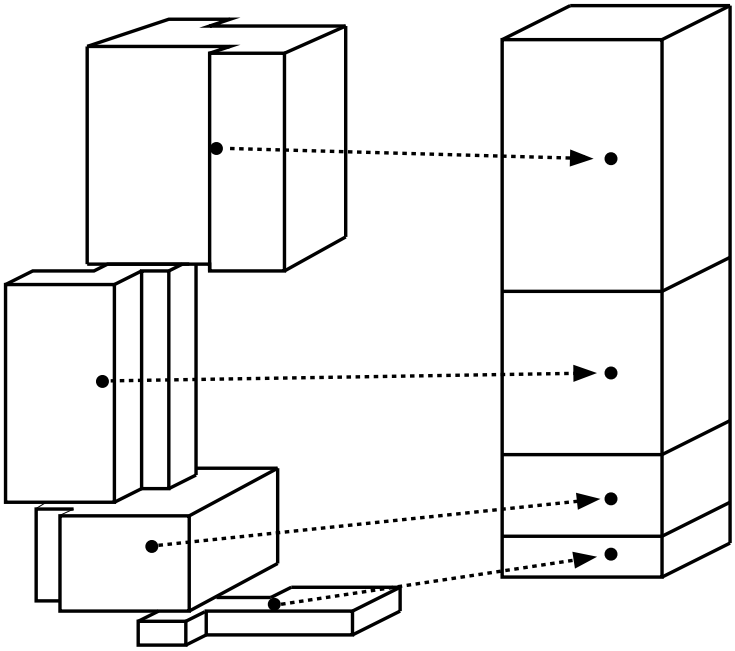
An inherently mass-conservative semi-implicit semi-Lagrangian model

by

Peter Hjort Lauritzen



ph.d. thesis

Department of Geophysics
University of Copenhagen
Denmark

September, 2005

Acknowledgments

This dissertation is a result of a collaboration between the University of Copenhagen and the Danish Meteorological Institute (DMI) under the Copenhagen Global Change Initiative (CoGCI). In addition to the research a CoGCI Ph.D. student must enroll in course activities corresponding to one semester of full-time study and be a teaching assistant in three university courses (the duration of each course is one semester and the assistant teaches two hours a week). The Ph.D. started in September 2002 and in the following three years a number of people contributed to this research.

First of all, I would like to thank my day-to-day advisers at DMI, Dr. Kaas and Dr. B. Machenhauer, for their energetic and collegial type of supervision on this project. Also thanks to my University adviser Prof. J.Ray Bates who unfortunately retired from the Department of Geophysics during this Ph.D. work. His overview and experience of the research topic helped shape the project, keep it doable and within the scope of a Ph.D. thesis. Dr. A.W. Hansen took over the supervision after Prof. J.Ray Bates had retired.

Part of this research was carried out while visiting the National Center of Atmospheric Research (NCAR) in Boulder, U.S.A. I am grateful to NCAR staff, in particular Dr. R. Nair, Dr. D. L. Williamson and Dr. C. Jablonowski, for helpful suggestions and useful discussions. Their motivation greatly accelerated the work on this thesis. The research carried out at NCAR resulted in an article accepted for publication in Monthly Weather Review.

The code development would not have been possible without the technical assistance of Dr. X. Yang and Dr. O.B. Christensen. Also thanks to Dr. K. Lindberg for discussions on the existing dynamical core of the Danish forecast model, and many thanks to Dr. R. Graversen and Dr. W. May for reading through parts of the manuscript.

Last, but not least, I would like to thank my future wife Silvia Agnona for her motivation and understanding during the past few years. My parents, Lone and Jens, and brother, Rasmus, receive my gratitude for supporting my choice of studies although they probably wonder what it is I am really doing !

Peter Hjort Lauritzen
Niels Bohr Institute
Department of Geophysics
University of Copenhagen
Denmark

Abstract

A locally mass-conservative dynamical core using a two-time-level, semi-Lagrangian semi-implicit integration scheme is presented. First a shallow water model is developed and tested, where after the approach is extended to a three-dimensional hydrostatic model.

The momentum equations are solved with the traditional semi-Lagrangian grid-point form. The explicit continuity equation is solved using a cell-integrated semi-Lagrangian (CISL) scheme and the semi-implicit part is designed such that the resulting elliptic equation is on the same form as for the traditional semi-Lagrangian grid-point system.

The accuracy of the shallow water model is assessed by running standard test cases adapted to a limited area domain. The accuracy and efficiency of the new model is comparable to traditional semi-Lagrangian methods and is not susceptible to noise problems for high Courant number flow over orography.

The shallow water model is extended to a baroclinic model by using a hybrid trajectory algorithm which is backward in the horizontal and forward in the vertical. In addition, the vertical part of the trajectory scheme is consistent with the discretized explicit CISL continuity equation. Since the vertical part of the trajectory is forward, the cells depart from model levels making the upstream integral two-dimensional and existing CISL schemes directly applicable. A price to pay is that the prognostic variables must be mapped back to the model grid at each time step. The problem is, however, one-dimensional.

Two model versions are derived. In the first version the thermodynamic equation is discretized as in traditional semi-Lagrangian models (only adapted to the hybrid trajectory). In the other model version the conversion term in the thermodynamic equation is discretized consistently with the semi-implicit CISL continuity equation. In both model versions the momentum equations are discretized in grid-point form.

The new dynamical cores are inherently mass conservative and can perform consistent online transport of tracers. They are implemented within the framework of HIRLAM and are tested using the Jablonowski-Williamson idealized baroclinic wave test case. The new dynamical cores run stably with long time steps and without the need for decentering or filtering of the non-linear terms in time as is needed in HIRLAM. In less active parts of the domain HIRLAM is noisy whereas the cell-integrated models produce smooth solutions. Compared to HIRLAM the baroclinic development is equally or more intense with the cell-integrated models. The CISL model version using a consistent energy conversion term has the strongest baroclinic development.

