

**A COLLECTION OF PARAMETERIZATION STUDIES HIGHLIGHTING THE  
IMPORTANCE OF EXTENDED FORECAST LENGTHS**

By

Nick P. Bassill

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of final oral examination: 06/03/2013

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

Michael Morgan, Professor, Atmospheric and Oceanic Sciences  
Larissa Back, Assistant Professor, Atmospheric and Oceanic Sciences  
James Kossin, Atmospheric Scientist, National Oceanic and Atmospheric  
Administration  
Jonathan Martin, Professor, Atmospheric and Oceanic Sciences  
Gregory Tripoli, Professor, Atmospheric and Oceanic Sciences

© Copyright by Nick P. Bassill 2013  
All Rights Reserved

## Abstract

For the foreseeable future, numerical weather prediction models will require *parameterizations* in order to represent sub-grid scale processes. There are many classes of parameterization (cumulus, microphysics, etc.), and for each class of parameterization a number of options are available which are generally differentiated by their complexity or formulation. Although in a broad sense, different options within a class of parameterizations perform similarly, their differences do have significant forecast implications. A considerable amount of research has been performed analyzing the impacts of varying parameterizations. However, the vast majority of this research has been of the short-term forecast (48 hours or less) and case-study variety. This dissertation discusses the importance of varying parameterizations on longer forecasting timescales (up to 5-50 days) over a wide variety of geographical regions and conditions. Additionally, care is taken to create sufficiently large amounts of data such that statistically significant conclusions can be reached. Since it is not feasible to examine all available parameterizations within this dissertation, the Grell-Dévényi cumulus parameterization will be a particular focus. This is in part due to its novel formulation as well as its increasing popularity for use in both short-term forecasting as well as climate scale forecasting. It is shown that the current construction of this parameterization causes deficiencies in the depiction of tropical convection, which has important implications for phenomena ranging from the Hadley Cell to tropical cyclones. Various statistical methods of optimizing this parameterization as well as other parameterizations are also discussed.

## Acknowledgements

First and foremost I would like to thank my advisor, Dr. Michael Morgan. Michael has always been extremely supportive both personally and professionally. He's made travel to conferences a priority and has shown confidence and a willingness to allow me the freedom to study what interests me most. Aside from being my advisor, I also consider him a good friend. I'd also like to thank the rest of my committee – Professor Jonathan Martin, Professor Larissa Back, Professor Gregory Tripoli, and Dr. Jim Kossin. Professor Martin in particular has been extremely welcoming of me within his group as well. I certainly could not have gotten this far without the work of many others in the Atmospheric and Oceanic Sciences department. Most notably, I would like to thank Pete Pokrandt for helping fix more computer problems than I could possibly remember. In a similar vein, I'd like to thank my past lab mates for assisting with problems, brainstorming, and their general camaraderie. In particular, Ross Lazear, Brett Hoover, and Dianna Nelson (among many others) have been great to work with.

I certainly could never have completed this dissertation (or achieved anything else of significance) without the great upbringing and care provided to me by my family. My mom (Deb Schenk), dad (Pat Bassill), step-dad (Steve Schenk), and step-mom (Donna Bassill) have all been instrumental in any success I've achieved. My grandparents, aunts, uncles, extended family, and sister have all assisted in ways great and small as well. Finally, I want to thank my wonderful fiancé, who has both kept me sane as well as been understanding of the late hours I've needed to keep in order to finish this dissertation. I'd also like to acknowledge the great city of Madison, which I've greatly enjoyed living in.

## Table Of Contents

<b>Abstract</b>	i
<b>Acknowledgements</b>	ii
<b>Chapter 1: Introduction</b>	
<i>Historical Overview</i>	1
CUMULUS PARAMETERIZATION	2
MICROPHYSICS PARAMETERIZATIONS	9
ENSEMBLE FORECASTING	11
<i>Motivation</i>	15
<i>Dissertation Synopsis</i>	22
<b>Chapter 2: The Grell-Dévényi Cumulus Parameterization</b>	
<i>Introduction</i>	24
<i>Model Description And Methodology</i>	26
RESULTS	29
<i>Description Of Precipitation Patterns</i>	29
<i>Description Of Related Atmospheric Variables</i>	34
<i>Potential Causes Of Observed Differences</i>	39
DISCUSSION	47
<i>Examination Of Parameterization Characteristics</i>	47
<i>Grell-3 CP Sensitivity Tests</i>	56
CONCLUSION	62
<b>Chapter 3: Parameterizations And Tropical Cyclones</b>	
<i>Introduction</i>	66

<i>Poleward Track Differences</i>	68
<i>Eta Ferrier Microphysics Properties</i>	76
<i>Additional Relevant Research: Tropical Cyclogenesis</i>	88
CONCLUSION	103
<b>Chapter 4: Statistical Optimization</b>	
<i>Introduction</i>	105
STATISTICAL PRECIPITATION CORRECTION	106
<i>A Statistical Approach For Correcting The Grell-Dévényi Cumulus Parameterization</i>	106
<i>Simplest Approach: Precipitation Reassignment Based On Climatology</i>	108
<i>Grell Cumulus Parameterization: Linear Regression Approach</i>	119
<i>Analysis Of Weights</i>	127
<i>Strengths And Weaknesses Of Both Approaches</i>	129
<i>Examination Of Forecast Differences</i>	131
A COMPARISON OF TWO OPTIMIZED ENSEMBLES	132
<i>Methodology</i>	132
CONCLUSION	145
<b>Chapter 5: Conclusion</b>	
<i>Chapter Overviews</i>	146
<i>Overarching Conclusions</i>	149
<i>Future Directions</i>	150
<b>Bibliography</b>	158

## **Chapter 1: Introduction**

### *Historical Overview*

One of the great scientific achievements of the 20th century was the development of numerical weather prediction (NWP) models. Given the fundamental governing equations and well-defined initial and boundary conditions, credible estimates of future atmospheric states are achieved through the numerical integration of these complex models forward-in-time. The models represent monuments to our current understanding (and lack of understanding) of the fundamental physical processes which govern the atmospheric circulation, and also represent useful tools to enhance the protection of lives, property, and commerce through skillful prediction of relevant weather factors.

Despite the great achievements in the development of NWP models, it is recognized that output from these models is deficient due to errors in the model input (initial condition errors), errors in the boundary conditions (lower, upper, and lateral for limited area models), errors in the model formulation (numerical), and errors in the model governing equations attributable to lack of understanding of fundamental physical processes, or how to best represent those processes.

Given that there are no known solutions to the prognostic equation set governing the atmosphere, approximate numerical solutions are sought. The governing equations are then approximated in a discrete form (e.g., grid point, spectral representation, finite element) and solved in a discrete fashion. As a consequence, there are a number of processes not well-represented in the discrete versions of the model equations. These phenomena, while not explicitly representable at the resolved scale of the model, can be important implicitly through scale interactions. They include moist convection as well as boundary layer moisture and

momentum fluxes. The representations of the collective effects of subgrid-scale processes in terms of the prognostic variables on the grid scale are known as *parameterizations*.

For each process not representable at grid scale in a model, there may be numerous schemes to parameterize those processes. For a given process, the variety of schemes reflects both the varying complexity of the possible representations and the lack of consensus on how to “properly” represent the process. Some schemes are simple because of convenience in implementation (for use in idealized or conceptual models), while others are necessarily “simple” because they must be used in operational prediction requiring reliably fast computation. Despite differences in formulation, for processes that are reasonably well understood, these seemingly different parameterizations will often produce very similar results. Over time, however, the differences in formulation can have clear, unambiguous impacts on the model state as the parameterized interactions influence larger-scale, grid resolvable phenomena. Below is a brief discussion of important classifications of parameterizations relevant to this dissertation as well as additional supplementary information.

## CUMULUS PARAMETERIZATION

Cumulus parameterization in numerical modeling is the problem of formulating the collective effects of subgrid-scale clouds in terms of the prognostic variables of grid scale (Arakawa, 1993). Given the interactions of moist convective clouds with their larger-scale environment (a scale interaction), cumulus parameterization is needed in any formulation of scale interactions in a moist atmosphere regardless of whether such interactions are being studied in numerical, theoretical, or conceptual models (Arakawa, 1993). To motivate an understanding of the interactions of convection with the larger-scale environment, we briefly consider the impacts of convection on the heat and moisture budgets of the atmosphere.

We begin by considering the Reynolds averaged thermodynamic, water vapor budget, and mass continuity equations:

$$c_p \left( \frac{\overline{\partial \theta}}{\partial t} + \bar{\mathbf{V}} \cdot \nabla \bar{\theta} + \bar{\omega} \frac{\overline{\partial \theta}}{\partial p} \right) = \left( \frac{p_0}{p} \right)^{c_p} (Q_R + L\bar{C} - \nabla \cdot (\overline{s'V'}) - \frac{\overline{\omega's'}}{\bar{p}}),$$

$$L \left( \frac{\overline{\partial q}}{\partial t} + \bar{\mathbf{V}} \cdot \nabla \bar{q} + \bar{\omega} \frac{\overline{\partial q}}{\partial p} \right) = -L\bar{C} - L\nabla \cdot (\overline{q'V'}) - \frac{\overline{(\omega'Lq')}}{\bar{p}}, \text{ and}$$

$$\nabla \cdot \bar{\mathbf{V}} + \frac{\partial \bar{\omega}}{\partial p} = 0,$$

where the overbar denotes the horizontal Reynolds average over an NWP model grid box,  $L$  is the latent heat of vaporization,  $C$  the Reynolds averaged net rate of condensation,  $Q_R$  is the Reynolds averaged radiation heating,  $q$  the water vapor mixing ratio,  $s$  is the dry static energy,  $c_p T + gz$ , and all other variables have their standard meteorological meanings. Typically, the horizontal transports associated with subgrid-scale processes are neglected and the equations governing the area-averaged potential temperature and water vapor as defined by Yanai (1973) are:

$$\frac{\partial \bar{\theta}}{\partial t} + \bar{\mathbf{V}} \cdot \nabla \bar{\theta} + \bar{\omega} \frac{\partial \bar{\theta}}{\partial p} = \frac{1}{c_p} \left( \frac{p_0}{p} \right)^{c_p} Q_1$$

$$\frac{\partial \bar{q}}{\partial t} + \bar{\mathbf{V}} \cdot \nabla \bar{q} + \bar{\omega} \frac{\partial \bar{q}}{\partial p} = -\frac{Q_2}{L}$$

where

$$Q_1 \equiv Q_R + L\bar{C} - \frac{\overline{\omega's'}}{\bar{p}}$$

and

$$Q_2 \equiv L\bar{C} + \frac{\overline{(\omega'Lq')}}{\bar{p}}$$

The terms  $Q_1$  and  $Q_2$ , are the apparent heat source and the apparent moisture sink, respectively. Arakawa (1993) notes that  $Q_1$  and  $Q_2$  differ from the actual heat source and moisture sink due to the Reynolds averaging. Defining  $Q_{1C} = Q_1 - Q_R = L\bar{C} - \frac{\overline{\omega s'}}{\bar{p}}$ ,  $Q_{1C}$  is that part of the apparent heating due exclusively to condensation and convective transport of sensible heat. The difference of  $Q_2$  from the net rate of condensation is the convective transport of latent heat. Removing the net rate of condensation between equations 1 and 2, and defining the moist static energy,  $h = s + Lq$ , yields:

$$Q_{1C} - Q_2 = -\frac{\overline{\omega h'}}{\bar{p}},$$

Given measurements of the horizontal wind field, water vapor mixing ratio, and temperature over a large area (possibly encompassing convection) and over time,  $Q_1$  and  $Q_2$  may be calculated as residuals. If we separate the local tendency of the area-averaged temperature and moisture from the contributions to the tendency by large-scale advective processes, we may rewrite 1 and 2 as:

$$\frac{\partial \bar{T}}{\partial t} = \left( \frac{\partial \bar{T}}{\partial t} \right)_{LS} + \frac{1}{c_p} Q_1$$

$$\frac{\partial \bar{q}}{\partial t} = \left( \frac{\partial \bar{q}}{\partial t} \right)_{LS} - \frac{1}{L} Q_2$$

$$\left( \frac{\partial \bar{T}}{\partial t} \right)_{LS} = - \left( \frac{p}{p_0} \right)^{\frac{R}{c_p}} \left( \bar{\mathbf{V}} \cdot \nabla \bar{\theta} + \bar{\omega} \frac{\partial \bar{\theta}}{\partial p} \right)$$

$$\left( \frac{\partial \bar{q}}{\partial t} \right)_{LS} = - \left( \bar{\mathbf{V}} \cdot \nabla \bar{q} + \bar{\omega} \frac{\partial \bar{q}}{\partial p} \right).$$

The parameterization problem may readily be seen from these four equations where, unlike in the diagnostic studies, the large-scale (area-averaged) temperature ( $\bar{T}$ ) and mixing ratio ( $\bar{q}$ ) are

now unknowns. In short, there are four unknowns (we assume we know the large-scale wind and vertical velocity) and just two equations. In order to close this system, (at least) two additional relationships are needed. These are known as closure assumptions.

Arakawa and Chen (1987) identified three basic types of closure assumptions: those that use a constraint on the (time tendency) of large-scale *states*  $\frac{\overline{T}}{\overline{t}}, \frac{\overline{q}}{\overline{t}}$  by coupling the net moistening and net warming terms (Type I); those that constrain moist convective processes by coupling  $Q_1$  and  $Q_2$  (Type II); and those that assume a coupling between large-scale processes  $(\frac{\overline{T}}{\overline{t}}_{LS}, \frac{\overline{q}}{\overline{t}}_{LS})$  and moist-convective *processes* ( $Q_1$  and  $Q_2$ ) (Type IV). From the preceding discussion, it is clear that cumulus parameterization requires the creation of subgrid-scale implicit clouds, which transport heat, water vapor, and other quantities vertically, generally in the absence of grid-scale saturation. Closure assumptions are required to define the relationship between these implicit clouds and the resolvable (i.e., grid-scale) variables. Given the complexity of the problem, it is perhaps not surprising that many different forms of cumulus parameterization exist. However, all cumulus parameterizations use the inherent assumption that the effects of cumulus convection can be approximated through the use of resolved variables within the model (Emanuel 1994). Briefly, a review of the history of cumulus parameterization schemes are described. While not discussed explicitly, many of these schemes can be derived from the basic closure type describe above.

Early attempts to tackle the cumulus parameterization problem largely revolved around simulating tropical cyclones. Charney and Eliassen (1964) estimated cumulus heating using column moisture convergence, while Ooyama (1964) used boundary layer convergence to estimate cumulus heating. Arakawa (2004) deems this class of parameterization part of the

“convergence school” in that both rely on larger, synoptic scale atmospheric control to aid in determining the atmospheric feedback, in this case in the form of latent heat release. Subsequent work by Ooyama (1969) related cumulus heating to the strength of boundary layer convergence and cloud entrainment (which was allowed to be a variable parameter). Using this parameterization, Ooyama was able to successfully simulate an intensifying tropical cyclone using an axisymmetric model. The work of Charney and Eliassen as well as Ooyama in understanding tropical cyclone development led to the concept of conditional instability of the second kind (commonly known as CISK), which is a theory linking cumulus heating and frictionally induced boundary layer convergence.

Conversely, Arakawa (2004) lists convective parameterizations such as those used by Manabe (1964) as part of the “adjustment” school of thought. Manabe’s (1964) parameterization assumed convection existed when relative humidity exceeded 100% *and* the stratification was conditionally unstable. Afterward, the parameterization adjusts profiles of temperature and moisture until a neutral atmospheric profile is created; which typically involved moistening and cooling of the atmosphere. In this class of cumulus parameterization, stabilization of the atmosphere is effectively the feedback to the model. Manabe (1964) was focused on simulating the larger tropical atmosphere (not specifically tropical cyclones), and it has been shown that his scheme performed reasonably. One criticism of this scheme is that it requires grid-scale saturation (which effectively never occurs in real conditions for grid spacings of the scale 1 km or greater) to represent subgrid scale convection (Cho, 1975).

Arakawa and Schubert (1974) produced a more sophisticated cumulus parameterization in that it allows for a variety of cumulus cloud sizes within a given grid cell, rather than a prescribed cloud type as in earlier models such as Ooyama (1969). This scheme uses a cloud

work function which relates the creation of buoyancy to the production of convection. Furthermore, rising vertical motion within a cloud is balanced by compensating subsidence surrounding the cloud, which warms and dries the atmosphere. The Arakawa-Schubert cumulus parameterization (1974) generally performs well in relation to earlier cumulus parameterizations, in part owing to its complexity. In its most advanced implementation, this cumulus parameterization allows for a variety of entrainment rates. In its current implementation in NCEP's Global Forecast System (GFS) model, it only assumes a single cloud type in order to be more computationally efficient (Stensrud, 2007).

The Kain-Fritsch cumulus parameterization (Kain and Fritsch, 1990 and 1993) is another similarly complex parameterization which produces convection when the following conditions are met: sufficient upward vertical motion at the lifted condensation level exists, CAPE exists in the updraft source layer, and the resultant cloud exceeds a given threshold. Afterward, the parameterization works to remove CAPE based upon a determined convective timescale. This scheme also incorporates variable entrainment and detrainment parameters along with an inclusion of convective downdrafts originating within the cloud. Parameterizations such as these are frequently described as "mass-flux" schemes, owing to their dependence on upward vertical motion to trigger the parameterization as well as their active nature in the rearrangement of mass in the vertical. Unlike the Arakawa-Schubert parameterization, cloud types are not allowed to vary within a grid box for the Kain-Fritsch scheme.

A newer version of adjustment type schemes is the Betts-Miller-Janjic cumulus parameterization (Betts and Miller (1986, 1993), Janjic (1994)). The concept behind the development of this parameterization was the observation noted by Betts (1985) that atmospheric soundings post-convection often exhibited similar structures. As with earlier adjustment

schemes, this scheme relaxes a conditionally unstable atmosphere to an atmosphere representative of the post-convective atmosphere, which is characterized by a convectively neutral environment. The Betts-Miller-Janjic cumulus parameterization uses a more advanced system than earlier adjustment schemes, in this case based upon optimal mixing lines calculated after the cloud base and cloud top are determined. This scheme allows for both shallow, non-precipitating convection as well as deep, precipitating convection. Due to the relative simplicity of the model and its favorable precipitation forecasts, it is a commonly used parameterization in operation models. However, it is known to produce unrealistic atmospheric structures, perhaps due to its simplicity (for example the effects of convective downdrafts are not explicitly included) as shown by Baldwin et al. (2002).

The Grell-Dévényi (Grell and Dévényi, 2002) cumulus parameterization is relatively new compared with the aforementioned parameterizations, and seeks to leverage the varying approaches described above in order to produce a superior parameterization. Grell and Dévényi recognized that in certain situations, some forms of convective parameterization perform well, while different parameterizations perform well in other situations. This scheme uses an ensemble of 144 different specific simplified cumulus parameterizations at each individual grid point, with the averaged result being the feedback to the model. These different members are comprised of differing closures such as a cloud-work function similar to that used in the Arakawa-Schubert scheme, moisture convergence schemes, CAPE relaxation schemes similar to the Kain-Fritsch, and others. Additionally, a number of parameters such as precipitation efficiency and entrainment rate are varied across all closures to produce the 144 total members. Theoretically this parameterization is optimizable such that for different grid points the ensemble member weighting can be adjusted based upon prior observations, although in practice this is not done.

## MICROPHYSICS PARAMETERIZATIONS

As touched upon at the beginning of the above section, accurately predicting clouds is supremely important for NWP models. Aside from the direct effect they have on the atmosphere through redistributions of heat and moisture, and the occasional production of precipitation, clouds have other indirect effects such as radiation-cloud interactions which alter the energy budget for places not directly impacted by the presence of clouds or precipitation. Generally speaking, microphysics parameterizations act to model cloud processes (including water phase conversions and droplet formation and growth) when the relative humidity nears 100% for a given location. This is different from most cumulus parameterizations in two primary ways – many cumulus parameterizations do not require grid scale saturation in order to form clouds and microphysics parameterizations are needed for atmospheric locations in which convection is not occurring, such as stratiform rain regions associated with extratropical cyclones. Certainly microphysics parameterizations are important for many aspects of NWP, but their importance increases as the horizontal grid spacing decreases which allows for more explicit resolution of convective features without using a cumulus parameterization (i.e. at grid scales where the model is able to successfully predict mesoscale processes as described by Weisman et al., 1997). Stensrud (2007) lists two primary challenges specific to microphysics parameterizations. The first is the challenge in properly simulating the conversion between different water phases (all six combinations between solid, liquid, and vapor) and the second is correctly modeling the interaction of cloud and aerosol particles which have various sizes and shapes. As is to be expected, significant assumptions and approximations are necessary; for example, the use of something similar to the Marshall-Palmer droplet size distribution.

There are two classes of microphysics parameterizations, “bulk” and “bin.” Bulk approaches use a predefined function to predict size distributions, while bin approaches divide predicted hydrometeors into discrete bins, as its name suggests (Stensrud, 2007). Due to the computational expense of the latter approach, bulk microphysics schemes tend to be the most common form of microphysics parameterizations. An early parameterization was that of Kessler (1969) which simplified the process by only allowing for rain, cloud water, and water vapor (thus it did not allow for frozen hydrometeors). The WRF Single Moment 3-class (WSM3) is a fairly simple bulk scheme which allows three arrays consisting of water vapor, rain/snow, and cloud water/ice (Hong et al., 2004). Frozen hydrometeors are allowed to exist in the event the temperature is below freezing, although cloud water/ice is considered to be a single quantity internally (as is rain/snow). More advanced versions of this parameterization include the WSM5 (and WSM6) which adds additional arrays such that supercooled water is allowed to exist internally, as is snow as it falls past the freezing level (the WSM6 adds graupel as well as to its associated processes).

As is likely intuitive, as the microphysics parameterization chosen increases in complexity, model performance often improves, although the trade-off frequently is in the form of slower computational time. One parameterization which seeks to maximize computational efficiency while maintaining a reasonable level of complexity is that of the Eta-Ferrier parameterization (Ferrier, 2005). This parameterization only uses two arrays – water vapor and total condensate, which represents the various species of hydrometeors including cloud water, cloud ice, snow, and graupel. The Eta-Ferrier uses prescribed look-up tables which allow the parameterization to partition the total condensate into various forms depending upon atmospheric variables such as temperature. A version of this scheme is currently used in many operational

models such as NCEP's GFS and North American Mesoscale Model (NAM) (Ferrier et al., 2001).

## ENSEMBLE FORECASTING

Thompson (1977) described the theoretical justification for using an ensemble of forecast models rather than a single model. Thompson showed while using a simple idealized situation that an optimal combination of two independent models produced more accurate forecasts on average over many forecasts than either forecast on its own. This is the concept behind ensemble approaches – the notion that several reasonably independent models, when combined in some manner for a given forecast event, should produce superior results on average than a single model forecast. Ensembles can also be used to form an understanding of the various potential outcomes for a given meteorological event. Implicit in the notion of ensemble forecasting is the necessity for many forecasts simulating the same event. This fact made early practical ensemble applications toward real-time NWP modeling quite restricted, as computational resources were limited.

Murphy (1998) notes that given imperfect representations of the observed atmospheric state at time of model initiation (combined with model limitations such as large grid spacing, etc.), individual forecasts are prone to significant error. He suggests that an ensemble of forecasts ideally represents a probability distribution function of the likely possible outcomes. Murphy relates this probability distribution function idea to the above noted imperfect observations of the atmospheric state at model initialization, and suggests using different initial conditions to initialize a model, wherein these differing initial conditions would theoretically represent a probability distribution function of the observed atmosphere at model initialization. Using a simple 5-layer general circulation model, Murphy demonstrates that a combination of eight

forecasts using different initial conditions gives a statistically significant increase in forecast skill of the mean winter atmosphere than does a single forecast.

As available computational resources increased, real-time ensemble use became more feasible. By 1992, the United States National Meteorological Center implemented real-time ensemble forecasting using fourteen members with differing initial conditions (Tracton and Kalnay, 1993). The European Centre for Medium-Range Weather Forecasts (ECMWF) also implemented ensemble modeling around this time, with noticeable improvements to forecasts (Buizza, 1997). One obvious application of ensemble forecasting was toward tropical cyclones, in part because of the importance of their societal impact, but also because tropical cyclones offer very discrete variables that are reasonably easily validated (such as track and maximum wind speed). Zhang and Krishnamurti (1997) studied tropical cyclone forecasting and used various techniques to create a fifteen member ensemble wherein the initial conditions were varied across members (including the location of the tropical cyclone). While their work wasn't in real-time, they demonstrated that a combination of the ensemble members improved tropical cyclone forecasts over several cases. They also discussed the usefulness of ensemble spread in understanding uncertainty – for example, if a given ensemble of tropical cyclone tracks were widely variable, one would expect to have less confidence in the correctness of a mean forecast track in that scenario than one would if that ensemble contained very tightly clustered storm tracks.

All of the above studies used an individual model with differing initial conditions, but a “poor man’s ensemble” is another method to create computationally efficient ensembles. A poor man’s ensemble is an ensemble combining the single forecasts of different operational centers. One benefit to this approach is that both the initial conditions and models are varied. In 2001,

Ebert demonstrated that such an ensemble allowed for increased skill in precipitation forecasting over 48 hour time periods for Australia. However, unlike with tropical cyclone track, Ebert points out that a simple average of precipitation forecasts is not necessarily optimal, as this frequently produces a forecast wherein many areas receive light precipitation, which is an unrealistic distribution. Therefore, for precipitation forecasting, often more advanced techniques are required.

Generally speaking, ensemble modeling has continued to advance along the above trajectory, albeit with greater computational resources allowing for more ensemble members with greater resolution. Advancements in data assimilation have also aided the creation of the initial conditions used in these ensembles. Currently, the Global Ensemble Forecasting System (GEFS) is comprised of 20 members with varying initial conditions. Additionally, endeavors such as the Hurricane Forecast Improvement Project (Gall et al., 2013) incorporate many independent high-resolution forecasts performed across a number of sites with differing models to improve tropical cyclone forecasting. As mentioned earlier, ideally an ensemble of forecasts contains the range of potential outcomes for a given forecast (for example, that the verified position of a tropical cyclone falls within the spread of forecasts outlined by the ensemble used). In practice, this is not always the case. Intuitively, if several ensemble members are better than a single forecast, it is reasonable that continuing to increase the size of the ensemble should aid in this problem. Several studies have attempted to quantify the impact of increasing ensemble size on the ensemble's effectiveness. Buizza and Palmer (1998) examined the effectiveness of the ECMWF ensemble over a 45 day period using ensembles of size 2, 4, 8, 16, and 32. They demonstrated that forecast improvements were found with each successive increase in ensemble size, implying that 32 members are insufficient. For example, the outlier percentage (i.e. the

likelihood that verification will be outside the ensemble spread) fell from 66% to 41% to 26% as ensemble size increased from 2 to 4 to 32.

The above ensemble summary generally consists of situations in which ensembles are generated primarily through varying initial conditions. However, several studies have suggested the usefulness of forecasting using an ensemble comprised of members using different parameterizations. Using differing parameterizations to comprise an ensemble is a comparatively new development. Upon finding considerable sensitivity to varying the cumulus parameterization while studying precipitation forecasting, Gallus Jr. (1999) suggested that many low resolution simulations could provide forecasters with more information than one or few high resolution simulations. Jankov et al. (2005) attempted this approach for forecasting warm-season rainfall with an eighteen member ensemble comprised of members using different parameterizations. They found that no single combination of parameterizations consistently outperformed the others, but rather that certain members performed best depending on rainfall intensity or time since initialization.

Jankov et al. (2007) expanded upon this study by using two sets of initial conditions (for a total of 36 members). They again concluded that no single member significantly outperformed another, while a combination of the members performed best. Both Jankov et al. (2005) and Jankov et al. (2007) found that this approach was effective at increasing model diversity. A separate study of cold-season precipitation using different parameterizations (Jankov et al. 2007) found that different microphysics parameterizations produced statistically significant differences, implying that a well-chosen ensemble could produce a superior forecast compared to a single high-resolution control simulation. The operational Short-Range Ensemble Forecasting (SREF) system has attempted to incorporate the above observations that added physics diversity within

an ensemble aids ensemble forecasting. Currently the SREF uses several different dynamical cores (including WRF) as well as different parameterizations to comprise its ensemble. The earlier mentioned Hurricane Forecast Improvement Project also seeks to incorporate an understanding of parameterization variation as a component of its ensemble.

### *Motivation*

The work presented in this dissertation will focus both on understanding the impacts and potential deficiencies of particular parameterization schemes and will also explore how the differences in parameterization formulation might be exploited to create effective ensemble members for ensemble prediction. It is demonstrated that such parameterization ensembles can produce forecasts with enhanced skill than the individual ensemble members and comparable skill to initial condition ensembles. While there are numerous extant published case studies of forecasts wherein parameterizations are varied and the authors discuss what differences arise in their forecasts it is often difficult to generalize conclusions from these studies and apply them broadly to other cases in large part due to how these studies were conducted, rather than the lack of potential attainable universal conclusions. To motivate the work to follow in this dissertation, the general characteristics of existing studies are described and the common weakness in the approach taken is identified.

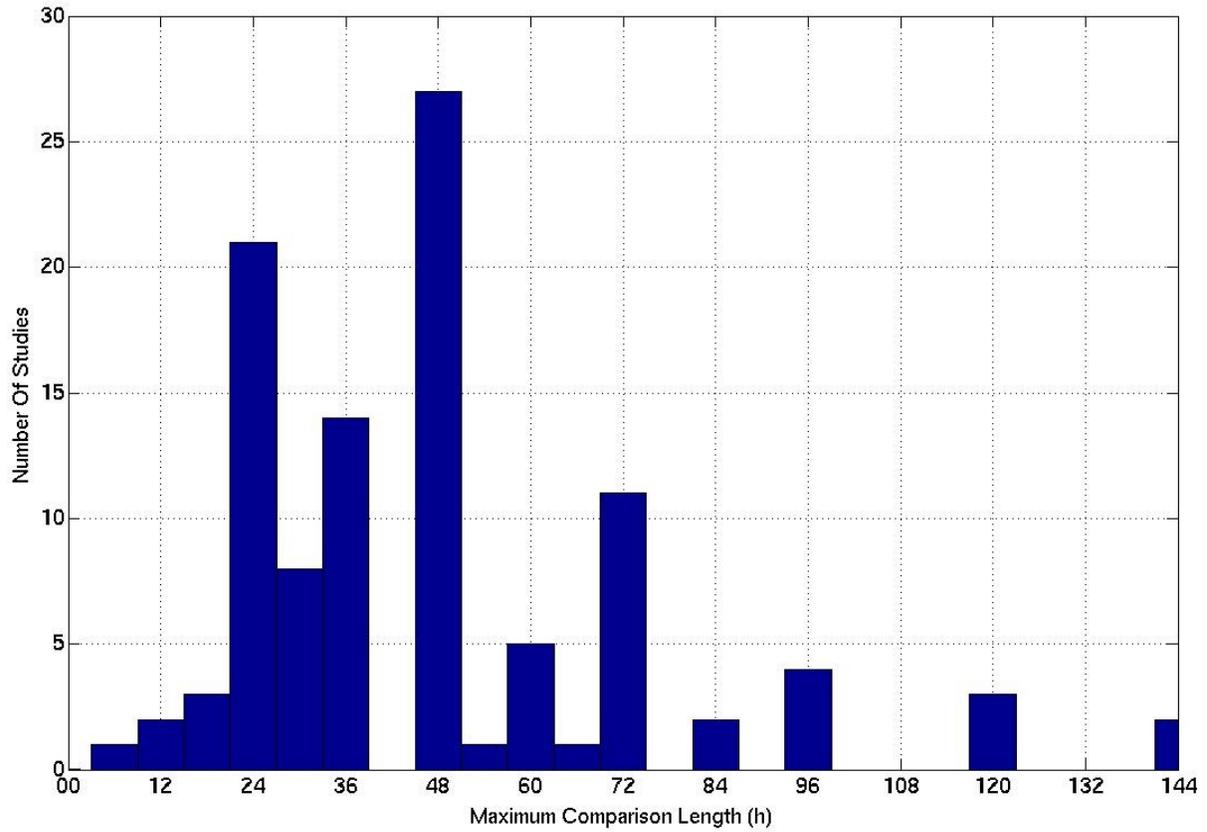
To achieve this goal, more than 200 peer-reviewed articles<sup>1</sup> published between the year 2000 and 2012 which in some manner compared the effect of varying parameterizations across a number of journals were studied. This time frame was chosen for two reasons. First, during this period several community models became widely accessible (such as the MM5 model and its

---

<sup>1</sup> The method used here consists of using two search terms – “Kain” and “Betts,” after two commonly used cumulus parameterizations. In articles in which both of these terms were found, characteristics of the study were noted if *any* parameterizations were compared, regardless of whether or not the parameterizations were these cumulus parameterizations.

successor, the WRF model). Second, computing resources greatly expanded during this period, which made sophisticated desktop modeling much more feasible. While this is in no means an exhaustive list of every peer-reviewed article on the subject, it should be of sufficient size to develop some general conclusions. One of the characteristics noted was the absolute maximum forecast length in which any two parameterizations was compared.

**Figure 1.1** depicts a histogram of maximum comparison length (in hours) as well as the total number of studies for all studies which were not climate studies. It is immediately clear that the vast majority of parameterization studies use approximately a 24-72 hour forecasting timeframe. The mean forecast length examined within these articles is 48 hours. (It should be noted that 33 climate studies were not included on this graph. These studies ranged from 10 day comparisons to 29 years. Several of these studies will be discussed in greater detail later). Of the studies which compared parameterizations over forecasting length beyond 72 hours, the vast majority were tropical cyclone case studies. In fact, the vast majority of all of these studies were of the case study variety in which only a few particular cases were examined. Of the different forms of case study, tropical cyclones, convection, and intense extratropical cyclones were the overwhelming majority of cases. Of those few instances of parameterization studies which were not limited to particular case studies, the vast majority were model evaluation or ensemble verification studies, commonly using precipitation or surface variables as forms of verification. Of all non climate-scale studies, only three focused on oceanic domains which were not related to tropical cyclones. Two of these three examined tropical convection while the third studied a polar low case. All studies (including climate scale studies) used a limited area model (meaning lateral boundaries were prescribed in some manner).



**Figure 1.1** A histogram of the maximum length (h) of model simulation in which any parameterizations were compared.

Collectively, these studies indicate that the most common form of parameterization study follows a familiar pattern – most are short term case studies using a limited area model over land (with the addition of several tropical cyclone case studies) of high-impact weather events. However, these studies largely leave many meteorological areas unstudied. For instance, most daily weather is not of the “case study” variety. Second, land only comprises approximately one-third of the Earth’s total surface area. Third, many operational forecast centers issue forecasts for much longer than the 24-72 hour forecast period most of these studies examine. For example, national agencies like the National Weather Service as well as private companies such as The Weather Channel routinely forecast up to 10 days in advance (Bickel et al., 2011). Additionally, more focused forecast activities such as the Hurricane Forecast Improvement Project, have a goal of improving tropical cyclone track, intensity, and structure forecasts out to a lead time of seven days (Gall et al., 2013). Extended range forecasts (between three and ten days) often greatly depend on global models such as the GFS and the ECMWF models. This means that the only method in which conclusions learned from the aforementioned studies can be transferred to long-term day-to-day forecasts is to make the following assumptions:

- (1) Parameterizations behave in a similar manner across multiple environments (e.g. winter/summer, land/ocean)
- (2) Parameterizations behave in a similar manner across varying forecast durations (e.g. their performance at hour 24 is effectively equivalent to later forecast hours)
- (3) Parameterizations perform similarly for day-to-day weather as they perform in case studies (typically conducted for “high-impact” weather events).

If any of these assumptions were invalid – and often at least one is, then conclusions derived from most parameterization studies are not necessarily applicable to many daily

forecasts. Even one of the most organized programs of model testing and evaluation performed by the Developmental Testbed Center (Bernardet (2008), Wolff (2012)) implicitly incorporates the above assumptions. The Developmental Testbed focuses primarily on model verification and parameterization testing (mostly with the WRF model). While the robustness of their endeavors is admirable, they suffer from many of the same potential problems associated with other studies. Their forecasts used for comparison are generally 36-48 hour forecasts of continental events, which is the same pattern noted for many other studies.

Warner (2011) provides a brief overview of current practices in general modeling studies while also providing suggestions for the improvement of such studies. While Warner does not specifically address the type of modeling studies performed (as above), he does note that the proliferation of easily accessible NWP models (like MM5 and WRF), along with users with limited meteorological or NWP backgrounds, in addition to little time invested in verification lends itself to poor practices. While Warner provides an extensive list for ways to improve generic modeling studies (not parameterization specific), those that are most pertinent to parameterization studies include:

- (1) Identifying clearly the proposed purpose and goals of the study in advance;
- (2) Determining what forms of model evaluation and verification will take place in advance, with the knowledge that using multiple forms of evaluation/verification is best;
- (3) Performing sensitivity tests for studies conducted in a limited domain to determine the optimal size and location for the domain; and
- (4) Testing the model for many different types of meteorological phenomena if the model is to be used for operational NWP.

Many of the parameterization studies alluded to above follow these best practices, while many do not. Certainly, studies of the form “Tropical cyclone X was modeled which demonstrated that parameterizations A, B, and C produced differing forecasts” do not satisfy the best practices suggested by Warner. Ideally, parameterization studies should seek to either provide some sort of generally applicable conclusion regarding a parameterization or class of parameterizations, or advance the community’s knowledge of atmospheric processes through parameterization studies (or both). Given this goal, and taking into account the types of parameterization studies currently lacking in the literature, broadly speaking this dissertation will seek generalizable conclusions regarding several parameterizations while also furthering the collective meteorological knowledge where possible through a series of WRF-ARW simulations examining various features of tropical and subtropical regions using both limited area and global domains covering topics ranging from tropical rainfall to tropical cyclones to the global circulation.

Due to the sheer volume of available parameterizations, an all-inclusive parameterization study would be impossible to complete due to the number of different configurations possible, as well as the rapidly growing number of available parameterizations. For example, in the latest release of the WRF-ARW (Skamarock et al., 2005), there are no fewer than 70 changeable parameterizations of different types (cumulus, microphysics, radiation, etc.). If the changeable constants (such as friction and dispersion coefficients) are included, there are a minimum of two trillion potential WRF-ARW model configurations. Given these obvious limitations, this research will focus on several commonly used cumulus parameterizations, and will incorporate two parameterizations of other processes (microphysical and boundary layer processes) for the purposes of ensemble generation. Furthermore, this research will begin with a detailed analysis

of one particular parameterization (the Grell-Dévényi cumulus parameterization) and will end with a slightly broader analysis of two physics ensembles. To the greatest extent possible, this research will follow the best practices as outlined above by Warner (2011).

Before a detailed description of the upcoming research is given, some general descriptions are warranted. For a given comparison of parameterizations, all other parameterizations, grid spacings, and coefficients will be held constant in order to isolate the impacts of the changeable parameterizations under study. In order to study parameterizations and their effects in a statistically meaningful way, a significant amount of data is required. Furthermore, it is important to limit the amount of degrees of freedom as much as possible, therefore this research will frequently employ a similar approach: for a given parameterization comparison, a meteorological forecast or simulation (for example, a tropical cyclone forecast) will be conducted repeatedly using the same initial conditions and WRF-ARW model configuration, except for the changeable parameterizations. All forecasts will be medium to long term forecasts in the 5 to 54 day range. Afterward, additional forecasts will be made in this same manner for different cases, until enough data is generated such that, where possible, statistically significant conclusions can be drawn. This strategy has an added benefit in that in some studies every forecast is essentially an ensemble of forecasts, with the different parameterization combinations being the different members of that ensemble. The approach taken allows this research on differing parameterizations to occasionally serve also as research involving the utility of parameterization ensembles versus more traditional ensembles such as NCEP's Global Ensemble Forecast System (GEFS), or a conventional multi-model ensemble. The cases chosen for this research will generally be geared towards forecasting tropical cyclones and their larger-scale tropical environment.

*Dissertation Synopsis*

This dissertation is organized into five subsequent chapters. The work contained within these chapters will seek to follow the best practices as outlined by Warner (2011). Additionally, they seek to begin addressing some deficiencies in current parameterization research. For instance, Chapter 2 deliberately eschews short-term land-based case studies in favor of many oceanic and global simulations. Chapter 2 examines the impact of the current WRF implementation of the Grell-Dévényi CP on the large-scale environment and the precipitation fields relative to the implementations of other cumulus schemes. Using the Grell- Dévényi CP as an example, it is demonstrated that (1) this parameterization performs very differently at different times as forecasts progress, and that (2) it performs quite differently over land versus over water, particularly when compared with other cumulus parameterizations. This example raises significant concerns regarding the universal applicability of many of the results of parameterization studies described earlier. Chapter 3 answers the question “*do different cumulus and microphysics parameterizations exhibit systemic biases when used in tropical cyclone track, intensity, structure, or genesis forecasts at time scales up to 180 hours?*” It is shown in Chapter 3 that there are indeed systematic biases in using particular (combinations of) parameterization schemes. Chapter 4 exploits the existence of the biases found in Chapters 2 and 3 and addresses two questions directly pertaining to ensembles. First, “*given the deficiencies present in the Grell-Dévényi CP, are there statistical applications that can ameliorate these deficiencies?*” and second, “*can a low-resolution physics-based ensemble produce comparable five day forecasts when compared with an initial condition ensemble of higher resolution?*” The former addresses an ensemble within a parameterization while the latter address an ensemble of different

parameterizations. Chapter 5 contains a summary of the key results and lessons form Chapters 2 through 4 and directions for further research.

## **Chapter 2: The Grell-Dévényi Cumulus Parameterization**

### *Introduction*

The Weather Research and Forecasting model using the Advanced Research Core (WRF-ARW) (Skamarock et al., 2008) currently provides eight distinct cumulus parameterization (CP) options, not including additional tunable parameters within each option. While the diversity of options is partially due to varying degrees of complexity, it is also due to a variety of uncertainties regarding how to properly address the cumulus parameterization problem (Arakawa, 2004). Many previous studies have aimed to understand which CPs perform best in certain situations using case studies of a variety of meteorological phenomena including severe thunderstorm outbreaks, midlatitude cyclones, and tropical cyclones. However, there is little agreement concerning the “best” choice of CP because this determination is (1) highly case-dependent, (2) somewhat subjective, and (3) the vast majority of existing studies employ examinations of short-term forecasts for continental regions (less than 48 hours, as described in Chapter 1), therefore the conclusions therein may not be representative of other environments or forecast lengths.

Certain CPs may perform better in certain situations than others depending on terrain, season, meteorological event, length of model integration, grid spacing, and other factors. However, given enough data, robust conclusions can be reached regarding the performance of different CPs. The work presented in this chapter will demonstrate that there are certain meteorological regimes in which the Grell-Dévényi CP creates deficient forecasts relative to two other commonly used CPs – the Kain Fritsch (KF) CP and the Betts-Miller-Janjic (BMJ) CP, when considering either realism of forecast precipitation or large-scale atmospheric variables. This will be shown using a series of five-day limited area WRF-ARW forecasts as well as a

comparison between different WRF-ARW global simulations. The limited area model simulations are used to generate a significant amount of data valid at a given forecast hour, which makes parameterization comparison more legitimate and less prone to a single extreme event. The global simulations are useful for comparing the cumulative effects that varying parameterizations have on the atmosphere over extended forecast periods. In either case, the use of extended forecasts times, as well as forecast domains not restricted to continental areas, allow the results presented herein to be applied more generally.

The results obtained in this paper appear to explain several unexplained results from other studies. One such study by Wehner (2011) examined the incidence of extreme precipitation over a number of North American Regional Climate Change Assessment Program (NARCCAP) ensemble members over North America for all seasons. Wehner found that, with respect to mean precipitation, the NARCCAP member using the Grell-Dévényi CP performed approximately equivalently as all other members, yet an examination of extreme precipitation events found that the ensemble member using Grell-Dévényi CP was a significant outlier that produced return rates for extreme precipitation which were far too long compared with reality (i.e. that the member using the Grell-Dévényi CP had greater difficulty producing heavy precipitation). Similarly, Kawazoe and Gutowski (publication pending) analyzed intense precipitation characteristics within the NARCCAP models and found that the member using the Grell-Dévényi CP produced infrequent consecutive heavy rainfall days relative to either other models or observations.

Another study by Mukhopadhyay et al. (2010) examined forecasts of the Indian monsoon using the Grell-Dévényi CP, the KF CP, and the BMJ CP. In a result similar to Wehner (2011), Mukhopadhyay et al. found that the Grell-Dévényi CP produced significantly more light

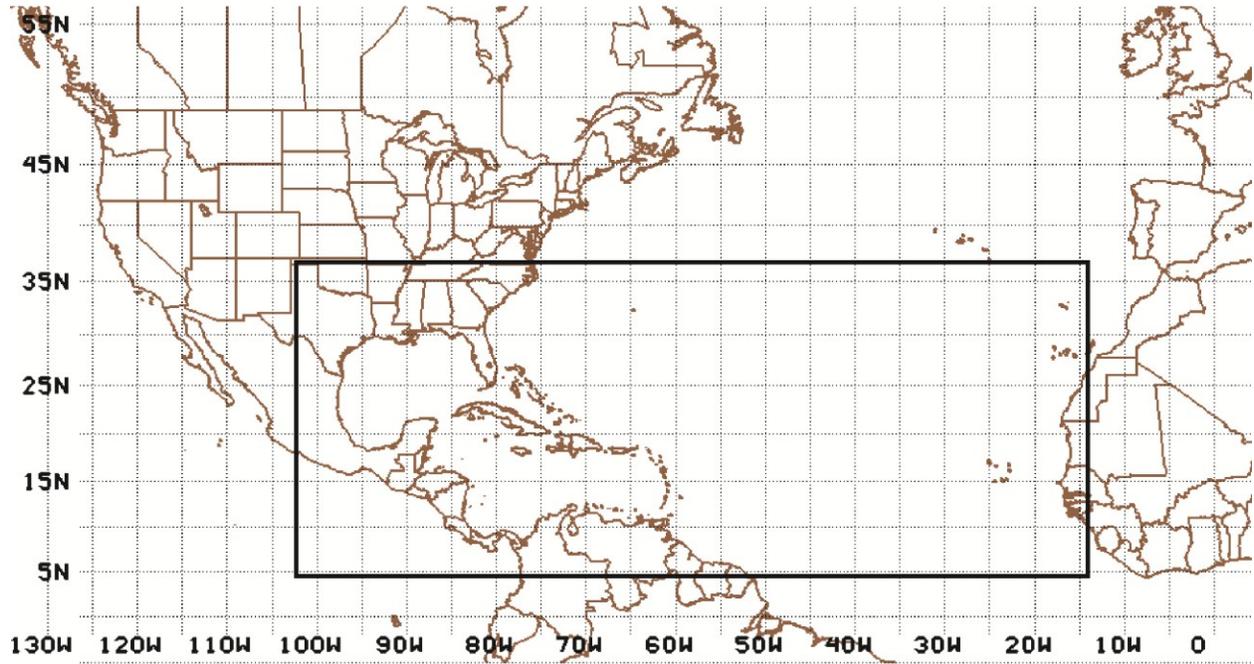
precipitation (rather than heavy precipitation) relative to either of the other CPs examined. Associated with an underproduction of intense rainfall with the Grell-Dévényi CP was less intense upward vertical motion when compared with other CPs. Several other studies such as Crétat and Pohl (2012) and Gilliland and Rowe (2007) found similar results when examining African convection and supercell development, respectively (i.e. the Grell-Dévényi CP underproduced intense precipitation). The commonality amongst these studies is that all of them were examining extreme rain events. Collectively, use of the Grell-Dévényi CP has been gaining popularity, and to date has been used in well over one hundred peer-reviewed articles.

