

Perspectives on warm conveyor belts - sensitivities to ensemble configuration and the role for forecast error

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M.Sc. Moritz Pickl
aus Regensburg

Tag der mündlichen Prüfung:	01.07.2022
Referent:	Jun.-Prof. Dr. Christian M. Grams
Korreferentin:	Prof. Dr. Corinna Hoose

Abstract

Warm conveyor belts (WCBs) are weather systems that substantially modulate the large-scale extratropical circulation. As they can amplify forecast errors and project them onto the Rossby wave pattern, they are of high relevance for numerical weather prediction. At the same time, the ascending motion of WCBs that transports air masses from the lower to the upper troposphere is strongly driven by latent heat release from cloud-condensational processes, whose representation in forecast models is prone to uncertainties. This thesis elaborates on two aspects of WCBs in the context of ensemble forecasts: (1) sensitivities of WCBs to the representation of initial condition and model uncertainties, and (2) the role of WCBs for forecast error growth.

Ensemble prediction systems account for forecast errors related to physical parametrizations by running multiple integrations of the forecast model with perturbed model physics. The stochastically perturbed parametrization tendencies (SPPT) scheme, a well-established technique to represent model uncertainty, introduces perturbations into the tendencies provided by the physical parametrizations. The amplitude of the local perturbations is proportional to the magnitude of the tendencies from the parametrization schemes, which are typically large in the ascent regions of WCBs. The first part of this thesis investigates the impact of the SPPT-scheme and other model uncertainty representations on diabatically driven, rapidly ascending air streams. We use the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) and perform a set of sensitivity experiments with different ensemble configurations to disentangle the impact of initial condition and model perturbations on rapidly ascending air streams, which are identified with trajectory analysis. Despite its zero-mean design, SPPT results in a systematic increase of the frequency of rapidly ascending air streams, without changing the physical properties of the trajectories. The magnitude of the effect depends on the integrated latent heat release along the air streams and is therefore more pronounced in the tropics than in the midlatitudes. A Eulerian perspective on the distribution of vertical velocities reveals that SPPT increases the occurrence frequency of strong upward motions, which is balanced by accelerated downward motions. In contrast to SPPT, perturbations of the initial conditions do not result in frequency changes of rapidly ascending air streams. This insight is used to substantiate the findings from the sensitivity experiments by comparing the perturbed and unperturbed forecasts of operational ECMWF ensemble forecasts across a large set of forecast initializations in different seasons. Experiments with two other model uncertainty schemes suggest that the impacts on the frequency of rapidly ascending air streams mainly result from the perturbations to the physical parametrizations, whereas perturbations to the dynamical core of the

forecast model have only minor impacts on the vertical velocities. Based on these results, a mechanism is introduced how stochastic, zero-mean perturbations can result in a unilateral effect on the frequency occurrence of rapidly ascending air streams. We hypothesize that symmetric perturbations can result in biased responses when they are applied to nonlinear systems that are characterized by a threshold, such that perturbations in one direction are more effective in triggering a process than perturbations of identical amplitude, but of opposite sign can suppress it.

WCBs and other rapidly ascending air streams are closely linked to the formation of precipitation and to the evolution of the large-scale extratropical flow. The previous results therefore raise the question whether the impact of the stochastic perturbations on the ascending motions is reflected in these weather phenomena. We show that stochastic model perturbations modulate the global precipitation distribution consistently with their impact on the vertical velocities and result in a shift of the distribution towards higher values. Also the impact on the large-scale flow in the midlatitudes is in accordance with the increased frequency of WCBs. Experiments with stochastic model uncertainty schemes are characterized by an increased amplitude of the upper-level Rossby wave pattern and by a poleward shift of the tropopause compared to forecasts with an unperturbed model. The results are corroborated by comparing the waviness of the dynamical tropopause of perturbed and unperturbed forecasts in a large reforecast data. The consistency of the modulation of the vertical velocities and the Rossby wave amplitude across different schemes and seasons suggests that WCBs project the unilateral effect of the perturbations onto the large-scale circulation. Although the magnitude of the effect is relatively small, it illustrates the role of WCBs in communicating signals across different scales and vertical levels.

The second overarching aspect of this thesis is the systematic investigation of the role of WCBs for errors in operational forecasts. By exploiting a unique data set of WCB trajectories in the North Atlantic domain in operational ECMWF ensemble forecasts initialized in three winter seasons, the analysis attempts to isolate the impact of WCBs on the degradation of forecast skill and to corroborate the findings of previous studies that approached the topic on a case study basis. We find that forecast errors are climatologically co-located with regions where WCBs occur, and forecasts that are characterized by high WCB activity have on average lower skill than forecasts with low WCB activity. The forecast time when the error growth over the North Atlantic domain is largest is characterized by anomalously high WCB activity. Composites of normalized forecast errors centered on WCB objects reveal that WCBs are associated with characteristic spatio-temporal patterns of increased forecast errors. The error patterns in the mid-troposphere are related to an upstream trough and a ridge developing downstream, but feature a large case-to-case variability. In the upper troposphere, in contrast, the degradation of the forecast skill is robust across many cases and is associated with the jet streak on the northern flank of the upper-level ridge. This analysis provides evidence that WCBs are involved in the growth and amplification of forecast errors and emphasizes the need for the improvement of their representation for skillful forecasts.

Kurzfassung

Warmluftförderbänder (engl. warm conveyor belts, WCBs) sind Wettersysteme, die einen erheblichen Einfluss auf die großräumige Zirkulation in den Außertropen ausüben. Da sie Vorhersagefehler verstärken und auf das Rossby-Wellenmuster projizieren können, sind sie von großer Bedeutung für numerische Wettervorhersagen. Gleichzeitig wird der Aufstieg von Luftmassen in WCBs von der unteren in die obere Troposphäre stark von der Freisetzung von latenter Wärme durch Wolkenkondensationsprozesse angetrieben, deren Darstellung in Vorhersagemodellen mit Unsicherheiten behaftet ist. In der vorliegenden Arbeit werden zwei Aspekte von WCBs im Zusammenhang mit Ensemblevorhersagen näher beleuchtet: (1) Sensitivitäten von WCBs auf die Darstellung von Unsicherheiten der Anfangsbedingungen und des Vorhersagemodells, und (2) die Rolle von WCBs für das Wachstum von Vorhersagefehlern.

Ensemble-Vorhersagesysteme berücksichtigen Vorhersagefehler, die mit physikalischen Parametrisierungen zusammenhängen, indem mehrere Vorhersageläufe mit gestörter Modellphysik durchgeführt werden. Das “stochastically perturbed parametrization tendencies“ (SPPT) Schema, eine bewährte Methode zur Darstellung von Modellunsicherheiten, führt Störungen in die von den physikalischen Parametrisierungen berechneten Tendenzen ein. Die Amplitude der lokalen Störungen ist proportional zum Betrag der Tendenzen aus den Parametrisierungsschemata, der in den Aufstiegsregionen von WCBs typischerweise groß ist. Im ersten Teil dieser Arbeit werden die Auswirkungen des SPPT-Schemas und anderer Methoden der Modellunsicherheitsdarstellung auf diabatisch getriebene, schnell aufsteigende Luftströmungen untersucht. Wir verwenden das Integrated Forecasting System (IFS) des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) und führen eine Reihe von Sensitivitätsexperimenten mit verschiedenen Ensemble-Konfigurationen durch, um die Auswirkungen von Anfangsbedingungsstörungen und Modellstörungen auf schnell aufsteigende Luftströmungen zu unterscheiden. Die aufsteigenden Luftströmungen werden mithilfe von Trajektorien detektiert. Trotz seines symmetrischen, null-zentrierten Designs führt das SPPT-Schema zu einer systematischen Erhöhung der Häufigkeit schnell aufsteigender Luftströmungen, ohne dass die physikalischen Eigenschaften der Trajektorien verändert werden. Das Ausmaß dieses Effekts hängt von der integrierten latenten Heizrate entlang der Luftströmungen ab und ist daher in den Tropen stärker ausgeprägt als in den mittleren Breiten. Eine Eulersche Perspektive auf die Verteilung der Vertikalgeschwindigkeiten zeigt, dass SPPT die Häufigkeit starker Aufwärtsbewegungen erhöht, was durch beschleunigte Abwärtsbewegungen ausgeglichen wird. Im Gegensatz zu SPPT führen Störungen der Anfangsbedingungen nicht zur Änderung der Häufigkeiten schnell aufsteigender Luftströmungen. Diese Erkenntnis wird genutzt, um die Ergebnisse der Sensitivitätsexpe-

rimente zu untermauern, indem die gestörten und ungestörten Vorhersagen von operationellen ECMWF-Ensemblevorhersagen über eine große Anzahl von einzelnen Vorhersagen in verschiedenen Jahreszeiten verglichen werden. Experimente mit zwei anderen Modellunsicherheitsschemata deuten darauf hin, dass die Auswirkungen auf das Auftreten schnell aufsteigender Luftströmungen hauptsächlich auf die Störungen der physikalischen Parametrisierungen zurückzuführen sind, während Störungen des dynamischen Kerns des Modells nur geringe Auswirkungen auf die Vertikalgeschwindigkeiten haben. Basierend auf diesen Ergebnissen wird ein Mechanismus vorgestellt, wie stochastische, null-zentrierte Störungen zu einem einseitigen Effekt auf die Häufigkeit des Auftretens schnell aufsteigender Luftströmungen führen können. Wir stellen die Hypothese auf, dass symmetrische Störungen zu verzerrten Reaktionen führen können, wenn sie auf nichtlineare Systeme angewandt werden, die durch einen Schwellenwert gekennzeichnet sind, sodass Störungen in eine Richtung effektiver darin sind, einen Prozess auszulösen als Störungen gleicher Amplitude, aber mit entgegengesetztem Vorzeichen den Prozess unterdrücken können.

WCBs und andere schnell aufsteigende Luftströmungen sind eng mit der Bildung von Niederschlag und der Entwicklung der großräumigen Strömung in den mittleren Breiten verbunden. Die bisherigen Ergebnisse der Arbeit werfen daher die Frage auf, ob sich die Auswirkungen der stochastischen Störungen auf die aufsteigenden Trajektorien in diesen Wetterphänomenen widerspiegeln. Wir zeigen, dass stochastische Modellstörungen die globale Niederschlagsverteilung konsistent mit ihren Auswirkungen auf die Vertikalengeschwindigkeiten modulieren und zu einer Verschiebung der Verteilung hin zu höheren Werten führen. Auch die Auswirkungen auf die großräumige Strömung in den mittleren Breiten stehen im Einklang mit der erhöhten Häufigkeit von WCBs. Im Vergleich zu Vorhersagen mit einem ungestörten Modell sind Experimente, die mit stochastischen Modellunsicherheitsschemata gestört werden, durch eine erhöhte Amplitude des Rossby-Wellenmusters in der oberen Troposphäre und durch eine polwärtige Verschiebung der Tropopause gekennzeichnet. Die Ergebnisse werden durch die Auswertung der Rossbywellenamplitude in einem großen Reforecast-Datensatz mit gestörten und ungestörten Simulationen bestätigt. Die Konsistenz der modifizierten Vertikalengeschwindigkeiten und der Rossbywellenamplitude über verschiedene Schemata und Jahreszeiten hinweg deutet darauf hin, dass die WCBs den einseitigen Effekt der stochastischen Störungen auf die großräumige Zirkulation projizieren. Obwohl die Größenordnung dieses Effekts relativ gering ist, verdeutlicht er die Rolle von WCBs bei der Kommunikation von Signalen zwischen verschiedenen räumlichen Skalen und vertikalen Leveln.

Der zweite übergreifende Aspekt dieser Arbeit ist die systematische Untersuchung der Rolle von WCBs für das Wachstum von Vorhersagefehlern. Durch die Nutzung eines einzigartigen Datensatzes von WCB-Trajektorien im Nordatlantik in operationellen ECMWF-Ensemblevorhersagen, die in drei Wintern initialisiert wurden, zielt die Analyse darauf ab, die negativen Auswirkungen von WCBs auf die Vorhersagequalität zu isolieren und die Ergebnisse früherer Studien zu bestätigen, die sich des Themas auf

der Basis von Fallstudien angenommen haben. Wir stellen fest, dass Vorhersagefehler klimatologisch mit Regionen zusammenhängen, in denen WCBs häufig auftreten, und dass Vorhersagen, die durch eine hohe WCB-Aktivität gekennzeichnet sind, im Durchschnitt eine geringere Qualität aufweisen als Vorhersagen mit geringer WCB-Aktivität. Die Zeit in der Vorhersage, zu der das Fehlerwachstum über dem Nordatlantik am größten ist, ist durch anomal hohe WCB-Aktivität gekennzeichnet. Composites von normalisierten Vorhersagefehlern, die auf WCB-Objekte zentriert sind, zeigen, dass WCBs mit charakteristischen raum-zeitlichen Mustern erhöhter Vorhersagefehler verbunden sind. Die Fehlermuster in der mittleren Troposphäre hängen mit einem stromaufwärts gelegenen Trog und einem sich stromabwärts entwickelnden Rücken zusammen, weisen aber eine große Variabilität von Fall zu Fall auf. In der oberen Troposphäre hingegen ist die Verschlechterung der Vorhersagegüte über viele Fälle robust und steht hauptsächlich mit erhöhten Windgeschwindigkeiten an der Nordflanke des Höhenrückens in Zusammenhang. Diese Analyse liefert Hinweise dafür, dass WCBs am Wachstum und der Amplifizierung von Vorhersagefehlern beteiligt sind, und unterstreicht die Notwendigkeit einer Verbesserung ihrer Darstellung.

Preface

The PhD candidate confirms that the research presented in this thesis contains significant scientific contributions by himself. This thesis, in particular the Abstract and Chapters 4 and 5, reuses material from the following publication:

Pickl, M., S. T. Lang, M. Leutbecher, and C. M. Grams, 2022: The effect of stochastically perturbed parametrisation tendencies (SPPT) on rapidly ascending air streams. *Quarterly Journal of the Royal Meteorological Society*, **148 (744)**, 1242–1261

The research leading to the results was performed within the Young Investigator Group “Sub-seasonal Predictability: Understanding the Role of Diabatic Outflow” (SPREADOUT), funded by the Helmholtz Association under the grant VH-NG-1243. The research proposal for the Young Investigator Group was written by Jun.-Prof. Dr. Christian M. Grams.

The IFS ensemble simulations for Chapters 4 and 5 and the implementation of the automated postprocessing suite were performed by the candidate, with advice from Dr. Simon T. K. Lang and Dr. Sarah-Jane Lock from the European Centre for Medium-Range Weather Forecasts (ECMWF). The real-time data retrieval of operational ECMWF ensemble forecasts and the computation of the trajectory data used in Chapters 4 and 6 is done by Prof. Dr. Christian Grams. The analyses in Pickl et al. (2022) were solely performed by the candidate, who also wrote the manuscript, with advice from Prof. Dr. Christian M. Grams, Dr. Simon T. K. Lang and Dr. Martin Leutbecher (ECMWF).

The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that this is copyright material and that no quotation from the thesis may be published without proper acknowledgment.

©2022, Karlsruhe Institute of Technology and Moritz Pickl

Contents

1	Introduction	1
2	Background and Research Questions	5
2.1	Ensemble forecasting	5
2.1.1	Historic perspective	5
2.1.2	Rationale	6
2.1.3	Stochastic parametrization	7
2.2	Warm conveyor belts (WCBs)	10
2.2.1	WCBs and the extratropical circulation	10
2.2.2	The potential vorticity framework	13
2.2.3	Large-scale flow modification of WCBs	15
2.2.4	The relevance of WCBs for numerical weather prediction	16
2.3	Research questions	19
3	Data and Methods	21
3.1	ECMWF's ensemble prediction system	21
3.1.1	Integrated Forecasting System (IFS)	21
3.1.2	Uncertainty representation and ensemble configuration	23
3.2	Data sets	27
3.2.1	Experiments	27
3.2.2	Forecast archives	31
3.2.3	Verification data sets	32
3.3	Lagrangian WCB detection	33
3.3.1	Lagranto	33
3.3.2	Procedure for WCB-detection	34
3.3.3	Automated / Real-time implementation	36
4	The impact of ensemble configuration on rapidly ascending air streams	39
4.1	Sensitivities to operational schemes	40
4.1.1	Lagrangian perspective	40
4.1.2	Eulerian omega-perspective	49
4.1.3	Exploiting operational forecasts	54

4.2	Sensitivities to other model uncertainty schemes	59
4.2.1	Lagrangian perspective	60
4.2.2	Eulerian perspective	65
4.3	Mechanism	67
4.4	Summary and Discussion	72
5	Impacts through stochastic perturbations	75
5.1	Precipitation	75
5.1.1	Operational schemes	75
5.1.2	Other schemes	79
5.1.3	Discussion	80
5.2	Rosby wave amplitude	81
5.2.1	Methodology	81
5.2.2	Operational schemes	82
5.2.3	Exploiting forecast archives	89
5.2.4	Other schemes	92
5.2.5	Discussion	94
5.3	Summary and Outlook	97
6	The role of WCBs for forecast error growth	99
6.1	Data and Methods	100
6.2	Spatial co-occurrence of WCBs and forecast error	103
6.3	Temporal relations of WCBs and forecast error	105
6.4	Perspective on WCB objects	111
6.5	Concluding discussion	116
7	Conclusions	121
A	Appendix for Chapter 3	127
B	Appendix for Chapter 4	129
C	Appendix for Chapter 5	131
D	Appendix for Chapter 6	133
E	Bibliography	139

1. Introduction

Since the beginning of modern numerical weather prediction (NWP), the performance of forecasts has steadily improved. Bauer et al. (2015) describe this evolution as "the quiet revolution of numerical weather prediction", because the forecast improvements cannot be attributed to fundamental breakthroughs, but are rather a result of many incremental steps. Within the last 40 years, a combination of scientific and technological progress enabled a shift of the lead time until deterministic forecasts are considered skillful by about one day per decade. The most important scientific developments along this evolution include, but are not limited to, a better understanding of the physical processes that are represented by the models, improvements in the numerical methods to solve the governing equations, and the development of sophisticated data assimilation techniques to estimate the initial state (Magnusson and Källén, 2013). This progress is fueled by technological advancements, first and foremost by the ever increasing computational resources that allow higher and higher model resolutions, but also by the exploitation of satellite-based products or through the application of machine-learning tools.

A milestone in the evolution of modern NWP was the recognition that forecasts of nonlinear systems are subject to intrinsic limits of predictability: founded on the seminal work on chaos theory by Lorenz (1963), the need for a probabilistic approach that represents uncertainties in the initial estimate of the atmospheric state and in the forecast model became evident. Nevertheless, it took almost 30 more years until the first operational ensemble forecast was issued by the European Centre for Medium-Range Weather Forecasts (ECMWF) in 1992 (Palmer, 2019a). Since then, the operationally issued probabilistic forecasts allow for a quantification of the forecast uncertainty and helped to extend the forecast horizon into (sub-) seasonal time scales. The probabilistic forecasts are commonly produced by running several integrations of the forecast model from slightly perturbed initial conditions and/or with perturbations to the model. The techniques for both the initial condition as well as for the model perturbations have been refined within the last decades, resulting in increasingly accurate estimates of the forecast uncertainty (Leutbecher et al., 2017).

Despite this significant progress, state-of-the-art weather prediction models still suffer from occasional poor forecasts of the large-scale atmospheric flow in the extratropics. It has been shown that such "forecast busts" are often linked to the onset of blocking anticyclones over Europe (Rodwell et al., 2013). Large-scale circulation patterns associated with blockings in the Euro-Atlantic sector tend to be less predictable than zonal flow configurations (Ferranti et al., 2015; Büeler et al., 2021), and Quinting and Vitart

(2019) demonstrate that current NWP-models systematically underestimate the occurrence frequency of blocking anticyclones. In the past few years, several studies pointed towards an important contribution of latent heat release in moist ascending air streams, so-called warm conveyor belts (WCB), to the onset and maintenance of blockings (e.g. Pfahl et al. 2015; Steinfeld and Pfahl 2019). Due to their transport of air masses from the lower to the upper troposphere, WCBs have the potential to modify the upper-level flow such that ridges are built or amplified, thereby serving as a communicator between the height levels and horizontal scales (Grams et al., 2011; Chagnon et al., 2013). However, the diabatic processes responsible for the latent heat release along the ascending WCBs are parametrized in NWP models, which poses a source of uncertainty for the subsequent forecast evolution (Grams et al., 2018). It is therefore hypothesized that deficiencies in the representation of diabatic processes in WCBs are related to systematic forecast errors of the upper-level flow (Gray et al., 2014).

Typically, the identification of WCBs in gridded data sets uses Lagrangian methods (Wernli and Davies, 1997) that require a high spatio-temporal resolution - this makes their systematic evaluation in (ensemble) forecasts very challenging, because the resources needed for the data storage and for the computations are immense. This is why most studies that investigate WCBs in the context of forecast uncertainty are either based on individual case studies (e.g. Grams et al. 2018; Berman and Torn 2019), or use Eulerian proxies to quantify the imprint of WCBs (e.g. Sánchez et al. 2020). Just recently, novel methods using artificial intelligence have been made available that allow for an accurate detection of WCBs also with a reduced data basis (Quinting and Grams, 2021, 2022), enabling a first systematic evaluation of WCBs in ensemble reforecasts (Wandel et al., 2021). Despite many new approaches emerging in the past few years, research on WCBs in ensemble forecasts is still in its infancy, and many aspects have not yet been covered.

A close collaboration with colleagues at ECMWF enabled a detailed investigation of WCBs in the context of ensemble forecasts in the framework of this thesis project, which employs two different perspectives:

1. The first perspective focuses on how WCBs and other moist, rapidly ascending air streams such as tropical convection are represented in ensemble forecasts that are perturbed with different techniques. The main attention lies on the assessment of the schemes that are used operationally at ECMWF, but also schemes that are currently under development are analyzed (Chapter 4). To its end, this process-oriented perspective on ensemble configuration will help to foster the understanding how uncertainty schemes impact the model climate (Chapter 5).
2. The analysis in the second part of the thesis flips the point of view and attempts to quantify the impact that WCBs exert on the performance of ensemble forecasts in a systematic way (Chapter 6).

Before these aspects are worked out in detail, Chapter 2 will give background information on ensemble forecasting and on WCBs and will assess the current state of research, from which the research goals of the thesis will be derived. Chapter 3 introduces the data sets and methods that are used throughout the thesis, and the results are presented in chapters 4, 5 and 6. Finally, the overarching findings of the thesis will be discussed in chapter 7.

2. Background and Research Questions

This chapter gives an overview of the basic concepts of ensemble forecasting (Section 2.1) and of warm conveyor belts (Section 2.2) that are important for the understanding of the subsequent chapters of the thesis. Both sections introduce some historical aspects and discuss recent scientific developments in the fields. To the end of the chapter, the topics are linked and the main research questions of the thesis are formulated (Section 2.3).

2.1. Ensemble forecasting

2.1.1. Historic perspective

Based on Newton's equations of motion, numerical weather prediction was considered a purely deterministic problem in its early days in the 1950s (Lewis, 2005). The first successful NWP forecast goes back to Charney et al. (1950), who simplified the Navier-Stokes equations using the quasi-geostrophic (QG) assumptions and produced a first successful 24-hour forecast of the large-scale flow. After a decade of scientific progress in the field, such as for example the exploration of the Primitive Equations which eased the implementation of moist processes and turbulence, it took a byproduct from a scientific experiment to question the deterministic principle of NWP (Lewis, 2005): in a study that aimed at disproving the adequacy of using linear regression methodologies to forecast the nonlinear evolution of the atmospheric flow, Lorenz (1963) rounded off intermediate output of a low-order QG-model for convenience and found that the subsequent forecast was drastically different from the one using the full precision. Lorenz quickly recognized the relevance of his findings and stated that "if the atmosphere behaved this way, then long-range forecasting was impossible because we certainly don't measure things as accurately as that" (Thompson and Lorenz, 1986)¹. He formulated the need for a stochastic-dynamic approach to represent uncertainties in the initial state of the atmosphere for skillful forecasts in the extended range (Lorenz, 1965) and envisaged a Monte-Carlo approach consisting of a set of deterministic forecasts starting from slightly different initial states that consider errors or inadequacies in the observations, thereby providing probabilistic information about the future state of the atmosphere.

¹This transcript is part of the American Meteorological Society Oral History Project Collection and used with permission from the American Meteorological Society. The American Meteorological Society Oral History Project was created as a joint program between the American Meteorological Society and the University Corporation for Atmospheric Research. It aims to capture the history of the atmospheric sciences as told by the researchers, scientists, administrators and others working in the field.

Although this idea bears a strong resemblance to modern ensemble prediction systems, it took almost another 30 years before Lorenz's vision could be operationalized (Palmer, 2019a). This was mainly due to the limited computational resources which did not allow the operational production of ensemble forecasts, but also because the deterministic forecasts had to be further improved, and appropriate perturbation methods had to be developed (Lewis, 2005). Pioneered by ECMWF, the first operational ensemble forecast was issued in the early 1990's and consisted of a 33-member ensemble with initial condition perturbations through a singular vector (SV) technique (see Chapter 3; Palmer et al. 1992). Since then, multiple prediction centers worldwide have adopted ensemble forecasts into their operations, leading to a continuous improvement of the approaches and to the development of advanced methodologies (Leutbecher and Palmer, 2008), such as the incorporation of model uncertainty representations (Buizza et al., 1999). Further, the range of applications of the approach has greatly expanded, so that ensembles are now not only used for NWP, but also in climate simulations and in other components of the Earth system than the atmosphere, for example for representing uncertainties in the ocean or the land surface (e.g. Strømme et al. 2019).

2.1.2. Rationale

The basic idea of ensemble forecasting is to sample the probability density function of the future atmospheric state (PDF on the right side in Figure 2.1) by running several perturbed integrations of a forecast model (blue arrows in Figure 2.1; e.g. Leutbecher and Palmer 2008). In order to obtain a skillful probabilistic estimate, the sources of uncertainty have to be represented by the forecasting system. In theory, an adequate representation of all sources of uncertainty allows for the prediction of the forecast skill based on the so-called spread-error relationship: in a perfectly reliable ensemble, the deviations between the individual ensemble members (i.e. spread) yield a quantitative estimate of the forecast error (e.g. Leutbecher et al. 2017). The design of a reliable ensemble is, however, a non-trivial task, because not all sources of forecast uncertainty can be adequately included into a model in practice. The dominating sources of forecast uncertainty are:

- errors arising from deficient estimates of the initial condition which grow and propagate in the nonlinear atmospheric system (as described by Lorenz 1963), and
- model errors resulting from the numerical approximation of the governing equations including the required spatial and temporal discretization, and the parametrization of sub-grid processes that are not resolved by the model grid.

Even though these two error sources are commonly treated separately both in terms of the methods and their implications for predictability, they are in practice very much related to each other, because the estimation of the initial conditions for the model integration relies on a forecast model (i.e. data assimilation) which is subject to model errors (Leutbecher and Palmer, 2008).

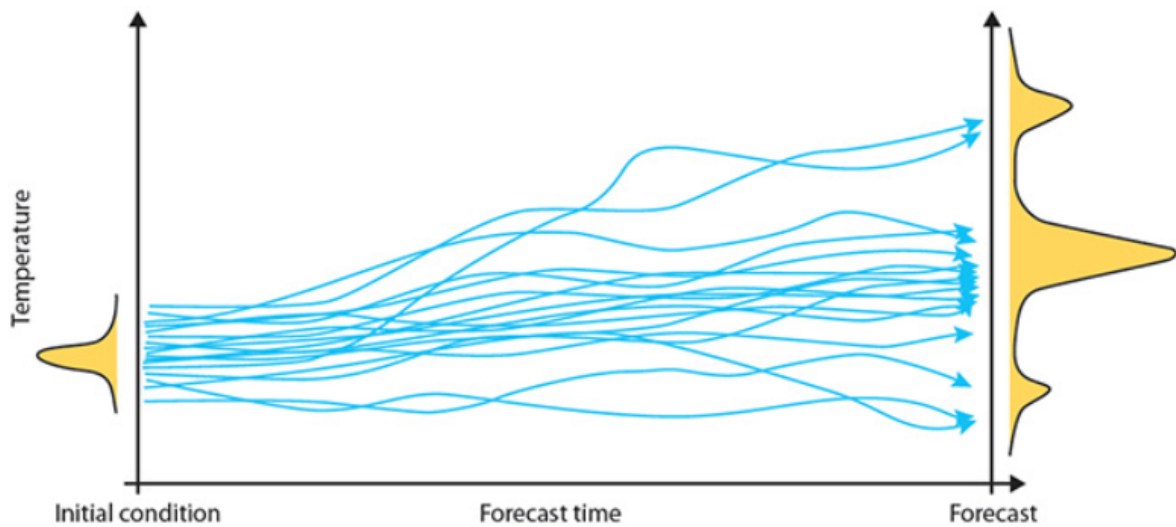


Figure 2.1.: Schematic illustration of a probabilistic temperature forecast that starts from slightly perturbed initial conditions and predicts a future probability density function. The blue arrows represent trajectories of individual forecast realisations. Figure reprinted from Buizza and Richardson (2017).

In the early days of ensemble forecasting, most prediction centers only considered initial condition uncertainties for their probabilistic forecasts (Lewis, 2005) and followed a "perfect-forecast-model-assumption" (Buizza et al., 1999) stating that uncertainties are mainly related to the initial condition errors, and that model errors are of minor relevance (Downton and Bell, 1988). However, it became evident that ensemble forecasts with only initial condition perturbations lack ensemble spread and are underdispersive even if the perturbations are consistent with the true distribution of initial condition errors (Wilks, 2005). Simply increasing the amplitude of the initial perturbations is not a solution to this problem, as this results in a substantial loss of forecast skill in the short range (Buizza et al., 1999). This issue motivated the development of a method to represent also the second major source of forecast error: model uncertainty. ECMWF implemented the stochastically perturbed parametrization tendencies (SPPT) scheme into their ensemble prediction system in 1998 (Buizza et al., 1999), and this scheme is still in operational use today (in a revised version, see Lock et al. 2019 for the latest update).

2.1.3. Stochastic parametrization

Even though Chapters 4 and 5 deal with sensitivities of weather phenomena to "ensemble configuration" in general, the main focus will lie on the impact of model uncertainty representations, for which reason the most important concepts of the latter are introduced here. A detailed description of the schemes used throughout this thesis is given in Chapter 3.

Classical parametrizations provide a deterministic estimate of the bulk effect of subgrid-scale processes on the resolved flow. However, there always exist multiple states of the subgrid space that are consistent with the model dynamics, such as different representations of clouds in a convectively active region (Palmer, 2019b). Further, the parametrizations are based on empirical formulae and parameters. Large parts of the uncertainty of forecast models therefore originate from the parametrization of processes that are not resolved by the model grid (Leutbecher et al., 2017). Most schemes that represent model uncertainty therefore do this by introducing stochastic perturbations to the physical parametrizations of the model (e.g. Berner et al. 2017, Leutbecher et al. 2017, Palmer 2019b, Wang et al. 2019). Additional to the stochastic treatment of the uncertain parametrization tendencies, such schemes intrinsically account for the probabilistic nature of parametrizations (i.e. multiple possible states of the sub-grid space). There exists a large variety of different approaches that are applied in different contexts (global NWP, high-resolution limited area simulations, climate simulations), but all have the common goal of increasing spread without reducing the skill of the ensemble - i.e. the increase of ensemble reliability. The most commonly used techniques are

- the stochastically perturbed parametrization tendencies (SPPT) scheme, which introduces state-dependent perturbations into the net tendencies provided by the physical processes (e.g. Buizza et al. 1999),
- the stochastic kinetic energy backscatter (SKEB) scheme, which accounts for interactions between resolved and unresolved scales (e.g. Shutts 2005),
- and many variants of stochastically perturbed parameter schemes, where selected parameters in individual physical parametrizations are perturbed (e.g. Ollinaho et al. (2017)).

ECMWF uses the SPPT-approach by perturbing the net tendencies from the physical parametrizations by a random number. These perturbations are state-dependent (or multiplicative), which means that their magnitude scales with the deterministic tendencies. Hence, situations with large parametrization tendencies (e.g. deep convection) are considered to be more uncertain than situations with only small physics tendencies (e.g. clear-sky radiation). This state-dependence leads to a larger impact of SPPT in the tropics than in the extratropics, where the net parametrization tendencies are smaller (e.g. Lock et al. 2019, see also Figure 3.4 in Chapter 3). More details on the SPPT-scheme and on its implementation follow in Chapter 3.

The main purpose of stochastic parametrization is an improvement of the ensemble reliability through an increase of spread, which results in improved probabilistic skill of the forecasts in the case of a well-designed scheme (Leutbecher et al., 2017). Despite its bulky approach of perturbing the net tendencies from all parametrizations, basic assumptions of the SPPT-approach have been shown to be consistent with the sub-grid uncertainty of forecast models (Shutts and Palmer, 2007). For example, Christensen

(2020) used high-resolution simulations to represent the "true atmosphere" and compared its coarse-grained output to the variability that is represented by the stochastic perturbations in coarse-resolution forecasts. With this approach, some previously made assumptions of the SPPT-scheme could be justified retrospectively, such as its multiplicative nature. This methodology can also be used to derive new stochastic parametrization schemes and test their physical consistency.

Apart from effects on the reliability, stochastic parametrization also affects other aspects of forecasts. There exists an extensive literature catalogue that deals with the impacts of SPPT on various aspects of the tropical model climate: Positive effects are for example reported for the representation of the El Niño Southern Oscillation (ENSO, Christensen et al. 2017; Yang et al. 2019), tropical precipitation (Subramanian et al., 2017; Strømme et al., 2019), the variability of the Madden-Julian-Oscillation (MJO, Weisheimer et al. 2014), the Asian summer monsoon (Strømme et al., 2018), and tropical cyclones (Stockdale et al., 2018; Vidale et al., 2021). In contrast, only few studies deal with the effect of SPPT on the extratropics, which is due to a more subtle influence of the perturbations at higher latitudes. Dawson and Palmer (2015) report an improved representation of North Atlantic weather regimes in climate models, and Christensen et al. (2015) show similar results in idealized simulations. A recent study investigated the impact of SPPT on atmospheric blocking in seasonal forecasts, but only minor effects were detected (Davini et al., 2021).

In the literature, the improved representation of large-scale flow regimes through SPPT is explained by the ability of a stochastic model to explore a larger portion of the system's variability compared to deterministic simulations (Christensen et al., 2015). Changes to the mean state of the model as a response to symmetric perturbations are, however, often assigned to the "noise-induced drift" of multiplicative, state-dependent perturbation techniques (Berner et al., 2017). This can be understood by considering an idealized system that is characterized by a single-potential well (i.e. a region around a local minimum of potential energy, see Figure 2.2). When the system is forced with additive (i.e. state-independent) stochastic perturbations (orange arrows), the mean state of the model will remain unchanged (Figure 2.2a). Multiplicative perturbations that depend on the state of the system (i.e. state-dependent noise, red arrows in Figure 2.2b) will, in contrast, result in a shift of the mean state of the model, which is indicated by the dashed line in the associated PDF of the system (Figure 2.2b). In a system with multiple stable states (i.e. a double-potential well), also additive noise can result in a shift of the mean state (Berner et al., 2017). Another way how symmetric perturbations can result in asymmetric responses is when they are applied to systems with nonlinear, irreversible dynamics, such as the formation of precipitation (Berner et al., 2015; Leutbecher et al., 2017).

Stochastic parametrization schemes are mostly developed by forecasting centers in the context of operational forecasting (such as by ECMWF, Buizza et al. 1999; Shutts 2005; Lang et al. 2021), and the

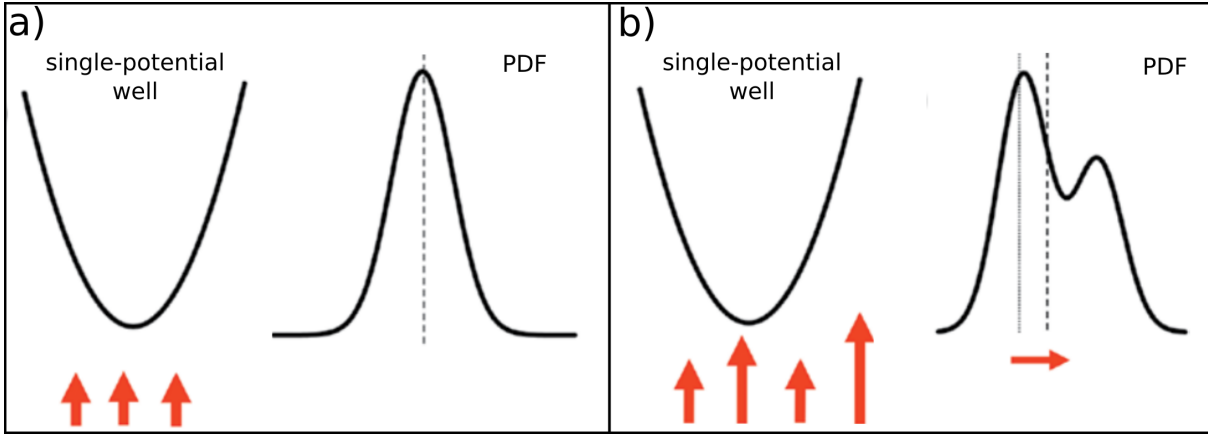


Figure 2.2.: System characterized by a single-potential well (left) and its associated PDF (right) in the case of (a) additive and (b) multiplicative noise (orange arrows). Modified from Berner et al. (2017).

schemes are tuned to increase the ensemble spread and are evaluated by standard verification methods and score cards. In this setting, alternative approaches to evaluate model uncertainty schemes are often neglected, because their implementation can be time-consuming and no common frameworks to ensure comparability among multiple institutions are defined. Therefore, process-based evaluations of stochastic perturbation schemes and how they interact with weather system are rare. Such approaches, however, have the potential to foster the understanding of the behaviour of stochastic noise in the nonlinear atmosphere and to broaden the perspective beyond skill scores. One weather phenomenon that is ideally suited for such an approach is the warm conveyor belt (WCB), which will be introduced in detail in the following section.

2.2. Warm conveyor belts (WCBs)

Warm conveyor belts (WCBs) are synoptic-scale flow features that are embedded into the midlatitude circulation. To better understand their dynamics and the relevance for forecasting, a short review on the large-scale extratropical circulation is given.

2.2.1. WCBs and the extratropical circulation

The large-scale extratropical atmospheric flow is governed by sequences of upper-level troughs and ridges, that are commonly referred to as Rossby waves (Rossby, 1939). The dynamics of these waves are based on the conservation of absolute vorticity:

$$\frac{D}{Dt}(f + \zeta) = 0, \quad (2.1)$$

where $f = 2\Omega \sin(\phi)$ is the planetary vorticity and ζ is the vertical component of the relative vorticity. A meridional displacement of an air parcel will modify its planetary vorticity, which has to result in a

changed relative vorticity in order to conserve the absolute vorticity. A northward displacement of an air parcel, for example, is equivalent to an increase of the planetary vorticity which leads to a decrease of the relative vorticity and results in an anticyclonic flow anomaly, and a southward displacement of an air parcel leads to an increase of relative vorticity and a cyclonic flow anomaly (Holton, 2004). The waves that result from this principle propagate westwards, but are advected with the mean westerly background winds.

The fundamental dynamics of Rossby waves can be explained in a purely barotropic framework, in which surfaces of constant pressure and temperature are parallel. The observed atmosphere, however, strongly differs from these strict assumptions, as it is characterized by strong horizontal temperature gradients (i.e. baroclinic zones), where isobars and isotherms intersect. Under geostrophic balance, such meridional temperature gradients result in zonal winds that increase with altitude (i.e. thermal wind, e.g. Holton 2004). When the temperature gradient exceeds a critical threshold, the thermal wind becomes unstable and baroclinic waves develop. This process, also referred to as baroclinic instability and first investigated by Charney (1947) and Eady (1949), is directly linked to the formation of extratropical cyclones (ETCs) that are a key component of the climate system, as they substantially contribute to the heat and momentum fluxes from the tropics to the poles (Holton, 2004). The cyclonic vorticity associated with ETCs advects cold and dry air equatorwards and warm and moist air polewards and thereby produces the well-known frontal systems, which characterize the day-to-day variability of the mid-latitude weather (e.g. Wernli and Schwierz 2006).

As the moist air in the warm sector of ETCs approaches the cold front, it ascends along the tilted isentropes, and clouds and precipitation form. Large parts of the precipitation associated with ETCs result from this coherent air stream (Pfahl et al., 2014) that ascends along the cold front into the upper troposphere - the so-called Warm Conveyor Belt (WCB). During the mostly slantwise ascent of WCBs, a large-scale cloud band forms and wraps cyclonically around the cyclone center, leading to the characteristic comma-shaped cloud spiral that is often visible on satellite images and depicted schematically in Figure 2.3 (Carlson, 1980). The term WCB goes back to Browning (1971), and first systematic analyses of WCBs used isentropic analysis to describe cyclone-relative air streams in a simplified framework (Harrold, 1973), consisting of the warm conveyor belt (red arrow in Figure 2.3), the cold conveyor belt (blue arrow in Figure 2.3), an air stream ascending on the polarward side of the warm front, and the dry intrusion (yellow arrow in Figure 2.3), a deeply descending air stream originating from the upper troposphere. In this idealized model, the air streams were identified by computing streamlines on surfaces of constant wet-bulb potential temperature (Harrold, 1973).

A milestone in the research of the conveyor belt model was the introduction of a Lagrangian method to compute kinematic trajectories following the 3-dimensional wind field (Wernli and Davies, 1997),

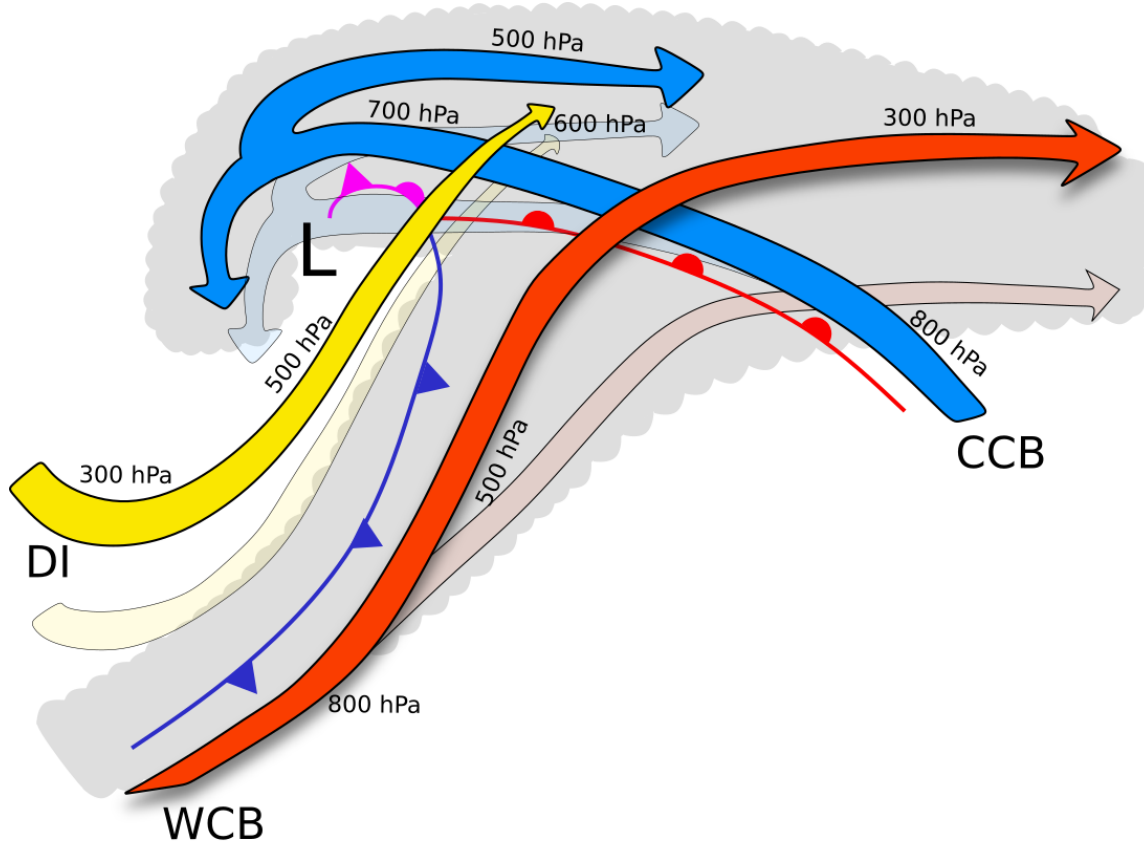


Figure 2.3.: Schematic of the conceptual conveyor belt model, consisting of the warm conveyor belt (WCB, red arrow), the cold conveyor belt (CCB, blue arrow), and the dry intrusion (DI, yellow arrow). The L denotes the center of the surface cyclone, and the blue, red and purple lines depict the cold, warm and occluded fronts. The gray areas illustrate clouds related to the air streams. The pressure values correspond to approximate height levels of the air streams. Inspired by Figure 9 of Carlson (1980).

which allowed to identify and quantitatively analyze air streams in transient weather systems. With this approach, WCBs have first been identified as coherent bundle of trajectories with a maximum decrease of specific humidity within a period of two days (Wernli, 1997). Other definitions use the integrated diabatic heating along or the ascent rate of trajectories as selection criteria (e.g. Madonna et al. 2014b), where all of these approaches reflect the diabatically driven, cross-isentropic ascent of WCBs. Since its introduction, the Lagrangian framework has been used extensively to study physical and dynamical processes along WCBs (examples are Eckhardt et al. 2004, Grams et al. 2011), which substantially fostered the understanding of WCBs and related processes. Madonna et al. (2014b) provide a comprehensive climatological analysis based on reanalysis data and show that WCBs mainly occur in the storm tracks of the extratropical ocean basins, predominantly in the cold season when baroclinic cyclone activity is largest. Trajectory analysis also helped to refine the classical conveyor belt model: based on high-resolution simulations it was shown that the ascent of WCBs is not solely of slantwise nature, but also occurs convectively (Rasp et al., 2016; Oertel et al., 2019).

A key aspect of WCBs is their cross-isentropic, diabatically enhanced ascent that results from latent heat release due to cloud-diabatic processes (Wernli, 1997). Originating from the planetary boundary layer in the cyclone's warm sector (label 1 in Figure 2.5), the WCB inflow is characterized by large amounts of specific humidity due to moisture flux convergence and strong surface fluxes from the ocean (Schäfler and Harnisch, 2015; Dacre et al., 2019). Under the influence of upper-level quasi-geostrophic forcing, this warm and moist air stream starts to ascend across the warm front (Binder et al., 2016) and cools adiabatically (label 2 in Figure 2.5). Once saturation is reached, the formation of clouds and precipitation results in strong latent heat release, which typically amounts to approximately 20 K during the entire ascent (Eckhardt et al., 2004; Madonna et al., 2014b) and further accelerates the vertical motions. The diabatically heated air stream thereby reaches the upper troposphere (label 3 in Figure 2.5), where it impinges on the tropopause and diverges horizontally (also referred to as diabatic outflow, Grams and Archambault 2016). Due to the strong diabatic heating along their ascent, WCBs have a substantial impact on the large-scale atmospheric flow, which will be explained in terms of potential vorticity (PV) in the subsequent section.

2.2.2. The potential vorticity framework

In order to describe the interaction of WCBs and the large-scale circulation, we employ a potential vorticity (PV) - potential temperature (θ) framework (Hoskins et al., 1985) which helps to understand the impact of the latent heat release on midlatitude dynamics. After Ertel (1942), PV is defined as

$$PV = \frac{1}{\rho} \boldsymbol{\omega} \cdot \nabla \Theta, \quad (2.2)$$

where ρ is density, $\boldsymbol{\omega}$ the 3-dimensional absolute vorticity, and θ potential temperature. PV is a quantity that combines dynamical (i.e. vorticity) and thermodynamical (i.e. the gradient of potential temperature) aspects of the flow. PV is usually expressed in PV units (PVU), where $1 \text{ PVU} = 10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$. High PV values are associated with cyclonic vorticity and/or a stable stratification of the atmosphere, for which reason the threshold of 2 PVU is frequently used to identify the dynamical tropopause. In contrast, low PV values are linked to anticyclonic vorticity. When frictional processes are neglected, the Lagrangian rate of change of PV is defined as (Hoskins et al., 1985)

$$\frac{D}{Dt} PV = \frac{1}{\rho} \boldsymbol{\omega} \cdot \nabla \dot{\Theta}. \quad (2.3)$$

In the absence of diabatic processes (i.e. under adiabatic conditions, where $\dot{\Theta} = 0$), PV is materially conserved. For the mostly large-scale and slantwise ascent of WCBs, Eq. 2.3 can be approximated to

$$\frac{D}{Dt} PV \approx \frac{1}{\rho} (f + \zeta) \cdot \frac{\partial \dot{\Theta}}{\partial z}. \quad (2.4)$$

2. Background and Research Questions

This is because the vertical components of both the vorticity as well as the diabatic heating gradient dominate on the corresponding scales of motion (Holton, 2004). Equation 2.4 shows that, apart from the advective tendencies hidden in the Lagrangian form of the equation, PV is generated below and destroyed above the maximum of diabatic heating. Assuming a local impulsive diabatic heating in the mid-troposphere without the advection of PV, a positive PV-anomaly will form below the heating maximum, and a negative PV-anomaly will form aloft (Figure 2.4a). In the case of continuous diabatic heating, as it is the case in WCB-like situations due to the steady condensational heating during the ascent from the lower to the upper troposphere, also the advection of PV has to be considered. This results in an upward displacement of the positive PV-anomaly, which is then co-located with the heating maximum in the lower- to mid-troposphere, and the negative PV-anomaly is advected to the upper levels, where it spreads out with the upper-tropospheric divergent wind (Figure 2.4b). This concept of PV-modification in WCBs has been derived based on a case study of an extratropical cyclone (Wernli and Davies, 1997; Wernli, 1997), and the findings have been generalized subsequently in a climatological framework (Madonna et al. 2014b). WCB-trajectories in the inflow stage are on average characterized by PV-values below 0.5 PVU. During the ascent, PV increases due to cloud-diabatic heating in the mid-troposphere to about 1 PVU, followed by a decrease in the outflow region to levels around 0.5 PVU (Wernli, 1997; Madonna et al., 2014b). Hence, the PV in the outflow stage of WCBs is similar to the one in the inflow region.

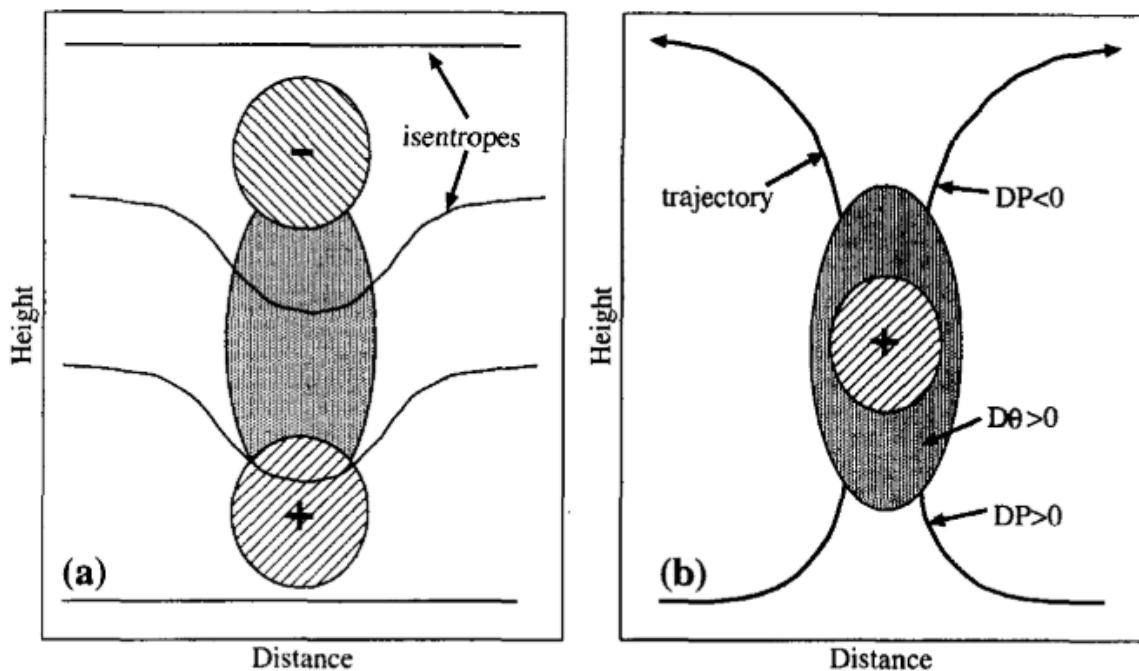


Figure 2.4.: Schematic vertical cross-sections showing diabatically produced PV anomalies (hatched regions with + or -) for the idealized cases of (a) local impulsive diabatic heating and (b) continuous diabatic heating. Shading indicates the region of diabatic heating. The solid lines in (a) are isentropes, the bold arrows are trajectories. $D\theta$ and DP denote material tendencies of potential temperature and potential vorticity, respectively. Reprinted from Wernli and Davies (1997).

These empirical findings are supported by theoretical considerations of Methven (2015), who argues that the net change of PV along WCB-trajectories is zero, under the assumption that there can be no PV-flux across isentropic surfaces (Haynes and McIntyre, 1987). The different climatological background states in the in- and outflow region of WCBs, however, result in the formation of positive PV-anomalies in the lower troposphere, which is usually characterized by low values of PV, and negative PV-anomalies in the upper troposphere, where on average high PV-values are predominant (Pomroy and Thorpe, 2000). The latent heat release along the ascending air streams therefore does not determine the magnitude of the PV-anomaly itself, but in which environment it is injected; the diabatically enhanced ascent results in higher outflow levels than it would be possible under adiabatic conditions (Saffin et al., 2021). In summary, WCBs provide an efficient mechanism to transport low-PV air from the lower troposphere cross-isentropically to the upper levels (Grams et al., 2013).

2.2.3. Large-scale flow modification of WCBs

Both the positive low-level and negative upper-level PV-anomaly resulting from the cloud-condensational processes during the WCB-ascent influence the mid-latitude flow. The former is associated with cyclonic vorticity and can contribute to the intensification of the surface cyclone (Binder et al., 2016; Martínez-Alvarado et al., 2016a). The divergent outflow of WCBs, associated with a negative (i.e. anticyclonic) upper-level PV-anomaly (blue shading in Figure 2.5), impinges on the upper-level jet (green line in Figure 2.5) and deflects it pole- and upward, which ultimately results in ridge building and amplification (Pomroy and Thorpe, 2000; Grams et al., 2011; Chagnon et al., 2013; Martínez-Alvarado et al., 2016b; Saffin et al., 2021). The low-PV outflow additionally enhances the PV-gradient across the tropopause, which results in the formation of a jet streak (Grams et al., 2013). The ridge amplification can subsequently lead to the downstream development and propagation of baroclinic Rossby wave packets (Röthlisberger et al., 2018) and contribute to Rossby wave breaking events (Madonna et al., 2014a). Recent studies further highlight the role of WCB-outflow for the formation and maintenance of atmospheric blockings (e.g. Pfahl et al. 2015; Steinfeld and Pfahl 2019; Steinfeld et al. 2020), which are a key ingredient of the midlatitude flow and related to high-impact weather (e.g. Matsueda 2011; Pfahl and Wernli 2012; Sousa et al. 2018).

In turn, the details of the WCB-ascent and the subsequent modulation of the large-scale flow are controlled by the environmental background flow, such as the low-level baroclinicity (Grams et al., 2018), the upper-level QG-forcing (Binder et al., 2016), or the moisture supply in the inflow region (Schäfler and Harnisch, 2015), which largely determines the integrated diabatic heating along the ascent and the characteristics of the outflow air mass. Because of the combination of the sensitivity of WCBs to the atmospheric background state and the strong influence of WCBs of the large-scale circulation, they can be seen as a dynamical link between the lower and upper troposphere.

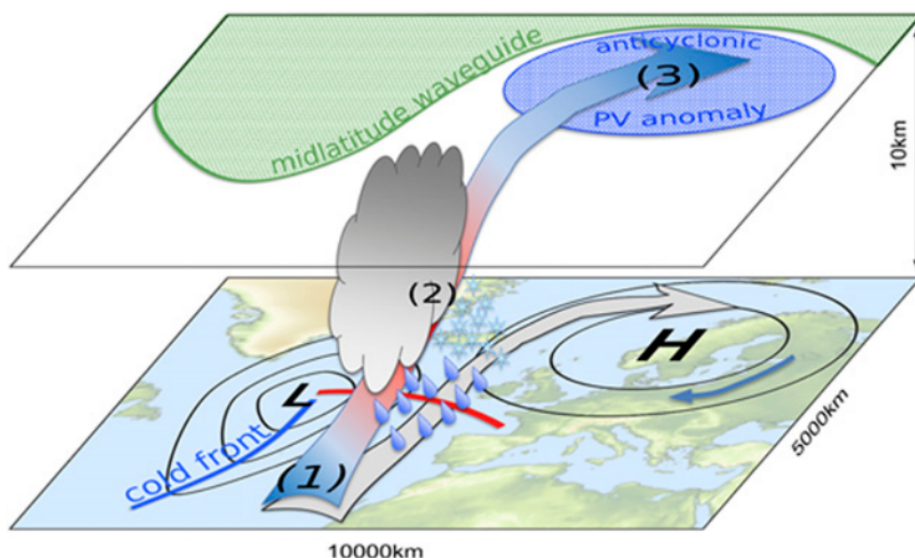


Figure 2.5.: Schematic illustration of a WCB and the surrounding synoptic and large-scale situation. The cyclone center is indicated with an **L** and the associated cold and warm fronts as blue and red line, respectively. The colored arrow depicts the typical trajectory of the WCB (blue and red colors denote low and high values of PV, respectively) that starts in the inflow region (1), ascends through the mid-troposphere in which diabatic processes are active (2, shown as cloud and precipitation) and reaches the outflow region in the upper troposphere (3). The diabatic outflow is characterized by anticyclonic PV-anomalies (blue hatching) and deflects the upper-level waveguide (green contour/shading). Reprinted from Quinting and Grams (2021).

Even though the tropical dynamics are fairly different from the midlatitude atmospheric circulation, also tropical weather phenomena, such as tropical cyclones (TCs), organized tropical convection or mesoscale convective systems (MCSs), are characterized by strongly ascending air streams, latent heat release and a cross-isentropic air mass transport. Due to their diabatic outflow, these systems can have a similar impact on the extratropical circulation as WCBs, provided that they approach the midlatitudes. When TCs, for example, undergo extratropical transition (ET, Evans et al. 2017), they can impact the Rossby wave pattern and enhance ridge building (Quinting and Jones, 2016; Grams and Archambault, 2016; Keller et al., 2019). Even though the focus of this thesis lies on the extratropics, parts of Chapter 4 deal with diabatically heated, rapidly ascending air streams in general, and therefore also tropical weather systems will be considered.

2.2.4. The relevance of WCBs for numerical weather prediction

For its future strategy, ECMWF has identified the correct representation of WCBs as an important element to achieve their target of improving the reliability and sharpness of their forecasts (Rodwell et al., 2018a). This shows the relevance of WCBs for forecasting, which results from the substantial impact that WCBs exert on the large-scale flow: erroneous environmental conditions determining details of the WCB-ascent and deficient representations of physical and dynamical processes in WCBs can lead to a mis-representation of the upper-level Rossby wave pattern and thereby result in downstream forecast

uncertainty and strongly reduced forecast skill. An example for such a "forecast bust" occurred in March 2016: a small-scale upstream error led to a mis-placed WCB over the North Atlantic, with a too weak poleward extent of the WCB outflow and a too strong southerly branch extending into Europe. As a consequence, the large-scale flow configuration in the forecast decoupled from the analysis, which resulted in historically low forecast skill at 6 days lead over Europe (Magnusson, 2017; Grams et al., 2018).

