Processes governing the amplification of forecast errors and forecast uncertainty in a quantitative potential-vorticity framework

Dissertation

zur Erlangung des Grades

„Doktor der Naturwissenschaften“

am Fachbereich Physik, Mathematik und Informatik

der Johannes Gutenberg-Universität

in Mainz

Marlene Annette Baumgart

geb. in Hannover

Mainz, den 10. Mai 2020

JG|U

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Date of oral exam: 08.05.2020

Abstract*

This thesis provides a quantitative view on the relative importance of individual processes to the flow-dependent amplification of forecast errors and forecast uncertainty. The focus is on the near-tropopause region of the midlatitudes, which is of particular importance to the evolution of midlatitude weather systems and the evolution of forecast errors. A potential-vorticity (PV) framework is developed to partition the contributions from individual processes. This partitioning builds on the PV perspective for midlatitude dynamics and includes the influence of near-tropopause dynamics (nonlinear Rossby-wave dynamics), tropospheric-deep interaction (baroclinic growth), upper-tropospheric divergence (often associated with latent heat release), and nonconservative processes (direct PV modification by diabatic heating and nonconservative momentum change).

In the first part of this thesis, the error growth of a state-of-the-art deterministic forecast is analyzed. In this case, localized mesoscale error maxima form near the tropopause until two forecast days. In the following days, this error pattern changes into a wave-like pattern that maximizes along the dynamical tropopause and reaches the scale of individual Rossby-wave anomalies after about six forecast days. An analysis based on the PV framework reveals that the error growth in this case is dominated by differences in the nonlinear Rossby-wave dynamics near the tropopause. Upper-tropospheric divergence makes a large contribution to error growth in localized regions where a surface cyclone is misrepresented. The direct impact of tropospheric-deep interaction, and thus baroclinic instability, is much smaller than the impact of near-tropopause dynamics. The mesoscale errors generated near the tropopause thus do not primarily project on differences in the baroclinic growth as stated in previous studies. Instead, they directly project on the tropopause evolution and amplify because of differences in the nonlinear Rossby-wave dynamics.

The second part of this thesis investigates upscale error growth from the grid scale up to the planetary scale using global simulations with a stochastic convection scheme. Using the PV

*This abstract summarizes the results presented in Chapters 2–4, which have been published in Baumgart et al. (2018), Baumgart et al. (2019), and Baumgart and Riemer (2019). A few phrases of this abstract coincide with phrases from these publications.

framework, a distinct sequence of the processes governing upscale error growth can be identified: In the first twelve hours, latent-heating differences induced by the convection scheme dominate the PV-error growth near the tropopause. In the following 1.5 days, the dominant error-growth contribution is given by upper-tropospheric divergence, which provides an effective mechanism to project errors from moist processes into the tropopause region. After two days, differences in the nonlinear near-tropopause dynamics dominate the error growth. A fourth stage of the error growth exists after about 14.5 days when the error grows from the synoptic scale of individual Rossby-wave anomalies up to the planetary scale of the Rossby-wave envelope. Compared to previous studies, a novel interpretation of the processes governing upscale error growth is provided.

In the last part of this thesis, the amplification of ensemble spread is investigated for a medium-range forecast with large forecast uncertainty. The focus is on two aspects of the ensemble behavior: i) the mean divergence of the ensemble members indicating the general amplification of forecast uncertainty and ii) the divergence of the best and worst member indicating extremes in possible error-growth scenarios. To quantify the amplification of forecast uncertainty, a tendency equation for the ensemble variance of PV is derived. Averaged over the midlatitudes of the Northern Hemisphere, the variance amplification of this case is dominated by near-tropopause dynamics. Locally, however, there can be large differences as in the region where tropical storm Karl interacts with the Rossby-wave pattern during extratropical transition. In this region, the variance amplification is mostly related to the moist baroclinic cyclone development. The differences between the error growth in the best and worst member can, to a large part, be traced back in time to differences in the representation of a cutoff evolution, which further amplify in the highly nonlinear region of a large-amplitude ridge.

This thesis provides novel insight into the processes governing the flow-dependent amplification of forecast errors and forecast uncertainty. This insight can be used to improve the interpretation, and possibly also the design, of future forecast systems. The results from this thesis can also be used to identify regions that are (intrinsically) prone to a large amplification of forecast errors and forecast uncertainty. As the first two stages of upscale error growth are not dominating error-growth mechanisms in the operational forecasts, one can expect that initial-condition errors are still of larger importance to error growth in operational forecasts than errors from moist processes with low intrinsic predictability.

List of Publications

This thesis is based on three lead-author publications, which have been accepted in peer-reviewed journals during the time of this Ph.D. project. The following list provides a short summary of the publications and specifies the author's contribution to the publications.

Baumgart et al. (2018)

Baumgart, M., M. Riemer, V. Wirth, F. Teubler, and S.T. Lang, 2018: Potential Vorticity Dynamics of Forecast Errors: A Quantitative Case Study. *Mon. Wea. Rev.*, **146**, 1405–1425, https://doi.org/10.1175/MWR-D-17-0196.1

This paper investigates forecast-error growth from a process-based perspective in an operational state-of-the-art deterministic forecast. A potential-vorticity (PV) framework is developed to quantify the relative importance of four processes to the error growth near the tropopause.

For this case, S.T. Lang set up experiments with the operational weather-prediction model at the European Centre for Medium-Range Weather Forecasts. The PV-inversion code underlying the PV framework was developed by F. Teubler who also provided guidance in using the code and interpreting its output. V. Wirth contributed to the conceptual interpretation of the results and suggested the use of the Eady model to better understand the limited importance of tropospheric-deep (baroclinic) interaction to the error growth. M. Riemer supervised the work leading to the results in the paper and helped substantially on writing the paper. It was my work to develop the PV framework for the analysis of forecast-error growth and to perform the data analysis. I prepared all figures and wrote the paper as lead author.

Baumgart et al. (2019)

Baumgart, M., P. Ghinassi, V. Wirth, T. Selz, G.C. Craig, and M. Riemer, 2019: Quantitative view on the processes governing the upscale error growth up to the planetary scale using a stochastic convection scheme, *Mon. Wea. Rev.*, **147**, 1713–1731, https://doi.org/10.1175/MWR-D-18-0292.1

This study investigates upscale error growth from the grid scale up to the planetary scale using global simulations with a stochastic convection scheme. Quantitative insight into the processes governing the error growth is gained using the PV framework.

The simulations were carried out by T. Selz with the stochastic Plant-Craig convection scheme.

P. Ghinassi applied an envelope diagnostic to quantify the planetary-scale error growth and prepared the figures and text about the envelope diagnostic. The work of P. Ghinassi was supervised by V. Wirth, while M. Riemer supervised my work. Both contributed substantially to interpreting the results and to writing the paper. T. Selz and G.C. Craig provided useful feedback on the interpretation of the results and contributed to writing the paper. It was my part to apply the PV framework to the upscale error-growth simulations and to lead the interpretation of these results and the synthesis of the overall results. I prepared all figures except for Figs. 3, 6b, and 9 (included in this thesis as Figs. 3.3, 3.6b, and 3.9) and wrote the paper as lead author.

Baumgart and Riemer (2019)

Baumgart, M., and M. Riemer, 2019: Processes governing the amplification of ensemble spread in a medium-range forecast with large forecast uncertainty, *Quart. J. Roy. Meteor. Soc.*, **145**, 3252–3270, `https://doi.org/10.1002/qj.3617`.
This paper provides a process-based perspective on the amplification of ensemble spread. Using a PV framework, the processes governing the amplification of forecast errors and forecast uncertainty are quantified in an ensemble forecast with large forecast uncertainty.
M. Riemer supervised this work and contributed substantially to interpreting the results and to writing of the paper. It was my part to derive the tendency equation for the ensemble variance of PV, to perform the data analysis, to prepare all figures, and to write the paper as lead author.

Other publications

The following publications, to which I have contributed as a co-author, have been published (or are currently in revision) during the time of this Ph.D. project, but are not discussed in this thesis:

- **de Vries et al. (2016):** de Vries, A.J., S.B. Feldstein, M. Riemer, E. Tyrlis, M. Sprenger, M. Baumgart, M. Fnais, and J. Lelieveld : Dynamics of tropical–extratropical interactions and extreme precipitation events in Saudi Arabia in autumn, winter and spring. *Q.J.R. Meteorol. Soc.*, **142**, 1862–1880. `https://doi.org/10.1002/qj.2781`.

- **Kumpf et al. (2018):** Kumpf, A., B. Tost, M. Baumgart, M. Riemer, R. Westermann, and M. Rautenhaus: Visualizing Confidence in Cluster-based Ensemble Weather Forecast Analyses, *IEEE Transactions on Visualization and Computer Graphics*, **24**, 109–119, `https://ieeexplore.ieee.org/document/8019883/`.

- **Ghinassi et al. (2020):** Ghinassi, P., M. Baumgart, F. Teubler, M. Riemer, and V. Wirth: A budget equation for the amplitude of Rossby wave packets based on finite amplitude local wave activity, **77**, 277–296, *J. Atmos. Sci.*, `https://doi.org/10.1175/JAS-D-19-0149.1`.

Contents

CHAPTER 1

Introduction

This chapter provides a general introduction to the topics discussed in Chapters 2–4, which have been published in Baumgart et al. (2018)[1], Baumgart et al. (2019)[2], and Baumgart and Riemer (2019)[3]. A few phrases of this introduction coincide with phrases from these publications.

1.1 Motivation

Our society depends strongly on the weather. This dependence starts from the personal need of a good forecast for, e.g, the planned weekend hike, continues with the financial need of good forecasts for, e.g., the energy sector, and ends with warnings for extreme weather events like flash floods or extreme wind gusts. Accurate and reliable weather forecasts can thus facilitate every-day live, prevent economic losses, and save lives.

The history of weather forecasting based on physical laws goes back to early ideas of Abbe (1901) and Bjerknes (1904) who proposed the use of a mathematical model for weather forecasting in two steps by first determining the initial state and then predicting the future weather by time integration of the governing equations. The first weather forecast was then performed by Richardson (1922) who manually solved the governing equations. After about six weeks of computation for a six-hour forecast, he got a result that was completely unrealistic due to an imbalance in the initial conditions. With the development of computers, the first attempt to calculate numerical weather forecasts based on an approximation of the equations for large-scale dynamics was presented by Charney et al. (1950) who were able to forecast the large-scale evolution reasonably well one day ahead. Today, more than hundred years after the early work of Abbe (1901) and Bjerknes (1904), weather forecasts are produced several times every day with a much larger amount of complexity and forecast capabilities

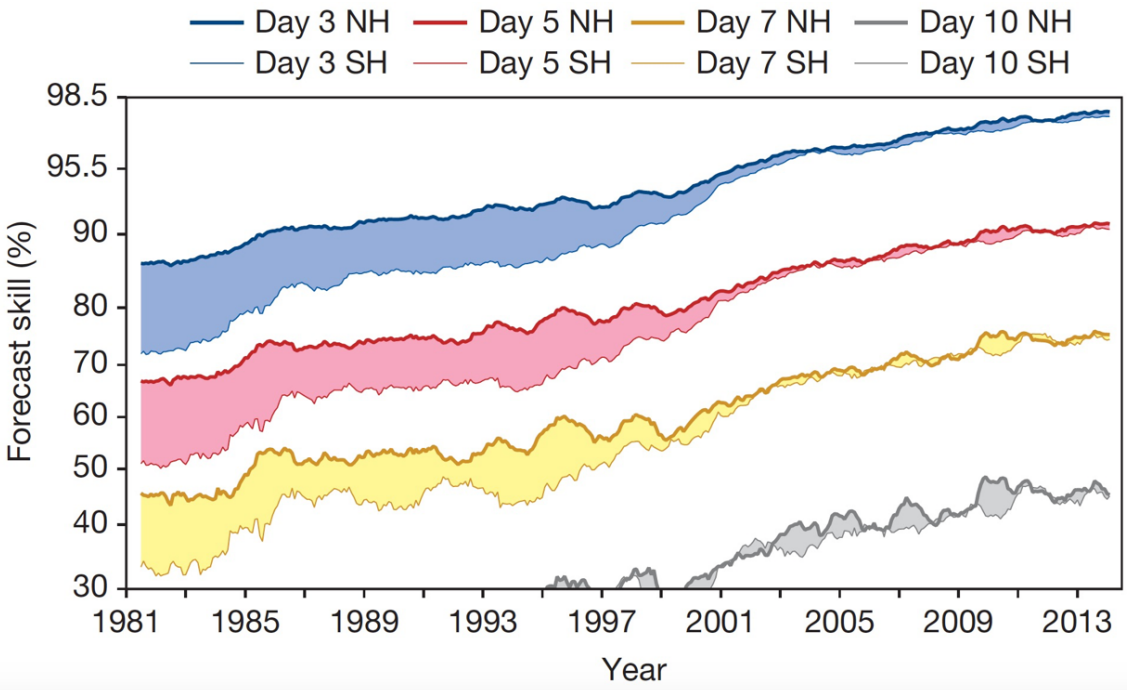

Figure 1.1: Evolution of forecast skill over the past decades for a lead time of three, five, seven and ten days. Forecast skill is measured as the anomaly correlation coefficient for 500 hPa geopotential height and computed over the extratropical Northern Hemisphere (NH) and Southern Hemisphere (SH). The figure is taken from Fig. 1 of Bauer et al. (2015).

have increased enormously (e.g., Bauer et al., 2015).

Weather forecasting is an initial-value problem. With the Navier-Stokes equation, the continuity equation, the first law of thermodynamics, the ideal gas law, and the conservation equation for water mass, there exists a full set of (nonlinear, partial-differential) equations that governs the atmospheric evolution (Bjerknes, 1904; Kalnay, 2002). Based on these equations changes in time and space can be described for individual atmospheric variables like temperature, wind, or pressure based on the initial state of the atmosphere. There exists, however, no exact solution of the governing equations. The equations thus have to be solved numerically using a discretization in space and time. Processes that are smaller and/or shorter than the discretization of the model have to be represented by parametrization schemes.

Over the past decades, the forecast skill of numerical weather-prediction models has improved significantly (e.g., Bauer et al., 2015). As a rough estimate, the predictability has improved by about 1 forecast day per decade, that means a today's 6-day forecast is as good as a 5-day forecast 10 years ago (Fig. 1.1). This "quiet revolution of numerical weather prediction", as it was named by Bauer et al. (2015), can be attributed to several improvements in terms of better observations and better weather-prediction models. There exists, however, still occasions where the forecast skill is exceptionally bad ("forecast busts", Rodwell et al., 2013). Because of the multi-stage behavior of atmospheric dynamics there may also always be an intrinsic limit of atmospheric predictability (Lorenz, 1969).

Forecast errors arise due to errors in the initial conditions of the forecast and due to model

errors in correctly representing the governing equations (e.g., Palmer and Hagedorn, 2006). These initially small errors grow both in magnitude and scale, until they influence the synoptic scale after a few forecast days (e.g., Davies and Didone, 2013; Martínez-Alvarado et al., 2014). In the following, the errors further amplify and propagate downstream (e.g., Langland et al., 2002; Anwender et al., 2008; Pantillon et al., 2013). The forecast horizon, until which a forecast can be interpreted as useful, depends on both the spatial and temporal scale under consideration (Buizza and Leutbecher, 2015). Importantly, there is also a pronounced flow dependence of atmospheric predictability (e.g., Palmer and Hagedorn, 2006; Slingo and Palmer, 2011; Ferranti et al., 2014). It is thus of large importance to better understand how forecast errors and forecast uncertainty amplify with time. For that purpose, it is crucial to gain insight into the processes governing the dynamical evolution of forecast errors and forecast uncertainty. This thesis contributes to this problem by quantifying the relative importance of individual processes, including the influence of near-tropopause Rossby-wave dynamics, baroclinic growth, and moist processes, to the flow-dependent amplification of forecast errors and forecast uncertainty.

1.2 Weather prediction

In this section we provide an overview of three aspects of weather prediction that are important for this thesis: First, a general introduction into atmospheric predictability is given both for practical and intrinsic predictability. Afterwards, the upscale propagation of forecast errors is discussed, which limits intrinsic predictability. Finally, we introduce the concept of ensemble forecasts, which provides a probabilistic framework for weather forecasting.

1.2.1 Atmospheric predictability

From the seminal work of Lorenz (1963, 1969), it is known that weather forecasts are very sensitive to the initial conditions and particularly challenging due to the interaction of multiple scales of motion. Atmospheric predictability still constitutes a major research area (e.g., Bauer et al., 2015) that has been discussed controversially in recent years (e.g., Melhauser and Zhang, 2012; Durran and Gingrich, 2014; Durran and Weyn, 2016; Sun and Zhang, 2016). One can define two types of predictability, practical and intrinsic predictability (e.g., Lorenz, 1995). Practical predictability describes the ability to predict the weather based on the best-known forecasting procedures (e.g., Lorenz, 1995). Practical predictability is thereby limited by uncertainties in the initial conditions and in the forecast model that arise from the representation of the governing equation in the model, the parametrization of subgrid-scale processes, the observations, and the data-assimilation procedures (e.g., Melhauser and Zhang, 2012; Sun and Zhang, 2016). Intrinsic predictability, in contrasts, describes the ability to predict the weather in the presence of optimal forecast procedures and nearly perfect initial conditions (e.g., Lorenz, 1995). This intrinsic predictability originates from the multi-scale, chaotic nature of the atmosphere and sets a finite intrinsic limit to weather forecasting that cannot be

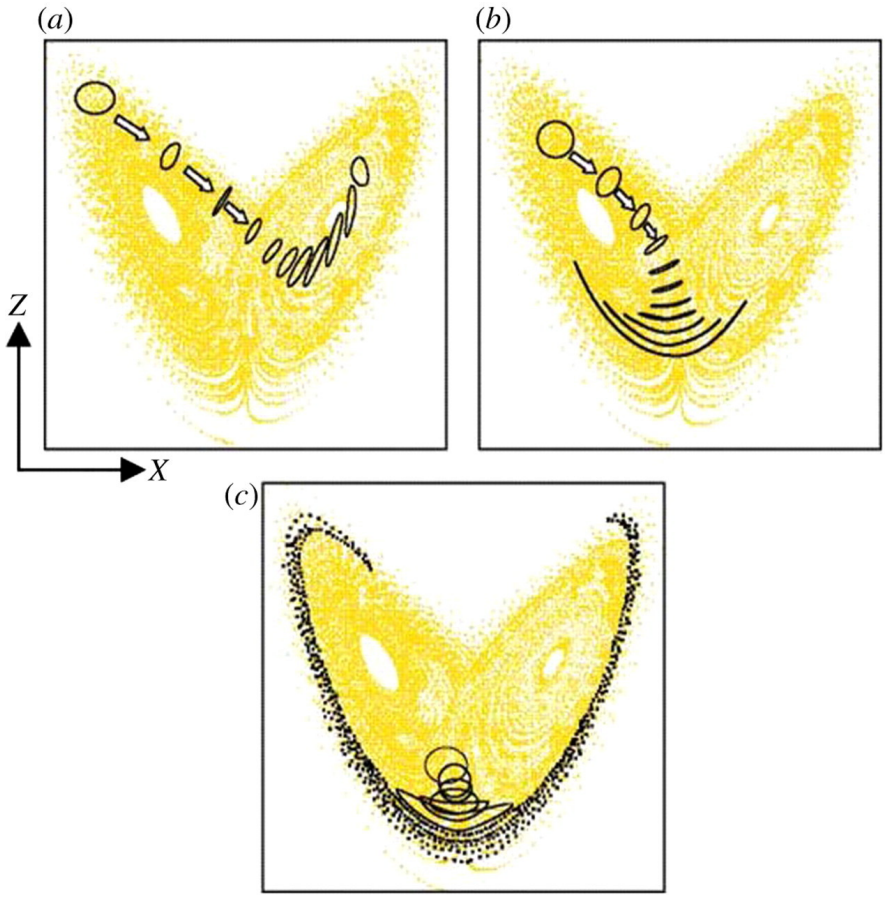

Figure 1.2: Finite-time error growth in the Lorenz (1963) system. The yellow shading denotes all possible states ("climatological distribution") in the Lorenz (1963) system, while the first black circle denotes the initial-condition uncertainty of an ensemble forecast. The arrows and further circles illustrate the evolution of forecast uncertainty. The initial forecast uncertainty is confined to different regimes of the system, yielding a) a high-predictability situation, b) a high-predictability situation in the short term, but a low-predictability situation in the long term, and c) a low-predictability situation. The figure is taken from Fig. 1 of Slingo and Palmer (2011).

extended by improving the forecast system (Lorenz, 1969). For current weather forecasts, predictability is often limited by practical predictability (e.g., Durran and Gingrich, 2014; Durran and Weyn, 2016). Predictability, however, is highly flow-dependent (e.g., Palmer and Hagedorn, 2006; Slingo and Palmer, 2011; Ferranti et al., 2014) and can in some cases also be limited by intrinsic predictability (e.g., Melhauser and Zhang, 2012; Sun and Zhang, 2016). The flow-dependent aspect of predictability can be illustrated with the simple Lorenz (1963) system as a low-order chaotic system consisting of a set of deterministic ordinary nonlinear partial-differential equations. Depending on the initial conditions, the predictability of this system differs largely: some flow situations yield high predictability also in the long term (Fig. 1.2a), other flow situations yield high predictability only in the short term (Fig. 1.2b), and again other flow situations yield low predictability even in the short term (Fig. 1.2c). The famous metaphor "Does the flap of a butterfly's wings in Brazil set off a tornado in

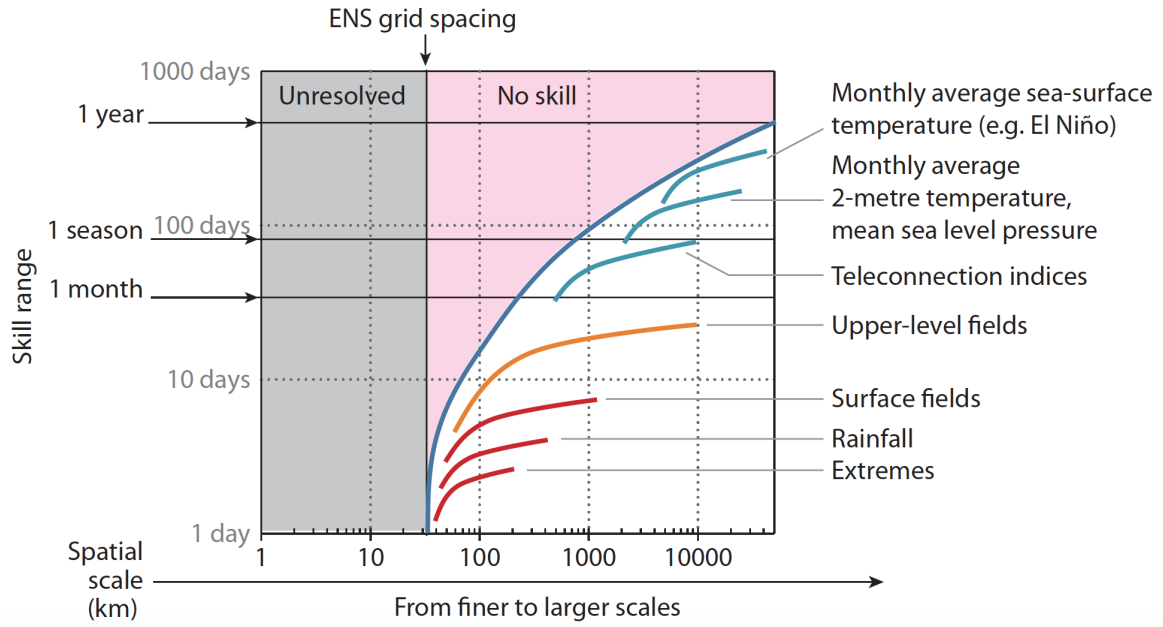

Figure 1.3: Schematic illustration of the forecast-skill horizon of the ECMWF operational forecasts in dependence of spatial scale. Forecast-skill horizon is defined using the continuous ranked probability score (Brown, 1974; Hersbach, 2000) as forecast metric for medium-range, monthly, and seasonal forecasts. The figure is taken from Fig. 5 of Buizza et al. (2015).

Texas?", which was the title of a presentation from Edward N. Lorenz in 1972[4], illustrates this sensitive dependence on the initial conditions. This so-called "butterfly effect" expresses also another important, and maybe even more important, characteristic of atmospheric predictability, namely that multiple scales of motion interact in the atmosphere and that errors on the small scale (in the metaphor the "butterfly") can be responsible for errors on the larger scale (in the metaphor the "tornado").

Buizza and Leutbecher (2015) analyzed the current practical predictability limit, which they refer to as the "forecast-skill horizon", for different spatial and temporal scales using ensemble forecasts (the concept of ensemble forecasting will be introduced in Sec. 1.2.3) from the European Centre for Medium-Range Weather Forecasts (ECMWF). The authors defined the forecast-skill horizon as the time when the ensemble forecast stops being more skillful than the climatological distribution. For instantaneous grid-point fields, the authors derived a forecast-skill horizon of 16–23 days, while they found a considerably longer forecast-skill horizon for spatially and temporally averaged fields (Fig. 1.3). In addition to the dependence on the spatio-temporal scale, the authors showed that the forecast-skill horizon depends also on the variable, the region, and the season under consideration.

Recently, Selz (2019) used upscale error-growth simulations (see Sec. 1.2.2) to estimate an intrinsic predictability limit of about 17 days, consistent with the intrinsic predictability limit of around 2–3 weeks that was derived by Judt (2018). This even shorter intrinsic predictability limit than the one derived by Buizza and Leutbecher (2015) for practical predictability

[4]Talk at the 139th meeting of the American Association for the Advancement of Science in Washington, D.C., on December 29, 1972.

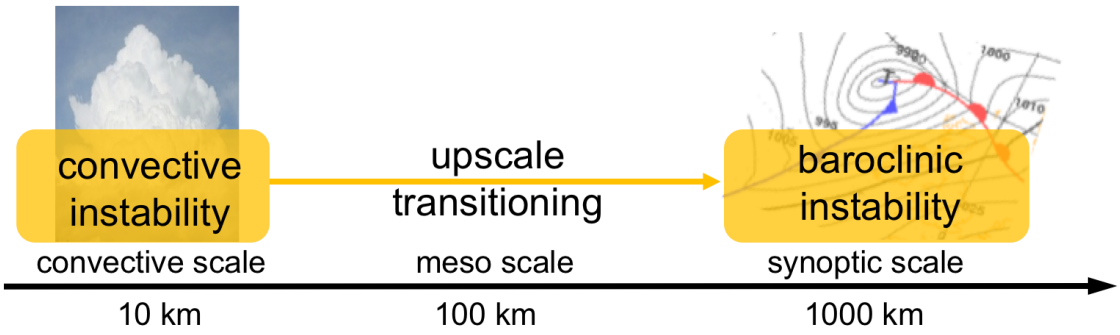

Figure 1.4: Schematic diagram illustrating the three-stage error-growth model derived by Zhang et al. (2007) (see text for details).

can be attributed to the use of another forecast-error metric and a smaller data set. By comparing the results from the upscale error-growth simulations with the initial-condition uncertainty of the current ECMWF data-assimilation system, Selz (2019) could nevertheless derive a potential improvement of about 3.5 days through perfecting the initial conditions in current weather-prediction models. In a further study, Zhang et al. (2019) found a similar potential improvement of 3–5 days. Both studies thus suggest that there is still a distinct gap between the intrinsic predictability limit and the current practical predictability limit. There is thus still large potential for further improvements in weather forecasting.
In this thesis, we investigate aspects of both practical and intrinsic predictability: In Chapter 2, we analyze the forecast-error growth in a deterministic forecast of a state-of-the-art weather-prediction model from a process-based perspective, which provides quantitative insight into practical predictability. Further insight into practical predictability is gained by analyzing the amplification of ensemble spread for an operational forecast associated with large forecast uncertainty (Chapter 4). In Chapter 3, we quantify the processes governing upscale error growth and focus thereby on intrinsic predictability.

1.2.2 Upscale error growth

Small-scale error growth associated with moist convection is about ten times faster than synoptic-scale error growth and errors saturate about ten times faster on the small scale than on the synoptic scale (Hohenegger and Schär, 2007). Due to the multi-scale nature of atmospheric flow, errors on the smallest resolved scale amplify both in magnitude and in scale and can thereby limit predictability on the synoptic scale (Lorenz, 1969). This upscale error growth sets an intrinsic limit to atmospheric predictability (see Sec. 1.2.1) and is thus of large importance to weather forecasting.
Using idealized baroclinic wave simulations, Zhang et al. (2007) derived a conceptual model for upscale error growth that was, in general, confirmed for realistic weather situations (Selz and Craig, 2015b; Judt, 2018). This model consists of three different stages as indicated in Fig. 1.4. In the first stage, errors grow on the small scale driven by convective instability. Error growth in this stage quickly saturates when individual convective cells become completely

decorrelated (Zhang et al., 2003, 2007). In the second stage, error growth is characterized by a transitioning from a small-scale unbalanced error pattern to a large-scale balanced error pattern. Recent work (Bierdel et al., 2017, 2018) shows that changes in the net divergence forced by convection are most important to the error growth in this stage, while the spreading of gravity waves has only a transient effect on the error growth. In the final stage, error growth is dominated by the balanced, large-scale part of the flow. This stage is related to baroclinic instability in the conceptual model of Zhang et al. (2007).
Previous studies on upscale error growth are mostly based on error-energy spectra (e.g., Zhang et al., 2007; Selz and Craig, 2015b; Judt, 2018; Selz, 2019). These spectra provide direct insight into the spatial scale of the error growth. The relation to the processes governing the error growth, however, is only based on qualitative arguments. In this thesis (Chapter 3), we provide a quantitative view on the processes governing upscale error growth and thereby novel insight into the upscale error-growth mechanisms.

1.2.3 Ensemble forecasts

In the past 25 years, weather forecasting has experienced a paradigm shift from a deterministic approach to a probabilistic approach (e.g., Buizza, 2019): Instead of just providing a single, deterministic weather forecast, current state-of-the-art weather-prediction models are producing ensembles of about 20–50 forecasts to provide an estimate of the future range of possible scenarios and of the uncertainty associated with a forecast. This paradigm shift towards probabilistic forecasts was triggered by the recognition that atmospheric predictability is limited due to uncertainties in the initial conditions and in the forecast model (see Sec. 1.2.1). Today's ensemble-prediction system account for both types of uncertainty by introducing perturbations to the initial conditions and to the model physics (e.g., Palmer and Hagedorn, 2006) as schematically illustrated in Fig. 1.5 (Slingo and Palmer, 2011): Starting from only small perturbations at initial time, the individual ensemble members diverge from each other, leading to a pronounced amplification of ensemble spread with forecast-lead time. As predictability is flow-dependent, the evolution of ensemble spread is also largely dependent on the flow situation (e.g., Palmer and Hagedorn, 2006; Slingo and Palmer, 2011; Ferranti et al., 2014). Going back to the simple Lorenz (1963) system, this characteristic can be illustrated schematically: Starting with the same amplitude of initial-condition uncertainty, the forecast uncertainty in the three examples indicated in Fig. 1.2 evolves rather differently with some flow situations having much higher predictability than other flow situations.
Defining perturbations for ensemble forecasts is a nontrivial task that has shown major achievements over the past 25 years, but also remains a key challenge for the future (e.g., Bauer et al., 2015; Buizza, 2019). At the different weather-prediction centers, there exists a wide range of methods to define ensemble perturbations (e.g., Buizza, 2019). Today's operational ensemble-forecasting model at ECMWF employs three different perturbation strategies (ECMWF, 2018): First, perturbations associated with the initial-condition uncertainty are introduced using the ensemble data assimilation (EDA) system and the leading singular vec-

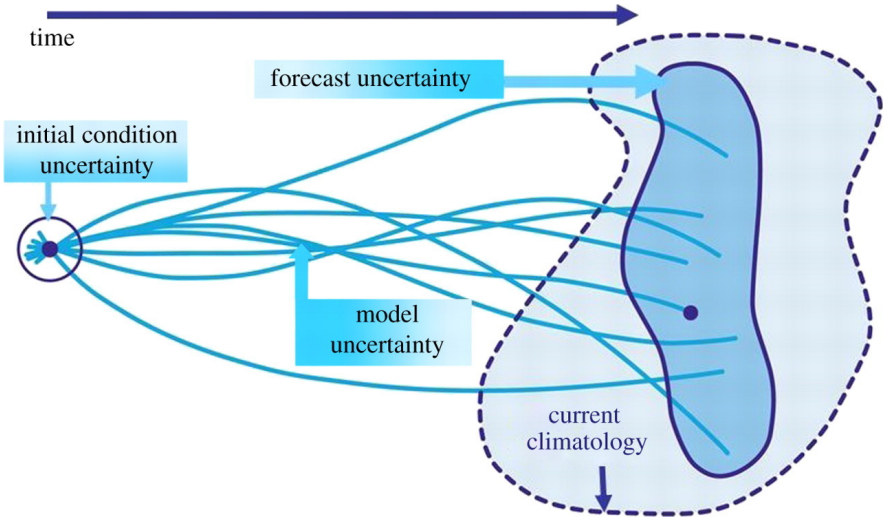

Figure 1.5: Schematic illustration of an ensemble forecast. Trajectories of individual forecasts (blue lines) diverge due to the uncertainty in the initial conditions and in representing subgrid-scale processes in the model. The range of possible states as estimated by the model predictions is represented by the solid envelope, while the dashed envelope represents the climatological range of possible states of the atmosphere. The figure is taken from Fig. 2 of Slingo and Palmer (2011).

tors (SVs). In the EDA system, an ensemble of analyses is generated by randomly perturbing the observation values in a way that is consistent with the statistical characteristics of the observation error (Buizza et al., 2008; Isaksen et al., 2010). These perturbed analyses should provide a reliable estimate of the initial-condition uncertainty if all relevant sources of uncertainty were represented. By only using EDA perturbations, ensemble spread is, however, underestimated (Buizza et al., 2008). ECMWF thus uses a combination of EDA perturbations and singular vectors, which are designed to maximize the perturbation growth over a finite time interval (Buizza and Palmer, 1995). Using this combination leads to a more skillful representation of the initial-condition uncertainty at ECMWF (Buizza et al., 2008). Uncertainties associated with the forecast model are represented using the Stochastically Perturbed Parameterization Tendencies (SPPT) scheme (ECMWF, 2018). The SPPT scheme randomly perturbs the tendencies from the physical parameterization schemes to represent uncertainties associated with the representation of the parameterized physical processes (Buizza et al., 1999).

Over the past 25 years, the accuracy of ensemble forecast has improved significantly (e.g., Buizza, 2019; Rodwell et al., 2018). In addition to forecast accuracy, a second important characteristic of the ensemble-forecast quality is forecast reliability, which measures whether the forecast distribution matches the observed frequency of occurrence. Concerning annual-mean averages over the Northern Hemisphere, forecast reliability has also improved significantly over the past decades (see, e.g., Fig. 1a of Rodwell et al., 2018). Concerning the day-to-day variability in more localized regions, there exists, however, still deficiencies in flow-dependent reliability (see, e.g., Fig. 1b of Rodwell et al., 2018). One crucial aspect to improve flow-dependent reliability is to better understand the dynamical evolution of ensemble spread. Rodwell et al. (2018) presented an approach to gain insight into deficiencies in the short-term

reliability, highlighting the importance of moist processes as sources for forecast uncertainty. In this thesis (Chapter 4), we provide a complementary approach and focus on the further amplification of initially small ensemble perturbations. Using a potential-vorticity framework, which will be introduced in the next section, we quantify the relative importance of individual processes governing the flow-dependent amplification of ensemble spread. This insight can be useful for the interpretation, and possible also the design, of future ensemble-prediction systems.

1.3 Potential vorticity

Potential vorticity is a key variable for dynamical meteorology as it provides a compact and coherent theoretical framework to analyze midlatitude dynamics. This section provides a general introduction into potential vorticity and introduces the potential-vorticity framework that is used in this thesis to partition the relative importance of individual processes to the amplification of forecast errors and forecast uncertainty.

1.3.1 Definition

Potential vorticity (PV) is a variable that combines thermodynamic and kinematic properties of atmospheric flow. The nonhydrostatic version of Ertel (1942) PV is defined as

$$P = \frac{1}{\rho}\boldsymbol{\eta} \cdot \boldsymbol{\nabla}\theta \ , \tag{1.1}$$

where ρ denotes the air density, $\boldsymbol{\eta}$ the absolute vorticity, $\boldsymbol{\nabla}$ the three-dimensional nabla operator, and θ the potential temperature. Absolute vorticity is defined as $\boldsymbol{\eta} = \boldsymbol{\nabla} \times \mathbf{u} + 2\boldsymbol{\Omega}$ where $\mathbf{u}$ denotes the three-dimensional wind field and $\boldsymbol{\Omega}$ the angular velocity of the Earth's rotation. PV is often measured in potential-vorticity units (PVU), which are defined as 1 PVU $= 10^{-6}$ K m^2 s^{-1} kg^{-1}. In the primitive-equation framework and isentropic coordinates, Ertel-PV is given by

$$P = -g\frac{\partial\theta}{\partial p}(\zeta_\theta + f) \ , \tag{1.2}$$

where g is the gravitational acceleration, θ the potential temperature, p the pressure, ζ_θ the vertical component of relative vorticity on an isentropic surface, and f the Coriolis parameter. In this thesis, we use this primitive-equation form of Ertel-PV.

1.3.2 Climatological distribution

From a climatological perspective, PV increases upward and poleward (Fig. 1.6). In the troposphere, PV is generally smaller than 1.5 PVU, while it is generally larger than 4 PVU in the stratosphere (e.g., Hoskins et al., 1985). The tropopause is characterized by a large PV gradient. An isosurface of PV such as the 2 PVU-surface can thus be used to define the dynamical tropopause.

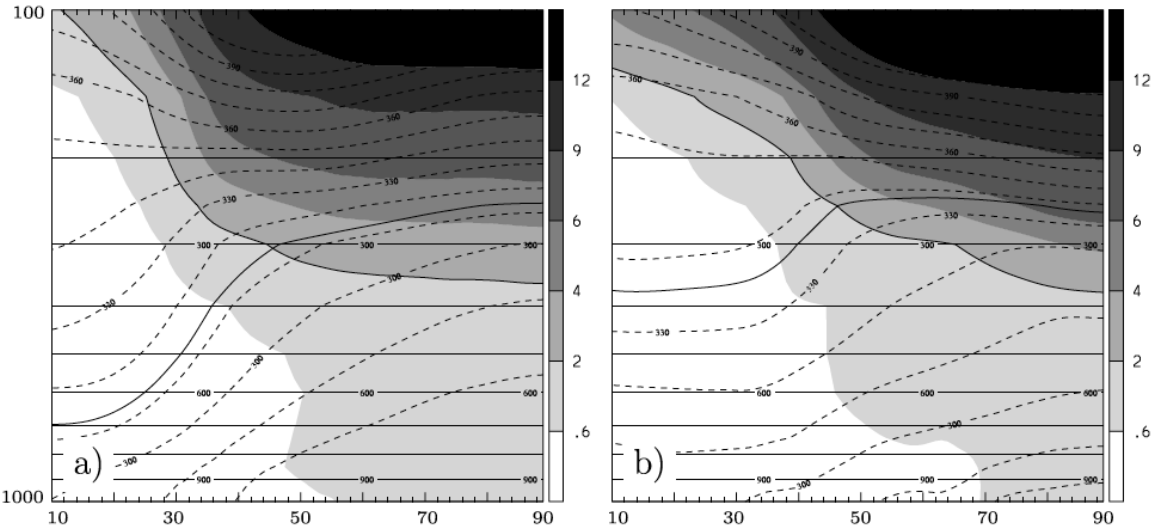

Figure 1.6: Latitude-height cross section of the zonal mean climatological distribution of PV (shading) and of potential temperature (dashed contours with an interval of 10 K) on the Northern Hemisphere averaged from 1979 to 2003 for a) January and b) July. The solid line indicates the isentrope that intersects the tropopause at 45° N. Pressure is used as height coordinates (solid horizontal lines with a spacing of 100 hPa). The figure is taken from Fig. 1 of Liniger and Davies (2004).

The zonal-mean distribution of PV and potential temperature exhibits a pronounced seasonality as shown in Fig. 1.6 (Liniger and Davies, 2004). For dynamical analyses in the midlatitudes, it is useful to analyze PV on those isentropes that intersect the tropopause in the midlatitudes (e.g., at 45°N) as the PV evolution near the tropopause is of particular importance to the evolution of midlatitude weather systems (e.g., Hoskins et al., 1985). In winter this intersection corresponds to isentropes around 310 K and in summer to isentropes around 335 K (Fig. 1.6).

1.3.3 Main characteristics

PV has three characteristics that are particularly useful for dynamical analyses (Hoskins et al., 1985):

1. **PV conservation:** Under conservative (adiabatic and frictionless) conditions, PV is conserved materially, that means
$$\frac{DP}{Dt} = 0 \ , \tag{1.3}$$
where $\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla$. Under such conditions, the PV of an air parcel is simply advected with the flow. As potential temperature is also conserved under adiabatic and frictionless conditions, the advection of PV in isentropic coordinates is (quasi-)horizontally under conservative conditions. Maps of PV on isentropic surfaces provide a sharp picture on atmospheric dynamics and can be used to identify salient flow features. This characteristic is a large advantage compared to other atmospheric variables as, e.g., geopotential, which exhibit a much smoother pattern than PV.

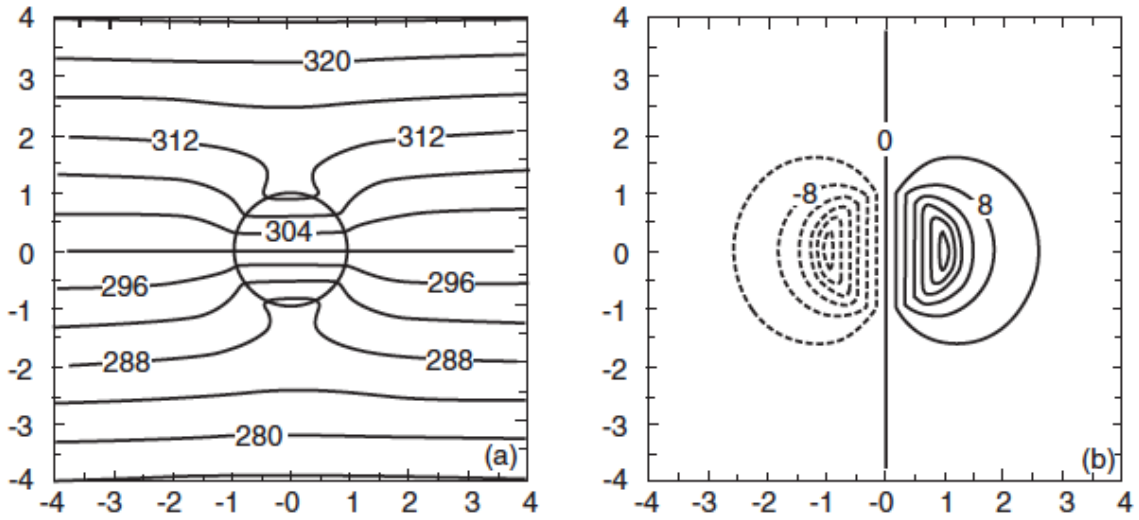

Figure 1.7: Vertical cross section of a) potential temperature (contour interval 4 K) and b) normal component of geostrophic velocity (contour interval 2 m s^{-1}) associated with a compact positive potential-vorticity anomaly (PV "ball" indicated by the circle in a)) that is added to a constant standard-atmosphere PV distribution. Solid and dashed contours in b) denote flow into and out of the page, respectively. Spatial distances are normalized by the radius of the PV anomaly and scaled in the vertical direction. For a PV anomaly with 250 km radius, the distances shown correspond to 2000 km in the horizontal direction and 20 km in the vertical direction. The figure is adapted from Fig. 1b of Thorpe and Bishop (1995) and taken from Fig. 6.10 of Holton and Hakim (2013).

The contributions from diabatic processes and friction to the PV evolution can be quantified, complementing the conservative view on the PV evolution. Including these two contributions, the material change of PV is given by

$$\frac{DP}{Dt} = \frac{1}{\rho}\boldsymbol{\eta} \cdot \nabla\dot{\theta} + \frac{1}{\rho}(\nabla \times \mathbf{F}) \cdot \nabla\theta \ , \tag{1.4}$$

where $\dot{\theta}$ denotes the diabatic heating rate and $\mathbf{F}$ the frictional forces. On the synoptic scale in isentropic coordinates, the diabatic PV modification can be approximated by (e.g., Davies and Didone, 2013)

$$\frac{DP}{Dt} = P\frac{\partial\dot{\theta}}{\partial\theta} \ . \tag{1.5}$$

In regions of latent heating, a positive PV anomaly is thus generally produced beneath the maximum heating rate and a negative PV anomaly above.

2. **PV inversion:** The PV distribution contains information about the flow and mass distribution. Together with a balance condition, a reference state, and suitable boundary conditions, the flow and mass distribution can be deduced from a PV distribution by the inversion of a Laplace operator. Qualitatively, the flow and temperature of a localized positive PV anomaly[5] is illustrated in Fig. 1.7 (Thorpe and Bishop, 1995): Positive PV anomalies are, in general, associated with a cyclonic circulation, a positive static stability

[5] Anomalies denote deviations from a reference state as, e.g., in Fig. 1.7 from the constant standard atmosphere.

anomaly inside the PV anomaly, and a negative static stability anomaly above and beneath the PV anomaly. For a negative PV anomaly, the corresponding circulation and static stability anomalies have opposite sign. Importantly, PV anomalies have a far-field effect on the mass and flow distribution, that means that the flow and static-stability anomalies associated with a PV anomaly are larger in scale than the PV anomaly itself. The largest PV anomalies are found near the tropopause and in terms of potential temperature anomalies just above the boundary layer. These temperature anomalies can be interpreted as infinitesimally shallow PV anomalies (Bretherton, 1966). A warm anomaly just above the boundary layer can thus be interpreted as a positive PV anomaly and a cold anomaly as a negative PV anomaly.

3. **PV partitioning:** The PV distribution exhibits characteristic patterns related to individual weather systems. By treating the PV distribution as a composite of a reference state and individual PV perturbations and performing separate (piecewise) PV inversions (PPVI), it is possible to obtain the flow and mass distribution associated with individual PV perturbations (Davis and Emanuel, 1991; Davis, 1992). It is thereby possible to quantify the influence of individual weather systems on the flow.
PPVI has become a valuable tool for dynamical analyses. Previous studies applied it, e.g., to analyze midlatitude cyclogenesis (e.g., Davis and Emanuel, 1991; Davis et al., 1996; Stoelinga, 1996) or to analyze Rossby-wave dynamics (e.g., Nielsen-Gammon and Lefevre, 1996; Teubler and Riemer, 2016). In addition, PPVI has often been applied in the context of the extratropical transition of tropical cyclones (e.g., McTaggart-Cowan et al., 2001; Agustí-Panareda et al., 2004, 2005; Riemer et al., 2008; Riemer and Jones, 2010; Grams et al., 2013).