### *Model Description And Methodology*

All results shown in this chapter were obtained using the WRF-ARW model (Skamarock et al., 2008). Some of the data shown were obtained from a series of 76 “real-time” WRF-ARW version 3.0 physics ensembles initialized using the 0000 UTC Global Forecast System (GFS) forecasts<sup>2</sup> for both initial and boundary conditions. One ensemble was constructed every other day between early June and late October 2009. These 76 ensembles employ ten different combinations of physics options, although for the purposes of this chapter only three will be discussed – those using the KF, and BMJ, and G3 CPs, with all other parameters being held constant. Additionally, four forecasts were created for each ensemble representing a unique sub-ensemble of the G3 CP (to be explained in greater detail below). Each simulation is run 120 hours using the domains shown in **Figure 2.1**. The outer (inner) domain uses a 90 km (30 km) horizontal grid spacing and both domains use 28 vertical levels and a Mercator projection. The outer (inner) domain has a lower-left location of 4.98° S, 134.17° W (4.98° N, 100.83° W) and an upper-right location of 56.02° N, 4.17° E (36.51° N, 15.15° W). The notable parameterizations which remain constant are the WSM3 microphysics, the Yonsei University

---

<sup>2</sup> Details can be found at: <http://www.emc.ncep.noaa.gov/GFS/impl.php>



**Figure 2.1** The area covered by the 90 km outer domain is shown, as well as the 30 km inner domain (denoted by the rectangular box).

boundary layer parameterization, the 5-layer thermal diffusion land surface model, the Rapid Radiative Transfer Model longwave parameterization, and the Dudhia shortwave parameterization (descriptions of these can be found in Skamarock et al., 2008). For the purposes of this study, only inner domain data will be used.

Other data is generated using global WRF-ARW, version 3.3.1. The global simulations have a  $.73^\circ \times .73^\circ$  grid spacing, which at the equator is equal to a horizontal grid spacing of 80 km. All global simulations have 28 vertical levels. The global simulations are run from a period of 0000 UTC 1 January through 19 0000 UTC February and 0000 UTC 1 July through 0000 UTC 24 August, 2005. These simulations are initialized with the appropriate GFS Final Analysis, and are updated once daily with observed sea surface temperatures. The 1 January - 19 February (1 July – 24 August) simulation will be referred to as the winter (summer) simulation.

The Grell-Dévényi (GD) CP and its updated version (G3) use a sophisticated ensembling approach employing multiple closure assumptions (moisture convergence, CAPE removal, vertical velocity, and a cloud work function) combined with differing changeable parameters (such as precipitation efficiencies and detrainment rates among others) as described in Grell and Dévényi (2002). For a given grid point, 144 separate ensemble members are run, with the averaged result being the feedback to the model. One of the philosophies behind this approach is the above-mentioned notion that different CP types perform well in some situations and not in others. By combining different forms of CPs, ideally a superior result would be obtained. The four additional G3 CP simulations referred to above use an individual closure variety (moisture convergence, CAPE removal, vertical velocity, and a cloud work function) while still allowing for other variations in constants such as precipitation efficiency. For each of these simulations, the effective ensemble size is nine members.

The KF CP employs a sophisticated cloud model that seeks to remove CAPE over a given convective time-scale. The CP is activated when low-level vertical motion exceeds the threshold such that any existing convective “cap” can be broken, CAPE exists for the parcel source layer, and the resulting cloud exceeds a prescribed depth. Both the G3 CP and the KF CP are mass-flux schemes. The BMJ CP is an adjustment scheme that seeks to adjust profiles of temperature and moisture toward reference profiles characteristic of post-convective atmospheres (i.e. moist neutral). The CP is activated if CAPE exists and the convective cloud depth is sufficiently large. While the KF and the BMJ CPs are not the focus of this chapter, they serve as useful points of comparison when discussing the G3 CP. The KF and BMJ CPs are chosen due to their extensive use in many studies examining parameterizations as well as their use in several operational models.

All model output is converted into GEMPAK format using the program *wrf2gem*<sup>3</sup>. Afterward, all statistical analysis is performed using MATLAB. Trajectory analysis is performed using GEMPAK.

## RESULTS

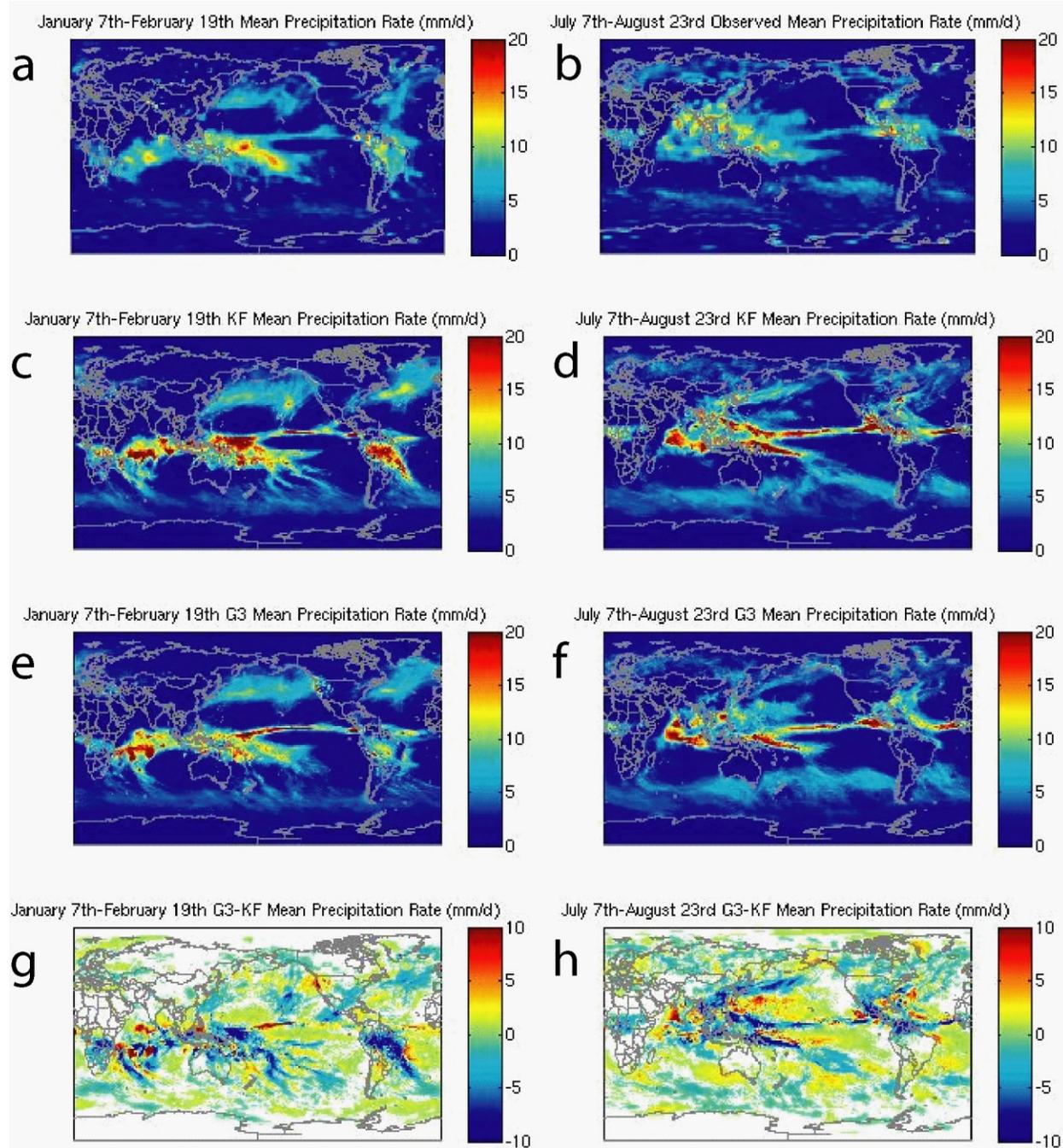
### *Description Of Precipitation Patterns*

**Figure 2.2** shows the mean rain rate for both the KF (**Fig. 2.2c,d**) and G3 CPs (**Fig. 2.2e,f**) for the winter and summer simulations as well as their difference (**Fig. 2.2g,h**) and the estimated rainfall<sup>4</sup> for this same period (**Fig. 2.2a,b**). While it is probably unwise to compare exact values of model generated precipitation to satellite estimated precipitation (due to the fact estimated rainfall is an imperfect representation of actual rainfall while model precipitation is a perfect representation of model rainfall), the estimated precipitation does serve as a helpful point of

---

<sup>3</sup> Downloadable from <http://envsci.rutgers.edu/~decker/wrf2gem/>

<sup>4</sup> Data provided by the NOAA/ESRL Physical Sciences Division, Boulder Colorado from their Web site at <http://www.esrl.noaa.gov/psd/>

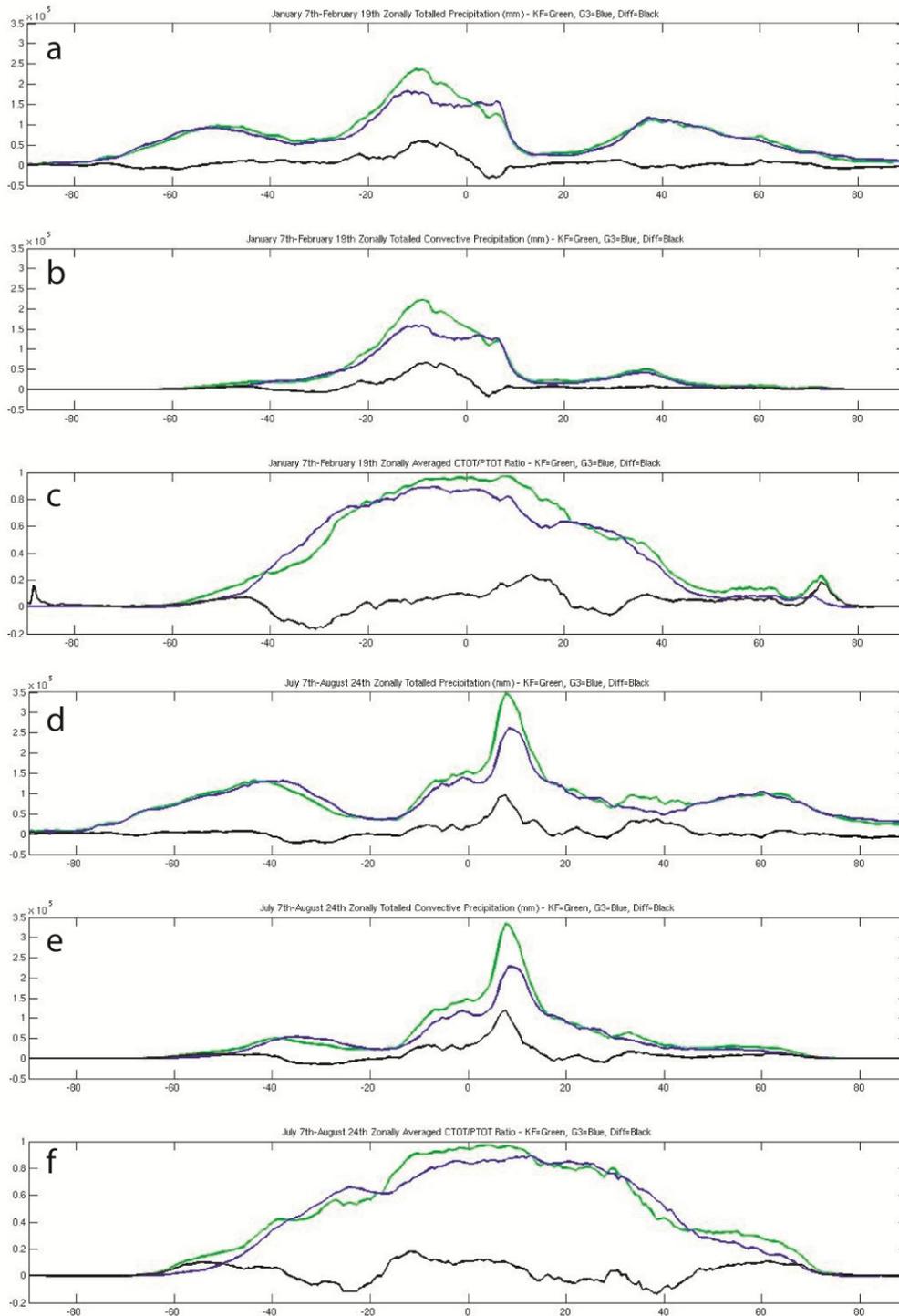


**Figure 2.2** (a) Daily satellite estimate of mean rain rate (mm/day) for the period 7 January – 19 February 2005, (c) forecast daily rain rates for the KF CP, (e) G3 CP, and (g) the difference between (c) and (e). (b), (d), (f), and (h) are as for (a), (c), (e), and (g) except for the period 7 July – 23 August 2005.

comparison for relative magnitudes and locations. Upon first glance, the different simulations seem fairly similar for a given season, in that precipitation amounts and patterns (ITCZ, extratropical storm tracks, etc.) are approximately equal. However, upon closer examination, several subtle differences are observed.

For the winter simulations, the G3 CP produces smaller rain rates than the KF CP over south central Africa, most of South America, and the western Pacific Ocean maritime continent region. Conversely, the G3 CP produces slightly heavier precipitation along the eastern edge of the Pacific Ocean and western North America, the subtropical regions of the Pacific Ocean north of the equator, as well as the majority of the tropical and subtropical Atlantic Ocean south of the equator. Generally speaking, the summer simulations reproduce the above pattern, with the primary difference being a northward progression for regions in which the KF CP produces heavier precipitation over land. The KF CP produces significantly heavier precipitation over northern South America, southern North America, southeastern Asia, and most of the maritime continent region of the western Pacific Ocean. A general conclusion is that the KF CP produces greater precipitation in areas in which there is already heavy precipitation (such as over the maritime continent) while the G3 CP produces heavier precipitation over areas in which precipitation is light (such as over the subtropical oceans).

**Figure 2.3** shows the zonally averaged totals of precipitation, CP generated precipitation, as well as the ratio with which all precipitation is generated by the CP for the winter and summer simulations. This indicates two things: (1) most CP generated precipitation falls between approximately  $-20^{\circ}$  S and  $10^{\circ}$  N ( $-10^{\circ}$  S and  $15^{\circ}$  N) for the winter (summer) simulation, and (2) this region also corresponds to the region where the largest fractions of rainfall are generated by CPs, with both CPs accounting for approximately 80% of total rainfall in this region. Although



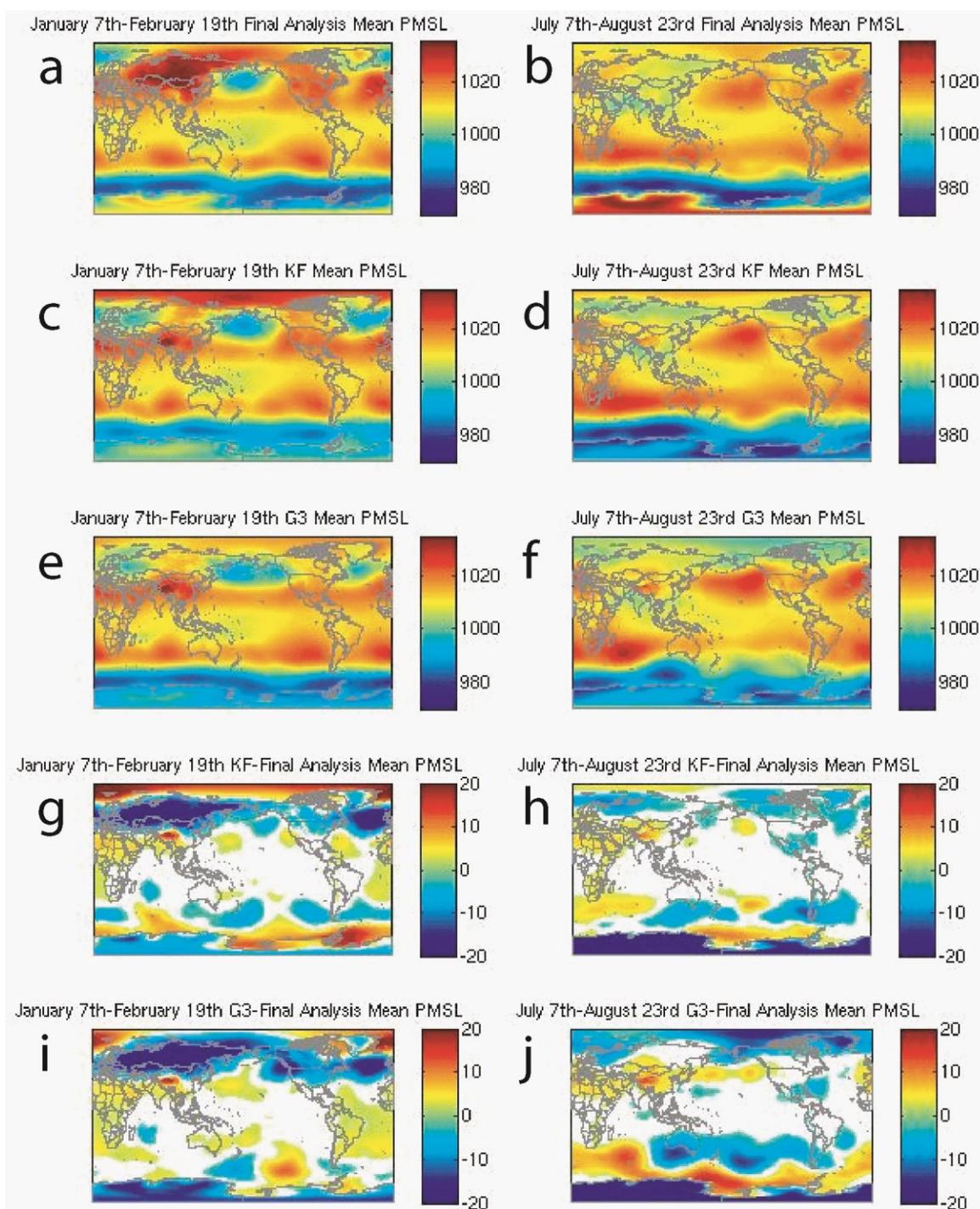
**Figure 2.3** (a) shows zonally totaled forecast precipitation for the KF CP (Green), G3 CP (Blue), and difference (Black) for the period 7 January – 19 February 2005, (b) same as (a) except only convectively parameterized precipitation, (c) as (a) except for the percentage of precipitation which is parameterized. (d), (e), and (f) same as (a), (b), and (c) except for the period 7 July – 23 August 2005.

the exact values are somewhat different for the above rainfall totals and convectively generated fraction of precipitation, both CPs demonstrate quite similar characteristics for a given latitude for either season, particularly in regions far from the equator, which indicates their relative inactivity for these regions.

Regarding the limited domain simulations, an initial examination of mean total rainfall patterns for the duration of the 120 hours forecasts yields similar patterns among the different CPs (not shown) – a primary maximum exists in the eastern Pacific Ocean, a secondary maximum exists along the ITCZ in the Atlantic Ocean, and a third maximum exists along the Gulf Stream off the coast of the southeastern United States, with decreasing amounts throughout the remainder of the domain. These patterns are similar to observed rainfall distributions during the summer and fall of 2009. Additionally, the mean domain-averaged precipitation is fairly similar for each CP (31.71 mm, 28.62 mm, and 28.34 mm for the KF CP, BMJ CP, and G3 CP respectively), as is the standard deviation for domain-averaged precipitation for all cases (4.57 mm, 4.57 mm, and 4.43 mm for the KF CP, BMJ CP, and G3 CP respectively). This indicates that for this time span (120 hours) the different CPs behave similarly in terms of total aggregate precipitation. However, an analysis of a shorter timescale – three hour precipitation ending at time 120, demonstrates a more substantial difference between the three CPs. The mean over all 76 cases of standard deviation of three hour precipitation within the domain gives values of .11 mm, .11 mm, and 0.06 mm for the KF CP, BMJ CP, and G3 CP, respectively. This indicates that over a long duration (120 hours) these three CPs produce a similar *total* amount of rainfall, albeit through differing *intensities* of rainfall. The smaller standard deviation found with the G3 CP is a statistically significance difference when compared with either the KF CP or BMJ CP over all cases.

### *Description Of Related Atmospheric Variables*

As described above, an examination of the global simulations as well as a collection of real-time simulations suggests that the G3 CP is more disposed to produce light precipitation relative to the KF CP (and BMJ CP in the case of the real-time simulations), and perhaps less able to produce very intense precipitation. Intuitively, differing distributions of rainfall intensities should affect the larger global circulation over longer timescales given the attendant differences in latent heat release. The global simulations provide an opportunity to examine this. **Figure 2.4** shows the observed sea level pressure (**Fig. 2.4a,b**) as well as the simulated sea level pressure (**Fig. 2.4c,d,e,f**) and the difference (**Fig. 2.4g,h,i,j**) between the two for each simulation and season. Although it isn't realistic to expect seven week forecasts to accurately reproduce the observed atmosphere, a comparison with reality serves as a useful benchmark for understanding the observed differences between the G3 CP and KF CPs. For both CPs, the greatest departures from observed generally occur in the winter hemisphere of the simulations. For instance, during northern winter, both CPs under represent the intensity of the Siberian high. However, the remainder of this description will focus on locations where only one of the CPs differ significantly from reality. This is for two reasons – (1) the focus of this paper is on what makes the G3 CP simulations different from other CPs, and (2) many of the shared differences between the simulations is due to other factors (such as surface parameterization deficiencies which lead to overly warm or cold continental regions over extended forecasts) which are not the focus of this paper. The G3 CP shows anomalously low pressure over the northeastern Pacific Ocean and northwestern North America during the northern winter (**Fig. 2.4i**) as well as an enhanced wave pattern over the southern hemisphere winter (**Fig. 2.4j**). Both of these are consistent with an

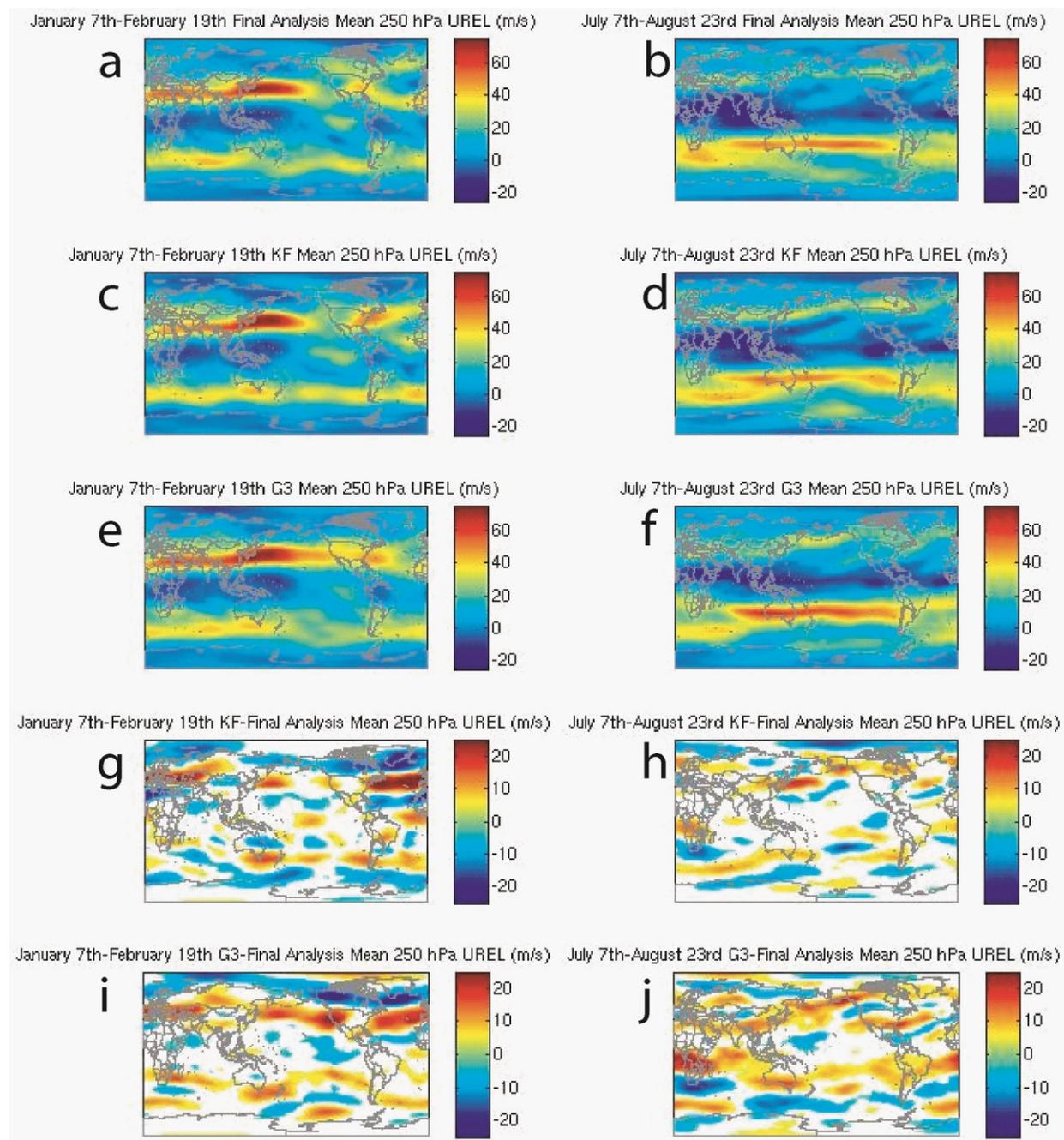


**Figure 2.4** (a) Final Analysis mean sea level pressure for the period 7 January – 19 February 2005, (c) mean forecast sea level pressure for the KF CP, (e) the G3 CP, (g) the difference between (a) and (c), (i) the difference between (a) and (e). (b), (d), (f), (h), and (j) are as for (a), (c), (e), (g) and (i) except for the period 7 July – 23 August 2005.

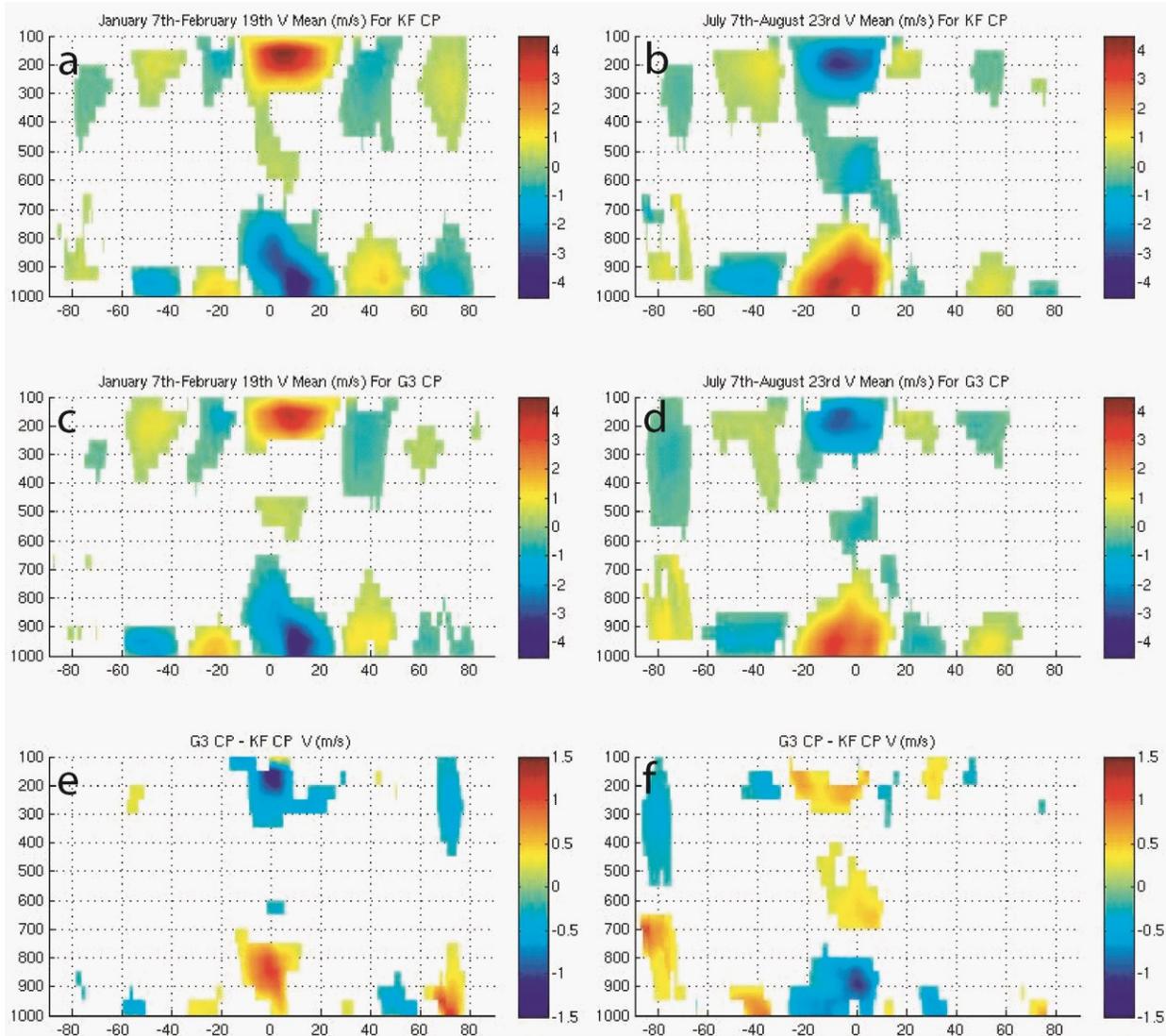
increase in precipitation for these regions when compared with the KF CP as seen in **Figure 2.2**, which does not possess these sea level pressure differences.

**Figure 2.5** shows the observed 250 hPa zonal wind (**Fig. 2.5a,b**) as well as the simulated 250 hPa zonal wind (**Fig. 2.5c,d,e,f**) and the difference (**Fig. 2.5g,h,i,j**) between the two for each simulation and season. Again, the largest departures from reality are seen in the winter hemisphere. The primary feature is an equatorward shift in the mean zonal jet in the winter hemisphere for the G3 CP simulations – which is most notable over the northeastern Pacific for the winter simulation as well as over Africa and Australia in the summer simulation. Although the latter pattern is also observed in the KF CP summer simulation, its magnitude is significantly muted. The KF CP summer simulation does have an anomalous jet structure near Japan, although a closer examination of the simulation reveals that this is due to the presence of two slow-moving tropical cyclones in this area (not shown) which did not exist in reality.

**Figure 2.6** shows a meridional cross-section of the zonal and time average of meridional winds (**Fig. 2.6a,b,c,d**) as well as the difference (**Fig. 2.6e,f**) between the two for each simulation and season. The latitudes of maximum precipitation for each season (as seen by **Figure 2.3**) serve as a proxy for the location of the mean ITCZ (roughly  $-8^{\circ}$  and  $8^{\circ}$  for the winter and summer simulations, respectively). The mean inflow to the latitude of the ITCZ can be easily seen as well as the mean outflow at upper levels. However, the difference plots indicate fairly substantial differences in these regions. In both seasons, the G3 CP has a weaker inflow/outflow couplet than the KF CP of approximately 1.5 m/s. Although this value is not particularly large in an absolute sense, it represents approximately 33% of the maximum value for these regions. These differences are located predominantly in the winter hemisphere. Due to mass continuity



**Figure 2.5** As in Figure 2.4, except for 250 hPa zonal wind (m/s).



**Figure 2.6** Zonally and temporally averaged cross-section of forecasted meridional wind (m/s) for the (a) KF CP during the period 7 January – 19 February 2005, (c) as for the G3 CP, (e) the difference between (a) and (c). (b), (d), and (f) are as for (a), (c), and (e) and except for the period 7 July – 23 August 2005.

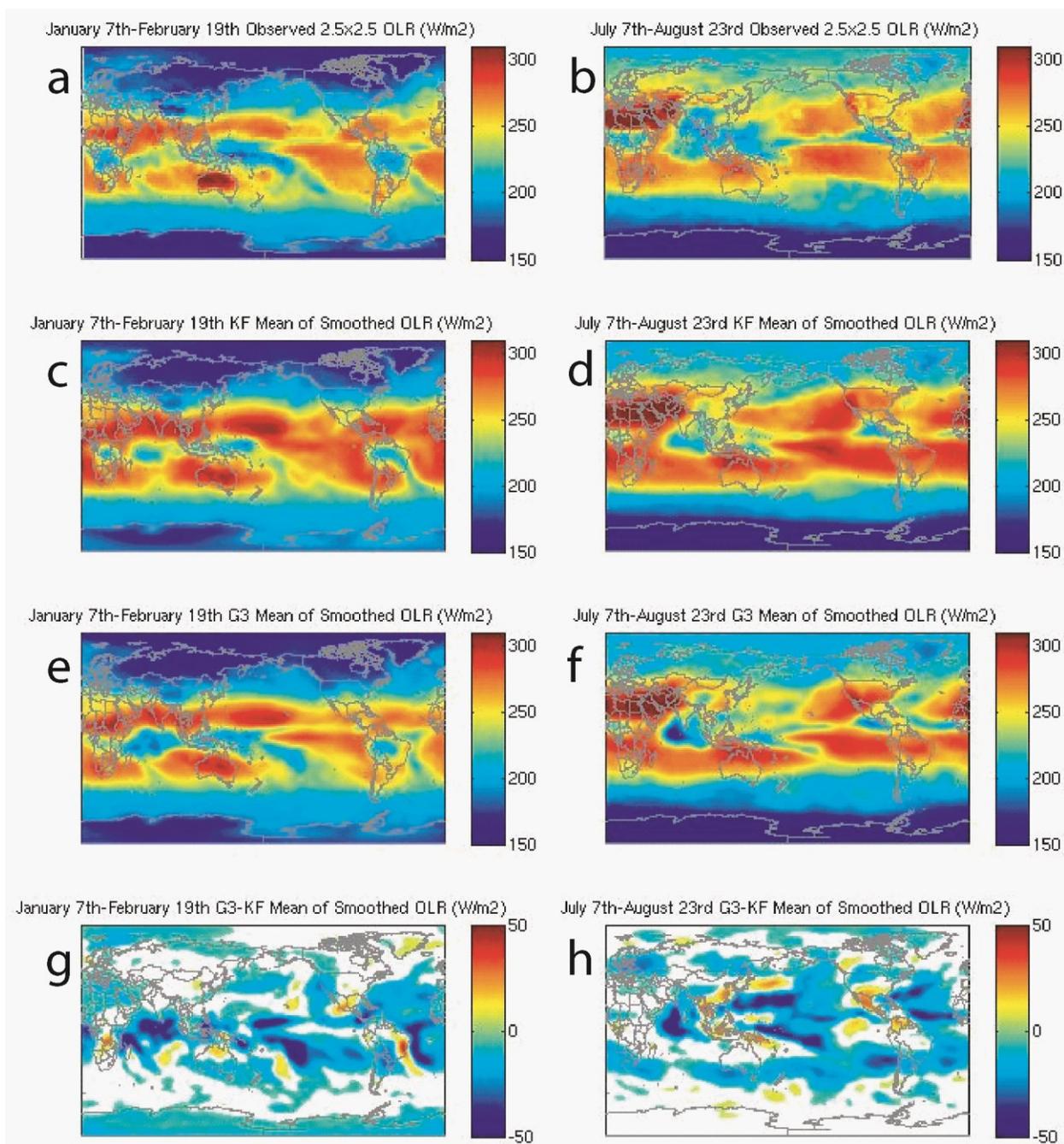
considerations, one would then expect the G3 CP to exhibit less mean upward vertical motion in the ITCZ region when compared with the KF CP, and this is what is seen (not shown).

A detailed description of atmospheric variables is not provided for the real-time simulations due to its redundancy with the above descriptions of the global simulation (although obviously on a smaller scale). Weaker inflow (outflow) at low levels (upper levels) is seen as well as weaker ascent (descent) in the ascending (descending) branch of the Hadley Cell is observed with the G3 CP in comparison to either the KF CP or BMJ CP.

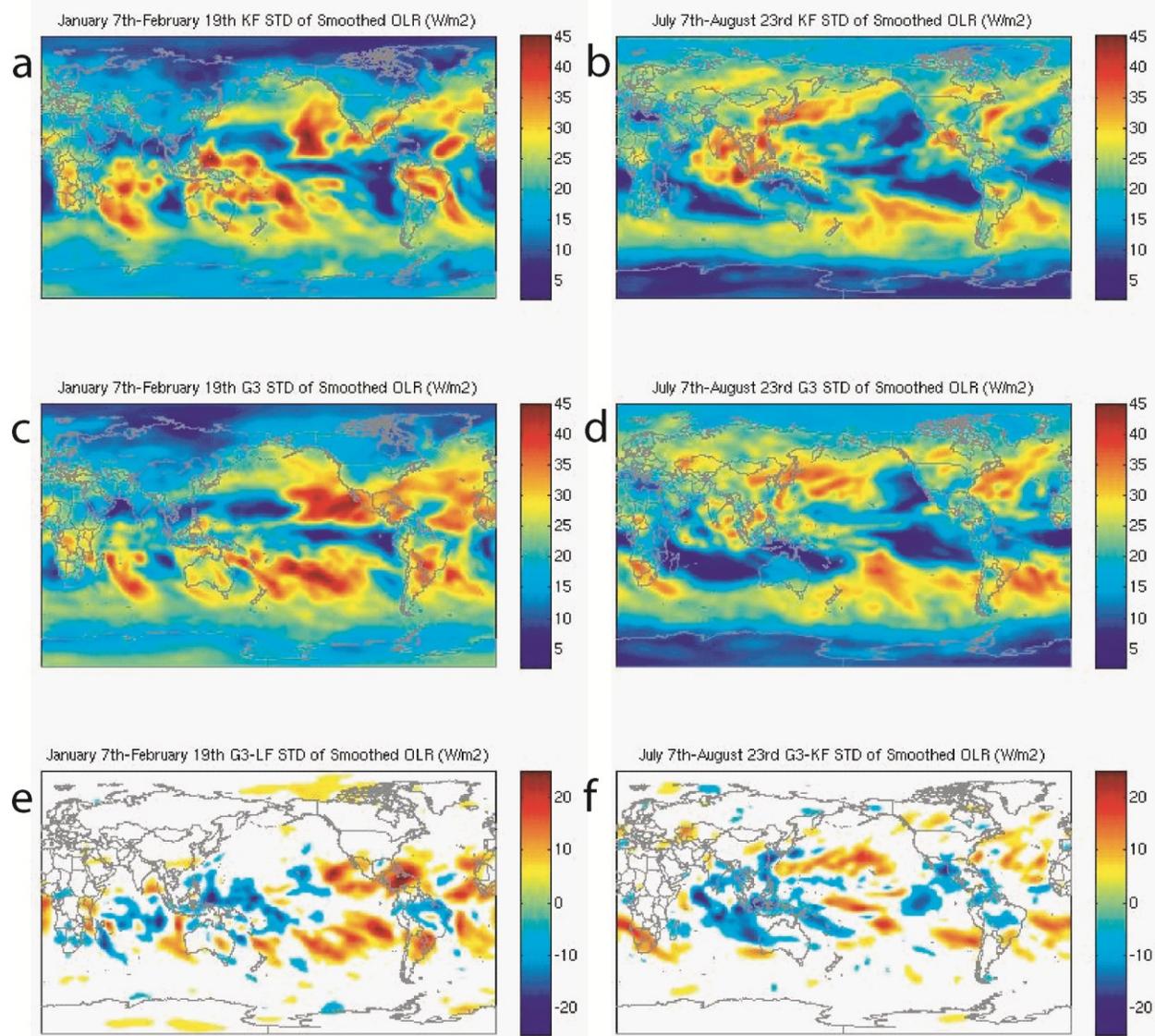
#### *Potential Causes Of Observed Differences*

The effect of the ensembling technique used in the G3 CP can be observed when comparing outgoing longwave radiation (OLR) between the G3 CP and KF CP simulations. **Figure 2.7** shows the observed mean OLR (**Fig. 2.7a,b**), mean smoothed OLR for both CP simulations (**Fig. 2.7c,d,e,f**), and the difference between both simulations for both seasons (**Fig. 2.7g,h**). In both seasons, it can be seen that the value of mean OLR is generally greater for the KF CP than the G3 CP in tropical and subtropical regions with the exception of land locations. Intuitively, smaller values of OLR may seem to indicate that the G3 CP has more intense convection over these regions, because stronger convection generally attains a greater height in the troposphere, and thus its cloud tops emit radiation at a lower effective temperature.

However, as shown by **Figure 2.8**, this is misleading. **Figure 2.8** shows the smoothed standard deviation of OLR for both simulations (**Fig 2.8a,b,c,d**) for both seasons as well as the difference between the G3 and KF CPs (**Fig 2.8e,f**). Examining the standard deviation of OLR is an effective proxy for judging the variability of convective intensity. By this metric, the G3 CP has significantly less variability over a 40° latitude x 120° longitude box centered approximately over the maritime continent for both seasons. Conversely, for most other oceanic locations in the



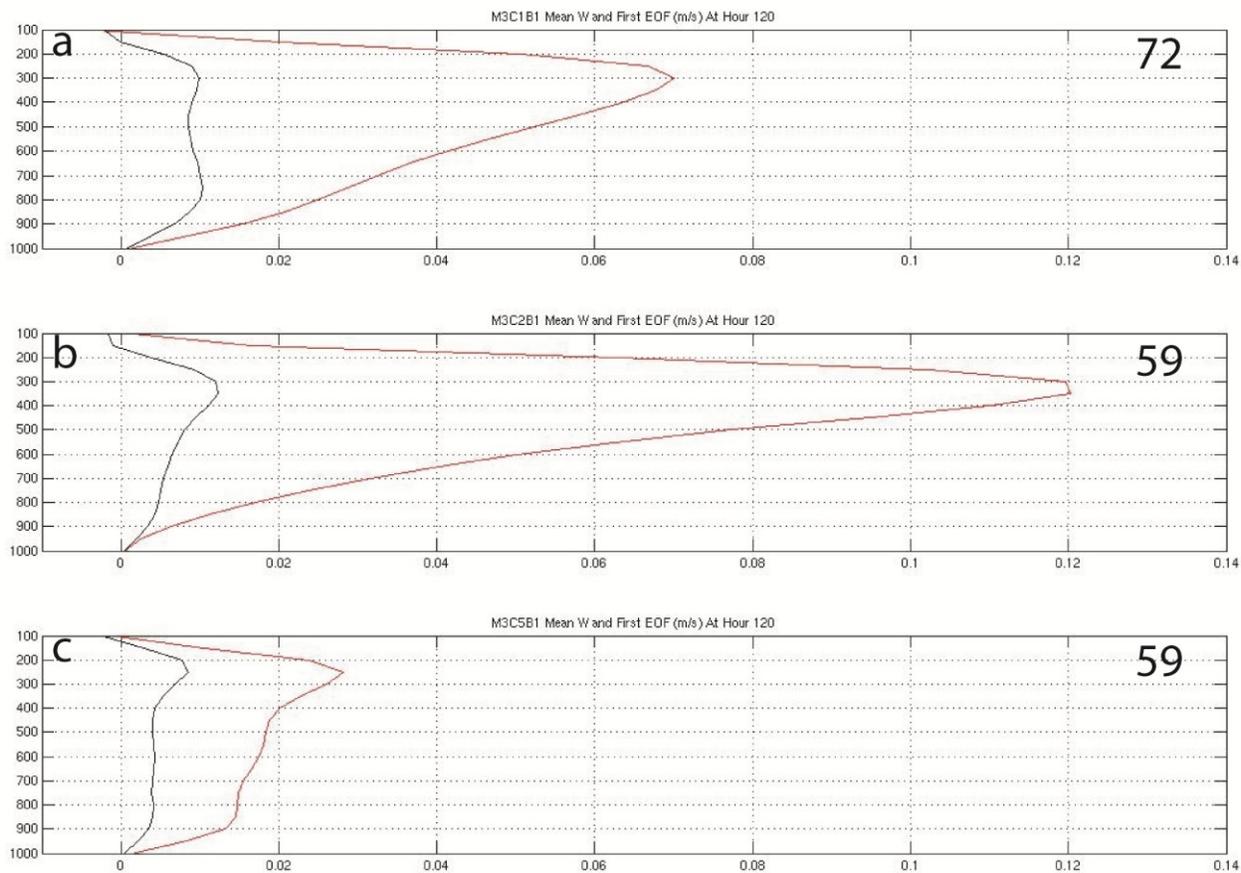
**Figure 2.7** As in Figure 2.2, except for outgoing longwave radiation ( $W/m^2$ ).



**Figure 2.8** As in Figure 2.6, except for the standard deviation of outgoing longwave radiation ( $W/m^2$ ) across the globe.

tropical and subtropical regions, the G3 CP has more variability in convective intensity than the KF CP. Taken together, **Figure 2.7** and **Figure 2.8** indicate that over the maritime continent the G3 produces *more* convection but convection which is *less* variable in its intensity in both seasons. Alternatively, for the majority of the remaining tropical and subtropical oceanic regions, the G3 again produces *more* convection *and* convection which is *more* variable in its intensity in both seasons, which is particularly obvious when comparing the subtropical eastern Pacific Ocean and subtropical Atlantic Ocean for either season, although this effect is most pronounced for the winter simulations. This phenomenon corroborates the differences in precipitation described earlier. For global regions in which the G3 CP produced higher (lower) rain rates than the KF CP, the variability in OLR is generally greater (lower) than that of the KF CP. Likewise, this corroborates the observed decreased variability observed in three hour rainfall seen with the G3 CP in the limited area domain simulations when compared with the KF CP.

An analysis of vertical motion profiles using the limited domain data is performed below using a dataset comprised of all profiles of vertical motion for a given parameterization at hour 120 for all cases over an ITCZ region (defined as all points between  $6.8^{\circ}$  N and  $15^{\circ}$  N). An effective way to do this is to analyze the dominant modes of vertical motion variability associated with precipitation. An empirical orthogonal function (EOF) analysis is performed on these vertical motion profiles. **Figure 2.9** shows, for each parameterization, the first EOF regressed onto the first principal component as well as the mean vertical motion profile. Only the first EOF is shown because (1) the first EOF explains at least 59% of the variance for all three CPs, and (2) it is the only EOF to be strongly correlated with rainfall. (This was determined by correlating the principal components with three hour rainfall ending at the time the vertical motion profile was produced. As a check, this same analysis was performed on vertical motion



**Figure 2.9** The leading EOF regressed onto the first principal component for each CP (red), as well as the mean vertical motion (black) for the (a) KF CP, (b) BMJ CP, (c) and G3 CP. The number shown is the percent of variance the first EOF explains.

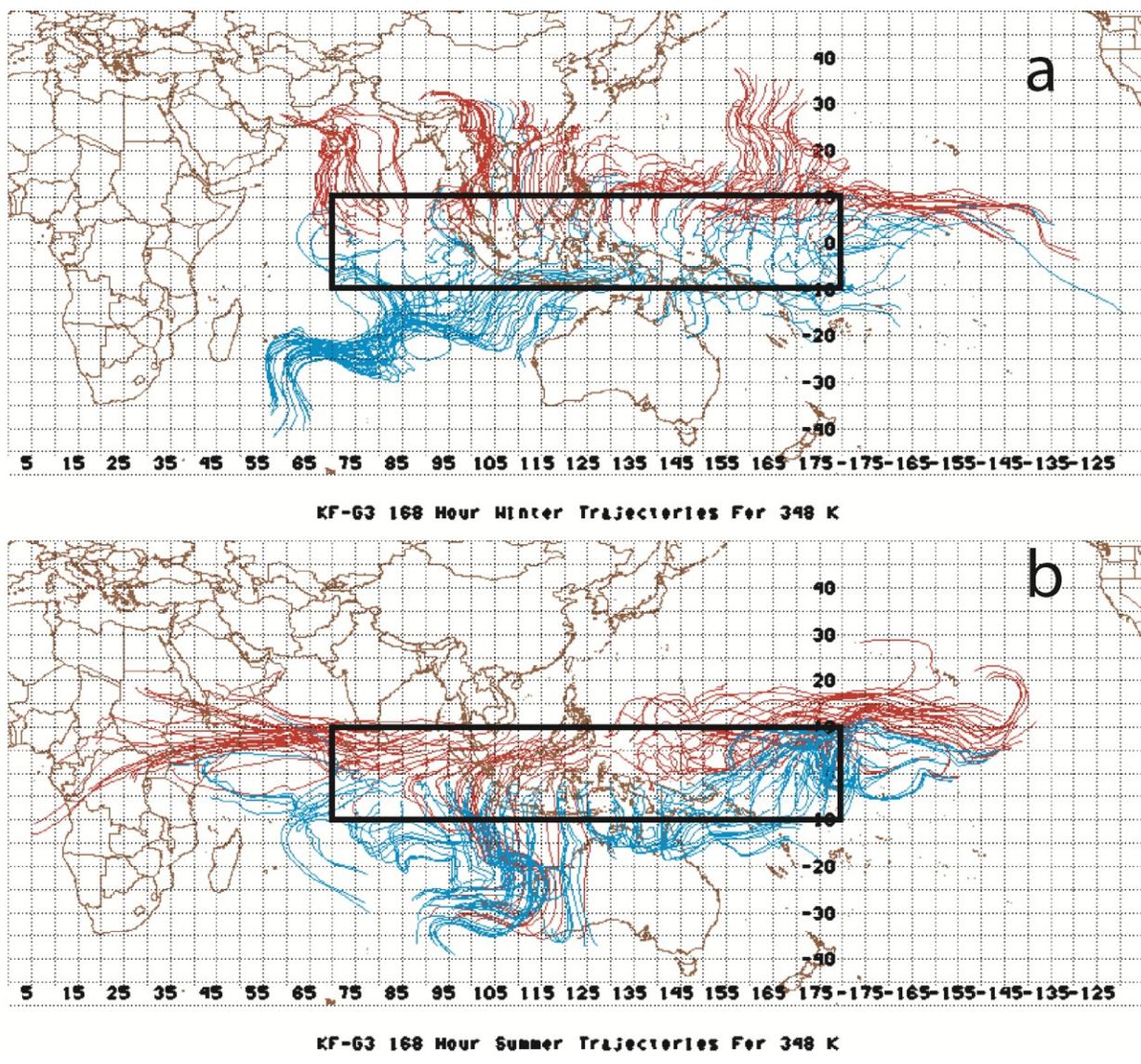
profiles at forecast hour 117 using the following three hour rainfall amounts. This analysis produced similar results). In the case of the KF CP (**Fig. 2.9a**) and BMJ CP (**Fig. 2.9b**), the leading EOF depicts a type of vertical motion profile associated with deep convection, with a maximum in upward vertical motion at middle to upper levels of the atmosphere. Noticeably, the G3 CP (**Fig. 2.9c**) produces a vertical motion profile which is significantly muted relative to the other CPs, which is consistent with weaker convection and smaller precipitation rates.