The systematic investigation of WCBs in (ensemble) forecast data sets and their relation to forecast errors, however, is very challenging, because the detection of WCBs is usually based on a Lagrangian approach (Wernli and Davies, 1997). This methodology requires a comparably high temporal ($\mathcal{O}(3 - 6h)$) and spatial resolution ($\mathcal{O}(1^\circ)$ in the horizontal, $\mathcal{O}(10hPa)$ in the vertical) of the input fields (Bowman et al., 2013), which is not provided in forecast data archives. Therefore, most of the existing studies are either based on detailed investigations of individual cases (e.g. Joos and Forbes 2016 or Grams et al. 2018), or make use of proxy metrics that quantify the impact of WCBs, such as the advection of PV by the upper-level divergent wind (e.g. Teubler and Riemer 2016). So far, the only trajectory-based systematic evaluation of WCBs in forecasts is provided by Madonna et al. (2015), who report that several poor forecasts were linked to erroneous representation of WCBs. Just recently, Wandel et al. (2021) used a novel, statistical approach trained on trajectory-based WCB-imprints (Quinting and Grams, 2021) to evaluate the forecast skill of WCBs in extended-range ECMWF reforecasts, and found that skillful predictions of WCBs are on average possible until lead times of 8-10 days.

In the existing literature, two pathways how WCBs can influence forecast skill have emerged:

Upscale error growth

WCBs can amplify pre-existing errors and project them on other scales of motion, as described in Grams et al. (2018) and Berman and Torn (2019). This mechanism is part of the conceptual 3-stage model of upscale error growth, first introduced by Zhang et al. (2007): during the first hours of the model integration, localized, small-scale errors from the initial conditions grow to the convective scale. In a second stage, the errors in convective-scale unbalanced flow are projected to the large scale via geostrophic adjustment (Bierdel et al., 2018), resulting in errors in the balanced flow field, which finally amplify with dry-dynamic baroclinic instability (Davies and Didone, 2013) and barotropic Rossby wave dynamics (Baumgart et al., 2019). The moist-diabatic processes associated with the ascent of WCBs act in the range of convective to synoptic scales, and are therefore related to stage two of the upscale error growth mechanism (Grams et al., 2018). Deterministic convection schemes, however, do not represent the convective-scale error growth adequately, which possibly results in underdispersive ensembles (Selz, 2019). Baumgart et al. (2019) investigate upscale error growth in a quantitative framework and demonstrate that the divergent outflow of moist-diabatic processes is of particular relevance during the first 2 days of a forecast, before the near-tropopause barotropic processes make the main contribution.

Error growth due to parametrizations

Forecast uncertainty associated with WCBs is also related to insufficient representations of the diabatic processes in their ascending stage. The cloud-diabatic processes that occur when saturation is reached determine the amount of latent heat release, which subsequently controls the outflow characteristics of the ascending trajectories (Madonna et al., 2014b). In global NWP-models, these moist-dynamic processes have to be parametrized, which is a source of forecast uncertainty (Leutbecher et al., 2017). Several studies investigated sensitivities of the ascent characteristics of WCBs and the downstream flow evolution on the choice of different parametrization schemes. It was found that the upper-level Rossby wave pattern is altered when different microphysics parametrizations are used (e.g. Joos and Wernli 2012; Joos and Forbes 2016; Mazoyer et al. 2021; Choudhary and Voigt 2022). Also the convective updrafts in WCBs cannot be resolved by coarse NWP-models. Similarly to the microphysics parametrizations, Rivière et al. (2021) report sensitivities of WCBs to deep convection parametrizations. Further, high-resolution, convection-permitting simulations showed that embedded convection creates small-scale PV-structures that are not captured by global NWP-models (Oertel et al., 2020), and that negative PV-anomalies and the response of the upper-level flow are erroneous with parametrized convection (Clarke et al., 2019). These sensitivities to different parametrizations show that WCBs can also act as a direct source of forecast uncertainty and error.

Weather systems that share similar characteristics with WCBs have been linked with increased forecast uncertainty, such as extratropical cyclones (Rodwell and Wernli, 2022) or extratropical transitions of tropical cyclones (Aiyyer, 2015; Torn, 2017). Other studies do not explicitly investigate specific weather systems or separate the two error growth mechanisms, but focus on periods of particularly bad forecast skill (i.e. predictability barriers and forecast busts) and analyze the concomitant flow situation. In a recent study by Sánchez et al. (2020), periods of rapid forecast error growth and with reduced ensemble reliability were linked to diabatic processes at the tropopause level. A composite analysis of the most severe European forecast busts revealed that MCSs over the North American continent with associated diabatic outflow act as precursors of low forecast skill downstream (Rodwell et al., 2013). These low-predictability flow situations are on average characterized by anticyclonic flow anomalies over the North Atlantic/Europe (Rodwell et al., 2018b). This large-scale flow configuration with an amplified Rossby wave pattern has been shown to be less predictable than zonal flow patterns (Ferranti et al., 2015; Büeler et al., 2021), which links to a long-lasting issue of NWP- and climate models: state-of-the-art forecasting systems have deficiencies in correctly representing the upper-level Rossby wave pattern (Gray et al., 2014; Martínez-Alvarado and Plant, 2014; Martínez-Alvarado et al., 2018) and systematically underestimate atmospheric blocking (e.g. Quinting and Vitart 2019, Davini and D’Andrea 2020). Due to the dynamical link of WCBs to these weather patterns, it has been argued that systematic errors in the representation of the large-scale mid-latitude flow are (partly) related to erroneous representations of diabatic

processes and WCBs (Gray et al., 2014; Saffin et al., 2017; Maddison et al., 2020).

2.3. Research questions

This literature overview demonstrates that WCBs are weather phenomena that connect different horizontal scales and vertical levels of the troposphere. The diabatic processes along their ascent exhibit sensitivities to parametrizations in models, and the response of the large-scale dynamics strongly depends on the diabatic heating and on the environmental conditions. This makes WCBs and other diabatically driven air streams highly relevant for numerical weather prediction. At the same time, the above mentioned characteristics suggest that WCBs can serve as a promising test bed for the evaluation of forecasts and of individual components of prediction systems, such as model uncertainty schemes. The latter are usually assessed by established verification techniques and score cards, but rarely from a process-oriented perspective.

In the following, the main research questions that have been derived from the background chapter and will be addressed in the thesis are formulated. They are classified by the chapter in which they are addressed (**C4**, **C5** and **C6**) and numbered consecutively.

In Chapter 4, a weather-system perspective is employed by evaluating the sensitivities of WCBs and other diabatically driven, rapidly ascending air streams to the configuration of ECMWF's ensemble forecasting system. As model uncertainty schemes are often designed to introduce large perturbations into regions where parametrizations are active, the question arises if and how such schemes affect WCBs. In particular, the following aspects are addressed:

- C4-1** Do the perturbation schemes that are in operational use at ECMWF affect the occurrence and characteristics of diabatically driven, rapidly ascending air streams?
- C4-2** Do other model uncertainty schemes result in similar effects as SPPT?
- C4-3** How can the observed effects be explained?

Motivated by the results from Chapter 4, the subsequent Chapter 5 examines if and how the effects that have been found in the previous chapter affect the model climate by modifying atmospheric phenomena related to rapidly ascending air streams. The main research questions for Chapter 5 are:

C5-1 Do stochastic model uncertainty schemes modify the global distribution of precipitation?

C5-2 Is the amplitude of the upper-level Rossby wave pattern affected by stochastic model uncertainty schemes?

It will be discussed in detail if the effects of model uncertainty schemes on precipitation and the Rossby wave amplitude are modulated by altered WCB-activity, or if this occurs independently from each other.

Chapter 6 changes the perspective and picks up the discussion if WCBs have a negative impact on the skill of numerical weather forecasts. This has been demonstrated previously, but only in case studies and/or with proxy metrics that describe the impact of WCBs on the large-scale flow. In this chapter, this gap is bridged by systematically analysing the relationship of WCBs and forecast error growth in a Lagrangian framework. The following questions are considered to establish a causal relationship:

C6-1 Do WCBs and forecast errors of the large-scale circulation co-occur?

C6-2 Can periods of increased forecast error growth be linked to WCB-activity?

C6-3 Is WCB-activity related to coherent spatio-temporal error patterns?

Before these research questions are tackled, the following chapter (Chapter 3) introduces the data sets and methods used throughout the thesis.

3. Data and Methods

In this Chapter, the data and methods for the investigation of WCBs in ensemble forecasts are presented. At first, the forecasting system that is used throughout the thesis is introduced, including a detailed description of the model uncertainty schemes (Section 3.1). Then, the experimental design and the employed data sets are described (Section 3.2), and Section 3.3 introduces the Lagrangian detection of WCBs and other diabatically driven, rapidly ascending air streams. Note that this chapter only introduces the main methods that are used in large parts of the thesis. Other diagnostics that are tailored to address specific aspects will be introduced in the corresponding results chapters.

3.1. ECMWF's ensemble prediction system

All analyses in this thesis are conducted with data produced by the ensemble prediction system (EPS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), which consists of a data assimilation suite, a numerical model, and representations of forecast uncertainty. The latter two are introduced in the following section.

3.1.1. Integrated Forecasting System (IFS)

Dynamical core

The Integrated Forecasting System (IFS) is ECMWF's operational global forecasting model that forms the basis of the ensemble prediction system. The prognostic variables of the model are the horizontal wind components (U and V), temperature (T), specific humidity (Q), and surface pressure (p_s). The IFS is a hydrostatic model, hence the vertical wind component (ω) is a diagnostic variable derived from hydrostatic balance. The prognostic variables are integrated using a semi-Lagrangian advection scheme, combined with a semi-implicit time integration scheme. In the vertical, the IFS is discretized using a finite-element scheme, resulting in a hybrid σ -coordinate that follows orography at lower levels and pressure surfaces in the free atmosphere. The semi-Lagrangian advection scheme computes air parcel trajectories and iteratively estimates the departure point (DP) of the trajectory arriving at a grid point at the subsequent time step (arrival point, AP). The model dynamics are represented in spectral space, which enables very high computational performance. The physical parametrizations are computed locally in grid-point space; to exchange information between the model dynamics and physics, the model therefore has to convert the variables U , V , T and p_s between grid-point and spectral space at every model time step. The IFS-versions used in this thesis (see Table 3.3) have a cubic octrahedral reduced Gaussian

grid (TCO), where the maximum wavenumber of the spectral truncation determines the horizontal resolution. For further details, we refer to the IFS documentation (ECMWF, 2019a).

Physical parametrizations

The sub-grid processes that are not resolved by the dynamical core are represented by the physical parametrizations. The parametrization schemes can be classified into 5 different groups: radiation, turbulent diffusion and subgrid orography, non-orographic gravity wave drag, convection, and clouds and large-scale precipitation. The order of the list reflects the sequence in which the parametrizations are called. The tendencies from the parametrizations are added to the dynamical tendencies sequentially, such that the subsequently called schemes use updated variables (fractional stepping, c.f. Wedi 1999).

The convection parametrization represents the effect of unresolved convective clouds and is based on a bulk mass flux scheme originally introduced by Tiedtke (1989) that was revised in 2008 (Bechthold et al., 2008). It describes three types of convection (shallow, mid-level and deep convection), where the choice of the convection type determines the formulation of certain cloud properties. The up- and downdrafts in the clouds are described by a pair of entraining and detraining plumes (i.e. mixing of dry environmental air into or out of the moist convective plume, respectively).

The parametrization of clouds and large-scale precipitation (originally termed "large-scale water processes" LSWP) accounts for sub-grid processes within clouds that are resolved by the model grid. The scheme builds on the original formulation of Tiedtke (1993) with two prognostic parameters representing the grid-box fraction of cloud coverage and the mass mixing ratio of total cloud water content, divided into liquid and ice categories. In this formulation, clouds are formed through large-scale ascent, diabatic cooling, boundary-layer turbulence, and horizontal transport of cloud water through convective updrafts, and dissipated by adiabatic and diabatic heating, turbulent mixing with unsaturated environmental air, and depletion of cloud water by precipitation. A major upgrade implemented in model cycle CY36R4 introduced three additional prognostic variables that enables a more physically based representation of mixed-phase clouds and precipitating rain and snow (Forbes et al., 2011). Each of the prognostic variables is represented by a grid point average mass of water, and number concentration, particle size distribution and properties are diagnosed (i.e. single-moment microphysics scheme).

As this thesis mainly addresses the role of moist-diabatic processes related to warm conveyor belts in ensemble forecasts, we omit the description of the other parametrizations and refer to the documentation of the IFS physics for detailed information (ECMWF, 2019b) and the references therein.

3.1.2. Uncertainty representation and ensemble configuration

ECMWF's EPS accounts for forecast uncertainty arising from the initial conditions and from the forecast model to produce probabilistic forecasts of the future state of the atmosphere. Figure 3.1 gives a schematic illustration of such an ensemble system.

Representing initial condition uncertainty

The estimation of the initial state of the atmosphere is of utmost importance for skillful weather forecasts. At ECMWF, the initial conditions for the forecast integration are generated by a 4-dimensional variational (4D-var) data assimilation that accounts for systematic and random observational errors and nudges short-term forecasts towards the observed state (Rabier et al., 2000). The 4D-var data assimilation is one of the core components of the IFS and is jointly responsible for the high quality forecasts of ECMWF (Magnusson and Källén, 2013). Uncertainties in the initial conditions are subsequently addressed by two approaches:

- At first, an ensemble of 4D-var data assimilations (EDA) with 25 independent members that introduce perturbations to the observations, to the physical processes in the short-term forecasts, and to the sea surface temperatures is run (Isaksen et al., 2010).
- Then, singular vector (SV) perturbations are added to the analysis fields. The SV-approach was developed by Molteni and Palmer (1993) and identifies those directions of initial uncertainty that are responsible for the largest forecast uncertainty within a specified time interval, and thereby generate a rapid dispersion of the individual ensemble members (Leutbecher and Palmer, 2008).

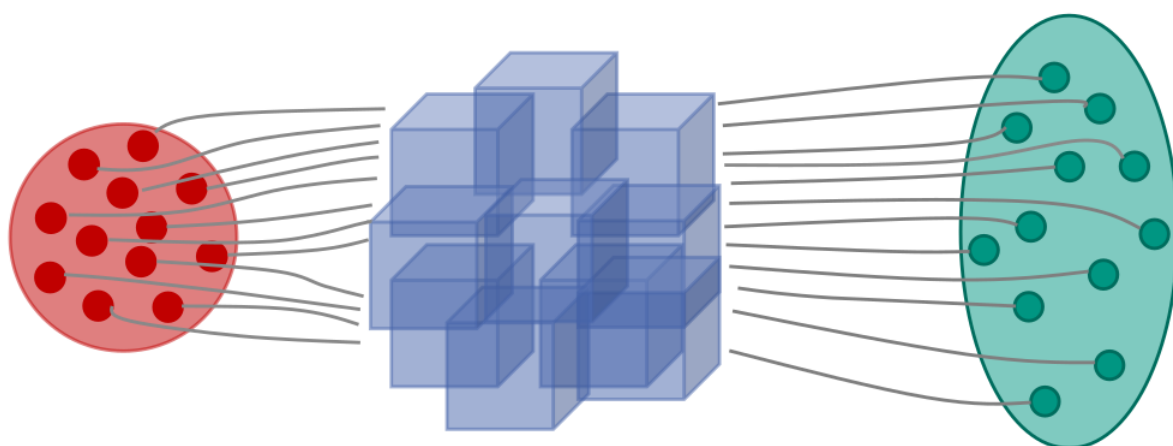


Figure 3.1.: Schematic illustration of an ensemble prediction system, consisting of initial condition perturbations (red dots) and forecast model perturbations (blue cubes) to obtain a probabilistic estimate of the future state of the atmosphere (green dots). The example consists of 11 ensemble members.

The EDA-approach is computationally very expensive. To save computational resources, the operational implementation at ECMWF uses pairs of opposite initial perturbations to double the number of members (Leutbecher, 2019). In contrast, the SV-technique is comparably cheap and serves as a pragmatic way to boost the divergence of the ensemble members. The combination of EDA and SVs is very efficient in generating ensemble spread, especially in the extratropics (see Figure 3.4), and will be referred to as initial condition perturbations (ICP) throughout this thesis.

Representing model uncertainty

In this section, those model uncertainty schemes that used throughout this thesis are introduced. Large parts of the subsequent chapters deal with the scheme that is currently in operational use at ECMWF - the SPPT-scheme, which will therefore be introduced in more detail. The two other scheme (SPP and STOCHDP) are introduced afterwards.

Stochastically perturbed parametrization tendencies (SPPT)

In order to increase the ensemble spread, the stochastically perturbed parametrization tendencies (SPPT) scheme has been implemented in 1998 (Buizza et al., 1999) and is still being used operationally in a revised form (see Lock et al. 2019 for the most recent version). SPPT is designed to introduce perturbations in regions where parametrizations are active and is based on the assumption that the uncertainty of the parametrized processes is proportional to the magnitude of the net tendencies of the parametrizations. SPPT multiplies the net physics tendencies (i.e. the sum of the tendencies from all parametrization schemes) of the variables temperature, specific humidity and the horizontal wind components with a 2-dimensional (horizontal) random field r :

$$\mathbf{p} = (1 + \mu r)\mathbf{p}_D, \quad (3.1)$$

where \mathbf{p} and \mathbf{p}_D are the perturbed and the deterministic parametrization tendencies, respectively, and μ is a tapering function depending on the model level with values in the range $[0,1]$.

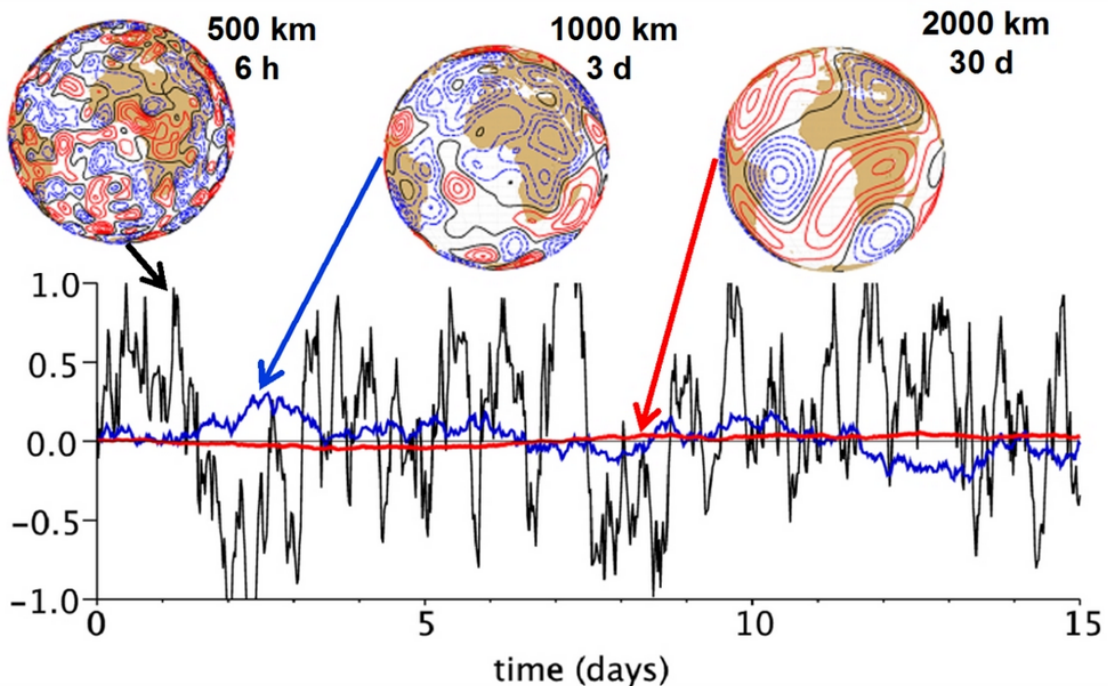
The random field r evolves in space and time with every model time step and is generated through a first-order auto-regressive process in spectral space. The random pattern is represented by three independent scales with corresponding spatial auto-correlation and temporal decorrelation scales (see Table 3.1). Each of the three scales is characterized by a standard deviation ($\sigma=0.52, 0.18$ and 0.06), which defines the underlying distribution and determines the fluctuations of the patterns (see Figure 3.2 for a visualization of the pattern). r samples a Gaussian distribution centered on the unperturbed tendency with variances corresponding to the aforementioned standard deviations, and is clipped to the range $[-1,1]$. For each grid point, r is applied throughout the depth of the atmosphere, which preserves the vertical structures of the physics tendencies. Exceptions are the boundary layer, where perturbations can result

Table 3.1.: Standard deviation (σ), horizontal decorrelation scale (L) and temporal decorrelation scale (τ) of the random field r .

pattern id	σ	L (km)	τ (d)
1	0.52	500	0.25
2	0.18	1000	3
3	0.06	2000	30

in numeric instabilities, and the stratosphere, where perturbations would introduce large uncertainties due to clear-sky radiative tendencies, which are well-constrained. Therefore, a tapering function μ is applied that is set to zero in the boundary layer and gradually increases to 1 when moving into the free troposphere.

Two things are important to note: Firstly, the perturbations by SPPT are symmetric and zero-mean; and secondly, that even though there is variability in the random field, the spatio-temporal correlations of the field generator ensure that neighboring grid points or subsequent model time steps at the same grid point do not receive unphysical large-amplitude perturbations of opposite signs.

**Figure 3.2.:** Example of the spatial structure (maps) and the 15-day temporal evolution on an arbitrary grid point (time series) of the 3-component random pattern r which is used to generate the perturbations of the parametrization tendencies in the SPPT-scheme (see Table 3.1). Figure adapted from Shutts et al. (2011).

From equation 3.1, it becomes evident that the maximum possible perturbation is a doubling of the net tendencies (for $r = 1$) or a vanishing of the physical tendencies (for $r = -1$), and the minimum possible perturbation is that the deterministic tendencies from the parametrizations are used (for $r = 0$). Apart from the stochasticity, the magnitude of the introduced perturbation is controlled by the net physics tendencies (\mathbf{p}_D), which makes the perturbations state-dependent: Averaged over many situations (i.e. neglecting the influence of the random pattern), SPPT will introduce larger perturbations into regions where processes are parametrized in the model than in situations that are resolved by the dynamics.

Even though SPPT has been used very successfully during the past decades, the scheme has some disadvantages. One of them is that SPPT can violate conservation properties of the parametrizations, as the physics tendencies are perturbed after they have been computed by the parametrization schemes. Furthermore, uncertainties of individual processes are not considered, but only the bulk uncertainty of all parametrized processes, expressed by the net parametrization tendencies. In situations where parametrization tendencies from different processes cancel each other out (i.e. the net tendency is close to 0), SPPT will therefore introduce only little perturbations. Hence, process-level uncertainty is not directly taken into account by SPPT.

Stochastically perturbed parametrizations (SPP)

A scheme which circumvents these issues is the "stochastically perturbed parametrizations" (SPP) scheme, that perturbs selected parameters and variables of different parametrizations instead of the net tendencies of all physics parametrizations (Ollinaho et al., 2017). Similarly to SPPT, it accounts for uncertainties in the parametrization schemes, but attempts to do this on a process level. This is achieved by selecting such parameters of the parametrization schemes that are considered uncertain, which allows for a direct representation of uncertainty at its source. In the setup used in this thesis, in total 27 parameters are perturbed in the following parametrizations (Lang et al., 2021):

- turbulent diffusion and subgrid orography,
- convection,
- clouds and large-scale precipitation,
- radiation.

The parameters are mostly sampled from log-normal distributions with the mean centered on the unperturbed, deterministic parameter, and the distributions represent a range of physically plausible values. A complete list of the perturbed parameters and the underlying distributions is given Table A.1 in Appendix A. Similarly to SPPT, a 2D random field generator is used to generate the perturbations, but only a single scale with a decorrelation length scale of 2000 km and a decorrelation time of 72 h is used. As the parameters are directly perturbed, physical consistency and local conservation properties of the

parametrizations are preserved (Ollinaho et al., 2017).

One scientific advantage of SPP compared to SPPT is that sensitivity experiments with different configurations of SPP can be conducted. For example, forecasts where only parameters in individual parametrizations (e.g. the convection scheme) are perturbed by SPP can help to understand the interactions of the scheme with processes in the model.

STOCHDP

Uncertainties in the forecast model are not restricted to the physical processes, but are also present in the dynamical core. Even though the dynamics are directly resolved by the model, the discretization and the numerical approximation introduce errors into the forecast. The "stochastically perturbed semi-Lagrangian departure points" (STOCHDP) scheme accounts for uncertainty in the dynamical core by introducing stochastic perturbations into the estimation of the semi-Lagrangian departure points. The magnitude of the perturbation scales with the complexity of the flow (i.e. large perturbations in situations with strong shear or vorticity) and is therefore flow-dependent (Leutbecher et al., 2017). STOCHDP is currently under development at ECMWF and therefore not (yet) in operational use. In this thesis, it is used to extend the analyses and diagnostics to a scheme which is not related to uncertainties in physical parametrizations, and thereby helps to generalize concepts of stochasticity and to understand how stochastic perturbations interact with specific processes.

3.2. Data sets

Throughout this thesis, a number of different forecast data sets have been used, that can be classified into two categories: (1) Research experiments with the ECMWF ensemble prediction system and (2) (re-)forecasts from operational archives.

3.2.1. Experiments

Experimental design

The research experiments have been conducted to investigate the effects of the ensemble configuration on rapidly ascending air streams and on the model climate in general (Chapters 4 and 5). This was done by running ensemble forecasts with different techniques to perturb the ensemble members, with a special focus on model uncertainty schemes.

The experiments have been initialized from 32 starting dates that are distributed in 2-day intervals between August 15 and October 15 2016. Each experiment was run out to a forecast lead time of 12 days (288 hours) and consists of 20 ensemble members which are perturbed by the techniques described in Table 3.2. The nomenclature of the experiments might raise confusion, as the experiments are partly named

Table 3.2.: Description of experiments. *SPP* and *STOCHDP* are available for 32 starting dates. As *SPP-CONV-ONLY* and *SPP-CONV-OFF* are only available for 11 dates, the diagnostics in chapter 4 are shown only for the reduced data set (11 initial times) to preserve comparability.

Abbreviation	Description	Initial times	Set
<i>SPPT</i>	ICP and SPPT	32	1
<i>IC-ONLY</i>	ICP	32	1
<i>SPPT-ONLY</i>	SPPT	32	1
<i>SPP</i>	ICP and SPP	32/11	2
<i>SPP-CONV-ONLY</i>	ICP and SPP only in conv. param.	11	2
<i>SPP-CONV-OFF</i>	ICP and SPP in all but conv. param.	11	2
<i>STOCHDP</i>	ICP and STOCHDP	32/11	2

by their perturbation technique; to clarify whether a term refers to the experiment or to the perturbation technique, the experiments are written in *italic*, whereas the perturbation techniques are written in normal font (e.g. *SPPT* refers to the experiment that is perturbed with SPPT and initial condition perturbations, whereas SPPT refers to the stochastically perturbed parametrization tendencies scheme). Overall, this sums up to approximately 22 simulation years per experiment and corresponds to 285 verification time steps at 6-hourly resolution. The forecasts are run at a resolution of TCo399, which is equivalent to a grid spacing of approximately 30 km, and 91 vertical levels, with the model top at 0.01 hPa. This setup has a coarser resolution than the operational one (TCo639 - ca 18km), but requires less computational resources and has therefore been used frequently for research and development purposes at ECMWF (e.g. in Lang et al. 2021). This setup is used frequently for research purposes at ECMWF. For further details of the setup of the experiments, please refer to Table 3.3. The study period during Autumn 2016 has been chosen to match the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) field campaign (Schäfler et al., 2018), which aimed at gathering observations from WCB outflow at the tropopause level. The period was characterized by enhanced WCB activity and several extratropical transitions of tropical cyclones in the North Atlantic sector. Subsequent to the measurement campaign, many case studies have been published, mostly combining observations with numerical simulations (see Schäfler et al. 2018 for a detailed description of the field campaign).

Description of model runs and evaluation techniques

The performance of the research experiments in the investigation period over the North Atlantic, expressed as the anomaly correlation coefficient (ACC) of Z500 on forecast day 6, is shown in Figure 3.3. The ACC is defined as

$$ACC = \frac{\sum_{i=1}^n (F_i - C_i)(O_i - C_i)}{\sqrt{\sum_{i=1}^n (F_i - C_i)^2 \sum_{i=1}^n (O_i - C_i)^2}}, \quad (3.2)$$

where F_i is the prediction, O_i the analysis and C_i the climatology at the grid point i . Typically, a value above 0.6 is considered skillful. The ensemble mean of the experiment with the quasi-operational setup (*SPPT*) is well above the threshold of 0.6 throughout the entire period. Despite this overall good performance, there are situations with reduced ensemble forecast skill at day 6, such as those initialized between September 24 and October 4, and individual members frequently drop below the critical threshold. This indicates that the predictability strongly depends on the large-scale flow situation (i.e. flow-dependent predictability; Ferranti et al. 2015). The experiment *IC-ONLY* is characterized by a very similar evolution of the ensemble-mean ACC. The interquartile range of *IC-ONLY* is slightly narrower than the one of *SPPT* and reflects that the deviations between the individual members are reduced without *SPPT*. In contrast, the evolution of *SPPT-ONLY* strongly differs from the other two experiments, as

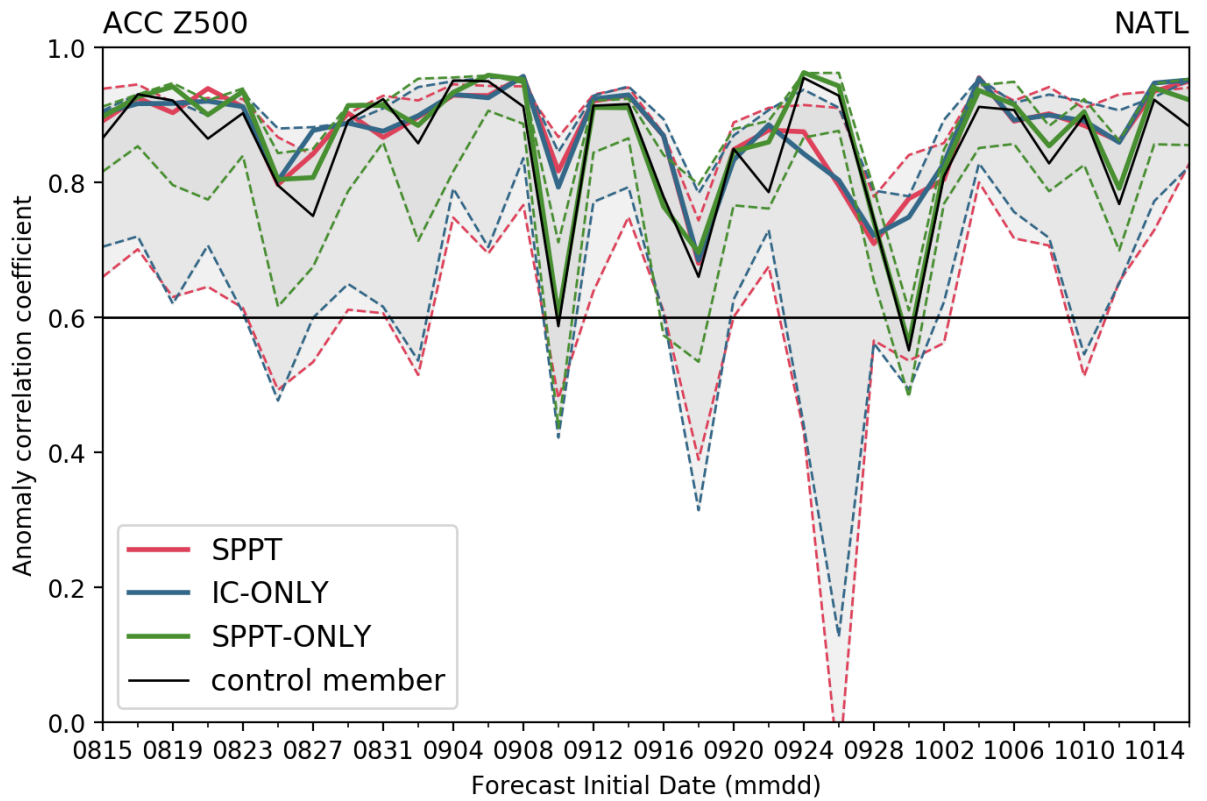


Figure 3.3.: 6-day (144h) anomaly correlation coefficient (ACC) of geopotential height at 500 hPa (Z500) in the North Atlantic domain (70° - 10° W, 25° - 65° N) of the experiments *SPPT* (red), *IC-ONLY* (blue) and *SPPT-ONLY* (green) of the 32 initial dates between August 15 and October 15 2016. The thick line represents the ensemble mean, and the grey range depicts the 5-95 interquartile range. The black line shows the unperturbed control member.

the ensemble mean has larger variations to both higher and lower values of the ACC, and by a strongly reduced interquantile range. The latter reflects the strong impact of initial condition perturbations on the generation of ensemble spread in the extratropics (Leutbecher and Palmer, 2008). The overall evolution of the forecast skill of the ensemble experiments resembles the one from the operational forecasts issued in 2016 reported in Schäfler et al. (2018), hence the experiments can be considered suitable for the investigations in this thesis.

The spread-error relationship goes a step further and evaluates the reliability of an ensemble forecasting system: the spatially averaged ratio of ensemble spread and ensemble-mean root-mean squared error (RMSE) shows whether the ensemble is overconfident (i.e. has too little spread compared to the error), or if it is underconfident (i.e. has too much spread). Averaged over many forecasts, a perfectly reliable ensemble (which does not exist) has a spread-error relationship of 1, which would enable a quantification of the forecast skill without a verification (Leutbecher et al., 2017). The ensemble spread is defined as the standard deviation of a variable with respect to the ensemble members, and the RMSE is computed as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (F_i - O_i)^2}, \quad (3.3)$$

with F_i being the predicted value at grid point i , O_i the corresponding verifying analysis, and N the number of grid points in the verification domain.

On average, all research experiments used in this study are to some extent underdispersive (i.e. have too little spread, see Figure 3.4). In the extratropics (panel a), the experiments with both initial condition perturbations and model physics perturbations through SPPT (red line) or SPP (yellow line) have the largest ensemble spread compared to the RMSE and are therefore the closest of all experiments to the diagonal line, which indicates perfect reliability. Switching off model physics perturbations results in slightly decreased values of spread (*IC-ONLY*, blue line), while switching off initial condition perturbations results a very large decrease of ensemble spread (*SPPT-ONLY*, green line). Compared to *SPPT*, the spread of *STOCHDP* is slightly decreased, but it is still larger than of *IC-ONLY*.

In the tropical regions, the spread-error relationships of the experiments are affected in a different way than in the extratropics: While the experiments with both initial and model physics perturbations are again the closest to the diagonal line, now the initial condition perturbations are not as effective in generating spread (*IC-ONLY*) as the stochastic physics perturbations. Perturbations to the dynamical core (*STOCHDP*) even lead to a reduction of ensemble spread. Further, despite the values of both spread and error are reduced by approximately one order of magnitude in the tropics, the spread-error relationship

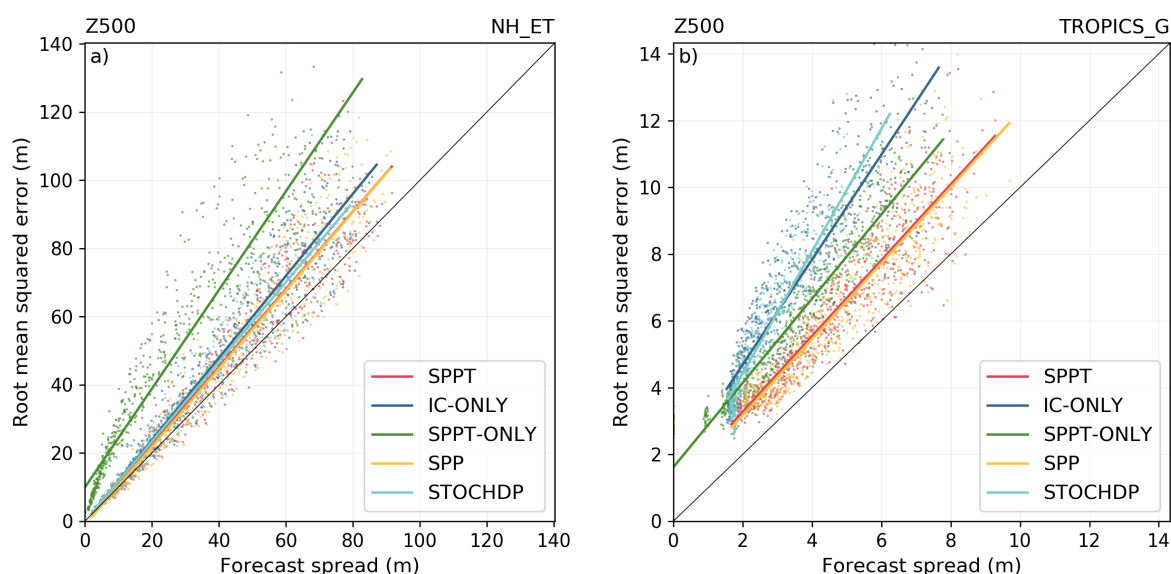


Figure 3.4.: Spread-error ratios (points) and linear fits (lines) of geopotential height at 500 hPa in the (a) Northern hemisphere extratropics (180°W - 180°E , 30° - 90°N) and (b) Tropics (180°W - 180°E , 20°S - 20°N) of the experiments *SPPT* (red), *IC-ONLY* (blue), *SPPT-ONLY* (green) and *SPP* (yellow).

is even more underdispersive than in the extratropics.

Consistently with literature (e.g. Leutbecher et al. 2017), model physics perturbations are very effective in generating additional ensemble spread in the tropical regions, where large parametrization tendencies occur. In the extratropics, initial condition perturbations are the dominant source of ensemble spread, whereas SPPT and SPP contribute only little to the overall ensemble spread. Note that the estimation of the ensemble reliability usually requires a much larger number of cases than what is shown in this thesis, and it is usually done with reforecasts covering multiple years of forecasts. The figures shown here shall solely illustrate the concept of ensemble reliability and evaluate their response to different perturbation techniques.

3.2.2. Forecast archives

During the course of the thesis project, it was found that the effects of SPPT on the rapidly ascending air streams and the model climate can also be investigated by comparing the unperturbed control member to the perturbed members of ECMWF's ensemble forecasts (explained in detail in Chapter 4). Therefore, also operational forecasts and data from forecast archives are used in this thesis. Further, data from operational forecasts are also used to analyze the impact of WCBs on forecast errors in Chapter 6.

Operational ECMWF ensemble forecasts

We use data from operational medium-range ECMWF ensemble forecasts, initialized twice daily (00 and 12 UTC) between December 1, 2018 and February 28, 2021. Hence, the data set contains 3 winter (DJF) seasons and each 2 seasons of spring (MAM), summer (JJA) and autumn (SON). The forecasts are retrieved for lead times up to 12 days in a domain ranging from the North American west coast to eastern Europe. Compared to the research experiments, the operational forecasts are run at a higher spatial resolution (TCo639, equivalent to approx. 18 km in the extratropics) and with 50 perturbed ensemble members. This data set is used for two different purposes: The perturbed and unperturbed forecasts are compared in Chapter 4 to corroborate the findings from the research experiments in a larger data set. In Chapter 6, the relationship between WCBs and forecast error growth is investigated. More details on the data set can be found in Table 3.3 and in Haiden et al. (2021).

Subseasonal to seasonal (S2S) reforecasts

20 years of ECMWF reforecast data from the subseasonal to seasonal (S2S) prediction project database (Vitart et al., 2017) is used to test the robustness of the effect of stochastic perturbations on the upper-level Rossby wave amplitude in Chapter 5. Until the forecast lead time considered in this thesis (up to 15 days), the model is run at the same resolution as the operational medium-range forecasts (TCo639). The forecasts are initialized twice per week and consist of 10 perturbed ensemble members. The model versions are not consistent across the whole data set, but no major differences (such as updated resolution, model physics or perturbation schemes) are apparent. Further details can be found in Table 3.3 and in Vitart et al. (2017).

3.2.3. Verification data sets

For verification purposes, a range of different data sets is used throughout the thesis. The rapidly ascending air streams in the IFS research experiments in Chapter 4 are compared against ECMWF's high-resolution analysis that is interpolated horizontally and vertically to the grid of the experiments

Table 3.3.: Detailed information on the data sets used in the thesis.

	Experiments	Operational forecasts	S2S reforecasts
IFS version	CY46R1	CY45R1/CY46R1/CY47R1	CY43R1/CY43R3/CY45R1
Model resolution	TCo399L91	TCo639L91	TCo639L91
Output grid spacing	1x1°	1x1°	1x1°
Temporal resolution	6 hours	6 hours	24 hours
Perturbed members	20	50	10
Simulation period	Aug 15 - Oct 15, 2016	Dec 1, 2018 – Feb 28, 2021	1997 - 2017
Start interval	48 hours	12 hours	Twice per week
Max. lead time	288 hours (12 days)	288 hours (12 days)	360 hours (15 days)
Analysis time steps	285	2,961	7,666 (including JJA)

(ECMWF, 2019c). The operational high-resolution analysis obtained from a 4D-var data assimilation (Rabier et al., 2000) is employed for the verification of WCB-trajectories in the operational ensemble forecasts in Section 4.1.3. The effect of model uncertainty schemes on the model climate (Chapter 5) is evaluated against the ERA5 global reanalysis (Hersbach et al., 2020). For the computation of the error metrics in Chapter 6, a "pseudo"-analysis is constructed from operational ECMWF short-range forecasts. Fields that are valid at 00 and 12 UTC are compared against the unperturbed initial conditions of the forecasts initialized at 00 and 12 UTC. The valid times 6 and 18 UTC are compared to the 6-hour forecasts initialized at 00 and 12 UTC, respectively. This is done to obtain consistent vertical levels between forecasts and analysis for the evaluation of variables on model levels.

3.3. Lagrangian WCB detection

Throughout the entire thesis, WCBs and other rapidly ascending air streams are detected in a Lagrangian framework by computing air parcel trajectories based on the 3-dimensional wind of model output.

3.3.1. Lagranto

We employ the **LAGR**angian **AN**alysis **TO**ol (Lagranto) developed by Wernli and Davies (1997) and updated by Sprenger and Wernli (2015). Lagranto iteratively solves the trajectory equation

$$\frac{D\mathbf{x}}{Dt} = \mathbf{u}(\mathbf{x}), \quad (3.4)$$

where $\mathbf{x} = (\lambda, \phi, p)$ denotes the position vector in geographical coordinates and $\mathbf{u} = (u, v, \omega)$ is the 3-dimensional wind. Starting at time t and position \mathbf{x} , the first iteration of the new position \mathbf{x}^* at time $t + \Delta t$ is computed as

$$\mathbf{x}^* = \mathbf{x} + \mathbf{u}(\mathbf{x}, t) \cdot \Delta t. \quad (3.5)$$

Hence, only the wind information of the starting position is used for the first iteration. For the subsequent iterations, the wind vector is averaged between the starting position and the in Equation 3.5 estimated ending position by

$$\mathbf{u}^* = \frac{1}{2}[\mathbf{u}(\mathbf{x}, t) + \mathbf{u}(\mathbf{x}^*, t + \Delta t)]. \quad (3.6)$$

The averaged wind vector is then used to obtain the updated trajectory position by computing

$$\mathbf{x}^{**} = \mathbf{x} + \mathbf{u}^* \cdot \Delta t. \quad (3.7)$$

The number of $*$ indicates the iteration step. In its default version, Lagranto uses 3 iterations to estimate the final position of the trajectory. The time step Δt is $\frac{1}{12}$ of the temporal increment between the input

data files. For 6-hourly data (as used in this thesis), the time step is 30 min. To ensure accurate estimates of the air parcel trajectories, the input fields must have a spatial resolution of at least 1° in the horizontal and 10 hPa in the vertical, and the time interval between two fields should not be longer than 6 hours (Bowman et al., 2013).

3.3.2. Procedure for WCB-detection

Trajectory computation

WCBs are qualitatively often described as "coherent bundle of rapidly ascending air streams". The Lagranto tool can be used to detect such ascending air streams by computing forward trajectories that start in the lower troposphere and reach the upper troposphere within a specific time interval (e.g. Madonna et al. 2014b). The first step of the procedure is the definition of appropriate trajectory starting points. As WCBs typically originate from the boundary layer, the height layer between 1000 and 700 hPa in 25 (experiments) and 50 hPa steps (operational data) were chosen. The horizontal starting positions are seeded on a 100 km equidistant grid distributed over the whole globe (experiments) and over an extended North Atlantic domain (operational data, see Table 3.4 for details), respectively. The trajectories are then computed forward in time for 48 hours, and only those trajectories that ascend by at least 600 hPa (i.e. $p_{\max} - p_{\min} \geq 600 \text{ hPa}$) are retained. An example of WCB-trajectories computed from forecast data and the surrounding synoptic situation is shown in Figure 3.5a. Along the trajectories, meteorological variables are traced by a 3-dimensional interpolation of the corresponding fields from the regular model output grid to the trajectory positions. This allows for the evaluation of quantities and processes (e.g. latent heating) along the ascending air streams. Trajectory computations are started every 6 hours. As the computation started at a specific forecast time needs 48 hours of subsequent data, the last time step at which trajectories can be started is the latest available lead time minus 48 hours (e.g. when 12 days of forecast data is available, trajectories can only be started up to lead times of 10 days).

Gridding

In order to ease the analysis of these Lagrangian features, we apply a technique to transform the trajectories from their 4-dimensional phase space to 2-dimensional objects that correspond to pre-defined height levels and to a particular time step. Thereby, each trajectory point is assigned to the inflow, ascent or outflow phase of a WCB, based on its pressure: Trajectory points with a pressure larger than 800 hPa are assigned to the inflow phase, air parcels with a pressure between 800 and 400 hPa are attributed to the ascent phase, and trajectory points with a pressure less than 400 hPa correspond to the outflow phase. The trajectory positions in the corresponding WCB phase are then interpolated to a regular $1^\circ \times 1^\circ$ grid, based on a radius of 100 km around the trajectory point. For each layer, grid points on the regular grid that are touched by this circle are assigned the value 1 (if multiple trajectories touch a grid point, this is

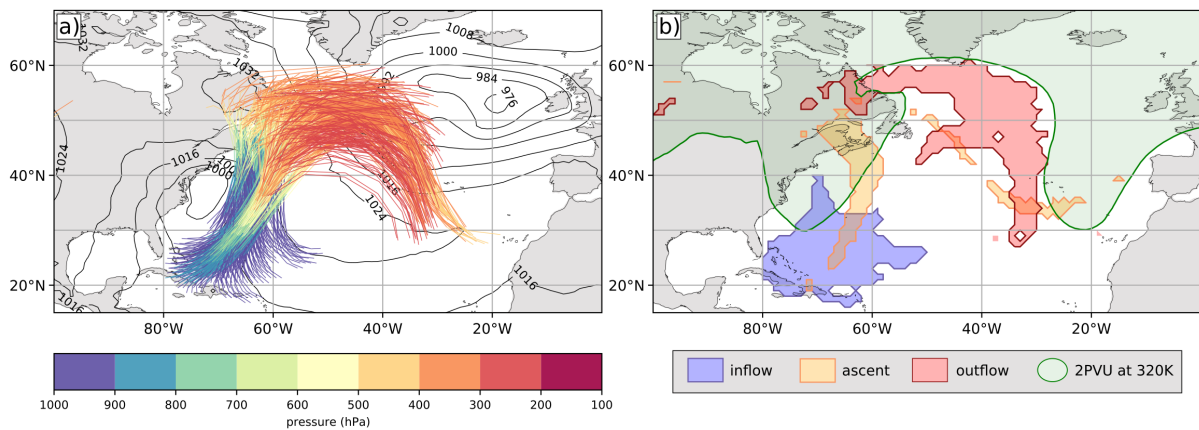


Figure 3.5.: a) 48-hour forward trajectories (colored lines) started at forecast leadtime 48 hours of a forecast initialized on Jan 31st 2021 at 00 UTC, and mean sea level pressure (black contours) at the time of trajectory start. Only trajectories that fulfill the ascent criterion of 600 hPa in 48 hours are plotted. b) WCB-masks of inflow (blue), ascent (orange) and outflow (red) valid at lead time 48 (inflow), 72 (ascent) and 96 hours (outflow) of the same forecast as in a). The green line shows the 2 PVU contour at the isentropic level of 320 K.

not considered).

To obtain imprints of the WCB-phases at a specific valid time, all trajectories started in a time period of 48 hours or less before the valid time have to be considered. For example, when the imprint of WCB-outflow at time t shall be evaluated, all trajectories that can possibly be located in the outflow layer at the corresponding time (i.e. trajectories that have been started in the time interval between $t-48$ hours and t) have to be considered. This implies that the data period for which the WCB imprints are calculated is reduced by another 48 hours both in the beginning and at the end of the forecast. The former can be avoided using the wind fields from (re)-analysis data for the computation of trajectories that have been started before the forecast initialization. Figure 3.5b gives an example of the gridding technique: the colored areas denote imprints of the different WCB phases that are obtained from trajectory data. Note that the three phases of the WCB are not evaluated at the same valid time, but are shifted by 24 hours relative to each other to match the inflow, ascent and outflow stages shown in panel a). The 2 PVU contour (valid at the same time as the outflow mask) is plotted to provide a context of the large-scale flow configuration associated with WCB outflow. The binary fields of WCB inflow, ascent and outflow obtained from the gridding technique are referred to as "masks" or "imprints" throughout the thesis.

Lagranto setup

The trajectory computation with Lagranto is sensitive to the resolution of the input fields. Sensitivity tests with different spatial and temporal resolutions of the input data as well as two different model resolutions showed that the trajectory count increases when the grid spacing and time increments of the input fields is small (Figure A.1c in Appendix A). The model resolution also affects the trajectory

count (more trajectories at a higher resolution), but the effect is less pronounced. The highest trajectory count is found for the combination of a TCo639-simulation with 0.25° and 1-hourly output, whereas the combination of a TCo399-simulation with 1.0° and 6-hourly output results in the lowest trajectory counts. Also the representation of physical quantities differs with the configuration of the input fields: higher spatial and temporal resolutions of the input fields result in higher latent heating rates (Figure A.1a) and higher outflow pressures (Figure A.1b) of the trajectories. This is probably due to a more accurate representation of the physical and dynamical processes along the ascending air streams on the finer grid. Despite these sensitivities, all trajectories in this thesis are computed with 6-hourly and 1.0° model output. The motivation for this choice is that the differences between the experiments with the different model uncertainty schemes (results shown in Chapter 4) do not depend on these settings. The higher resolution for the trajectory computation is therefore not beneficial for the research purpose in this project, but is actually disadvantageous, as the computational effort associated with both the computation of the trajectories as well as their analysis is largely increased. To ensure comparability among the different data sources, all data sets are regridded to the same 1.0° grid for the trajectory computation, independent of the original model resolution. For other research questions, however, it might be beneficial to use a setup with a higher resolution.

3.3.3. Automated / Real-time implementation

The Lagrangian detection of WCBs in large data sets requires a very efficient workflow, mainly because of two reasons:

- Prior to the selection of those trajectories that fulfill the chosen criterion (e.g. ascent of at least 600 hPa in 2 days), trajectories have to be initialized from all possible starting positions. Considering the standard setup used in this thesis (starting points seeded on a global 100 km equidistant grid every 25 hPa between 1000 and 700 hPa), roughly 850.000 trajectories have to be computed per timestep, which sums up to more than 730.000.000 trajectories for a 21-member ensemble forecast up to 10 days lead.
- An accurate estimation of the trajectories is only assured when the input fields are provided in sufficient spatial and temporal resolution. Especially the vertical resolution is crucial, and the data is ideally provided on model levels. The IFS-versions used in this thesis have 91 vertical levels, of which 53 levels represent the height levels between 1000 and 100 hPa (i.e. the levels that are required to represent WCBs). Considering the 5 output variables needed for the trajectory computation (u , v and ω) and for the computation of physical quantities (T and Q), a single time step on a global $1^\circ \times 1^\circ$ grid with 53 levels requires about 35 MB in grib2-format, resulting in about 30 GB per ensemble forecast.

Table 3.4.: Details about the trajectory data sets.

	Sensitivity experiments	Operational forecasts
Trajectory starting points	100 km equidistant grid	100 km equidistant grid
Trajectory starting region	Globally	-130°W - 60°E, 15°-60°N
Vertical levels	1000 to 700 hPa every 25 hPa	1000 to 700 hPa every 50 hPa
Selection criterion	600 hPa in 48 hours	550 and 600 hPa in 48 hours

Automated postprocessing of ensemble experiments

In order to tackle the high demands for both computational and disk storage resources associated with the WCB-detection in the ensemble experiments, a postprocessing workflow has been implemented into the IFS-suite: every time the integration of an individual forecast (i.e. a single member) of an experiment is finished, the trajectory computation is automatically started. The postprocessing is performed on the same HPC-system where the ensemble forecasts run, such that no raw data has to be transferred to and stored on an offline file system. The trajectory computation is implemented in an embarrassingly parallel¹ framework which enables to obtain the WCB-data shortly after the forecasts have been produced. Apart from the trajectory computation, the postprocessing consists of several other steps, such as the preprocessing of the data files for Lagranto, the gridding procedure and the retrieval and computation of variables on surface, pressure and isentropic levels. The last step of the workflow consists of a transfer of the postprocessed data to a permanent data directory and the cleanup of the HPC cluster.

Real-time processing of operational forecasts

Model level data of operational ensemble forecasts is not archived in the Meteorological Archiving and Retrieval System (MARS) of ECMWF, but is only available for a few days after the dissemination of the forecast. Therefore, model level data has to be collected in real-time in order to build an ensemble data archive. This is done twice daily (00 and 12 UTC) in the "Large-Scale Dynamics and Predictability" group at KIT since December 2018, yielding a unique data set of more than 3 years of operational ensemble forecasts with a high vertical resolution. After the scheduled data retrieval, the trajectory computation is performed automatically and starts when the data retrieval is finished. To save disk space and computational effort, the data is only retrieved in a domain ranging from the North American west coast to eastern Europe (see Table 3.4 for details).

¹The tasks are manually splitted into individual sub-tasks that are executed independently.

4. The impact of ensemble configuration on rapidly ascending air streams

Ensemble prediction aims to quantify uncertainties in the initial condition estimation and model formulation. Most of the commonly used operational model uncertainty schemes account for deficiencies related to parametrized processes (e.g. SPPT and SPP, see Chapter 3 Data and Methods). These schemes, especially SPPT, are designed to introduce stochastic noise where parametrizations are active and output large tendencies. During the ascending phase of WCBs, several processes at different spatio-temporal scales, including sub-grid diabatic processes, result in large parametrization tendencies. This becomes evident when mid-tropospheric (700–500 hPa) temperature tendencies from physical processes are overlaid with masks of rapidly ascending air streams (Figure 4.1). Therefore, the question arises whether the configuration of the ensemble affects the representation of WCBs in ensemble forecasts. Here we present a systematic evaluation of the sensitivity of WCBs on details of the initial condition and model error representation in the ECMWF EPS. First, the ECMWF’s operationally used ensemble setup is investigated with research experiments and data from operational forecast in Section 4.1. In a second part (Section 4.2), the analysis is extended to other schemes that are currently under development and not

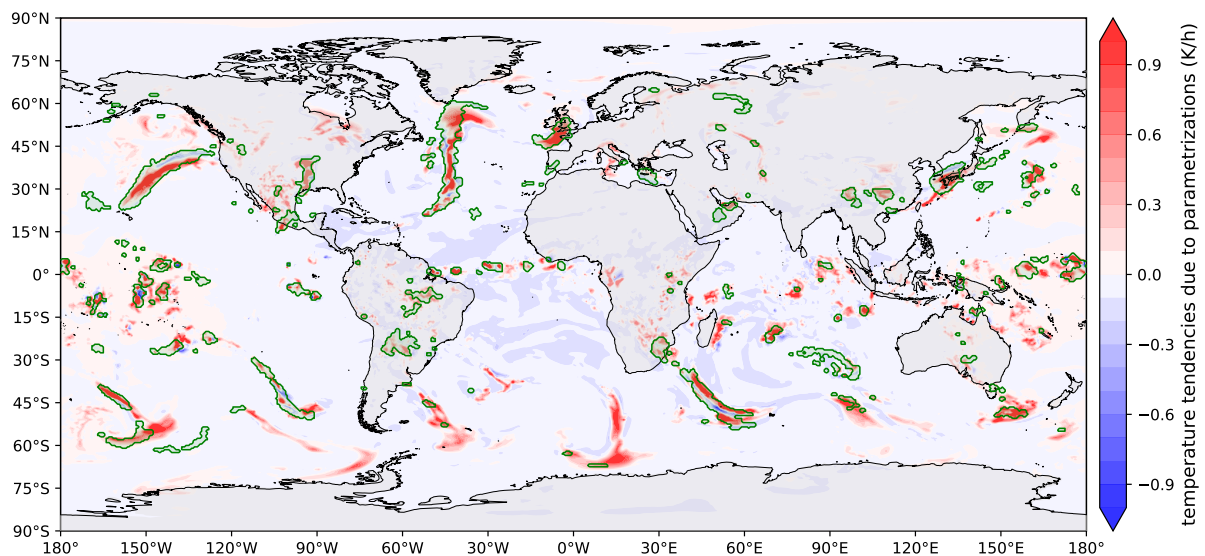


Figure 4.1.: One-hour accumulated temperature tendencies due to parametrizations (shading every $0.1 \text{ K}\cdot\text{h}^{-1}$) averaged over the model levels 105–96 (approx. 700–500 hPa) from the ERA5 short-term forecast initialised at 18 UTC on March 8, 2016 at lead time 6 hours and imprints of rapidly ascending air streams (ascent of at least 600 hPa in 48 hr) in their ascending stage between 800 and 400 hPa from ERA5 reanalysis data valid at 0000 UTC on March 9, 2016 (green contour). Reprinted from Pickl et al. (2022).

(yet) in operational use, and finally a mechanism explaining the observed effects is introduced (Section 4.17).

4.1. Sensitivities to operational schemes

We here present insights from sensitivity experiments with uncertainty representations used operationally at ECMWF (SPPT and initial condition perturbations) regarding the representation of rapidly ascending air streams. At first, a trajectory-based diagnostic is used to analyse the experiments, followed by an Eulerian perspective. Afterwards, the robustness of the findings are evaluated by investigating operational ensemble forecasts.

4.1.1. Lagrangian perspective

Gridded frequencies

Figure 4.2a) shows the mean gridded frequencies of trajectories reaching the upper troposphere above 400 hPa during the simulated period in the interpolated analysis (*ANA*). The threshold of 400 hPa reflects the outflow stage of WCBs in the Extratropics (e.g. Madonna et al. 2014b) and other diabatically enhanced air streams (termed "diabatic outflow" or simply "outflow" in the following, cf. Grams and Archambault 2016). Consider that the experimentation period covers the period from mid-August to the end of October 2016, corresponding to boreal summer/autumn and austral winter/spring. In the northern hemisphere Extratropics, two regions of enhanced diabatic outflow are visible in the storm tracks of the North Atlantic and North Pacific, with maximum frequencies of up to 20% in the North Atlantic and up to 15% in the North Pacific. Enhanced outflow is also present in the South Pacific and Atlantic basins, with occurrence frequencies comparable to those in the Northern Hemisphere. Additionally to these typical WCB regions, also the Tropics show enhanced frequencies of diabatic outflow, especially the tropical Pacific and Atlantic, with widespread regions exceeding frequencies of 25%. Due to the global distribution of trajectory starting points, also rapidly ascending air streams related to tropical convection and tropical cyclones are detected (and not only WCBs).

The outflow frequencies in the experiment *SPPT*, computed as average over all forecast initialisations, lead times, and members, resemble the patterns from the analysis (Figure 4.2b), even though the fields are much smoother due to the 20-member ensemble. However, the difference between *SPPT* and *ANA* reveals that *SPPT* underestimates the outflow frequencies in the North Atlantic sector as well as in most parts of the Tropics, especially over the Maritime Continent (Figure 4.2c). In the extratropical North Pacific and the southern hemisphere storm tracks, the outflow frequency is overestimated in *SPPT*. Deactivating SPPT (i.e. *IC-ONLY*) leads to improved outflow frequencies where *SPPT* has too much outflow, while the simulations with activated SPPT perform better in regions where the frequencies are

underestimated by the model (e.g. in the North Atlantic and the Tropics, see Figure 4.2d).