1.3.4 PV thinking for midlatitude dynamics

Based on the seminal work of Hoskins et al. (1985), "PV thinking" has become an established concept to describe midlatitude dynamics. PV thinking combines the main characteristics of PV (see Sec. 1.3.3): Based on PV inversion, one can derive the flow and temperature associated with a PV distribution. Under conservative conditions, one can then calculate the local change of PV based on PV conservation and thus determine the time evolution of PV. Finally, with PV partitioning one can identify the importance of different weather systems to the flow evolution.
The concept of PV thinking can be used to describe several atmospheric processes, including the downstream development of Rossby waves and baroclinic instability. These two processes will be discussed in more detail in the following, together with the influence of diabatic processes and upper-tropospheric divergence on Rossby-wave dynamics. In this thesis, these four processes will be quantified using a PV framework (see Sec. 1.3.5).
From the PV perspective, the barotropic downstream development of Rossby waves can be explained based on qualitative arguments (Hoskins et al., 1985): As illustrated schematically in Fig. 1.8 (Wirth et al., 2018), a trough (southward deflection of a PV contour) can be inter-

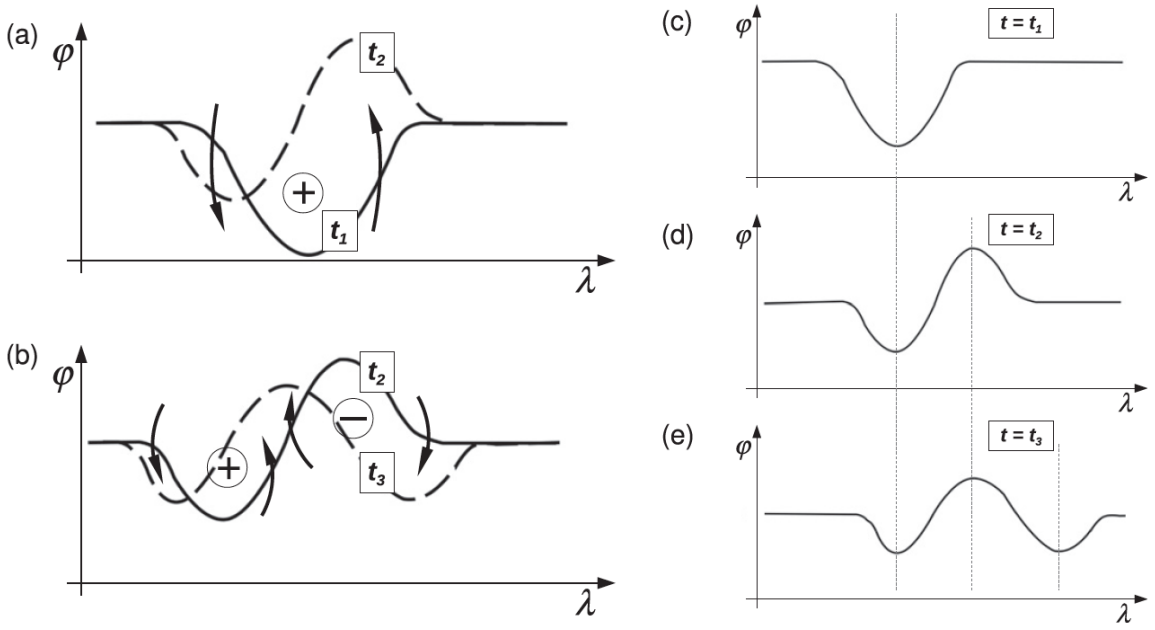

Figure 1.8: Schematic diagram illustrating the barotropic downstream development from the PV perspective in longitude-latitude maps for a background atmsophere with northward PV gradient. Left: Black lines indicate lines of constant PV at different time steps (earlier time in solid, later time in dashed) for a) the time interval $t_1 \rightarrow t_2$ and b) the time interval $t_2 \rightarrow t_3$. The signs + and − indicate the individual PV anomalies in the Rossby wave. Arrows denote the circulation associated with the PV anomalies. Right: Stationary scenario for the phase of troughs and ridges when a westerly basic flow is added that exactly opposes the westerly phase propagation. The figure is taken from Fig. 8 of Wirth et al. (2018).

preted as a positive PV anomaly in a background atmosphere with northward PV gradient (as indicated at time t_1 in Fig. 1.8a). This PV anomaly is associated with a cyclonic flow, which leads to a southward deflection of the PV contour on the western side of the PV anomaly and a northward deflection on the eastern side. The positive PV anomaly is thereby advected westward, while a new ridge (negative PV anomaly) develops downstream of the positive PV anomaly at time t_2 (Fig. 1.8a). These two PV anomalies are again associated with a flow pattern that leads to a westward shift of the existing PV anomalies and to the generation of a new trough at time t_3 (Fig. 1.8b). Adding a westerly flow of exactly the same magnitude as the westward phase speed of the troughs and ridges, yields a stationary scenario for the phase of the troughs and ridges (Fig. 1.8 right panel). In this scenario, the generation of new PV anomalies downstream of the existing PV anomalies, as manifested in the classical picture of downstream development (e.g., Hoskins et al., 1985), becomes clearly evident.

In a baroclinic atmosphere, vertical interaction occurs between upper- and lower-level PV anomalies. The mutual interaction between PV anomalies near the tropopause and near the surface can be explained in the simple two-layer model of Eady (1949), which consists of one upper-level and one lower-level layer. As indicated in Fig. 1.9a, an upper-level PV anomaly approaching a lower-level baroclinic zone will be associated with a circulation at lower levels. This circulation leads to the generation of a warm anomaly at the lower level, which can be

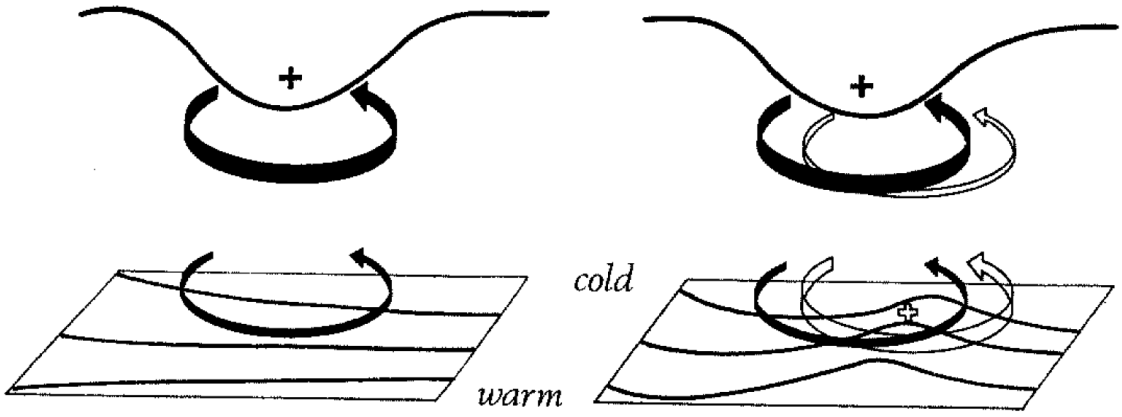

Figure 1.9: Schematic diagram illustrating baroclinic growth from the PV perspective. The approach of an upper-level positive PV anomaly (solid line at top) over a baroclinic zone (lower panel) is illustrated with the corresponding wind field associated with the upper-level PV anomaly (solid arrows). b) A warm anomaly is created from the low-level circulation of the upper-level PV anomaly, which in turn is associated with a circulation (open arrows). The figure is taken from Fig. 21 of Hoskins et al. (1985).

interpreted as a positive PV anomaly (Bretherton, 1966). This anomaly is in turn associated with a cyclonic circulation at upper-levels. In this scenario, the vertical interaction between the upper- and lower-level anomalies leads to a mutual amplification of the upper- and lower-level anomalies, which is the essence of baroclinic instability (Eady, 1949; Hoskins et al., 1985; Heifetz et al., 2004).

As noted in Sec. 1.3.3, diabatic processes can lead to a material change of PV and thereby modify the PV distribution directly. The role of diabatic processes for Rossby-wave dynamics was the focus of recent research (e.g., Grams et al., 2011; Chagnon et al., 2013). As illustrated schematically in Fig. 1.10, diabatic processes generate positive PV anomalies on the stratospheric side of the tropopause and negative PV anomalies on the tropospheric side, resulting in an enhancement of the PV gradient near the tropopause (Chagnon et al., 2013). The positive PV anomalies originate mostly from longwave radiative cooling, whereas the negative PV anomalies originate mostly from latent heating (Chagnon et al., 2013). The role of diabatic processes has also attained recent attention in the context of forecast-error growth (e.g., Gray et al., 2014; Martínez-Alvarado et al., 2016; Grams et al., 2018, see also Sec. 1.3.6).

Diabatic processes play also an important indirect role for Rossby-wave dynamics (e.g., Davis et al., 1993) due to (i) balanced PV anomalies generated by diabatic processes and (ii) upper-tropospheric divergence, which is often associated with latent release below. Previous work shows that the indirect diabatic effect from upper-tropospheric divergence is often of larger importance to Rossby-wave dynamics than the direct diabatic PV modification (Davis et al., 1993; Riemer and Jones, 2010; Teubler and Riemer, 2016). As illustrated schematically in Fig. 1.11, upper-tropospheric divergence leads to an area increase of Rossby-wave ridges and is thus of large importance during ridge-building (e.g., Davis et al., 1996; Riemer and Jones, 2010; Teubler and Riemer, 2016).

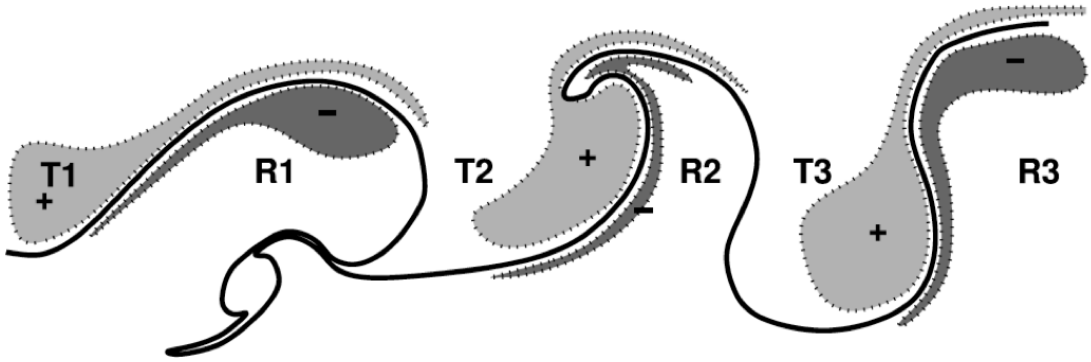

Figure 1.10: Schematic diagram illustrating the influence of diabatic processes on the PV evolution near the tropopause. The solid contour denotes the dynamical tropopause, while the shading denotes the net diabatic PV anomalies in individual troughs (index T) and ridges (index R). The figure is taken from Fig. 10 of Chagnon et al. (2013).

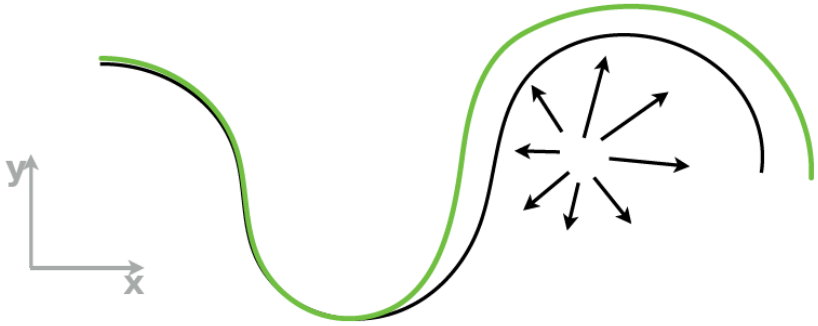

Figure 1.11: Schematic diagram illustrating the influence of upper-tropospheric divergence on the Rossby-wave evolution near the tropopause. The black contour denotes the Rossby-wave pattern at initial time, while the green contour denotes the Rossby-wave pattern at a later time. Arrows depict the divergent outflow near the tropopause. The figure is taken from Fig. 2c of Teubler and Riemer (2016).

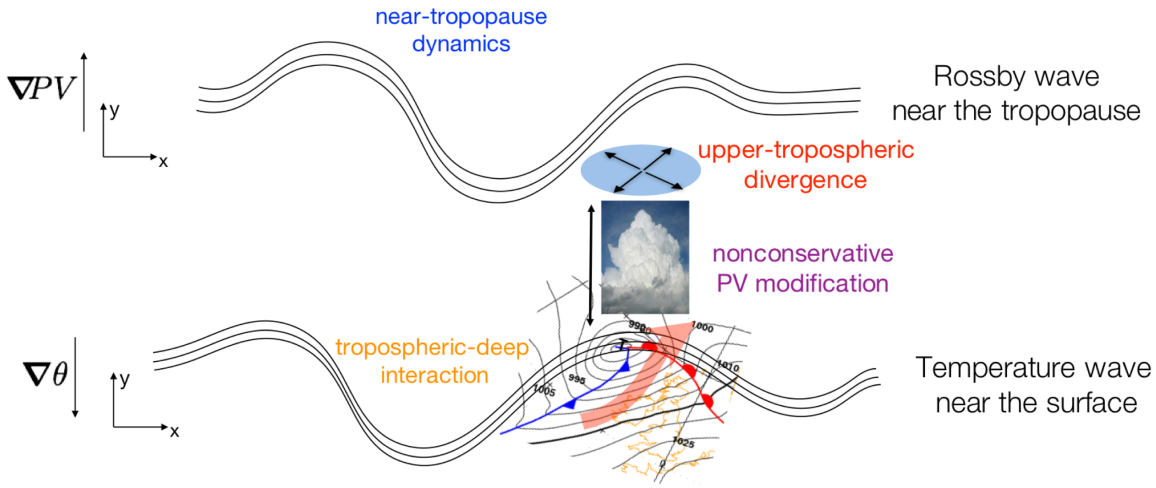

Figure 1.12: Schematic diagram illustrating the four processes partitioned in the PV framework (see text for details). Courtesy of F. Teubler with slight modifications.

1.3.5 PV framework to quantify the influence of individual processes

In this study, we use a PV framework to quantify the relative importance of individual processes to the amplification of forecast errors and forecast uncertainty. For that purpose, we use the same partitioning of processes as in a previous analysis of Rossby-wave dynamics (Teubler and Riemer, 2016). This PV framework is based on the PV perspective of midlatitude dynamics (as described in the previous subsection, Hoskins et al., 1985) and partitions the influence of four processes, namely near-tropopause dynamics, tropospheric-deep interaction, upper-tropospheric divergence, and nonconservative PV modification (Fig. 1.12).
The influence of near-tropopause dynamics, tropospheric-deep interaction, and upper-tropospheric divergence is determined using a flow partitioning: First, the nondivergent and divergent flow is partitioned using a Helmholtz partitioning (Lynch, 1989). The nondivergent flow is then further partitioned into the flow associated with upper- and lower-level PV anomalies, respectively, using PPVI (Davis and Emanuel, 1991; Davis, 1992). The influence of upper-level (near-tropopause) PV anomalies on the upper-level (near-tropopause) evolution provides a quasi-two-dimensional view on the dynamics and is mostly related to nonlinear Rossby-wave dynamics. The influence of lower-level PV anomalies on the near-tropopause evolution describes the influence of the vertical interaction between upper- and lower-level dynamics and includes the influence of baroclinic instability. The influence of upper-tropospheric divergence can be related to balanced dynamics and moist processes (see, e.g., Chapter 6.4 in Holton and Hakim, 2013). For Rossby-wave dynamics, a strong influence of upper-tropospheric divergence is often associated with latent heat release below (e.g., Davis et al., 1993; Riemer et al., 2014; Quinting and Jones, 2016).
The nonconservative processes associated with radiative transfer, turbulent mixing, subgrid-scale orographic drag, non-orographic gravity wave drag, moist convection, clouds, and surface/soil processes (see Fig. 1.13 for an overview of all processes) exhibit generally a small

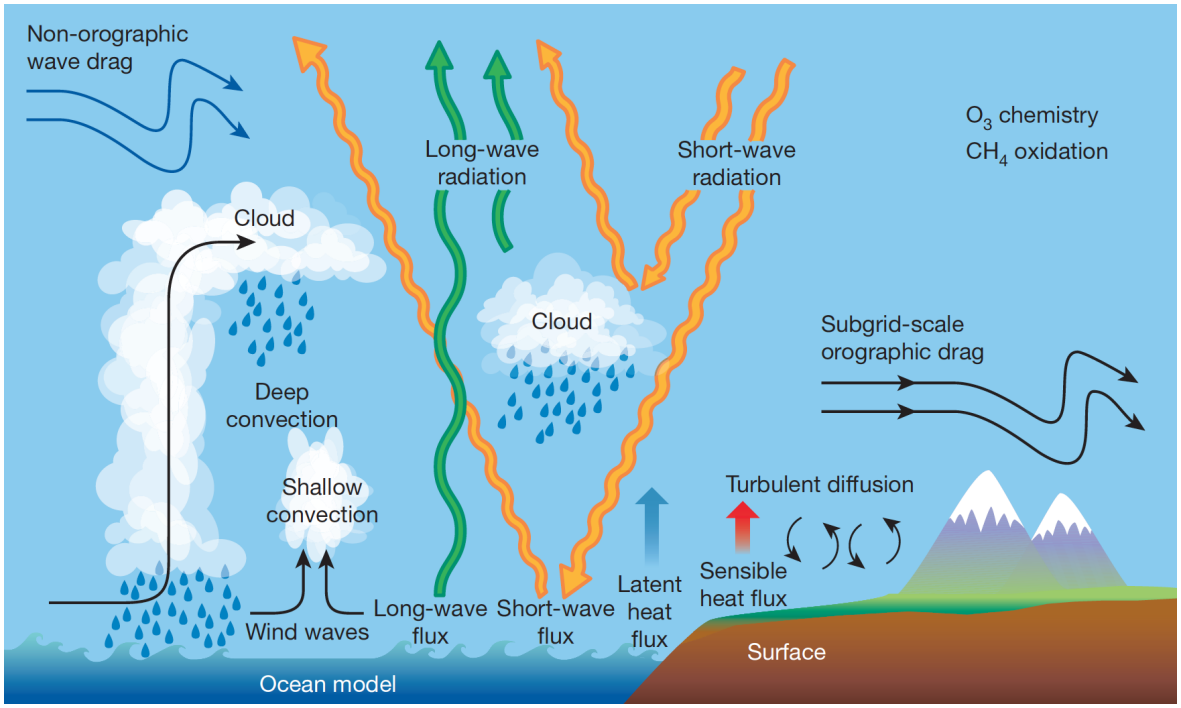

Figure 1.13: Schematic diagram illustrating the different physical processes that are represented in current numerical weather-prediction models via parametrization schemes. The figure is taken from Fig. 2 of Bauer et al. (2015).

spatial scale and are therefore not explicitly resolved in global models (ECMWF, 2013). These processes, however, are of relevant importance to weather prediction as they drive the heat and momentum budget (ECMWF, 2013). Their influence on the grid scale is thus represented by parametrization schemes. The parametrization schemes in the Integrated Forecasting System (IFS) model at ECMWF, for example, provide heating rates from the cloud, convection, radiation, and turbulence parameterization, and wind tendencies from the convection and turbulence parameterization (ECMWF, 2013). In this study, we use these tendencies from the parametrization schemes to quantify the direct PV modification by nonconservative processes due to diabatic heating and nonconservative momentum change.

1.3.6 PV perspective on error growth

The PV perspective has become a valuable perspective to gain insight into the dynamics of forecast errors. Early work demonstrated that a misrepresentation of the PV distribution near the tropopause can result in rapid error growth in individual subsynoptic and synoptic weather systems (e.g., Fehlmann and Davies, 1997, 1999). The PV error maximizes, in general, in regions with a strong PV gradient, as it occurs near the tropopause (e.g., Snyder et al., 2003; Davies and Didone, 2013). By analyzing a case study and a single winter climatology, Dirren et al. (2003) showed that the PV error ("forecast minus analysis" PV) provides insight into the nature and evolution of forecast errors.

Climatological studies of the PV error can be used to identify systematic errors of numerical

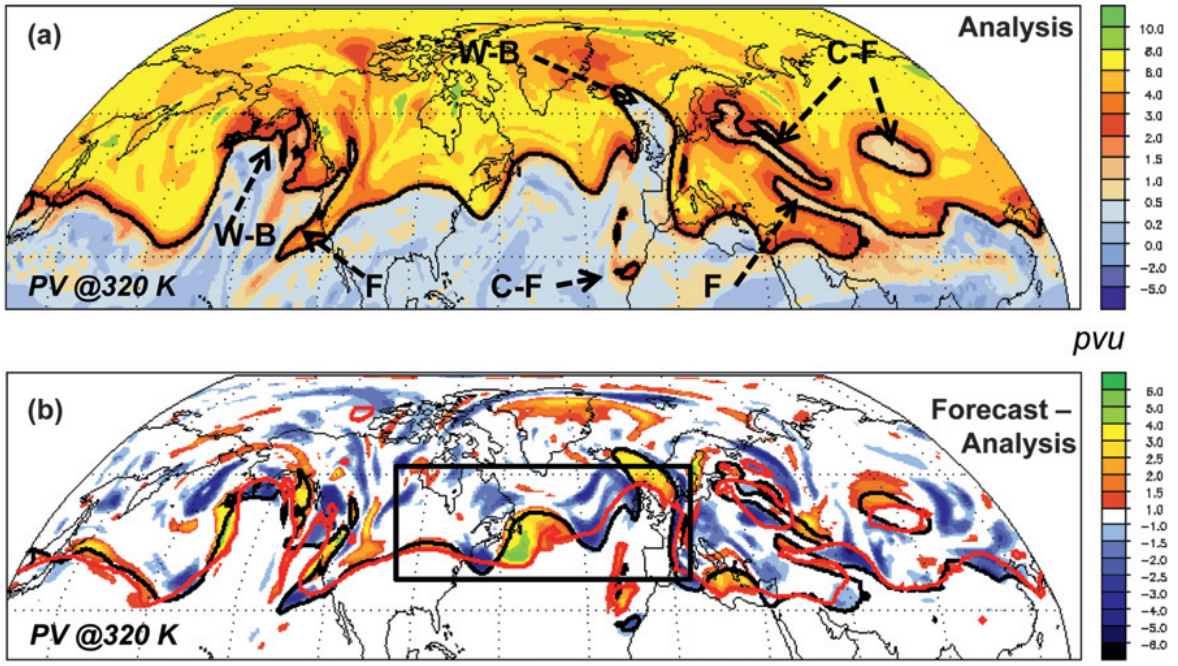

Figure 1.14: Example of a PV pattern on the 320 K isentropic surface from a) the analysis field, and b) the error (forecast minus analysis) field of the corresponding 4 day forecast. The black (red) contour denotes the 2 PVU contour of the analysis (forecast) and the symbols W-B, F, and C-F refer to wave-breaking, filament, and cutoff structures, respectively. The figure is taken from Fig. 1 of Davies and Didone (2013).

weather prediction models. By focusing on PV, Gray et al. (2014) showed that the amplitude of Rossby-wave ridges and the tropopause sharpness is systematically underestimated in the forecast. Furthermore, PV has been used to identify model deficiencies in forecasting Rossby-wave breaking (Wiegand and Knippertz, 2014) and the Rossby waveguide (Giannakaki and Martius, 2016).

Based on the PV perspective, Davies and Didone (2013) provided an outline to diagnose the dynamics of forecast errors in the medium range. For an example-weather situation, the authors showed that the PV-error field of this case reveals key-error features that are related to a misrepresentation of the Rossby-wave pattern near the tropopause and of surface-weather systems (Fig 1.14). Furthermore, the authors derived a tendency equation for the PV error that provides qualitative insight into the dynamical evolution of forecast errors and into possible error-growth mechanisms.

This study builds on the study of Davies and Didone (2013). Here, we extend their work significantly by providing not only a qualitative, but also a quantitative view on the dynamics of forecast errors and forecast uncertainty. Using the partitioning of processes as introduced in the previous subsection, we quantify the relative importance of individual processes to the amplification of forecast errors and forecast uncertainty.

1.4 Research questions and thesis outline

The central aim of this thesis is to provide dynamical insight into the processes governing the amplification of forecast errors and forecast uncertainty. For that purpose, we use a PV framework and focus on the near-tropopause region of the midlatitudes, which is of particular importance to the evolution of midlatitude weather systems and the evolution of PV errors. In essence, we address the following research questions:

1. **Error growth in an operational state-of-the-art deterministic forecast**
   a) How can the PV perspective be used to quantify the processes governing the amplification of forecast errors?
   b) What are the most important processes governing the error growth in an operational state-of-the-art forecast?
   c) How can the error-growth mechanisms be explained from a theoretical perspective?
2. **Error growth in upscale error-growth simulations**
   a) What are the most important processes governing upscale error growth from the grid scale up to the planetary scale?
   b) Are the quantitative results from the PV framework consistent with the conceptual model for upscale error growth?
   c) Do the processes governing upscale error growth differ from the processes governing error growth in operational forecasts?
3. **Amplification of ensemble spread in an operational state-of-the-art forecast**
   a) How can the PV perspective be used to quantify the processes governing the amplification of forecast uncertainty?
   b) What are the most important processes governing the amplification of ensemble spread in an operational state-of-the-art forecast?
   c) Are there flow-dependent differences in the amplification of ensemble spread?

These three research-question items serve as an outline for this thesis: In Chapter 2, we derive a tendency equation for forecast-error growth and provide a quantitative view on the processes governing the error growth in a current state-of-the-art operational forecast. Furthermore, we discuss individual error-growth mechanisms. Chapter 3 investigates upscale error growth from the grid scale up to the planetary scale using global simulations with a stochastic convection scheme. We quantify the processes governing the upscale error growth and compare our results with the conceptual model for upscale error growth and with the error growth in operational forecasts. In Chapter 4, we derive a tendency equation for the ensemble variance of PV. The processes governing the amplification of ensemble spread are quantified for a current state-of-the-art operational forecast and flow-dependent aspects of this amplification are highlighted. Finally, we conclude with a summary and outlook in Chapter 5.

CHAPTER 2

Potential vorticity dynamics of forecast errors: A quantitative case study

Marlene Baumgart[a], Michael Riemer[a], Volkmar Wirth[a], Franziska Teubler[a], and Simon T. K. Lang[b]

[a]Institut für Physik der Atmosphäre, Johannes Gutenberg-Universität Mainz, Mainz, Germany

[b]European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

This work has been published in *Monthly Weather Review* in May 2018, Volume **146**, Pages 1405—1425, `https://doi.org/10.1175/MWR-D-17-0196.1` (Baumgart et al., 2018).[1]

Abstract

Synoptic-scale error growth near the tropopause is investigated from a process-based perspective. Following previous work, a potential vorticity (PV) error tendency equation is derived and partitioned into individual contributions to yield insight into the processes governing error growth near the tropopause. Importantly, we focus here on the further amplification of preexisting errors and not on the origin of errors. The individual contributions to error growth are quantified in a case study of a 6-day forecast. In this case, localized mesoscale error maxima have formed by forecast day 2. These maxima organize into a wavelike pattern and reach the Rossby wave scale around forecast day 6. Error growth occurs most prominently within the Atlantic and Pacific Rossby wave patterns. In our PV framework, the error growth is dominated by the contribution of upper-level, near-tropopause PV anomalies (near-tropopause dynamics). Significant contributions from upper-tropospheric divergent flow (prominently associated with latent heat release below) and lower-tropospheric anomalies [tropospheric-deep

[1] Slight modifications have been added to the text for consistency.

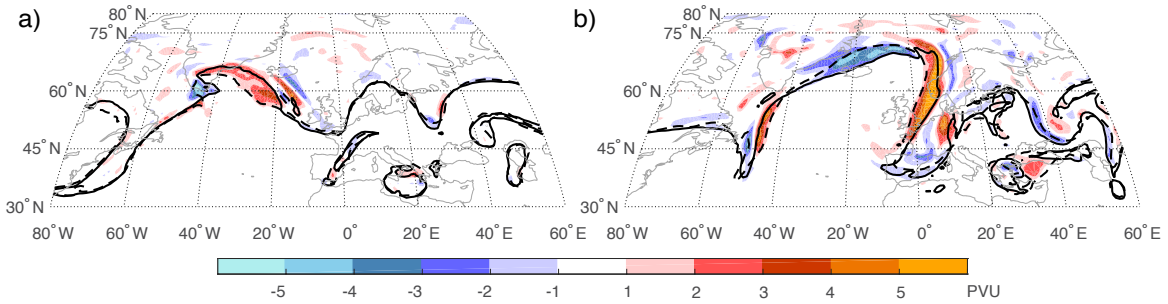

Figure 2.1: PV error (colored) on an isentropic surface (320 K) for the forecast started at 00 UTC, 12 November 2013 in the Atlantic region for a) forecast day 2, and b) forecast day 3. The black contour denotes the dynamical tropopause (2 PVU surface; 1 PVU = 10^{-6} K m^2 kg^{-1} s^{-1}) of the analysis (solid) and of the forecast (dashed) on 320 K.

(i.e., baroclinic) interaction] are associated with a misrepresentation of the surface cyclone development in the forecast. These contributions are, in general, of smaller importance to error growth than near-tropopause dynamics. This result indicates that the mesoscale errors generated near the tropopause do not primarily project on differences in the subsequent baroclinic growth, but instead directly project on the tropopause evolution and amplify because of differences in the nonlinear Rossby wave dynamics.

2.1 Introduction

Numerical weather prediction has improved remarkably over the last decades (e.g., Bauer et al., 2015). Occasionally, however, very poor medium-range forecasts do still occur (Rodwell et al., 2013). Forecast errors arise due to errors in the initial conditions and due to model deficiencies (e.g., Palmer and Hagedorn, 2006). After 1–2 forecast days, localized errors may form that start to affect the synoptic-scale flow (e.g., Davies and Didone, 2013; Martínez-Alvarado et al., 2016). Subsequently, these errors further amplify and propagate downstream (e.g., Langland et al., 2002; Anwender et al., 2008; Pantillon et al., 2013).

The focus of this study is on the amplification of existing errors as illustrated in Fig. 2.1. At forecast day 2, a localized PV error is in our case generated in the Atlantic ridge (Fig. 2.1a). Within the next day, significant error amplification occurs and the localized error evolves into a wave-like pattern at forecast day 3 (Fig. 2.1b). This study aims at understanding such error growth by quantifying the processes governing the error amplification, rather than identifying the initial source of the error growth. Gaining deeper insight into the dynamics of error growth, and thus into the atmospheric conditions that exhibit high or low intrinsic predictability (e.g., Melhauser and Zhang, 2012) can be expected to improve the future interpretation of the forecast uncertainty observed in ensemble forecast systems.

Synoptic-scale error growth is often related to baroclinic instability (e.g., Tribbia and Baumhefner, 2004; Hakim, 2005; Hohenegger and Schär, 2007; Zhang et al., 2007; Boisserie et al., 2014; Selz and Craig, 2015b; Sun and Zhang, 2016). Baroclinic instability leads to exponential growth of wave amplitudes (e.g., Eady, 1949). Therefore, a small difference between the

baroclinic growth in the forecast and the analysis can be expected to yield exponential error growth. A number of other studies have highlighted the importance of processes that are unrelated to baroclinic instability: Snyder (1999) and Plu and Arbogast (2005) emphasize that errors in the propagation characteristics and nonlinear error growth mechanisms can be as important for the error growth as barotropic and baroclinic instabilities. In addition, more recent studies indicate that synoptic-scale error growth can be induced by misrepresented deformation fields or potential vorticity (PV) anomalies near the tropopause (Davies and Didone, 2013), a systematic underestimation of the background PV gradient in the forecast (Harvey et al., 2016), and bifurcation-like behavior (Riemer and Jones, 2014). A quantitative analysis of the relative importance of individual processes, however, is still missing.

This study adopts the potential-vorticity (PV) perspective of forecast errors to gain insight into the processes governing the error growth. This perspective provides a useful framework to analyze error growth from a dynamical, process-oriented perspective (Davies and Didone, 2013). The largest values of the PV error occur near the midlatitude tropopause and are related to erroneous displacements of the tropopause (e.g., Dirren et al., 2003; Gray et al., 2014). It is these near-tropopause errors that we will focus on in this study. Davies and Didone (2013) have derived a tendency equation for the PV error and analyzed it qualitatively. Our study considers a similar tendency equation in combination with a partitioning of the PV tendencies as applied in an investigation of Rossby wave packet dynamics (Teubler and Riemer, 2016). This partitioning allows us to quantify the relative importance of four individual processes to PV error growth near the tropopause: 1) near-tropopause dynamics, 2) tropospheric-deep interaction, 3) upper-tropospheric divergent flow, and 4) nonconservative processes.

The contribution of near-tropopause dynamics quantifies the influence of upper-level, near-tropopause PV anomalies on the tropopause evolution, which is mostly related to Rossby wave dynamics, while the contribution of tropospheric-deep interaction quantifies the influence of lower-tropospheric anomalies on the tropopause evolution, which signifies baroclinic interaction in the PV framework (Eady, 1949; Hoskins et al., 1985; Davis and Emanuel, 1991; Heifetz et al., 2004). Baroclinic instability can be described by the mutual amplification of upper- and lower-level PV anomalies. Its potential influence near the tropopause can therefore be assessed by the influence of tropospheric-deep interaction. In addition to these contributions from balanced dynamics, we quantify the influence of upper-tropospheric divergent flow. Upper-tropospheric divergence can be associated with balanced dynamics and diabatic processes (see considerations with the omega equation, e.g., Holton and Hakim, 2013). For Rossby wave dynamics, a few cases indicate that strong upper-tropospheric divergence is mostly associated with latent heat release below (e.g., Davis et al., 1993; Riemer et al., 2014; Quinting and Jones, 2016). Finally, the direct impact of nonconservative processes is quantified.

This Chapter is organized as follows. In Section 2.2, we describe the data and tendency equation used to quantify PV error growth and the relative importance of individual processes. Section 2.3 presents our results for a case study of medium-range forecast error growth. In Section 2.4, we discuss some conceptual aspects of error growth based on the error patterns

observed in our case study. Finally, Section 2.5 contains our summary and conclusions of the results.

2.2 Data and diagnostics to quantify the individual contributions to PV error growth

2.2.1 Data

We use analysis and deterministic-forecast data from the Atmospheric Model high resolution 10-day forecast (HRES) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Error growth is analyzed for the forecast starting at 00 UTC, 12 November 2013[2]. This forecast covers an extreme precipitation event over Sardinia on 18-19 November 2013, which was closely connected with the preceding Rossby wave packet, wave breaking, and the formation of a cutoff low. The error growth within this Rossby wave packet is therefore of particular interest.

Six-hourly data are used on pressure levels and a $1^\circ \times 1^\circ$ grid. The piecewise PV inversion used to partition the PV tendencies is performed on pressure levels as in Davis and Emanuel (1991) and Davis (1992). For the further analysis of PV error growth, all variables are interpolated to isentropic levels. Similar to YOTC ("Year of Tropical Convection") data (Moncrieff et al., 2012), accumulated temperature and wind tendencies from the physical parametrization schemes were saved for our case study[3]. Note that the computations of the physical parametrization schemes are performed in the vertical only (ECMWF, 2013), in particular they account only for vertical diffusion and not for horizontal mixing. Instantaneous tendencies are approximated by centered differences of the 3 h accumulated tendency data[4]. Tendency data for the analysis are estimated with accumulated data from short-range forecasts.

2.2.2 PV error and tendency equation

The hydrostatic form of Ertel PV (Ertel, 1942) is used in isentropic coordinates,

$$P = -g\frac{\partial\theta}{\partial p}\left(\zeta_\theta + f\right) , \tag{2.1}$$

where g is the gravitational acceleration, θ the potential temperature, p the pressure, ζ_θ the vertical component of the isentropic relative vorticity, and f the Coriolis parameter.

Forecast errors (index *) are defined as the difference between the forecast (index fc) and

[2]The error growth was also analyzed for the forecasts starting at 00 UTC, 9 November 2013 and 12 UTC, 10 November 2013. These forecasts showed similar results with respect to the relative importance of the processes contributing to the error growth.

[3]Note that we do not use the tendencies for specific humidity as they do not directly affect the PV evolution.

[4]A comparison between temperature tendency data with 1 h- and 3 h-resolution did not show a significant dependence on the time resolution. Therefore, we assume that this approximation yields reasonable results for our diagnostics.

the analysis (index an), e.g. $P^* := P_{\text{fc}} - P_{\text{an}}$. Following the approach of Davies and Didone (2013), we derive a PV-error tendency equation. Our tendency equation is based on the local tendency equation for PV in isentropic coordinates, which is given by (e.g., Davies and Didone, 2013)

$$\frac{\partial P}{\partial t} = -\mathbf{v} \cdot \nabla_\theta P + N + \text{res.} \,, \tag{2.2}$$

where $N = -\dot{\theta}\frac{\partial P}{\partial \theta} + P\frac{\partial \dot{\theta}}{\partial \theta} - g\frac{\partial \theta}{\partial p}\left(\mathbf{k} \times \frac{\partial \mathbf{v}}{\partial \theta}\right) \cdot \nabla\dot{\theta} - g\frac{\partial \theta}{\partial p}\mathbf{k} \cdot (\nabla \times \dot{\mathbf{v}})$ describes the nonconservative PV tendency due to diabatic heating and nonconservative momentum change as measured by the parametrization schemes. The heating rate $\dot{\theta}$ comprises heating due to the cloud, convection, radiation, and turbulence parametrization, while the horizontal wind tendencies $\dot{u}$ and $\dot{v}$ comprise tendencies from the convection and turbulence parametrization. The residual term res. comprises the influence of those processes that can not be quantified with the available data, i.e. numerical diffusion, analysis increments due to data assimilation, and numerical inaccuracies due to the discretization and interpolation of data.

As the error is defined by the difference between the forecast and analysis, the local tendency for the PV error is given by (see Davies and Didone, 2013)

$$\frac{\partial P^*}{\partial t} = \frac{\partial P_{\text{fc}}}{\partial t} - \frac{\partial P}{\partial t} = -\mathbf{v}^* \cdot \nabla_\theta P - \mathbf{v}^* \cdot \nabla_\theta P^* - \mathbf{v} \cdot \nabla_\theta P^* + N^* + \text{res.}, \tag{2.3}$$

where variables without an index refer to analysis variables. Note that the PV error is not materially conserved under conservative conditions and that the tendency equation for the PV error is thereby more complicated than Eq. 2.2.

Similar to the enstrophy error by Boer (1984), we consider the local tendency of the potential enstrophy error (the squared PV error, P^{*2}), which is a positive-definite error metric and thus ensures that a positive error tendency is associated with error amplification. The according tendency equation is given by

$$\begin{aligned}
\frac{\partial}{\partial t}\frac{P^{*2}}{2} &= P^*\frac{\partial P^*}{\partial t} = -P^*\mathbf{v}^* \cdot \nabla_\theta P - (\mathbf{v}^* + \mathbf{v}) \cdot \nabla_\theta \left(\frac{P^{*2}}{2}\right) + P^* N^* + \text{res.} \\
&= -P^*\mathbf{v}^* \cdot \nabla_\theta P - \nabla_\theta \cdot \left(\frac{P^{*2}}{2}(\mathbf{v}^* + \mathbf{v})\right) + \frac{P^{*2}}{2}\nabla_\theta \cdot (\mathbf{v}^* + \mathbf{v}) \\
&\quad + P^* N^* + \text{res.}\,.
\end{aligned} \tag{2.4}$$

To provide a succinct quantitative view on the error growth, we consider the error evolution integrated over a (potentially) time-dependent area A

$$\begin{aligned}
\frac{d}{dt}\int_A \frac{P^{*2}}{2} dA &= \int_A \frac{\partial}{\partial t}\frac{P^{*2}}{2} dA + \oint_S \frac{P^{*2}}{2}\mathbf{v_S} \cdot \mathbf{n}\, dS \\
&= -\int_A P^*\mathbf{v}^* \cdot \nabla_\theta P dA - \oint_S \frac{P^{*2}}{2}(\mathbf{v}^* + \mathbf{v}) \cdot \mathbf{n}\, dS + \int_A \frac{P^{*2}}{2}\nabla_\theta \cdot (\mathbf{v}^* + \mathbf{v}) dA \\
&\quad + \int_A P^* N^* dA + \oint_S \frac{P^{*2}}{2}\mathbf{v}_S \cdot \mathbf{n}\, dS + \text{res.}\,,
\end{aligned} \tag{2.5}$$

where $\mathbf{v}_S$ describes the motion of the integration area. As in Boer (1984), the first term on the right hand side of Eq. 2.5 can be interpreted as a (nonlinear) production term. Since the second term in Eq. 2.4 merely redistributes existing errors, it can be evaluated in terms of an error flux across the boundary of the integration domain (second term in Eq. 2.5). Not explicitly included in the barotropic framework of Boer (1984), the third term in Eq. 2.5 constitutes an error source due to the divergence of the quasi-horizontal (adiabatic) flow. The remaining terms describe the influence of nonconservative processes (term 4), the boundary contribution due to changes in the integration area (term 5), and the residual (term 6).
We evaluate the PV-error tendency equation on an isentropic level intersecting the midlatitude tropopause, namely on 320 K. PV anomalies on this isentropic level are representative for midlatitude Rossby waves during the time of year under consideration (Liniger and Davies, 2004; Martius et al., 2010). A comparison with the error growth on nearby isentropes (315 K and 325 K) showed qualitatively very similar results.

2.2.3 Further separation of the advective term

To gain deeper insight into the dynamics of the PV error growth, we further partition the nonlinear production term (term 1 on the right hand side in Eq. 2.5), following Teubler and Riemer (2016). This partitioning builds on the PV perspective of midlatitude dynamics. In essence, the evolution of PV anomalies near the tropopause can be described by advective tendencies associated with 1) the upper-level (near-tropopause) PV anomalies themselves, 2) θ anomalies just above the boundary layer and "interior" PV anomalies in the lower troposphere, and 3) upper-tropospheric divergent flow. In addition, upper-level PV anomalies are modified by 4) nonconservative processes.
The influence of upper-level PV anomalies on upper-level dynamics can be related to Rossby wave dynamics. As mentioned in the introduction, baroclinic instability is in the PV framework described by the mutual amplification of upper- and lower-level anomalies (Eady, 1949; Hoskins et al., 1985; Davis and Emanuel, 1991; Heifetz et al., 2004). Since our focus is on error growth near the tropopause, the contribution of tropospheric-deep interaction is quantified by the influence of lower-level errors on the upper-level errors. This part of the interaction, however, provides a necessary condition that amplification due to baroclinic instability does occur. Upper-tropospheric divergent flow can significantly influence Rossby wave dynamics, in particular ridge building (e.g., Grams et al., 2011; Teubler and Riemer, 2016). It can be associated with dry balanced dynamics and diabatic processes (see considerations with the omega equation as discussed in, e.g., Chapter 6.4 in Holton and Hakim, 2013). A few cases (e.g., Davis et al., 1993; Riemer and Jones, 2010; Riemer et al., 2014) indicate that pronounced upper-level divergence is mostly related to latent heat release below. This indirect diabatic impact can be of larger importance for Rossby wave dynamics than direct diabatic PV modification (e.g., Riemer and Jones, 2010; Teubler and Riemer, 2016).
Figure 2.2 illustrates that the errors in our case concentrate in two regions: Errors exhibit their strongest signal in PV close to the tropopause and, equivalently, as θ errors at the top

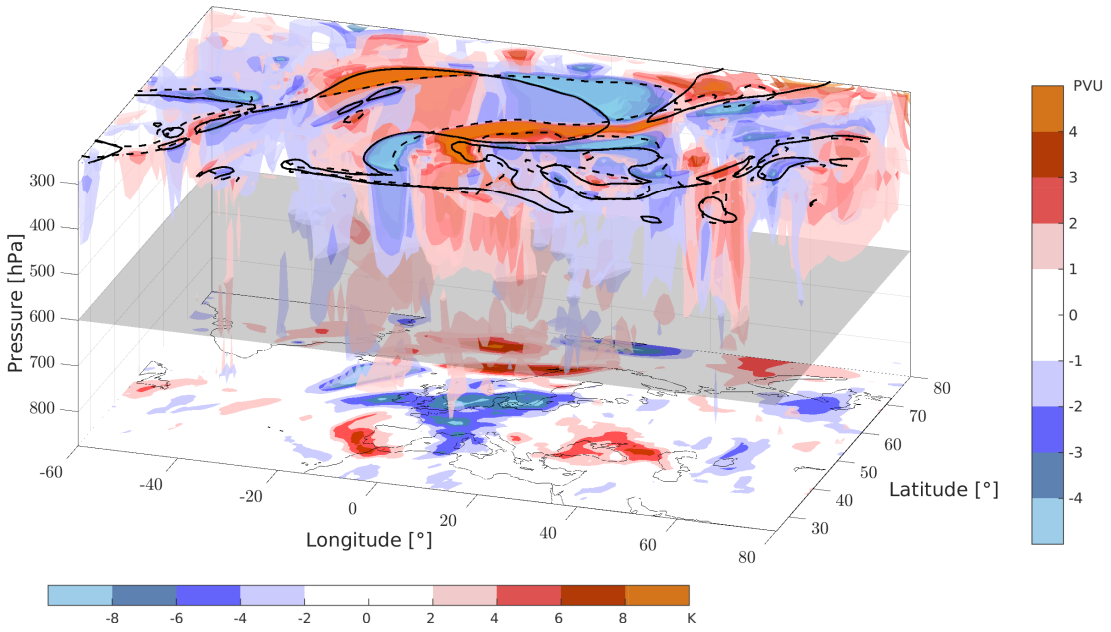

Figure 2.2: 3D picture of the PV error between 850 and 250 hPa (colored) and the θ error at the top of the boundary layer (875 hPa, colored) for the forecast started at 00 UTC, 12 November 2013 for forecast day 4. The black contour denotes the dynamical tropopause (smoothed 2 PVU surface) of the analysis (solid) and of the forecast (dashed) on 250 hPa. The gray section at 600 hPa indicates the pressure level where the upper and lower PV anomalies are separated.

of the boundary layer (875 hPa). There is a clear vertical separation between these upper- and lower-level PV and θ errors. Therefore, it is reasonable to separate the error features at a mid-level where the error features are in general of small amplitude. Our results are not sensitive to the exact choice of the separation level (tested for 700, 600, and 500 hPa). The following results are shown for a separation level of 600 hPa.

The upper-level wind associated with the lower-tropospheric (PV and θ) anomalies will be referred to as "tropospheric-deep" ($\mathbf{v}_{\mathrm{TPd}}$), the upper-level wind associated with upper-level anomalies as "near-tropopause" ($\mathbf{v}_{\mathrm{nTP}}$), and the upper-level wind associated with upper-level divergence as "divergent" ($\mathbf{v}_{\mathrm{div}}$). Thus, the full wind field can be partitioned as

$$\mathbf{v} = \mathbf{v}_{\mathrm{nTP}} + \mathbf{v}_{\mathrm{TPd}} + \mathbf{v}_{\mathrm{div}} + \mathbf{v}_{\mathrm{uct}}. \tag{2.6}$$

The uncertainty term, $\mathbf{v}_{\mathrm{uct}}$, describes the uncertainties associated with the partitioning of the wind field. It is calculated as the difference between the wind field given by ECMWF and the sum of the near-tropopause, tropospheric-deep and divergent wind field. This uncertainty turns out to be small and does not affect the physical interpretation of our results. Technical details on our flow partitioning are provided in Appendix 2.6.1.

Quantitatively, we separate the relative importance of the individual processes by spatially-

integrating their contribution to the error tendency (Eq. 2.5)

$$\begin{aligned}
\text{OBS} &= \text{nTP+TPd+DIV+NONCONS+BND+UCT+RES} \,, \qquad (2.7)\\
\text{where}\quad \text{OBS} &:= \frac{d}{dt}\int_A \frac{P^{*2}}{2} dA \approx \frac{P^{*2}_{t+6\,h} - P^{*2}_{t}}{6\,h} \,,\\
\text{nTP} &:= -\int_A P^* \mathbf{v}^*_{\text{nTP}} \cdot \nabla_\theta P dA \,,\\
\text{TPd} &:= -\int_A P^* \mathbf{v}^*_{\text{TPd}} \cdot \nabla_\theta P dA \,,\\
\text{DIV} &:= -\int_A P^* \mathbf{v}^*_{\text{div}} \cdot \nabla_\theta P dA + \int_A \frac{P^{*2}}{2} \nabla_\theta \cdot (\mathbf{v}^*_{\text{div}} + \mathbf{v}_{\text{div}}) dA \,,\\
\text{NONCONS} &:= \int_A P^* N^* dA \,,\\
\text{BND} &:= -\oint_S \frac{P^{*2}}{2} (\mathbf{v}^* + \mathbf{v} - \mathbf{v}_S) \cdot \mathbf{n}\, dS \,,\\
\text{UCT} &:= -\int_A P^* \mathbf{v}^*_{\text{uct}} \cdot \nabla_\theta P dA \,,\\
\text{RES} &:= \text{OBS} - (\text{nTP+TPd+DIV+NONCONS+BND+UCT}) \,.
\end{aligned}$$

The residual term is determined by the difference between the observed error evolution (OBS; approximated by 6 h finite differences) and the sum of the quantified contributions to error growth (averaged over the two time steps used for the finite differences).
We emphasize that Eq. 2.7 describes the influence of individual processes to error growth at an instantaneous time step. This approach is different from Lagrangian approaches that investigate the accumulated effect of, e.g., diabatic processes on the PV distribution (e.g., Gray, 2006; Chagnon et al., 2013) or that track errors back to their initial source (e.g., Martínez-Alvarado et al., 2016). In particular, we do not explicitly analyze the indirect effect of diabatic processes by having created PV anomalies at earlier times (see summary and discussion by Davis et al., 1993). As the processes governing the error growth may differ at different stages of the error growth (e.g., Zhang et al., 2007; Hohenegger and Schär, 2007; Selz and Craig, 2015b), analysis of the instantaneous tendencies provides helpful insight into the error-growth dynamics.

2.3 Error growth in an illustrative case

In this section, the error growth in our case is first illustrated by the time evolution of the PV error. Then we show the spatial patterns of the individual contributions to net error amplification. Finally, the relative importance of these contributions to error growth in specific regions is quantified by time series of spatially-integrated tendencies.

2.3.1 Error evolution

Fig. 2.3 illustrates the PV-error evolution of our case study on 320 K together with the location of the dynamical tropopause in the forecast and the analysis (here defined as the 2 PVU surface; 1 PVU = 10^{-6} K m^2 kg^{-1} s^{-1}), which serves as an indication of the synoptic evolution in the forecast and in the analysis. The tropopause is, in general, undulated by Rossby wave patterns. Two regions, the Atlantic and Pacific region, turn out to be associated with a significant error evolution and will be discussed in more detail.

The synoptic evolution in the Atlantic starts with a ridge building around 30°W (Fig. 2.3a), which is in association with cyclogenesis off the coast of Newfoundland. The cyclone's intensity and ridge building were underestimated in the forecast, leading to a localized positive error in surface pressure and PV around 40–20°W at forecast day 2. In the following two days, the downstream trough breaks anti-cyclonically and a cutoff is formed at day 4 (10°W–20°E, Fig. 2.3b). This wave breaking occurred too late in the forecast and was associated with significant error amplification between day 2 and 4. At day 6 (Fig. 2.3c), the wave pattern in the forecast and in the analysis exhibit large differences, including local phase shifts of approx. half a wavelength and significant amplitude errors.

