atmosphere (Fig. 2.6a). We focus on the four southern subdomains (5-8) that are far from the synoptic wave pattern to the north. These regions feature rather scattered precipitation, which is likely forced by surface heating, and indeed show a clear maximum in precipitation rates in the afternoon (Fig. 2.6c). In subdomain 8, the influence of land-sea breezes is apparent, leading to stronger, more organized precipitation that starts in the early morning. The combination of broad subsidence in the high-pressure area, combined with surface heating, leads to a strong CIN that inhibits convective initiation and large CAPE values build-up

(Fig. 2.6c). CAPE increases until convection starts to remove the instability. The τ_c values reflect the buildup and slow removal of CAPE, with values exceeding 6 hours for almost all subdomains and times. This is the behavior expected for nonequilibrium convection, where convection is inhibited and cannot act fast enough to remove CAPE. The exception is subdomain 5, where strong orography is present, and CAPE and τ_c values remain low. In this mountainous region, even relatively weak convection is sufficient to remove the instability as fast as it is created, leading to equilibrium conditions.

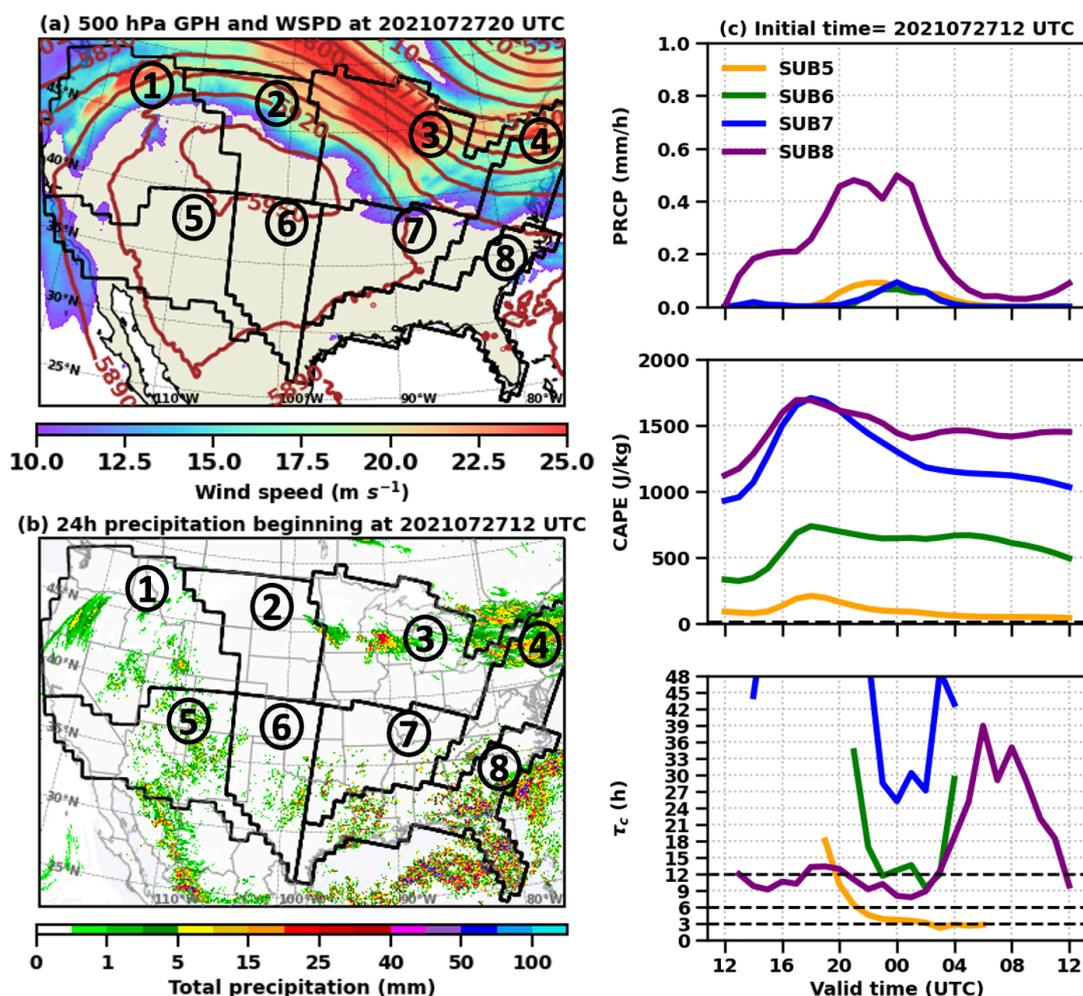


Figure 2.6: (a) 500 hPa geopotential height (contour, m) and wind speed (shaded, $m s^{-1}$) valid at 2000 UTC 27 July 2021. (b) 24h accumulated precipitation (shaded, mm) beginning at 1200 UTC on 27 July 2021. (c) Time series of hourly precipitation (upper, mm), CAPE (middle, $J kg^{-1}$), and convective adjustment time scale (lower, hours) averaged over different subdomains. The dashed line in the convective adjustment time scale (τ_c) plot marks the 6h threshold between equilibrium and nonequilibrium convection.

The two days highlighted in this section show examples of equilibrium convection, with strong forcing by Rossby waves facilitating favorable conditions for convection that removes CAPE as fast as it is produced, and nonequilibrium convection, with precipitation following the diurnal cycle and convection that is unable to remove CAPE. However, the examples of orographic convection show that simply looking for strong synoptic forcing is not a reliable indication of convective behavior.

2.3.2 Convective adjustment time scale climatology

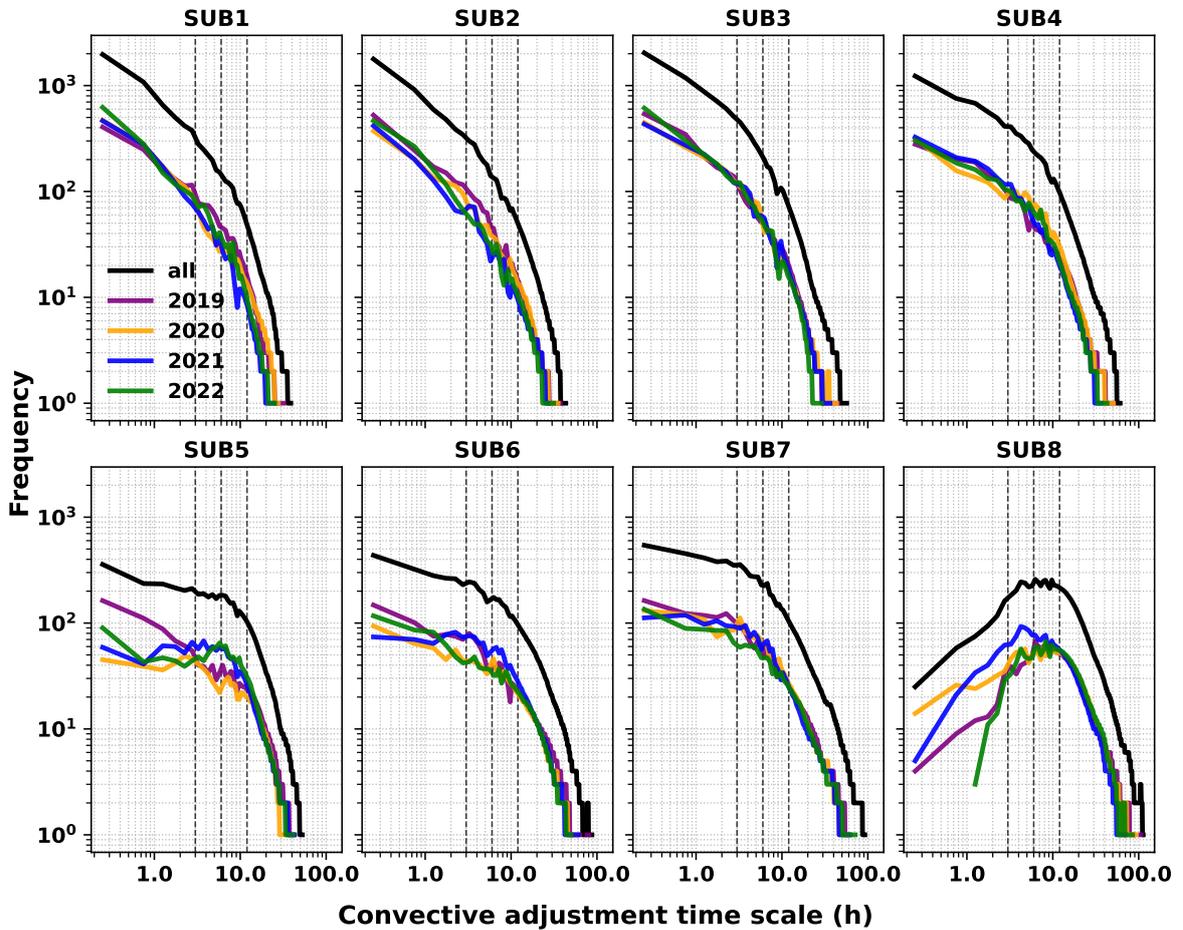


Figure 2.7: Frequency distributions of the convective adjustment time scale shown annually (colors) and 4y-mean (black) for each subdomain. The frequency is computed with a bin size of 0.2h, and lines are smoothed by a Gaussian filter. The dashed vertical lines indicate 3h, 6h, and 12h, respectively.

The climatological frequencies of convective adjustment time scales are separately shown for all subdomains in Figure 2.7. The frequency distribution follows roughly

two power law lines with a scale break between 3 to 12 h. The distributions can be divided into three different categories. The convection in the northern subdomains (1 to 4) is characterized by a smooth distribution with a high occurrence of small τ_c , indicating that equilibrium convection is relatively common in the northern part of CONUS. In comparison, τ_c values larger than 6 h occur more often in the southern subdomains (5 to 8), indicating that nonequilibrium convection is dominant. In subdomain 8, τ_c values shorter than about 6 h are much rarer than in other subdomains.

Previous work investigated the existence of a scale break in the frequency distributions, which would be a clear indication of two distinct convection regimes. While Flack et al. (2016) identified a distinct scale break at three hours over the British Isles, Zimmer et al. (2011) find a continuous distribution and propose $\tau_c = 6h$, a value somewhere between 3 and 12 h to distinguish both weather situations. The results in Fig. 2.7 show strong evidence of a scale break between 6 h to 12 h. Notably, this is the timescale for which the frequency distributions in subdomain 8 peaks. While the convective regimes should be regarded as a continuum with two extremes, we use a τ_c value of 6 hours as the threshold to categorize equilibrium and nonequilibrium convection in this study. To get a sense of the statistical robustness of our results, Figure 2.7 also shows the frequency distribution for each of the four years 2019 to 2022. While there is some year-to-year variability - especially for the southern domains - all findings hold qualitatively.

The fraction of cases attaining τ_c values smaller than certain cumulative thresholds represents another test of robustness (Fig. 2.8). Applying a threshold of $\tau_c = 6h$ classifies 53% of the cases as equilibrium convection for the northern subdomains versus 22% equilibrium for the southern subdomains. This is consistent with the case studies presented earlier, where the northern part of CONUS is frequently under the influence of Rossby waves. The prevalence of equilibrium conditions in the northern subdomains closely aligns with the fraction observed in Germany. During the summer months (June, July, August), approximately 52% of convective precipitation cases in Germany were classified as an equilibrium regime (Zimmer et al., 2011).

Next, we investigate the monthly variation of τ_c for warm-season convection (Fig. 2.9). In all subdomains, May is characterized by shorter τ_c than the other months, leading to a higher proportion of equilibrium conditions. There is a transition in the precipitation pattern over the central plains and southeast coastal region from May to June (see Fig. 2.3). The increase in rainfall over the central Plains in late spring is primarily driven by the Great Plains low-level jet (Wang and Chen, 2009) and the passage of eastward-moving systems (Carbone and Tuttle, 2008). Accordingly, the convection in May is predominantly forced by synoptic flow and contributes to a high percentage of small τ_c values, particularly in the central four subdomains 2, 3, 6, and 7. In June, the total precipitation in the central Plains is reduced due to the development of an upper-level anticyclone (Wang and Chen, 2009).

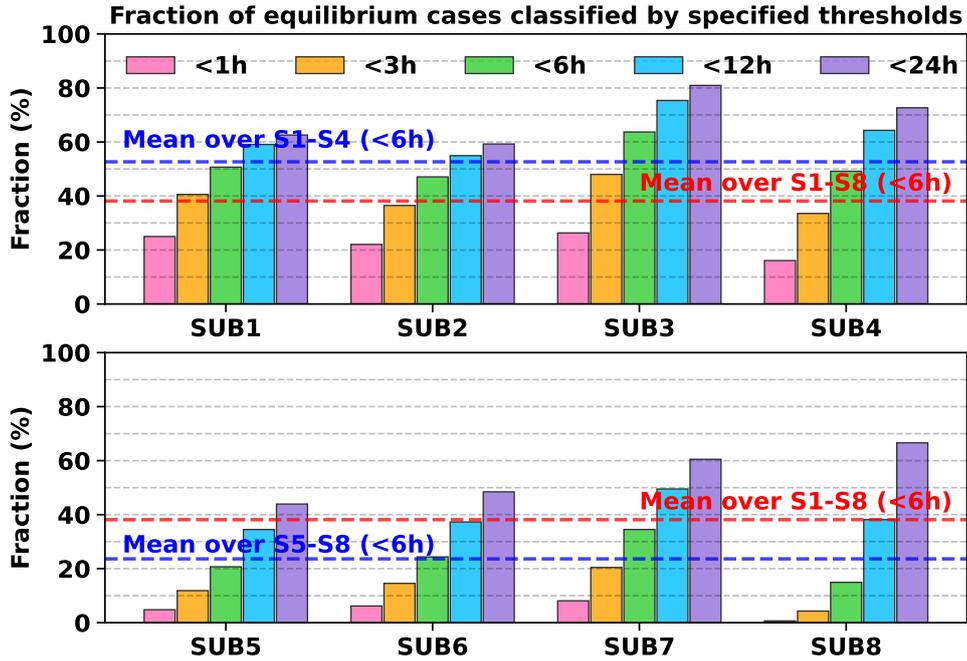


Figure 2.8: Cumulative percentage of convective cases with τ_c less than 1, 3, 6, 12, and 24h for each subdomain. τ_c not meeting the thresholding criteria in section 2a are excluded. The red dashed lines indicate the percentage of equilibrium cases ($\tau_c < 6h$) averaged over all subdomains. The blue dashed line in the upper and lower panel indicates the percentage of equilibrium cases ($\tau_c < 6h$) averaged over northern and southern subdomains, respectively.

In the North, the climatology of convective regimes is similar throughout June to August. In the South, in contrast, there is a strong monthly variation in the frequency of equilibrium conditions, with short τ_c values more common in May and June, and particularly rare in July. Much of this variability may be associated with changes in the mean latitude of the jet stream where synoptic waves propagate, but the seasonal cycle of solar heating may also play a role. For example, the increase in coastal precipitation from May to June is primarily due to the seasonal transition of thermodynamic forcing (Rickenbach et al., 2015).

2.3.3 Diurnal cycle of precipitation under different convective regimes

To further corroborate that our classification mirrors the underlying basic dynamics, we classify all forecasts within each subdomain into two types: equilibrium and nonequilibrium regimes. This classification is done using a τ_c threshold of 6 hours. Next, we analyze the subdomain-averaged rainfall time series to determine the presence of a

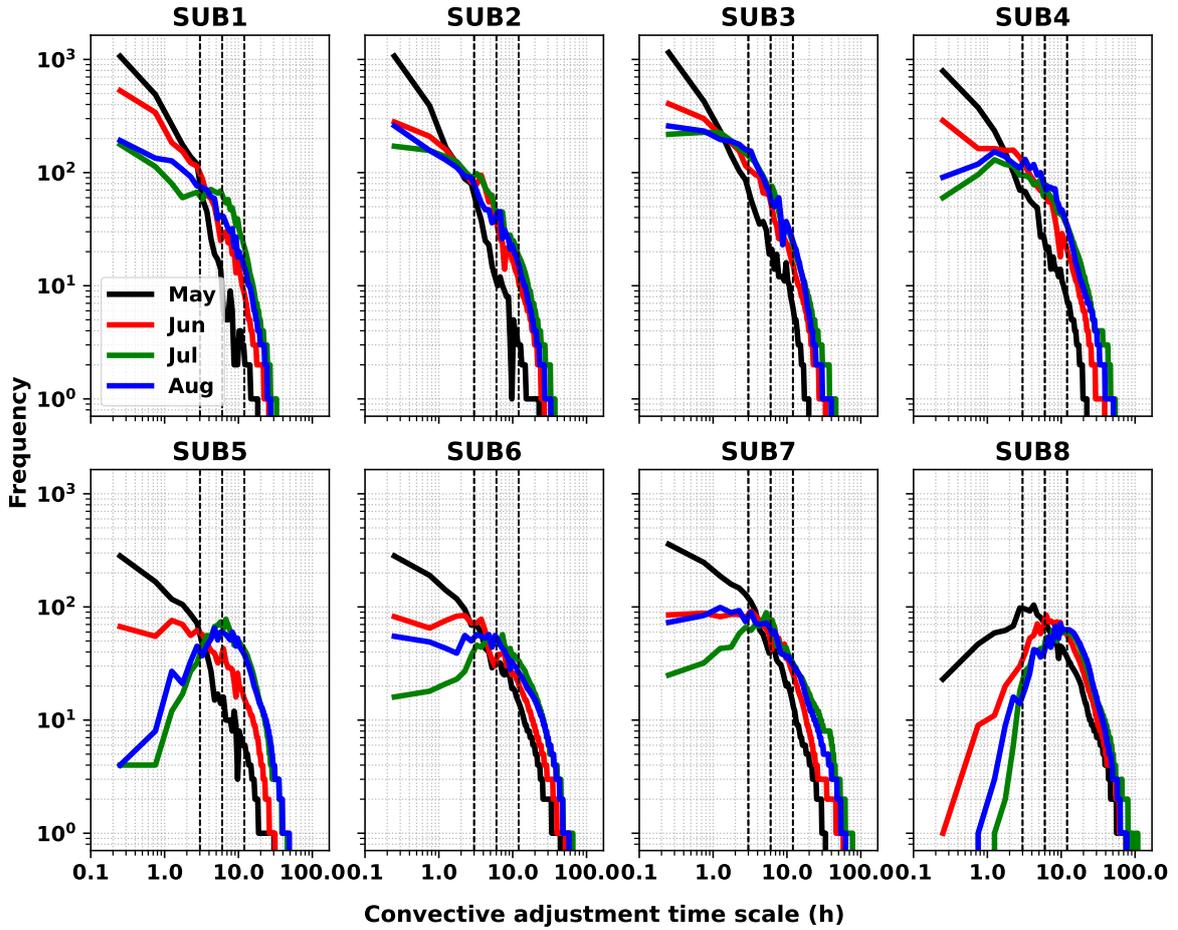


Figure 2.9: Frequency distributions of convective adjustment time scale stratified by calendar month. All other details as in Fig.2.7.

diurnal cycle. We hypothesize that nonequilibrium convection should exhibit a higher proportion of diurnal precipitation. This is based on the assumption that under weak forcing conditions, convection relies more on diurnal solar heating as the triggering mechanism.

To objectively identify diurnal and nondiurnal precipitation systems, we use a supervised learning algorithm, namely the linear Support Vector Machine (SVM), which has been applied to real-world applications and is reported to have significant accuracy (Cortes and Vapnik, 1995). In particular, the SVM algorithm is suitable for solving data regression and classification problems. For data classification, an SVM classifier functions by identifying the optimal decision boundary that maximizes the separation between classes. Since supervised learning is possible only if data are labeled, a hundred diurnal and a hundred nondiurnal precipitation time series are labeled subjectively in this study. The 200 samples with subjectively assigned labels are drawn from the eight subdomains, with an effort to distribute them as evenly as possible. For the SVM classifier, we chose to

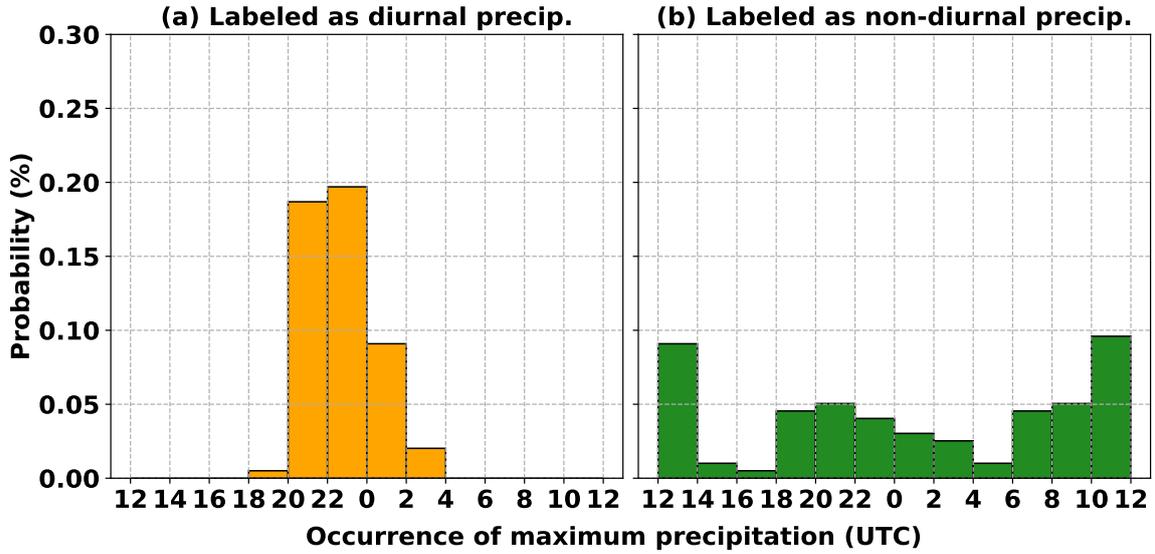


Figure 2.10: Probability of precipitation maxima in a hundred of precipitation time series subjectively labeled as (a) diurnal precipitation and (b) nondiurnal precipitation. The dataset used in this figure consists of 200 time series data.

reduce the number of features by employing a 2-hour precipitation interval. This feature reduction helps maintain a more manageable classifier dimensionality while retaining valuable information capable of capturing time series patterns. As shown in Fig. 2.10, the time series labeled as diurnal precipitation exhibit precipitation maxima primarily during the afternoon hours. In contrast, the time series labeled as nondiurnal precipitation do not display a distinct peak in precipitation timing. Among the sampled datasets, 70% is utilized for training the model, while the remaining 30% is used for testing purposes. The final SVM classifier used in this study has an accuracy of 0.94, a precision (1 - false alarm ratio) of 0.95, and a recall (probability of detection) of 0.96. As shown in Fig. 2.11, the classification result supports the effectiveness of the SVM classifier. Overall, the identified diurnal precipitation time series have precipitation maxima in the afternoon to the early evening. In contrast, the precipitation maxima of identified nondiurnal precipitation time series tend to distribute across various hours.

As shown in Fig. 2.12, the statistical results support that about 61.9% of nonequilibrium convection is associated with a diurnal cycle in precipitation, while only about 33.8% of equilibrium cases. This is consistent with our expectation that the nonequilibrium convection should encompass more diurnal precipitation than equilibrium convection. As the partition between equilibrium and nonequilibrium cases was determined using a τ_c threshold of 6 hours, these results provide evidence that the convective adjustment time scale, along with the 6-hour threshold, can yield a classification that aligns with our physical understanding of the underlying dynamics. While somewhat less conclusive, it is evident that the nonequilibrium precipitation maxima tend to be confined within a shorter

timeframe of a few hours (Fig. 2.12). Based on visual examination, one explanation for the noon peak in equilibrium cases is that precipitation often intensifies in the afternoon during summertime, even under strong synoptic forcing.

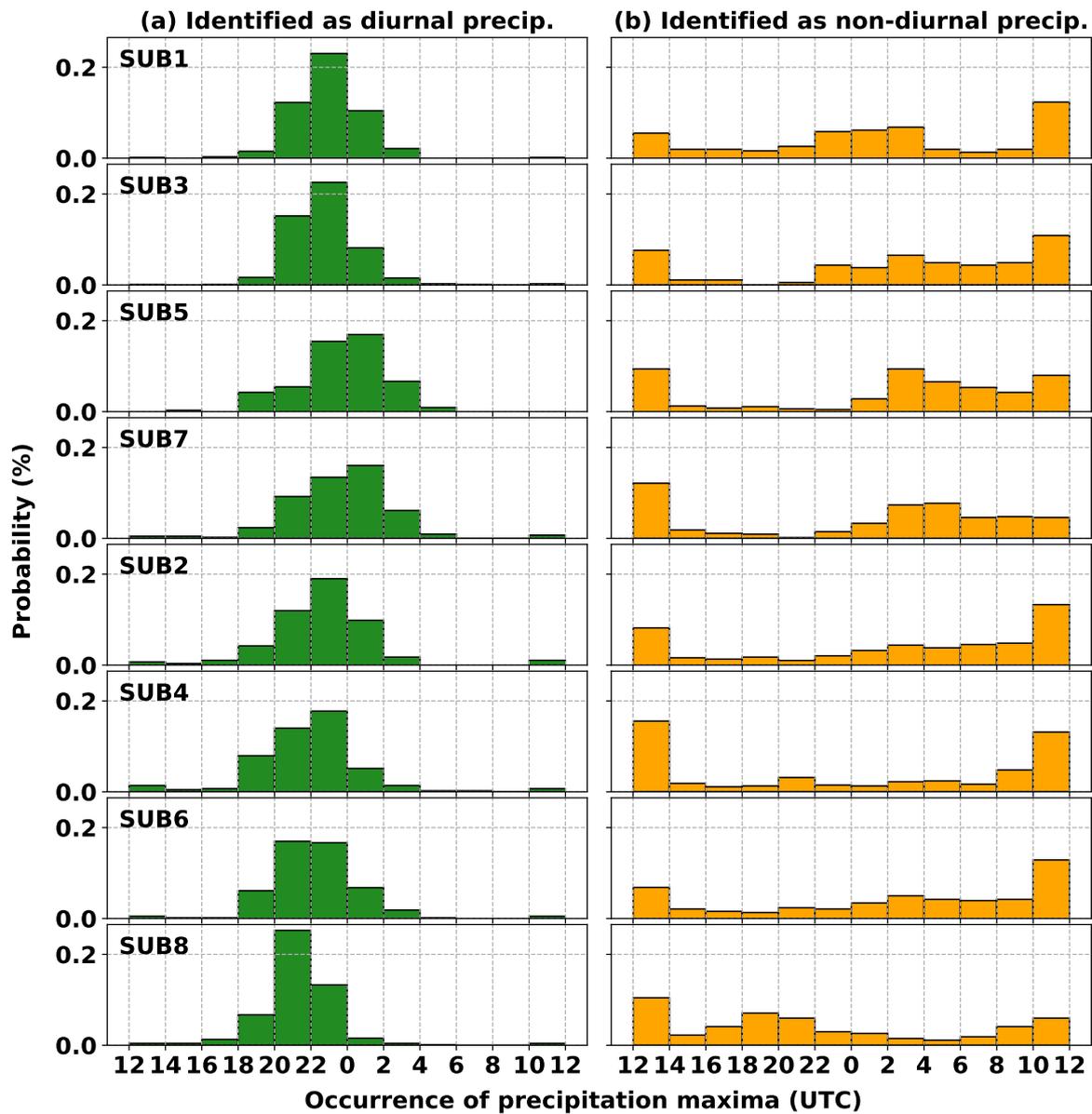


Figure 2.11: Probability of precipitation maxima in precipitation time series classified as (a) diurnal precipitation and (b) non-diurnal precipitation for each subdomain. The figure comprises a complete four-year MJJA dataset excluding those associated with tropical systems, resulting in slight variations in the exact number across different subdomains.

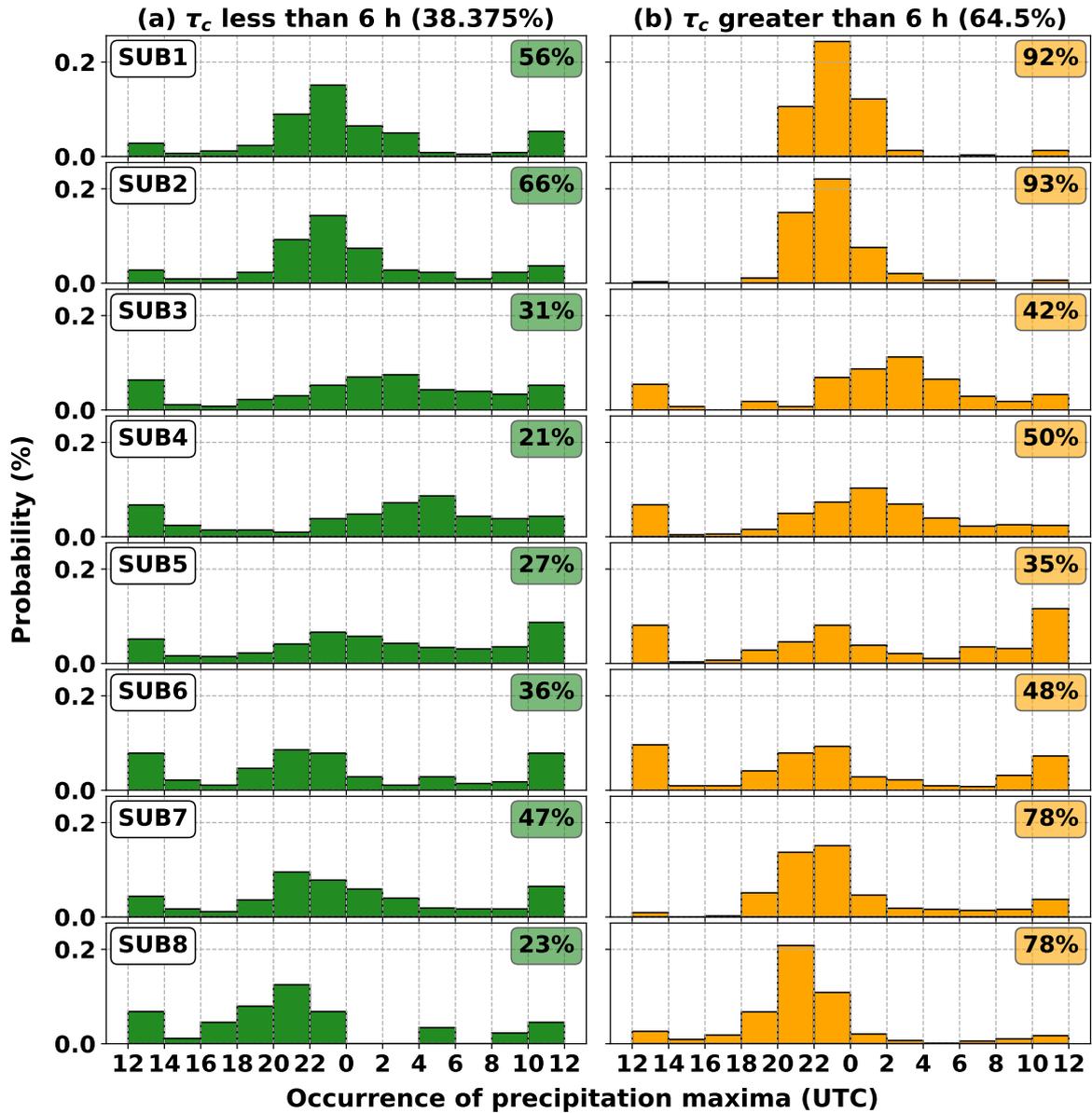


Figure 2.12: Identification of diurnal precipitation pattern in subdomain-averaged rainfall time series. The (a) equilibrium convection and (b) nonequilibrium convection are classified by a τ_c threshold of 6 hours. Bars represent probabilities of precipitation maxima, while annotations correspond to fractions of cases identified as having a diurnal pattern. The figure comprises a complete four-year MJJA dataset excluding those associated with tropical systems, resulting in slight variations in the exact number across different subdomains.

2.4 Convective adjustment time scale as a predictor of forecast skill

One central aim of this study is to assess if the convective adjustment time scale derived from HRRR forecasts provides some indication of their QPF skill over CONUS, as suggested by Keil and Craig (2011) and Keil et al. (2014) for central Europe. To avoid the double penalty problem of rainfall, where the predicted storms are realistic but in a slightly wrong location, we complement the Gilbert Skill Score (GSS) (Doswell et al., 1990) with a neighborhood-based skill score, the Fractional Skill Score (FSS) (Roberts and Lean, 2008), to assess the forecast quality of 24-hour accumulated precipitation. The GSS, also called equitable threat score, is modified from the threat score to account for the correct hits due to random chance. This verification is performed using the Model Evaluation Tools (Brown et al., 2021).

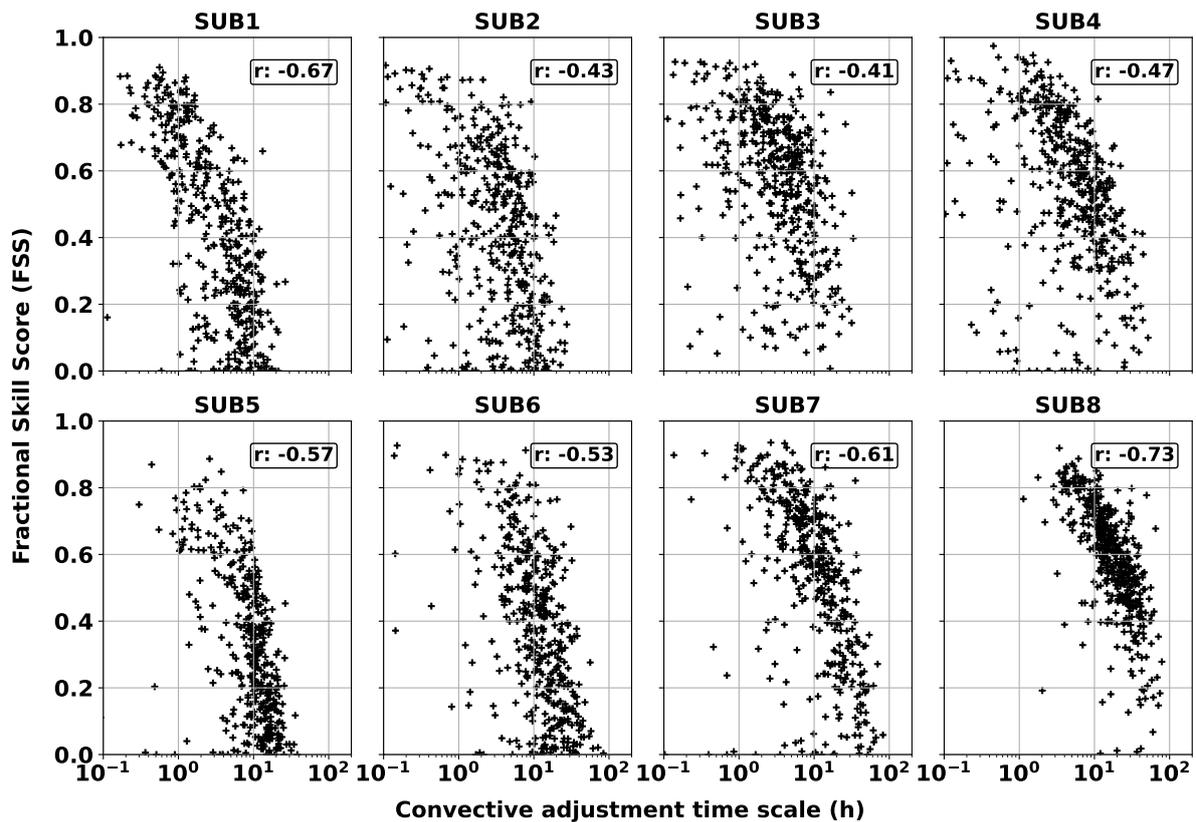


Figure 2.13: Scatterplots of convective adjustment time scale (h) and Fractional Skill Score (FSS) for the eight subdomains. Annotations denote the correlation coefficient computed for each subdomain. The FSS presented here used a squared neighborhood of 25 grid points and a binary threshold of $5mm\ 24h^{-1}$

Figure 2.13 shows the FSS for a neighborhood of 25 grid points and an absolute threshold of 5 mm in 24 hours against τ_c values, where the hourly convective time scale is averaged over 24 hours to cover the accumulation window. In general, the FSS and τ_c are negatively correlated with coefficients between 0.4 and 0.7, depending on the region. Since the FSS cannot distinguish between false alarms and misses, we use the GSS as a second skill metric. Again, we see a distinct anticorrelation between GSS and convective time scale of similar amplitude (Fig. 2.14).