The comparison of *IC-ONLY* and *SPPT* reveals that the outflow frequency is systematically reduced globally without model uncertainty representations through *SPPT* (Figure 4.2e). The reduction is largest in the regions where the absolute frequencies are highest, i.e. in the tropical regions (exceeding 5%) and in the North Atlantic and Pacific (2-4%). The sign of the signal is negative everywhere except for

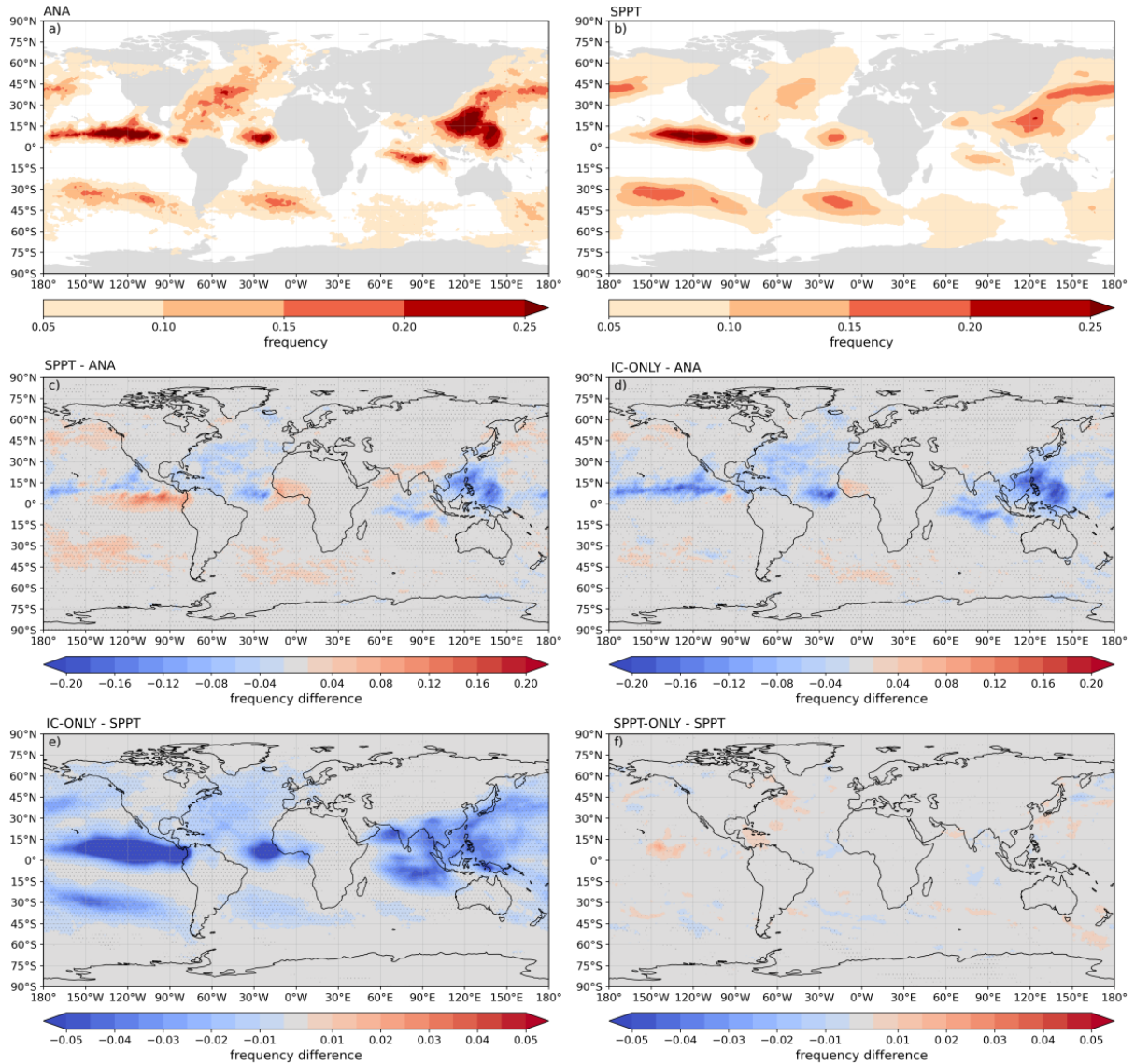


Figure 4.2.: Frequency maps of trajectories reaching the upper troposphere above 400 hPa over all analysis times (285 time steps) for the interpolated analysis (a), and over all forecasts (32), lead times (41), and ensemble members (20) for experiment *SPPT* (b), the difference between *SPPT* and the interpolated analysis (c), the difference between *IC-ONLY* and the interpolated analysis (d), the difference between *IC-ONLY* and *SPPT* (e), and the difference between *SPPT-ONLY* and *SPPT* (f). The stippling in c)-f) denotes statistically significant differences between the datasets at a confidence level of 0.99 based on a χ^2 -test. Note that the displayed differences refer to absolute differences of the frequencies. Adapted from Pickl et al. (2022).

a few grid points at high latitudes in regions with very low absolute frequencies. Hence, the reduced frequencies in *IC-ONLY* compared to *SPPT* are systematic. The comparison of *SPPT* and *SPPT-ONLY* reveals that no systematic frequency changes are introduced by the IC-perturbations (Figure 4.2f). The effect of SPPT on the other stages of rapidly ascending air streams (inflow and ascent) is similar to the outflow stage (see Figure B.1 in Appendix B).

Trajectory counts

To further investigate differences between the experiments, the number of trajectories in different regions on the globe are counted. The regions are chosen to separate tropical (tropical convection) from extratropical trajectories (WCBs). Further, a special focus is set on the North Atlantic, as it is a very prominent WCB region impacting European weather. A trajectory is assigned to a region when its starting position lies within the boundaries of the region, which are indicated in Table 4.1. The counts are normalized by the maximum possible number of trajectories in each region (i.e. the number of all starting points in that region before only those trajectories are retained that ascend). The numbers in the different regions thereby become comparable, as they do not rely on the size of the region.

The distributions of the normalized trajectory counts over all forecasts, lead times, and perturbed members in all regions are depicted in Figure 4.3. Globally, the median trajectory count in *IC-ONLY* is around 37% less than in *SPPT*. In absolute numbers, this is a reduction from around 2009 trajectories in *SPPT* to 1271 in *IC-ONLY* per time step. In the Tropics, the difference between the medians is much larger (reduction by 60%), while the effect is less prominent in the Extratropics (21 % in the northern hemisphere Extratropics and 32% in the North Atlantic). Trajectory counts in *SPPT-ONLY* do not differ from the ones in *SPPT*; not only the medians, but also the distributions agree very well, showing that IC-perturbations do not systematically affect the occurrence of WCBs. Moreover, the median counts of the unperturbed control member (dashed lines) are indistinguishable to the medians of *IC-ONLY*.

The variability in the different regions strongly depends on the size of the region: globally and in the Tropics, the variability is very low, as there are always sub-regions with enhanced vertical velocities on the globe. In the smaller North Atlantic domain, WCB activity depends on single transient weather

Table 4.1.: Boundaries of regions for which trajectory counts and characteristics are computed.

Region	Abbreviation	Region boundaries
Global	G	-180° -180° E, -90° - 90° N
Global Tropics	TROPICS_G	-180° -180° E, -20° - 20° N
Northern hemisphere Extratropics	NH_ET	-180° -180° E, 30° - 90° N
North Atlantic	NATL	-70° -10° E, 25° - 65° N

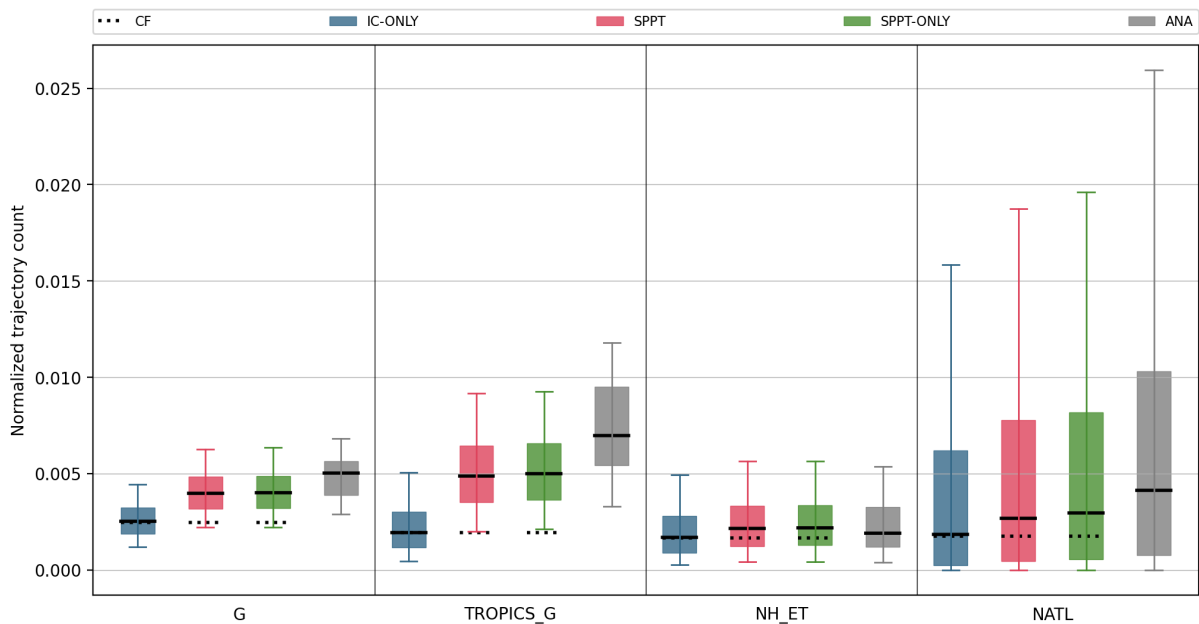


Figure 4.3.: Trajectory counts starting in the global domain (G), the tropical belt (TROPICS_G), the northern hemisphere Extratropics (NH_ET), and the North Atlantic sector (NATL) for the experiments *SPPT*, *IC-ONLY*, *SPPT-ONLY*, and the interpolated analysis (*ANA*). Counts are normalized by the maximum number of possible trajectories in each domain (i.e. all trajectories before the selection procedure). The solid line is the median of the distribution, the boxes denote the inter-quartile range and whiskers the 5-95 inter-quantile range. The dashed line is the median of the unperturbed control member (CF). Adapted from Pickl et al. (2022).

systems. Hence, the lower edge of the distribution has the value 0, as there are situations without any trajectories fulfilling the ascent criterion in the domains. On the other hand, synoptic situations with strong WCB activity lead to much higher relative trajectory numbers in the smaller regions. Additionally, the reduced sample size in the smaller regions leads to larger variability.

Note that the differences in the trajectory counts are given in relative numbers, whereas the differences of the gridded frequencies in Figure 4.2 are expressed as absolute values. Thus both approaches qualitatively show the same picture and are of the same order of magnitude. Remaining quantitative differences are due to the trajectory gridding algorithm that accounts for trajectories only once per corresponding grid point.

The median trajectory counts of the unperturbed control member (CF) are very similar to the ones in *IC-ONLY* in all regions. This is a logical consequence of the fact that the control member does not receive any perturbations, and that the *IC*-perturbations do not systematically influence the frequency of rapidly ascending air streams. This is a very useful information, because most modelling centers issue an unperturbed forecast as part of their operational ensemble forecast. Therefore, studies investigating some of the effects of model uncertainty schemes may not necessarily rely on expensive experiments,

but can also be conducted by exploiting operational ensemble forecasts.

The systematic increase of rapidly ascending air streams by SPPT raises the question whether the NWP model with unperturbed or with perturbed physics produces a better representation of WCBs compared to a verifying analysis. To investigate this, the same trajectory count diagnostic has been applied to the interpolated analysis (*ANA*, grey bars in Figure 4.3). Globally, all experiments underestimate the occurrence of rapidly ascending air streams during the investigated period compared to the interpolated analysis (gray bars in Figure 4.3). This underestimation arises predominantly from the Tropics, where the trajectory counts are strongly reduced in the experiments. Also in the North Atlantic, the experiments simulate too few rapidly ascending air streams. In contrast, the trajectory counts are slightly overestimated in the northern hemisphere Extratropics. As SPPT systematically increases the trajectory count, the offset between the simulations and the interpolated analysis is increased in the experiments without SPPT in regions where the trajectory counts are underestimated by *SPPT* (Tropics, North Atlantic). In regions where the counts are overestimated by *SPPT* (northern and southern hemisphere Extratropics), the deactivation of SPPT makes the trajectory frequencies more consistent with the interpolated analysis. Due to these regional differences, a general statement whether SPPT improves or deteriorates the representation of rapidly ascending air streams cannot be made.

The number of trajectories decreases with lead time in all experiments (Figure 4.4; note that the time axis corresponds to the time of the trajectory start). This explains the reduced trajectory counts in most parts of the globe, compared to the verifying analysis. In the global domain, starting at a level close to the counts in the interpolated analysis, the number of trajectories in *SPPT* and *SPPT-ONLY* rapidly decreases by about 25% during the first 48 hours of the forecast and then stays constant. The counts in the *IC-ONLY* experiment show a very similar behaviour, but the curve has an offset to lower values. The number of trajectories in the simulations decreases with lead time and consequently the underestimation compared to the analysis increases. This lead-time dependency is similar in the other investigated regions and shown exemplary for the northern hemisphere Extratropics (dotted lines). In that region, however, the simulations with perturbed model physics through SPPT initially overestimate the trajectory counts and level off close to the interpolated analysis at later lead times, while *IC-ONLY* initially matches the analysis and drops below during the forecast integrations.

The differences between simulations perturbed with SPPT (*SPPT* and *SPPT-ONLY*) and *IC-ONLY* stay almost constant with lead time. This indicates that SPPT directly influences the ascending air streams and does not change the mean state during the course of the forecast. If the deviation was time-dependent (i.e. very small or not existing in the beginning and increasing with forecast lead time), it would suggest that the perturbations alter the atmospheric state in a favourable way for the development of rapidly ascending air streams during the forecast - such as moisture accumulation in the boundary layer.

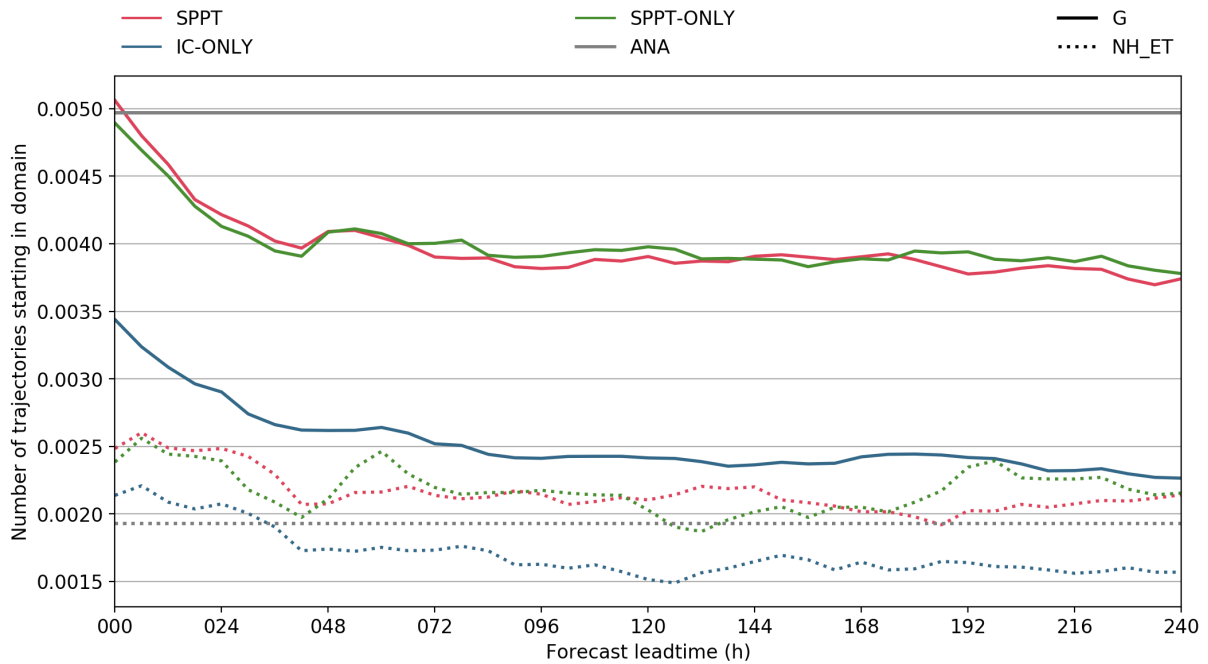


Figure 4.4.: Medians of trajectory counts starting in the global (G, solid lines) and northern hemisphere extratropical domains (NH_ET, dashed lines) with forecast lead time, normalized by the maximum number of possible trajectories in each domain, of the experiments *SPPT* (red), *IC-ONLY* (blue), *SPPT-ONLY* (green), and the interpolated analysis (*ANA*, grey). The time axis corresponds to the time of the trajectory start. Reprinted from Pickl et al. (2022).

Trajectory characteristics

As SPPT systematically changes the frequency of rapidly ascending air streams, the question arises whether the perturbations also influence their physical properties and characteristics. The Lagrangian perspective allows for the calculation of the evolution of meteorological fields along the trajectories. In this section, we focus on the latent heating rate of the trajectories, mainly because of two reasons: firstly, because the intense latent heat release in the ascending air streams is the reason why the SPPT-scheme affects them; and secondly, because the diabatic heating lifts the air parcel cross-isentropically and is therefore a key parameter which determines the height of the outflow and its impact on the large-scale circulation.

Figure 4.5 shows the frequency counts of latent heating rates (i.e. difference between maximum and minimum potential temperature) of all detected trajectories within the regions listed in Table 4.1. The distribution in the global domain shows that all analysed trajectories are heated by at least 10 K, which emphasizes the importance of diabatic processes for these ascending air streams. The global distribution of latent heating rates is dominated by two regimes in all experiments: one with a maximum occurrence centred around 45 K, and another one at lower values between 20 and 35 K. By looking at the other regions, it becomes clear that the regime associated with larger values is strongly dominated by the Tropics,

while the regime with lower heating rates is represented more by the extratropical regions. However, the heating rates in the northern hemisphere Extratropics and in the North Atlantic are mostly characterized by values which lie in between the two distinct global regimes, indicating that the lower heating regime is not explicitly depicted by the selected regions, and predominant in e.g. the southern hemisphere storm tracks (not shown).

Comparing the global frequency counts of *IC-ONLY* to *SPPT* shows that the median heating rate is strongly reduced when *SPPT* is deactivated (reduction of 7 K). However, the shapes of the histograms indicate that *IC-ONLY* primarily underestimates the frequency of strongly heated trajectories in the upper regime, while the frequencies of the less heated trajectories are similar in *SPPT* and *IC-ONLY*. Splitting the trajectories into sub-regions makes the distributions more comparable: In the Tropics, the median differences between *SPPT* and *IC-ONLY* is about 1.5 K, and in the northern hemisphere Extratropics and in the North Atlantic about 1 K. Comparing the simulated heating rates in the different regions to the interpolated analysis shows a similar effect as for the trajectory count: heating rates are underestimated globally by all experiments, but *SPPT* leads to a more realistic representation compared to simulations with unperturbed physics. The medians of the unperturbed control member (*CF*) again resemble the ones from *IC-ONLY*, and the distributions of *SPPT-ONLY* are very similar to the ones of *SPPT*. This cor-

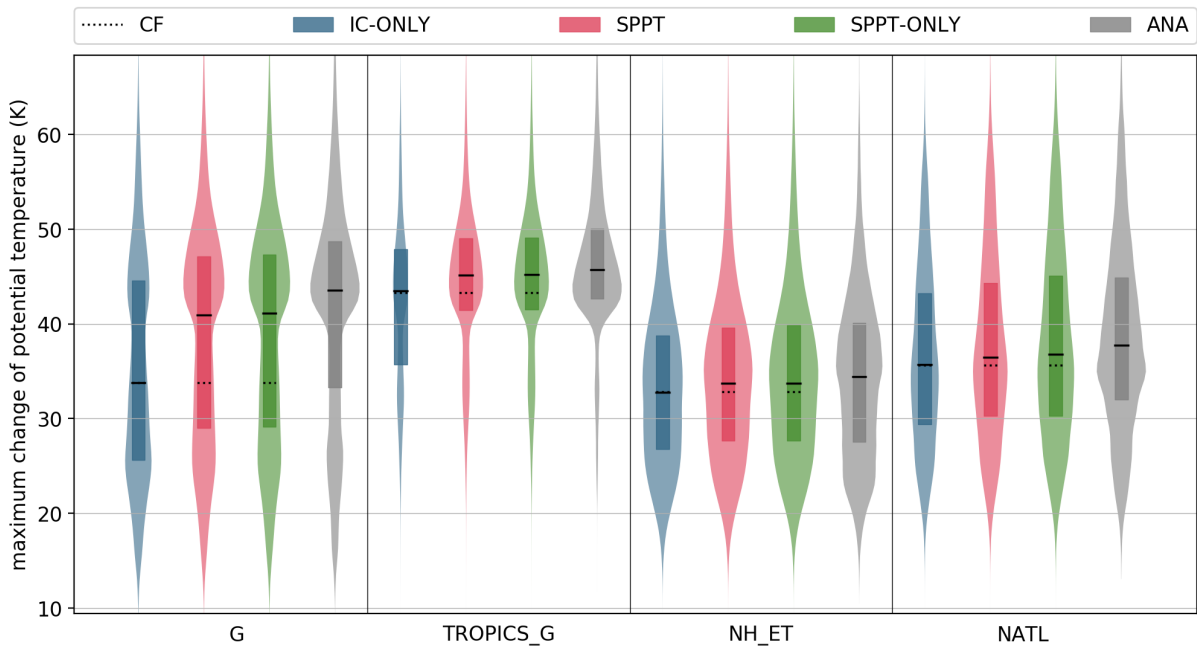


Figure 4.5.: Frequency distributions of latent heating rates along trajectories starting in the global (G), tropical (TROPICS_G), northern hemisphere extratropical (NH_ET), and North Atlantic (NATL) domains for the experiments *SPPT*, *IC-ONLY*, *SPPT-ONLY*, and the interpolated analysis (*ANA*). The black lines indicate the median and the boxes the inter-quartile ranges of the distributions. The dotted line represents the median value of the unperturbed control member. The scales of the frequency distributions are identical for all experiments within each region, allowing for a quantitative comparison between the experiments. Adapted from Pickl et al. (2022).

roborates the inferences from the trajectory count diagnostics regarding the potential use of the control member from operational forecasts.

A more physical classification is to assign the trajectories to one of the two global heating regimes, independent of their origin. The local minimum between the two heating maxima lies approximately at the value of 38 K in *SPPT* (see Figure 4.5); hence, this value is chosen as threshold to classify each trajectory either into the upper or lower category. For each of the two regimes, several characteristics have been computed, and the trajectory counts and heating rates are shown in Figure 4.6. For the lower heating rates, the deviation of the trajectory count in *IC-ONLY* from *SPPT* is much smaller (11%) than for higher heating rates (52%; panel a). Comparing the experiments to *ANA* shows that the trajectory count is slightly overestimated in the lower heating regime, with a larger positive bias in the experiments with *SPPT* than without *SPPT* (*IC-ONLY*), while it is strongly underestimated in the upper heating regime, where the experiments with *SPPT* are closer to *ANA* than without *SPPT*. The temporal evolution of the trajectory counts (Figure 4.6c) illustrates that the number of the strongly heated trajectories rapidly drops during the first 2 days of the forecast, and *SPPT* helps to keep the counts closer to *ANA*. In contrast, the number of the less heated trajectories is maintained throughout the forecast, and *IC-ONLY* accurately represents the counts of *ANA*, while the experiments with *SPPT* overestimate the counts.

A substantially different behaviour between the two heating regimes is not observed for the characteristics of the trajectories: The difference in latent heating between *SPPT* and *IC-ONLY* is very small and does not depend on the heating regime (Figure 4.6b). Table 4.2 summarizes the median values of trajectory characteristics (trajectory count, latent heating rate, isentropic and isobaric outflow level, and specific humidity at trajectory start) derived from the trajectories and gives an overview over the effects of *SPPT* on different aspects of rapidly ascending air streams in the two heating regimes. Similarly to the heating rates, the trajectory characteristics are only changed marginally or even remain unchanged when *SPPT* is activated. For example, the minimum pressure of the trajectories increases by 1.7 hPa in the lower regime and decreases by 1.5 hPa in the upper regime when *SPPT* is deactivated.

The large difference between the heating rates in *SPPT* and *IC-ONLY* (see Figure 4.5) is therefore mainly a result of changed trajectory frequencies in the different regions. Globally, rapidly ascending air streams occur more frequently with *SPPT* than without, and this effect is stronger for large heating rates than for small heating rates. Hence, the global statistics of trajectory characteristics will take the shape of stronger heated air streams, which results in larger heating rates and outflow heights (not shown). Also regionally (e.g. in the North Atlantic sector), *SPPT* increases the WCB frequency more strongly in the (sub-)tropical part of the domain than in the poleward part. However, comparing trajectories with similar heating rates (and not based on their origin) shows that individual trajectories are not substantially modified by *SPPT*. This is in agreement with the conclusions drawn from Figure 4.4: *SPPT* does not change the environ-

4. The impact of ensemble configuration on rapidly ascending air streams

mental conditions of rapidly ascending air streams or affect the trajectory characteristics significantly, but directly acts on the ascending motions and helps to initiate the rapid ascents.

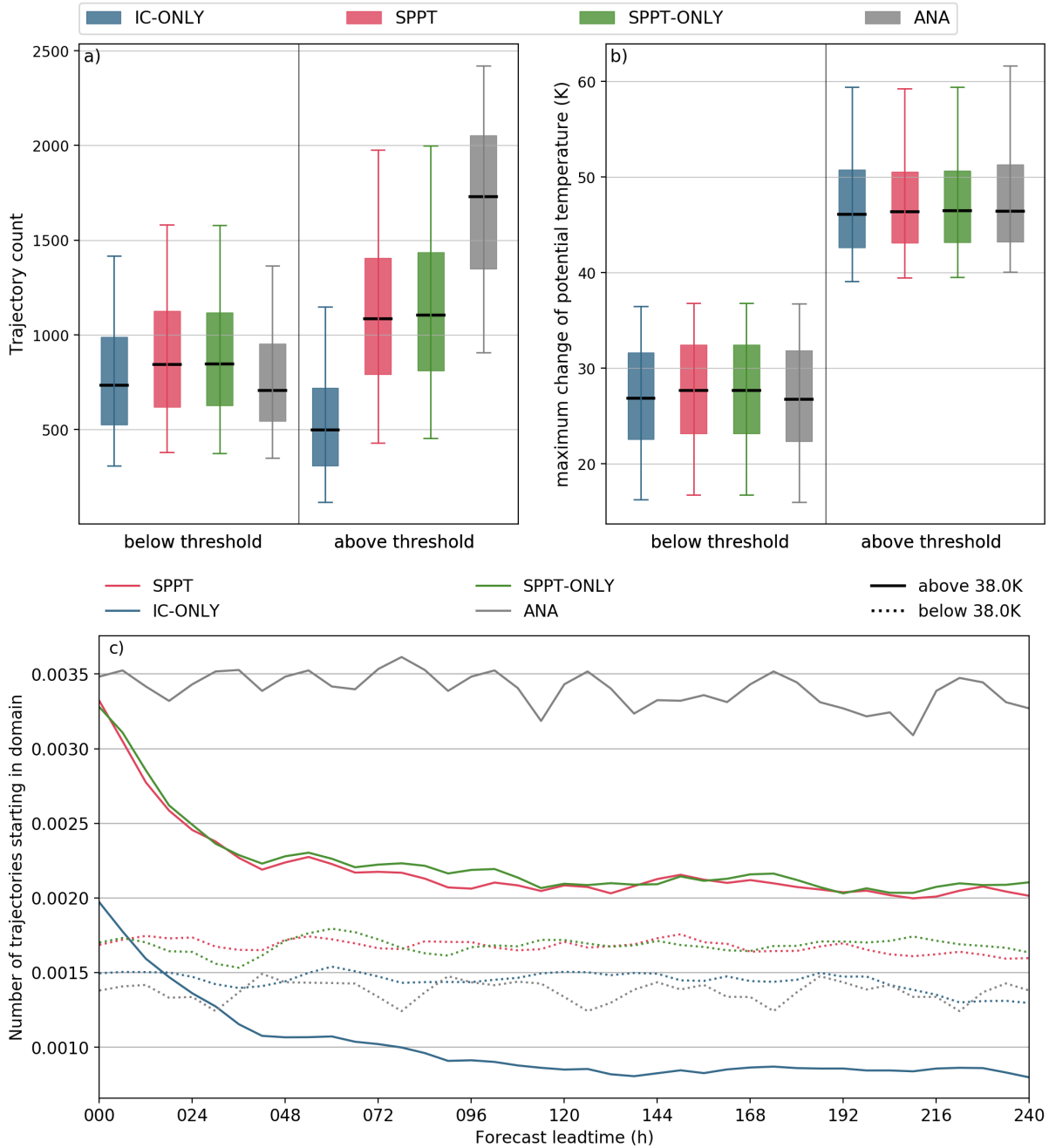


Figure 4.6.: a) Trajectory count, b) latent heating rate, and c) trajectory count as a function of forecast leadtime of all trajectories (global) in the experiments *SPPT*, *IC-ONLY*, *SPPT-ONLY*, and the interpolated analysis (*ANA*) separated for the two latent heating regimes below (left) and above (right) the threshold of 38 K. Adapted from Pickl et al. (2022).

Even though the physical properties of the trajectories do not substantially differ, the latent heating rates help to differentiate between strongly and weakly heated trajectories. The frequency of strongly heated rapidly ascending air streams is generally underestimated in the experiments compared to the analysis, while the trajectories in the lower heating regime are slightly overestimated. Therefore, by increasing the frequency of rapidly ascending air streams, SPPT makes the trajectory counts more consistent with ANA in the upper heating regime and less consistent in the lower heating regime.

Table 4.2.: Median values of various variables calculated along trajectories for two heating regimes "Lower" with latent heating rates below 38 K and "Upper" with latent heating rates above 38 K of the experiments *SPPT*, *IC-ONLY*, *SPPT-ONLY*, and the interpolated analysis (*ANA*)

Variable	Heating regime	SPPT	IC-ONLY	SPPT-ONLY	ANA
Trajectory count	Lower	851	735	854	707
	Upper	1109	497	1129	1729
Latent heating rate (K)	Lower	27.7	26.9	27.7	26.74
	Upper	46.4	46.1	46.5	46.4
Isentropic outflow level (K)	Lower	324.6	323.7	324.7	323.7
	Upper	349.4	349.9	349.5	349.7
Outflow pressure (hPa)	Lower	288.1	289.8	288.0	291.6
	Upper	206.4	204.9	205.8	201.8
Inflow specific humidity (g/kg)	Lower	11.2	11.0	11.2	10.7
	Upper	14.9	15.1	14.9	15.1

4.1.2. Eulerian omega-perspective

The trajectory-based analysis in the previous section helps to detect and understand sensitivities of rapidly ascending air streams to the perturbations introduced by SPPT. However, the analysed trajectories consider only the most rapid ascents and therefore cover only a very small fraction of the whole spectrum of (upward) vertical velocities. The slowest vertical velocity which is needed to fulfill the ascent criterion of 600 hPa within 2 days corresponds to constant 0.35 Pa/s, but about 95% of the mid-tropospheric upward velocities are slower than this threshold (Figure 4.7a). Additional to this scale-sensitivity, only (net) upward velocities can be detected by the trajectory diagnostics and instantaneous downward motion is not considered at all. In order to take into account the slower upward velocity scales as well as downward motion, distributions of mid-tropospheric (i.e. at 500 hPa) vertical velocities for all forecasts, leadtimes and perturbed ensemble members are computed based on all grid points in the global domain (Figure 4.7a). The shape of the frequency distributions indicates that slow velocities occur far more often than fast velocities (note the logarithmic scale), and that ascending (negative ω) motions are typically faster than descending (positive ω) motions (skewed distribution).

Comparing the distributions among the experiments shows that *IC-ONLY* has less occurrences of fast upward velocities in the range from -3 to -0.5 Pa/s compared to *SPPT* (Figure 4.7a and b). This re-

fects the reduced number of trajectories fulfilling the ascent criterion of 600 hPa in 2 days shown in the previous chapter. This increased occurrence of very rapid ascents is balanced by a decreased number of grid points of moderate ascents between -0.5 and -0.1 Pa/s (Figure 4.7b), showing that very rapid ascents occur more often with SPPT than without SPPT, at the expense of slower ascents. Additionally, this acceleration of the upward velocities is balanced by increased occurrences of downward motions faster than approximately 0.05 Pa/s.

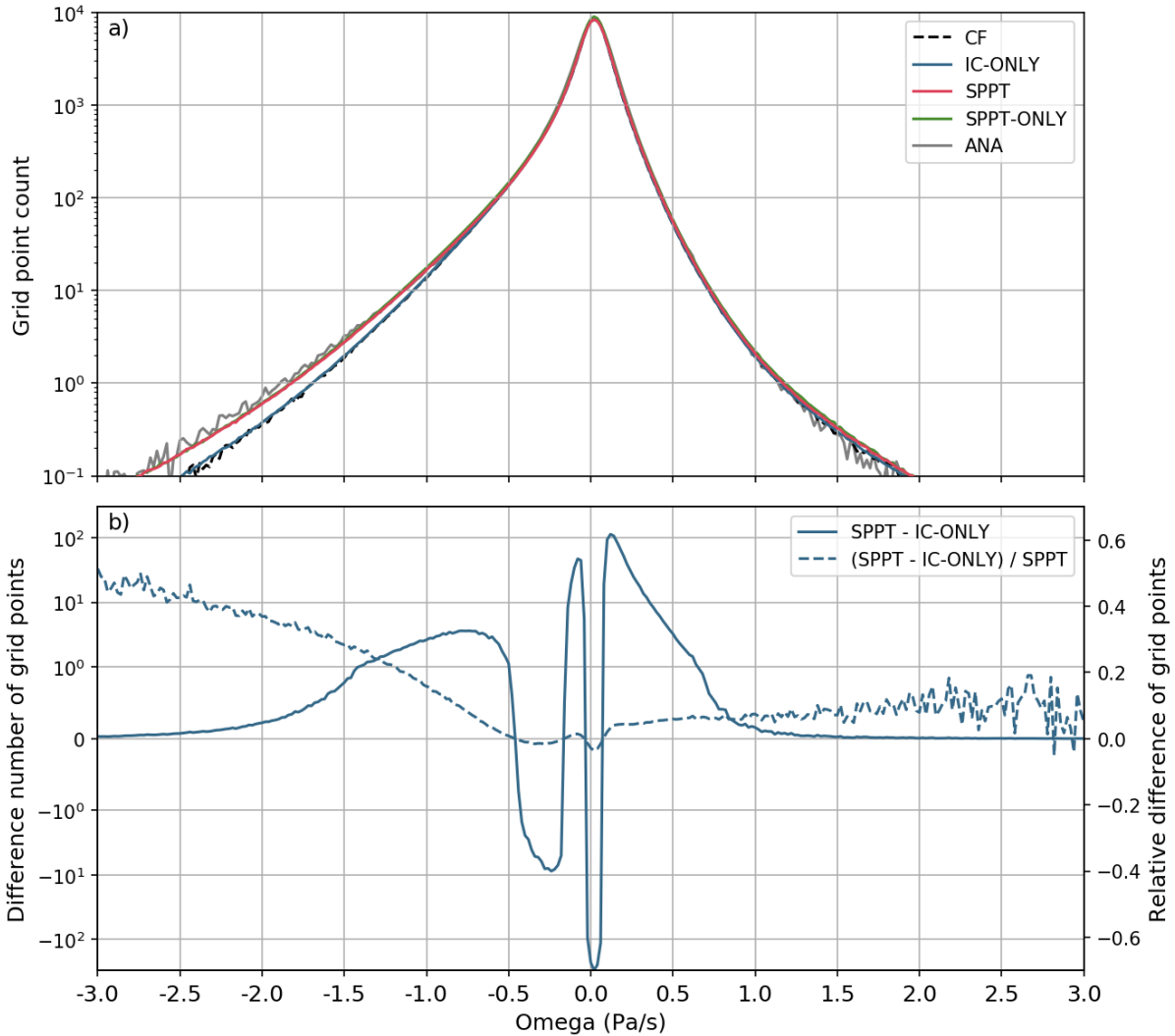


Figure 4.7.: a) Histogram of global (G) vertical velocities at 500 hPa in bins of width 0.02 Pa/s per forecast, lead time, and ensemble member for the experiments *SPPT*, *IC-ONLY* and *SPPT-ONLY*, the unperturbed control member (CF), and the interpolated analysis (*ANA*). b) shows the absolute (solid, left axis) and relative (dashed, right axis) differences between the histograms of *SPPT* and *IC-ONLY* in panel a). Negative (positive) omega values correspond to upward (downward) motion. Note that the left y-axis has a linear scale between the values -1 to 1 , and a log scale for values smaller than -1 and larger than 1 . Adapted from Pickl et al. (2022).

Additionally to these changes of the fast vertical motions, also the occurrence of grid points with very slow vertical motions are affected: The number of grid points that are in rest with respect to vertical mo-

tions (i.e. with vertical velocities close to 0 Pa/s) is reduced, and both upward and downward velocities of small magnitudes (i.e. upward motions between -0.1 and -0.01 Pa/s and downward motions faster than 0.05 Pa/s) occur more often with SPPT than without. In other words, SPPT accelerates air parcels without vertical velocities both up- and downwards and thereby makes the atmosphere more active.

In relative numbers (dashed line in panel b), the frequency changes of vertical velocities through SPPT are mainly relevant for the very rapid ascents: Upward velocities faster than -0.75 Pa/s occur approximately 10% more often with than without SPPT, whereas the relative decrease of moderate ascents through SPPT amounts to approximately 1%.

Summing up, the effect of SPPT on vertical motions is threefold:

1. The occurrence of very rapid ascents (faster than -0.5 Pa/s), mainly related to tropical convection and WCBs, is increased at the expense of moderate ascents.
2. Air parcels that do not have any or only very small vertical velocities are accelerated such that there are more up- and downward motions of slow velocities with than without SPPT
3. The downward velocities are accelerated throughout the whole spectrum of velocities.

This Eulerian diagnostic using instantaneous fields of ω at 500 hPa helps to understand that not only the very rapid upward motions are affected (i.e. which was previously diagnosed using trajectories), but all scales of vertical motion are influenced by SPPT. Similarly to the increased occurrence of rapid ascents, downward motions are enhanced systematically by SPPT. These changes in the descending part of the omega spectrum are likely to be a consequence of the altered upward motions and occur for reasons of mass balancing: if there is an increased mass flux from the lower to the upper troposphere, the model has to balance this by an increased downward mass flux. Apart from these considerations on mass conservation, the increased occurrence of vertical motions of specific velocities has to happen at the expense of other velocities for logical reasons, as the number of considered grid points is fixed in both experiments (i.e. the integral of the difference number of grid points has to be 0).

The hypothesis that the accelerated downward motions are not directly induced by SPPT, but a response to the accelerated upward motions is motivated as follows: SPPT introduces perturbations of large magnitude mainly in regions where parametrizations are active. In regions of ascending motions, parametrizations are more often active (for example when convection or cloud processes occur) than in regions of descending motions, which often takes place adiabatically. This results in larger perturbations in regions of ascents than of descents, which in turn favors non-linear responses of the model in the ascending part of the ω -spectrum. However, at the current stage of our research, this is only a hypothesis, and a better understanding requires additional experimentation.

These results based on a Eulerian perspective on vertical velocities are consistent with the previous results regarding the trajectory counts: SPPT increases the occurrence of very rapid ascents. Further, the initial condition perturbations (*SPPT-ONLY* vs *SPPT*) do not systematically affect the vertical velocities (panel a, green line almost hidden behind the red line), and the unperturbed control member (CF) behaves similarly to the experiment *IC-ONLY*. Finally, both the Eulerian and Lagrangian diagnostics agree on the underestimation of the most rapid ascents by all experiments compared to the analysis *ANA* (panel a, grey line).

Temporal behaviour of perturbations

When considering only the first time step (i.e. lead time 0h) of the global ω -distribution at 500 hPa (Figure 4.8a), a different behaviour is observed as for later lead times: at the forecast initialization, the SPPT-perturbations are not yet active; hence, the ω -distributions of *IC-ONLY* and *SPPT* are identical and indistinguishable. In contrast, the IC-perturbations have already been applied and lead to a different ω -distribution with increased vertical velocities of large magnitude compared to *SPPT-ONLY* (similar to *IC-ONLY* in 4.7a). This behaviour is apparent also for lead time 6h (not shown), but less pronounced, and vanishes subsequently. This becomes obvious when computing the lead-time dependent evolution of the area below the histograms from Figure 4.7a in the interval between -3 and -0.5 Pa/s (i.e. integrated high vertical velocities), normalized by the median (over all forecasts and lead times between 0 and 48 h) of the unperturbed control member (Figure 4.8b). Values larger than 1 denote that the experiment has on average more grid points with upward velocities larger than -0.5 Pa/s than the unperturbed control member. The threshold of -0.5 Pa/s is chosen, because the mean effect of SPPT on vertical velocities faster than -0.5 Pa/s is uniform (see Figure 4.7b).

At lead time 0h, the values of the unperturbed simulation (CF) and *SPPT-ONLY* are identical and close to 1. In both CF and *SPPT-ONLY*, no initial condition perturbations are active. As the SPPT-perturbations have not yet been applied to *SPPT-ONLY*, the two simulations are both unperturbed and result in the same distribution. In contrast, *SPPT* and *IC-ONLY* have already received (identical) initial condition perturbations, leading to larger areas below the histogram at lead time 0h (compare to panel a). After 6 hours, perturbations introduced by the SPPT-scheme increase the integrated fast updrafts in *SPPT-ONLY*, leading to larger values than the unperturbed control member. In contrast, the area below the histogram in *IC-ONLY* decreases, pointing towards a diminishing impact of the IC-perturbations. In *SPPT*, both effects (SPPT and IC-perturbations) are superimposed, resulting in a slower reduction of the histogram area than in *IC-ONLY*. After 12 hours, the lines of *SPPT* and *SPPT-ONLY* intersect, just as *IC-ONLY* and CF, indicating that the effect of the IC-perturbations has dissipated. At subsequent lead times, the experiments with SPPT (*SPPT* and *SPPT-ONLY*) as well as the experiments without SPPT (*IC-ONLY* and CF) evolve identically, which shows that the IC-perturbations exert no more systematic impact on

the vertical velocities, and that the SPPT-induced effect is constant after 6-12 hours.

The temporal evolution of the effect of IC- and SPPT-perturbations shows that both types of perturbations influence vertical velocities in a similar way. However, the IC-perturbations are applied only at the very beginning of the forecast, and the resulting instabilities leading to a shifted ω -distribution are quickly dissipated within the first hours of the integration. Hence, the effect vanishes at later lead times and is not visible in the trajectories, which are computed over a time period of 48 hours. In contrast, the perturbations introduced by SPPT are not active at the forecast initialization, but are repeatedly applied at every subsequent model time step, constantly renewing perturbations which affect the vertical velocities.

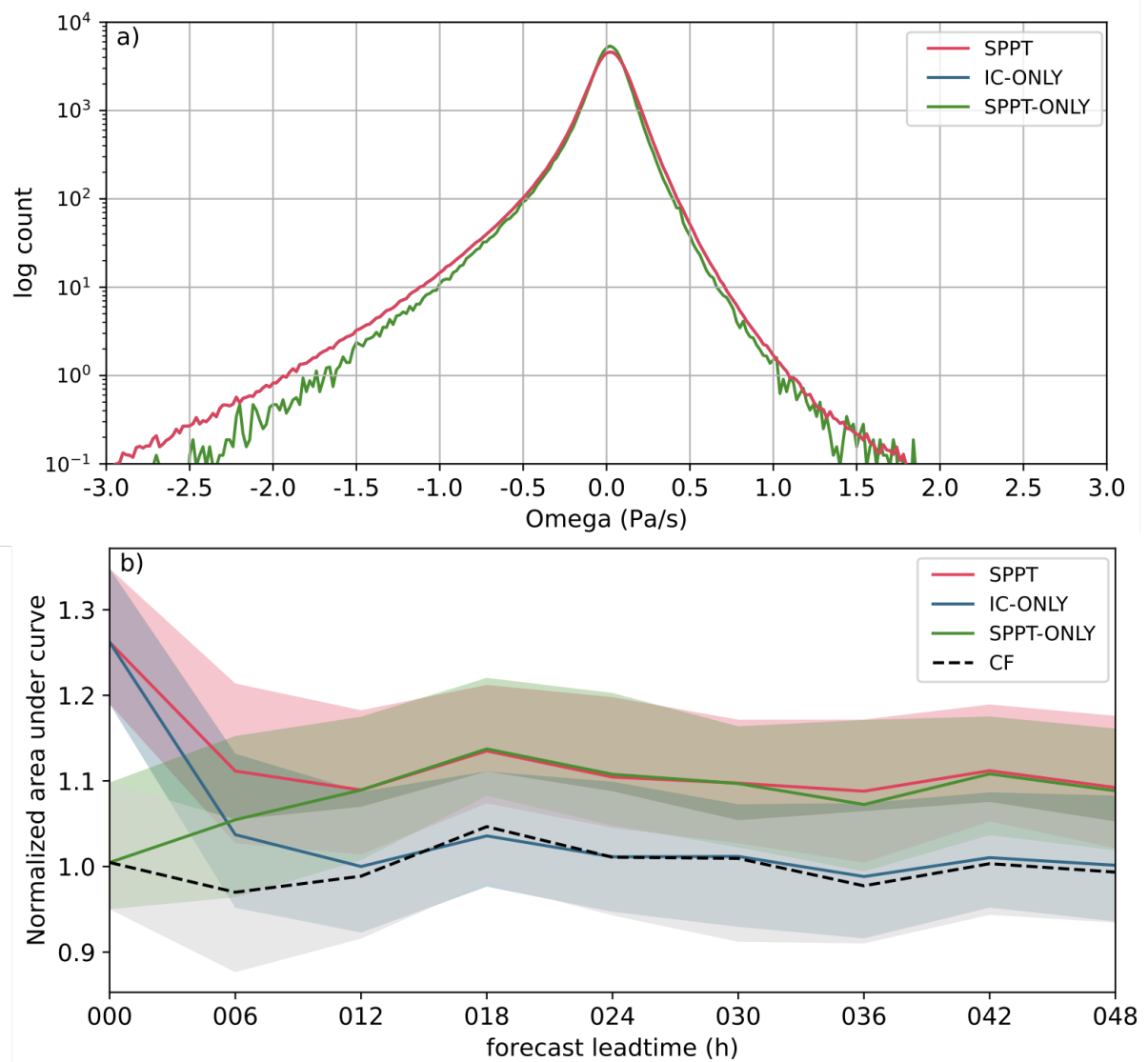


Figure 4.8.: a) As Figure 4.7a, but only for lead time 0h. b) Evolution of the median area under the curves from Figure 4.7a in the range -3 to -0.5 Pa/s with forecast lead time in 6-hourly time steps of *SPPT* (red), *IC-ONLY* (blue), *SPPT-ONLY* (green) and the unperturbed control member CF (dotted black line), normalized with the median of CF. The colored shading denotes the inter-quartile range.

This suggests that the applied perturbations and their effects on the ω -distribution are characterized by a representative life cycle. To investigate this more quantitatively, however, a higher temporal resolution of the ω -fields would be required (1-hourly).

4.1.3. Exploiting operational forecasts

In this section, we make use of the finding that the unperturbed control member (CF) shows the same behaviour regarding rapidly ascending air streams as the experiment which does not have any model perturbations (*IC-ONLY*). As the initial condition perturbations only impact the presented diagnostics in the very beginning of the simulations (see Figure 4.8b), the comparison of the control member and the perturbed (IC-perturbations and SPPT) ensemble members of operational ECMWF ensemble forecasts yields a comparable experimental design as the comparison of the experiments *SPPT* and *IC-ONLY*. This insight is very powerful, as it enables us to evaluate the effects of SPPT on the trajectories not only in research experiments, which provide only a limited data period and are computationally expensive, but also in operational forecasts across many more initial times.

A prerequisite for such an analysis is the availability of trajectory data: as the computation of trajectories requires a high spatial (especially vertical) resolution, it is not possible to use wind fields outputted on coarsely distributed pressure levels, as commonly provided in forecast archives. Ensemble forecast data on model levels has been retrieved from ECMWF servers in real time and archived locally since December 2018, which serves as a unique data set that enables the investigation of ensemble forecasts in a Lagrangian way.

This archive that is analyzed here consists of forecasts initialized twice daily within the period from December 2018 to November 2020 with 50 perturbed and 1 unperturbed ensemble member. It is run at a spatial resolution of TCo639, corresponding to about 18 km in the extratropics. The data is regrided from its original model resolution to a regular $1^\circ \times 1^\circ$ grid and retrieved in the North Atlantic region for lead times up to 12 days (288h) in 6-hourly time steps (for further details refer to Table 3.3). In this dataset, WCBs are been detected by computing 2-day forward trajectories that ascend by at least 550 hPa. Additionally to the ensemble data, trajectory data computed from the deterministic high-resolution forecast (TCo1279) is analyzed. As the trajectories are started only in the North Atlantic sector, this analysis does not include ascending air streams related to tropical weather systems, but only considers WCB-trajectories (see Chapter 3 for details).

Trajectory counts

Trajectory counts of the operational forecasts are plotted for the different years and split up into seasons according to forecast initialisation date in Figure 4.9. The seasonal cycle (higher trajectory counts in winter and autumn than in spring and summer) arises from a stronger baroclinicity and concomitant cyclone activity in the North Atlantic in the cold season (Madonna et al., 2014b). The comparison of the perturbed and control members reveals that the unperturbed forecasts have a lower median number of WCB trajectories in all seasons and in both years compared to the perturbed members. With activated SPPT, median values in winter increase by 8% in 2019 (panel a) and 14% in 2020 (panel b) and by about 20% in spring and autumn in both years; in summer, the WCB frequency in the North Atlantic is very low, and the relative changes are not meaningful. The offset between the perturbed and unperturbed members depends on the season investigated, with larger relative deviations in autumn and spring, where latent heating rates are larger than in winter. The unperturbed forecasts run at high resolution (HRES) slightly change the WCB frequencies compared to the control forecast: while the trajectory counts are decreased in winter 2019 (panel a), they are slightly increased in all other seasons and years, especially in autumn. Thus, both model uncertainty representations through SPPT as well as a higher model resolution increase WCB activity and brings it closer to the verifying analysis.

Comparing the medians of the operational analysis (unfilled boxes) with the ones from the forecasts reveals that the trajectory counts are underestimated by the forecasts for most of the initializations. In all seasons in 2020 and in summer and autumn 2019, the forecasts have lower trajectory counts than the analysis, and the offsets are largest in autumn. During winter and spring 2019, the perturbed forecasts overestimate the trajectory counts, whereas the unperturbed forecasts match the trajectory frequencies of the analysis. Nevertheless, the perturbed forecasts are on average more consistent with the analysis. This becomes evident in panel c, where the mean evolution of the trajectory counts with forecast lead time for the different seasons, averaged over both years, is shown. In winter and spring, the perturbed forecasts match the trajectory counts in the analysis, while it is systematically underestimated by the unperturbed forecasts. In autumn, where the underestimation of the unperturbed forecasts is largest, also the perturbed forecasts have too little WCB trajectories, but are closer to the analysis. Compared to the unperturbed forecasts, the high-resolution runs reduce the WCB bias in all seasons in 2020 and in autumn 2019. Nevertheless, SPPT is more effective in reducing the negative WCB bias than the increased resolution of the forecast. Despite the overall robust patterns, the different behaviour between the years 2019 and 2020 demonstrates that the impact of the ensemble configurations on WCBs is to some extent exposed to interannual variability.

When analyzing the evolution of the trajectory counts with lead time, it is evident that they differ from season to season: in the transition seasons (i.e., spring and autumn), the trajectory counts of the oper-

4. The impact of ensemble configuration on rapidly ascending air streams

ational analysis (solid lines) are not independent of the forecast lead time, because later lead times are closer to the subsequent season. Therefore, the WCB frequencies in the analysis decrease with lead time in spring, as the WCB activity is weaker in summer than in winter, and vice versa in autumn. In winter

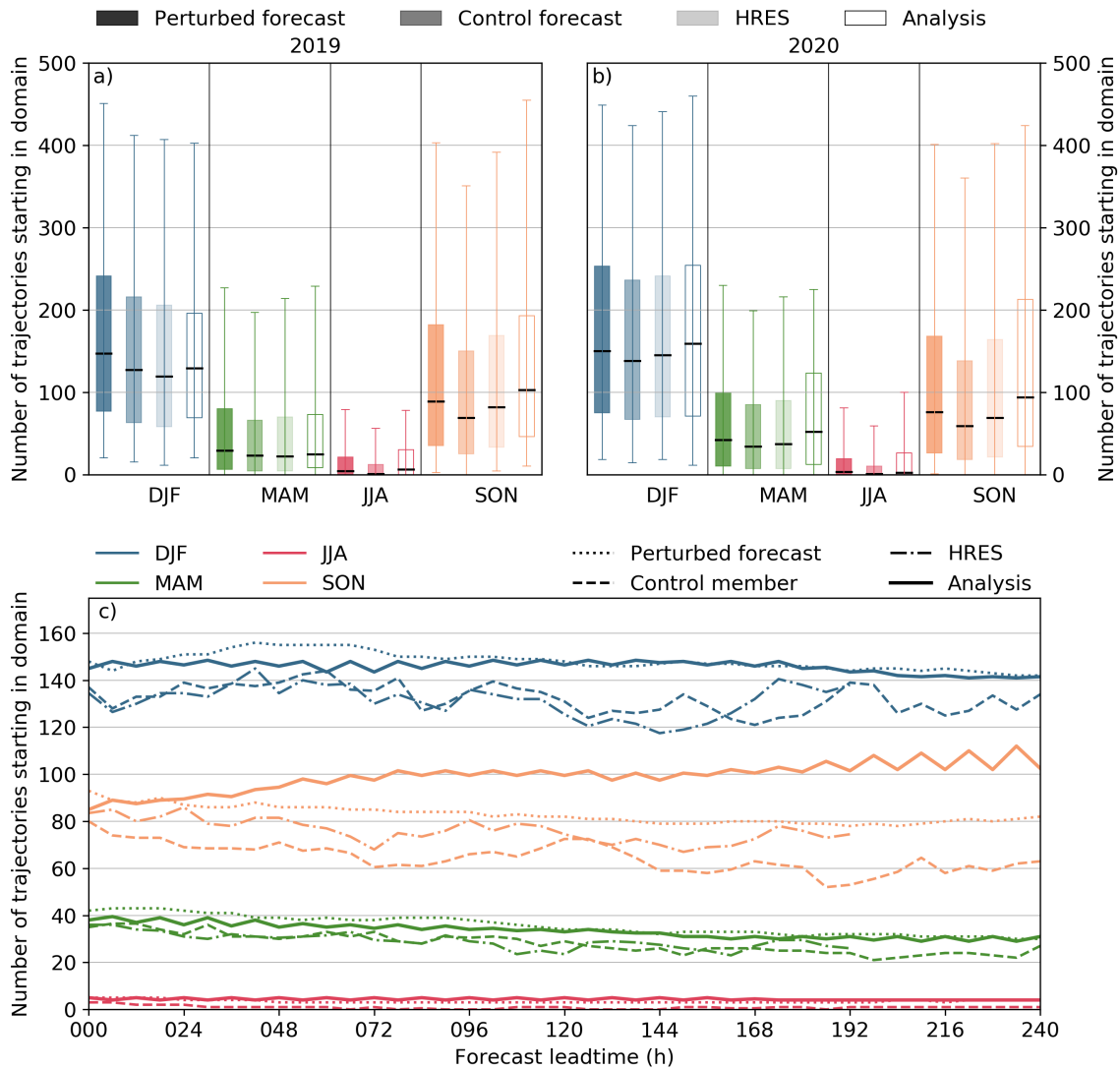


Figure 4.9.: Number of trajectories starting in the North Atlantic from operational ECMWF ensemble forecast over the years 2019 and 2020. Panel a and b show the median values (solid line) and the inter-quartile range (box) over 50 ensemble members, all lead times, and forecasts initialised in winter (blue), spring (green), summer (red) and autumn (yellow) averaged over the year 2019 (a) and 2020 (b). The darkest boxes represent the perturbed forecasts (i.e. IC-perturbations and SPPT), the dark pale boxes the unperturbed control forecasts, the light pale boxes show the high-resolution forecast and the unfilled boxes the operational analysis. Panel c is equivalent to panel a, but shows the evolution of the trajectory counts with lead time, averaged over both years. Dotted lines show the median of the perturbed ensemble members, dashed lines show the median of the unperturbed control member, dash-dotted lines show the high-resolution forecasts, and the thick solid lines represent the median of the operational high-resolution analysis. Note that for season JJA no high-resolution forecast is shown. Forecasts are classified into the seasons by the initialisation date. For winter, the indicated year refers to the year of January and February. The percental changes described in the text always refer to the median values of the data sets. Adapted from Pickl et al. (2022).

and spring, the trajectory counts of the forecasts evolve similarly to the analysis, when considering the mean differences from the operational analysis.

The evolution of the trajectory count in autumn between forecasts and analysis is contrary: while the WCB frequency increases in the analysis (due to the seasonal cycle), it decreases in the forecasts, leading to a large deviation at late lead times. The different seasonal behaviour is related to the latent heating rates along the trajectories, which are largest in autumn (and summer), reduced in spring, and lowest in winter (see Figure 4.10). This is in line with the results from the research experiments, where the differences between perturbed and unperturbed runs are larger when the heating rates are high. The offset between the perturbed (dotted lines) and unperturbed (dashed lines) forecasts is on average constant with lead time, similar to the results presented in the previous section. Note that the evolution of the trajectory counts is noisier in the unperturbed than in the perturbed case, as the time series of the former are computed with one member only.

The role of latent heating

The previous results suggest that the latent heating rate along the trajectories play a major role in modulating the impact of SPPT on the trajectories, which is further analyzed in this section. Separated for the season of forecast initialization, Figure 4.10 shows the average ratio of the number of trajectories in the perturbed and unperturbed forecasts as a function of integrated heating along the trajectories (colored lines, right axis). Values larger than 1 indicate that the trajectory frequency is larger in the perturbed than in the unperturbed forecasts, and vice versa. Additionally, the frequency counts in each heating bin are plotted (colored bars, left axis). This demonstrates that the largest heating rates occur in autumn and the lowest ones in winter. The dependency of the ratios to the heating rate is very similar for all seasons: while the ratio lies slightly below 1 for small heating rates, it slowly grows to values above 1 (17.5-20 K in winter, 20-22.5 K in autumn) and eventually increases exponentially with heating rates larger than about 25 K. The ratio is shifted to higher values in spring, where the curve lies below 1 only for heating rates below 12.5 K. For heating rates that are very large in the context of classical WCBs in the North Atlantic and mainly occur in autumn (i.e. larger than 30 K), the ratio reaches values around 1.5, corresponding to a 50% increase of the trajectory count in perturbed forecasts. Nevertheless, the ratio for a large fraction of trajectories that are heated less than 17.5 K (mainly in winter) is below 1, which reflects decreased trajectory counts in the forecasts with SPPT. This behaviour is also apparent in the gridded frequency maps, where a few grid points at high latitudes in the southern hemisphere extratropics have values larger than 1 (see Figure 4.2e).