The synoptic evolution in the Pacific starts with anti-cyclonic wave-breaking and cutoff formation around 140°W before forecast day 2 (not shown). The error pattern at day 2 is dominated by an error in the meridional location of the tropopause (Fig. 2.3a). In the following two days, ridge building associated with cyclogenesis governs the evolution (150°W–170°E, Fig. 2.3b). The cyclogenesis and ridge building were much stronger in the forecast than in the analysis leading to significant error amplification in surface pressure and PV. Further error amplification occurs between day 4 and 6 leading to large differences between the wave pattern in the forecast and the analysis at day 6 (Fig. 2.3c).

Consistent with, e.g., Dirren et al. (2003) and Davies and Didone (2013), large absolute values of the PV error are related to a displacement of the dynamical tropopause. Both in the Atlantic and the Pacific, a first localized error pattern changes into a wave-like error pattern between day 2 and 4 (Figs. 2.3a and b). Error extrema in both regions are associated with an erroneous representation of the Rossby wave pattern. Compared to the PV anomalies (defined by the deviation of the PV from a background PV, see appendix A for details), the PV error exhibits, in general, a smaller horizontal scale, in particular in the meridional direction (Figs. 2.3b and d).

2.3.2 Spatial patterns of the individual contributions to error growth

In Fig. 2.4 the individual contributions to the net error amplification in Eq. 2.7 are illustrated for forecast day 4 (note the different colorbars in this figure). The advective tendencies are further discussed with focus on the Pacific region in Fig. 2.5. All individual tendencies exhibit their largest amplitude and a relatively high degree of spatial coherence within the Rossby wave pattern in the Atlantic and in the Pacific. The Rossby wave patterns thus stand out as important regions of the error evolution, corroborating our notion from the previous

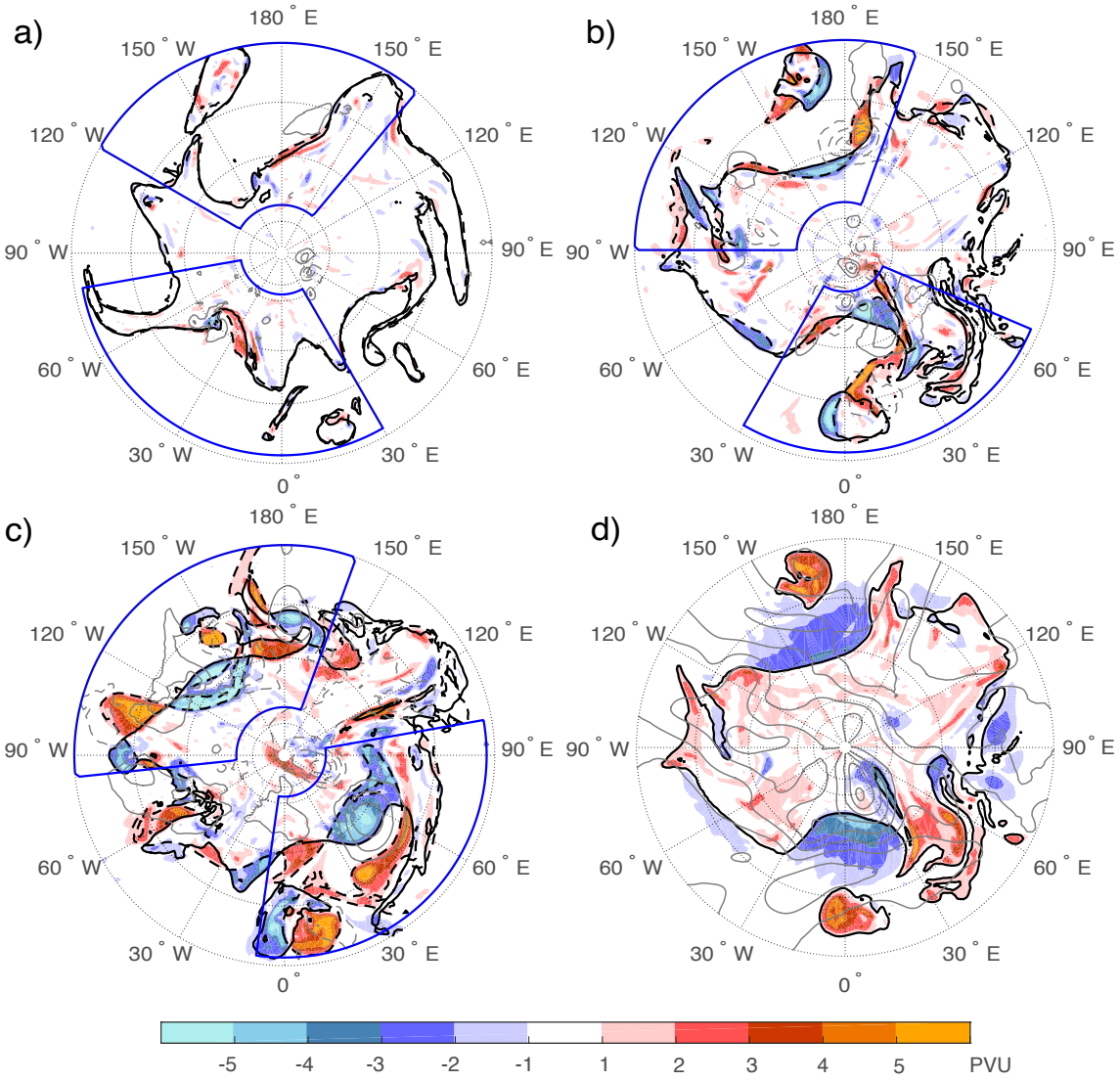

Figure 2.3: PV error (colored) on an isentropic surface (320 K) for the forecast started at 00 UTC, 12 November 2013 in the Northern Hemisphere for a) forecast day 2, b) forecast day 4, and c) forecast day 6. For comparison, d) shows the PV anomalies of the analysis, i.e. deviations of the analysis PV from the background PV (see appendix A for details), at day 4. The black contour denotes the dynamical tropopause (2 PVU surface) of the analysis (solid) and of the forecast (dashed) on 320 K. The gray contours depict the error in mean sea level pressure every 5 hPa in a), every 7.5 hPa in b), and every 10 hPa in c) without 0 hPa (positive: solid, negative: dashed), and the mean sea level pressure in the analysis every 10 hPa (smoothed over 3 gridpoints) in d). Blue boxes in a)–c) indicate the regions used as integration areas.

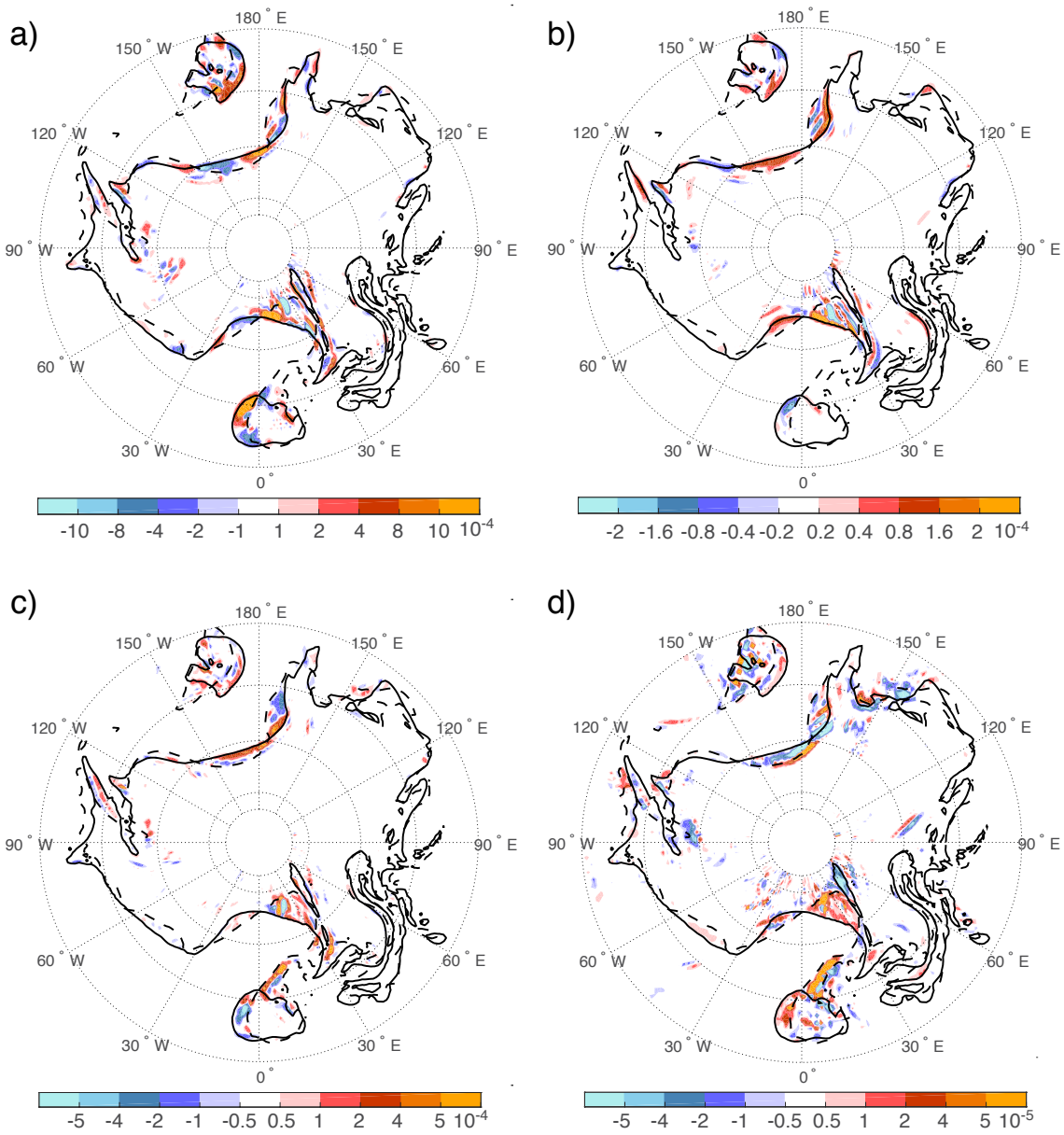

Figure 2.4: Spatial distribution of individual contributions to net error amplification as indicated by Eq. 2.7 (colored, in $\mathrm{PVU}^2/\mathrm{s}$) for forecast day 4 on 320 K by a) near-tropopause dynamics (nTP in Eq. 2.7), b) tropospheric-deep interaction (TPd in Eq. 2.7), c) upper-tropospheric divergence (DIV in Eq. 2.7), and d) nonconservative processes (NONCONS in Eq. 2.7). The black contour denotes the dynamical tropopause (2 PVU surface) of the analysis (solid) and of the forecast (dashed) on 320 K. Note the different scales of the colorbars.

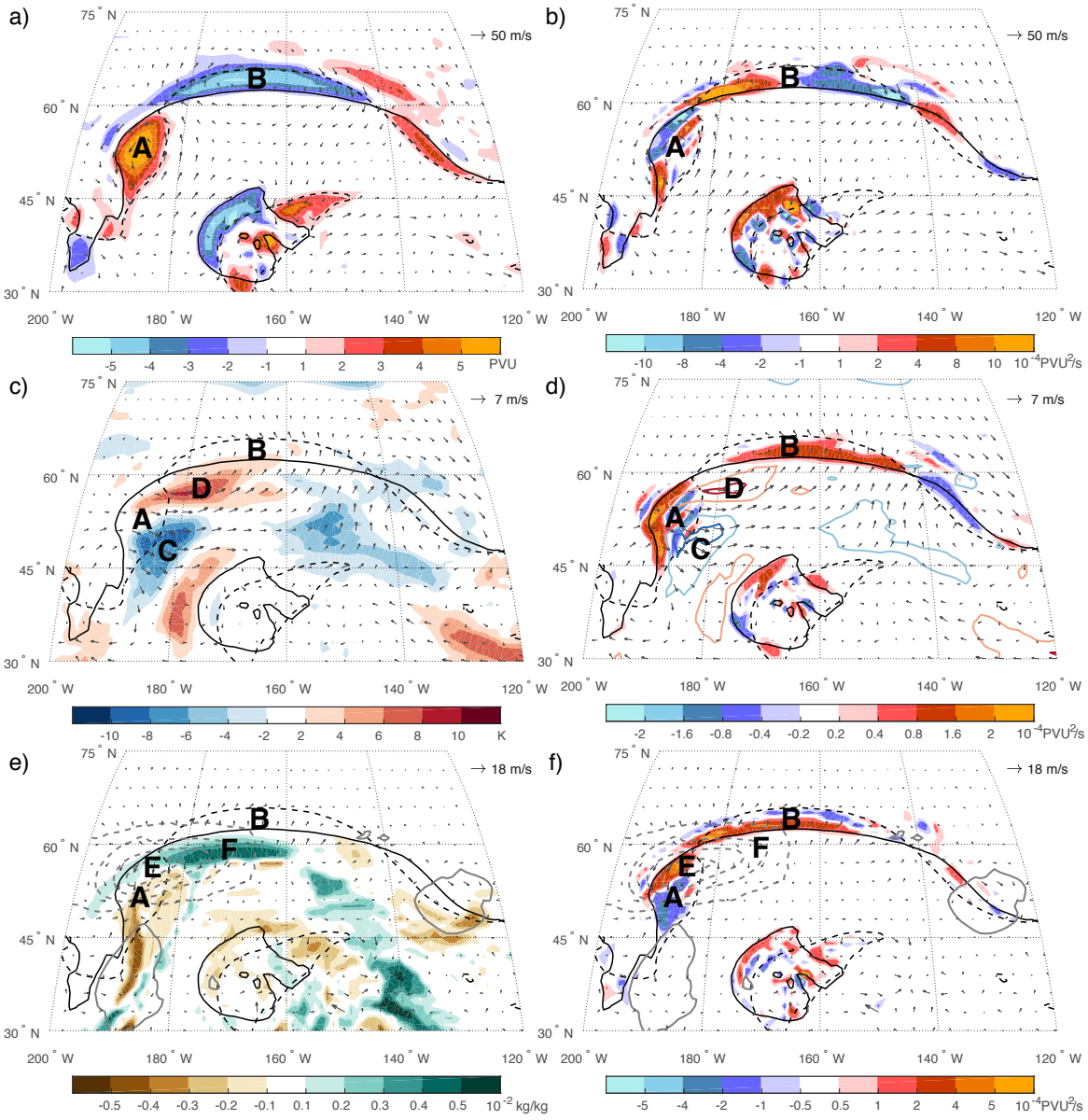

Figure 2.5: Focus on the Pacific region. a) PV error on 320 K, c) potential temperature error on 875 hPa, e) specific humidity error on 700 hPa, and net error amplification associated with b) near-tropopause dynamics, d) tropospheric-deep interaction, and f) upper-tropospheric divergence on 320 K for forecast day 4. The black contour denotes the 2 PVU surface of the analysis (solid) and of the forecast (dashed) on 320 K. Contours in d) depict the θ error on 875 hPa at -4 and -8 K (blue) and at 4 and 8 K (red). Gray contours in e) and f) show the error in mean sea level pressure every 7.5 hPa without 0 hPa (positive: solid, negative: dashed). Labels point to individual PV errors near the tropopause (A and B), θ errors on 875 hPa (C and D), a cyclone error (E), and a warm conveyor belt error (F). Arrows denote the error wind associated with near-tropopause dynamics (a and b), tropospheric-deep interaction (c and d), and upper-tropospheric divergence (e and f) on 320 K.

subsection.

The near-tropopause tendency (Figs. 2.4a and 2.5b) exhibits the largest absolute values compared to the other individual tendencies. The spatial pattern of the total tendency (not shown) is almost entirely determined by the near-tropopause tendency. The dominant feature of the spatial pattern are dipoles of positive and negative values. A comparison with the PV error (Figs. 2.5a and b) indicates that the near-tropopause error wind and near-tropopause error tendency are clearly related to the PV error: Positive PV errors (error feature A in Fig. 2.5a) are associated with a cyclonic error circulation, whereas negative PV errors (error feature B in Fig. 2.5a) are associated with an anti-cyclonic error circulation. The associated error tendency, which is defined by the advection of the analysis PV by the near-tropopause error wind (Eq. 2.7), leads to error amplification on the western side and error decay on the eastern side of the individual PV error features (see error tendency within error features A and B in Fig. 2.5b). The tendency pattern is thus dominated by a displacement of the PV error along the dynamical tropopause. From Figs. 2.4a and 2.5b it is not directly clear whether the net effect of the near-tropopause tendency is positive or negative.

The tropospheric-deep tendency (Figs. 2.4b and 2.5d) leads to error amplification in the Pacific ridge (140–180°W) and possibly also in the Atlantic ridge (near 10°W). In the Pacific region, temperature error features at the top of the boundary layer and PV error features near the tropopause have a phase shift of about $\pi/4$ to $\pi/2$ (see Figs. 2.5a and c). The tropospheric-deep error wind associated with the positive potential temperature error (error feature D in Fig. 2.5c) exhibits a cyclonic pattern leading to the tropospheric-deep error amplification observed in the Pacific ridge (Fig. 2.5d). The absolute values of the tropospheric-deep tendency are, however, almost one order of magnitude smaller than those of the near-tropopause tendency.

The divergent tendency (Figs. 2.4c and 2.5f) leads to error amplification over coherent regions, in particular in ridges and cutoffs. A comparison with the error in mid-tropospheric specific humidity and mean sea level pressure in the Pacific region (Figs. 2.5e and f) indicates that the divergent tendency is directly linked to the misrepresentation of the surface cyclone around 190°W and the associated latent heat release: This cyclone (error feature E in Fig. 2.5e), its accompanying warm conveyor belt (error feature F in Fig. 2.5e), and the associated upper-tropospheric divergent flow and divergent ridge amplification were too strong in the forecast. The divergent tendency thus increases the difference between the ridge amplification in the forecast and the analysis leading to error amplification in the Pacific region, which can be seen by coherent region of positive error tendency in the ridge (Fig. 2.5f). A similar investigation shows that strong divergent tendencies in other regions can be related to latent heat release also (not shown).

The largest values of the nonconservative tendency (approximated by the wind and temperature tendencies from the parametrization schemes, Fig. 2.4d) are found in ridges and cutoffs. A further partitioning of this tendency (not shown) reveals that error amplification in ridges and cutoffs is in this case mostly governed by radiation, whereas the turbulence parametrization leads to the error decay in regions of a strong PV gradient, e.g., in the Pacific ridge.

These two processes have mostly opposite effects. Compared to the advective tendencies, the nonconservative tendency exhibits the smallest absolute values.

2.3.3 Spatially-integrated metric of the error growth

Insight into the relative importance of the individual processes can be gained by considering spatially-integrated tendencies. Integration over large, sensibly-defined regions diminishes the compensation between positive and negative tendencies prominent in Fig. 2.4. The areas used for integration are chosen such that the distinct error patterns in the Atlantic and Pacific are included during the whole forecast time. The areas are indicated by the blue boxes in Fig. 2.3. Time series of the spatially-integrated tendencies are depicted in Fig. 2.6.

Comparing the observed error evolution with the quantified contributions to the error growth (RES in Eq. 2.7), reveals a systematic negative residual of considerable amplitude, in particular after day 2. From the processes contributing to the residual (see Sec. 2.2.2), numerical diffusion can be expected to have a systematic negative effect on the error evolution. Nonconservation of PV by the dynamical core, which is related to numerical diffusion, has been shown to yield PV tendencies of similar magnitude as diabatic processes (Chagnon and Gray, 2015; Saffin et al., 2016). From a physical point of view, numerical diffusion is associated with a smoothing of small-scale PV features. PV forms rich filamentary structures near the tropopause (e.g., Appenzeller et al., 1996), i.e. the cascade of potential enstrophy (P^2) is directed towards the smaller scales. Errors that are associated with such filaments, e.g. displacement errors, are thus eliminated when the scale of the filament reaches the resolution of our data. Effectively, this elimination constitutes irreversible horizontal mixing (e.g., Nie et al., 2016), a nonconservative process that cannot be quantified in this study.

A comparison of the individual tendencies clearly demonstrates that the error growth, integrated over the Northern Hemisphere from 30°N to 80°N, is dominated by the near-tropopause tendency, which yields the largest contribution to the error growth at all time steps (Fig. 2.6a). The most conservative estimate of the near-tropopause tendency would be to add all uncertainty of our diagnostic (residual and uncertainty of PPVI, i.e. terms RES and UCT in Eq. 2.7) to the near-tropopause tendency. However, even together with this error sink, the near-tropopause tendency still makes by far the largest contribution to error growth. The divergent tendency makes the second largest contribution to the error growth, but is at all times much smaller than the near-tropopause tendency. Furthermore, we find only a minor contribution from the tropospheric-deep tendency to the error growth. This tendency contributes, on average, only 14 % to the observed error growth. The nonconservative tendency makes a small, but continuously negative contribution to the error evolution. A further partitioning of this tendency reveals that it is mostly governed by the radiation and turbulence tendency (Fig. 2.7). As mentioned in Sec. 2.3.2, these tendencies have rather opposite effect on the error evolution. While the radiation tendency makes a significant positive contribution, the turbulence tendency makes a contribution of similar magnitude, but of opposite sign.

As demonstrated above, the error growth in this case most prominently occurs within the

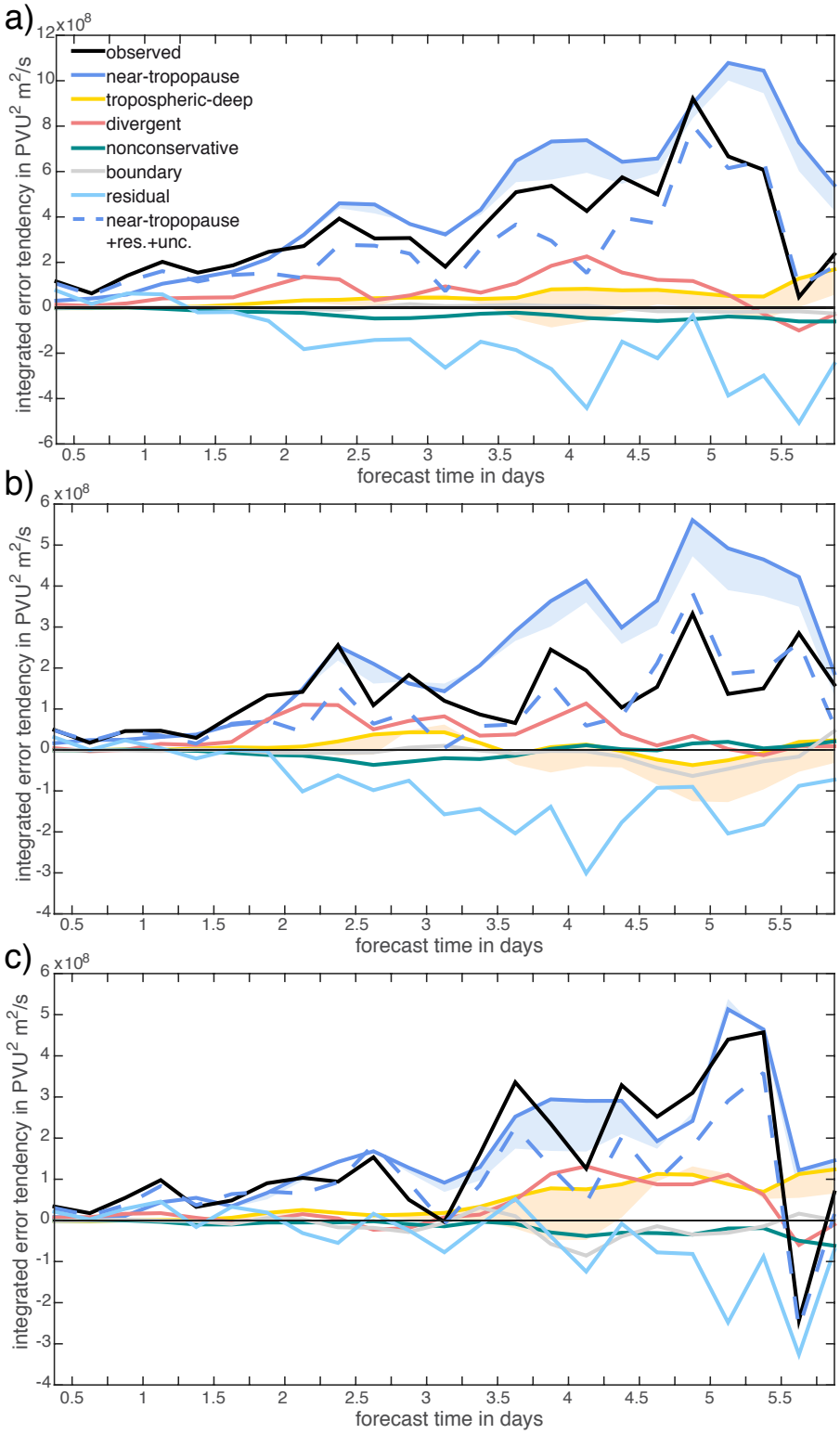

Figure 2.6: Individual contributions (colored lines) to spatially-integrated error growth on 320 K as indicated by Eq. 2.7 in dependence of forecast time. The integration area is a) Northern Hemisphere (30°N–80°N), b) Rossby wave pattern in the Atlantic, and c) Rossby wave pattern in the Pacific. The integration areas in the Atlantic and Pacific are indicated by the blue boxes in Fig. 2.3. The shading next to the near-tropopause and tropospheric-deep tendencies indicates the uncertainty of the result due to the uncertainty of PPVI (UCT in Eq. 2.7).

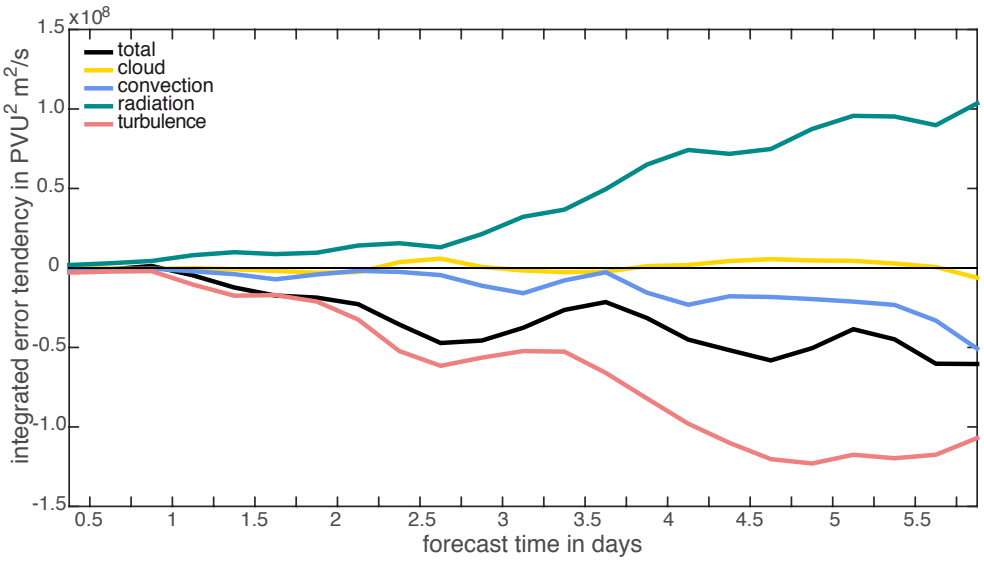

Figure 2.7: Individual contributions (colored lines) to the spatially-integrated nonconservative error growth (NONCONS in Eq. 2.7) on 320 K in dependence of forecast time. The integration area is the Northern Hemisphere (30°N–80°N).

Atlantic and Pacific Rossby wave pattern (Figs. 2.3 and 2.4). The error features in the Pacific and Atlantic are clearly separated and their amplification may exhibit different characteristics. In the following, these two regions are therefore investigated in isolation.

Within the Atlantic region, error growth during wave breaking and the highly nonlinear stage of the wave pattern (after forecast day 4) is clearly dominated by the near-tropopause tendency (Fig. 2.6b). Preceding the wave breaking, a significant contribution is also given by the divergent tendency, which is of similar magnitude as the conservative estimate of the near-tropopause tendency. During this time, the relative importance of the divergent tendency is larger than when integrated hemispherically (30–80°N, Fig. 2.6a). This large divergent tendency can be attributed to the surface cyclone located around 40°W at forecast day 2 and its accompanying warm conveyor belt that were underestimated in the forecast (surface pressure error shown in Fig. 2.3a). The forecast of the associated ridge building was of particular importance to the subsequent wave breaking, cutoff formation and high-impact weather over Sardinia. Thereby, the error growth associated with this high-impact weather was strongly related to near-tropopause dynamics and upper-tropospheric divergence.

Within the Pacific region (Fig. 2.6c), the largest contribution to error growth is again given by the near-tropopause tendency. The tropospheric-deep tendency contributes more prominently to error growth than when integrated hemispherically (30–80°N, Fig. 2.6a). Averaged over forecast day 4–5, it accounts for about one third of the observed error growth. Interestingly, this significant contribution from the tropospheric-deep tendency is accompanied by a contribution of similar magnitude from the divergent tendency. These significant contributions from the tropospheric-deep and divergent tendency can be attributed to the cyclone development on the upstream side of the Pacific ridge that was overestimated in the forecast (surface pressure error around 170°E in Figs. 2.3b and 2.5e and f).

2.4 Some insight into error-growth dynamics

The previous section has shown that PV-error growth in our case is not primarily governed by baroclinic instability (tropospheric-deep interaction) as assumed by some previous studies. Instead, we find a large contribution of near-tropopause dynamics to the error growth. This section provides a plausible explanation for the limited importance of tropospheric-deep interaction to error growth. Furthermore, we discuss the importance of differences between the nonlinear Rossby wave dynamics in the analysis and in the forecast, which can explain a large part of the observed error evolution in our case. Finally, we compare the error evolution with the underlying waviness evolution in the forecast and the analysis.

2.4.1 Limited importance of tropospheric-deep interaction to error growth

To better understand the limited importance of tropospheric-deep interaction to the error growth in our case, we consider the Eady (1949) model in the reformulation of Davies and Bishop (1994) as a well-known and arguably the simplest model that describes baroclinic instability. Details about the model setup are provided in Appendix 2.6.2. The basic idea is to prescribe idealized error patterns in this model and to apply our error tendency derived in Sec. 2.2 to quantify the tropospheric-deep contribution to error growth in the Eady model. Due to the linearity of the model, it turns out that error growth in this model exhibits the same characteristics as generic baroclinic growth.

The Eady model is nondivergent and conservative, and we will consider error growth globally, i.e., integrated over the whole domain. In the integrated PV-error tendency equation (Eq. 2.5) all terms on the right-hand-side therefore vanish, except for the first term. The dynamics of the Eady model are completely governed by the distribution of θ-edge waves at the upper and lower boundary, effectively representing the PV distribution, from which the wind field can be derived directly (Eqs. 2.17 and 2.18). Evaluating the error tendency equation in the Eady model requires specification of the upper-level analysis PV and of the PV error at the upper and lower level. A zonal background superposed by a wavenumber-6 wave (representing a synoptic-scale Rossby wave pattern) constitutes our idealized upper-level analysis state. Due to the linearity of the Eady model, arbitrary error patterns can be prescribed by the superposition of individual Fourier modes. It is therefore sufficient to consider error growth in the presence of one arbitrary wave mode in the analysis.

The dependence of the tropospheric-deep error tendency on the wavenumber and phase shift of the upper- and lower-level errors is investigated by varying these two parameters. These investigations show that tropospheric-deep error growth is only possible for a rather specific configuration, namely for identical wavenumbers of the upper and lower error wave and for a favorable phase shift between 0 and π (Fig. 2.8a). Furthermore, the error growth shows a prominent scale dependence. For the optimal growth configuration (identical wavenumbers and phase shift of $\pi/2$), the error tendency decreases exponentially with increasing wavenumber (Fig. 2.8b). Tropospheric-deep error growth in the Eady model thus shows the same

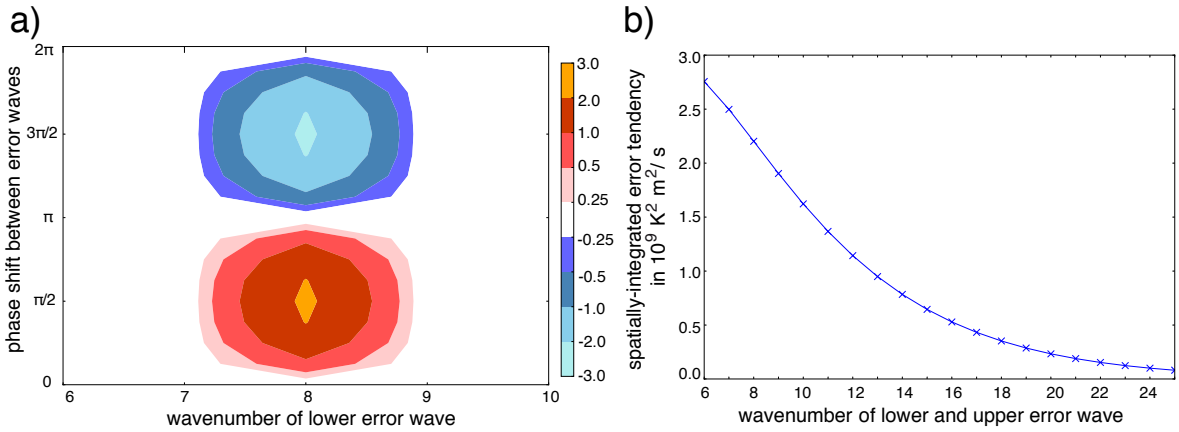

Figure 2.8: Error growth by tropospheric-deep interaction in the Eady model. a) Dependency of spatially-integrated tropospheric-deep error tendency (colored) for a fixed upper error wave (wavenumber 8, initial phase 0π) on the lower error wavenumber and phase shift between the lower and upper error wave, and b) Dependency of maximum of spatially-integrated tropospheric-deep tendency (identical wavenumbers of upper and lower error wave and phase shift of $\pi/2$) on the error wavenumber.

behavior as baroclinic growth in general.

In light of these results from an idealized model, we now examine in more detail the error pattern observed in our case study. In the Pacific, the region in which tropospheric-deep error growth is most prominent, the error pattern is indeed as suggested by Fig. 2.8a: Error features at upper- and lower-level have a similar wavelength and a phase shift of $\pi/4 - \pi/2$ (indicated by labels in Fig. 2.5c). It is evident that the associated tropospheric-deep error tendency is mostly positive in this region (Fig. 2.5d). Apart from this region, a *pronounced* coupling between the upper- and lower-level errors is not noticeable (Fig. 2.4b). From a qualitative perspective, a favorable configuration for tropospheric-deep interaction is evident in the Atlantic ridge (similar wavelength of upper- and lower-level error features and phase shift of approx. $\pi/2$) and positive tendencies prevail in the western part of the ridge. In the integrated sense of Fig. 2.6b, however, the tropospheric-deep error tendency is much weaker than in the Pacific due to cancellation with negative values in the eastern part of the ridge and near the cutoff.

Evidently, the favorable configuration for tropospheric-deep error growth is achieved only over a limited region. In our case, the organization of the error pattern into such a configuration requires 3–4 forecast days, which is significantly longer than assumed in previous studies (e.g., Zhang et al., 2007; Hohenegger and Schär, 2007). In addition, we note that the spatial scale of the error features is smaller than that of the anomalies governing baroclinic development of the midlatitudes (Figs. 2.3b and d). Due to the dependence of vertical penetration on the horizontal scale (Fig. 2.8b), the tropospheric-deep interaction of PV and θ errors can be expected to be less prominent than the tropospheric-deep interaction of respective anomalies in a baroclinically-unstable Rossby wave. Our considerations with the Eady model thus provide theoretical support for our observation that tropospheric-deep growth, i.e. error growth due to differences in the release of baroclinic instability, is not a dominant error-growth mechanism

in our case study.

2.4.2 Importance of nonlinear near-tropopause dynamics

Further insight into the important role of near-tropopause dynamics to error growth can be gained by considering the differences in the underlying linear dynamics and nonlinear dynamics. Recall from Eq. 2.3 that the error tendency is defined as the difference between the PV tendency in the forecast and in the analysis. Introducing a stationary background state and deviations thereof (see Appendix A for details), we further partition the advective part of the respective tendency as

$$\left.\frac{\partial P}{\partial t}\right|_{\text{adv}} = -\mathbf{v}\cdot\nabla_\theta P = -\overline{\mathbf{v}}\cdot\nabla_\theta\overline{P} - \mathbf{v}'\cdot\nabla_\theta\overline{P} - \overline{\mathbf{v}}\cdot\nabla_\theta P' - \mathbf{v}'\cdot\nabla_\theta P' \ , \tag{2.8}$$

where variables with an overbar denote the background state and variables with a prime the deviations thereof. The first three terms in Eq. 2.8 therefore denote the linear contributions to the near-tropopause dynamics and the last term the nonlinear contribution. Using this partitioning and noting that the background state is identical for the forecast and for the analysis, it is straight-forward to rewrite the advective error tendency as

$$\begin{aligned}\left.\frac{\partial P^*}{\partial t}\right|_{\text{adv}} &= \left.\frac{\partial P_{\text{fc}}}{\partial t}\right|_{\text{adv}} - \left.\frac{\partial P_{\text{an}}}{\partial t}\right|_{\text{adv}} \\ &= -\mathbf{v}^*\cdot\nabla_\theta\overline{P} - \overline{\mathbf{v}}\cdot\nabla_\theta P^* - \mathbf{v}^*\cdot\nabla_\theta P' - (\mathbf{v}'+\mathbf{v}^*)\cdot\nabla_\theta P^* \ ,\end{aligned} \tag{2.9}$$

where P' and $\mathbf{v}'$ refer to the PV anomalies and the associated wind in the analysis, respectively. In this equation, the first two terms denote error evolution due to differences in the linear dynamics and the last two terms error evolution due to differences in the nonlinear dynamics. We here focus on the near-tropopause dynamics and consider $\mathbf{v}'$ and $\mathbf{v}^*$ in Eq. 2.9 as associated with the near-tropopause wind. For nondivergent wind, as the near-tropopause wind, the second and fourth term in Eq. 2.9 merely redistribute the error, but do not contribute to a net amplification. In the integrated sense of Eq. 2.7, the partitioning of the near-tropopause contribution to error growth thus reads

$$\text{nTP} = -\int_A P^*\mathbf{v}^*_{\text{nTP}}\cdot\nabla_\theta P dA = -\int_A P^*\mathbf{v}^*_{\text{nTP}}\cdot\nabla_\theta\overline{P} dA - \int_A P^*\mathbf{v}^*_{\text{nTP}}\cdot\nabla_\theta P' dA \ , \tag{2.10}$$

where the first and the second term are the contributions due to differences in the underlying linear and nonlinear dynamics, respectively. Spatially integrated over the Northern Hemisphere, error growth is clearly dominated by differences in the nonlinear dynamics (Fig. 2.9). The contribution due to differences in the linear dynamics is essentially negligible in our case. It is thus not possible to explain the error evolution based on simple linear arguments for Rossby wave dynamics.

For nonlinear Rossby wave dynamics, we are not aware of an established conceptual framework that could guide a further systematic analysis of the nonlinear processes underlying the error

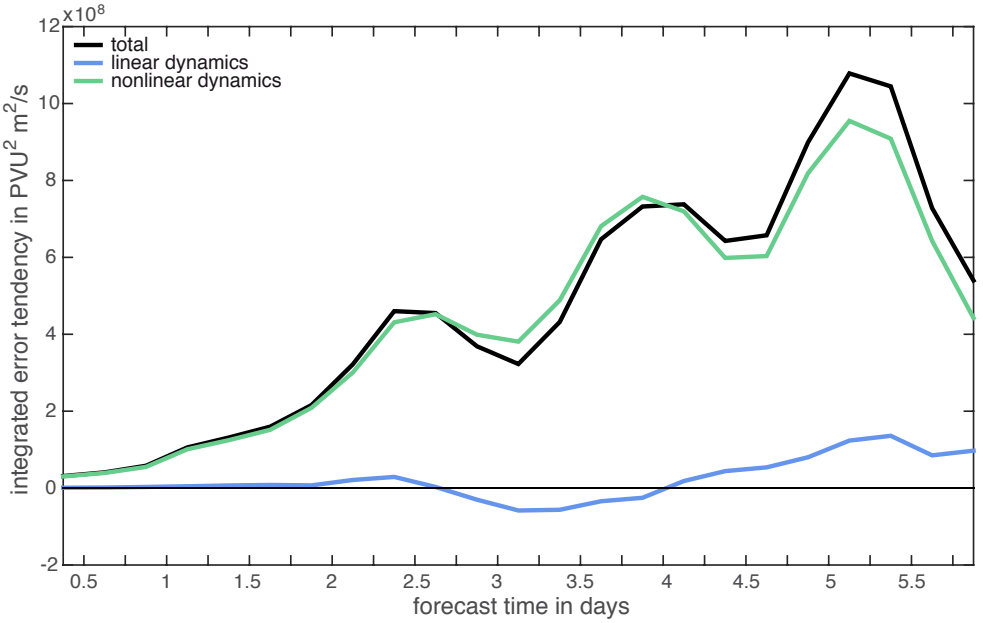

Figure 2.9: Partitioning of the spatially-integrated near-tropopause error growth (nTP in Eq. 2.7) on 320 K into the contributions due to differences in the underlying linear and non-linear dynamics, respectively, (see Eq. 2.10). The integration area is the Northern Hemisphere (30°N–80°N).

growth. Instead, we here illustrate exemplary how differences between the near-tropopause PV tendency in the analysis and in the forecast can be used to explain a large part of the observed error evolution. Specifically, we consider the localized error pattern in the Atlantic at forecast day 2 and the wave-like error pattern in the Pacific at forecast day 4.

The Atlantic error pattern at forecast day 2 is dominated by an amplitude error of the ridge, which is too weak in the forecast compared to the analysis (Figs. 2.1a and 2.10a). This difference is associated with differences in the near-tropopause PV tendency (Figs. 2.10b and c): The negative tendency on the eastern side of the ridge is of larger amplitude in the analysis than in the forecast. The tendency in the analysis is thus associated with a faster shift (phase speed) of the ridge to the east and an enhanced thinning of the downstream trough as compared to the forecast. Consistently, at day 3 the ridge in the analysis is farther to the east than in the forecast, leading to a distinct phase error generated between day 2 and 3 (Figs. 2.1b and 2.10a). Furthermore, the downstream trough in the analysis is thinner than in the forecast and exhibits a more pronounced positive tilt, indicating that anticyclonic wave breaking occurs earlier in the analysis than in the forecast. Differences in the near-tropopause tendency at day 2 thereby help to explain the observed error evolution in the Atlantic between day 2 and 3.

The Pacific wave pattern in the forecast exhibits a larger amplitude and smaller wavelength at day 4 than the analysis wave pattern leading to the observed wave-like error pattern at that day (Figs. 2.5a and 2.10d). Comparing the near-tropopause PV tendency in the analysis and in the forecast clearly indicates that these differences are associated with large differences in

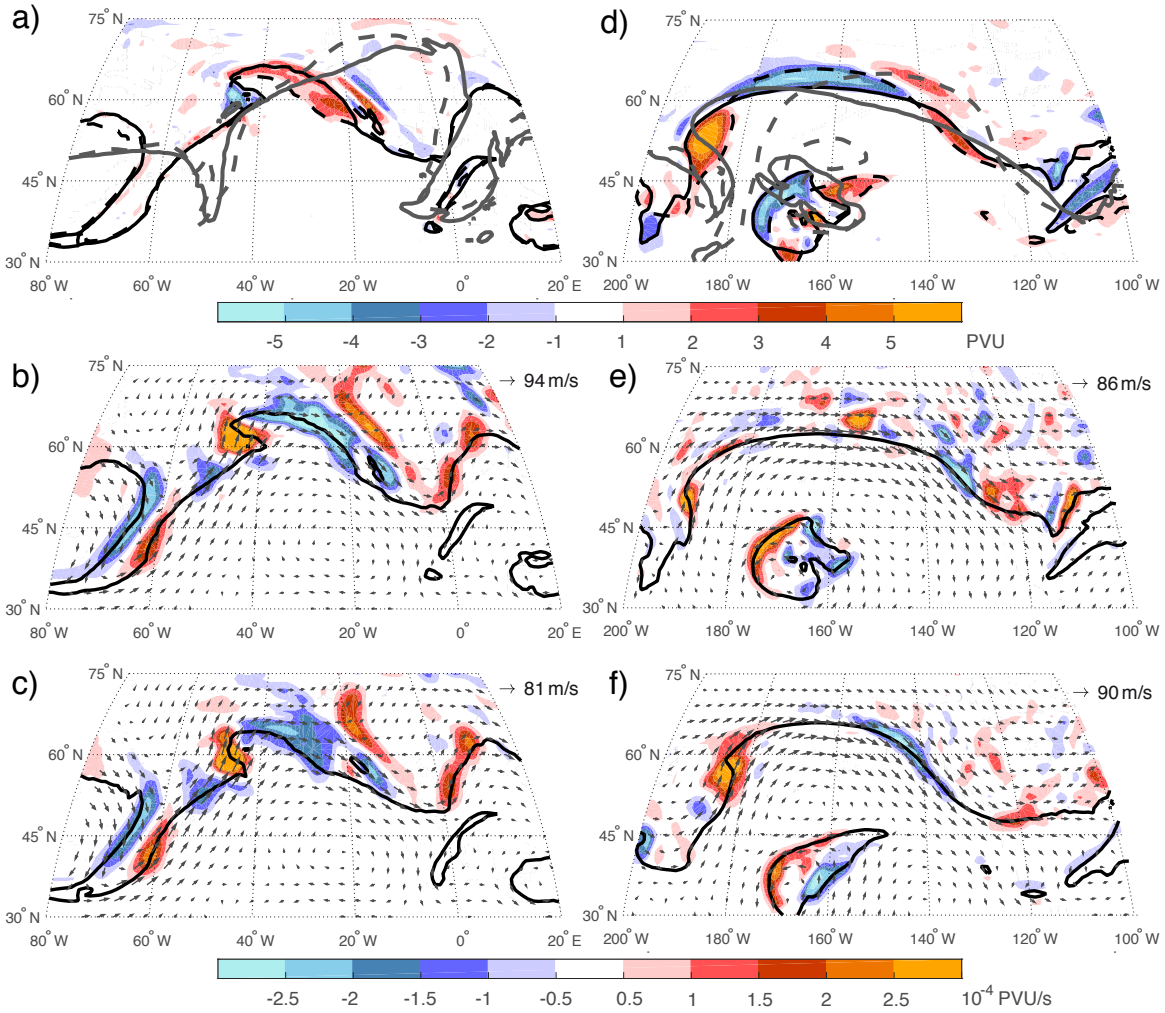

Figure 2.10: Near-tropopause error growth in the Atlantic region at forecast day 2 (a–c) and in the Pacific region at forecast day 4 (d–f). a) and d) PV error (colored) on 320 K with the black contour denoting the dynamical tropopause (2 PVU surface) of the analysis (solid) and of the forecast (dashed) on 320 K and the respective gray contour denoting the dynamical tropopause one forecast day later. b), c), e), and f) Near-tropopause PV tendency (colored, smoothed over 5 grid points in b and c and over 3 grid points in e and f) in the analysis (b and e) and in the forecast (c and f). The black contour denotes the dynamical tropopause (2 PVU surface) on 320 K of the analysis (b and e) and of the forecast (c and f). The arrows depict the near-tropopause wind of the analysis (b and e) and of the forecast (c and f) on 320 K.

the subsequent evolution of the wave pattern (Figs. 2.10e and f): While the forecast tendency exhibits a distinct signal of Rossby wave propagation with positive and negative tendencies on the western and the eastern flank of the ridge, respectively, the analysis tendency exhibits instead a patchy signal. The forecast ridge is thus associated with eastward phase propagation between day 4 and 5, while the analysis ridge is still at a similar location at day 5, but tilted northwestward (gray contours in Fig. 2.10d). A large part of the observed error evolution between day 4 and 5 can thus be explained by these differences in the nonlinear Rossby wave dynamics.

2.4.3 Underlying waviness evolution in the forecast and analysis

In this final subsection, we briefly compare the processes contributing to error amplification with those contributing to the evolution of the underlying Rossby wave patterns. We focus on wave amplitude, or waviness, which we here measure by potential eddy enstrophy, i.e. the square of the PV anomaly. Starting from the PV equation (Eq. 2.2), applying some manipulations as in the derivation of the spatially-integrated error tendency (Eq. 2.5), and using the partitioning of the wind field (Eq. 2.6), it is straight-forward to derive a tendency equation for waviness that quantifies the contributions from near-tropopause dynamics, tropospheric-deep interaction, upper-tropospheric divergent flow, and nonconservative processes:

$$\begin{aligned}\frac{d}{dt}\int_A \frac{P'^2}{2} dA \;&=\; -\int_A P'(\mathbf{v}_{\mathrm{nTP}} + \mathbf{v}_{\mathrm{TPd}} + \mathbf{v}_{\mathrm{div}} + \mathbf{v}_{\mathrm{uct}}) \cdot \nabla_\theta \overline{P} dA \\ &\quad + \int_A \frac{P'^2}{2} \nabla_\theta \cdot \mathbf{v}_{\mathrm{div}}\; dA + \int_A P' N\; dA - \oint_S \frac{P'^2}{2} (\mathbf{v} - \mathbf{v}_S) \cdot \mathbf{n}\; dS + \text{res.}\;,\end{aligned} \tag{2.11}$$

with $N = -\dot{\theta}\frac{\partial P}{\partial \theta} + P\frac{\partial \dot{\theta}}{\partial \theta} - g\frac{\partial \theta}{\partial p}\left(\mathbf{k} \times \frac{\partial \mathbf{v}}{\partial \theta}\right) \cdot \nabla\dot{\theta} - g\frac{\partial \theta}{\partial p}\mathbf{k} \cdot (\nabla \times \dot{\mathbf{v}})$ and residual res.. $\overline{P}$ here denotes the background PV and P' the PV anomaly, i.e. the deviation from the background PV (see appendix A for details). For our case, the individual contributions to the hemispheric (30–80°N) waviness evolution in the analysis and in the forecast are shown in Figs. 2.11a and b, respectively.