Several observations are consistent across multiple analyses at this point. First, relative to other CPs, the G3 CP generally produces less intense precipitation in regions where intense precipitation would be expected, and slightly more precipitation in regions where weak precipitation would be expected. Second, the weaker and less variable convection produced with the G3 CP is associated with weaker upward vertical motion, as would be expected. This corroborates other studies such as Mukhopadhyay et al. (2010). This was observed both in the mean global circulation (as can be inferred from mass conservation arguments in **Figure 2.5**) as well as a regional analysis at a constant forecast time (as seen in **Figure 2.9**). Finally, as shown above, these differences in the handling of convection are observed in conjunction with non-trivial differences in the mean global circulation. **Figure 2.5** demonstrated the fact that the G3 CP produced a dramatically weaker Hadley Cell circulation (about 33% weaker when comparing ITCZ low-level inflow or upper level outflow, particularly for the winter hemisphere) while **Figure 2.8** showed that the G3 CP produced much less variability in convection than the KF CP over a region centered on the maritime continent.

In order to relate the differences in convection to the differences in the mean global circulation, an isentropic trajectory is performed over an area approximately centered on the maritime continent. The 348 K surface is chosen, as that is a level which roughly corresponds to

an upper-tropospheric level associated with convection outflow (and one which is firmly centered in the largest ITCZ outflow differences seen in **Figure 2.5**). **Figure 2.10** shows forward trajectories for both seasons over the first 168 hours for the global simulations beginning at locations within the highlighted box. However, rather than standard model trajectories, the trajectories shown are created using the *difference* in observed wind between the KF CP and the G3 CP, and therefore are not actually an “observable” wind. This is done to more easily highlight the difference between these two schemes as it relates to upper level outflow from convection. The first 168 hours of the simulation are chosen because (1) the two CPs begin with identical initial conditions, and it is therefore a clean comparison, and (2) generally the simulations are otherwise similar across the global extratropics, so changes observed over this time can be easily related to local differences, rather than imposed circulation differences. Red (blue) trajectories indicate initial points north (south) of the equator.

An analysis of **Figure 2.10** indicates that nearly all trajectories diverge from their starting points, with most trajectories moving poleward, or in some cases spreading out along the equator. As these trajectories are created using the difference of two winds, this indicates that in virtually all locations the KF CP winds are more strongly divergent from their point of origin than the G3 CP winds. Put another way, the zonal and meridional components of the observed KF CP winds are greater in magnitude than the G3 CP winds on this level. This can be seen by virtue of the fact that the red trajectories generally move toward the north pole while the blue trajectories generally move toward the south pole, especially notable for the winter simulation (**Fig. 2.10a**). While this is approximately true of the summer simulation as well, a significant zonal movement is seen, with the trajectories originating at points at the eastern (western) edge of the box generally continuing further east (west).



**Figure 2.10** Forward trajectories created by using the difference of the wind between the KF and G3 CP (KF-G3) on the 348 K level, where red (blue) depicts trajectories initiating north (south) of the equator. (a) Shows trajectories for the period 1 January – 7 January and (b) 1 July – 7 July.

Cumulatively, the above analyses suggest a relationship between the manner in which a CP produces tropical rainfall and the resultant global circulation produced by that CP. As shown above, the G3 CP produces less convection variability (particularly over the maritime continent) than the KF CP, as seen in an examination of the differences in the standard deviation of either three hour rainfall or OLR in this region (or an examination of vertical motion profiles ). More specifically, the G3 CP is less able to produce intense convection spanning the depth of the troposphere. It is also observed that the G3 CP produces a weaker Hadley Cell when compared with the KF CP when examining upward vertical motion, low-level inflow, or upper-level outflow. A trajectory analysis indicates that upper tropospheric convective outflow generally expands to cover a greater area with the KF CP than with the G3 CP, and it is also seen that the G3 CP generally produces mean zonal jets which are generally too close to the equator. Collectively, these observations suggest that weaker convection, as produced by the G3 CP, leads to less intense upper tropospheric outflow, which causes an equatorward displacement in the mean zonal jet as the largest meridional potential vorticity gradient occurs nearer to the equator (not shown).

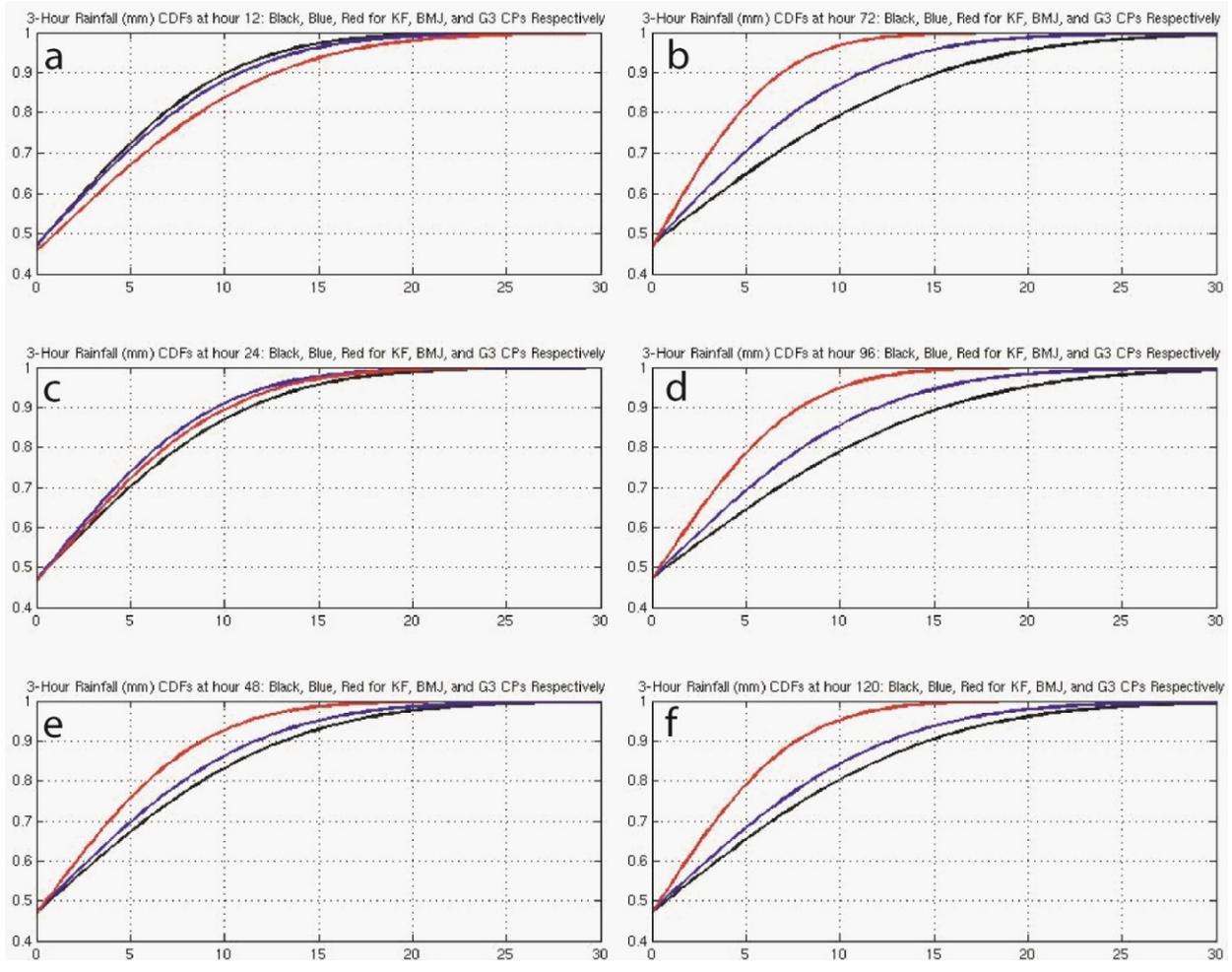
## DISCUSSION

### *Examination Of Parameterization Characteristics*

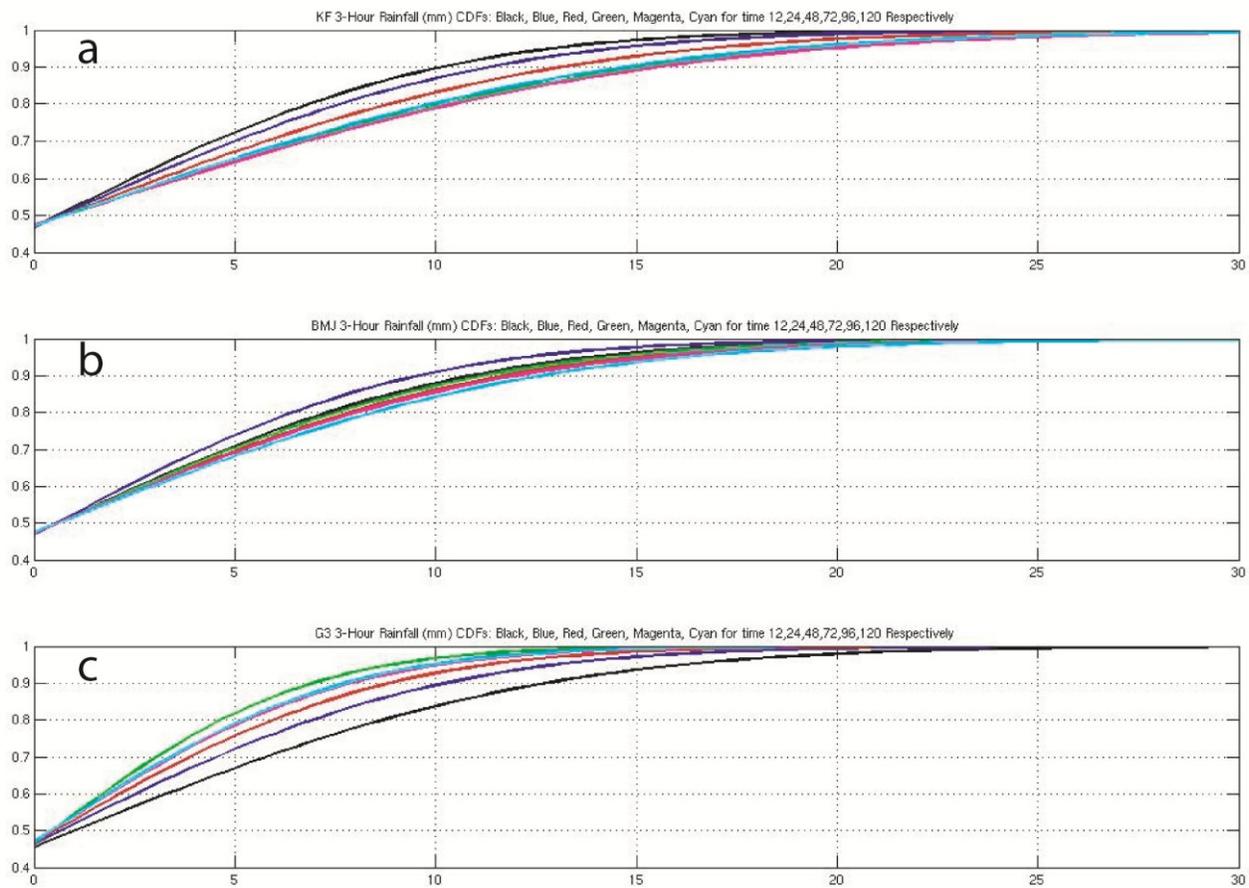
Thus far, the focus of this paper has been mostly on the differences in rainfall distributions and atmospheric variables between the G3 CP and other CPs in medium to long term forecasts (5+ days). However, given that it has been well established that (1) the G3 CP does seem to possess certain deficiencies in its handling of tropical rainfall and (2) these deficiencies appear to affect the large-scale circulation, it is appropriate to inspect this parameterization in greater detail.

Using the limited domain simulations, a comparison of three hour rainfall cumulative distribution functions (CDFs) for the KF CP, BMJ CP, and G3 CP at different forecast times is performed. This is accomplished by accumulating all grid points for all cases at a given time before determining the CDFs. **Figure 2.11** shows CDFs for these parameterizations at forecast hours 12, 24, 48, 72, 96, and 120. At forecast hour 12, an analysis of the three CDFs indicates that the G3 CP produces slightly more light precipitation *and* more heavy precipitation than the other CPs. This result is different from prior rainfall analyses conducted above, which indicated that the G3 generally produced less intense rainfall than either the KF CP or BMJ CP. However, at forecast hour 24 the G3 CP produces an equivalent amount of light precipitation as the other CPs and less heavy precipitation than the KF CP. By forecast hour 48, the G3 CP produces less heavy precipitation than both the KF CP and BMJ CP. At forecast hour 72, this disparity has increased, and the disparity remains relatively constant between forecast hours 72 and 120. These CDFs indicate that the evolution of rainfall with forecast time is different for the G3 CP compared with other CPs.

This can be seen more easily in **Figure 2.12**, which shows rainfall CDFs through forecast times binned according to CP. **Figure 2.12** makes it clear that as forecast time increases, the ability of the G3 CP to generate heavy precipitation decreases, eventually becoming roughly constant after forecast hour 72. To a lesser extent, both the KF CP and BMJ CP also change their rainfall CDF evolution, although in their case it is to increase the likelihood of heavy precipitation, with a steadiness of rainfall CDFs at forecast hour 48 and 72 for the BMJ CP and KF CP respectively. This may be related to a greater propensity for the KF and BMJ CPs to produce tropical cyclones in their forecasts relative to the G3 CP (not shown, but discussed in Chapter 3). The changing rainfall CDFs track very well with the change in mid-level relative



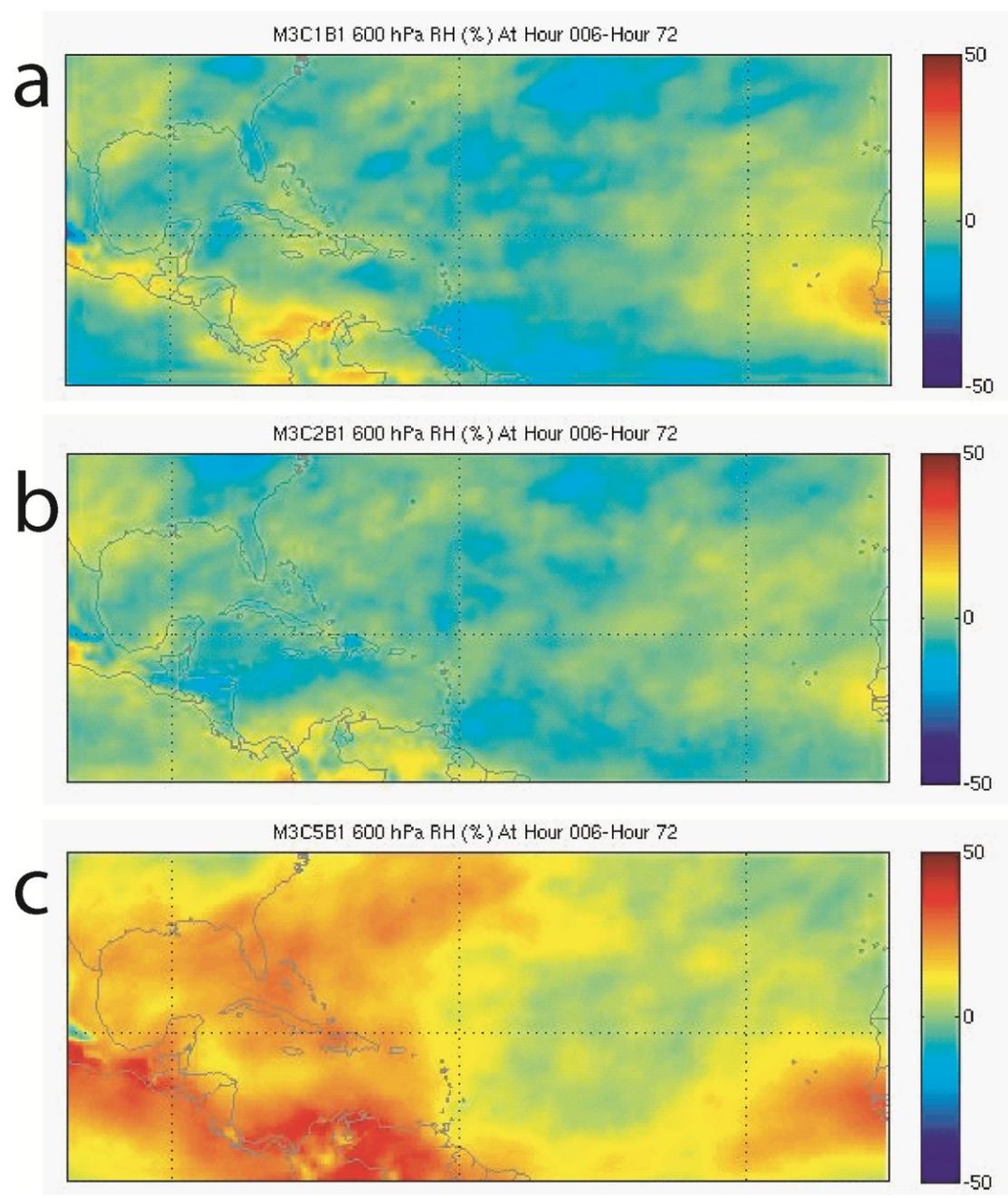
**Figure 2.11** Cumulative Distribution Functions are shown for the KF CP (black), BMJ CP (blue), and G3 CP (red). Forecast hour 12, 24, 48, 72, 96, and 120 are shown in panels (a), (b), (c), (d), (e), and (f) respectively.



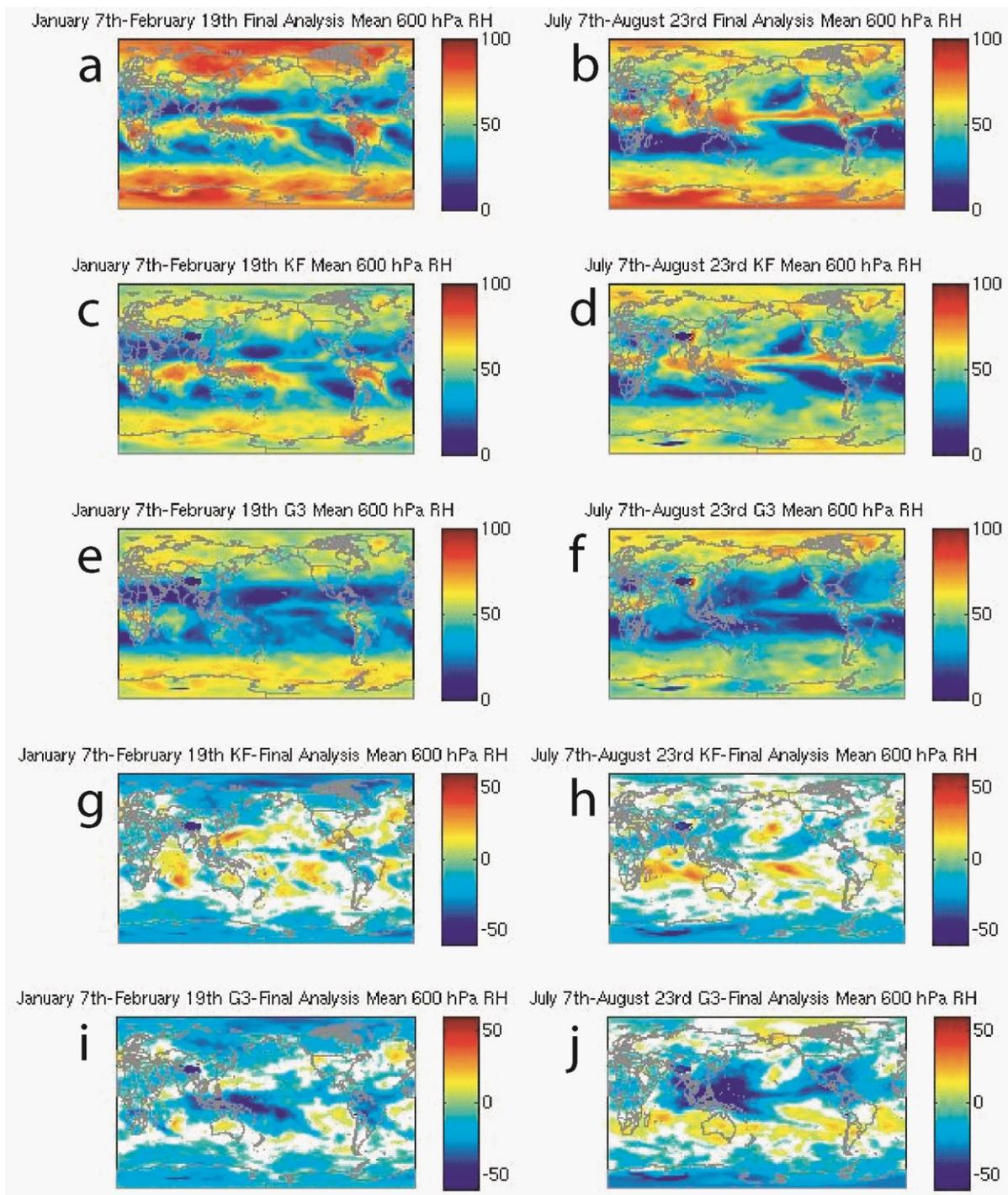
**Figure 2.12** Cumulative Distribution Functions are shown for the KF CP (a), BMJ CP (b), and G3 CP (c). Forecast hour 12, 24, 48, 72, 96, and 120 are shown in black, blue, red, green, magenta, and cyan respectively.

humidity, as shown in **Figure 2.13**, which shows the mean difference in relative humidity for these three CPs between forecast hour 6 and forecast hour 72 at 600 hPa. The KF CP (**Fig. 2.13a**) and BMJ CP (**Fig. 2.13b**) are very similar and show little relative humidity change over this timeframe. However, the G3 CP (**Fig. 2.13c**) produces increasing mid-level dryness as indicated by relative humidity decreases by as much as 50%. Furthermore, the areas with the greatest decrease in relative humidity correspond to the locations which produce the most convectively generated precipitation (not shown) which further demonstrates that the CP is the reason for these differences, as excessive atmospheric drying is often a symptom of an overactive CP.

Relative humidity differences are observable with the global simulations as well as shown by **Figure 2.14**, which shows observed mean 600 hPa relative humidity (**Fig. 2.14a,b**), the mean 600 hPa relative humidity for the KF CP (**Fig. 2.14c,d**) and G3 CP (**Fig. 2.14e,f**), and the difference between simulated and observed for both seasons (**Fig. 2.14g,h,i,j**). The KF CP exhibits relatively small differences and no obvious systemic errors, with most values close to observed for either season. However, the G3 CP exhibits significant differences in tropical regions for both seasons, with the G3 CP underestimating relative humidity by as much as 60% for some areas over the western Pacific Ocean and maritime continent. A comparison of the pattern of these relative humidity differences to rain rates as shown in **Figure 2.2** indicate a significant correlation between the two for the G3 CP, which was also observed with the limited area simulations above. This indicates that the G3 CP is overactive in these regions – when the CP is active, compensating drying is produced at adjacent grid points. This drying is enhanced by the nature of the two CP schemes – the G3 CP assumes detrainment occurs at cloud top, while the KF CP allows detrainment to occur throughout the depth of the cloud. **Figure 2.3** indicates



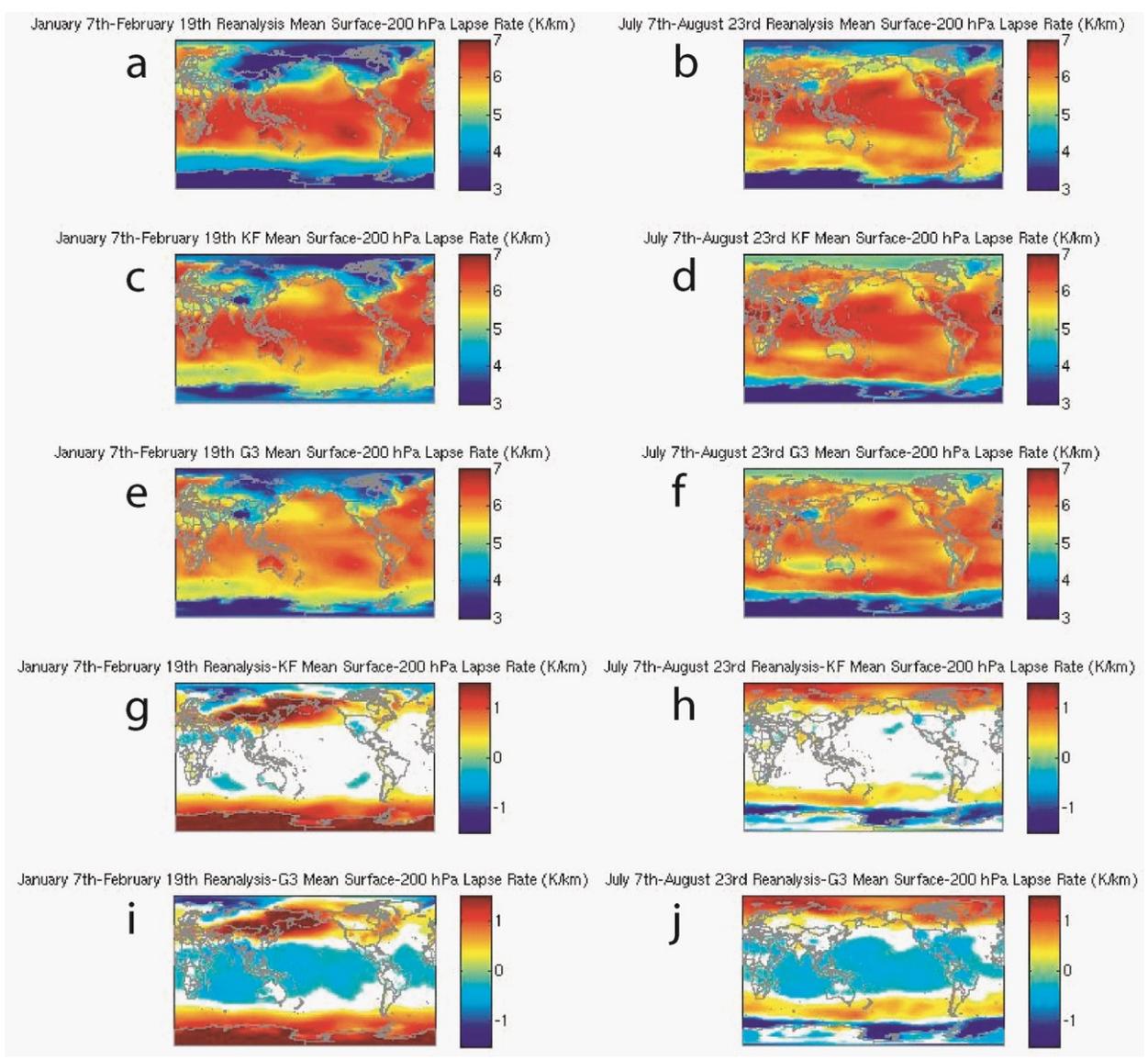
**Figure 2.13** The mean relative humidity (%) difference between forecast hour 6 and forecast hour 72 is shown for the KF CP (a), BMJ CP (b) and G3 CP (c).



**Figure 2.14** As in Figure 2.4, except for 600 hPa relative humidity (%).

that the latitudinal bands which encompass the most significant mid-level relative humidity differences in the G3 CP global simulations are also the same areas in which the parameterization produces the majority of the rainfall. Areas in which the parameterization produces little rainfall are generally not characterized by excessive drying, which further indicates that the parameterization is the cause of this extreme drying in tropical locations. Given that the G3 CP produces all cloud detrainment at cloud top, one would perhaps expect anomalously large values of relative humidity at the cloud top height over the same areas which exhibited anomalously low relative humidity at mid-levels. This is seen when comparing mean relative humidity values between 200-300 hPa (not shown), which again indicates that the nature of the G3 CP is the cause of these differences.

An additional consequence of the fact that the G3 CP produces all cloud detrainment at cloud top is that cloud heating and moistening is maximized at the cloud top, which means that locations which produce significant amounts of convective activity (as outlined easily by the mid-level relative humidity differences seen in the G3 CP in **Figure 2.14**) also produce significant warming at cloud top height relative to the KF CP or reality. **Figure 2.15** gives an indication of this phenomenon by showing the mean surface to 200 hPa lapse rate for reality (**Fig. 2.15a,b**), the KF CP (**Fig. 2.15c,d**) and G3 CP (**Fig. 2.15e,f**), and the difference between simulated and observed for both seasons (**Fig. 2.15g,h,i,j**). Both CPs greatly over-forecast mean lapse rates over northeastern Asia in the winter simulation due the surface parameterization scheme, which keeps this location too warm in the mean. Over tropical and subtropical regions the KF CP is very similar to observed for either season – exhibiting a mean lapse rate that is approximately the same as reality. However, over these same regions, the G3 CP produces a lapse rate that is approximately .5 K/km too small. This is a function of upper tropospheric



**Figure 2.15** As in Figure 2.4, except for mean lapse rate between the surface and 200 hPa (K/km).

warming in these locations due to the aforementioned nature of the G3 CP cloud detrainment rather than any significant differences regarding surface temperatures.

### *Grell-3 CP Sensitivity Tests*

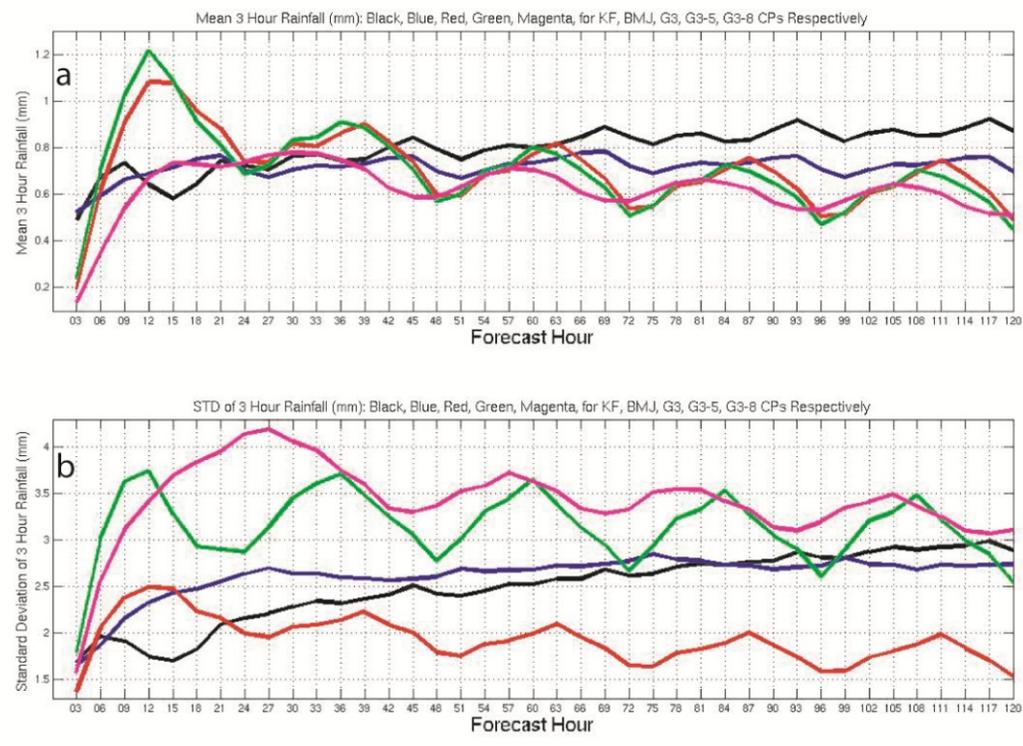
As described earlier, the G3 CP employs a sophisticated ensembling technique which seeks to make use of multiple closure assumptions as well as other constant variations to create a variety of members, which are then averaged to provide the model with one outcome. However, this appears to be problematic for certain weather regimes. Consider this simplification – if half of the ensemble members predict a thunderstorm, and half predict no thunderstorm, what is the result? The answer is something approximating half of a thunderstorm. However, there are a number of weather regimes in which this result may be unsatisfactory, in particular regions which are characterized by a moderately unstable atmosphere with little external synoptic scale forcing for ascent. Much of the tropical and subtropical oceans exhibit a significant amount of low-level moisture, moderate amounts of conditional instability, and weak ascent (due to the Hadley circulation), although not necessarily all occurring in significant amounts simultaneously.

Another way to state this is that low-level moisture, conditional instability, and ascent are directly or indirectly analogs to three of the closure assumptions used in the G3 scheme – moisture convergence, CAPE relaxation, and low-level omega thresholds. Given sufficient quantities of moisture convergence, CAPE, or upward vertical motion, at least several of the G3 ensemble members will produce convection. However, in order to produce convection that is instantaneously of the magnitude of other single closure schemes (such as the KF CP or BMJ CP), a significant fraction of the G3 members must produce convection, which will not happen if all of the above conditions are not sufficiently met. This appears to be the phenomenon described above for ITCZ locations and some other similar tropical locations. It should be noted that in

certain rare instances, such as the development of an intense tropical cyclone in either the global or limited area G3 CP simulations, intense rain rates and deep convection comparable to other CPs are produced. However, in these cases the CP contributes very little to the rainfall produced due to the strong dynamical forcing for upward vertical motion.

In order to test the hypothesis that the ensemble technique used in the G3 CP is producing the unusual rainfall intensity distributions, the limited domain simulations were rerun with two modified versions of the G3 CP – one employing *only* moisture convergence closures (G3-5 CP henceforth) and one employing *only* omega closures (G3-8 CP henceforth). For more information on these closures, refer to Grell and Dévényi (2002). This greatly reduces the number of ensemble members for each modified CP relative to the original G3 CP and creates a much less diverse ensemble within the CP.

**Figure 2.16a** shows mean three hour precipitation (calculated by totaling precipitation across all grid points and all cases and then dividing by the total number of grid points and cases) over the forecast period, as well as the standard deviation of mean three hour precipitation for the KF CP, BMJ CP, and all variations of the G3 CP (**Fig. 2.16b**). Forecast hour 00 (and therefore 24, 48, 72, 96, and 120) represents 0000 UTC. **Figure 2.16** shows that after an initial adjustment period (on the order of 24-36 hours) the KF CP and BMJ CP are fairly steady through time when comparing either parameter, although the KF CP does show a slight increase with time of the standard deviation of three hour precipitation with time. This appears to be due to an increase in the number of intense tropical cyclones developed by this CP compared with other CPs. All three G3 CPs exhibit a marked diurnal cycle in mean three hour precipitation and to a lesser extent the standard deviation of three hour precipitation – with a peak near 1200 UTC and a minimum near 0000 UTC, although the exact maximum and minimum may fall at a slightly different time. It is

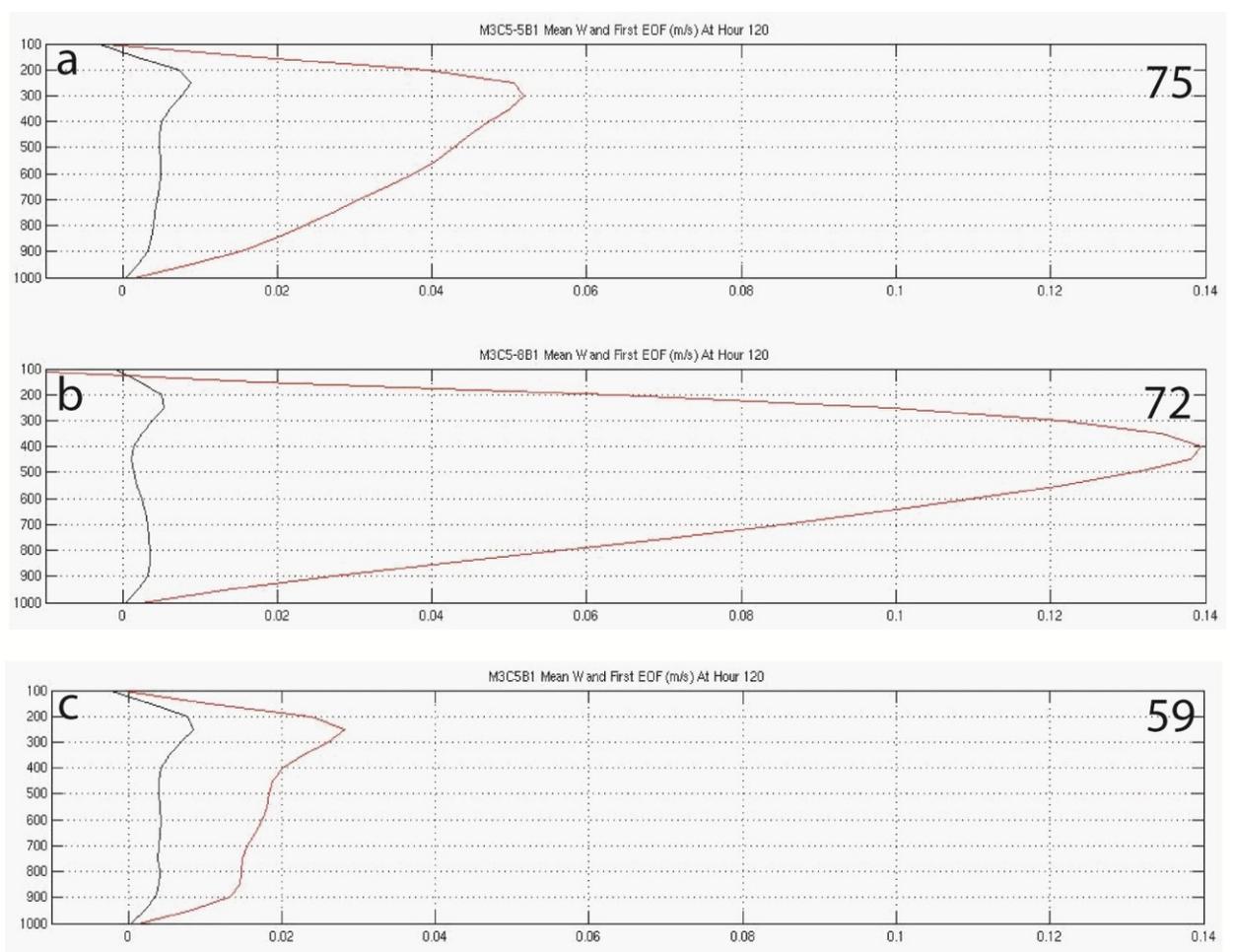


**Figure 2.16** (a) mean and (b) standard deviation 3 hour rainfall properties with forecast time for the KF CP (black), BMJ CP (blue), G3 CP (red), omega G3 CP (green), and moisture convergence G3 CP (magenta).

extremely apparent that the overall trend with the three G3 CPs is for lower mean precipitation with time. This is consistent with **figures 2.11** and **2.12**.

However, the three G3 CPs do not behave similarly when examining the standard deviation of three hour precipitation. While the standard G3 CP produces less variability, with a decreasing trend with forecast time, the G3-5 CP and G3-8 CP produce significantly more variable precipitation and end with similar values as either the KF CP or the BMJ CP. This seems to confirm the hypothesis that it is the ensembling technique used in the G3 CP that causes a decrease in the variability of rainfall intensity. However, since all three G3 CPs produce similar patterns regarding mean precipitation, something different must cause this phenomenon. A brief examination of mid-level relative humidity indicates that the G3-5 CP and G3-8 CP produce similar drying as seen with the G3 CP (although not of the same magnitude), so this is a potential cause. **Figure 2.17** depicts the results of an EOF analysis performed on profiles of vertical motion in the same manner as shown in **Figure 2.9**, except using the G3-5 CP (**Fig. 2.17a**) and G3-8 CP (**Fig. 2.17b**), with the original G3 CP (**Fig. 2.17c**) shown for comparison. As would be suggested by the greater variability in rainfall observed in **Figure 2.16**, **Figure 2.17** indicates vertical motion profiles associated with rainfall in the G3-5 CP and G3-8 CP are more characteristic to those observed earlier with the KF CP and BMJ CP. It should be noted that the greater realism in precipitation forecasts when using a modified G3 CP are similar to other studies such as Gianotti et al. (2012) which found that using the GD CP with only one closure improved precipitation forecasting over the maritime continent.

As the model sensitivity experiments described above suggest, the ensembling nature of the G3 CP causes the CP to become overactive in tropical and subtropical regions. As **Figure 2.16** indicates, as the number of closures (and therefore ensemble members are decreased) the



**Figure 2.17** As in Figure 2.9, except for omega G3 CP (a), and moisture convergence G3 CP (b) rather than KF CP and BMJ CP. (c) is identical to (c) in Figure 2.9.

G3 CP produces more intense convection (as seen by the standard deviation of three hour precipitation), but the same amount of total precipitation as the conventional G3 CP. The only way both of these statements can be true is if the modified G3 CPs produce fewer total locations with rainfall than the standard G3 CP. This further suggests that the G3 CP is overactive in its conventional form for many tropical and subtropical locations. As described above, when the G3 CP is overactive, substantial mid-level drying and upper-level warming occurs.

Collectively, the G3 CP's handling of the distribution of moisture and heat in the atmosphere appears related to both its difficulty in producing intense convection and its observed decrease in total precipitation with model forecast hour. As the model forecasts integrate forward, an overactive G3 CP in these regions produces compensating drying at mid-levels of the atmosphere, which makes further intense precipitation more difficult as there is less total atmospheric moisture in a given column. Similarly, as the model forecast integrates forward the lapse rate decreases, meaning that the atmosphere becomes more stable and less prone to convection. Additionally, the anomalous heating and drying are out of phase in a way which further suppresses convection. At upper levels, the G3 CP produces anomalously large values of relative humidity (which would be favorable for more intense precipitation), yet these are the same locations which exhibit a substantial decrease in the atmospheric lapse rate, which acts to stabilize this level. Conversely, at mid-levels of the atmosphere over regions with an overactive CP, the G3 CP actually possesses a greater lapse rate than is observed in reality (which would be favorable for convection), yet the extreme compensating dryness in these regions depresses the likelihood of convection. In these ways, the G3 CP itself is a break on intense convection with time, as can be seen in the downward trend in both the mean and standard deviation of three hour rainfall shown in **Figure 2.16**.

## CONCLUSION

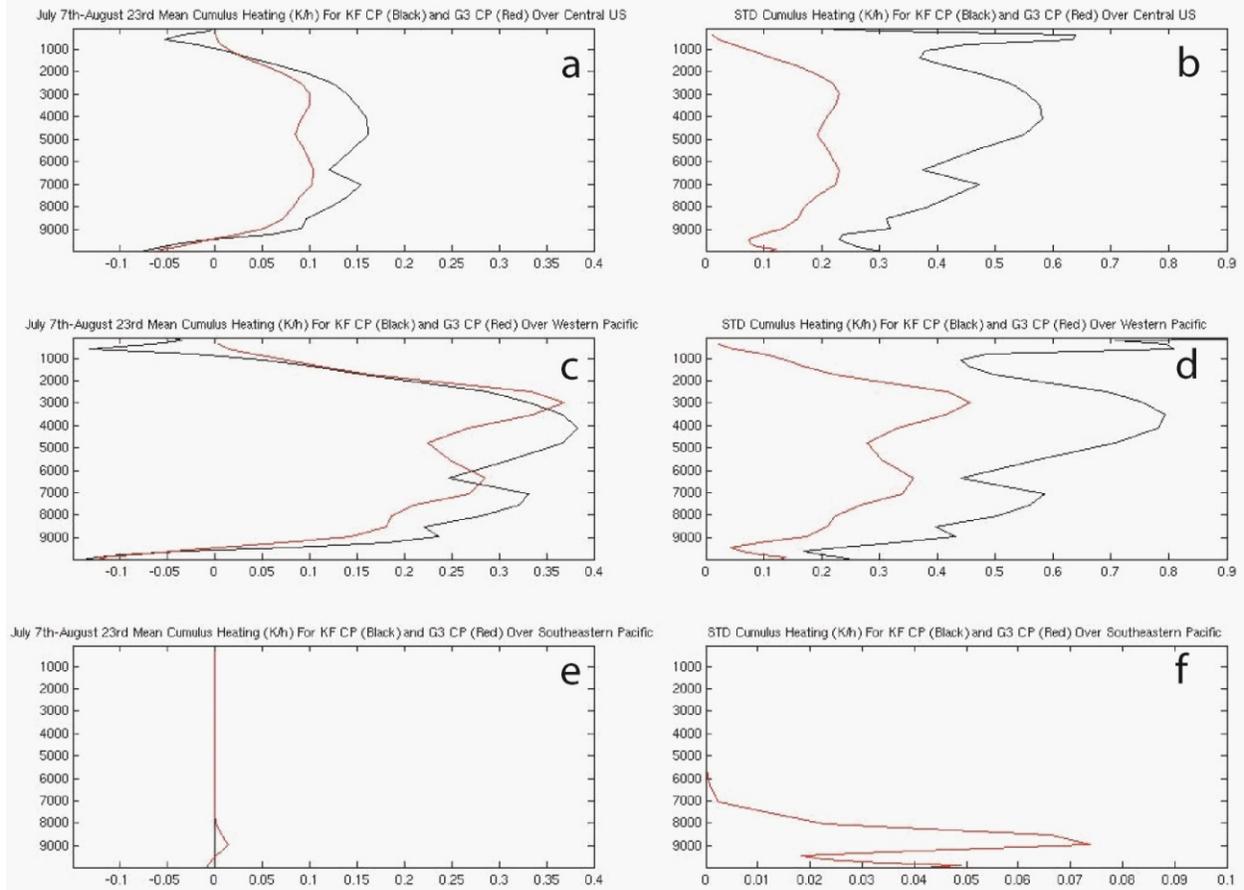
Several caveats apply to this research. Primarily, no high resolution modeling (the highest resolution used in this study is 30 km) was performed in this study. It is possible that the differences observed within this study would be lessened at higher resolutions. However, since it was the goal of this study to analyze medium and long term forecasts, medium resolution forecasts were considered appropriate as the resolutions chosen here are similar to current effective resolutions seen in global models and climate models. Additionally, the simulations Grell and Dévényi (2002) used to test their parameterization used a 27 km horizontal grid spacing, which is comparable to the 30 km grid spacing used in many of the simulations presented here. Second, with the exception of the modified G3 CP simulations described above, in all other details the G3 CP was used as it comes “out of the box” with the WRF-ARW model. In addition to closure choice, it is also possible that other modifications would change the observed patterns seen here. Changing parameters such as the entrainment rate and drying rate, among others, would almost certainly alter the results discussed here, potentially giving an improvement in forecasts. However, it was the goal of this research to test the impacts of the ensemble nature of the G3 CP, and therefore changing these parameters was not relevant to this study.