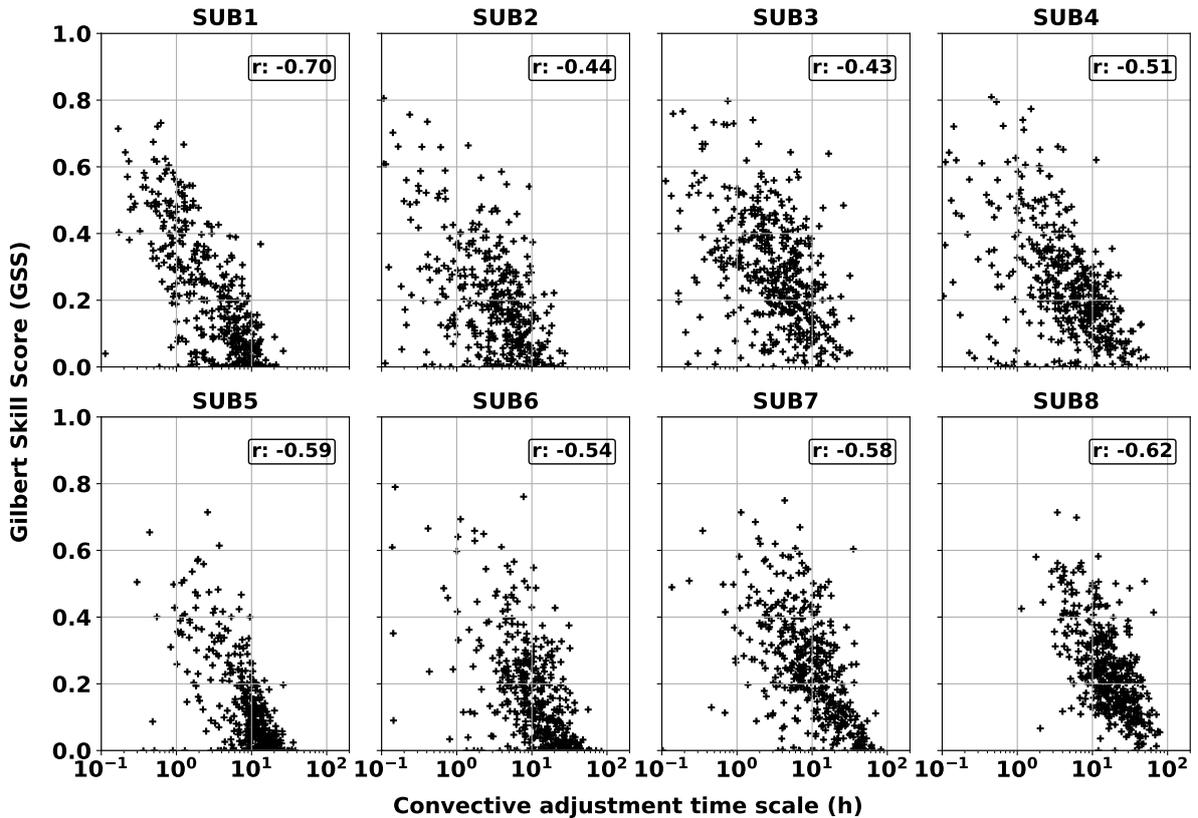


Figure 2.14: As Fig. 13, but for Gilbert Skill Score (GSS).

The sensitivity to different neighborhood sizes and thresholds was also investigated. While there is sensitivity, especially for high precipitation thresholds, our findings hold qualitatively across a large range of neighborhood sizes and thresholds. In this chapter, we present correlations between the FSS of 24-hour accumulated precipitation and the 24-hour average τ_c . Reducing the window to 12- or 6-hour windows yields similar conclusions. These results consistently demonstrate that forecast skill is reduced for nonequilibrium convection and decreases with increasing values of τ_c .

2.5 Summary and discussion

Previous work proposed the use of the convective adjustment time scale (τ_c) to classify convective weather regimes over different parts of Europe and link them to various properties like forecast skill and predictability of precipitation. The hypothesis is that equilibrium convection (small τ_c) is governed by more predictable synoptic flow conditions and, therefore, is expected to have higher skill and above-average predictability of convective systems. In contrast, nonequilibrium convection (large τ_c) is assumed to be less predictable since it is subject to local variations (e.g., topography, cold pools).

This chapter identifies the climatological frequency of equilibrium and nonequilibrium convection in different parts of the Contiguous United States (CONUS) domain. Here, we apply these concepts to operational forecasts from the convection-permitting High-Resolution Rapid Refresh (HRRR) model to see if these findings also hold over CONUS. Since this domain is much larger and often comprises various weather systems simultaneously, we develop a τ_c climatology over eight subdomains north and south of $\sim 40^\circ\text{N}$ latitude and focus on precipitation systems during May, June, July, and August from 2019 to 2022.

The results show that τ_c is a good measure for identifying equilibrium (synoptically forced) and nonequilibrium (locally forced) convection. Unlike Flack et al. (2016), which points out a distinct scale break at 3 h, the climatological frequency of τ_c generally has a scale break between 6 and 12 h across the CONUS domain. In general, the northern domains exhibit more equilibrium convection than the southern domains. Synoptically, this reflects the fact that the North is dominated by a trough-ridge pattern of the Rossby waveguide, while the South is dominated by the subtropical anticyclone with descending air masses leading to a more stable stratification and build-up of CAPE. These results collectively support that τ_c can provide a physically meaningful regime classification. For the southern regions, a scale break of the distribution allows for a clear classification into two distinct regimes, while it is more continuous in the North. In this study, we apply a τ_c of 6 hours as the threshold to categorize equilibrium and nonequilibrium convection.

One hypothesis of this study is that the precipitation of equilibrium convection depends on the propagation and intensity of synoptic weather systems. In contrast, nonequilibrium convection is expected to follow the diurnal cycle since it depends more on radiative heating and local dynamic processes. Using a supervised learning algorithm, we find that precipitation patterns with a diurnal cycle are about twice as likely to be identified in nonequilibrium regimes based on τ_c values.

Finally, we investigate if the convective adjustment time scale can be used as a predictor of forecast skill for precipitation systems over the CONUS, as suggested by Keil and Craig (2011) and Keil et al. (2014) for a domain over Germany. To this extent, we examine the relationship between the Fractional Skill Score and the Gilbert Skill Score

for 24 h accumulated precipitation and τ_c values and find negative correlations between 0.4-0.7 for the different subdomains. Regional factors like maritime versus continental conditions suggest why previous studies chose different thresholds for the convective adjustment time scale. Convection over the British Isles and Germany is characterized by synoptic flow, more like the northern CONUS domain that has, on average, $\sim 55\%$ of equilibrium convection. In contrast, only $\sim 22\%$ of convection is classified as strongly forced in southern CONUS.

This is the first study to derive a climatology of the convective adjustment time scale over the CONUS domain using data from a convection-permitting forecast model. By studying convective systems over several subdomains and for a wide range of synoptic situations, our conclusions largely confirm previous work conducted on much smaller domains. Our results indicate that the convective adjustment time scale is an effective objective measure to classify equilibrium and nonequilibrium convection regimes, enabling the flow-dependent assessment of the relative impact of different sources of uncertainty as well as various aspects of predictability. Ultimately, this study offers guidance on how to utilize the convective adjustment time scale to classify weather regimes across the CONUS domain.

Chapter 3

Regime-Dependent Forecast Error Growth in Convection-Permitting Models: The Roles of Stochastic Parameterizations

3.1 Background and overview

In this chapter, the emphasis is on analyzing regime-dependent predictability and forecast error growth, especially focusing on understanding how different stochastic parameterizations affect forecast error growth for different weather systems. With the ever-increasing computer resources, current state-of-the-art NWP forecasts use horizontal resolutions of a few kilometers, which enables the explicit representation of convective storms. Since convective parameterization schemes are arguably the leading source of model error in convection-parameterizing models (Rodwell et al., 2013), the question arises of which parameterized processes contribute most to model uncertainties in convection-permitting forecasts. Two primary candidates are 1) uncertainties in the formulation of the PBL parameterization, which triggers the initiation of convective clouds through subgrid-scale processes, and 2) uncertainties in the formulation of microphysics processes, which affects clouds, precipitation, and the radiation budget. In this study, the Physically based Stochastic Perturbation (PSP) scheme, which was originally developed in the Consortium for Small-scale Modeling (COSMO) and ICOSahedral Nonhydrostatic (ICON) models and has been extensively tested over central Europe (Kober and Craig, 2016; Hirt et al., 2019; Puh et al., 2023), is implemented into the Weather Research and Forecasting (WRF) model. Next, we assess the effects of stochastic tendencies perturbations applied to the PBL parameterization with the PSP to the effects of perturbing internal parameters in the microphysics parameterization using a Stochastic Perturbed Parameter (SPP) method, separately and in combination. Since the two schemes represent different sources of model uncertainty, they will tend to be active in different flow situations. If the perturbations are

completely independent, their impact on the overall forecast uncertainty may be additive. On the other hand, if the growth of perturbations is governed by the flow in which they are embedded, as suggested by the three-stage error growth model (Zhang et al., 2007), the effects may be independent of the detailed properties of the stochastic perturbations, and the result is independent of which scheme or combination of schemes is used. In the chapter, we seek to answer the following questions:

1. What is the impact of stochastic schemes on precipitation and forecast spread across various weather situations?
2. Is the growth of forecast error primarily governed by the flow dynamics, or are there inherent differences arising from the formulation of model error schemes?
3. Does introducing perturbations to both the PBL and microphysics schemes lead to an additive effect, and what are the underlying reasons for this outcome?

The last question is of practical importance since several studies have found that multiple error sources may not consistently result in an additive forecast spread (Berner et al., 2011; Matsunobu et al., 2022; Chen et al., 2023). This study will shed light on the underlying reasons for their findings. This chapter is organized as follows: Chapter 3.2 introduces the model configuration, selected cases, and diagnostic metrics. Two contrasting cases are presented in Chapter 3.3, followed by general results in Chapter 3.4 and 3.5. Finally, the summary and discussion are presented in Chapter 3.6.

3.2 Methodology

This study investigates the impact of different model error formulations on convection-permitting forecasts through the three ensemble experiments described below. Since the effects of model error schemes may differ across weather situations and time of day, this study selected twelve cases characterized by different forcing regimes and initialization times. The configuration of the model, stochastic perturbation schemes, experimental design, and diagnostic metrics are described as follows.

3.2.1 Model configuration and experimental design

This study used the non-hydrostatic WRF model version 4.3.1 (Skamarock et al., 2019). The 3 km computational domain (Fig. 3.1, 1023 x 1023 grid points) covers the east part of the CONUS domain and has 51 vertical levels. The model top is positioned at 50 hPa. All simulations use similar physics parameterization options (Table. 3.1) as the operational HRRR system (Dowell et al., 2022). The cumulus parameterization scheme is turned off since models with a horizontal resolution between 1 and 4 km can resolve convective processes arguably explicitly. In this study, all simulations utilize initial and boundary

conditions obtained from the High-Resolution Rapid Refresh (HRRR) operational forecast system, accessible on the NOAA open data platform (see Chapter 2.2). Given that the operational HRRR system employs an hourly cycled DA strategy, it is anticipated to have higher forecast skills than those initialized with downscaled initial conditions.

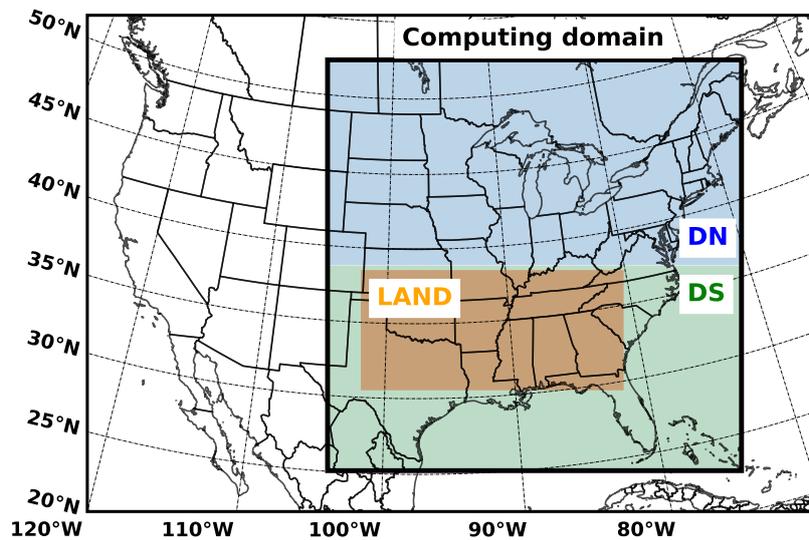


Figure 3.1: Configuration of the domain (black box) for the WRF simulations in this study. Diagnostic domains DN (light blue, 511x1023 grid points) and DS (light green, 512x1023 grid points) are partitioned along $\sim 38^\circ N$ parallel. A smaller diagnostic domain LAND (light orange, 300x600 grid points) is chosen to cover land points only.

Parameterization	Options	References
PBL	Mellor–Yamada–Nakanishi–Niino (MYNN) scheme	Nakanishi and Niino (2006)
Microphysics	Thompson aerosol aware scheme	Thompson and Eidhammer (2014)
Long- and short-wave radiation	Rapid Radiative Transfer Model for General Circulation Models	Iacono et al. (2008)
Surface layer	MYNN surface layer scheme	Olson et al. (2019)
Land surface model	Rapid Update Cycle Land-Surface Model	Benjamin et al. (2004)

Table 3.1: Configuration of physics parameterization schemes for all simulations.

Based on an unperturbed control run (abbreviated as CNTL), three ensemble experiments are designed to investigate the effect of various model error formulations

on forecast error growth. The experiment abbreviations (SPPMP, PSP, and BOTH) mirror the schemes used, making interpretation simpler. Each experiment has seven ensemble members, and no perturbations are introduced to the initial and boundary conditions. These members vary solely based on the realization of model uncertainty, determined by the random seeds assigned during the generation of stochastic random patterns. All simulations have a length of 48 hours. To minimize the impact of model spin-up, model perturbations are introduced after the first 3 h forecast, after which the ensemble members begin to diverge. The SPPMP and PSP experiments activate the Stochastically Perturbed Parameter applied to Microphysics (Thompson et al., 2021) and the WRF implementation of the PSP schemes (Kober and Craig, 2016; Hirt et al., 2019; Puh et al., 2023), respectively. The BOTH experiment combines both stochastic schemes simultaneously to evaluate their combined effect. Details of stochastic perturbation schemes are described in Chapter 3.2.2 and Chapter 3.2.3, respectively.

3.2.2 Formulation of the SPPMP scheme

The SPPMP scheme developed in Thompson and Eidhammer (2014) adopted the SPP approach to sample uncertainty in the formulation of the Thompson aerosol-aware microphysics scheme. It perturbs key internal parameters used in bulk microphysics representations, including the cloud droplet shape parameter and the graupel/hail intercept parameters. Also, it perturbs vertical velocity, which determines how many aerosols and mineral dust concentrations will be activated as cloud condensation and ice nuclei, respectively.

Parameters	Magnitudes	Spatial and temporal scales
Cloud droplet shape parameter	± 2.0	150 km and 2 hours
Graupel and hail intercept parameters	± 0.75	150 km and 2 hours
Vertical velocity for cloud condensation nuclei activation	+0.375	150 km and 2 hours
Activation of ice nuclei concentration	+13.5	150 km and 2 hours

Table 3.2: Details of the four perturbed parameters in the SPPMP scheme. The listed perturbation magnitudes originated from stochastic random patterns with a standard deviation of 0.75.

Aiming to represent realistic and physically justifiable uncertainty for each parameter, the SPPMP settings closely adhere to the methodology outlined by Thompson et al. (2021) and insights gained from personal discussions with the physics developer. As shown in Fig.3.2a, the SPPMP scheme uses a stochastic pattern generator with larger decorrelation time (2 hours) and length scales (200 km) compared to those used in the PSP

scheme (Fig.3.2b). This stochastic pattern is coherent throughout each vertical column. While the random pattern used in SPPMP scheme possesses larger spatial and temporal scales, the effective microphysics perturbations may not inherit this property since microphysical processes are naturally more intermittent and exhibit smaller scales than the PBL processes. The stochastic pattern generator uses a standard deviation of 0.75 and a cutoff threshold of 2.5 standard deviations, which leads to different perturbation amplitudes for each parameter, as summarized in the Table 3.2 (see also Thompson et al. (2021)).

3.2.3 The WRF implementation of the PSP scheme

This study implements the PSP scheme into the MYNN PBL scheme (Nakanishi and Niino, 2006), which is part of the WRF model. To distinguish it from the original implementation, the current section refers to this as the WRF implementation of the PSP scheme (WRF-PSP) to elucidate its formulation. The WRF-PSP scheme considers missing variability in resolved potential temperature (θ) and water vapor (q) fields due to partially resolved subgrid turbulence processes. This is achieved by adding stochastic perturbations into the PBL tendencies (RTHBLTEN and RQVBLTEN terms in the WRF model). At each time step, the magnitude of stochastic perturbations undergoes scaling based on the turbulence variance. This scaling reflects the concept that uncertainties are more pronounced at more turbulent grid points. The mathematical formulation of the WRF-PSP scheme is shown as (3.1).

$$\frac{\partial \phi^*}{\partial t}^i = \frac{\partial \phi}{\partial t}^i + \eta_i \cdot \alpha_{tuning} \cdot \frac{l_{mixing}}{5dx} \frac{1}{\tau_{eddy}} r_i \sqrt{\bar{\Phi}_l^2}, \quad (3.1)$$

where terms with subscript i vary every time step. $\frac{\partial \phi^*}{\partial t}^i$ and $\frac{\partial \phi}{\partial t}^i$ are perturbed and unperturbed tendency terms, respectively. ϕ is the resolved state variable potential temperature (θ) or water vapor mixing ratio (q) fields. $\bar{\Phi}_l^2$ is the variance for θ or q diagnosed in the turbulence closure. Similar to Hirt et al. (2019), the mixing length (l_{mixing}) and eddy turnover time (τ_{eddy}) were set to 1000 m and 600 seconds, respectively. The inclusion of the effectively resolved gridscale ($5dx$) makes the scheme scale-dependent by yielding it more effective at higher resolution. The two-dimensional variable k_{pbl} represents the model level where the PBL top is located and is a diagnostic variable in the MYNN scheme. As defined in (3.2), the perturbations linearly decrease from k_{pbl} to the fourth vertical level above k_{pbl} . The model-level dependent η_i linearly tapers the perturbations to avoid discontinuities near the PBL top:

$$\eta_i = \begin{cases} 1 & \text{if } k \leq k_{pbl} \\ 1 - 0.33x & k = k_{pbl} + x, x \in \{1, 2, 3\} \\ 0 & k \leq k_{pbl} + 4 \end{cases}, \quad (3.2)$$

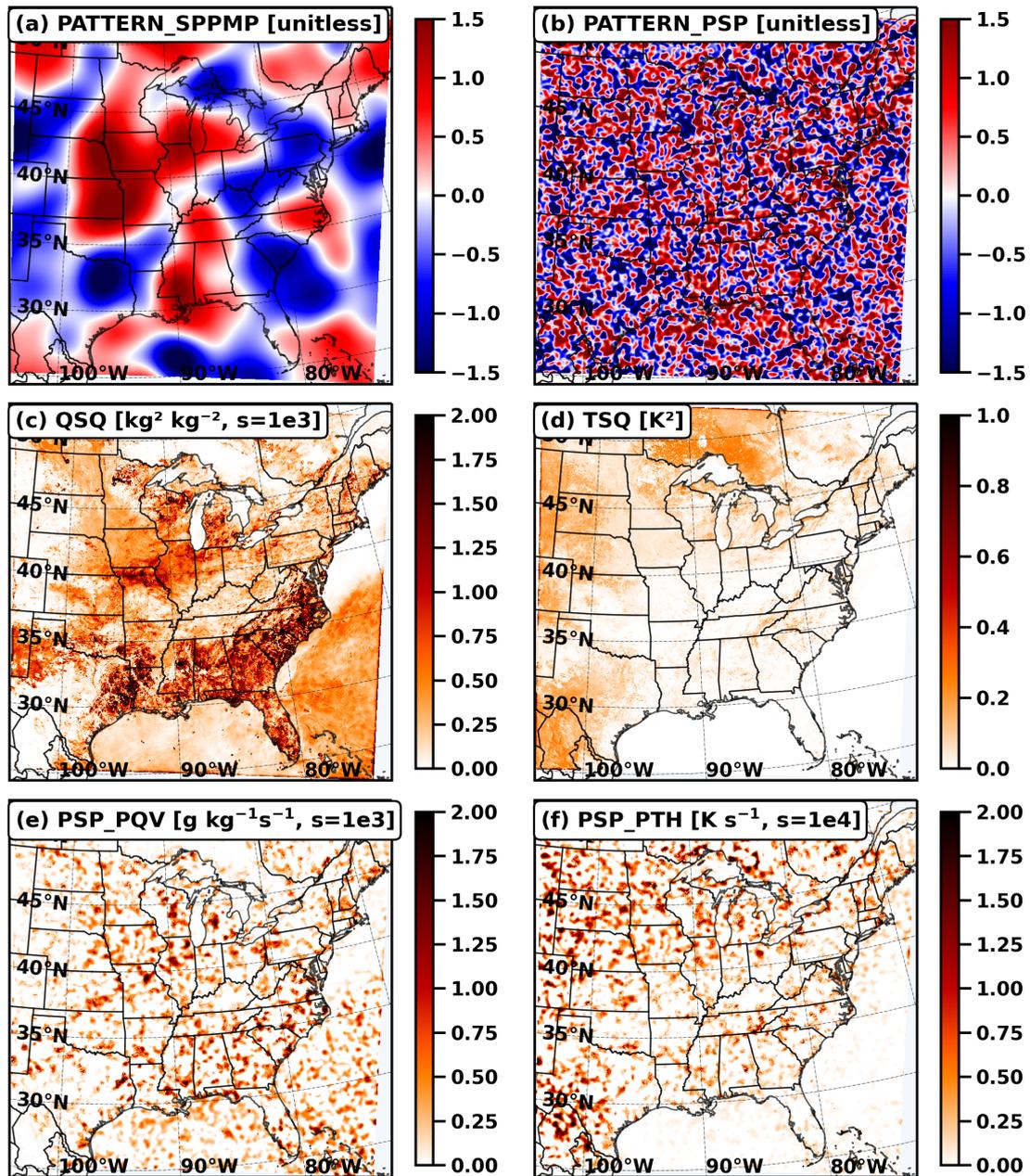


Figure 3.2: Stochastic random patterns (used in all vertical levels) for the (a) SPPMP and (b) WRF-PSP schemes. (c) Water vapor turbulence variance QSQ ($\text{kg}^2 \text{kg}^{-2}$), (d) potential temperature turbulence variance (K^2), (e) perturbed water vapor tendency ($\text{kg}^2 \text{kg}^{-2}$), and (f) perturbed potential temperature tendency (K) at the lowest model level and valid at 18 UTC on 21 July 2021.

The tuning parameter α_{tuning} scales the perturbation amplitude and should be an order of unity given realistic estimations for other parameters (Kober and Craig, 2016). In this study, the α_{tuning} is set to a fixed value of 2, a choice determined through sensitivity experiments akin to those illustrated in Figure 7 of Kober and Craig (2016). The stochastic pattern generator adheres to the implementation outlined by Berner et al. (2011), producing a two-dimensional random field (r_i) at each time step with predefined amplitude, spatial, and temporal correlations. From the spectral space ($r_{k,l}$) to the physical space (r_i), the random field is given as (3.3) and (3.4):

$$r_i = \sum_{k=-\frac{K}{2}}^{\frac{K}{2}} \sum_{l=-\frac{L}{2}}^{\frac{L}{2}} r_{k,l}(t) \exp^{2\pi i(\frac{kx}{X} + \frac{ly}{Y})}, \quad (3.3)$$

$$r_{k,l}(t + \delta t) = \exp^{\frac{-2\delta t}{\tau}} \cdot r_{k,l}(t + \delta t) + g_{k,l} \cdot \epsilon_{k,l}, \quad (3.4)$$

where X and Y are the model grid points in the x and y direction. Terms k and l are the zonal and meridional wavenumber components, respectively. As shown in (3.4), the coefficient $r_{k,l}(t)$ evolves as a autoregressive process where τ is temporal decorrelation time. In this study, τ is set to 600 seconds, as same as eddy turnover time (τ_{eddy}). $\epsilon_{k,l}$ is a white-noise process following Gaussian distribution with a zero mean and a standard deviation of one. We don't allow for perturbations above a threshold of 2.5 standard deviations since the Gaussian distribution is theoretically unbounded. The noise amplitude ($g_{k,l}$) is a function of assigned length scale, which is set to 15 km to represent variability at the smallest effectively resolved scale ($5dx$). Like the SPPMP scheme, the WRF-PSP scheme used a vertically coherent random pattern since eddies often extend through several vertical levels, especially for convective boundary layers Kober and Craig (2016).

Figure. 3.2b shows a random field from the stochastic pattern generator configured for the WRF-PSP. Since the random fields are scaled with the turbulence variance (Fig. 3.2c and Fig. 3.2d), the resulting perturbation patterns (Fig. 3.2e and Fig. 3.2f) show larger perturbations added at grid points where the turbulence variances are larger. Although θ and q use the same random field, the scaling with the variance results in distinct tendency perturbations that reflect the underlying weather situation. It is worth mentioning that the largest perturbation amplitudes are about an order smaller than the total PBL tendencies. In the following section, the abbreviation PSP is used to indicate the WRF-PSP formulation introduced above. However, it's important to note that this implementation is not identical to the implementation in the ICON and COSMO models.

3.2.4 Case description

To investigate the effect of model perturbations across different weather situations, we carefully pick cases typical for strong forcing convection, weak forcing convection, and winter storm conditions. For each case, we initialize forecasts at 00 and 12 UTC to account for diurnal error growth behavior. The abbreviation of cases and their corresponding initial times are summarized in Table 3.3. The precipitation events are objectively classified using the convective adjustment time scale τ_c (Keil et al., 2014; Flack et al., 2016). As detailed in Chapter 2, this index measures the degree of synoptic forcing by estimating the timescale at which convective instability is consumed by convective heating.

Abbreviation	Initial time	Abbreviation	Initial time
STRONG1.00	2021070100 UTC	STRONG1.12	2021070112 UTC
STRONG2.00	2021070700 UTC	STRONG2.12	2021070712 UTC
WEAK1.00	2021072100 UTC	WEAK1.12	2021072112 UTC
WEAK2.00	2021072700 UTC	WEAK2.12	2021072712 UTC
WINTER1.00	2022122200 UTC	WINTER1.12	2022122212 UTC
WINTER2.00	2023030200 UTC	WINTER2.12	2023030212 UTC

Table 3.3: Abbreviation and model initial time for each simulation.

Following the methodology outlined in the previous section, factor b is set to 0.5 to account for the reduction of CAPE (Keil and Craig, 2011; Keil et al., 2014). As a recap, Chapter 2 illustrates that equilibrium and nonequilibrium convection over the CONUS can be properly classified by a convective adjustment time scale threshold of 6 hours. In alignment with previous studies, it is confirmed that strong forcing convection ($\tau_c < 6h$) has higher predictability than weak forcing ($\tau_c > 6h$) convection in the CONUS domain. This study examines two diagnostic domains, denoted as DN (north of $\sim 38^\circ$ longitude) and DS (south of $\sim 38^\circ$ longitude), as illustrated in Figure 3.1. The two domains are analyzed separately due to the prevalence of distinct synoptic flows in each subdomain (see Chapter 2). While it is challenging to have a completely homogeneous weather situation over the entire domain for all forecast lead times, we picked the cases that exemplify typical conditions.

The STRONG1.00 and STRONG1.12 cases (Fig. 3.3a and Fig. 3.3b) have a northeast-southwest elongated rainband. The organized rainband is located at the leading edge of the upper-level trough, which provides large-scale rising air motion supporting the development of precipitation systems. For this case, τ_c is clearly less than 6 h in the DN domain, and precipitation does not exhibit a diurnal pattern (Fig. 3.4a and Fig. 3.4b). In contrast, the DS domain has a τ_c larger than 6 h with a precipitation peak in the afternoon, which is caused by scattered precipitation in coastal regions and over the

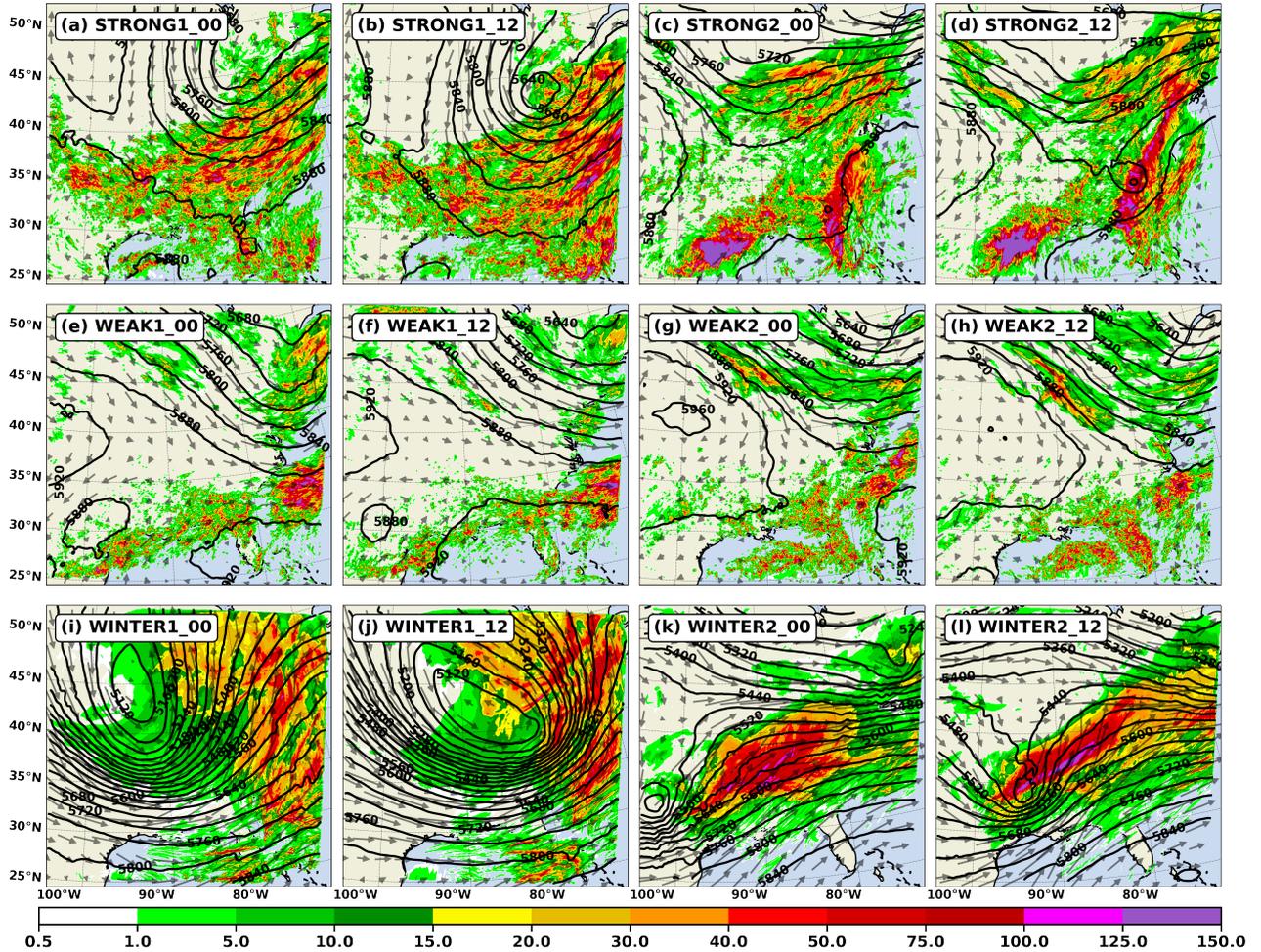


Figure 3.3: 500-hPa Geopotential height (contours, m) and two-day accumulated precipitation (shaded, mm) under different flow regimes: strong forcing (top), weak forcing (middle), and winter storm (bottom). Panels show 00 UTC and 12 UTC states for six different days (see panel labels). Results are from the unperturbed forecasts.

ocean. Similarly, the STRONG2.00 and STRONG2.12 cases have precipitation systems located at the leading edge of a 500 hPa trough (Fig. 3.3c and Fig. 3.3d). The heavy precipitation systems over the DS domain are more subtropical: part of the precipitation is associated with Tropical Storm Elsa. As shown in Fig. 3.4, this case has $\tau_c < 6$ h and τ_c slightly larger than 6 h for the DN and DS domains, respectively. Since Chapter 2 demonstrated that more than half of the convection cases in the DS domain exhibit τ_c larger than 12 h, the two precipitation cases in the DS can be viewed as a relatively strong forcing case in comparison to climatology.

Overall, the weak forcing convection cases (Fig. 3.3e-Fig. 3.3h) have less precipitation than strong forcing convection cases (Fig. 3.3a-Fig. 3.3d). As shown in Fig. 3.4e-Fig.

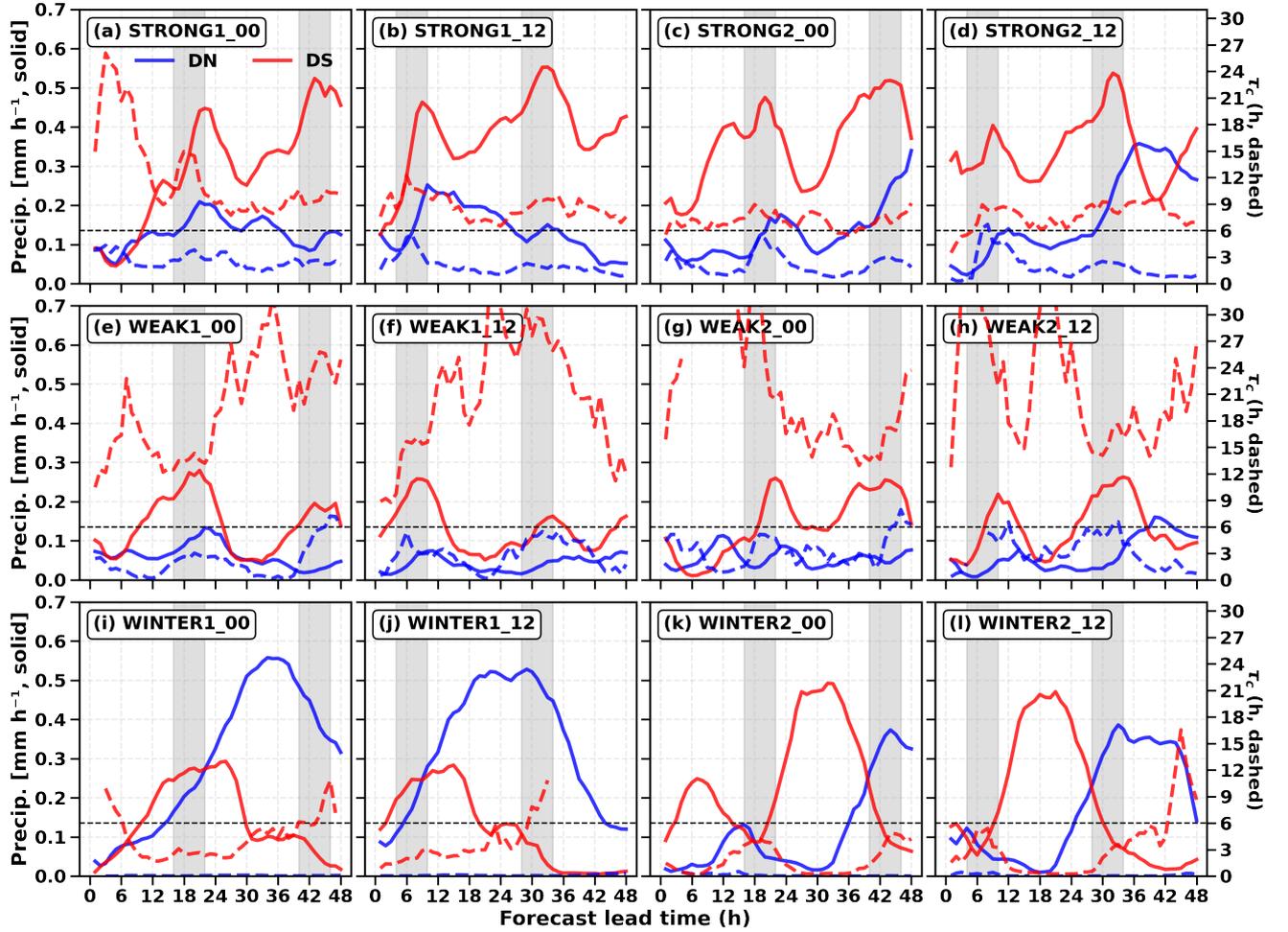


Figure 3.4: Hourly accumulated precipitation rate (mm h^{-1} ; solid lines; left y-axis) and convective adjustment time scale τ_c (h; dashed lines; right y-axis) for cases shown in Fig. 3.3 in the Northern (DN; blue) and Southern (DS; red) domains, respectively. Results are from the unperturbed forecasts.