Tropospheric-deep interaction leads to amplification of the waviness in both the analysis and the forecast (Figs. 2.11a and b). Our case is thus characterized by baroclinic amplification of the wave patterns. In contrast, tropospheric-deep interaction has only a small influence on the hemispheric error evolution (Fig. 2.6a). In our case, there is consequently a difference between the processes contributing to the waviness evolution and the processes contributing to the error evolution. Previous studies (e.g., Zhang et al., 2007; Boisserie et al., 2014) often related synoptic-scale error growth to baroclinic instability since the largest error growth was found in baroclinic active regions. Our case indicates, however, that differences in the release of baroclinic instability are not necessarily a primary error growth mechanism and that it is important to distinguish the processes contributing to the error evolution from those contributing to the underlying Rossby wave evolution.

The differences between the individual contributions to the waviness evolution in the forecast

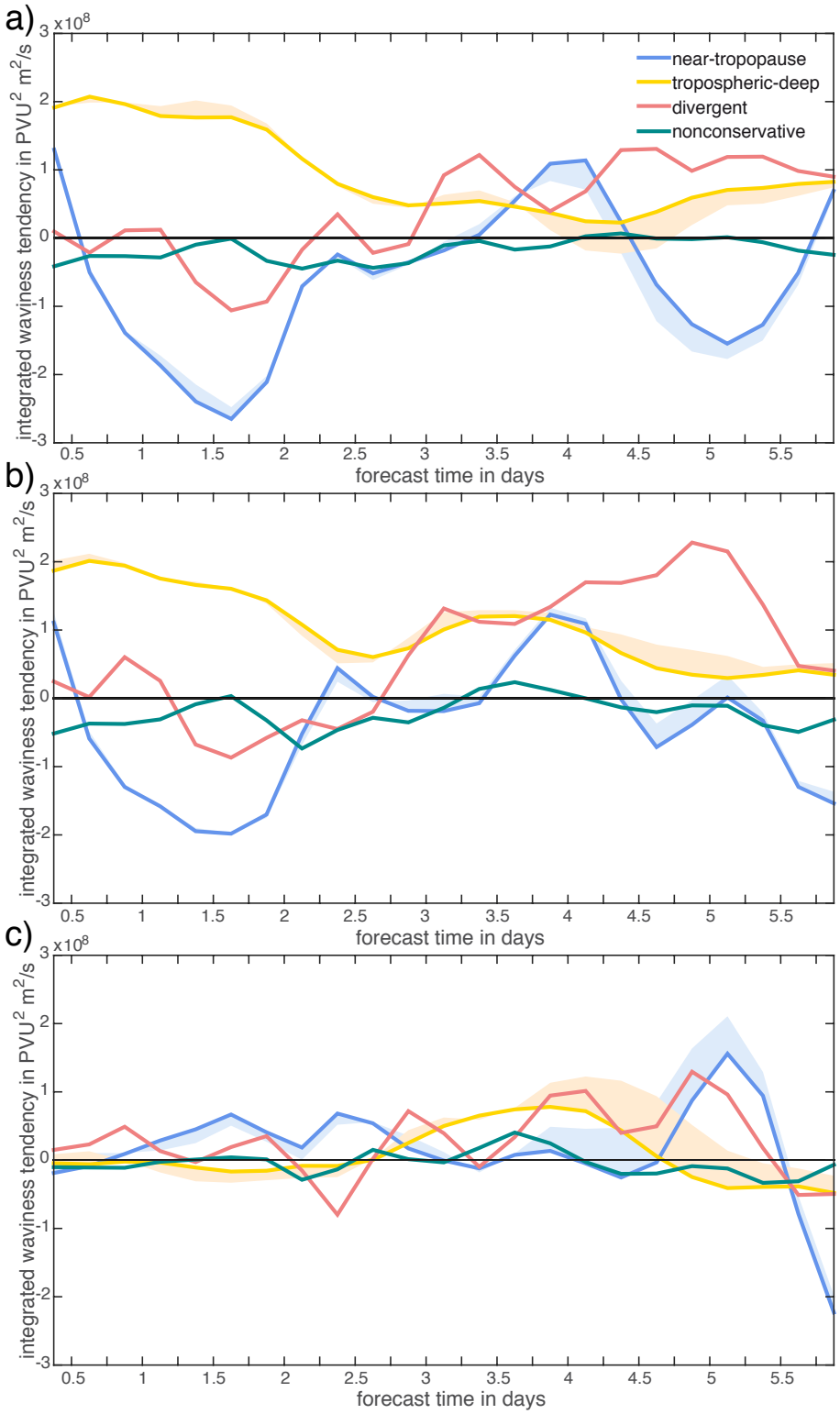

Figure 2.11: Individual contributions (colored lines) to spatially-integrated waviness evolution on 320 K as indicated by Eq. 2.12 for a) the analysis, and b) the forecast. The difference between the spatially-integrated waviness evolution in the forecast and the analysis (i.e. values of b) minus values of a)) is shown in c). The integration area is the Northern Hemisphere (30°N–80°N). The shading next to the near-tropopause and tropospheric-deep tendencies indicates the uncertainty of the result due to the uncertainty of PPVI.

and the analysis are not exceeding a magnitude of 3×10^8 PVU^{2}m^2/s (Fig. 2.11c), while the largest values of the error tendency are much larger (10×10^8 PVU^{2}m^2/s, Fig. 2.6a). Differences in the waviness evolution can thus only account for a small part of the error evolution. Thereby, this comparison supports our result that the error growth in our case is not primarily governed by differences in the (baroclinic) growth rate of the wave patterns. Since near-tropopause dynamics makes the largest contribution to the error growth, we rather suggest that the largest part of the error evolution is related to errors in the phase and the shape of the wave patterns. We hypothesize that errors generated near the tropopause, e.g. due to differences in upper-level divergence associated with latent heat release below, directly project onto the tropopause dynamics and then further amplify due to differences in the nonlinear Rossby wave dynamics (see Sec. 2.4.2). Our results thus indicate that the processes amplifying the errors can differ from the processes generating the initial error.

2.5 Summary and discussion

This study provides a quantitative analysis of error growth from a PV perspective. A tendency equation for the squared PV error is combined with the partitioning of PV tendencies as used in a previous analysis of Rossby wave dynamics (Teubler and Riemer, 2016). This allows us to quantify the relative importance of four individual contributions to the PV-error evolution near the tropopause, namely: near-tropopause dynamics (influence of upper-level PV anomalies), tropospheric-deep (baroclinic) interaction (influence of lower-tropospheric PV and θ anomalies), upper-tropospheric divergent flow, and PV modification by nonconservative processes.

The impact of upper-level anomalies on the tropopause evolution can be related to Rossby wave dynamics (e.g., Hoskins et al., 1985). In the PV framework, baroclinic instability is signified by the mutual amplification of anomalies in the upper and lower troposphere (Eady, 1949; Hoskins et al., 1985; Davis and Emanuel, 1991; Heifetz et al., 2004) and is thus captured by the tropospheric-deep tendency in our diagnostic. Upper-tropospheric divergent flow is mostly associated with latent heat release below and can thus be interpreted as an indirect impact of diabatic processes, which is of particular importance to ridge building (e.g., Davis et al., 1993; Riemer and Jones, 2010; Grams et al., 2011; Riemer et al., 2014; Teubler and Riemer, 2016). Indirect diabatic effects may be included also in the near-tropopause and tropospheric-deep tendency due to the diabatic generation of PV anomalies at earlier times. The partitioning applied in this study therefore allows a quantification of whether baroclinic instability is a dominating error growth mechanism, as assumed in many previous studies.

Error growth is analyzed in a case study of a 6-day forecast for the Northern Hemisphere. The Atlantic and Pacific Rossby wave pattern turn out to be associated with the most pronounced error growth. After 2–3 forecast days, the pattern of the PV error changes from localized error maxima to wave-like error patterns that are located along the dynamical tropopause. The largest error amplification is found within the Rossby wave patterns.

In our case, near-tropopause dynamics clearly dominates the PV-error growth near the tro-

popause, even in the conservative estimate when all uncertainties of our diagnostic are attributed to the near-tropopause contribution. The near-tropopause error amplification in our case can almost entirely be related to differences in the nonlinear dynamics. During the prominent ridge building in the Atlantic (forecast day 2–3) and in the Pacific (forecast day 4–5), the divergent flow is of comparable importance to error growth as the conservative estimate of near-tropopause dynamics. Within the Pacific wave pattern during forecast day 4–5, tropospheric-deep interaction is of comparable importance also. It is of only minor importance within the Atlantic wave pattern. Significant tropospheric-deep and divergent tendencies can be related to a misrepresentation of the respective surface cyclone development in the forecast.

Considerations with the Eady model (Eady, 1949; Davies and Bishop, 1994) provide a plausible explanation for the limited importance of tropospheric-deep interaction to the error growth in our case. This idealized model demonstrates that significant tropospheric-deep error growth may occur only when the upper- and lower-level error patterns are in a specific, favorable configuration. However, the error patterns observed in our case do not exhibit the general tendency to develop towards such a configuration. A favorable configuration develops only in a limited region around forecast day 4, which is significantly later than assumed in previous studies (e.g., Zhang et al., 2007; Hohenegger and Schär, 2007). Furthermore, the Eady model demonstrates that error growth due to tropospheric-deep interaction decreases with increasing wavenumber of the error pattern. Within the 6-day forecast period investigated here, the spatial scale of the error features is significantly smaller than that of anomalies governing baroclinic development of the midlatitudes. The small importance of tropospheric-deep interaction to error growth is thus also conceptually supported by the idealized considerations with the Eady model.

Our diagnostic reveals a systematic negative residual of considerable magnitude. In accordance with previous authors (e.g., Nie et al., 2016), we argue that this residual can be interpreted as an additional physical process related to the downscale enstrophy cascade of quasi-horizontal flow: the filamentation of PV error features and their effective diffusion at the grid scale. This process constitutes an error sink that cannot be quantified from forecast models' parameterization schemes. We are currently designing idealized numerical experiments to further investigate the role of the enstrophy cascade in the error evolution.

We emphasize that our analysis quantifies the direct impact of processes on the amplification of existing errors, in contrast to several previous studies, which tracked the accumulation of specific error sources with time using Lagrangian diagnostics (e.g., Martínez-Alvarado et al., 2016). Our approach provides useful insight into the error-growth dynamics as the processes governing the error growth may differ from the processes generating the error (e.g., Zhang et al., 2007; Hohenegger and Schär, 2007; Selz and Craig, 2015b). Using this approach, we demonstrate that at any given time of the 6-day forecast analyzed here it is mostly differences in the nonlinear near-tropopause dynamics that amplifies existing errors in the tropopause region. Differences in baroclinic instability have instead only a weak (direct) impact on the error growth in our case. Our results are in general agreement with Snyder (1999) and Plu

and Arbogast (2005), who showed that nonlinear dynamics can be as important to the error growth as baroclinic and barotropic instabilities. It is clear, however, that a single case study is not sufficient to make robust, general statements about the processes governing the error growth. Preliminary results from a few additional cases (including the "forecast bust" from Rodwell et al. (2013)), however, indicate that near-tropopause dynamics is in all of these cases of large importance to the error growth.

Arguably, the relative importance of the processes governing error growth is regime dependent. A systematic investigation of different synoptic regimes is planned to gain deeper insight into this dependence. Furthermore, we plan to adopt the method presented herein to investigate the divergence of the members in an ensemble forecast system. Thereby, the processes that govern the evolution of the ensemble spread, i.e. forecast uncertainty, could be quantified.

An improved understanding of error growth and the amplification of forecast uncertainty may help to improve the interpretation, and possibly the design, of future forecast systems. Based on our results, we hypothesize that the nonlinear dynamics of high-amplitude Rossby wave patterns is most susceptible to rapid growth of forecast errors and uncertainty. Baroclinic instability and diabatic processes, including the effect of upper-level divergence associated with latent heat release below, may play important roles by setting up such high-amplitude Rossby wave patterns. In that sense, tropospheric-deep interaction and diabatic processes may significantly influence error growth on the synoptic scale, albeit in an indirect way.

Acknowledgments

M. Baumgart and F. Teubler acknowledge support from the DFG Collaborative Research Center TRR 165, project A1: Upscale impact of diabatic processes from convective to near-hemispheric scale. Additionally, M. Baumgart appreciates funding from Internal University Research Funding (phase I) from the JGU, Mainz, and F. Teubler from the DFG research unit PANDOWAE (FOR 896). We thank Huw Davies, Tobias Selz, and George Craig for inspiring discussions. Furthermore, we are grateful for valuable comments on the manuscript from Martin Leutbecher and Nils Wedi. We thank the editor Ron McTaggart-Cowan for fostering a highly constructive review process. Comments from Gwendal Rivière and two anonymous reviewers helped to significantly improve the manuscript. Gwendal Rivière's comments were of particular importance to improve the discussion in Section 4.

2.6 Appendix

2.6.1 Partitioning of the individual processes

For the partitioning of the wind field (Eq. 2.6), we employ a Helmholtz partitioning (following Lynch, 1989) to calculate the nondivergent and irrotational wind components. The nondivergent flow is further partitioned using piecewise PV inversion (PPVI) under nonlinear balance (Charney, 1955; Davis, 1992). The PPVI is performed on the Northern Hemisphere between

25°N and 85°N on 15 evenly-spaced pressure levels, with vertical boundary conditions specified at 875 and 125 hPa. Anomalies are defined as deviations from a background state, which is here defined as the 30-day-mean, centered on 00 UTC, 14 November 2013, of the analysis. We use a mid-tropospheric pressure level (600 hPa) as the separation level between upper- and lower-level PV anomalies. The partitioning yields a residual component, which is due to the harmonic flow component, nonlinear effects in the PPVI, and uncertainties in the PPVI's boundary conditions. This residual does not exhibit a systematic impact on upper-level PV tendencies and is small enough such that it does not affect the physical interpretation of our results.

2.6.2 Eady model in the reformulation of Davies and Bishop (1994)

The Eady model (Eady, 1949) in the reformulation of Davies and Bishop (1994) comprises quasi-geostrophic channel flow with walls at $y = (0, \pi/l)$. We set $l = 6$ and the height of the domain to $d = 10^4$ m. The dynamics of the model are governed by the distribution of θ at the upper ($z = d$) and the lower ($z = 0$) boundary.
It is straight-forward to derive an error tendency equation for the Eady model analogous to Eq. 2.4, but for θ instead of PV. As the setup of the Eady model is adiabatic and nondivergent, the error tendency equation is simply given by

$$\frac{\partial}{\partial t}\frac{\theta^{*2}}{2} = -\theta^* \mathbf{v}^* \cdot \nabla\theta - \nabla \cdot \left(\frac{\theta^{*2}}{2}(\mathbf{v}^* + \mathbf{v})\right) \ . \tag{2.12}$$

Again, the second term merely redistributes the error pattern and is therefore not further considered.
Since we are interested in error growth by tropospheric-deep interaction near the tropopause (represented by the upper boundary in the Eady model), it is sufficient to analyze the tropospheric-deep error tendency at $z = d$

$$\left.\frac{\partial}{\partial t}\frac{\theta^{*2}}{2}\right|_{\mathrm{TPd},z=d} = -\theta^*_{z=d}\ \mathbf{v}^*_{TPd,z=d} \cdot \nabla\theta_{z=d} \ . \tag{2.13}$$

At $z = d$ an idealized analysis state is represented by a zonal background state and a wavenumber 6 pattern

$$\theta_{z=d} = \overline{\theta}_{z=d} + \theta'_{T,z=d} = -\frac{f\theta_S}{g}\Lambda y + \theta_{0,T}\sin(k_T x + \delta_T)\sin(ly) \ , \tag{2.14}$$

where $\Lambda = 3 \times 10^{-3}$ s^{-1} describes the uniform vertical shear, $f = 10^{-4}$ s^{-1} the Coriolis parameter, $\theta_s = 300$ K the reference temperature and g the gravitational acceleration. We choose $\theta_{0,T} = 10$ K and $\delta_T = 0$.
We introduce error patterns at the upper and lower model levels as Eady edge waves

$$\theta^*_T(z) = \frac{\sinh(\mu^*_T z)}{\sinh(\mu^*_T d)}\theta^*_{0,T}\sin(k^*_T x + \delta^*_T)\sin(ly), \tag{2.15}$$

$$\theta_B^*(z) = \frac{\sinh(\mu_B^*(d-z))}{\sinh(\mu_B^* d)} \theta_{0,B}^* \sin(k_B^* x + \delta_B^*) \sin(ly). \tag{2.16}$$

with $\mu_{T/B}^* = \frac{N}{f}\sqrt{k_{T/B}^{*2} + l^2}$ and Brunt-Väisälä frequency $N = 10^{-2}$ s^{-1}. By definition, θ_B^* and θ_T^* are zero at the respective other boundary. Since the amplitude and scale of the error patterns observed in our case study is, in general, smaller than that of the analysis pattern, we choose $\theta_{0,T}^* = \theta_{0,B}^* = 5$ K, $k_T^* \geq k_T$ and $k_B^* \geq k_T$. The tropospheric-deep contribution to the error wind is given by the wind associated with the lower error wave θ_B^*

$$u_{\mathrm{TPd},z=d}^* = \frac{gl}{f\theta_S \mu_B^*} \frac{1}{\sinh(\mu_B^*)} \theta_{0,B}^* \sin(k_B^* x + \delta_B^*) \cos(ly), \tag{2.17}$$

$$v_{\mathrm{TPd},z=d}^* = -\frac{gk_B^*}{f\theta_S \mu_B^*} \frac{1}{\sinh(\mu_B^*)} \theta_{0,B}^* \cos(k_B^* x + \delta_B^*) \sin(ly). \tag{2.18}$$

For Fig. 2.8a we consider error growth by tropospheric-deep interaction (spatial integration of Eq. 2.13) as a function of k_B^* and δ_B^* for a fixed upper-level error pattern ($k_T^* = 8$ and $\delta_T^* = 0$). Varying the parameters of the upper error wave reveals that the maximal baroclinic error growth is always attained for $k_B^* = k_T^*$ and $\delta_B^* - \delta_T^* = \pi/2$ (not shown). The magnitude of this maximum is displayed in Fig. 2.8b as a function of the error wavenumber.

CHAPTER 3

Quantitative view on the processes governing the upscale error growth up to the planetary scale using a stochastic convection scheme

Marlene Baumgart[a], Paolo Ghinassi[a], Volkmar Wirth[a], Tobias Selz[b], George C. Craig[b], and Michael Riemer[a]

[a]Institut für Physik der Atmosphäre, Johannes Gutenberg-Universität Mainz, Germany

[b]Meteorologisches Institut, Ludwig-Maximilians-Universität München, Germany

This work has been published in *Monthly Weather Review* in May 2019, Volume **147**, Pages 1713-–1731, `https://doi.org/10.1175/MWR-D-18-0292.1` (Baumgart et al., 2019).[1]

Abstract

Two diagnostics based on potential vorticity and the envelope of Rossby waves are used to investigate upscale error growth from a dynamical perspective. The diagnostics are applied to several cases of global, real-case ensemble simulations, in which the only difference between the ensemble members lies in the random seed of the stochastic convection scheme. Based on a tendency equation for the enstrophy error, the relative importance of individual processes to enstrophy-error growth near the tropopause is quantified. After the enstrophy error is saturated on the synoptic scale, the envelope diagnostic is used to investigate error growth up to the planetary scale. The diagnostics reveal distinct stages of the error growth: in the first 12 h, error growth is dominated by differences in the convection scheme. Differences in the upper-tropospheric divergent wind then project these diabatic errors into the tropopause region (day 0.5-–2). The subsequent error growth (day 2—14.5) is governed by differences

in the nonlinear near-tropopause dynamics. A fourth stage of the error growth is found up to 18 days when the envelope diagnostic indicates error growth from the synoptic up to the planetary scale. Previous ideas of the multiscale nature of upscale error growth are confirmed in general. However, a novel interpretation of the governing processes is provided. The insight obtained into the dynamics of upscale error growth may help to design representations of uncertainty in operational forecast models and to identify atmospheric conditions that are intrinsically prone to large error amplification.

3.1 Introduction

Weather prediction has improved significantly in the past decades (Bauer et al., 2015). Forecast dropouts, however, do still occur in operational numerical weather prediction models (Rodwell et al., 2013, 2018). Due to the multi-scale nature of atmospheric dynamics, there may always be an intrinsic limit of predictability even if model errors and initial-condition errors occur only on the smallest resolved scale (Lorenz, 1969). Small-scale errors associated with moist processes grow much faster than errors on the synoptic scale and saturate already after about one day (Hohenegger and Schär, 2007). Due to this scale-dependence, small-scale errors can initiate a multi-stage sequence of upscale error growth and thereby affect the forecast skill on much larger scales (e.g., Zhang et al., 2003, 2007; Selz and Craig, 2015b; Judt, 2018).

Based on an idealized moist baroclinic wave simulation, Zhang et al. (2007) derived a conceptual three-stage model for upscale error growth. A similar upscale error-growth behavior could also be found in realistic weather events (Selz and Craig, 2015b; Judt, 2018). In the first stage, errors grow quickly especially in regions of moist convection due to convective instability. This growth quickly slows down and saturates on the convective scale with a complete displacement of the individual cells (Zhang et al., 2003, 2007). The following stage is characterized by the transitioning from an unbalanced small-scale error pattern to a balanced larger-scale error pattern. Although spreading gravity waves are a prominent feature of this stage (Selz and Craig, 2015b), Bierdel et al. (2017, 2018) have argued that these have only transient effects that are small in comparison to changes in the net divergence forced by convection. In the last stage, the model suggests that errors grow on the synoptic scale with the background baroclinic instability.

Upscale error growth has often been investigated using error-energy spectra (e.g., Zhang et al., 2007; Selz and Craig, 2015b; Judt, 2018; Selz, 2019). These spectra provide insight into the scale of the error, but do not provide insight into the processes governing the error growth. One diagnostic to provide such insight into the processes governing the error growth is based on potential vorticity (PV, Snyder et al., 2003; Dirren et al., 2003; Davies and Didone, 2013; Baumgart et al., 2018). Recently, Baumgart et al. (2018) applied this diagnostic to a case study of an operational forecast. In the presence of the initial-condition and model errors of the investigated forecast of a state-of-the-art operational model, the error growth was dominated by differences in the nonlinear near-tropopause dynamics. This work suggests that

these near-tropopause interactions might be more important to synoptic-scale error growth than the direct impact of baroclinic instability. For upscale error growth, one expects that the processes dominating the error growth are more directly related to moist processes than in the operational forecast investigated by Baumgart et al. (2018). To gain deeper insight into the role of moist processes to upscale error growth, we here apply the same diagnostic framework as in Baumgart et al. (2018) to simulations, in which the initial differences stem only from the representation of convection.

On the synoptic scale, forecast errors saturate when troughs and ridges in an upper-level Rossby wave pattern are completely out of phase. Beyond this synoptic-scale phase saturation at long lead times (beyond 2 weeks), the results of Buizza and Leutbecher (2015) indicate that there is still forecast skill for large-scale fields. To investigate error growth up to the planetary scale, we employ a complementary diagnostic that filters out phase information and identifies the envelope of the upper-level Rossby waves. This diagnostic is based on finite amplitude local wave activity (LWA) in the primitive-equation, isentropic-coordinates framework following Ghinassi et al. (2018). It has been shown recently that the LWA-based diagnostic is particularly suitable in identifying finite amplitude wave packets and related nonlinear phenomena such as wave-breaking and blocking (Ghinassi et al., 2018). Using this phase-filtered envelope diagnostic, we are capable to investigate error growth up to the planetary scale and long lead times.

Simulating explicitly the upscale error growth from the convective up to the planetary scale requires convection-permitting simulations on very large, preferably global, domains and for long lead times. Such an approach is computationally extremely costly, but has recently been performed for a single case by Judt (2018). This approach, however, is currently not feasible to be applied to a large number of cases with several ensemble members to ensure the robustness of the results. Using coarser resolution with a deterministic scheme to parameterize convection, upscale error growth turns out to be much slower (Zhang et al., 2003; Selz and Craig, 2015a). A way out of this problem was suggested by Selz and Craig (2015a) who showed that a stochastic convection scheme, namely the scheme of Plant and Craig (2008), could generate error amplitudes on meso- and synoptic scales similar to a convection-permitting simulation. The scheme models the individual updrafts in a grid box and draws their magnitude randomly from a distribution. A different realization of the convection can thus be achieved by changing the random seed of the scheme. The scheme thus simulates the outcome of the first error-growth stage and projects it adaptively onto the model grid. For upscale error-growth experiments, this characteristic is a crucial advantage compared to simulations with a deterministic scheme.

Recently, Selz (2019) investigated the intrinsic predictability on larger scales that originate from the short intrinsic predictability of convection using the stochastic convection scheme of Plant and Craig (2008) to represent the first stage of error growth. The coarse resolution allowed for the simulation of 12 cases, regularly distributed over one year, with 5 ensemble members each. In the current work we use this data set to apply the advanced diagnostics described above to gain more insight into the processes governing the upscale error growth.

This Chapter is organized as follows: Section 3.2 describes the upscale error-growth simulations and introduces the PV and envelope diagnostics that are used to quantify the error growth from a dynamical, process-based perspective. Section 3.3 illustrates the results of our diagnostic for one representative case study. The relative importance of the individual processes to upscale error growth is then discussed in Section 3.4. Finally, we provide a summary of our results in Section 3.5.

3.2 Data and methods

3.2.1 Global simulations with a stochastic convection scheme

We use the icosahedral non-hydrostatic global forecasting model ICON of the German weather service (Zängl et al., 2015). The model was run with a horizontal resolution corresponding to about 40 km. Convection is parametrized using the stochastic Plant-Craig (PC) convection scheme (Plant and Craig, 2008). The PC scheme uses a random value of the convective mass flux, which is drawn from a quasi-equilibrium distribution that uses the ensemble mean mass flux given by a standard CAPE closure. For each ensemble member a different seed is used to generate the random draw, resulting in a different realization of convective variability (see Fig. 1 in Selz (2019) for illustration). At the initial time, this is the only difference between ensemble members, but as the simulations evolve, differences in the resolved flow will also lead to differences in the closure.

Twelve different cases are considered, one for each month from October 2016 to September 2017. These cases are initialized 12 h before the start of the month, using initial conditions from ECMWF. After a spin-up period of 12 h, we create an ensemble of five members by changing the random seed of the PC scheme. Initially, the only difference between these members thus lies in the convective scheme. All simulations are run for a duration of 31 days. Further information on the set up of the simulations are provided in Selz (2019).

For the PV inversion, we use data with a horizontal resolution of 1°x1° on pressure levels from 900–100 hPa with a grid distance of 50 hPa. Data on pressure levels below the ground are created by horizontal linear interpolation from valid neighboring grid points. We use hourly data for the first 10 days and three-hourly data afterwards. For the further analysis of the error growth, data are interpolated from pressure to isentropic levels. Different isentropes are used for each case to account for the seasonal dependence of the tropopause location. For each case, the isentrope is chosen such that the tropopause is located at approximately 45°N (see climatology of Liniger and Davies, 2004), resulting in the choice of 315 K for Dec., Jan., and Feb., 320 K for Nov., Mar., and Apr., 325 K for Oct. and May, 330 K for Jun. and Sep., and 335 K for Jul. and Aug..

3.2.2 PV error and tendency equation

We use Ertel-PV (Ertel, 1942) as our key variable. In the framework of the primitive equations and with potential temperature θ as vertical coordinate, it is defined as

$$P = \frac{1}{\sigma}\left(\zeta_\theta + f\right) , \tag{3.1}$$

where ζ_θ is the vertical component of the isentropic relative vorticity, f is the Coriolis parameter, and $\sigma = -g^{-1}(\partial p/\partial\theta)$ is the isentropic layer density with gravitational acceleration g and pressure p.

The PV tendency is given by advection and nonconservative PV modification,

$$\frac{\partial P}{\partial t} = -\mathbf{v}\cdot\nabla_\theta P + N + \text{res.} , \tag{3.2}$$

where $N = -\dot{\theta}\frac{\partial P}{\partial\theta} + P\frac{\partial\dot{\theta}}{\partial\theta} + \frac{1}{\sigma}\mathbf{k}\cdot(\nabla\times\dot{\mathbf{v}})$ describes the nonconservative PV modification due to diabatic heating and nonconservative momentum change. The heating rates $\dot{\theta}$ and horizontal wind tendencies $\dot{\mathbf{v}}$ are approximated by centered differences of accumulated tendencies[2] from the parametrization schemes (convection, microphysics, long- and shortwave radiation, turbulence, subgrid-scale orography, and gravity wave drag). The residual (res.) includes the influence of those nonconservative processes that cannot be quantified with the available data as, e.g., the contribution of numerical diffusion, and intrinsic numerical errors due to the discretization and interpolation of the data.

Our focus is on the difference in PV between the individual ensemble members, ΔP, defined as

$$\Delta P = P_\text{n} - P_\text{m} , \tag{3.3}$$

where P_n and P_m denote the PV of two different ensemble members n and m. Our ICON ensemble contains five members. It is thus possible to define the PV difference for ten different member pairs: member 2 – member 1, member 3 – member 1, member 3 – member 2, ...[3] In the following, we use the terminology 'error' for the difference between two members although, strictly speaking, none of the members describes the true evolution.

In order to have a positive definite metric, we consider the (potential) enstrophy error, $(\Delta P)^2/2$. The mean enstrophy error $\mathscr{P}$ is then defined as the spatial average of the enstrophy error over the midlatitudes of the Northern Hemisphere (30–80°N),

$$\mathscr{P} = \frac{1}{A}\int_A \frac{(\Delta P)^2}{2} dA , \tag{3.4}$$

[2]Data are saved with an hourly resolution for the first ten days and with a three-hourly resolution afterwards (see Sec. 3.2.1), so the heating rates and horizontal wind tendencies are approximated by 2 h- and 6 h-centered differences, respectively.

[3]Note that for the further analysis of the error amplification it is not necessary to investigate the other possible combination of two members (e.g., one does not need to investigate member 1 – member 2 in addition to member 2 – member 1) as enstrophy and LWA are positive definite variables.

where A is the area of integration, $dA = a^2 \cos\phi d\lambda d\phi$ is the area element in spherical coordinates, a denotes the radius of the Earth, λ is longitude, and ϕ is latitude. $\mathscr{P}$ is then averaged over the 12 cases and 10 member pairs within each case, which allows us to obtain reasonably robust statistics (a sample of 120), yielding $\overline{\mathscr{P}}$. The statistical uncertainty of $\overline{\mathscr{P}}$ associated with the month-to-month variability is estimated as the standard error for the 12 cases, i.e., $\sigma(\overline{\mathscr{P}}) = \sigma/\sqrt{12}$, where σ denotes the standard deviation of the 12 cases.

To further investigate the enstrophy-error evolution, we follow the derivation of Baumgart et al. (2018) and derive a tendency equation for ΔP

$$\frac{\partial \Delta P}{\partial t} = -\Delta\mathbf{v} \cdot \nabla_\theta \overline{P} - \overline{\mathbf{v}} \cdot \nabla_\theta \Delta P + \Delta N + \text{res.}\ , \tag{3.5}$$

where

$$\Delta N = -\Delta\dot{\theta}\frac{\partial \overline{P}}{\partial \theta} - \bar{\dot{\theta}}\frac{\partial \Delta P}{\partial \theta} + \Delta P \frac{\partial \bar{\dot{\theta}}}{\partial \theta} + \overline{P}\frac{\partial \Delta\dot{\theta}}{\partial \theta} + \frac{1}{\overline{\sigma}}\mathbf{k} \cdot (\nabla \times \Delta\dot{\mathbf{v}}) + \frac{1}{\Delta\sigma}\mathbf{k} \cdot \left(\nabla \times \bar{\dot{\mathbf{v}}}\right) \tag{3.6}$$

denotes the tendency of ΔP due to nonconservative processes. Variables with a Δ denote the difference between the two members, while variables with an overbar denote the mean of the two members. We further extend the tendency for the PV error to a tendency for the enstrophy error

$$\frac{\partial}{\partial t}\frac{(\Delta P)^2}{2} = -\Delta P\ \Delta\mathbf{v} \cdot \nabla_\theta \overline{P} - \nabla_\theta \cdot \left(\frac{(\Delta P)^2}{2}\overline{\mathbf{v}}\right) + \frac{(\Delta P)^2}{2}\nabla_\theta \cdot \overline{\mathbf{v}} + \Delta P\ \Delta N + \text{res.}\ . \tag{3.7}$$

We partition the advective part of the error tendency based on the PV perspective for mid-latitude dynamics (Hoskins et al., 1985) as in Teubler and Riemer (2016) and Baumgart et al. (2018). We analyze the contributions to error growth near the tropopause by decomposing the horizontal velocity at that near-tropopause level. For that purpose, we first apply a Helmholtz partitioning to separate the divergent flow from the nondivergent flow (Lynch, 1989). The nondivergent flow is then further partitioned into the flow associated with upper- and lower-level PV anomalies, respectively, by using piecewise PV inversion (PPVI, Davis and Emanuel, 1991; Davis, 1992). Anomalies are defined as deviations from a background state, which for each case is defined as the temporal mean over the 31 days of simulation time averaged over the five members. Piecewise PV inversion is performed between 25–85°N and 850–150 hPa. Vertical boundary conditions are specified as potential temperature anomalies at 875 and 125 hPa. 600 hPa is used as the separation level between upper- and lower-level PV anomalies as PV errors are, in general, only small at this mid-level (Fig. 2 in Baumgart et al., 2018). The partitioning includes uncertainties due to the harmonic flow component, the horizontal boundary conditions, and the nonlinearities of the piecewise PV inversion. In general, these uncertainties are small and do not affect the physical interpretation of the results.

The influence of upper-level PV anomalies (by PV advection due to the associated winds) on the near-tropopause evolution is here interpreted as the contribution from near-tropopause

dynamics ($\mathbf{v}_{\mathrm{nTP}}$). As shown by Baumgart et al. (2018), this contribution to near-tropopause error growth is mostly related to differences in the nonlinear dynamics.[4] The influence of lower-level PV and θ anomalies on the near-tropopause evolution is here interpreted as the contribution from tropospheric-deep interaction ($\mathbf{v}_{\mathrm{TPd}}$). This contribution describes the vertical interaction between upper- and lower-levels and includes the basic mechanism for baroclinic instability (Eady, 1949; Hoskins et al., 1985; Davis and Emanuel, 1991; Heifetz et al., 2004). The influence of upper-tropospheric divergence ($\mathbf{v}_{\mathrm{div}}$) can be associated with balanced dynamics and diabatic processes (as one would obtain from an omega equation, see, e.g., Chapter 6.4 in Holton and Hakim, 2013). A strong divergence contribution is often associated with latent heat release below (Davis et al., 1993; Riemer et al., 2014; Quinting and Jones, 2016). In summary, our flow partitioning results in

$$\mathbf{v} \approx \mathbf{v}_{\mathrm{nTP}} + \mathbf{v}_{\mathrm{TPd}} + \mathbf{v}_{\mathrm{div}} \ . \tag{3.8}$$

We insert this flow partitioning in Eq. 3.7 and spatially integrate over the midlatitude Northern Hemisphere (30–80°N). Thereby, we can quantify the relative importance of near-tropopause dynamics (term 1 on the right-hand side of Eq. 3.9), tropospheric-deep interaction (term 2), upper-tropospheric divergence (term 3 and 4), and nonconservative processes (term 5) to enstrophy-error growth

$$\begin{aligned}\frac{d\mathscr{P}}{dt} &= \frac{1}{A}\int_A \frac{\partial}{\partial t}\frac{(\Delta P)^2}{2} dA = \frac{1}{A}\Big[-\int_A \Delta P\ \Delta\mathbf{v}_{\mathrm{nTP}}\cdot\nabla_\theta\overline{P} dA - \int_A \Delta P\ \Delta\mathbf{v}_{\mathrm{TPd}}\cdot\nabla_\theta\overline{P} dA \\ &\quad - \int_A \Delta P\ \Delta\mathbf{v}_{\mathrm{div}}\cdot\nabla_\theta\overline{P} dA + \int_A \frac{(\Delta P)^2}{2}\nabla_\theta\cdot\overline{\mathbf{v}_{\mathrm{div}}}\ dA + \int_A \Delta P\ \Delta N dA\Big] \\ &\quad + \mathrm{bnd.} + \mathrm{res.} \ .\end{aligned} \tag{3.9}$$

The boundary term, bnd. $= -\frac{1}{A}\oint_S \frac{(\Delta P)^2}{2}\overline{\mathbf{v}}\cdot\mathbf{n}\ dS$, describes the boundary effect due to errors being advected into or out of the integration domain. Due to the choice of the integration domain (hemispheric integration from 30–80°N), this boundary contribution is in general very small and will not be considered hereafter.

In addition to the enstrophy-error tendency, we consider the growth rate associated with the individual processes

$$\alpha_{\mathrm{process}} = \frac{1}{\mathscr{P}}\left(\frac{d\mathscr{P}}{dt}\right)\Big|_{\mathrm{process}} \ . \tag{3.10}$$

In contrast to the enstrophy-error-tendency diagnostic (Eq. 3.9), which provides a quantitative metric for the absolute error growth, this growth-rate diagnostic provides a quantitative metric for the error growth relative to the existing error. By combining both diagnostic, we thus have a quantitative metric for both the absolute and relative error growth.

As for the mean enstrophy error ($\overline{\mathscr{P}}$), the enstrophy-error tendency (Eq. 3.9) and growth

[4] In Baumgart et al. (2018), the partitioning into the contribution of linear and nonlinear dynamics, respectively, was based on the difference between the full near-tropopause tendency and its linearized part where the background state of the PPVI (30 day-temporal mean) was used as reference state for the linearization (see their Sec. 4b).

rates (Eq. 3.10) associated with the individual processes are averaged over the 12 cases and the 10 member pairs within each case to obtain reasonably robust statistics.[5]

3.2.3 Rossby wave envelope error

The Rossby wave envelope is obtained by, first, quantifying Rossby "waviness" through Local Finite Amplitude Wave Activity (LWA) and then extracting the envelope through a spatial filter. In our work we decided to use the envelope of LWA instead of the more established envelope of the meridional wind (Zimin et al., 2003). While both yield very similar results during the linear stage of a Rossby wave, the former was shown to be more appropriate in the case of finite amplitude waves (Ghinassi et al., 2018).
Our LWA diagnostic is defined in the framework of the primitive equations with potential temperature θ as vertical coordinate. Details can be found in Ghinassi et al. (2018). Broadly speaking, LWA is a positive definite quantity related to the displacement of PV contours from zonal symmetry. Consider a specific isentropic layer at a specific time. For any given latitude ϕ one defines an associated PV contour $Q(\phi)$ through the relation

$$\iint_{P\geq Q} \sigma \, dA = \iint_{\phi'\geq\phi} \sigma \, dA \, , \tag{3.11}$$

where dA denotes surface integration on the considered isentrope, the integration on the left hand side extends over those areas where PV is larger than Q, and the integration on the right hand side extends poleward of the considered latitude ϕ. LWA is then defined as

$$\begin{aligned} \mathrm{LWA}(\lambda,\phi) &= \frac{1}{\cos\phi}\left(\int_{l_S}[P(\lambda,\phi')-Q(\phi)]\sigma\, a\cos\phi' d\phi' \right. \\ &+ \left. \int_{l_N}[Q(\phi)-P(\lambda,\phi')]\sigma\, a\cos\phi' d\phi'\right) \, , \end{aligned} \tag{3.12}$$

where ϕ' denotes the integration variable in the meridional direction, and l_S and l_N are the arcs along the meridian at any given λ satisfying the following conditions

$$l_S: \quad P \geq Q,\ \phi' \leq \phi \, , \tag{3.13}$$

$$l_N: \quad P \leq Q,\ \phi' \geq \phi \, , \tag{3.14}$$

(for illustration see Figure 1 in Ghinassi et al. 2018). LWA has units of m s^{-1} and quantifies the strength of Rossby waves in terms of their pseudo-momentum. The zonal average of LWA exactly recovers the finite-amplitude wave activity of Nakamura and Solomon (2011). LWA is Eulerian in longitude and partly Lagrangian in latitude. This characteristic has to be kept in mind when comparing LWA-based diagnostics with our earlier-defined PV-based diagnostics, since the latter is fully Eulerian. In addition, the LWA-based diagnostic is nonlocal in the

[5] If a PV inversion at one time step did not converge for a specific case and member (about 5 % of all inversions), the average of the near-tropopause and tropospheric-deep tendency shown in Figs. 3.7 and 3.8 is taken only over the converged cases and member pairs. Due to the small number of not-converged time steps, we still expect our diagnostic to be reasonably robust.

sense that (possibly remote) nonconservative processes affect the relation between ϕ and Q. This effect renders the relation $Q(\phi)$ dependent on time; however, this time-dependence turns out to be small in the sense that $Q(\phi)$ varies by less than 10% within a one-month period and, by consequence, the difference between P and Q in the expression for LWA is dominated by variations in P. Regarding our ensembles, it is appropriate to define a common $Q(\phi)$-relation for all members, and this is achieved by using the corresponding relation at initial time where all ensemble members are identical. This guarantees that LWA at one latitude ϕ is computed with respect to the same PV contour $Q(\phi)$ for each ensemble member. For computational efficiency, we use a spatial resolution of 2°x2° for the computation of LWA.

LWA as defined above still contains the phase information of the wave. We now go on to filter out this phase dependence through a convolution with a Hann window in the zonal direction, which results in the wave envelope E. The filter is allowed to depend on longitude and its width is estimated through a wavelet analysis as described in Ghinassi et al. (2018).

In our analysis we consider the difference ΔE in the envelope between two ensemble members n and m,

$$\Delta E = E_{\mathrm{n}} - E_{\mathrm{m}} \ . \tag{3.15}$$

Note that ΔE (in contrast to E) is a signed variable and includes both amplitude and position errors of the envelope.

As for the enstrophy error, we consider the squared envelope error, $(\Delta E)^2/2$, which is a positive definite quantity. From this, the mean squared envelope error is calculated by a spatial average yielding

$$\mathscr{E} = \frac{1}{A}\int_A \frac{(\Delta E)^2}{2} dA \ , \tag{3.16}$$

where the integration area A extends over the midlatitudes of the Northern Hemisphere (20–80°N). Finally, $\mathscr{E}$ is averaged over the 12 cases and 10 member pairs within each case yielding $\overline{\mathscr{E}}$. The statistical uncertainty of $\overline{\mathscr{E}}$ associated with the month-to-month variability is estimated as the standard error for the 12 cases, i.e., $\sigma(\overline{\mathscr{E}}) = \sigma/\sqrt{12}$.

3.3 Illustrative case study

The October ensemble is used to illustrate the spatial patterns of our error-growth diagnostics. This section starts with a discussion of the differences between the first two members. Recall that the difference between two ensemble members is here referred to as error. The contributions from the individual processes to error growth are then discussed in terms of their spatial pattern. Finally, the error growth in the two-member ensemble is compared to that in the five-member ensemble.

3.3.1 Evolution of error patterns

The PV and envelope error (ΔP and ΔE) derived from the first two members of the October ensemble at selected time steps are shown in Figs. 3.1, 3.2, and 3.3. At day 1, the PV error

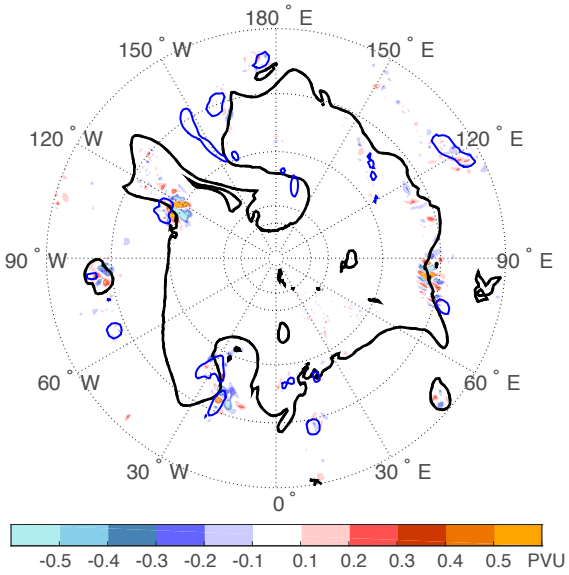

Figure 3.1: PV error derived from the first two members ($\Delta P = P_2 - P_1$) of the October ensemble at day 1. Black contours denote the dynamical tropopause of member 1 (solid) and member 2 (dashed). The blue contours depict the 1 mm/h-precipitation averaged over member 1 and 2.

occurs near the grid scale and exhibits considerable values only in localized regions (Fig. 3.1). These regions are spatially correlated with regions of precipitation. Between day 1 and 5 (Figs. 3.1 and 3.2a), the PV error amplifies by an order of magnitude (note the different scales of the colorbars) and starts to develop spatially more coherent patterns. Differences in the position of the dynamical tropopause become apparent by day 5. During these first five days the envelope error is mostly related to amplitude errors, but overall only of very small magnitude (not shown).

Between day 5 and 10, the PV error further amplifies (Fig. 3.2a and b). Importantly, the rather localized PV errors change to a larger-scale and more coherent error pattern that maximizes in regions where the tropopause is highly distorted. The PV error now projects more strongly on the envelope error (Fig. 3.3g), which shows a dipole error around $120 - 150$°W, suggesting that the error consists in a shift of the envelope. The second envelope-error maximum is a monopole located around $30 - 60$°E, suggesting instead an error in the amplitude of the envelope.

At day 15, the PV error exhibits approximately the same scale and magnitude as the troughs and ridges within the Rossby wave pattern (Fig. 3.2c). At the same time, the envelope error affects an extended part of the midlatitudes of the Western Hemisphere, with two large negative errors over the North East Pacific and North Atlantic sector and a positive error in between (Fig. 3.3h). The region with a more zonally oriented tropopause (70–140°E) shows only small errors both in PV and the envelope.

Between day 15 and 20, the PV error amplification is less pronounced than previously (Fig. 3.2c and d), indicating that the PV error starts to saturate. The Rossby wave pattern in the West-

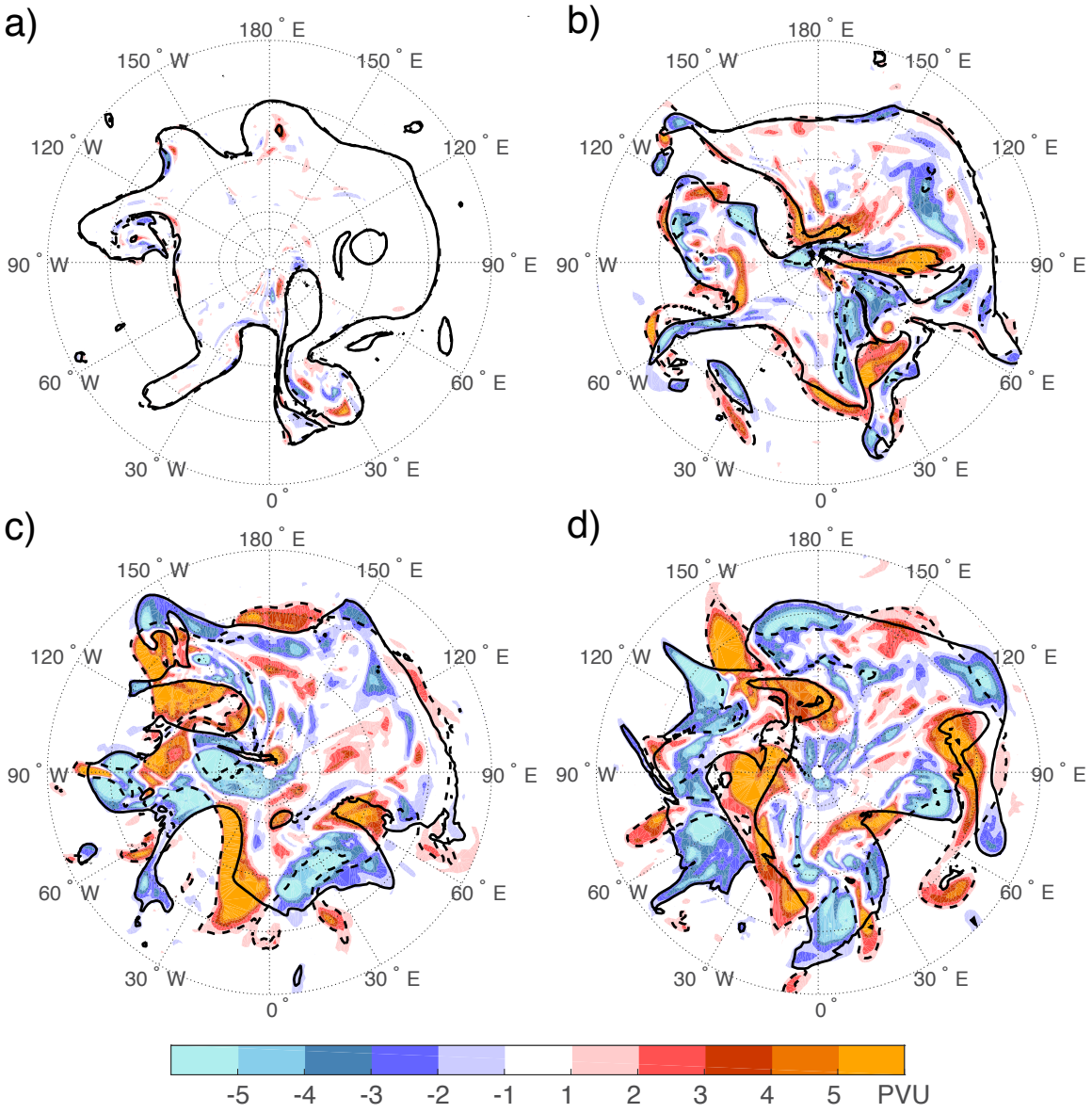

Figure 3.2: PV error derived from the first two members ($\Delta P = P_2 - P_1$) of the October ensemble at a) day 5, b) day 10, c) day 15, and d) day 20. Black contours denote the dynamical tropopause of member 1 (solid) and member 2 (dashed).