While this study describes the GD/G3 CP as being inadequate in tropical locations (particularly for longer forecasts), the existing framework of the GD/G3 CP actually lends itself to significant improvement (potentially) without a considerable investment. As originally presented, this CP was given the capability for different ensemble members to be trained using a Bayesian framework in such a way as to improve the resultant forecast of precipitation and associated redistribution of heat and moisture in the atmosphere. While no study is known to the

author attempting to accomplish this, it is a worthwhile endeavor which would presumably greatly improve forecasts using the GD/G3 CP (Grell, personal communication, 2013).

This paper also raises potential caveats with other parameterization research. When comparing all three CPs studied herein, relative CP characteristics changed throughout time. For example, at forecast hour 6, the G3 CP was more likely to produce intense precipitation than either the KF CP or BMJ CP, while the opposite was true by forecast hour 48. A state of effective steadiness was not reached for time scales as long as 72 hours. However, the vast majority of current parameterization-based studies incorporate a forecasting timeframe of 24-48 hours. This suggests that results demonstrated in studies within this timeframe are not necessarily valid for all cases as might otherwise be implicitly inferred. For instance, Grell and Dévényi (2002) used a series of 12 hour model simulations to test their parameterization, whereas the results demonstrated here reveals that this is likely an insufficient time period for parameterization testing.

One other potential caveat with current parameterization research is that most of the cases involved take place over land, as did the test simulations for Grell and Dévényi (2002). However, as with model forecast time, conclusions developed from land-based simulations do not necessarily hold true for tropical ocean simulations. In the case of the G3 CP, the ensembling nature of the G3 CP led to an overestimation of light precipitation over tropical and subtropical oceans in a way that would not be observed in short-term studies over the central United States. One final caveat is that when comparing parameterizations, *it is important to examine the variability in those parameterizations rather than simply examining the mean state*. The ability for a parameterization to produce reasonable extremes in atmospheric variables is of obvious importance. **Figure 2.18** shows the mean and standard deviation of cumulus heating for three



**Figure 2.18** Profiles of mean (a,c,e) and standard deviation (b,d,f) cumulus heating for the KF CP (black) and G3 CP (red) for the period 7 July – 23 August 2005 for the central United States (a,b), Western Pacific (c,d), and Southeastern Pacific (e,f).

very different geographic regions for the northern summer global WRF simulation for the KF CP and the G3 CP. Notice that the mean traits for these three regions are remarkably similar. For example, both produce significant heating over the western Pacific Ocean (**Fig. 2.18a**). However, an examination of the standard deviations demonstrates that in strongly convecting regions, such as the central United States (**Fig. 2.18b**) and the western Pacific Ocean (**Fig. 2.18c**) the KF CP is significantly more variable, which is consistent with the greater rain rate variability observed with the KF CP compared with the G3 CP. Conversely, in weakly convecting regions, such as the southeastern Pacific Ocean (**Fig. 2.18e,f**), the G3 CP is more variable. If one were looking only at mean parameterization traits (whether that variable is cumulus heating, precipitation, etc.) little difference would be observed in this study.

Finally, the comparisons of global WRF simulations imply an important relationship between the characteristic of tropical convection and downstream midlatitude effects. The importance of tropical convection has been noted before – for example, Simmons et al. (1983) demonstrated the importance of tropical forcing on downstream development. However, this study suggests the character of that tropical forcing is as important as its location. While a detailed examination of the relationship between tropical convection and extratropical circulation patterns is outside the scope of this study, based on results presented herein, this topic warrants further study with a specific focus on the role cumulus parameterization plays in this interaction.

## Chapter 3: Parameterizations And Tropical Cyclones

### *Introduction*

Research presented in a prior Master's thesis titled *Tropical Cyclone Forecast Track and Intensity Sensitivities To Various Parameterizations Using The WRF-ARW Model* (Bassill, 2009 (B09 hereafter)) examined forecasts of several notable tropical cyclones across three basins in an attempt to determine whether certain cumulus or microphysics parameterizations exhibited systemic biases in tropical cyclone track, intensity, or structure. Although it is well known that changing model parameterizations impacts tropical cyclone forecasts (recent studies include Tao et al. (2011), Bao et al. (2012), Nasrollahi et al. (2012), Osuri et al. (2012), and Xiang (2012)), this work demonstrated that certain parameterization biases were reproducible across many disparate scenarios. However, while serious attempts were made to explain the cause of the observed parameterization behavior, detailed explanations were occasionally deficient. Where possible, this chapter will provide a more rigorous exploration of the parameterization biases identified, while also expanding upon that research where appropriate using additional research and data previously unavailable. Although a complete explanation of B09 will not be provided here, an overview of the significant findings are provided below.

Three cumulus parameterizations (CPs) and three microphysics parameterizations (MPs) representing a range of complexity were examined in B09 using the WRF-ARW version 2.2 (Skamarock et al. 2005). Every possible combination was performed using the Kain-Fritsch (KF) CP, Betts Miller-Janjic (BMJ) CP, and the Grell-Devenyi (GD) CP in conjunction with the Kessler MP, Eta-Ferrier (EF) MP, WSM6 MP, and an updated version of the Eta-Ferrier (nEF) MP after it was found to have an error in its original release (this results in a total of twelve possible combinations). Three notable tropical cyclones were studied which exhibited significant

track and intensity forecasting challenges – Typhoon Cimaron (2006), a western Pacific cyclone which impacted the Philippines, Hurricane Ernesto (2006), which was a North Atlantic cyclone, Hurricane Lenny (1999), which was a North Atlantic cyclone notable for an unusually extreme November intensity and very long-lived eastward motion throughout the duration of its lifetime. Hurricane Ioke (2006), which was a very long-lived powerful hurricane traversing the central and western Pacific, was also studied due to its relative real-time ease of track and intensity prediction. Cumulatively, the above combinations of parameterizations were used in conjunction with these four storms over the course of fourteen different forecasts consisting of different initialization times and initial conditions. Each simulation used equally sized domains with a 30 km horizontal grid spacing and 31 vertical levels. Generally speaking, two broad conclusions were found:

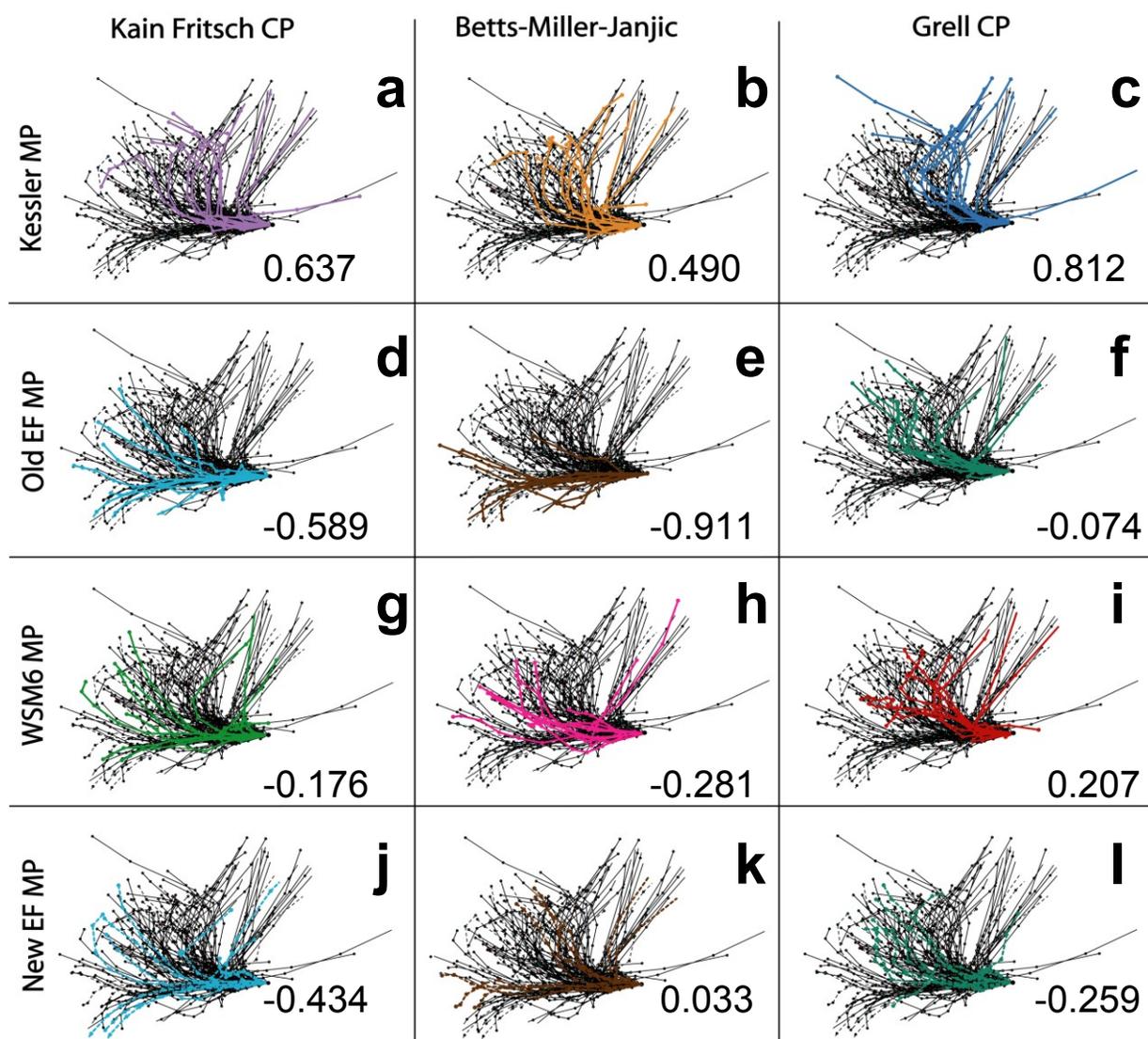
- (1) The GD CP and the Kessler MP exhibited significantly more poleward tracks than the other CPs and MPs respectively, with the split occurring at approximately 48 hours into a given forecast.
- (2) The EF MP, and to a slightly lesser extent the nEF MP, predicted much weaker cyclones than either of the other MPs while also possessing an unusual vertical potential vorticity structure.

Each of these major conclusions as well as related smaller findings will be discussed using further analysis of the data created for B09. Additionally, some pertinent related research will be discussed and where possible additional data will be incorporated. Specifically, forecasts from the “real-time” ensemble simulations performed during the 2009 North Atlantic hurricane season discussed in Chapter 2 will be used.

### *Poleward Track Differences*

**Figure 3.1** reproduces a figure from B09 which shows a compilation of all westward moving cyclones (therefore excluding Hurricane Lenny simulations) normalized to an initial westward movement and binned according to parameterization combination. The number denoted in each panel represents a value of mean normalized 120 hour meridional position, where a value of 1 (-1) would mean that particular combination was always the most poleward (equatorward) cyclone among the forecasts for that particular set of initial conditions. As demonstrated both visually according to the color-coding as well as numerically, it is readily apparent that the Kessler MP (**Fig. 3.1a-c**) produces storm tracks with a greater poleward bias than the other MPs.

While not explained in B09, subsequent research by Fovell, Corbosiero, and Kuo (2009, F09 hereafter) demonstrated a likely cause for this behavior. F09 used idealized simulations to examine the impact of varying MPs using the WRF-ARW model, which depicted an idealized simulation of Hurricane Rita (2005). Rita had significantly greater poleward movement when forecast using the Kessler MP compared with any other MP (representing more than twice the latitudinal gain of any other MP). A detailed examination of the various simulated storms showed that the Kessler MP cyclones were significantly larger (although not necessarily more intense) when examined by virtually any metric. Further analysis led F09 to ascribe the significant poleward movement of the Kessler MP cyclone to a result shown in Fiorino and Elsberry (1989), which demonstrated that TC movement was impacted more strongly by changing Beta gyres related to the TC wind field far from the center (300-1000 km) compared with inner storm differences. A series of modifications to the Kessler MP allowed F09 to

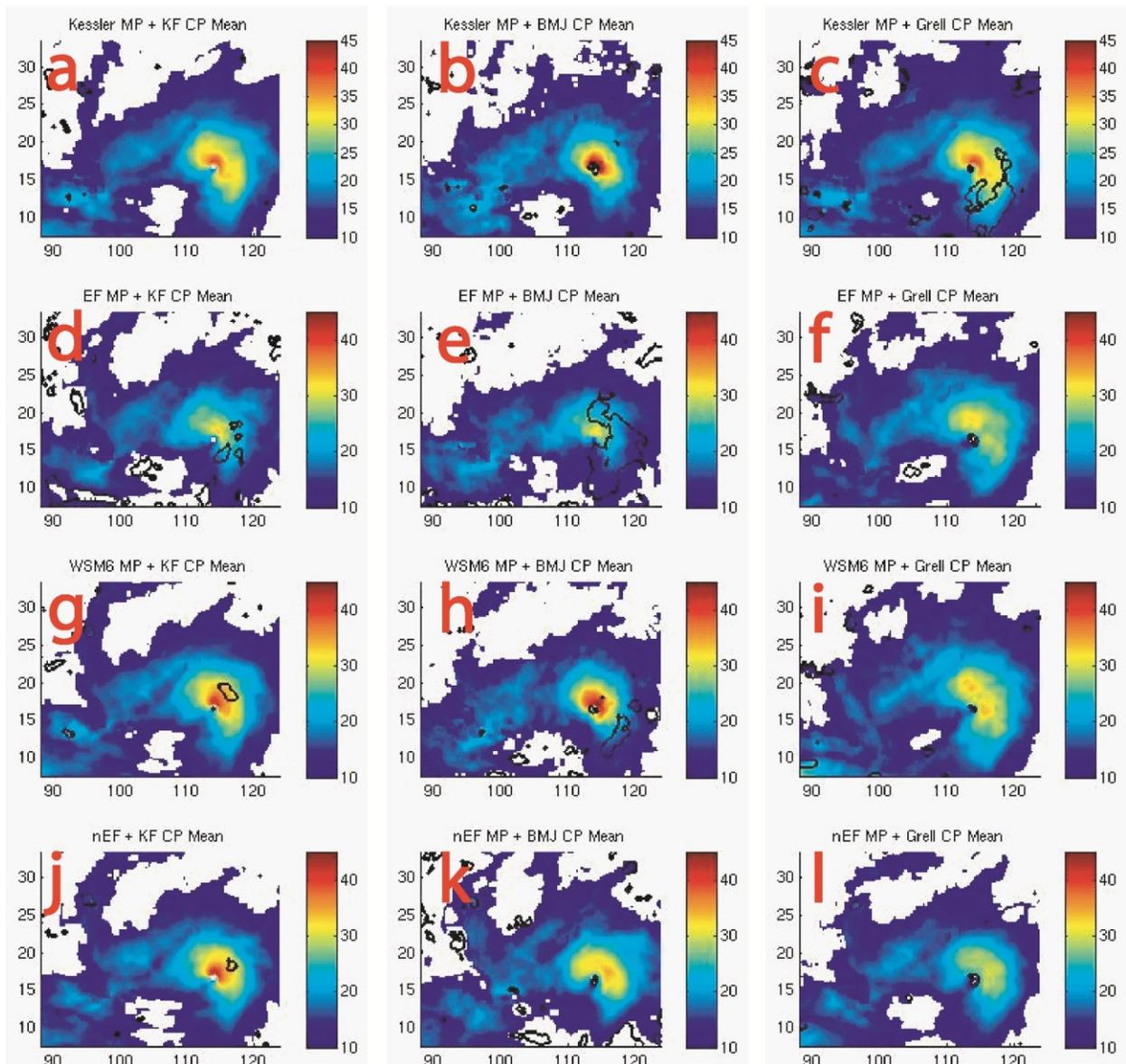


**Figure 3.1** Overlaid storm tracks for cyclones using the (a,b,c) Kessler MP, the (d,e,f) EF MP, the WSM6 MP (g,h,i), and the nEF MP (j,k,l). (a,d,g,j) use the KF CP, (b,e,h,k) use the BMJ CP, and (c,f,i,l) use the GD CP. The number value corresponds to each parameterization combination's relative poleward displacement at forecast hour 120, where -1 (1) corresponds to a southerly (northerly) position.

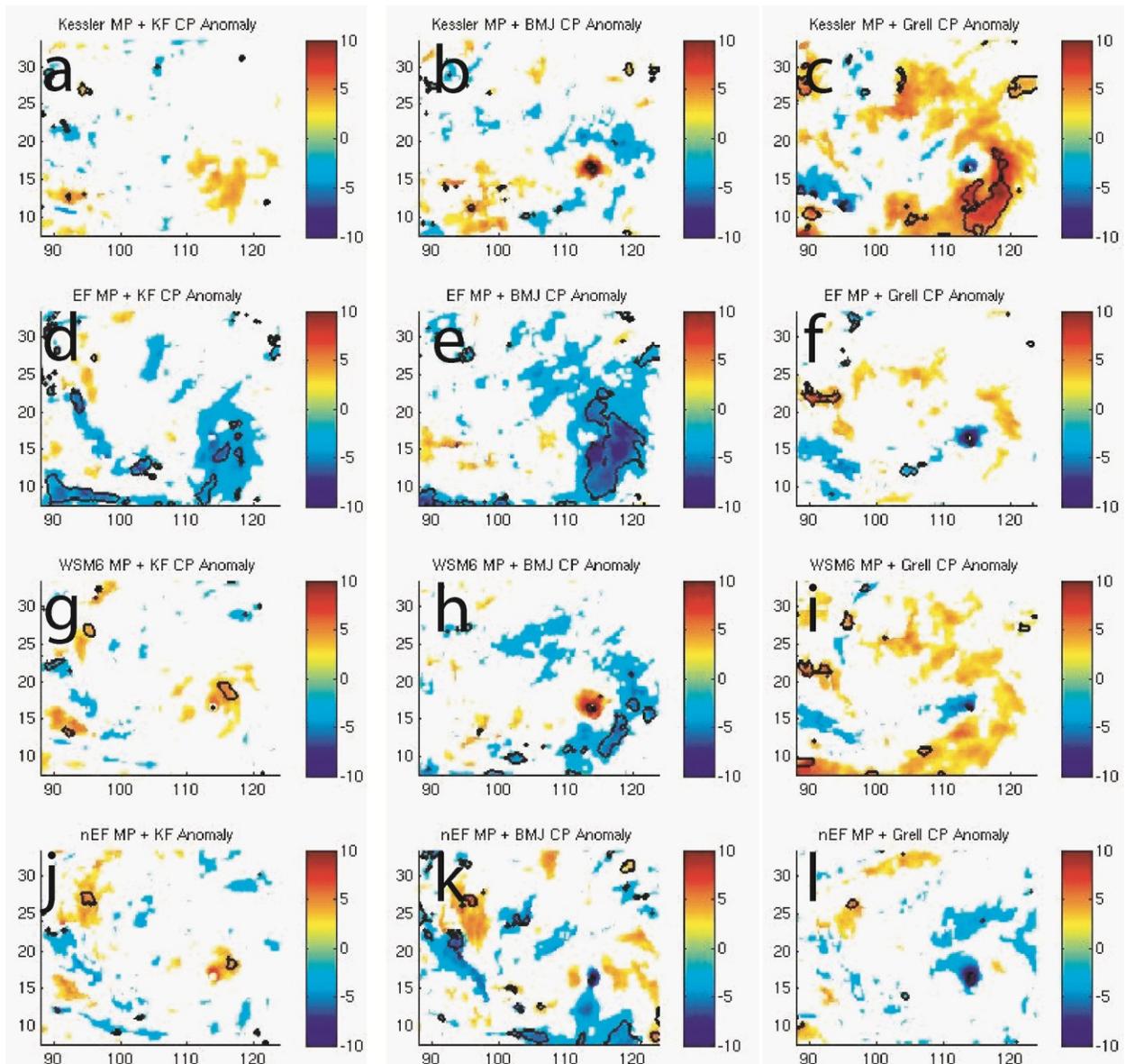
conclusively demonstrate that its large storm size was responsible for an increase in poleward movement.

Although it would be redundant to entirely reproduce F09's analysis here, it should be noted that their results are entirely consistent with those observed in B09. However, a brief analysis of the mean TC circulation is informative. **Figure 3.2** shows mean 850 wind speed (kt) composited at forecast hour 48 for all ten westward moving cyclones (centered in the lower right of the plot). The top row uses the Kessler MP (**Fig. 3.2a-c**), the second row uses the EF MP (**Fig. 3.2d-f**), the third row uses the WSM6 MP (**Fig. 3.2g-i**), and the bottom row uses the nEF MP (**Fig. 3.2j-l**). Solid contours indicate statistical significance using a 95% threshold. Hour 48 is chosen because significant track splitting has not yet happened, although storm-scale changes have had time to occur. While not the focus of this particular section of the chapter, clear intensity differences can be seen amongst the various parameterization combinations when examining mean wind speed near the centers of circulation. **Figure 3.3** shows the anomalies of 850 hPa wind speed for each combination relative to a mean of all combinations. When comparing the Kessler MP (**Fig. 3.3a-c**) with any other MPs within a given CP, it is fairly clear that the Kessler MP has greater winds a significant distance from the center, as shown when used in conjunction with the KF CP (**Fig. 3.3a,d,g**) or GD CP (**Fig. 3.3c,f,i**). In the case of the BMJ CP (**Fig. 3.3b,e,h**), the Kessler MP has wind anomalies which are less weak than the other MPs.

Comparisons of other variables to F09 (not shown) such as cross-sections of relative humidity are likewise similar to those presented in F09. Ultimately, the authors traced the larger storm size when the Kessler MP was used to a slow fall speed for hydrometeors, which increased the areal coverage of cloud cover (particularly in the anvil region). This then created a greater deep-layer mean virtual temperature at large radii from the storm center, which acted to extend



**Figure 3.2** Mean 48 hour 850 hPa wind (kt) for cyclones using the (a,b,c) Kessler MP, the (d,e,f) EF MP, the WSM6 MP (g,h,i), and the nEF MP (j,k,l). (a,d,g,j) use the KF CP, (b,e,h,k) use the BMJ CP, and (c,f,i,l) use the GD CP.

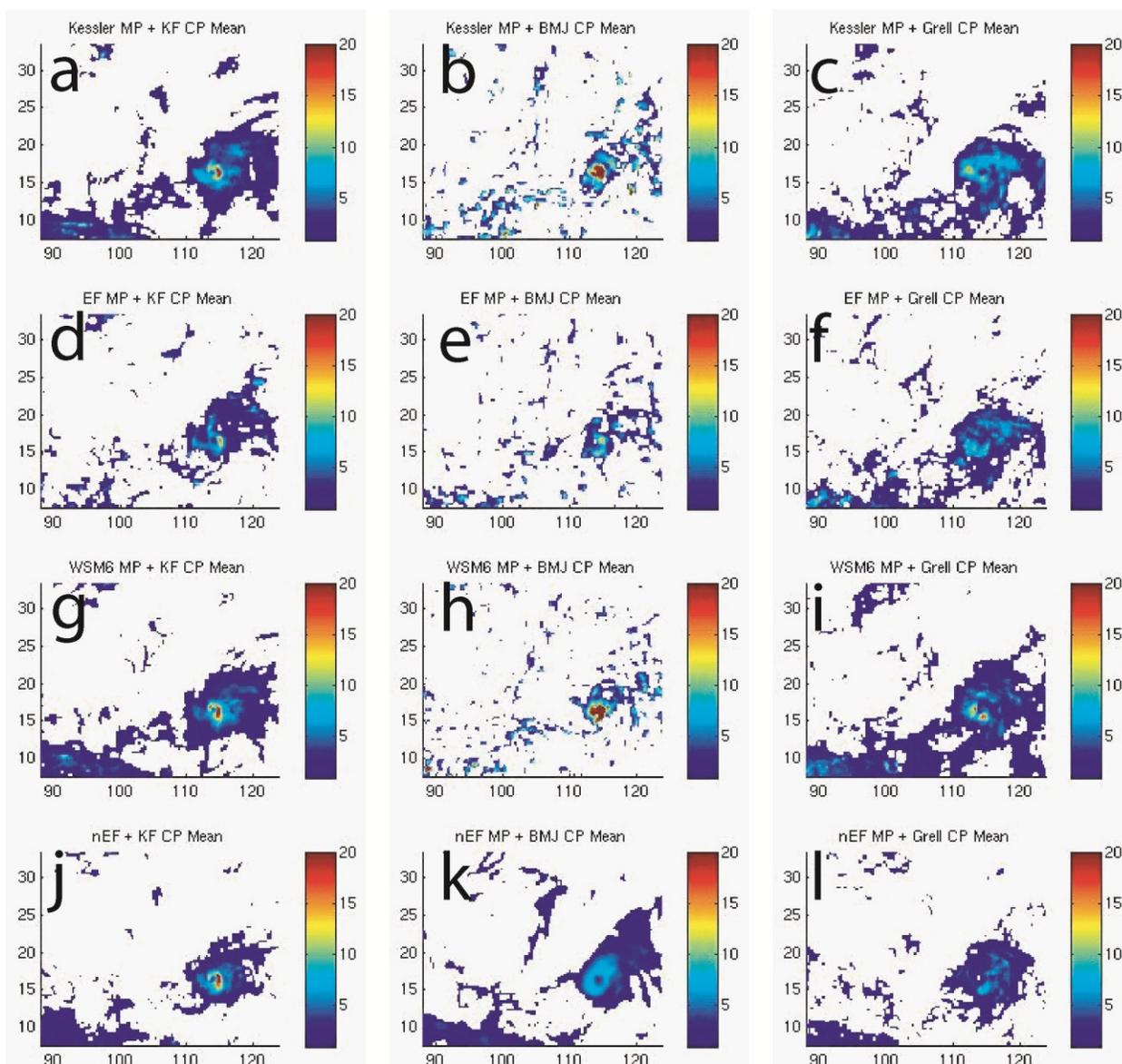


**Figure 3.3** As in Figure 3.2, but for mean wind anomaly (kt) relative to a mean of all combinations.

the horizontal gradient of pressure to greater distances and therefore created increased winds at large radii from the center. F09 also demonstrated that this effect could be removed in the Kessler MP (or created in other MPs) by altering various microphysical parameters such as droplet fall speed.

Further analysis of **Figure 3.3** suggests the GD CP (**Fig. 3.3c,f,i**) cyclones' excessive poleward movement is very similar to that of the Kessler MP cyclones. The large positive wind anomalies at roughly  $5^\circ$  from the storm center for the GD CP relative to any other CP indicates a much larger circulation which would induce a larger "beta drift" than storms produced by other CPs, all else being equal. However, the mechanism by which a larger storm is generated is almost certainly different than that found in F09 given that the phenomenon is present across all MPs. Chapter 2 extensively documented problems relating to the GD and G3 CP's depiction of tropical convection. Specifically, the GD CP overproduces light precipitation and underproduces intense precipitation. It was determined that this resulted from the ensembling nature of the parameterization – specifically due to the fact the GD CP employs 144 members which are run individually and averaged to provide feedback to the model. Given the variety of different ensemble members, it is uncommon for a significant fraction to produce intense precipitation (while being reasonably likely that several may produce intense precipitation). The nature of tropical cyclones provides additional challenges for this parameterization, given that TCs maintain themselves through the development and continuation of intense convection reasonably near the cyclone center, which leads to an increase in vorticity, a warming of the air column, and a decrease in surface pressure.

**Figure 3.4** shows mean three hour precipitation (mm), at the same time and format as **Figure 3.2**. A comparison of GD CP (**Fig. 3.4c,f,i**) mean rainfall to other CPs within a given MP



**Figure 3.4** As in Figure 3.2, but for mean three hour precipitation (mm).

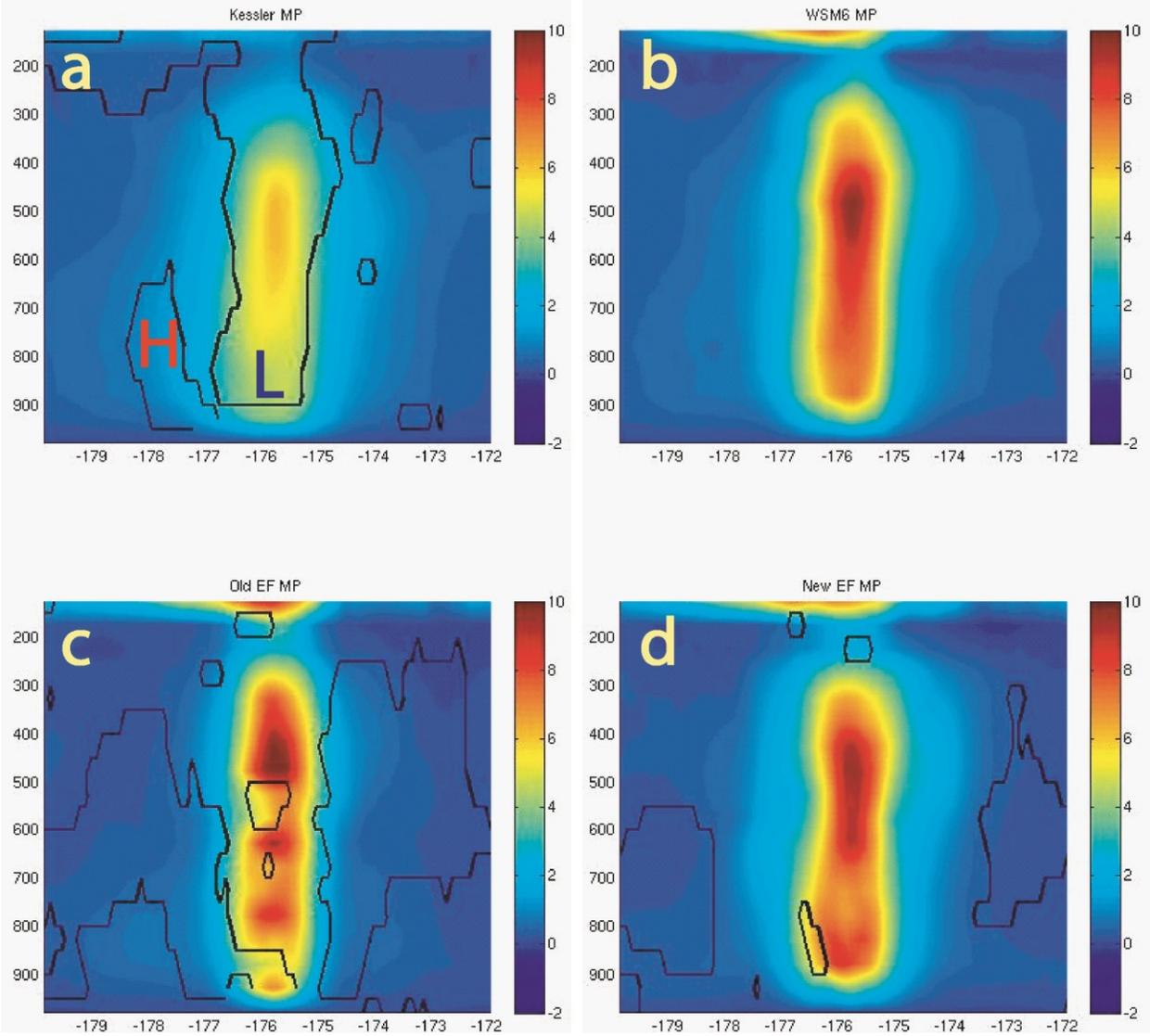
demonstrates that the GD produces as much or more precipitation coverage surrounding the cyclone center as any other CP, which would certainly corroborate **Figure 3.3** indicating a larger mean cyclone. Generally speaking, TCs are characterized by warm, moderately unstable moist air converging towards a central circulation. Since the GD CP is comprised of members employing moisture convergence closures, CAPE removal closures, omega thresholds, etc., it is not unreasonable to conclude that the environment within the 300-1000 km radius of the storm center (as identified in Fiorino and Elsberry) would produce frequent precipitation as several of these closures are activated. Bister (2001) demonstrated that the effect of peripheral convection is to (1) increase a cyclone's horizontal wind field, (2) decrease the likelihood of subsequent intense inner core convection, and (3) decrease the likelihood of rapid intensification.

A further examination of **Figure 3.4** indicates that the mean inner-core rain rates of the GD CP (**Fig. 3.4c,f,i**) cyclones are generally much less than either other CP, which is consistent with Bister (2001). In another study, Jiang (2012) used satellite estimates of rainfall to establish a statistically significant relationship between intense inner-core convection and intensifying TCs. Given the above, one might not expect the GD CP to produce cyclones as intense as the other CPs. While not a statistically significant relationship for all cases given the small sample size, the mean minimum sea level pressure for the GD CP at hour 48 is 999.02 hPa while for the KF CP and BMJ CP it is 998.13 hPa and 998.01 hPa respectively. Although they did not examine this discrepancy, Nasrollahi et al. (2012) showed that the precipitation biases within TCs using the GD CP extend to high-resolution models as well. Nasrollahi et al. (2012) forecast Hurricane Rita (2005) using 20 different combinations of MPs and CPs using the same WRF-ARW version as B09 and found that the GD CP consistently produced a smaller coverage of intense precipitation (> 10 mm/h) than other CPs, which is consistent with **Figure 3.4**.

*Eta Ferrier Microphysics Properties*

Among the nine original parameterization combinations examined in B09 (excluding the nEF simulations), the three weakest mean cyclones at forecast hour 48 were the three using the EF MP, with sea level pressure minima between 3 and 9 hPa weaker than either other MP depending on the choice of CP. The same is true when peak mean wind is examined, as shown in **Figure 3.2** and **Figure 3.3**. If the minimum pressure attained at any point in the cyclone's lifetime is considered, the EF MP also consistently produces weaker peak intensities. Additionally, B09 found that the EF MP created an unusual vertical distribution of potential vorticity in simulated TCs using that parameterization. A composite zonal cross-section of all times between forecast hour 24 and 168 for each MP (all using the KF CP for consistency) for Hurricane Ioke initialized 0000 UTC 26 August 26 is shown in **Figure 3.5**. This particular set of forecasts was chosen for comparison purposes because it was one of the few situations in which the EF MP produced a cyclone of comparable intensity to that of the other MPs, which made comparisons between MPs more appropriate. Although there is no universally accepted standard vertical structure of potential vorticity within intense cyclones, a number of studies (Wu and Kurihara (1996), Molinari et al. (1998), Wu and Wang (2000), Wu and Wang (2001a), Wu and Wang (2001b), Hausman, Ooyama, and Schubert (2006), Kieu and Zhang (2010), Hill and Lackmann (2011)) depict a relatively uniform tower of potential vorticity with a maximum value roughly in the mid-levels of the atmosphere with perhaps a secondary maximum near the top of the boundary layer, similar to that depicted in **Figure 3.5** with the WSM6 MP (**Fig. 3.5b**) or nEF MP (**Fig. 3.5d**).

However, as shown in **Figure 3.5**, the EF MP (**Fig. 3.5c**) forms a potential vorticity structure characterized by a series of alternating maxima and minima. The contours denote



**Figure 3.5** Ioke zonal mean potential vorticity (PVU) for forecast hours 24 through 168 for (a) Kessler MP, (b) WSM6 MP, (c) EF MP, and (d) nEF MP. Contour indicates locations of statistical significance at the 99.9% threshold computed relative to the WSM6 MP. ‘L’ and ‘H’ in panel (a) indicates regions with lower and higher values of potential vorticity relative to panel (b).

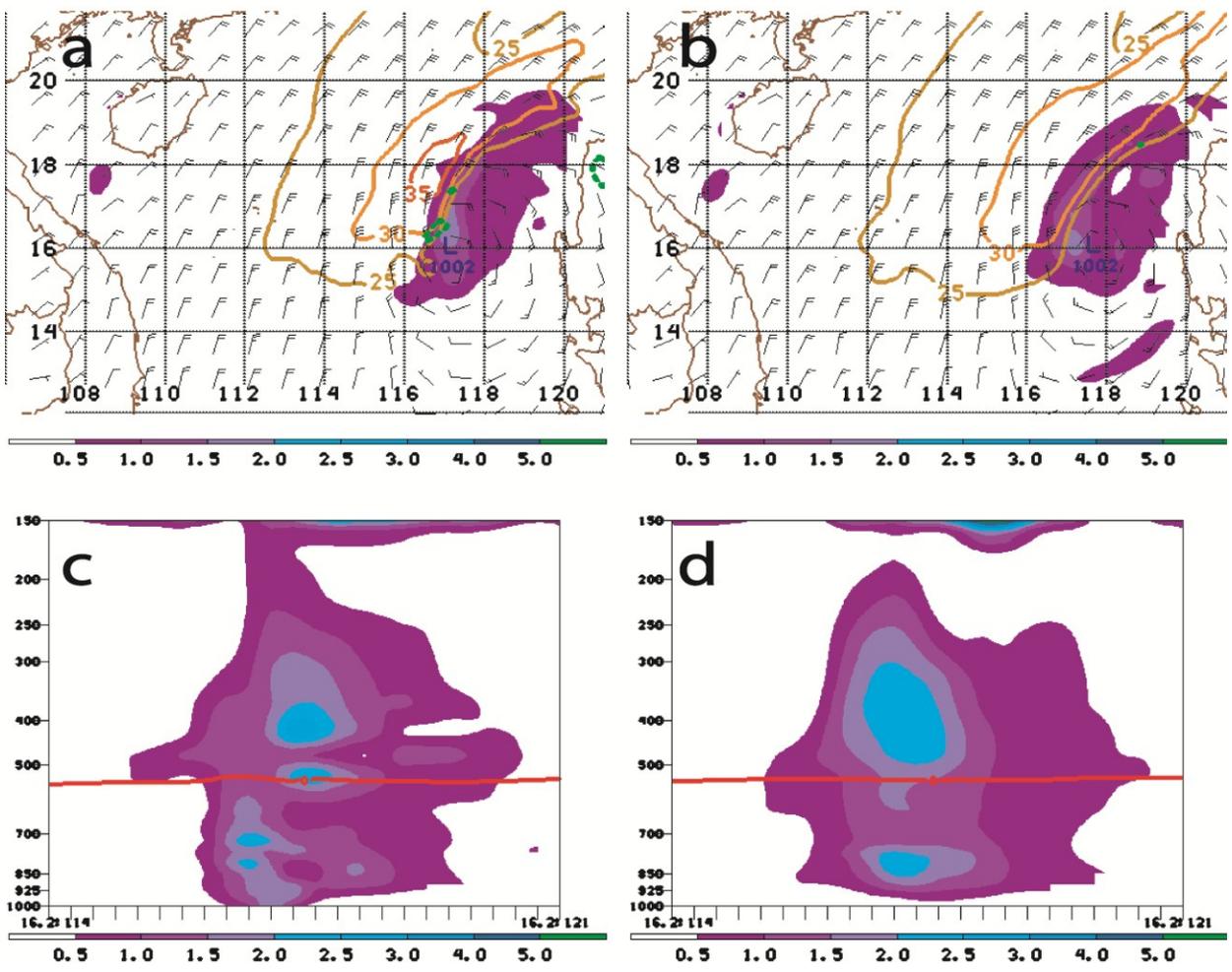
statistical significance using a 99.9% threshold when compared with the WSM6 MP. The EF MP minima are found to be statistically significant anomalies. This unusual structure seems to be significantly ameliorated (although not completely eliminated) when the nEF MP is used. Incidentally, **Figure 3.5** also demonstrates that the Kessler MP (**Fig. 3.5a**) produces a cyclone which is statistically significantly larger than other MPs, as discussed above, and marked by the 'L' and 'H' denoting locations where the potential vorticity is lower and greater, respectively, than the WSM6 MP.

The reason the EF MP simulations were recreated with the nEF MP was due to a bug revealed with the original EF MP in which frozen hydrometeors were treated as liquid by radiation, which was corrected in the nEF MP (Haglund, 2007). Given that the mean freezing level sits immediately below the uppermost potential vorticity maximum (near 500 hPa, not shown), it is possible the minimum directly below that is related to a sudden transition from frozen to liquid hydrometeors, which therefore produced an incorrect profile of heating near this level. While this is possible, this theory does not explain the alternating maxima and minima below this level. Regardless of the exact cause of this phenomenon, closer examination reveals that this potential vorticity structure is a likely reason for the EF MP cyclones' underintensification relative to the other MPs which do not produce such a structure.

As mentioned earlier, these four cyclones were chosen for study specifically due to their inherent real-time forecasting difficulty, specifically where it concerned track forecasts. The one exception to this was Hurricane Ioke, which was very well predicted due in part to a deep easterly flow in which it was embedded. One particular commonality among the remaining three cyclones was variable steering winds with height or time. For instance, in the case of Hurricane Ernesto and Typhoon Cimaron, it was unclear whether a digging trough would accelerate the

cyclone to the northeast, or whether the cyclone would instead be steered westward by the low-level easterly flow. Hurricane Ernesto was forecast to continue moving generally west-northwestward, while Typhoon Cimaron was initially forecast to recurve ahead of a digging trough. In reality, the opposite happened for these respective cyclones. Hurricane Lenny posed similar problems, as it was being steered eastward by a deep trough despite low level easterly flow. The commonality with these storms (except Hurricane Ioke) is a reasonable amount of vertical wind shear. Since Hurricane Ioke is the one storm where the EF MP intensifies the cyclone a similar amount as the other MPs, the logical implication is that the potential vorticity structure produced by the EF MP leaves the cyclone more susceptible to vertical wind shear, and therefore leads to a weaker cyclone.

The best way to test this phenomenon would be to examine a situation where the EF MP cyclone and the nEF MP cyclone are both affected by the same shearing environment while remaining in similar positions relative to one another during which the EF MP cyclone weakens while the nEF MP cyclone intensifies. Finding an ideal example is slightly problematic, as changes in cyclone structure impact the cyclone's movement (or steering flow) which frequently leads to slowly growing track differences (and structures) with track splits that are difficult to pinpoint. However, the set of forecasts initialized for Typhoon Cimaron at 0000 UTC 30 October 2006 provides a suitable EF/nEF MP pair of forecasts for comparison. The top row of **Figure 3.6** depicts the 12 hour 200-1000 hPa potential vorticity (PVU), 10 m wind speed (contoured above 25 kt), 3 hour rainfall above 25 mm, and minimum sea level pressure associated with each MP. The two forecasts are broadly similar, although one might be inclined to believe the EF MP (**Fig. 3.6a**) cyclone is slightly stronger based on stronger surface winds and more intense precipitation near the center of circulation. However, both have similar magnitudes and distributions of



**Figure 3.6** (a) and (b) show 200-1000 hPa potential vorticity (PVU, fill), 10 m wind (kt, magnitude contoured above 25 kt), and 3 hour precipitation above 25 mm (heavy green dashed contour) for the EF MP and nEF MP, respectively. (c) and (d) depict a zonal cross-section of potential vorticity (PVU, fill) as well as the 0° C isotherm for the EF MP and nEF MP, respectively.

potential vorticity and the same minimum surface pressure. It should be noted that these simulations initialized Cimaron too weak, which is reflected in the relative high surface pressures shown in **Figure 3.6a,b**. Officially, Cimaron was a typhoon with maximum sustained winds of 90 kt at this time.

When a zonal cross-section is taken through each cyclone, as shown by the bottom row of **Figure 3.6**, a much different structure of potential vorticity emerges than one might presume based on the 200-1000 hPa distribution of potential vorticity seen in the top row of **Figure 3.6**. While the outline of the 0.5 PVU surface is effectively the same for each cyclone, the EF MP (**Fig. 3.6c**) is characterized by an alternating series of maxima and minima similar to that depicted with Hurricane Ioke in **Figure 3.5** while the nEF MP (**Fig. 3.6d**) depicts a more uniform distribution with a maximum in the mid levels of the atmosphere and one at the top of the boundary layer. The freezing level is also plotted, which again shows a maximum-minimum couplet of potential vorticity straddling the freezing level in the EF MP (**Fig. 3.6c**) cyclone as with Ioke (and to a lesser extent in the nEF MP cyclone (**Fig. 3.6d**)). At this time, the cyclone was slowly moving westward and entering an area characterized by northeasterly flow at low levels (as seen by **Fig. 3.6a,b**) and southeasterly flow at upper levels (not shown). In reality, Cimaron moved slowly westward and intensified into a strong typhoon with surface winds exceeding 100 kt by 1 November before stalling near 18° N, 116° E and slowly weakening as it upwelled cooler water.

**Figure 3.7** follows the forecasted cyclones through time, showing low-level potential vorticity (800-900 hPa), mid- and upper-level potential vorticity (350-550 hPa), the difference in mean wind for these layers, 3 hour precipitation greater than 25 mm/3 h, and minimum sea level pressure at twelve hour intervals beginning at forecast hour 36 and ending at hour 72. In the EF

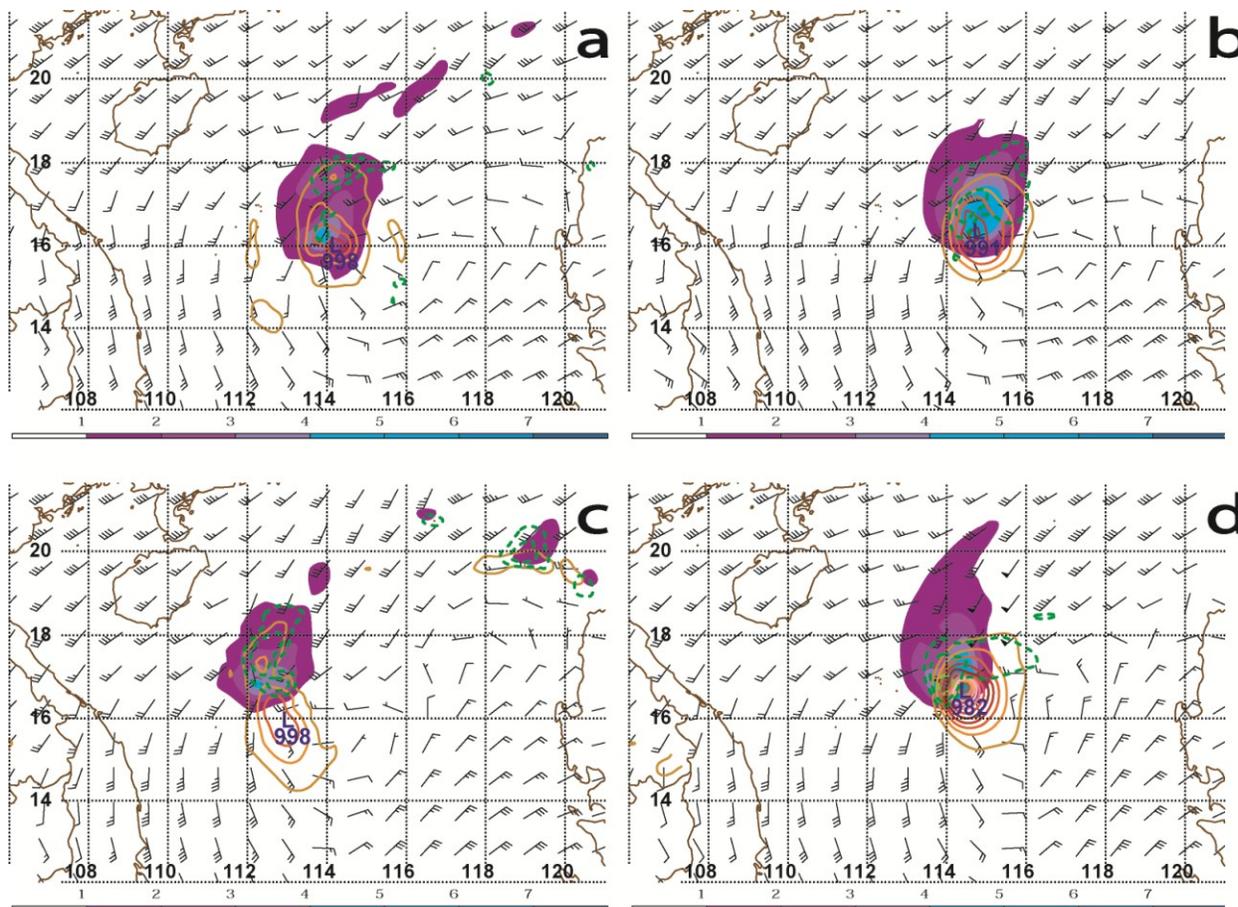
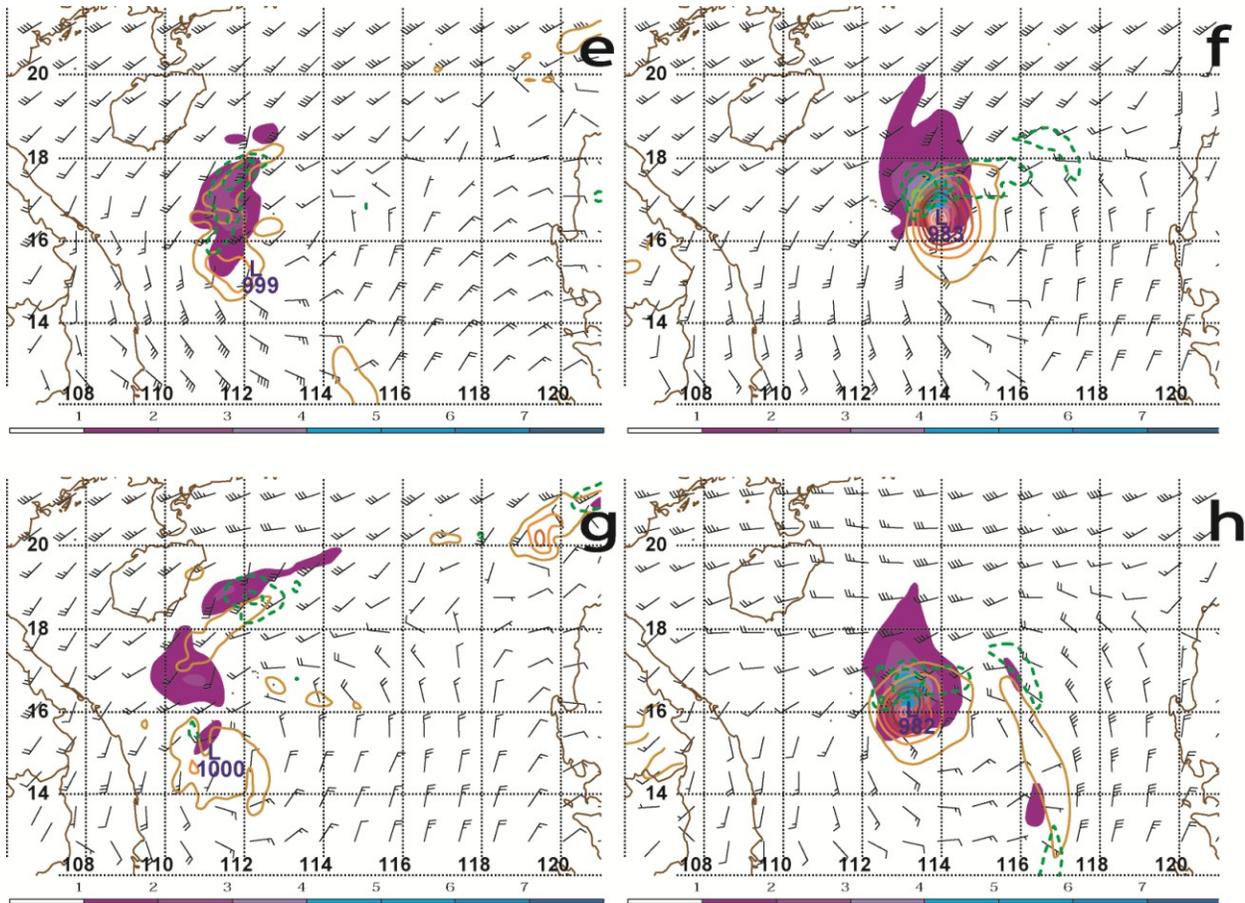


Figure 3.7 (continued on next page).