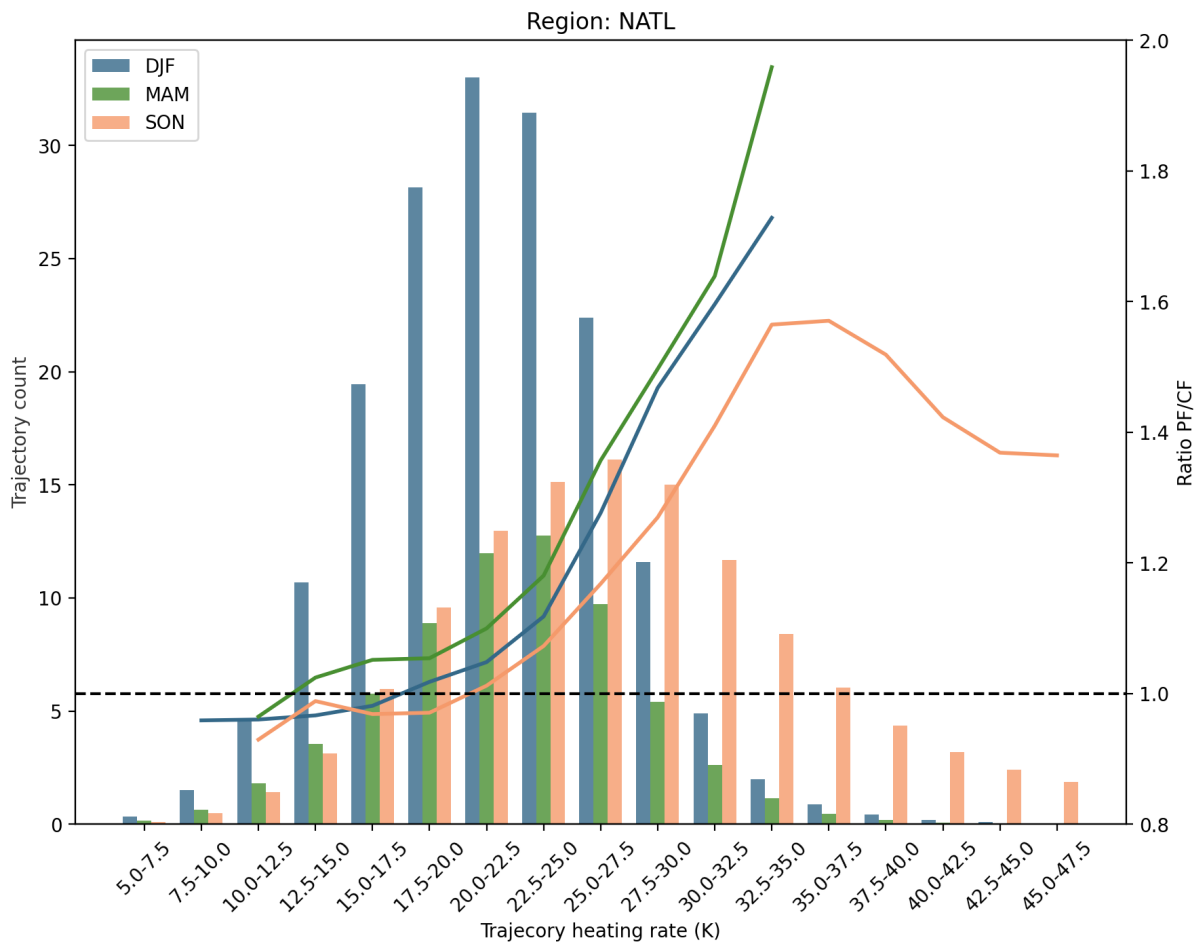


Figure 4.10.: Number of trajectories from the perturbed forecasts (bars, left axis) per time step and ratio of the trajectory count of the perturbed to the unperturbed members (lines, right axis) starting in the North Atlantic from operational ECMWF ensemble forecast over the years 2019 and 2020 separated by the integrated latent heating rate along the trajectories for the seasons DJF (blue), MAM (green) and SON (orange). The bin width is 2.5 K. Forecasts are classified into the seasons by the initialization date. Note that the ratios are only plotted when at least one trajectory with the corresponding heating rate is present.

Summary

The results presented in this section are consistent with the results obtained from the sensitivity experiments and corroborate the main findings. The differences of the trajectory counts between the experiments *SPPT* and *IC-ONLY*, with initialisations in late summer and early autumn, amount to about 30% in the North Atlantic region. This is higher than the offset in the operational data in autumn, but the qualitative signal is similar, and quantitative discrepancies arise from interannual variability and time periods considered for the analysis. The sensitivity to the latent heating rate is consistent between the operational data set and the experiments: in the experiments, the role of the latent heating rate was demonstrated indirectly by analyzing different starting regions of the trajectories (e.g., larger differences between *SPPT* and *IC-ONLY* in the Tropics than in the Extratropics) and directly by classifying the trajectories into two heating regimes. In the operational data set, the seasonal variations of the frequency differences between

perturbed and unperturbed forecasts support the hypothesis that the magnitude of the latent heating controls the effect of SPPT on the trajectories, and the effect was also directly demonstrated by stratifying the ratio of trajectory counts in perturbed and unperturbed forecasts according to their latent heating (see Figure 4.10). Finally, the deviations from the verifying analyses are also in accordance, with a larger underestimation of the trajectory frequencies for large heating rates (in the experiments) and in the warm season (operational data set), respectively.

The consistent results across the two data sets reveal that the effect of SPPT on rapidly ascending air streams is very robust: it is detectable globally (in the experiments) and within two years in all seasons (operational data). Further, the analysis shows that the effect is not sensitive to the resolution of the forecast model, at least for the resolutions investigated here. When unperturbed, both the experimental resolution of TCo399 as well as the operational resolution of TCo639 underestimate the trajectory frequencies in a quantitatively consistent way compared to the perturbed forecasts. The analysis of the operational data set further revealed that a higher resolution (HRES) slightly reduces the frequency bias of the WCB trajectories in the North Atlantic; however, the bias reduction through SPPT is larger than the bias reduction resulting from the increase of resolution. Finally, the synthesis of the two data sets showed that the effect of SPPT on rapidly ascending air streams is not sensitive to the details of the trajectory computation and the detection of the air streams: differences in the starting positions (every 25 hPa in the experiments vs every 50 hPa in the operational data) of the trajectories as well as in the selection criterion (600 hPa vs 550 hPa in 2 days) do not result in changes of the observed signal.

4.2. Sensitivities to other model uncertainty schemes

The previous sections have shown that model uncertainty representations through SPPT systematically affect the occurrence of rapidly ascending air streams and alter the distribution of vertical velocities in the model. This effect is robust across many forecast initializations, different setups of trajectory computation and for two model resolutions. In the following section, it will be investigated whether this effect occurs solely when the forecast model is perturbed with SPPT, or if it is also apparent for other model uncertainty schemes. In doing so, it will be investigated if the design of the SPPT-scheme is responsible for the observed effect, or if such a behaviour is triggered intrinsically by stochastic model perturbations and hence can be generalized.

Not all aspects from the previous sections will be repeated here. We will rather present those results which are the most insightful and synthesize the findings. It is important to note that the data basis in this section is not identical to the one of the previously shown experiments. As the main focus of the thesis was on the schemes used operationally (SPPT, IC-perturbations), the experiments with the other model uncertainty schemes (i.e. set 2) were only run for a subset of initial times (11 instead of 32). Even

though the results are robust, some of the diagnostics from the section with SPPT and IC-perturbations are omitted here (especially the ones that are confined to specific regions), as the reduced data basis results in noisy signals. Despite being slightly repetitive, the results from *IC-ONLY* and *SPPT* are again shown to enable a better comparison of the schemes.

4.2.1. Lagrangian perspective

Trajectory counts

In Figure 4.11 the trajectory counts are shown for set 2 of experiments in the different regions, equivalent to Figure 4.3. For reference, the experiment *IC-ONLY* and the analysis *ANA* are also shown. Similarly to *SPPT*, *SPP* has higher trajectory counts than *IC-ONLY* in all domains. The effect is largest (larger than in *SPPT*) in the tropics, and it is smaller, but still present in the extratropics (NH_ET and NATL). Maps of the differences of the gridded trajectories between *SPP* and *SPPT* is given in Figure B.2 in Appendix B. The *SPP* sensitivity experiments (*SPP-CONV-ONLY* and *SPP-CONV-OFF*) also show higher frequencies of rapidly ascending air streams than the unperturbed experiment in all regions; interestingly, the effect is larger in the experiment with perturbations only to the parameters in the convection

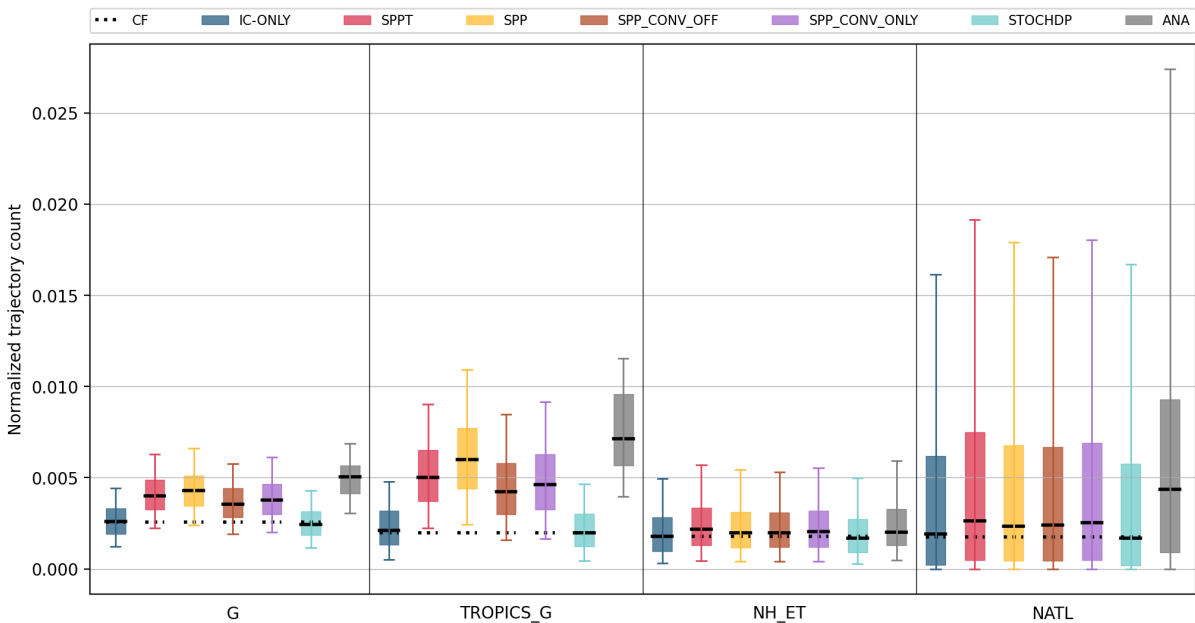


Figure 4.11.: Trajectory counts starting in the global domain (G), the tropical belt (TROPICS_G), the northern hemisphere Extratropics (NH_ET), and the North Atlantic sector (NATL) for the experiments *IC-ONLY*, *SPPT*, *SPP*, *SPP-CONV-OFF*, *SPP-CONV-ONLY*, *STOCHDP*, and the interpolated analysis (*ANA*). Counts are normalized by the maximum number of possible trajectories in each domain (i.e. all trajectories before the selection procedure). The solid line is the median of the distribution, the boxes denote the inter-quartile range and whiskers the 5-95 inter-quantile range. The dashed line is the median of the unperturbed control member (CF). Averaged over 11 initial times.

parametrizations (*SPP-CONV-ONLY*) than in the experiment with perturbations in all parameters except for convection (*SPP-CONV-OFF*). This indicates that the perturbations in the convection scheme are more efficient in triggering rapidly ascending air streams than perturbations to all other parametrizations. In the extratropics and the North Atlantic, the counts in *SPP-CONV-ONLY* are even larger than in *SPP*, pointing towards the dominant role of perturbations in the convection parametrization. The added differences of the trajectory counts of the schemes *SPP-CONV-ONLY* and *SPP-CONV-OFF* to the unperturbed experiment (*IC-ONLY*) are larger than the counts of *SPP*. This shows that the effects of the two schemes partly act on the same trajectories (i.e. are superimposed), or even cancel each other. Other than all other model uncertainty schemes analyzed in this thesis, *STOCHDP* does not increase the count of rapidly ascending trajectories, but even reduced it slightly.

The temporal evolution of the counts in the global domain with forecast lead time shows that the number of trajectories decreases during the forecasts, especially within the first 2 to 3 days, in all experiments of set 2 (Figure 4.12). The curves do not follow the exact evolution of the experiments in set 1 (c.f. Figure 4.4), as the data set contains less initial dates (11 instead of 32), which results in a noisier signal. As the pattern of the reduction is very similar in all experiments, this behaviour is intrinsic to the forecast model,

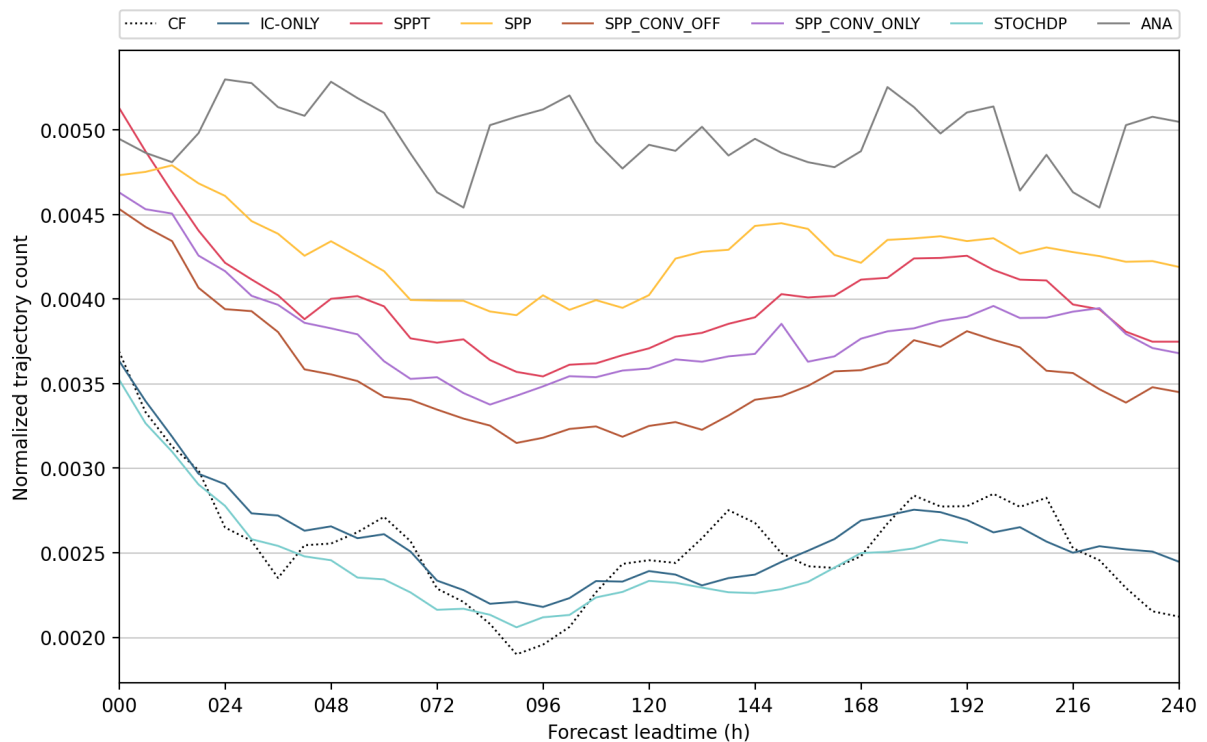


Figure 4.12.: Medians of trajectory counts starting in the global domains with forecast lead time, normalized by the maximum number of possible trajectories in each domain, of the experiments *IC-ONLY*, *SPPT*, *SPP*, *SPP-CONV-ONLY*, *SPP-CONV-OFF*, *STOCHDP* and of the interpolated analysis (*ANA*). The time axis corresponds to the time of the trajectory start. Averaged over 11 initial times.

independent from the perturbation techniques. The figure nicely depicts that, despite the offset between the experiments, the differences remain almost constant throughout the forecasts. Hence the effect of the schemes on the trajectory counts does not establish during the forecast, but is present from the time step when the perturbations are introduced. Therefore, the conclusion that the ascending motions are directly affected by the perturbations drawn from the analysis of the operational schemes also applies for the schemes in set 2.

Trajectory characteristics

Equivalent to the analysis of the operational schemes, also the trajectory characteristics, expressed by the latent heating rate along the ascents, are computed for the experiments in set 2 (Figure 4.13). Very similarly to SPPT, SPP increases the median heating rate along the trajectories globally, and this occurs mainly by a strong increase of the number of trajectories with large heating rates (i.e. larger than 38 K). Compared to SPPT, the global-mean effect is slightly larger with SPP than with SPPT, whereas it vanishes in the Northern Hemisphere Extratropics and the North Atlantic. The latent heating rates of the experiments with perturbed parameters only in specific parametrizations show a similar behaviour as the trajectory counts: perturbations to the parameters in the convection scheme result in a larger increase of strongly heated trajectories than perturbations to the other parametrization schemes, especially in the Tropics. In the extratropics, there are no detectable differences between the heating rates in *SPP-CONV-OFF* and *SPP-CONV-ONLY*. The heating rates of trajectories in *STOCHDP* are equivalent to the ones

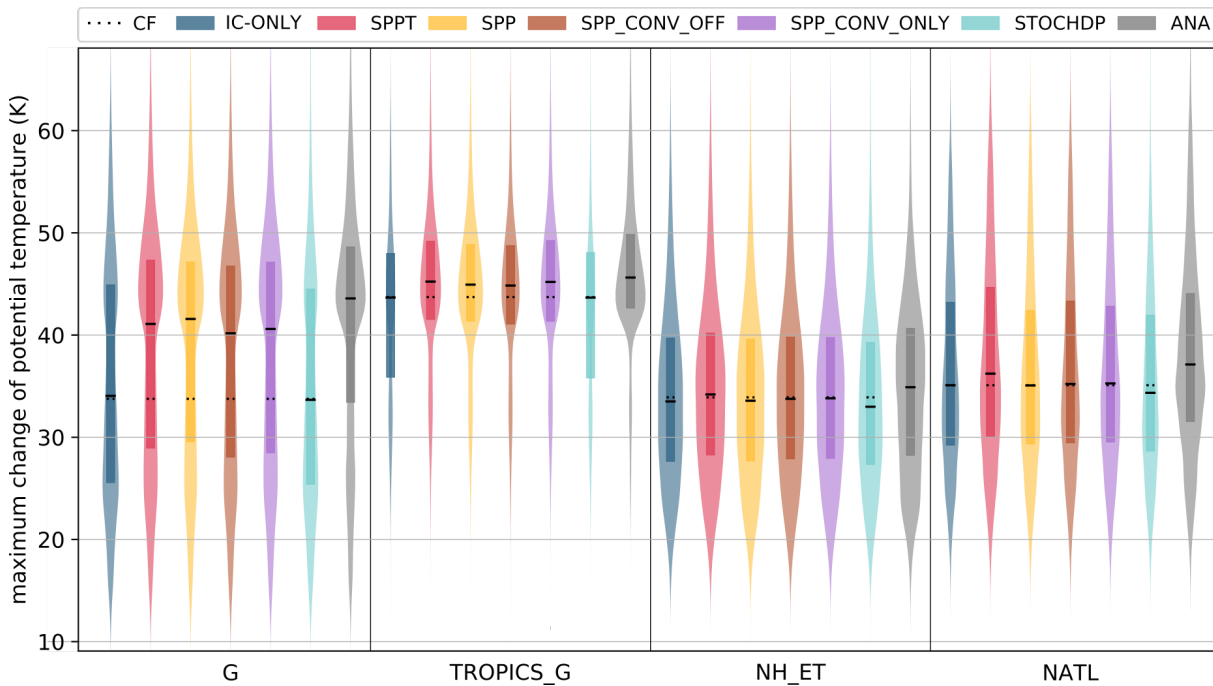


Figure 4.13.: Frequency distributions of diabatic heating rates of trajectories from the experiments *IC-ONLY*, *SPPT*, *SPP*, *SPP-CONV-ONLY*, *SPP-CONV-OFF*, *STOCHDP* and of the interpolated analysis (*ANA*) in the regions G, TROPICS_G, NH_ET and NATL. Averaged over 11 initial times.

from *IC-ONLY*, which shows that not only the trajectory counts within predefined domains remain unchanged through *STOCHDP*, but also their integrated diabatic heating.

Yet again, a separation into different heating regimes is required to provide a meaningful statement whether the trajectory characteristics are changed by the schemes. The trajectory counts and the latent heating rates are therefore shown for the two regimes with $\Delta\theta < 38\text{K}$ and $\Delta\theta > 38\text{K}$ (Figure 4.14). The qualitative picture regarding the effect of *SPP* is again comparable to the results obtained from the operationally used schemes: in the upper heating regime, *SPP* results in a strongly increased trajectory count compared to *SPPT* and is even higher than the experiment with *SPPT*. The trajectory count in the lower heating regime is, however, only slightly increased compared to the unperturbed model. The effect of *SPP* on the heating rates are, in contrast to the trajectory count, independent from the heating regime, resulting in heating rates comparable to the other schemes as well as to the unperturbed model in both categories. The same holds true for *SPP-CONV-ONLY* and *SPP-CONV-OFF*, which do not affect the heating rates of the trajectories within the two bins, but have a larger impact on the number of trajectories in the upper than in the lower heating category. Again, the latter effect is more pronounced for the experiment where only parameters in the convection scheme are perturbed, pointing towards the dominant role of those perturbations for the change of the trajectory frequencies. Consistent with the previous results, the trajectories computed from the experiment perturbed with *STOCHDP* have similar trajectory counts in the lower regime and reduced counts in the upper regime, and as for the other

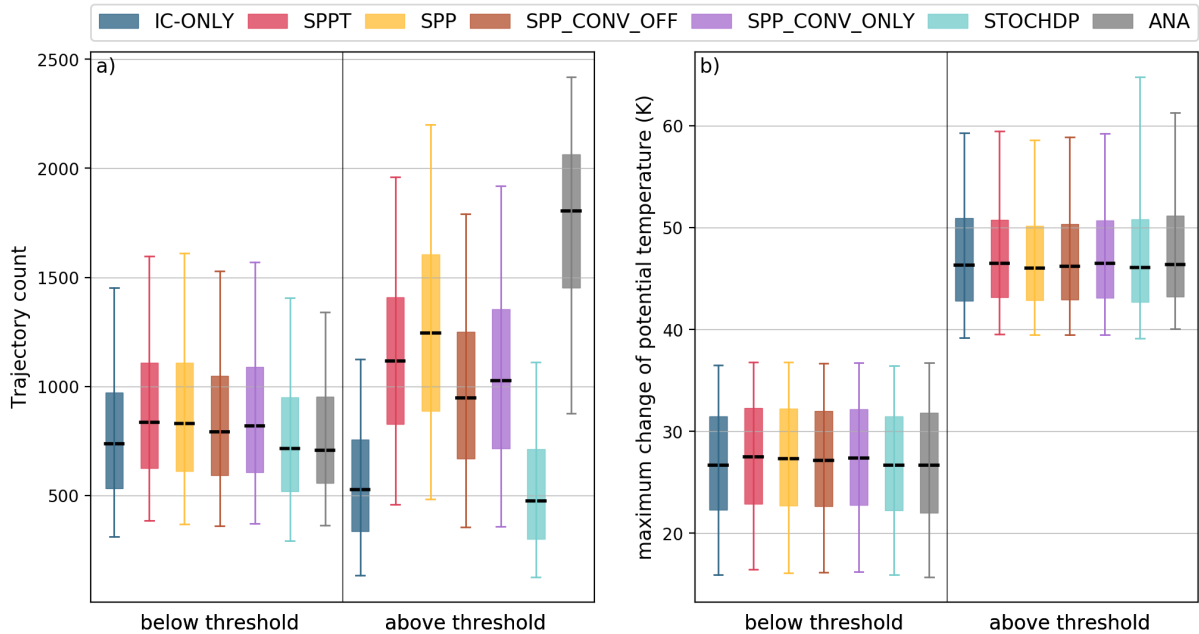


Figure 4.14.: a) Trajectory count and b) latent heating rate of all trajectories (global) in the experiments *IC-ONLY*, *SPPT*, *SPP*, *SPP-CONV-ONLY*, *SPP-CONV-OFF*, *STOCHDP*, and the interpolated analysis (*ANA*) separated for the two latent heating regimes below (left) and above (right) the threshold of 38 K. Averaged over 11 initial times.

schemes, the heating rates are in line with the unperturbed experiment *IC-ONLY* in both heating regimes.

The role of latent heating

To further highlight the importance of the latent heat release and to simultaneously discuss the differences between the model uncertainty schemes, the ratio of trajectory counts in the perturbed experiments and in *IC-ONLY* is computed (similar to Figure 4.10 of the operational data set) and shown in Figure 4.15. In contrast to the operational data set, which only contains trajectories in the North Atlantic, the global distribution of the heating rates in the experiments (grey histogram, based on *SPPT*) is characterized by two frequency maxima. The colored lines depict the ratio of the number of trajectories for each heating rate bin of the corresponding experiment to the unperturbed experiment (*IC-ONLY*). Similar to the perturbed forecasts in the operational data set, the ratio of *SPPT* lies below 1 for trajectories that are heated less than 17.5K. For heating rates higher than that, the ratio is exclusively above 1, increases slowly to a value of about 1.1 at the secondary frequency maximum at 22.5 - 25 K, and grows exponentially until heating rates of 50 K, where it exceeds the value of 2 (corresponding to a doubling of the trajectory number). The qualitative behaviour of *SPP* is very similar to *SPPT*, as the effect generally increases with the latent heating rate between approximately 15 K and 50 K. However, the ratio of *SPP* to the reference (*IC-ONLY*) is above 1 for all heating rates without exception, which is especially relevant for the weakly heated trajectories that are affected in the opposite way by *SPPT*. Further, the ratio of *SPP* is smaller than the one of *SPPT* in the range around the secondary frequency maximum between 22.5 to 35 K, which is of particular importance for WCBs in the extratropics. This behaviour is also reflected in the trajectory counts in the northern hemisphere Extratropics and the North Atlantic, where the experiment *SPP* has less trajectories than *SPPT*. For even higher heating rates ($\Delta\theta > 35\text{ K}$), the ratio of *SPP* overtakes the one of *SPPT* and results in substantially higher values in the range of the global frequency maximum at 42.5-45 K. The *SPP* sensitivity experiments behave according to the previously shown diagnostics, with the perturbed parameters in the convection parametrizations (*SPP-CONV-ONLY*) resulting in larger ratios than *SPP-CONV-OFF* across the whole range of heating rates. Finally, the previously reported unchanged or decreased trajectory counts in the experiment *STOCHDP* also appear in this diagnostic, where especially trajectories with very large heating rates are underrepresented compared to *IC-ONLY*.

This analysis shows again that the impact of the uncertainty schemes which introduce perturbations into the physical parametrizations (i.e. all schemes except *STOCHDP*) on the trajectory frequency counts is strongly modulated by the integrated diabatic heating rate along the trajectories. Despite some quantitative differences, the qualitative behaviour is very similar among those schemes. *STOCHDP*, in contrast, behaves very differently from the other schemes in this diagnostic.

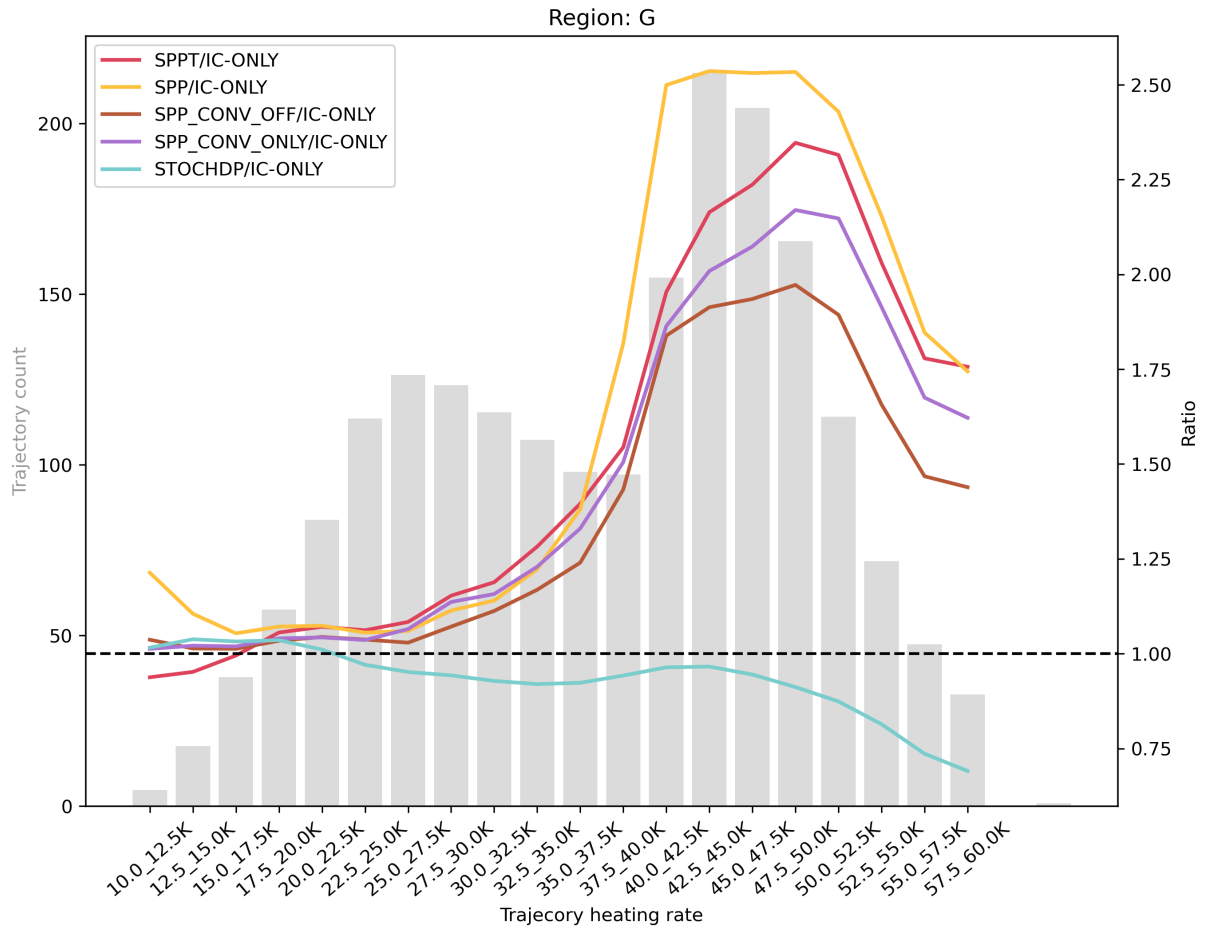


Figure 4.15.: Number of trajectories in the experiment *SPPT* started globally (histogram, left axis) per time step and ratio of the trajectory counts (right axis) in the experiments *SPPT* (red), *SPP* (yellow), *SPP-CONV-OFF* (brown), *SPP-CONV-ONLY* (purple) and *STOCHDP* (lightblue) and the unperturbed experiment *IC-ONLY* for the integrated latent heating rate along the trajectories. The bin width is 2.5 K, and the ratios are only plotted for those heating rates that occur at least once per time step.

4.2.2. Eulerian perspective

In this section, the Eulerian analysis of vertical velocities is repeated to further evaluate the effects of the perturbation techniques in the experiment set 2. Other than in the analysis of the operational schemes, only the difference of the grid point counts of ω -values between the corresponding experiments and the experiment with the unperturbed model is shown (Figure 4.16).

Qualitatively, the differences between the three experiments *SPP*, *SPP-CONV-ONLY* and *SPP-CONV-OFF* to *IC-ONLY* show a similar behaviour as the experiment *SPPT*: in the upward spectrum of the histogram, the lines are characterized by a double structure that corresponds to increased occurrences of very rapid ascents at the expense of moderate ascents, and an acceleration of air parcels without vertical motion to slow ascents. Also the downward motions behave equivalent to *SPPT*, with a uniform acceleration of descending motions. The detailed effect on the upward motions is very similar in *SPP* and

4. The impact of ensemble configuration on rapidly ascending air streams

SPP-CONV-ONLY, but is different in *SPP-CONV-OFF*, where the structure related to moderate ascents is shifted towards smaller velocities (local minimum at -0.25 Pa/s compared to -0.4 Pa/s in *SPP*). Further, also the increased occurrence of very slow upward velocities is more pronounced in *SPP-CONV-ONLY* than in *SPP-CONV-OFF*. This indicates that perturbations to parameters in the convection scheme have overall a larger impact on the distribution of vertical motions than parameter perturbations in the other parametrization schemes.

The good qualitative agreement between the experiments *SPPT* and *SPP* suggests that the underlying mechanism how the perturbations influence the vertical velocities is similar for the two schemes. However, despite these similarities, it is evident that the detailed bimodal structure in the experiment *SPP* is differs to the one of *SPPT*. For example, the difference plot of *SPP* is characterized a broader range of positive values and larger frequencies of slow upward velocities (the x-axis is intersected at -0.2 Pa/s in *SPPT* and at -0.25 Pa/s in *SPP*). Contrarily, the very rapid ascents are affected more strongly by *SPPT* than through *SPP* (see the higher frequency differences for upward velocities faster than -0.5 Pa/s of *SPPT* than of *SPP*).

The effect of perturbations to the dynamical core (*STOCHDP*) on the vertical velocities differs substantially from the one described for the schemes acting on the model physics: *STOCHDP* only results in an acceleration of air parcels that have no or only very slow vertical motions both up- and downwards, and does not feature the double structure on the positive side of the ω -spectrum. This indicates that *STOCHDP* does not systematically affect the very rapid ascents, which is in agreement with

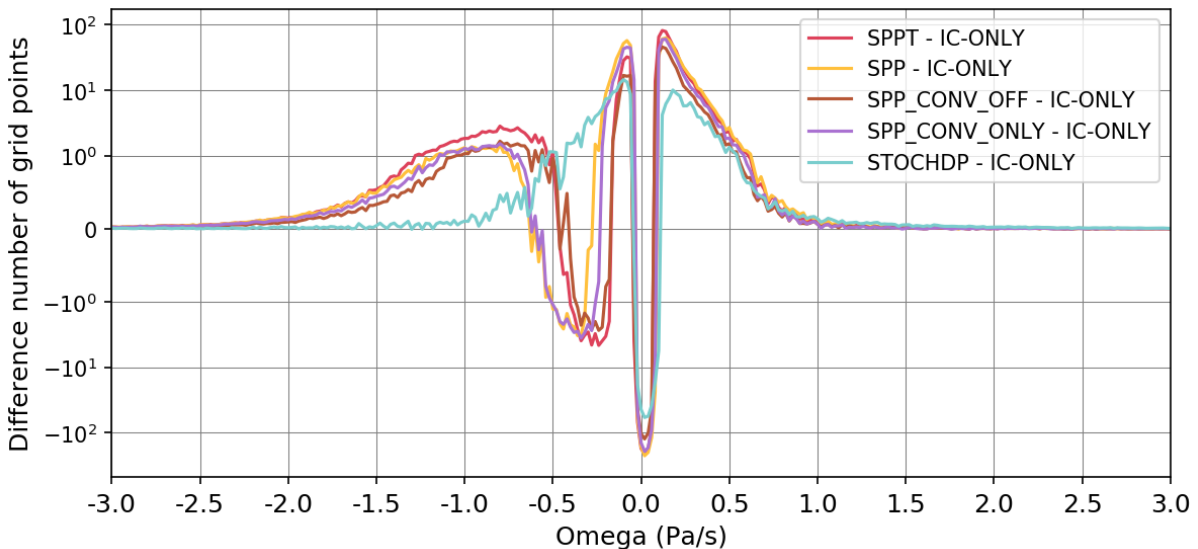


Figure 4.16.: Difference number of grid points of vertical velocities at 500 hPa between the experiments *SPPT*, *SPP* (yellow), *SPP-CONV-OFF* (brown), *SPP-CONV-ONLY* (purple), and *STOCHDP* (lightblue), and the reference experiment *IC-ONLY*. Averaged over 11 initial times.

the Lagrangian analysis, where the number of rapidly ascending trajectories was not increased through STOCHDP. Nevertheless, also STOCHDP causes a unidirectional response of the model, as it results in less grid points without or with only very slow vertical motions.

The large discrepancy between *STOCHDP* and the other experiments in the range between about -0.25 and -0.5 Pa/s, where the difference number of grid points is positive for *STOCHDP* and negative for all other experiments, can be explained as follows: All experiments have a similar effect on the very slow vertical motions and initiate both slow up-and downward motions. For *STOCHDP*, this increases also the number of grid points in the aforementioned range of values, which does not occur in the other experiments because of the acceleration to even larger velocities at the expense of grid points in this velocity range. This superposition of the effects leads to a net decrease of grid points with moderate vertical velocities.

The mechanism behind the unilateral effect of stochastic perturbations on vertical velocities is explored in more detail in the following section.

4.3. Mechanism

The previous sections gave a detailed description of how perturbations of the model physics result in an increased number of rapidly ascending air streams, while initial condition perturbations only exert such an impact at the very beginning of the forecast, and perturbations to the dynamical core do not trigger this behaviour at all, but only accelerate very slow vertical motions. Despite the symmetric and zero-mean design of the perturbation schemes analyzed in the framework of this thesis, all of them result in a unidirectional response of the model, even though these responses differ between the individual schemes. This section discusses a potential mechanism how this behaviour can be explained. Before doing so, the key findings from the chapter which are important for the reasoning are summarized:

- All model uncertainty schemes except STOCHDP result in an increased frequency of rapidly ascending air streams without substantially changing of the trajectory characteristics.
- The effect of the model physics perturbations is apparent from the beginning of the forecast and is constant with lead time.
- The distributions of the vertical velocities are affected unilaterally by all perturbation techniques: the stochastic physics perturbations result in a bimodal modulation of the ω -distribution, and also STOCHDP changes the vertical velocities, even though only the slowest ones. Finally, also the initial condition perturbations affect the structure of the ω -distribution, but only within the first few time steps of the model integration.

Building on these findings, it can be excluded that the impact of the perturbations on the rapidly ascending air streams is a result from pre-existing or developing differences between the perturbed and unperturbed simulations, such as differing moisture distributions. This would impact both the trajectory characteristics as well as the temporal evolution of the trajectory count/characteristics. The effect is rather of an instantaneous nature, which indicates that the perturbations must have a direct impact on the thermodynamic state of the air parcel.

One pathway how symmetric, zero-mean perturbations can lead to a unilateral response can be understood with help of Figure 4.17. It shows a hypothetical and simplified probability distribution of a quantity (x) in a non-linear system, which is characterized by a critical value (x_{crit}), below which the system is in a stable state, while it resides in an unstable state above it. The phase space below the threshold x_{crit} (blue area) is therefore more densely populated than the phase space directly above the threshold (red area). When residing in the unstable phase space, a process will be triggered to bring the system back into equilibrium. This can be brought into a meteorological context with the following example: the quantity x could be temperature, and x_{crit} could be the temperature above which the air parcel becomes lighter than its surrounding. This triggers ascending motion, leading to a cooling of the air parcel until it

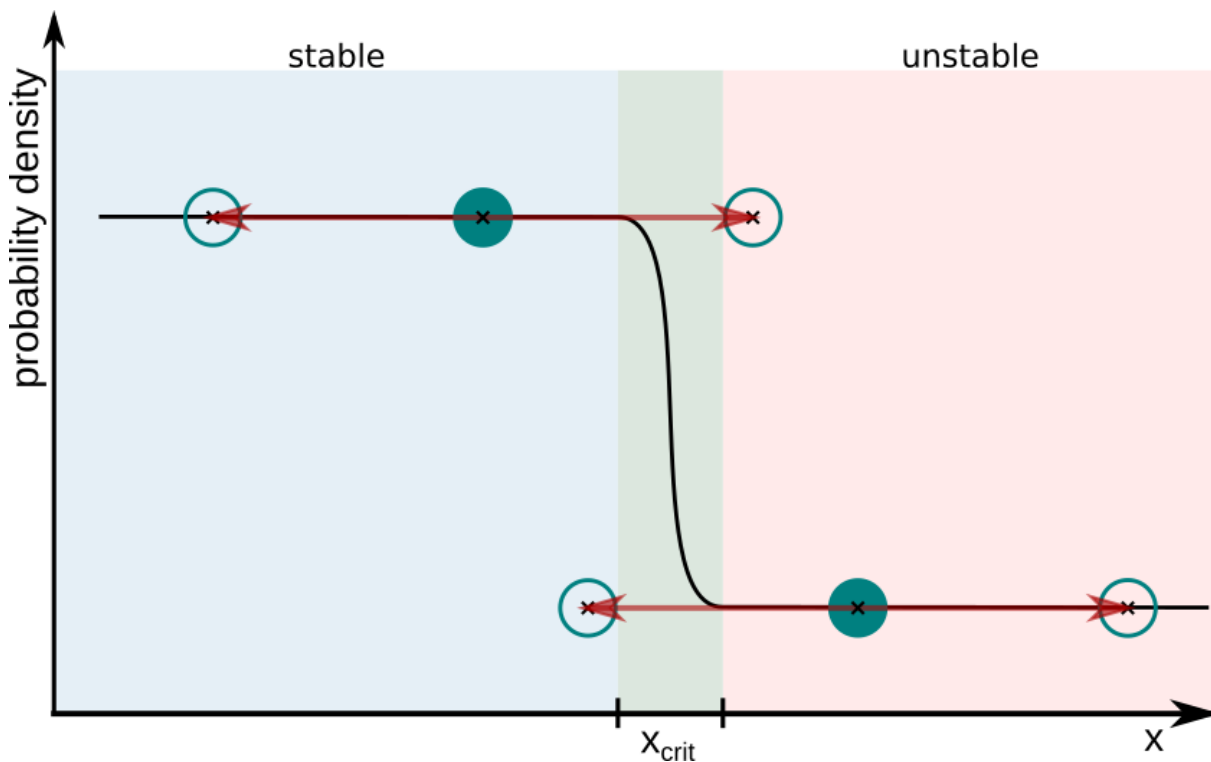


Figure 4.17.: Schematic illustrating the asymmetric effect of zero-mean perturbations on variables in a nonlinear system. The black line is the probability density, the black crosses surrounded by the filled circles illustrate values which are perturbed symmetrically with an amplitude given by the red arrows. The black crosses with the unfilled circles show the values after the perturbation has been applied. The green area separates the stable (blue) from the unstable (red) phase space. Adapted from Pickl et al. (2022).

is in a thermodynamic equilibrium with its environment.

When such a system is perturbed with symmetric, finite-amplitude perturbations, there will be more instances of values below the threshold being pushed above the critical value by positive perturbations than values above the threshold being pushed below the critical value, as the distribution around the threshold is not uniform. Prerequisite is that the perturbations are large enough, that they have the same magnitude on both sides of the distribution, and that two subsequent or neighboring perturbations that are introduced shortly after each other acting on the same air parcel do not have the opposite sign (i.e. the air parcel is directly brought back to its original position in the phase space)¹. If these criteria are fulfilled, perturbations in such a nonlinear system will result in a unidirectional effect, despite their symmetric zero-mean design.

In order to test this hypothesis, ascending trajectories from a forecast with a perturbed model and from an unperturbed simulation that start from identical initial conditions are compared. Additionally to the trajectories that have been selected by the strict ascent criterion of 600 hPa within 48 hours, also such trajectories that ascent by only 300 hPa in 2 days have been retained². As all experiments start from identical initial conditions (except for *SPPT-ONLY*), it is possible to compare those trajectories that started at lead time 0 h and were detected as "rapidly ascending" (i.e. 600 hPa ascent) in the one experiment, but did not fulfill this criterion in the other experiment. This has been done for one member of one individual forecast in the experiments *SPPT* and *IC-ONLY*, where 2971 trajectories ascended by at least 600 hPa in 2 days in *SPPT*, while only 1628 trajectories fulfilled the criterion in *IC-ONLY*. From the remaining 1343 trajectories, 1092 trajectories ascended by at least 300 hPa. Figure 4.18 shows selected characteristics of those trajectories for the two experiments. Starting from an identical initial state, the trajectories diverge on average at a lead time of 6 hours. Per definition, the trajectories from *SPPT* ascent to higher pressure levels than the ones from *IC-ONLY*. This is also reflected in a higher potential temperature (b) and a stronger decrease of specific humidity along the trajectories (c). The evolution of potential vorticity shows a diabatically generated dipole of PV with high values in the mid-troposphere and low values in the upper troposphere, which is not present for the weakly ascending trajectories in *IC-ONLY* (d).

From this analysis, the location and the time when the individual trajectories begin to diverge between the experiments can be identified. When considering only the first possible time step when the trajectories can be different (i.e. 6 hours lead time), a dominant imprint of the perturbations introduced by *SPPT* at that time can be expected. We define the diverging point of the trajectories as that time when the pressure difference between the trajectories of the two experiments is larger than 10 hPa. This procedure is done

¹All perturbation techniques discussed here introduce perturbations that are spatially and temporally correlated. This prevents such a behaviour

²Ideally this analysis would be done with all trajectories independently from their ascent. To save disk space, however, only those trajectories that ascend by at least 300 and 600 hPa have been retained

4. The impact of ensemble configuration on rapidly ascending air streams

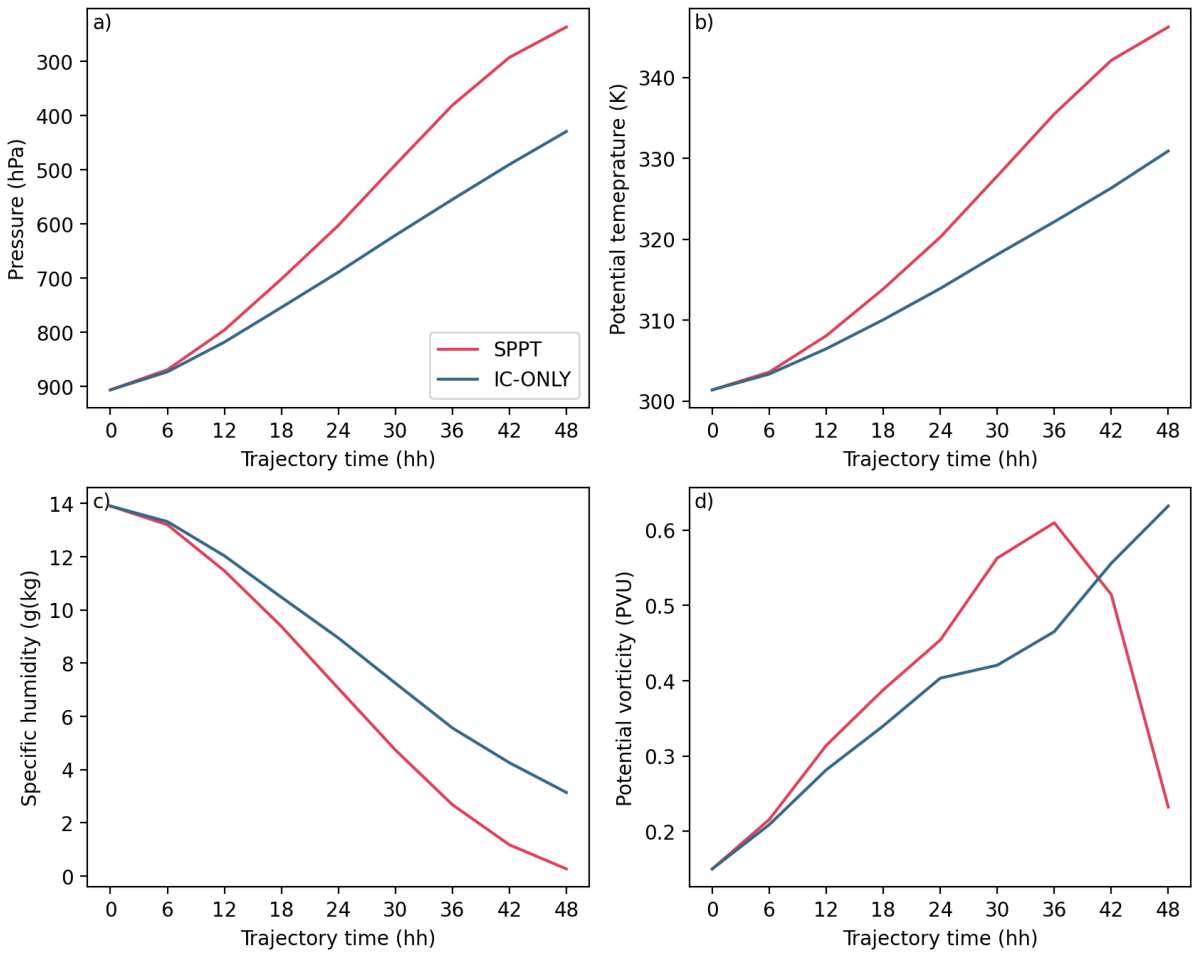


Figure 4.18.: Mean evolution of pressure (a), potential temperature (b), specific humidity (c) and potential vorticity (d) of trajectories in the experiments *SPPT* (red) and *IC-ONLY* (blue) started globally at lead time 0h of the ensemble member 01 of the forecast initialized on September 24 2016. Shown are those trajectories that fulfill the ascent criterion of 600hPa in 48 hours in *SPPT*, but fail to do so in *IC-ONLY*, where the ascent is less than 600 hPa, but more than 300 hPa in 48 hours (1092 trajectories). The diagnostic is shown only for one forecast, but the signal is representative for many forecasts.

for all these trajectory pairs in *SPPT* and *IC-ONLY* in the 32 ensemble forecast with each 20 members, resulting in overall around 85000 of such cases. For each of these cases, the difference of meteorological fields between the two experiments *SPPT* and *IC-ONLY* is computed and centered spatially and temporally on the corresponding trajectory point. The composite-mean of temperature at 850 hPa for these cases is shown in Figure 4.19a. Centered on the position where the trajectories diverge, the forecasts that are perturbed with *SPPT* are on average around 0.2 K warmer than the unperturbed forecasts. The warm anomaly has an elliptic geometry, a meridional extent of about 10° and a zonal extent of about 15° , and its magnitude decreases with increasing distance from the composite center. The horizontal extent is of the order of magnitude of the length scales of the random perturbation patterns of *SPPT* (between 500 and 2000 km, see table 3.1); together with the early stage in the forecast (lead time of 6 hours), it is likely that this signal is directly shaped by the perturbations introduced by *SPPT*, and not due to a

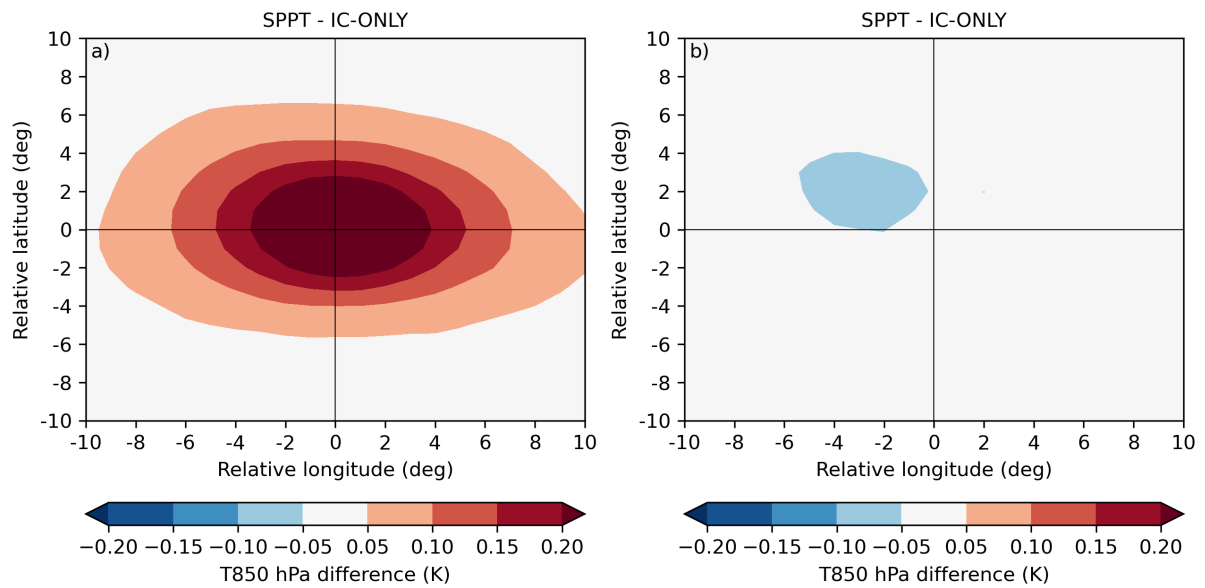


Figure 4.19.: Difference of composite-mean temperature at 850 hPa between the experiments *SPPT* and *IC-ONLY* centered on the diverging points of trajectories that (a) ascend by at least 600 hPa in 48 hours in *SPPT* and between 300 and 600 hPa in 48 hours in *IC-ONLY* (84857 cases), and (b) ascend by at least 600 hPa in 48 hours in *IC-ONLY* and between 300 and 600 hPa in 48 hours in *SPPT* (10302 cases).

chaotic, nonlinear drift of the perturbed and unperturbed forecasts.

The same analysis can be conducted the other way round: the composite is again computed centered on those points where trajectories from the experiment diverge; however, we now look for trajectories in *IC-ONLY* that ascend by at least 600 hPa in 2 days and do not in *SPPT*. We thereby obtain those trajectories that would ascend without the perturbations, but are hindered to do so due to *SPPT*. This occurs less frequently than the other way round, as *SPPT* tends to increase the occurrence of ascending trajectories, but still results in about 10000 cases. The composite-mean of temperature at 850 hPa centered on these diverging points is shown in Figure 4.19b, and the signal is of opposite sign than its counterpart and of smaller magnitude. Hence, negative temperature perturbations tend to suppress ascending motions when they are applied at critical stages of the trajectory inflow. This analysis supports the aforementioned hypothesis that positive (temperature) perturbations are more effective in modifying the thermodynamic state of an air parcel such that it can fulfill a strict ascent criterion, than negative perturbations in preventing this process.

Most likely, the idealized mechanism explained in Figure 4.17 explains the one-sided response of the rapidly ascending air streams to the model physics perturbations. The fact that the latent heating rate controls the strength of the response goes along with the described process chain: large perturbations, which occur more likely in regions where latent heating takes place and is large, are more effective in pushing data points across the critical value than small perturbations.

Our analysis suggests that this behaviour is to some part intrinsic to stochastic perturbations, as it appears in all experiments that are perturbed stochastically: while the effects behave similarly in the experiments with model physics perturbations (i.e. *SPPT*, *SPP*, *SPP-CONV-ONLY*, *SPP-CONV-OFF*), there are substantial differences to the experiment with only initial condition perturbations (*IC-ONLY*), where such an effect can only be detected in the first model time steps, and especially to the experiment that is perturbed with *STOCHDP*. This shows that, despite the general applicability of the concept to stochastic perturbations in a nonlinear system, the detailed configuration of the perturbations matters for the effect on ascending motions. The investigated model physics perturbations are designed to introduce large perturbations into regions where parametrizations produce large physics tendencies - which are often co-located with rapid ascents. Therefore, the comparably large perturbations are likely to trigger the aforementioned process chain. In contrast, *STOCHDP* tends to introduce largest perturbations in regions where the large-scale flow configuration is complex. These regions do not correspond to areas of diabatically driven, rapid ascents, but rather to the position of the upper-level jet stream, which is characterized by pronounced wind shear. Hence, one reason why *STOCHDP* affects the slow scales of vertical motion and not the larger ones could be that the perturbations that are large enough to trigger the nonlinear behaviour are located only in regions of large-scale, dry-dynamic ascents which are forced quasi-geostrophically by the upper-level jet and not in regions where moist-diabatic processes are very active.

4.4. Summary and Discussion

This chapter gave an in-depth investigation on how different ensemble perturbation techniques affect the vertical motions of ECMWF's ensemble prediction system. In the first part, research experiments with perturbation schemes that are currently in operational use (*SPPT* and initial condition perturbations) have been analyzed in a Lagrangian and an Eulerian framework. It was found that *SPPT* increases the occurrence of rapidly ascending air streams, that have been detected with trajectory analysis, in a systematic way without altering the trajectory characteristics, such as the latent heating rate. Apart from the most rapid vertical velocities, the frequency distributions of other scales of vertical motion are also affected, leading to an overall acceleration of vertical motions both up- and downwards. In contrast, the initial condition perturbations do not exert such a unilateral behaviour throughout the forecast, even though they have a comparable effect as *SPPT* in the first few time steps of the forecast.

The results from these experiments further showed that the unperturbed control member shows a very similar behaviour in the diagnostics as the research experiment without *SPPT*. We made use of this insight and investigated the differences between perturbed and unperturbed ensemble members of archived operational forecasts of two years. With this data set, we were able to show that the results are very robust across a long time period, different model configurations (e.g. the model resolution) and different

setups of the trajectory calculation.

Furthermore, it was shown that the impact of SPPT on the trajectories strongly correlates with the integrated latent heat release along the ascending air parcels. This is reflected in regional differences, as the frequency increase of the trajectories is larger in the tropics than in the extratropics, and in seasonal variations, as the effect is larger in autumn and spring than in winter.

In the second part of the chapter, a set of other model uncertainty schemes was analyzed in order to advance the understanding how perturbations interact with the forecast model. The SPP-scheme, which, similarly to SPPT, accounts for uncertainties in the physics parametrizations, results in qualitatively very similar effects. STOCHDP, a scheme that introduces uncertainties into the dynamical core of the forecast model, does not affect the very rapid ascents, but still impacts vertical velocities of smaller magnitude in a unidirectional way.

Finally, we introduce a mechanism how stochastic, symmetric and zero-mean perturbations can result in such a biased response. Based on the analyses, we hypothesize that the perturbations are more effective in triggering fast ascents than suppressing them, as such motions are often characterized by a nonlinear behaviour with a nonuniform probability distribution around a threshold.

In the scientific community, biased responses to zero-mean forcing are not new: for example, Tompkins and Berner (2008) observe that positive humidity perturbations in experiments with a stochastic convection scheme are more likely to trigger convection than negative perturbations can suppress it, and justify their results with the highly nonlinear nature of atmospheric convection. Leutbecher et al. (2017) argue that nonlinear physical processes, such as the saturation of humidity that results in the formation of clouds and precipitation, can be the reason for asymmetric responses of the model to zero-mean perturbations. However, to the authors knowledge, no study has so far broken down a detailed description how symmetric perturbations result in a unilateral response on a process-level, like it is done in this thesis. Potential impacts of these findings for other processes in the model are discussed in the following chapter.

5. On the impacts of modulated vertical motions through stochastic perturbations

In the previous chapter, it was shown that stochastic perturbations in the forecast model have a systematic impact on rapidly ascending air streams and on vertical velocities in general. Up- and downward motions are a very important component of the atmospheric circulation and are linked to atmospheric phenomena on different spatio-temporal scales. Therefore, imprints of the previously observed modulations of vertical velocities should also be reflected in weather activity that is directly or indirectly linked to vertical motions. In this chapter, the impact of stochastic model uncertainty schemes is evaluated for two such phenomena: precipitation (Section 5.1) and the representation of the upper-level Rossby wave amplitude (Section 5.2).

5.1. Precipitation

Weather systems characterized by distinct upward motions, such as tropical convection, tropical cyclones or WCBs, contribute substantially to the global precipitation budget (e.g. Jiang and Zipser 2010; Pfahl et al. 2014). Due to the cooling during the ascent, water vapor condenses and cloud droplets form, which ultimately leads to precipitation. When the characteristics, the spatial extent or the frequency occurrence of such weather systems is changed through stochastic perturbations of the forecast model, this should also affect the precipitation in the forecasts. It was shown in Chapter 4 that the occurrence of rapidly ascending air streams is systematically increased with stochastic physics perturbations. Hence, it will be evaluated whether precipitation is changed through the representation of model uncertainty, and if this can be linked to the changes in the vertical motions.