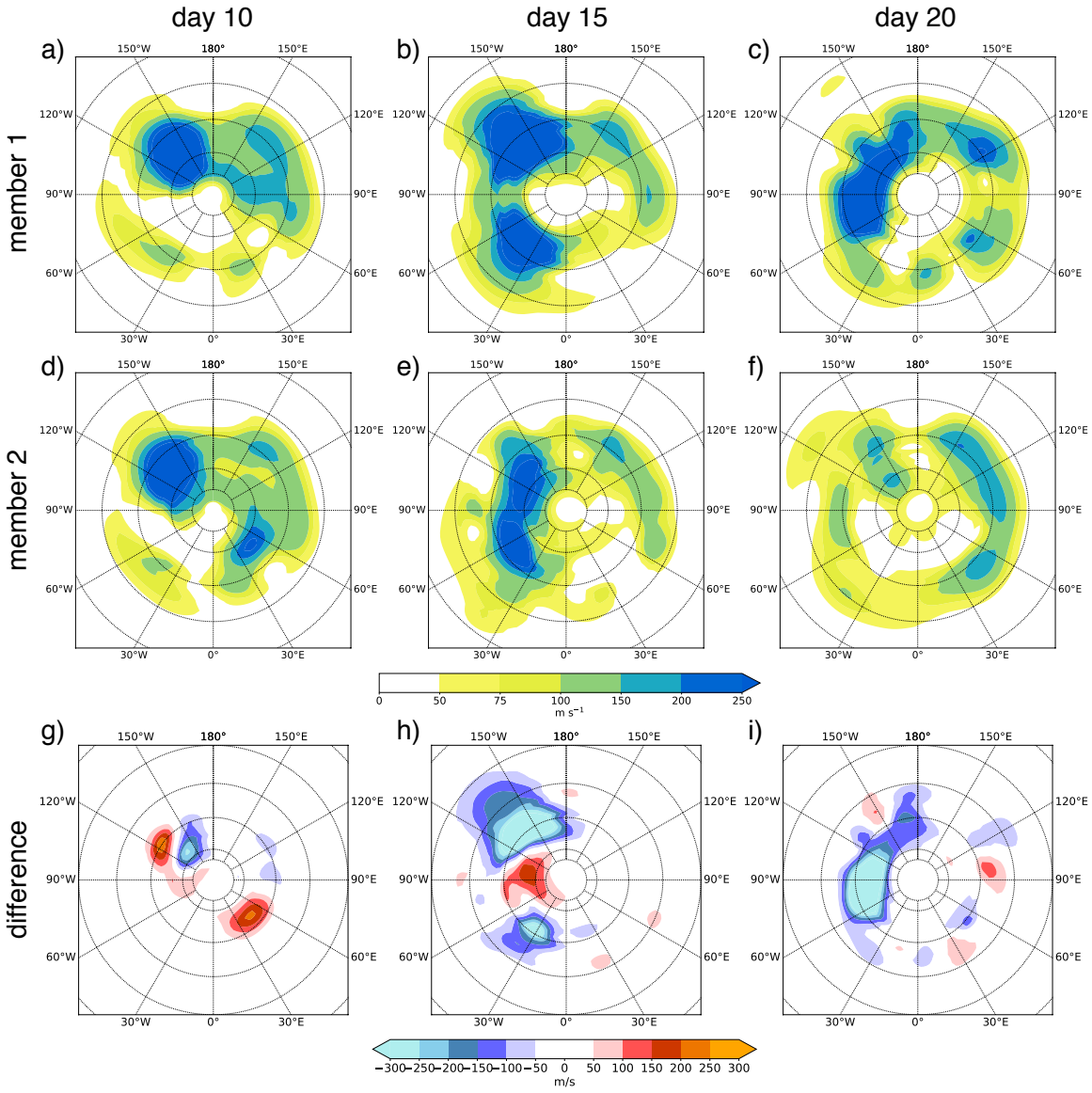

Figure 3.3: Rossby wave envelope, E, from member 1 of the October ensemble at a) day 10, b) day 15, and c) day 20 . The same for member 2 in d)–f) and for the envelope difference, ΔE, in g)–i).

ern Hemisphere seems to be completely decorrelated. This decorrelation appears as a large error in the envelope diagnostic, mostly in the form of an amplitude error (Fig. 3.3i). At the same time, more pronounced PV errors start to occur in the Eastern Hemisphere. These errors also become apparent as small errors in the envelope diagnostic.

3.3.2 Error tendencies associated with individual processes

In terms of PV, we can go one step further and relate the error evolution to individual processes. The spatial patterns of the individual contributions to the enstrophy-error tendency (Eq. 3.9) are illustrated in Fig. 3.4. The time steps and regions are chosen such that the respective process makes a dominant contribution to the error growth.
The nonconservative tendency at 18 h (Fig. 3.4a) occurs near the grid scale in localized regions of precipitation. In these regions, there is also a significant precipitation error (Fig. 3.4b). A partitioning of the nonconservative tendency into the contributions from the individual parametrization schemes (see below in Sec. 3.4.2, Fig. 3.7b) reveals that the contribution of the convection scheme dominates at this time, suggesting that differences in the latent heating induced by the convection scheme govern the error growth at that time.
The divergence tendency at day 1.5 (Fig. 3.4c) leads to error amplification along the dynamical tropopause, particularly in the two ridges and the cutoff. A large divergence tendency is found in the vicinity of cyclones where also the vertical velocity exhibits a pronounced error (Fig. 3.4d). At this early lead time, the vertical-velocity error and the associated divergence error tendency are still rather small-scale and mainly within an envelope of expected mesoscale ascent, as in the warm conveyor belt region around 20°W. This observation suggests that the differences in latent heating induced by the convection scheme (see previous paragraph) result in differences in the vertical motion within existing features of ascent, which are projected into the tropopause region by upper-tropospheric divergence. This pattern is in contrast to later lead times, when differences in vertical velocity are rather associated with a displacement of the features themselves (not shown).
The near-tropopause tendency at day 4 (Fig. 3.4e) maximizes along the dynamical tropopause on the flanks of the narrow trough and the high-amplitude ridge. Large error amplification is found in regions where the PV error aligns with a strong PV gradient (Fig. 3.4f), consistent with the observation of Snyder et al. (2003) that the PV error maximizes in regions where the PV gradient of the reference flow is large. The regions of nonlinear Rossby wave behavior are associated with large near-tropopause error growth, consistent with the large importance of nonlinear Rossby wave dynamics to the error growth found by Baumgart et al. (2018).
The tropospheric-deep tendency at day 22 (Fig. 3.4g) leads to error amplification in regions of large-scale PV errors (as can be inferred by the respective 2 PVU contours). In these regions, potential temperature errors near the surface have a similar scale and are in vicinity of the near-tropopause PV errors (Fig. 3.4h). The Eady (1949) model for baroclinic instability suggests that such a configuration is a necessary condition for tropospheric-deep error amplification (Baumgart et al., 2018). Amplification occurs when the lower-level potential

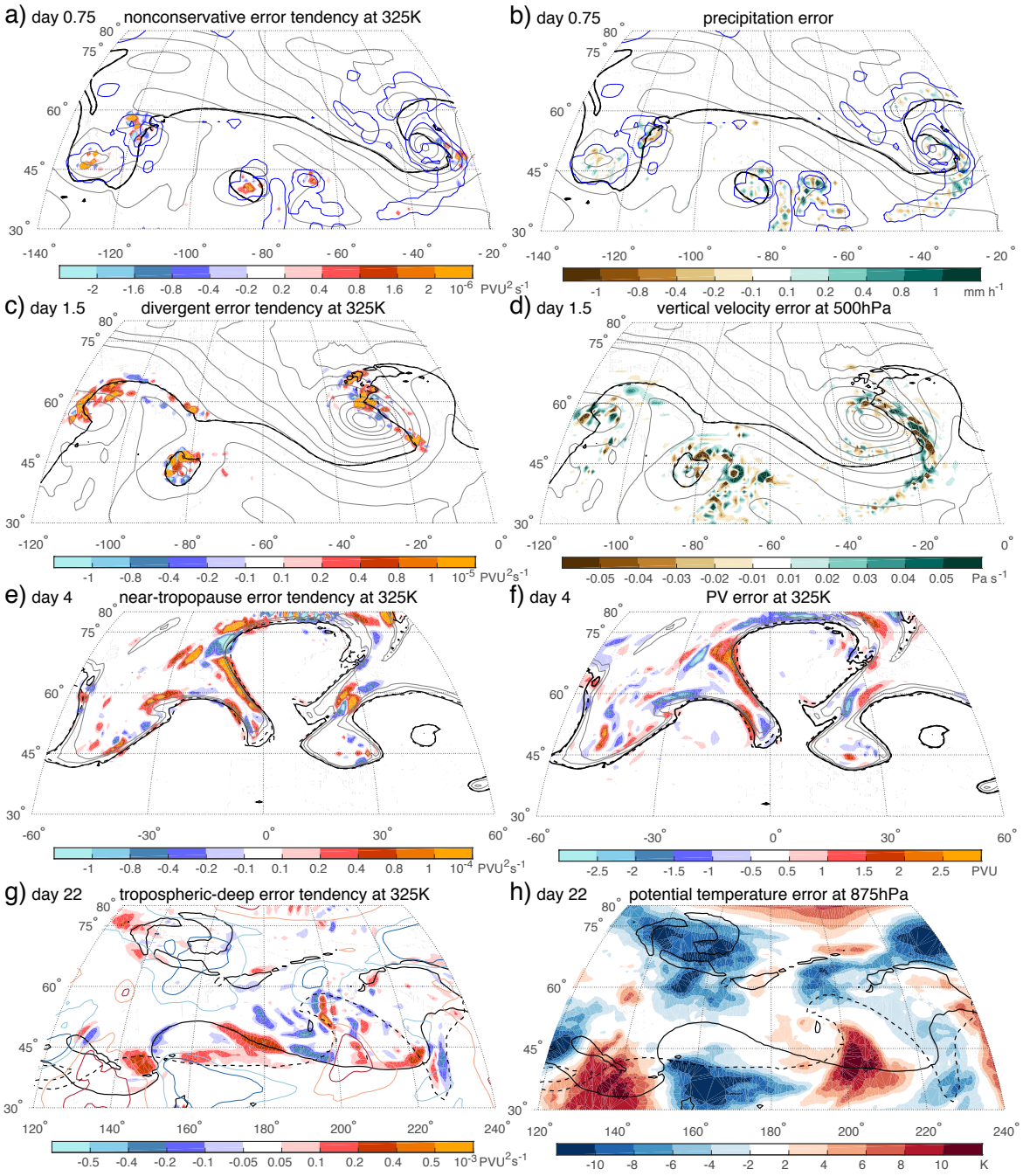

Figure 3.4: Individual contributions to the enstrophy-error tendency (Eq. 3.9, left column) together with the error field it is mostly related to (right column). The time and region of the panels are chosen such that the individual processes make a dominating contribution and thus differ for the individual panels. Thick black contours denote the dynamical tropopause at 325 K of member 1 (solid) and member 2 (dashed). Thin contours depict the mean over member 1 and 2 of a)-d) geopotential at 850 hPa every 500 m^2/s^2 (smoothed over 3 grid points) in gray, a)-b) precipitation at 0.1 and 1 mm/h (smoothed over 3 grid points) in blue, and e)-f) PV at 3 and 4 PVU in gray. Contours in g) denote the potential temperature error at 875 hPa at -4 and -8 K (blue) and at 4 and 8 K (red).

temperature errors and upper-level PV errors have a phase shift between 0 and π. The tropospheric-deep tendency thus quantifies error growth due to differences in the baroclinic growth. Before day 15 when the the error has a smaller horizontal scale than the troughs and ridges, the tropospheric-deep error amplification is small (not shown). This small importance of tropospheric-deep interaction to the error growth during the first 15 days is plausible as the vertical penetration of temperature anomalies, and thereby also of temperature errors, depends on their horizontal scale (Eady, 1949; Baumgart et al., 2018).

3.3.3 Ensemble of all five members

In the following subsection, we consider the differences within the whole five-member ensemble to investigate how representative the previous error-growth discussion of the two-member ensemble is for the five-member ensemble. For that purpose, Fig. 3.5 shows the 2 PVU contour (dynamical tropopause) of all five members for the same time steps as in Fig. 3.2.
At day 1 (not shown), no difference in the location of the dynamical tropopause is apparent between the individual members, indicating that the error occurs near the grid scale and does not yet affect the synoptic scale. At day 5 (Fig. 3.5a), differences in the tropopause location of the individual members start to occur. These differences increase significantly until day 10 (Fig. 3.5b) and an increased spread of the 2 PVU contours is evident in regions of a highly distorted tropopause. Individual troughs and ridges show both phase and amplitude errors, but are still located at similar positions. By day 15, the troughs and ridges are almost completely decorrelated (Fig. 3.5c), indicating that the PV error becomes saturated on the scale of individual troughs and ridges. The region with a rather zonal tropopause pattern ($70 - 140°$E), in contrast, stands out as a region with only small spread of the tropopause location. At day 20 (Fig. 3.5d), the differences in the ensemble are apparently not related to individual troughs and ridges any longer.
Compared to the error evolution in the two-member ensemble, the differences between all five members evolve in a similar way both in terms of the typical scale, typical magnitude, and typical location of the error: Errors generated near the grid scale start to affect the synoptic scale at around day 5 and reach the scale of the Rossby wave troughs and ridges at around day 15. This change in the scale of the error is also associated with a strong increase in the magnitude of the error. Regions with a distorted tropopause are associated with large error growth, whereas regions with a more zonal tropopause exhibit only weak error growth. We conclude that the error-growth characteristics discussed in the previous subsections for the two-member ensemble are representative for the five-member ensemble.

3.4 Quantitative view on the processes governing the error growth

To quantify the relative importance of the processes governing the error growth, we consider all twelve cases including all pairwise combinations of the five-member ensembles. The large

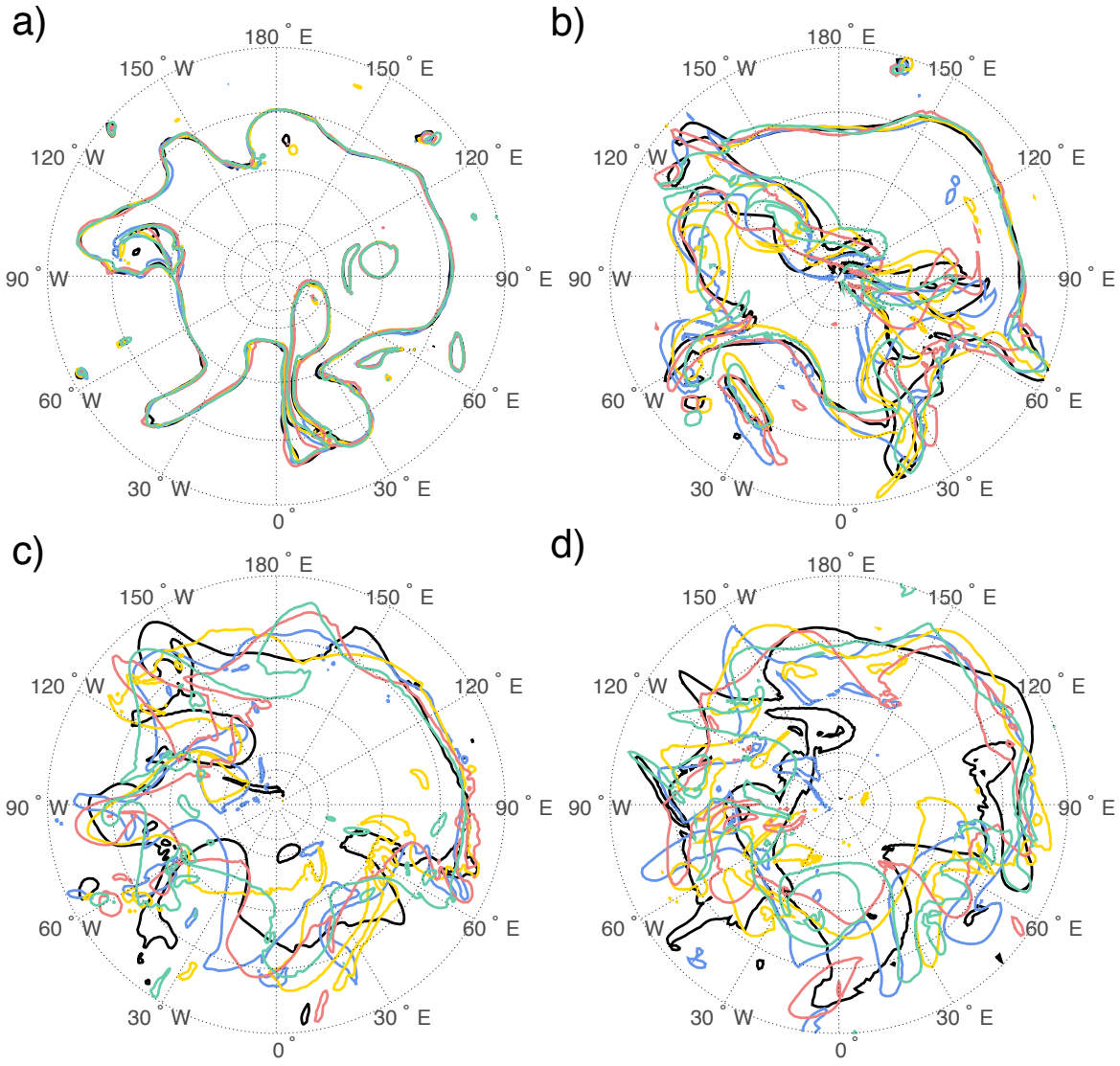

Figure 3.5: Dynamical tropopause of member 1 (black contour), member 2 (blue contour), member 3 (yellow contour), member 4 (red contour), and member 5 (green contour) at a) day 5, b) day 10, c) day 15, and d) day 20.

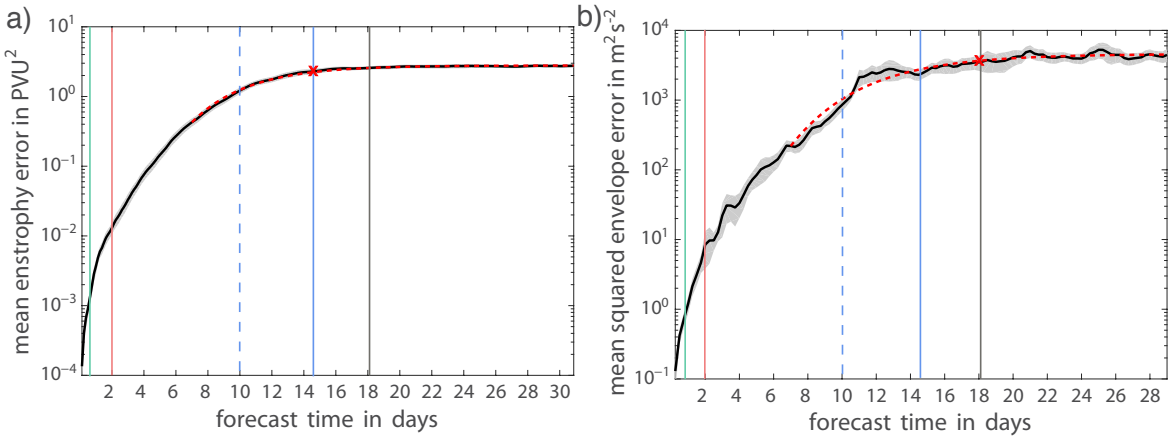

Figure 3.6: Time series of a) mean enstrophy error ($\overline{\mathscr{P}}$), and b) mean squared envelope error ($\overline{\mathscr{E}}$) over the midlatitudes of the Northern Hemisphere. $\mathscr{P}$ and $\mathscr{E}$ are averaged over the 12 cases and 10 member pairs within each case. These averages are shown as thick black lines, while the shading next to the lines denotes the (statistical) standard error for the 12 cases. The red dashed line shows the fit of the temporal development of the mean enstrophy error and mean squared envelope error to Eq. 3.17 and the red cross indicates the predictability time (as measured by Eq. 3.18) of the diagnostics. For later references, vertical solid lines denote the end of the individual stages, while the dashed blue line denotes the time of maximal near-tropopause error growth (see Figs. 3.7 and 3.8).

number of cases and member pairs provides a large enough sample for statistics to be reasonably robust, namely a sample of 120. This section starts with a discussion of time series of the mean enstrophy error and mean squared envelope error. Afterward, the contributions of individual processes to error growth are quantified. Spatial maps of the squared envelope error are then discussed to identify 'error hot spots'. Finally, our results are compared with the conceptual three-stage error growth model of Zhang et al. (2007).

3.4.1 Time series of error evolution

Time series of the mean enstrophy error ($\overline{\mathscr{P}}$) and mean squared envelope error ($\overline{\mathscr{E}}$) over the midlatitudes of the Northern Hemisphere are shown in Fig. 3.6 together with the (statistical) standard error for the cases.

Both the enstrophy and the envelope diagnostic show the typical form of an error-growth function (e.g., Selz, 2019): The initially very fast error growth seems to follow a power-law behavior in the first two days (as seen on a log-log plot, not shown), which is followed by quasi-exponential growth until day 6–7. Afterwards, error growth slows down significantly, until the error eventually reaches saturation after a certain time.

Two main differences exist between the enstrophy and the envelope diagnostic. First, the envelope diagnostic shows larger differences between the 12 cases as measured by the standard error. Second and more importantly, the envelope error saturates at a later time than the enstrophy error.

To quantify the different saturation levels in the enstrophy and envelope diagnostic, we cal-

culate the predictability time τ in the two diagnostics, according to Selz (2019). For that purpose, the temporal development of $\overline{\mathscr{P}}$ and $\overline{\mathscr{E}}$ after day 7 is fitted to the following function with the fitting parameters a and b

$$\overline{\mathscr{X}} = \overline{\mathscr{X}}(t_0) \exp\left(a\left(1 - \exp(-b(t - t_0))\right)\right) \ , \tag{3.17}$$

where t_0 is set to 7 days as in Selz (2019) and $\overline{\mathscr{X}}$ denotes either $\overline{\mathscr{P}}$ or $\overline{\mathscr{E}}$. From Fig. 3.6 it is evident that Eq. 3.17 provides a very good fit for the temporal development of $\overline{\mathscr{P}}$ and $\overline{\mathscr{E}}$, respectively.

The predictability time is defined by the time when 80% of the error magnitude at $t \to \infty$ is reached, that means

$$\overline{\mathscr{X}}(\tau) = 0.8\overline{\mathscr{X}}(t_\infty) = 0.8\overline{\mathscr{X}}(t_0) \exp(a) \ . \tag{3.18}$$

With this definition, we derive a predictability time of about 14.5 days for the enstrophy diagnostic and of about 18 days for the envelope diagnostic. For difference kinetic energy, Selz (2019) derived a predictability time of 17 days, which lies in between the result for the enstrophy and envelope diagnostic. This intermediate time can be explained by the fact that the enstrophy spectrum has larger weight on small scales than kinetic energy, while small scales are filtered out in the envelope diagnostic.

Based on Fig. 3.6, we could thus quantitatively derive a time lag of 3.5 days between the enstrophy diagnostic maximizing on the synoptic scale and the envelope diagnostic maximizing on the planetary scale. The predictability of individual ridges and troughs in a Rossby wave pattern is thus about 3.5 days shorter than the predictability of the wave-pattern envelope.

3.4.2 Relative error growth

For the early stage of the error evolution, we consider the growth rate associated with the individual processes (Eq. 3.10) to quantify the error growth relative to the existing error. Time series of these growth rates, averaged over all cases and member pairs, are shown in Fig. 3.7 together with the (statistical) standard error.

During the first 12 h, the by far largest growth rate is associated with the nonconservative processes (Fig. 3.7a). A further partitioning of this growth rate (Fig. 3.7b) indicates that it is mostly related to the convection scheme. The growth rate associated with the convection scheme is at the beginning about 7×10^{-5} s^{-1} and then sharply decreases by one order of magnitude within the first 12 h. During the first 12 h, error growth is thus dominated by the latent-heating differences that were induced by the convection scheme. In the first 12 h, the longwave radiation and microphysics scheme are associated with error growth, while the turbulence scheme is associated with error decay. These growth rates, however, are much smaller than the growth rate associated with convection in the first 12 h. After about 18 h, when the growth rate associated with the convective scheme is small, these parameterized nonconservative processes become of larger importance to the error growth than the parameterized convection.

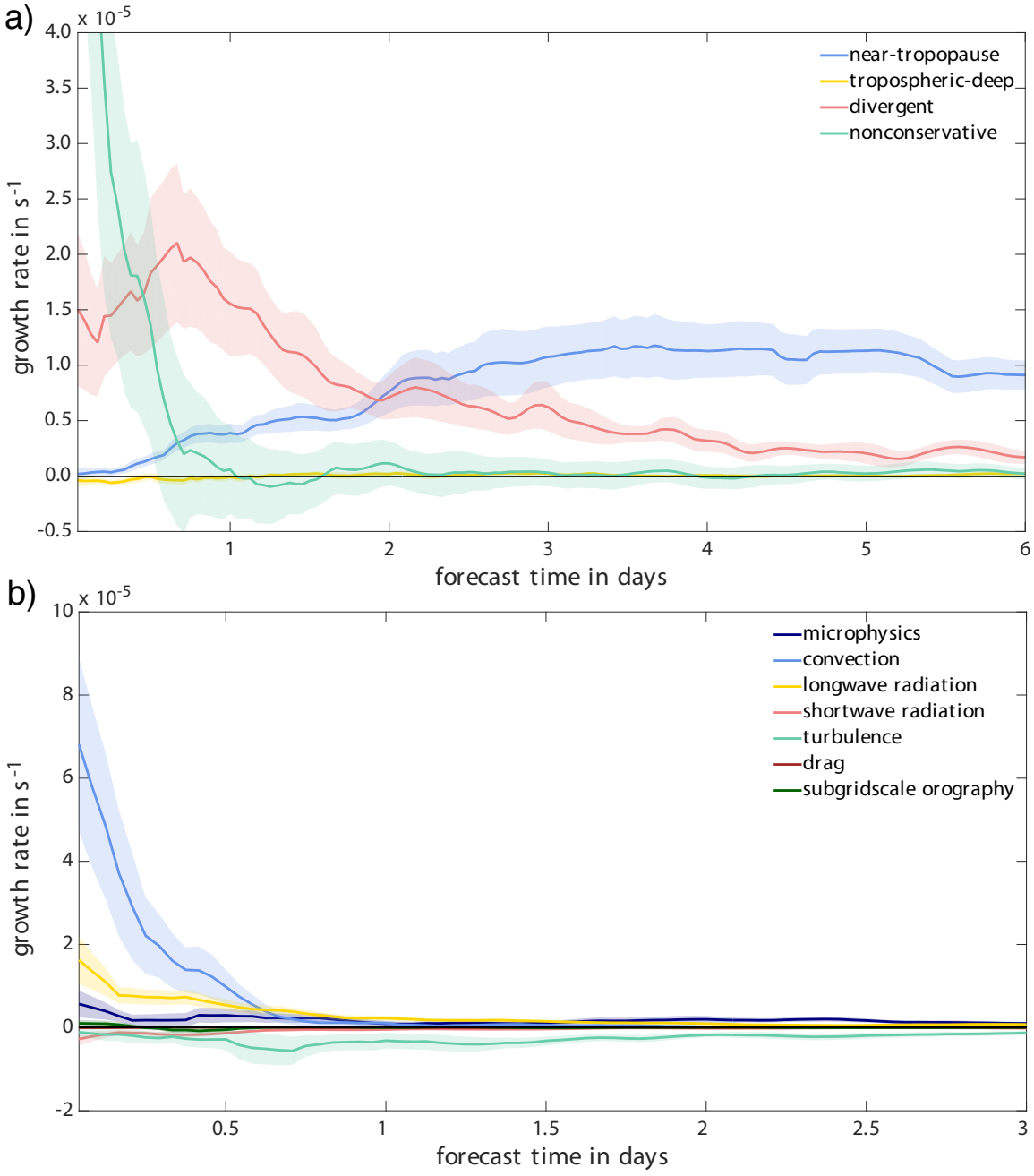

Figure 3.7: a) Growth rate of enstrophy error associated with the individual processes as indicated by Eq. 3.10, and b) further partitioning of the nonconservative contribution into the contributions of the individual parametrization schemes. Results are averaged over the 12 cases and the 10 member pairs within each case. These averages are shown as thick black line, while the shading next to the line denotes the (statistical) standard error for the 12 cases. The time series are smoothed by a running mean over 5 time steps.

A substantial positive residual exists in our budget equation for the enstrophy error in the first 12 h. This residual translates to a growth rate of similar magnitude as the growth rate associated with the convection scheme (not shown) and can be attributed to the fact that our diagnostic measures the growth of existing errors, but not the generation of new errors, which makes a dominating contribution in the very beginning of the simulations. We do not see any indication that this residual associated with the generation of new errors compromises our interpretation of the relative importance of the processes governing the error growth.
The growth rate associated with upper-tropospheric divergence exhibits large values already at the beginning of the simulations and becomes the dominating error-growth contribution after about 12 h when the growth rate associated with the convection scheme has rapidly decreased. The divergence error growth during this stage is, in general, associated with pronounced vertical-velocity errors embedded within regions of meso-scale ascent in the vicinity of cyclones (see Sec. 3.3.2, Fig. 3.4c and d). Upper-tropospheric divergence thus provides the dominating error-growth mechanism to project the latent-heating differences induced by the convection scheme into the tropopause region.
After about 2 days, the growth rate associated with near-tropopause dynamics becomes larger than the growth rate associated with upper-tropospheric divergence, indicating that differences in the nonlinear Rossby wave dynamics near the tropopause become more important to the error growth than differences in the moist processes and upper-tropospheric divergence. The growth rate associated with near-tropopause dynamics is about 1×10^{-5} s^{-1} and thereby one order of magnitude smaller than the growth rate associated with the convection scheme at the beginning of the simulations. Hohenegger and Schär (2007) found a similar magnitude for error-growth rate on the synoptic scale and also assumed that balanced dynamics are governing the synoptic-scale error growth.

3.4.3 Absolute error growth

For the later stage of the error growth, we consider the individual contributions to the enstrophy-error tendency (Eq. 3.9) to quantify the error growth in an absolute sense. Time series of the individual error-tendency contributions, averaged over all cases and member pairs, are shown in Fig. 3.8 together with the (statistical) standard error.
From the growth-rate discussion (see Sec. 3.4.2), it is clear that near-tropopause dynamics makes the largest contribution to error growth after about 2 days. Until day 10, the contribution of near-tropopause dynamics to the error growth is strongly increasing with lead time (Fig. 3.8a). In the following 7 days, it is decreasing with time, until it is eventually fluctuating around zero after about 17 days. After about 15 days, the contribution of near-tropopause dynamics is no longer the largest error-growth contribution. Interestingly, 15 days is also about the time when the predictability time of enstrophy is reached (see Sec. 3.4.1, Fig. 3.6). Large error amplification by near-tropopause dynamics thus occurs until the error becomes saturated on the synoptic scale of individual troughs and ridges.
After about 15 days, the nonconservative, tropospheric-deep, and divergence contributions be-

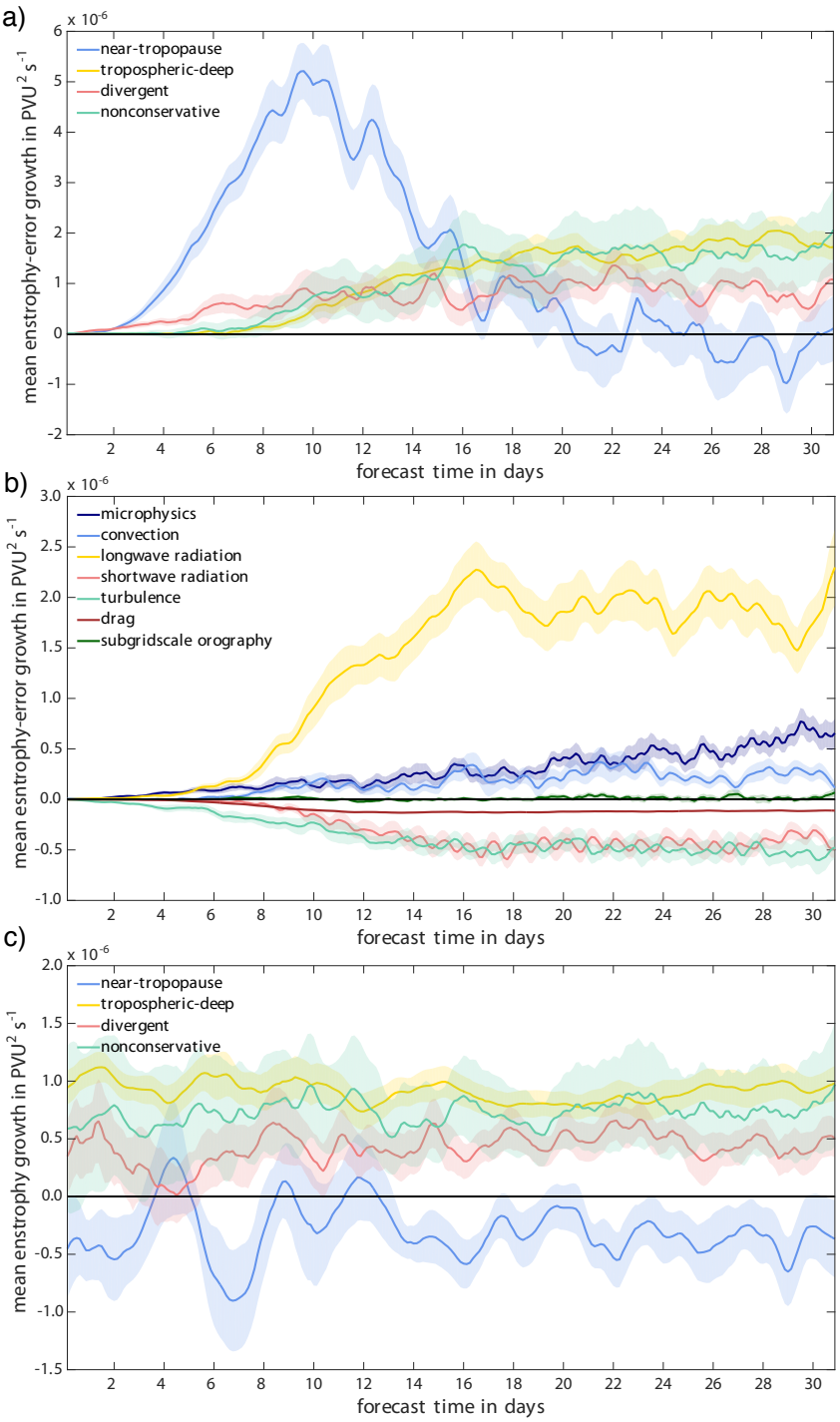

Figure 3.8: a) Partitioning of the mean enstrophy-error tendency into the contributions from the individual processes, b) partitioning of the mean nonconservative enstrophy-error tendency into the contributions from the individual parametrization schemes, and c) partitioning of the mean enstrophy tendency into the contributions from the individual processes. The individual error-tendency contributions in a) and b) are averaged over the 12 cases and the 10 member pairs within each case, while the individual enstrophy-tendency contributions in c) are averaged over the 12 cases and the 5 members within each case. These averages are shown as thick lines, while the shading next to the lines denotes the (statistical) standard error for the 12 cases. The time series are smoothed by a running mean over 5 time steps.

come more important to the error growth than the near-tropopause contribution. The large contribution of nonconservative processes is mostly related to the longwave radiation scheme (Fig. 3.8b). The longwave radiative PV tendency correlates with the PV anomalies of troughs and ridges. Once these anomalies start to decorrelate, this correlation leads to a systematic radiative error amplification. Tropospheric-deep error amplification occurs due to differences in the release of baroclinic instability when the PV and θ errors near the tropopause and near the surface, respectively, establish a favorable configuration for baroclinic growth (see Sec. 3.3.2, Fig. 3.4g and h). The divergence error amplification can be related to a misrepresentation of surface cyclones and their associated divergent outflow in warm conveyor belts (not shown).

To gain insight into the processes governing underlying Rossby wave evolution, we evaluate an analogue tendency equation for the enstrophy evolution as derived in Baumgart et al. (2018) (their Sec. 4c). The enstrophy evolution in our simulation (averaged over all cases and members, Fig. 3.8c) is governed by enstrophy amplification by the nonconservative, divergence, and tropospheric-deep contribution, while the near-tropopause contribution is weakly negative. After the enstrophy error reaches the scale of individual Rossby wave troughs and ridges (around day 15), the processes governing the error evolution are thus very similar to the processes governing the underlying Rossby wave evolution. The magnitude of the individual error-tendency contributions is about twice as large as the magnitude of the respective enstrophy-tendency contributions, which is consistent with the double penalty effect of phase errors during error saturation. In this stage, the processes governing the error growth thus resemble the climatological signal of the underlying Rossby wave evolution.

From Fig. 3.8c it is clear that the sum of all contributions to enstrophy growth is positive, suggesting that enstrophy would grow during the whole simulation time. Such continuous growth, however, is not physically meaningful as the Rossby wave amplitude is, on average, not amplifying with time. Rossby wave patterns amplify by (moist) baroclinic instability and dissipate at a similar rate by (barotropic) dissipation on the small scales due to the downscale cascade of enstrophy. In numerical weather prediction models, this dissipation appears as numerical diffusion, which is an additional enstrophy sink that cannot be quantified with the available data. The residual in the enstrophy tendency is about -2×10^{-6} PVU^2 s^{-1} and almost constant over time (not shown). A negative residual exists also in the enstrophy-error tendency. This residual increases in magnitude from day 2 to 15, until it reaches an almost constant value of about -4×10^{-6} PVU^2 s^{-1} after day 15 (not shown). This value is about twice as large as the residual in the enstrophy tendency, consistent with the observed double penalty effect of phase errors during error saturation. The important role of dissipation was also discussed in previous studies (Snyder et al., 2003; Zhang et al., 2007; Saffin et al., 2016, 2017; Baumgart et al., 2018). For instance, Zhang et al. (2007) quantified the contribution of horizontal and vertical diffusion to their difference kinetic energy budget and found that this kind of dissipation is of similar relative importance to the error evolution as the residual observed in the current study.

3.4.4 Spatial distribution of the envelope error

The spatial distribution of the squared envelope error, $(\Delta E)^2/2$, at day 5, 10, 15, and 20 is shown in Fig. 3.9. At day 5, the envelope error appears still very localized and small in magnitude. At day 10 (Fig. 3.9b), two distinct error regions are found around $120 - 90°$W and around 30°W. By day 15 (Fig. 3.9c), large error amplification occurs in the Western Hemisphere: an extended region of large error is found around $150 - 110°$W and secondary maxima are found around 180° and over an "error belt" extending from North America towards Europe. Between day 15 and 20 (Fig. 3.9c and d), error growth continues further both in magnitude and in its spatial extension. At day 20 (Fig. 3.9d), the error affects almost the whole mid-to-high latitudes of the Northern Hemisphere. At this stage, the predictability time in the envelope diagnostic is already reached (see Sec. 3.4.1, Fig. 3.6b).
From Fig. 3.9 it is evident that the envelope error tends to appear and amplify more in the western portion of the Northern Hemisphere, whereas in the eastern portion its magnitude is lower and its spatial extent more confined. This difference may be due to the fact that in our cases the Rossby wave amplitude was, in general, weaker in the eastern than in the western Northern Hemisphere.

3.4.5 Comparison with the conceptual three-stage error growth model

Our quantitative results from the PV- and envelope-based diagnostic reveal that upscale error growth can be divided into distinct stages based on the processes governing the error growth (as indicated by the vertical lines in Fig. 3.6). In the first stage (0–12 h), differences in the convection scheme dominate the error growth, while upper-tropospheric divergence dominates in the second stage (0.5–2 days). In the third stage (2–14.5 days), near-tropopause dynamics dominates the error growth. A fourth stage is found between 14.5 and 18 days, which is characterized by the planetary-scale error growth in the envelope diagnostic. In the following, we will compare these four stages with the conceptual three-stage error growth model derived by Zhang et al. (2007).
Stage 1 of the conceptual model is characterized by error growth on the convective scale in regions of moist convection (Zhang et al., 2007; Selz and Craig, 2015b; Judt, 2018). The stochastic convective scheme in our study represents explicitly resulting variability by reshuffling of the individual clouds.[6] The net effect of this stage on the grid scale in the first few hours of the simulations is dominated by differences in the diabatic PV modification from the convection scheme.
In the conceptual model, stage 2 is related to the transitioning towards error growth on the synoptic scale (Zhang et al., 2007; Selz and Craig, 2015b). In our simulations, the latent-heating differences induced by the convection scheme project on differences in the vertical motion and the associated upper-tropospheric divergent PV advection that appear within regions of meso-scale ascent in the vicinity of cyclones. The transition is in our simulation

[6]Note that this is different from a deterministic convection scheme where the individual updrafts are not simulated, but only the cloud distribution mean.

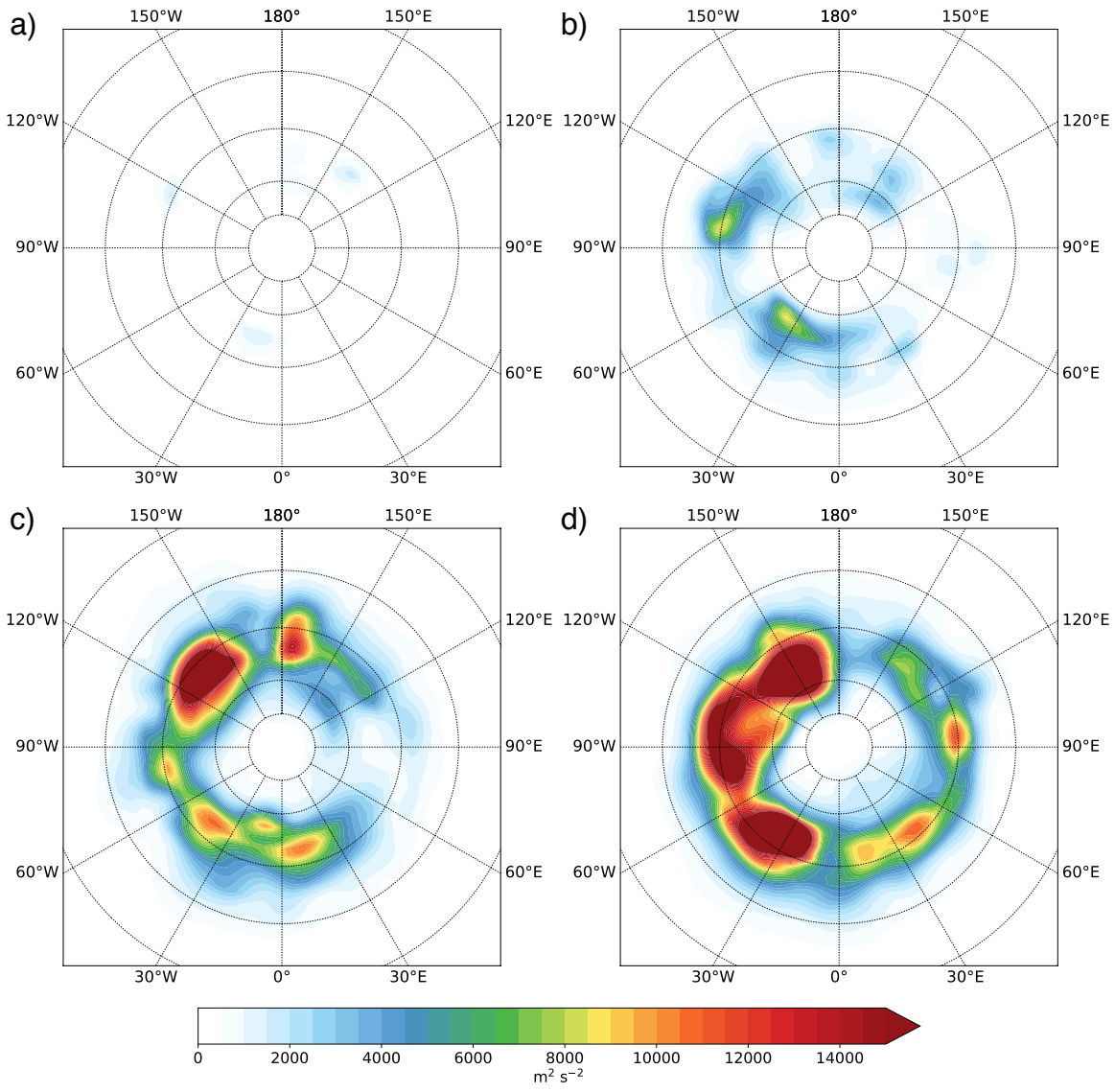

Figure 3.9: Squared envelope error $(\Delta E)^2/2$ averaged over the 12 cases and 10 member pairs within each case at a) day 5, b) day 10, c) day 15, d) day 20.

thus characterized by the divergent flow making a dominant contribution to the error growth. Thereby, upper-tropospheric divergence provides an effective mechanism to project errors from moist processes into the tropopause region.
The third stage in the conceptual model is related to error growth by the background baroclinic instability (Zhang et al., 2007). Although our diagnostic shows that the underlying Rossby wave evolution during this stage is governed by (moist) baroclinic amplification (Fig. 3.8c), we find that the error growth is governed by differences in the nonlinear near-tropopause dynamics (Fig. 3.8a). Our diagnostic thereby provides a new interpretation of the processes governing the error growth during this stage and emphasize the importance to distinguish the processes governing the underlying Rossby wave evolution from the processes governing the error growth (see also Sec. 4c in Baumgart et al., 2018).
We find a fourth stage of the error growth between about 14.5 and 18 days when the predictability time is reached for the enstrophy error, but not for the envelope error. This fourth stage could not be identified by Zhang et al. (2007) and Selz and Craig (2015b) because these studies did not have a long enough lead time to identify this stage.[7] The enstrophy error maximizes when the phase of the individual troughs and ridges in the Rossby wave pattern becomes decorrelated. The envelope diagnostic, instead, filters phase information and its error can continue to grow beyond the saturation of phase errors. The fourth stage is thus characterized by error growth from the synoptic scale of individual Rossby wave troughs and ridges up to the planetary scale.
In summary, previous ideas of a multi-scale behavior of upscale error growth are confirmed by our diagnostics. Furthermore, our results provide a novel interpretation of the governing processes and suggest a fourth stage of upscale error growth.

3.5 Summary

This study provides a quantitative view on the processes governing upscale error growth. Employing the stochastic Plant-Craig convection scheme, global ICON simulations of real cases are used to study upscale error growth induced by the convection scheme up to the planetary scale (Selz, 2019). The use of a stochastic convection scheme enables us to represent convective uncertainty and its upscale propagation at relatively low computational cost (Selz and Craig, 2015a). Thereby, it is possible to simulate 12 cases each consisting of a 5 member ensemble, in which the only difference between the ensemble members lies in the random seed of the stochastic convection scheme.
We investigate upscale error growth with a PV diagnostic (Teubler and Riemer, 2016; Baumgart et al., 2018) that allows us to quantify the relative importance of individual processes to error growth near the tropopause. The partitioning of the processes builds on the PV perspective for midlatitude dynamics and includes the influence of Rossby wave dynamics, baroclinic growth, and moist processes. Additionally, we use a diagnostic for the (phase-filtered) enve-

[7] In addition, there are differences in the experimental setup (e.g., the f-plane approximation and the periodic symmetry in the study of Zhang et al. (2007) or the limited domain size in Selz and Craig (2015b)).

lope of Rossby waves based on local finite amplitude wave activity (Ghinassi et al., 2018) to investigate error growth up to the planetary scale.