**Figure 3.7** All panels depict 350-550 hPa potential vorticity (PVU, fill), 800-900 hPa potential vorticity (solid contour, interval 1 PVU), difference in mean wind between 350-550 hPa and 800-900 hPa (kt, barbs), and 3 hour precipitation above 25 mm (heavy green dashed contour for 25 mm, light green dashed contour for 50 mm, bright green contour for 100 mm). Forecast time 36, 48, 60, 72 shown in panels (a,b), (c,d), (e,f), (g,h) respectively. Panels (a,c,e,g) depict EF MP while panels (b,d,g,h) depict nEF MP.

MP simulation (**Fig. 3.7a,c,e,g**), the southerly vertical wind shear over the cyclone slowly displaces and advects the mid-level potential vorticity maxima northward away from the remaining low-level circulation, which conversely drifts southwestward. Although convection remains associated with the low-level circulation as shown by the green contours, the majority of precipitation follows the mid- and upper-level potential vorticity northward and northeastward with time. In contrast, the nEF MP (**Fig. 3.7b,d,f,h**) cyclone does become noticeably tilted, with mid-upper level potential vorticity slanted north-northwestward relative to the low-level potential vorticity, yet the two never fully decouple as with the EF MP simulation. Similarly, the vast majority of the precipitation associated with the cyclone remains near the center (albeit displaced northward), with rain rates up to 100 mm/3 h. Given these two scenarios, it is not surprising that the EF MP cyclone slowly weakens while the nEF cyclone intensifies even while experiencing moderate wind shear (which as noted occurred in reality). The enhanced vertical coupling of the nEF MP cyclone causes a more westward motion for the surface cyclone relative to the EF MP, while also moving at a slower speed, which is more consistent with the observed motion of Cimaron.

To further analyze the relationship between vertical wind shear and the decoupling (or lack thereof) of the potential vorticity structure within the two cyclones, a series of cross-sections are taken for the times depicted in **Figure 3.7**. The cross-sections shown in **Figure 3.8** are chosen such that the cross-section is roughly aligned so that both the low-level potential vorticity maxima and the mid-upper level potential vorticity maxima are included within the cross-section (e.g. the direction of the cross-section is roughly parallel to the wind shear vector). Every cross-section is perfectly oriented south-southeast to north-northwest, with each end point representing a location  $4^\circ$  north and  $2^\circ$  west of the starting location. The approximate midpoint between the

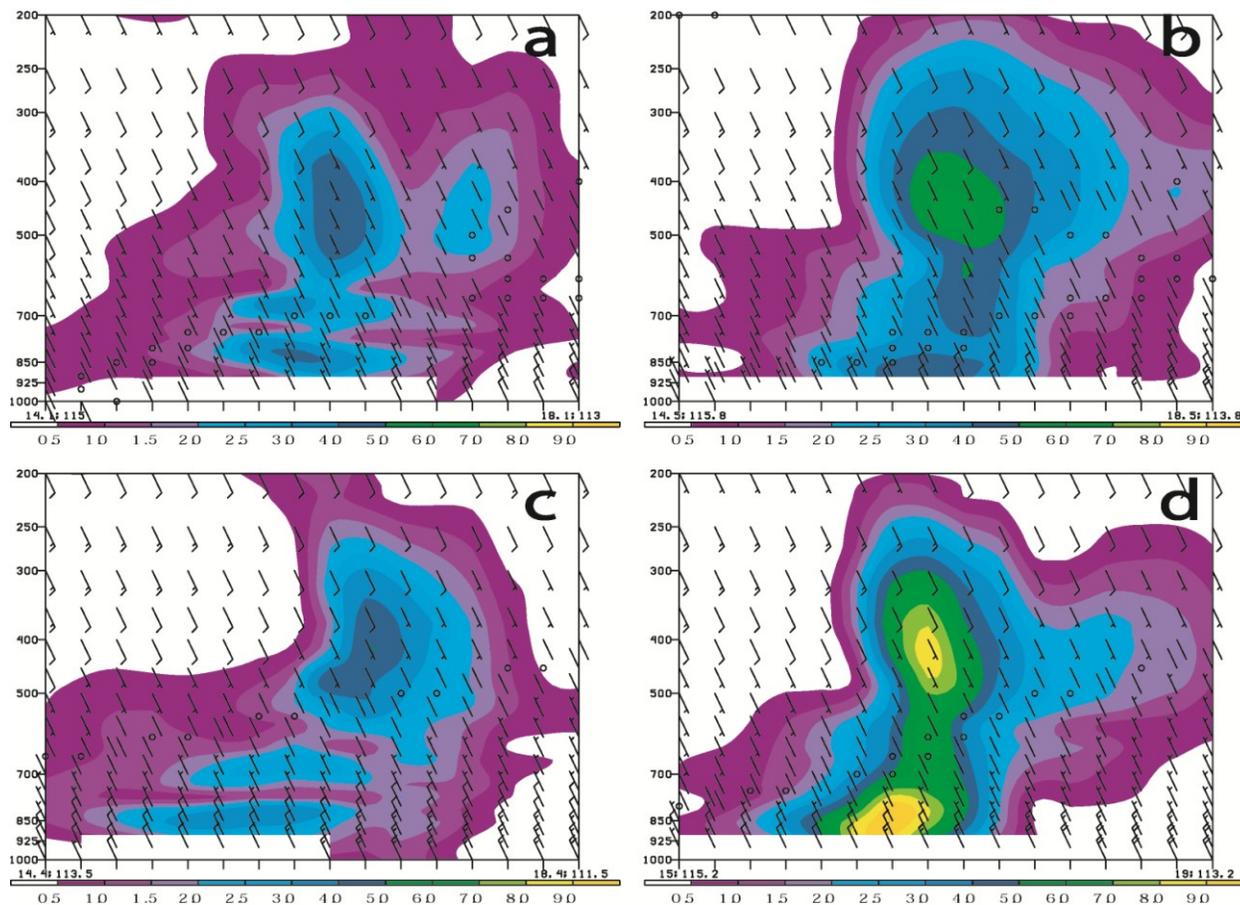
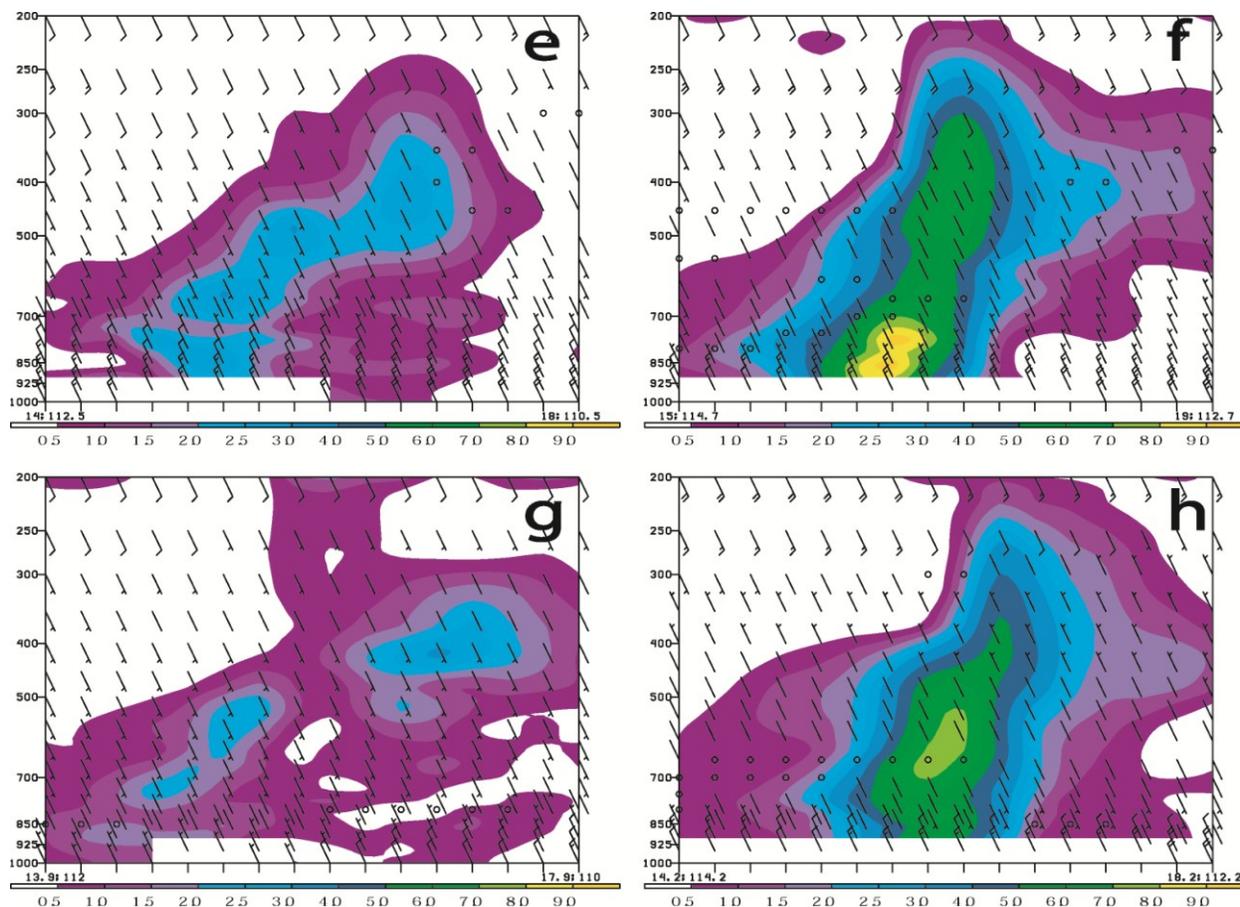


Figure 3.8 (continued on next page).



**Figure 3.8** All panels depict a cross-section of potential vorticity (PVU, fill) and smoothed tangential (to cross-section) wind (kt, barbs). Forecast time 36, 48, 60, 72 shown in panels (a,b), (c,d), (e,f), (g,h) respectively. Panels (a,c,e,g) depict EF MP while panels (b,d,g,h) depict nEF MP.

surface circulation and the mid-level circulation for each cyclone is the midpoint of each cross-section. Finally, a heavily smoothed wind vector (to best eliminate the influence of the cyclone's circulation) tangential to the cross-section is shown. Interpreting wind barbs on cross-sections can be challenging, although since these are just the winds tangential to the cross-section, the orientation of the wind barbs also is the orientation of the cross-section (e.g. parallel to the tilt of the cyclones).

For both cyclones and for all times, it can be seen that the wind vectors switch direction at roughly 650 hPa, with approximately 20 kt of low level flow to the south-southeast and approximately 10 kt of upper level flow to the north-northwest, which represents a steady 30 kt of wind shear affecting either cyclone within a roughly 300-900 hPa layer. However, while the nEF MP (**Fig. 3.8b,d,f,h**) cyclone begins to tilt downshear slightly, the EF MP (**Fig. 3.8a,c,e,g**) cyclone quite clearly separates into a low- to mid-level potential vorticity tower extending upward to roughly 500 hPa and a separate maxima above which is slowly advected away from the low-level maxima. By forecast hour 72 (**Fig. 3.8h,g**), the circulations imposed by the lower and upper potential vorticity maxima in the EF MP cyclone would be destructively interfering with one another, which would also act to increase the likelihood of weakening.

Considering both cyclones began to experience roughly equivalent wind shear while possessing an equal intensity (although with unique potential vorticity distributions), a compelling case can be made that the unusual potential vorticity structure produced by the EF MP creates a serious impediment to strengthening (or maintenance) in the presence of wind shear. One might conjecture that the differing potential vorticity structures are related to weaker convection associated with the EF MP. However, as shown by **Figure 3.6**, the EF MP cyclone actually possesses greater initial rainfall rates near the center of circulation. It isn't until the

cyclones have begun to diverge in their intensity that significant precipitation differences become apparent.

As mentioned earlier, the one set of forecasts in which the EF MP does not exhibit significant underintensification relative to other MPs are those involving Hurricane Ioke, which was situated in an environment of fairly uniform deep-layer easterly flow. Conversely, similar evolutions can be seen in virtually all Typhoon Cimaron, Hurricane Ernesto, and Hurricane Lenny simulations as are described in the specific Cimaron example above. In these cases, the low-level and upper-level circulations slowly decouple from one another in the EF MP. As the disparity grows, the destructive interference of the circulations adds a further deterrent to intensification, as low-level convergence decreases which makes additional intense convection less likely. While resulting in generally weaker storms relative to other MPs, it also appears to be a likely cause of the storms' more equatorward position at forecast hour 120. By this forecast time, the shallower EF MP cyclones are generally steered by low-level easterly flow rather than being significantly influenced by upper-level flow.

*Additional Relevant Research: Tropical Cyclogenesis*

Extending the GD CP conclusions discussed earlier to TC genesis and intensity can be achieved by studying the 2009 real-time ensembles described in the previous chapter, which allows for a more independent comparison between CPs. Only forecasts using the 30 km inner domain are used for this analysis. First, a subjective examination of all simulations was undertaken beginning at forecast hour 48 to determine location and intensity of all TCs within any forecast using the Yonsei University boundary layer parameterization and the WSM3 MP. This was done to allow clean comparisons, and leaves three ensemble members – those using the

KF CP, the BMJ CP, or the G3 CP. In order to determine the existence of a tropical cyclone, the following criteria were used:

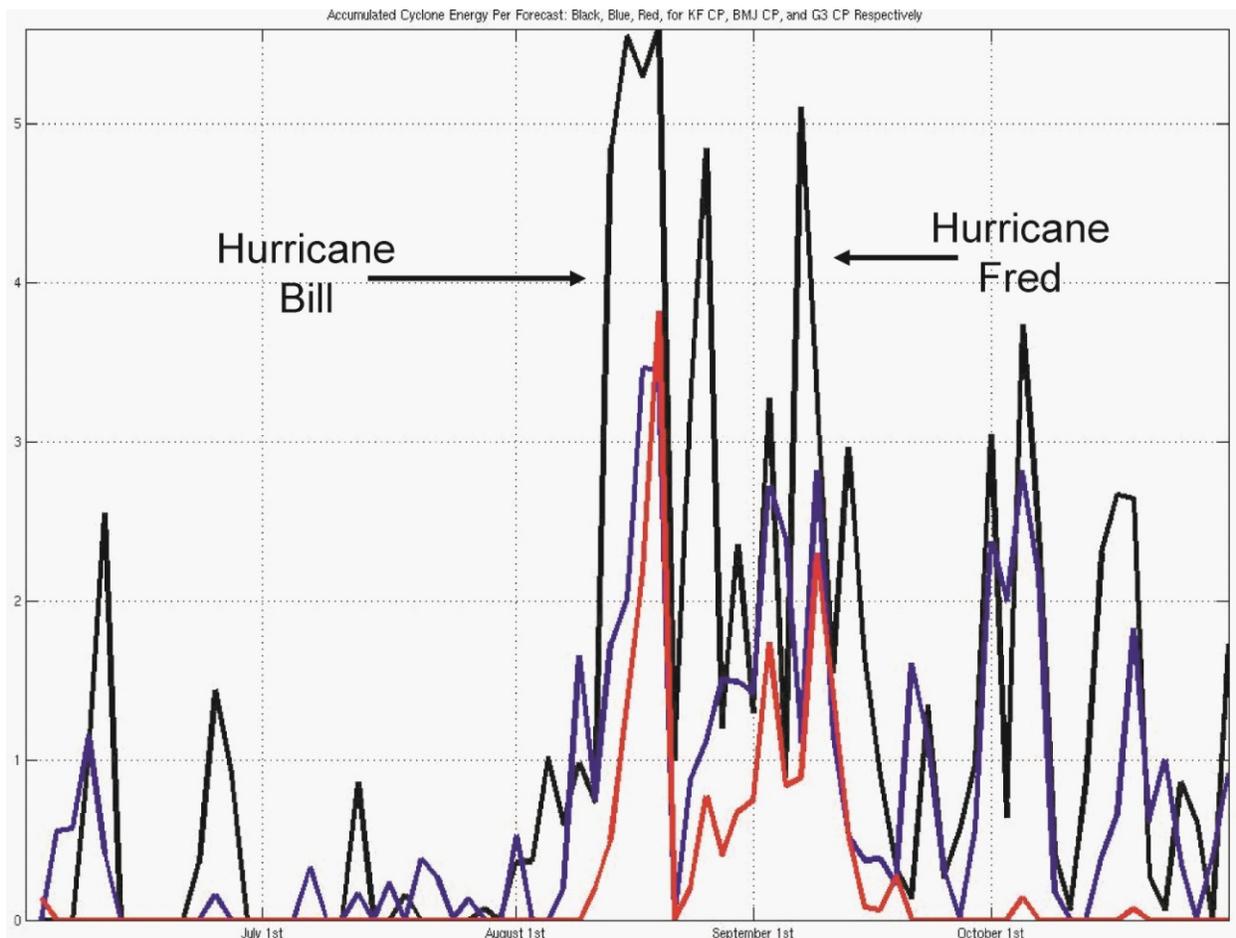
- (1) A closed surface and 850 hPa cyclonic circulation must be present for a minimum of six hours
- (2) The cyclone must have a warm-core at 850 hPa
- (3) 10 m winds must exceed 25 kt at some location within the circulation
- (4) Precipitation must be associated with the circulation at time of declaration

The decision to begin at forecast hour 48 is based on research described earlier which marked forecast hour 48 as a point of significant forecast divergence within a physics-based ensemble. This appears to be a rough threshold point at which a given parameterization has shaped the atmosphere in such a way that forecasts start significantly diverging from one another. In addition, as mentioned in the introduction, most current parameterization research examines forecast timescales less than 48 hours.

For all circulations conforming to the above criteria, the minimum sea level pressure, maximum 10 m wind, and latitude/longitude coordinates are noted at every six hour interval after hour 48 until the end of each forecast at hour 120. Once this is done, the accumulated cyclone energy (ACE) is calculated for each forecast. Traditionally, ACE is defined as the square of maximum surface wind (multiplied by  $10^{-4}$ ) for each storm at tropical storm intensity or greater and at every synoptic time summed over a storm's or season's lifetime. Two slight modifications to this definition are made here. First, ACE will be calculated as the accumulated value of all storms within a given forecast after hour 48. Second, the minimum threshold is lowered from tropical storm intensity to tropical depression intensity in order to incorporate more data (because ACE is calculated as the square of the maximum wind speed, this affects very low

values of ACE much more than high values). **Figure 3.9** shows the progression of ACE for each parameterization throughout the 2009 season. For reference, a weak tropical storm lasting for several synoptic times would create a value of ACE roughly equal to 1, while a weak short-lived tropical depression would generate an ACE value slightly above 0. Generally speaking, ACE is maximized for all parameterizations during August and September, which is intuitive considering that the 2009 season had nine systems of tropical depression strength or greater between mid-August and early October (one system in May and one system in November are outside of the time these forecasts were performed). The maximum in all three CPs is associated with Hurricane Bill in mid-August, and the secondary maximum in mid-September is associated with Hurricane Fred.

Earlier in this chapter, composite analysis of several targeted cyclones at forecast hour 48 indicated that the GD CP cyclones were generally weaker (either when measured by minimum sea level pressure or maximum surface winds) than the cyclones modeled using the KF CP or BMJ CP. **Figure 3.9** reaffirms that conclusion due to the G3 CP's ACE generally being weaker than either other CP for virtually every forecast. Cumulatively, the mean ACE generated per forecast is 1.214, 0.733, and 0.255 for the KF, BMJ, and G3 CPs respectively. Additionally, the G3 CP seems much less likely to produce a TC when that TC was not already present in the initial conditions, as seen by the frequent 0 values associated with the G3 CP relative to the other CPs. In some respects, this is a good thing because the rate of false alarms is dramatically reduced (note all of the spurious values of ACE during June and July in the KF CP and BMJ CP when no storm existed). On the other hand, this dramatically reduces the ability of the G3 CP to provide valuable advance information about tropical cyclogenesis. No specific judgment is applied to this distinction, although potential causes will be described.

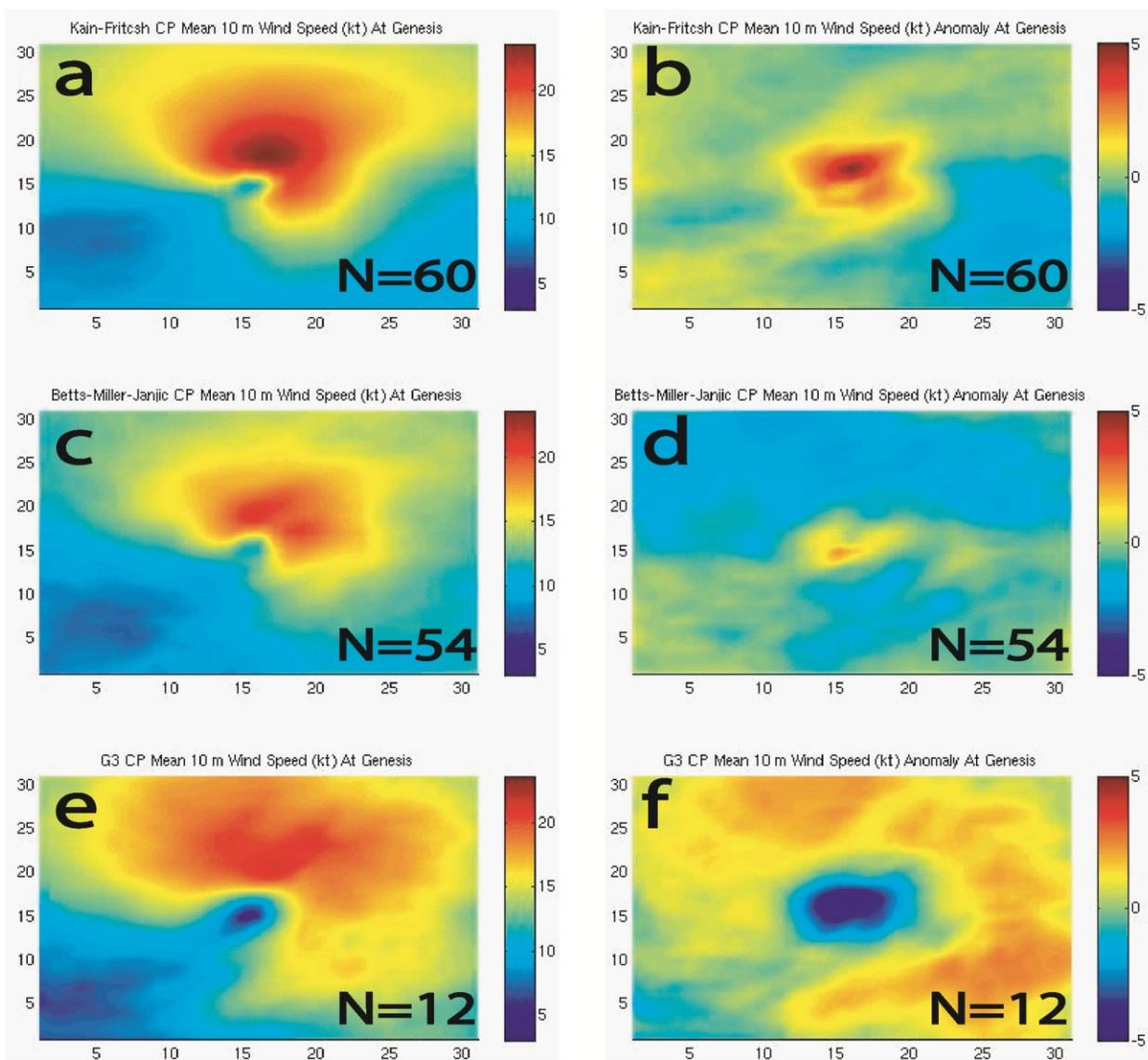


**Figure 3.9** ACE is show for KF CP (black), BMJ CP (blue), and G3 CP (red) for the 2009 real-time forecasts.

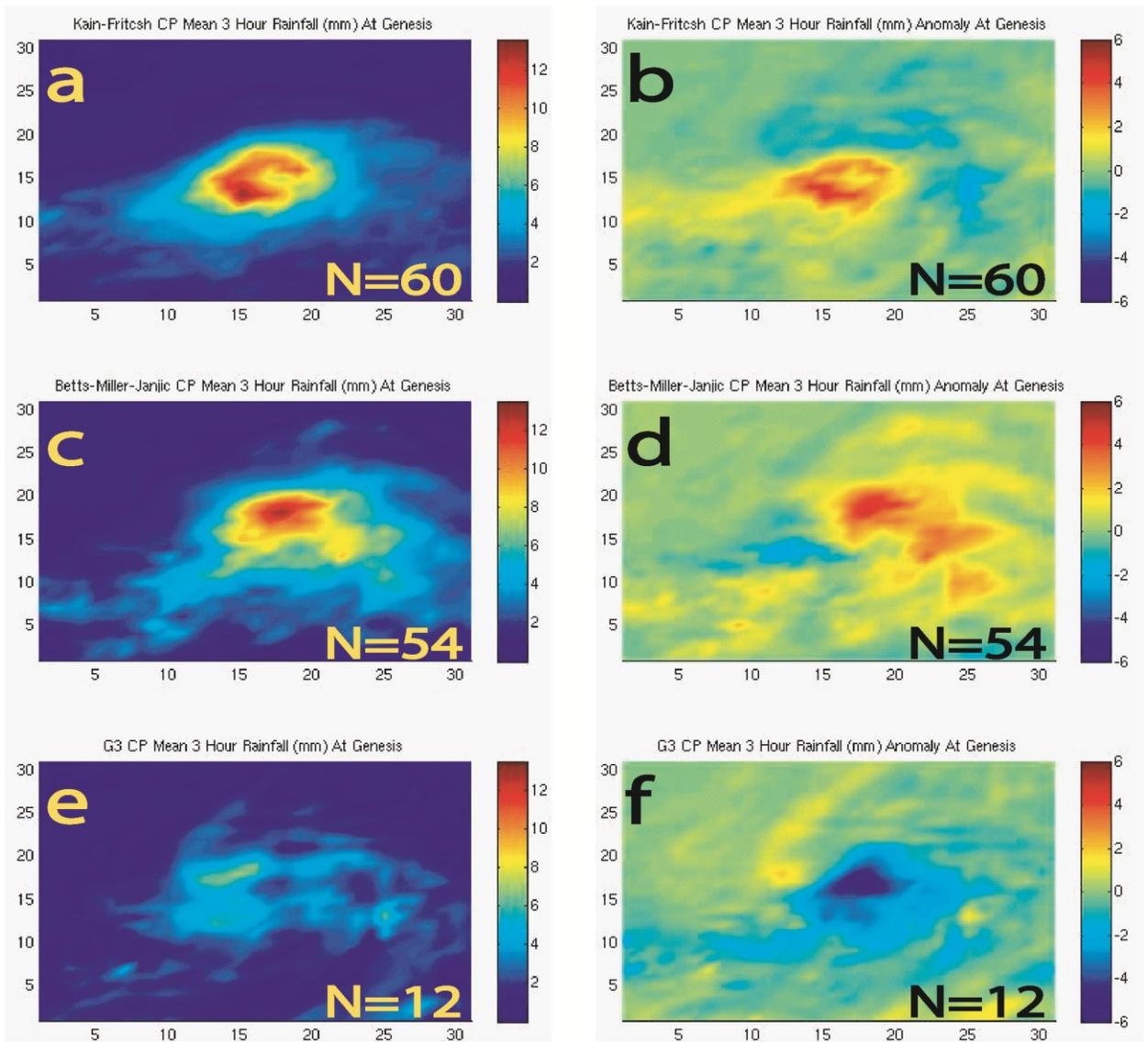
Tropical cyclogenesis is further examined by compositing the incipient disturbances at time of declaration, provided that declaration occurred after forecast hour 48. This allows for the disturbances to be entirely a product of each individual forecast, rather than being constrained to exist based upon their presence in the initial conditions used. A total of 12 cases occur with the G3 CP, while 60 and 54 occur with the KF CP and BMJ CP respectively. This partially illustrates the disparity between the likelihood of cyclogenesis between the CPs. Each composite is created by centering the minimum sea level pressure within a 31 x 31 grid point grid. This corresponds to a box roughly 10° of latitude or longitude on a side.

**Figure 3.10** indicates the mean magnitude of 10 m wind (kt) at time of declaration as well as the anomaly for each CP computed relative to a mean of all three CPs. All three possess maximum winds north of the cyclone center, which is a fairly common location for westward moving cyclones. However, an examination of the anomalies show that the G3 CP (**Fig. 3.10e,f**) produces sprawling cyclones with much lighter winds near the center and greater winds at a distance from the center relative to the other CPs. This is very consistent with **Figure 3.3**, which showed larger circulations (and weaker near-center winds) associated with the GD CP. **Figure 3.11** shows mean 3 hour rainfall (mm) for each CP at time of genesis, which demonstrates that the G3 CP (**Fig. 3.11e,f**) exhibits significantly lower rain rates near the center of circulation relative to the KF CP (**Fig. 3.11a,b**) and BMJ CP (**Fig. 3.11c,d**), while possessing slightly greater rain rates far from the center which is similar to that shown in **Figure 3.4**.

The lack of intense precipitation near the center is likely related to the G3 CP's inability to develop cyclones (either to the stage of genesis or in terms of developing intense TCs after genesis), as there would be much less latent heating concentrated near the center, and thus a decreased ability to generate concentrated low surface pressure and related spin-up of vorticity.



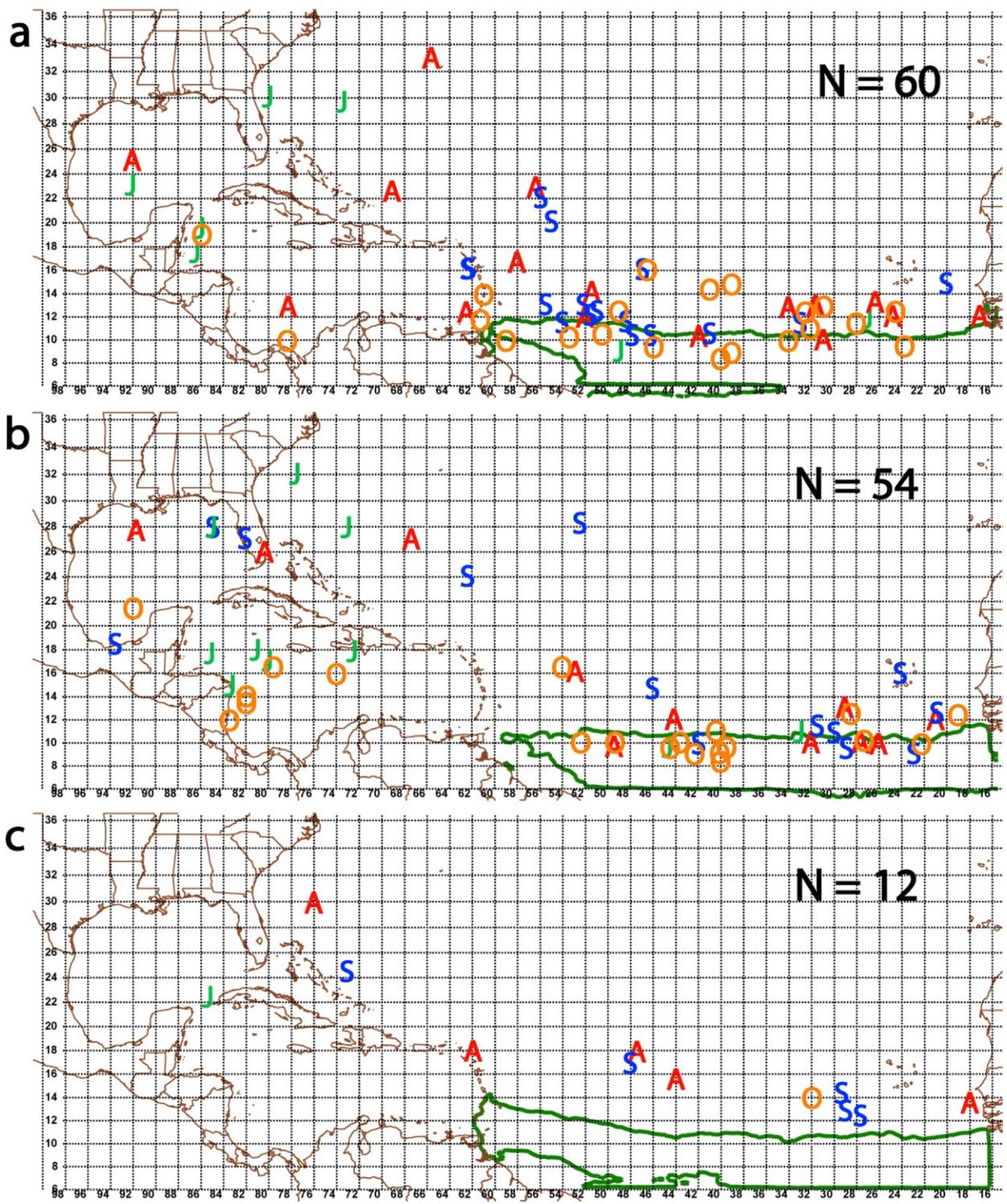
**Figure 3.10** Composite mean 10 m winds (kt) for (a) KF CP, (c) BMJ CP, (e) G3 CP, and 10 m wind anomalies relative to a mean of (a,b,c) for (b) KF CP, (d) BMJ CP, and (f) G3 CP. ‘N’ indicates number of events per CP.



**Figure 3.11** As in Figure 3.10 but for three hour precipitation (mm).

Lunney (1988) examined developing and non-developing cloud clusters and found that the cases in which genesis occurred were characterized by deep convection within  $2^\circ$  of the center of circulation, while non-developing systems had deep convection occurring outside this radius. Other studies such as those by Tory, Montgomery, and Davidson (2006) demonstrated that TC genesis is aided primarily by an increase in vorticity accomplished through low-level convergence and intense updrafts within deep convection. The G3 CP exhibits weaker low-level convergence (to be discussed later) and substantially weaker upward vertical motion over the center, so it is therefore unsurprising that the G3 CP is less likely to produce TCs than other CPs. As mentioned previously, the inability of the G3 CP to produce deep convection is in part due to a preference for greater coverage of light rainfall rather than small areas of intense rainfall as well as excessive mid-level drying occurring due to an overactive parameterization.

**Figure 3.12** notes the genesis points of all cyclones included within the composites for each parameterization. The distributions of KF CP (**Fig. 3.12a**) and BMJ CP (**Fig. 3.12b**) cyclones is generally similar, although the KF CP has significantly more cyclones develop just east of the Lesser Antilles, while the BMJ CP develops many more cyclones west of approximately  $70^\circ$  W. An exact cause of this disparity is unknown, although the KF CP does produce more cyclones in the eastern Pacific (not shown) which may lead to an increase in vertical wind shear across the western Caribbean and Gulf of Mexico, negatively affecting potential cyclones in the KF CP. This region is also characterized by lower values of upper-level potential vorticity in the BMJ CP (not shown), although it is not clear whether this is a result of, or aid to, tropical cyclone development. One noteworthy difference between the G3 CP (**Fig. 3.12c**) and the other two CPs (besides the obvious difference in number of developing cyclones) is their latitudinal location. All the G3 CP cyclones develop north of  $12.5^\circ$  N. However, for the

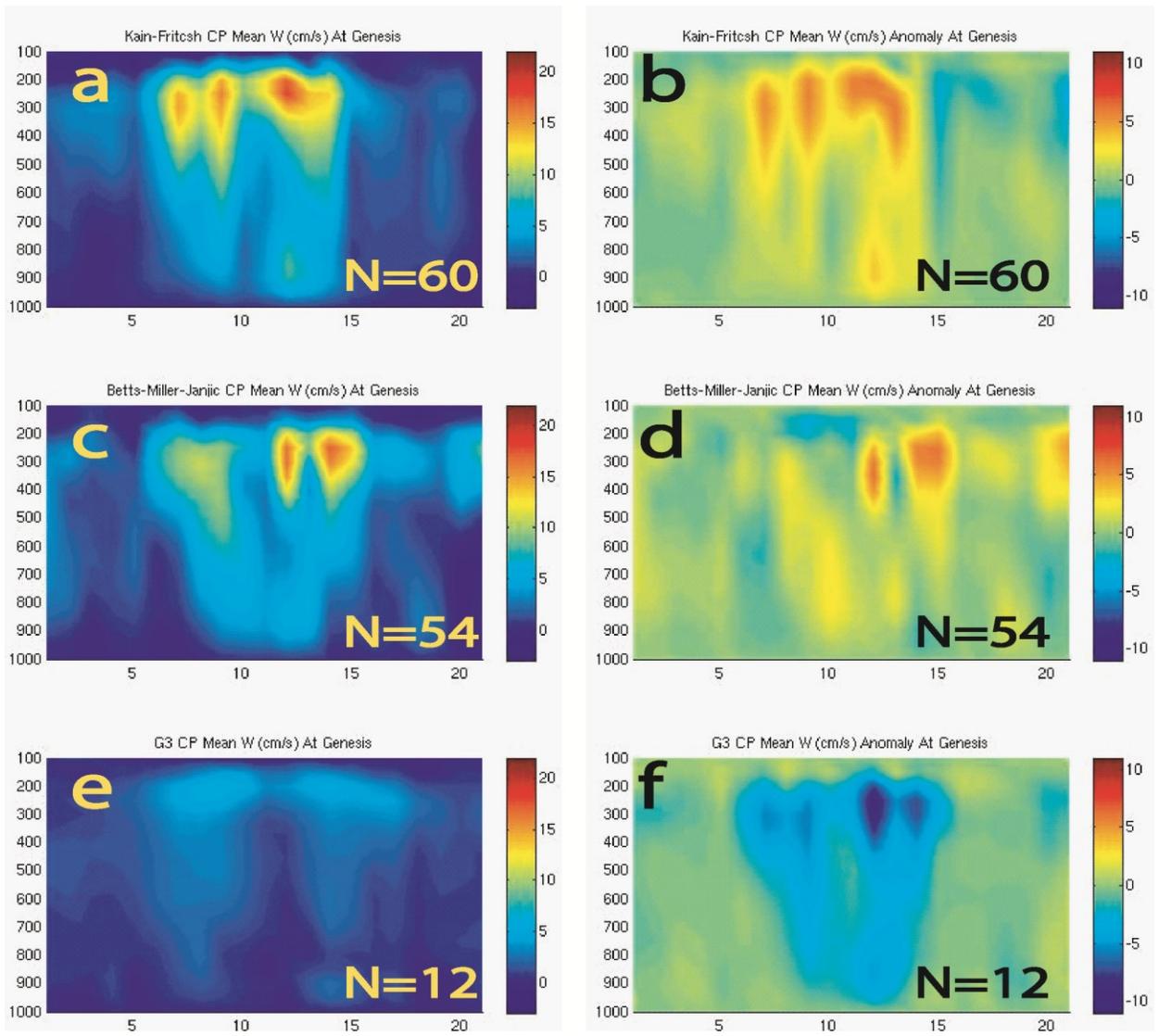


**Figure 3.12** Genesis locations for (a) KF CP, (b) BMJ CP, and (c) G3 CP. ‘N’ indicates number of events while ‘J’, ‘A’, ‘S’, ‘O’ indicate occurrence in June/July, August, September, and October, respectively. The green contour denotes the location of mean 5 day rainfall exceeding 50 mm.

KF CP and BMJ CP cyclones, the majority developing east of the Lesser Antilles develop south of  $12.5^{\circ}$  N.

An examination of mean total rainfall over the entire domain (as shown by the green contour, which denotes locations with rainfall greater than 50 mm over 5 days) for the three CPs yields a similar distribution, depicting an ITCZ located south of roughly  $12^{\circ}$  N. Mean rainfall quickly decreases for locations north of  $12^{\circ}$  N, with these areas receiving less than 50 mm over 5 days. This suggests that the KF CP and BMJ CP are much more likely to develop TCs within the ITCZ than the G3 CP. Although not quantified, it was noted when doing an analysis of genesis locations that the KF CP and BMJ CP were able to produce (and destroy) circulations much more easily than the G3 CP. This is likely due to the more intense convective updrafts noted with these CPs (as shown in Chapter 2), which would be more conducive to enhancing low-level vorticity. Since these intense updrafts are characteristic of convection within the ITCZ, it is understandable that these two CPs would produce more cyclones within this region.

Conversely, the G3 CP cyclones are characterized by large circulations which produce weak updrafts in the presence of weakly positive low-level vorticity rather than single intense updraft cores near or over the cyclone center characterized by large values of vorticity. This has the effect in the G3 CP cyclones of slower development. The combination of a slower development and larger circulation may be the cause of the northward shift in G3 CP development – the incipient circulation may form within the ITCZ, but with time the circulation drifts northward before developing, enhanced by a greater Beta drift due the size of the circulation. The mean vertical motion can be seen in **Figure 3.13**, which shows a zonal cross-section of mean vertical motion at time of genesis for the three CPs as well as the vertical motion anomalies. The KF CP (**Fig. 3.13a**) and BMJ CP (**Fig. 3.13c**) cyclones have clear areas of

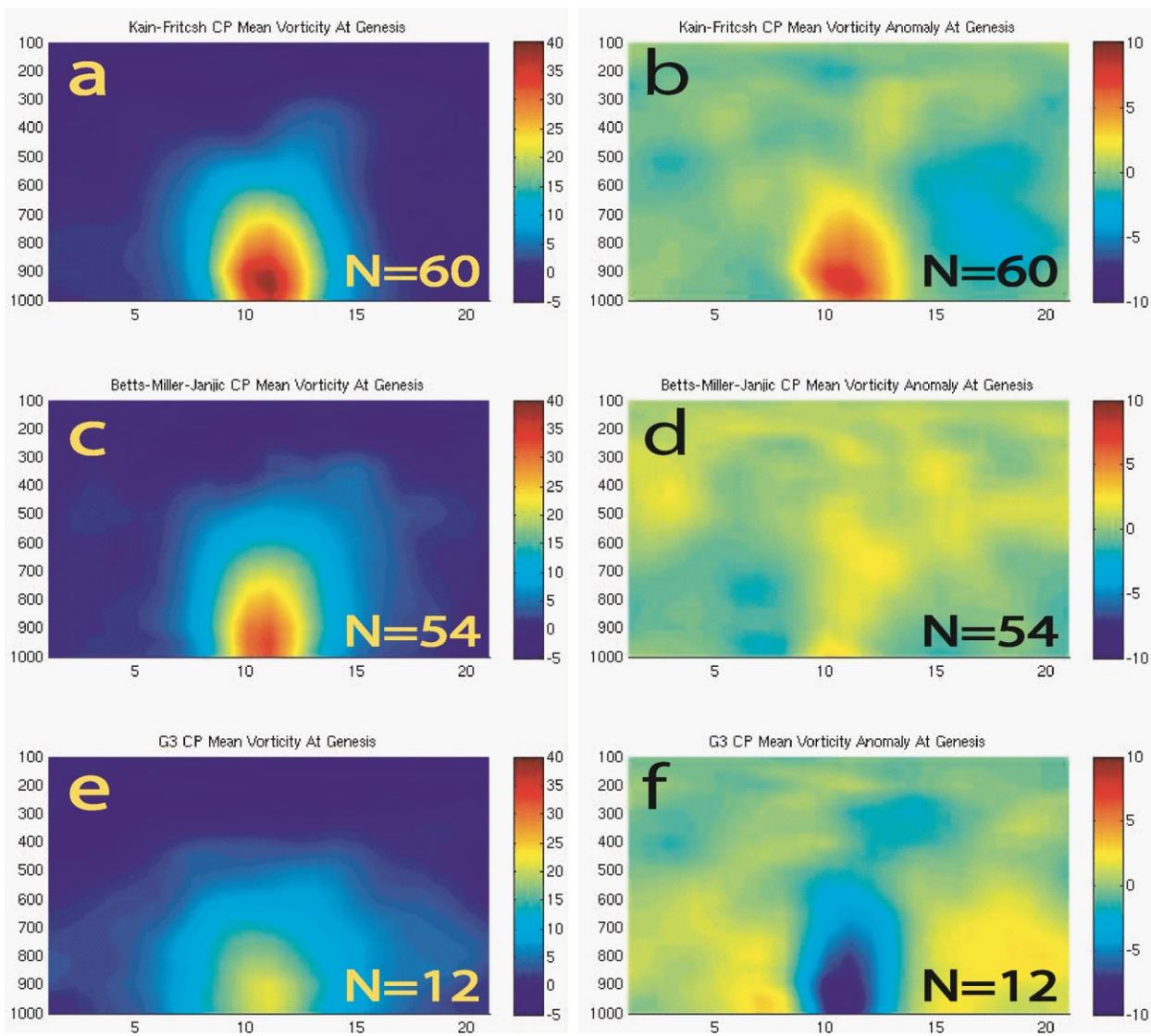


**Figure 3.13** As in Figure 3.10, except depicting a zonal cross-section of vertical motion (cm/s).

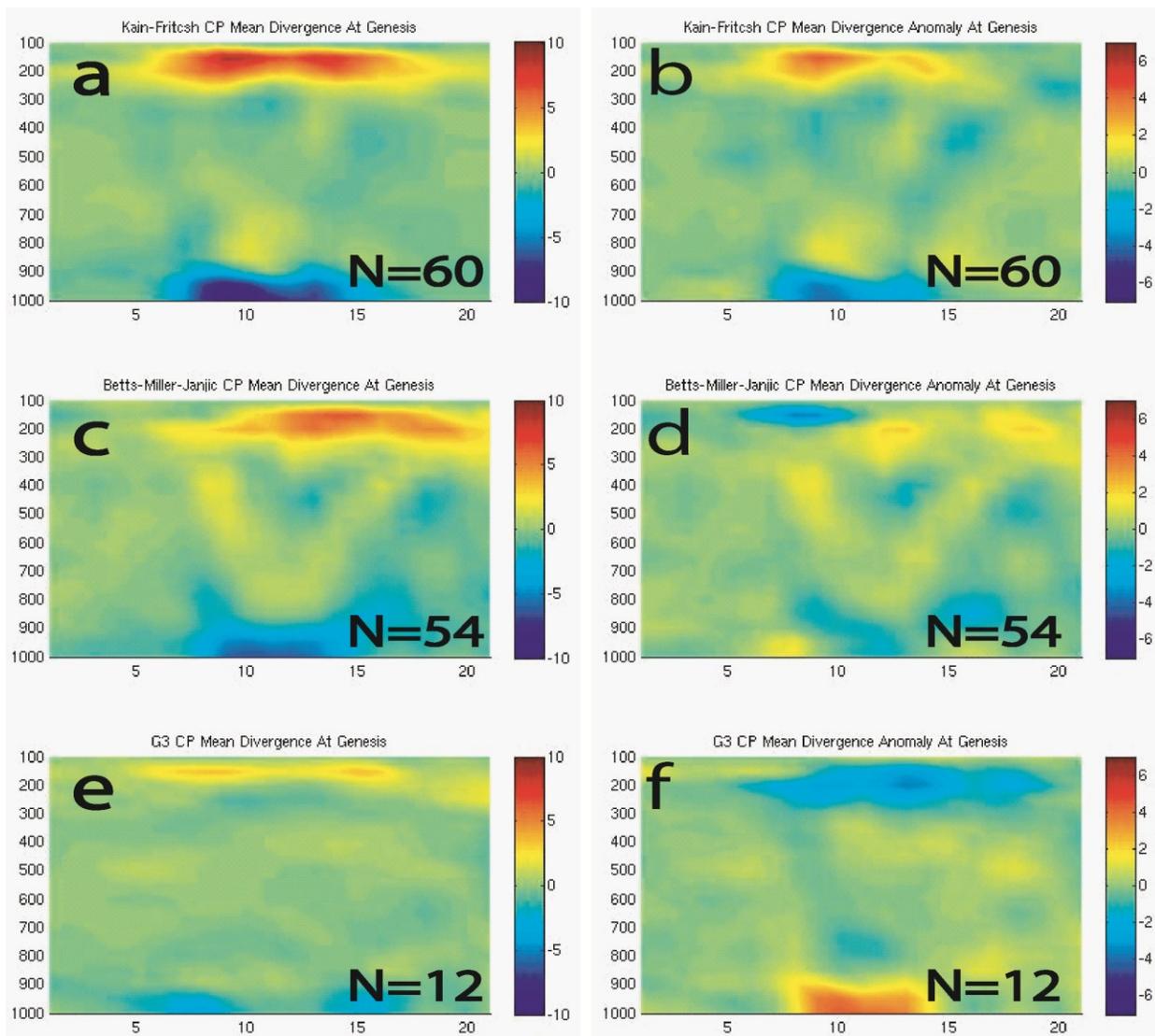
intense vertical motion over or near the cyclone center, with mean velocities around 2 m/s in the mid-upper atmosphere. Conversely, the G3 CP (**Fig. 3.13e,f**) produces peak updrafts which are weaker and further away from the center of circulation, which is consistent with a larger circulation and greater rain rates far from the center of circulation.