5.1.1. Operational schemes

Figure 5.1a shows the average daily precipitation sums derived from the experiment *SPPT* over all forecasts, members and lead times during the study period from August 15th to October 27th. The largest precipitation sums occur in the tropical belt, mainly along the intertropical convergence zone (ITCZ), which is located slightly north of the equator due to the study period in summer/autumn, and over the Maritime Continent. Further, the southern and northern hemisphere storm tracks are characterized by increased precipitation sums, whereas the subtropics receive little precipitation. Comparing these global precipitation patterns to the frequencies of rapidly ascending air streams (contours in panel a) clearly

indicates the linkage of precipitation to these weather systems.

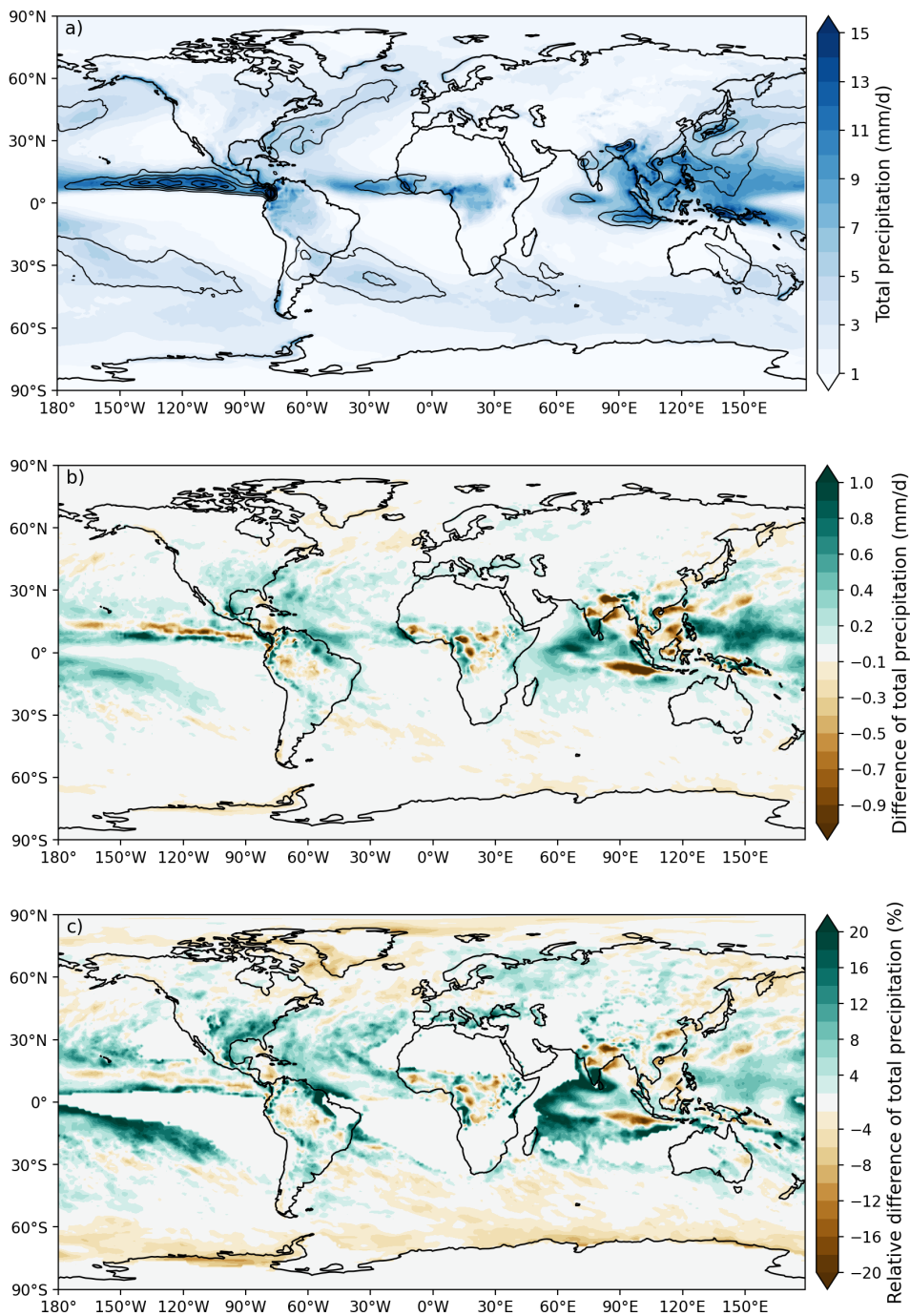


Figure 5.1.: (a) Daily total precipitation (large-scale plus convective precipitation, shading) and ascent frequencies of rapidly ascending air streams (contours from 4% to 20% in 4% steps) averaged over all forecasts (32), lead times (41) and ensemble members (20) in the experiment *SPPT*, (b) the absolute difference of total precipitation and (c) the relative difference of total precipitation between *SPPT* and *IC-ONLY*. In (c), only grid points with an absolute precipitation of at least 1 mm/day are shown.

The difference of daily precipitation sums between *SPPT* and *IC-ONLY* (Figure 5.1b and c) shows that *SPPT* has a substantial impact on precipitation, especially in the tropical regions. Over large parts of the tropical Indian and West Pacific Ocean, *SPPT* increases the precipitation frequencies (locally up to 1 mm/d and 20%, respectively), even though there are also regions with substantially decreased rainfall. In the tropical East Pacific, *SPPT* results in a southward shift of the ITCZ and leads to a patchy signal over tropical Africa. These patterns are, however, rather noisy, which likely arises from the short study period and biased different representations of single events. Apart from the tropical regions, also differences in the Extratropics, especially in the northern hemisphere, are apparent. For example, *SPPT* increases the daily precipitation sums in the North Atlantic and over south-eastern North America locally by up to 0.4 mm/d or 10% until latitudes of about 45° N. Over the North Pacific, a similar, yet less pronounced pattern is observed. The overarching pattern is that precipitation sums are generally increased by *SPPT*, especially in the (sub-)tropical regions. Nevertheless, the patchy signals in Figure 5.1b indicate that the study period is too short for sophisticated analyses of the impact of stochastic physics perturbations on the details of the climatological precipitation patterns.

Visually, these main differences in precipitation, apart from the smaller-scale changes along the ITCZ or over the Maritime Continent, compare well to the changes of rapidly ascending air streams induced by *SPPT* (see Figure 4.2e for outflow and Figure B.1b in Appendix B for ascent), which are strongest in the tropics, persist into the storm tracks and decrease/vanish for high latitudes. To investigate the linkage between the altered vertical motions and precipitation, the precipitation frequency occurrences are analyzed analogous to the vertical velocity distributions in the previous chapter. Figure 5.2a shows the differences of the number of grid points for different hourly precipitation rates between *SPPT* and *IC-ONLY*. Very small (< 0.5 mm/h) and large precipitation rates (> 1.8 mm/h) occur more often with than without *SPPT*. In contrast, precipitation rates between 0.5 and 1.8 mm/h are reduced by *SPPT*. The relative change of the number of grid points is particularly apparent for high precipitation rates above around 4 mm/h, with an increase of frequency well above 20%. These frequency changes are consistent with the impact of *SPPT* on vertical velocities: differences between *SPPT* and *IC-ONLY* of precipitation rate frequencies strongly resemble the signal corresponding to the upward side of the omega frequency changes (c.f. Figure 4.7b). Rapid ascents which produce large amounts of precipitation occur more often with than without *SPPT*, while intermediate ascents with smaller precipitation rates are reduced.

To further elaborate on the causality between the increased number of rapidly ascending air streams and the changed precipitation frequencies, precipitation sums are computed where the ascent (i.e. masks of gridded trajectories between 800 and 400 hPa) of rapidly ascending air streams takes place; grid points which are not associated with ascent masks are omitted. Note that these values, shown in Figure 5.2b, are normalized by the number of grid points in each region; this allows to directly compare the values between the regions, but the values cannot be compared to the grid-point based values in panel a. The

precipitation sums associated with the ascent masks are higher with SPPT than without in all regions. In the tropics, precipitation within the ascent masks is more than double in *SPPT* compared to *IC-ONLY*, and the increase amounts to about 10% in the Northern Hemisphere Extratropics and in the North Atlantic. In contrast, the local rain rate within the masks is actually slightly reduced compared to the experiment with unperturbed model physics. This may be counter-intuitive at the first sight, but makes sense when recalling how SPPT changes the upward motions: SPPT increases the frequency of trajectories which fulfill the criterion of 600 hPa ascent within 2 days without changing their physical properties,

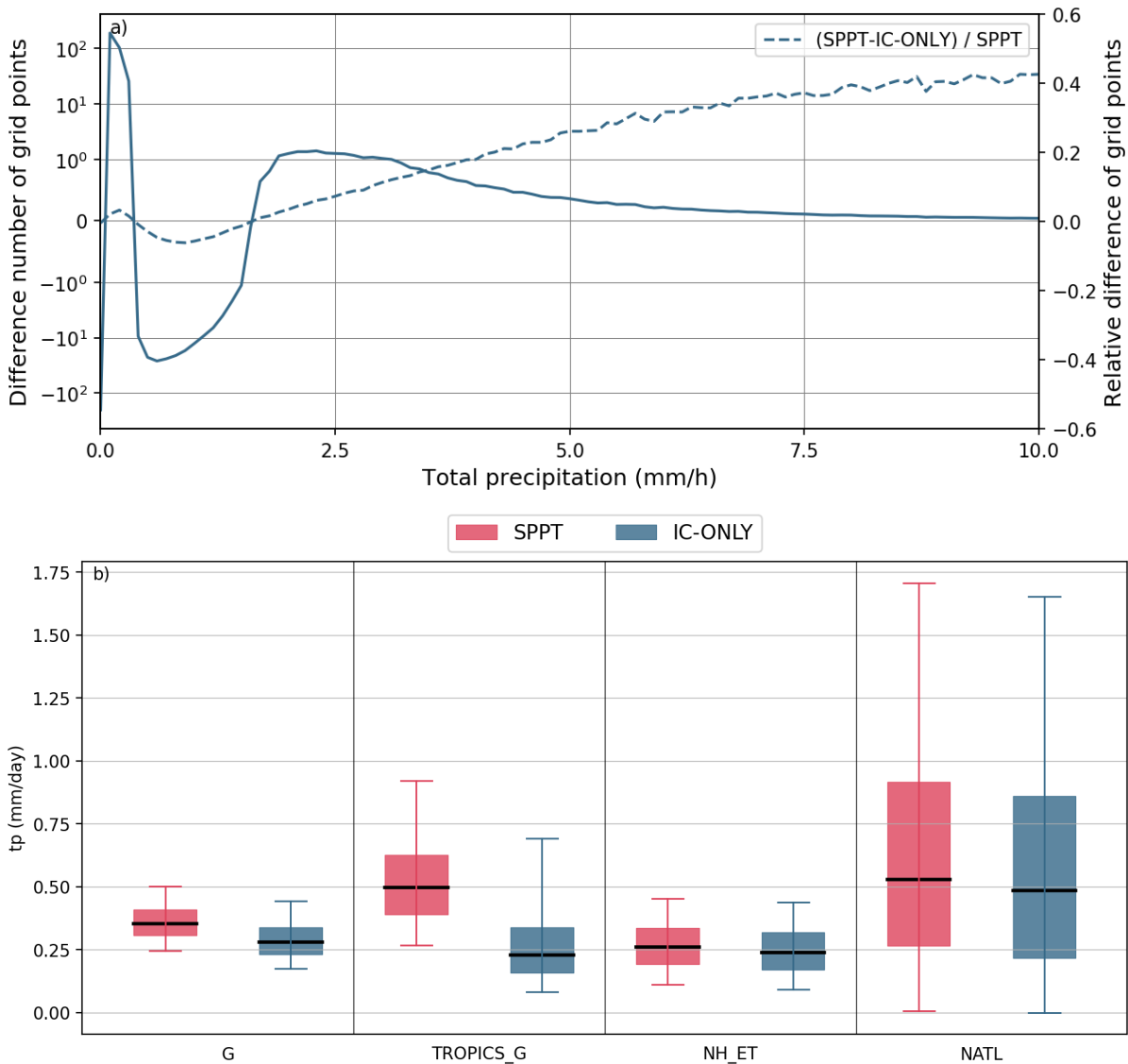


Figure 5.2.: Absolute (solid, left axis) and relative (dashed, right axis) differences between the frequency counts of precipitation rates (bin width of 0.1 mm/h) of *SPPT* and *IC-ONLY* per forecast, lead time, and ensemble member (a). The left y-axis has a linear scale between the values -1 and 1 , and a log scale for values between -1 and 1 . Precipitation sums within masks of rapidly ascending air streams in the regions global (G), global tropics (TROPICS_G), Northern Hemisphere Extratropics (NH_ET) and the North Atlantic (NATL), scaled by the number of grid points within the corresponding region (b). Panel (a) is adapted from Pickl et al. (2022).

such as latent heat release. Hence, air streams that have already reached the criterion remain unaffected by SPPT, but others that previously did not fulfill the criterion now do. This will result in larger (i.e. more grid points) masks of ascent, in which precipitation occurs, and hence the precipitation sum over all grid points corresponding to the mask will be increased. The local precipitation rate at individual grid points associated with ascent is, however, not increased (but even decreased), as the moisture supply is not changed through SPPT. Hence, same amount of moisture is distributed among an increased number of grid points associated with ascending motion, resulting in a reduced local rain rate. Consistently with the impact of SPPT on the trajectories, the increased precipitation sums associated with rapidly ascending air streams does not result from an intensification of local precipitation, but from an enlargement of the ascent regions.

5.1.2. Other schemes

In this section, the impact of the model uncertainty schemes in the experiment set 2 on the global precipitation distribution is analyzed. In Figure 5.3, the frequency differences of precipitation rates in these experiments to *IC-ONLY* are displayed (*SPPT* is again plotted for reference). The experiment with model perturbations through *SPP* results in a similar bimodal response of precipitation as the experiment with *SPPT*: grid points with precipitation rates smaller than 0.5 mm/h occur more often with than without *SPP*, and the same is observed for large precipitation rates (≥ 2 mm/h). This increase comes at the expense of decreased precipitation frequencies of moderate intensity (i.e. between 0.5 and 2 mm/h). Compared to *SPPT*, *SPP* has a larger impact on the intense precipitation rates, as indicated by the more pronounced tail of the distribution difference with respect to *IC-ONLY*. This is balanced by a larger decrease of moderate precipitation events in *SPP*. The shape of the differences of the frequency distributions again reflects the impact of the scheme on the upward motions (compare to Figure 4.16). The very large precipitation rates (i.e. ≥ 2 mm/h) mainly occur in tropical regions; hence this shows again that *SPP* has a larger impact in the tropics than *SPPT*.

Both *SPP* sensitivity experiments (*SPP-CONV-ONLY* and *SPP-CONV-OFF*) have a qualitatively similar impact on the mean precipitation distributions as *SPP*, but analogous to the previous diagnostics, perturbations to the convection scheme alone are more effective in shifting the distribution to higher values than perturbations to all other parametrizations. The summed differences of these two experiments exceed the magnitude of the differences between *SPP* and *IC-ONLY*, which indicates that the perturbations to the parameters in the different schemes partly affect the same grid points and are superimposed when applied simultaneously in *SPP*. Finally, the impact of *STOCHDP* on the precipitation frequencies is different from the experiments with perturbed model physics: the frequencies of weak to intermediate precipitation rates are slightly increased, whereas grid points with no or very low precipitation rates are reduced.

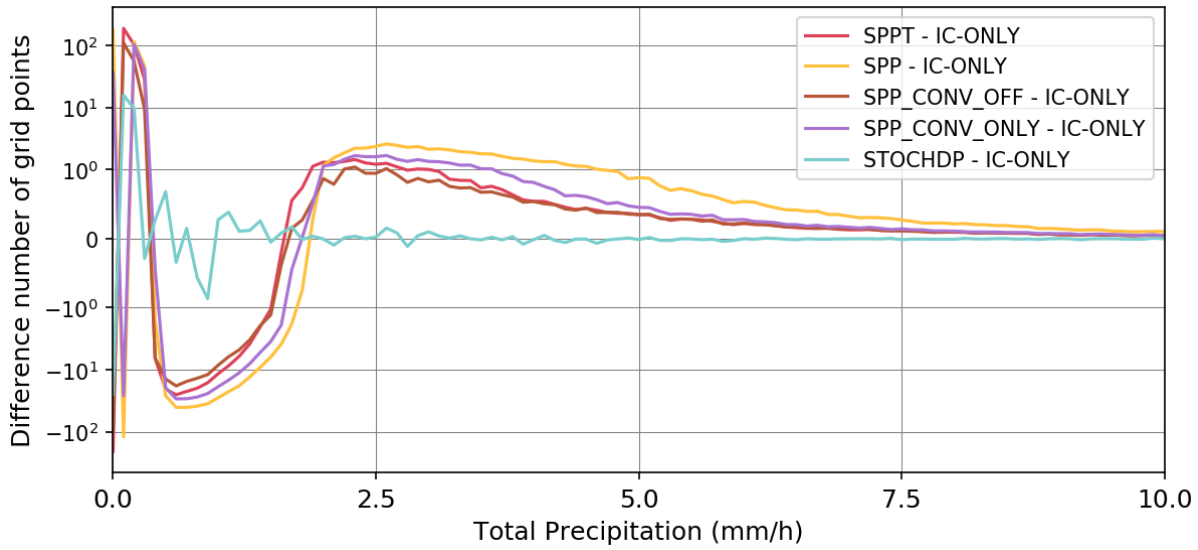


Figure 5.3.: Absolute differences of the frequency counts of precipitation rates (bin width of 0.1 mm/h) of *SPPT* (red), *SPP* (yellow), *SPP-CONV-ONLY* (purple), *SPP-CONV-OFF* (brown) and *STOCHDP* (lightblue) with *IC-ONLY* per forecast, lead time, and ensemble member (analogous to Figure 5.2). The left y-axis has a linear scale between the values -1 to 1 , and a log scale for values between -1 and 1 . Averaged over 11 initial times.

5.1.3. Discussion

This analysis demonstrates that the impact that stochastic model uncertainty schemes exert on the distribution of vertical velocities are directly reflected in the distribution changes of precipitation. The changes of precipitation with *SPPT* found in this study are in line with the results from Subramanian et al. (2017), who found qualitatively similar signals in seasonal IFS simulations: they report a similar, bimodal pattern of precipitation frequency changes with *SPPT*, where the frequency of small and large precipitation rates is increased, and the frequencies of intermediate precipitation rates is reduced. However, there are substantial quantitative differences between the studies, that most likely arise from several aspects, such as a different model version of the IFS (including different formulations of the *SPPT*-scheme) and a different investigation period. Regarding the analysis of the statistics of precipitation, the time period from mid-August until end of October 2016 in our study is comparably short, leading to a strong imprint of individual weather systems (see Figure 5.1b).

In their work, Subramanian et al. (2017) explain the changes of precipitation frequencies through *SPPT* in a purely stochastic way and argue that Gaussian perturbations to a Gamma-distributed random variable (which is a good approximation for precipitation) result in changes comparable to the observed signal. The comparison of different schemes in this thesis, however, shows that the modulation of the precipitation frequencies is directly tied to the changes in the ω -distributions. The bimodality in the differences between the ω -distributions reported for the experiments with perturbations in the physics parametrizations (*SPPT* and *SPP*) are reflected by a bimodal modification of precipitation frequencies,

whereas the unimodal change of upward velocities in *STOCHDP* is reflected in an equivalent signal in the precipitation differences. Hence, this modulation is not a pure mathematical result of stochastic perturbations to specific distribution, but is directly tied to the interaction of the perturbations with the underlying dynamical processes.

5.2. Rossby wave amplitude

In the extratropics, rapidly ascending air streams in the form of WCBs play an important role for the large-scale circulation. Their diabatically driven, cross-isentropic transport of low-PV air from the lower into the upper troposphere contributes to the amplification of the upper-level Rossby wave pattern, as the divergent outflow close to the tropopause strengthens the upper-level PV gradient and thereby deflects the dynamical tropopause pole- and upward (see Chapter 2 for more details on the role of WCBs for the large-scale circulation). As the frequency of WCBs is systematically increased through stochastic model physics perturbations, the question arises if this signal is propagated upscale and whether the large-scale circulation is affected by model uncertainty schemes. This question is tackled by comparing the areas of objectively detected upper-level ridges and troughs in the experiments with different model uncertainty representations. After the methodology is introduced in Section 5.2.1, the impact of the SPPT-scheme on the Rossby wave amplitude is investigated by analyzing research experiments and data from a forecast archive (Section 5.2.2). The alternative model uncertainty schemes from set 2 are analyzed subsequently (Section 5.2.4), followed by a detailed discussion of the results (Section 5.2.5) and an outlook (Section 5.3).

5.2.1. Methodology

To assess the impact of SPPT on the large-scale upper-level flow, and more specifically on the amplitude of the upper-level Rossby wave pattern, we employ a slightly modified¹ technique from Gray et al. (2014) to classify each grid point on an isentropic surface into one of the following categories: Polar vortex, subtropics, trough or ridge (see Figure 5.4). The classification is based on the structure of the dynamical tropopause, which we define as the 2 PVU contour on an isentropic surface (PV_{tp}). The dynamical tropopause is used to derive the equivalent latitude (Φ_{eq}), which was originally introduced by Butchart and Remsberg (1986) and is defined as the perimeter of a circle centered on the pole that encloses the same area as the instantaneous 2 PVU contour on an isentropic surface. This zonally symmetric background state is also referred to as the Modified Lagrangian Mean and encloses the same mass and circulation as the full instantaneous field (Methven and Berrisford, 2015). To determine Φ_{eq} , the sum of the areas of every grid point exceeding PV_{tp} is computed individually for each isentropic surface and valid time; the equivalent latitude is then obtained from the ratio of this area ($A_{PV_{tp}}$) to the area of the whole hemisphere (A_{hem}) by

¹Details on the modified technique follow in subsection 5.2.2.

$$\Phi_{\text{eq}} = \arcsin\left(1 - \frac{A_{\text{PV}_{\text{tp}}}}{A_{\text{hem}}}\right). \quad (5.1)$$

For the classification procedure, the PV-value of each grid point is at first compared to PV_{tp} ; subsequently, the latitude of the grid point (Φ) is compared to Φ_{eq} . A grid point on the northern hemisphere is then classified as

- Polar vortex, when $PV > PV_{\text{tp}}$ and $\Phi > \Phi_{\text{eq}}$
- Subtropics, when $PV < PV_{\text{tp}}$ and $\Phi < \Phi_{\text{eq}}$
- Trough, when $PV > PV_{\text{tp}}$ and $\Phi < \Phi_{\text{eq}}$
- Ridge, when $PV < PV_{\text{tp}}$ and $\Phi > \Phi_{\text{eq}}$.

An example of the identification of the four categories on the northern hemisphere is shown in Figure 5.4. On the 320 K isentrope, Φ_{eq} (blue contour) of the 2 PVU contour (black contour) has the value 55° N. Grid points that are located north of 55° N with PV-values below 2 PVU are classified as ridge (green areas), and grid points located south of Φ_{eq} and with values above 2 PVU are classified as trough. Note that we do not additionally consider cut-off lows as done in Gray et al. (2014), which are enclosed areas with $PV > PV_{\text{tp}}$ that are not directly connected to the pole (e.g. off the west coast of Ireland). Cut-off lows are classified into the category "trough".

To assess the impact of model uncertainty schemes on the structure of the large-scale circulation and the amplitude of the upper-level jet undulations, the areas of the different categories will be compared between the experiments introduced in the data and methods chapter (Chapter 3). Other than Gray et al. (2014), we compute Φ_{eq} in every ensemble member of each experiment rather than using an independent value for Φ_{eq} . In the following, the term "(Rossby wave) amplitude" refers to the sum of the hemispheric ridge and trough areas, and not to the meridional displacement of ridges or troughs (following Gray et al. 2014). Similarly to Chapter 4, at first the experiments with the operational schemes will be analyzed in depth. As the initial condition perturbations do not systematically affect the distribution of the vertical velocities and of WCBs, only the experiments *SPPT* and *IC-ONLY* are considered. Then, the perturbed and unperturbed ensemble members of a reforeccast data set will be compared, and finally also the model uncertainty schemes in the experiment set 2 will be considered.

5.2.2. Operational schemes

Equivalent latitude Φ_{eq}

The classification of each grid point into one of the four categories depends on the data set which is used for the identification of the equivalent latitude Φ_{eq} . Therefore, its behaviour and differences between the

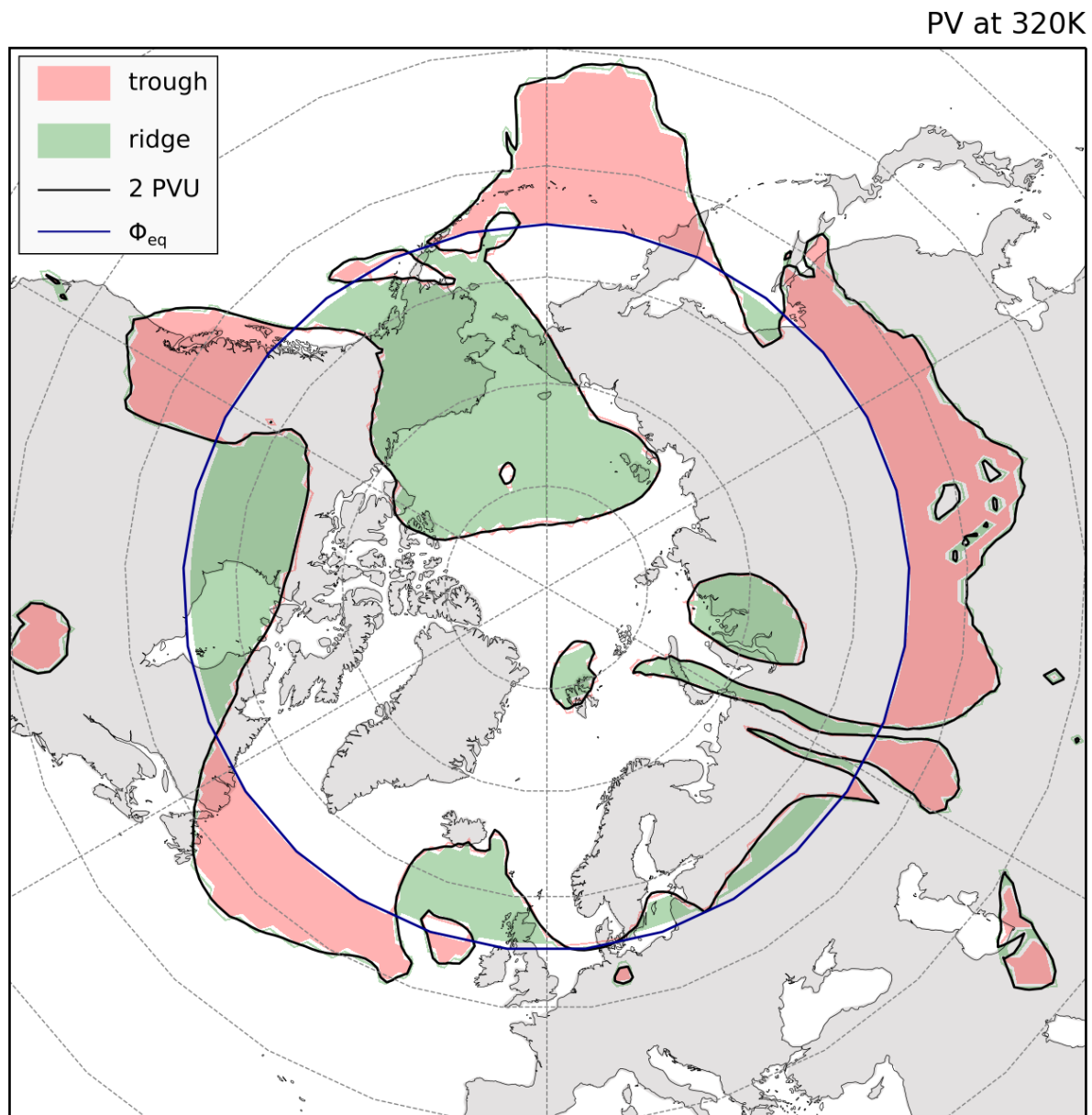


Figure 5.4.: Identification of troughs (red) and ridges (green) on the 320 K isentropic based on the 2 PVU contour (black contour) in the experiment *SPPT* initialized on September 26 00UTC at a forecast lead time of 7 days. The equivalent latitude Φ_{eq} is shown as blue line. The polar vortex and the subtropics are not indicated in the Figure, but can be identified as uncoloured areas poleward (polar vortex) and equatorward (subtropics) of the equivalent latitude.

simulations will be analyzed before evaluating the areas of upper-level ridges and troughs.

The average evolution of Φ_{eq} in the reanalysis and in the different experiments on three isentropic levels is shown in Figure 5.5. All data sets and isentropes have in common that the corresponding Φ_{eq} decreases with forecast lead time: For example, Φ_{eq} on 320 K starts above 61° N and ends up with values below 59° N, which corresponds to a southward shift of the hemisphere-mean 2 PVU line by about 3° within 12

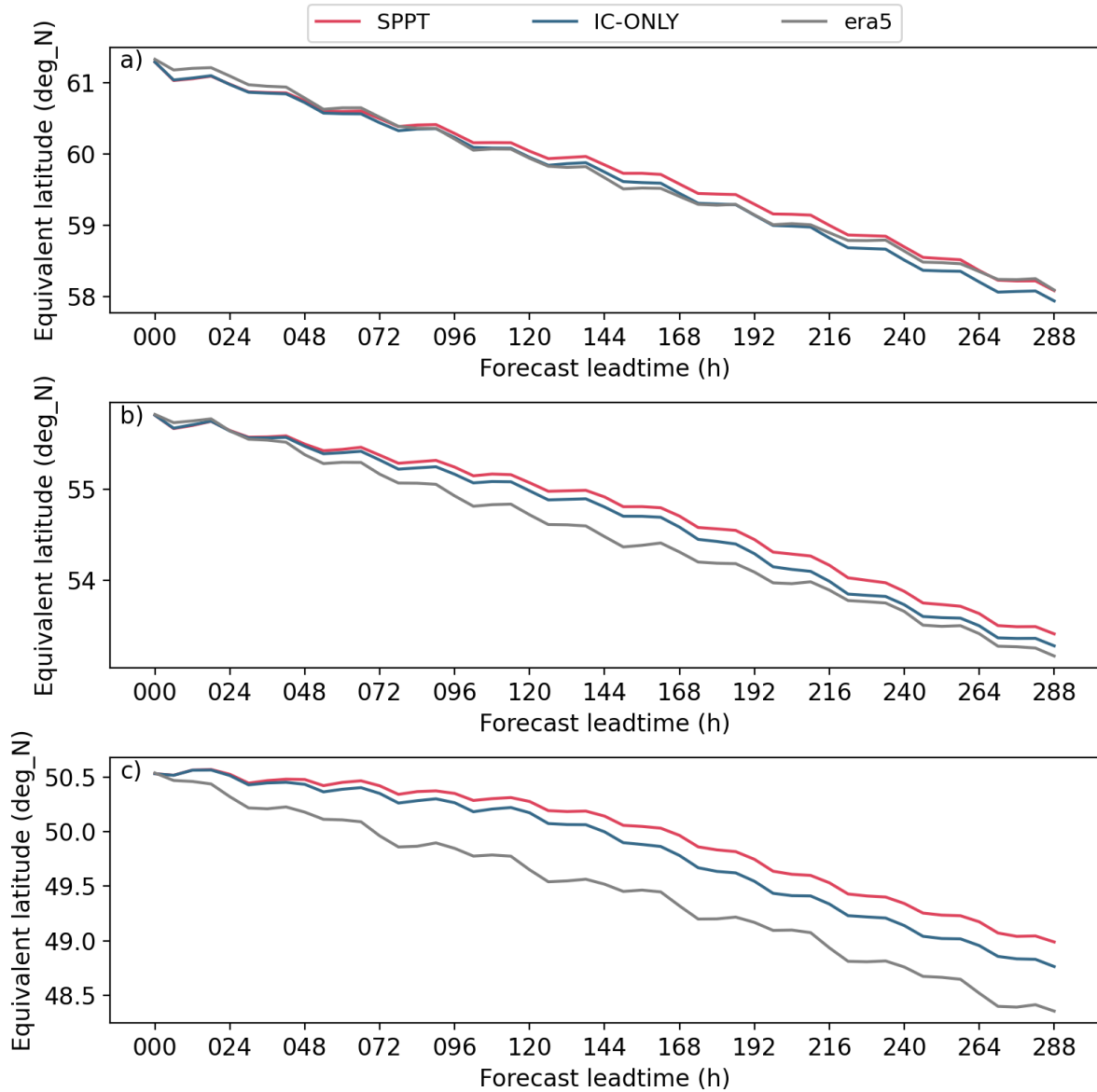


Figure 5.5.: Equivalent latitude Φ_{eq} with forecast lead time averaged over 32 forecast initializations and 20 members of the experiments *SPPT* (red), *IC-ONLY* (blue) and of *ERA5* (grey) on 320 K (a), 325 K (b) and 330 K (c).

days. The systematic southward drift in all data sets is a result of the experimentation period, which lies in the transition between summer and autumn, during which temperatures on the northern hemisphere decrease. Note that the diurnal cycle in all lines originates from temperature-induced fluctuations of PV in regions of high topography (e.g. over the Himalayas), where the isentropic surfaces intersect topography. Comparing the experiment *SPPT* to *IC-ONLY* shows that, despite identical initial conditions, the southward drift of Φ_{eq} is more pronounced in the experiment without SPPT on all isentropes: the differences between *SPPT* and *IC-ONLY* amount to 0.14° on 320 K, 0.13° on 325 K and 0.23° on 330 K at forecast lead time 288h. This displacement of Φ_{eq} in *SPPT* with respect to *IC-ONLY* corresponds

to a pole- and upward shift of the hemispheric tropopause through SPPT. Φ_{eq} in the reanalysis data set decreases at a faster rate than in the experiments (except for 320 K, panel a), hence *IC-ONLY* is more consistent with reanalysis than *SPPT*.

In the study of Gray et al. (2014), Φ_{eq} of an independent data set (i.e. ERA-Interim) is used to determine the areas of upper-level ridges of different models. However, when applying the same technique to the experiments shown here, the following problem emerges: as Φ_{eq} of the experiments differ from each other, an independent, constant Φ_{eq} for both experiments would result in biased ridge and trough areas. The hemisphere-mean 2 PVU line, and hence the whole Rossby wave pattern, is located further southward in *IC-ONLY* than in *SPPT*. Even if there is no difference in the ridge and trough amplitude between the experiments, a fixed Φ_{eq} for the computation of the areas would result in different areas of troughs and ridges. This problem is illustrated in Figure 5.8a: two wave patterns with different amplitudes are shifted relative to each other (red and blue lines) and are evaluated with the same value of Φ_{eq} (black line). This results in reduced ridge areas of the blue wave pattern, but similar trough areas in both data sets (see colored vertical lines as illustration for ridge and trough areas). In contrast, taking Φ_{eq} of each individual experiment for the computation of the ridge and trough areas considers the hemisphere-mean state of the upper-level Rossby wave pattern in each data set and results in both decreased ridge and trough areas for the blue pattern. This approach thus yields dynamically coherent and balanced ridge and trough areas (Figure 5.8b). For this reason, we use Φ_{eq} of each individual data set, computed separately for each valid time and ensemble member, to classify each grid point into one of the four categories. This is done for both the experiments as well as for the reforecast data set.

Ridge and trough areas

Figure 5.6 shows how the areas of all grid points identified as upper-level ridges and troughs on different isentropic levels evolve on average with forecast lead time. The analyzed ridge and trough areas (*ERA5*, grey line) on the 325 K isentrope (panels c and d) amount to $1.175 - 1.18 \cdot 10^7 \text{ km}^2$, which corresponds to about 4.5% of the area of the Northern Hemisphere. On 320 K (panel a and b), the areas increase with "lead time", whereas they decrease on 330 K (panel e and f). This is again due to the experimentation period in the transition time from summer to autumn, when ridge and trough areas become larger on lower isentropes. Note that the cyclic behaviour, especially at 320 K comes from the short data period with forecast initializations every second day, which leads to a distinct impact of single events.

Analyzing the evolution of the experiment *SPPT* (red lines in Figure 5.6) clearly indicates that the forecasts underestimate the areas of both the ridges (left) and troughs (right) on all isentropes. On 320 and 325 K, the areas of the troughs and ridges are only slightly underestimated until forecast day 6-7 (144-168h). Afterwards the difference to *ERA5* becomes larger, ending up with a reduction of the ridge area of approximately $1.3 \cdot 10^6 \text{ km}^2$ (10%) at 320 K at forecast day 12. On the other isentropic levels, the

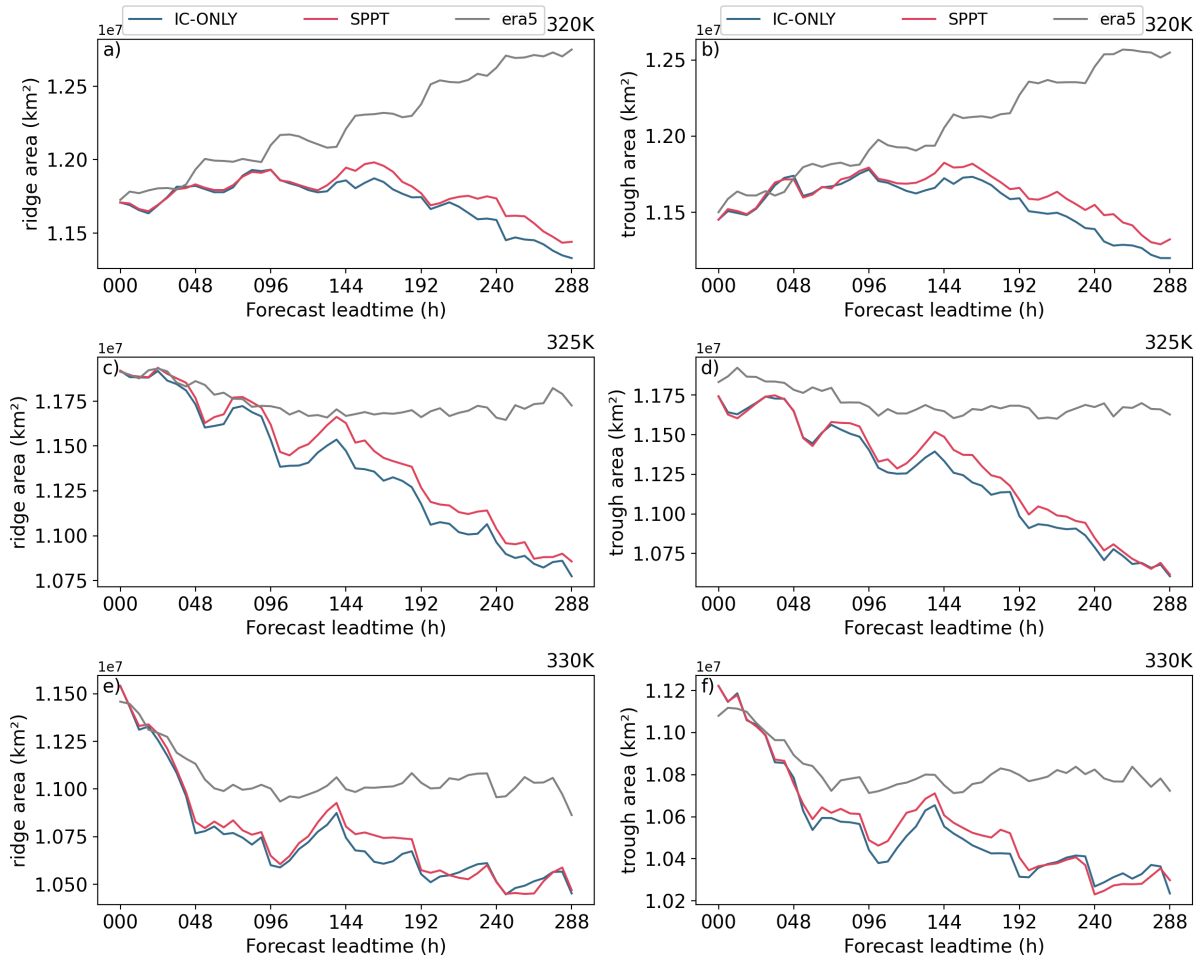


Figure 5.6.: Evolution of the area of upper-level ridges (left) and troughs (right) on the 320 K (top), 325 K (middle) and 330 K (bottom) isentropes with forecast lead time averaged over 32 forecasts and 20 ensemble members of the experiments *SPPT* (red) and *IC-ONLY* (blue). The grey line shows the corresponding areas derived from *ERA5*. The equivalent latitudes are taken from each individual experiment.

reduction is not as pronounced as on 320 K, but still amounts to 7% on 325 K and 4% on 330 K. The reduction of the trough areas is equivalent to the one of the ridges. This simultaneous decrease of ridge and trough areas corresponds to a reduction of the amplitude of the upper-level Rossby wave pattern: with forecast lead time, undulations of the upper-tropospheric wave guide become less pronounced and result in an overall more zonal flow configuration. This behaviour is in line with the result of previous studies: Gray et al. (2014) and Martínez-Alvarado et al. (2018) report a systematic decrease of the ridge areas on the northern hemisphere in several winter seasons and in different models.

The ridge and trough areas in the experiment without *SPPT* (*IC-ONLY*, blue line Figure 5.6) are reduced compared to *SPPT*. At the forecast initialization, *SPPT* and *IC-ONLY* have identical values (both start at the same initial conditions), but the areas gradually decrease with forecast lead time in *IC-ONLY* compared to *SPPT*. This is especially visible at 320 K, where the ridge area on forecast day 12 is decreased by

approximately 1% compared to *SPPT* (equivalent to an area of $1.1 \cdot 10^5 \text{ km}^2$). On 325 K, this behaviour is also apparent, but less pronounced. At 330 K, the areas of *IC-ONLY* are reduced up to forecast day 8 (192h), but are then very similar until the end of the forecast.

This shows that the activation of *SPPT* leads to an increased amplitude of the upper-level Rossby wave pattern compared to experiments, especially at 320 and 325 K, and hence also to a reduced underestimation of the waviness compared to *ERA5*. In other words, *SPPT* helps to maintain the upper-level ridge and trough areas against the systematic lead-time dependent degeneration of the Rossby wave amplitude. The magnitude of the effect is, however, relatively small and results in only minor improvements of the upper-level tropopause structure.

Polar vortex and subtropics areas

Figure 5.7 shows the evolution of the areas of the grid points which are classified as polar vortex (left) and subtropics (right). These two categories provide information on the upper-level PV budget and are therefore directly linked to Φ_{eq} , but also depend on the waviness of the upper-level jet: increased ridge and trough areas will result in decreased areas of the polar vortex and subtropics, and vice versa. The areas of the polar vortex increase with forecast lead time in the reanalysis (grey line) as well as in all experiments on all isentropes, whereas the area of the subtropics decreases. This behaviour is again a result of the experimentation period, where temperatures on the northern hemisphere decrease, resulting in a south- and downward propagation of the 2 PVU line on a defined isentrope. This leads to a growth of the area of the polar vortex at the expense of the subtropics, without affecting the size of the troughs and ridges.

Comparing the forecast experiments to *ERA5* shows that the increase of the polar vortex area takes place faster, and the decrease of the subtropics area is slower in all experiments. This occurs on all isentropes, except for the polar vortex at 330 K, where the growth of the polar vortex area is faster in *ERA5* than in the experiments. This mimics the evolution of the ridge and trough areas in Figure 5.6: both ridge and trough areas are reduced in the experiments compared to *ERA5*. This has to result in larger polar vortex and subtropics areas in the experiments, as the reduced waviness is filled with grid points associated with the polar vortex and the subtropics.

For the same reason, the polar vortex areas are larger in *IC-ONLY* than in *SPPT* (1.7% on 320 K, 0.9% on 325 K and 1.4% on 330 K): the increased ridge areas with *SPPT* come at the expense of shrunk polar vortex areas. Counterintuitively, the subtropics areas are larger in *SPPT* compared to *IC-ONLY*. This is due to the fact that not only the amplitude of the Rossby waves is modulated by *SPPT*, but also the hemisphere-mean position of the 2 PVU line, which is further to the south in *IC-ONLY* than in *SPPT*. This southward displacement of the hemispheric wave pattern in *IC-ONLY* results in a decrease of the

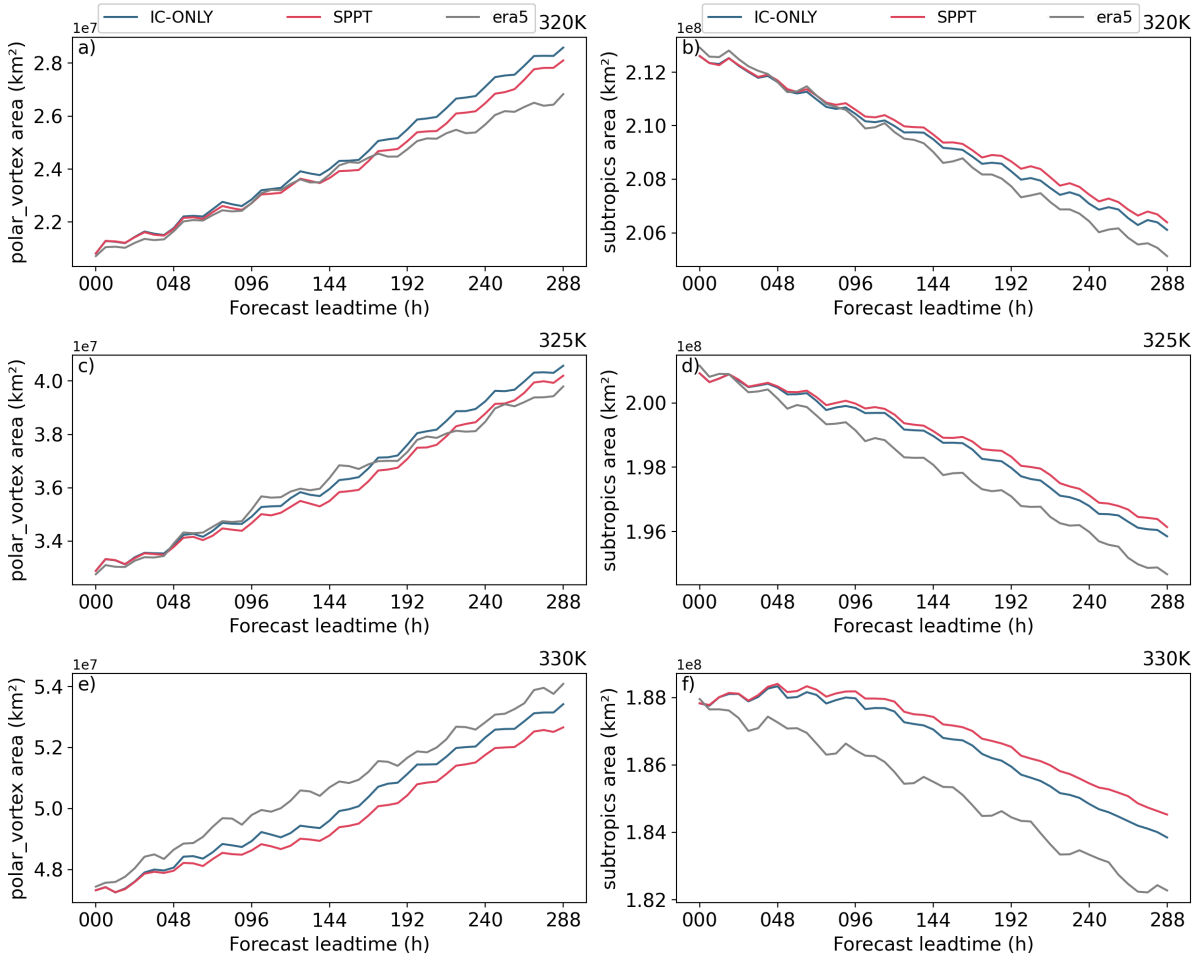


Figure 5.7.: As Figure 5.6, but for the categories polar vortex and subtropics

subtropical areas compared to *SPPT*, opposing the impact of the reduced waviness of the 2 PVU line. The impact of the southward displacement of Φ_{eq} without *SPPT* outweighs the effect of the decreased waviness. Hence, the superposition of these two independent effects leads to the opposite effect on the subtropics than on the polar vortex areas.

The twofold effect of *SPPT* on the structure of the upper-level Rossby-wave pattern is illustrated in Figure 5.8b and summarized as follows:

1. The amplitude of upper-level ridges and troughs (i.e. the Rossby wave pattern) is systematically increased with *SPPT* compared to experiments without physics perturbations. The effect is on the order of magnitude of $\mathcal{O}(1\%)$, but depends on the chosen isentropic level.
2. The hemispheric-mean latitude of the 2 PVU line (i.e. Φ_{eq}) is located further northward and at a higher potential temperature level with *SPPT* than without.

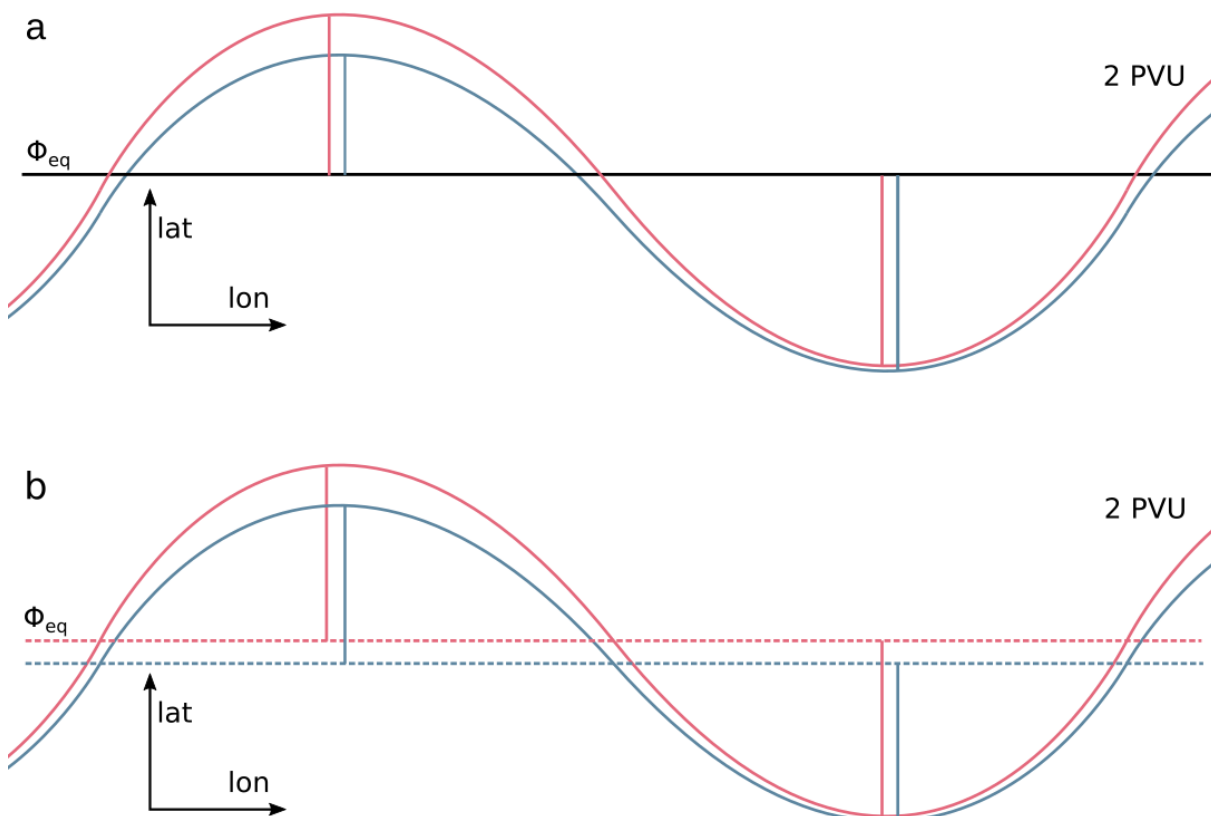


Figure 5.8.: Simplified schematic of a ridge-trough pattern to illustrate the effect of the choice of Φ_{eq} for the ridge and trough area detection. Panel a) shows the ridge and trough identification with a fixed, independent Φ_{eq} , e.g. from ERA5 (black line). Panel b) shows the case where Φ_{eq} is taken from each data set individually (dashed red (*SPPT*) and blue (*IC-ONLY*) lines). The solid red and blue lines indicate the 2 PVU line of the experiments *SPPT* and *IC-ONLY*, respectively. The coloured vertical lines illustrate the amplitude of the ridge/trough, which is here assumed to be proportional to the area of the feature.

As there are two different mechanisms at play, it is crucial to consider both of them for the evaluation of the ridge and trough areas. Ignoring the effect of *SPPT* on Φ_{eq} would lead to artefacts in the results (i.e. increased ridge areas at the expense of trough areas), as shown in Figure 5.8a.

5.2.3. Exploiting forecast archives

Similarly to the trajectory analysis in the previous chapter, we will make use of the finding that the unperturbed control forecast of routinely issued forecasts can be interpreted as experiment without physics perturbations (i.e. *IC-ONLY*), and the perturbed forecasts can be regarded as sensitivity experiment with *SPPT* (i.e. *SPPT*). For the evaluation of the Rossby wave amplitude, this is very useful, because the methodology to detect upper-level troughs and ridges only requires 2D-fields of isentropic potential vorticity, which is typically available in forecast archives. The data basis that has been used in the previous section to analyze the effect of *SPPT* on the Rossby wave amplitude in the experiments is rather sparse: as the spatial information of the forecasts gets lost during the computation of the ridge and trough areas,

only $n_{\text{forecasts}} \cdot n_{\text{members}} = 620$ data points are available for every forecast lead time, which makes the results noisy. Further, the experimentation period between August and October imposes some problems related to seasonality, as described above.

Therefore the analysis from above is repeated with data taken from the Sub-seasonal to Seasonal Prediction (S2S) data base (Vitart et al., 2017), where reforecasts from ECMWF are considered until lead times of 15 days. Until this lead time, the forecasts are run at the same resolution as the operational medium-range forecasts (about 18km), but consist of only 10 perturbed ensemble members. The perturbations are identical to the ones from the operational forecasts (IC-perturbations and SPPT), hence the comparison of the perturbed and unperturbed members is equivalent to the comparison of the experiments *SPPT* and *IC-ONLY*, apart from the first time step that is influenced by initial condition perturbations. In the data base, only one isentropic level ($\Theta=320$ K) is available. Ensemble forecasts initialized twice-weekly between 1997 and 2017 are analyzed, where each forecast is classified by the season of the forecast initialisation (as in Figure 4.9). The number of forecasts is 920 forecasts in the winter season, 1060 forecasts in spring, and 1040 forecasts in autumn, respectively. Summer is omitted here, because the 320 K isentrope does not represent the upper levels. The value of the equivalent latitude Φ_{eq} for the classification procedure is taken from the individual forecasts, as described previously. For further details regarding the data set, see Table 3.3.

The relative differences of the ridge (a) and trough (b) areas between the perturbed and unperturbed forecasts are shown as a function of forecast leadtime for forecasts initialized in winter, spring and autumn in Figure 5.9. Despite the large data basis, the curves are noisy because the unperturbed forecasts consist of only one ensemble member. In all seasons, the ridge and trough areas are on average smaller in the unperturbed forecasts than in the perturbed forecasts, with the largest signal in autumn (on average 0.65% difference), followed by winter (0.45%) and spring (0.2%). In autumn, the difference ridge and trough areas increase with forecasts leadtime, reaching differences of up to 0.8-1.2% after lead times of 12-15 days. This increase is, however, not strictly monotonic. In winter, the differences between perturbed and unperturbed members are similar to the ones in autumn until lead times of about 10 days (240 hours) and reach a maximum area difference of about 1% between forecast days 5 and 7; after day 10, the Rossby wave amplitude is only marginally larger in the perturbed forecasts and even reaches the one of the unperturbed forecasts on day 15. Perturbed forecasts initialized in spring are characterised by only slightly increased Rossby wave amplitude compared to the unperturbed forecasts (maximum of about 0.6 - 0.8%) at day 11. Similarly to winter, the differences vanish for later lead times.

The qualitative behaviour of the area differences is quite similar across the three seasons: an increase of the Rossby wave amplitude in the perturbed forecasts during the first few days is followed by a dip between forecast days 5-10. Afterwards, the differences again increase, but are subsequently reduced to-

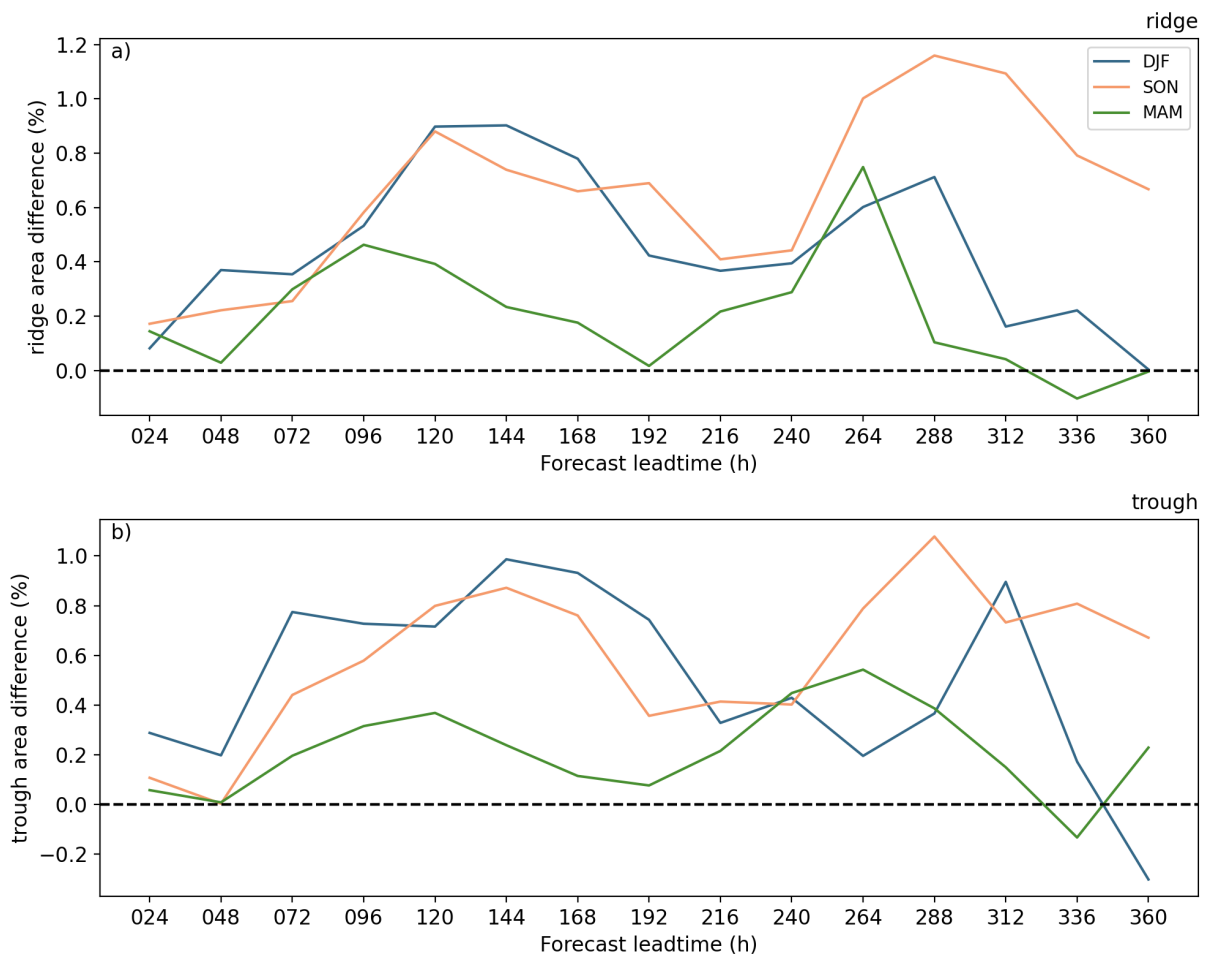


Figure 5.9.: Averaged ridge (a) and trough (b) area differences between perturbed and unperturbed forecasts with forecast lead time in winter (DJF, blue), spring (MAM, green) and autumn (SON, orange) of S2S ECMWF reforecasts initialized 1997 to 2017. Note that the first time step is not shown, as it includes effects from the initial condition perturbations.

wards the end of the investigated forecast period of 15 days. Despite the shift of the maxima and minima to different lead times, the similarities of the wave-like patterns between the seasons suggests a common underlying mechanism that exerts a stronger impact in autumn than in winter and spring.