Based on the PV and envelope diagnostic, we find a distinct sequence of the processes governing upscale error growth: During the first 12 h, error growth is dominated by differences in the convection scheme. This growth is related to differences in the direct PV modification by diabatic processes resulting from their latent heating. Between 12 h and 2 days, the largest contribution to error growth is related to upper-tropospheric divergence, which projects errors from moist processes into the tropopause region. This divergence error growth is associated with vertical-motion errors in regions of meso-scale ascent in the vicinity of cyclones that result from the latent-heating differences. After 2 days, error amplification up to the synoptic scale is dominated by differences in the nonlinear near-tropopause dynamics. During this stage, the dynamics of error amplification differ distinctly from the underlying dynamics, which are that of the classical life cycle of moist baroclinic Rossby waves. The envelope diagnostic indicates that a fourth stage of the error growth exists between about day 14.5 and 18 when the predictability time in the PV diagnostic is reached, but when there is still some predictability in the envelope diagnostic. This stage is thus characterized by error growth from the synoptic scale of individual Rossby wave troughs and ridges up to the planetary scale of the (phase-filtered) wave envelope. The processes governing the error evolution during this stage, as identified by the PV diagnostic, are basically indistinguishable from those governing the Rossby wave evolution itself.

The results of our PV and envelope diagnostic confirm the existence of a multi-stage behavior of upscale error growth as observed in previous studies (e.g., Zhang et al., 2007; Selz and Craig, 2015b; Judt, 2018). Importantly, new insight is gained into the processes that govern error growth: By displacing the strong PV gradient associated with the tropopause, upper-level divergence provides an effective way to project errors associated with moist processes directly into the tropopause region. Subsequently, these errors amplify by nonlinear near-tropopause dynamics (the importance of nonlinear dynamics was shown explicitly by Baumgart et al., 2018). This notion is distinct from the projection of errors on differences in the release of baroclinic instability, which is in our simulations only of minor importance to the error growth during this stage. Our results suggest an additional fourth stage of planetary-scale error growth that was missing in the conceptual model of Zhang et al. (2007).

Due to the initial-condition and model errors contained in current state-of-the-art operational weather prediction models, one can expect a different error-growth behavior in current operational forecasts than in the upscale error-growth simulations investigated here. For instance, the error growth in a recently investigated operational case (Baumgart et al., 2018) did not show the first two stages of upscale error growth that we found in the present study. By contrast, it was dominated by differences in the (balanced) near-tropopause dynamics right from the beginning. For the same data as investigated here, Selz (2019) calculated explicitly that the error level at day 3.5 of our upscale error-growth simulations corresponds to the initial-condition uncertainty of the current ECMWF data assimilation system. Indeed, at day 3.5 upscale error growth in our simulations is already in stage 3 (error growth by

near-tropopause dynamics). These results provide evidence that the first two stages do not dominate the error growth in operational forecasts. Nevertheless, stochastic perturbations are necessary to increase the spread in current ensemble forecasting systems that would be underdispersive otherwise. In that sense, upscale error growth is of practical importance to current weather forecasts, albeit not as the dominant mechanism in an average sense. The insight into the processes governing upscale error growth obtained in this study may help to generate stochastic perturbations that effectively initiate an upscale error-growth cascade and thereby help to increase spread in current ensemble forecasts.

An improved understanding of upscale error growth may also help to identify regions that are prone to large error growth. As upscale error growth provides a rather effective mechanism to seed an error into the tropopause region within 2 days, one can expect that localized errors along the tropopause establish predominantly in regions with latent heat release below. The dominant contribution of near-tropopause dynamics in the third stage of the error growth suggests that these errors then further amplify along the dynamical tropopause, particularly in regions with a highly distorted tropopause. In that sense, the results of this study point to those regions where one can expect large forecast error growth and thus large forecast uncertainty.

Acknowledgments

The research leading to these results was carried out as part of the Collaborative Research Center SFB/TRR 165 “Waves to Weather” within project A1: “Upscale impact of diabatic processes from convective to near-hemispheric scale” funded by the German Research Foundation (DFG). We thank Franziska Teubler for providing the piecewise PV inversion code and her assistance with the code. We are grateful to Chris Snyder and two anonymous reviewers for constructive comments that helped to improve the presentation of our results.

CHAPTER 4

Processes governing the amplification of ensemble spread in a medium-range forecast with large forecast uncertainty

Marlene Baumgart[a] **and Michael Riemer**[a]
[a]Institut für Physik der Atmosphäre, Johannes Gutenberg-Universität Mainz, Germany

This work has been published in *Quarterly Journal of the Royal Meteorological Society*, Volume **145**, Pages 3252—3270, `https://doi.org/10.1002/qj.3617` (Baumgart and Riemer, 2019).[1]

Abstract

This study provides a process-based perspective on the amplification of forecast uncertainty and forecast errors in ensemble forecasts. A case from the North Atlantic Waveguide and Downstream Impact Experiment that exhibits large forecast uncertainty is analyzed. Two aspects of the ensemble behaviour are considered: (a) the mean divergence of the ensemble members, indicating the general amplification of forecast uncertainty, and (b) the divergence of the best and worst members, indicating extremes in possible error-growth scenarios. To analyze the amplification of forecast uncertainty, a tendency equation for the ensemble variance of potential vorticity (PV) is derived and partitioned into the contributions from individual processes. The amplification of PV variance is, on average for the midlatitudes of the Northern Hemisphere, dominated by near-tropopause dynamics. Locally, however, other processes can dominate the variance amplification, for example, in the region where tropical storm Karl

interacts with the Rossby-wave pattern during extratropical transition. In this region, the variance amplification is dominated by upper-tropospheric divergence and tropospheric-deep interaction and is thereby mostly related to (moist baroclinic) cyclone development. The differences between the error growth in the best and worst ensemble members can, to a large part, be attributed to differences in the representation of cut-off evolution around 3 days, which subsequently amplifies substantially in the highly nonlinear region of the Rossby-wave pattern until 5 days. In terms of the processes, the differences in error growth are dominated by differences in the error growth by near-tropopause dynamics. The approach presented provides flow-dependent insight into the dynamics of forecast uncertainty and forecast errors and helps to understand better the different contributions of specific weather systems to the medium-range amplification of ensemble spread.

4.1 Introduction

4.1.1 Flow-dependent forecast uncertainty

Ensemble forecasts have become an essential component of operational weather forecasts in the past 25 years. Weather prediction has thereby changed from a deterministic to a probabilistic approach, i.e., from providing a single forecast to providing an estimate of the uncertainty associated with a forecast and of the range of possible future scenarios (e.g. Buizza, 2019). On average, the forecast accuracy of ensemble forecasts has improved substantially over the past decades (for the European Centre for Medium-Range Weather Forecasts (ECMWF) see, e.g., Fig. 1a of Rodwell et al., 2018). In addition to forecast accuracy, forecast reliability is a second important characteristic of the quality of an ensemble forecast. A forecast system is reliable if the forecast distribution matches the observed frequency of occurrence. For a reliable forecast system, the ensemble standard deviation should match the root-mean-square error of the ensemble mean when averaged over many cases (as measured by the so-called spread–error relationship, Leutbecher and Palmer, 2008). Considering the annual mean of the spread-error relationship for the Northern Hemisphere, large improvements have been achieved in the last decades, resulting in almost indistinguishable spread-error curves in 2014 (see, e.g., Fig. 1a of Rodwell et al., 2018).

Due to the chaotic nature of atmospheric flow and the associated sensitive dependence on the initial conditions (Lorenz, 1963), atmospheric predictability exhibits a pronounced flow dependence. A central aim of ensemble forecasts is to provide a reliable estimate of this day-to-day variability of forecast uncertainty. On a day-to-day basis, and averaged over a local domain, however, the spread-error relationship holds only partially (see, e.g., Fig. 1b of Rodwell et al., 2018). While the spread-error relationship can never hold perfectly on a day-to-day basis (Whitaker and Loughe, 1998), there is arguably still room for improvement (e.g., Rodwell et al., 2018). In addition, a better understanding of the flow dependence of atmospheric predictability and associated local forecast reliability may help to improve the interpretation of ensemble forecasts.

The design of ensemble perturbations to generate appropriate ensemble spread is a nontrivial task and remains a major area of research (e.g., Bauer et al., 2015; Berner et al., 2017; Buizza, 2019). Ensemble forecast systems account for both uncertainties in the initial conditions and in the model formulation (Palmer and Hagedorn, 2006). At ECMWF, e.g., initial-condition uncertainty is accounted for by ensemble data assimilation (EDA) in combination with singular-vector perturbations, and model uncertainties by stochastically perturbed parameterization tendencies (ECMWF, 2018). There is, however, a wide range of approaches used at different operational weather prediction centers (e.g., Buizza, 2019). An improved understanding of the processes governing the amplification of ensemble spread is deemed important to design the best suitable ensemble perturbations.

Rodwell et al. (2018) presented an approach to study flow-dependent deficiencies in short-term reliability using a budget equation of ensemble variance in observation space (Rodwell et al., 2016). Deficiencies in short-term reliability could be identified for a composite of initial flow type that is associated with large forecast uncertainty, highlighting the importance of moist processes in mesoscale convective systems and in warm conveyor belts as sources for forecast uncertainty. In addition, the authors used a (spatio-temporally filtered) growth rate for the ensemble standard deviation of potential vorticity (PV) to highlight uncertainty growth that projects on the synoptic scale.

The current study considers a PV framework also. The focus here, in contrast to Rodwell et al. (2018), is not on identifying model deficiencies, but rather on quantitatively diagnosing the processes governing the flow-dependent amplification of initially small ensemble perturbations. Our PV framework builds on the PV perspective of midlatitude dynamics (Hoskins et al., 1985) and provides a quantitative partitioning of the dynamics into individual processes, including the influence of near-tropopause Rossby wave dynamics, baroclinic growth, and moist processes. The PV framework has recently been applied to quantify the governing processes of Rossby wave packets (Piaget et al., 2015; Teubler and Riemer, 2016) and a large-amplitude ridge (Schneidereit et al., 2017). The PV framework has been applied also to a case study of error growth in an operational ECMWF forecast (Baumgart et al., 2018) and to study upscale error growth in dedicated numerical experiments (Baumgart et al., 2019). Applying this framework to an operational ensemble forecast using a tendency equation for the ensemble variance of PV is a key novelty of this study. The current study focuses on the medium-range amplification of forecast uncertainty. Notwithstanding the importance of the short-range sources of forecast uncertainty, we emphasize that the amplification of forecast errors is highly nonlinear, i.e., ensemble members with relatively small errors at short lead times may have relatively large errors at medium-range lead times, and vice versa (illustrated in Fig. 4.1, see discussion below).

4.1.2 A case of large forecast uncertainty

We analyze one case from the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX, Schäfler et al., 2018), namely the extratropical transition of tropical storm Karl

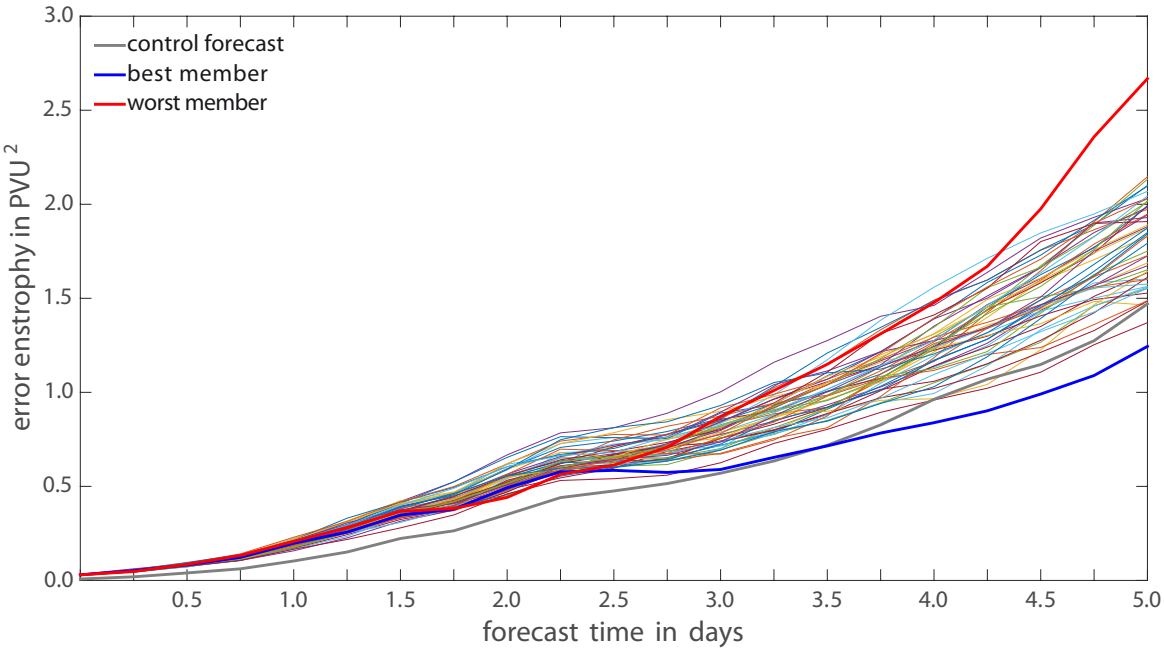

Figure 4.1: Error enstrophy of each ensemble member (spatially-averaged over the midlatitudes of the Northern Hemisphere, 30–80°N) as a function of forecast time. Thick lines denote error enstrophy of the control forecast (gray), the best member (blue), and the worst member (red).

to Ex-Karl. This case was associated with large forecast uncertainty (Schäfler et al., 2018) and the medium-range ensemble forecast showed very different developments for the interaction between Ex-Karl and the waveguide with only a few ensemble members capturing this interaction correctly (Kumpf et al., 2018). Here, the same ensemble forecast is investigated as in Kumpf et al. (2018). Besides the extratropical transition of Karl in the North Atlantic, the ensemble forecast indicates a second region of very large forecast uncertainty, which is associated with the interaction of a cutoff and a high-amplitude ridge over North America. A hemispheric perspective on the ensemble evolution is therefore provided first, before discussing important local differences from this hemispheric perspective in more detail.

To provide an overview of the ensemble evolution in our case, Fig. 4.1 shows error enstrophy of each ensemble member (spatially-averaged over the midlatitudes of the Northern Hemisphere, 30–80°N) as a function of forecast time. Error enstrophy is here defined as the squared PV error, i.e. $\frac{1}{2}P^{*2}$ with P^* denoting the PV error. Note that error enstrophy is thus directly related to a standard-error metric: the root-mean-square error. At initial time, error enstrophy of the control forecast is close to zero, whereas the perturbed members start with a notable error enstrophy that is of similar magnitude for each perturbed forecast.[2] Error enstrophy amplifies steadily for approximately 2 days, both in the perturbed forecasts and the control forecast. The control forecast maintains the smallest error until 3 day. Around 2 days, nonlinearities in the error growth become apparent also: some members with smaller error exhibit larger error-growth rates than members with larger errors, i.e., there are prominent

[2]Note that the control forecast is at the same resolution as each member.

intersections of the error curves of the individual members.
At the end of the period considered here (5 days), two ensemble members are distinct from the rest of the ensemble member in terms of error enstrophy. One of the members (in the following referred to as "worst member") exhibits a substantially larger error than the other ensemble members, whereas another member ("best member") exhibits a substantially smaller error than the other ensemble members. Interestingly, error enstrophy of the best and worst member is very similar in the first 2.5 days with both members having a relatively small error. Afterwards, the worst member exhibits, however, much larger error growth than the best member and the two members diverge prominently until they become the apparent outliers of the ensemble at 5 days.
The comparison of the error growth of the individual ensemble members illustrates two main aspects of interest: i) the mean divergence of the ensemble members indicating the general amplification of forecast uncertainty and ii) the divergence of the best and worst member indicating extremes in possible error-growth scenarios. These two aspects are discussed in more detail in the following. Section 4.2 first describes the data and methods used to quantify the amplification of forecast uncertainty and forecast errors. Then, we will discuss the amplification of forecast uncertainty both from a hemispheric and a localized perspective (Sec. 4.3). Subsequently, the error growth of the best and worst member will be compared in Sec. 4.4. We will conclude with a summary and discussion of the results in Sec. 4.5.

4.2 Data and methods to quantify the amplification of forecast uncertainty and forecast errors

4.2.1 Ensemble data

We use real-time data from the Atmospheric model Ensemble 15-day forecast (ENS) of ECMWF with a 3 h-temporal resolution in the first 3 days and 6 h-resolution afterwards. These data were archived manually on model levels in spectral space during the NAWDEX campaign. Operationally archived data on pressure levels are not sufficient for our diagnostic as the vertical resolution is too low. For the PV inversion, we interpolate the manually-archived data to a 1°x1°-grid and from model to pressure levels from 900 to 100 hPa with a grid distance of 50 hPa. For the further analysis, all variables are interpolated to isentropic levels.
We analyze the ensemble forecast initialized on September 22, 2016 00 UTC. This case is related to one of the observational highlights during the NAWDEX campaign, namely the extratropical transition of Hurricane Karl, which was related to a heavy-precipitation event in Norway (Schäfler et al., 2018; Kumpf et al., 2019). Forecasting this event was rather difficult as the evolution was very sensitive to uncertainties in the timing and location of the interaction between Ex-Karl and the midlatitude waveguide (Schäfler et al., 2018). The ensemble forecast investigated here shows very different developments of individual members in the medium-range forecast of Ex-Karl (Kumpf et al., 2018). A few members capture correctly

the interaction between Ex-Karl and the waveguide, whereas the majority of members show a distinct error during the interaction (Kumpf et al., 2018).

4.2.2 Quantitative PV framework for the amplification of forecast uncertainty and forecast errors

To attribute the evolution of forecast uncertainty to individual processes, we extend our recently developed potential-vorticity diagnostic for error growth (Baumgart et al., 2018, 2019) to the evolution of ensemble variance. In isentropic coordinates and the primitive-equations framework, Ertel (1942) PV is defined as follows

$$P = -g\frac{\partial\theta}{\partial p}\left(\zeta_\theta + f\right) , \tag{4.1}$$

where g is the gravitational acceleration, θ the potential temperature, p the pressure, ζ_θ the vertical component of the isentropic relative vorticity, and f the Coriolis parameter. PV changes locally due to advection and nonconservative processes

$$\frac{\partial P}{\partial t} = -\mathbf{v}\cdot\nabla_\theta P + \text{Nres.} . \tag{4.2}$$

Nonconservative tendencies from the parametrization schemes are not archived for ensemble forecasts so the direct PV modification by nonconservative processes due to diabatic heating and nonconservative momentum change, which is given by

$$N = \dot{\theta}\frac{\partial P}{\partial\theta} + P\frac{\partial\dot{\theta}}{\partial\theta} - g\frac{\partial\theta}{\partial p}\left(\mathbf{k}\times\frac{\partial\mathbf{v}}{\partial\theta}\right)\cdot\nabla\dot{\theta} - g\frac{\partial\theta}{\partial p}\mathbf{k}\cdot\left(\nabla\times\dot{\mathbf{v}}\right) , \tag{4.3}$$

is here interpreted as part of the residual Nres.[3]. The residual also includes the influence of other processes that cannot be quantified even if the tendencies from all parameterization schemes were available, such as numerical diffusion, analysis increments due to data assimilation, and numerical inaccuracies due to the discretization and interpolation of data. Near the tropopause, diabatic processes can play an important role for Rossby wave dynamics (e.g., Chagnon et al., 2013; Martínez-Alvarado et al., 2014). Relative to the advective tendencies, however, the direct diabatic PV changes are only of second-order importance (Teubler and Riemer, 2016). The impact of latent heat release is arguably most prominently communicated to the tropopause region by the associated upper-tropospheric divergent outflow (e.g., Davis et al., 1996; Teubler and Riemer, 2016) and furthermore by enhancing baroclinic coupling (e.g., Gutowski et al., 1992). Both processes are included in the advective tendencies analyzed here. In addition, a case study of error growth in an operational ECMWF forecast including nonconservative tendencies from the parameterization schemes (Baumgart et al., 2018) clearly demonstrated that these nonconservative tendencies are negligible, even at very short lead times. For the operational ensemble forecast investigated here, neglecting direct

[3]For convenience, we will denote the residual in all tendency equations as Nres., although the mathematical forms of the terms in the respective tendency equation differ.

diabatic tendencies is therefore not expected to affect the analysis.
The PV perspective has proven very helpful to gain insight into the dynamics of forecast errors (e.g., Snyder et al., 2003; Davies and Didone, 2013; Baumgart et al., 2018). Based on the difference between the PV tendency in the forecast and the analysis, Baumgart et al. (2018) derived a tendency equation for error (potential) enstrophy

$$\frac{\partial}{\partial t}\frac{P^{*2}}{2} = -P^*\mathbf{v}^* \cdot \nabla_\theta P - \nabla_\theta \cdot \left(\frac{P^{*2}}{2}(\mathbf{v}^* + \mathbf{v})\right) + \frac{P^{*2}}{2}\nabla_\theta \cdot (\mathbf{v}^* + \mathbf{v}) + \text{Nres.}\,, \tag{4.4}$$

where variables with index * denote error fields (forecast – analysis) and variables without index denote analysis fields. This equation will be used in Sec. 4.4 to quantify the differences between the error growth in best and worst member, respectively.
Deriving a similar equation for the amplification of forecast uncertainty in ensemble forecasts is a natural extension of our previous work. For that purpose, we use the ensemble variance in PV as our metric for forecast uncertainty

$$V_P := \frac{1}{n-1}\sum_{i=1}^{n}(P_i - \overline{P})^2\ , \tag{4.5}$$

where n describes the number of ensemble members (50 perturbed forecasts and 1 control forecast in the ECMWF system) and $\overline{P}$ the ensemble-mean PV.
The local change of the PV variance can be written as

$$\frac{\partial V_P}{\partial t} = \frac{2}{n-1}\left(\sum_{i=1}^{n}(P_i - \overline{P})\frac{\partial}{\partial t}(P_i - \overline{P})\right) = \frac{2}{n-1}\left(\left(\sum_{i=1}^{n}P_i\frac{\partial P_i}{\partial t}\right) - n\overline{P}\frac{\partial \overline{P}}{\partial t}\right)\ . \tag{4.6}$$

For the second step in Eq. 4.6, we made use of the fact that the terms $\sum_{i=1}^{n}\overline{P}\frac{\partial \overline{P}}{\partial t}$, $\sum_{i=1}^{n}\overline{P}\frac{\partial P_i}{\partial t}$, and $\sum_{i=1}^{n}P_i\frac{\partial \overline{P}}{\partial t}$ can be written as $n\overline{P}\frac{\partial \overline{P}}{\partial t}$.
To gain further insight into the variance amplification, we insert the PV tendency of the individual ensemble members $\left(\frac{\partial P_i}{\partial t} = -\mathbf{v_i}\cdot\nabla P_i + \text{Nres}_i\right)$ and of the ensemble mean $\left(\frac{\partial \overline{P}}{\partial t} = -\frac{1}{n}\sum_{i=1}^{n}(\mathbf{v_i}\cdot\nabla P_i + \text{Nres}_i)\right)$ into Eq. 4.6. By expressing P_i as the sum of the ensemble mean value and the perturbation thereof (i.e., $P_i = \overline{P} + \delta P_i$ with δ denoting the perturbation from the ensemble mean) and rearranging and combining terms, Eq. 4.6 can be written as

$$\frac{\partial V_P}{\partial t} = \frac{2}{n-1}\left[-\sum_{i=1}^{n}\delta P_i\mathbf{v_i}\cdot\nabla\overline{P} - \sum_{i=1}^{n}\mathbf{v_i}\cdot\nabla\frac{(\delta P_i)^2}{2}\right] + \text{Nres.}\ . \tag{4.7}$$

Using $\mathbf{v_i} = \overline{\mathbf{v}} + \boldsymbol{\delta}\mathbf{v_i}$ for the first term on the right-hand side together with noting that $-\sum_{i=1}^{n}\delta P_i\overline{\mathbf{v}}\cdot\nabla\overline{P} = -\overline{\mathbf{v}}\cdot\nabla\overline{P}\sum_{i=1}^{n}\delta P_i = 0$ and using the chain rule for the second term

finally leads to

$$\frac{\partial V_P}{\partial t} = \frac{2}{n-1}\Bigg[\underbrace{-\sum_{i=1}^{n} \delta P_i \boldsymbol{\delta}\mathbf{v_i} \cdot \nabla \overline{P}}_{\text{nonlinear production term}} \underbrace{-\sum_{i=1}^{n} \nabla \cdot \left(\frac{(\delta P_i)^2}{2} \mathbf{v_i} \right)}_{\text{flux term}} \underbrace{+\sum_{i=1}^{n} \frac{(\delta P_i)^2}{2} \nabla \cdot \mathbf{v_i}}_{\text{area change term}} \Bigg] + \text{Nres.} \, . \tag{4.8}$$

In analogy to the tendency equation for error enstrophy (Eq. 4.4), the first term on the right-hand side of Eq. 4.8 can be interpreted as a nonlinear production term. The flux term (second term in Eq. 4.8) merely redistributes variance and does not contribute to a global variance amplification. The third term in Eq. 4.8 is associated with an area change of PV variance due to the divergent flow. This term leads to variance amplification (decay) when the quasi-horizontal flow is divergent (convergent).

By using $\mathbf{v_i} = \overline{\mathbf{v}} + \boldsymbol{\delta}\mathbf{v_i}$, the flux and area-change term can be combined to $-\overline{\mathbf{v}} \cdot \nabla V_p - \frac{1}{n-1} \sum_{i=1}^{n} \boldsymbol{\delta}\mathbf{v_i} \cdot \nabla (\delta P_i)^2$. Rodwell et al. (2018) used a similar form of an equation for ensemble spread, but using standard deviation instead of variance. They derived a "material" derivative for ensemble spread following the ensemble mean flow, which would correspond to the term $-\overline{\mathbf{v}} \cdot \nabla V_p$ in our diagnostic. For this study, we decided to use, instead, the local derivative of PV variance as indicated in Eq. 4.8 to have an exact budget equation for PV variance. The use of Eq. 4.8 has also the advantage that the second term can be interpreted as a boundary term when spatially integrating over a specific domain.

To gain further insight into the processes governing the variance amplification, we use the same partitioning of processes as in the previous work by Teubler and Riemer (2016) and Baumgart et al. (2018, 2019). This partitioning is based on the PV perspective of midlatitude dynamics (Hoskins et al., 1985). From this perspective, the evolution of PV anomalies near the tropopause can be described by advective tendencies associated with 1) upper-level (near-tropopause) dynamics, 2) mid-tropospheric PV anomalies and potential temperature anomalies just above the boundary layer, and 3) upper-tropospheric divergent flow. The influence of upper-level (near-tropopause) PV anomalies on the upper-level (near-tropopause) evolution describes the influence of nonlinear Rossby wave dynamics (Hoskins et al., 1985) and will here be referred to as the contribution from near-tropopause dynamics (index nTP). The influence of lower-level anomalies on the upper-level evolution describes the influence of vertical interaction, including baroclinic instability (Eady, 1949; Hoskins et al., 1985; Heifetz et al., 2004), and will here be referred to as tropospheric-deep interaction (index TPd). Upper-tropospheric divergence (index div) can be associated with dry balanced dynamics and diabatic processes (see, e.g., Chapter 6.4 in Holton and Hakim, 2013) and is of particular importance during ridge building (Grams et al., 2011; Teubler and Riemer, 2016). Pronounced upper-tropospheric divergence is often associated with latent heat release below (e.g. Davis et al., 1993; Riemer et al., 2014; Quinting and Jones, 2016) and can usually be expected of larger importance to Rossby wave dynamics than direct diabatic PV modification (e.g., Davis et al., 1993; Riemer and Jones, 2010; Teubler and Riemer, 2016).

The technicalities of the flow partitioning are also the same as in Teubler and Riemer (2016)

and Baumgart et al. (2018, 2019): We use a Helmholtz partitioning to separate the divergent flow from the nondivergent flow following Lynch (1989). The nondivergent flow is further partitioned into those parts associated with upper- and lower-level PV anomalies, respectively, using piecewise PV inversion (PPVI) under nonlinear balance (Charney, 1955) following Davis and Emanuel (1991) and Davis (1992). PPVI is performed on the Northern Hemisphere from 25–85°N and 850–150 hPa. Potential temperature anomalies at 875 and 125 hPa serve as vertical boundary conditions for the inversion. Anomalies are defined as deviations from a background state, which is here defined as the 30 day-temporal mean centered on September 23, 2016 00 UTC of the analysis. A mid-tropospheric pressure level (600 hPa) is used as the separation level between upper- and lower-level anomalies. The flow partitioning yields an uncertainty, $\mathbf{v}_{\text{unc}}$, due to the harmonic flow component, the uncertainty in the horizontal boundary conditions, and nonlinearities of the the piecewise PV inversion. This uncertainty is calculated as the difference between the wind field in the ECMWF data and the sum of the near-tropopause, tropospheric-deep and divergent wind field. It is, in general, small and does not affect the physical interpretation of the results.
In summary, our flow partitioning yields

$$\mathbf{v} = \mathbf{v}_{\text{nTP}} + \mathbf{v}_{\text{TPd}} + \mathbf{v}_{\text{div}} + \mathbf{v}_{\text{unc}} \,. \tag{4.9}$$

The flow partitioning is performed separately for each individual ensemble member and then inserted into Eq. 4.8, yielding

$$\frac{\partial V_P}{\partial t} = \frac{2}{n-1}\Bigg[\underbrace{-\sum_{i=1}^{n} \delta P_i (\boldsymbol{\delta}\mathbf{v}_{\text{i, nTP}} + \boldsymbol{\delta}\mathbf{v}_{\text{i, TPd}} + \boldsymbol{\delta}\mathbf{v}_{\text{i, div}} + \boldsymbol{\delta}\mathbf{v}_{\text{i, unc}}) \cdot \nabla \overline{P}}_{\text{nonlinear production term}} \underbrace{-\sum_{i=1}^{n} \nabla \cdot \left(\mathbf{v}_{\text{i}} \frac{(\delta P_i)^2}{2} \right)}_{\text{flux term}} + \underbrace{\sum_{i=1}^{n} \frac{(\delta P_i)^2}{2} \nabla \cdot \mathbf{v}_{\text{i, div}}}_{\text{area change term}} \Bigg] + \text{Nres.} \,. \tag{4.10}$$

Note that the divergent wind contributes to both the nonlinear production term and to the area change term. These two contributions will be considered together as the divergent contribution in our discussion below. For a quantitative view on the relative importance of the individual processes, we spatially integrate Eq. 4.10

$$\int_A \frac{\partial V_P}{\partial t} dA = \frac{2}{n-1}\Bigg[-\int_A \Big(\sum_{i=1}^{n} \delta P_i (\boldsymbol{\delta}\mathbf{v}_{\text{i, nTP}} + \boldsymbol{\delta}\mathbf{v}_{\text{i, TPd}} + \boldsymbol{\delta}\mathbf{v}_{\text{i, div}}) \cdot \nabla \overline{P} \Big) dA + \int_A \Big(\sum_{i=1}^{n} \frac{(\delta P_i)^2}{2} \nabla \cdot \mathbf{v}_{\text{i, div}} \Big) dA \Bigg] + \text{bnd.} + \text{unc.} + \text{Nres.} \,, \tag{4.11}$$

where $dA = a^2 cos\phi d\lambda d\phi$ is the area element in spherical coordinates with Earth radius a, longitude λ, and latitude ϕ. The uncertainty term unc. $= -\frac{2}{n-1} \int_A \big(\sum_{i=1}^{n} \delta P_i \boldsymbol{\delta}\mathbf{v}_{\text{i, unc}} \cdot \nabla \overline{P} \big) dA$ describes the uncertainty of the flow partitioning, while the boundary term bnd.

describes the contribution from the flux term, bnd. $= \frac{2}{n-1}\int_A\big(-\sum_{i=1}^{n}\nabla\cdot(\mathbf{v}_\mathrm{i}\frac{(\delta P_i)^2}{2})\big)dA = \frac{2}{n-1}\oint_S\big(-\sum_{i=1}^{n}\frac{(\delta P_i)^2}{2}\mathbf{v}_\mathrm{i}\cdot\mathbf{n}\big)dS$, where dS denotes the boundary of the area and $\mathbf{n}$ the normal vector of the boundary pointing outward. The variance change observed between consecutive time steps acts as an indication for the representativeness of our diagnostic and is calculated by centered differences:[4]

$$\text{obs.} = \int_A \frac{\partial V_P}{\partial t}dA \approx \frac{\int_A V_p(t+\Delta t)dA - \int_A V_p(t-\Delta t)dA}{2\Delta t} \tag{4.12}$$

with Δt being 3 h in the first 3 forecast days and 6 h afterwards. Eq. 4.11 thus provides a novel diagnostic to quantify the relative importance of near-tropopause dynamics, tropospheric-deep interaction, and upper-tropospheric divergence to the amplification of forecast uncertainty.

4.3 Quantitative view on the variance amplification

4.3.1 Synoptic overview and variance evolution

Before discussing the variance amplification in more detail, we provide a synoptic overview of our case together with a description of the variance evolution (Fig. 4.2). The synoptic evolution of our case is characterized by a large-amplitude Rossby wave pattern spanning from (counter-clockwise) 180° to 60°E (as seen by the blue and black contour denoting the 2 PVU-surface of the analysis and the ensemble mean, respectively, in Fig. 4.2). Most interesting for the Rossby wave evolution is the evolution of several ridges (labeled R1–R3 in Fig. 4.2). At 2 days, a large-amplitude ridge exists around 150–75°W (label R1). Its upstream trough interacts with a cutoff around 110°W (2–3 days, Fig. 4.2a and b) and completely reabsorbs this cutoff shortly after 3 days (Fig. 4.2b). Ridge R1 is characterized by a large extension in the meridional direction and a contraction in the zonal direction and thereby exhibits a highly nonlinear evolution (Fig. 4.2c and d). Another ridge exists around 20°W–10°E (label R3), which is characterized by ridge building between 2–3 days (Fig. 4.2a and b). This ridge building was associated with the development of cyclone "Vladiana" (labeled V in Fig. 4.2), which was another observational highlight during the NAWDEX period due to the occurrence of pronounced warm-conveyor-belt ascent (Schäfler et al., 2018; Oertel et al., 2019). In the following, ridge R3 is characterized by a large amplitude and a similar nonlinear evolution as ridge R1, albeit its spatial extent is smaller than that of ridge R1 (Fig. 4.2c and d). In between these larger-amplitude ridges, a smaller-amplitude ridge exists (label R2, Fig. 4.2). Between 4–5 days, this ridge is influenced by the interaction with Ex-Karl around 40°W (labeled K in Fig. 4.2), which leads to cyclonic wave-breaking of the ridge (as seen by the PV wrap up of the analysis ridge around 4.5 days, blue contour in Fig. 4.2d).

The ensemble variance of PV, which we use as our metric for forecast uncertainty, is not homogeneously distributed over the hemisphere (colored shading in Fig. 4.2), manifest of the well-known flow dependence of forecast uncertainty (e.g., Palmer and Hagedorn, 2006; Rod-

[4]For the first and last time step, we use instead forward and backward differences, respectively.

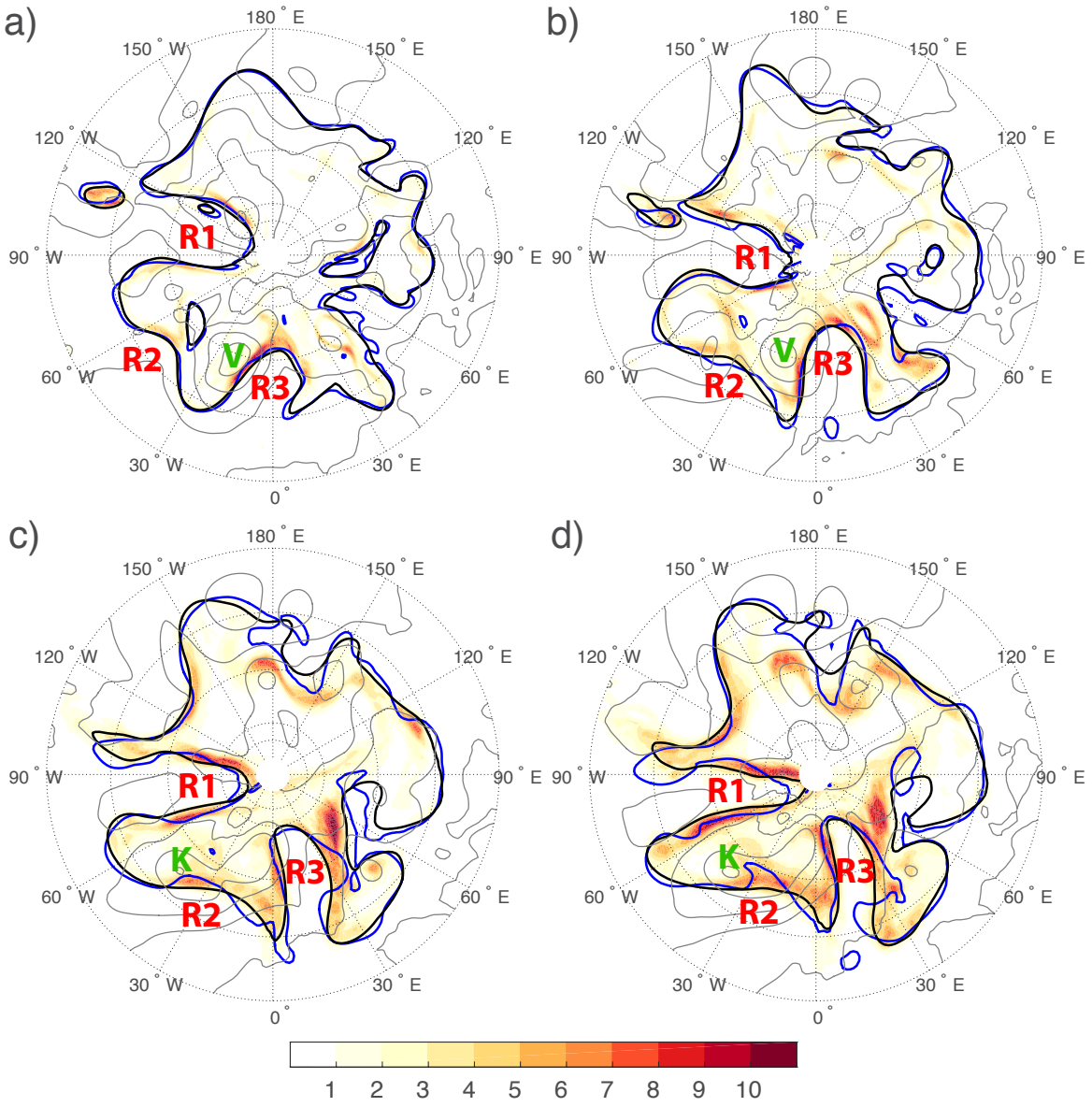

Figure 4.2: PV variance (colored shading in PVU^2) on 325 K at a) 2 days, b) 3 days, c) 4 days, and d) 4.5 days. The black and blue contour denote the 2 PVU contour (smoothed over a box of 5x5 grid points using a mean filter) of the ensemble mean and the analysis, respectively. Gray contours denote the ensemble mean of mean sea level pressure every 10 hPa (smoothed over a box of 3x3 grid points using a mean filter). Labels refer to individual ridges (prefix R) and the cyclones Vladiana (label V) and Ex-Karl (label K).

well et al., 2018). PV variance maximizes, in general, along the dynamical tropopause and exhibits a larger amplitude in the region of the Rossby wave pattern (counter-clockwise from 180°–60°E) than in the more zonally-oriented region of the tropopause. In the time range investigated here, variance amplification occurs both in amplitude and in scale. Several local maxima of variance amplification can be identified: i) the cutoff evolution and its reabsorption by the waveguide around 2–3 days, ii) ridge-building events in association with cyclone development, e.g., the ridge building of ridge R3 around 2–3 days, iii) highly nonlinear regimes of the wave pattern, e.g., the large-amplitude ridge R1, and iv) the interaction between Ex-Karl and ridge R2 around 4–5 days.

4.3.2 Individual contributions to variance amplification: spatial illustration

The previous subsection revealed several local maxima of variance amplification. To spatially illustrate the mechanisms that govern this variance amplification, we partition the variance tendency into the contributions of the individual processes as detailed in Sec. 4.2. Before looking at these individual contributions, the representativeness of our diagnostic for the actual variance amplification is assessed by comparing the advective variance tendency (right-hand side of Eq. 4.8) with the observed variance tendency (left-hand side of Eq. 4.8 approximated by centered differences) at two forecast lead times (2 and 4.5 days, Fig. 4.3). The observed variance tendency is characterized by dipole patterns along the tropopause that are associated with an eastward displacement of PV variance. This displacement is consistent with the eastward phase propagation of the Rossby wave anomalies, in which the maxima of PV variance are located. In most of the dipole patterns, the positive part is larger than the negative part, leading to an overall amplification of PV variance. The main patterns of the observed variance tendency are captured well by the advective tendency both in terms of the variance displacement and the overall variance amplification. The magnitude of the observed variance tendency, however, is smaller than that of the advective tendency, in particular at 4.5 days. One reason for the smaller magnitude might be that the observed variance change has to be approximated by centered differences of 3-h data for lead times smaller than 3 days and 6-h data afterwards, which yields a smoothing and thereby a reduction of maxima and minima of the observed tendency. There exists thus a physically meaningful explanation for the differences between the observed and advective tendency.

To remove the dipole patterns that are only associated with a displacement of variance, but not with a net amplification of variance, we exclude the flux term of Eq. 4.8 from our investigation as it is only associated with a redistribution of PV variance. The main pattern of this net advective tendency (Figs. 4.4a and 4.5a) is no longer characterized by dipole patterns as it was the case for the full advective tendency (Fig. 4.3b and d). Regions of variance amplification are thus much easier to identify. For the partitioning into the contributions from individual processes (Eq. 4.10), we will thus only discuss the net advective tendency and drop the prefix “net” for brevity.

At 2 days (Fig. 4.4), a large amplitude of the advective tendency is, in general, found within

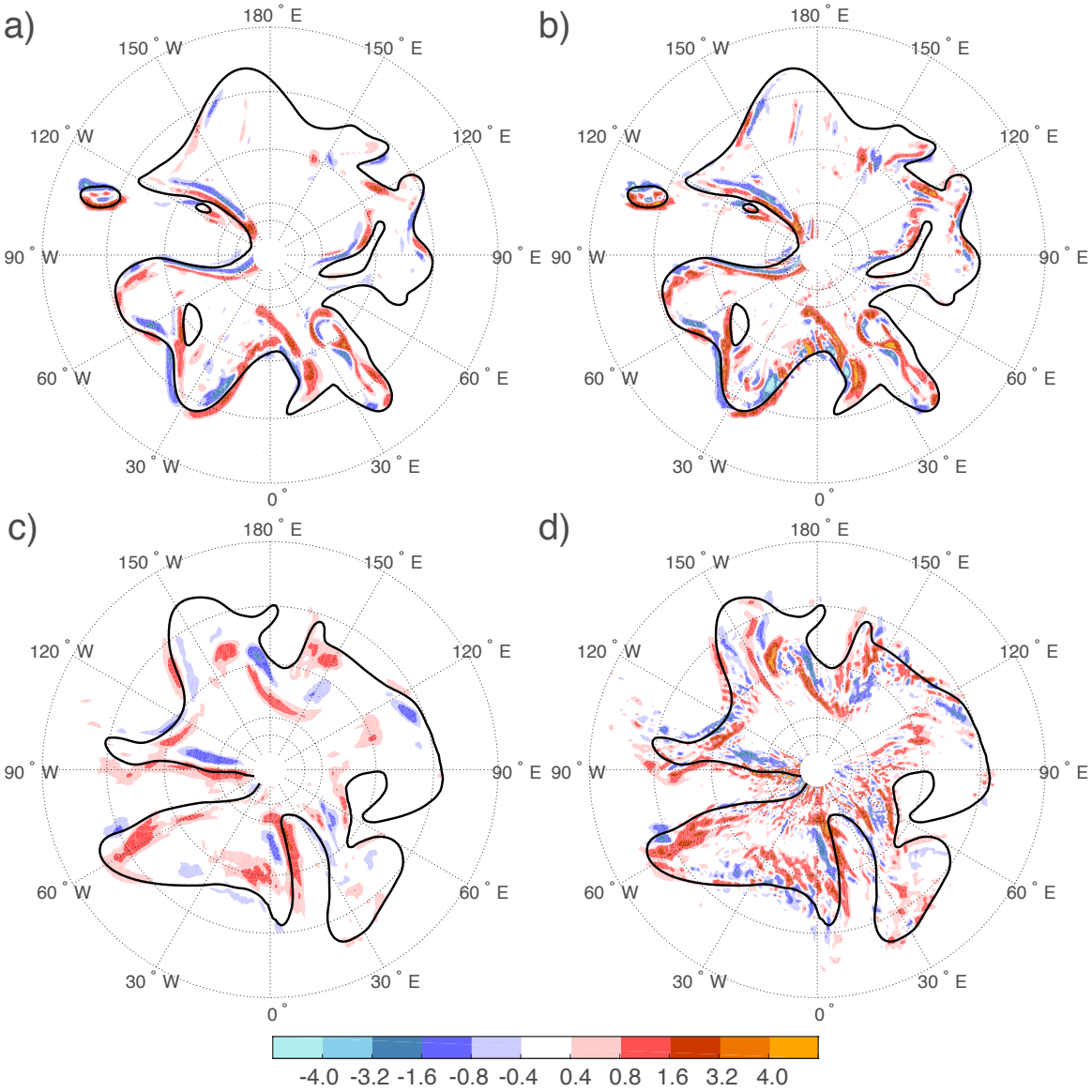

Figure 4.3: Comparison between the observed variance tendency (left-hand side of Eq. 4.8 approximated by centered differences in a and c) and the advective variance tendency (right-hand side of Eq. 4.8 in b and d) on 325 K at 2 days (a and b) and 4.5 days (c and d). The colored shading shows the respective variance tendency in 10^{-4} PVU^2 s^{-1} and the black contour denotes the 2 PVU contour (smoothed over a box of 5x5 grid points using a mean filter) of the ensemble mean.

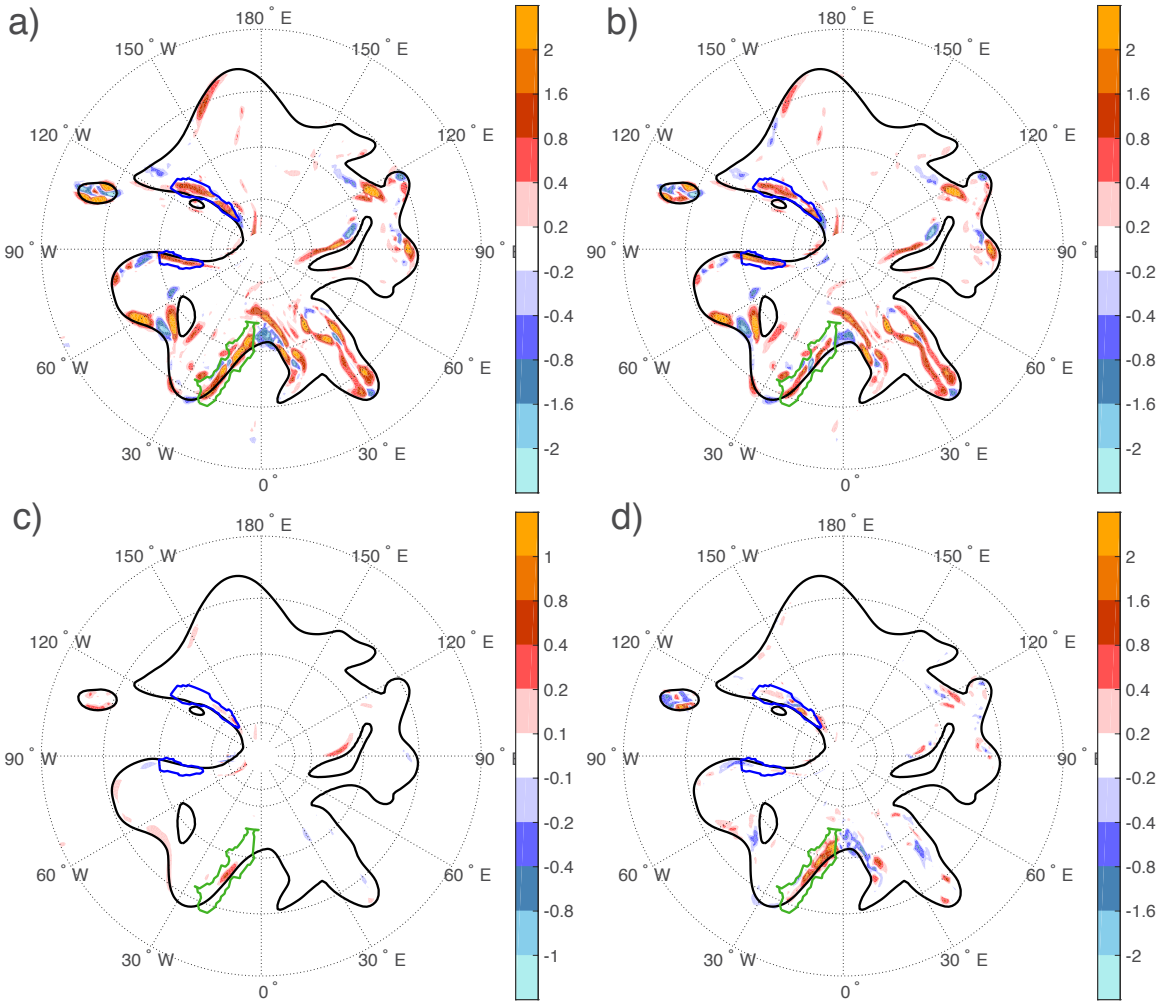

Figure 4.4: Variance tendency (colored shading in 10^{-4} PVU^2 s^{-1}) on 325 K at 2 days by a) full wind, b) near-tropopause wind, c) tropospheric-deep wind, and d) divergent wind. The black contour denotes the 2 PVU contour (smoothed over a box of 5x5 grid points using a mean filter) of the ensemble mean. The blue and green contours indicate the regions used for the localized perspective on the variance amplification (Fig. 4.7a/b).

the Rossby wave pattern around 180°–60°E (counter-clockwise) with a particular large variance amplification on the western flanks of the ridges R1 and R3 around 130°W and 20°W, respectively, and in the trough around 40°E (Fig. 4.4a). The individual contributions to this tendency are shown in Fig. 4.4b–d. The advective tendency is mostly dominated by the near-tropopause tendency (Fig. 4.4b). This tendency is particularly large where the Rossby wave pattern exhibits a large amplitude as, e.g, in ridge R1 around 120°W or in the trough around 40°E. One region, in which not only the near-tropopause tendency makes a dominating contribution to the variance amplification, is ridge R3 around 20°W. This ridge is characterized by ridge building in association with cyclone development (Fig. 4.2a and b) and its variance amplification is largely governed by both the near-tropopause and divergent tendency (Fig. 4.4b and d). Compared to the near-tropopause and divergent tendency, the

tropospheric-deep tendency is much smaller in amplitude (Fig. 4.4c, note the different scale of the associated colorbar).
At 4.5 days (Fig. 4.5), the variance tendency has become larger in scale than at 2 days and now exhibits a large amplitude almost everywhere along the dynamical tropopause. The large-amplitude ridge R1 (120–60°W) is still associated with large variance amplification (Fig. 4.5a). Another region that exhibits large variance amplification is the smaller-amplitude ridge R2 where Ex-Karl interacts with the Rossby wave pattern around 40°W. The processes that govern the variance amplification in the two ridges are distinct: While the near-tropopause tendency dominates the variance amplification in the large-amplitude ridge R1, it cannot explain the variance amplification in the smaller-amplitude ridge R2 (Fig. 4.5b). In this ridge, the divergent and tropospheric-deep tendency make dominating contributions to the variance amplification (Fig. 4.5c and d), which suggests that the variance amplification in this ridge is mostly related to the (moist baroclinic) cyclone development of Ex-Karl. A locally dominating contribution from the tropospheric-deep and divergent tendency is also found in the smaller-amplitude ridge around 130°W. The main pattern of the advective tendency, however, is dominated by the near-tropopause tendency as at 2 days. Interestingly, the variance tendency in ridge R3 (around 10–40°W), which was dominated by the divergent tendency at 2 days (Fig. 4.4), is now dominated by the near-tropopause tendency, indicating a multi-stage behaviour of variance amplification as it was observed for upscale error growth (Baumgart et al., 2019).
From the previous discussion it is evident that the variance tendency is in most regions dominated by the near-tropopause tendency, but that there are also localized regions, in which other processes dominate the variance tendency. In the following two subsections, we will thus quantify the variance amplification both from a hemispheric-averaged perspective for the midlatitudes of the Northern Hemisphere and from a local perspective for contrasting regions exhibiting i) large near-tropopause variance amplification and ii) large divergent and/or tropospheric-deep variance amplification.