Similarly, **Figure 3.14** shows a cross-section of mean vorticity. The G3 CP cyclone (**Fig. 3.14e**) is characterized by weaker cyclonic vorticity near the storm center and larger vorticity far from the center when compared with the KF CP (**Fig. 3.14a**) or BMJ CP (**Fig. 3.14c**) cyclone. **Figure 3.15** shows mean divergence for the same time and locations as **Figures 3.13** and **3.14**, and shows the expected couplet of convergence at low-levels of the atmosphere and divergence at upper-levels of the atmosphere. The combination of **Figures 3.14** and **3.15** indicate vorticity generation through column stretching would generally be confined to near or over the center for the KF CP and BMJ CP cyclones given that low-level convergence is coincident with positive low-level vorticity. This is an effective combination for cyclone intensification. However, the maximum convergence locations within the G3 CP mean circulation are not coincident with the maximum vorticity with the G3 CP, meaning that (1) increasing vorticity associated with the low-level center will be slow to occur, and (2) vorticity generation also occurs at a distance from the storm center, which would act to increase storm size.

Rotunno and Emanuel (1987, RE87 hereafter) demonstrated that incipient vortices developed more slowly if the radius of maximum wind was larger than vortices with an equivalent maximum wind at smaller radii, which is consistent with the observed behavior shown here. They showed that although vortices with different radii of maximum winds would eventually reach an equal steady-state (all else being equal), a threshold value of roughly 160 km existed at which point the vortex would no longer be able to intensify. While there are certainly



**Figure 3.14** As in Figure 3.10, except depicting a zonal cross-section of vorticity ( $10^{-5}/s$ ).



**Figure 3.15** As in Figure 3.10, except depicting a zonal cross-section of divergence ( $10^{-5}/s$ ).

many differences between RE87's simplified model and the more complex WRF-ARW model used here, a number of parallels exist between their study and the behavior of the G3 CP cyclones after genesis. First, their cyclone experiments studying the importance of radius of maximum wind used a maximum tangential wind of 12 m/s which is effectively equivalent to the 25 kt threshold used here to determine genesis. Second, the 160 km radius threshold by which a cyclone did not intensify with time is also effectively equal to the radius of maximum wind associated with the G3 CP cyclones at genesis. Finally, of all twelve genesis cases when using the G3 CP, none developed to an intensity of greater than 31 kt during their lifetime or before the simulation ended. In RE87, the final intensity of their equivalently sized cyclone was approximately 15 m/s, which is quite consistent with the development of G3 CP cyclones. One additional impediment toward development of the G3 CP cyclones is the relatively large mean radius of deformation (due both to a larger size and slightly smaller maximum wind) associated with them relative to the comparatively smaller radii found in the KF CP and BMJ CP mean cyclone. A smaller radius of deformation would favor a more beneficial materialization of latent heating into kinetic energy with the latter two CPs.

Although the G3 CP is capable of producing stronger cyclones (including to hurricane intensity), this only seems possible when the initial conditions already contain a relatively compact system. For example, the G3 CP was eventually able to produce a TC with hurricane force winds in the case of Hurricane Bill when initialized with a smaller circulation than is depicted for the circulations at genesis with the G3 CP. A related problem is an increased inability to correctly predict TC genesis for real TCs. As shown in **Figure 3.9**, the peak in ACE associated with Hurricane Bill is delayed for the G3 CP. This is a function of the KF CP (and to a lesser extent the BMJ CP) developing the cyclone at extended forecast times before it was

officially declared a storm, which the G3 CP could not effectively do until the system was initialized in the model. A similar phenomenon can be seen with Hurricane Fred.

## CONCLUSION

B09 attempted to determine whether certain MPs and CPs exhibited systemic biases when forecasting TC track, intensity, or structure. This chapter continues that effort and extends it toward tropical cyclogenesis. Cumulatively, they demonstrate that certain parameterizations do display clear biases when analyzing these fields (or processes). The biases noted with convection (particularly over tropical oceans) with the G3 CP from Chapter 2 have been shown here to have significant impacts on TCs. Most prominently, difficulty with producing intense convection decreases the likelihood of TC genesis, and decreases the potential for intensification afterward. This is in part caused by larger circulations associated with the G3 CP and GD CP, which have the added effect of creating a poleward track bias with these cyclones when compared with other common CPs. Although not a surprising result, these conclusions reinforce the need for accurate depictions of TC circulations and distributions of heating in order to produce realistic forecasts within models. Additionally, it demonstrates that the ensemble nature of the GD CP and G3 CP poses significant challenges to TC prediction in their present configuration.

It was also found that the cause of track and intensity differences between the EF MP and the nEF MP were due to differences in the vertical distribution of potential vorticity. Although this is a problem unlike the above GD/G3 CP problem in that it has been largely solved in the current iteration of the EF MP, it does demonstrate the forecasting value of correctly distributed potential vorticity within a cyclone (even if the column total is correct). Additionally, it suggests potential utility in the development of a TC steering approach guided by the interaction between the magnitude of potential vorticity and the steering flow in that layer. Velden and Leslie (1991)

demonstrated that the depth of a vortex significantly impacted the choice of an optimal steering flow, so perhaps this approach could be extended to the distribution of potential vorticity.

Finally, virtually all of the findings within B09 and this chapter reinforce the necessity of parameterization studies which incorporate forecast lengths of greater than 48 hours. This is of particular importance for TCs, considering the obvious usefulness of forecasting them correctly as well as the fact most TCs are forecast out to 120 hours (or longer) among the primary forecasting agencies.

## **Chapter 4: Statistical Optimization**

### *Introduction*

As discussed in Chapter 1, ensemble forecasting is commonly used to understand uncertainty and improve forecasts. Frequently, ensembles are generated by varying one or more of the following parameters amongst ensemble members: initial conditions, parameterizations, grid spacings, or dynamical core. Two topics are discussed within this chapter pertaining specifically to parameterization-based ensemble forecasting (i.e. where choice of parameterization differentiates ensemble members). The first is within the context of a specific parameterization - the Grell-Dévényi cumulus parameterization (and the related G3 cumulus parameterization), which is an ensemble-based parameterization that frequently produces anomalous precipitation forecasts as shown in Chapters 2 and 3. Using tropical precipitation as an example, several statistical approaches will be discussed to improve these forecasts. The second topic considers a comparison between an optimized ensemble of forecasts produced by the Global Ensemble Forecasting System and a low resolution ensemble where members are differentiated by choice of parameterization. Linear regression techniques will be used to demonstrate that a parameterization-based ensemble contains additional information which, when leveraged properly, can allow such an ensemble to produce much more improved forecasts relative to the improvement shown when a similar technique is used with the Global Ensemble Forecasting System. The result is that the two ensembles produce comparable forecasts when optimized.

## STATISTICAL PRECIPITATION CORRECTION

### *A Statistical Approach For Correcting The Grell-Dévényi Cumulus Parameterization*

As described in Chapter 2 and Chapter 3, the Grell-Dévényi (GD) cumulus parameterization, as well as its updated version (G3), suffers from two general problems. First, forecasts using the GD CP frequently overproduce light precipitation and underproduce intense precipitation relative to other CPs, particularly over tropical locations. Second, this effect becomes magnified with forecast time due to the nature of the parameterization, which acts to enhance mid-troposphere drying and further decrease the model's (and parameterization's) ability to produce heavy rainfall rates relative to other parameterizations. Both of these problems stem from the ensemble nature of the parameterization, wherein 144 distinct members are averaged to yield feedback to the model on moisture and temperature tendencies as well as precipitation. In the original introduction of the parameterization (Grell and Dévényi, 2002), Grell and Dévényi suggest using a Bayesian approach whereby rainfall data over the first few hours of a simulation are assimilated and used to determine a proper weighting for each member.

However, despite being used in operational models such as the Rapid Refresh (RAP<sup>5</sup>) and its predecessor, the Rapid Update Cycle (RUC<sup>6</sup>), models, neither this approach nor any alternative approach has been implemented. Grell notes that primary reasons include lack of funding and manpower (Grell, personal communication, 2013). Specifically, in order to properly implement the approach outlined in his paper (or other similar approaches), a significant number of simulations need to be performed (each individual member needs to be run many times) while simultaneously having excellent observations of rainfall and the atmosphere in order to properly determine the appropriate weighting for each member. Additional problems include the

---

<sup>5</sup> A RAP technical note can be found at: [http://www.nws.noaa.gov/os/notification/tin11-53ruc\\_rapaae.htm](http://www.nws.noaa.gov/os/notification/tin11-53ruc_rapaae.htm)

<sup>6</sup> A description of the RUC can be found at <http://www.meted.ucar.edu/nwp/pcu2/rucintro.htm>

susceptibility to model resolution when determining weights as well as the issue of what observations to use to train the model (precipitation, vertical profiles of temperature or moisture, upward vertical motion, etc.). Finally, this approach becomes virtually impossible over oceanic environments, as high temporal and spatial observations of rainfall or atmospheric data are rare.

Given the above limitations, this work will focus on alternate methods of improvement, while allowing for future implementation in the GD CP or G3 CP. The data used will be the “real-time” simulations using a 30 km grid spacing described in Chapter 2 (along with additional simulations described below). Due to the inherent difficulty in obtaining high spatial and temporal resolution observational data over oceanic environments, all subsequent approaches will seek not to fit forecasts to observed data, but rather to fit the general characteristics of model output to observed characteristics of the real atmosphere. For example, it was noted in Chapter 2 that the cumulative distribution function (CDF) of the G3 CP both underforecast intense precipitation and also changed significantly with model forecast time. The approaches presented below seek to convert forecast model CDFs of a given field into realistic distributions of that field (additional detail will be provided below). For the purposes of this study, precipitation will be the variable analyzed, although in practice any variable could be used.

In the absence of a significant source of observational data, the approaches used below will use a substitute dataset in the form of rain rates estimated every three hours obtained using observations from the Tropical Rainfall Measuring Mission (TRMM) dataset available every three hours over tropical locations (Liu et al., 2012). These instantaneous rain rate estimates are delivered as one hour rain rates, and were locally converted to three hour rain rates. It should be noted that one of these three hour rainfall totals is *not* the equivalent of an actual rain gauge accumulating rainfall over an equivalent three hour period. While the latter would be more

desirable (and more accurate), the former is the best option available. TRMM data is available at  $.25^\circ \times .25^\circ$  resolution, which is comparable to the 30 km horizontal grid spacing used in the real-time forecasts. Data is obtained every three hours between 0000 UTC 1 June 2009 and 0000 UTC 1 November 2009 over a domain approximately equal to the real-time forecasts. In total, this constitutes 53,517,800 unique “observations”.

*Simplest Approach: Precipitation Reassignment Based On Climatology*

This approach follows a fairly simple supposition: given that the range of possible precipitation intensities in the model forecast is smaller than the range of precipitation intensities in reality, *given an infinite amount of model data, and an infinite amount of real data (over the same region), one could relate the most intense model output precipitation with the most intense observed precipitation.* That is to say, that if the greatest model forecasted precipitation over 3 hours was 15 cm (as an example) while in reality the greatest atmospheric precipitation over 3 hours is 50 cm, then we can say that in the event the model forecast 15 cm of precipitation over 3 hours, what it *should be* forecasting is 50 cm over 3 hours. Given a large dataset, one could construct two CDFs of precipitation – one for model data and one for TRMM-based precipitation rate estimate data. Once created, one could reassign the observed precipitation to the model forecast, using the same percentile for both. For example, if the model forecasts 4 mm of rainfall over 3 hours, one could look-up the percentile for this amount on the model CDF and reassign the precipitation to have a value equivalent to that percentile on the observed CDF.

Rather than implement this method directly, a slightly simplified method will be used. Instead of a continuous CDF (with infinite possibilities), a binning approach will be used. Nine different bins are chosen representing a range of values. While additional bins could be added to enhance effectiveness, this would decrease the number of observations per bin. Conversely,

fewer bins would unnecessarily increase the range within each bin. Nine bins is chosen as a compromise of these options. **Table 4.1** gives the range for each bin, along with the equivalent range of TRMM estimated precipitation rate for forecast hour 120 as an example. First, both precipitation datasets are sorted from lowest to highest rain rate. Using the example of the zero precipitation bin (bin 1), the percentage of zero precipitation model grid points is calculated, and an equivalent percentage of the lowest real rain rates is assigned to bin 1 (in this case, they are zero as well). Next, bin 2 is calculated, which includes non-zero model rainfall. The real precipitation values assigned to this range are zero, because there are more zero precipitation grid points in the real data (as can be seen by comparing row 2 and row 4 of **Table 4.1**).

However, as this process is continued to bin 4, it can be seen that model values between 0.01 mm/3h and .1 mm/3h are reassigned to a range from zero to 0.030 mm/3h. This process is continued for each bin, which consists of all rainfall data within the given range. A series of bins and ranges are created for every forecast time and saved. A comparison of the percentage of data falling into a given bin again demonstrates that the G3 CP produces excess light precipitation than the real data and less intense precipitation (as shown by bins 7, 8, and 9). This can also be seen in **Table 4.2**, which shows the percentage of total accumulated precipitation which falls for each bin relative to all precipitation. This is done by adding all rain rates in a given bin and dividing by the total of all rain rates. **Table 4.2** clearly demonstrates that the majority of total rainfall produced by the G3 CP is light (more than 89% of total rainfall falls as rain rates between 0.1 /3h and 10 mm/3h), while the real dataset produces the majority of its rainfall in the form of intense precipitation (it should be noted that TRMM frequently has difficulty observing very light precipitation, which likely skews the exact percentages somewhat, as shown by Berg et al. (2010) as well as Rapp et al. (2013)).

Bins	1	2	3	4	5	6	7	8	9
Model Range (mm/3h)	$X=0$	$0 < X < 0.01$	$0.01 \leq X < 0.1$	$0.1 \leq X < 1$	$1 \leq X < 5$	$5 \leq X < 10$	$10 \leq X < 25$	$25 \leq X < 50$	$50 \leq X$
% of Model Data per Bin	60.911	4.706	9.771	12.834	8.884	2.522	0.3633	0.0070	0.0009
TRMM Range	0	0	0	0 to 0.030	0.030 to 5.010	5.010 to 22.59	22.59 to 93.57	93.57 to 149.22	149.22 to 214.80
% of Real Data per Bin	69.949	5.238	10.615	9.262	2.847	0.7744	0.7254	0.3996	0.1889

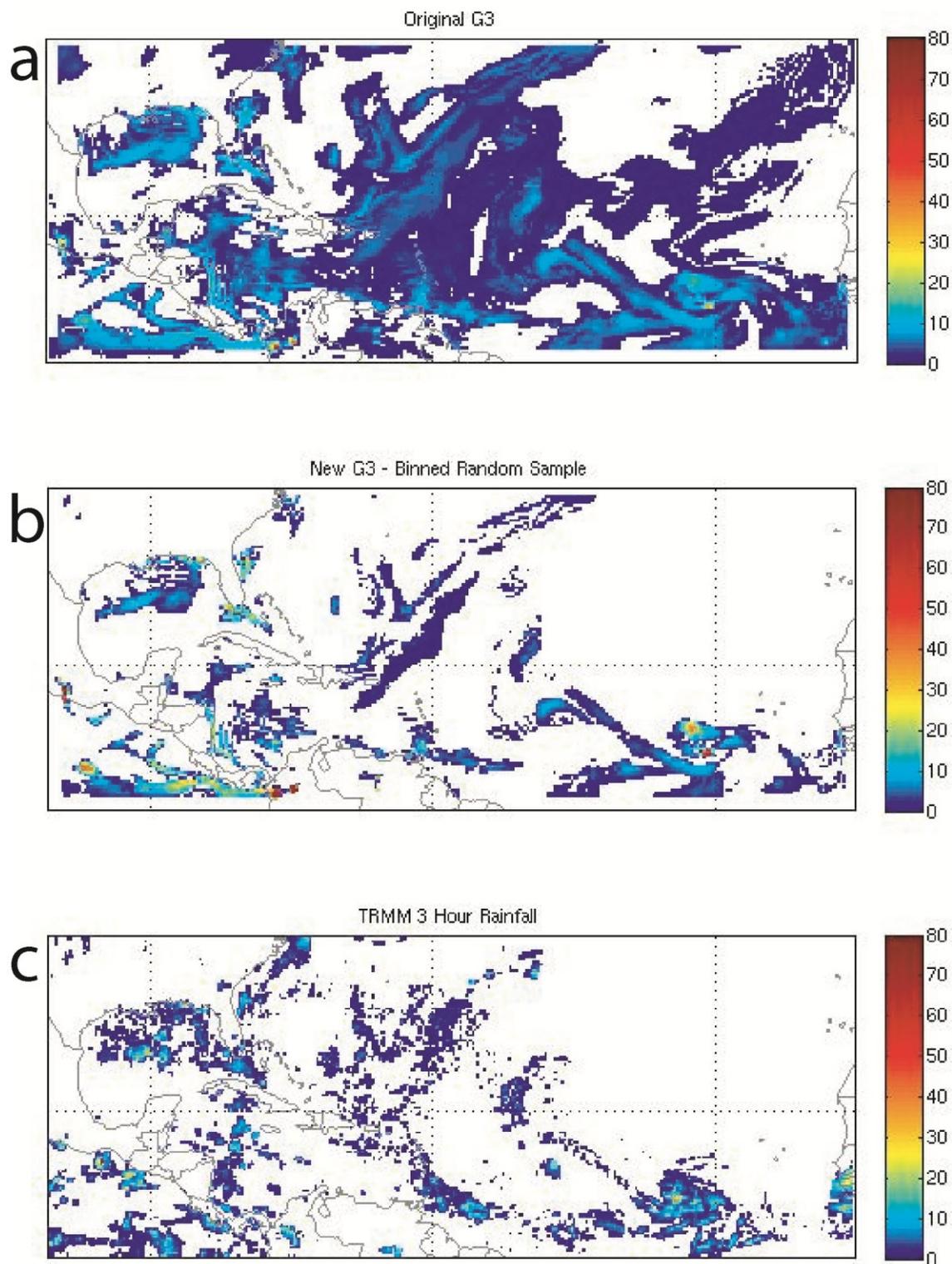
**Table 4.1:** This table details the range of rain rates prescribed to each bin within the G3 CP model at forecast hour 120, as well as the percent of total grid points which fall within each bin. Also shown is the equivalent range produced by TRMM observations if they were distributed according to the percentages shown per G3 CP bin. Finally, the percentage of TRMM observations which fall into the model bin ranges is shown.

Bins	1	2	3	4	5	6	7	8	9
Range (mm/3h)	X=0	0<X<0.01	0.01≤X<.1	0.1≤X<1	1≤X<5	5≤X<10	10≤X<25	25≤X<50	50≤X
% of Model Total Rainfall Per Bin	0	0.0314	0.9331	9.156	45.523	34.690	9.086	0.4562	0.1199
% of TRMM Total Rainfall Per Bin	0	0	0	11.505	22.897	12.259	17.947	17.338	18.055

**Table 4.2:** This table details the fraction of rainfall produced within each bin as a fraction of total liquid water produced across all bins for both the G3 CP at forecast hour 120 as well as all TRMM observations.

Once all bins are created, it becomes possible to correct any G3 CP forecast at any forecast hour. This is done by first determining how many grid points with rain rates within the bins shown in **Table 4.1** and **Table 4.2** exist. Once this is determined, a random sample of an equivalent number of rain rates is taken from the TRMM dataset in the given bin. For example, if a given forecast at hour 120 possessed 3000 grid points with rain rates falling in bin 5 (consisting of rain rates between 1 mm/3h and 5 mm/3h), a random sample of 3000 rain rates would be taken from bin 5 of the real dataset (which contains rain rates between 0.030 mm/3h and 5.010 mm/3h). After this is done for each bin, the model rain rates are reassigned to the newly derived rain rates by assigning the lowest model value to the lowest real value, continuing through the greatest model rain rate, which is assigned the value of the greatest real rain rate obtained in the sample. This method allows the forecast precipitation to be corrected while also maintaining some of the original forecast characteristics. For instance, forecasts with very little intense precipitation are not treated equivalently to those with significant amounts of precipitation, because the number of rain rates sampled per bin will vary (and may include zero samples taken from certain bins in some situations). The beneficial effects of this approach are best demonstrated with a series of examples from a given forecast.

**Figure 4.1** shows an uncorrected 12 hour forecast initialized at 0000 UTC 15 August 2009 (**Fig. 4.1a**), the corrected version using the above approach (**Fig. 4.1b**), as well as the closest TRMM observed rainfall (**Fig. 4.1c**). There are a number of interesting features to note. Most notably, the original forecast produces a rather unphysical large region of extremely light rainfall (approximately 1 mm/3hr or less) covering much of the tropical and subtropical Atlantic ocean. A comparison of the uncorrected and corrected plots immediately demonstrates that this problem is greatly ameliorated using this technique. Second, sea-breeze convection can be seen

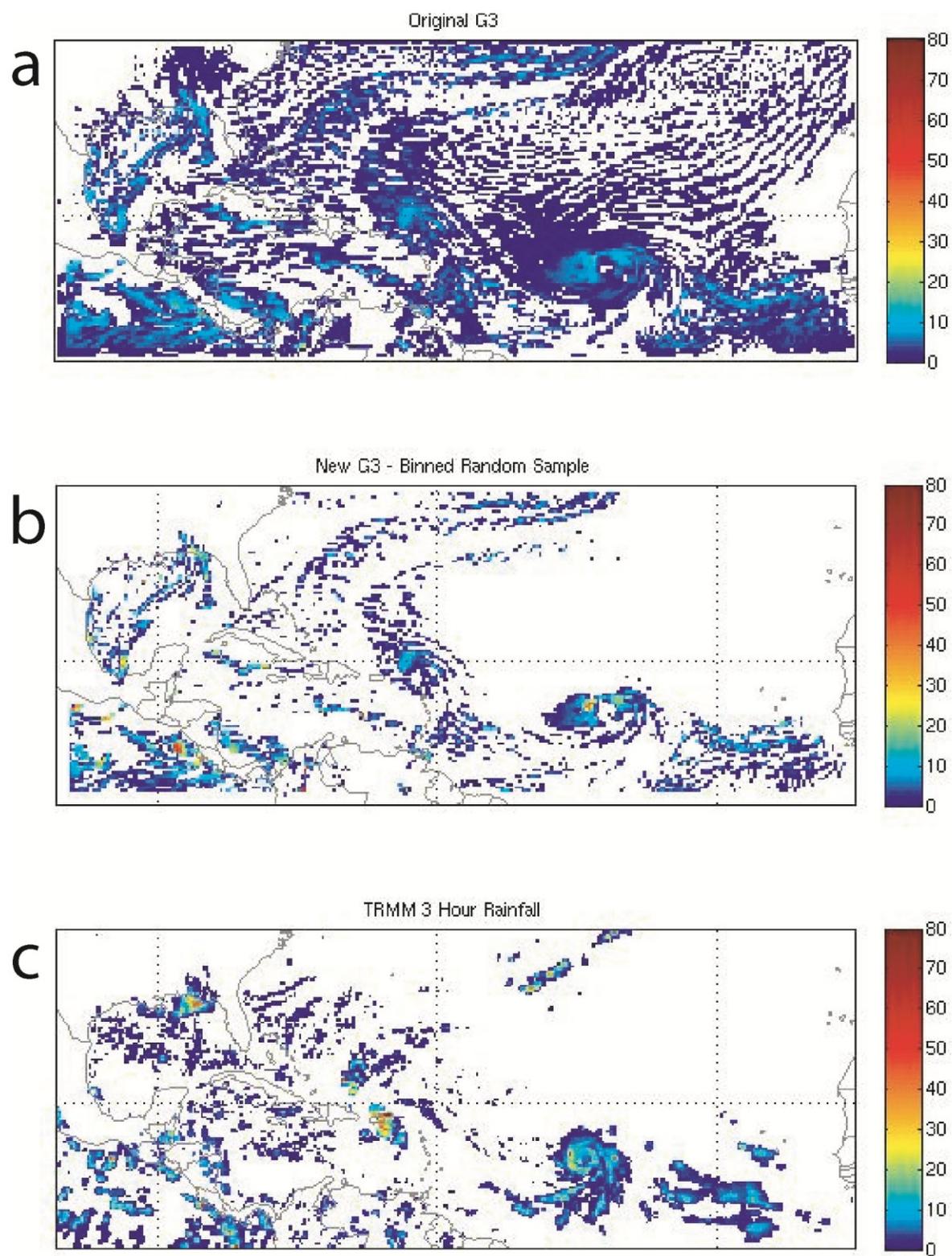


**Figure 4.1** Forecast three hour precipitation (mm) for forecast hour 12 for (a) uncorrected G3 CP, (b) corrected G3 CP, and (c) nearest corresponding TRMM data. Values less than .01 mm/3 h are not plotted. All forecasts initialized 0000 UTC 15 August.

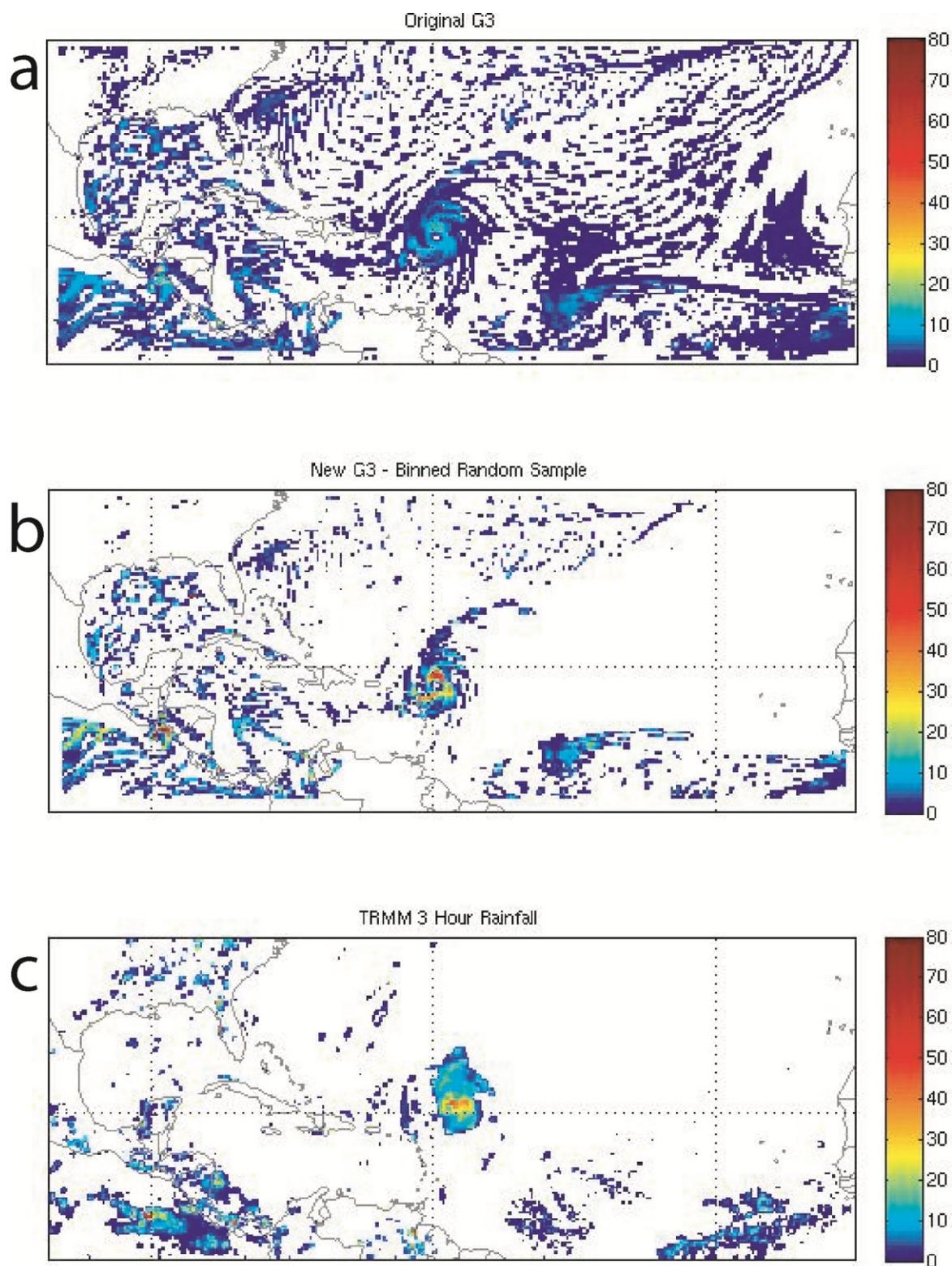
paralleling the coastlines of Central America as well as portions of Florida. The corrected forecast portrays larger rain rates in these locations, as would be more typical of sea-breeze convection seen by TRMM. Finally, the area of precipitation to the southwest of the Cape Verde Islands was recently declared a tropical depression that would eventually become Hurricane Bill (2009). As with the previous example, this technique produces enhanced rain rates in this region, as are commonly found with intensifying tropical cyclones. In this forecast, the system center was slightly north of the area of greatest rain rate.

**Figure 4.2** is similar to **Fig. 4.1** except for forecast hour 60. At this time the model continues to produce very light precipitation across a significant portion of the domain (**Fig. 4.2a**), albeit scattered across a larger area. It is immediately obvious that the corrected (**Fig. 4.2b**) forecast drastically reduces the quantity of very light precipitation, particularly across regions to the northwest of the Cape Verde Islands. The tropical cyclone noted at the previous forecast time can now be seen approximately equidistant between the west coast of Africa and the Lesser Antilles. While in reality this storm had just achieved hurricane status, in this forecast it is a borderline tropical storm (this underintensification can be ascribed in part to systemic problems with the G3 CP described in Chapter 3 concerning tropical cyclone forecasts). A comparison of the uncorrected and corrected forecast once again demonstrates an enhancement of rain rates very near the center of circulation, as denoted by the eye-like feature observed in the rainfall distribution (**Fig. 4.2c**).

At forecast hour 120, shown by **Figure 4.3**, noteworthy differences in the coverage of very light precipitation exist between the uncorrected and corrected forecast. At this forecast time the tropical cyclone has advanced to a position just east of the Virgin Islands. In actuality this cyclone was a powerful storm with maximum sustained winds of 115 kts. However, in this



**Figure 4.2** As in Figure 4.1 but for forecast hour 60.

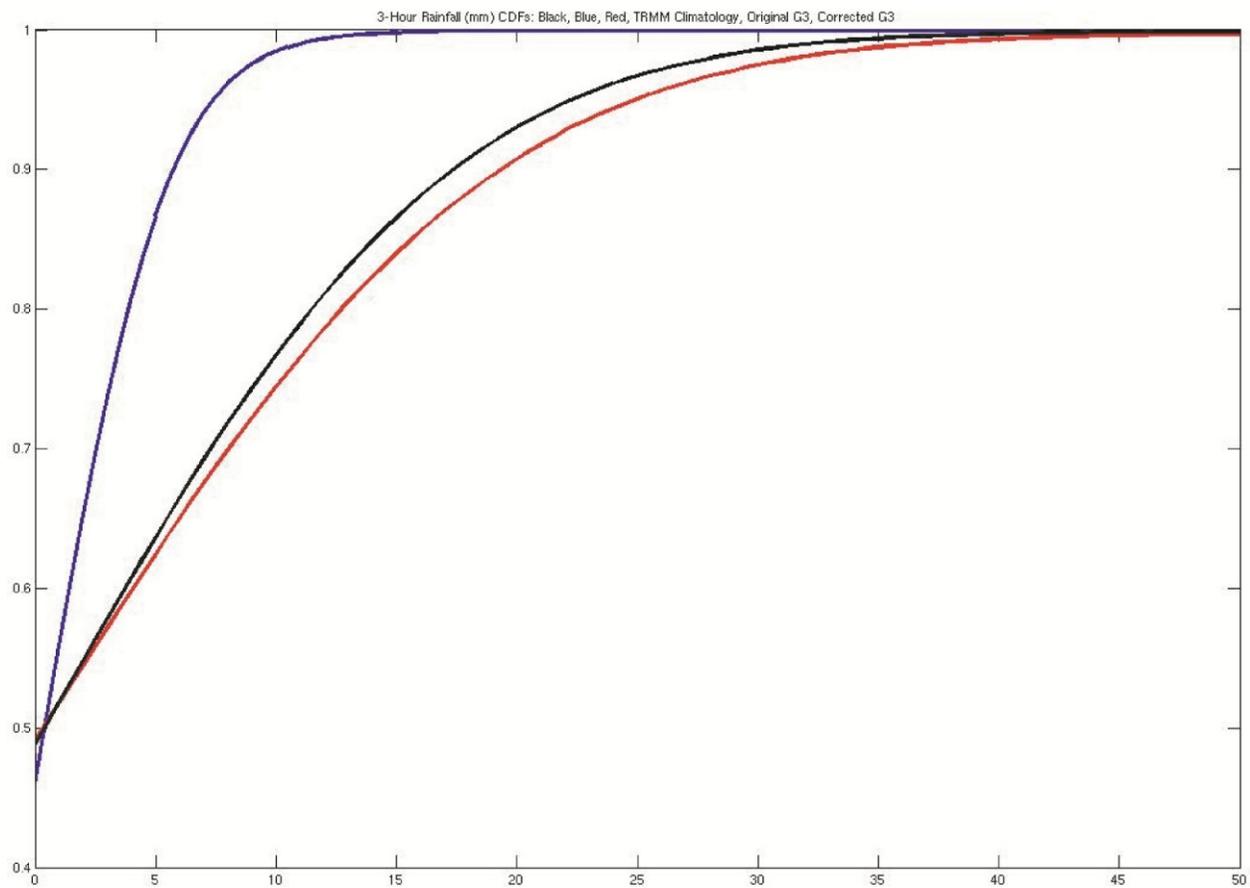


**Figure 4.3** As in Figure 4.1 but for forecast hour 120.

forecast the tropical storm was of moderate tropical storm intensity undergoing a modest intensification phase. It can be seen that maximum rain rates associated with this storm are roughly 20 mm/3 hr (**Fig. 4.3a**), while in the corrected forecast the rain rates near the center of circulation approach 80 mm/3 hr (**Fig. 4.3b**), which corresponds well to the observed rates using TRMM (**Fig. 4.3c**).

To demonstrate the impact this precipitation reassignment approach has on the CDFs of forecast rainfall, **Figure 4.4** was created. **Figure 4.4** displays the original and corrected G3 CP forecast rainfall for the above time as well as the climatological (2009 North Atlantic hurricane season) value of TRMM precipitation for comparison. A comparison of the two G3 CP forecasts demonstrates several important features. First, very light values of precipitation are reduced in frequency in the corrected forecast compared with the original forecast. Second, the frequency of intense rainfall rates is dramatically increased. Finally, it can be seen that the CDF of corrected precipitation is similar, yet not identical, to the CDF of TRMM climatology for this reason. This discrepancy is due to the fact this approach does *not* simply randomly sample the TRMM data, but rather uses a binning approach in order to retain some character of the original distribution (i.e., whether a given forecast contained more or less rainfall than normal). If this approach used a pure random sample, one would expect the corrected G3 CP CDF and TRMM climatological CDF to perfectly overlap. The fact that the corrected G3 CP CDF indicates a greater frequency of intense precipitation is due in part to the presence of a reasonably intense tropical cyclone, which includes a slightly greater amount of intense rain rates in the unaltered G3 CP forecast than is representative of the entire season within the G3 CP.

Although this approach is presented here using precipitation data from forecasts using the G3 CP, it should be noted that theoretically this can be used with any model forecast and for any



**Figure 4.4** Cumulative distribution functions of TRMM climatology (black), precipitation shown in Figure 4.3a (blue) and Figure 4.3b (red).

variable, provided a sufficient amount of prior data exists such that the statistical framework this is based upon can be constructed.

*Grell Cumulus Parameterization: Linear Regression Approach*

As noted earlier, the G3 CP utilizes a 144 member ensemble of individual cumulus parameterizations, which are weighted equally to produce a final result. This frequently yields unphysical depictions of tropical rain rates, as shown in the previous section as well as prior chapters. For reasons described earlier, it is not possible to employ the Bayesian approach suggested by the authors due to lack of observational data as well as a lack of computer resources. A substitute to this in keeping with the authors' original intent will be used here. The original 144 members broadly fall into four classes of cumulus parameterization closure: moisture convergence types, CAPE relaxation types, Arakawa-Schubert types, and omega-triggered types (all of which are described in Grell and Dévényi (2002), as well as the sub-permutations such as precipitation efficiency, entrainment rate assumptions, etc.). Although it is not feasible within the scope of this project to analyze each member, four representative members are chosen for analysis (representing each closure variety mentioned above) with the assumption that these four represent an approximate range of the total distribution of ensemble members. For each of the 76 real-time forecasts used for the above G3 CP precipitation correction, four forecasts were recreated using a representative member of each closure as choice of CP.

As with the prior example, rather than attempt to fit the forecast precipitation directly to the spatial distribution of observed precipitation, the goal was to produce forecast weightings such that the distribution of forecast precipitation more closely aligned with the characteristics of observed precipitation, as above. However, rather than using precipitation reassignment, this

approach will determine optimal weights for the four closures using linear regression, in the form:

$$\boldsymbol{\beta} = (\mathbf{X}' * \mathbf{X})^{-1} * (\mathbf{X}' * \mathbf{Y}) \quad (1)$$

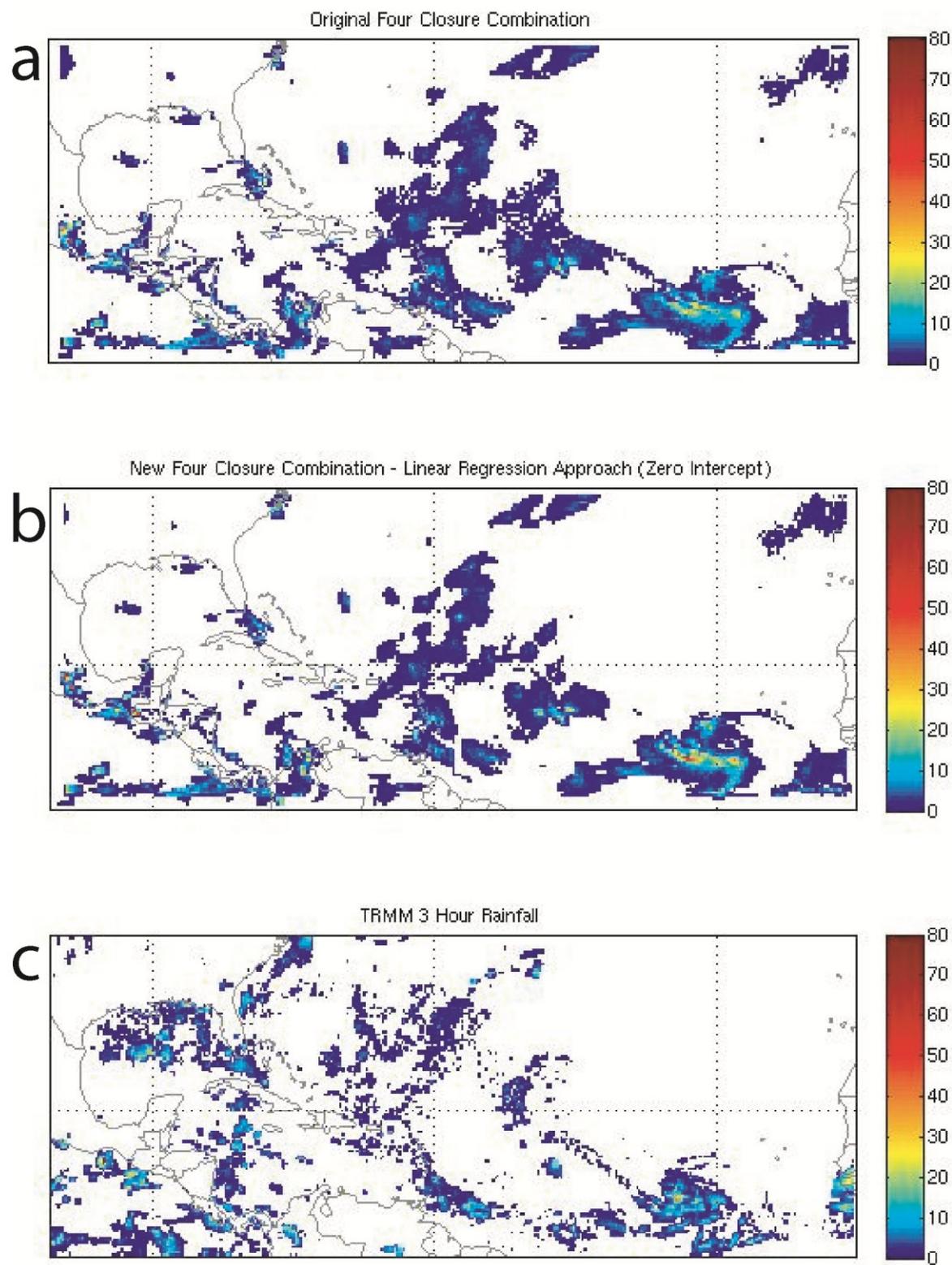
Here  $\boldsymbol{\beta}$  represents a vector of coefficients of length equal to the number of predictors (and intercept, if included). Once it is obtained, an optimal combination ( $\mathbf{Y}_i$ ) for a given forecast can be calculated by:

$$\mathbf{Y}_i = \boldsymbol{\beta}_1 \mathbf{X}_{i1} + \dots + \boldsymbol{\beta}_n \mathbf{X}_{in} + \mathbf{I} \quad (2)$$

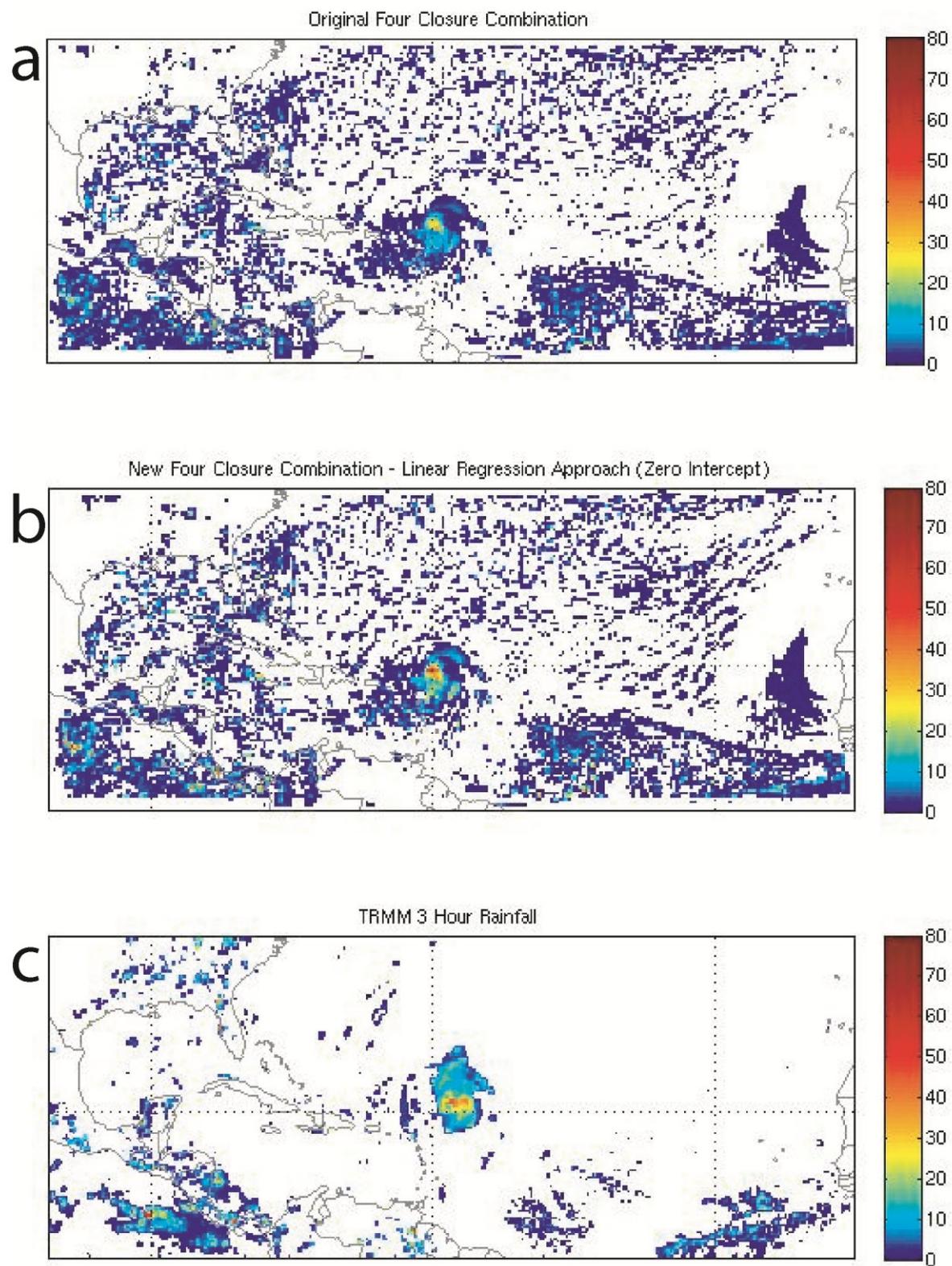
For each forecast time, a vector of precipitation forecasts consisting of each closure's forecast as well as the mean of the four is created. This is done for each grid point across all 76 forecasts, and combined to form one 2328336 x 5 matrix. Next the matrix is sorted from lowest mean precipitation value to greatest mean precipitation value. Afterward, the mean precipitation column is removed, resulting in a 2328336 x 4 matrix, which is  $\mathbf{X}$  in equation 1 (with the four columns representing four predictors). The TRMM data is used again as the observed dataset. A random sample (of size 76 x 30636) is taken from this dataset and sorted from lowest value to highest value, which represents  $\mathbf{Y}$  in equation 1. Subsequently, linear regression is performed to determine the optimal weights to apply to each closure. It should be noted that no intercept parameter is calculated ( $\mathbf{I} = 0$ ), due to the very frequent likelihood of precipitation forecasts being 0 mm/3h (and the non-physical nature of negative precipitation in the event of a negative intercept parameter and zero forecast precipitation by each closure). This technique generally follows the simple assumption that the heaviest model forecast precipitation should most often occur when the mean of the ensemble members' predicted precipitation is greatest. The above procedure is performed for each forecast time between hour three and 120 to produce optimal coefficients which are unique to each forecast time.