The comparison of perturbed and unperturbed members of ECMWF S2S forecasts corroborates the main findings from the sensitivity experiments in the previous section. Even though only one isentropic level ($\theta=320$ K) is investigated, the large sample size in comparison to the experiments allows for an assessment of the robustness of the observed patterns. In the experimentation period, which is mainly in autumn 2016, the amplitude of the upper-level wave pattern on 320 and 325 K is increased when perturbations through SPPT are active. A quantitatively very similar pattern is found for 21 years of S2S reforecasts initialized in autumn, where the order of magnitude ($\mathcal{O}(1\%)$) of the observed effect is similar. This behaviour can therefore be considered as robust, even though it does not pass a test of statistical

significance². On the other isentropic levels in the experiment data set as well as in other seasons than autumn in the S2S data set, the signal is not as distinct as in autumn and on 320 K.

5.2.4. Other schemes

In this section, the evolution of the Rossby wave amplitude is investigated in the experiments with the perturbation schemes in set 2. The diagnostics related to the Rossby wave amplitude require a larger data basis than the trajectory- and grid-point based diagnostics in the previous chapters, as the dimensions longitude and latitude are lost. We therefore omit the experiments *SPP-CONV-ONLY* and *SPP-CONV-OFF* and only consider the experiments *SPP* and *STOCHDP*, where 32 initial dates are available. This analysis aims to investigate whether the effect on the upper-level Rossby wave structure observed in the previous section is exclusive to SPPT, or if it can be generalized for stochastic model uncertainty schemes. Afterwards, we attempt to link the different magnitudes of the effect on the Rossby wave amplitude between the model uncertainty schemes to the results from the trajectory analysis in Chapter 4. Consistently with Chapter 4, only a reduced selection of diagnostics is presented for the experiments in set 2.

The evolution of the ridge and trough areas with forecast lead time for the experiments in set 2 shows that *SPP* and *STOCHDP* underestimate the Rossby wave amplitude compared to *ERA5*, especially for lead times later than 6 days and on the 320 and 325 K isentrope (Figure 5.10). This behaviour is qualitatively similar to the evolution of the ridge and trough areas of *SPPT*, which is plotted again for reference. Compared to *IC-ONLY*, the model perturbations through *SPP* increase both the ridge and trough areas on 320 and 325 K at late lead times, but the magnitude is slightly smaller than in the experiment with *SPPT* (particularly on 320 K). Also, the amplitudes of *SPP* and *IC-ONLY* diverge at later lead times (around day 7-8) than it is the case for *SPPT* (day 5-6). This is in accordance with the weaker impact on the forecasts in the extratropics of *SPP* than *SPPT* (Leutbecher et al., 2017), that was also observed for the trajectory count in the extratropics and in the North Atlantic region. The ridge and trough areas in the experiment *STOCHDP* evolve similarly to the ones of *SPP* until forecast lead times of about 6-9 days. After that, the Rossby wave amplitude decreases relative to *SPP* and approaches the one of the unperturbed forecast experiment (especially on 325 and 330 K). The qualitative evolution of the equivalent latitude Φ_{eq} behaves accordingly: In both *SPP* and *STOCHDP*, the equivalent latitude decreases with forecast leadtime, but the decrease with *STOCHDP* is faster than with *SPP*. Hence, Φ_{eq} in *STOCHDP* is the closest of all experiments to the one without physics perturbations (see Figure C.1 in Appendix C); the same holds true for the polar vortex and subtropics areas, where *SPP* is closest to the experiment

²A student's t-test was applied to the ridge and trough area differences between the perturbed and unperturbed members of the S2S reforecast data set. Based on a 95% confidence interval, only the differences in autumn at lead times of 288 hours are statistically significant.

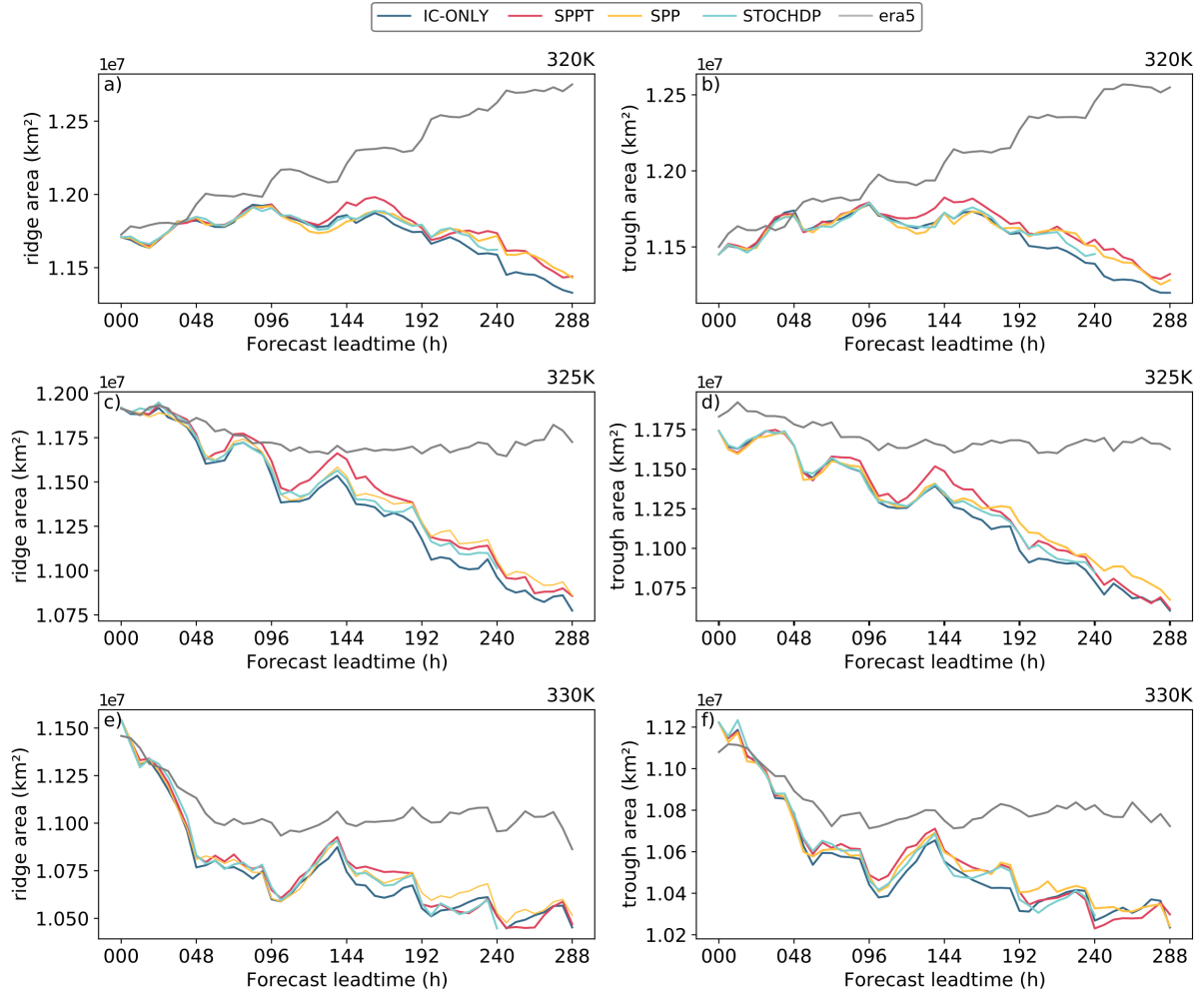


Figure 5.10.: Ridge and trough areas with forecast leadtime, averaged over 32 forecasts and 20 ensemble members of the experiments *IC-ONLY*, *SPPT*, *SPP*, *STOCHDP* and *ERA5*. Note that the time series of *STOCHDP* is limited to forecast lead times up to 240 hours.

with *SPPT*, and *STOCHDP* is closest to the unperturbed simulations (see Figure C.2 in Appendix C).

The analysis of the experiments in set 2 shows that not only *SPPT*, but all the examined stochastic model perturbation schemes have an impact on the representation of the upper-level Rossby wave pattern by amplifying its waviness and by shifting its mean latitude northwards. This indicates that the modulation of the amplitude and the mean position of the tropospheric wave guide does not solely result from perturbations through *SPPT*, but also occurs through other types of model uncertainty schemes. *SPPT* has a slightly larger impact on the Rossby wave amplitude than *SPP*, and *STOCHDP* exerts the weakest influence of the investigated schemes. However, the differences between the techniques are minor, and also the offset to the unperturbed simulations is rather small.

5.2.5. Discussion

This section demonstrated that the upper-level Rossby wave structure is modified by stochastic perturbations of the forecast model: the areas of troughs and ridges are increased when stochastic model perturbations are active during the forecasts, leading to a more amplified wave pattern. Furthermore, the mean position of the 2 PVU line is pushed north- and upwards through the perturbations. The magnitude of this effect is, however, rather small and statistically not significant. Here it will be discussed whether the behaviour can be related to the increased frequency of rapidly ascending air streams through model uncertainty representations that has been described in Chapter 4.

The initial hypothesis that the increased diabatic outflow from WCBs through stochastic model uncertainty schemes should result in a more amplified upper-level flow can be confirmed insofar that the ridge and trough areas are larger with than without model perturbations. The order of magnitude of the effect of model physics perturbations on the trajectories is, however, one order of magnitude larger than the effect on the upper-level troughs and ridges: with SPPT, the counts of trajectories that are detected as WCBs are increased by approximately 10-20% in the northern hemisphere extratropics (exact numbers depend on the season, see Chapter 4), while the ridge and trough areas are increased by up to 1%. Comparing these two numbers with each other has to be done with caution, because of the following considerations:

- For the trajectory counts, only such trajectories that ascend by at least 600 hPa within 2 days are considered. Air streams that fail to fulfill this criterion (maybe by only a few hPa) are not considered in this diagnostic, even though they might exert a similar impact on the upper levels. Further, the perturbation schemes do not trigger WCB-like air streams in situations where there is no ascending motion at all, but tend to push trajectories across a threshold. The additional diabatically induced divergent outflow at the tropopause level responsible for the ridge amplification is therefore likely not of the same order of magnitude as the trajectory count.
- The ω -distributions of the experiments showed that SPPT increases the occurrence of downward motions of all velocities, which is not captured by the trajectory diagnostic. These downward motions balance the increased upward motions, and tend to decay upper-level ridges (i.e. upper-level convergence).
- The divergent outflow of diabatically enhanced ascents is not the only process which contributes to the formation and amplification of upper-level ridges. Teubler and Riemer (2016) and Teubler and Riemer (2021) highlight that barotropic dynamics at the tropopause, the interaction of baroclinic cyclones with the upper-level jet, and other diabatic processes (such as long-wave radiation) influence troughs and ridges. Without the knowledge of how these processes are changed by stochastic perturbations, it is not possible to make a quantitative statement on the role of altered WCB frequencies for the ridge and trough areas.

- The dynamical response of the upper-level jet to the divergent outflow of WCBs depends on the distance of the outflow to the jet and on the outflow height relative to the tropopause level (Grams and Archambault, 2016). The tropopause height and the latitude of the 2 PVU line (i.e. Φ_{eq}) are both increased through the perturbations, which makes it more "difficult" for the WCB outflow to impinge on the jet stream and to initiate or amplify ridge building, assuming similar characteristics of the WCB outflow (e.g. outflow height and latitude) with and without perturbations.

From these aspects, it becomes clear that the modulated WCB-occurrences cannot be translated directly (i.e. one-to-one) to the changes of the ridge and trough areas. Apart from these discrepancies, a causal relationship between the effect of stochastic perturbation schemes on both the WCB-occurrence and the upper-level Rossby wave patterns is supported by the following considerations:

- The analyses of the operational forecast archives have shown that the trajectories and the Rossby-wave amplitude are modulated strongest in autumn and weaker in winter. This points towards a direct interrelation and a common underlying mechanism of the processes. The discrepancies in spring (strong impact on trajectories, but weak impact on Rossby wave amplitude) are due to the fact that despite the large relative increase of WCB trajectory counts, the absolute numbers are rather small (see Figure 4.9 and Figure 4.10). The resulting differences in the upper-level divergent outflow are therefore expected to be smaller than in winter and spring, where the absolute WCB-frequencies are larger.
- SPP alters the vertical velocities and the frequency of WCB-trajectories in a similar way as SPPT, but the magnitude in the northern hemisphere extratropics and in the North Atlantic is slightly reduced compared to SPPT (see comparison of Figure 4.3 and Figure 4.11). Consistently, the impact on the ridge and trough areas is comparable, but slightly weaker in *SPP*, which could be due to the reduced upper-level divergent outflow. In *STOCHDP*, the frequency of WCBs is not changed compared to the unperturbed forecasts, and *STOCHDP* is the scheme which has the weakest impact on the ridge and trough areas. Nevertheless, the Rossby wave amplitude differs slightly from the one in *IC-ONLY*, even though their trajectory counts are equivalent. However, *STOCHDP* exerts a systematic impact on the slowly ascending and descending motions (see Figure 4.16), that is not reflected in the trajectory diagnostics. This acceleration of the vertical motions might also influence the waviness of the tropospheric wave guide, but to a smaller extent than the other investigated schemes.
- In contrast to the immediate and constant effect that the perturbations have on the vertical velocities, the ridge and trough areas are identical at the beginning of the forecasts with and without model perturbations (only in the experiments, not in the operational forecasts, where the control member does not have initial condition perturbations) and diverge very slowly with forecast lead time. This points towards a weak, yet constant forcing of the increased mass flux through the

enhanced WCB activity on the ridge and trough areas. The saturation of the process (i.e. the decreasing and vanishing differences of ridge and trough areas between perturbed and unperturbed forecasts, see Figure 5.6 and Figure 5.10) as well as the wave-like pattern of the evolution of the differences (Figure 5.9) could arise from compensating effects of the model in the unperturbed forecasts, without which the differences would constantly grow with forecast lead time.

The question if and how representations of model uncertainties modify the large-scale extratropical circulation has been investigated in previous studies (see Chapter 2 for details). It has been shown that stochastic model perturbations improve the representation of Euro-Atlantic weather regimes across different model hierarchies (Dawson and Palmer, 2015; Christensen et al., 2015; Dorrington, 2021), especially for such regimes that are characterized by blocking anticyclones. However, the mechanisms behind the improved behaviour, which is mainly related to more a realistic persistence, are not clear and partly contradictory between the studies. Further, the reported impacts on the large-scale circulation are mostly very subtle, especially in numerical models of high complexity (e.g. Davini et al. 2021, Dorrington 2021). Christensen et al. (2015) argue that stochastic forcing enables a more realistic sampling of Lorenz-like attractors in models of reduced complexity, as the introduced noise helps to transition between stable states of the system (i.e. noise-induced regime transitions; Berner et al. (2015)). Dorrington (2021) mention that improved representations of the Atlantic ridge regime in fully coupled simulations with SPPT might be driven by improved tropical modes of variability (i.e. ENSO) whose signal is transferred to the extratropics via teleconnections. Martínez-Alvarado et al. (2018) state that differences in the sharpness of the tropospheric wave guide between perturbed and unperturbed forecasts are directly induced by vorticity perturbations along the large gradients at the dynamical tropopause. This latter argumentation contrasts our results, as the STOCHDP perturbations exert the weakest impact on the ridge and trough areas, even though it is likely to introduce large perturbations directly into regions close to the jet stream. Nevertheless, it is difficult to compare the approaches in the studies, as many different forecast models, perturbation techniques, and diagnostic methods have been applied to study the effects of stochasticity on the large-scale circulation.

With the analysis provided in this thesis, we contribute to the discussion how the large-scale extratropical flow is modified through stochastic model perturbations on a process level: we hypothesize that the stochastic perturbation schemes amplify the upper-level Rossby wave pattern by modulating the occurrence frequency of rapidly ascending, moist air streams, which occurs due to the nonlinear nature of systems that are characterized by threshold behaviour (see Figure 4.17). We thereby put forward a process chain from the random perturbation patterns which are projected by ascending air streams from the small to the large scale, where they systematically affect the Rossby wave structure.

5.3. Summary and Outlook

Summary

In this chapter, we showed that model uncertainty schemes do not only change the distribution of vertical motions in the atmosphere and thereby alter the occurrence of rapidly ascending air streams, but also have an impact on other aspects of the model climate as well. In a first part, it was demonstrated that the global distribution of precipitation is affected through the representation of model uncertainty, and these changes are clearly linked to the changes of the vertical motions. The precipitation sum associated with coherent masks of rapidly ascending air streams is increased with SPPT compared to simulations with unperturbed model physics. Further, the complex, bimodal pattern of the modified ω -distribution through SPPT and SPP (see Chapter 4) is directly visible in the frequency changes of precipitation, with increased occurrences of precipitation rates in ranges corresponding to increased occurrences of vertical velocities, and vice versa. Accordingly, the effect of STOCHDP on precipitation is characterized by only one single pair of positive and negative values, which directly reflects the changes in the vertical velocities.

In the second part, it was shown that the model uncertainty schemes analyzed in this study have a subtle, yet systematic impact on the upper-level Rossby wave structure: experiments with model perturbations are characterized by a slightly increased wave amplitude, and the whole wave pattern is shifted polewards compared to unperturbed simulations. These changes are stronger with SPPT and SPP than with STOCHDP, but are overall of small magnitude. By comparing the detailed effects of the different schemes on the WCB-frequencies and on the Rossby wave amplitude (such as differences between the schemes or common seasonal characteristics), we attempted to establish a causality between the two processes. The consistent process chain which starts with the small-scale localized perturbation, modulates the mass flux of diabatically driven weather systems into the upper troposphere and finally modifies the large-scale flow, is a strong hint towards causality. However, it is not possible to prove this causal relationship with the available setup, but further ideas how this can be achieved are given in the following section.

Outlook

In the following, a list of promising approaches is given how a direct relation between the two observed effects could be established. All of these methods require a tailored experimentation with specific model output, which goes beyond the scope of this thesis and was therefore omitted.

- Teubler and Riemer (2016) developed a method to account for different processes that contribute to the PV-budget of dynamic features, such as upper-level ridges and troughs. With their approach, it is possible to differentiate between PV-tendencies emerging from barotropic processes at the

tropopause level, from baroclinic processes due to low-level PV-anomalies, and from the divergent wind component related to diabatic processes, such as in WCBs. Comparing these PV-tendency terms among experiments with and without model perturbations could not only yield a better understanding if diabatic processes are relevant for the observed changes of Rossby wave amplitude, but also provide valuable information about the impact of stochastic model perturbations on other processes.

- Computing backward trajectories from objectively identified troughs and ridges, equivalent to the studies of Pfahl et al. (2015), Steinfeld and Pfahl (2019) and Steinfeld et al. (2020), could give a direct Lagrangian quantification of the fraction of diabatically heated air streams contributing to upper-level ridges (and troughs) in different experiments.
- Finally, sensitivity experiments with perturbations activated regionally, such as only in the extratropics, could provide information if the signal originates from remote (tropical) teleconnections or if it is induced locally. An equivalent approach with perturbations only active in pre-defined height layers of the atmosphere (e.g. by modifying the tapering function) could be used to determine whether the Rossby wave pattern is altered due to local, upper-level perturbations at the tropopause, or if the perturbations are projected by WCBs from the lower to the upper levels.

6. The role of WCBs for forecast error growth

This chapter switches perspective and gives a detailed investigation on the impact of WCBs on forecast error and error growth in ensemble forecasts. Other than the two previous chapters, in which ensemble techniques have been investigated by analyzing sensitivities of WCBs and related phenomena to different model uncertainty schemes, the following chapter focuses on the role of WCBs for the performance of operational ensemble forecasts. Several studies have shown that WCBs can be involved in the degradation of forecast skill, and two pathways are discussed in the literature:

1. As WCBs have a substantial impact on the evolution of the large-scale flow, they can project previously existing, small-scale errors from the inflow region onto the large-scale circulation. Examples for studies elaborating on this process chain are Grams et al. (2018) and Berman and Torn (2019).
2. The ascending phase of WCBs is characterized by moist-diabatic processes which are mostly parametrized in current global NWP-models. The parametrizations of these processes are subject to uncertainties, which are projected onto the outflow and subsequently on the upper-level Rossby wave structure. The forecast uncertainty related to parametrized processes in WCBs has been studied extensively in recent years, mainly by running simulation experiments of individual WCB events with different parametrization schemes. Recent examples for such analyses include the work of Mazoyer et al. (2021), Rivière et al. (2021) and Choudhary and Voigt (2022).

However, all of these analyses are based on case studies, which means that no systematic relationship can be inferred. On the other hand, there are studies that systematically investigate the origin and dynamics of forecast errors and found that diabatic processes in synoptic-scale weather systems tend to reduce forecast skill (e.g. Rodwell et al. 2013 or Sánchez et al. 2020). Nevertheless, these studies do not use the "classical" Lagrangian approach to detect WCBs (Wernli and Davies, 1997), but employ proxy metrics to characterize the imprint of the divergent outflow of WCBs.

Thanks to a unique data set that has been assembled by means of real-time retrievals of ensemble forecasts on model levels suitable for trajectory computations (see Chapter 3 Data and Methods), it is possible to close this gap and systematically investigate the link between WCBs and forecast error (growth) in a Lagrangian framework. This is done in the following way: after Section 6.1, which introduces the data set and methodology of the chapter, the climatological co-occurrence of WCBs and forecast errors is investigated (Section 6.2). The temporal relationship of WCBs and forecast error growth is analyzed

in Section 6.3, after which a composite approach is employed to obtain spatio-temporal error patterns centered on WCB-objects (Section 6.4).

6.1. Data and Methods

For this study, we use imprints of WCBs computed from operational ECMWF forecasts (see Table 3.3). In contrast to Chapter 4, where trajectory data was analyzed in all seasons of 2 years of forecast data, here we only consider forecasts that are initialized in winter (DJF), but include one additional year, resulting in three winter seasons of ensemble forecasts (2018/2019, 2019/2020 and 2020/2021). With 50 perturbed and one unperturbed ensemble member and two initializations per day (00 UTC and 12 UTC), this sums up to more than 27000 individual forecasts, in which the data is analyzed until lead times of 10 days. WCBs are detected in the North Atlantic region by selecting 2-day forward trajectories that ascend by at least 600 hPa. For each forecast lead time, we obtain binary masks of WCB inflow, ascent and outflow, based on the pressure of the trajectory points (for details of the methodology refer to Chapter 3).

The forecast performance is evaluated based on two different fields: geopotential height at 500 hPa (Z500), and the wind speed at the hybrid model level $\sigma=52$, which corresponds to a pressure of approximately 250 hPa (WS250). The former (Z500) is a well-established variable to evaluate forecast skill and is located in a height layer of the troposphere where WCB ascent takes place (i.e. trajectories between 800 and 400 hPa). The latter (WS250) is used to evaluate the forecast performance in the upper troposphere. From the available data set, it is non-trivial to compute geopotential height at other pressure levels than 500 hPa. As data manipulation for such a large data set is very time-consuming, we use WS250 for practical reasons, as the field is archived in a local data base. Z500 is used throughout the whole chapter, whereas WS250 will mainly be used at a later stage of the analysis. The analysis is restricted to the North Atlantic domain (60°W-0°E, 35°-75°N, see white boxes in Figure 6.3), as WCB-imprints are only available regionally.

The forecast error is expressed in terms of the root-mean squared error (RMSE) of Z500 and WS250, respectively, has been introduced in Chapter 3. The RMSE is expressed in the unit of the input field and refers to a perfect forecast for the value 0. In this chapter, we will analyze the RMSE spatially averaged, but also in grid-point space. As verifying analysis, we use the forecast field of the unperturbed control member at lead time 0 h for valid times at 00 and 12 UTC, and 6 h lead times for the valid times 06 and 18 UTC. This is done to ensure consistent vertical levels when evaluating the wind speed on a defined model level (which would not be possible with the high-resolution analysis or ERA5, which both have 137 levels instead of 91).

In the last part (Section 6.4), forecast errors related to WCBs are evaluated by computing RMSE-composites centered on WCB-objects. The used approach is introduced in some detail in the following.

At first, WCB-objects are detected as coherent regions of grid points that are associated with imprints of WCB-ascent or outflow. These imprints are based on the trajectory gridding approach described in Chapter 3. The size distribution of the WCB ascent (yellow) and outflow (blue) objects shows that most of the detected objects are rather small: for example, 50% of the outflow objects are smaller than 81.000 km^2 (i.e. $0.081 \cdot 10^6 \text{ m}^2$, see Figure 6.1). Further, the ascent objects are on average smaller than the outflow objects, which is mainly reflected in the longer tail of the distribution of the outflow objects. In order to consider only such WCB-objects that are relevant for the modification of the large-scale flow, only objects with a minimum size of 300.000 km^2 are considered¹ (which is equivalent to about 35 grid cells at a mean latitude of 45°N). This threshold is to some extent arbitrary and was chosen based on a visual inspection of individual cases. The value of 300.000 km^2 corresponds to the 78th percentile of the ascent objects and to the 75th percentile of the outflow objects, respectively, resulting in about 550.000 ascent and about 750.000 outflow objects distributed across the entire data set for the forecast lead times 72-168 hours.

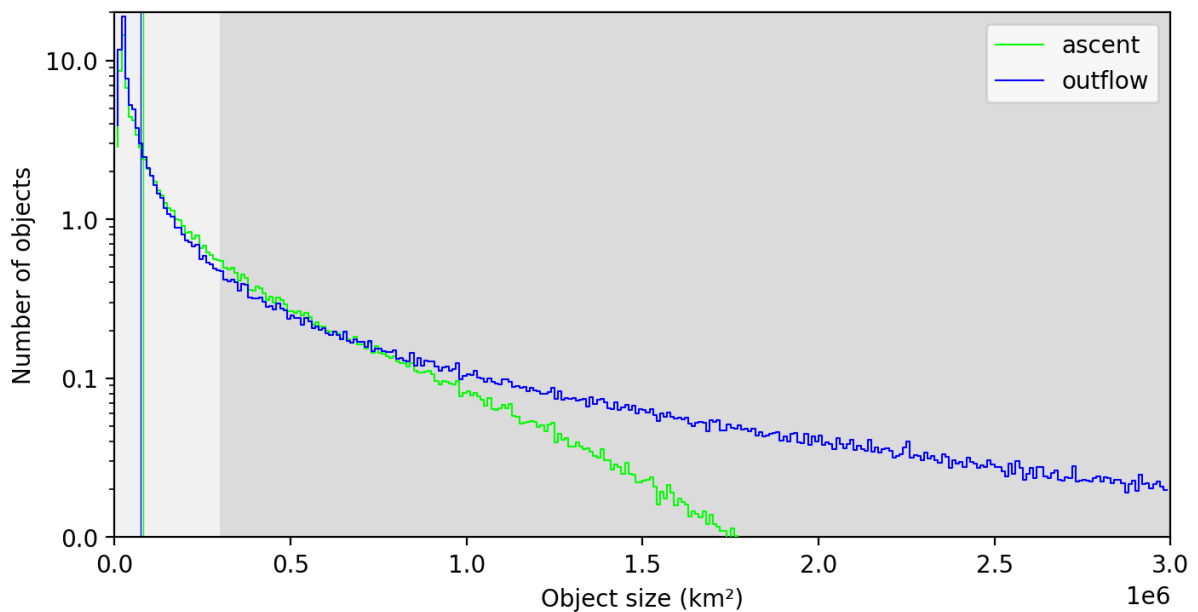


Figure 6.1: Object size distribution of WCB ascent (green) and outflow (blue) objects in the North Atlantic domain (60°W - 0°E , 35° - 75°N), detected in the ensemble data set. The dark gray shading denotes the size range in which WCB objects are considered. The bin width is 10.000 km^2 . The colored vertical lines show the median size of the ascent (green) and outflow (blue) masks. The counts are normalized by the number of forecasts in the data set, resulting in the average number of objects of the corresponding size per 168 h-forecast. Note the logarithmic scale on the y-axis.

¹For comparison: the area of Germany is ca. 350.000 km^2

For every object, the center of mass is then determined. Note that the grid point to which the center of mass is assigned does not necessarily lie within a WCB object. Subsequently, the coordinates of a $60^\circ \times 40^\circ$ longitude-latitude box around the center of mass are extracted. This procedure is illustrated in Figure 6.2 for outflow objects in one arbitrarily chosen situation. In the shown example, only the objects with the ID 2 and 6 are retained, all other objects are smaller than the selection threshold. Note that no object tracking has been performed, which means that objects associated with the same feature can be identified multiple times at subsequent time steps. The coordinates of the $60^\circ \times 40^\circ$ lon-lat box are then used to compute composites of different variables centered on the WCB objects with time lags 3 days prior to until 3 days after the detection².

The composites are calculated for three different types of variables: WCB imprints of inflow, ascent and outflow, meteorological variables (mean sea level pressure (MSLP), Z500, WS250) and the RMSE of Z500 and WS250. The WCB imprints are absolute frequencies (0-1 masks) and can therefore be

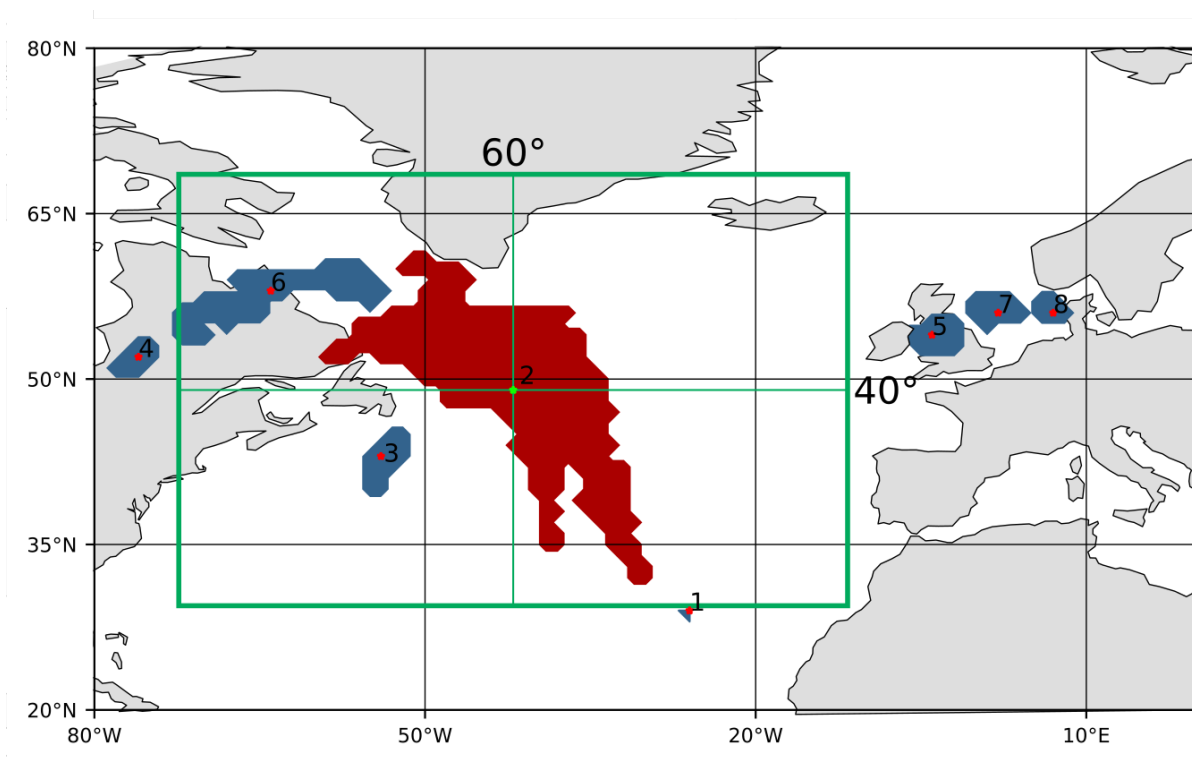


Figure 6.2.: Illustration of the WCB object detection and selection, and the subsequent construction of object-centered composites. The shading shows WCB outflow objects that have been detected by the object identification algorithm at a randomly chosen valid time in the forecast data set. The colored dots mark the center of mass of WCB-outflow objects, and numbers indicate the ID of the individual object. The masks with the IDs 2 and 6 fulfill the size criterion ($\geq 300.000 \text{ km}^2$), all other masks are too small and therefore omitted. The green rectangle visualizes the region centered on the outflow object with the ID 2 which is extracted for the composite calculation (60° longitude, 40° latitude).

²Only such objects that are detected between forecast lead time 72 and 168 hours are considered to ensure equal data availability for all time lags.

composited without any further steps. The meteorological fields as well as the RMSE, however, vary spatially (and the RMSE also with forecast lead time), which requires a normalization when composite members with different coordinates and lead times are compared. For the meteorological fields, this is done by subtracting the model climatology (seasonal (DJF) climatology over three years averaged for all lead times from 0-240 h) at the corresponding grid points from the full field, which yields an anomaly. For the RMSE, the instantaneous RMSE is divided by the corresponding model climatology (determined for the grid point and lead time) of the RMSE. This results in a range between 0 and ∞ , where values larger than 1 correspond to anomalously large errors, and values below 1 denote errors smaller than usual. This method follows the one from Aiyyer (2015) who performed a similar analysis with extratropical transitions of tropical cyclones (ETs). The main difference to their approach is that Aiyyer (2015) detect ETs in (re-) analysis data sets and compute the associated ensemble standard deviation, whereas we detect WCBs in the forecasts and investigate the associated error in every ensemble member. Please note that this approach is not used until Section 6.4; in the analysis steps before, all variables (WCB-frequencies, Z500 and WS250 fields and error fields) are used without spatial centering or normalization.

6.2. Spatial co-occurrence of WCBs and forecast error

As a first step towards the question if WCBs systematically influence the forecast performance, we analyze the average spatial patterns of forecast errors and WCB-occurrence (see Figure 6.3). The 3-year DJF model climatology of Z500 (panel a) is characterized by a stationary wave with a trough over the western North Atlantic and a ridge over the eastern North Atlantic and over Europe. Especially the central to eastern North Atlantic is associated with large variability, as displayed by the black contours. An elongated band of large climatological forecast error is apparent across the northern North Atlantic, ranging from the North American east coast towards the west coast of Scandinavia (panel c). This pattern has similarities to climatological occurrence frequencies of extratropical cyclones (Sprenger et al., 2017), often also referred to as the storm track region. The largest errors occur in the eastern North Atlantic in the region of climatological enhanced geopotential heights and of large Z500 variability, south of Iceland. WCB ascent (green contours in panel c) occurs mainly from the western to the central North Atlantic, where frequencies exceed the 5% level. The region of enhanced climatological Z500 errors is co-located with the north-western edge of the WCB ascent region, even though a large part of the WCB ascents occurs more southerly, where the Z500 errors are relatively low. The maximum WCB-outflow occurrence (blue contours) is located over the central North Atlantic and extends north-eastwards into the region where the Z500-errors are largest.

The 250 hPa wind speed climatology in the North Atlantic domain is characterized by a distinct maximum (jet) ranging from eastern North America into the central North Atlantic (panel b). Another max-

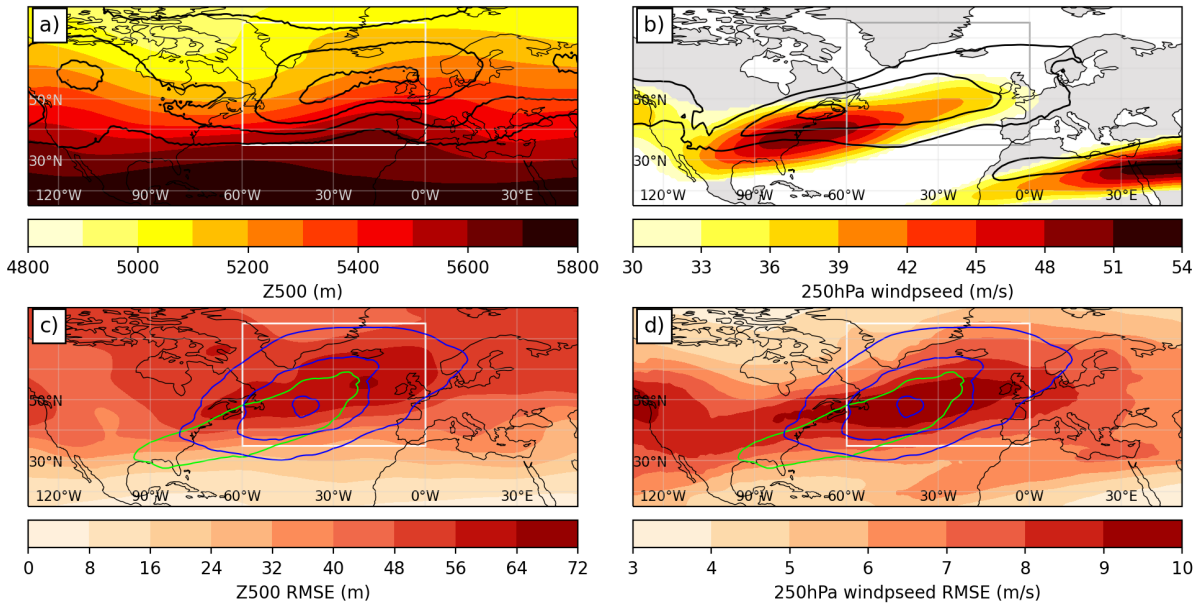


Figure 6.3.: Model climatology derived from forecasts initialized in winter (DJF) 2018/2019, 2019/2020 and 2020/2021, averaged over lead times from 0-240 hours and over 50 perturbed ensemble members of (a) Z500 mean (shading) and standard deviation (contours from 100 to 200 m in 25 m-intervals), (c) Z500 mean RMSE (shading) and WCB ascent (green contours) and outflow (blue contours) frequencies, (b) wind speed at model level 52 (approx. 250 hPa) mean (shading) and standard deviation (contour lines at 15, 17.5, 20 and 22.5 m/s) and (d) WS250 mean RMSE (shading) and WCB ascent and outflow frequencies (contours). The contours in (c) and (d) are 5, 10 and 15%. The white box shows the domain which is used for spatial averages in subsequent analyses (60°W-0°E, 35°-75°N).

imum appears in the south-east of the plotted domain, which corresponds to the subtropical jet. The largest climatological errors of WS250 are shifted eastwards with respect to the maximum wind speeds and occur over the central North Atlantic and reach into western Europe (panel d). In that region, the standard deviation of WS250 (contours in panel b) is relative large, which indicates that the model has deficiencies in forecasting the correct WS250 variability. The region of largest WS250 errors is co-located with the maximum WCB outflow frequencies, that reach values above 15% in the core region of WS250 errors.

This analysis shows that forecast errors and WCBs appear on average in similar regions, especially when considering the upper-levels (WS250) and the outflow stage of WCBs in the North Atlantic. In a next step, the temporal relationship between forecast performance and the occurrence of WCBs will be investigated.

6.3. Temporal relations of WCBs and forecast error

Forecast error during high and low WCB activity

The area-mean evolution of Z500 RMSE, averaged for the forecasts with the 20% highest (red) and 20% lowest (blue) WCB activity in the North Atlantic domain, are shown in Figure 6.4. Both the WCB-activity and the Z500 RMSE are evaluated in the same region ($60^{\circ}\text{W}-0^{\circ}\text{E}$, $35^{\circ}\text{N}-75^{\circ}\text{N}$; see white box in Figures 6.3c and d). The WCB-activity is computed as the area-mean value of the gridded WCB-masks in the outflow stage. In panel (a), all lead times are considered for the classification of the forecasts into the "high activity" and "low activity" groups. The general evolution of the RMSE with forecast lead time is characterized by a monotonic increase, reflecting the inevitable growth of errors during the course of the forecast. The differences of the two sub-groups clearly indicate that forecasts with high WCB activity have generally less skill than forecasts with low WCB activity (e.g. RMSE of 80 m with high WCB activity and 71 m with low WCB activity on day 7), which is in line with Wandel et al. (2021).

The subsequent panels of Figure 6.4 also show the average Z500 error for forecasts with high and low WCB activity, but with the difference that the classification is based on the WCB-activity in a specified time interval of the forecasts. In panel (b), the WCB-activity is evaluated during the first two days of the forecasts. In that case, the forecast error between the two groups does not differ (except for lead times around 8 days, which cannot be explained here). When the evaluation window is shifted to forecast days 2-4 (panel c), the two sub-groups show a very similar error growth behaviour until the evaluation period starts (at lead times of 48 hours). Subsequently, the forecasts with high WCB activity are characterized by slightly increased values of RMSE until about day 7, after which the differences between the classes vanish. Moving the 2-day evaluation window even further results in an equivalent behaviour: when reaching the start of the evaluation period, the errors in the category with high WCB-activity increase relative to the low-activity forecasts. Before that time, the lines are hardly distinguishable. The differences between the groups are largest when the WCB-activity is evaluated in the intervals 96-144 hours (d) and 144-192 hours (e). For 192-240 hours (f), the differences are again smaller than in the two previous windows. After the evaluation time, the RMSE of the low-WCB-activity forecasts slowly catches up to the forecasts with high WCB activity (all panels except a and f).

This analysis suggests that the occurrence of WCBs on average dilutes the forecast performance. This impact is weaker when the WCB-occurrence is in an early stage of a forecast, indicating that a previously existing error source is favorable for the subsequent error growth due to WCBs. The largest impact occurs between day 4 and 8, where the average slope of the RMSE-curve is the steepest. At that forecast stage, it is also likely that forecast errors have already emerged which can be propagated or amplified by WCBs. At later lead times (between day 8 and 10), the slope starts to level off, as the forecast errors start to saturate, and the impact of the WCBs decreases. We therefore hypothesize that the impact of WCBs

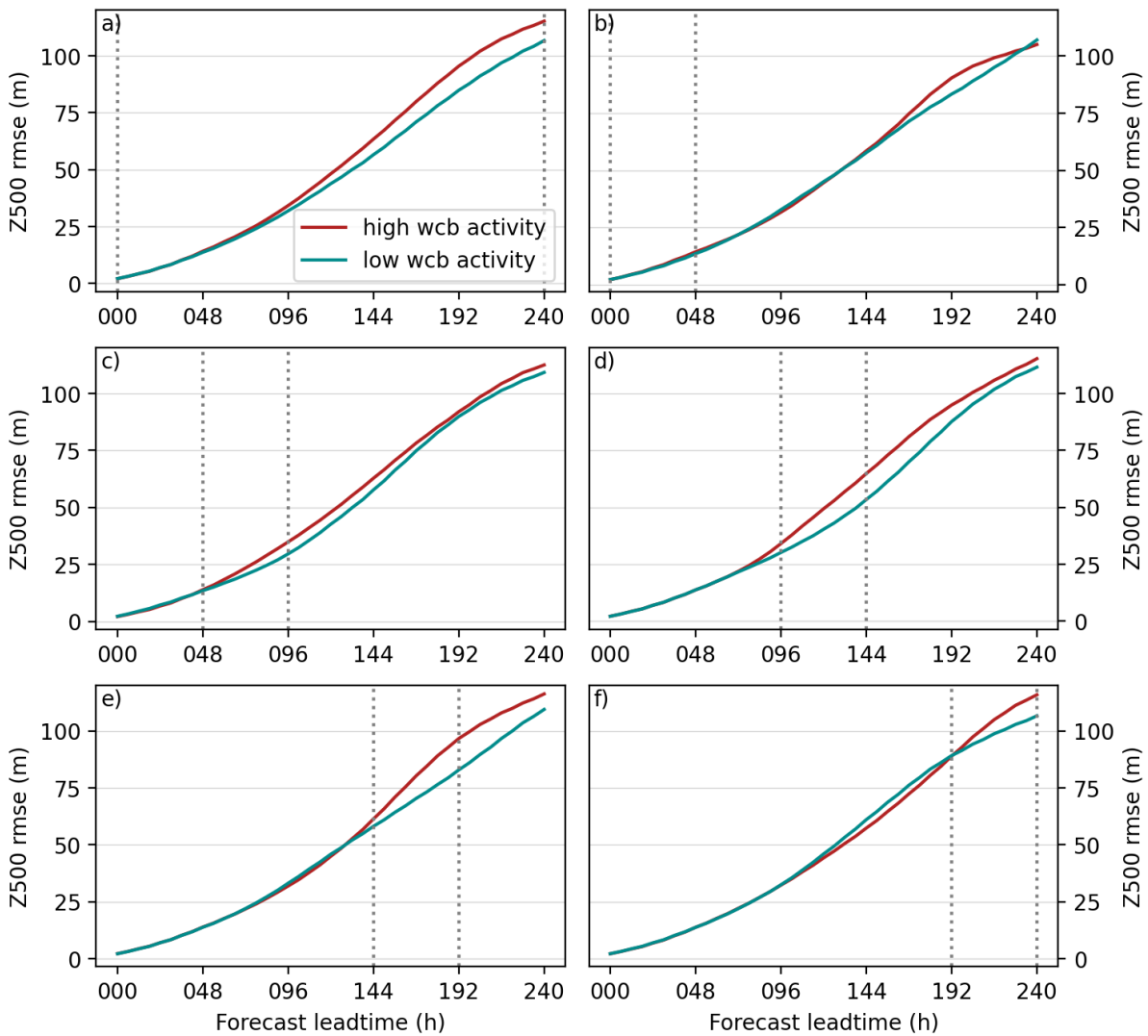


Figure 6.4.: Mean evolution of root-mean squared error (RMSE) of Z500 with forecast lead time in the North Atlantic domain (60°W - 0°E , 35°N - 75°N) for subsets of forecasts with the 20% highest (red) and lowest (blue) WCB outflow activity at all forecast lead times (a) and between lead times 0-48 hours (b), 48-96 hours (c), 96-144 hours (d), 144-192 hours (e) and 192-240 hours (f). The dotted gray lines mark the evaluation period of the WCB activity.

at a later stage of the forecast (beyond day 10) gets even weaker. This analysis can unfortunately not be performed, because the data archive only includes WCB data until day 10.

Focus on maximum error growth

Even though it was shown that anomalous WCB-activity co-occurs with reduced forecast skill (i.e. flow-dependent predictability), this does not imply that the forecasts are degraded by the WCBs, or that the error growth is induced by processes within the WCB. To further elaborate on a potential causal relationship, we now focus on the time in the forecast when the error grows the fastest. In doing so, the slope of

the area-averaged Z500-RMSE-evolutions of all individual forecasts is computed, and RMSE as well as the WCB-activity within the same region are lagged on that lead time when the RMSE growth is largest. Only lead times up to 192h lead time are considered to ensure the availability of data at time lags up to 2 days after the maximum error growth occurred. Note that not only bad or good forecasts, but all forecasts are considered.

Typically, the maximum error growth of Z500 in the North Atlantic occurs between day 5 and 7 of the forecast, with the median of the distribution at a lead time of 150 hours (Figure 6.5b). However, some forecasts experience their strongest degradation at earlier times: 5% of the forecasts have the strongest forecast error growth before day 4.

The mean evolution of the area-averaged RMSE is characterized by a constant, but slow increase 72 to 12 hours before the time of maximum error growth (blue line in Figure 6.5a). By definition, the RMSE grows rapidly around time lag 0, while its slope flattens out after time lags larger than 12-18 hours. The reason for this saturation is that these time lags mainly correspond to lead times larger than 8 days, during which the upper-level Rossby wave patterns in the forecast and analysis start to become out of phase and errors begin to saturate (e.g. Baumgart et al. 2019).

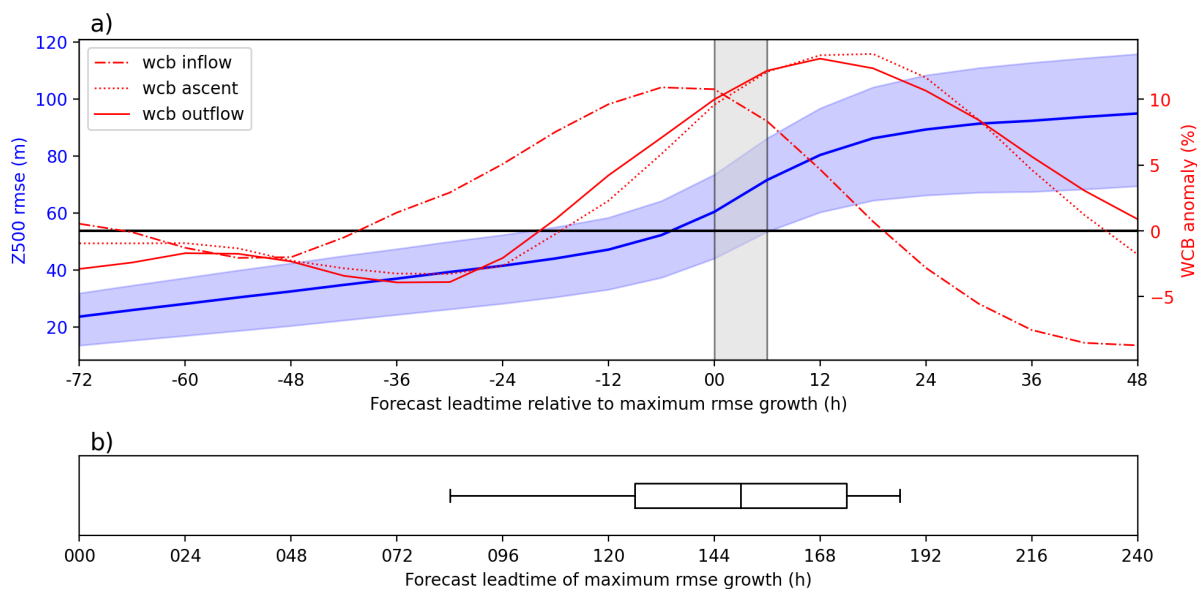


Figure 6.5.: (a) Mean (blue line) and interquartile range (blue shading) of the evolution of the domain-integrated Z500 RMSE and the domain-integrated relative WCB-anomaly (dash-dotted: inflow; dotted: ascent; solid: outflow) lagged on the lead time of maximum error growth in the North Atlantic region ($60^{\circ}\text{W}-0^{\circ}\text{E}$, $35^{\circ}\text{N}-75^{\circ}\text{N}$). The gray bar highlights the section with the largest error growth between two time steps. (b) Distribution (box indicates the interquartile range, whiskers the 5-95 interquartile range) of forecast lead times when the maximum error growth occurs.

6. The role of WCBs for forecast error growth

Around the time of maximum error growth, the area-averaged WCB-activity is systematically increased (red lines in Figure 6.5): On average, the WCB inflow anomaly (dashed line) reaches its maximum of 10% 12 hours prior to the strongest increase of forecast error, but starts to be above the climatological mean already about 1-2 days before. 18-24 hours after the maximum RMSE growth rate, the WCB inflow occurrence drops below climatological occurrence frequencies. WCB ascent and outflow anomalies almost evolve simultaneously and show a similar pattern as the inflow, but shifted towards later time lags: values start to be higher than the model climatology 18 hours before and reach their maximum of about 12% on average 6-18 hours after the forecasts experience their strongest degradation.

The synoptic evolution around the time of maximum error growth is depicted in Figure 6.6. Originating from a situation with a slightly amplified flow configuration 2 days prior to the largest increase of

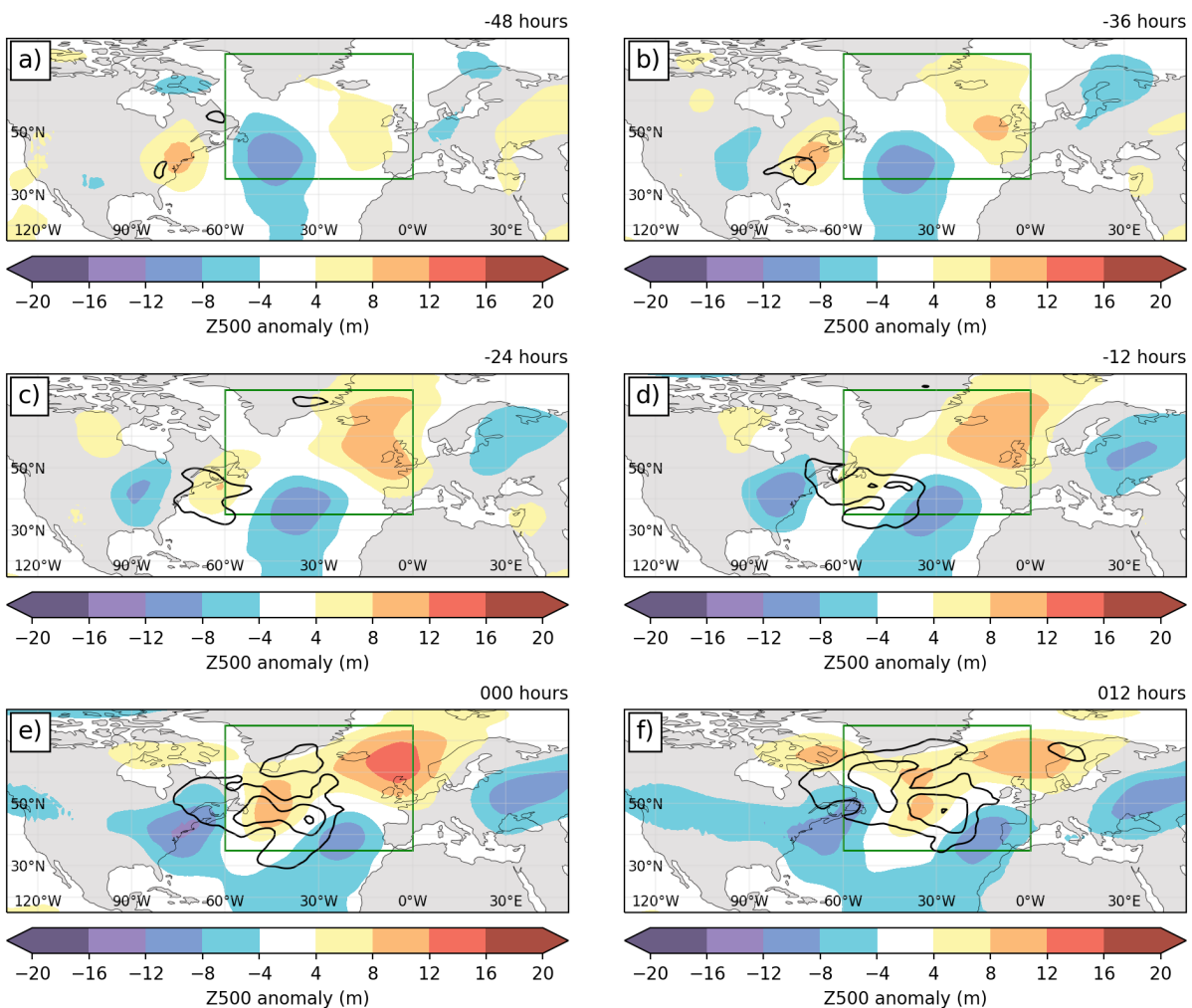


Figure 6.6.: Composites of Z500 anomalies (shading) and absolute WCB outflow frequency anomalies (contours of 1, 2 and 3%) with time lags of (a) -48 hours, (b) -36 hours, (c) -24 hours, (d) -12 hours, (e) 0 hours and (f) 12 hours to the maximum error growth in the North Atlantic region (60°W-0°E, 35°N-75°N) outlined by the green box.

Z500 RMSE (panel a), the Rossby wave pattern further amplifies (panels b-d) and reaches its maximum stage on the day of maximum error growth (panel e), with a wave pattern extending from the North American east coast towards northern Europe. Within the ridge over the western/central North Atlantic, higher-than-usual WCB-outflow occurrences are apparent (2-3%, absolute anomalies), which shows that the build-up of the ridge is linked to diabatic outflow in that region. Overall, the magnitudes of both the Z500- and the WCB-anomalies are rather small. However, it has to be taken into account that no classification into subgroups has been done (i.e. good vs bad forecasts) and all different kinds of weather situations are considered for the composites. Despite this circumstance, the analysis shows that forecast skill deteriorations over the North Atlantic on average coincide with enhanced WCB-activity and upper-level ridge building.

A sharper picture emerges when the same analysis is done with forecasts classified into a "good" and a "bad" group. This classification is done equivalently to the previous categorization into forecasts with high and low WCB activity (see Figure 6.4): the 20% forecasts with the largest spatio-temporal RMSE are labelled as "bad", while the 20% forecasts with the lowest RMSE are classified as "good". The evolution of the RMSE lagged on the maximum error growth for these two categories shows that the slope of the bad forecasts is much steeper than the ones of the good forecasts, resulting in much larger RMSE values at the end of the time series (see dashed and dotted lines in Figure 6.8). The WCB activity is also fairly different in the two subsets: the bad forecasts are characterized by a high level of WCB outflow frequencies, already 3 days prior to the maximum error growth. At time lag -24h, the WCB anomalies increase to even larger values and reach their maximum of more than 20% around 12 hours after the strongest error growth occurred. Subsequently, the values decrease and approach climatological values. The forecasts with low RMSE values, in contrast, have a generally much lower level of WCB-activity. At time lags -72 to -24 hours, the area-averaged outflow anomalies are just above -20%. When approaching

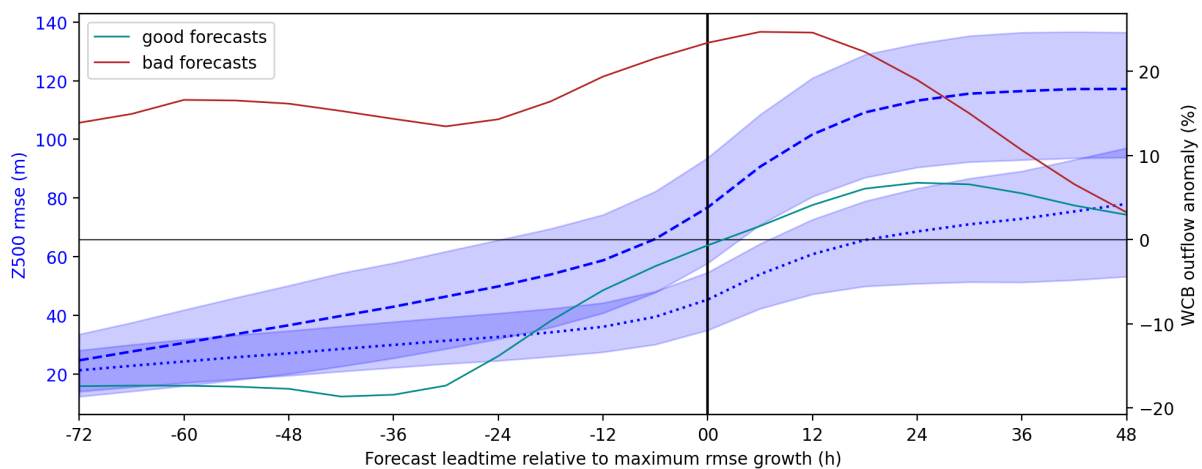


Figure 6.7.: As Figure 6.5, but for the 20% best (blue dotted line: RMSE, green solid line: WCB activity) and 20% worst forecasts (blue dashed line: RMSE, red solid line: WCB activity) and for WCB outflow only.

6. The role of WCBs for forecast error growth

the time of the strongest error growth, however, the WCB-frequencies also increase in the group of good forecasts and reach their maximum of about 5% at time lag 24 hours. Hence, not only bad forecasts worsen in the presence of WCBs, but also the good ones.