3.4h, all cases have weakly forced precipitation over the DS domain as indicated by their large τ_c and diurnal precipitation pattern. In this study, the discussion of different forcing regimes is primarily focused on the DS domain since precipitation over the DN domain is mainly governed by the synoptic flow, albeit with varying degrees of influence. This is a common precipitation feature over the CONUS since climatologically, more than half of warm-season precipitation cases are classified as strong forcing in the northern CONUS (see Chapter 2). Although the persistent jet stream shifts poleward during the warm season, it still modulates precipitation over the northern regions most of the time.

In addition to summertime convection, two winter storm cases (depicted in Fig. 3.3i-3.3l) are incorporated to account for their unique synoptic environment and

boundary layer stability. The Winter1 cases formed on 21 Dec, reached peak intensity on 23 Dec, and dissipated on 26 Dec 2022. Its associated winter storm produces total precipitation of up to 100 mm over the entire forecast. As shown in Fig. 3.4i and Fig. 3.4j, the precipitation system dominantly settles in the DS domain during the early- to mid-forecast hours. After that, the major rainband moved to the DN domain following the counterclockwise cyclone flow. The Winter2 case formed on 01 March and dissipated on 03 March 2023. Widespread storm complexes moved throughout the DS domain and provided favorable synoptic conditions for thunderstorm development. The two winter storm cases have stronger synoptic forcing compared to the summer convection cases. As shown in Fig. 3.4, their τ_c values remain small during the whole forecast period and have an average of 0.05 h and 4.06 h for the DS and DN domains, respectively.

3.2.5 Diagnostic metric

This study evaluates how different model error formulations affect forecast errors of precipitation, wind, and temperature. Four diagnostic regions are analyzed separately: land points situated within the DN domain (abbreviated to as DN_LAND), land points situated within the DS domain (abbreviated as DS_LAND), water points situated within the DN domain (abbreviated to as DN_WATER), and water points situated within the DS domain (abbreviated as DS_WATER) domain. Furthermore, a smaller subdomain predominantly encompassing land points (abbreviated as LAND in Fig. 3.1) is employed in the spectral analysis. Three diagnostic metrics are introduced as follows:

Domain-averaged precipitation intensity and normalized spread

In this study, precipitation intensity is assessed using the domain-averaged precipitation of the ensemble mean. To minimize the effect of comparing absolute precipitation spreads across different events, we quantify the spread using the domain-average normalized standard deviation (Keil et al., 2019):

$$S_n(x, y) = \frac{1}{\overline{R(x, y)}} \sqrt{\frac{1}{N-1} \sum_{m=1}^N \{\overline{R(x, y)} - R_m(x, y)\}^2}, \quad (3.5)$$

where $S_n(x, y)$ denotes the normalized precipitation spread at each grid point. $\overline{R(x, y)}$ and $R_m(x, y)$ are the precipitation of the ensemble mean and each ensemble member, respectively. N (equal to 7 in this study) is the number of ensemble members.

Root Mean Difference Total Energy (RMDTE)

This study calculates the total energy difference between the unperturbed forecast and each ensemble member, taking into account both thermodynamic and kinematic fields

(Zhang et al., 2003). For a horizontal grid point and an ensemble member, the vertically integrated energy difference is measured by the pressure-weighted RMDTE:

$$RMDTE = \sqrt{\sum_{k=z_1}^{z_2} \frac{p(k+1) - p(k)}{p(z_1) - p(z_2)}} \cdot \frac{1}{2} [\Delta u^2 + \Delta v^2 + \frac{c_{pd}}{T_r} \Delta T^2], \quad (3.6)$$

where Δu , Δv , and ΔT are zonal wind, meridional wind, and temperature departure from the unperturbed forecast, respectively. p is the pressure field at each vertical level from the surface (denoted as z_1) up to the 300 hPa (denoted as z_2). In this study, the constant c_{pd} and T_r are set to $1004.9 J kg^{-1} K^{-1}$ and $270K$, respectively. Similar to normalized precipitation spread, this study compared energy differences across different cases using Normalized Root Mean Difference Total Energy (NRMDTE), which is defined as the standard RMDTE divided by the pressure-weighted ensemble mean KE (Nielsen and Schumacher, 2016):

$$NRMDTE = \frac{RMDTE}{\sqrt{\sum_{k=z_1}^{z_2} \frac{p(k+1) - p(k)}{p(z_1) - p(z_2)} (\bar{u}^2 + \bar{v}^2)}}, \quad (3.7)$$

where \bar{u}^2 and \bar{v}^2 are the zonal and meridional wind of the ensemble mean at each grid point. Both RMDTE and NRMDTE derived above are two-dimensional total energy differences for each ensemble member. For each experiment, RMDTE and NRMDTE are averaged over all ensemble members. Like precipitation, domain-averaged RMDTE and NRMDTE time series are computed over each diagnostic subdomain.

Kinetic Energy (KE) and precipitation spectra

Following Lorenz (1969) and subsequent error growth and predictability studies (Selz and Craig, 2015; Selz et al., 2022), we analyze the spectra of total (or background) and perturbation (or difference) fields to quantify the growth of model perturbations in the ensemble simulations. As laid out in Selz and Craig (2015), a complete phase decorrelation between two fields occurs when the energy of the forecast differences reaches a level twice the magnitude of the background energy. The specific value of the energy threshold depends on the underlying flow conditions. In line with this approach, we assess the predictability level of wind and precipitation at various spectral wavelengths through the ratio of perturbation energy to the background energy. The energy spectrum of background and difference kinetic energy (BKE and DKE, respectively) and background and difference precipitation (BPR and DPR) are derived using 500 hPa horizontal wind and hourly precipitation forecasts, respectively.

The power spectra are calculated using normalized wavenumbers described below to accommodate the rectangular diagnostic subdomain (denoted as LAND in Fig. 3.1). Two-dimensional hourly precipitation (r_{ij}), meridional (u_{ij}), and zonal wind (v_{ij}) fields are transformed to spectral space using Discrete Cosine Transform (DCT). This choice is made because the DCT is capable of obtaining more realistic spectra on small scales without the need for detrending, particularly for limited-area atmospheric fields (Denis et al., 2002). For a grid mesh with N_i and N_j grid points, the one-dimensional power spectrum has wavenumber bands equally distributed from 0 to 1 with N intervals, where $N = \min(N_i, N_j)$. The wavenumber (m, n) is normalized using $\sqrt{\frac{m^2}{N_i^2} + \frac{n^2}{N_j^2}}$ before computing their contribution to each specific wavenumber band. Take the KE spectrum as an example. The one-dimensional power spectra for the background wind and difference wind are averaged across all ensemble members, as indicated by the numerator and denominator (3.8). In each experiment, $R(l)$ gauges the saturation level at each wavenumber considering all ensemble members, utilizing their ensemble mean as the reference background flow.

$$R(l) = \frac{DKE(l)}{2 * BKE(l)} = \frac{\frac{1}{N} \sum_{m=1}^N \frac{1}{2} [E_{\Delta u}(l, m) + E_{\Delta v}(l, m)]}{\frac{1}{N} \sum_{m=1}^N [E_u(l, m) + E_v(l, m)]}, \quad (3.8)$$

where l, m, N is the spectral mode, ensemble member, and ensemble size ($N=7$ in this study). $E_u(l)$, $E_v(l)$, $E_{\Delta u}(l)$, and $E_{\Delta v}(l)$ are the one-dimensional variance spectra of background zonal, background meridional, difference zonal, and difference meridional wind fields for each ensemble member, respectively.

3.3 Two contrasting cases

Before proceeding to a more quantitative analysis, we consider two case studies, which illustrate how the two stochastic parameterizations respond differently in different weather situations. In particular, we compare a summer case (WEAK2_12) with a winter case (WINTER2_00), since these are situations where the PSP and SPPMP schemes might be expected to respond differently. In summertime over land, surface heating leads to intense boundary layer turbulence, and the PSP scheme will produce large perturbations. On the other hand, the widespread regions of cloud and precipitation in winter storms lead to large microphysical conversion terms in the model that will be perturbed by the SPPMP scheme.

Maps of RMDTE and precipitation are shown in Fig. 3.5. For the WEAK2_12 case, the RMDTE of the PSP surpasses that of SPPMP in the beginning of the forecast. However, they are remarkably similar for lead times of more than 24 hours. The high values of RMDTE (shading) typically fall in the areas with high precipitation (blue contours), signifying that error growth is organized by the atmospheric flow. As shown in Fig. 3.6a, it is apparent that in the WEAK2_12 case, SPPMP consistently exhibits a

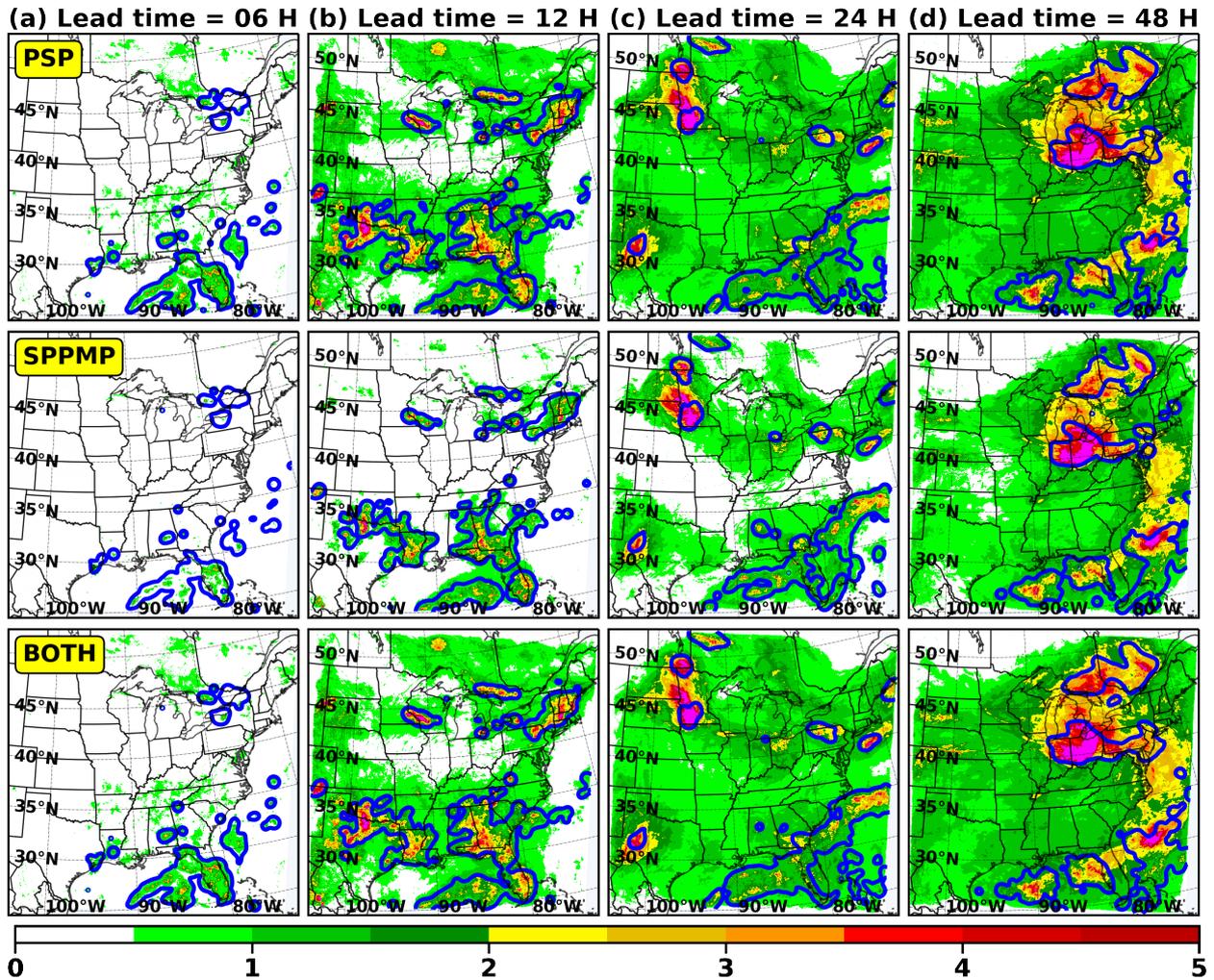


Figure 3.5: RMDTE ($m s^{-1}$) and hourly precipitation ($> 0.1mm$, blue contours) at lead times of (a) 6 h, (b) 12 h, (c) 24 h, (d) 36 h, and (e) 48 h for forecast initialized at 12 UTC on 27 July 2021. Rows are PSP, SPPMP, and BOTH experiments, respectively

lower RMDTE ratio than PSP until around 18 to 30 hours into the forecast. With the increasing lead times, the ratio tends to approach one, indicating a convergence in spread between the two schemes. It is noteworthy that this behavior does not show a significant association with the domain-average precipitation intensity. When comparing BOTH with PSP and SPPMP (Fig. 3.5), it is evident that combining both schemes does not lead to an additive effect in this case.

The WINTER2.00 case stands out as an illustration where microphysics perturbations play a more substantial role compared to their impact on the summertime convection case (Fig. 3.7). Examining the RMDTE indicates that the influence of SPPMP can surpass

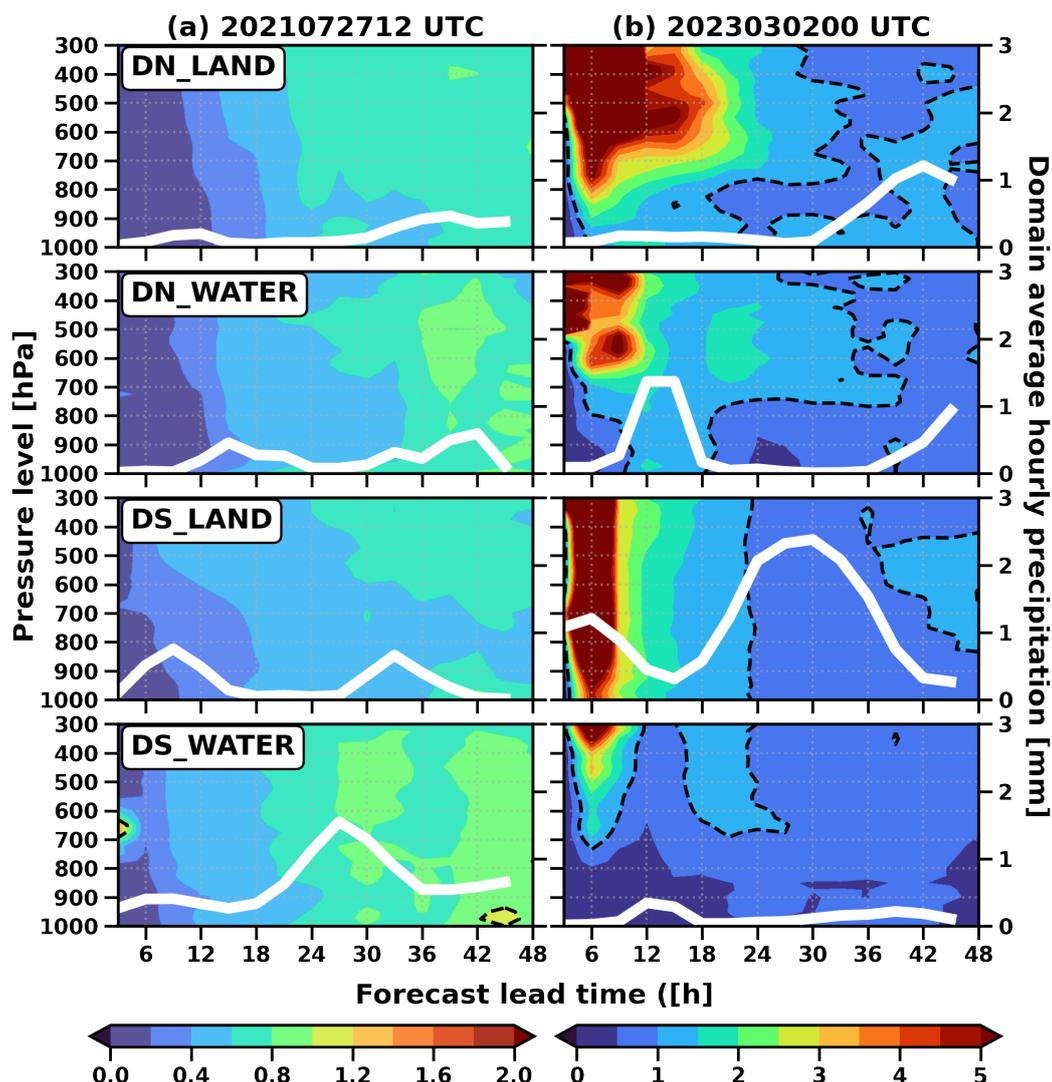


Figure 3.6: Time series of averaged RMDTE ratio (shaded, SPPMP against PSP) at pressure levels and domain average hourly precipitation rate (mm, white lines) for the 48-h forecast of the (a) WEAK2_12 and (b) WINTER2_00 cases. Rows show ratios averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively.

that of PSP within a lead time of 12-24 hours (Fig. 3.7b). As shown in Fig. 3.6b, this is especially notable in the DN_LAND and DS_LAND domains, where SPPMP significantly takes over PSP during nighttime (0 to 12 hours forecast for 00 UTC run corresponds to 7 pm to 7 am LT). The effects are most pronounced at higher vertical levels where cloud formation is prevalent, and the influence of the boundary layer processes diminished. Following the early morning period, the normalized ratios drop swiftly, indicating that the PSP perturbation grows effectively in the morning and can catch up with or even surpass SPPMP quickly, even during winter.

3. Regime-Dependent Forecast Error Growth in Convection-Permitting Models: The Roles of Stochastic Parameterizations

60

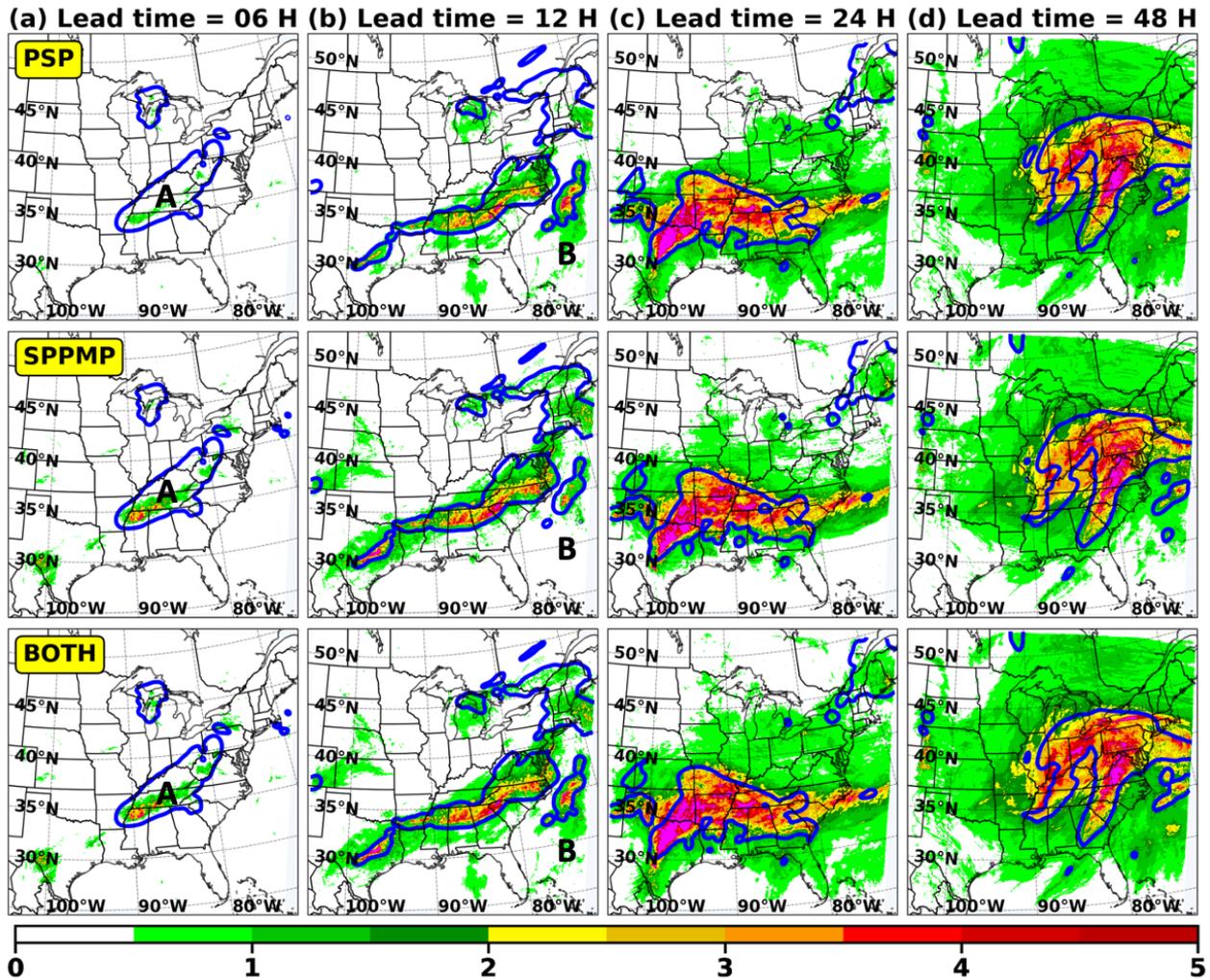


Figure 3.7: Same as Fig. 5., but for forecast initialized at 00 UTC on 02 Mar 2023.

For the WINTER2.00 case, we further study two pronounced convective systems labeled A and B at short lead times (Fig. 3.7). In system A, SPPMP generates a larger RMDTE than PSP. This suggests that the SPPMP scheme produces effective perturbations in areas with existing convection. Conversely, during this period, the PSP scheme has little impact, likely attributable to the stable nocturnal boundary layer over land. In contrast to the results over land, PSP generates a higher RMDTE than SPPMP at convective system B. In the BOTH experiment, there is a notable RMDTE for all three convective systems, as it is influenced by both microphysics and PBL perturbations.

The two case studies show different responses to the two stochastic schemes at early lead times. As expected, the PSP scheme has a stronger impact in the WEAK2.12 case, while SPPMP dominates the WINTER2.00 case. When both schemes are included, the

effects are not additive but rather reflect the impact of the dominant scheme. But as time goes by, the results become more similar, and after 6-12 h the differences appear negligible. The RMDTE ratio converges toward 1 in both cases, and the forecasts become less distinguishable, not just in amplitude but also in spatial distribution, showing the importance of the dynamical processes that amplify errors over time. In the following section, we will explore when and where the two schemes are effective and how their effects combine in different weather situations.

3.4 Impact of stochastic parameterizations in different weather situations

In this section, we quantify the impact of stochastic perturbations on forecast error growth for STRONG, WEAK, and WINTER conditions, focusing on the diagnostic domains defined in Chapter 3.2.5. Domain-averaged normalized precipitation spread S_n and NRMDTE are summarized for all cases in Fig. 3.8. We note that the stochastic perturbations are least effective for WINTER cases as it has the smallest S_n and NRMDTE values overall. In general, WEAK cases have larger S_n than STRONG cases over DS_LAND (Fig. 3.8a), which is consistent with previous studies which find that nonequilibrium precipitation is more sensitive to model perturbations (Keil et al., 2019; Puh et al., 2023). Interestingly, WEAK cases have smaller NRMDTE than strong cases (Fig. 3.8c). This suggests that a larger precipitation spread does not necessarily translate to a proportionately larger error energy when considering wind and temperature. In the DN_LAND, there are no systematic differences between STRONG and WEAK (Fig. 3.8b and Fig. 3.8d). This aligns with the diagnostics that, even in the WEAK cases, the DN domain retains equilibrium convection, as indicated by its τ_c being smaller than 6 hours (blue dashed line in Fig. 3.4e-h). Regarding the influence of different model error formulations, it is clear that PSP and BOTH demonstrate higher values of S_n and NRMDTE in comparison to SPPMP. The combination of both stochastic schemes (BOTH) does not lead to a proportional increase in NRMDTE and precipitation spread when compared to PSP. Similar conclusions can be drawn from results over water grid points (Fig. 6.1). As these findings apply to individual cases within each weather type, the subsequent diagnostics evaluate the average properties associated with three specific weather situations.

Following Lorenz (1969), previous work studied upscale error growth by analyzing time-evolving spectra (Koshyk et al., 1999; Selz and Craig, 2015; Selz et al., 2022). To link our study with this body of work, we compute precipitation and KE error spectra over the LAND diagnostic domain (denoted in Fig.1) at various lead times (Fig. 3.9). For the winter storm case, the strong synoptic forcing organizes precipitation along synoptic disturbances (Fig. 3.3i-l), leading to large amplitudes in the background precipitation spectra at the larger scales (Fig. 3.9a). For STRONG and WEAK cases, the background precipitation

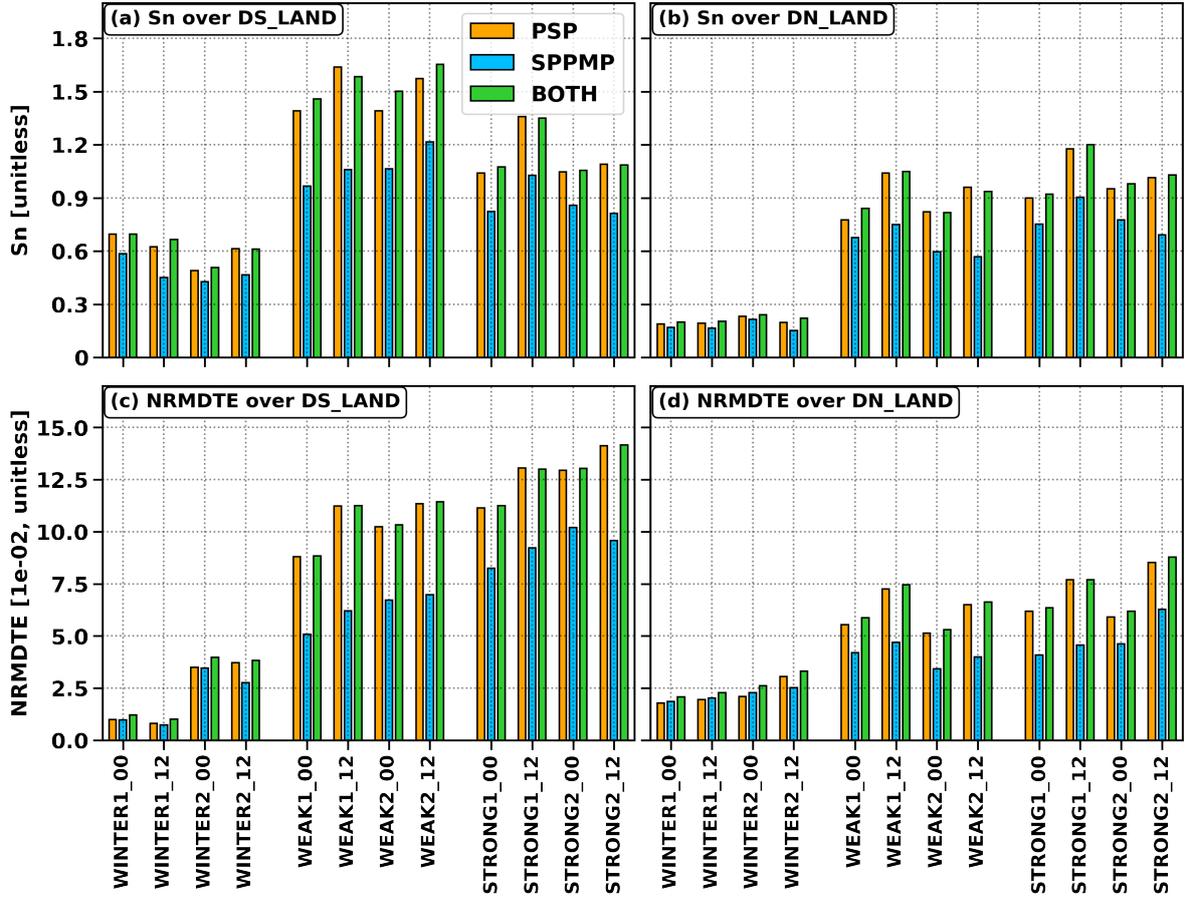


Figure 3.8: Domain-averaged normalized precipitation spread (S_n , unitless) over (a) DS_LAND and (b) DN_LAND subregions for each 48-hour forecast. Panels (c) and (d) are domain-averaged NRMDTE (unitless) over DS_LAND and DN_LAND, respectively.

spectra peak at wavelengths between 30 and 50 km (Fig. 3.9b-c), reflecting the fact that summertime convective precipitation systems have the most variability at smaller scales. In particular, the large-scale saturation amplitudes are the smallest for the WEAK cases, consistent with locally triggered nonequilibrium convection (Fig. 3.3e-h). Shifting our attention to the difference KE spectra (Fig. 3.9d-f), we note that the background KE is largest for synoptic wavelengths above 500 km and decreases for smaller wavelengths for all weather situations. Generally, PSP possesses faster precipitation and KE error growth than SPPMP for all wavenumbers. An exception is the WINTER storm situation where SPPMP stands out with higher amplitudes in both KE and precipitation at a lead time of 6 hours (Fig. 3.9a and 3.9d), mirroring the findings in the case study (section 3.3). We note that although PSP and SPPMP schemes perturb model variables with energy peaked at 15 and 150 km horizontal scales, respectively, the difference spectra after 6 hours have peaks in similar wavenumber bands. This confirms that the spatial scale of the perturbations does not imprint itself onto the forecasts.

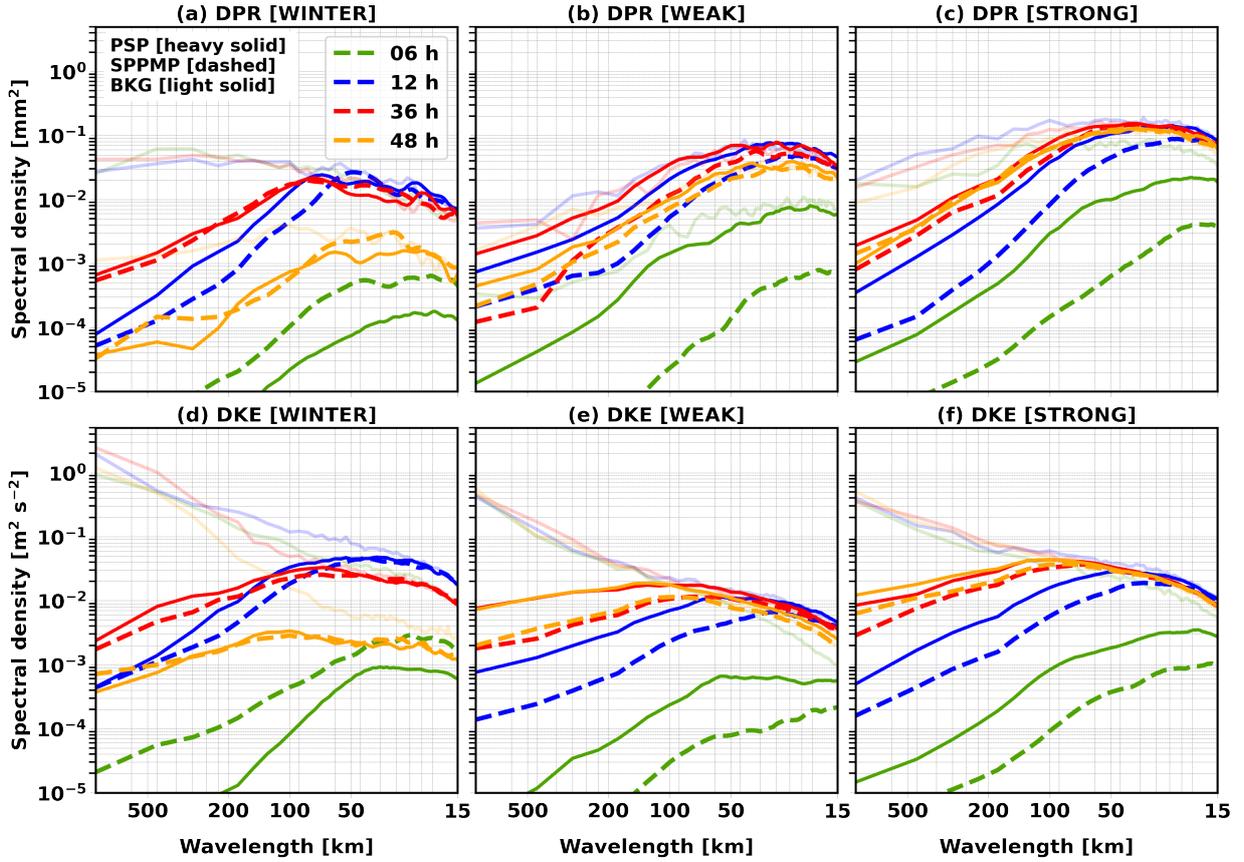


Figure 3.9: Spectra of difference hourly precipitation (heavy solid lines for PSP and dashed lines for SPPMP) and background hourly precipitation (light solid lines) per unit log k for (a) winter storm, (b) weak forcing, and (c) strong forcing cases. Figures (d), (e), and (f) are the same but for background KE and difference KE. Each line shows the average of four spectra (two 00 UTC and two 12 UTC runs). The difference fields are computed against unperturbed forecasts (see section 3.2.5 for details).