4.3.3 Hemispheric variance amplification

For a quantitative perspective on the hemispheric-averaged variance amplification, Fig. 4.6 shows the spatially-averaged variance amplification for the midlatitudes of the Northern Hemisphere (30–80°N) and its partitioning into the individual processes (Eq. 4.11) as a function of lead time. The observed variance tendency (as approximated by centered differences, Eq. 4.12) is mostly increasing with lead time until around 4.5 days and then mostly decreasing until it is almost zero after around 8 days. A particular large growth of the observed tendency is found in the first 1.5 days and between 3.5 and 4.5 days. Between these two periods (day 1.5–3.5), the observed tendency is almost constant. The initially large variance amplification might be related to the singular-vector approach at ECMWF that maximizes the growth of ensemble perturbations in the first 2 days, whereas the later nonhomogeneous growth of the observed tendency points to a flow dependence of the variance amplification with a period of lower

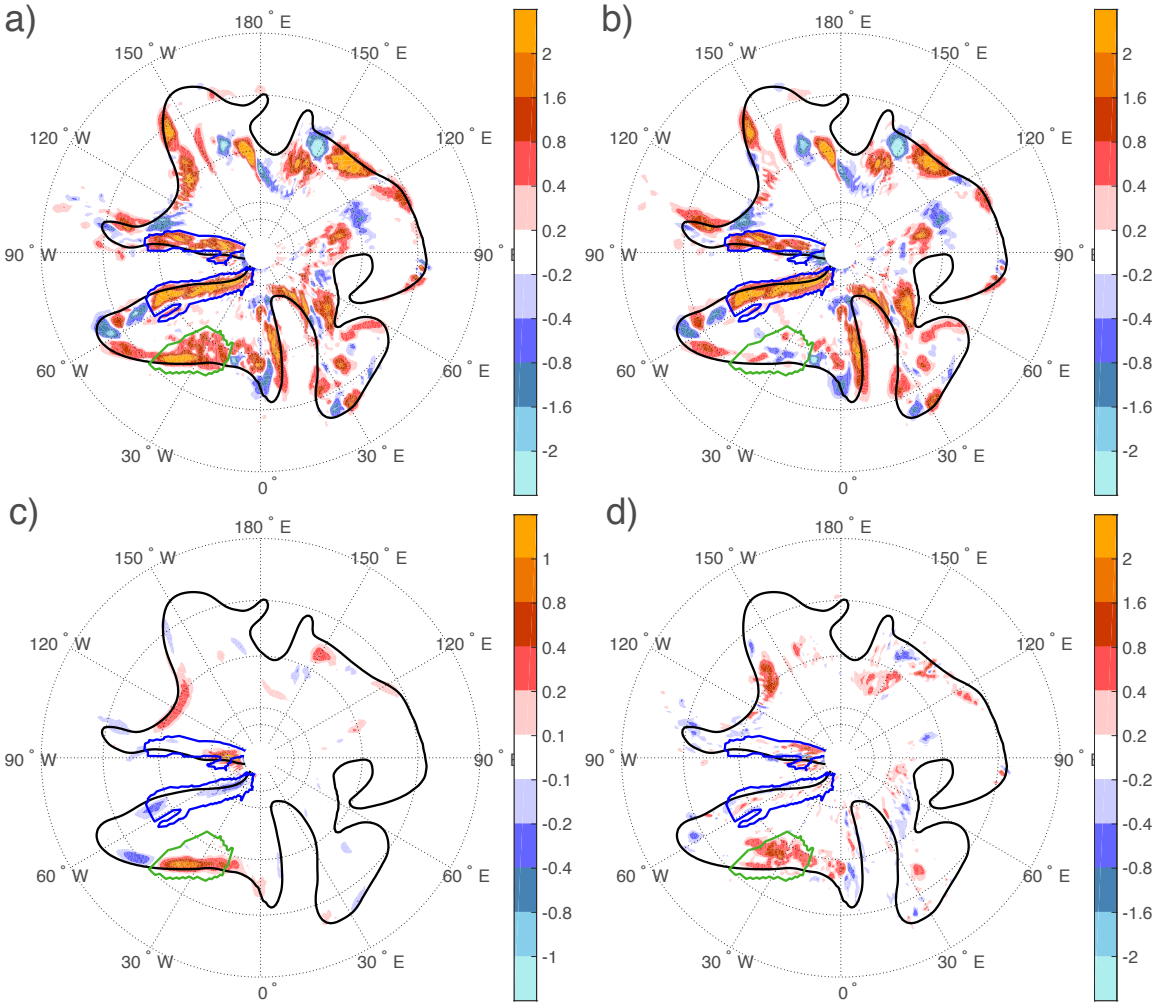

Figure 4.5: Variance tendency (colored shading in 10^{-4} PVU^2 s^{-1}) on 325 K at 4.5 days by a) full wind, b) near-tropopause wind, c) tropospheric-deep wind, and d) divergent wind. The black contour denotes the 2 PVU contour (smoothed over a box of 5x5 grid points using a mean filter) of the ensemble mean. The blue and green contours indicate the regions used for the localized perspective on the variance amplification (Fig. 4.7c/d).

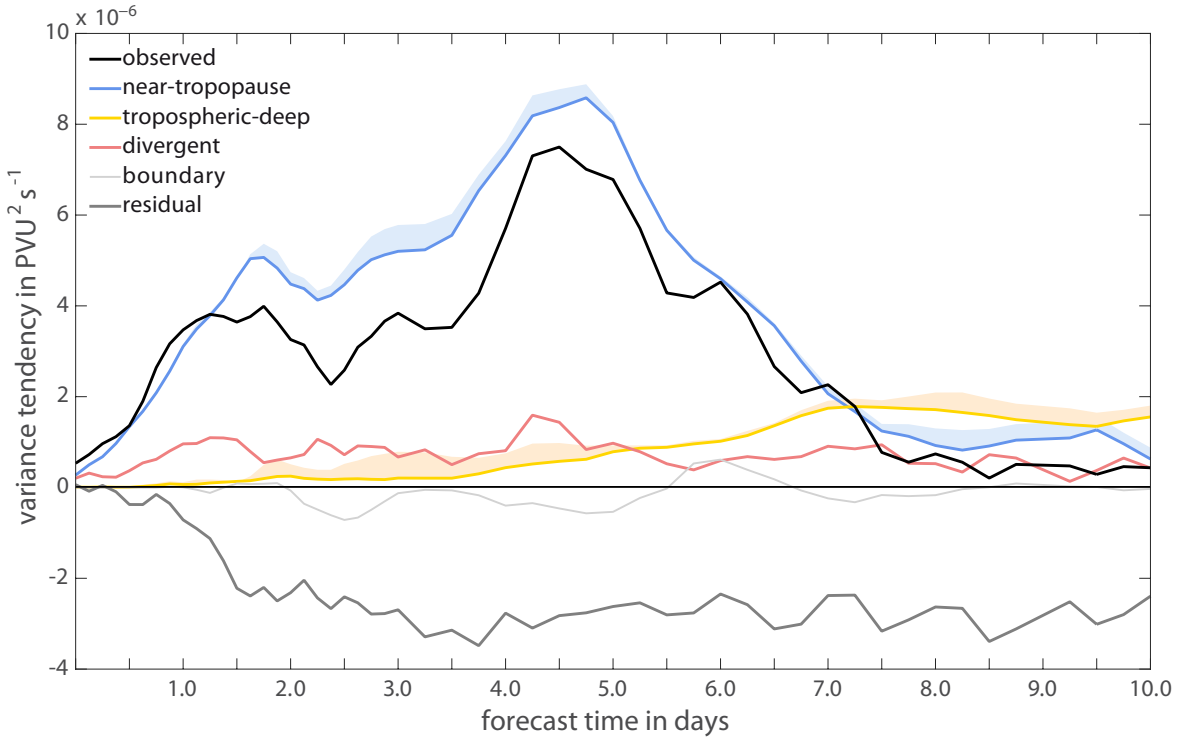

Figure 4.6: Individual contributions (colored lines) to the variance tendency (spatially-averaged over the midlatitudes of the Northern Hemisphere, 30–80°N) on the 325-K surface as indicated by Eq. 4.11 as a function of forecast time. The observed tendency is approximated by centered differences in time of the spatially-averaged variance (obs. in Eq. 4.12). The shading next to the near-tropopause and tropospheric-deep tendencies indicates the uncertainty of the result due to the uncertainty of PPVI (unc. in Eq. 4.11).

predictability and thus larger variance amplification between 3.5 and 4.5 days.

Our diagnostic yields a residual, which increases in magnitude in the first 2 days and is then almost constant at a value of around -3×10^{-6} PVU^2 s^{-1}. This residual is related to those processes that cannot be measured with the available data (see Sec. 4.2). Based on the results of Baumgart et al. (2018), dissipation can be expected to make a dominant negative contribution. Dissipation yields a smoothing of small-scale PV features in association with the downscale cascade of enstrophy. In this way, dissipation also leads to a smoothing of the PV variance associated with small-scale PV features and thereby provides a variance sink that explains the negative residual observed in Fig. 4.6. This important role of dissipation was also discussed in further previous studies (e.g., Zhang et al., 2007; Saffin et al., 2016, 2017; Baumgart et al., 2019).

The near-tropopause tendency clearly dominates the hemispheric variance amplification. This tendency makes the by far largest contribution to the observed variance tendency in the first 7 days and its general time evolution corresponds well with that of the observed variance tendency. In the first two days, the near-tropopause tendency increases almost linearly with lead time, suggesting that the singular-vector approach, which maximizes the perturbation growth in the first 2 days, projects most strongly onto the near-tropopause tendency.

The divergent tendency is almost as large as the near-tropopause tendency at the first time step, but much smaller afterwards. This small importance suggests that the divergent variance amplification observed in the spatial maps (Fig. 4.4d and 4.5d) is only important in a localized sense, but not in a hemispheric sense. Interestingly, the divergent tendency shows a small peak around 4–4.5 days, which corresponds to the time when Ex-Karl interacts with the waveguide.
The tropospheric-deep tendency is small in the first 4 days. Afterwards, it increases slowly and becomes larger than the near-tropopause tendency after 7 days when the near-tropopause tendency is small. This slow increase of the tropospheric-deep tendency suggests that the tropospheric-deep variance amplification depends on the horizontal scale of the potential-temperature uncertainty at the lower boundary (875 hPa), which has to become large enough in scale to vertically penetrate up to the tropopause region. Such a scale dependence would be consistent with the scale dependence found for error growth in the Eady (1949) model (Baumgart et al., 2018).

4.3.4 Variance amplification in localized regions

The hemispheric variance amplification investigated in the previous subsection was dominated by near-tropopause dynamics (Fig. 4.6). In the spatial maps (Figs. 4.4 and 4.5), we found, however, that in localized regions the variance amplification can be dominated by other processes. In the following, we will thus consider the two time steps used for the spatial maps (2 and 4.5 days) and quantify the variance amplification for two localized regions at each time step that exhibit different characteristics in terms of the processes governing the variance amplification. The localized regions of variance amplification are chosen by a threshold of the smoothed advective tendency ($> 0.2 \times 10^{-4}$ PVU^2 s^{-1}, smoothing is performed over a box of 5x5 grid points using a mean filter). These integration regions are indicated by the blue contours in Figs. 4.4 and 4.5.
At 2 days, there are two localized regions of variance amplification that exhibit different characteristics in terms of the processes, the large-amplitude ridge R1 around 150–75°W and the smaller-amplitude ridge R3 around 20–0°W (Fig. 4.4). From a quantitative perspective, the variance amplification in the large-amplitude ridge is dominated by the near-tropopause tendency, which contributes about 90% to the advective variance amplification (Fig. 4.7a). The variance amplification in the smaller-amplitude ridge, instead, is dominated by upper-tropospheric divergence, which contributes about two thirds to the advective variance amplification (Fig. 4.7b).
At 4.5 days, there are again two localized regions of variance amplification that exhibit different characteristics in terms of the processes, the large-amplitude ridge R1 around 120–60°W, which was also investigated at 2 days, and the smaller-amplitude ridge R2 around 45–15°W (Fig. 4.5). In the large-amplitude ridge, the largest contribution to the variance amplification is again given by the near-tropopause tendency (almost 100% of the full advective amplification, Fig. 4.7c). In the small-amplitude ridge, instead, the near-tropopause tendency makes

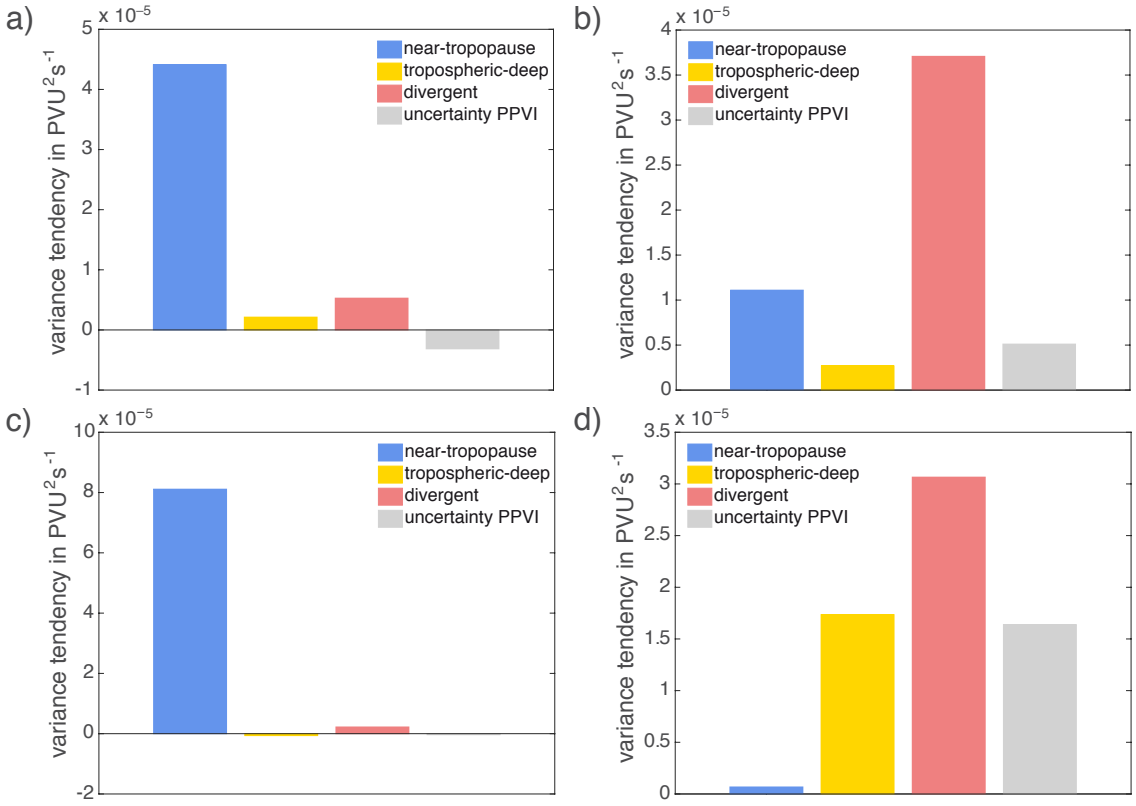

Figure 4.7: Individual contributions (colored bars) to the variance tendency on the 325-K surface. The variance tendency is averaged over localized regions of variance amplification: a) in the flanks of large-amplitude ridge R1 around 150–75°W at day 2 (integration area indicated by blue contours in Fig. 4.4), b) on the western flank of the smaller-amplitude ridge R3 around 20–0° at day 2 (green contour in Fig. 4.4), c) in the flanks of large-amplitude ridge R1 around 120–60°W at day 4.5 (blue contours in Fig. 4.5), and d) in the smaller-amplitude ridge R2 around 45–15°W at day 4.5 (green contour in Fig. 4.5).

only a very small contribution to the variance amplification (about 1% of the full advective amplification, Fig. 4.7d). In this ridge, the variance amplification is largely governed by the divergent tendency (almost 50% of the full advective amplification) and the tropospheric-deep tendency (about 25% of the full advective amplification).[5]

This comparison of localized variance amplification shows that there can be large flow-dependent differences in the processes that govern variance amplification. The large importance of near-tropopause dynamics in the large-amplitude ridge R1 at 2 and 4.5 days suggests that this contribution is particularly large when the flow is highly nonlinear. The large importance of the divergent and tropospheric-deep tendency in the smaller-amplitude ridges R2 and R3, in which ridge building occurs in association with cyclone development, instead, suggests that the variance amplification during the earlier phase of ridge building is

[5]Note that there is also a large uncertainty of the piecewise PV inversion (about 25% of the full advective amplification), which could be attributed to both the near-tropopause and the tropospheric-deep tendency. The general observation that the near-tropopause tendency does not make the largest contribution to the variance amplification, however, holds even when the whole uncertainty of the diagnostic is attributed to the near-tropopause tendency.

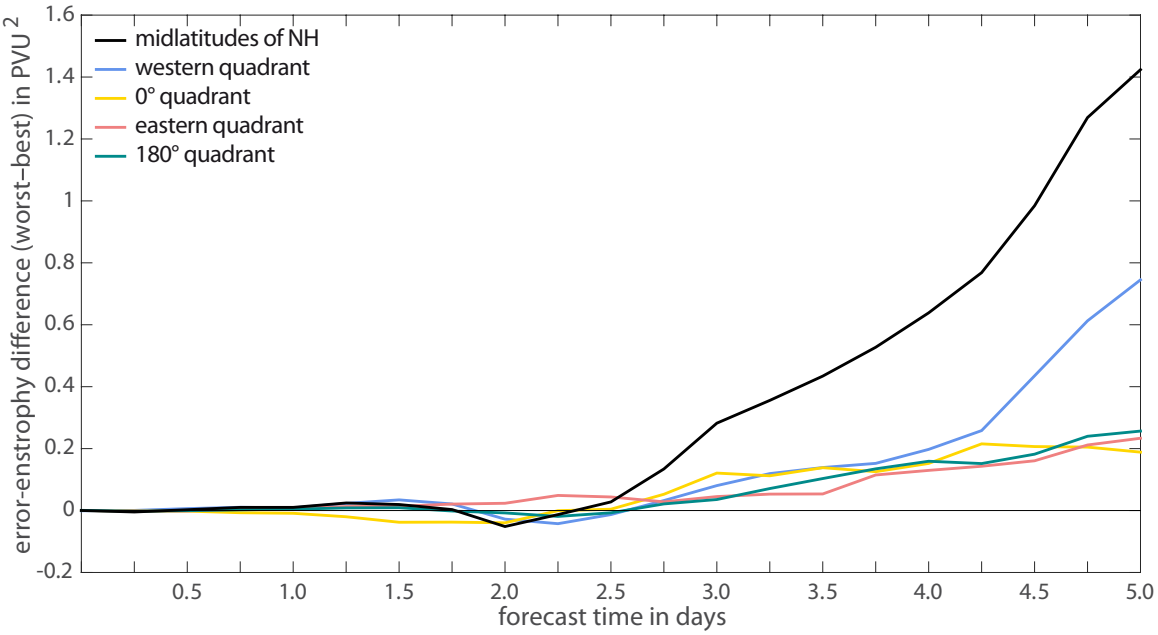

Figure 4.8: Difference between error enstrophy of the worst and best member on 325 K (spatially-averaged over the midlatitudes of the Northern Hemisphere, 30–80°N, black line) partitioned into the contribution from individual segments of the hemisphere with each segment spanning a longitude range of 90° (colored lines).

mainly related to uncertainties in the development of the respective cyclones both in terms of the associated warm conveyor belt and divergent outflow (large divergent contribution at 2 and 4.5 days) and in terms of the associated baroclinic growth (large tropospheric-deep contribution at 4.5 days).

4.4 Comparison between the error growth in the best and worst member

As noted in the introduction, two members are distinct from the rest of the ensemble members at 5 days in terms of error enstrophy being exceptionally small and large, respectively (Fig. 4.1). In addition to our analysis of variance amplification (Sec. 4.3), which provides a mean picture about the divergence of ensemble members, it is also interesting to analyze the divergence of these two individual members that are associated with the largest differences in error enstrophy of the ensemble. This section thus investigates in more detail the mechanisms leading to the pronounced differences in the error evolution of the two members.

The difference between the hemispheric error enstrophy of the worst and best members is shown in Fig. 4.8 (black line) as a function of forecast time. In the first 2.5 days, this difference is small, but increases prominently afterwards, until the two members become the best and worst member of the ensemble at 5 days (Fig. 4.1). To gain insight into the origin of this large difference in the hemispheric error enstrophy of the two members, the hemisphere

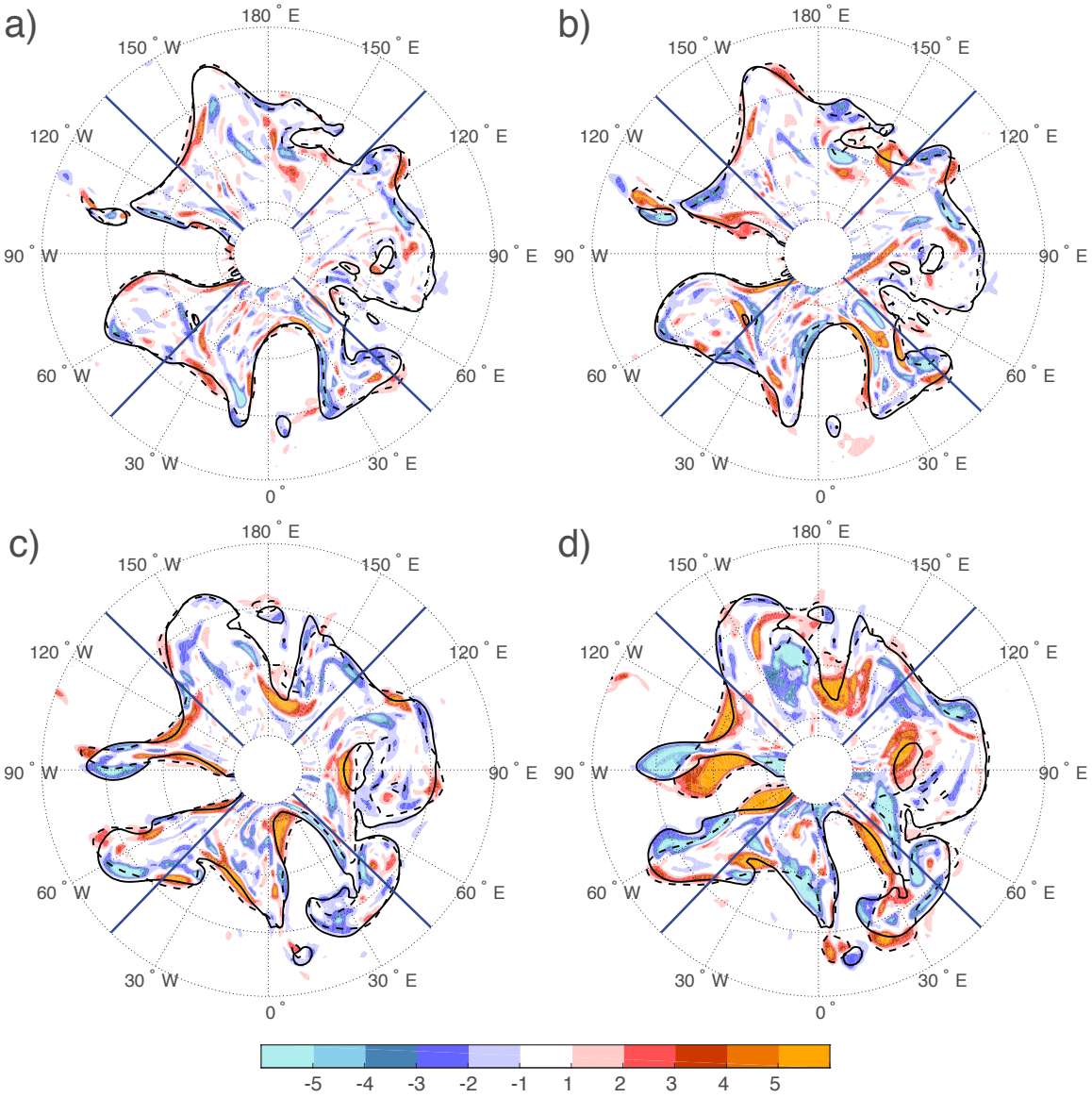

Figure 4.9: Comparison between the PV error (colored shading in PVU) in the best member (member 9 in a and c) and worst member (member 18 in b and d) on 325 K at 3 days (a and b) and 4.5 days (c and d). The solid (dashed) black contour denote the 2 PVU contour (smoothed over a box of 5x5 grid points using a mean filter) of the analysis (respective ensemble member). The blue lines indicate the regions used for the partitioning of error-enstrophy difference between the worst and best member in Fig. 4.8.

is split into four quadrants and the difference in error enstrophy is calculated separately for each quadrant (colored lines in Fig. 4.8). From this partitioning it is evident that all four quadrants contribute in a similar way to the hemispheric error-enstrophy difference between 2.5 and 4 days. Between 4–5 days, however, the western quadrant contributes much more strongly to the hemispheric error-enstrophy difference than the other three quadrants. Finally, this quadrant contributes more than 50% to the hemispheric error-enstrophy difference at 5 days, whereas the other three quadrants contribute individually only 13–18%.

Comparing spatial maps of the PV error of the best and worst member (Fig. 4.9) reveals that the large error-enstrophy difference in the western quadrant is mostly related to the highly nonlinear region of the wave pattern (120–60°W). At 4.5 days (Fig. 4.9c and d), this region is captured rather well by the best member, while the worst member exhibits large errors that are related to both the phase and the shape of the wave pattern. This difference in the error pattern can be traced back in time (a manual error tracking is applied as in Magnusson, 2017) to around 3 days when the cutoff interacts with the Rossby wave pattern (around 110°W, Fig. 4.9a and b). The worst member exhibits a distinct phase error in the trough that interacts with the cutoff, whereas the error of the best member is small in this region. This difference between the error pattern of the two members then amplifies largely in the highly nonlinear region of the wave pattern, leading to the pronounced differences in the error pattern around 4–5 days (Fig. 4.9c and d). The important role of a misrepresented cutoff evolution leading to large error growth in the further evolution was also discussed by Grams et al. (2018) for a forecast-bust event in March 2016.

For a quantitative and process-based perspective on the error-growth differences between the best and worst member, Fig. 4.10 compares the spatially-averaged error-enstrophy tendency (Eq. 4.4) in the midlatitudes of the Northern Hemisphere (30–80°N) for both members together with its partitioning into the contribution from individual processes. In the first 2.5 days, the error tendency of the best and worst member is rather similar in terms of the observed error-tendency magnitude, consistent with the fact that error enstrophy of the two members first diverges after 2.5 days (Figs. 4.1 and 4.8). During this time, the error growth of the two members is also similar in terms of the processes governing the error growth with the near-tropopause tendency dominating the error growth in both members.

Around 2.5 days, the best member shows a distinct drop in the observed error tendency with a magnitude around zero between 2.5–2.75 days. The error tendency in the worst member does not show such a distinct drop, but instead increases largely with forecast time, starting from a tendency of about 0.3×10^{-5} PVU^2 s^{-1} around day 2.5 and ending at a tendency of about 1.3×10^{-5} PVU^2 s^{-1} between 4–5 days. A particularly large error growth is found between 4–5 days, which corresponds to the time when the error in the worst member becomes much larger than that of all other ensemble member (Fig. 4.1). Between 4–5 days, the observed error tendency of the best member is only about 0.5×10^{-5} PVU^2 s^{-1} and thereby less than half of that of the worst member.

In terms of the processes, the difference between the error growth in the best and worst member can be mostly attributed to the near-tropopause tendency. While the near-tropopause

tendency between 2.5–5 days is about 0.4×10^{-5} PVU^2 s^{-1} for the best member, it is about 1.4×10^{-5} PVU^2 s^{-1} for the worst member. The magnitude of the divergent tendency, instead, is very similar and both the best and worst member show a modest maximum around 4.5 days when Ex-Karl interacts with the waveguide. Between 2.5 and 5 days, the tropospheric-deep tendency is positive for the best member, whereas it is negative for the worst member. This tendency, however, is much smaller than the near-tropopause tendency and thus not of large importance to the error-growth differences between the two members. Baumgart et al. (2018) showed that the near-tropopause error tendency is mostly related to differences in the nonlinear Rossby wave dynamics. The large importance of near-tropopause tendency to the error-growth differences between the best and worst member is thus consistent with the observation that the largest differences between the best and worst member occur in the western quadrant of the hemisphere where the Rossby wave pattern is highly nonlinear (Figs. 4.8 and 4.9).

4.5 Summary and discussion

This study provides a quantitative framework to investigate the processes governing the amplification of forecast uncertainty and forecast errors in ensemble forecasts. A tendency equation for the ensemble variance of PV is derived and partitioned into the contributions from individual processes. The framework is applied to a case from the North Atlantic Waveguide and Downstream Impact Experiment, namely the interaction of tropical storm Karl and the Rossby wave pattern during extratropical transition. In the medium range, this interaction was associated with large forecast uncertainty (Schäfler et al., 2018) and only a few members captured the evolution correctly (Kumpf et al., 2018). Here, the same ensemble forecast is investigated as in Kumpf et al. (2018). We focus, however, not only on the region of Ex-Karl, but provide a hemispheric perspective on the ensemble evolution, before local differences, including the evolution of Ex-Karl, are highlighted. Two aspects of the ensemble behaviour are our main interests: i) the mean divergence of the ensemble members indicating the general amplification of forecast uncertainty and ii) the divergence of the best and worst member indicating extremes in possible error-growth scenarios.

The synoptic evolution of our case is characterized by a large-amplitude Rossby wave pattern spanning from (counter-clockwise) 180° to 60°E. This Rossby wave pattern is, in general, associated with larger PV variance than the other part of the hemisphere, in which the tropopause is more zonally oriented. Several local maxima of PV variance exist, including a cutoff, a large-amplitude ridge, a ridge-building event, and the interaction between Ex-Karl and the strong midlatitude PV gradient. The variance amplification is, on average for the midlatitudes of the Northern Hemisphere, dominated by near-tropopause dynamics. This contribution is particularly large in highly nonlinear regions of the wave pattern as, e.g, in the large-amplitude ridge around 120–60°W. Locally, however, the variance amplification can also be dominated by other processes. One prominent example is the region in which Ex-Karl interacts with the Rossby wave pattern around 4.5 days lead time. In this region, the variance amplification is dominated by upper-tropospheric divergence and tropospheric-deep interaction and is thereby

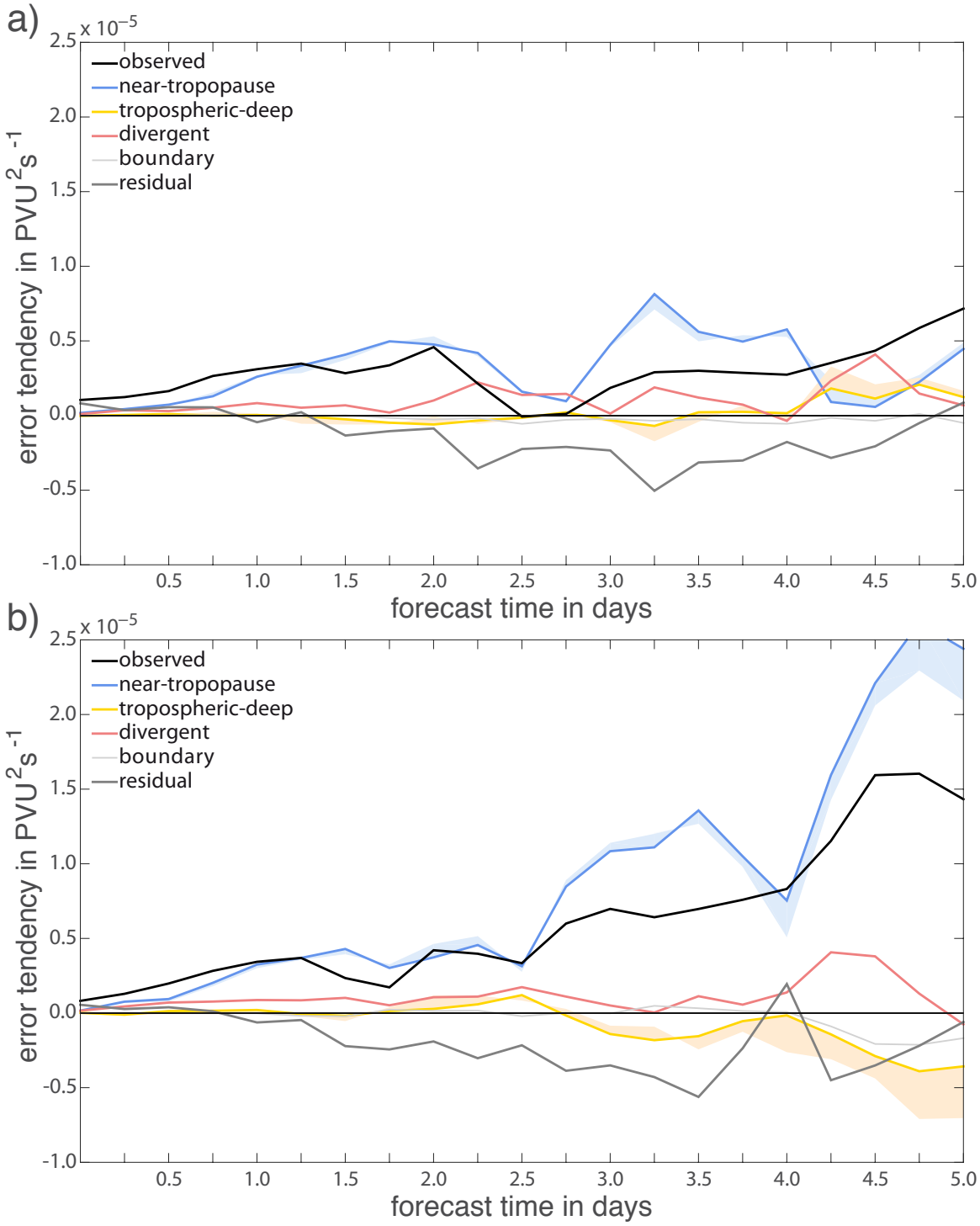

Figure 4.10: Individual contributions (colored lines) to error-enstrophy tendency (spatially-averaged over the midlatitudes of the Northern Hemisphere, 30–80°N) on the 325-K surface as indicated by Eq. 4.4 as a function of forecast time for a) best member (member 9), and b) worst member (member 18). The observed tendency is approximated by centered differences in time of the spatially-averaged error enstrophy. The shading next to the near-tropopause and tropospheric-deep tendencies indicates the uncertainty of the result due to the uncertainty of PPVI.

mostly related to uncertainties in the (moist baroclinic) cyclone development.
The differences between the error growth in the best and worst member can, to a large part, be attributed to the highly nonlinear evolution of the large-amplitude ridge around 120–60°W. Around 4–5 days, this region is captured well by the best member, whereas the worst member exhibits large errors associated with both the phase and the shape of the wave pattern. This different error pattern can be traced back in time until 3 days when a cutoff interacts with the upstream trough. During this interaction, the worst member is characterized by a distinct phase error, which largely amplifies in the highly nonlinear region of the wave pattern until 5 days. In terms of the processes, the differences in error growth between the two members are dominated by differences in the error growth by near-tropopause dynamics, which manifests the large importance of nonlinear Rossby wave dynamics.
The large importance of near-tropopause dynamics to the variance amplification and the error-growth differences between the best and worst member is consistent with the large importance of near-tropopause dynamics found in a recent case study of error growth in a deterministic ECMWF operational forecast (Baumgart et al., 2018) and in numerical error-growth experiments (Baumgart et al., 2019). In the current study, however, we identify also distinct local differences in the variance amplification as, e.g., in the region of Ex-Karl. In addition, the large (hemispherically-averaged) error-growth differences between the best and worst ensemble member around 5 days could be linked to a specific local event, namely the cutoff interaction around 2–3 days. The presented approach is thus able to identify flow features that are of large importance for the amplification of ensemble spread. Because both the interaction of Ex-Karl and of the cutoff with the strong midlatitude PV gradient can be interpreted as vortex-wave interactions, we speculate that such interactions are particularly sensitive to perturbations and are therefore prominent "amplifiers" of uncertainty in medium-range forecasts.
To extend our case study, future work could apply the presented approach in a more systematic way, e.g., to the trough/CAPE-type flow situations identifed by Rodwell et al. (2013, 2018) as associated with increased forecast uncertainty in the medium range. Rodwell et al. (2018) demonstrated that there are deficiencies in the short-range reliability associated with this flow situation by investigating a budget equation of ensemble variance in observation space. Complementary, our approach is able to identify the processes governing the flow-dependent amplification of ensemble spread in the medium range. As illustrated in Fig. 4.1, the amplification of forecast errors and forecast uncertainty can be highly nonlinear in the sense that ensemble members with relatively small errors at short lead times may have relatively large errors at medium-range lead times, and vice versa. Notwithstanding the importance of the short-range sources of forecast uncertainty, a better understanding of the flow-dependent nature of the medium-range amplification of forecast errors and forecast uncertainty (as also highlighted in previous studies, e.g., Ferranti et al., 2014) is of importance for the design and interpretation of ensemble forecasts also.
Even though the focus of this study is on the medium range, it is of interest to briefly consider the evolution of ensemble spread in the short range also. During the first 1-2 days,

the near-tropopause tendency dominates on average the amplification of PV variance. This result is in contrast to results from upscale-error-growth simulations, in which the ensemble members differ only in terms of the stochastic seed of the convection scheme (Selz, 2019). In these simulations, error growth in the early phase of the simulations is dominated by moist processes (Baumgart et al., 2019). This difference indicates i) that the ensemble perturbations representing initial-condition uncertainty in the ECMWF operational system in 2016 project most strongly on the near-tropopause tendency and ii) that this initial-condition uncertainty dominates on average over the (direct) generation of ensemble spread by processes with low intrinsic predictability, most notably moist processes. The latter notion is consistent with the lack of ensemble spread in low-predictability scenarios identified by Rodwell et al. (2018). Extending the focus of our analysis to the spread evolution at short lead times in low-predictability scenarios will require nonconservative-tendency data that are not available in the archives of operational centres. Results from such an analysis, however, may yield important insight into the deficiencies of flow-dependent reliability and may be of practical relevance for the design of stochastic parameterization schemes.

Acknowledgments

The research leading to these results was carried out as part of the Collaborative Research Center SFB/TRR 165 “Waves to Weather” within project A1: “Upscale impact of diabatic processes from convective to near-hemispheric scale” funded by the German Research Foundation (DFG). We thank Franziska Teubler for providing the piecewise PV inversion code and her assistance with the code. Furthermore, we are grateful to Marc Rautenhaus for archiving the real-time ECMWF data during the NAWDEX campaign and to both Alexander Kumpf and Marc Rautenhaus for insightful discussions about the forecast uncertainty associated with Ex-Karl. We thank the NAWDEX community for fruitful discussions during the NAWDEX campaign and the NAWDEX workshops. Furthermore, we are grateful for useful comments from two anonymous reviewers that helped to improve the manuscript.

CHAPTER 5

Summary and outlook

This chapter provides a general summary and outlook of the results presented in Chapters 2–4, which have been published in Baumgart et al. (2018)[1], Baumgart et al. (2019)[2], and Baumgart and Riemer (2019)[3]. A few phrases of this chapter coincide with phrases from these publications.

5.1 Summary

This thesis provides novel insight into the dynamics of forecast errors and forecast uncertainty by quantifying the relative importance of individual processes to the flow-dependent amplification of forecast errors and forecast uncertainty. The focus of this thesis is on the near-tropopause region of the midlatitudes, which is of particular importance to the evolution of midlatitude weather systems and the evolution of PV errors.

Potential-vorticity framework

A PV framework is developed to quantify the relative importance of individual processes to the amplification of forecast errors and forecast uncertainty. This PV framework builds on the PV perspective of midlatitude dynamics and makes use of the three key characteristics of PV (conservation, inversion, and attribution, Hoskins et al., 1985). Using the same partitioning of processes as in a previous study on Rossby-wave dynamics (Teubler and Riemer, 2016), we partition the contributions from four processes, namely 1) near-tropopause dynamics, 2) tropospheric-deep interaction, 3) upper-tropospheric divergent flow, and 4) nonconservative processes. The contribution of near-tropopause dynamics describes the influence of upper-level anomalies on the upper-level evolution and is mostly related to nonlinear Rossby-wave

dynamics near the tropopause (e.g., Hoskins et al., 1985). The contribution of tropospheric-deep interaction describes the influence of lower-level anomalies on the upper-level evolution and thereby comprises the influence of vertical interaction, including the influence of baroclinic instability (Eady, 1949; Hoskins et al., 1985; Davis and Emanuel, 1991; Heifetz et al., 2004). Upper-tropospheric divergent flow can be associated with dry dynamics and diabatic processes (see, e.g., Holton and Hakim, 2013). For Rossby-wave dynamics, a strong divergent contribution is often associated with latent-heat release below (Davis et al., 1993; Riemer et al., 2014; Quinting and Jones, 2016). The contribution from nonconservative processes describes the direct PV modification due to diabatic heating and nonconservative momentum changes. This contribution is usually of smaller importance than the indirect diabatic effect from upper-tropospheric divergence (e.g., Davis et al., 1996; Teubler and Riemer, 2016). Following the ideas from Davies and Didone (2013), we derive a tendency equation for the PV error and the ensemble variance of PV. We use these equations in combination with the partitioning of processes to quantify the relative importance of individual processes to the amplification of forecast errors and forecast uncertainty.

Error growth in an operational state-of-the-art deterministic forecast

In the first part of this thesis, we investigate the error growth in an operational state-of-the-art deterministic forecast. To quantify the processes governing the amplification of forecast errors, we derive a tendency equation for error-potential enstrophy, defined here as the squared PV error, and partition it into the influence of individual processes. The PV-error growth of this case is characterized first by localized error maxima that develop around 2 forecast days. These maxima change into wave-like error patterns that reach the scale of individual Rossby-wave anomalies until 6 forecast days. Error growth maximizes along the dynamical tropopause and occurs most prominently within the Rossby-wave patterns in the Pacific and the Atlantic, respectively. Our quantitative PV framework reveals that the error growth of this case is clearly dominated by differences in the near-tropopause dynamics. This near-tropopause error growth can almost entirely be related to the influence of *nonlinear* Rossby-wave dynamics. A large localized importance of upper-tropospheric divergence is found during two ridge-building events in the Atlantic and the Pacific, respectively, when the development of a cyclone and its associated warm conveyor belt is misrepresented in the forecast. Tropospheric-deep interaction makes a large contribution to error growth in the Pacific-wave pattern around 4–5 forecast days. Otherwise, the impact of tropospheric-deep interaction and thereby baroclinic instability is small. Theoretical considerations with the Eady (1949) model provide a plausible explanation for this small importance of baroclinic instability. The results from the quantitative PV framework thus demonstrate that the meso-scale errors generated near the tropopause do not primarily project on differences in the baroclinic instability as stated in previous studies (e.g., Zhang et al., 2007). Instead, they amplify most prominently because of differences in the nonlinear Rossby-wave dynamics.

Error growth in upscale error-growth simulations

In the second part of this thesis, we investigate upscale error growth from the grid scale up to the planetary scale using global simulations with a stochastic convection scheme. Twelve cases of real-case ensemble simulations are considered, in which the ensemble members differ only in the random seed of the convection scheme (Selz, 2019). Using the PV framework, we find a distinct sequence of the processes governing the upscale error growth near the tropopause (as summarized in the schematic diagram in Fig. 5.1): In the first 12 h, error growth is dominated by latent-heating differences induced from the convection scheme. Differences in the upper-tropospheric divergence dominate the error growth in the following 1.5 days. Upper-tropospheric divergence thus provides an effective mechanism to project an error from moist processes into the tropopause region. After about 2 days, error growth is dominated by differences in the nonlinear Rossby-wave dynamics. By comparing the results from the PV framework with a framework that identifies the envelope of Rossby waves based on local finite-amplitude wave activity (Ghinassi et al., 2018), a fourth stage of the error growth can be identified between 14.5 and 18 days. During this stage, the error grows from the synoptic scale of individual troughs and ridges up to the planetary scale of the Rossby-wave envelope. Compared to a conceptual model for upscale error growth (Zhang et al., 2007), our PV framework provides novel insight into the processes governing upscale error growth. In particular, our quantitative results highlight the importance of differences in the tropopause displacement induced by the upper-tropospheric divergent flow and the importance of differences in the nonlinear Rossby-wave dynamics to upscale error growth.

Amplification of ensemble spread in an operational state-of-the-art forecast

In the third part of this thesis, we investigate the amplification of ensemble spread for an operational forecast during the NAWDEX campaign that was associated with large forecast uncertainty. The focus is on two aspects of the ensemble behavior: i) the mean divergence of the ensemble members indicating the general amplification of forecast uncertainty and ii) the divergence of the best and worst member indicating extremes in possible error-growth scenarios. To quantify the processes governing the amplification of forecast uncertainty, we derive a tendency equation for the ensemble variance of PV. Averaged over the midlatitudes of the Northern Hemisphere, the variance amplification of this case is dominated by the contribution of near-tropopause dynamics, which is particularly large in highly nonlinear regions of the Rossby-wave pattern. Locally, however, there can be large differences in the variance amplification as, for example, in the region where the extratropical transition of tropical storm Karl interacts with the Rossby-wave pattern. In this region, the variance amplification is dominated by upper-tropospheric divergence and tropospheric-deep interaction and thus mostly related to the moist baroclinic cyclone development of Ex-Karl. The differences between the error growth in the best and worst member can, to a large part, be attributed to the highly nonlinear evolution of a large-amplitude ridge around 5 days, which is captured well by the best member, while it is associated with large error growth in the worst member. These diffe-

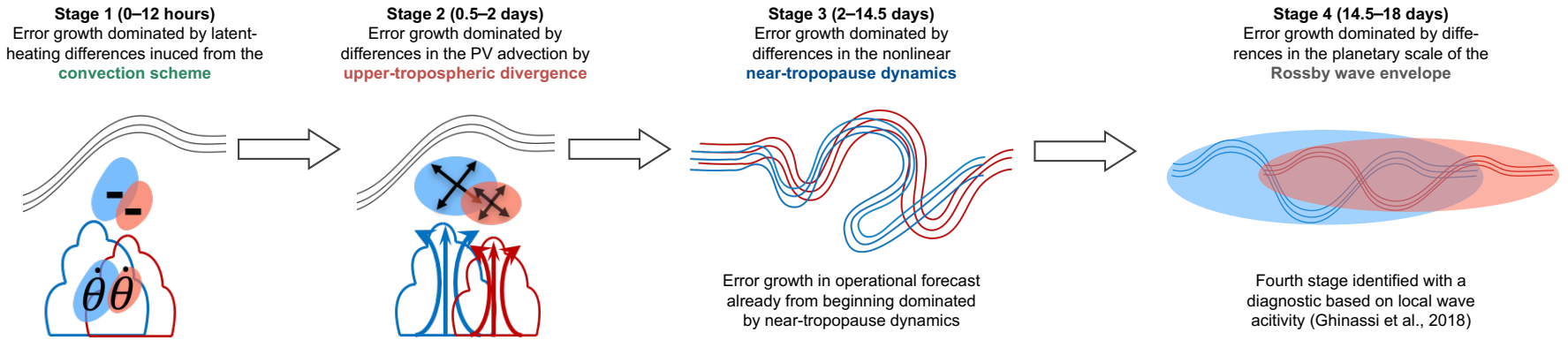

Figure 5.1: Schematic diagram illustrating the four stages of upscale error growth from a process-based perspective (see text for details).

rences can be traced back in time until 3 days when a cutoff interacts with the upstream trough. In terms of the processes, the error-growth differences can be mostly attributed to differences in the error growth by near-tropopause dynamics. This study is thus able to identify flow-dependent differences in the amplification of ensemble spread that result from, e.g., the extratropical transition of a tropical storm or a cutoff reabsorption.