For comparison purposes, the forecast case presented earlier initialized at 0000 UTC 15 August, will be used as an example of this technique as well. **Figure 4.5a** shows the mean four-closure precipitation, the optimally weighted precipitation forecast (**Fig. 4.5b**), and the same TRMM precipitation (**Fig. 4.5c**) shown in **Figure 4.1**. Although the mean of the four closures is similar to the standard G3 CP forecast shown in **Figure 4.1**, the coverage of very light precipitation is slightly reduced while the maximum precipitation rates are increased. This is due to the fact there are fewer ensemble members in this forecast combination. Skipping ahead to forecast hour 120 depicted in **Figure 4.6**, it is immediately apparent that while this linear regression approach generally correctly increases maximum rain rates (as depicted by the warm colors near Hurricane Bill), this approach does nothing to decrease the widespread overproduction of light rainfall. Barring a negative or zero weight for any closure, this will always be true in the absence of an intercept parameter.

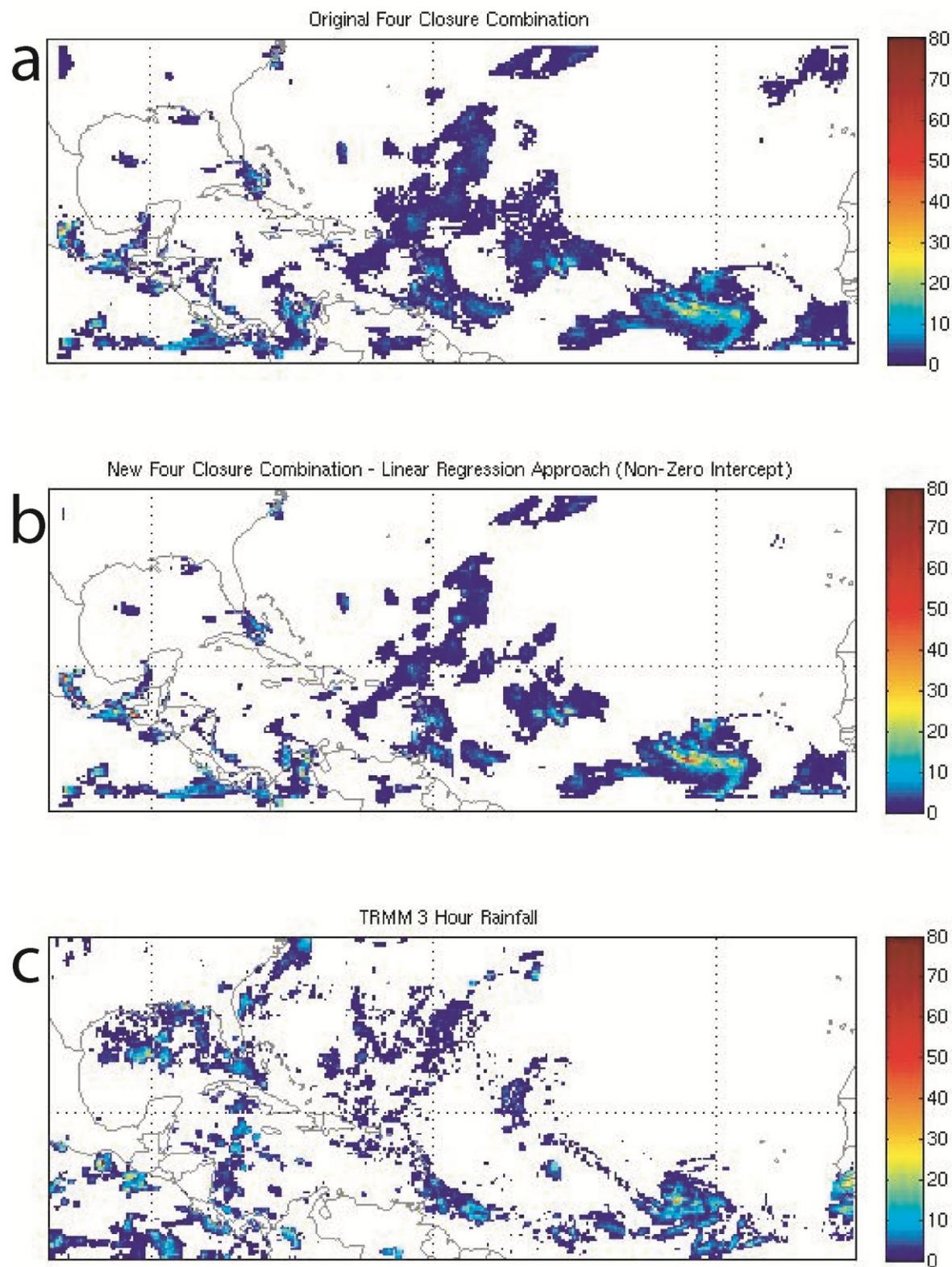
**Figures 4.7** and **4.8** reproduce **Figures 4.5** and **4.6**, except with recalculated coefficients without constraining the intercept to zero. This allows a negative intercept, which reduces the light precipitation bias. It should be noted again that TRMM likely tends to underestimate or miss very light precipitation, so it is difficult to know what constitutes the “correct” amount of light precipitation coverage. Another option is to incorporate both of the above approaches. This is done by first using linear regression (as above, with a zero intercept) and following with a reassignment approach as described in the first approach. This method both allows for more fidelity regarding appropriate weights for the various closures while also removing to a large extent the light precipitation bias. This can be seen in **Figure 4.9**, which uses this technique for the same 120 forecast shown in **Figures 4.6** and **4.8**.



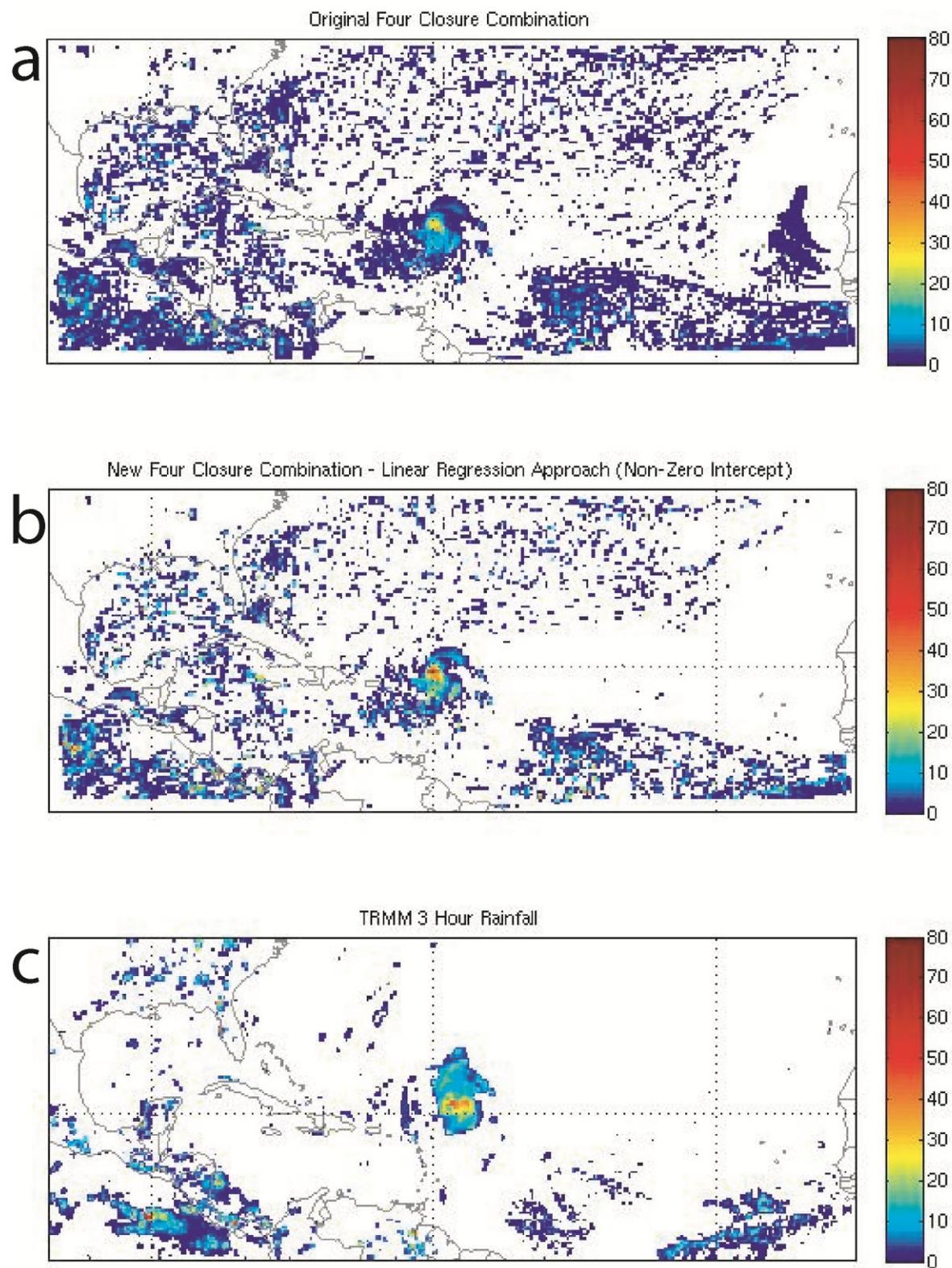
**Figure 4.5** As in Figure 4.1 but for (a) a mean of four G3 CP forecasts and (b) an optimal combination of four G3 CP forecasts constrained with a zero intercept.



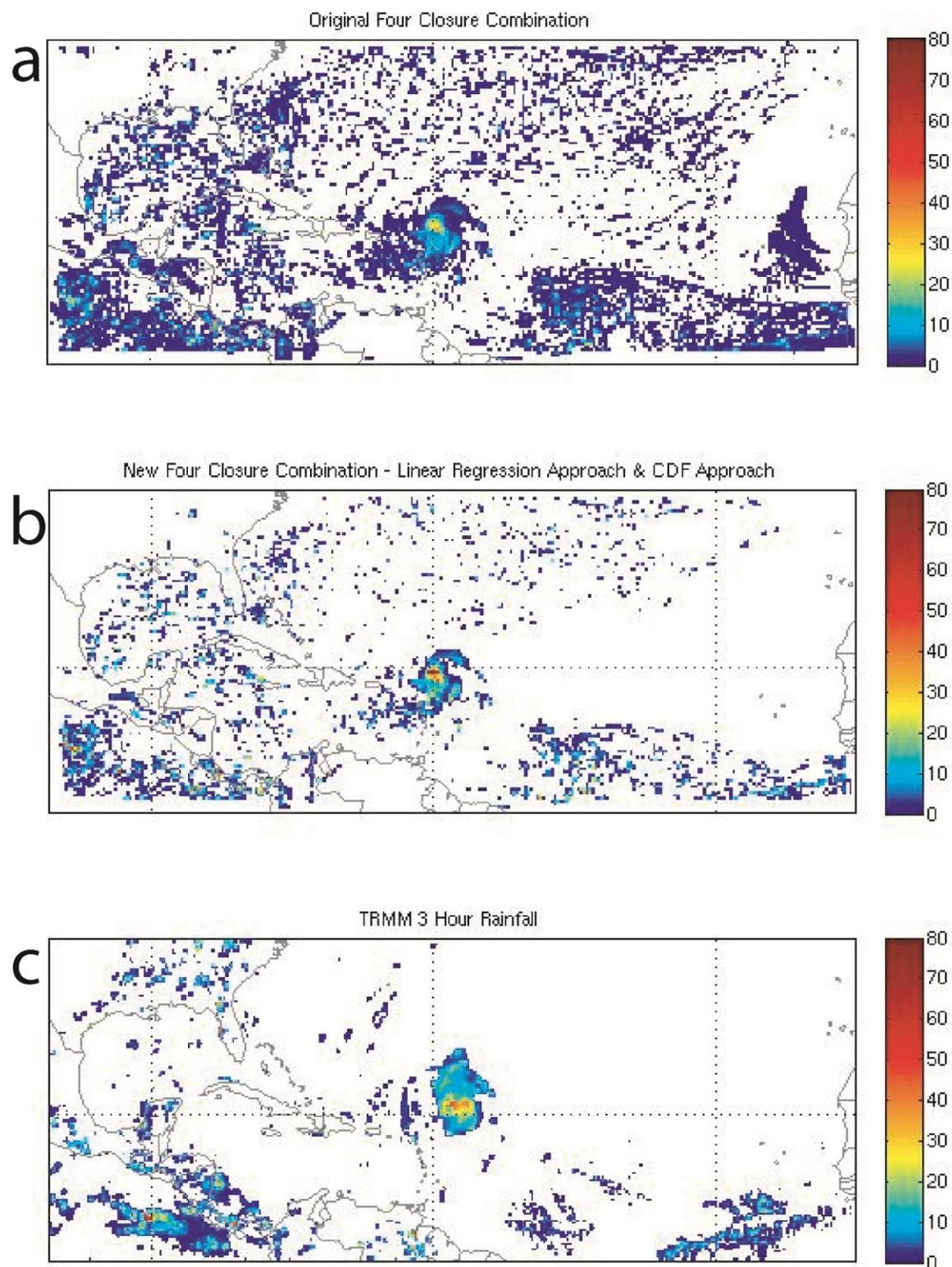
**Figure 4.6** As in Figure 4.3 but for (a) a mean of four G3 CP forecasts and (b) an optimal combination of four G3 CP forecasts constrained with a zero intercept.



**Figure 4.7** As in Figure 4.1 but for (a) a mean of four G3 CP forecasts and (b) an optimal combination of four G3 CP forecasts not constrained with a zero intercept.



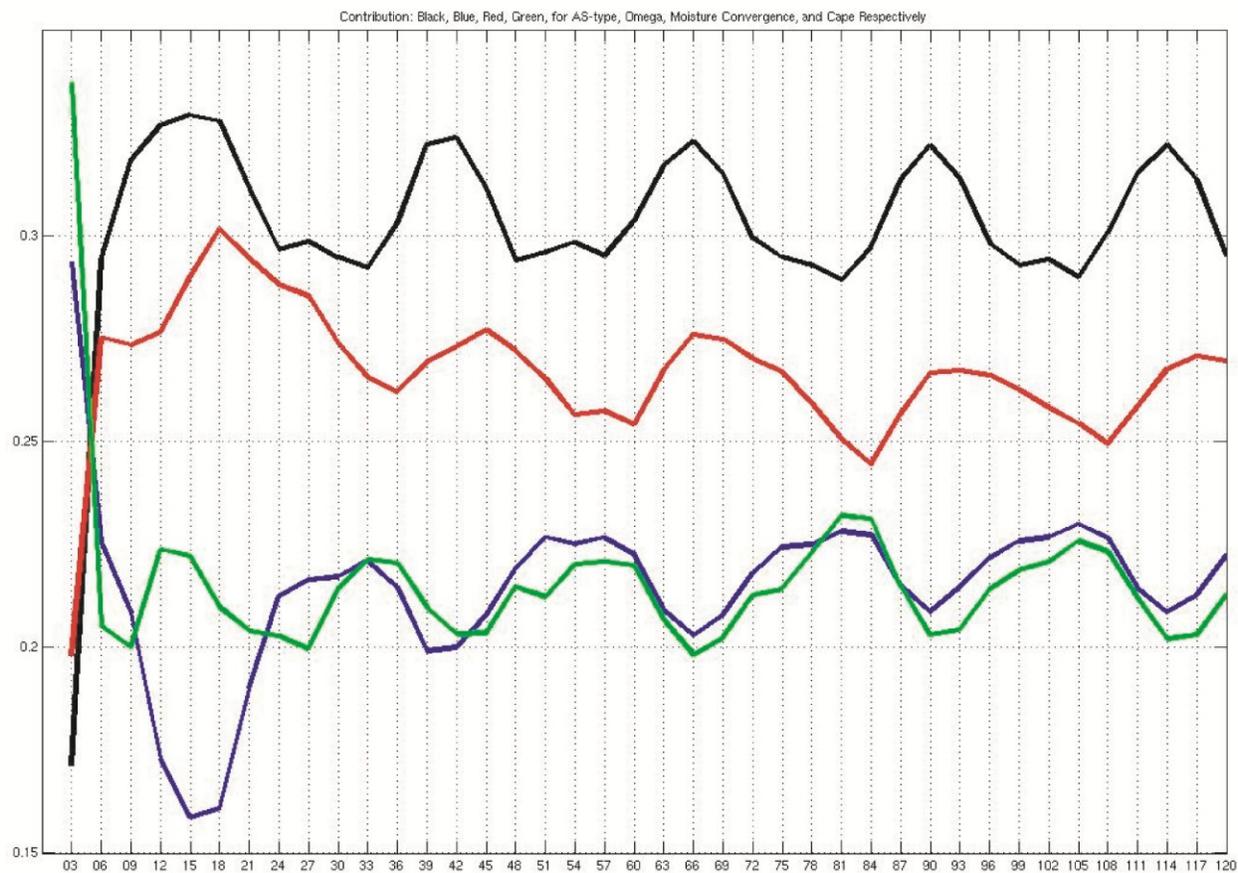
**Figure 4.8** As in Figure 4.3 but for (a) a mean of four G3 CP forecasts and (b) an optimal combination of four G3 CP forecasts not constrained with a zero intercept.



**Figure 4.9** As in Figure 4.3 but for (a) a mean of four G3 CP forecasts and (b) an optimal combination of four G3 CP forecasts not constrained with a zero intercept followed by precipitation reassignment.

### *Analysis Of Weights*

Since the ensemble member weights are calculated independently at each forecast hour, it is possible to analyze the relative importance of each member as a function of forecast hour. **Figure 4.10** displays the relative contribution of each member by dividing each member's weight by the total of all weights for a given forecast hour. One obvious observation is that a prominent diurnal cycle emerges, which is nearly steady after approximately forecast hour 30. This somewhat corroborates the finding in Chapter 2 that the precipitation characteristics associated with the G3 CP do not become approximately steady-state until after forecast hour 24. The four closure types are distinctly paired in their behavior – the Arakawa-Schubert type and moisture convergence type are most influential during the afternoon hours of forecasts, while the CAPE-removal type and omega type are relatively more important during the nighttime hours of the domain. Several factors likely contribute to the patterns seen here. The Arakawa-Schubert type closure is strongly related to large-scale buoyancy, so it is perhaps intuitive that it would be of greatest relative importance during the afternoon hours when buoyancy is often maximized. Its position of greatest weight at nearly all times is likely due to its enhanced ability to produce intense rainfall relative to the other closures (not shown). Although the moisture convergence closure is also of relatively great importance, its slow decline after the first 6 to 30 hours of model integration is probably a result of the modeled atmosphere slowly becoming dryer with time, as described in Chapter 2. Since the precipitation produced in this closure is proportional to the vertical integral of moisture convergence, as the mid-level atmosphere dries this vertical integral would be expected to decrease with time, producing less intense rain rates with time (as described in Chapter 2).



**Figure 4.10** Percent contribution (with forecast hour) to an optimal combination with Arakawa-Schubert type (black), Omega type (blue), Moisture Convergence type (red), and Cape removal type (green).

*Strengths And Weaknesses Of Both Approaches*

The primary strength of both approaches is their extreme computational efficiency. Once the climatological precipitation bins and closure weights are created for the precipitation reassignment and linear regression approaches are created, improved precipitation forecasts can be created within seconds. The clear weakness of both approaches is that they occur outside of any model integration; therefore any enhancement of model accuracy resulting from a (presumably) more realistic depiction of convection goes unrealized. Intuitively, one could alter the existing G3 CP cumulus heating, drying, and other atmospheric variable profiles with the weights described above. However, it is not certain that doing so within the parameterization would produce a superior or realistic result. While precipitation is an observable model variable with obvious societal and economic importance, it is also generally a final product of a cumulus parameterization integration. Given this, determining appropriate weights based on this final product may not be optimal. For example, perhaps variables such as the vertical profiles of temperature and moisture would be more appropriate for calculating weights rather than precipitation. The difficulty with this approach using the G3 CP is that the various ensemble members determine precipitation based upon different criteria. This leaves precipitation as the primary variable to use for weighting, which as stated earlier may be problematic. It should be noted that Grell and Dévényi (2002) used their optimal weighting approach to improve precipitation forecasts rather than parameterization feedback to the model.

Although a linear regression approach ameliorates the G3 CP's difficulty with intense precipitation, at best it is neutrally effective at reducing widespread spurious light rainfall. Barring a zero or negative coefficient for any ensemble member, this is true by definition. For that reason, another statistical approach might be optimal. The simplest option is to simply

combine the above approaches – use the new member weighting to produce a corrected precipitation, and then apply the precipitation reassignment approach to that corrected precipitation. However, this still does not address potential changes to real-time modeled interactions the convection has with the grid scale variables. One option to address this issue would be to first run the parameterization as is at a given time-step, then determine what the corrected precipitation value should be based upon either above approach. Afterward, the parameterization could be rerun with the appropriate weighting applied to any heating, cooling, moistening, or drying predicted by that ensemble member. In the event that the corrected precipitation changes from a positive value to a zero value, the parameterization could be deactivated for that particular grid point. While an attempt at this approach is beyond the scope of this work, this method would effectively cause the CP to be called twice per time-step, which would certainly increase the length of model integration. However, assuming no significant destabilization of the model results, this approach could presumably lead to improved forecasts.

Another obvious strength of either approach is the fact that they do not treat all forecast hours as equal. Since it has been demonstrated that the precipitation characteristics of the G3 CP change with forecast time, forecast time-dependent approaches are preferable. Although this may be possible to incorporate within the parameterization itself, it may be difficult to account for both diurnal variations as well as factors related to the time since model initialization (i.e. whether it is forecast hour 12 or 120). Finally, either of these approaches are effective for any model configuration and virtually any location on the globe at any time of year, provided a sufficiently large database of climatological values exists.

*Examination Of Forecast Differences*

Although incorporating these approaches within a parameterization is outside the scope of this study, some brief speculation of the likely impacts is appropriate within the context of previous work. For instance, Chapter 2 showed that the G3 CP produced weaker (yet more frequent) convection over tropical locations relative to other CPs, which ultimately led to a weaker Hadley Cell circulation. Either of these approaches could help ameliorate excess precipitation produced by the G3 CP over subtropical regions while enhancing the intensity of convection within the ITCZ region, which may help resolve the weak Hadley Cell problem with this CP.

Chapter 3 demonstrated that the GD CP was less inclined to produce tropical cyclogenesis due to a larger circulation when compared with other CPs. This large circulation was the result of more peripheral convection (rather than intense near-center convection). Similarly, it was shown that the GD CP was less able to intensify TCs due to a similar problem – latent heating was spread over a large circulation rather than concentrated near the center of circulation. As shown earlier in this chapter, when either of these techniques are applied to a G3 CP forecast including a hurricane, the result is to (1) decrease light rainfall distant from the storm center, and (2) dramatically enhance convection near the center. Given the problems described earlier, if a similar outcome occurred upon implementation within the parameterization, it is likely that the GD CP and G3 CP would be more able to generate TCs and more able to realistically intensify them after genesis. Similarly, a lack of peripheral convection (and associated latent heating) would likely decrease the generated mean circulation size, which may decrease the notable poleward track biases produced by these CPs shown in Chapter 3.

## A COMPARISON OF TWO OPTIMIZED ENSEMBLES

Global forecast models such as NCEP's Global Forecast System (GFS) employ ensemble techniques to improve an understanding of the potential future state of the atmosphere. The Global Ensemble Forecast System (GEFS) is a collection of twenty forecasts run in conjunction with the official GFS control forecast performed at a lower resolution and differentiated by choice of initial conditions. The variance within these initial conditions is designed to represent analysis uncertainty, with differences constrained to be orthogonal to one another (Wei et al., 2008). Afterward, stochastic perturbations are applied to the different initial conditions<sup>7</sup>. The discussion below will compare the benefits of optimizing an ensemble such as this (to produce one superior forecast) to the benefits of optimizing an ensemble comprised purely of members differentiated by choice of parameterization (while using the same initial conditions).

### *Methodology*

Two ensemble datasets representing two different forms of ensemble generation will be compared. A parameterization ensemble constructed of 10 members with perturbed physics initialized using the 0000 UTC GFS control simulation for initial and lateral boundary conditions represents the parameterization ensemble. This is also the outer domain of the real-time forecasts described earlier, and uses a 90 km horizontal grid spacing and 28 vertical levels. The various parameterizations used in each member are shown in **Table 4.3**. The first nine GEFS ensemble members as well as the control simulation with a resolution matching the other nine GEFS members represent the initial condition ensemble. All members use the same combination of parameterizations while each member uses uniquely perturbed initial conditions. Consequently, each ensemble represents a significantly different way of generating model spread. It should be

---

<sup>7</sup> Additional information can be found at [http://www.dtcenter.org/ensemble\\_presentations/2-10\\_Toht-NAEFS\\_GEFS.pdf](http://www.dtcenter.org/ensemble_presentations/2-10_Toht-NAEFS_GEFS.pdf)

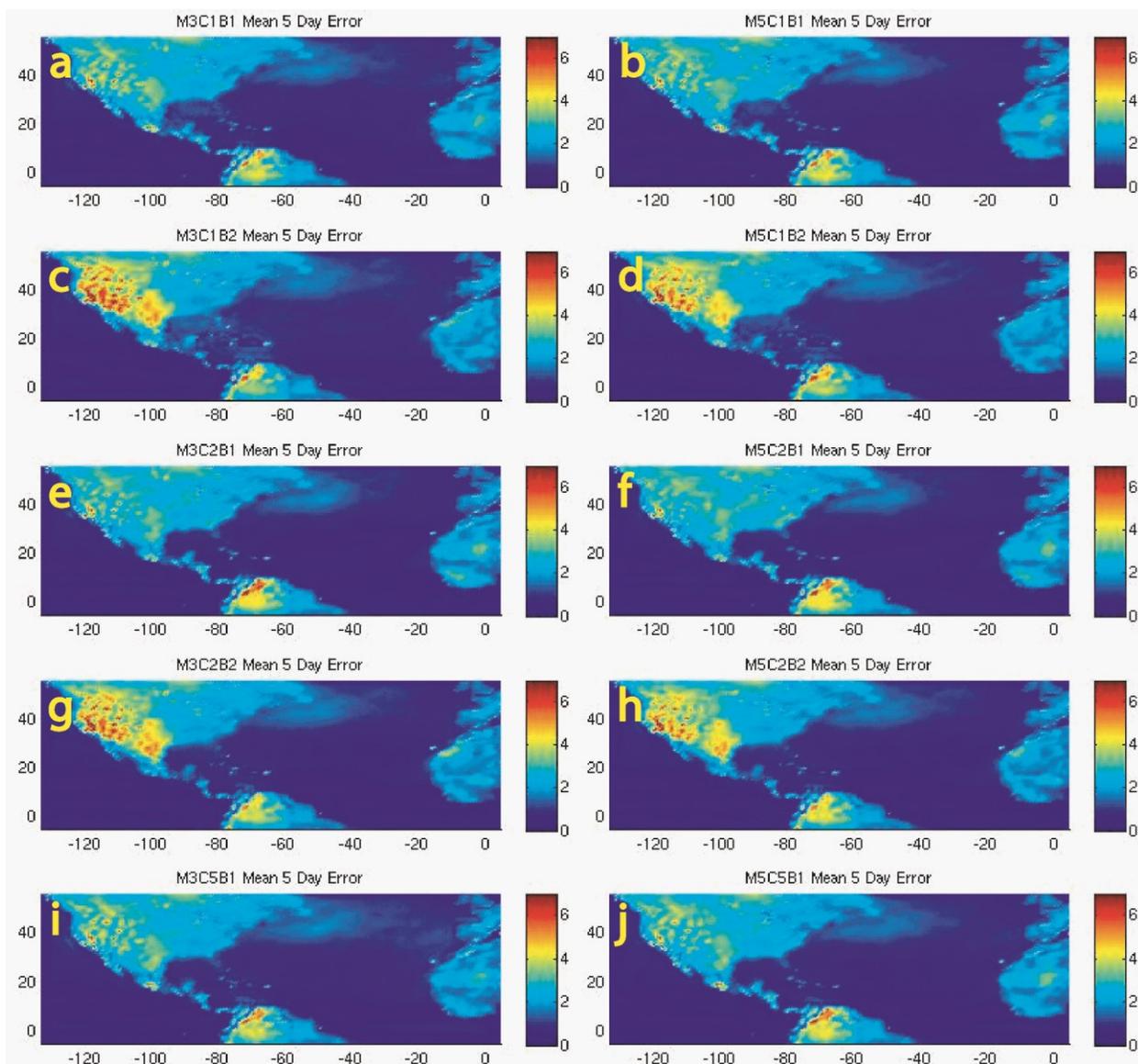
Member	Microphysics Parameterization	Cumulus Parameterization	Boundary Layer Parameterization
1	WSM3	Kain-Fritsch	Yonsei University
2	WSM3	Kain-Fritsch	Mellor-Yamada-Janjic
3	WSM3	Betts-Miller-Janjic	Yonsei University
4	WSM3	Betts-Miller-Janjic	Mellor-Yamada-Janjic
5	WSM3	Grell-3	Yonsei University
6	Eta-Ferrier	Kain-Fritsch	Yonsei University
7	Eta-Ferrier	Kain-Fritsch	Mellor-Yamada-Janjic
8	Eta-Ferrier	Betts-Miller-Janjic	Yonsei University
9	Eta-Ferrier	Betts-Miller-Janjic	Mellor-Yamada-Janjic
10	Eta-Ferrier	Grell-3	Yonsei University

**Table 4.3:** The parameterizations for each ensemble member within the parameterization ensemble is shown.

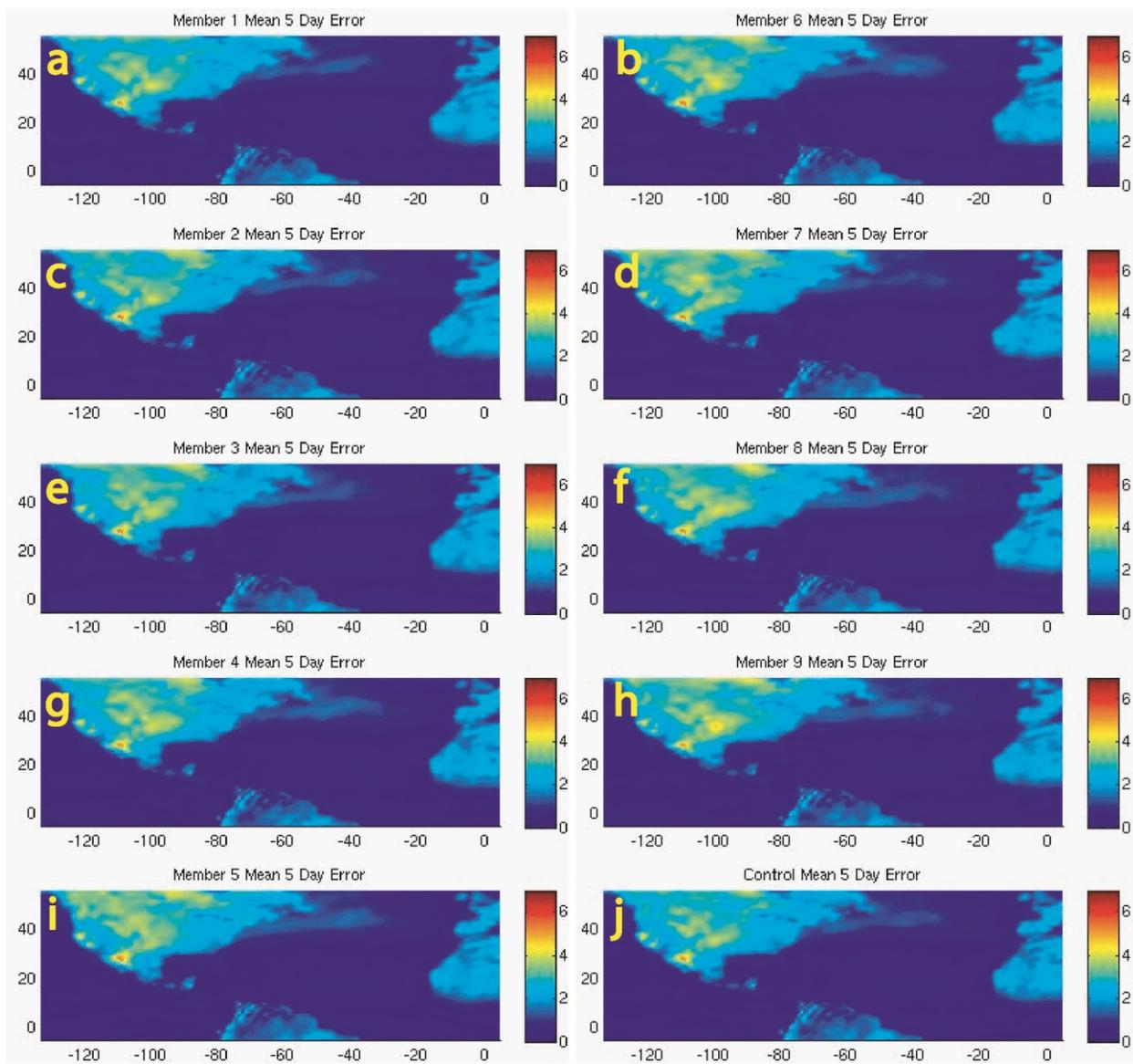
noted that the effective horizontal resolution of the GEFS is approximately 30% greater than that of the parameterization ensemble, and would not suffer from any potential lateral boundary errors considering it is a global model.

For each ensemble and for a given variable, a matrix of 76 five-day forecasts are created, representing one every two days of the 2009 North Atlantic Hurricane season (this is  $\mathbf{X}$  in equation 1 shown earlier). For each matrix, forecast hour 120 is analyzed using the native grid of the parameterization ensemble; the higher resolution initial condition ensemble is mapped onto this grid. Finally, for every grid point linear regression is used to determine an optimal weighting for each member of both ensembles such that a linear combination of the ten members within a given ensemble produce a single best forecast. Truth is considered to be the valid 0000 UTC initial conditions from the GFS control simulation mapped onto the same 90 km grid (this is  $\mathbf{Y}$  in equation 1 shown earlier). Once appropriate weights are determined for each ensemble at each grid point, a single forecast is recreated for each of the 76 cases for each ensemble. It should be noted that traditionally two independent samples are created, whereby the statistics generated from a training sample are tested on the other sample. However, since the same method is used for both the parameterization ensemble and the initial condition ensemble, this is deemed unnecessary since both use the identical statistical analysis, making comparisons between the two impartial.

This process can be performed for virtually any variable and statistics easily calculated. **Figures 4.11 and 4.12** show, respectively, the mean 2 m temperature error ( $^{\circ}\text{C}$ ) per forecast (without any form of optimization) for each member of the physics ensemble and initial condition ensemble. A comparison of the two shows that generally the physics ensemble members have greater mean errors than the initial condition ensemble members, particularly over



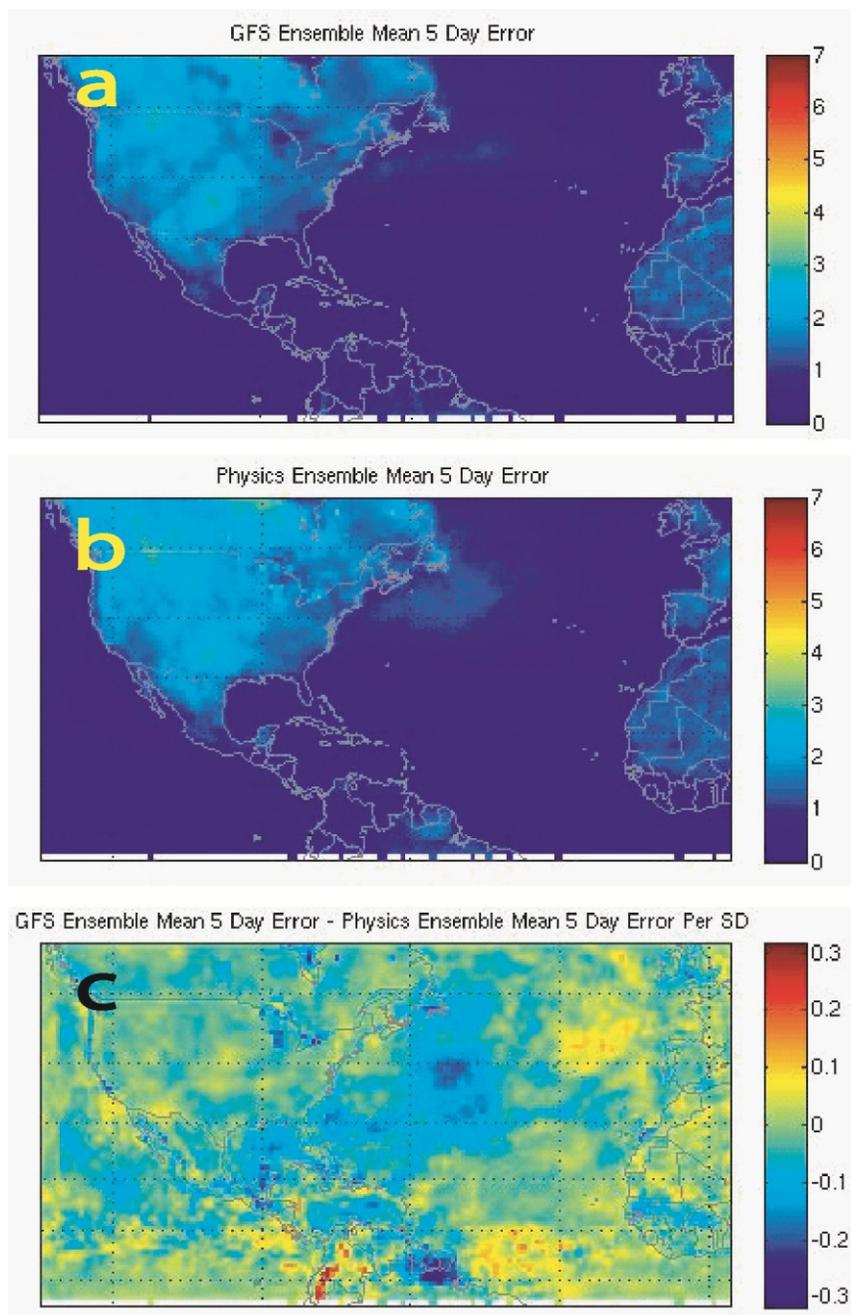
**Figure 4.11** Mean 2 m temperature error (°C) at forecast hour 120 for ten parameterization ensemble members.



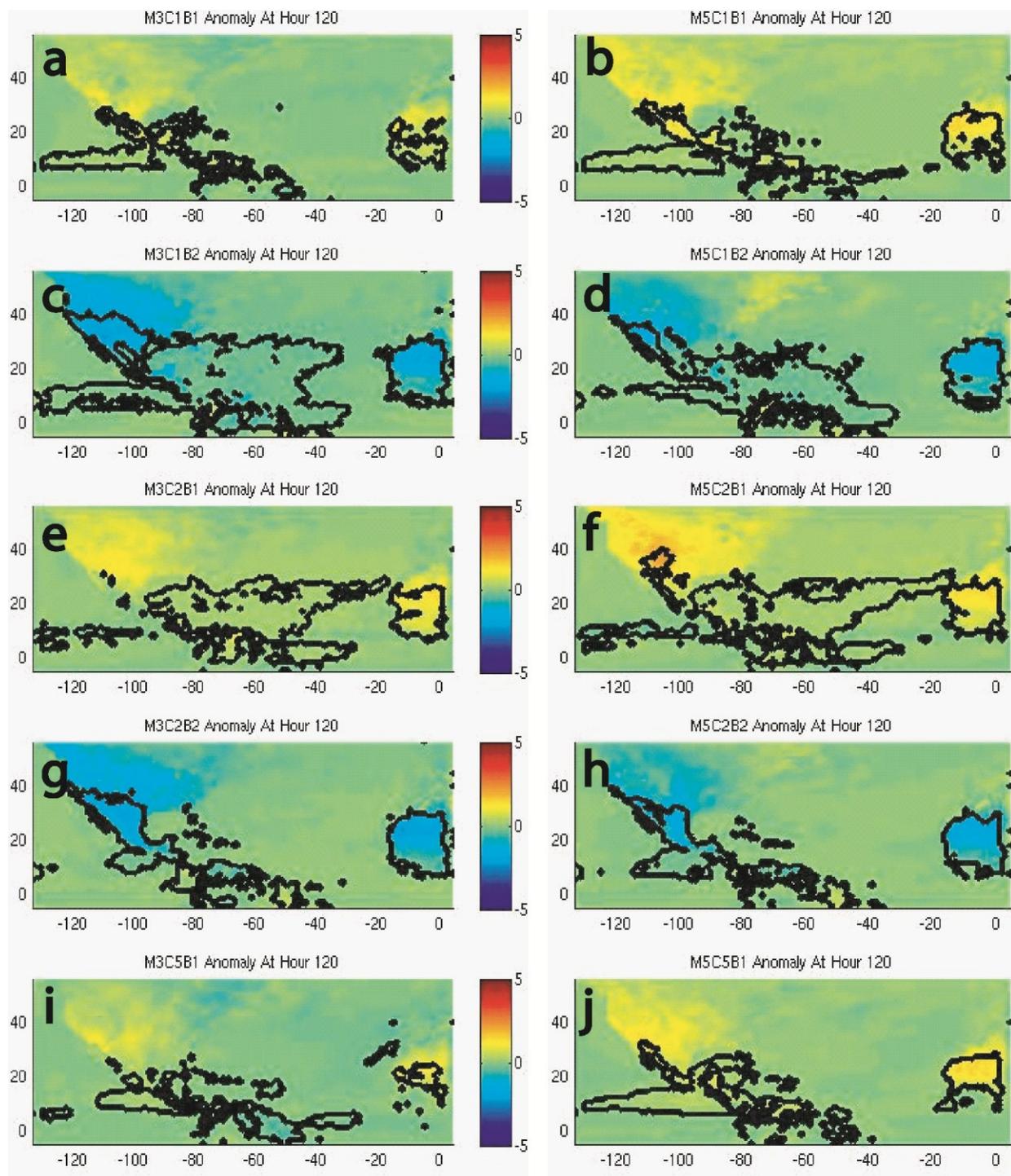
**Figure 4.12** As in Figure 4.11 but for ten GEFS members.

land. Although some of the differences can be attributed to the elevation differences between the native grid and the verification grid, the majority of the difference is simply due to the fact that on average the GFS is a better, well-tested, higher-resolution model than the physics members. However, once both ensembles are optimized to produce a single forecast the errors for both are reduced, as can be seen in **Fig. 4.13a** (GFS) and **Fig. 4.13b** (parameterization ensemble), which shows the mean errors for each ensemble. Both ensembles are fairly comparable – maximum mean errors are approximately  $3^{\circ}$  C over North America. **Figure 4.13c** shows the difference between the two mean errors normalized relative to the standard deviation of the analyzed 2 m temperature per grid point, which allows for an easier comparison of relative skill across different locations. Warm (cool) colors represent locations where the physics ensemble performed relatively better (more poorly) than the initial condition ensemble. There is a roughly equivalent amount of area where one performed better than the other. Overall, the initial condition ensemble outperformed the physics ensemble by an average of 0.0257, which is effectively parity. Similar results can be found when examining any number of common variables such as 2 m dewpoint, 10 m wind speed, sea level pressure, 500 hPa height, and others.

In all examples, a crude low-resolution physics ensemble initially more poorly predicts the atmosphere than an equivalent-sized collection of GEFS members when examining several variables. Yet when an identical statistical improvement is applied to both, the effect is that the two ensembles perform effectively equally. What accounts for the significant improvement in the physics ensemble relative to the initial condition ensemble? The answer lies in the type of biases associated with each ensemble. **Figure 4.14** shows the mean 2 m temperature anomaly ( $^{\circ}$ C) for each physics ensemble member relative to the mean of all members. The solid contour denotes locations where that particular member is statistically significantly different from the other



**Figure 4.13** Optimized ensemble mean 2 m temperature error ( $^{\circ}\text{C}$ ) for (a) GEFS ensemble, (b) parameterization ensemble, and (c) the difference between (a) and (b) normalized by the standard deviation of 2 m temperature at each grid point.

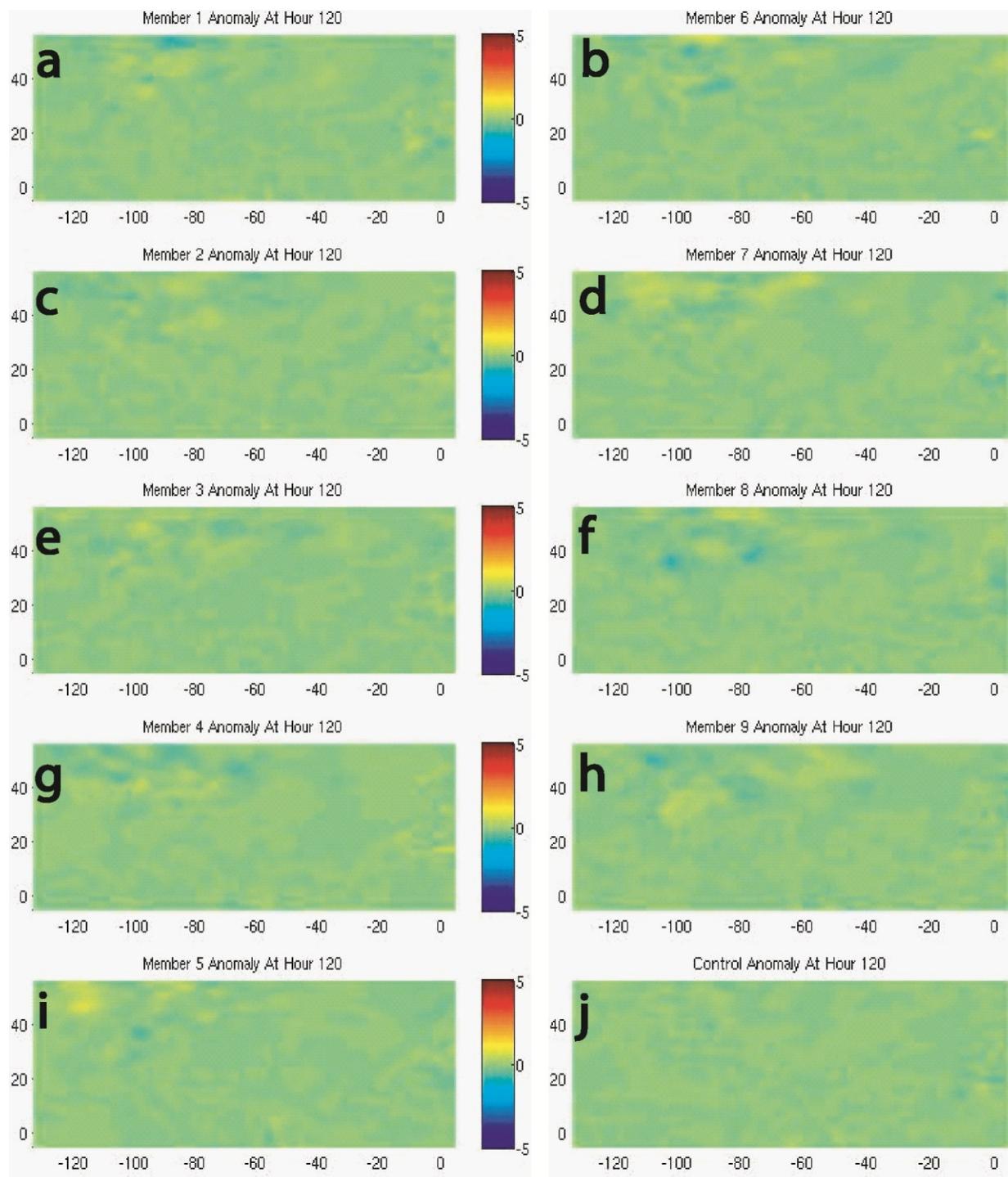


**Figure 4.14** Fill depicts 2 m temperature anomaly ( $^{\circ}\text{C}$ ) per member relative to a mean of all members. Contour denotes regions with anomalies surpassing a 95% statistical significance threshold.

members using a 95% confidence threshold. These locations are generally in the lower center of each panel or over Africa on the eastern portion of each panel. It can be immediately seen that certain members are cooler or warmer than other members on average, particularly over continental locations.