As mentioned in Selz and Craig (2015), the lead time after which the perturbation spectrum matches that of the background notifies the wavenumber-dependent predictability horizon. We expect small spatial scales to saturate faster and thus have less predictability than larger scales, which is evident in Fig. 3.9. Roughly speaking, all wavelengths of 100 km or less are close to saturated – and thus unpredictable – after a lead time of 36 hours no matter which scheme is used (Fig. 3.9). To go into more detail, we plot the saturation ratio of the difference kinetic energy against difference precipitation for various forecast leads and wavelengths (Fig. 3.10). Over time (colors progressing from blue to red), the forecasts progress from small perturbation amplitude (lower left corner of the plot) to fully saturated (upper right corner). For small scales (<50 km), precipitation and KE errors saturate (defined as a ratio of 0.8 as Selz and Craig (2015)) within one day in WEAK and STRONG cases. In contrast, the saturation

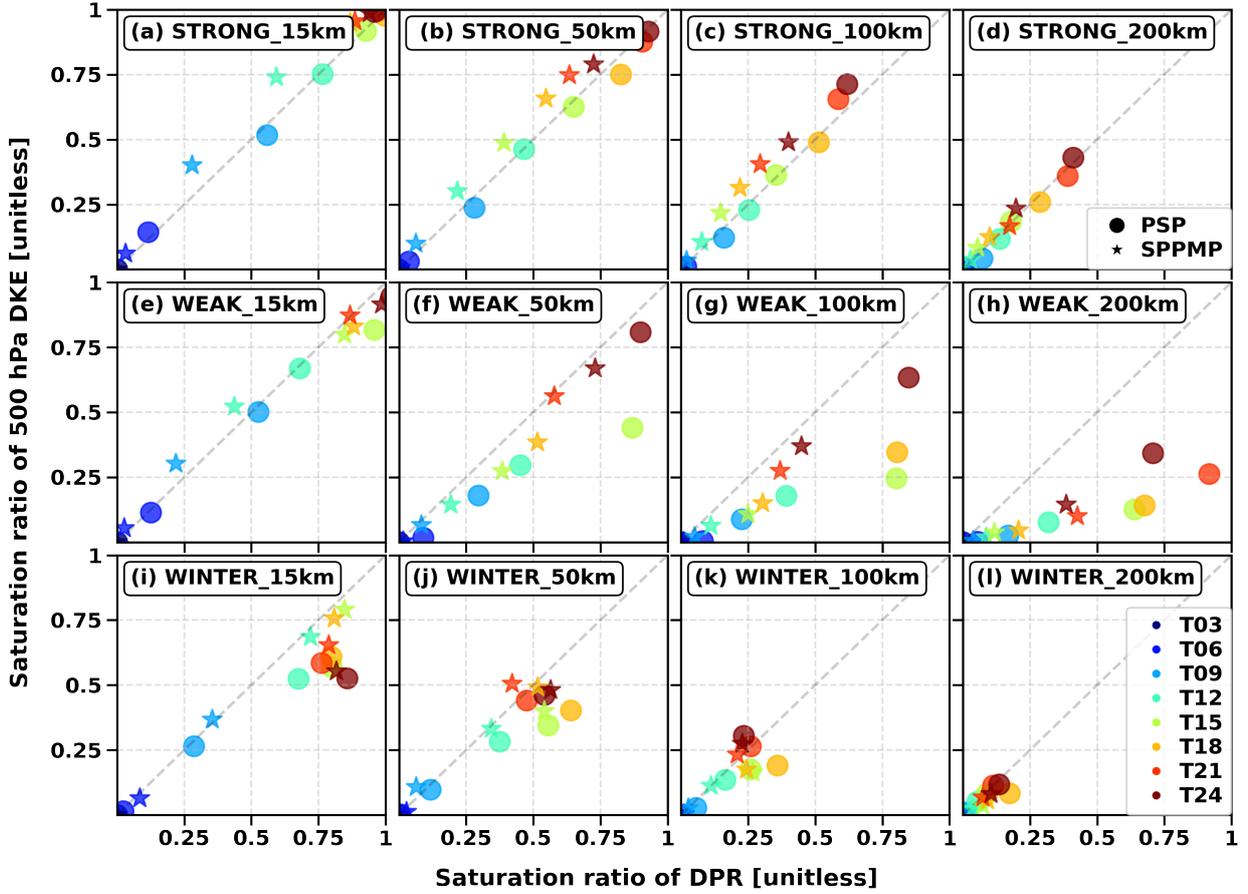


Figure 3.10: Saturation ratios of difference KE against difference precipitation at different forecast lead times (h, colors) for strong forcing (top), weak forcing (middle), and winter storm (bottom) cases, respectively. Each dot shows the average of four spectra (two 00 UTC and two 12 UTC runs). Columns are saturation ratios at 15 km, 50 km, 100 km, and 200 km wavelengths.

levels are significantly lower in WINTER cases, indicating increased predictability and suggesting that the small-scale flow is organized by the larger scales. In both the WEAK and STRONG cases, the PSP scheme has higher saturation levels in DKE and DPR than SPPMP for all wavelengths and lead times. The only circumstance in which SPPMP has a higher difference error energy than PSP occurs during WINTER cases with a short lead time, as mentioned above. This tendency is primarily observed at wavelengths smaller than 50 km (Fig. 3.10i,j). For scales exceeding 100 km, the PSP scheme has faster error growth than SPPMP (Fig. 3.10k,l).

In addition to understanding the saturation level, this analysis helps assess the extent to which the increase in KE error is associated with the error in precipitation. If the KE and precipitation errors proceed toward saturation at the same rate, the points will always

lie on the diagonal. Zhang et al. (2007) propose that fast error growth in convective scale is linked to local saturation due to displacement. In general, for the small scale (< 50 km), the saturation ratio of precipitation and KE grows at the same rate and is close to the diagonal in Fig. 3.10. For wavelengths larger than 50 km, STRONG and WINTER cases still have error saturation ratios close to the diagonal, although progress towards saturation of both variables is slower on larger scales. In contrast, for these wavelengths, WEAK cases have higher precipitation saturation ratios than KE, as indicated by the markers falling below the diagonal (Fig. 3.10 g,h). This is particularly noticeable with the PSP scheme, pointing to its more effective decorrelation of the precipitation field for weak forcing convection. Notably, for the weak forcing situation, the precipitation error nearly saturates at 200 km within 24 h, while the KE error has not reached saturation yet (Fig. 3.10h). Although the predictability of the weakly forced precipitation is less than a day on all scales shown here, this is not reflected in rapid error growth in the larger-scale flow (wavelengths of 100 km or more).

3.5 Role of the diurnal cycle

This section investigates the impact of stochastic schemes on domain average precipitation intensity, precipitation spread, and RMDTE at various lead times. To identify potential variations in diurnal error growth patterns, the ensemble forecasts initialized at 00 and 12 UTC are treated separately. Unlike the complete case stratification in the previous section, this section presents the influence of stochastic schemes on average properties for summer and winter situations. It's important to note that each case has been individually assessed, and the conclusions regarding error growth remain qualitatively consistent across summertime cases. However, there remain some regime-dependent differences, which will be further clarified in the text.

Concentrating on summertime convection cases (Fig. 3.11), the PSP scheme leads to precipitation difference relative to the control with a distinct diurnal cycle in DS_LAND (Fig. 3.11g, dashed lines). This behavior is primarily attributed to the prevalent nonequilibrium convective regime in this domain. Specifically, for 00 UTC initializations, PSP increases the precipitation rate between 18-24 hours and 42-48 hours, corresponding to 1 pm until 7 pm LT. This pattern is similarly observed for 12 UTC initializations during the respective local times. This aligns with the previous PSP studies using the ICON and COSMO models, in that the PSP scheme yields slightly higher precipitation rates and an earlier onset of precipitation for nonequilibrium convection. Following the peak precipitation event, the precipitation rate is reduced because of the depletion of precipitable water later in the diurnal cycle. In contrast to DS_LAND, it is evident that DN_LAND does not show this intensity change (Fig. 3.11a), proving that this precipitation feature primarily arises from weak forcing convection that is sensitive to PBL perturbations. Likewise, PSP scheme does not lead to large variation in the precipitation intensity for winter storm cases as they are strongly forced (Fig. 6.2). Over the DS_WATER region (Fig. 3.11j), PSP

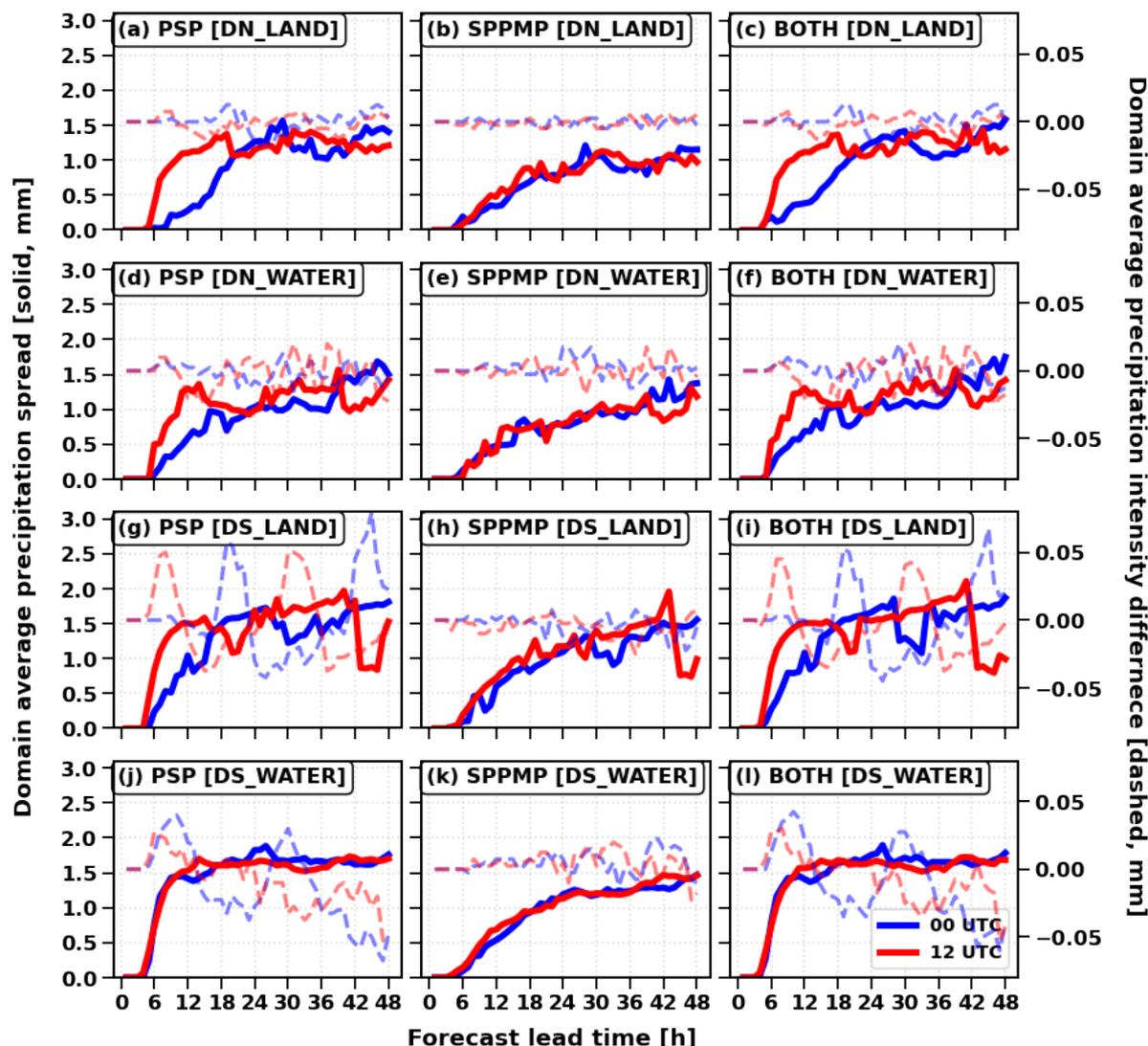


Figure 3.11: Time series of domain-averaged normalized hourly precipitation spread (solid, mm) and domain-averaged hourly precipitation difference against CNTL forecasts (dashed, mm) for PSP, SPPMP, and BOTH. The rows depict results averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Each line within the figure corresponds to four summertime convection cases.

scheme intensifies nighttime precipitation while diminishing it during the daytime. Unlike the stable nocturnal PBL over land, the PBL over the ocean remains turbulent during nighttime, creating conditions that enable the generation and the subsequent amplification of PSP perturbations (see Fig. 6.3). This elucidates the reason why PSP can result in a larger RMDTE than SPPMP in precipitation area B, as demonstrated in the case study (Fig. 3.7). In Fig. 3.11, one notable difference between the PSP and SPPMP schemes is evident: SPPMP has minimal impact on domain average precipitation intensity.

The temporal changes in precipitation intensity during summertime are reflected in the precipitation spread (Fig. 3.11, solid lines). For instance, in DS_WATER, the rapid increase of precipitation spread in 00 UTC forecasts mirrors the fluctuations in precipitation intensity caused by the turbulent ocean boundary layer during nighttime. In the other three diagnostic domains, the precipitation spreads of PSP and BOTH are significantly influenced by local time, showing faster growth in the spread for the 12 UTC run compared to the 00 UTC run. One contributing factor is the increased activity of PSP during turbulent daytime PBL conditions. It can modify or trigger additional convective cells during peak convective periods, resulting in rapid precipitation error growth. It is important to note that while the PSP scheme does not significantly alter precipitation intensity for equilibrium convection, it can decorrelate them in spatial space and increase precipitation spread, thereby explaining the diurnal precipitation spread growth in DN_LAND (Fig. 3.11a). In contrast to the local-time dependent spread characteristics for all summer cases, the winter storm cases exhibit precipitation spreads more closely as a function of lead time (Fig. 6.2). Similarly, SPPMP shows consistent spread growth behavior for both initialization times across all cases (Fig. 3.11 and Fig. 6.2).

In summary, the summertime precipitation spread induced by the PSP scheme grows rapidly during the local daytime as boundary layer turbulence and convection intensify, whereas the spread associated with SPPMP and winter cases increases more slowly and steadily with forecast lead time. For all cases studied here, the PSP scheme generally leads to a larger precipitation spread compared to that of the SPPMP scheme (Fig. 3.11 and Fig. 6.2). Further combining the two stochastic schemes (BOTH) does not result in a proportional increase in precipitation spread compared to using a single scheme alone.

Next, we investigate the forecast error growth in temperature and wind fields as measured by the RMDTE. For summer convection cases, all experiments demonstrate faster error growth rates during the local afternoon hours regardless of the scheme employed, except for DS_WATER (Figure 3.12). The reason behind this is that the evolution of RMDTE is primarily modulated by the activity of moist convection, which peaks during the afternoon in summer, rather than by the details of the perturbations introduced by the stochastic schemes (also indicated by Chapter 3.3). While PSP and BOTH exhibit comparable RMDTE growth rates across the entire forecast, SPPMP displays significantly lower error growth rates during the first local afternoon period. Looking more closely, SPPMP exhibits slightly larger growth rates in DN_LAND and DS_LAND following the peak error growth (Fig. 3.12). This suggests that it may catch up with the spread of PSP during the stable nighttime period. Over DS_WATER, PSP leads to a rapid RMDTE growth rate at short lead times for both 00 UTC and 12 UTC runs, which is consistent with its large precipitation spread. This phenomenon is again attributed to the substantial water vapor variability over the ocean during nighttime. An important finding in Fig. 3.12 is that the two schemes demonstrate distinct error growth patterns mainly during the first afternoon period. Following this, the differences in error growth rates between schemes decrease.

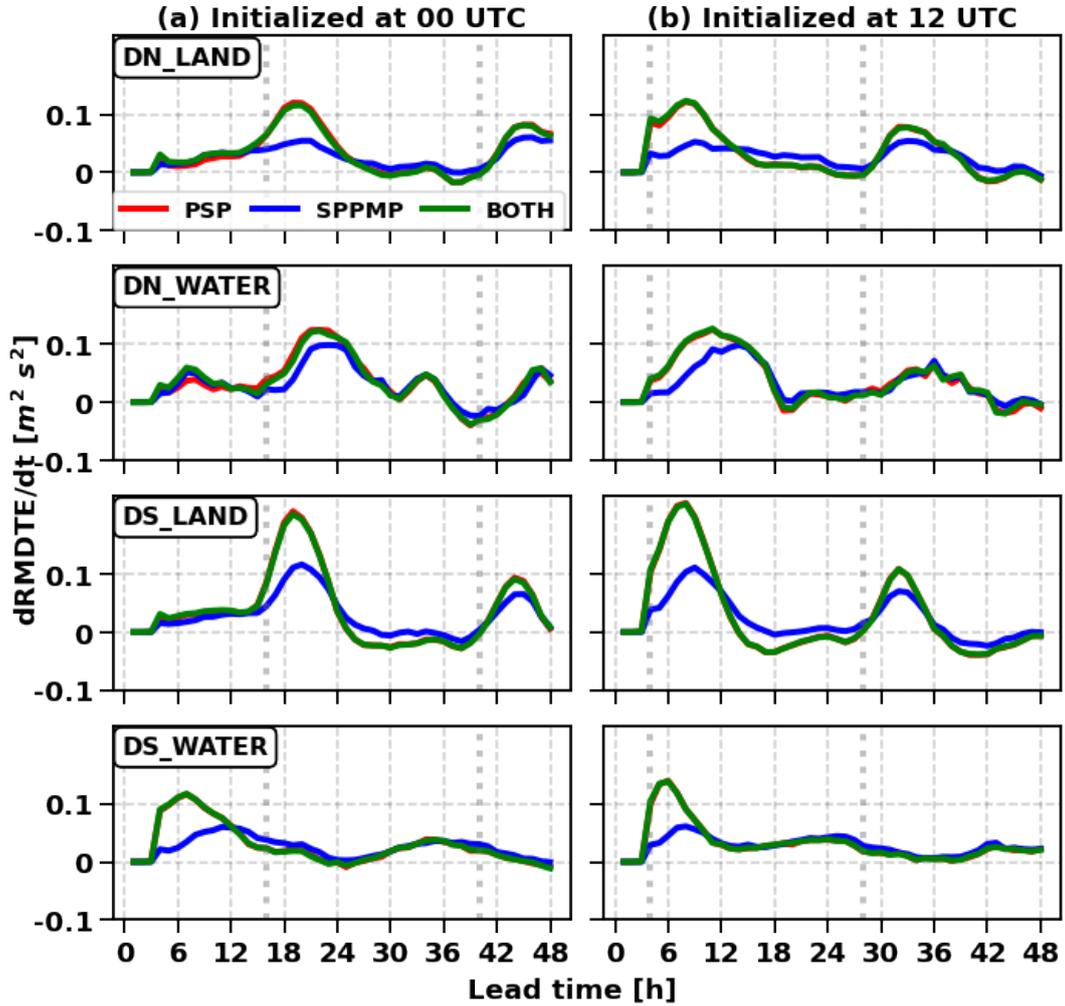


Figure 3.12: Time series of domain-averaged root-mean difference total energy (RMDTE, $m^2 s^2$) growth rate for PSP (red), SPPMP (blue), and BOTH (green). Rows are $dRMDTE/dt$ averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show forecasts initialized at (a) 00 UTC and (b) 12 UTC. Each line within the figure corresponds to four summer convection cases. Dashed lines indicate the 12 LST corresponding to each initial time.

The growth rate of RMDTE in the winter storm cases depicts a different scenario (Fig. 6.4). Although PSP still exhibits slightly larger error growth rates when considering all lead times, the distinction between the three schemes is smaller compared to their differences in summer convection cases. Put differently, the similarity in error growth among the three experiments suggests that the forecast error growth is not significantly affected by the choice of scheme in the winter storm situations. The error growth evolution is less influenced by the diurnal cycle, as the evolution of winter storms is predominantly modulated by synoptic flow, with negligible impact from local forcing.

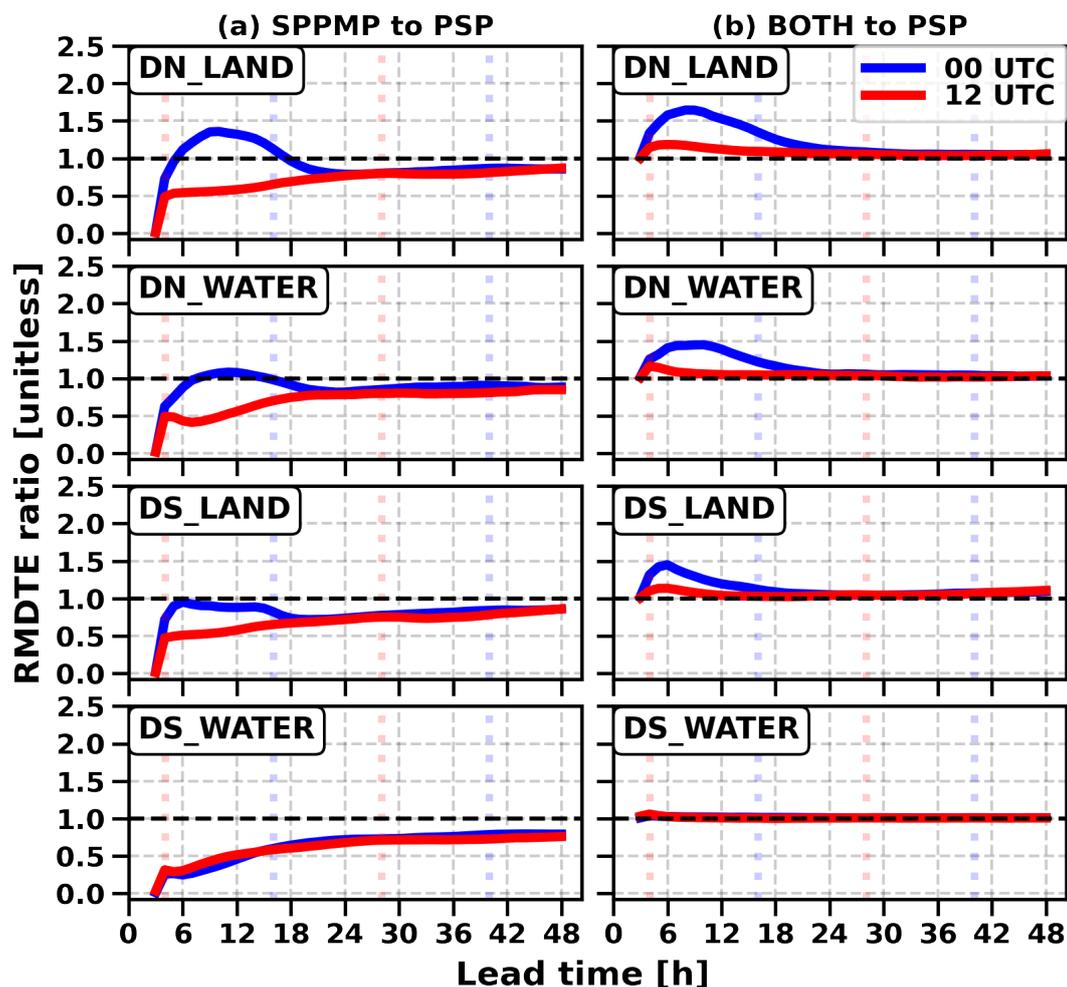


Figure 3.13: Time series of domain-averaged RMDTE ratios averaged over all cases initialized at 00 UTC (blue) and 12 UTC (red). Rows show averages over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show ratios computed as (a) SPPMP divided by PSP and (b) BOTH divided by PSP. Dashed lines indicate the 12 LT corresponding to each initial time (blue for 00 UTC and red for 12 UTC). Each line within the figure corresponds to six cases.

The growth rate of RMDTE suggests an overall faster error growth for the PSP scheme compared to SPPMP. This contrast is particularly noticeable during the peak convection afternoon period in the summertime. Despite that, the SPPMP scheme can have a larger impact after peak convection hours and to some degree catch up with the spread of PSP. Hence, it is worthwhile to examine the total RMDTE to understand how the two schemes influence the overall spread at various lead times (Fig. 3.13). We opted to present the RMDTE ratios between the two experiments as they directly illustrate the relative impact of the two schemes on the total forecast error. Since smaller RMDTE val-

ues appear at shorter lead times than at longer lead times, this normalized quantification facilitates a visual comparison of the two schemes regardless of their absolute amplitudes. One might question the significance of exploring small RMDTE values. However, it's important to note that even though the RMDTE values are small at short lead times, they represent the evolution of forecast errors at the convective scale, which aspect is crucial for the convective-scale prediction. To provide a clearer illustration, the RMDTE ratios presented in Fig. 3.13 are computed across all weather types. The regime-dependent RMDTE ratios will be discussed later as the qualitative conclusions remain valid.

As depicted in Fig. 3.13a, the PSP scheme generally yields a higher total RMDTE compared to SPPMP, with their ratios mostly falling below one across various lead times. However, SPPMP may be more important under the right conditions. For forecasts initialized at nighttime, microphysics perturbations can exert a larger impact at shorter forecast lead times, typically before the first local noon. As illustrated in Fig. 3.13a, the RMDTE ratio exceeds one before 16 h for the 00 UTC runs, corresponding to 11 am LT. This conclusion remains applicable in situations where existing precipitation systems are observed, as demonstrated in the winter storm event examined in the case study (see Chapter 3.3). After the first noon time, the influence of PSP catches up and can surpass SPPMP quickly. For forecasts initialized in the morning (e.g., the 12 UTC runs), PSP outpaces SPPMP swiftly due to the activation of PBL perturbations.

It is evident that the RMDTE ratios between the two schemes stabilize over time, approaching a value close to, though not exactly, 1 as the lead time increases (Fig. 3.13), suggesting a diminishing distinction between the amplitudes of the two schemes. Examining each weather type individually reveals that the two schemes may have an equal impact at longer lead times for winter storm cases (a ratio of 1), whereas the ratios reach approximately 0.75 to 0.9 for summertime convection cases (Fig. 6.5a). Additionally, the comparison of BOTH to PSP (Fig. 3.13b and Fig. 6.5b) reinforces the notion that the impact of different stochastic schemes is not additive overall, especially for longer lead times. The additional spread can sometimes be achieved in winter storm cases by activating SPPMP in addition to PSP, resulting in an increased forecast error of approximately 10-20% (Fig. 6.5). In contrast, the combination of the two schemes has a negligible impact on summertime convection cases, except for short lead times as explained above.

In summary, this section investigates the impact of different model error formulations on precipitation intensity change, precipitation spread, and RMDTE. While differences exist, particularly noticeable at shorter lead times, they gradually diminish as lead time increases, indicating a convergence in their behavior over time. However, within the 48-hour lead time, PSP still exhibits larger RMDTE than SPPMP, particularly in summertime convection cases. A likely explanation for their amplitude differences at longer lead times is that their convective-scale errors saturate at different times, resulting in varied amplitudes in the mesoscale error growth mechanism simultaneously. During

the 48-hour lead time, SPPMP fails to fully match the spread of PSP, but we expect this to happen with longer lead times. While differences in total error amplitudes may exist, it's essential to note that the influence of perturbation details is minor compared to the dominant impact of moist convection, which significantly regulates when and where error growth occurs. The notable similarity in error distribution and growth rate at later stages indicates that their effects are not orthogonal, explaining why adding them together does not result in additive effects (as demonstrated in Chapter 3.3). The current study indicates that the difference between the two schemes primarily lies within the short lead time. Here, their relative error amplitude, which significantly influences error growth at the convective scale, can vary considerably between the two schemes.

3.6 Summary and discussion

Current regional and next-generation global NWP models can represent convection explicitly. With convection being resolved, the leading sources of model uncertainty move toward the representation of microphysical and PBL processes. In this context, we contrast two stochastic model error representations formulated with distinct theoretical foundations in separate parameterization processes. One aims to address unresolved PBL processes due to finite grid size, while the other considers an incomplete understanding of microphysics processes. Firstly, this study utilizes the Thompson aerosol-aware microphysics scheme with stochastic parameter perturbations (SPPMP). Secondly, this chapter integrates a physically based stochastic boundary layer perturbation (PSP) scheme within the WRF model. In particular, we are interested in the question of whether forecast error growth on the convective and mesoscale scale is sensitive to the details of these model error formulations or is primarily organized by the dynamics of atmospheric flow. Since studies have shown that error growth characteristics can depend on the weather conditions, we study the impact of stochastic schemes across three flow regimes: weakly forced convection, strongly forced convection, and winter storm conditions. Overall, winter storm cases have higher predictability than summertime convection cases, as illustrated by their lower saturation levels in both precipitation and kinetic energy errors. On the contrary, weakly forced convection exhibits the lowest predictability, especially for precipitation errors that saturate within a day over a range of about 100-200 km. However, such perturbations do not necessarily result in significant variations in the temperature and wind fields (Fig. 3.10).

Confirming earlier results in other modeling systems (COSMO and ICON), we find that the effect of PSP scheme is more pronounced in nonequilibrium (weak forcing) precipitation. In this situation, the PSP scheme tends to trigger additional precipitation cells and shifts the onset of precipitation slightly earlier over land, leading to more pronounced precipitation error growth. In addition, it slightly increases nighttime precipitation over the ocean due to a turbulent marine boundary layer. In contrast, the

SPPMP scheme does not modify precipitation intensity. For summertime convection over land, the precipitation spread of PSP is more dependent on local time, whereas, for winter storm cases and SPPMP, it relies more on the forecast lead time. In general, PSP leads to a larger precipitation spread than SPPMP, and combining both schemes does not result in a proportional increase in spread.

A central finding of this study is that the differences resulting from the perturbation schemes are more pronounced at shorter forecast lead times, approximately until the first afternoon time in the diurnal cycle. As the lead time increases, the relative impact of the two schemes on temperature and wind fields tends to converge. In most cases and lead times, the PSP scheme results in a larger spread compared to those of the SPPMP scheme. Nevertheless, the SPPMP scheme has the potential for a greater impact under the right conditions. This study reveals that for forecasts initialized at night, the SPPMP scheme can have a discernible greater influence than the PSP. Specifically, the SPPMP scheme can respond to existing precipitation systems and trigger the instability earlier, especially at higher vertical levels (Fig. 3.6). In contrast, the PBL perturbations cannot grow effectively at certain times for reasons such as the stable nocturnal boundary layer. As the lead time increases, the ratios of forecast errors between the two schemes tend to converge in magnitude with remarkably similar distributions, although they do not necessarily approach 1 in every case. Despite differences in amplitude, the alignment of large forecast errors with precipitating areas implies that the evolution of forecast errors is primarily influenced by the underlying atmospheric flow rather than the details of the model perturbation. As the effects of the two schemes are not orthogonal, applying the PSP and SPPMP schemes simultaneously (experiment “BOTH”) generally does not result in proportional error growth compared to those with a single scheme. Instead, forecast errors tend to reflect those of the dominant scheme. This finding aligns with prior research and provides further insight into the underlying reasons (Berner et al., 2011; Matsunobu et al., 2022; Chen et al., 2023).

Within the two-day lead time, the SPPMP scheme generally has a smaller impact in terms of amplitude, except in winter storm cases where its influence is more pronounced than in summer convective events. Given its smaller spreads compared to PSP and the effects of SPPMP and PSP are not entirely orthogonal, one might question the worthiness of developing microphysics perturbation schemes. Our result implies that employing a microphysics perturbation scheme in addition to a PBL perturbations scheme can be particularly important for NWP systems focusing on nowcasting and short-range applications. Its potential advantage is contingent on the model initiation time. For forecasts initialized at night, microphysics perturbations can potentially contribute by leading to an earlier onset of error growth. For forecasts initialized in the morning, microphysics perturbation can have a minor impact since PSP perturbation can activate the error growth process within short lead times. The model uncertainty schemes studied here were developed to represent physically justifiable uncertainties, either derived from first principles (as in PSP) or true parameter uncertainty (as in SPPMP), in close collaboration with the parameterization developers. An attempt was made to scale the perturbation

amplitudes so that the two schemes introduced similar amounts of perturbation energy, but this was not successful owing to the constantly evolving flow. However, based on a few case studies, we confirmed that the findings reported here are still valid qualitatively when the amplitude of the microphysics perturbations is doubled. It is conceivable that the SPPMP scheme could have a comparable impact to the PSP scheme, even during the afternoon convective peak, but this would involve parameters exceeding a reasonable range within the framework of our current scheme.

Here, we study the impact of stochastic schemes on short-range forecast error growth and find that for this metric, perturbations to the PBL have more impact than to the microphysics. However, representing uncertainties in microphysics might play a much bigger role in other aspects, e.g., climatological systematic errors or response to an external forcing (Berner et al., 2017). Furthermore, the SPPMP scheme introduces variability in cloud variables, which might help ingest satellite or radar observations in DA systems. Our finding shows that even for weak forcing situations, upscale error growth is primarily dominated by the underlying flow, which explains why ad-hoc stochastic perturbation schemes - such as the SKEB and SPPT schemes- have been so successful. Small differences, either introduced by microphysics or PBL perturbations, will naturally grow where the underlying flow has moist instabilities. While we think that our findings can inform the development of EPSs, the impact on forecast skill should be verified in a comprehensive EPS, which includes initial and boundary condition uncertainties.

Chapter 4

Predictability of a Heavy Precipitation Event: The Impacts of Initial and Model Uncertainty

4.1 Background and overview

To improve the comprehension of forecast error growth in convection-permitting models, the third study explores the relative impact of initial and model uncertainties. In addition, this study serves as a follow-up to Kuo et al. (2023) and therefore concentrates on a Meiyu front precipitation case in accordance with their study. The Meiyu front, also referred to as the Baiu front in Japan, plays a substantial role in shaping the weather and climate of East Asia during the spring-to-summer transition (Chang, 2004; Qiao et al., 2021). The Meiyu front typically manifests in the lower troposphere and is characterized by an east-west-oriented, baroclinic narrow belt that often extends from eastern China through Taiwan to southern Japan. This quasistationary baroclinic zone is commonly associated with a series of eastward-propagating mesoscale convective systems, often leading to heavy precipitation events (Takahashi and Fujinami, 2021; Wang et al., 2022; Chien and Chiu, 2023). Accurately predicting the evolution of the Meiyu front and its associated precipitation is challenging due to the collective influence of multiscale processes. As summarized by Luo and Chen (2015), the Meiyu front is influenced by the convergence of tropical and polar air masses, the subtropical high, the low-level jet in southwesterly monsoon flows, and the low-level vortex in the southwest part of China. As demonstrated in Luo and Chen (2015), small errors can introduce significant forecast uncertainty in predicting the Meiyu front-associated rainfall.

Over the past decades, improvements in convection-permitting NWP models, DA techniques, and observation systems have collectively bolstered the practical predictability of convective storms. Aside from in-situ observations, there has been a growing emphasis on utilizing remote sensing observations, with radar observations being a notable example

(Sun, 2005; Bachmann, 2019). Here, another remote sensing measurement technique known as the Global Navigation Satellite System (GNSS) Radio Occultation (RO) is introduced (Kursinski et al., 1997). The RO technique can retrieve vertical profiles of temperature and water vapor, which potentially can cut through the PBL. Many studies have demonstrated the positive impact of assimilating RO data in global and regional NWP systems (Kuo et al., 2000; Healy and Thépaut, 2006; Cucurull et al., 2007; Anthes et al., 2008; Aparicio and Deblonde, 2008; Kuo et al., 2008). Despite this, the influence of assimilating RO data on the prediction of meso- and convective-scale precipitation systems has received limited attention, possibly due to its perception as a large-scale measurement. Recently, Kuo et al. (2023) investigated the impact of RO DA on the prediction of a pre-frontal squall line associated with a Meiyu front. The study realistically predicted the frontal evolution and rainfall by assimilating RO data with a nonlocal excess phase operator. Simulations using a local refractivity operator or excluding RO data below 2.5 km failed to make accurate predictions. These findings imply that proper assimilation of RO data in convection-permitting NWP systems can improve the practical predictability of convection. However, the findings in Kuo et al. (2023) were based on a deterministic forecast system and a single case study, requiring additional assessments of their robustness. In particular, it is important to acknowledge the predictability limit of any NWP forecast due to intrinsic predictability of nonlinear atmospheric flow (Zhang et al., 2003, 2007; Markowski, 2020), as well as practical predictability imposed by imperfect initial conditions, boundary conditions, and model processes (Melhauser and Zhang, 2012; Zhang et al., 2015; Flora et al., 2018). The information regarding the forecast uncertainty is crucial, as Palmer (2000) states: “no weather or climate prediction can be considered complete without a forecast of the associated flow-dependent predictability”.

Using a convection-permitting EPS, this study investigates forecast error growth from initial and model errors, and utilizes the results to evaluate the robustness of the findings in Kuo et al. (2023). For better comprehension, below provides a detailed explanation of the three experiments and hypotheses. The first experiment was designed to identify the most crucial analysis correction among wind, temperature, and water vapor fields made by RO DA. Following this, the second experiment examines the forecast error growth from random initial and model errors, respectively. This experiment has two research goals: comparing the impacts of initial and model errors, and gauging the robustness of the findings in Kuo et al. (2023) when subjected to different random perturbations. If RO DA establishes a dominant flow, its forecasts are expected to withstand and have limited error growth from small random errors, statistically outperforming forecasts without RO DA. Alternatively, if the successful prediction lacks robustness, occurring coincidentally or within an unpredictable regime transition, small perturbations can result in substantial forecast error growth (Palmer, 1993). Thus, there may be no statistical difference between forecasts initialized with and without RO DA.

Additionally, this study intends to measure the practical predictability of this event by taking into account the realistic analysis uncertainty. As Kuo et al. (2023) employed

a deterministic DA system, which does not provide uncertainty measures, this study utilized the operational Global Ensemble Forecast System (GEFS) dataset (Zhou et al., 2022) to represent analysis uncertainty as Selz et al. (2022). Specifically, the GEFS analysis uncertainty was downscaled to provide initial and boundary uncertainty for the convection-permitting EPS. Motivated by Selz et al. (2022), this study further explores how predictability evolves with the reduction of initial uncertainties, expecting an increase in skill if the forecast falls within the realm of practical predictability (Melhauser and Zhang, 2012).