The synoptic situations of the two sub-groups differ substantially (Figure 6.8): the good forecasts (panels a-c) are characterized by a strong cyclonic anomaly over the whole North Atlantic, which splits up into an eastern and western part during the maximum error growth. Along the leading edge of the trough in the western North Atlantic, WCB outflow is directed into the central North Atlantic, where it reduces the cyclonic anomaly. In contrast, the bad forecasts are dominated by a distinct anticyclonic anomaly that is located over the north-eastern North Atlantic (panels d-e). This persistent anomaly is constantly fed by overly frequent WCB outflow, also prior to time lag -48h (not shown). Moving closer to the maximum error growth, the anticyclonic anomaly intensifies and expands towards the central North

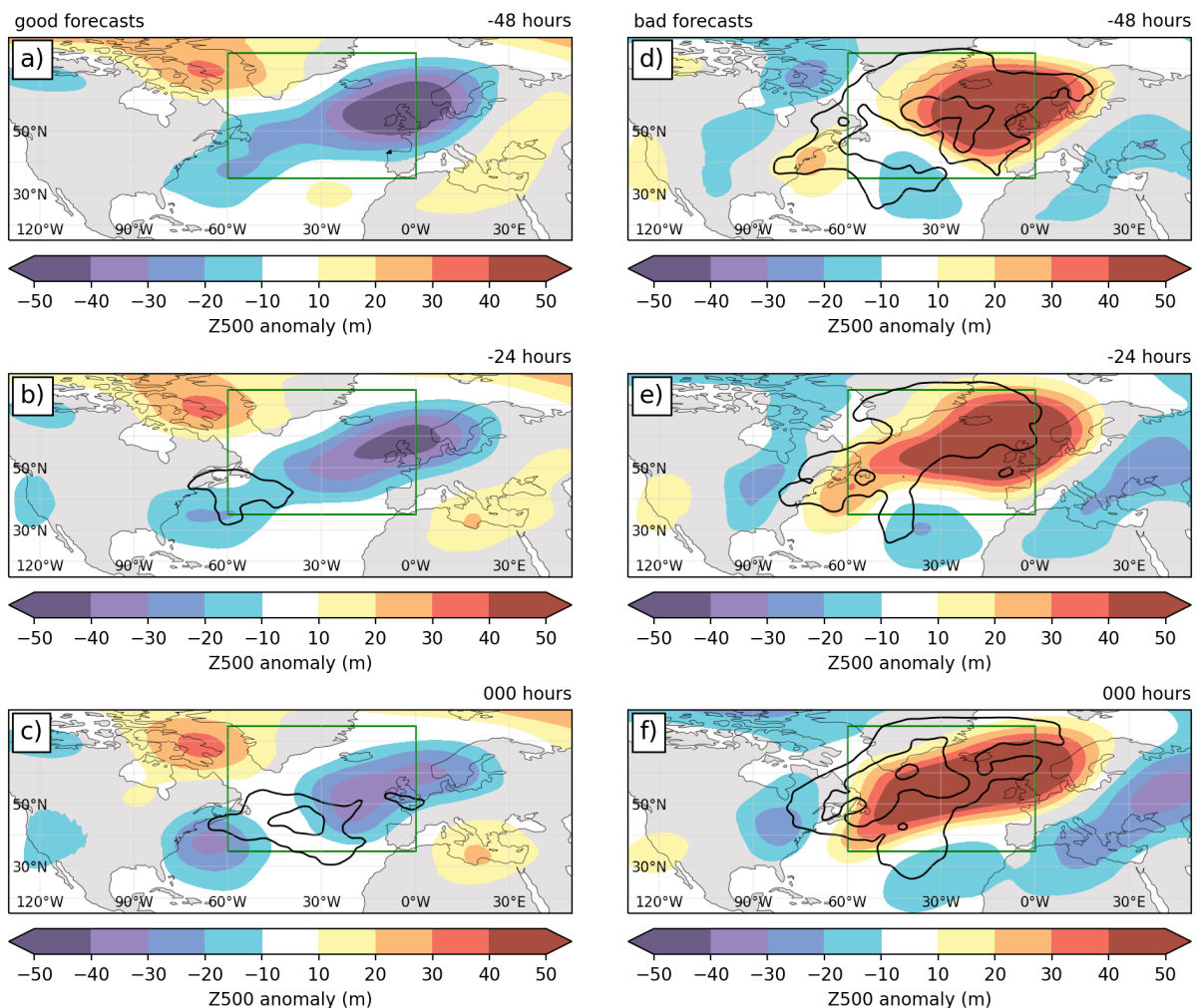


Figure 6.8.: As Figure 6.6, but separated for the 20% best (a-c) and the 20% worst forecasts (d-f) and only for time lags -48 hours (a,d), -24 hours (b,e) and 00 hours (c,f). WCB outflow anomalies are shown in contours from 2.5 to 10% in 2.5% intervals.

Atlantic, as a small positive Z500-anomaly from upstream merges into the main anomaly. This process is accompanied by an even stronger WCB outflow activity, especially in the western North Atlantic.

A very similar pattern is evident when the forecast error is not evaluated in the North Atlantic domain, but further downstream over Europe. Prior to the maximum error growth rate, the WCB-activity is systematically increased in the upstream region. Compared to the in-situ perspective, the maximum WCB-occurrence in the upstream box occurs one day earlier, which reflects the propagation of the signal with the background winds (see Figure D.1 in Appendix D). The large-scale flow configuration around the time of maximum forecast degradation over Europe resembles the one from the North Atlantic region, as it is characterized by a positive geopotential anomaly and by anomalously high diabatic outflow frequencies over Europe (see Figure D.2 in Appendix D). Also the classification into "good" and "bad" forecasts over Europe results in similar patterns as the composites around the maximum error growth over the North Atlantic (see Figure D.3 in Appendix D).

With this analysis, it was shown that the occurrence of WCBs is systematically increased around the time of the strongest reduction of forecast skill. This is particularly the case for forecasts of poor quality, where WCB-frequencies are generally increased and the large-scale circulation is characterized by anticyclonic flow anomalies. These results are in line with the findings of Rodwell et al. (2013), who found that situations of low predictability over Europe are often associated with atmospheric blockings, and stated that diabatic processes might play an important role for the degradation of the forecasts.

6.4. Perspective on WCB objects

So far, the analysis focused on composites anchored on the time of maximum forecast error growth in a large, predefined domain, which makes it difficult to establish a causal relationship between WCBs and forecast error growth, as also other processes that occur simultaneously could be involved.. To further elaborate on a possible causality, we adopt one further perspective that focuses on WCB objects in the forecasts by computing WCB-centered composites of meteorological variables and error metrics (see subsection 6.1 for a detailed description of the methodology).

Meteorological composites

Composites of anomalies of Z500 (shading), MSLP (black contours), and WS250 (purple contours) centered on WCB outflow objects for time lags 3 days prior to until 3 days after the detection of WCB outflow objects are shown in Figure 6.9. For clarity, no WCB-frequencies are shown in this Figure. It is, however, important to note that WCB-frequencies (especially of the outflow phase) appear already before the actual time of detection (i.e. lag 0h, panel d). This can be seen in the subsequent Figures (Figures 6.10 and 6.11). The composites are averages over the forecast lead times 72-168 hours centered

6. The role of WCBs for forecast error growth

on lag 0 hours, which ensures a full data coverage at all time lags, and each panel is a temporal average over 4 time steps with 6-hourly increments (e.g. -3 - -2 days lag is the average over 72-54 hours prior to the outflow event).

At lag day three to two prior to the evaluation time (a), the Z500-field is very similar to climatology, with a slight tendency of a ridge in the south-eastern quadrant and a trough in the north-western quadrant, both accompanied by a corresponding MSLP anomaly. Two to one days prior to the detected outflow event (b), the upstream trough intensifies, and a weak signal of enhanced wind speed emerges upstream of the composite center. On lag days -1 to 1 (c and d), both the downstream ridge and the upstream trough as well as the MSLP anomalies further intensify, and the WS250 anomaly located north of the composite center is maximized on the day of the WCB-detection. One to two days after the event (e), the downstream ridge and the wind speed anomaly are still pronounced, and the upstream signal is only present as negative MSLP anomaly, but vanishes in the mid-troposphere; the positive wind speed and Z500 anomalies are still apparent one day later (f). The synoptic sequence of ridge building and amplification and

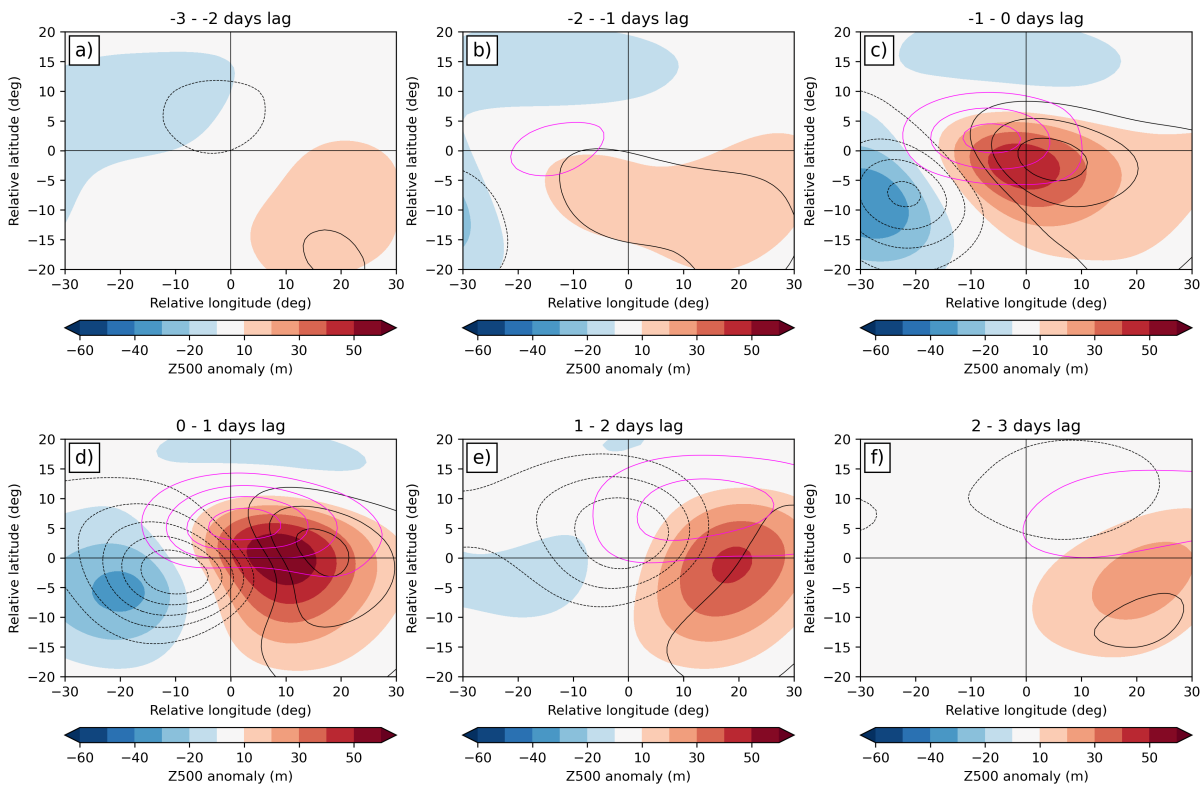


Figure 6.9.: Composite-means ($n=541.033$) of anomalies of Z500 (shading), MSLP (black contours from -10 to 10 hPa in 1 hPa-steps) and WS250 (pink contours at 3, 6, 9 and 12 m/s) centered on objects of WCB outflow. Anomalies are differences of the instantaneous field and the model climatology at the corresponding grid points. The spatial averages are means over the forecast lead times 72-168 hours (17 forecast lead times). Shown are temporal means over (a) 3-2 days (72-54 hours) before, (b) 2-1 days (48-30 hours) before, (c) 1-0 days (24-6 hours) before, (d) 0-1 days (0-18 hours) after, (e) 1-2 days (24-42 hours) after, and (f) 2-3 days (48-66 hours) after the WCB-outflow event.

the formation of a jet streak associated with WCB outflow events is well understood (see e.g. Grams and Archambault 2016 for reference) and matches the previously shown composites of Z500 around the time of maximum forecast error growth (see Figure 6.6). A similar sequence is apparent when the composites are centered on the ascent phase of WCBs; the pattern is, however, shifted north-eastwards, as the WCB ascent typically occurs slightly upstream of the outflow (see Figure D.4 in Appendix D).

Error composites

We here show similar composites, but now for the normalized RMSE in the mid-troposphere (Z500) and in the upper troposphere (WS250). As the ascent phase of WCBs occurs in the mid-troposphere, the Z500-errors are centered and lagged on WCB ascent objects (Figure 6.10). For further guidance, the WCB-inflow (orange contours), ascent (green contours) and outflow (blue contours) frequencies relative to the ascent event are plotted. 3-2 days prior to a WCB ascent event, the RMSE in the region is similar to the climatological forecast error (panel a). 2-1 days before the event (panel b), increased errors (15-20% larger than climatological errors) emerge in the south-western quadrant of the composite, and some WCB outflow takes place in the composite center. On the two days around the WCB-ascent event (panels c and d), the error pattern further amplifies (> 30%) and propagates eastwards; simultaneously, WCB-ascent and outflow reach their highest frequencies, with the ascent located slightly upstream of the outflow. 1-3 days after the WCB-ascent occurred (panels e and f), the errors are still present and are advected north-eastwards into the region of WCB-outflow.

The increased errors upstream of the ascent are co-located with the approaching trough in the Z500-field (see Figure D.4 in Appendix D), which indicates that errors are already present before the WCB-event. Nevertheless, on the two days around the WCB-ascent, large errors are also associated with the maximum occurrence frequencies of ascents, that might be amplified by the diabatic processes in the ascending motion. The downstream errors in the region of the WCB-outflow are clearly associated with the developing ridge in the north-eastern quadrant. Interestingly, this region is not associated with above-normal errors before the WCB-event occurs, which points towards the propagation and amplification of forecast errors from the trough into the ridge by WCBs.

Despite the clear error patterns related to the upstream trough and the (downstream) ridge, the variability among the composite members is very large: an investigation of the individual composite members with a k-means clustering algorithm (Hartigan and Wong, 1979) shows that about 50% of the cases are characterized by errors that correspond to or lie below the climatological mean (see Figure D.5 row c in Appendix D). The other half of the composite members is subject to substantial case-to-case variability. This reflects that the structure of the mid-tropospheric error patterns in the vicinity of WCBs is rather complex and cannot be entirely attributed to the WCB itself, but is also affected by the Rossby wave

6. The role of WCBs for forecast error growth

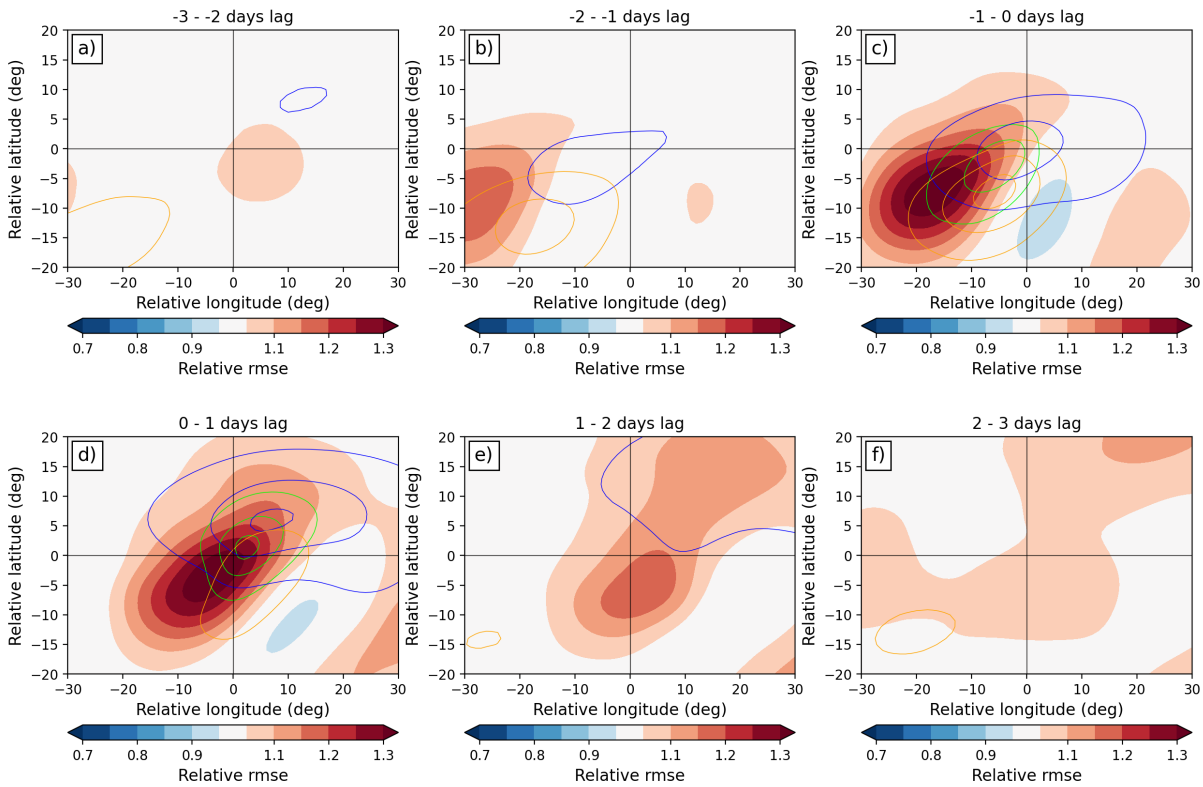


Figure 6.10.: As Figure 6.9, but centered on objects of WCB-ascents ($n=316,769$) and for the RMSE of Z500 normalized by the climatological RMSE (shading), and for WCB-inflow (orange contours), ascent (green contours) and outflow (blue colors). Contours show WCB-frequencies of 12.5, 25 and 37.5%.

dynamics.

Next, composites of the relative RMSE of 250 hPa wind speeds centered on WCB outflow events are analyzed (Figure 6.11). In contrast to the mid-tropospheric errors, no upstream errors are apparent for the upper-level wind speed. Errors first become visible about 2 days before the WCB-outflow occurs west of the composite center (panel b), propagate eastwards and intensify to maximum values of more than 30% on lag day 0-1 (panel d). The elliptical shape of the error pattern strongly resembles the shape of the jet streak that accompanies the WCB outflow and is located north of the main WCB-outflow and at the north-western edge of the developing ridge. After lag day 0-1 (panel d), also downstream errors emerge in the upper-level wind field. These errors could be related to anticyclonic Rossby wave breaking or to downstream development of a trough. In contrast to the mid-tropospheric errors, the individual composite members of WS250 errors are mostly characterized by spatially coherent error patterns, and the largest part of the cases is associated with above-climatological errors.

A qualitatively similar picture emerges when the variables for the composites centered on the WCB-stages are swapped (i.e. Z500-errors centered on WCB-outflow or WS250-errors centered on WCB-

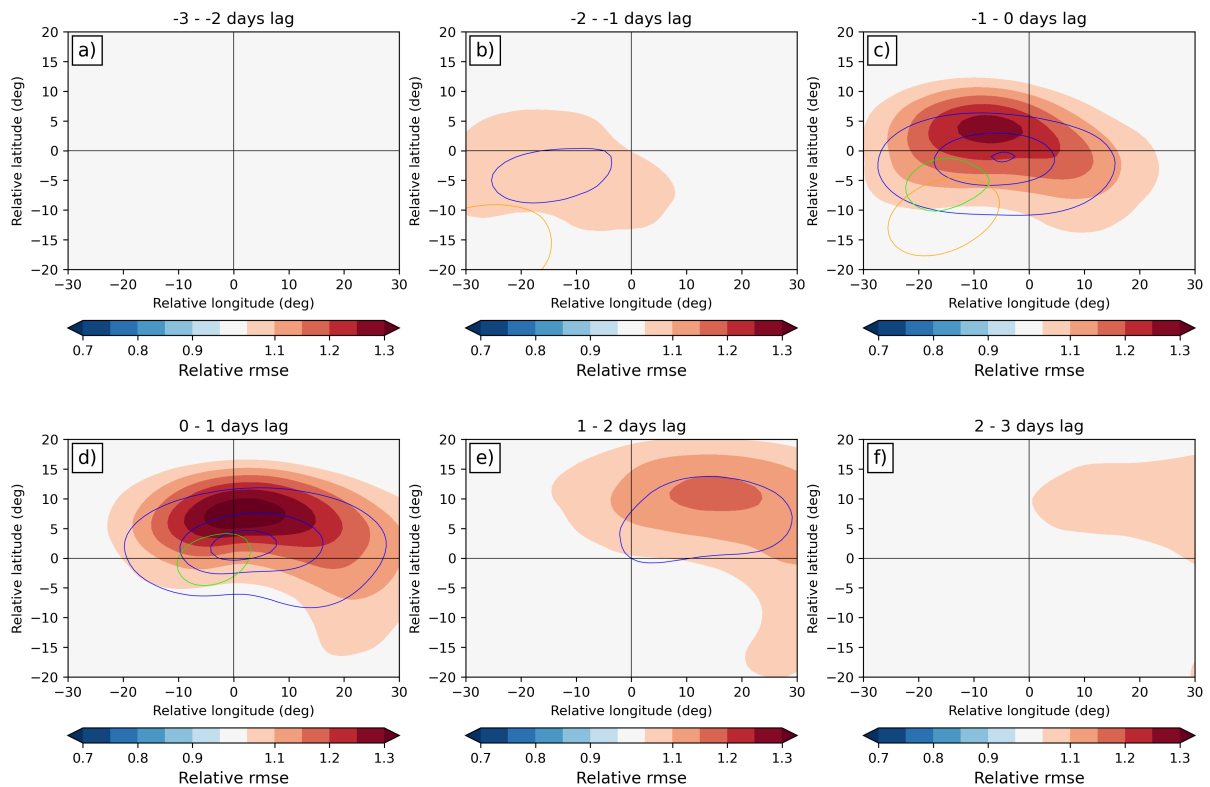


Figure 6.11.: As Figure 6.10, but centered on objects of WCB-outflow ($n=541.033$) and for the RMSE of WS250 normalized by the climatological RMSE (shading).

ascent), when the spatial displacement of the WCB-ascent relative to the outflow is considered: the spatial variability of mid-tropospheric errors persists even when centered on the outflow, and the error patterns of the upper-level wind speed remain spatially homogeneous when centered on the ascent (not shown). This consistent response of forecast errors across the two WCB-stages and variables suggests that the imprint of WCBs is more robust in the upper levels than in the mid-troposphere.

Lead-time dependency of error composites

The previously shown error composites centered on WCB-ascent and outflow objects (Figures 6.10 and 6.11) are averages over forecast lead times ranging from 3 to 7 days. However, it was shown in the beginning of the chapter that the error differences between forecasts with high and low WCB-activity depend on the time when the WCB-activity is evaluated (see Figure 6.4). To further investigate this lead-time dependency of WCB-related errors, the temporal evolution of the spatial averages of the Z500 and WS250 error composites centered on WCB-ascent (Figure 6.12a) and outflow (Figure 6.12b) are computed for different lead times when the WCB object is detected. While both variables follow a similar shape at all lead times, with maximum mean errors slightly after the WCB event, there are some differences between the mid-tropospheric and upper-level errors: the relative Z500-errors are on average smaller when

the event occurs early in the forecast (pale colors in panel a), increase for forecast days 3-5 and again decrease for late lead times (dark colors). This is substantially different from the signal observed for the upper-level wind speed errors (panel b), which are largest when the WCB-event occurs within the first 42 h after the initialization (pale colors), and subsequently decrease continuously (dark colors). Despite the larger peak of the signal for WS250, the domain-integrated errors remain at a higher level after the WCB-event in the mid-troposphere. Except for the earliest lead times, the normalized errors are higher after the WCB-event than before, which shows that the forecast skill is permanently reduced after a WCB event occurred. However, it is important to note that the evaluation domain does not move with the object, for which reason enhanced errors that are advected outside the domain are not considered here.

The larger magnitude of mid-tropospheric errors related to WCB-ascents that occur at later forecast stages compared to early WCBs is in line with the analysis of the skill of forecasts with high WCB-activity at different lead times (Figure 6.4) and indicates that WCBs amplify and propagate pre-existing errors. On the other hand, the opposite is the case in the upper troposphere, where errors are particularly large for WCB outflow at early forecast stages, when pre-existing errors are mainly of small magnitude and localized. This indicates that WCBs introduce errors during their ascent, but can also amplify small-scale errors and project them onto the large-scale flow. The different complexities of the lead-time dependencies between the two variables emphasize that the mid-tropospheric errors are potentially affected by interacting and superimposed processes, such as diabatic processes and baroclinic dynamics, whereas the the upper-level wind errors follow a rather simple pattern that can be fully attributed to the diabatic outflow of WCBs.

6.5. Concluding discussion

With the presented analysis, it was shown that errors in medium-range forecast of ECMWF's ensemble prediction system are associated spatially and temporally with the occurrence of WCBs. Different perspectives have been adopted to investigate the role of WCBs for forecast errors: at first, a simple stratification of forecasts by the WCB-activity was used to demonstrate that forecasts with high WCB-activity are characterized by reduced forecast skill compared to forecasts with low WCB-activity. The separation into "good" and "bad" forecasts further indicates that the large-scale flow configurations of the two forecast classes are on average very different from each other: in the North Atlantic, forecasts with low skill are characterized by a highly amplified Rossby wave pattern and anticyclonic flow anomalies, whereas good forecasts are featured by flow configurations with strong geopotential height gradients. This serves as an illustrative example for flow-dependent predictability, which is particularly pronounced in the North Atlantic region (e.g Ferranti et al. 2015, Büeler et al. 2021). Focusing on the time in the forecast where the error growth is largest demonstrates that the WCB-activity is systematically increased

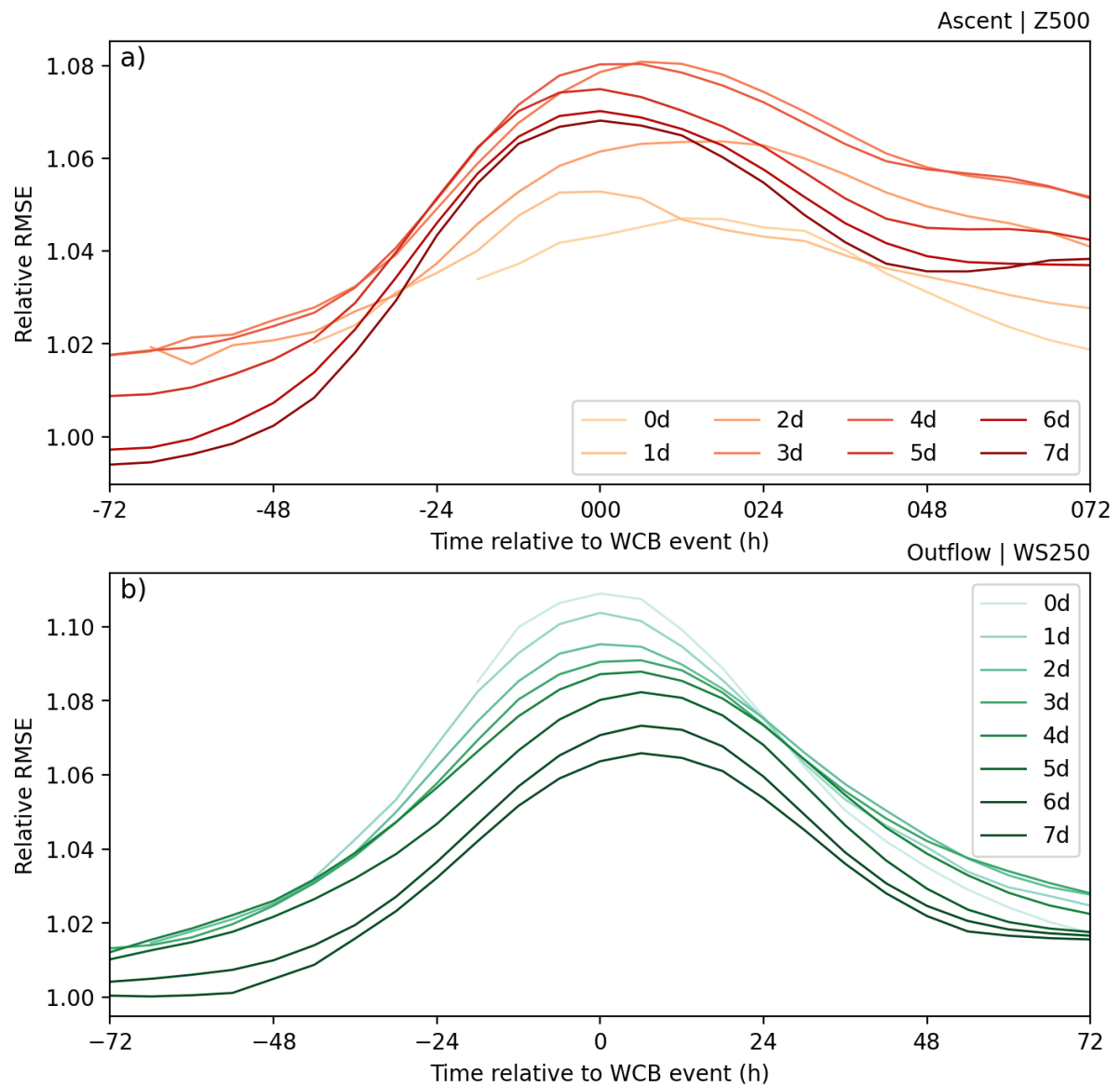


Figure 6.12.: Time-lagged evolution of domain-integrated composites of the relative RMSE of (a) Z500 centered on WCB-ascents (see Figure 6.10) and (b) WS250 centered on WCB-outflow (see Figure 6.11). Different color shades denote the forecast lead time at which the WCB-event occurs ("0d" corresponds to the average of 0, 6, 12 and 18 hours, "1d" to 24, 30, 36 and 42 hours, etc.).

around the time of strongest skill reduction.

Composites of normalized forecast errors centered spatially and lagged temporally on WCB ascent and outflow objects further substantiate the finding that WCBs are involved in the growth and amplification of errors: coherent patterns of WCB-occurrences and increased forecast errors that are associated with the representation of the Rossby wave structure, especially in the developing jet streak north of and in the ridge downstream of WCBs, suggest a direct relationship between WCBs and forecast errors. The error composites in the mid-troposphere are characterized by large case-to-case variability, whereas the

signal in the upper-tropospheric wind speed is robust across a large number of cases. One possible reason for these differences is that the diabatic outflow of WCBs has a more pronounced impact on the large-scale flow than the ascent. The outflow transports low-PV air into the upper troposphere, where it diverges and sharpens the PV-gradient across the tropopause. As a consequence, a jet streak forms and the waveguide is deflected northwards (Ahmadi-Givi et al., 2004; Grams and Archambault, 2016). Forecast errors that arrive in that region are quickly advected by the strong jet and project onto the Rossby wave pattern. In contrast, the WCB-ascent is more confined than the outflow and quickly passes through the mid-troposphere. Therefore, its impact on the large-scale flow is not as direct as the one of the outflow. Another aspect could be that the two variables Z500 and WS250 are not directly comparable, even though they are both commonly used to characterize the large-scale flow. Using the zonal (U) and meridional (V) components of the wind vector at 250 hPa, however, results in very similar patterns as the wind speed (not shown). As the geopotential heights and the wind components are directly linked to each other in the free troposphere (i.e. geostrophic balance, Holton (2004)), we hypothesize that the signal would be equivalent when using Z250. Nevertheless, a cleaner way would be the direct comparison of the RMSE of Z500 and Z250, which was not feasible in the context of this thesis due to data availability and computing performance issues.

The three-dimensional view on the forecast errors associated with WCBs nicely illustrates that errors are propagated both horizontally and vertically: in the mid-troposphere, errors that are spatially related to the upstream trough are on average present already before the WCB-event occurs. In the upper levels, in contrast, no pre-existing errors are apparent, but large errors emerge co-located with the WCB outflow in the region of the developing jet streak and at the north-western edge of the downstream ridge. Hence, WCBs act as a communicator between a region of enhanced errors and a region with originally low (or climatological) errors. The ascending motion associated with WCBs involves strong diabatic heating and cross-isentropic transport of air masses, which is reflected by substantially different levels of Θ in the different stages of WCBs, and hence also in the source and destination regions of WCBs. Under adiabatic conditions, such a material transport of mass (and errors) from lower to higher isentropic levels is not possible (Saffin et al., 2021). The diabatic processes involved in the WCB-dynamics are therefore crucial for the growth and amplification of forecast errors, even if their direct contribution to forecast error growth might be small.

Finally, it was shown that the error patterns associated with WCBs vary with lead time, and that this lead-time dependency is different for Z500-errors than for WS250-errors. While WCBs that happen early in forecasts exert the largest impact on the domain-integrated errors of WS250, the magnitude continuously decreases with progressing forecast lead times and are half as much as in the beginning of the forecast. Apart from the first 2 days, this also occurs for errors in the Z500-field and indicates that WCBs will have no immediate impact on forecast errors at lead times well beyond the analyzed forecast times (e.g.

at lead times in the extended range). At a stage where forecast error growth is dominated by barotropic Rossby wave dynamics (i.e. when the Rossby wave patterns of forecast and analysis are out of phase, Baumgart et al. 2019), WCBs will not result in an additional skill reduction. These results fit into the conceptual model of upscale error growth (Zhang et al., 2007), that describes a 3-stage sequence how small-scale initial errors are propagated across the scales. In the second stage, errors on the convective scale are projected onto the synoptic scale, and it is hypothesized that diabatic processes in convection and WCBs substantially contribute to this error growth (Grams et al., 2018; Selz, 2019; Selz et al., 2022). Our analysis shows that WCBs project and amplify errors to the large-scale flow, in particular in the early stages of the forecast, and thereby corroborate the findings from previous studies.

Outlook

Future work could generalize the findings from this thesis in multiple aspects: for example, the analysis could be extended to the North Pacific region and to other seasons (especially autumn could be interesting, when strongly heated WCBs appear in the northern hemispheric ocean basins). Further, the analysis could be performed in different forecast models, such as for the models in the TIGGE³ archive. Comparing the WCB-related error structures across a range of different forecast systems could yield valuable information on how uncertainties due to diabatic processes depend on different model formulations. Such an intercomparison project based on forecast archives, however, is not feasible with a Lagrangian approach, and requires different techniques to detect WCBs in the forecasts, such as the newly developed machine-learning based approach by Quinting and Grams (2022).

In our study, all WCB-objects have been considered in the same way. However, we found that the error patterns in the mid-troposphere are rather complex and feature a large variability between individual cases. It could therefore be worthwhile to quantify several aspects of WCB objects and subsequently determine the related error structures. Promising characteristics are the object size, the latent heating rate along the ascent serving as a proxy for the diabatic processes, the curvature of the outflow air mass (cyclonic vs anticyclonic), the geographical position, or the large-scale flow configuration. Such an analysis could advance the understanding of the error origins and which WCB-properties and/or flow configurations are particularly prone to error growth.

Finally, the question whether WCBs amplify pre-existing errors and project them onto different scales, or if they act as a source of errors, could not be fully resolved. Tackling this issue requires a sophisticated diagnostic setup, including error metrics that can be compared in different atmospheric levels. A promising approach could be to compute errors along WCB-trajectories, compare errors in the inflow and outflow region, and link the error evolution to processes that occur during the ascent of WCBs.

³THORPEX Interactive Grand Global Ensemble

7. Conclusions

This thesis provides a detailed investigation of warm conveyor belts in ensemble forecasts by adopting two perspectives on the subject matter: sensitivities of WCBs to the configuration of ensemble prediction systems and the role of WCBs for forecast errors. This chapter summarizes the main findings of the project by addressing the research questions posed in Chapter 2 and discusses them in a broader context.

Ensemble weather prediction attempts to estimate the future probability distribution of the atmospheric state by representing the governing sources of forecast uncertainty. Apart from uncertainties related to the erroneous estimation of the initial conditions, forecast errors also arise from deficiencies in the forecast model, especially from the parametrization of physical processes. Most schemes which represent model errors, for example the widely used SPPT-scheme, therefore introduce stochastic perturbations into the physical parametrizations, and the amplitude of the introduced perturbations scales with the local magnitude of the parametrization tendencies. WCBs and other rapidly ascending air streams, such as tropical convection, are characterized by large amounts of latent heat release due to cloud-condensational processes. Therefore, regions in which such weather systems occur are prone to perturbations from model uncertainty schemes. This co-occurrence motivated the systematic investigation of rapidly ascending, diabatically driven air streams in the light of ensemble configuration in general, but specifically in the context of model uncertainty representations. By conducting sensitivity experiments with the ECMWF ensemble prediction system and by exploiting an operational forecast archive, the following research questions were addressed in Chapter 4:

C4-1 Do the perturbation schemes that are in operational use at ECMWF affect the occurrence and characteristics of diabatically driven, rapidly ascending air streams?

ECMWF's operational model uncertainty scheme, the SPPT-scheme, systematically increases the occurrence frequencies of rapidly ascending air streams compared to experiments with an unperturbed model. The characteristics of the trajectories, such as the integrated latent heat release, however, are unaffected by the scheme. The magnitude of this effect, expressed as the ratio of rapidly ascending trajectories in perturbed and unperturbed simulations, strongly correlates with the diabatic heating rate along the trajectories, and is therefore larger in the (sub-) tropics than in the extratropics, and more pronounced in autumn and spring than in winter. A Eulerian diagnostic of the occurrence frequencies of vertical velocities reveals that the increased occurrence of rapid ascents is balanced by accelerated downward motions. Further, SPPT also accelerates air parcels that are at rest both up- and downwards, and thereby enhances the vertical air mass transport.

In contrast to SPPT, the initial condition perturbations do not systematically affect the trajectory counts and the distribution of vertical velocities. This information can be exploited, as it implies that the comparison of operational ensemble forecasts that are perturbed with both initial condition and model perturbations to unperturbed forecasts mimics the experimental design of expensive sensitivity simulations.

C4-2 Do other model uncertainty schemes result in similar effects as SPPT?

The SPP-scheme, which represents uncertainties in the individual parameters of physical parametrizations, results in very similar effects on the trajectories and the distribution of the vertical velocities as SPPT. Sensitivity experiments with perturbations only to the parameters in the convection parametrization reveal that changes to the vertical velocities are especially sensitive to perturbations to the convection scheme. Perturbations to the model dynamics through the STOCHDP-scheme, in contrast, exert a much less pronounced impact on the diagnostics related to ascending motions and only affect the slow spectrum of vertical velocities.

C4-3 How can the observed effects be explained?

Despite the symmetric and zero-mean design of the schemes, the perturbations result in a unilateral response of the vertical velocities. The detailed analysis of the trajectories shows that the effect on the frequency occurrence of rapidly ascending air streams is apparent from the first time step when the model uncertainty schemes are active, and remains constant throughout the course of the forecast integration. This indicates that the effect is not a result of changing environmental conditions (e.g. SPPT-induced biases in the low-level specific humidity that feeds into the ascending air streams), but acts directly and immediately on the ascending motions. Based on these findings, a mechanism is hypothesized that explains the observed unilateral response. The triggering of the rapid ascent of an air parcel can be described as a nonlinear process that is characterized by a threshold behaviour. In the phase space below a critical value, the system resides in a stable, densely populated state. When the threshold is exceeded, the system becomes unstable and is quickly brought to an equilibrium state. Symmetric, finite-amplitude perturbations across such a threshold result in an asymmetric response of the system, because the unstable phase space is less densely populated than the stable phase space. Applied to the concept of rapidly ascending air streams, positive perturbations are more effective in triggering ascent than negative perturbations in preventing ascending motion.

Model uncertainty schemes are mostly developed and tested in the framework of operational forecasting, where mainly established and standardized verification tools are employed (see e.g. Lang et al. 2021 for the testing suite of the SPP-scheme). Weather system oriented evaluations, as the one presented in this thesis, are, in contrast, often explorative, and their development requires a high level of effort. Such alternative approaches are therefore rarely adopted. By applying such an approach, however, we were

able to highlight a well-known, yet poorly understood aspect of stochastic parametrization: biased responses to symmetric, zero-mean forcing. The concept of WCBs, and generally of diabatically driven, rapidly ascending air streams, turned out to be an ideal test bed to investigate this aspect of stochastic perturbations, mainly because of two reasons: firstly, because WCBs have been shown to be sensitive to physical parametrizations, especially to the choice of the microphysics (e.g. Joos and Forbes 2016; Mazoyer et al. 2021) and convection scheme (e.g. Rivière et al. 2021), which makes them prone to perturbation schemes that target uncertainties in parametrizations. And secondly, because their ascent follows to some extent nonlinear dynamics, which is a prerequisite for a unilateral response to symmetric perturbations. Such effects have been observed previously (e.g. for tropical cyclones, see Stockdale et al. 2018 or Vidale et al. 2021), but their origin is either not discussed in detail, or is simply referred to as a consequence of the nonlinear dynamics of the atmosphere. The study presented in this thesis, however, gives a process-based explanation of this behaviour, and thereby contributes to the understanding of the impacts of stochastic perturbations on the mean state of the model.

These findings are of particular interest when considering the importance of diabatically driven, rapidly ascending air streams for several weather phenomena: for example, large amounts of precipitation are associated with tropical convection and tropical cyclones (Jiang and Zipser, 2010) and with WCBs (Pfahl et al., 2014), and WCBs play a major role in shaping the large-scale extratropical circulation through ridge building (e.g. Grams and Archambault 2016; Teubler and Riemer 2016). This motivated the investigation of the impact of stochastic model uncertainty schemes on the global distribution of precipitation and on the amplitude of the upper-level Rossby wave pattern in Chapter 5:

C5-1 Do stochastic model uncertainty schemes modify the global distribution of precipitation?

The distribution of precipitation is altered by stochastic model perturbations. From a global perspective, the precipitation sums are increased in experiments with SPPT compared to unperturbed simulations, and this effect is particularly pronounced in the tropical and subtropical regions, and decreases with increasing latitudes. Frequency distributions of precipitation sums reveal that the changes induced by SPPT are consistent with the modifications of vertical velocities described in the previous chapter. Further, the comparison between different model uncertainty schemes reveals that their characteristic impact on the vertical velocities is directly reflected in the precipitation distributions. This suggests that the unilateral effect of the perturbations on the vertical velocities is directly passed on precipitation.

The reported changes to precipitation are in line with previous studies (e.g. Subramanian et al. 2017; Strømme et al. 2019). The focus and the experimental setup of these studies is, however, rather different from the one in this thesis: while both Subramanian et al. (2017) and Strømme et al. (2019) used lower-resolution climate models to simulate long time periods and to evaluate systematic biases and precipitation variability, our approach aimed to link the altered precipitation distribution to modified

vertical velocities, and thereby to provide a process-based understanding of changes to the mean state of the model climate.

C5-2 Is the amplitude of the upper-level Rossby wave pattern affected by stochastic model uncertainty schemes?

Compared to experiments with an unperturbed forecast model, simulations with SPPT have systematically increased areas of both upper-level ridges and troughs, resulting in a reduction of systematic biases of the waviness of the upper-level wave pattern. Additional to this increased Rossby wave amplitude, also the entire Rossby wave pattern is shifted polewards with SPPT. Even though the magnitude of these changes is rather small, it qualitatively reflects the increased mass flux of lower-tropospheric air into the upper troposphere by the modified activity of WCBs. The analysis of the Rossby wave amplitude in experiments with other model uncertainty schemes and in forecast archives, where perturbed and unperturbed forecasts are compared against each other, supports the hypothesis that the changes to the Rossby wave amplitude are directly linked to the increased WCB-frequencies.

Changes to the large-scale extratropical circulation through stochastic parametrizations have been documented before: for example Dawson and Palmer (2015) report an improved representation of weather regimes in climate simulations with stochastic physics, and Martínez-Alvarado et al. (2018) found changes to the upper-level PV-gradient at the tropopause in perturbed forecasts. The signals in the midlatitudes are, however, mostly of small magnitude and rather subtle, which makes an interpretation challenging. The authors of the aforementioned studies argue that the changes are either introduced by noise-induced changes to the mean state by multiplicative forcing (see Berner et al. 2017), or by a direct impact of the perturbations to the large-scale flow. Our results contribute to this discussion and provide a new perspective on how the large-scale circulation can be altered through stochastic forcing. We introduce a coherent, process-based mechanism that explains how the perturbations result in the observed behaviour. Similarly to the analysis of precipitation, the linkage of the impact of stochastic perturbations on the WCB-frequency and on the Rossby wave amplitude enables to describe how the perturbations are projected to the large-scale flow. WCBs thereby act as a communicator between the perturbations and the Rossby wave pattern and propagate the unilateral response onto the midlatitude circulation.

The property of connecting different scales of atmospheric motion and different levels of the troposphere that is characteristic to WCBs makes them highly relevant for the predictability of the large-scale midlatitude circulation: case studies have shown that the sensitivities of WCBs to the environmental conditions can result in large forecast uncertainty in the case of preexisting forecast errors (upscale error growth, Grams et al. 2018). Ascent characteristics of WCBs and their impact on the large-scale flow have been shown to depend on the choice of the physics parametrizations in models (e.g. Joos and Forbes 2016,

Rivière et al. 2021), which further underlines the potential role of WCBs for forecast uncertainties. Therefore, several studies argued that systematic errors in the representation of the large-scale circulation could be linked to the upscale error growth mechanism and to the diabatic processes associated with WCBs (e.g. Gray et al. 2014; Martínez-Alvarado et al. 2016b; Sánchez et al. 2020). These findings are, however, mostly based on individual case studies, or employ Eulerian metrics to quantify the impact of WCBs on forecast errors. This motivated the systematic investigation of the role of WCBs for forecast errors based on a unique data set of trajectory-based imprints of WCBs in ensemble forecasts. In Chapter 6, the following questions were addressed:

C6-1 Do WCBs and forecast errors of the large-scale circulation co-occur?

The regions with the largest climatological forecast errors over the North Atlantic are co-located with the frequency maxima of WCB occurrence. This is especially pronounced for errors of the upper-level wind speed and for the outflow phase of WCBs. On average, the skill of forecasts with a high activity of WCBs is reduced compared to forecasts with low WCB activity.

C6-2 Can periods of increased forecast error growth be linked to WCB-activity?

Around the time in the forecast when the spatially averaged error growth in the North Atlantic domain is largest, the mean WCB-activity is systematically increased. Composites lagged on the maximum error growth indicate that the synoptic situation is on average characterized by an amplified Rossby wave pattern. The WCB-activity and the anticyclonic flow anomalies become especially pronounced when only forecasts with low skill are considered.

C6-3 Is WCB-activity related to coherent spatio-temporal error patterns?

Composite fields of forecast errors centered on a large number of WCB objects reveal that WCB ascent and outflow are on average both associated with larger-than-normal forecast errors. In the mid-troposphere, these errors are co-located with an upstream trough and with a downstream ridge, but the patterns are characterized by a large case-to-case variability. In the upper troposphere, WCB outflow is clearly linked to errors in the representation of the jet streak which forms on the poleward flank of the upper-level ridge, and the error pattern is very robust across a large number of cases. The analysis suggests that WCBs project and amplify forecast errors from lower levels to the upper-level flow.

These findings are in line with previous studies that dealt with the impact of diabatic processes for forecast uncertainty (e.g. Rodwell et al. 2013; Martínez-Alvarado et al. 2016b; Grams et al. 2018; Berman and Torn 2019; Sánchez et al. 2020) and highlight the role of WCBs as "predictability barriers". Further, the results emphasize that WCBs are involved in the conceptual upscale error growth mechanism by projecting small-scale errors to the large scale (Zhang et al., 2007; Selz and Craig, 2015). Due to the diabatic heating during their ascent, WCBs reach height levels that are not accessible for adiabatic motions (e.g. Saffin et al. 2021) and can thereby project forecast errors of smaller scale that originate from the lower-

or mid-levels of the troposphere onto the Rossby wave pattern. At the upper levels, the errors are propagated along the tropopause by the jet and grow via barotropic wave dynamics (e.g. Baumgart et al. 2019).

The process-oriented evaluation of model uncertainty schemes and of systematic forecast errors by using a Lagrangian detection of rapidly ascending air streams that has been adopted throughout the thesis is new to the community. This allowed to build a bridge between the disciplines of atmospheric dynamics and model development / evaluation. The findings from the thesis could contribute to the improvement of forecasting systems in the following ways:

The frequency occurrence of WCBs and other diabatically driven, rapidly ascending air streams is sensitive to model uncertainty schemes that introduce perturbations into the physical parametrizations. The nonlinear dynamics of the ascending air streams translate the stochastic, zero-mean forcing of the perturbations into a unilateral response which affects the mean state of the model climate. This is a valuable insight for model developers and can be taken into account for the development process and evaluation of future model uncertainty schemes.

The systematic degradation of forecast skill by WCBs suggests that efforts should be undertaken to improve the representation of WCBs and their effect on the large-scale circulation in numerical models. Promising directions for this challenging task could be the improvement of the observations in critical regions (e.g. low-level moisture in typical WCB inflow regions) to minimize the errors that can be amplified by WCBs, and improved representations of the diabatic processes during the ascent of WCBs (Rodwell et al., 2018a).

A. Appendix for Chapter 3

Table A.1.: List of parameters that are perturbed in the SPP scheme. The first column gives the parameter ID, the second explains the role of the parameter, the third denotes the type of the underlying distribution, and the fourth gives the standard deviation of the underlying distribution from which the parameter is sampled. The table is adapted from Lang et al. (2021).

	Parameter ID	Role of Parameter	Dist	Standard deviation
Turbulent Diffusion and subgrid orography	CFM_OC	Transfer coefficient for momentum over ocean	LN	0.26 / 0.33
	CFM_LA	Transfer coefficient for momentum over land	LN	0.65 / 0.78
	RKAP	von Kármán constant	LN	0.065 / 0.26
	ROFDC	Coefficient in turbulent orographic form drag scheme	LN	0.78
	HSDT	Std. Dev. of subgrid orography	LN	0.52
	VDEXC_LEN	Length-scale for vertical mixing in stable boundary layer	LN	1.04
Convection	ENTRORG	Entrainment rate	LN	0.39
	ENTSHALP	Shallow entrainment rate	LN	0.39
	DETRPEN	Detrainment rate for penetrative convection	LN	0.39
	RPRCON	Conversion coefficient cloud to rain	LN	0.52
	CUDU	Zonal convective momentum transport, deep convection	N	1.22
	CUDV	Meridional convective momentum transport, deep convection	N	1.22
	CUDUS	Zonal convective momentum transport, shallow convection	N	1.33
	CUDVS	Meridional convective momentum transport, shallow convection	N	1.33
	RTAU	Adjustment time-scale in CAPE closure	LN	0.78
	ENTSTPC1	Shallow entrainment test parcel entrainment	LN	0.39
Cloud and large-scale precipitation	RAMID	Relative humidity threshold for the onset of stratiform condensation	LN	0.13
	RCLDIFF	Diffusion coefficient for the evaporation of cloud at subgrid cloud edges	LN	1.04
	RLCRITSNOW	Cloud ice threshold for autoconversion to snow	LN	0.78
	RAINEVAP	Rain evaporation rate	LN	0.65
	SNOWSUBLIM	Snow sublimation rate	LN	0.65
	QSATVERVEL	Vertical velocity used to calculate the adiabatic temperature change for saturation adjustment	LN	0.39
Radiation	ZDECORR	Cloud vertical decorrelation height	Ln	0.78
	ZSIGQCW	Fractional st. dev. of horiz. distribution of water content	LN	0.52
	ZRADEFF	Effective radius of cloud water and ice	LN	0.78
	ZHS_VDAERO	Scale height of aerosol normal vertical distribution	LN	1.04
	DELTA_AERO	Optical thickness of aerosol	LN	0.78

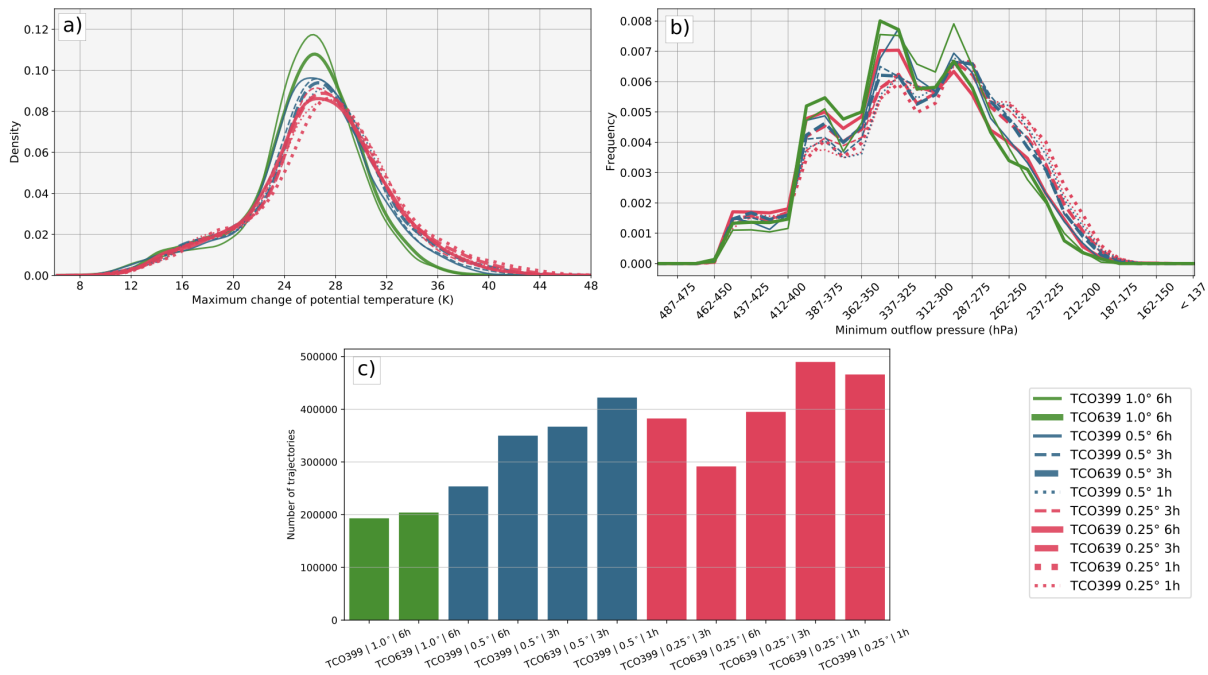


Figure A.1.: (a) Maximum change of potential temperature, (b) minimum outflow pressure and (c) count of WCB-trajectories in a set of test experiments with different configurations for one case. The color of the bars/lines denotes the spatial resolution of the input files (green: 1° , blue: 0.5° , red: 0.25°), the style of the lines in (a) and (b) shows the temporal resolution of the input files (solid: 6-hourly, dashed: 3-hourly, dotted: 1-hourly), and the width of the lines represents the native resolution of the model (thin: TCo399, thick: TCo639). The data basis of the Figure is a 50-member ensemble forecast initialized on March 7, 2016, that is run for 120 hours. The trajectories are computed in a domain covering the North Atlantic and Europe (100°W - 50°E , 15° - 80°N), and the counts in (c) represent sums over all dimensions of the forecast (i.e. over all members and forecast lead times). Note that only 11 of 18 possible combinations (2 native resolutions, 3 grid spacings and temporal resolutions each) are shown.

B. Appendix for Chapter 4

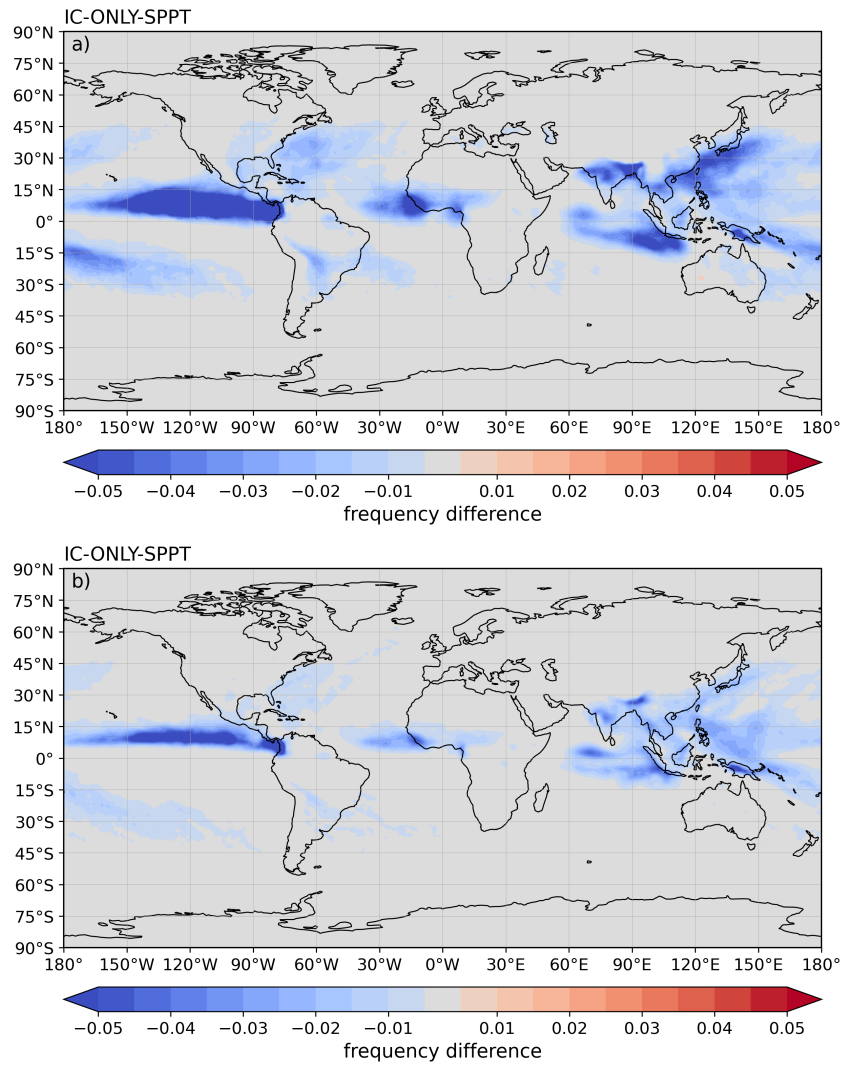


Figure B.1.: Absolute frequency differences of WCB inflow (a) and ascent (b) of the experiments *SPPT* and *IC-ONLY*, averaged over 32 forecast, 20 members and 41 lead times.

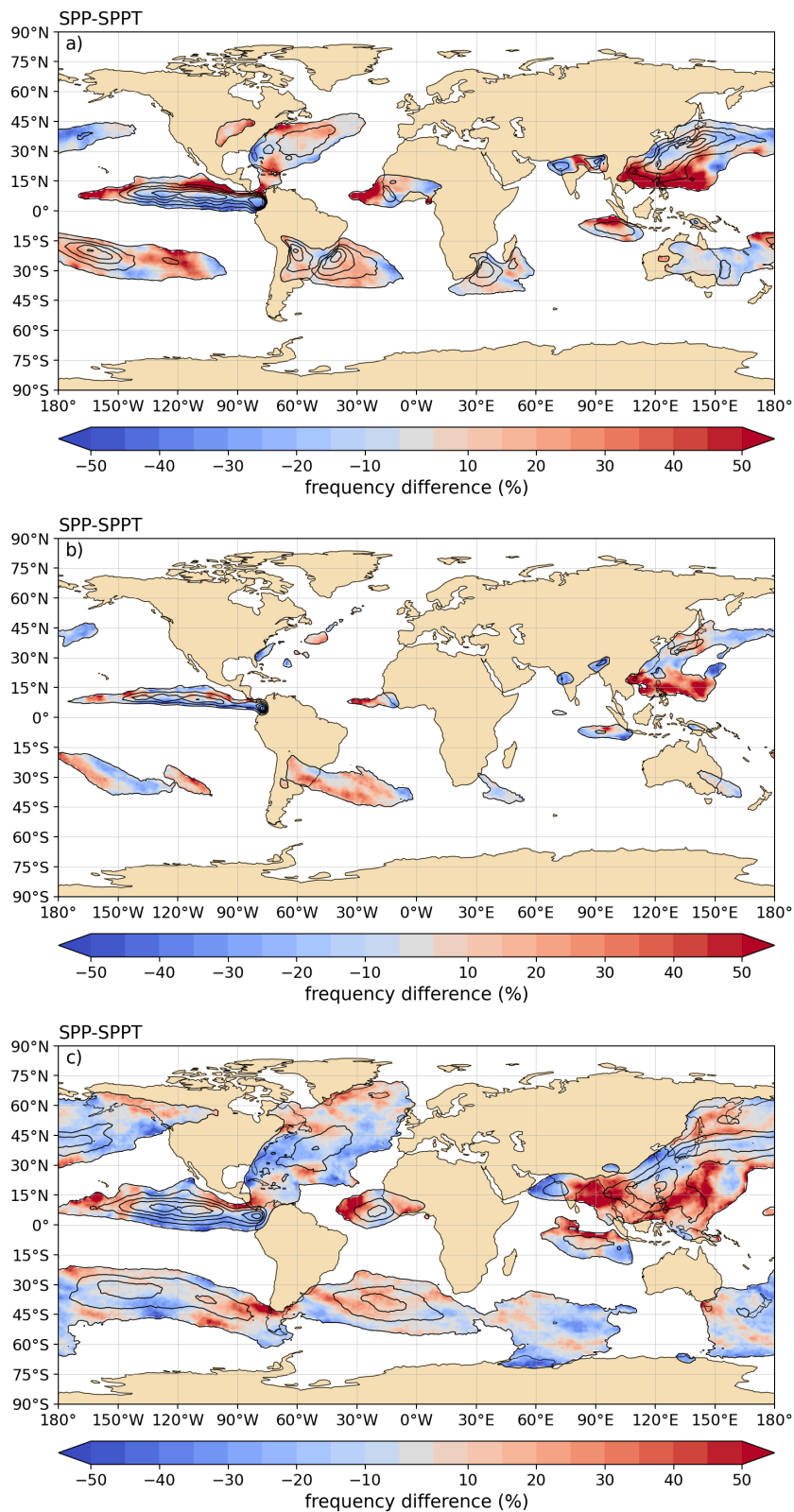


Figure B.2.: Relative frequency differences of WCB (a) inflow, (b) ascent and (c) outflow of the experiments *SPP* and *SPPT*, averaged over 32 forecast, 20 members and 41 lead times. Plotted are only differences where the absolute frequencies of *SPP* (black contours in 5%-intervals) are larger than 5%.