Synthesis of results

This thesis provides process-based insight into the amplification of forecast errors and forecast uncertainty. This insight can be used to identify regions that are prone to a large amplification of forecast errors and forecast uncertainty. Based on the result that the amplification of forecast errors and forecast uncertainty in the operational forecasts is dominated by nonlinear Rossby-wave dynamics near the tropopause, we hypothesize that the highly-nonlinear, large-amplitude regions of the Rossby-wave pattern are most susceptible to rapid growth of forecast errors and forecast uncertainty in current weather-prediction models. Baroclinic instability and diabatic processes, including the indirect effect associated with upper-tropospheric divergence, play important roles in setting up such large-amplitude Rossby-wave patterns and can thus play important indirect roles for the amplification of forecast errors and forecast uncertainty.

The results from the upscale error-growth simulations, on the other hand, can be used to identify regions with low intrinsic predictability. As upper-tropospheric divergence provides an effective mechanism to project errors from moist processes into the tropopause region, one can expect that localized errors along the tropopause establish predominantly in regions with latent-heat release below. The dominance of nonlinear Rossby-wave dynamics for the further error growth suggests that the highly-nonlinear, large-amplitude regions of the Rossby-wave pattern are most susceptible for the further error growth.

One of the areas, from which the largest benefits in numerical weather prediction have been achieved over the past 25 years, represents the introduction of ensemble forecasts (e.g., Bauer et al., 2015; Buizza, 2019). The central aim of ensemble forecasts is to provide an accurate and reliable estimate of the future range of possible scenarios and of the uncertainty associated with a forecast. Averaged over the Northern Hemisphere, the annual-mean forecast uncertainty is now represented reasonably well at ECMWF (see, e.g., Fig. 1a of Rodwell et al., 2018). On a day-to-day basis, ensemble forecasts, however, have still problems in representing forecast uncertainty (see, e.g., Fig. 1b of Rodwell et al., 2018). It is thus important to gain a better understanding of the processes governing the amplification of forecast uncertainty. The extension of our process-based PV framework from deterministic forecasts (Chapter 2) to ensemble forecasts (Chapter 4) thus represents an important step to gain further insight not only into the dynamics of forecast errors, but also into the dynamics of forecast uncertainty. This insight might help to improve the representation of forecast uncertainty in future ensemble-forecasting systems.

Using the PV framework, flow-dependent differences in the amplification of ensemble spread

can be identified. For the ensemble forecast investigated in this thesis, the interaction of a cutoff and of Ex-Karl with the strong midlatitude PV gradient are of large importance to the amplification of ensemble spread, which suggests that vortex-wave interactions can act as "amplifiers" for ensemble spread. The amplification of ensemble spread can be highly nonlinear in the sense that ensemble members with relatively small error at short lead times can have relatively large errors at medium-range lead times, and vice versa. In that sense, a better understanding of the flow-dependent amplification of forecast uncertainty in the medium range is important for the interpretation of ensemble forecasts.
The first two stages of upscale error growth do not occur as dominating error-growth mechanisms in the operational forecasts. This result provides evidence that initial-condition errors are, on average, still of larger importance to error growth in operational forecasts than errors from moist processes with low intrinsic predictability. Nevertheless, stochastic perturbations are necessary to increase the spread in current ensemble-forecasting systems that would be underdispersive otherwise. The insight into upscale error growth obtained in this thesis can thus be used to help designing suitable ensemble perturbations that effectively initiate an upscale error-growth cascade.

5.2 Insight into research questions

In the following, the research questions raised in Sec. 1.4 will be repeated and answered based on the results of this thesis.

1. **Error growth in an operational state-of-the-art deterministic forecast**

   a) *How can the PV perspective be used to quantify the processes governing the amplification of forecast errors?*
   Following Davies and Didone (2013), we derive an error-tendency equation based on the difference between the PV tendency of the forecast and the analysis. Here, we extend the work of Davies and Didone (2013) by providing not only a qualitative, but also a quantitative view on the dynamics of forecast-error growth.
   Using the same partitioning of processes as in a previous analysis of Rossby-wave dynamics (Teubler and Riemer, 2016), we partition the error tendency into the contributions from individual processes. We are thereby able to quantify the relative importance of near-tropopause dynamics, tropospheric-deep interaction, upper-tropospheric divergence, and nonconservative processes to the error growth near the tropopause.

   b) *What are the most important processes governing the error growth in an operational state-of-the-art forecast?*
   Averaged over the midlatitudes of the Northern Hemisphere, the operational forecast analyzed in this thesis is dominated by differences in the nonlinear Rossby-wave dynamics near the tropopause. Upper-tropospheric divergence can be of large localized importance to the error growth in regions of a misrepresented cy-

clone development, while the direct importance of tropospheric-deep interaction (and thereby also baroclinic instability) is limited.

c) *How can the error-growth mechanisms be explained from a theoretical perspective?*
Theoretical considerations with the Eady (1949) model help to explain the limited importance of baroclinic instability to the error growth in our case: A specific favorable configuration between the upper- and lower-level error pattern is needed for baroclinic error growth to occur, which is in our case only observed over a limited region and duration. Another reason for the limited importance of baroclinic instability is that the magnitude of baroclinic error growth increases with the spatial scale of the error pattern, which is significantly smaller in our case than that of anomalies governing baroclinic growth in the midlatitudes.
In our case, the by far largest error-growth contribution is given by near-tropopause dynamics. A further separation of this contribution demonstrates that it is dominated by differences in the nonlinear dynamics, while differences in the linear dynamics are only of small importance to the error growth. For two example regions, we illustrate schematically how such differences in the nonlinear Rossby-wave dynamics lead to large error growth.

2. **Error growth in upscale error-growth simulations**

a) *What are the most important processes governing upscale error growth from the grid scale up to the planetary scale?*
Using the PV framework, we quantify the processes governing the upscale error growth in global simulations with a stochastic convection scheme: In the first 12 h, upscale error growth is dominated by latent-heating differences induced from the convection scheme. In the following 1.5 days, error growth is dominated by differences in the upper-tropospheric divergence, which provides an effective mechanism to project an error from moist processes into the tropopause region. Afterwards, error growth is dominated by differences in the nonlinear Rossby-wave dynamics near the tropopause up to 14.5 days. Based on a diagnostic for the Rossby-wave envelope (Ghinassi et al., 2018), error growth up to the planetary scale can be identified until 18 days.

b) *Are the quantitative results from the PV framework consistent with the conceptual model for upscale error growth?*
The results from the PV framework confirm the multi-stage behavior of upscale error growth in the conceptual model as the error growth can be divided into distinct stages based on the governing processes (see previous research question). Compared to the conceptual model, our results, however, provide novel insight into the processes governing upscale error growth, especially in the second and third stage: In the second stage, the conceptual model assumes that differences in the diabatic PV generation dominate the error growth, while our quantitative results show that the error growth during this stage is dominated by differences

in the displacement of the strong midlatitude PV gradient induced by the upper-tropospheric divergent flow. In the third stage, the conceptual model relates error growth to baroclinic instability, while our quantitative results show that the error growth during this stage is dominated by differences in the near-tropopause dynamics. Furthermore, an additional fourth stage of planetary-scale error growth is identified that was missing in the conceptual model.

c) *Do the processes governing upscale error growth differ from the processes governing the error growth in operational forecasts?*
In the operational forecast investigated in this thesis, the first two stages of upscale error growth do not occur. Instead, the error growth is dominated by differences in the near-tropopause dynamics from the beginning. This result indicates that errors in the initial conditions are still of larger importance to the error growth in current state-of-the-art forecasting systems than errors from moist processes with low intrinsic predictability. Nevertheless, stochastic perturbations are needed in ensemble forecasting systems to increase ensemble spread. In that sense, upscale error growth is also of practical importance to current weather forecasts, albeit not as the dominating error-growth mechanism.

3. **Amplification of ensemble spread in an operational state-of-the-art forecast**

a) *How can the PV perspective be used to quantify the processes governing the amplification of forecast uncertainty?*
In this thesis, we derive a tendency equation for the ensemble variance of PV and partition it into the contributions from individual processes to quantify the general amplification of forecast uncertainty. Furthermore, the error growth of the ensemble members with smallest and largest error growth is compared to indicate two extremes in possible error-growth scenarios.

b) *What are the most important processes governing the amplification of ensemble spread in an operational state-of-the-art forecast?*
Averaged over the midlatitudes of the Northern Hemisphere, the amplification of ensemble variance is dominated by near-tropopause dynamics. For two localized regions, in which ridge building occurs in association with a cyclone development, we show, however, that the variance amplification can also be dominated by other processes, namely upper-tropospheric divergence and/or tropospheric-deep interaction. The differences between the error growth in the best and the worst member can, to a large part, be attributed to differences in the error growth by near-tropopause dynamics, which can be traced back to differences in the representation of a cutoff reabsorption.

c) *Are there flow-dependent differences in the amplification of ensemble spread?*
Distinct flow-dependent differences in the amplification of ensemble spread are highlighted in this thesis. Prominent examples include the interaction of the ex-

tratropical transition of tropical storm Karl with the Rossby-wave pattern, in which the variance amplification is related to the moist baroclinic cyclone development, and the representation of a cutoff reabsorption, which leads to the large error-growth differences between the best and the worst member, respectively. Flow features that are particularly prone to a large amplification of ensemble spread can thus be identified.

5.3 Outlook

A wide range of ideas for future research arises from the work of this thesis. These ideas will be outlined in the following.

Climatological perspective on forecast-error growth

In Chapter 2, we analyze the error growth in a case study of an operational forecast. This case provides the novel insight that synoptic-scale error growth can be dominated by differences in the nonlinear Rossby-wave dynamics and not, as expected from previous studies (e.g., Zhang et al., 2007), by differences in baroclinic instability. To draw more general conclusions, it is necessary to perform a more systematic investigation. One possibility would be to apply the PV framework used in this thesis to a climatological study with a large set of data. Such a climatological investigation could, for example, be carried out with the ERA-Interim reanalyses and reforecasts (Berrisford et al., 2011) or the operational analyses and forecasts archived at ECMWF (`www.ecmwf.int/en/forecasts/datasets`). These data sets provide several years of analysis and forecast data for a robust investigation of forecast-error growth. The hypothesis from this thesis that nonlinear Rossby-wave dynamics are more important to synoptic-scale error growth than baroclinic instability could thereby be verified from a climatological perspective.

Composite view on situations with small/large error growth

One very interesting aspect of the dynamics of forecast-error growth is whether there are systematic differences between situations with small and large error growth. A composite of such cases could be drawn from a climatological data set, taking for example the 10th and 90th percentile of the climatological error-magnitude distribution to define composites with small and large error growth, respectively. Based on the PV framework, one could then investigate whether there are systematic differences in the processes governing the error growth. This process-based insight could be used to identify flow situations that are prone to small and large error growth, respectively.

Another possibility to gain process-based insight into situations with large error growth would be to apply the PV framework to the composite of forecast-bust events investigated by Rodwell et al. (2013). Rodwell et al. (2013) showed that the largest forecast errors over Europe are related to initial errors in representing mesoscale convective systems over the U.S. By applying

the PV framework to these forecast-bust events, one could highlight the processes that are responsible for the further amplification of these initial errors from moist processes. Such an analysis would thus provide a complementary perspective on the dynamics of forecast busts.

Localized perspective on individual weather systems

In this thesis, the amplification of forecast errors and forecast uncertainty is mostly investigated in terms of spatial averages over large regions. The results from the case study in Chapter 4, however, indicate that locally there can be pronounced differences in the processes governing the amplification of ensemble spread. A more localized perspective can thus help to identify flow-dependent differences. One could for example focus on individual weather systems and quantify the amplification of forecast errors and forecast uncertainty separately for these weather systems. Such a separation could help to identify, which weather systems are most susceptible to a large amplification of forecast errors and forecast uncertainty. The PV framework could furthermore provide insight into the processes governing the amplification of forecast errors and forecast uncertainty in different weather systems.

Comparison between upscale error growth and operational forecast error growth

From the comparison between the case study of operational forecast-error growth (Chapter 2) and the upscale error-growth simulations (Chapter 3) it is evident that there are still significant differences between operational forecast-error growth and upscale error growth. The first two stages of upscale error growth do not appear as dominating error-growth mechanisms in the operational case study, at least when averaging over large regions as the midlatitudes of the Northern Hemisphere or whole Rossby-wave patterns. In contrast, the error growth in the operational forecast is dominated by differences in the near-tropopause dynamics from the beginning. This result indicates that initial-condition errors contained in operational forecasts are still of larger importance to the error growth than errors from moist processes with low intrinsic predictability. For future research, it is an interesting question how much initial-condition errors have to be reduced, until the first two stages of upscale error growth become apparent.

An identical-twin simulation has been carried out by T. Selz (LMU Munich), which contains perturbations from the ensemble data assimilation (EDA) system of ECMWF (Isaksen et al., 2010) in addition to the stochastic perturbations from the stochastic convection scheme. This simulation thus includes realistic perturbations for the initial-condition uncertainty. Using the PV framework, we can show that the error growth in this simulation is dominated by differences in the near-tropopause dynamics from the beginning (Fig. 5.2a). The first two stages of upscale error growth do not appear, as it is the case for the operational forecast investigated in Chapter 2. A second identical-twin simulation has been carried out, in which the perturbations from the EDA system have been rescaled to 10% of their original amplitude. In this simulation, the initial error growth is dominated by latent-heating differences induced from the convection scheme, followed by a short period, in which differences in the

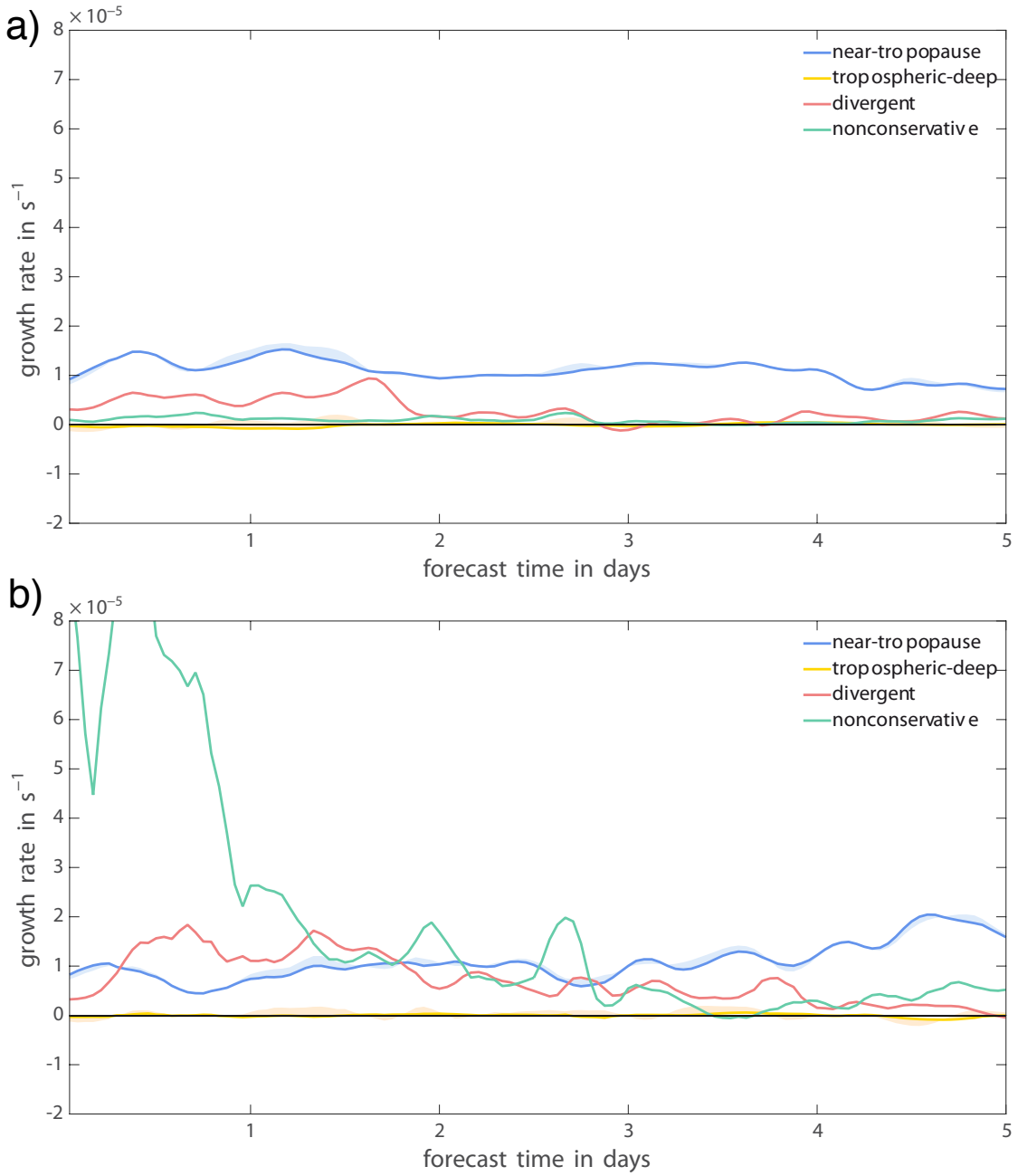

Figure 5.2: Comparison of upscale error-growth simulations with simulations that include additional perturbations from the EDA system: a) original EDA perturbations, and b) EDA perturbations rescaled to 10% of their original amplitude.

upper-tropospheric divergent flow dominate the error growth (Fig. 5.2b). In this simulation, the first two stages of upscale error growth do thus appear. Errors from moist processes with low intrinsic predictability can thus play a substantial role for error growth when the initial-condition uncertainty is reduced to 10%.
Additional simulations are needed to make robust statements about what reduction of initial-condition errors is needed before the first two stages of upscale error growth become dominating error-growth mechanisms. One possibility would be to simulate the same set of cases as analyzed in Chapter 3 with additional perturbations from the EDA system. By systematically rescaling the amplitude of the EDA perturbations (e.g., starting from 90% and then going stepwise down to 10%), one could identify the amount of rescaling that is needed, until the first two stages of upscale error growth become dominating error-growth mechanisms. Such an analysis would provide process-based insight into the practical importance of upscale error growth.

Predictability of weather regimes

The results from the upscale error-growth simulations in Chapter 3 show that there exists a final error-growth stage, in which errors grow from the synoptic scale of individual troughs and ridges up to the planetary scale of the Rossby-wave envelope. On the planetary scale, predictability is closely related to the predictability of weather regimes. Due to their large spatial scale and long life cycle, these weather regimes provide possible sources for predictability on the subseasonal to seasonal time scale (e.g., Grams et al., 2017; Ferranti et al., 2018). The individual weather regimes, however, differ in terms of their predictability with the blocking regime having lowest predictability (Ferranti et al., 2014). By applying the PV framework to individual weather regimes, one could highlight systematic differences between the amplification of forecast errors and forecast uncertainty in individual weather regime. This process-based insight could help to explain why some weather regimes have higher or lower predictability than other weather regimes.

Future work within Waves to Weather

This thesis is part of project A1 “Upscale impact of diabatic processes from convective to near-hemispheric scale” from the first phase of the transregional collaborative research center 165 “Waves to Weather”. Building on the results from phase 1, this project is split into two new projects in phase 2, namely A1 “Multiscale analysis of the evolution of forecast uncertainty” and A8 (N) “Dynamics and predictability of blocked regimes in the Atlantic-European region”.
Compared to phase 1, the A1 project in phase 2 will provide a more localized perspective on upscale error growth and analyze different sources for convective-scale uncertainty. Based on asymptotic theory, a theoretical model will be developed for upscale error growth from the convective scale up to the synoptic scale. This project will moreover provide a systematic comparison between upscale error growth and operational forecast-error growth based on

simulations with realistic initial-condition uncertainty.

The A8 (N) project in phase 2 will investigate the dynamics and predictability of weather regimes. The focus of this project will be on blocked weather regimes, which are often associated with high-impact events, but also most difficult to represent in numerical weather prediction models (Ferranti et al., 2014). In this project, the PV framework will be applied to a large set of data (reanalysis- and reforecast-data sets and operational analysis- and forecast-data sets) to gain deeper insight into the underlying dynamics and the predictability of blocked weather regimes.

Bibliography

Abbe, C., 1901: The physical basis of long-range weather forecasts. *Monthly Weather Review*, **29 (12)**, 551–561, doi:10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2.

Agustí-Panareda, A., S. L. Gray, G. C. Craig, and C. Thorncroft, 2005: The extratropical transition of tropical cyclone Lili (1996) and its crucial contribution to a moderate extratropical development. *Monthly Weather Review*, **133 (6)**, 1562–1573, doi:10.1175/MWR2935.1.

Agustí-Panareda, A., C. D. Thorncroft, G. C. Craig, and S. L. Gray, 2004: The extratropical transition of hurricane Irene (1999): A potential-vorticity perspective. *Quarterly Journal of the Royal Meteorological Society*, **130 (598)**, 1047–1074, doi:10.1256/qj.02.140.

Anwender, D., P. A. Harr, and S. C. Jones, 2008: Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Case studies. *Monthly Weather Review*, **136 (9)**, 3226–3247, doi:10.1175/2008MWR2249.1.

Appenzeller, C., H. C. Davies, and W. A. Norton, 1996: Fragmentation of stratospheric intrusions. *Journal of Geophysical Research: Atmospheres*, **101 (D1)**, 1435–1456, doi:10.1029/95JD02674.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525 (7567)**, 47–55, doi:10.1038/nature14956.

Baumgart, M., P. Ghinassi, V. Wirth, T. Selz, G. C. Craig, and M. Riemer, 2019: Quantitative view on the processes governing the upscale error growth up to the planetary scale using a stochastic convection scheme. *Monthly Weather Review*, **147 (5)**, 1713–1731, doi:10.1175/MWR-D-18-0292.1.

Baumgart, M., and M. Riemer, 2019: Processes governing the amplification of ensemble spread in a medium-range forecast with large forecast uncertainty. *Quarterly Journal of the Royal Meteorological Society*, **145 (724)**, 3252–3270, doi:10.1002/qj.3617.

Baumgart, M., M. Riemer, V. Wirth, F. Teubler, and S. T. K. Lang, 2018: Potential vorticity dynamics of forecast errors: A quantitative case study. *Monthly Weather Review*, **146 (5)**, 1405–1425, doi:10.1175/MWR-D-17-0196.1.

Berner, J., and Coauthors, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, **98 (3)**, 565–588, doi:10.1175/BAMS-D-15-00268.1.

Berrisford, P., and Coauthors, 2011: The ERA-Interim archive Version 2.0. Report 1, ECMWF, Shinfield Park, Reading, 23 pp. https://www.ecmwf.int/node/8174.

Bierdel, L., T. Selz, and G. Craig, 2017: Theoretical aspects of upscale error growth through the mesoscales: an analytical model. *Quarterly Journal of the Royal Meteorological Society*, **143 (709)**, 3048–3059, doi:10.1002/qj.3160.

Bierdel, L., T. Selz, and G. C. Craig, 2018: Theoretical aspects of upscale error growth on the mesoscales: Idealized numerical simulations. *Quarterly Journal of the Royal Meteorological Society*, **144 (712)**, 682–694, doi:10.1002/qj.3236.

Bjerknes, V., 1904: Das Problem der Wettervorhersage betrachtet vom Standpunkt der Mechanik und Physik. *Meteorologische Zeitschrift*, **21**, 1–7.

Boer, G. J., 1984: A spectral analysis of predictability and error in an operational forecast system. *Monthly Weather Review*, **112 (6)**, 1183–1197, doi:10.1175/1520-0493(1984)112<1183:ASAOPA>2.0.CO;2.

Boisserie, M., P. Arbogast, L. Descamps, O. Pannekoucke, and L. Raynaud, 2014: Estimating and diagnosing model error variances in the Météo-France global NWP model. *Quarterly Journal of the Royal Meteorological Society*, **140 (680)**, 846–854, doi:10.1002/qj.2173.

Bretherton, F. P., 1966: Critical layer instability in baroclinic flows. *Quarterly Journal of the Royal Meteorological Society*, **92 (393)**, 325–334, doi:10.1002/qj.49709239302.

Brown, T. A., 1974: *Admissible Scoring Systems for Continuous Distributions.* Manuscript P-5235, The Rand Corporation: Santa Monica, CA, https://www.rand.org/pubs/papers/P5235.html.

Buizza, R., 2019: Introduction to the special issue on "25 years of ensemble forecasting". *Quarterly Journal of the Royal Meteorological Society*, **145 (S1)**, 1–11, doi:10.1002/qj.3370.

Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quarterly Journal of the Royal Meteorological Society*, **141 (693)**, 3366–3382, doi:10.1002/qj.2619.

Buizza, R., M. Leutbecher, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **134 (637)**, 2051–2066, doi:10.1002/qj.346.

Buizza, R., M. Leutbecher, and A. Thorpe, 2015: Living with the butterfly effect: a seamless view of predictability. *ECMWF Newsletter*, **145**, 18–23, doi:10.21957/x4h3e8w3.

Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ecmwf ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **125 (560)**, 2887–2908, doi:10.1002/qj.49712556006.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *Journal of the Atmospheric Sciences*, **52 (9)**, 1434–1456, doi:10.1175/1520-0469(1995)052<1434:TSVSOT>2.0.CO;2.

Chagnon, J. M., and S. L. Gray, 2015: A diabatically generated potential vorticity structure near the extratropical tropopause in three simulated extratropical cyclones. *Monthly Weather Review*, **143 (6)**, 2337–2347, doi:10.1175/MWR-D-14-00092.1.

Chagnon, J. M., S. L. Gray, and J. Methven, 2013: Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Quarterly Journal of the Royal Meteorological Society*, **139 (674)**, 1270–1282, doi:10.1002/qj.2037.

Charney, J., 1955: The use of the primitive equations of motion in numerical prediction. *Tellus*, **7 (1)**, 22–26, doi:10.1111/j.2153-3490.1955.tb01138.x.

Charney, J. G., F. R., and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2 (4)**, 237–254, doi:10.1111/j.2153-3490.1950.tb00336.x.

Davies, H. C., and C. H. Bishop, 1994: Eady edge waves and rapid development. *Journal of the Atmospheric Sciences*, **51 (13)**, 1930–1946, doi:10.1175/1520-0469(1994)051<1930:EEWARD>2.0.CO;2.

Davies, H. C., and M. Didone, 2013: Diagnosis and dynamics of forecast error growth. *Monthly Weather Review*, **141 (7)**, 2483–2501, doi:10.1175/MWR-D-12-00242.1.

Davis, C. A., 1992: Piecewise potential vorticity inversion. *Journal of the Atmospheric Sciences*, **49 (16)**, 1397–1411, doi:10.1175/1520-0469(1992)049<1397:PPVI>2.0.CO;2.

Davis, C. A., and K. A. Emanuel, 1991: Potential vorticity diagnostics of cyclogenesis. *Monthly Weather Review*, **119 (8)**, 1929–1953, doi:10.1175/1520-0493(1991)119<1929:PVDOC>2.0.CO;2.

Davis, C. A., E. D. Grell, and M. A. Shapiro, 1996: The balanced dynamical nature of a rapidly intensifying oceanic cyclone. *Monthly Weather Review*, **124 (1)**, 3–26, doi:10.1175/1520-0493(1996)124<0003:TBDNOA>2.0.CO;2.

Davis, C. A., M. T. Stoelinga, and Y. H. Kuo, 1993: The integrated effect of condensation in numerical simulations of extratropical cyclogenesis. *Monthly Weather Review*, **121 (8)**, 2309–2330, doi:10.1175/1520-0493(1993)121<2309:TIEOCI>2.0.CO;2.

de Vries, A. J., S. B. Feldstein, M. Riemer, E. Tyrlis, M. Sprenger, M. Baumgart, M. Fnais, and J. Lelieveld, 2016: Dynamics of tropical–extratropical interactions and extreme precipitation events in Saudi Arabia in autumn, winter and spring. *Quarterly Journal of the Royal Meteorological Society*, **142 (697)**, 1862–1880, doi:10.1002/qj.2781.

Dirren, S., M. Didone, and H. C. Davies, 2003: Diagnosis of "forecast-analysis" differences of a weather prediction system. *Geophysical Research Letters*, **30 (20)**, 2060, doi:10.1029/2003GL017986.

Durran, D. R., and M. Gingrich, 2014: Atmospheric predictability: Why butterflies are not of practical importance. *Journal of the Atmospheric Sciences*, **71 (7)**, 2476–2488, doi:10.1175/JAS-D-14-0007.1.

Durran, D. R., and J. A. Weyn, 2016: Thunderstorms do not get butterflies. *Bulletin of the American Meteorological Society*, **97 (2)**, 237–243, doi:10.1175/BAMS-D-15-00070.1.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1 (3)**, 33–52, doi:10.1111/j.2153-3490.1949.tb01265.x.

ECMWF, 2013: *Part IV: Physical Processes*. IFS Documentation CY38R1, ECMWF, https://www.ecmwf.int/node/9245.

ECMWF, 2018: *Part V: Ensemble prediction system*. IFS Documentation CY45R1, ECMWF, https://www.ecmwf.int/node/18715.

Ertel, H., 1942: Ein neuer hydrodynamischer Erhaltungssatz. *Meteorologische Zeitschrift*, **30 (36)**, 543–544, doi:10.1007/BF01475602.

Fehlmann, R., and H. C. Davies, 1997: Misforecasts of synoptic systems: Diagnosis via PV retrodiction. *Monthly Weather Review*, **125 (9)**, 2247–2264, doi:10.1175/1520-0493(1997)125<2247:MOSSDV>2.0.CO;2.

Fehlmann, R., and H. C. Davies, 1999: Role of salient potential-vorticity elements in an event of frontal-wave cyclogenesis. *Quarterly Journal of the Royal Meteorological Society*, **125 (557)**, 1801–1824, doi:10.1002/qj.49712555716.

Ferranti, L., S. Corti, and M. Janousek, 2014: Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, **141 (688)**, 916–924, doi:10.1002/qj.2411.

Ferranti, L., L. Magnusson, F. Vitart, and D. S. Richardson, 2018: How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe? *Quarterly Journal of the Royal Meteorological Society*, **144 (715)**, 1788–1802, doi:10.1002/qj.3341.

Ghinassi, P., M. Baumgart, F. Teubler, M. Riemer, and V. Wirth, 2020: A budget equation for the amplitude of Rossby wave packets based on finite-amplitude local wave activity. *Journal of the Atmospheric Sciences*, **77 (1)**, 277–296, doi:10.1175/JAS-D-19-0149.1.

Ghinassi, P., G. Fragkoulidis, and V. Wirth, 2018: Local finite-amplitude wave activity as a diagnostic for Rossby wave packets. *Monthly Weather Review*, **146 (12)**, 4099–4114, doi:10.1175/MWR-D-18-0068.1.

Giannakaki, P., and O. Martius, 2016: An object-based forecast verification tool for synoptic-scale Rossby waveguides. *Weather and Forecasting*, **31 (3)**, 937–946, doi:10.1175/WAF-D-15-0147.1.

Grams, C. M., R. Beerli, S. Pfenninger, I. Staffell, and H. Wernli, 2017: Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, **7**, 557–562, doi:10.1038/nclimate3338.

Grams, C. M., S. C. Jones, C. A. Davis, P. A. Harr, and M. Weissmann, 2013: The impact of Typhoon Jangmi (2008) on the midlatitude flow. Part I: Upper-level ridgebuilding and modification of the jet. *Quarterly Journal of the Royal Meteorological Society*, **139 (677)**, 2148–2164, doi:10.1002/qj.2091.

Grams, C. M., L. Magnusson, and E. Madonna, 2018: An atmospheric dynamics perspective on the amplification and propagation of forecast error in numerical weather prediction models: A case study. *Quarterly Journal of the Royal Meteorological Society*, **144 (717)**, 2577–2591, doi:10.1002/qj.3353.

Grams, C. M., and Coauthors, 2011: The key role of diabatic processes in modifying the upper-tropospheric wave guide: a North Atlantic case-study. *Quarterly Journal of the Royal Meteorological Society*, **137 (661)**, 2174–2193, doi:10.1002/qj.891.

Gray, S. L., 2006: Mechanisms of midlatitude cross-tropopause transport using a potential vorticity budget approach. *Journal of Geophysical Research: Atmospheres*, **111 (D17)**, doi:10.1029/2005JD006259.

Gray, S. L., C. M. Dunning, J. Methven, G. Masato, and J. M. Chagnon, 2014: Systematic model forecast error in Rossby wave structure. *Geophysical Research Letters*, **41 (8)**, 2979–2987, doi:10.1002/2014GL059282.

Gutowski, W. J., L. E. Branscome, and D. A. Stewart, 1992: Life cycles of moist baroclinic eddies. *Journal of the Atmospheric Sciences*, **49 (4)**, 306–319, doi:10.1175/1520-0469(1992)049<0306:LCOMBE>2.0.CO;2.

Hakim, G. J., 2005: Vertical structure of midlatitude analysis and forecast errors. *Monthly Weather Review*, **133 (3)**, 567–578, doi:10.1175/MWR-2882.1.

Harvey, B. J., J. Methven, and M. H. P. Ambaum, 2016: Rossby wave propagation on potential vorticity fronts with finite width. *Journal of Fluid Mechanics*, **794**, 775–797, doi:10.1017/jfm.2016.180.

Heifetz, E., C. H. Bishop, B. J. Hoskins, and J. Methven, 2004: The counter-propagating Rossby-wave perspective on baroclinic instability. I: Mathematical basis. *Quarterly Journal of the Royal Meteorological Society*, **130 (596)**, 211–231, doi:10.1256/qj.02.184.

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, **15 (5)**, 559–570, doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bulletin of the American Meteorological Society*, **88 (11)**, 1783–1793, doi:10.1175/BAMS-88-11-1783.

Holton, J., and G. Hakim, 2013: *An Introduction to Dynamic Meteorology*. Academic Press, Elsevier Science, https://books.google.de/books?id=-ePQ6x6VbjgC.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society*, **111 (470)**, 877–946, doi:10.1256/smsqj.47001.

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. Technical memorandum 636, ECMWF, 45 pp. doi:10.21957/obke4k60.

Judt, F., 2018: Insights into atmospheric predictability through global convection-permitting model simulations. *Journal of the Atmospheric Sciences*, **75 (5)**, 1477–1497, doi:10.1175/JAS-D-17-0343.1.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, doi:10.1017/CBO9780511802270.

Kumpf, A., M. Rautenhaus, M. Riemer, and R. Westermann, 2019: Visual analysis of the temporal evolution of ensemble forecast sensitivities. *IEEE Transactions on Visualization and Computer Graphics*, **25 (1)**, 98–108, doi:10.1109/TVCG.2018.2864901.

Kumpf, A., B. Tost, M. Baumgart, M. Riemer, R. Westermann, and M. Rautenhaus, 2018: Visualizing confidence in cluster-based ensemble weather forecast analyses. *IEEE Transactions on Visualization and Computer Graphics*, **24 (1)**, 109–119, doi:10.1109/TVCG.2017.2745178.

Langland, R. H., M. A. Shapiro, and R. Gelaro, 2002: Initial condition sensitivity and error growth in forecasts of the 25 January 2000 east coast snowstorm. *Monthly Weather Review*, **130 (4)**, 957–974, doi:10.1175/1520-0493(2002)130<0957:ICSAEG>2.0.CO;2.

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics*, **227 (7)**, 3515 – 3539, doi:10.1016/j.jcp.2007.02.014.

Liniger, M. A., and H. C. Davies, 2004: Seasonal differences in extratropical potential vorticity variability at tropopause levels. *Journal of Geophysical Research: Atmospheres*, **109 (D17)**, doi:10.1029/2004JD004639.

Lorenz, E., 1995: Predictability: a problem partly solved. *Seminar on Predictability, 4-8 September 1995*, ECMWF, Shinfield Park, Reading, Vol. 1, https://www.ecmwf.int/node/10829.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20 (2)**, 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21 (3)**, 289–307, doi:10.1111/j.2153-3490.1969.tb00444.x.

Lynch, P., 1989: Partitioning the wind in a limited domain. *Monthly Weather Review*, **117 (7)**, 1492–1500, doi:10.1175/1520-0493(1989)117<1492:PTWIAL>2.0.CO;2.

Magnusson, L., 2017: Diagnostic methods for understanding the origin of forecast errors. *Quarterly Journal of the Royal Meteorological Society*, **143 (706)**, 2129–2142, doi:10.1002/qj.3072.

Martínez-Alvarado, O., H. Joos, J. Chagnon, M. Boettcher, S. L. Gray, R. S. Plant, J. Methven, and H. Wernli, 2014: The dichotomous structure of the warm conveyor belt. *Quarterly Journal of the Royal Meteorological Society*, **140 (683)**, 1809–1824, doi:10.1002/qj.2276.

Martínez-Alvarado, O., E. Madonna, S. L. Gray, and H. Joos, 2016: A route to systematic error in forecasts of Rossby waves. *Quarterly Journal of the Royal Meteorological Society*, **142 (694)**, 196–210, doi:10.1002/qj.2645.

Martius, O., C. Schwierz, and H. C. Davies, 2010: Tropopause-level waveguides. *Journal of the Atmospheric Sciences*, **67 (3)**, 866–879, doi:10.1175/2009JAS2995.1.

McTaggart-Cowan, R., J. R. Gyakum, and M. K. Yau, 2001: Sensitivity testing of extratropical transitions using potential vorticity inversions to modify initial conditions: Hurricane Earl case study. *Monthly Weather Review*, **129 (7)**, 1617–1636, doi:10.1175/1520-0493(2001)129<1617:STOETU>2.0.CO;2.

Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *Journal of the Atmospheric Sciences*, **69 (11)**, 3350–3371, doi:10.1175/JAS-D-11-0315.1.

Moncrieff, M. W., D. E. Waliser, and J. Caughey, 2012: Progress and direction in tropical convection research: YOTC international science symposium. *Bulletin of the American Meteorological Society*, **93 (8)**, ES65–ES69, doi:10.1175/BAMS-D-11-00253.1.

Nakamura, N., and A. Solomon, 2011: Finite-amplitude wave activity and mean flow adjustments in the atmospheric general circulation. Part II: Analysis in the isentropic coordinate. *Journal of the Atmospheric Sciences*, **68 (11)**, 2783–2799, doi:10.1175/2011JAS3685.1.

Nie, Y., Y. Zhang, G. Chen, and X. Yang, 2016: Delineating the barotropic and baroclinic mechanisms in the midlatitude eddy-driven jet response to lower-tropospheric thermal forcing. *Journal of the Atmospheric Sciences*, **73 (1)**, 429–448, doi:10.1175/JAS-D-15-0090.1.

Nielsen-Gammon, J. W., and R. J. Lefevre, 1996: Piecewise tendency diagnosis of dynamical processes governing the development of an upper-tropospheric mobile trough. *Journal of the Atmospheric Sciences*, **53 (21)**, 3120–3142, doi:10.1175/1520-0469(1996)053<3120:PTDODP>2.0.CO;2.

Oertel, A., M. Boettcher, H. Joos, M. Sprenger, H. Konow, M. Hagen, and H. Wernli, 2019: Convective activity in an extratropical cyclone and its warm conveyor belt - a case study combining observations and a convection-permitting model simulation. *Quarterly Journal of the Royal Meteorological Society*, **145 (721)**, 1406–1426, doi:10.1002/qj.3500.

Palmer, T., and R. Hagedorn, 2006: *Predictability of Weather and Climate*. Cambridge University Press, https://books.google.de/books?id=c-rY28QQCj8C.

Pantillon, F., J. P. Chaboureau, C. Lac, and P. Mascart, 2013: On the role of a Rossby wave train during the extratropical transition of hurricane Helene (2006). *Quarterly Journal of the Royal Meteorological Society*, **139 (671)**, 370–386, doi:10.1002/qj.1974.

Piaget, N., P. Froidevaux, P. Giannakaki, F. Gierth, O. Martius, M. Riemer, G. Wolf, and C. M. Grams, 2015: Dynamics of a local Alpine flooding event in October 2011: moisture source and large-scale circulation. *Quarterly Journal of the Royal Meteorological Society*, **141 (690)**, 1922–1937, doi:10.1002/qj.2496.

Plant, R. S., and G. C. Craig, 2008: A stochastic parameterization for deep convection based on equilibrium statistics. *Journal of the Atmospheric Sciences*, **65 (1)**, 87–105, doi:10.1175/2007JAS2263.1.

Plu, M., and P. Arbogast, 2005: A cyclogenesis evolving into two distinct scenarios and its implications for short-term ensemble forecasting. *Monthly Weather Review*, **133 (7)**, 2016–2029, doi:10.1175/MWR2955.1.

Quinting, J. F., and S. C. Jones, 2016: On the impact of tropical cyclones on Rossby wave packets: A climatological perspective. *Monthly Weather Review*, **144 (5)**, 2021–2048, doi:10.1175/MWR-D-14-00298.1.

Richardson, L., 1922: *Weather prediction by numerical process*. University Press, https://books.google.de/books?id=cWXWhffnUokC.

Riemer, M., M. Baumgart, and S. Eiermann, 2014: Cyclogenesis downstream of extratropical transition analyzed by Q-vector partitioning based on flow geometry. *Journal of the Atmospheric Sciences*, **71 (11)**, 4204–4220, doi:10.1175/JAS-D-14-0023.1.

Riemer, M., and S. C. Jones, 2010: The impact of tropical cyclones on a developing baroclinic wave in idealized scenarios of extratropical transition. *Quarterly Journal of the Royal Meteorological Society*, **136 (648)**, 617–637, doi:10.1002/qj.605.

Riemer, M., and S. C. Jones, 2014: Interaction of a tropical cyclone with a high-amplitude, midlatitude wave pattern: Waviness analysis, trough deformation and track bifurcation. *Quarterly Journal of the Royal Meteorological Society*, **140 (681)**, 1362–1376, doi:10.1002/qj.2221.

Riemer, M., S. C. Jones, and C. A. Davis, 2008: The impact of extratropical transition on the downstream flow: An idealized modelling study with a straight jet. *Quarterly Journal of the Royal Meteorological Society*, **134 (630)**, 69–91, doi:10.1002/qj.189.

Rodwell, M. J., S. T. K. Lang, N. B. Ingleby, N. Bormann, E. Hólm, F. Rabier, D. S. Richardson, and M. Yamaguchi, 2016: Reliability in ensemble data assimilation. *Quarterly Journal of the Royal Meteorological Society*, **142 (694)**, 443–454, doi:10.1002/qj.2663.

Rodwell, M. J., D. S. Richardson, D. B. Parsons, and H. Wernli, 2018: Flow-dependent reliability: A path to more skillful ensemble forecasts. *Bulletin of the American Meteorological Society*, **99 (5)**, 1015–1026, doi:10.1175/BAMS-D-17-0027.1.

Rodwell, M. J., and Coauthors, 2013: Characteristics of occasional poor medium-range weather forecasts for Europe. *Bulletin of the American Meteorological Society*, **94 (9)**, 1393–1405, doi:10.1175/BAMS-D-12-00099.1.

Saffin, L., S. L. Gray, J. Methven, and K. D. Williams, 2017: Processes maintaining tropopause sharpness in numerical models. *Journal of Geophysical Research: Atmospheres*, **122 (18)**, 9611–9627, doi:10.1002/2017JD026879.

Saffin, L., J. Methven, and S. L. Gray, 2016: The non-conservation of potential vorticity by a dynamical core compared with the effects of parametrized physical processes. *Quarterly Journal of the Royal Meteorological Society*, **142 (696)**, 1265–1275, doi:10.1002/qj.2729.

Schäfler, A., and Coauthors, 2018: The North Atlantic Waveguide and Downstream Impact Experiment. *Bulletin of the American Meteorological Society*, **99 (8)**, 1607–1637, doi:10.1175/BAMS-D-17-0003.1.

Schneidereit, A., and Coauthors, 2017: Enhanced tropospheric wave forcing of two anticyclones in the prephase of the January 2009 major stratospheric sudden warming event. *Monthly Weather Review*, **145 (5)**, 1797–1815, doi:10.1175/MWR-D-16-0242.1.

Selz, T., 2019: Estimating the intrinsic limit of predictability using a stochastic convection scheme. *Journal of the Atmospheric Sciences*, **76 (3)**, 757–765, doi:10.1175/JAS-D-17-0373.1.

Selz, T., and G. C. Craig, 2015a: Simulation of upscale error growth with a stochastic convection scheme. *Geophysical Research Letters*, **42 (8)**, 3056–3062, doi:10.1002/2015GL063525.

Selz, T., and G. C. Craig, 2015b: Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Monthly Weather Review*, **143 (3)**, 813–827, doi:10.1175/MWR-D-14-00140.1.

Slingo, J., and T. Palmer, 2011: Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **369 (1956)**, 4751–4767, doi:10.1098/rsta.2011.0161.

Snyder, C., 1999: Error growth in flows with finite-amplitude waves or coherent structures. *Journal of the Atmospheric Sciences*, **56 (4)**, 500–506, doi:10.1175/1520-0469(1999)056<0500:EGIFWF>2.0.CO;2.

Snyder, C., T. M. Hamill, and S. B. Trier, 2003: Linear evolution of error covariances in a quasigeostrophic model. *Monthly Weather Review*, **131 (1)**, 189–205, doi:10.1175/1520-0493(2003)131<0189:LEOECI>2.0.CO;2.

Stoelinga, M. T., 1996: A potential vorticity-based study of the role of diabatic heating and friction in a numerically simulated baroclinic cyclone. *Monthly Weather Review*, **124 (5)**, 849–874, doi:10.1175/1520-0493(1996)124<0849:APVBSO>2.0.CO;2.

Sun, Y. Q., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *Journal of the Atmospheric Sciences*, **73 (3)**, 1419–1438, doi:10.1175/JAS-D-15-0142.1.

Teubler, F., and M. Riemer, 2016: Dynamics of Rossby wave packets in a quantitative potential vorticity–potential temperature framework. *Journal of the Atmospheric Sciences*, **73 (3)**, 1063–1081, doi:10.1175/JAS-D-15-0162.1.

Thorpe, A. J., and C. H. Bishop, 1995: Potential vorticity and the electrostatics analogy: Ertel—Rossby formulation. *Quarterly Journal of the Royal Meteorological Society*, **121 (526)**, 1477–1495, doi:10.1002/qj.49712152612.

Tribbia, J. J., and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Monthly Weather Review*, **132**, 703–713, doi:10.1175/1520-0493(2004)132<0703:SIAAPA>2.0.CO;2.

Whitaker, J. S., and A. F. Loughe, 1998: The relationship between ensemble spread and ensemble mean skill. *Monthly Weather Review*, **126 (12)**, 3292–3302, doi:10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2.

Wiegand, L., and P. Knippertz, 2014: Equatorward breaking Rossby waves over the North Atlantic and Mediterranean region in the ECMWF operational Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society*, **140 (678)**, 58–71, doi:10.1002/qj.2112.

Wirth, V., M. Riemer, E. K. M. Chang, and O. Martius, 2018: Rossby wave packets on the midlatitude waveguide—a review. *Monthly Weather Review*, **146 (7)**, 1965–2001, doi: 10.1175/MWR-D-16-0483.1.

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, **141 (687)**, 563–579, doi:10.1002/qj.2378.

Zhang, F., N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *Journal of the Atmospheric Sciences*, **64 (10)**, 3579–3594, doi:10.1175/JAS4028.1.

Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *Journal of the Atmospheric Sciences*, **60 (9)**, 1173–1185, doi:10.1175/1520-0469(2003)060<1173:EOMCOM>2.0.CO;2.

Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *Journal of the Atmospheric Sciences*, **76 (4)**, 1077–1091, doi:10.1175/JAS-D-18-0269.1.

Zimin, A. V., I. Szunyogh, D. Patil, B. R. Hunt, and E. Ott, 2003: Extracting envelopes of Rossby wave packets. *Monthly Weather Review*, **131 (5)**, 1011–1017, doi: 10.1175/1520-0493(2003)131<1011:EEORWP>2.0.CO;2.