In summary, the third study will tackle the question of: **What is the impact of initial and model uncertainty on precipitation predictability and forecast error growth within a convection-permitting model?** In addition, it supplements Kuo et al. (2023) by answering the following research questions:

- What is the most crucial analysis correction from the RO DA for the successful prediction of this case?
- Is the improvement made by the RO DA robust or merely coincidental when considering random errors?
- What is the practical predictability of this event given the current GEFS analysis uncertainty?

This chapter is organized as follows: Section 4.2 introduces the Meiyu case and presents the simulation results of Kuo et al. (2023). Section 4.3 describes the methodology, including model configuration, experimental design, and diagnostic metrics. Results are presented in Sections 4.4 to 4.6. Finally, the summary and discussion are presented in Section 4.7.

4.2 Case description and numerical simulations

This study selected an intense rainfall period from 00 to 12 UTC on May 22, 2020. Synoptically, the westerly wind prevails at the upper troposphere (Fig. 4.1a), while in the lower troposphere, the southwesterly flow transports abundant moisture from the open ocean to Taiwan Island (Fig. 4.1c-d). In particular, the ample moisture ahead of the Meiyu front favors the development of convective systems. The east–west-oriented frontal convergence zone is shallow and clearly discernible only at 950 hPa (green dashed line in Fig. 4.1d). A series of mesoscale convective systems form along this convergence boundary and move from west to east, leading to heavy precipitation in southern Taiwan. Figure 4.2a presents the Quantitative Precipitation Estimation (QPE) drawn from the Quantitative Precipitation Estimation and Segregation Using Multiple Sensors (QPESUMS) system (Chang et al., 2021). Alongside an empirical relation tailored for the Taiwan region, the

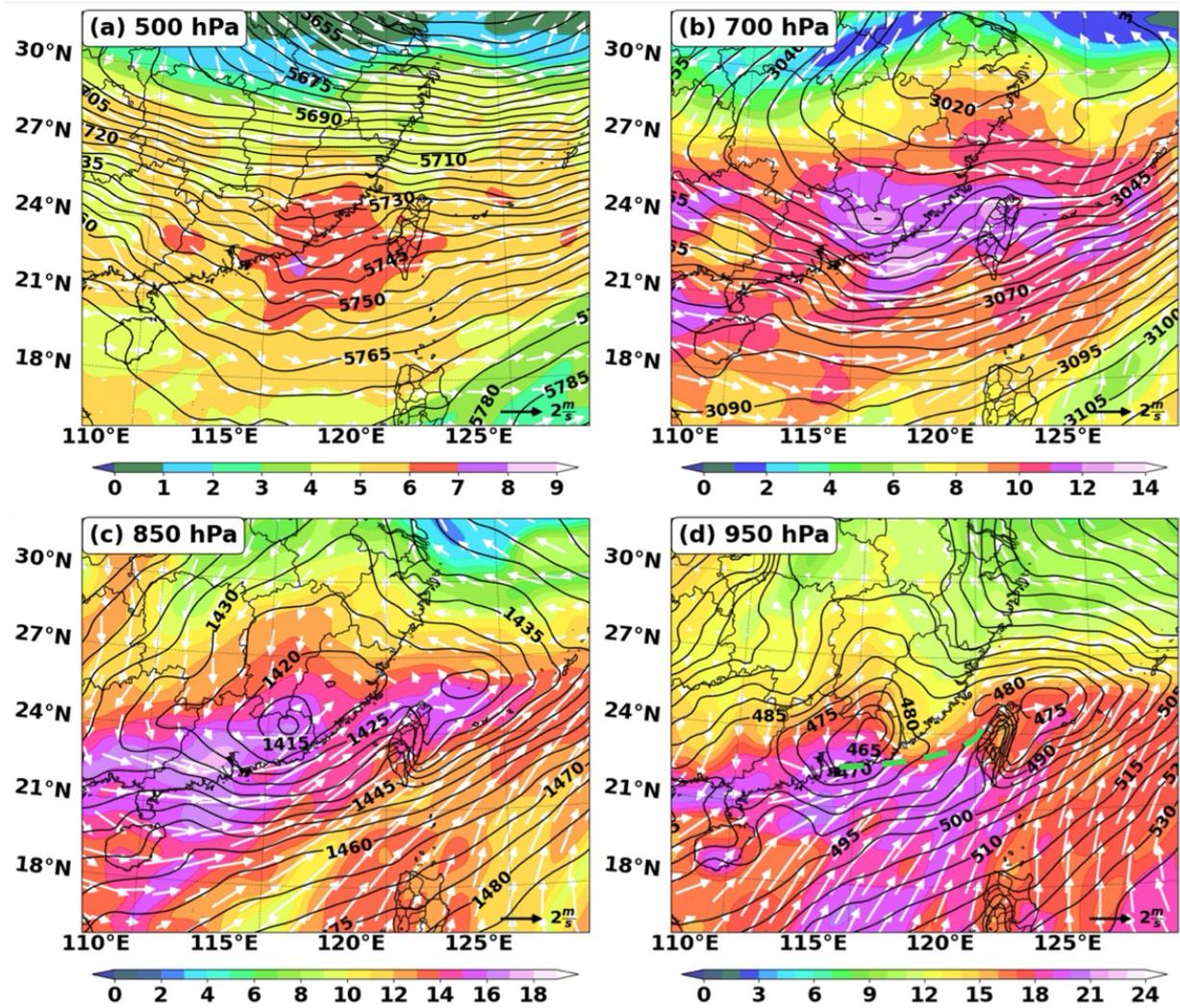


Figure 4.1: Geopotential height (contour, m), horizontal wind (vector $m\ s^{-1}$), and water vapor mixing ratio (shaded, $g\ kg^{-1}$) at (a) 500 hPa, (b) 700 hPa, (c) 850 hPa, and (d) 950 hPa, derived from the ECMWF Reanalysis v5 valid at 00 UTC on May 22, 2020.

QPESUMS QPE undergoes a local bias correction based on rain gauge observations. The 12-hour accumulated rainfall clearly illustrates an east–west-oriented rainband (Fig. 4.2a). In this case, the southern part of Taiwan receives the heaviest rainfall ($> 700\ mm$), a phenomenon shaped by the interaction of the Meiyu front and the Central Mountain Range (CMR) of Taiwan island.

As mentioned earlier, Kuo et al. (2023) explored the effect of assimilating RO data in a convection-permitting model on the prediction of this Meiyu front case. In this chapter, we utilized two analysis datasets from Kuo et al. (2023): one assimilating only conventional observations (abbreviated as CNTL), and the other additionally assimilating RO data

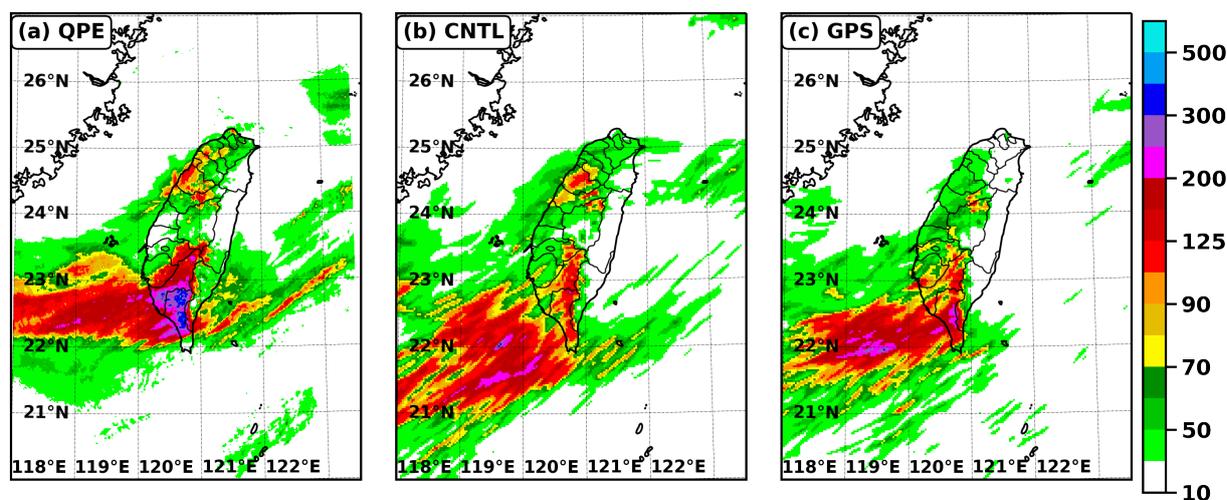


Figure 4.2: The 12-hour accumulated precipitation (shaded, mm) of (a) QPESUMS QPE, (b) CNTL, and (c) GPS ends at 12 UTC on May 22, 2020. The CNTL and GPS forecasts depicted here are based on data from Kuo et al. (2023).

using a nonlocal excess phase operator (abbreviated as GPS). The latter was named the “GPS2 experiment” by Kuo et al. (2023). Without RO DA, the CNTL experiment failed to predict a structured rainband but instead two isolated precipitation areas—one over the CMR and the other over the ocean (Fig. 4.2b). In contrast, the GPS forecast had a significant improvement as it realistically predicted the elongated rainband (Fig. 4.2c). Although both forecasts fell short in predicting extreme rainfall compared to the QPE, the GPS forecast had a closer rainfall amount with the QPE than those in the CNTL forecast. It should be noted that in Kuo et al. (2023), the 3-hour DA cycle began at 00 UTC on May 18, 2020, and the GPS forecast presented here (initialized at 00 UTC on May 22, 2020) benefited from four days of continuous RO DA. In addition to the 12-hour QPF, the hourly QPF is presented to illustrate the squall line evolution (Fig. 4.3). As observed in the QPE, the squall line evolves southward as influenced by the Meiyu front and the CMR, resulting in heavy rainfall in southern Taiwan (Fig. 4.3a). The CNTL experiment failed to predict the organized squall line and its subsequent landfall (Fig. 4.3b). Instead, it predicted a weaker system over the ocean with a tilted orientation in comparison to the QPE. With additional RO DA, the GPS experiment realistically captured the squall line evolution (Fig. 4.3c), which requires an accurate representation of flow motion and terrain effects. It should be noted that while the GPS experiment is successful in capturing the squall line evolution, there is a difference in timing. As the predicted squall line occurred three hours earlier than those in the QPE, Fig. 4.3 has been adjusted for visual comparison. Given that the GPS experiment realistically predicts the squall line, its forecast is considered the ground truth in this study. Based on the GPS experiment, the key analysis correction from RO DA is identified (Chapter 4.4), the robustness of RO DA is discussed (Chapter 4.5), and the practical predictability of this case is explored (Chapter 4.6).

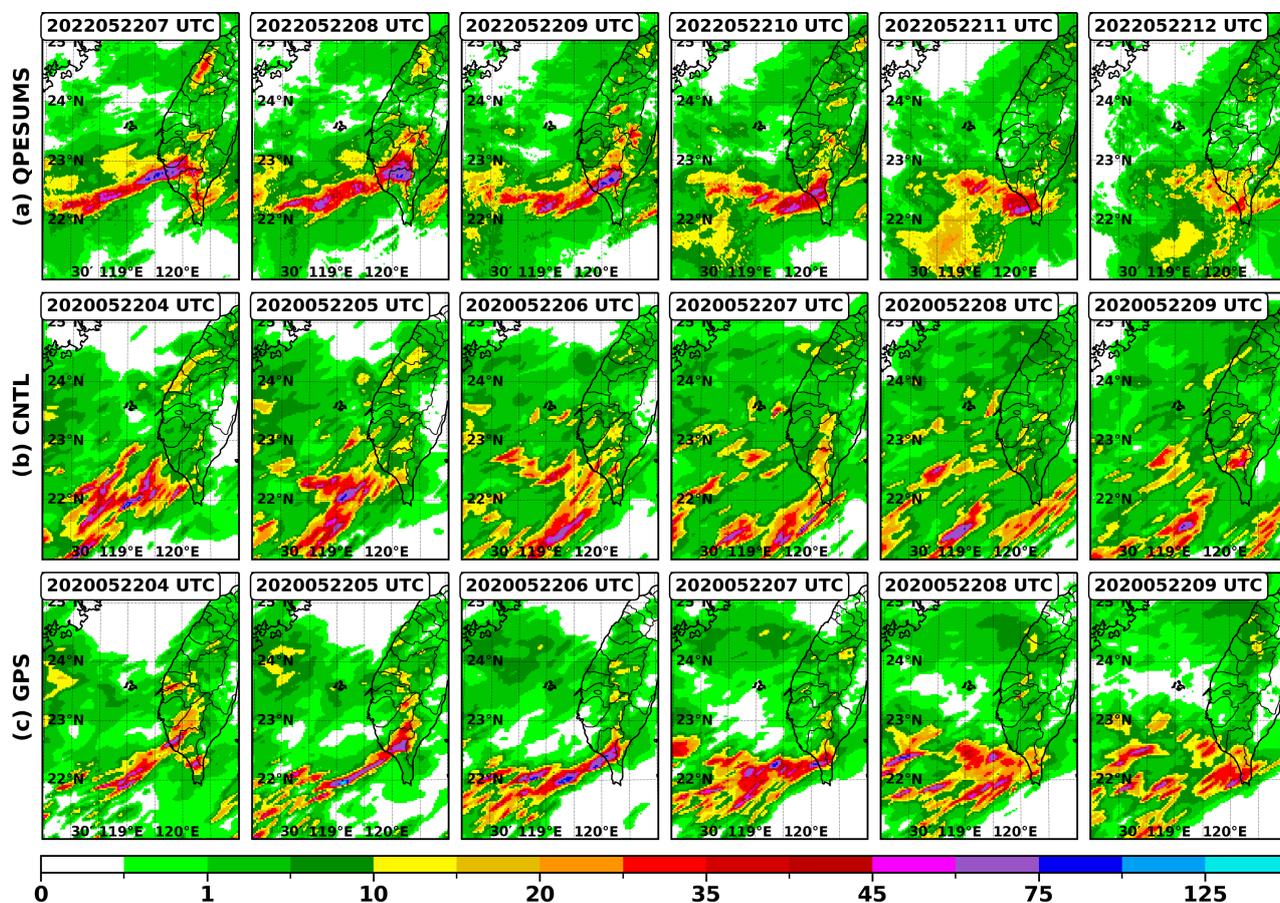


Figure 4.3: The hourly precipitation (shaded, mm) of (a) QPESUMS QPE, (b) CNTL, and (c) GPS at various forecast lead times (columns) initialized at 12 UTC on May 22, 2020. The CNTL and GPS forecasts depicted here are based on data from Kuo et al. (2023).

4.3 Methodology

This section outlines the model configuration, experimental design, perturbation methods, and diagnostic metrics utilized in the study.

4.3.1 Model configuration

This study employed the non-hydrostatic WRF model version 3.8.1 (Skamarock et al., 2008). The 15 km (denoted as D01) and 3 km (denoted as D02) computing domains were employed with one-way grid nesting (Fig. 4.4a). Both of them have 52 vertical model levels with the model top set at 20 hPa. As summarized in Table. 4.1, all simulations

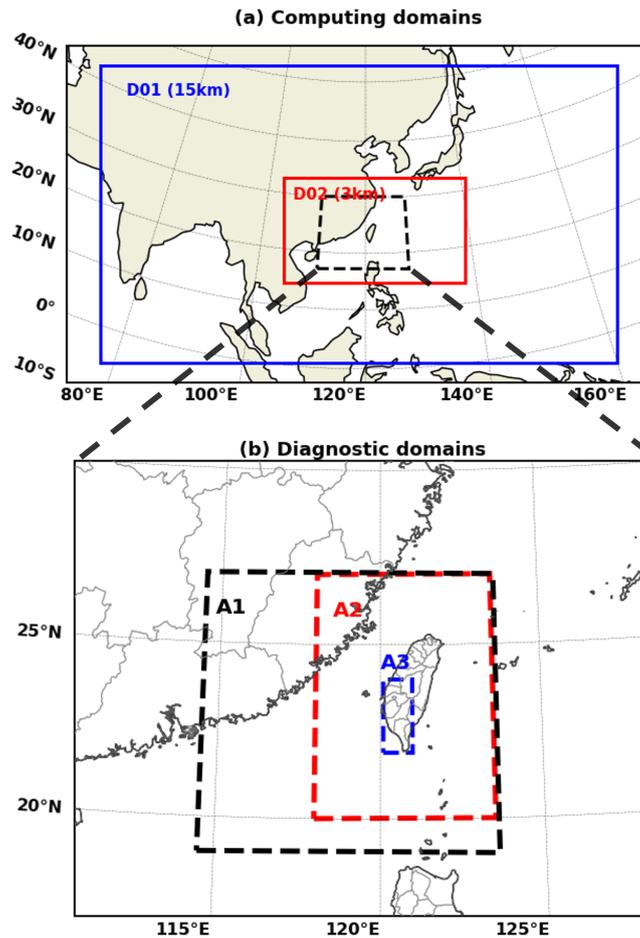


Figure 4.4: (a) Configuration of the 15 km (D01, blue box) and 3km (D02, red box) domains for the WRF simulations in this study. (b) Configuration of three diagnostic domains (denoted as A1, A2, and A3).

employed physics parameterization schemes that closely follow the operational regional forecast systems developed at the Central Weather Administration of Taiwan (Hsiao et al., 2020; Chen et al., 2023). The cumulus parameterization scheme was disabled in the 3 km domain as models with a horizontal resolution between 1 and 4 km can realistically resolve convective storms. In this study, the outermost boundary conditions were supplied by the NCEP Global Forecast System (GFS) forecast dataset, while the initial conditions were adopted from Kuo et al. (2023).

4.3.2 Experimental design

In alignment with Kuo et al. (2023), this study concentrates on lead times of up to 12 hours, during which the convection-permitting NWP systems with appropriate initial conditions can often provide satisfactory forecasts for convective storms. Table 4.2

4. Predictability of a Heavy Precipitation Event: The Impacts of Initial and Model Uncertainty

Parameterization	Options	References
PBL	YSU scheme	Hong et al. (2006)
Microphysics	Goddard Scheme	Tao et al. (2016)
Long- and shortwave radiation	Rapid Radiative Transfer Model for General Circulation Models	Iacono et al. (2008)
Surface Layer	MM5 Similarity Scheme	Paulson (1970)
Land surface model	Unified Noah Land Surface Model	Tewari et al. (2004)

Table 4.1: Configuration of physics parameterization schemes for all simulations.

summarizes the three tests (denoted as Test 1, Test 2, and Test 3) designed to address the three research questions outlined in Chapter 4.1. Except for the deterministic setup in Test 1, all experiments in Test 2 and Test 3 comprise ten ensemble members. While all experimental settings mentioned below were applied to both the 15 and 3 km domains, additional sensitivity experiments confirmed that the heavy rainfall in 3 km forecasts is largely unaffected by the solutions of the outer domain in the one-way nesting configuration (not shown). Since the RO DA corrects temperature, wind, and water vapor, the Test 1 experiments aim to identify the most critical analysis correction leading to differences between the CNTL and GPS forecasts. Using the GPS run as the reference, four deterministic experiments were conducted by substituting the initial temperature, moisture, wind, or all of them with those from the CNTL run. These deterministic runs are denoted as T1_UV, T1_T, T1_Q, and T1_UVTQ, respectively. These experiments allow an assessment of the degree to which each variable contributes to the changes in forecast skill.

Group	Experiment	Description
Test 1	T1_{UV, T, Q, and UVTQ}	Substitute the initial wind, temperature, water vapor, or all above of GPS with those from CNTL.
Test 2	T2_IC_{CNTL, GPS, Q}	Randomly perturb the initial moisture field of {CNTL, GPS, T1_Q}.
	T2_MP_{CNTL, GPS, Q}	Activate the SPPT scheme during the {CNTL, GPS, T1_Q} forecast.
Test 3	T3_IC_p{x}	Decrease the initial spread to x% of the current level. Here, $x \in \{100\%, 50\%, 10\%\}$

Table 4.2: A summary of the three sensitivity tests with detailed experimental designs.

The Test 2 experiments generate ten ensemble members by sampling random initial (referred to as T2_IC_) and model uncertainty (referred to as T2_MP_). These random errors were applied to CNTL, GPS, and also T1_Q as it has been identified as the

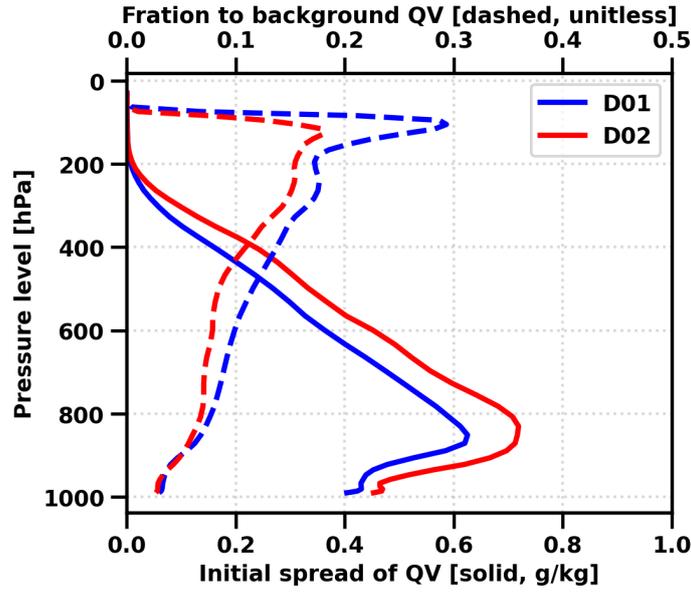


Figure 4.5: Vertical profiles of the initial moisture spread (solid) and their fraction relative to the background moisture (dashed) for D01 (blue) and D02 (red) domains.

key analysis correction in Test 1 (results shown in Chapter 4.4). To represent initial uncertainty, this study perturbed the initial moisture field, given its heightened forecast sensitivity in this case (Chapter 4.4). The random initial perturbations were generated using the random CV option 3 provided by the WRF Data Assimilation (WRFDA) system (Barker et al., 2012). In the random CV package, control variable vectors are randomly generated from a normal distribution. Next, these vectors are transformed into the model space and are further scaled to obtain perturbations for the state variables. As shown in Fig. 4.5, the initial moisture spread is more pronounced at the lower model levels, making up roughly 5% to 10% of the background states. The values at lower levels resemble those incorporated in Luo and Chen (2015) and align closely with the uncertainties observed in the ERA-Interim analysis (Dee et al., 2011). The model uncertainty was represented by the Stochastically Perturbed Parameterization Tendency (SPPT) scheme, formulated as follows:

$$\frac{\partial x^*}{\partial t} = \frac{\partial x}{\partial t} + \sum_i^n (1.0 + \gamma) \cdot \frac{\partial x}{\partial t_i} \quad (4.1)$$

where the expressions $\frac{\partial x^*}{\partial t}$ and $\frac{\partial x}{\partial t}$ are the perturbed and unperturbed total parameterized tendencies at each model grid point, respectively. The perturbed variables x included the x and y components of wind, potential temperature, and water vapor mixing ratio. The term

$\frac{\partial x}{\partial t_i}$ represents the tendencies originating from each parameterization process, such as the PBL and radiation schemes. As mentioned in Lupo et al. (2020), the SPPT scheme does not apply to the microphysics process in the WRF model. The term γ , which is fixed and applied to all parameterization tendencies, represents the perturbation amplitude. In this study, a value of 0.35 was set, aligning with the setting of Lupo et al. (2020) for the Taiwan region. The random patterns employed in this study have a coherent vertical structure.

The Test 3 experiments aim to assess the practical predictability of the heavy rainfall event when considering the analysis uncertainty of the NCEP GEFS system. To generate ensemble members, the perturbations of the first ten GEFS members were interpolated onto the 15 and 3 km domains. Next, the ten sets of initial perturbations (wind, temperature, and moisture) were re-centered using the GPS as their ensemble mean. Inspired by Selz et al. (2022), the Test 3 experiments further explore how predictability evolves when initial perturbations are rescaled with three different factors (100%, 50%, and 10%). Here, T3_IC_p100 signifies no change to the original sample, and T3_IC_p10 scales the initial spreads down to a relatively small amount to estimate how predictable this case is with nearly perfect initial conditions, akin to the definition of intrinsic predictability.

4.3.3 Diagnostic metric

This study focuses on the predictability of precipitation forecast, with two diagnostic metrics described as follows:

Precipitation spectra

To examine scale-dependent predictability and forecast error growth (Selz and Craig, 2015; Selz et al., 2022), this study analyzes the spectra of total (background) and perturbed (difference) precipitation fields at various lead times. As detailed by Selz and Craig (2015), a complete loss of predictability is signified when the energy of forecast differences matches that of the background energy. The calculation of spectra follows the same procedure as described in Chapter 3. In this study, spectral analysis was performed on the rectangular domain labeled as A1 in Fig. 4.4.

Fractional Skill Score (FSS)

Much like the verification presented in Chapter 2, this study performs QPF verification using the Model Evaluation Tools (Brown et al., 2021). Two analysis domains (denoted as A2 and A3 in Fig. 4.4) were defined to examine the Meiyu rainband at a larger scale and to focus on localized heavy precipitation over land, respectively. The FSS with a squared neighborhood consisting of 25 grid points was chosen to avoid the double penalty problem in the verification of high-resolution forecasts. Events and

non-events were binary classified based on absolute rainfall thresholds for both hourly and 12-hour accumulated rainfall. As a recap, the FSS values range from 0, indicating no forecasting skill, to 1, representing a perfect forecast. As mentioned by Skok and Roberts (2016), an FSS larger than 0.5 can indicate a useful forecast in most situations.

4.4 The key analysis correction from RO DA

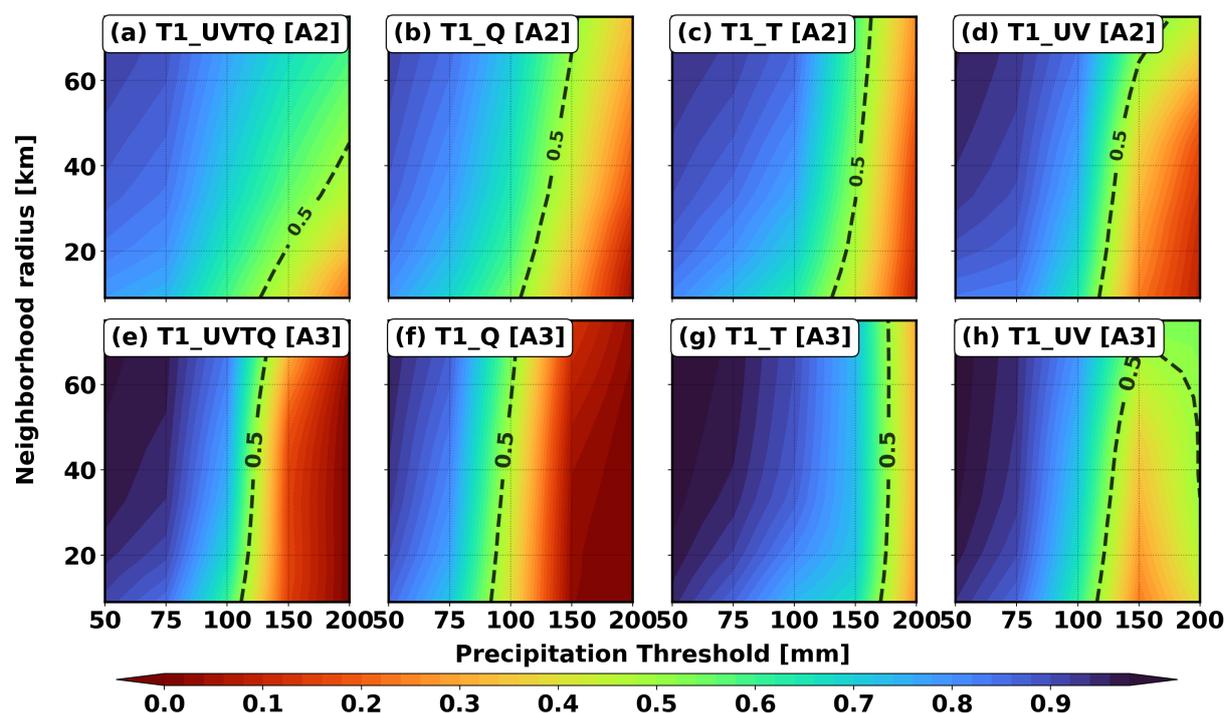


Figure 4.6: Fractional skill scores of different precipitation thresholds ($\text{mm } 12 \text{ h}^{-1}$) and neighborhood sizes (km) in the A2 domain for (a) T1_UVTQ, (b) T1_Q, (c) T1_T, and (d) T1_UV. Panels e-h represent the same analysis but for the A3 domain.

This section aims to identify the key analysis correction from RO DA that leads to the forecast differences between CNTL and GPS experiments. Here, the 12-hour QPF is verified over A2 (Fig. 4.6a-d) and A3 (Fig. 4.6e-h) domains using the GPS forecast as the ground truth. It's worth noting that a faster drop in the FSS indicates a higher sensitivity. The results indicate that the 12-hour QPF is sensitive to changes in initial conditions, particularly for heavy rainfall exceeding approximately 100 mm, where their FSSs largely fall below 0.5. It is evident that T1_Q and T1_UV lead to larger FSSs decrease than T1_T (Fig. 4.6b-d). Although RO DA does not correct the wind field directly, the correction of other variables can alter the wind field through the background error covariance and model adjustment during DA cycles. The T1_UV experiment indicates that this indirect

4. Predictability of a Heavy Precipitation Event: The Impacts of Initial and Model Uncertainty

impact can actually lead to a larger effect compared to T1_T, which is directly corrected by RO DA. Interestingly, T1_UVTQ exhibits the least sensitivity (Fig. 4.6a), which can be explained by the maintenance of precipitation intensity when compared to the GPS (Fig. 4.7a and Fig. 4.2c), whereas other experiments demonstrate significant intensity changes (Fig. 4.7b-d). In this case, the precipitation intensity weakened when either temperature or water vapor was substituted. It is worth noting that T1_UVTQ (Fig. 4.7a) results in a forecast closely resembling the CNTL run (Fig. 4.2b), implying the dominant role of initial conditions on the prediction of this particular event. Here, we acknowledge that an imbalanced initial condition can lead to spin-up and contribute partly to the outcome. Despite that, Fig. 4.6a-d and Fig. 4.7 collectively imply that the 12-hour QPF is sensitive to and may weaken due to incorrect initial water vapor and temperature fields.

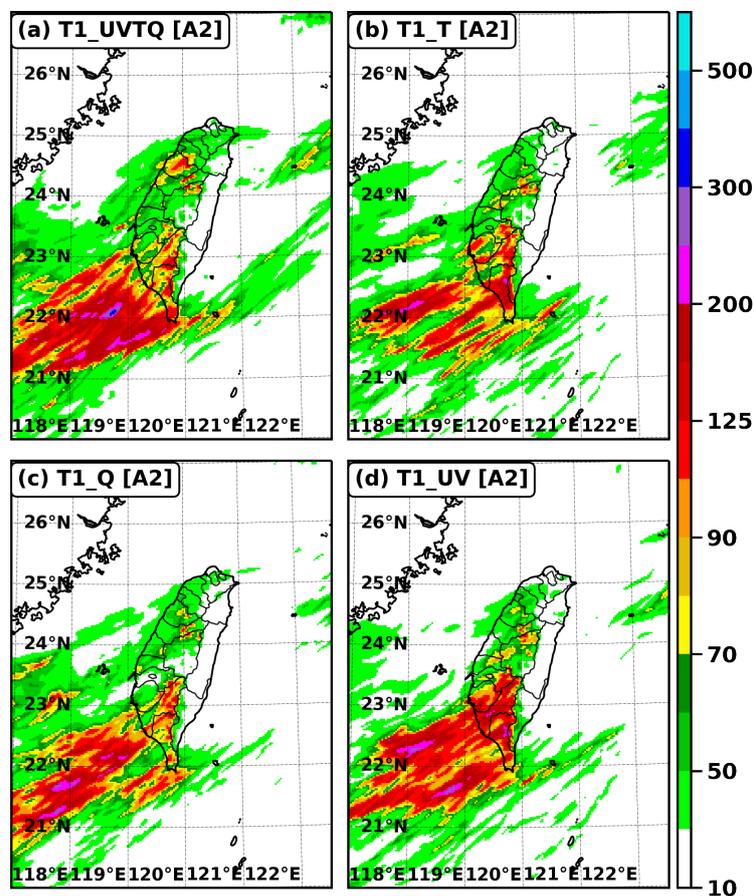


Figure 4.7: Distribution of 12-hour QPF for (a) T1_UVTQ, (b) T1_T, (c) T1_Q, and (d) T1_UV initialized at 00 UTC on May 22, 2020.

Moving on now to consider localized precipitation over land only, it becomes apparent that T1_Q exhibits the highest sensitivity, as its FSSs drop below 0.5 for thresholds exceeding 100 mm even with large neighborhood sizes (Fig. 4.6e-h). The time-latitude

diagram (computed over A3) illustrates the hourly evolution of the precipitation system (Fig. 4.8). While the CNTL run lacks a distinct evolution (Fig. 4.8a), the GPS forecast (Fig. 4.8b) displays a clear southward precipitation evolution. As mentioned earlier, predicting this movement requires a proper representation of flow motion and its interaction with the CMR. Despite T1_UVTQ showing the highest FSS in the 12-hour QPF compared to other experiments, its rainband evolution deviates significantly from the GPS (Fig. 4.8c). Substituting either temperature or wind fields still allows the forecast to capture the rainband evolution, while this evolution is seriously disrupted in T1_Q (Fig. 4.8f). In summary, the results suggest that RO DA improves moisture analysis, which is the most critical factor leading to differences between CNTL and GPS forecasts. This finding aligns with Luo and Chen (2015), which underscores the crucial role of the moisture field in predicting Meiyu-associated precipitation. However, it is important to acknowledge that the robustness of this moisture improvement remains uncertain, as this conclusion is based on a single deterministic forecast.

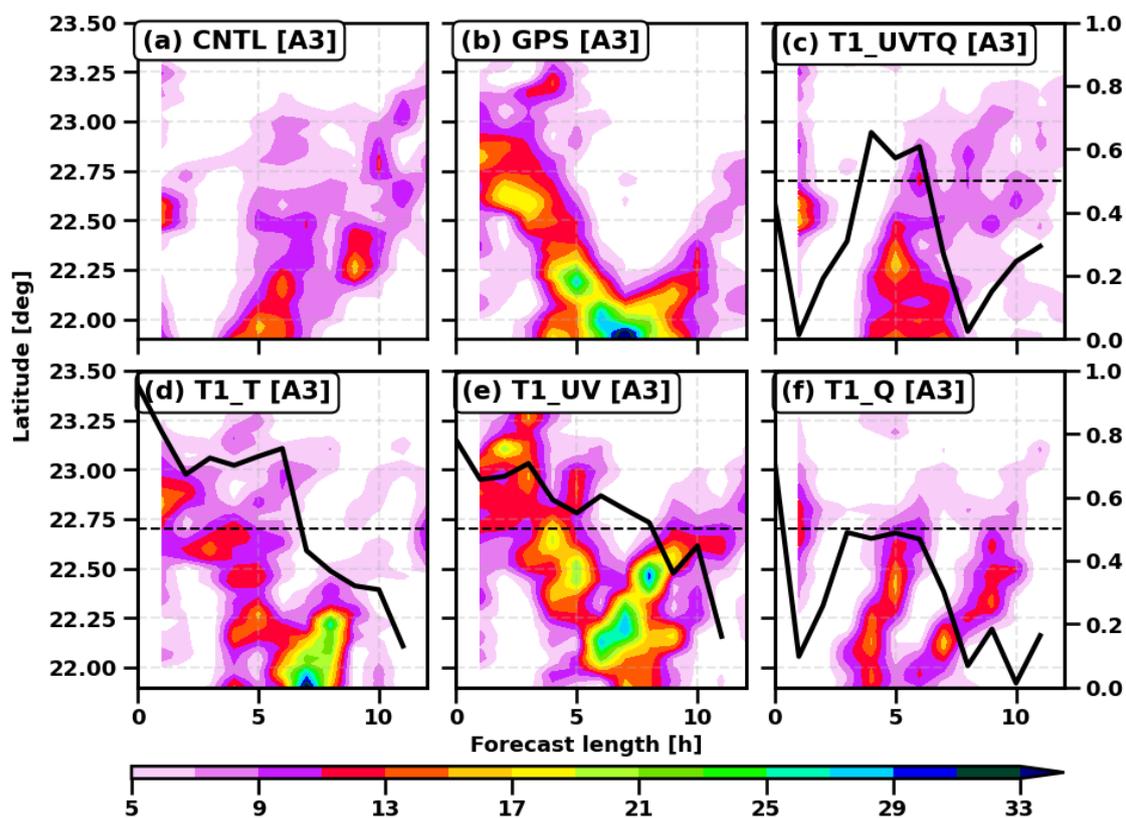


Figure 4.8: The time-latitude diagram computed over the A3 domain for hourly precipitation of (a) CNTL, (b) GPS, (c) T1_UVTQ, (d) T1_T, (e) T1_UV, and (f) T1_Q. The black solid line indicates the FSSs of hourly precipitation verified against GPS.