C. Appendix for Chapter 5

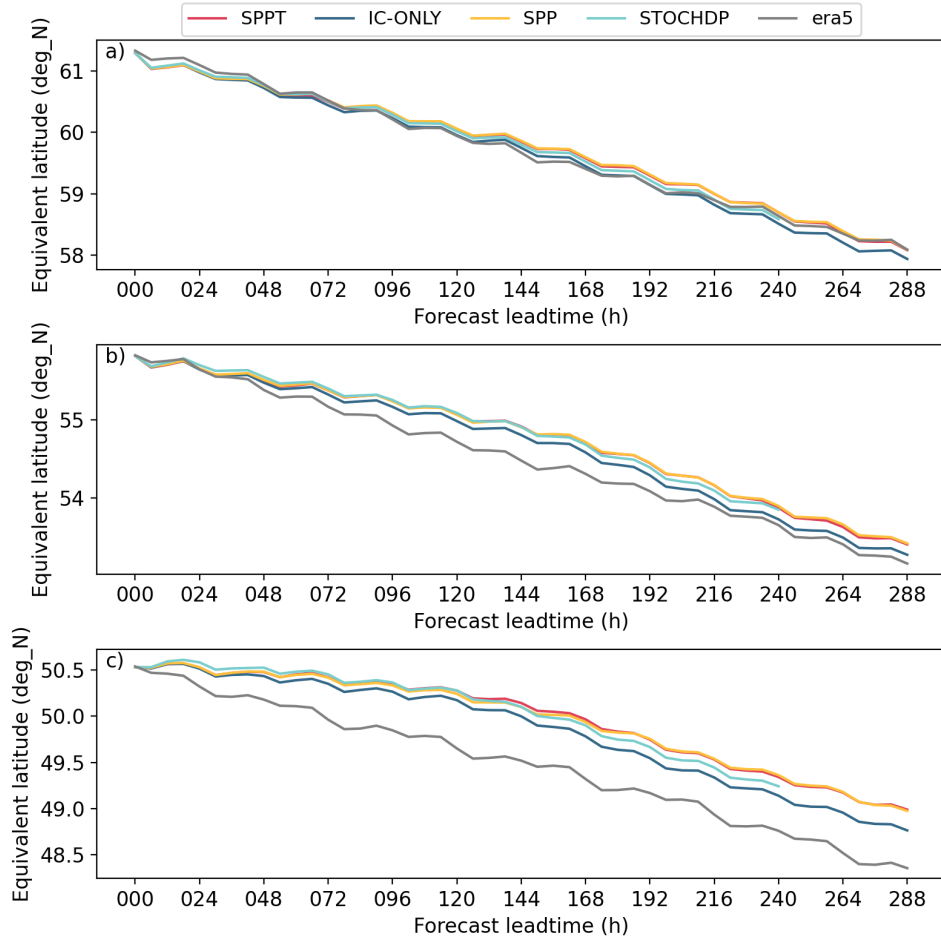


Figure C.1.: Equivalent latitude Φ_{eq} with forecast lead time averaged over 32 forecast initializations and 20 members of the experiments *IC-ONLY*, *SPPT*, *SPP* and *STOCHDP*, and of *ERA5* on 320 K (a), 325 K (b) and 330 K (c).

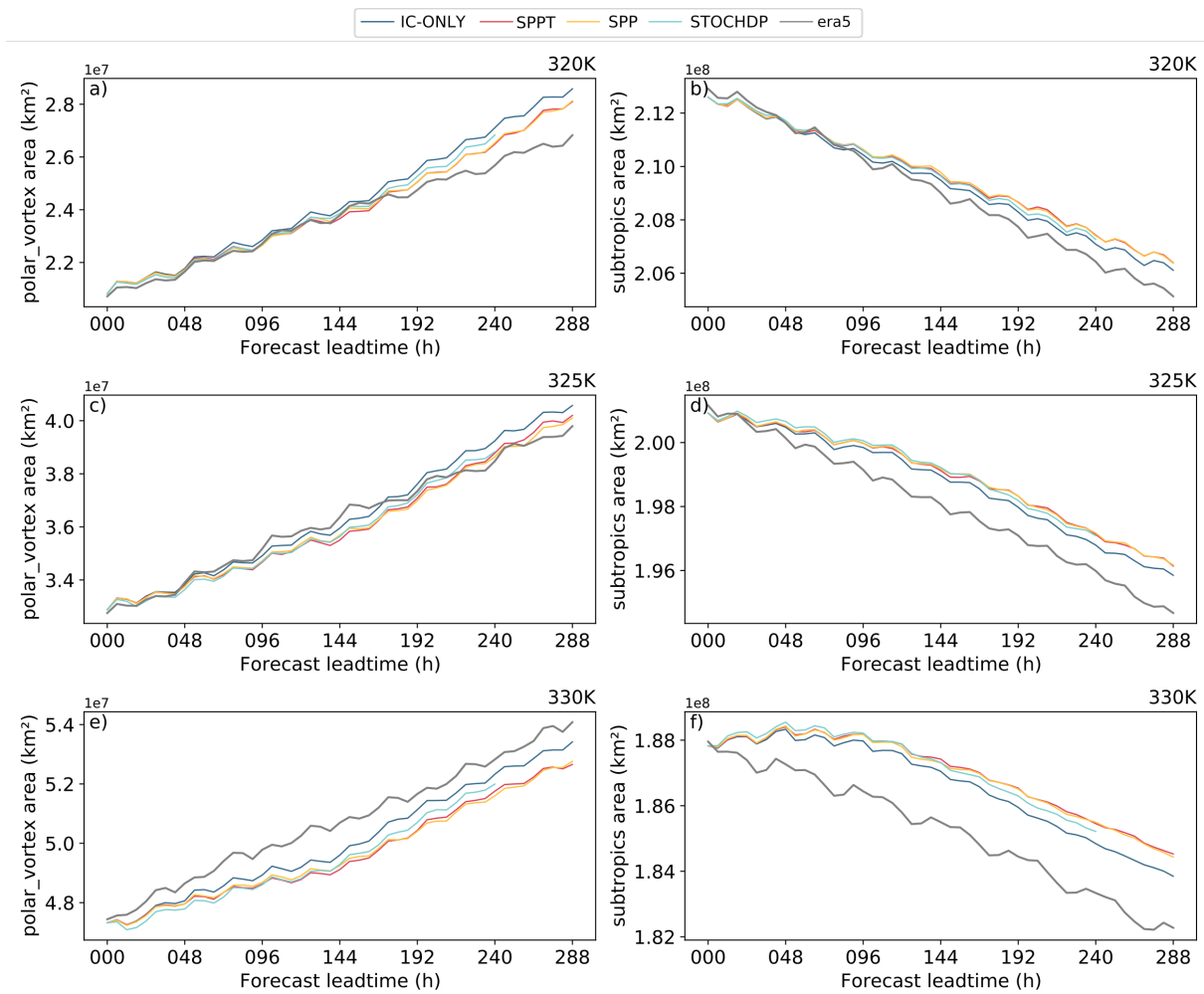


Figure C.2.: Evolution of the area of the polar vortex (left) and subtropics (right) on the 320 K (top), 325 K (middle) and 330 K (bottom) isentropes with forecast lead time averaged over 32 forecasts and 20 ensemble members of the experiments of the experiments *IC-ONLY*, *SPPT*, *SPP* and *STOCHDP*, and for *ERA5*. The equivalent latitudes are taken from each individual experiment.

D. Appendix for Chapter 6

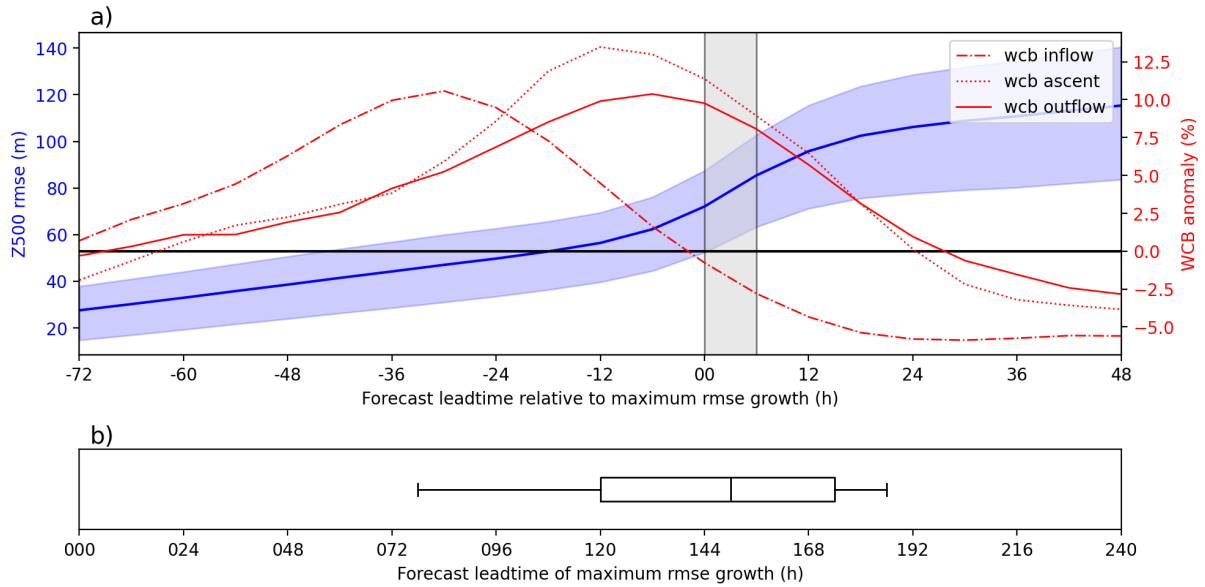


Figure D.1.: (a) Mean (blue line) and interquartile range (blue shading) of the evolution of the domain-integrated Z500 RMSE over Europe ($-12.5^{\circ}\text{W}-42.5^{\circ}\text{E}$, $35^{\circ}-75^{\circ}\text{N}$) and the domain-integrated relative WCB-anomaly (dash-dotted: inflow; dotted: ascent; solid: outflow) over the North Atlantic ($60^{\circ}\text{W}-0^{\circ}\text{E}$, $35^{\circ}-75^{\circ}\text{N}$) lagged on the lead time of maximum error growth over Europe ($-12.5^{\circ}\text{W}-42.5^{\circ}\text{E}$, $35^{\circ}-75^{\circ}\text{N}$). The gray bar highlights the section with the largest error growth between two time steps. (b) Distribution (box indicates the interquartile range, whiskers the 5-95 interquartile range) of forecast lead times when the maximum error growth occurs.

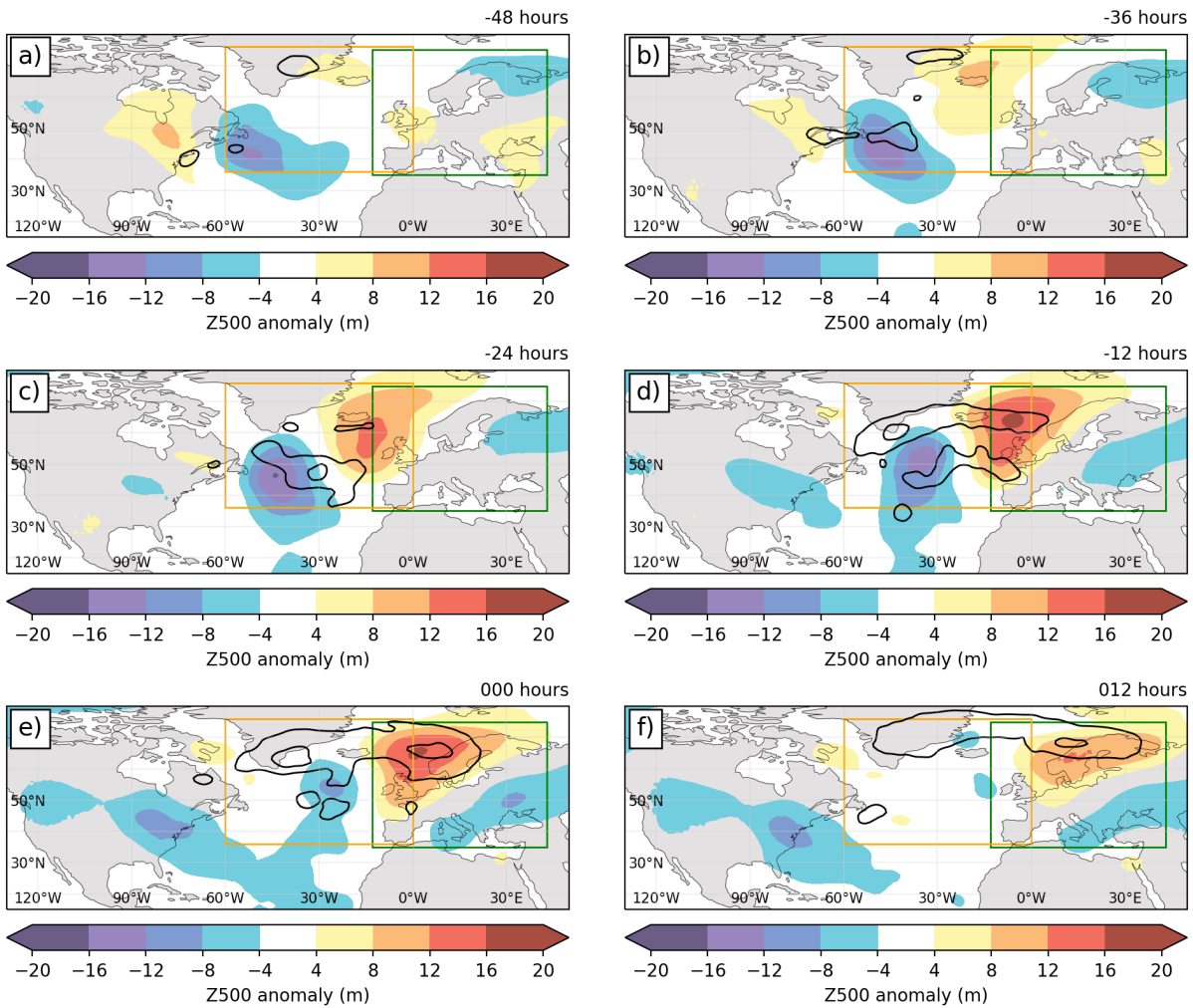


Figure D.2.: Composites of Z500 anomalies (shading) and absolute WCB outflow frequency anomalies (contours of 1, 2 and 3%) with time lags of (a) -48 hours, (b) -36 hours, (c) -24 hours, (d) -12 hours, (e) 0 hours and (f) 12 hours to the maximum error growth in the North Atlantic region ($-12.5^{\circ}\text{W}-42.5^{\circ}\text{E}$, $35^{\circ}-75^{\circ}\text{N}$) outlined by the green box. The orange box shows the region where the WCB-activity is computed for Figure D.1.

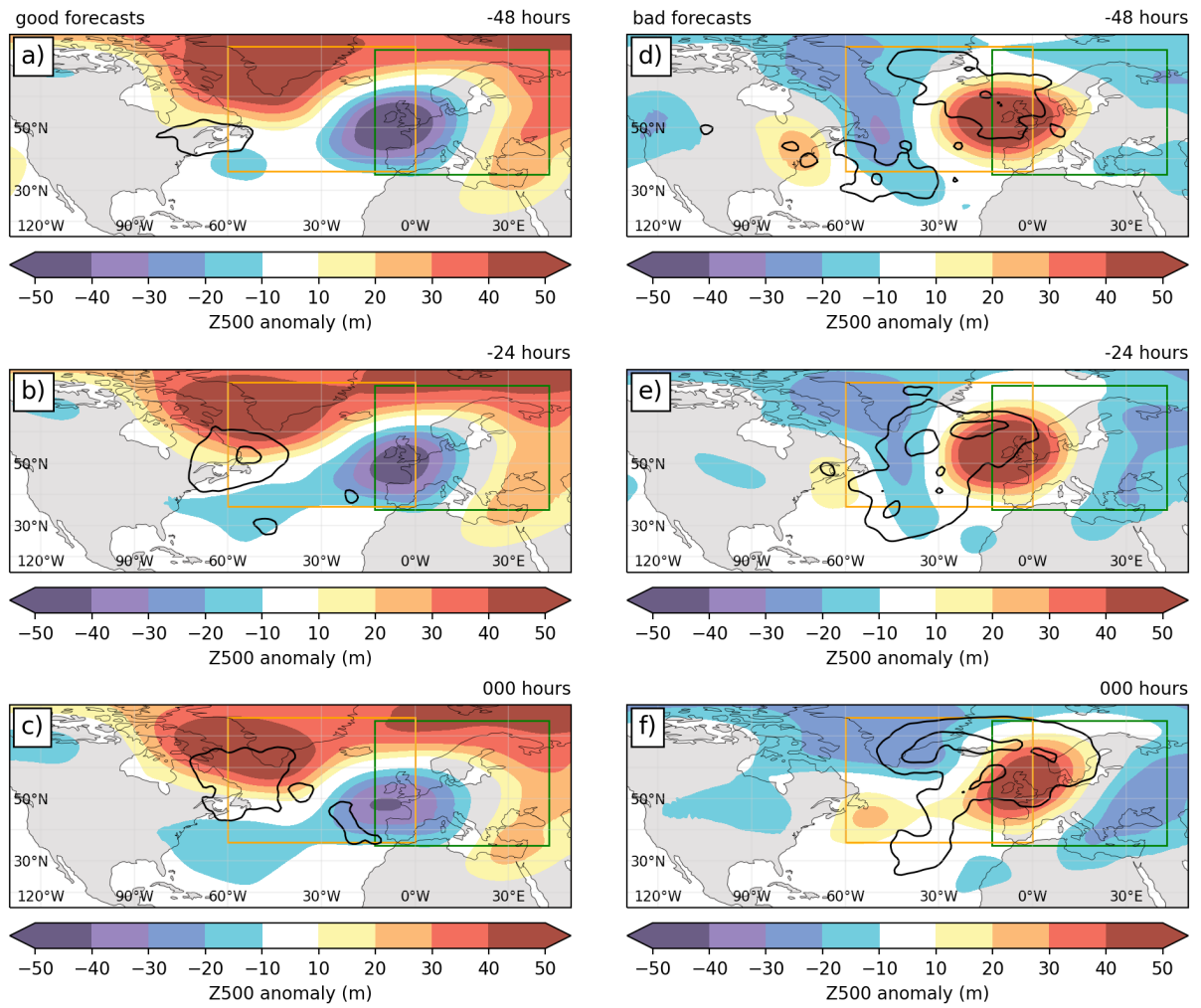


Figure D.3.: As Figure D.2, but for the 20% best (a-c) and 20% worst forecasts (d-f) in the North Atlantic region (-12.5°W-42.5°E, 35°-75°N, green box) for time lags -48 hours (a,d), -24 hours (b,e) and 0 hours (c,f) relative to the maximum error growth. WCB outflow anomalies are shown in contours from 2.5 to 10% in 2.5% intervals.

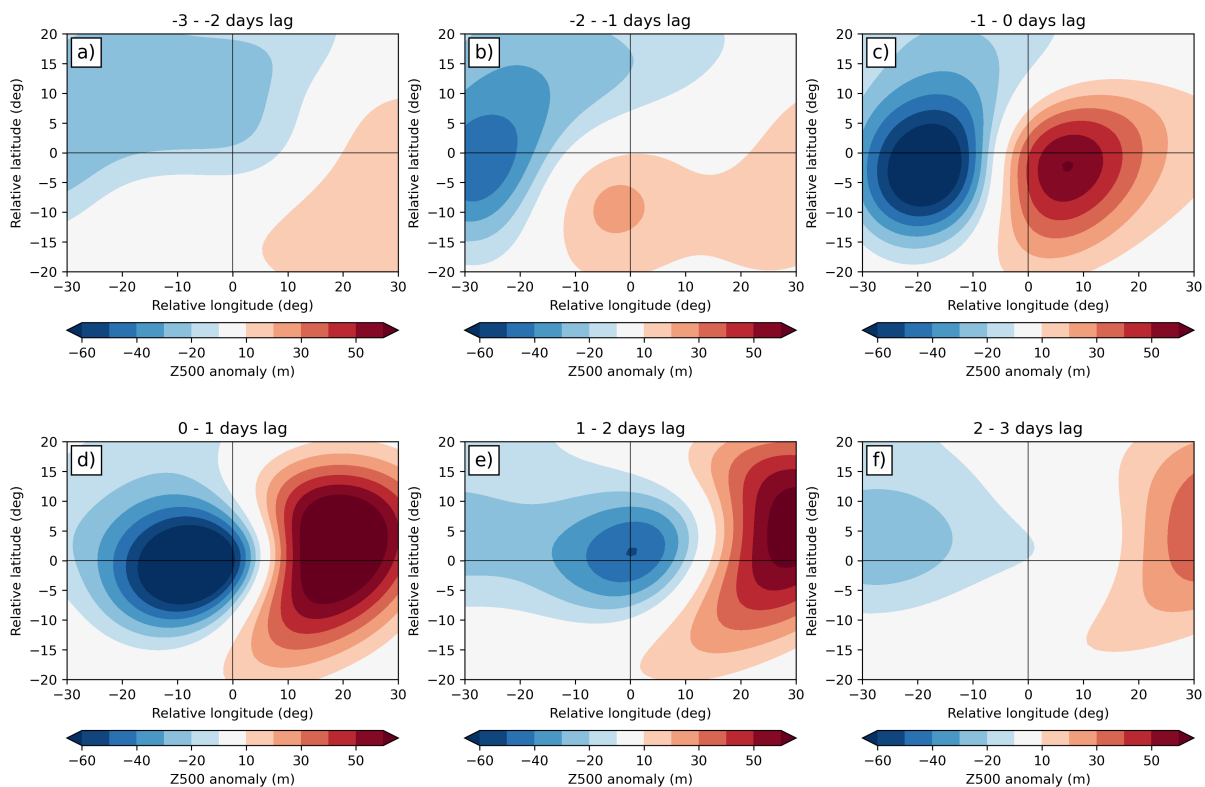


Figure D.4: Composite-means of Z500-anomalies on objects of WCB ascent. Anomalies are differences of the instantaneous field and the model climatology at the corresponding grid points. The spatial averages are means over the forecast lead times 72-168 hours (17 forecast lead times). Shown are temporal means over (a) 3-2 days (72-54 hours) before, (b) 2-1 days (48-30 hours) before, (c) 1-0 days (24-6 hours) before, (d) 0-1 days (0-18 hours after), (e) 1-2 days (24-42 hours) after, and (f) 2-3 days (48-66 hours) after the WCB-outflow event.

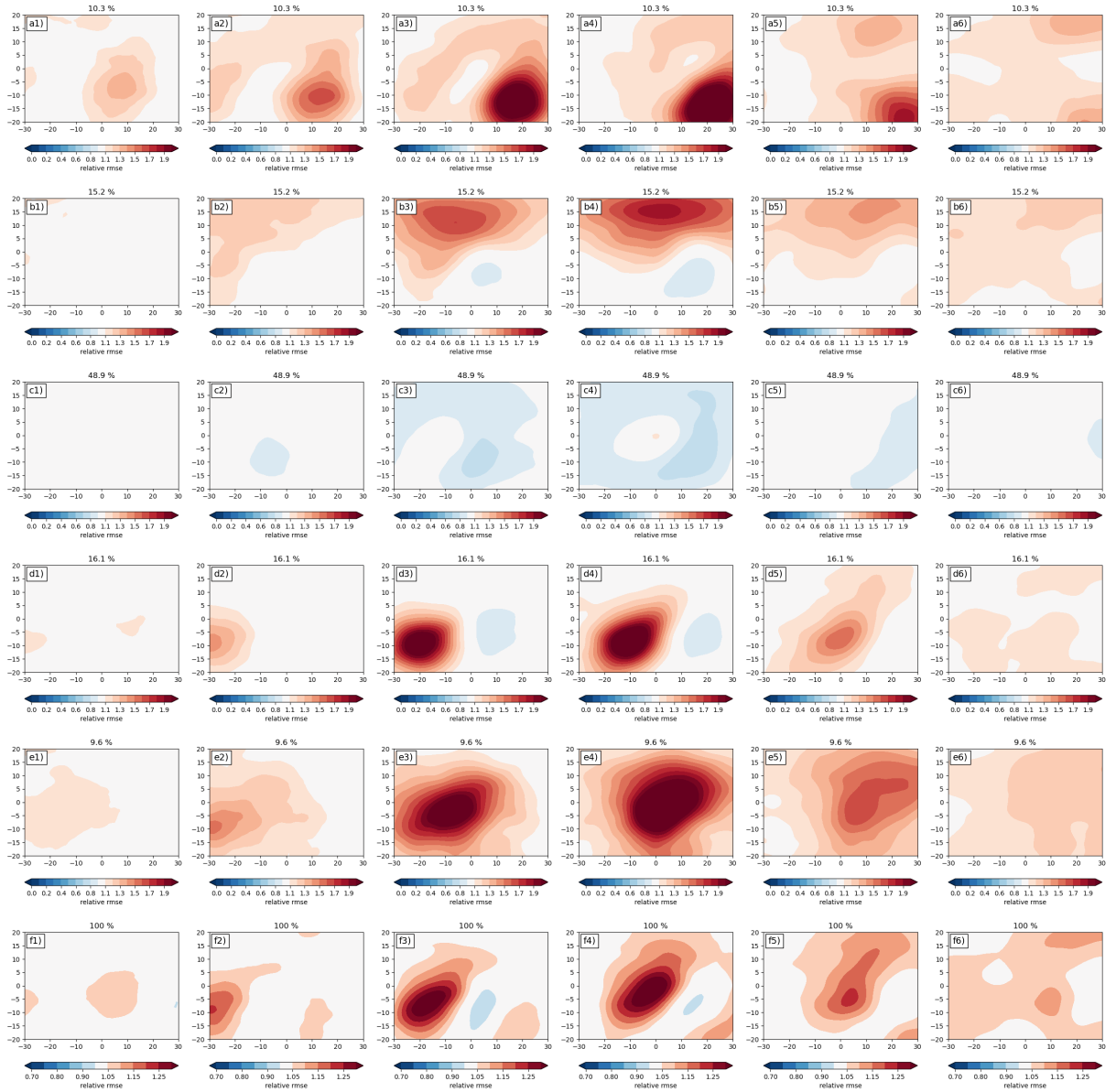


Figure D.5.: Composite-means of individual members obtained from a k-means cluster algorithm ($k=5$, Hartigan and Wong 1979) performed on Z500 RMSE centered on WCB ascent (i.e. Figure 6.10). Only WCB-ascent that occur at lead time 96 hours are considered. The clusters are evaluated based on the pattern of the centered Z500-field at time lag 0h. The rows a)-e represent the composite-means of the 5 clusters, and row f) shows the average of all clusters. The numbers 1)-6) show the time lags 72-54h (1), 48-30h (2) and 24-6 (3) prior to and 0-18h (4), 24-42h (5) and 48-66h (6) after the WCB-ascent. The percentage in every panel shows the fraction of individual cases in each cluster.

E. Bibliography

- Ahmadi-Givi, F., G. C. Graig, and R. S. Plant, 2004: The dynamics of a midlatitude cyclone with very strong latent-heat release. *Quarterly Journal of the Royal Meteorological Society*, **130** (596), 295–323.
- Aiyyer, A., 2015: Recurving western North Pacific tropical cyclones and midlatitude predictability. *Geophysical Research Letters*, **42** (18), 7799–7807.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525** (7567), 47–55.
- 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.
- Bechthold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, **134**, 1337–1351.
- Berman, J. D. and R. D. Torn, 2019: The impact of initial condition and warm conveyor belt forecast uncertainty on variability in the downstream waveguide in an ECWFM case study. *Monthly Weather Review*, **147** (11), 4071–4089.
- Berner, J., U. Achatz, L. Batté, L. Bengtsson, A. De La Cámara, H. M. Christensen, M. Colangeli, D. R. Coleman, D. Crommelin, S. I. Dolaptchiev, C. L. Franzke, P. Friederichs, P. Imkeller, H. Järvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajaajaajan, T. N. Palmer, C. Penland, M. Sakradz-ijaja, J. S. Von Storch, A. Weisheimer, M. Weniger, P. D. Williams, and J. I. Yano, 2017: Stochastic parameterization toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, **98** (3), 565–587.
- Berner, J., K. R. Fossell, S. Y. Ha, J. P. Hacker, and C. Snyder, 2015: Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations. *Monthly Weather Review*, **143** (4), 1295–1320.
- 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.

- Binder, H., M. Boettcher, H. Joos, and H. Wernli, 2016: The role of warm conveyor belts for the intensification of extratropical cyclones in Northern Hemisphere winter. *Journal of the Atmospheric Sciences*, **73** (10), 3997–4020.
- Bowman, K. P., J. C. Lin, A. Stohl, R. Draxler, P. Konopka, A. Andrews, and D. Brunner, 2013: Input data requirements for Lagrangian trajectory models. *Bulletin of the American Meteorological Society*, **94** (7), 1051–1058.
- Browning, K. A., 1971: Radar Measurements of Air Motion Near Fronts. *Weather*, **26** (8), 320–340.
- Büeler, D., L. Ferranti, L. Magnusson, J. F. Quinting, and C. M. Grams, 2021: Year-round sub-seasonal forecast skill for Atlantic-European weather regimes. *Quarterly Journal of the Royal Meteorological Society*, **147** (741), 4283–4309.
- Buizza, R., M. Miller, 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.
- Buizza, R. and D. Richardson, 2017: 25 Years of ensemble prediction at ECMWF. *ECMWF Newsletter*, **153**, 20–31.
- Butchart, N. and E. E. Remsburg, 1986: The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface. *Journal of the Atmospheric Sciences*, **43** (13), 1319–1339.
- Carlson, T. N., 1980: Airflow through midlatitude cyclones and the comma cloud pattern. *Monthly Weather Review*, **108** (10), 1498–1509.
- 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.
- Charney, J. G., 1947: The Dynamics of Long Waves in a Baroclinic Westerly Current. *Journal of Meteorology*, **4** (5), 135–162.
- Charney, J. G., R. Fjørtoft, and J. V. Neumann, 1950: Numerical Integration of the Barotropic Vorticity Equation. *Tellus*, **2** (4), 237–254.
- Choudhary, A. and A. Voigt, 2022: ICON simulations of cloud diabatic processes in the warm conveyor belt of North Atlantic cyclone Vladiana. *Weather and Climate Dynamics Discussions*.
- Christensen, H. M., 2020: Constraining stochastic parametrisation schemes using high-resolution simulations. *Quarterly Journal of the Royal Meteorological Society*, **146** (727), 938–962, 1904.04503.
- Christensen, H. M., J. Berner, D. R. Coleman, and T. N. Palmer, 2017: Stochastic parameterization and El Niño-southern oscillation. *Journal of Climate*, **30** (1), 17–38.

- Christensen, H. M., I. M. Moroz, and T. N. Palmer, 2015: Simulating weather regimes: impact of stochastic and perturbed parameter schemes in a simple atmospheric model. *Climate Dynamics*, **44**, 2195–2214.
- Clarke, S. J., S. L. Gray, and N. M. Roberts, 2019: Downstream influence of mesoscale convective systems. Part 1: influence on forecast evolution. *Quarterly Journal of the Royal Meteorological Society*, **145 (724)**, 2933–2952.
- Dacre, H. F., O. Martínez-Alvarado, and C. O. Mbengue, 2019: Linking atmospheric rivers and warm conveyor belt airflows. *Journal of Hydrometeorology*, **20 (6)**, 1183–1196.
- Davies, H. C. and M. Didone, 2013: Diagnosis and dynamics of forecast error growth. *Monthly Weather Review*, **141 (7)**, 2483–2501.
- Davini, P. and F. D’Andrea, 2020: From CMIP3 to CMIP6: Northern hemisphere atmospheric blocking simulation in present and future climate. *Journal of Climate*, **33 (23)**, 10 021–10 038.
- Davini, P., A. Weisheimer, M. Balmaseda, S. J. Johnson, F. Molteni, C. D. Roberts, R. Senan, and T. N. Stockdale, 2021: The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, **147 (735)**, 1344–1363.
- Dawson, A. and T. N. Palmer, 2015: Simulating weather regimes: impact of model resolution and stochastic parameterization. *Climate Dynamics*, **44**, 2177–2193.
- Dorrington, J., 2021: On the variability and forced response of atmospheric regime systems. Ph.D. thesis, University of Oxford.
- Downton, R. A. and R. S. Bell, 1988: The impact of analysis differences on a medium-range forecast. *Meteorological Magazine*, **117**, 279–285.
- Eady, E. T., 1949: Long Waves and Cyclone Waves. *Tellus*, **1 (3)**, 33–52.
- Eckhardt, S., A. Stohl, H. Wernli, P. James, C. Forster, and N. Spichtinger, 2004: A 15-year climatology of warm conveyor belts. *Journal of Climate*, **17 (1)**, 218–237.
- ECMWF, 2019a: IFS Documentation CY46R1 - Part III: Dynamics and numerical procedures. *IFS documentation Cy46r1*, 1–29.
- , 2019b: IFS Documentation CY46R1 - Part IV: Physical Processes. *IFS documentation CY46R1*, 1–223.
- , 2019c: IFS Documentation CY46R1 - Part VI: Technical and Computational Procedures. *IFS documentation CY46R1*, 1–239.

- Ertel, H., 1942: Ein neuer hydrodynamischer Erhaltungssatz. *Die Naturwissenschaften*, **30** (36), 543–544.
- Evans, C., K. M. Wood, S. D. Aberson, H. M. Archambault, S. M. Milrad, L. F. Bosart, K. L. Corbosiero, C. A. Davis, J. R. Pinto, J. Doyle, C. Fogarty, T. J. Galarneau, C. M. Grams, K. S. Griffin, J. Gyakum, R. E. Hart, N. Kitabatake, H. S. Lentink, R. Mctaggart-Cowan, W. Perrie, J. F. Quinting, C. A. Reynolds, M. Riemer, E. A. Ritchie, Y. Sun, and F. Zhang, 2017: The extratropical transition of tropical cyclones. Part I: Cyclone evolution and direct impacts. *Monthly Weather Review*, **145** (11), 4317–4344.
- Ferranti, L., S. Corti, and M. Janousek, 2015: Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, **141** (688), 916–924.
- Forbes, R. M., A. M. Tompkins, and A. Untch, 2011: A new prognostic bulk microphysics scheme for the ECMWF forecast model. *ECMWF Technical Memorandum*, **649**, 1–28.
- Grams, C. M. and H. M. Archambault, 2016: The key role of diabatic outflow in amplifying the midlatitude flow: A representative case study of weather systems surrounding western North Pacific extratropical transition. *Monthly Weather Review*, **144** (10), 3847–3869.
- 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.
- 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.
- Grams, C. M., H. Wernli, M. Böttcher, J. Čampa, U. Corsmeier, S. C. Jones, J. H. Keller, C. J. Lenz, and L. Wiegand, 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.
- 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.
- Haiden, T., M. Janousek, F. Vitart, Z. Ben Bouallegue, L. Ferranti, F. Prates, and D. Richardson, 2021: Evaluation of ECMWF forecasts, including the 2020 upgrade. *ECMWF Technical Memorandum*, **880**, 1–54.
- Harrold, T. W., 1973: Mechanisms influencing the distribution of precipitation within baroclinic disturbances. *Quarterly Journal of the Royal Meteorological Society*, **99** (420), 232–251.

- Hartigan, J. and M. Wong, 1979: Algorithm AS 136: a k-means clustering algorithm. *Journal of the Royal Statistical Society Series C*, **28** (1), 100–108.
- Haynes, P. H. and M. E. McIntyre, 1987: On the Evolution of Vorticity and Potential Vorticity in the Presence of Diabatic Heating and Frictional or Other Forces. *Journal of the Atmospheric Sciences*, **44** (5), 828–841.
- Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J. N. Thépaut, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146** (730), 1999–2049.
- Holton, J. R., 2004: *An Introduction to Dynamic Meteorology*, Vol. 88. Academic Press, 1–535 pp.
- Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *The Quarterly Journal of the Royal Meteorological Society*, **111** (466), 877–946.
- Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Technical Memorandum*, **636**, 1–41.
- Jiang, H. and E. J. Zipser, 2010: Contribution of Tropical Cyclones to the Global Precipitation from Eight Seasons of TRMM Data: Regional, Seasonal, and Interannual Variations. *Journal of Climate*, **23** (6), 1526–1543.
- Joos, H. and R. M. Forbes, 2016: Impact of different IFS microphysics on a warm conveyor belt and the downstream flow evolution. *Quarterly Journal of the Royal Meteorological Society*, **142** (700), 2727–2739.
- Joos, H. and H. Wernli, 2012: Influence of microphysical processes on the potential vorticity development in a warm conveyor belt: A case-study with the limited-area model COSMO. *Quarterly Journal of the Royal Meteorological Society*, **138** (663), 407–418.
- Keller, J. H., C. M. Grams, M. Riemer, H. M. Archambault, L. Bosart, J. D. Doyle, J. L. Evans, T. J. Galarneau, K. Griffin, P. A. Harr, N. Kitabatake, R. McTaggart-Cowan, F. Pantillon, J. F. Quinting, C. A. Reynolds, E. A. Ritchie, R. D. Torn, and F. Zhang, 2019: The extratropical transition of tropical cyclones. Part II: Interaction with the midlatitude flow, downstream impacts, and implications for predictability. *Monthly Weather Review*, **147** (4), 1077–1106.

- Lang, S. T., S. J. Lock, M. Leutbecher, P. Bechtold, and R. M. Forbes, 2021: Revision of the Stochastically Perturbed Parametrisations model uncertainty scheme in the Integrated Forecasting System. *Quarterly Journal of the Royal Meteorological Society*, **147 (735)**, 1364–1381.
- Leutbecher, M., 2019: Ensemble size: How suboptimal is less than infinity? *Quarterly Journal of the Royal Meteorological Society*, **145 (S1)**, 107–128.
- Leutbecher, M., S. J. Lock, P. Ollinaho, S. T. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, S. English, M. Fisher, R. M. Forbes, J. Goddard, T. Haiden, R. J. Hogan, S. Juricke, H. Lawrence, D. MacLeod, L. Magnusson, S. Malardel, S. Massart, I. Sandu, P. K. Smolarkiewicz, A. Subramanian, F. Vitart, N. Wedi, and A. Weisheimer, 2017: Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, **143 (707)**, 2315–2339.
- Leutbecher, M. and T. N. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics*, **227 (7)**, 3515–3539.
- Lewis, J. M., 2005: Roots of ensemble forecasting. *Monthly Weather Review*, **133 (7)**, 1865–1885.
- Lock, S. J., S. T. Lang, M. Leutbecher, R. J. Hogan, and F. Vitart, 2019: Treatment of model uncertainty from radiation by the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme and associated revisions in the ECMWF ensembles. *Quarterly Journal of the Royal Meteorological Society*, **145 (S1)**, 75–89.
- Lorenz, E., 1963: Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, **20**, 130–141.
- Lorenz, E. N., 1965: On the possible reasons for long-period fluctuations of the general circulation. *WMO-IUGG Symposium on Research and Development Aspects of Long-range Forecasting*, World Meteorological Organization, WMO, Boulder, CO, 203–211.
- Maddison, J. W., S. L. Gray, O. Martínez-Alvarado, and K. D. Williams, 2020: Impact of model upgrades on diabatic processes in extratropical cyclones and downstream forecast evolution. *Quarterly Journal of the Royal Meteorological Society*, **146 (728)**, 1322–1350.
- Madonna, E., M. Boettcher, C. M. Grams, H. Joos, O. Martius, and H. Wernli, 2015: Verification of North Atlantic warm conveyor belt outflows in ECMWF forecasts. *Quarterly Journal of the Royal Meteorological Society*, **141 (689)**, 1333–1344.
- Madonna, E., S. Limbach, C. Aebi, H. Joos, H. Wernli, and O. Martius, 2014a: On the co-occurrence of warm conveyor belt outflows and PV streamers. *Journal of the Atmospheric Sciences*, **71 (10)**, 3668–3673.

- Madonna, E., H. Wernli, H. Joos, and O. Martius, 2014b: Warm conveyor belts in the ERA-Interim Dataset (1979-2010). Part I: Climatology and potential vorticity evolution. *Journal of Climate*, **27** (1), 3–26.
- Magnusson, L., 2017: Diagnostic methods for understanding the origin of forecast errors. *Quarterly Journal of the Royal Meteorological Society*, **143** (706), 2129–2142.
- Magnusson, L. and E. Källén, 2013: Factors influencing skill improvements in the ECMWF forecasting system. *Monthly Weather Review*, **141** (9), 3142–3153.
- Martínez-Alvarado, O., S. L. Gray, and J. Methven, 2016a: Diabatic processes and the evolution of two contrasting summer extratropical cyclones. *Monthly Weather Review*, **144** (9), 3251–3276.
- Martínez-Alvarado, O., J. W. Maddison, S. L. Gray, and K. D. Williams, 2018: Atmospheric blocking and upper-level Rossby-wave forecast skill dependence on model configuration. *Quarterly Journal of the Royal Meteorological Society*, **144** (716), 2165–2181.
- Martínez-Alvarado, O., E. Madonna, S. L. Gray, and H. Joos, 2016b: A route to systematic error in forecasts of Rossby waves. *Quarterly Journal of the Royal Meteorological Society*, **142** (694), 196–210.
- Martínez-Alvarado, O. and R. S. Plant, 2014: Parametrized diabatic processes in numerical simulations of an extratropical cyclone. *Quarterly Journal of the Royal Meteorological Society*, **140** (682), 1742–1755.
- Matsueda, M., 2011: Predictability of Euro-Russian blocking in summer of 2010. *Geophysical Research Letters*, **38** (6), 1–6.
- Mazoyer, M., D. Ricard, G. Rivière, J. Delanoë, P. Arbogast, B. Vié, C. Lac, Q. Cazenave, and J. Pelon, 2021: Microphysics Impacts on the Warm Conveyor Belt and Ridge Building of the NAWDEX IOP6 Cyclone. *Monthly Weather Review*, **149** (12), 3961–3980.
- Methven, J., 2015: Potential vorticity in warm conveyor belt outflow. *Quarterly Journal of the Royal Meteorological Society*, **141** (689), 1065–1071.
- Methven, J. and P. Berrisford, 2015: The slowly evolving background state of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **141** (691), 2237–2258.
- Molteni, F. and T. Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. *Quarterly Journal of the Royal Meteorological Society*, **119** (510), 269–298.
- 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.
- Oertel, A., M. Boettcher, H. Joos, M. Sprenger, and H. Wernli, 2020: Potential vorticity structure of embedded convection in a warm conveyor belt and its relevance for large-scale dynamics. *Weather and Climate Dynamics*, **1 (1)**, 127–153.
- Ollinaho, P., S. J. Lock, M. Leutbecher, P. Bechtold, A. Beljaars, A. Bozzo, R. M. Forbes, T. Haiden, R. J. Hogan, and I. Sandu, 2017: Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble. *Quarterly Journal of the Royal Meteorological Society*, **143 (702)**, 408–422.
- Palmer, T., 2019a: The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Quarterly Journal of the Royal Meteorological Society*, **145 (S1)**, 12–24.
- Palmer, T. N., 2019b: Stochastic weather and climate models. *Nature Reviews Physics*, **1 (7)**, 463–471.
- Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, C. P., and J. Tribbia, 1992: Ensemble Prediction. *ECMWF Technical Memorandum*, **188**, 21–66.
- Pfahl, S., E. Madonna, M. Boettcher, H. Joos, and H. Wernli, 2014: Warm conveyor belts in the ERA-Interim Dataset (1979-2010). Part II: Moisture origin and relevance for precipitation. *Journal of Climate*, **27 (1)**, 27–40.
- Pfahl, S., C. Schwiertz, M. Croci-Maspoli, C. M. Grams, and H. Wernli, 2015: Importance of latent heat release in ascending air streams for atmospheric blocking. *Nature Geoscience*, **8 (8)**, 610–614.
- Pfahl, S. and H. Wernli, 2012: Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-)daily time scales. *Geophysical Research Letters*, **39 (12)**, 1–6.
- Pickl, M., S. T. Lang, M. Leutbecher, and C. M. Grams, 2022: The effect of stochastically perturbed parametrisation tendencies (SPPT) on rapidly ascending air streams. *Quarterly Journal of the Royal Meteorological Society*, **148 (744)**, 1242–1261.
- Pomroy, H. R. and A. J. Thorpe, 2000: The evolution and dynamical role of reduced upper-tropospheric potential vorticity in intensive observing period one of FASTEX. *Monthly Weather Review*, **128 (6)**, 1817–1834.
- Quinting, J. F. and C. Grams, 2022: Eulerian Identification of ascending Air Streams (ELIAS 2.0) in Numerical Weather Prediction and Climate Models. Part I: Development of deep learning model. *Geoscientific Model Development*, **15 (2)**, 715–730.

- Quinting, J. F. and C. M. Grams, 2021: Toward a systematic evaluation of warm conveyor belts in numerical weather prediction and climate models. Part I: Predictor selection and logistic regression model. *Journal of the Atmospheric Sciences*, **78** (5), 1465–1485.
- 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.
- Quinting, J. F. and F. Vitart, 2019: Representation of Synoptic-Scale Rossby Wave Packets and Blocking in the S2S Prediction Project Database. *Geophysical Research Letters*, **46** (2), 1070–1078.
- Rabier, F., H. Järvinen, E. Klinker, J. F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society*, **126** (564), 1143–1170.
- Rasp, S., T. Selz, and G. C. Craig, 2016: Convective and slantwise trajectory ascent in convection-permitting simulations of midlatitude cyclones. *Monthly Weather Review*, **144** (10), 3961–3976.
- Rivière, G., M. Wimmer, P. Arbogast, J.-M. Piriou, J. Delanoë, C. Labadie, Q. Cazenave, and J. Pelon, 2021: The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during NAWDEX IOP6. *Weather and Climate Dynamics*, **2** (4), 1011–1031.
- Rodwell, M. J., R. Forbes, and H. Wernli, 2018a: Why warm conveyor belts matter in NWP. *ECMWF Newsletter*, **154**, 21–28.
- Rodwell, M. J., L. Magnusson, P. Bauer, P. Bechtold, M. Bonavita, C. Cardinali, M. Diamantakis, P. Earnshaw, A. Garcia-Mendez, L. Isaksen, E. Källén, D. Klocke, P. Lopez, T. McNally, A. Persson, F. Prates, and N. Wedi, 2013: Characteristics of occasional poor medium-range weather forecasts for Europe. *Bulletin of the American Meteorological Society*, **94** (9), 1393–1405.
- Rodwell, M. J., D. S. Richardson, D. B. Parsons, and H. Wernli, 2018b: Flow-dependent reliability: A path to more skillful ensemble forecasts. *Bulletin of the American Meteorological Society*, **99** (5), 1015–1026.
- Rodwell, M. J. and H. Wernli, 2022: The Cyclogenesis Butterfly: Uncertainty growth and forecast reliability during extratropical cyclogenesis. *Weather and Climate Dynamics Discussions*, 1–32.
- Rossby, C.-G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *Journal of Marine Research*, **2** (1), 38–55.
- Röthlisberger, M., O. Martius, and H. Wernli, 2018: Northern Hemisphere Rossby Wave initiation events on the extratropical jet-A climatological analysis. *Journal of Climate*, **31** (2), 743–760.

- 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.
- Saffin, L., J. Methven, J. Bland, B. Harvey, and C. Sanchez, 2021: Circulation conservation in the outflow of warm conveyor belts and consequences for Rossby wave evolution. *Quarterly Journal of the Royal Meteorological Society*, **147** (740), 3587–3610.
- Sánchez, C., J. Methven, S. Gray, and M. Cullen, 2020: Linking rapid forecast error growth to diabatic processes. *Quarterly Journal of the Royal Meteorological Society*, **146** (732), 3548–3569.
- Schäfler, A., G. Craig, H. Wernli, P. Arbogast, J. D. Doyle, R. Mctaggart-Cowan, J. Methven, G. Rivière, F. Ament, M. Boettcher, M. Bramberger, Q. Cazenave, R. Cotton, S. Crewell, J. Delanoë, A. DörnbrAck, A. Ehrlich, F. Ewald, A. Fix, C. M. Grams, S. L. Gray, H. Grob, S. Groß, M. Hagen, B. Harvey, L. Hirsch, M. JAcob, T. Kölling, H. Konow, C. Lemmerz, O. Lux, L. Magnusson, B. Mayer, M. Mech, R. Moore, J. Pelon, J. Quinting, S. Rahm, M. Rapp, M. Rautenhaus, O. Reitebuch, C. A. Reynolds, H. Sodemann, T. Spengler, G. Vaughan, M. Wendisch, M. Wirth, B. Witschas, K. Wolf, and T. Zinner, 2018: The north atlantic waveguide and downstream impact experiment. *Bulletin of the American Meteorological Society*, **99** (8), 1607–1637.
- Schäfler, A. and F. Harnisch, 2015: Impact of the inflow moisture on the evolution of a warm conveyor belt. *Quarterly Journal of the Royal Meteorological Society*, **141** (686), 299–310.
- Selz, T., 2019: Estimating the intrinsic limit of predictability using a stochastic convection scheme. *Journal of the Atmospheric Sciences*, **76** (3), 757–765.
- Selz, T. and G. C. Craig, 2015: Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Monthly Weather Review*, **143** (3), 813–827.
- Selz, T., M. Riemer, and G. Craig, 2022: The transition from practical to intrinsic predictability of midlatitude weather. *Journal of the Atmospheric Sciences*, Early online release.
- Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, **131** (612), 3079–3102.
- Shutts, G., M. Leutbecher, A. Weisheimer, T. Stockdale, L. Isaksen, and M. Bonavita, 2011: Representing model uncertainty: stochastic parametrizations at ECMWF. *ECMWF Newsletter*, **129**, 19–24.
- Shutts, G. J. and T. N. Palmer, 2007: Convective forcing fluctuations in a cloud-resolving model: Relevance to the stochastic parameterization problem. *Journal of Climate*, **20** (2), 187–202.
- Sousa, P. M., R. M. Trigo, D. Barriopedro, P. M. Soares, and J. A. Santos, 2018: European temperature responses to blocking and ridge regional patterns. *Climate Dynamics*, **50**, 457–477.

- Sprenger, M., G. Fragkoulidis, H. Binder, M. Croci-Maspoli, P. Graf, C. M. Grams, P. Knippertz, E. Madonna, S. Schemm, B. Škerlak, and H. Wernli, 2017: Global climatologies of Eulerian and Lagrangian flow features based on ERA-Interim. *Bulletin of the American Meteorological Society*, **98** (8), 1739–1748.
- Sprenger, M. and H. Wernli, 2015: The LAGRANTO Lagrangian analysis tool - Version 2.0. *Geoscientific Model Development*, **8** (8), 2569–2586.
- Steinfeld, D., M. Boettcher, R. Forbes, and S. Pfahl, 2020: The sensitivity of atmospheric blocking to upstream latent heating – numerical experiments. *Weather and Climate Dynamics*, **1** (2), 405–426.
- Steinfeld, D. and S. Pfahl, 2019: The role of latent heating in atmospheric blocking dynamics: a global climatology. *Climate Dynamics*, **53**, 6159–6180.
- Stockdale, T., M. Alonso-Balmaseda, S. Johnson, L. Ferranti, F. Molteni, L. Magnusson, S. Tietsche, F. Vitart, D. Decremer, A. Weisheimer, C. D. Roberts, G. Balsamo, S. Keeley, K. Mogensen, H. Zuo, M. Mayer, and B. M. Monge-Sanz, 2018: SEAS5 and the future evolution of the long-range forecast system. *ECMWF Technical Memorandum*, **835**, 1–81.
- Strømme, K., H. M. Christensen, J. Berner, and T. N. Palmer, 2018: The impact of stochastic parametrisations on the representation of the Asian summer monsoon. *Climate Dynamics*, **50**, 2269–2282.
- Strømme, K., H. M. Christensen, D. Macleod, S. Juricke, and T. N. Palmer, 2019: Progress towards a probabilistic Earth system model: Examining the impact of stochasticity in the atmosphere and land component of EC-Earth v3.2. *Geoscientific Model Development*, **12** (7), 3099–3118.
- Subramanian, A., A. Weisheimer, T. Palmer, F. Vitart, and P. Bechtold, 2017: Impact of stochastic physics on tropical precipitation in the coupled ECMWF model. *Quarterly Journal of the Royal Meteorological Society*, **143** (703), 852–865.
- 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.
- , 2021: Potential-vorticity dynamics of troughs and ridges within Rossby wave packets during a 40-year reanalysis period. *Weather and Climate Dynamics*, **2** (3), 535–559.
- Thompson, P. and E. Lorenz, 1986: Transcript of Dialogue between Phil Thompson and Ed Lorenz. *Amer. Meteor. Soc. Tape Recorded Interview Project*, 19 pp.
- Tiedke, M., 1993: Representation of Clouds in Large-Scale Models. *Monthly Weather Review*, **121** (11), 3040–3061.
- Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, **117** (8), 1779–1800.

- Tompkins, A. M. and J. Berner, 2008: A stochastic convective approach to account for model uncertainty due to unresolved humidity variability. *Journal of Geophysical Research Atmospheres*, **113** (18), 1–12.
- Torn, R. D., 2017: A comparison of the downstream predictability associated with ET and baroclinic cyclones. *Monthly Weather Review*, **145** (11), 4651–4672.
- Vidale, P. L., K. Hodges, B. Vanni re, P. Davini, M. J. Roberts, K. Strommen, A. Weisheimer, E. Plesca, and S. Corti, 2021: Impact of stochastic physics and model resolution on the simulation of tropical cyclones in climate GCMs. *Journal of Climate*, **34** (11), 4315–4341.
- Vitart, F., C. Ardilouze, A. Bonet, A. Brookshaw, M. Chen, C. Codorean, M. D equ , L. Ferranti, E. Fucile, M. Fuentes, H. Hendon, J. Hodgson, H. S. Kang, A. Kumar, H. Lin, G. Liu, X. Liu, P. Malguzzi, I. Mallas, M. Manoussakis, D. Mastrangelo, C. MacLachlan, P. McLean, A. Minami, R. Mladek, T. Nakazawa, S. Najm, Y. Nie, M. Rixen, A. W. Robertson, P. Ruti, C. Sun, Y. Takaya, M. Tolstykh, F. Venuti, D. Waliser, S. Woolnough, T. Wu, D. J. Won, H. Xiao, R. Zaripov, and L. Zhang, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, **98** (1), 163–173.
- Wandel, J., J. F. Quinting, and C. M. Grams, 2021: Toward a systematic evaluation of warm conveyor belts in numerical weather prediction and climate models. Part II: Verification of operational reforecasts. *Journal of the Atmospheric Sciences*, **78** (12), 3965–3982.
- Wang, J. W. A., P. D. Sardeshmukh, G. P. Compo, J. S. Whitaker, L. C. Slivinski, C. M. McColl, and P. J. Pegen, 2019: Sensitivities of the NCEP global forecast system. *Monthly Weather Review*, **147** (4), 1237–1256.
- Wedi, N., 1999: The numerical coupling of the physical parametrizations to the "dynamical" equations in a forecast model. *ECMWF Technical Memorandum*, **174**, 1–37.
- Weisheimer, A., S. Corti, T. Palmer, and F. Vitart, 2014: Addressing model error through atmospheric stochastic physical parametrizations: Impact on the coupled ECMWF seasonal forecasting system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **372** (2018).
- Wernli, H., 1997: A Lagrangian-based analysis of extratropical cyclones. II: A detailed case-study. *Quarterly Journal of the Royal Meteorological Society*, **123** (542), 1677–1706.
- Wernli, H. and H. C. Davies, 1997: A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications. *Quarterly Journal of the Royal Meteorological Society*, **123** (538), 467–489.
- Wernli, H. and C. Schwierz, 2006: Surface cyclones in the ERA-40 dataset (1958-2001). Part I: Novel identification method and global climatology. *Journal of the Atmospheric Sciences*, **63** (10), 2486–2507.

- Wilks, D. S., 2005: Effects of stochastic parametrizations in the Lorenz '96 system. *Quarterly Journal of the Royal Meteorological Society*, **131** (606), 389–407.
- Yang, C., H. M. Christensen, S. Corti, J. von Hardenberg, and P. Davini, 2019: The impact of stochastic physics on the El Niño Southern Oscillation in the EC-Earth coupled model. *Climate Dynamics*, **53**, 2843–2859.
- 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.

Acknowledgments

First and foremost I would like to thank my supervisor and first reviewer Christian Grams, who gave me the opportunity to work on this exciting research topic and always supported me with scientific discussions and advice. His positive attitude always gave me inspiration and motivation boosts, for which I am very grateful. He further helped me to build a scientific network, established the collaboration with colleagues at ECMWF and gave me the freedom to develop and explore my own research ideas.

Many thanks also go to Corinna Hoose, who agreed on acting as a second reviewer and contributed to the progress of the project with fruitful discussions in regular meetings. Special thanks go to Simon Lang for numerous scientific discussions, advice for the setup of the ensemble experiments, support during the implementation of the IFS postprocessing tool, and especially for the nice and straightforward communication. Further I would like to thank Martin Leutbecher for his input for the initial project idea, for his scientific advice and for providing the computational resources for the project.

I am very grateful to have been part of the SPREADOUT team during my PhD - everyone in the group contributed to a great atmosphere and team spirit, that we were able to keep up even during Covid and times of home office. Thank you also for valuable scientific discussions, honest and encouraging feedback, and mental support when it was needed. I very much enjoyed our trips to Kloster Seeon, Japan and Klingenmünster and also our various freetime activities (Bowling, group hikes, Glühwein, ...).

In particular, I want to thank:

Julian for his tireless technical support and valuable scientific input,

Dominik for various discussions off- and on topic,

my fellow PhD-colleagues Sera and Jan for sharing the ups and downs that come along with a PhD, for listening ears, a good atmosphere in the office and for fun discussions about "Der Bachelor",

all the other group members, namely Annika, Fabian, Josh, Maria, Marisol and Marta,

Sarah-Jane Lock for setting up the STOCHDP-experiments,

Doris, Rosi and Roswitha for their support in administrative matters and Gabi for her help with IT affairs, the team of the "Werkstatt" for building a customized clothes rack for our office,

the Steinbuch Computing Center (SCC) for providing computing facilities,

and Deutscher Wetterdienst for providing access to ECMWF data.

Last but not least, I want to thank my friends who made the past few years very worthwhile also outside of the scientific world, my parents, my sister Julia and my brother Christoph for their continuous support, and, of course Clara for always being there for me; without you, this would not have been possible. Thank you.