**Figure 4.15** depicts the same plot as **Figure 4.14**, but for the initial condition ensemble. It is immediately apparent that any particular member does not exhibit the same magnitude of anomalies relative to the mean of all members. No statistical significance contour is plotted because no anomalies are significant at the 95% threshold. The clear differences between **Figures 4.14** and **4.15** are due to the nature of the two ensembles. The various parameterizations used in the physics ensemble have unique biases. For example, rows one, three, and five in **Figure 4.14** are comprised of members using the Yonsei University boundary layer parameterization while rows two and four use the Mellor-Yamada-Janjic boundary layer parameterization. For these forecasts it is quite clear that the Yonsei University boundary layer parameterization is warmer for continental locations than the Mellor-Yamada-Janjic boundary layer parameterization on average. When examining other atmospheric variables, similar patterns emerge. For instance, an examination of 500 hPa height (not shown) demonstrates that the various member biases relative to one another are sorted by cumulus parameterization, which intuitively seems reasonable considering this level is frequently modified by the effects of convection.

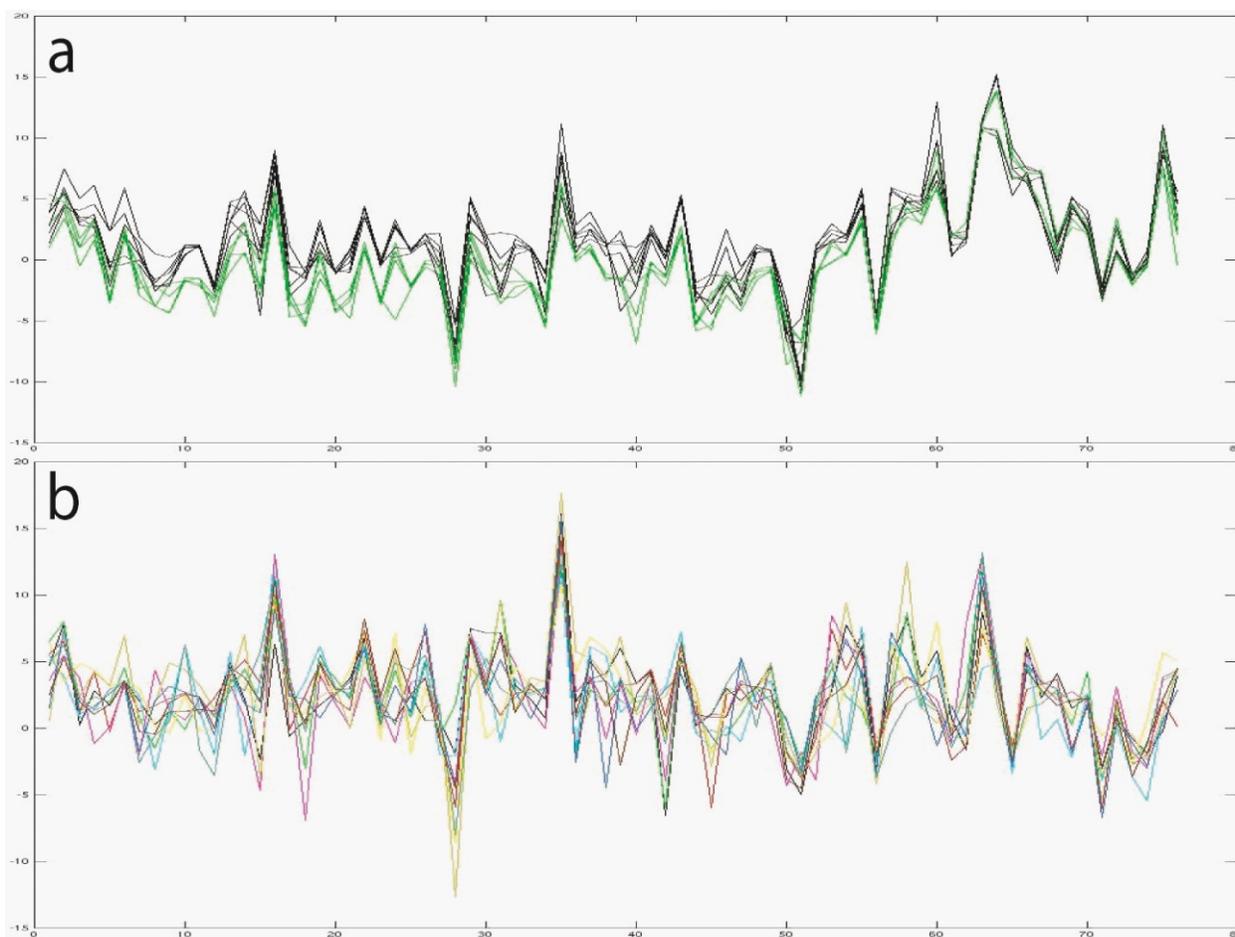
The GEFS uses the same combination of parameterizations for each member, because it differentiates members by choice of initial condition, so by definition any biases exhibited between members cannot be a product of differing parameterizations. However, unlike with a parameterization, there is no scientific reason that any apparent bias in a given member



**Figure 4.15** As in Figure 4.14 but for ten GEFS members.

differentiated by choice of initial conditions for one run would be reproduced in the following cycles, as is made clear by the lack of statistically significant differences. In this manner a physics ensemble possesses additional information to potentially leverage with statistical applications than does a purely initial condition ensemble. This is made more apparent by examining the absolute error for each ensemble member. **Figure 4.16** shows the absolute 2 m temperature error ( $^{\circ}\text{C}$ ) for each member of both ensembles for a grid point at approximately  $44^{\circ}\text{N}$ ,  $-108^{\circ}\text{W}$ , which is located over the western high plains of the United States. This is a location characterized by relative large 5 day forecast errors, given its frequent frontal passages and proximity to mountainous locations. The physics ensemble errors (**Fig 4.16a**) are highlighted according to choice of boundary layer parameterization while the initial condition ensemble errors (**Fig 4.16b**) are colored according to member. It is easily apparent that the physics ensemble's errors are frequently grouped according to choice of boundary layer parameterization, with the Yonsei University parameterization frequently being warmer than the Mellor-Yamada-Janjic parameterization, which confirms **Figure 4.14** for this region. However, no such distinction can be noted among any grouping of members within the initial condition ensemble.

The various biases which exist among different parameterizations is what allows the physics ensemble to be improved relatively more than the initial condition ensemble, leaving them of comparable skill when optimally combined into a single forecast for each ensemble. On a member by member basis, a relatively untested limited domain forecast employing a 90 km horizontal grid spacing would be expected to be considerably worse than a well-tested, higher resolution global forecast such as a GEFS member forecast. The fact that a collection of ten low resolution 90 km forecasts can perform on par with an equivalent collection of GEFS forecasts



**Figure 4.16** Absolute 2 m temperature ( $^{\circ}\text{C}$ ) errors are shown per case for (a) parameterization ensemble members and (b) GEFS members for a grid point at approximately  $44^{\circ}\text{N}$ ,  $-108^{\circ}\text{W}$ . Parameterization members are colored according to choice of boundary layer parameterization, with Yonsei University (Mellor-Yamada-Janjic) colored black (green). Each GEFS member is uniquely colored.

suggests there could be further utility in this approach, particularly if the physics ensemble members were constructed in such a way as to produce maximum (yet reasonable) forecast spread. As the physics ensemble is currently constructed, there is significant redundancy among the 10 members (combinations of only three cumulus parameterizations, two microphysics parameterizations and two boundary layer parameterizations). This suggests the possibility that a more well-chosen collection of members with more parameterization-independence would provide yet greater value.

It is not being suggested here that a collection of low resolution forecasts can effectively improve upon or replicate the GEFS skill in more difficult forecasting challenges such as tropical cyclogenesis, mode of severe weather, mountainous weather, or other scenarios (although it is also not being ruled out). However, given that the vast majority of day-to-day forecasts consist of surface temperature, moisture, and wind speed forecasts, it does seem that there is perhaps some utility to this approach. Given a finite amount of computer resources, it is perhaps more useful to create a very large ensemble of low-resolution forecasts than a small ensemble of high resolution forecasts (again, potentially dependent on the type of weather one is forecasting), particularly if the low resolution ensemble consists of members with easily correctable biases as is the case with the physics ensemble presented herein. An obvious potential solution would be to construct an ensemble similar to the GEFS, but which incorporates differing parameterizations as well as initial conditions. This is partially done within the current ECMWF ensemble, which uses both differing initial conditions as well as slightly different model equations to integrate the model forward (Buizza et al., 1999).

## CONCLUSION

Several statistical techniques to improve parameterization ensemble forecasts were examined within this chapter. First, precipitation deficiencies inherent in the ensemble-based G3 CP were corrected in two different ways. The first, a precipitation reassignment approach, was shown to effectively re-map distributions of forecast precipitation within a given forecast towards an expected climatological distribution while still retaining the forecast's original character. This approach is applicable to any model (and potentially other atmospheric variables). The second approach involved a linear regression framework with which the individual members within the G3 CP might be optimized to produce a more reasonable precipitation forecast. Although this was performed in post-processing, the approach is broadly within the context of an original proposal by Grell and Dévényi (2002). Both approaches were shown to have varying degrees of success eliminating spurious light precipitation while enhancing intense precipitation.

The second section of this chapter explored a simple approach whereby a low-resolution parameterization-based ensemble as well as an equally-sized GEFS ensemble were optimized using a linear regression approach. Despite the fact that the unoptimized members of the parameterization ensemble were shown to be significantly worse predictors than the GEFS members, after optimization both performed comparably. This was shown to be possible because member errors between forecasts were correlated within the parameterization ensemble due to the inherent biases among parameterizations (several of which were discussed in a different context in Chapters 2 and 3). Conversely, the GEFS members' errors are uncorrelated from one forecast to the next. This difference provides the parameterization ensemble with additional information to be leveraged upon optimization.

## Chapter 5: Conclusion

Parameterization remains a necessary component of numerical weather prediction models in order to represent the important scale interactions between subgrid-scale (i.e. unresolved) processes and grid scale processes. It is now widely recognized that there is value in using ensemble prediction techniques to develop forecasts rather than attempting to forecast from one single deterministic model. This dissertation has explored the use of ensemble parameterization (e.g., the GD and G3 schemes) and parameterization ensembles). Chapter 2 (and chapters 3 and 4 to a lesser degree) examined an ensemble within a (cumulus) parameterization, describing unintended consequences (largely negative) introduced by incorporating an ensemble approach to a cumulus parameterization scheme. Chapter 3 and a predecessor Master's thesis compared parameterizations within an ensemble of tropical cyclone forecasts chosen for their dispersive nature. Finally, chapter 4 demonstrated several methods in which an ensemble-based parameterization as well as an ensemble of different parameterization-based forecasts could be improved through statistical techniques. All of these chapters incorporated a specific focus on forecasting timescales extending beyond 48 hours, as this is a timescale largely lacking in current parameterization research. This chapter will briefly summarize the primary takeaways from the previous chapters, some broad conclusions, and will also identify potential avenues to improve and advance the research presented within this dissertation.

### *Chapter Overviews*

Chapter 2 examined the sophisticated G3 cumulus parameterization (CP) which attempts to improve performance relative to other CPs by incorporating several variations of different CPs. A number of constants are varied as well to produce a total of 144 members, which are run individually and averaged to provide feedback to the model. It was shown that this approach

becomes problematic over tropical regions with little strong forcing for ascent. The primary reason for this is that due to the averaging over ensemble members inherent in the scheme, it becomes very difficult to produce intense rain rates unless a significant fraction of the 144 members are also producing large rain rates. Conversely, it becomes relatively easy to produce widespread light precipitation, since at a given time step only a fraction of the members need to be producing rainfall for a grid point to report precipitation. The problem with this is that much of the tropical world (ITCZ and monsoon locations for example) are largely characterized by either intense precipitation, or no precipitation. This model behavior results in less intense convection and convective precipitation on average distributed over a greater area when compared with the Kain-Fritsch CP. A side-effect of this behavior is increased mid-level subsidence drying, which acts to make convection less likely as forecast length increases. These factors were shown to have important feedbacks on downstream extratropical weather through a weaker Hadley Cell. It was shown that artificially reducing the ensemble size within the G3 CP allowed for more intense convection.

Chapter 3 and a prior Master's thesis studied several tropical cyclone cases notable for their real-time forecasting difficulty relative to both track and intensity. The aim was to determine whether certain parameterizations exhibited track, intensity, or structural biases related to tropical cyclones. This work was extended to tropical cyclogenesis through an examination of real-time forecasts created during the 2009 North Atlantic hurricane season. The most prominent group of biases related to the GD and G3 CPs, which were found to produce large, weaker cyclones which tended to have a poleward track bias. Additionally, these CPs were less able to initiate tropical cyclogenesis than other CPs. The cause of all of these differences appears to be related to the conclusions described in Chapter 2, which found that the GD and G3 CPs produced

widespread weak convection, but were much less likely to produce intense convection. Chapter 3 showed that the use of the GD and G3 CPs result in model forecasts that produced larger nascent vortices which were less able to achieve tropical cyclogenesis or to develop as intensely once genesis was achieved. It was also analyzed that differing vertical distribution of potential vorticity, as diagnosed with a pair of EF MPs, produced significant forecast differences relating to track and intensity. Specifically, it was found that a less variable (in terms of value at each level of the atmosphere) vertical distribution of potential vorticity aided the maintenance of the vertical integrity of tropical cyclones when confronted with moderate vertical wind shear.

Chapter 4 studied several approaches whereby the forecast precipitation deficiencies noted with the GD and G3 CPs in Chapters 2 and 3 could be ameliorated using statistical techniques applied in post-processing. This was done using approaches which sought to transform the predicted CDF of rainfall into one more representative of the atmosphere as found in TRMM data. These approaches effectively decreased the widespread light rain rate bias shown in these CPs while allowing for more realistic intense rain rates to be produced. In a separate study, a low-resolution parameterization based ensemble was compared with an equivalently sized ensemble comprised of GEFS members. Although the parameterization ensemble members perform worse than the GEFS members at predicting the atmosphere five days in advance, upon ensemble optimization it was found that the parameterization ensemble produced forecasts with roughly the same error characteristics as an optimized GEFS ensemble. It was shown that this extra improvement was possible due to the inherent correlation of biases among the parameterization ensemble members from one forecast to the next, which provided this ensemble additional information upon optimization that the GEFS ensemble was lacking.

### *Overarching Conclusions*

As described in the introductory chapter (Chapter 1) most extant research into the development, testing, and behavior of parameterizations has been focused on short-term (less than 48 hours) periods. The work performed within this dissertation specifically avoided this research model, and the results shown certainly indicate the importance of examining parameterizations at extended forecast lengths for a variety of scales and situations. For example, it was shown that rainfall CDFs produced by the three CPs studied are not stationary with respect to forecast time until about hour 48, which was largely due to their slow modification of the simulated atmosphere which acted to alter the likelihood of forecast rainfall distributions. Similarly, research presented in a prior Master's thesis and continued here demonstrated that when forecasting tropical cyclones, significant track, intensity, structural, and genesis differences among different parameterizations became increasingly apparent after forecast hour 48 (at the resolutions analyzed here).

While this dissertation did not set out to explore this, it should be noted that occasionally a consequence of parameterization-based research is the creation of de facto “what if” case studies. The analysis presented in Chapter 3 is a good example of this possibility – the unintentional defects within the EF MP allowed for a “what if the potential vorticity within a cyclone were distributed differently?” scenario to be examined. Although the analysis presented is not a comprehensive study of potential vorticity distributions, it does suggest the utility in additional research into this area. Chapter 2 detailed a similar unintentional “what if” – “what if the amount of tropical convection were increased, but its intensity were decreased?” Again, the answer to this question was not necessarily sought within this dissertation, but it does imply that this may be an interesting avenue of research. Both of these questions were asked as a result of

parameterization research, yet the questions themselves could be more thoroughly examined independently of parameterization research in a future study.

### *Future Directions*

A significant fraction of this dissertation has detailed the various deficiencies within the GD and G3 CPs. This might give the impression that it is the author's opinion that these CPs are destined to be inferior to other CPs. This is not the case. The ensembling technique these CPs employ, while detrimental to forecasts currently, provide an avenue for potential significant improvement. The motivation for their creation – recognition that no single parameterization approach was superior in all situations combined with a desire to incorporate multiple techniques is still a worthwhile endeavor. While this thesis provides a variety of reasons as to why the current implementation of this concept is not ideal, it does not imply future efforts at improvement would not be valuable. Grell and Dévényi (2002) studied one possible technique to improve their parameterization, and this dissertation briefly examined an additional technique. A future attempt to implement these strategies could provide favorable results.

The scope of the precipitation reassignment technique discussed in Chapter 4 could also be greatly expanded. For example, instead of correcting precipitation in the forecasts used here with TRMM precipitation data, one could develop a more comprehensive database of observed precipitation from rain gauges (for example, over the continental United States) and use that to correct precipitation forecasts from global models such as the GFS. A similar approach has been used within a very narrow context thus far, primarily in order to drive hydrologic or agricultural models. Hwang et al. (2011) used a CDF mapping approach for stations around Tampa Bay, FL using a combination of MM5 forecasts and rain gauge observations. Their technique further

differs from those presented here by virtue of the fact the model data they used was specific to immediate Tampa Bay vicinity, while Chapter 4 used the entirety of output model precipitation.

Although they did not use model data, Tobin and Bennett (2010) used statistical techniques to correct overland TRMM precipitation data based upon ground observations. Both of these studies found substantial benefits to the use of their techniques, which also suggests continued research within this realm of applications could be useful. A similar application aimed to benefit forecasters would be to not output corrected model precipitation, but to instead output forecast rainfall percentile. For an ensemble such as the GEFS (comprised of 21 members) a significant amount of precipitation data is generated. It could be reasonably easy to gather the considerable precipitation data generated from the 21-member GFS ensemble for a region, construct a rainfall CDF for a given forecasting interval, and produce forecasting plots indicating the percentile of precipitation. One potential problem with this approach is that frequent updates<sup>8</sup> to the GFS model make comparing or combining statistics from one implementation to another problematic. However, with the recent creation of the GFS reforecast<sup>9</sup> dataset (which includes a sufficiently large amount of raw data to generate statistics), some limited implementations of similar statistical corrections have begun to be attempted. Given the earlier documented problems with model precipitation within this dissertation and among other studies (Brown et al., 2012), this may be a more useful product for forecasters than raw precipitation output.

Finally, a considerable portion of this thesis has detailed relative biases attributable to choice of parameterization scheme on time-scales beyond 48 hours. Chapter 4 also demonstrated that these biases could aid forecasting when properly utilized. This combination naturally lends itself to expanded research in this area of understanding how to properly utilize parameterization

---

<sup>8</sup>A list of recent changes can be found at: <http://www.emc.ncep.noaa.gov/GFS/impl.php>

<sup>9</sup>A description of this dataset can be found at:

[http://www.esrl.noaa.gov/psd/forecasts/reforecast2/README.GEFS\\_Reforecast2.pdf](http://www.esrl.noaa.gov/psd/forecasts/reforecast2/README.GEFS_Reforecast2.pdf)

biases. The Short-Range Ensemble Forecasting (SREF) System already does this for certain variables (Yuan et al., 2007), but a statistical correction of parameterization ensembles could expand to object-based forecasting, such as TC genesis, track, or intensity. For example, Chapter 3 demonstrated that different CPs produce TCs of different intensities, structures, and tracks. However, none of the forecasted behavior was clearly atypical of TC behavior. Likewise, an examination of TC genesis during the 2009 North Atlantic Hurricane season showed significant differences in TC structure at genesis, as well as frequency and location of genesis. Again, none of the forecasted behavior was clearly atypical of TC behavior.

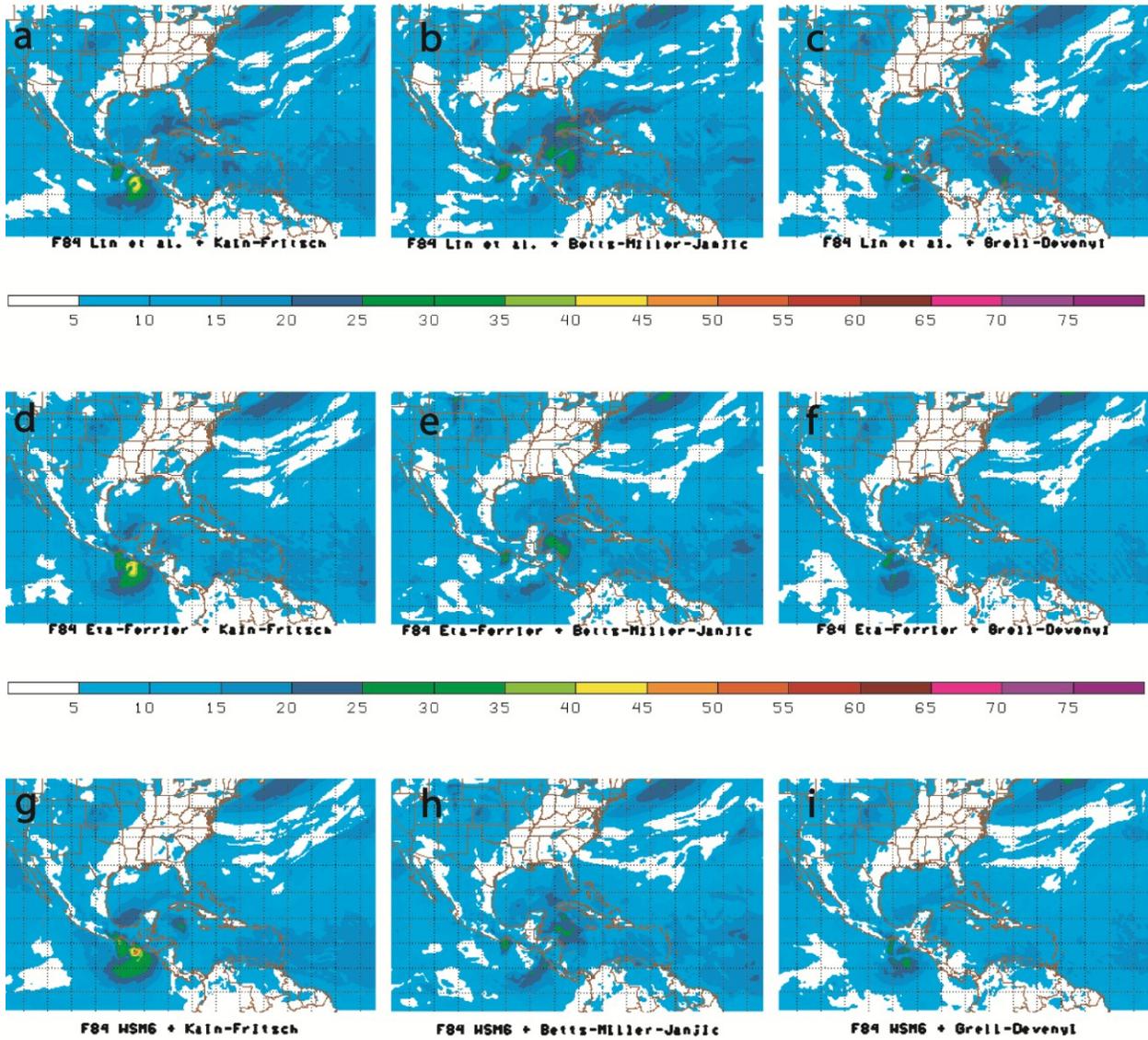
A combination of a significant amount of data and an understanding of the tendencies among different parameterizations could provide an opportunity to develop a useful parameterization-based ensemble forecasting system. Some work has already been done on the optimal combination of forecast tracks (Goerss, 2007). Additionally, the Hurricane Forecast Improvement Project (HFIP) is a wide-ranging, coordinated effort to improve TC forecasting. Of particular focus within HFIP is on extended forecast lengths (up to 7 days lead time) and improvement upon rapid intensity change (Gall et al., 2013). This is a noble undertaking, but a parallel endeavor, or one working within the HFIP framework may be appropriate for a few reasons. First, three of the four global models in HFIP use an *identical combination of parameterizations and grid spacings*. When regional models are included, six of the eleven models use the Arakawa-Schubert CP and Ferrier MP combination (the remaining CP choices are two KF CP, two Emmanuel CP, and one GD CP). This lack of diversity among model physics may pose a significant challenge toward achieving a realistic diversity of forecasts, and thus potentially limit the scope of further forecast understanding. Second, a primary focus of HFIP is to use very high-resolution models in order to successfully predict inner-core dynamics

as well as rapid intensification. The high computational cost of doing this in real-time greatly limits the number of forecasts one is able to generate (and thus decrease ensemble size), while also necessitating the distribution of their creation across the world.

Given the above two constraints, there may be great utility in a large parameterization ensemble of moderate resolution to aid in the goal of hurricane forecast improvement. There is already some evidence for the usefulness of this approach. **Figure 5.1** shows 84 hour forecasts of 10 m wind (kt) for a prototype version of the earlier described 2009 real-time ensemble, initialized 1800 UTC 26 May 2008. This ensemble used a slightly different domain, and slightly different parameterization combinations. At the time of this initialization, there was great uncertainty regarding the potential genesis of a tropical cyclone in either the Eastern Pacific Ocean or western Caribbean. As detailed by Jeff Masters<sup>10</sup> at the time: *“It is uncertain which ocean basin such a storm might form in ... the GFS, NOGAPS, and Canadian models all predict a tropical depression might form in the Western Caribbean near Mexico's Yucatan Peninsula. In contrast, the UKMET shows development in the Eastern Pacific.”* **Figure 5.1** shows that several ensemble members (the KF CP members (**Fig. 5.1a,d,g**) and to a lesser extent the GD CP (**Fig. 5.1b,f,i**) members) predict genesis in the Eastern Pacific, while several predict genesis in the northwest Caribbean (the BMJ CP members (**Fig. 5.1b,e,h**)). This distribution effectively duplicates the forecast distribution found in the available global models. Since the genesis location was differentiated by choice of CP, it is not inherently obvious that the models with identical parameterizations used in HFIP would be capable of producing such an appropriate diversity of forecasts. Incidentally, Tropical Storm Alma formed in the Eastern Pacific on 29

---

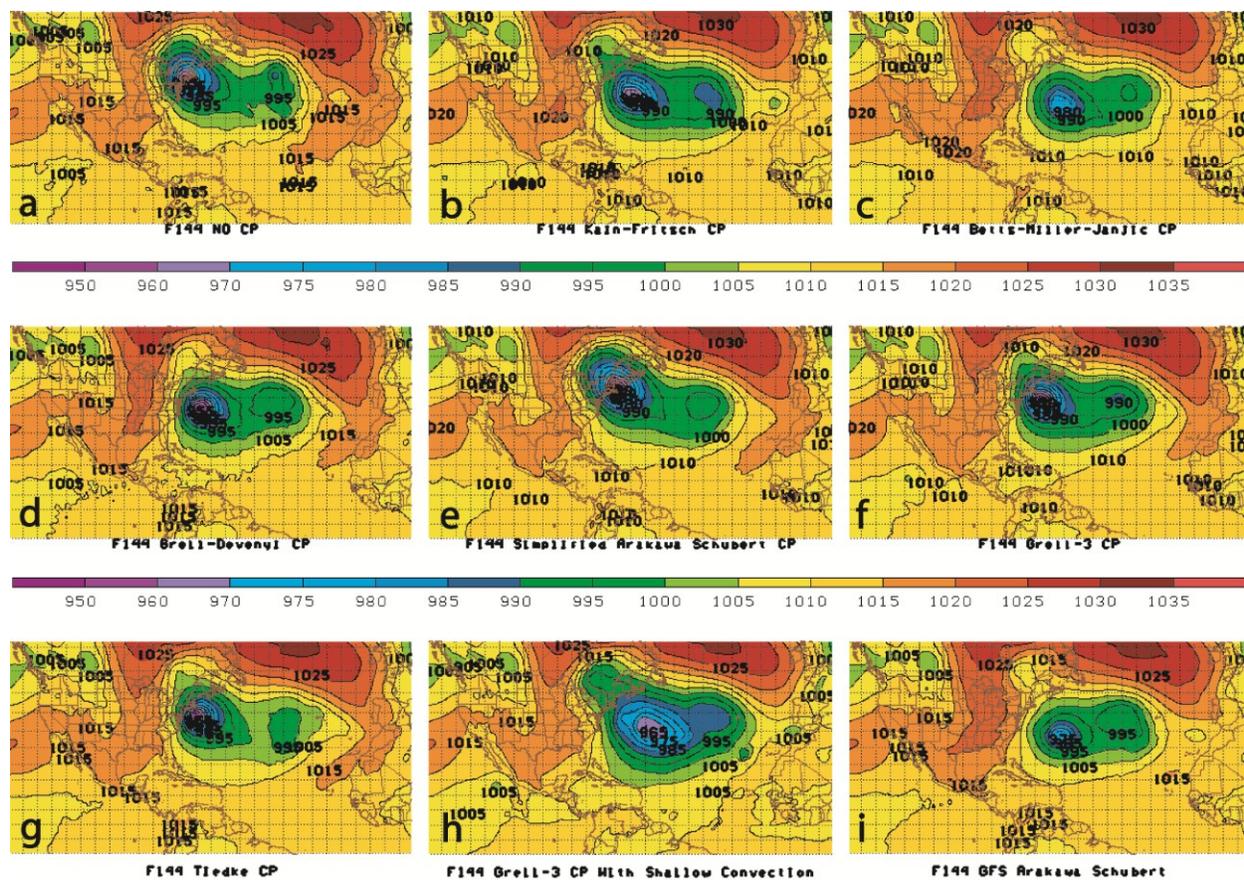
<sup>10</sup> <http://www.wunderground.com/blog/JeffMasters/comment.html?entrynum=947&tstamp=200805>



**Figure 5.1** 84 hour forecast 10 m wind (kt) is shown for nine parameterization ensemble members initialized 1800 UTC 26 May 2008.

May, eventually crossed Central America, while Tropical Storm Arthur formed near the Belize coastline on 30 May, so in a sense both genesis possibilities were correct.

A more recent example can be seen in **Figure 5.2**, which depicts nine 144 hour sea level pressure forecasts of Hurricane Sandy (2012) from an ensemble of members differentiated by choice of CP (all other parameterizations are held constant) initialized 1200 UTC 23 October 2012 using the GFS for initial and boundary conditions. The domain and grid is identical to the 90 km real-time outer domain described in previous chapters. Although a wide variety of potential solutions exists, it is clear that a significant fraction of the ensemble members predict the location and intensity of Sandy on day six very accurately relative to what occurred (in actuality Sandy made landfall in New Jersey approximately at the time these forecasts were valid). The range of forecasts shown here effectively reproduces the observed spread of forecasts in operational models at the time, as models such as the ECMWF consistently predicted a Mid-Atlantic landfall while the GFS consistently (at this time) forecast Sandy to recurve out to sea. Incidentally, the fact that this distribution is possible among forecasts using initial conditions from a model which produced a recurving track means that this result *was not* a product of the initial conditions. Rather, as shown by the ensemble member using the GFS version of the Arakawa-Schubert CP (**Fig 5.2i**), the forecast track produced by the official GFS model at that time was likely much more related to its choice of model physics given that that track is reproduced here in a different model when that CP is used. This may have potential implications for HFIP research, as it was earlier mentioned that a significant fraction of their models use a version of this parameterization.



**Figure 5.2** 144 hour forecast sea level pressure is shown for nine parameterization ensemble members initialized 1200 UTC 23 October 2012.

Several examples of a parameterization ensemble replicating observed trends in operational forecast models has been presented at the 14<sup>th</sup> Cyclone Workshop in 2008<sup>11</sup> as well as at a Unidata User's Meeting in 2012<sup>12</sup>. Although this has not been comprehensively analyzed to the extent where inclusion in this dissertation is appropriate, the accumulation of evidence certainly suggests that a great utility exists in parameterization-based ensembles, as shown by the previous two examples.

---

<sup>11</sup> Presentation viewable at: <http://aurora.aos.wisc.edu/~bassill/2008-Cyclone%20Workshop.ppt>

<sup>12</sup> Poster viewable at: <http://aurora.aos.wisc.edu/~bassill/2012%20-%20Unidata%20Workshop.ppt>

## Bibliography

- Arakawa, Akio, 2004: The Cumulus Parameterization Problem: Past, Present, and Future. *J. Climate*, **17**, 2493–2525.
- , 1993: Closure assumptions in the cumulus parameterization problem. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No. 46, Amer. Meteor. Soc., 1–16.
- , and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment. Part I. *J. Atmos. Sci.*, **31**, 674–701.
- , and J.-M. Chen, 1987: Closure assumptions in the cumulus parameterization problem. *Short- and Medium-Range Numerical Prediction: Collection of Papers Presented at the WMO/IUGG NWP Symposium*, T. Matsuno, Ed., Japan Meteorological Society, 107–131.
- Baldwin, Michael E., John S. Kain, Michael P. Kay, 2002: Properties of the Convection Scheme in NCEP's Eta Model that Affect Forecast Sounding Interpretation. *Wea. Forecasting*, **17**, 1063–1079.
- Bao, J.-W., S. G. Gopalakrishnan, S. A. Michelson, F. D. Marks, M. T. Montgomery, 2012: Impact of Physics Representations in the HWRF on Simulated Hurricane Structure and Pressure–Wind Relationships. *Mon. Wea. Rev.*, **140**, 3278–3299.
- Bassill, N., 2009: Forecast track and intensity sensitivities of tropical cyclones to various parameterizations using the WRF-ARW model, Univ. of Wisconsin – Madison, M.S. thesis.
- Bernardet, Lígia, and Coauthors, 2008: The Developmental Testbed Center and its Winter Forecasting Experiment. *Bull. Amer. Meteor. Soc.*, **89**, 611–627.
- Bickel, J. Eric, Eric Floehr, Seong Dae Kim, 2011: Comparing NWS PoP Forecasts to Third-Party Providers. *Mon. Wea. Rev.*, **139**, 3304–3321.
- Berg, Wesley, Tristan L'Ecuyer, John M. Haynes, 2010: The Distribution of Rainfall over Oceans from Spaceborne Radars. *J. Appl. Meteor. Climatol.*, **49**, 535–543.
- Betts, A. K., 1985.: Mixing Line Analysis of Clouds and Cloudy Boundary Layers. *J. Atmos. Sci.*, **42**, 2751–2763.
- , 1986: A new convective adjustment scheme. Part I: Observational and theoretical basis. *Quart. J. Roy. Meteor. Soc.*, **112**, 677–691.
- , and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709.

- , and ——, 1993: The Betts–Miller scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No. 46, Amer. Meteor. Soc., 107–122.
- Bister, Marja, 2001: Effect of Peripheral Convection on Tropical Cyclone Formation. *J. Atmos. Sci.*, **58**, 3463–3476.
- Brown, James D., Dong-Jun Seo, Jun Du, 2012: Verification of Precipitation Forecasts from NCEP’s Short-Range Ensemble Forecast (SREF) System with Reference to Ensemble Streamflow Prediction Using Lumped Hydrologic Models. *J. Hydrometeorol.*, **13**, 808–836.
- Buizza, Roberto, 1997: Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **125**, 99–119.
- , and T. N. Palmer, 1998: Impact of Ensemble Size on Ensemble Prediction. *Mon. Wea. Rev.*, **126**, 2503–2518.
- , M. Miller and T. N. Palmer, 1999. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.
- Charney, Jule G., Arnt Eliassen, 1964: On the Growth of the Hurricane Depression. *J. Atmos. Sci.*, **21**, 68–75.
- Cho, Han-Ru. "Cumulus cloud population and its parameterization." *pure and applied geophysics* 113.1 (1975): 837-849.
- Crétat, Julien, Benjamin Pohl, 2012: How Physical Parameterizations Can Modulate Internal Variability in a Regional Climate Model. *J. Atmos. Sci.*, **69**, 714–724.
- Ebert, Elizabeth E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480.
- Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP Eta model. Preprints, 15th Conf. on Numerical Weather Prediction, San Antonio, TX, Amer. Meteor. Soc., 280–283.
- , 2005: An efficient mixed-phase cloud and precipitation scheme for use in Operational NWP Models. *EOS Trans. AGU*, **86**, Jt. Assem. Suppl., Abstract A42A-02.
- Fiorino, M. J., and R. L. Elsberry, 1989: Some aspects of vortex structure related to tropical cyclone motion. *J. Atmos. Sci.*, **46**, 975–990.
- Fovell, Robert G., Kristen L. Corbosiero, Hung-Chi Kuo, 2009: Cloud Microphysics Impact on Hurricane Track as Revealed in Idealized Experiments. *J. Atmos. Sci.*, **66**, 1764–1778.
- Gall, Robert, James Franklin, Frank Marks, Edward N. Rappaport, Frederick Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343.

- Gallus, W.A., 1999: Eta Simulations of Three Extreme Precipitation Events: Sensitivity to Resolution and Convective Parameterization. *Wea. Forecasting*, **14**, 405–426.
- Gianotti, Rebecca L., Dongfeng Zhang, Elfatih A. B. Eltahir, 2012: Assessment of the Regional Climate Model Version 3 over the Maritime Continent Using Different Cumulus Parameterization and Land Surface Schemes. *J. Climate*, **25**, 638–656.
- Gilliland, E.K. and C.M. Rowe, 2007. A comparison of cumulus parameterization schemes in the WRF model. Paper read at Proceedings of the 87th AMS Annual Meeting & 21st Conference on Hydrology. San Antonio, TX.
- Goerss, James S., 2007: Prediction of Consensus Tropical Cyclone Track Forecast Error. *Mon. Wea. Rev.*, **135**, 1985–1993.
- Grell, G., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1693, doi:10.1029/2002GL015311.
- Haglund, N. L., 2007: Quantitative Precipitation Forecast Sensitivity to Microphysics Parameterization and Sea Surface Temperature Source over North Carolina during Two Cold Season Events, North Carolina State University, 221 pp. [Available online at: <http://www.lib.ncsu.edu/resolver/1840.16/1042>]
- Hausman, Scott A., Katsuyuki V. Ooyama, Wayne H. Schubert, 2006: Potential Vorticity Structure of Simulated Hurricanes. *J. Atmos. Sci.*, **63**, 87–108.
- Hill, Kevin A., Gary M. Lackmann, 2011: The Impact of Future Climate Change on TC Intensity and Structure: A Downscaling Approach. *J. Climate*, **24**, 4644–4661.
- Hong, Song-You, Jimy Dudhia, Shu-Hua Chen, 2004: A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Mon. Wea. Rev.*, **132**, 103–120.
- Hwang, Syewoon, Wendy Graham, José L. Hernández, Chris Martinez, James W. Jones, Alison Adams, 2011: Quantitative Spatiotemporal Evaluation of Dynamically Downscaled MM5 Precipitation Predictions over the Tampa Bay Region, Florida. *J. Hydrometeor*, **12**, 1447–1464.
- Jankov, I., and W.A. Gallus, M. Segal, B. Shaw, and S.E. Koch, 2005: The Impact of Different WRF Model Physical Parameterizations and Their Interactions on Warm Season MCS Rainfall. *Wea. Forecasting*, **20**, 1048–1060.
- \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, and S.E. Koch, 2007: Influence of Initial Conditions on the WRF–ARW Model QPF Response to Physical Parameterization Changes. *Wea. Forecasting*, **22**, 501–519.

- Jiang, Haiyan, 2012: The Relationship between Tropical Cyclone Intensity Change and the Strength of Inner-Core Convection. *Mon. Wea. Rev.*, **140**, 1164–1176.
- Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.
- , and ———, 1993: Convective parameterization for mesoscale models: The Kain–Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models*, Meteor. Monogr., No. **46**, Amer. Meteor. Soc., 165–170.
- Kawazoe, S., and W. Gutowski, 2013: Regional, Very Heavy Daily Precipitation in NARCCAP Simulations. *J. Hydrometeor.* doi:10.1175/JHM-D-12-068.1, in press.
- Kessler, E., 1969: On the distribution and continuity of water substance in atmospheric circulation., *Meteor. Monogr.*, **32**, Amer Meteor. Soc., 84 pp.
- Kieu, Chanh Q., Da-Lin Zhang, 2010: A Piecewise Potential Vorticity Inversion Algorithm and Its Application to Hurricane Inner-Core Anomalies. *J. Atmos. Sci.*, **67**, 2616–2631.
- Liu, Zhong, Dana Ostrenga, William Teng, Steven Kempler, 2012: Tropical Rainfall Measuring Mission (TRMM) Precipitation Data and Services for Research and Applications. *Bull. Amer. Meteor. Soc.*, **93**, 1317–1325.
- Lunney, P., 1988: Environmental and convective influence on tropical cyclone development vs. non-development. Dept. of Atmos. Sci. Paper No. 436, Colorado State University, 105 pp.
- Manabe, Syukuro, Robert F. Strickler, 1964: Thermal Equilibrium of the Atmosphere with a Convective Adjustment. *J. Atmos. Sci.*, **21**, 361–385.
- Molinari, John, Michael Dudek, 1992: Parameterization of Convective Precipitation in Mesoscale Numerical Models: A Critical Review. *Mon. Wea. Rev.*, **120**, 326–344.
- , Steven Skubis, David Vollaro, Frank Alsheimer, Hugh E. Willoughby, 1998: Potential Vorticity Analysis of Tropical Cyclone Intensification. *J. Atmos. Sci.*, **55**, 2632–2644.
- Mukhopadhyay, P., S. Taraphdar, B. N. Goswami, K. Krishnakumar, 2010: Indian Summer Monsoon Precipitation Climatology in a High-Resolution Regional Climate Model: Impacts of Convective Parameterization on Systematic Biases. *Wea. Forecasting*, **25**, 369–387.
- Murphy, Allan H., 1998: The Early History of Probability Forecasts: Some Extensions and Clarifications. *Wea. Forecasting*, **13**, 5–15.

- Nasrollahi, Nasrin, Amir AghaKouchak, Jialun Li, Xiaogang Gao, Kuolin Hsu, Soroosh Sorooshian, 2012: Assessing the Impacts of Different WRF Precipitation Physics in Hurricane Simulations. *Wea. Forecasting*, **27**, 1003–1016.
- Ooyama, Katsuyuki, 1964: A Dynamical Model for the Study of Tropical Cyclone Development. *Geofis. Intern.*, **4**, 187-198.
- , and ——, 1969: Numerical Simulation of the Life Cycle of Tropical Cyclones. *J. Atmos. Sci.*, **26**, 3–40.
- Osuri K, Mohanty U, Routray A, et al. Customization of WRF-ARW model with physical parameterization schemes for the simulation of tropical cyclones over North Indian Ocean. *Nat Hazards*, 2011, doi: 10.1007/s11069-011-9862-0
- Rapp, A. D., M. D. Lebsock, and T. S. L'Ecuyer, 2013: Low cloud precipitation climatology in the southeastern Pacific marine stratocumulus region using CloudSat. *Environ. Res. Letters*, **8**, doi:10.1088/1748-9326/8/1/014027
- Rotunno, Richard, Kerry A. Emanuel, 1987: An Air–Sea Interaction Theory for Tropical Cyclones. Part II: Evolutionary Study Using a Nonhydrostatic Axisymmetric Numerical Model. *J. Atmos. Sci.*, **44**, 542–561.
- Simmons, A. J., J. M. Wallace, G. W. Branstator, 1983: Barotropic Wave Propagation and Instability, and Atmospheric Teleconnection Patterns. *J. Atmos. Sci.*, **40**, 1363–1392.
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF version 2, NCAR Tech. Note, NCAR/TN—468+STR, 100 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
- , ——, ——, ——, ——, ——, ——, 2008: A description of the advanced research WRF version 3. NCAR Technical Note NCAR/TN-475+STR, 113 pp., Natl. Cent. for Atmos. Res., Boulder, Colo.
- Stensrud, D. J., 2009: *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press, 459 pp.
- Tao, Wei-Kuo, et al. "The impact of microphysical schemes on hurricane intensity and track." *Asia-Pacific Journal of Atmospheric Sciences* 47.1 (2011): 1-16.
- Thompson, Philip Duncan, 1977: How to Improve Accuracy by Combining Independent Forecasts. *Mon. Wea. Rev.*, **105**, 228–229.
- Tobin, Kenneth J., Marvin E. Bennett, 2010: Adjusting Satellite Precipitation Data to Facilitate Hydrologic Modeling. *J. Hydrometeor.*, **11**, 966–978.

- Tory, K. J., M. T. Montgomery, N. E. Davidson, 2006: Prediction and Diagnosis of Tropical Cyclone Formation in an NWP System. Part I: The Critical Role of Vortex Enhancement in Deep Convection. *J. Atmos. Sci.*, **63**, 3077–3090.
- Tracton, M. Steven, Eugenia Kalnay, 1993: Operational Ensemble Prediction at the National Meteorological Center: Practical Aspects. *Wea. Forecasting*, **8**, 379–398.
- Velden, Christopher S., Lance M. Leslie, 1991: The Basic Relationship between Tropical Cyclone Intensity and the Depth of the Environmental Steering Layer in the Australian Region. *Wea. Forecasting*, **6**, 244–253.
- Warner, Thomas T., 2011: Quality Assurance in Atmospheric Modeling. *Bull. Amer. Meteor. Soc.*, **92**, 1601–1610.
- Wehner, M., 2012: Very extreme seasonal precipitation in the NARCCAP ensemble: model performance and projections. *Climate Dyn.*, **22**, doi: 10.1007/s00382-012-1393-1
- Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79.
- Weisman, Morris L., William C. Skamarock, Joseph B. Klemp, 1997: The Resolution Dependence of Explicitly Modeled Convective Systems. *Mon. Wea. Rev.*, **125**, 527–548.
- Wolff, Jamie K., Brad S. Ferrier, Clifford F. Mass, 2012: Establishing Closer Collaboration to Improve Model Physics for Short-Range Forecasts. *Bull. Amer. Meteor. Soc.*, **93**, ES51–ES53.
- Wu, Chun-Chieh, Yoshio Kurihara, 1996: A Numerical Study of the Feedback Mechanisms of Hurricane–Environment Interaction on Hurricane Movement from the Potential Vorticity Perspective. *J. Atmos. Sci.*, **53**, 2264–2282.
- Wu, Liguang, Bin Wang, 2000: A Potential Vorticity Tendency Diagnostic Approach for Tropical Cyclone Motion. *Mon. Wea. Rev.*, **128**, 1899–1911.
- , and ———, 2001a: Movement and vertical coupling of diabatic baroclinic tropical cyclones. *J. Atmos. Sci.*, **58**, 1801–1814.
- , and ———, 2001b: Effects of Convective Heating on Movement and Vertical Coupling of Tropical Cyclones: A Numerical Study. *J. Atmos. Sci.*, **58**, 3639–3649.
- Li, Xiang. "Sensitivity of WRF simulated typhoon track and intensity over the Northwest Pacific Ocean to cumulus schemes." *Science China Earth Sciences* 56.2 (2013): 270-281.

- Yanai, Michio, Steven Esbensen, Jan-Hwa Chu, 1973: Determination of Bulk Properties of Tropical Cloud Clusters from Large-Scale Heat and Moisture Budgets. *J. Atmos. Sci.*, **30**, 611–627.
- Yuan, Huiling, Xiaogang Gao, Steven L. Mullen, Soroosh Sorooshian, Jun Du, Hann-Ming Henry Juang, 2007: Calibration of Probabilistic Quantitative Precipitation Forecasts with an Artificial Neural Network. *Wea. Forecasting*, **22**, 1287–1303.
- Zhang, Z., T. N. Krishnamurti, 1997: Ensemble Forecasting of Hurricane Tracks. *Bull. Amer. Meteor. Soc.*, **78**, 2785–2795.