4.5 Impact of random initial and model errors

As explained earlier, the second experiment aims to compare the effects of initial and model errors. Additionally, it seeks to evaluate the robustness of the distinction between GPS, CNTL, and T1_QV in the presence of random errors. Since the primary objective is to ascertain whether these errors would degrade the forecast skill, this study evaluates the FSS of each member. Afterward, the average FSS of the ten members is used to represent the performance of each experiment. In the following discussion, we consider a forecast to be useful if it achieves an FSS of 0.5 or higher (Skok and Roberts, 2016).

4.5.1 12-hour Quantitative Precipitation Forecasting

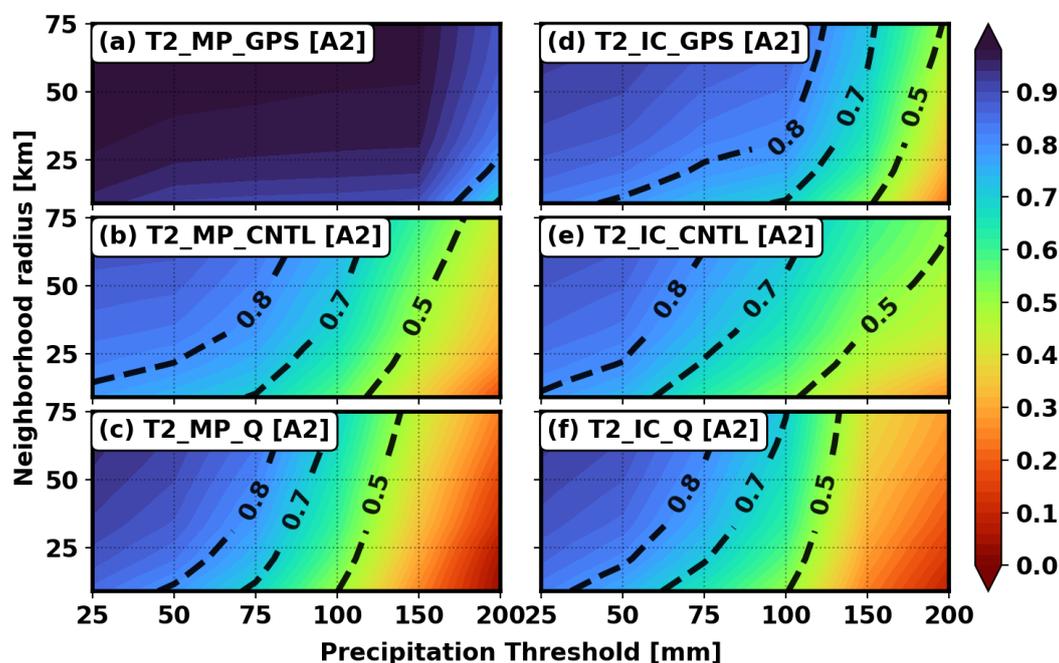


Figure 4.9: Fractional Skill Scores of different precipitation thresholds ($\text{mm } 12 \text{ h}^{-1}$) and neighborhood sizes (km) computed over the A2 domain for (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. Dashed lines denote FSS of 0.5, 0.7 and 0.8.

As shown in Fig. 4.9a, T2_MP_GPS has FSSs predominantly above 0.8, indicating that the SPPT scheme has little impact on the 12-hour QPF. In contrast, T2_IC_GPS has FSSs falling below 0.5 for thresholds above 150 mm when considering small neighborhood sizes (Fig. 4.9d). The results suggest that uncertainty arising from initial conditions grows much faster than that from model errors. Due to the small impact of model errors, T2_MP_GPS outperforms T2_MP_CNTL and T2_MP_Q (Fig. 4.9a-c). Despite the degradation caused by random initial errors, T2_IC_GPS still outperforms the other two

experiments across all thresholds and neighborhood sizes (Fig. 4.9d-f). Figure 4.10 displays the probability map for precipitation exceeding 150 mm. When compared to their unperturbed counterparts (Fig. 4.2), the rainfall distributions remain largely unchanged due to model errors (Fig. 4.10a-c), consistent with their high FSSs (Fig. 4.9a-c). In contrast, T2_IC_GPS suggests that initial perturbations may disrupt the elongated rainband between land and ocean (Fig. 4.10e), resembling those observed in the CNTL forecast. Overall, the 12-hour QPFs initialized from the CNTL, GPS, and T1.Q initial conditions have distinct rainfall distributions, suggesting that the differences between the three experiments are not coincidental but rather robust in the presence of random errors.

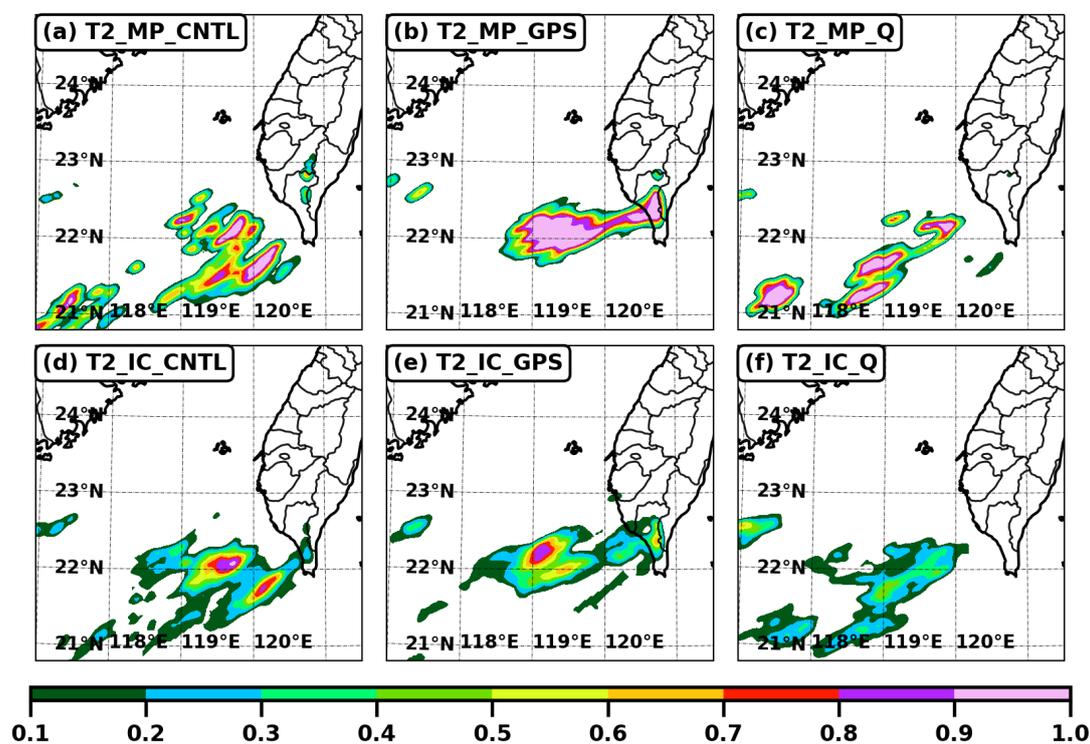


Figure 4.10: Probability of 12-hour QPF larger than 150 mm, derived from the 10 members of (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. The neighborhood probability is computed using 25 grids.

4.5.2 Hourly Quantitative Precipitation Forecasting

As illustrated in Fig. 4.11a, T2_MP_GPS has FSSs largely above 0.8, except for the later stages of the forecast. Its high FSSs are explainable as the SPPT scheme results in limited error growth without reaching saturation within the first 9 hours (Fig. 4.12a and Fig. 4.12c). Even at the end of the forecast, it only leads to error saturation at small scales. In contrast, the FSSs drop rapidly in T2_IC_GPS (Fig. 4.11d) as initial errors lead to

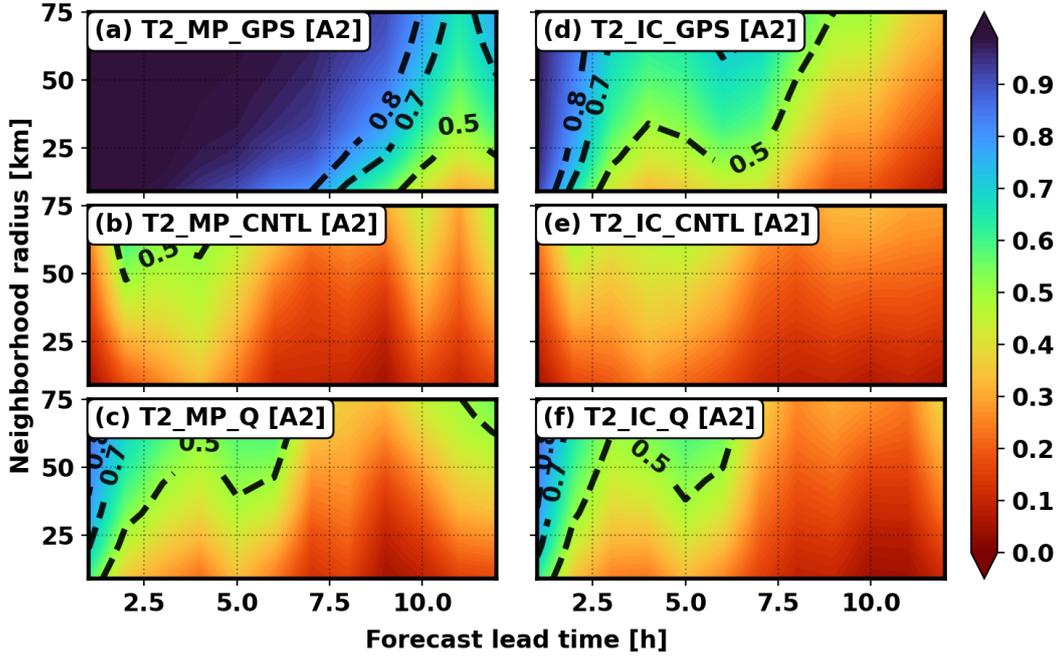


Figure 4.11: Fractional Skill Scores (10 mm threshold) at different forecast lead times and with different neighborhood sizes (km) in the A2 domain for (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. Dashed lines denote FSS values of 0.5, 0.7, and 0.8.

rapid error growth (Fig. 4.12b and 4.12d). The rapid saturation of errors at small scales limits the predictability of localized precipitation, explaining why T2_IC_GPS has FSSs drop below 0.5 for large precipitation thresholds since they are localized in nature (Fig. 4.9d). Except for T2_MP_GPS and T2_IC_GPS, the other four experiments have low FSSs right from the beginning of the forecast (Fig. 4.11), confirming the earlier conclusion that the RO DA has a positive impact that can withstand random errors.

In summary, the Test 2 experiments demonstrate that the impact of initial condition uncertainty is much larger than model uncertainty within the 12-hour timeframe. Additionally, this study supports the conclusions of Kuo et al. (2023) that assimilating RO data enhances short-range QPFs. Despite the sensitivity to random initial errors, forecasts with RO DA statistically outperform those without.

4.6 Assessment of practical predictability

The previous section has shown that the positive impact of RO DA can tolerate small random errors. However, it raises the question of whether this scenario is practically predictable given realistic analysis uncertainty. To address this question, this section explores the predictability of this case given the current uncertainty in the GEFS analysis. The

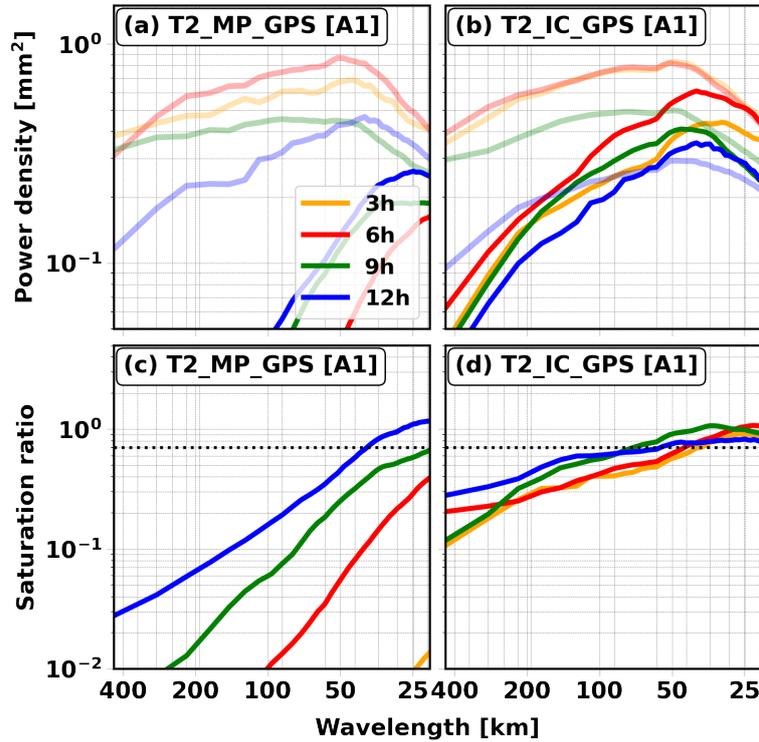


Figure 4.12: Background precipitation (light solid) and perturbation precipitation spectra (dark solid) averaged over ten ensemble members of (a) T2_IC_GPS and (b) T2_MP_GPS at lead times of 3, 6, 9, and 12 h. The spectra are computed over the A1 domain. Lower panels (c and d) depict the corresponding saturation ratios at different lead times.

discussion also covers how decreasing analysis uncertainty will lead to changes in predictability. The experiments formulated here echo those conducted by Zhang et al. (2019), who utilized initial perturbations from an ensemble DA system to represent current realistic initial condition uncertainties. They allocated 10% of these perturbations to approximate nearly perfect initial conditions, aiming to elucidate the intrinsic predictability limit of the atmosphere.

As shown in Fig. 4.13a, the precipitation intensifies over the first 3 to 6 hours, reaching its intensity peak at 6 hours, and weakening afterward — a pattern in alignment with our visual observations (Fig. 4.3). In T3_IC_p100 (Fig. 4.13d), error saturation occurs at around 50 km at the peak precipitation hour (6h), implying that localized precipitation smaller than this scale tends to be unpredictable given the current analysis uncertainty. Due to the small-scale nature of extreme rainfall, the model statistically cannot deliver useful 12-hour QPF for heavy rainfall exceeding 150 mm (Fig. 4.14a). Halving the initial uncertainty slows down the upscale error growth (Fig. 4.13e), leading to a significant uptick in the 12-h QPF skills. As shown in Fig. 4.14b, T3_IC_p50 can make useful predictions for all thresholds except those around 200 mm with small neighborhood sizes. Further reducing the initial spread to 10% of the current level sustains predictability for scales

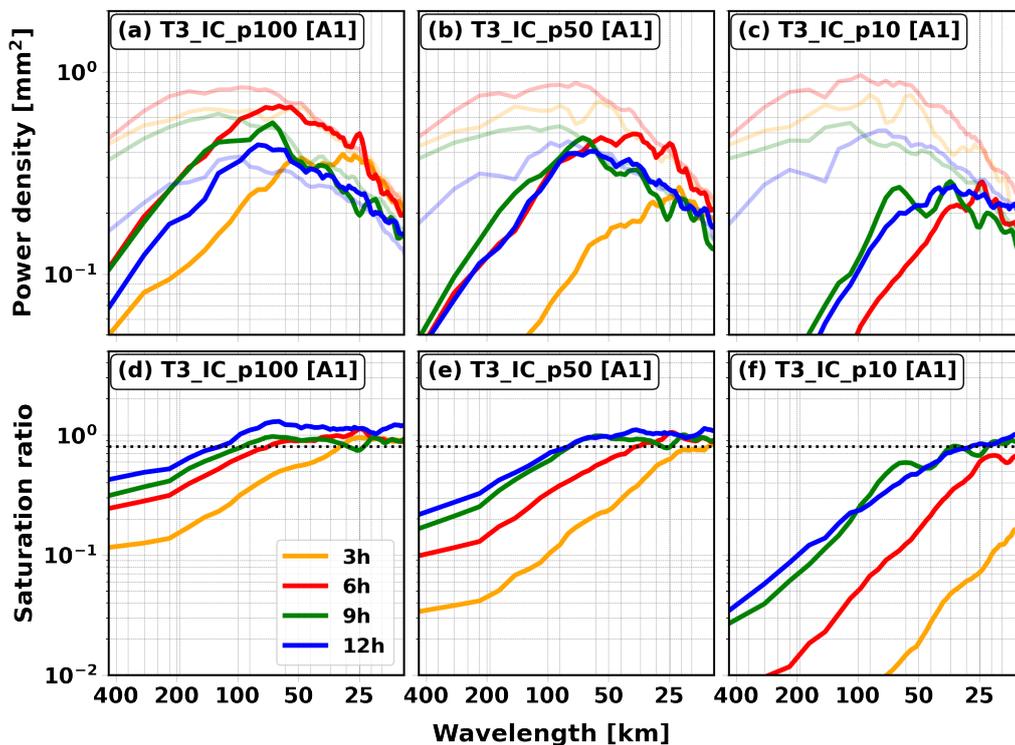


Figure 4.13: Same as Fig. 4.12 but for (a, d) T3.IC_p100, (b, e) T3.IC_p50, and (c, f) T3.IC_p10.

above approximately 30 to 40 km within the 12-hour timeframe. As a result, the T3.IC_p10 yields useful predictions across all thresholds considered (Fig. 4.14c). In summary, given the current analysis uncertainty, localized precipitation with scales smaller than about 50 km has limited practical predictability during the peak precipitation hour. However, the results indicate that the 12-hour QPF may be highly predictable, as even halving the analysis uncertainty significantly improves the forecast accuracy.

4.7 Summary and discussion

In a recent study, Kuo et al. (2023) demonstrated that assimilating RO data (referred to as GPS) into a convection-permitting model improves the prediction of a prefrontal squall line associated with a Meiyu front. While their results are promising, it is based on a case study, and the robustness of the conclusion requires further investigation. As a follow-up to Kuo et al. (2023), this study found that the moisture correction made by RO DA is the key factor leading to the successful prediction of the GPS in contrast to the CNTL forecast (which did not assimilate RO data). This aligns with the conclusion drawn by Luo and Chen (2015), which suggests the crucial role of the moisture field in predicting Meiyu front-associated precipitation.

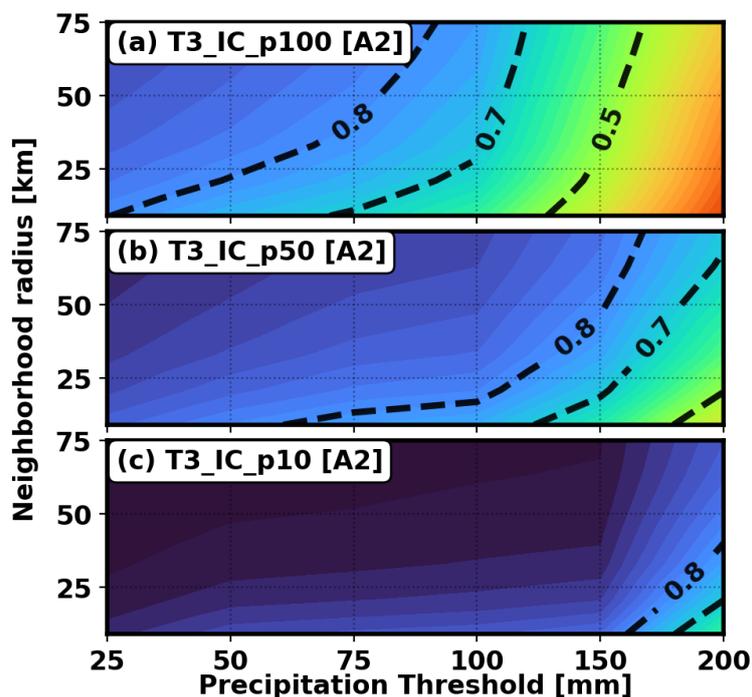


Figure 4.14: Same as Fig. 4.9 but for (a) T3_IC_p100, (b) T3_IC_p50, and (c) T3_IC_p10.

To measure if the conclusion drawn by Kuo et al. (2023) is robust or merely coincidental, this study samples random initial and model uncertainty to generate ten ensemble members. Within the 12-hour timeframe, precipitation forecasts are largely unaffected by model uncertainty but are sensitive to initial condition uncertainty. Despite the potential degradation caused by random initial errors, forecasts initialized with RO DA consistently outperform those without it. This experiment suggests that the positive impact of RO DA is not merely coincidental; rather, the correction induces a distinct rainband evolution that remains robust when accounting for random errors in the forecast procedures. Additionally, this study assesses the practical predictability of this event by taking into account the realistic analysis uncertainty measured by the operational Global Ensemble Forecast System (GEFS). Given the current analysis uncertainty, the predictability of localized precipitation is constrained within a short timeframe. Specifically, the precipitation error saturates at 50 km at 6 hours and above 100 km at the end of the forecast, respectively. The verification of 12-hour QPF skills confirms that the model cannot generate useful forecasts for heavy rainfall exceeding 150 mm due to the inherently small-scale nature of such events. As we reduce the initial spread to 50% and 10% of the current level, the results suggest that this event can be highly predictable when the analysis uncertainty is sufficiently small (e.g., to 10% of the current level).

In conclusion, this study affirms the positive impact of assimilating RO data for the

prediction of convective precipitation. Due to computational limitations, our results are derived from ensemble forecast experiments rather than full ensemble DA experiments. Further exploration is warranted as ensemble DA can provide flow-dependent initial uncertainty. Most importantly, this chapter reveals that in short-range forecasts, perturbations in initial conditions wield more influence than uncertainty in the model processes. In particular, the SPPT scheme utilized in this study functions as a more holistic, bulk representation of model error compared to the process-level PSP and SPPMP schemes discussed in Chapter 3. This finding suggests that the difference between PSP and SPPMP in short lead times (illustrated in Chapter 3) may not be significant in the presence of random initial errors.

Chapter 5

Summary and conclusions

5.1 Summary

Although the progress of Numerical Weather Prediction (NWP) models has been undeniably successful in recent decades, inevitably, every forecast will eventually encounter its predictability limit. In convection-permitting forecasts, this predictability limit becomes increasingly evident as their forecast errors grow several times faster than those of synoptic-scale forecasts, posing considerable uncertainty in predicting convective storms. The overarching goal of this dissertation is to “**advance the comprehension of regime-dependent predictability and forecast error growth in convection-permitting models**”. To achieve the overarching goal, three individual topics have been thoroughly investigated, and their findings are summarized as follows:

(1) **Characterization of Convective Regimes and Their Predictability**

The first study aims to classify different weather regimes and identify regime-dependent predictability in an objective way. As demonstrated by studies focusing on Europe, the convective adjustment time scale (τ_c) can serve as an effective classification measure, which can be further linked with properties such as precipitation predictability. However, the value of τ_c may vary based on the geographical region and the complexities involved in its computation, meaning that the classification thresholds suggested in the literature may not be directly applicable for categorizing weather regimes in different regions. To bridge this gap, the first study systematically evaluated the effectiveness of this diagnostic metric across the Contiguous United States (CONUS) domain, utilizing four-year summertime forecasts from the convection-permitting High-Resolution Rapid Refresh (HRRR) system. This exploration answers the question of “**How to effectively classify convective regimes over the vast CONUS domain and can this classification reflect physical understanding and regime-dependent predictability?**”, as outlined in Chapter 1.4. This study establishes the first convective adjustment time scale climatology for the CONUS that reveals the climatological frequency of equilibrium and nonequilibrium convection. Due to the vast size of the CONUS domain that can encompass multiple weather regimes

simultaneously, eight subdomains divided into north and south of approximately 40°N latitude were designed. Distinct climatological differences between subdomains reveal regional factors influencing regime classification, which can explain differences observed in previous studies. Climatologically, the northern subdomains exhibit more equilibrium convection due to the dominance of trough-ridge patterns in the planetary waves. In general, around 53% of convection is synoptic-driven, a proportion similar to that observed over Germany but lower than that over the British Isles. In contrast, the southern subdomains are characterized by subtropical anticyclones and exhibit more nonequilibrium convection, with only 22% of convection being linked with synoptic forcing. The classification demonstrates physical consistency when examining its seasonal and diurnal dependencies. Precipitation patterns with a diurnal cycle are approximately twice as likely to be identified in nonequilibrium regimes, aligning with the hypothesis that convection classified as nonequilibrium is modulated by local factors such as thermal forcing. As the τ_c value indicates where the case falls along the continuum between the equilibrium and nonequilibrium extremes, this study links the τ_c value with precipitation skill metrics (e.g., Fractional Skill Score and Gilbert Skill Score) to explore whether forecasts from the convection-permitting HRRR model exhibit regime-dependent precipitation predictability. The results demonstrate modest negative correlations between the two vectors, indicating that convection with stronger synoptic coupling is more predictable than those without.

(2) Regime-Dependent Forecast Error Growth in Convection-Permitting Models: The Roles of Stochastic Parameterizations

The second study delves into the question: “**Does forecast error growth significantly differ across various weather regimes and with different stochastic parameterizations in a convection-permitting model?**” Using the regime classification diagnostic established in the prior study, this study investigates error growth characteristics across different weather conditions, including winter storms, strongly forced convection, and weakly forced convection. In general, winter storm cases demonstrate higher predictability compared to summertime convection cases, as evidenced by their notably lower saturation levels in both precipitation and kinetic energy errors, even at small scales (< 50 km). On the contrary, forecast errors saturate rapidly at small scales in all summer convection cases. Especially in nonequilibrium situations, the precipitation error saturates within one day for all scales below about 200 km, while the kinetic energy error only saturates for scales below 100 km. This difference highlights the challenging nature of precipitation prediction in scenarios where there is no dominant synoptic flow. Additionally, this study raises the question of whether forecast error growth at meso- and convective scales is primarily influenced by the details of model error formulations or by the dynamics of atmospheric flow. To explore this question, the study compares forecast error growth resulting from two model error formulation schemes: (1) Stochastically Perturbed Parameter applied to Microphysics (SPPMP) scheme, which is designed to capture uncertainty stemming from the limited understanding of atmospheric processes, and (2) Physically based Stochastic Perturbation (PSP) scheme, which aims to represent uncertainty arising from finite grid

size. This dissertation implemented the PSP scheme, originally developed for the ICOSahedral Nonhydrostatic (ICON) and Consortium for Small-scale Modeling (COSMO) models, into the Weather Research and Forecasting (WRF) model and has confirmed its consistent behavior across the three NWP systems. In general, forecast error growth at longer lead times is less sensitive to the details of model error formulation and is primarily influenced by atmospheric flow. Specifically, both schemes yield similar forecast error distributions at longer lead times, albeit with differences in amplitude. This elucidates why previous studies argue that the spreads of multiple model perturbation schemes are usually not additive, as their effects are not orthogonal. Overall, the PSP scheme leads to larger forecast error growth, and its combination with the SPPMP scheme does not introduce additional forecast error compared to using the PSP alone. Exceptions can occur in certain situations, for example, the SPPMP scheme may facilitate additional forecast error growth for winter storm cases as the boundary layer is less turbulent. In contrast to their high similarity at longer lead times, the forecast errors induced by the two schemes exhibit pronounced differences at short lead times, approximately before the forecast reaches the first local afternoon time. Most of the time, the PSP scheme induces faster error growth than the SPPMP scheme at short lead times, with its impact converging with that of SPPMP as the lead time increases. For forecast initialized during nighttime, the SPPMP scheme can potentially induce instability earlier in regions with existing precipitation when the PSP scheme is constrained by the stable nocturnal boundary layer. This has practical implications for the development of stochastic physics suites, emphasizing the importance of microphysics uncertainty for those focusing on rapid-update ensemble Data Assimilation (DA) or short-range ensemble forecast systems. The model uncertainty schemes studied were configured to represent physically justifiable uncertainties. It is conceivable that the SPPMP scheme could have a comparable impact to the PSP scheme, even during the afternoon convective peak, but this would require parameters surpassing a reasonable range within the framework of our current scheme.

(3) Predictability of a Heavy Precipitation Event: The Impacts of Initial and Model Uncertainty

The previous study concentrates on discussing two process-level model errors. Yet, to be truly comprehensive, an ensemble system must address initial condition uncertainty and model uncertainty simultaneously. As outlined in Chapter 1.4, the third study aims to identify: **What is the impact of initial and model uncertainty on precipitation predictability and forecast error growth in a convection-permitting model?** This study serves as a follow-up study to Kuo et al. (2023) and focuses on a heavy rainfall event associated with a Meiyu front. Using the initial conditions provided by Kuo et al. (2023), this study generated ten ensemble members by considering random initial and model errors. Overall, within a 12-hour lead time, introducing random initial moisture errors results in markedly faster error growth compared to representing model uncertainty using the Stochastically Perturbed Parameterization Tendency (SPPT) scheme. In this case, the model 12-hour and hourly Quantitative Precipitation Forecasting (QPF) skills

are not sensitive to model perturbations, but the impact of initial perturbations is noticeable. Since the SPPT scheme is a bulk and more holistic model error representation scheme, it is anticipated to be more effective than the process-level SPPMP and PSP schemes used in Chapter 3. In this study, the limited sensitivity of the SPPT scheme implies that while different model error formulations may indeed result in distinct forecast error growth at short lead times (as discussed in Chapter 3), their distinctions may not be significant compared to the uncertainty introduced by initial conditions. To validate this speculation, further error growth studies with a more comprehensive representation of errors are warranted. Another research objective of this study is to complement the findings of Kuo et al. (2023), which demonstrated that assimilating Radio Occultation (RO) data can improve the model QPF skills for convective precipitation. While the improvement is encouraging, it is based on a deterministic NWP system and a case study, requiring further investigation to ascertain its robustness. Despite potential degradation caused by random initial errors, this study confirmed that forecasts with RO DA on average outperform those without, suggesting that the RO DA leads to a distinct flow that remains robust when accounting for random errors in the forecast procedures. While the positive impact of RO DA is proved to withstand small random errors, it raises the question of whether this scenario is practically predictable when considering realistic analysis uncertainty. Given the current large-scale analysis uncertainty as measured by the Global Ensemble Forecast System, the predictability of localized precipitation ($< 50\text{km}$) is limited within the short timeframe. In general, the model cannot provide useful forecasts for heavy rainfall exceeding 150 mm due to the small-scale nature of such events. When reducing the initial uncertainty to 50% and 10% of the current level, the rapid increase in predictability suggests that it falls within the realm of practical predictability, and the intrinsic predictability of this case could be relatively high.

5.2 Conclusions

The dissertation addresses three research topics, each offering unique insights into the predictability and forecast error growth in convection-permitting models. These findings can have a concrete impact on the research and operational community focusing on high-resolution short-range NWP forecasts for convection prediction. Specifically, the impacts of this dissertation include:

1. Offering a detailed guidance on how to employ the convective adjustment time scale to classify weather regimes over a vast domain. For example, this dissertation provides a valuable framework for future studies focusing on the United States.
2. Demonstrating the potential of using the convective adjustment time scale as an indicator of convection predictability, which typically assessed through ensemble spread in contemporary analyses.

3. Advancing our understanding of regime-dependent predictability and forecast error growth characteristics, including strong forcing convection, weak forcing convection, and winter storms.
4. Advancing our knowledge on the effects of different stochastic parameterizations in convection-permitting models for various weather systems and forecast lead times.
5. Supporting the concept that assimilating large-scale observations in convection-permitting models can be crucial in enhancing the predictability of convection.
6. Highlighting the significance of initial perturbations compared to model perturbations for forecast error growth at short lead times.

Despite the contributions outlined here, there are several limitations associated with the results presented. Due to the computational limitation, all ensemble experiments in this dissertation are conducted with a limited ensemble size, meeting the minimum requirement for ensemble sensitivity testing as suggested in the literature. Ideally, employing a larger ensemble size would improve the robustness of the statistics obtained. While this dissertation selected cases covering various weather types and regions, the number of cases presented here may not encompass all situations. For example, it may be worthwhile to explore the impact of different model perturbations for non-cloudy cases or cases with intermittent precipitation occurring later in the forecast period. Additionally, the PSP and SPPMP schemes discussed in this dissertation are employed in specific parameterization schemes, and their effects may not be identical when implemented in other parameterization processes. While this dissertation examines the predictability and forecast error growth in various model states, there are still other aspects that warrant further investigation. For example, exploring forecast error growth of cloud variables could be valuable as this metric is important for satellite and radar observation assimilation. While the findings presented in this dissertation can inform the development of convective-scale EPSs, the impact of the stochastic schemes should be further validated within a more comprehensive system that incorporates both initial and boundary condition uncertainties. Furthermore, activating the stochastic schemes during the DA cycling is advisable, as they are expected to address the under-dispersion issue commonly observed in convective-scale ensemble DA systems.

Chapter 6

Appendix

6.1 Normalized precipitation spread and NRMDTE

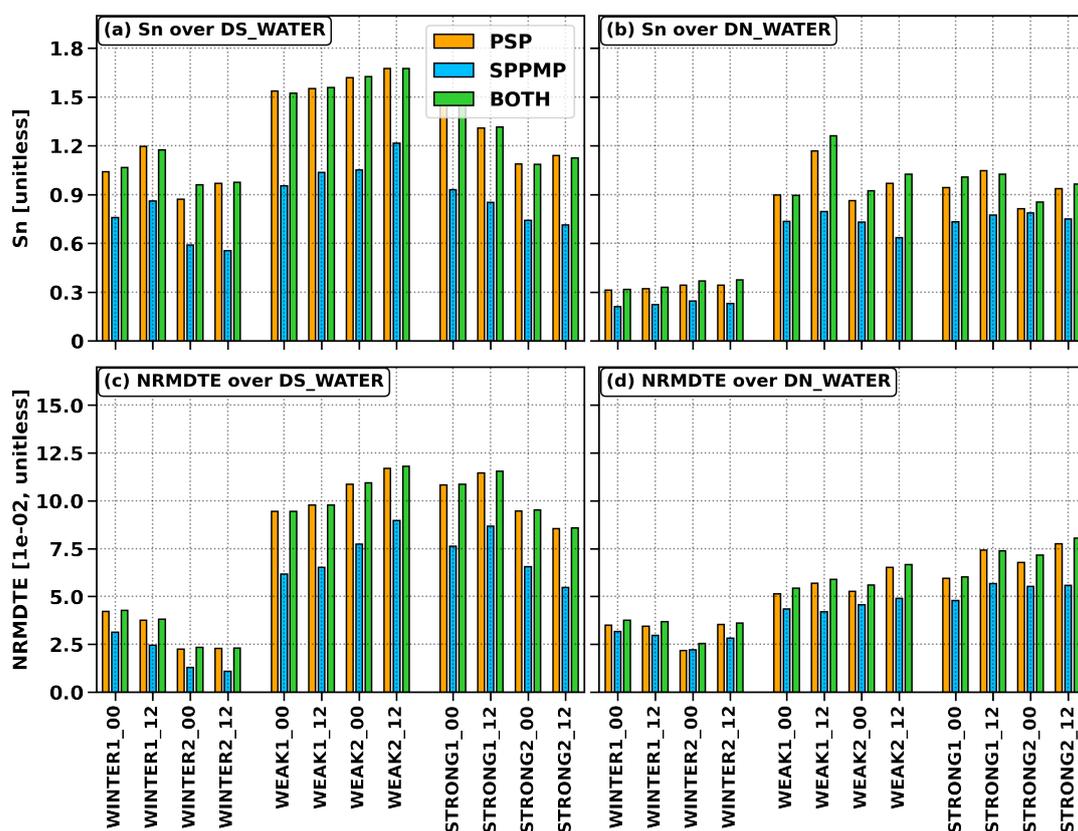


Figure 6.1: Domain-averaged normalized precipitation spread (S_n , unitless) over (a) DS_WATER and (b) DN_WATER subregions for each 48-hour forecast. Panels (c) and (d) are domain-averaged NRMDTE (unitless) for DS_WATER and DN_WATER, respectively.

6.2 Time series of precipitation spread and intensity for winter storm cases

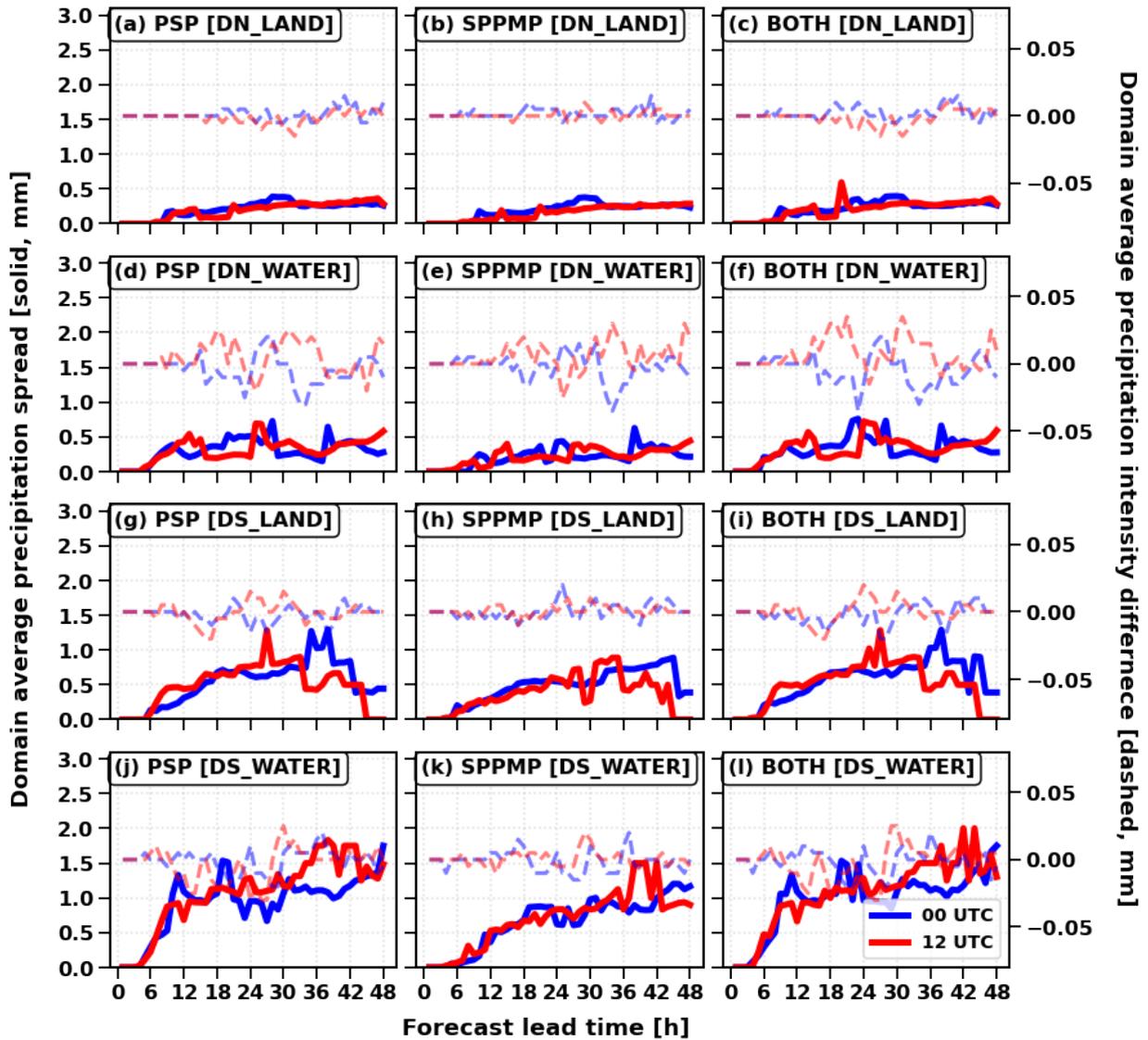


Figure 6.2: Time series of domain-averaged normalized hourly precipitation spread (solid, mm) and domain-averaged hourly precipitation difference against CNTL forecasts (dashed, mm) for PSP, SPPMP, and BOTH. The rows depict results averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Each line within the figure corresponds to two winter storm cases.

6.3 Snapshots of water vapor perturbation generated by the PSP scheme

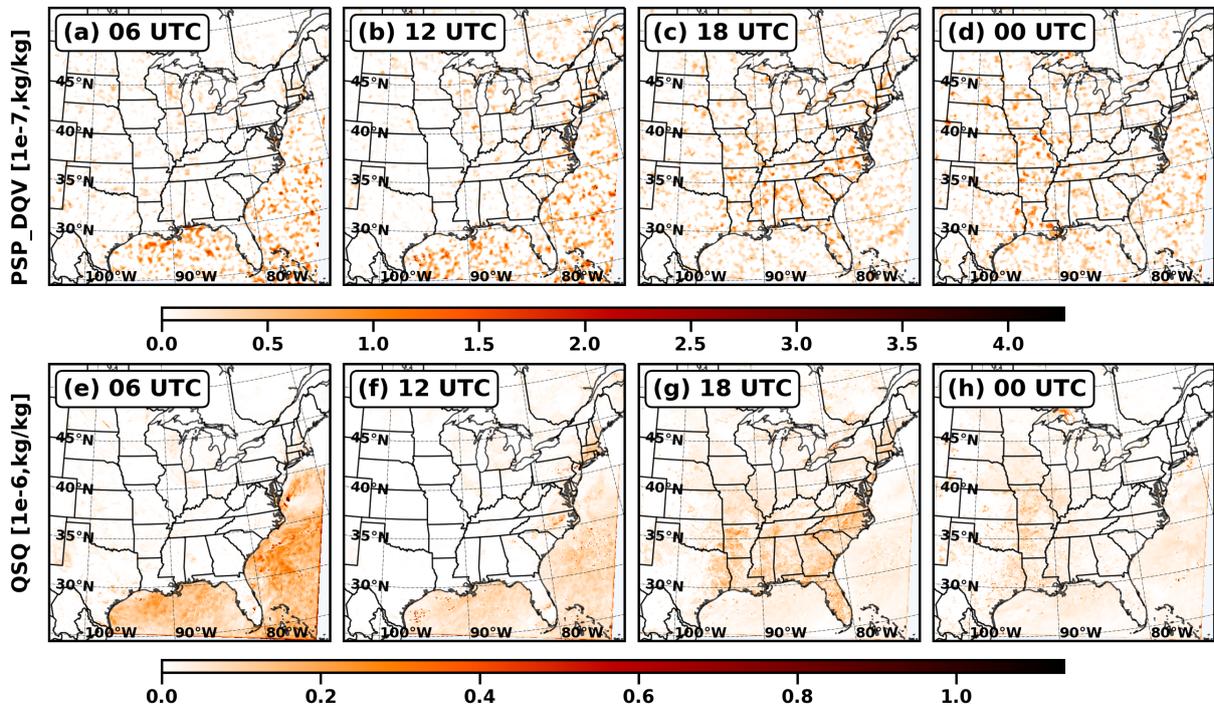


Figure 6.3: The upper panels depict the lowest-level water vapor perturbation fields ($g \text{ kg}^{-1} \text{ s}^{-1}$) generated by the PSP scheme, valid at (a) 06 UTC, (b) 12 UTC, (c) 18 UTC, and (d) 00 UTC on July 21, 2021. The lower panels depict their corresponding water vapor turbulence variance fields ($\text{kg}^2 \text{ kg}^{-2}$). These forecasts are from the first member in the PSP experiment initialized at 00 UTC on the same day.

6.4 Time series of RMDTE growth rate for winter storm cases

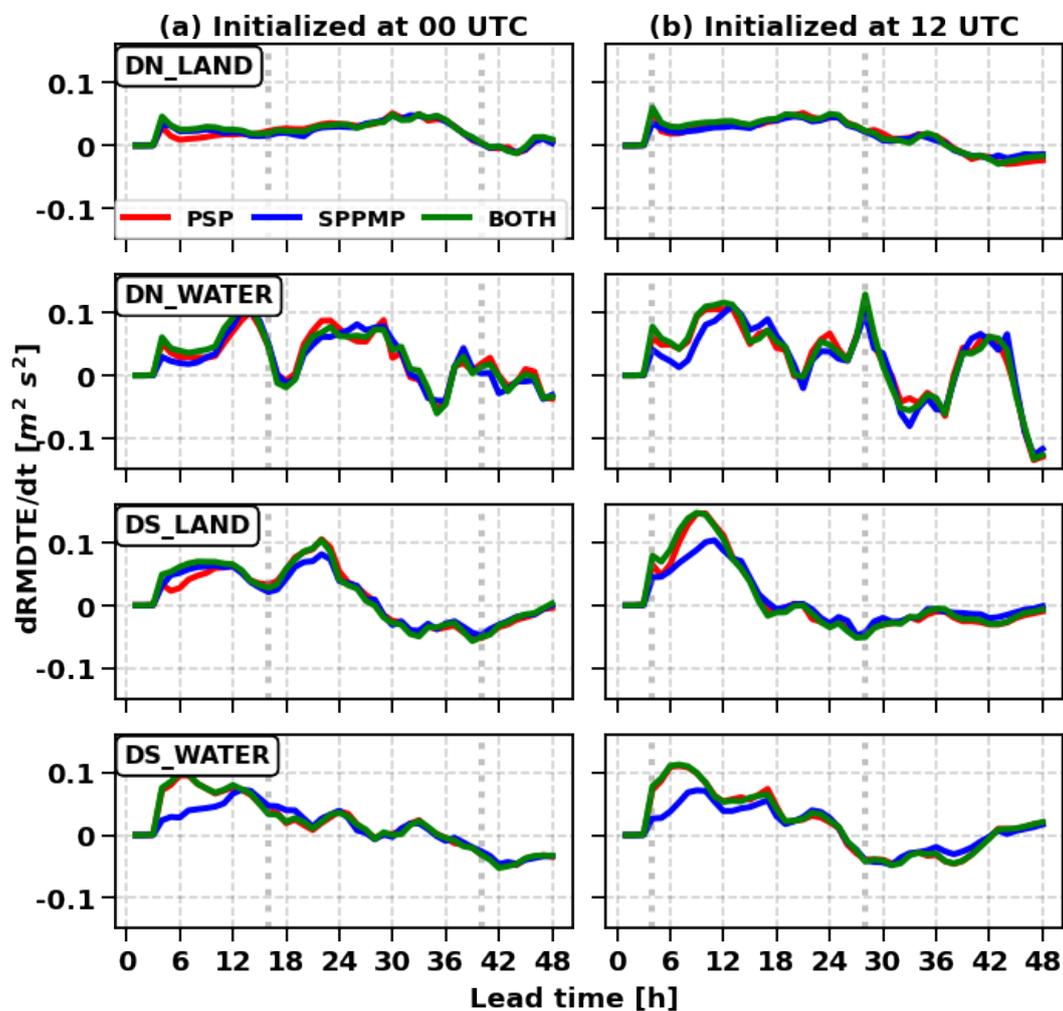


Figure 6.4: Time series of domain-averaged root-mean difference total energy (RMDTE, $m^2 s^2$) growth rate for PSP (red), SPPMP (blue), and BOTH (green). Rows are $dRMDTE/dt$ averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show forecasts initialized at (a) 00 UTC and (b) 12 UTC. Each line within the figure corresponds to two winter storm cases. Dashed lines indicate the 12 LST corresponding to each initial time.

6.5 RMDTE ratios for strong forcing convection, weak forcing convection, and winter storm cases

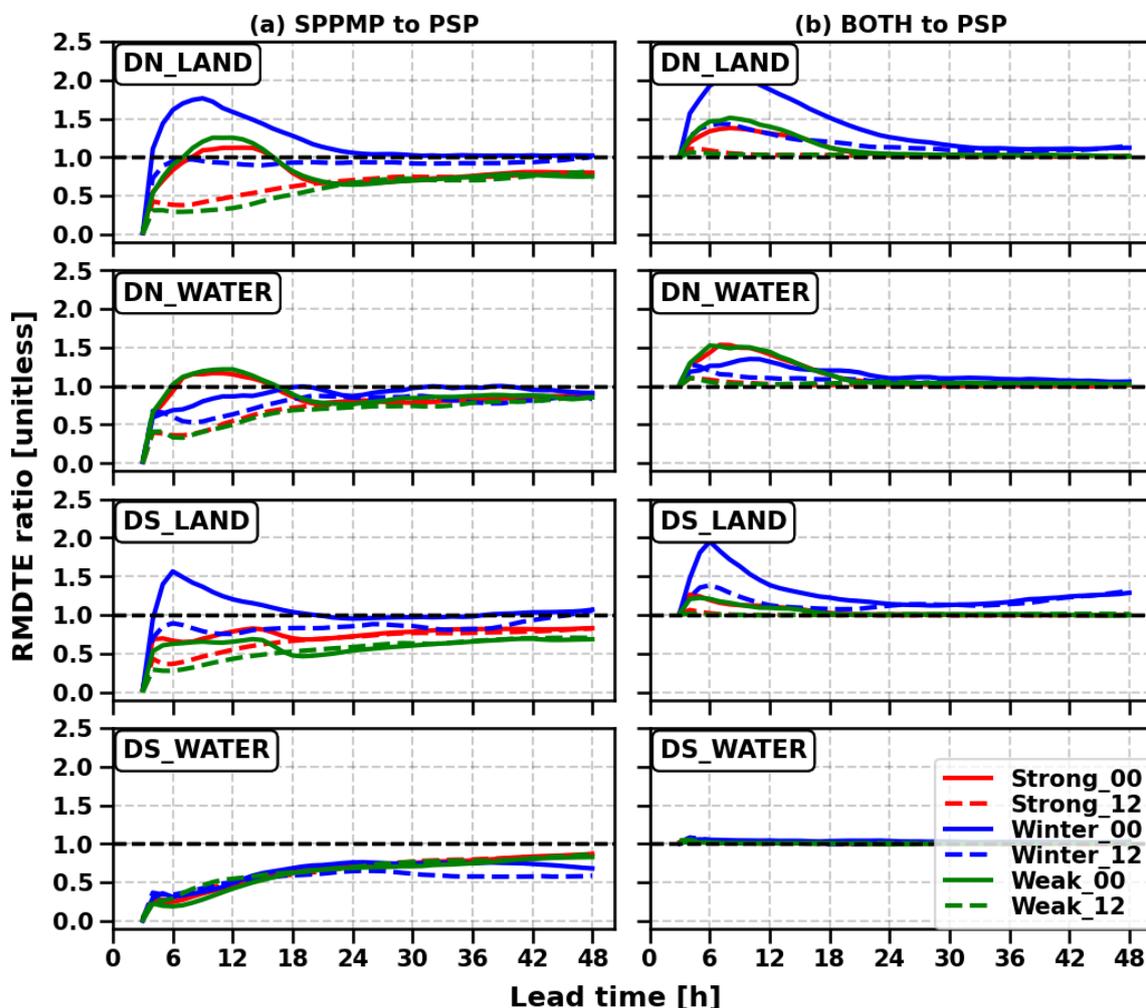


Figure 6.5: Time series of domain-averaged RMDTE ratios, averaged over forecasts of each weather type (colors) initialized at 00 UTC (solid lines) and 12 UTC (dashed lines). Rows show averages over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show ratios computed as (a) SPPMP divided by PSP and (b) BOTH divided by PSP. Each line within the figure corresponds to two cases.

Acknowledgements

My doctoral research involves scientific collaboration between Dr. George Craig and Dr. Christian Keil from LMU Munich, along with Dr. Bill Kuo and Dr. Judith Berner from the University Corporation for Atmospheric Research (UCAR). I am grateful for the opportunity to conduct research with all of you guiding me over the years. Furthermore, I am grateful to the Waves to Weather and the UCAR COSMIC program for the opportunities they have given me.

I want to express my gratitude to Dr. George Craig for his invaluable mentorship and consistent support for my doctoral research plan over the years. I also want to thank Dr. Christian Keil for addressing the numerous questions I've reached out to him with. I appreciate you and George for accommodating the time difference between Munich and Boulder, even though it often resulted in late meetings. Although the pandemic made things tough in the beginning, I got to meet many great colleagues in the LS-Craig team. Their research expertise is impressive, and their assistance has made my transition to Munich much smoother.

During my time at UCAR, many scientists there gave me really helpful advice. I want to thank Dr. Bill Kuo for his invaluable guidance in my professional development, for allowing me to explore research interests, and for sharing his life experiences with everyone. I am deeply grateful to have Dr. Judith Berner as an exceptional academic mentor, and moreover, as an incredibly supportive friend. In the past two years, my friends in Boulder have become my go-to companions. The regular hangouts have brought so much happiness to our research-filled lives.

The expertise I gained at the Central Weather Administration and the National Central University significantly benefits my doctoral research. Most importantly, the friendships formed at both places have been incredibly strong, providing abundant support and joy over the years. Above all, I am thankful for every one of my family members. My parents have always been there for me, supporting my choices, even when I haven't shared all the details with them. I'm grateful for my brother being there and keeping me informed about what's happening while I'm abroad. I am thankful for all their efforts in creating an environment that allows me to pursue my goals without any concerns during the past years. **Thank you.**

List of Abbreviations

3DEnVar	Three-Dimensional Ensemble-Variational
CAPE	Convective available potential energy
CIN	Convective inhibition
CMR	Central Mountain Range
CONUS	Contiguous United States
COSMO	Consortium for Small-scale Modeling
DA	Data Assimilation
DCT	Discrete Cosine Transform
ECMWF	European Centre for Medium-Range Weather Forecasts
EL	Equilibrium Level
EPS	Ensemble Prediction System
FSS	Fractional Skill Score
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
GNSS	Global Navigation Satellite System
GSI	Gridpoint Statistical Interpolation
GSS	Gilbert Skill Score
HRRR	High-Resolution Rapid Refresh
ICON	ICOsahedral Nonhydrostatic
KE	Kinetic Energy
LCL	Lifting Condensation Level
LFC	Level of Free Convection
LT	Local Time
MJJA	May, June, July, and August
MYNN	Mellor–Yamada–Nakanishi–Niino

NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NRMDTE	Normalized Root Mean Difference Total Energy
NWP	Numerical Weather Prediction
PBL	Planetary Boundary Layer
PDF	Probability Density Function
PSD	Particle Size Distribution
PSP	Physically based Stochastic Perturbation
QPE	Quantitative Precipitation Estimation
QPESUMS	Quantitative Precipitation Estimation and Segregation Using Multiple Sensors
QPF	Quantitative Precipitation Forecasting
RMDTE	Root Mean Difference Total Energy
RO	Radio Occultation
SKEB	Stochastic Kinetic Energy Backscatter
SPP	Stochastic Perturbed Parameter
SPPMP	Stochastically Perturbed Parameter applied to Microphysics
SPPT	Stochastically Perturbed Parameterization Tendency
SVM	Support Vector Machine
WRF	Weather Research and Forecasting
WRFDA	WRF Data Assimilation
WRF-PSP	WRF implementation of the PSP scheme
YSU	Yonsei University

List of Figures

1.1	A schematic diagram illustrating typical time and length scales of various weather phenomena. The figure is sourced from Brisch and Kantz (2019), and the original publication can be accessed at https://doi.org/10.1088/1367-2630/ab3b4c	4
1.2	An idealized schematic illustrating the distinction between (a) practical and (b) intrinsic predictability. The figure is sourced from Melhauser and Zhang (2012), and the original publication can be accessed at https://doi.org/10.1175/JAS-D-11-0315.1 . ©American Meteorological Society. Used with permission.	5
1.3	Amplitude of background KE (the thick gray line) and KE error (dashed lines) as a function of horizontal scale for (a) Lorenz’s experiment A and (b) Lorenz’s experiment B at various lead times. The figure is sourced from Durran and Weyn (2016), and the original publication can be accessed at https://doi.org/10.1175/BAMS-D-15-00070.1 . ©American Meteorological Society. Used with permission.	7
1.4	A probabilistic forecasting method that considers uncertainties stemming from initial conditions and model formulations. The figure is sourced from Slingo and Palmer (2011), and the original publication can be accessed at https://doi.org/10.1098/rsta.2011.0161	9
1.5	(a) A thermodynamic diagram illustrating atmospheric stability. (b) The yellow line illustrates the environmental sounding, with a black circle marking the initial position of an air parcel near the surface. Red solid and blue dashed lines delineate the dry and moist adiabatic lapse rates. Shaded regions denote CIN and CAPE, respectively.	11
1.6	System characterized by (a), (c) double-potential or (e),(g) single-potential well. (b), (d), (f), and (h) are their associated PDFs when encountering forcing. The figure is sourced from Berner et al. (2017), and the original publication can be accessed at https://doi.org/10.1175/BAMS-D-15-00268.1 . ©American Meteorological Society. Used with permission.	21
2.1	Frequency distributions of the convective adjustment time scale shown annually (colors) and 4y-mean (black) for the entire CONUS domain. The dashed vertical lines indicate 3h, 6h, and 12h, respectively.	27

2.2	Verification and diagnostic subdomains used in the NOAA model QPF verification by regions (shaded) and this study (contour). The black numbers denote the subdomain indexes.	28
2.3	Monthly total precipitation (shaded, mm) in May and June from 2019 to 2022 derived from the HRRR 3-27h forecasts (denoted as F) and NCEP StageIV products (denoted as O). Rows are the result of different years, and columns are the result of forecasts (the first and the third columns) and observations (the second and the fourth columns) at different months. Black contours are the subdomain boundaries used in this study.	29
2.4	Same as Fig.2 but for July and August.	30
2.5	(a) 500 hPa geopotential height (contour, m) and wind speed (shaded, $m s^{-1}$) valid at 1200 UTC 01 July 2021. (b) 24h accumulated precipitation (shaded, mm) beginning at 1200 UTC 01 July 2021. Black contours are subdomain boundaries. (c) Time series of hourly precipitation (upper, mm), CAPE (middle, $J kg^{-1}$), and convective adjustment time scale (lower, hours) averaged over different subdomains. The dashed line in the convective adjustment time scale (τ_c) plot marks the 6h threshold between equilibrium and nonequilibrium convection.	31
2.6	(a) 500 hPa geopotential height (contour, m) and wind speed (shaded, $m s^{-1}$) valid at 2000 UTC 27 July 2021. (b) 24h accumulated precipitation (shaded, mm) beginning at 1200 UTC on 27 July 2021. (c) Time series of hourly precipitation (upper, mm), CAPE (middle, $J kg^{-1}$), and convective adjustment time scale (lower, hours) averaged over different subdomains. The dashed line in the convective adjustment time scale (τ_c) plot marks the 6h threshold between equilibrium and nonequilibrium convection.	33
2.7	Frequency distributions of the convective adjustment time scale shown annually (colors) and 4y-mean (black) for each subdomain. The frequency is computed with a bin size of 0.2h, and lines are smoothed by a Gaussian filter. The dashed vertical lines indicate 3h, 6h, and 12h, respectively.	34
2.8	Cumulative percentage of convective cases with τ_c less than 1, 3, 6, 12, and 24h for each subdomain. τ_c not meeting the thresholding criteria in section 2a are excluded. The red dashed lines indicate the percentage of equilibrium cases ($\tau_c < 6h$) averaged over all subdomains. The blue dashed line in the upper and lower panel indicates the percentage of equilibrium cases ($\tau_c < 6h$) averaged over northern and southern subdomains, respectively.	36
2.9	Frequency distributions of convective adjustment time scale stratified by calendar month. All other details as in Fig.2.7.	37
2.10	Probability of precipitation maxima in a hundred of precipitation time series subjectively labeled as (a) diurnal precipitation and (b) nondiurnal precipitation. The dataset used in this figure consists of 200 time series data.	38

2.11	Probability of precipitation maxima in precipitation time series classified as (a) diurnal precipitation and (b) nondiurnal precipitation for each subdomain. The figure comprises a complete four-year MJJA dataset excluding those associated with tropical systems, resulting in slight variations in the exact number across different subdomains.	39
2.12	Identification of diurnal precipitation pattern in subdomain-averaged rainfall time series. The (a) equilibrium convection and (b) nonequilibrium convection are classified by a τ_c threshold of 6 hours. Bars represent probabilities of precipitation maxima, while annotations correspond to fractions of cases identified as having a diurnal pattern. The figure comprises a complete four-year MJJA dataset excluding those associated with tropical systems, resulting in slight variations in the exact number across different subdomains.	40
2.13	Scatterplots of convective adjustment time scale (h) and Fractional Skill Score (FSS) for the eight subdomains. Annotations denote the correlation coefficient computed for each subdomain. The FSS presented here used a squared neighborhood of 25 grid points and a binary threshold of $5mm\ 24h^{-1}$	41
2.14	As Fig. 13, but for Gilbert Skill Score (GSS).	42
3.1	Configuration of the domain (black box) for the WRF simulations in this study. Diagnostic domains DN (light blue, 511x1023 grid points) and DS (light green, 512x1023 grid points) are partitioned along $\sim 38^\circ N$ parallel. A smaller diagnostic domain LAND (light orange, 300x600 grid points) is chosen to cover land points only.	47
3.2	Stochastic random patterns (used in all vertical levels) for the (a) SPPMP and (b) WRF-PSP schemes. (c) Water vapor turbulence variance QSQ ($kg^2\ kg^{-2}$), (d) potential temperature turbulence variance (K^2), (e) perturbed water vapor tendency ($kg^2\ kg^{-2}$), and (f) perturbed potential temperature tendency (K) at the lowest model level and valid at 18 UTC on 21 July 2021.	50
3.3	500-hPa Geopotential height (contours, m) and two-day accumulated precipitation (shaded, mm) under different flow regimes: strong forcing (top), weak forcing (middle), and winter storm (bottom). Panels show 00 UTC and 12 UTC states for six different days (see panel labels). Results are from the unperturbed forecasts.	53
3.4	Hourly accumulated precipitation rate ($mm\ h^{-1}$; solid lines; left y-axis) and convective adjustment time scale τ_c (h; dashed lines; right y-axis) for cases shown in Fig. 3.3 in the Northern (DN; blue) and Southern (DS; red) domains, respectively. Results are from the unperturbed forecasts.	54
3.5	RMDTE ($m\ s^{-1}$) and hourly precipitation ($> 0.1mm$, blue contours) at lead times of (a) 6 h, (b) 12 h, (c) 24 h, (d) 36 h, and (e) 48 h for forecast initialized at 12 UTC on 27 July 2021. Rows are PSP, SPPMP, and BOTH experiments, respectively	58

3.6	Time series of averaged RMDTE ratio (shaded, SPPMP against PSP) at pressure levels and domain average hourly precipitation rate (mm, white lines) for the 48-h forecast of the (a) WEAK2_12 and (b) WINTER2_00 cases. Rows show ratios averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively.	59
3.7	Same as Fig. 5., but for forecast initialized at 00 UTC on 02 Mar 2023. . .	60
3.8	Domain-averaged normalized precipitation spread (S_n , unitless) over (a) DS_LAND and (b) DN_LAND subregions for each 48-hour forecast. Panels (c) and (d) are domain-averaged NRMDTE (unitless) over DS_LAND and DN_LAND, respectively.	62
3.9	Spectra of difference hourly precipitation (heavy solid lines for PSP and dashed lines for SPPMP) and background hourly precipitation (light solid lines) per unit log k for (a) winter storm, (b) weak forcing, and (c) strong forcing cases. Figures (d), (e), and (f) are the same but for background KE and difference KE. Each line shows the average of four spectra (two 00 UTC and two 12 UTC runs). The difference fields are computed against unperturbed forecasts (see section 3.2.5 for details).	63
3.10	Saturation ratios of difference KE against difference precipitation at different forecast lead times (h, colors) for strong forcing (top), weak forcing (middle), and winter storm (bottom) cases, respectively. Each dot shows the average of four spectra (two 00 UTC and two 12 UTC runs). Columns are saturation ratios at 15 km, 50 km, 100 km, and 200 km wavelengths.	64
3.11	Time series of domain-averaged normalized hourly precipitation spread (solid, mm) and domain-averaged hourly precipitation difference against CNTL forecasts (dashed, mm) for PSP, SPPMP, and BOTH. The rows depict results averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Each line within the figure corresponds to four summertime convection cases.	66
3.12	Time series of domain-averaged root-mean difference total energy (RMDTE, m^2s^2) growth rate for PSP (red), SPPMP (blue), and BOTH (green). Rows are $d \text{ RMDTE}/dt$ averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show forecasts initialized at (a) 00 UTC and (b) 12 UTC. Each line within the figure corresponds to four summer convection cases. Dashed lines indicate the 12 LST corresponding to each initial time.	68
3.13	Time series of domain-averaged RMDTE ratios averaged over all cases initialized at 00 UTC (blue) and 12 UTC (red). Rows show averages over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show ratios computed as (a) SPPMP divided by PSP and (b) BOTH divided by PSP. Dashed lines indicate the 12 LT corresponding to each initial time (blue for 00 UTC and red for 12 UTC). Each line within the figure corresponds to six cases.	69

4.1	Geopotential height (contour, m), horizontal wind (vector $m s^{-1}$), and water vapor mixing ratio (shaded, $g kg^{-1}$) at (a) 500 hPa, (b) 700 hPa, (c) 850 hPa, and (d) 950 hPa, derived from the ECMWF Reanalysis v5 valid at 00 UTC on May 22, 2020.	78
4.2	The 12-hour accumulated precipitation (shaded, mm) of (a) QPESUMS QPE, (b) CNTL, and (c) GPS ends at 12 UTC on May 22, 2020. The CNTL and GPS forecasts depicted here are based on data from Kuo et al. (2023).	79
4.3	The hourly precipitation (shaded, mm) of (a) QPESUMS QPE, (b) CNTL, and (c) GPS at various forecast lead times (columns) initialized at 12 UTC on May 22, 2020. The CNTL and GPS forecasts depicted here are based on data from Kuo et al. (2023).	80
4.4	(a) Configuration of the 15 km (D01, blue box) and 3km (D02, red box) domains for the WRF simulations in this study. (b) Configuration of three diagnostic domains (denoted as A1, A2, and A3).	81
4.5	Vertical profiles of the initial moisture spread (solid) and their fraction relative to the background moisture (dashed) for D01 (blue) and D02 (red) domains.	83
4.6	Fractional skill scores of different precipitation thresholds ($mm 12 h^{-1}$) and neighborhood sizes (km) in the A2 domain for (a) T1_UVTQ, (b) T1_Q, (c) T1_T, and (d) T1_UV. Panels e-h represent the same analysis but for the A3 domain.	85
4.7	Distribution of 12-hour QPF for (a) T1_UVTQ, (b) T1_T, (c) T1_Q, and (d) T1_UV initialized at 00 UTC on May 22, 2020.	86
4.8	The time-latitude diagram computed over the A3 domain for hourly precipitation of (a) CNTL, (b) GPS, (c) T1_UVTQ, (d) T1_T, (e) T1_UV, and (f) T1_Q. The black solid line indicates the FSSs of hourly precipitation verified against GPS.	87
4.9	Fractional Skill Scores of different precipitation thresholds ($mm 12 h^{-1}$) and neighborhood sizes (km) computed over the A2 domain for (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. Dashed lines denote FSS of 0.5, 0.7 and 0.8.	88
4.10	Probability of 12-hour QPF larger than 150 mm, derived from the 10 members of (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. The neighborhood probability is computed using 25 grids.	89
4.11	Fractional Skill Scores (10 mm threshold) at different forecast lead times and with different neighborhood sizes (km) in the A2 domain for (a) T2_MP_GPS, (b) T2_MP_CNTL, (c) T2_MP_Q, (d) T2_IC_GPS, (e) T2_IC_CNTL, and (f) T2_IC_Q. Dashed lines denote FSS values of 0.5, 0.7, and 0.8.	90

4.12	Background precipitation (light solid) and perturbation precipitation spectra (dark solid) averaged over ten ensemble members of (a) T2_IC_GPS and (b) T2_MP_GPS at lead times of 3, 6, 9, and 12 h. The spectra are computed over the A1 domain. Lower panels (c and d) depict the corresponding saturation ratios at different lead times.	91
4.13	Same as Fig. 4.12 but for (a, d) T3_IC_p100, (b, e) T3_IC_p50, and (c, f) T3_IC_p10.	92
4.14	Same as Fig. 4.9 but for (a) T3_IC_p100, (b) T3_IC_p50, and (c) T3_IC_p10.	93
6.1	Domain-averaged normalized precipitation spread (S_n , unitless) over (a) DS_WATER and (b) DN_WATER subregions for each 48-hour forecast. Panels (c) and (d) are domain-averaged NRMDTE (unitless) for DS_WATER and DN_WATER, respectively.	101
6.2	Time series of domain-averaged normalized hourly precipitation spread (solid, mm) and domain-averaged hourly precipitation difference against CNTL forecasts (dashed, mm) for PSP, SPPMP, and BOTH. The rows depict results averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Each line within the figure corresponds to two winter storm cases.	102
6.3	The upper panels depict the lowest-level water vapor perturbation fields ($g\ kg^{-1}\ s^{-1}$) generated by the PSP scheme, valid at (a) 06 UTC, (b) 12 UTC, (c) 18 UTC, and (d) 00 UTC on July 21, 2021. The lower panels depict their corresponding water vapor turbulence variance fields ($kg^2\ kg^{-2}$). These forecasts are from the first member in the PSP experiment initialized at 00 UTC on the same day.	103
6.4	Time series of domain-averaged root-mean difference total energy (RMDTE, m^2s^2) growth rate for PSP (red), SPPMP (blue), and BOTH (green). Rows are d RMDTE/dt averaged over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show forecasts initialized at (a) 00 UTC and (b) 12 UTC. Each line within the figure corresponds to two winter storm cases. Dashed lines indicate the 12 LST corresponding to each initial time.	104
6.5	Time series of domain-averaged RMDTE ratios, averaged over forecasts of each weather type (colors) initialized at 00 UTC (solid lines) and 12 UTC (dashed lines). Rows show averages over DN_LAND, DN_WATER, DS_LAND, and DS_WATER, respectively. Columns show ratios computed as (a) SPPMP divided by PSP and (b) BOTH divided by PSP. Each line within the figure corresponds to two cases.	105

List of Tables

3.1	Configuration of physics parameterization schemes for all simulations. . . .	47
3.2	Details of the four perturbed parameters in the SPPMP scheme. The listed perturbation magnitudes originated from stochastic random patterns with a standard deviation of 0.75.	48
3.3	Abbreviation and model initial time for each simulation.	52
4.1	Configuration of physics parameterization schemes for all simulations. . . .	82
4.2	A summary of the three sensitivity tests with detailed experimental designs.	82

Bibliography

- Ahmadalipour, A. and H. Moradkhani, 2019: A data-driven analysis of flash flood hazard, fatalities, and damages over the CONUS during 1996–2017. *Journal of Hydrology*, **578**, 124–106.
- Ancell, B. C. and C. F. Mass, 2006: Structure, growth rates, and tangent linear accuracy of adjoint sensitivities with respect to horizontal and vertical resolution. *Monthly Weather Review*, **134** (10), 2971–2988.
- Anthes, R. A., P. Bernhardt, Y. Chen, L. Cucurull, K. Dymond, D. Ector, S. Healy, S.-P. Ho, D. Hunt, Y.-H. Kuo, and others, 2008: The COSMIC/FORMOSAT-3 Mission: Early Results. *Bulletin of the American Meteorological Society*, **89** (3), 313–334.
- Aparicio, J. M. and G. Deblonde, 2008: Impact of the assimilation of CHAMP refractivity profiles on Environment Canada global forecasts. *Monthly Weather Review*, **136** (1), 257–275.
- Atkinson, B. W. and J. Wu Zhang, 1996: Mesoscale shallow convection in the atmosphere. *Reviews of Geophysics*, **34** (4), 403–431.
- Bachmann, K., 2019: *Sources of predictability for deep convection: radar data assimilation, orography and weather regime*. Dissertation, Ludwig-Maximilians-Universität München.
- Bachmann, K., C. Keil, G. C. Craig, M. Weissmann, and C. A. Welzbacher, 2020: Predictability of deep convection in idealized and operational forecasts: Effects of radar data assimilation, orography, and synoptic weather regime. *Monthly Weather Review*, **148** (1), 63–81.
- Bao, J.-W., S. Michelson, and E. Grell, 2021: Chapter 10 - Uncertainties in the parameterization of cloud microphysics: An illustration of the problem. *Uncertainties in Numerical Weather Prediction*, Elsevier, 265–296.
- Barker, D., X.-Y. Huang, Z. Liu, T. Auligné, X. Zhang, S. Rugg, R. Ajjaji, A. Bourgeois, J. Bray, Y. Chen, and others, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bulletin of the American Meteorological Society*, **93** (6), 831–843.

- Barthlott, C., A. Zarboon, T. Matsunobu, and C. Keil, 2022a: Impacts of combined microphysical and land-surface uncertainties on convective clouds and precipitation in different weather regimes. *Atmospheric Chemistry and Physics*, **22** (16), 10 841–10 860.
- Barthlott, C., A. Zarboon, T. Matsunobu, and C. Keil, 2022b: Importance of aerosols and shape of the cloud droplet size distribution for convective clouds and precipitation. *Atmospheric Chemistry and Physics*, **22** (3), 2153–2172.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525** (7567), 47–55.
- Beck, J., F. Bouttier, L. Wiegand, C. Gebhardt, C. Eagle, and N. Roberts, 2016: Development and verification of two convection-allowing multi-model ensembles over Western Europe. *Quarterly Journal of the Royal Meteorological Society*, **142** (700), 2808–2826.
- Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004: Mesoscale weather prediction with the RUC hybrid isentropic–terrain-following coordinate model. *Monthly Weather Review*, **132** (2), 473–494.
- Berner, J., S.-Y. Ha, J. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Monthly Weather Review*, **139** (6), 1972–1995.
- Berner, J., G. Shutts, M. Leutbecher, and T. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences*, **66** (3), 603–626.
- Berner, J., U. Achatz, L. Batte, L. Bengtsson, A. De La Camara, H. M. Christensen, M. Colangeli, D. R. Coleman, D. Crommelin, S. I. Dolaptchiev, and others, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, **98** (3), 565–588.
- Boucher, O., H. Le Treut, and M. B. Baker, 1995: Precipitation and radiation modeling in a general circulation model: Introduction of cloud microphysical processes. *Journal of Geophysical Research: Atmospheres*, **100** (D8), 16 395–16 414.
- Bouttier, F., B. Vié, O. Nuissier, and L. Raynaud, 2012: Impact of stochastic physics in a convection-permitting ensemble. *Monthly Weather Review*, **140** (11), 3706–3721.
- Brázdil, R., K. Chromá, L. Dolák, J. Řehoř, L. Řezníčková, P. Zahradníček, and P. Dobrovolný, 2021: Fatalities associated with the severe weather conditions in the Czech Republic, 2000–2019. *Natural Hazards and Earth System Sciences*, **21** (5), 1355–1382.
- Brisch, J. and H. Kantz, 2019: Power law error growth in multi-hierarchical chaotic systems—a dynamical mechanism for finite prediction horizon. *New Journal of Physics*, **21** (9), 093 002.

- Brown, B., T. Jensen, J. H. Gotway, R. Bullock, E. Gilleland, T. Fowler, K. Newman, D. Adriaansen, L. Blank, T. Burek, and others, 2021: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society*, **102** (4), 782–807.
- Buizza, R., P. Houtekamer, G. Pellerin, Z. Toth, Y. Zhu, and M. Wei, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Monthly Weather Review*, **133** (5), 1076–1097.
- Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **125** (560), 2887–2908.
- Cannon, F., C. W. Hecht, J. M. Cordeira, and F. M. Ralph, 2018: Synoptic and mesoscale forcing of Southern California extreme precipitation. *Journal of Geophysical Research: Atmospheres*, **123** (24), 13–714.
- Carbone, R. and J. Tuttle, 2008: Rainfall occurrence in the US warm season: The diurnal cycle. *Journal of Climate*, **21** (16), 4132–4146.
- Chang, C.-P., 2004: *East Asian Monsoon*. World Scientific, 564 pp.
- Chang, P.-L., J. Zhang, Y.-S. Tang, L. Tang, P.-F. Lin, C. Langston, B. Kaney, C.-R. Chen, and K. Howard, 2021: An operational multi-radar multi-sensor QPE system in Taiwan. *Bulletin of the American Meteorological Society*, **102** (3), 555–577.
- Chen, I.-H., J. Berner, C. Keil, Y.-H. Kuo, and G. Craig, 2024: Classification of Warm-Season Precipitation in High-Resolution Rapid Refresh (HRRR) Model Forecasts over the Contiguous United States. *Monthly Weather Review*, **152** (1), 187–201.
- Chen, I.-H., Y.-J. Su, H.-W. Lai, J.-S. Hong, C.-H. Li, P.-L. Chang, and Y.-J. Wu, 2023: Performance of a Convective-Scale Ensemble Prediction System on 2017 Warm-Season Afternoon Thunderstorms over Taiwan. *Weather and Forecasting*, **38** (4), 517–538.
- Chien, F.-C. and Y.-C. Chiu, 2023: Factors Leading to Heavy Rainfall in Southern Taiwan in the Early Mei-Yu Season of 2020. *Monthly Weather Review*, **151** (7), 1885–1908.
- Choudhury, D. and S. Das, 2017: The sensitivity to the microphysical schemes on the skill of forecasting the track and intensity of tropical cyclones using WRF-ARW model. *Journal of Earth System Science*, **126**, 1–10.
- Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteorological Applications*, **23** (2), 165–181.
- Cortes, C. and V. Vapnik, 1995: Support-vector networks. *Machine Learning*, **20**, 273–297.

- Craig, G. C., C. Keil, and D. Leuenberger, 2012: Constraints on the impact of radar rainfall data assimilation on forecasts of cumulus convection. *Quarterly Journal of the Royal Meteorological Society*, **138** (663), 340–352.
- Cucurull, L., J. Derber, R. Treadon, and R. Purser, 2007: Assimilation of global positioning system radio occultation observations into NCEP’s global data assimilation system. *Monthly Weather Review*, **135** (9), 3174–3193.
- Davies, T., 2014: Lateral boundary conditions for limited area models. *Quarterly Journal of the Royal Meteorological Society*, **140** (678), 185–196.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, d. P. Bauer, and others, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137** (656), 553–597.
- Denis, B., J. Côté, and R. Laprise, 2002: Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT). *Monthly Weather Review*, **130** (7), 1812–1829.
- Descamps, L. and O. Talagrand, 2007: On some aspects of the definition of initial conditions for ensemble prediction. *Monthly Weather Review*, **135** (9), 3260–3272.
- Done, J., G. Craig, S. Gray, P. A. Clark, and M. Gray, 2006: Mesoscale simulations of organized convection: Importance of convective equilibrium. *Quarterly Journal of the Royal Meteorological Society*, **132** (616), 737–756.
- Done, J., G. C. Craig, S. Gray, and P. A. Clark, 2012: Case-to-case variability of predictability of deep convection in a mesoscale model. *Quarterly Journal of the Royal Meteorological Society*, **138** (664), 638–648.
- Doswell, C. A., 2001: Severe convective storms—an overview. *Severe convective storms*, 1–26.
- Doswell, C. A., 2014: *Severe Storms*, 767–780. Springer.
- Doswell, C. A. and L. F. Bosart, 2001: Extratropical synoptic-scale processes and severe convection. *Severe convective storms*, 27–69.
- Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Weather and Forecasting*, **5** (4), 576–585.
- Doswell, C. A. and E. N. Rasmussen, 1994: The effect of neglecting the virtual temperature correction on CAPE calculations. *Weather and Forecasting*, **9** (4), 625–629.

- Dowell, D. C., C. R. Alexander, E. P. James, S. S. Weygandt, S. G. Benjamin, G. S. Manikin, B. T. Blake, J. M. Brown, J. B. Olson, M. Hu, and others, 2022: The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part I: Motivation and system description. *Weather and Forecasting*, **37** (8), 1371–1395.
- Durrán, D. R. and M. Gingrich, 2014: Atmospheric predictability: Why butterflies are not of practical importance. *Journal of the Atmospheric Sciences*, **71** (7), 2476–2488.
- Durrán, D. R. and J. A. Weyn, 2016: Thunderstorms do not get butterflies. *Bulletin of the American Meteorological Society*, **97** (2), 237–243.
- Flack, D. L., S. L. Gray, R. S. Plant, H. W. Lean, and G. C. Craig, 2018: Convective-scale perturbation growth across the spectrum of convective regimes. *Monthly Weather Review*, **146** (1), 387–405.
- Flack, D. L., R. S. Plant, S. L. Gray, H. W. Lean, C. Keil, and G. C. Craig, 2016: Characterisation of convective regimes over the British Isles. *Quarterly Journal of the Royal Meteorological Society*, **142** (696), 1541–1553.
- Flora, M. L., C. K. Potvin, and L. J. Wicker, 2018: Practical predictability of supercells: Exploring ensemble forecast sensitivity to initial condition spread. *Monthly Weather Review*, **146** (8), 2361–2379.
- Fraley, C., A. E. Raftery, and T. Gneiting, 2010: Calibrating multimodel forecast ensembles with exchangeable and missing members using Bayesian model averaging. *Monthly Weather Review*, **138** (1), 190–202.
- Frogner, I.-L., U. Andrae, P. Ollinaho, A. Hally, K. Hämäläinen, J. Kauhanen, K.-I. Ivarsson, and D. Yazgi, 2022: Model uncertainty representation in a convection-permitting ensemble—SPP and SPPT in HarmonEPS. *Monthly Weather Review*, **150** (4), 775–795.
- Frogner, I.-L., U. Andrae, J. Bojarova, A. Callado, P. Escribà, H. Feddersen, A. Hally, J. Kauhanen, R. Randriamampianina, A. Singleton, and others, 2019: HarmonEPS—the HARMONIE ensemble prediction system. *Weather and Forecasting*, **34** (6), 1909–1937.
- Gallo, B. T., A. J. Clark, I. Jirak, J. S. Kain, S. J. Weiss, M. Coniglio, K. Knopfmeier, J. Correia Jr, C. J. Melick, C. D. Karstens, and others, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/hazardous weather testbed spring forecasting experiment. *Weather and Forecasting*, **32** (4), 1541–1568.
- García-Ortega, E., J. Lorenzana, A. Merino, S. Fernández-González, L. López, and J. Sánchez, 2017: Performance of multi-physics ensembles in convective precipitation events over northeastern Spain. *Atmospheric Research*, **190**, 55–67.

- Gebhardt, C., S. Theis, M. Paulat, and Z. B. Bouallègue, 2011: Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmospheric Research*, **100 (2-3)**, 168–177.
- Grazzini, F., G. C. Craig, C. Keil, G. Antolini, and V. Pavan, 2020: Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques. *Quarterly Journal of the Royal Meteorological Society*, **146 (726)**, 69–85.
- Groenemeijer, P., T. Pucik, I. Tsonevsky, and P. Bechtold, 2019: An overview of convective available potential energy and convective inhibition provided by NWP models for operational forecasting. Tech. rep., European Centre for Medium-Range Weather Forecasts.
- Gustafsson, N., E. Källén, and S. Thorsteinsson, 1998: Sensitivity of forecast errors to initial and lateral boundary conditions. *Tellus A*, **50 (2)**, 167–185.
- Gustafsson, N., T. Janjić, C. Schraff, D. Leuenberger, M. Weissmann, H. Reich, P. Brousseau, T. Montmerle, E. Wattrelot, A. Bučánek, and others, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 1218–1256.
- Hanley, K. E., R. S. Plant, T. H. Stein, R. J. Hogan, J. C. Nicol, H. W. Lean, C. Halliwell, and P. A. Clark, 2015: Mixing-length controls on high-resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, **141 (686)**, 272–284.
- Healy, S. and J.-N. Thépaut, 2006: Assimilation experiments with CHAMP GPS radio occultation measurements. *Quarterly Journal of the Royal Meteorological Society*, **132 (615)**, 605–623.
- Hirt, M., 2020: *Convective initiation-relevant processes and their representation in convection-permitting models*. Dissertation, Ludwig-Maximilians-Universität München.
- Hirt, M. and G. C. Craig, 2021: A cold pool perturbation scheme to improve convective initiation in convection-permitting models. *Quarterly Journal of the Royal Meteorological Society*, **147 (737)**, 2429–2447.
- Hirt, M., G. C. Craig, S. A. Schäfer, J. Savre, and R. Heinze, 2020: Cold-pool-driven convective initiation: Using causal graph analysis to determine what convection-permitting models are missing. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 2205–2227.
- Hirt, M., S. Rasp, U. Blahak, and G. C. Craig, 2019: Stochastic parameterization of processes leading to convective initiation in kilometer-scale models. *Monthly Weather Review*, **147 (11)**, 3917–3934.

- Hohenegger, C., D. Lüthi, and C. Schär, 2006: Predictability mysteries in cloud-resolving models. *Monthly Weather Review*, **134** (8), 2095–2107.
- Hohenegger, C. and C. Schar, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bulletin of the American Meteorological Society*, **88** (11), 1783–1794.
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, **134** (9), 2318–2341.
- Honnert, R., G. A. Efstathiou, R. J. Beare, J. Ito, A. Lock, R. Neggers, R. S. Plant, H. H. Shin, L. Tomassini, and B. Zhou, 2020: The atmospheric boundary layer and the “gray zone” of turbulence: A critical review. *Journal of Geophysical Research: Atmospheres*, **125** (13), e2019JD030317.
- Houze Jr, R. A., 2004: Mesoscale convective systems. *Reviews of Geophysics*, **42** (4).
- Hsiao, L.-F., D.-S. Chen, J.-S. Hong, T.-C. Yeh, and C.-T. Fong, 2020: Improvement of the numerical tropical cyclone prediction system at the Central Weather Bureau of Taiwan: TWRF (Typhoon WRF). *Atmosphere*, **11** (6), 657.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, **113** (D13).
- Jankov, I., J. Beck, J. Wolff, M. Harrold, J. B. Olson, T. Smirnova, C. Alexander, and J. Berner, 2019: Stochastically perturbed parameterizations in an HRRR-based ensemble. *Monthly Weather Review*, **147** (1), 153–173.
- Jankov, I., J. Berner, J. Beck, H. Jiang, J. B. Olson, G. Grell, T. G. Smirnova, S. G. Benjamin, and J. M. Brown, 2017: A performance comparison between multiphysics and stochastic approaches within a North American RAP ensemble. *Monthly Weather Review*, **145** (4), 1161–1179.
- Jin, H.-G. and J.-J. Baik, 2023: Do Double-Moment Microphysics Schemes Make Reliable Predictions on the Raindrop Number Concentration?: A Squall-Line Case Study. *Journal of Geophysical Research: Atmospheres*, **128** (9), e2022JD038394.
- Jucker, M., T. Lane, C. Vincent, S. Webster, S. Wales, and V. Louf, 2020: Locally forced convection in subkilometre-scale simulations with the Unified Model and WRF. *Quarterly Journal of the Royal Meteorological Society*, **146** (732), 3450–3465.
- Kain, J. S. and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. *The representation of cumulus convection in numerical models*, 165–170.

- Kain, J. S., S. J. Weiss, D. R. Bright, M. E. Baldwin, J. J. Levit, G. W. Carbin, C. S. Schwartz, M. L. Weisman, K. K. Droegemeier, D. B. Weber, and others, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Weather and Forecasting*, **23** (5), 931–952.
- Kasahara, A., 1974: Various vertical coordinate systems used for numerical weather prediction. *Monthly Weather Review*, **102** (7), 509–522.
- Keil, C., F. Baur, K. Bachmann, S. Rasp, L. Schneider, and C. Barthlott, 2019: Relative contribution of soil moisture, boundary-layer and microphysical perturbations on convective predictability in different weather regimes. *Quarterly Journal of the Royal Meteorological Society*, **145** (724), 3102–3115.
- Keil, C., L. Chabert, O. Nuissier, and L. Raynaud, 2020: Dependence of predictability of precipitation in the northwestern Mediterranean coastal region on the strength of synoptic control. *Atmospheric Chemistry and Physics*, **20** (24), 15 851–15 865.
- Keil, C. and G. C. Craig, 2011: Regime-dependent forecast uncertainty of convective precipitation. *Meteorologische Zeitschrift*, **20** (2), 145.
- Keil, C., F. Heinlein, and G. C. Craig, 2014: The convective adjustment time-scale as indicator of predictability of convective precipitation. *Quarterly Journal of the Royal Meteorological Society*, **140** (679), 480–490.
- Khain, A., K. Beheng, A. Heymsfield, A. Korolev, S. Krichak, Z. Levin, M. Pinsky, V. Phillips, T. Prabhakaran, A. Teller, and others, 2015: Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization. *Reviews of Geophysics*, **53** (2), 247–322.
- Kirshbaum, D. J., B. Adler, N. Kalthoff, C. Barthlott, and S. Serafin, 2018: Moist orographic convection: Physical mechanisms and links to surface-exchange processes. *Atmosphere*, **9** (3), 80.
- Kober, K. and G. C. Craig, 2016: Physically based stochastic perturbations (PSP) in the boundary layer to represent uncertainty in convective initiation. *Journal of the Atmospheric Sciences*, **73** (7), 2893–2911.
- Köcher, G., T. Zinner, and C. Knote, 2023: Influence of cloud microphysics schemes on weather model predictions of heavy precipitation. *Atmospheric Chemistry and Physics*, **23** (11), 6255–6269.
- Koshyk, J. N., K. Hamilton, and J. Mahlman, 1999: Simulation of the κ -5/3 mesoscale spectral regime in the GFDL SKYHI general circulation model. *Geophysical Research Letters*, **26** (7), 843–846.

- Kühnlein, C., C. Keil, G. Craig, and C. Gebhardt, 2014: The impact of downscaled initial condition perturbations on convective-scale ensemble forecasts of precipitation. *Quarterly Journal of the Royal Meteorological Society*, **140** (682), 1552–1562.
- Kuo, Y.-H., H. Liu, Y.-R. Guo, C.-T. Terng, and Y.-T. Lin, 2008: Impact of FORMOSAT-3/COSMIC data on typhoon and Mei-yu prediction. *Recent Progress in Atmospheric Sciences: Applications to the Asia-Pacific Region*, World Scientific, 458–483.
- Kuo, Y.-H., S. V. Sokolovskiy, R. A. Anthes, and F. Vandenberghe, 2000: Assimilation of GPS radio occultation data for numerical weather prediction. *Terrestrial Atmospheric and Oceanic Sciences*, **11** (1), 157–186.
- Kuo, Y.-H., J. Sun, Y. Zhang, and I.-H. Chen, 2023: Impact of GNSS radio occultation data on the prediction of a pre-frontal squall line associated with a Mei-Yu front: I. Sensitivity Experiment. *In American Geophysical Union (AGU) Annual Meeting 2023. San Francisco, CA, US*.
- Kursinski, E., G. Hajj, J. Schofield, R. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, **102** (D19), 23 429–23 465.
- Lang, S. T., S.-J. Lock, M. Leutbecher, P. Bechtold, and R. M. Forbes, 2021: Revision of the stochastically perturbed parametrisations model uncertainty scheme in the integrated forecasting system. *Quarterly Journal of the Royal Meteorological Society*, **147** (735), 1364–1381.
- Leutbecher, M., S.-J. Lock, P. Ollinaho, S. T. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, and others, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, **143** (707), 2315–2339.
- Liu, C., K. Ikeda, R. Rasmussen, M. Barlage, A. J. Newman, A. F. Prein, F. Chen, L. Chen, M. Clark, A. Dai, and others, 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, **49**, 71–95.
- Lorenç, A. C., 1986: Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112** (474), 1177–1194.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, **20** (2), 130–141.
- Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21** (3), 289–307.
- Lorenz, E. N., 1996: Predictability: A problem partly solved. *Seminar on predictability*, Reading, UK.

- Luo, Y. and Y. Chen, 2015: Investigation of the predictability and physical mechanisms of an extreme-rainfall-producing mesoscale convective system along the Meiyu front in East China: An ensemble approach. *Journal of Geophysical Research: Atmospheres*, **120** (20), 10–593.
- Lupo, K. M., R. D. Torn, and S.-C. Yang, 2020: Evaluation of stochastic perturbed parameterization tendencies on convective-permitting ensemble forecasts of heavy rainfall events in New York and Taiwan. *Weather and Forecasting*, **35** (1), 5–24.
- Lupo, K. M., R. D. Torn, and S.-C. Yang, 2022: Process-Based Evaluation of Stochastic Perturbed Microphysics Parameterization Tendencies on Ensemble Forecasts of Heavy Rainfall Events. *Monthly Weather Review*, **150** (1), 175–191.
- Lynch, P., 2008: The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, **227** (7), 3431–3444.
- Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Monthly Weather Review*, **136** (11), 4092–4104.
- Markowski, P. and Y. Richardson, 2011: *Mesoscale meteorology in midlatitudes*. John Wiley & Sons.
- Markowski, P. M., 2020: What is the intrinsic predictability of tornadic supercell thunderstorms? *Monthly Weather Review*, **148** (8), 3157–3180.
- Matsunobu, T., C. Keil, and C. Barthlott, 2022: The impact of microphysical uncertainty conditional on initial and boundary condition uncertainty under varying synoptic control. *Weather and Climate Dynamics*, **3** (4), 1273–1289.
- McTaggart-Cowan, R., L. Separovic, R. Aider, M. Charron, M. Desgagné, P. L. Houtekamer, D. Paquin-Ricard, P. A. Vaillancourt, and A. Zadra, 2022: Using stochastically perturbed parameterizations to represent model uncertainty. Part I: Implementation and parameter sensitivity. *Monthly Weather Review*, **150** (11), 2829–2858.
- Melhauser, C. and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *Journal of the Atmospheric Sciences*, **69** (11), 3350–3371.
- Molini, L., A. Parodi, N. Rebora, and G. Craig, 2011: Classifying severe rainfall events over Italy by hydrometeorological and dynamical criteria. *Quarterly Journal of the Royal Meteorological Society*, **137** (654), 148–154.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, **122** (529), 73–119.

- Nakanishi, M. and H. Niino, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, **119**, 397–407.
- Nelson, B. R., O. P. Prat, D.-J. Seo, and E. Habib, 2016: Assessment and implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Weather and Forecasting*, **31** (2), 371–394.
- Nielsen, E. R. and R. S. Schumacher, 2016: Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. *Monthly Weather Review*, **144** (10), 3651–3676.
- Nutter, P., M. Xue, and D. Stensrud, 2004: Application of lateral boundary condition perturbations to help restore dispersion in limited-area ensemble forecasts. *Monthly Weather Review*, **132** (10), 2378–2390.
- Ollinaho, P., S.-J. Lock, M. Leutbecher, P. Bechtold, A. Beljaars, A. Bozzo, R. M. Forbes, T. Haiden, R. J. Hogan, and I. Sandu, 2017: Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble. *Quarterly Journal of the Royal Meteorological Society*, **143** (702), 408–422.
- Olson, J. B., J. S. Kenyon, W. Angevine, J. M. Brown, M. Pagowski, K. Sušelj, and others, 2019: A description of the MYNN-EDMF scheme and the coupling to other components in WRF–ARW. Tech. rep., National Oceanic and Atmospheric Administration.
- Palmer, T. N., 1993: Extended-range atmospheric prediction and the Lorenz model. *Bulletin of the American Meteorological Society*, **74** (1), 49–66.
- Palmer, T. N., 2000: Predicting uncertainty in forecasts of weather and climate. *Reports on Progress in Physics*, **63** (2), 71.
- Paulson, C. A., 1970: The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of Applied Meteorology and Climatology*, **9** (6), 857–861.
- Pearson, K., G. Lister, C. Birch, R. Allan, R. Hogan, and S. Woolnough, 2014: Modelling the diurnal cycle of tropical convection across the ‘grey zone’. *Quarterly Journal of the Royal Meteorological Society*, **140** (679), 491–499.
- Pincus, R. and B. Stevens, 2013: Paths to accuracy for radiation parameterizations in atmospheric models. *Journal of Advances in Modeling Earth Systems*, **5** (2), 225–233.
- Puh, M., C. Keil, C. Gebhardt, C. Marsigli, M. Hirt, F. Jakub, and G. C. Craig, 2023: Physically based stochastic perturbations improve a high-resolution forecast of convection. *Quarterly Journal of the Royal Meteorological Society*, **149** (757), 3582–3592.

- Qiao, S., D. Chen, B. Wang, H.-N. Cheung, F. Liu, J. Cheng, S. Tang, Z. Zhang, G. Feng, and W. Dong, 2021: The longest 2020 meiyu season over the past 60 years: Subseasonal perspective and its predictions. *Geophysical Research Letters*, **48** (9), e2021GL093596.
- Reinert, D., F. Prill, H. Frank, M. Denhard, M. Baldauf, C. Schraff, C. Gebhardt, C. Margigli, and G. Zängl, 2020: DWD database reference for the global and regional ICON and ICON-EPS forecasting system. Tech. rep., Deutscher Wetterdienst.
- Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 236 pp.
- Rickenbach, T. M., R. Nieto-Ferreira, C. Zarzar, and B. Nelson, 2015: A seasonal and diurnal climatology of precipitation organization in the southeastern United States. *Quarterly Journal of the Royal Meteorological Society*, **141** (690), 1938–1956.
- Roberts, B., B. T. Gallo, I. L. Jirak, A. J. Clark, D. C. Dowell, X. Wang, and Y. Wang, 2020: What does a convection-allowing ensemble of opportunity buy us in forecasting thunderstorms? *Weather and Forecasting*, **35** (6), 2293–2316.
- Roberts, B., I. L. Jirak, A. J. Clark, S. J. Weiss, and J. S. Kain, 2019: Postprocessing and visualization techniques for convection-allowing ensembles. *Bulletin of the American Meteorological Society*, **100** (7), 1245–1258.
- Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Monthly Weather Review*, **136** (1), 78–97.
- Rodwell, M. J., L. Magnusson, P. Bauer, P. Bechtold, M. Bonavita, C. Cardinali, M. Diamantakis, P. Earnshaw, A. Garcia-Mendez, L. Isaksen, and others, 2013: Characteristics of occasional poor medium-range weather forecasts for Europe. *Bulletin of the American Meteorological Society*, **94** (9), 1393–1405.
- Romine, G. S., C. S. Schwartz, J. Berner, K. R. Fossell, C. Snyder, J. L. Anderson, and M. L. Weisman, 2014: Representing forecast error in a convection-permitting ensemble system. *Monthly Weather Review*, **142** (12), 4519–4541.
- Schneider, L., C. Barthlott, A. I. Barrett, and C. Hoose, 2018: The precipitation response to variable terrain forcing over low mountain ranges in different weather regimes. *Quarterly Journal of the Royal Meteorological Society*, **144** (713), 970–989.
- Schwartz, C. S., J. Poterjoy, G. S. Romine, D. C. Dowell, J. R. Carley, and J. Bresch, 2022: Short-term convection-allowing ensemble precipitation forecast sensitivity to resolution of initial condition perturbations and central initial states. *Weather and Forecasting*, **37** (7), 1259–1286.

- Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2015: NCAR's experimental real-time convection-allowing ensemble prediction system. *Weather and Forecasting*, **30** (6), 1645–1654.
- Schwartz, C. S., J. S. Kain, S. J. Weiss, M. Xue, D. R. Bright, F. Kong, K. W. Thomas, J. J. Levit, and M. C. Coniglio, 2009: Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Monthly Weather Review*, **137** (10), 3351–3372.
- Selz, T., 2019: Estimating the intrinsic limit of predictability using a stochastic convection scheme. *Journal of the Atmospheric Sciences*, **76** (3), 757–765.
- Selz, T. and G. C. Craig, 2015: Upscale Error Growth in a High-Resolution Simulation of a Summertime Weather Event over Europe. *Monthly Weather Review*, **143** (3), 813–827.
- Selz, T., M. Riemer, and G. C. Craig, 2022: The transition from practical to intrinsic predictability of midlatitude weather. *Journal of the Atmospheric Sciences*, **79** (8), 2013–2030.
- Shin, H. H. and S.-Y. Hong, 2013: Analysis of resolved and parameterized vertical transports in convective boundary layers at gray-zone resolutions. *Journal of the Atmospheric Sciences*, **70** (10), 3248–3261.
- Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, **131** (612), 3079–3102.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, J. G. Powers, and others, 2008: A description of the advanced research WRF version 3. *NCAR technical note*, **475**, 113.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, Z. Liu, J. Berner, W. Wang, J. G. Powers, M. G. Duda, D. M. Barker, and others, 2019: A description of the advanced research wrf version 4. *NCAR technical note*, **145**.
- Skok, G. and N. Roberts, 2016: Analysis of fractions skill score properties for random precipitation fields and ECMWF forecasts. *Quarterly Journal of the Royal Meteorological Society*, **142** (700), 2599–2610.
- Slingo, J. and T. Palmer, 2011: Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A*, **369**, 4751–4767.
- Sobash, R. A., D. J. Gagne, C. L. Becker, D. Ahijevych, G. N. Gantos, and C. S. Schwartz, 2023: Diagnosing storm mode with deep learning in convection-allowing models. *Monthly Weather Review*, **151** (8), 2009–2027.

- Stanford, M. W., H. Morrison, A. Varble, J. Berner, W. Wu, G. McFarquhar, and J. Milbrandt, 2019: Sensitivity of simulated deep convection to a stochastic ice microphysics framework. *Journal of Advances in Modeling Earth Systems*, **11** (11), 3362–3389.
- Stensrud, D. J., 2007: *Parameterization schemes: keys to understanding numerical weather prediction models*. Cambridge University Press.
- Stevens, B., 2005: Atmospheric moist convection. *Annu. Rev. Earth Planet. Sci.*, **33**, 605–643.
- Stull, R. B., 1988: *An introduction to boundary layer meteorology*, Vol. 13. Springer Dordrecht.
- Sun, J., 2005: Convective-scale assimilation of radar data: progress and challenges. *Quarterly Journal of the Royal Meteorological Society*, **131** (613), 3439–3463.
- Sun, Y. Q. and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *Journal of the Atmospheric Sciences*, **73** (3), 1419–1438.
- Surcel, M., I. Zawadzki, M. Yau, M. Xue, and F. Kong, 2017: More on the scale dependence of the predictability of precipitation patterns: Extension to the 2009–13 CAPS Spring Experiment ensemble forecasts. *Monthly Weather Review*, **145** (9), 3625–3646.
- Takahashi, H. G. and H. Fujinami, 2021: Recent decadal enhancement of Meiyu–Baiu heavy rainfall over East Asia. *Scientific Reports*, **11** (1), 13665.
- Tao, W.-K. and X. Li, 2016: The relationship between latent heating, vertical velocity, and precipitation processes: The impact of aerosols on precipitation in organized deep convective systems. *Journal of Geophysical Research: Atmospheres*, **121** (11), 6299–6320.
- Tao, W.-K., D. Wu, S. Lang, J.-D. Chern, C. Peters-Lidard, A. Fridlind, and T. Matsui, 2016: High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: Further improvements and comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, **121** (3), 1278–1305.
- Tewari, M., F. Chen, W. Wang, J. Dudhia, M. LeMone, K. Mitchell, M. Ek, G. Gayno, J. Wegiel, and others, 2004: Implementation and verification of the unified Noah land-surface model in the WRF model. In *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*. American Meteorological Society: Seattle, WA, US.
- Thompson, G., J. Berner, M. Frediani, J. A. Otkin, and S. M. Griffin, 2021: A stochastic parameter perturbation method to represent uncertainty in a microphysics scheme. *Monthly Weather Review*, **149** (5), 1481–1497.

- Thompson, G. and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the atmospheric sciences*, **71** (10), 3636–3658.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, **136** (12), 5095–5115.
- Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, **125** (12), 3297–3319.
- Troen, I. and L. Mahrt, 1986: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Boundary-Layer Meteorology*, **37** (1-2), 129–148.
- Turner, D. D., H. Cutler, M. Shields, R. Hill, B. Hartman, Y. Hu, T. Lu, and H. Jeon, 2021: Evaluating the economic impacts of improvements to the high-resolution rapid refresh (HRRR) numerical weather prediction model. *Bulletin of the American Meteorological Society*, 1–36.
- Unger, D. A., H. van den Dool, E. O’Lenic, and D. Collins, 2009: Ensemble regression. *Monthly Weather Review*, **137** (7), 2365–2379.
- Wang, C.-C., P.-Y. Chuang, C.-S. Chang, K. Tsuboki, S.-Y. Huang, and G.-C. Leu, 2022: Evaluation of Mei-yu heavy-rainfall quantitative precipitation forecasts in Taiwan by a cloud-resolving model for three seasons of 2012–2014. *Natural Hazards and Earth System Sciences*, **22** (1), 23–40.
- Wang, S.-Y. and T.-C. Chen, 2009: The late-spring maximum of rainfall over the US central plains and the role of the low-level jet. *Journal of Climate*, **22** (17), 4696–4709.
- Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bulletin of the American Meteorological Society*, **78** (11), 2599–2618.
- Wastl, C., Y. Wang, A. Atencia, and C. Wittmann, 2019: Independent perturbations for physics parametrization tendencies in a convection-permitting ensemble (pSPPT). *Geoscientific Model Development*, **12** (1), 261–273.
- Weigel, A. P., M. Liniger, and C. Appenzeller, 2008: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quarterly Journal of the Royal Meteorological Society*, **134** (630), 241–260.
- Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Monthly Weather Review*, **138** (9), 3418–3433.

- Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP global forecast system. *Monthly Weather Review*, **136** (2), 463–482.
- Wicker, L. J. and W. C. Skamarock, 2002: Time-splitting methods for elastic models using forward time schemes. *Monthly Weather Review*, **130** (8), 2088–2097.
- Winkler, J., M. Denhard, H. Frank, A. Rhodin, H. Anlauf, A. Fernandez del Rio, A. Cress, and R. Potthast, 2018: ICON-EPS: The operational global ensemble forecasting system of DWD. *EGU General Assembly Conference Abstracts*, 13813.
- Yano, J.-I. and R. Plant, 2012: Interactions between shallow and deep convection under a finite departure from convective quasi equilibrium. *Journal of the atmospheric sciences*, **69** (12), 3463–3470.
- Yu, X. and T.-Y. Lee, 2010: Role of convective parameterization in simulations of a convection band at grey-zone resolutions. *Tellus A*, **62** (5), 617–632.
- Zeng, Y., T. Janjić, A. de Lozar, U. Blahak, H. Reich, C. Keil, and A. Seifert, 2018: Representation of model error in convective-scale data assimilation: additive noise, relaxation methods, and combinations. *Journal of Advances in Modeling Earth Systems*, **10** (11), 2889–2911.
- Zeng, Y., T. Janjić, M. Sommer, A. de Lozar, U. Blahak, and A. Seifert, 2019: Representation of model error in convective-scale data assimilation: Additive noise based on model truncation error. *Journal of Advances in Modeling Earth Systems*, **11** (3), 752–770.
- Zhang, F., N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *Journal of the Atmospheric Sciences*, **64** (10), 3579–3594.
- Zhang, F., A. M. Odins, and J. W. Nielsen-Gammon, 2006: Mesoscale predictability of an extreme warm-season precipitation event. *Weather and Forecasting*, **21** (2), 149–166.
- Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *Journal of the Atmospheric Sciences*, **60** (9), 1173–1185.
- Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *Journal of the Atmospheric Sciences*, **76** (4), 1077–1091.
- Zhang, L., J. Min, X. Zhuang, S. Wang, and X. Qiao, 2023: The Lateral Boundary Perturbations Growth and Their Dependence on the Forcing Types of Severe Convection in Convection-Allowing Ensemble Forecasts. *Atmosphere*, **14** (1), 176.

- Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2015: Practical predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma: Sensitivity to synoptic timing and topographical influence. *Monthly Weather Review*, **143** (8), 2973–2997.
- Zhao, X. and R. D. Torn, 2022: Evaluation of Independent Stochastically Perturbed Parameterization Tendency (iSPPT) Scheme on HWRP-Based Ensemble Tropical Cyclone Intensity Forecasts. *Monthly Weather Review*, **150** (10), 2659–2674.
- Zhou, X., Y. Zhu, D. Hou, B. Fu, W. Li, H. Guan, E. Sinsky, W. Kolczynski, X. Xue, Y. Luo, and others, 2022: The development of the NCEP global ensemble forecast system version 12. *Weather and Forecasting*, **37** (6), 1069–1084.
- Ziehmann, C., 2000: Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models. *Tellus A*, **52** (3), 280–299.
- Zimmer, M., G. Craig, C. Keil, and H. Wernli, 2011: Classification of precipitation events with a convective response timescale and their forecasting characteristics. *Geophysical Research Letters*, **38** (5).