Table of Contents

Chapter I

Introduction

The science of climate change and weather prediction makes massive use of numerical models as they are far the most important tools for quantitative predictions. Despite the increasing use and confidence in models their projections are subject to significant uncertainties and the research efforts within the modeling community strive toward a common goal: *to reduce the uncertainties*. The methods used to aim at this goal are diverse, ranging from understanding the dynamics and physics of the climate system using observations and simulations to improving the numerics, "initialization" and data assimilation in existing models. The method considered in this thesis is model improvement in terms of the numerical methods rather than understanding the dynamics of the climate; why the numerical methods should be improved for more accurate projections is discussed in detail in the following sections.

State-of-the-art general circulation models (GCMs, see Appendix C for a complete list of acronyms) involve the atmosphere, ocean, biosphere, land-surface processes etc. and obviously GCMs engage a wide range of scientific disciplines. The monumental task of attempting to model a system as complex as the climate system is brought about by dividing the latter system into modules associated with different components of the system, and then simplifying them in terms of space and time resolution. For example, a crude division of the climate system could be to separate it into a module for the atmosphere, ocean and cryosphere, respectively. Single modules are less comprehensive than the full system and are therefore more tractable. Hence model development is often confined to a single module or even only a part of a module so that initially the complex interaction between model components is avoided. Only when each module is tested on its own they are coupled to produce a complete model. This study is confined to the atmospheric module.

The atmosphere is typically represented with two sub-modules: One representing the dry and adiabatic atmosphere and one the diabatic processes. The latter sub-module is referred to as the physics module. This separation enables the developer to focus on the adiabatic solution without complex interactions with the physics package and other modules which make it difficult to determine cause and effect of model behavior. The sub-module dealing with the numerical solution to the dry, adiabatic, primitive equations of the atmosphere is referred to as the *dynamical core* and is found in the "heart" of every atmospheric GCM. This thesis focuses on this sub-module. Of course by not including the physics package a major cause of uncertainty in atmospheric GCMs is not considered; however, as will be discussed in the following, the dynamical core is worth of more attention before improving and tuning the parameterizations. The argumentation is initiated by considering only a small part of the dynamical core: tracer advection or equivalently the solution of the continuity equation for a tracer. This will lead to consideration of the full dynamical core.

1.1 Motivation

1.1.1 *Accurate transport of tracers*

The accurate transport of tracers is important for a wide range of applications such as data assimilation, pollution, and earth system modeling. Here the discussion focuses on tracer transport and the atmospheric climate, but the problems encountered here do in principle also apply to a much wider set of applications.

A significant part of the uncertainty in the estimation of past and future climates is due to changes in greenhouse gases and other atmospheric constituents (e.g., Watterson and Dix 2005). Changes in constituent concentrations induce atmospheric and surface feedbacks on the radiative forcing, which is a key component of the climate system. Obviously, an understanding of the distribution and fluxes of various atmospheric trace constituents is needed. This involves proper knowledge of atmospheric transport as well as the relevant physical and chemical transformation and deposition processes for the trace species considered.

The transport by the winds is the dominant process for so-called "long-lived tracers" such as nitrous oxide (NO_x), methane (CH_4), and the chlorofluorocarbons (CFCs) that have lifetimes on the order of years in the troposphere and lower stratosphere. For long-lived tracers the choice of advection scheme used to solve the continuity equation for the constituent in question is decisive (e.g., Eluszkiewicz et al. 2000). For more reactive tracers, which have shorter lifetimes, the chemical and physical parameterizations make it difficult to identify the numerical errors introduced by the advection scheme. It is likely that the parameterizations reduce the dependency on the advection scheme, but it is evident that an inaccurate transport scheme provides erroneous data for the parameterizations. Even for short lived tracers strong numerical dispersion is a key problem if the spatial gradients are large, e.g. O_3 near the tropopause.

The problem of performing accurate tracer transport or, equivalently, the approximation of the solution to the continuity equation for a given tracer has received considerable attention in the literature for many decades. For the development of advection schemes for use in atmospheric models Rasch and Williamson (1990) have defined seven properties which a transport scheme should posses. Among these are: the method should be local, transportive, monotonic, computationally efficient and conserve as many analogs of invariants of the continuous equations as possible. It is widely accepted within the meteorological community that a transport scheme should fulfill these requirements. Especially the first moment invariant, mass conservation, is regarded as being increasingly important the longer the simulation.

Mass conservation Apart from being an extremely useful property for eliminating errors during model implementation, the question is how important mass conservation is. As an example consider traditional semi-Lagrangian models (see below) which are well known for being non-

conservative. The use of such models to simulate the climate may result in a drift in the global mass field; e.g., Moorthi et al. (1995) found a monotonic increase in the global mean surface pressure of 37 hPa in a 17-month integration using a semi-Lagrangian model. Over time this error will accumulate and cause a significant drift in the global mass field. To restore global mass conservation *ad hoc a-posteriori* algorithms are used (e.g., Priestley 1993, Gravel and Staniforth 1994, Bermejo and Conde 2002). The simplest, and quite common, mass-restoration algorithm is to periodically add or subtract the same amount of mass at every point in the model domain so that global mass conservation is regained. The advantage of this algorithm is that the pressure gradient force is not affected and thus spurious gradients in the pressure field are not introduced. Moorthi et al. (1995) concluded that the simulated climate is not affected by this mass-restoration algorithm but the mass-restoration algorithm was also designed to have a minimal effect on the dynamics. However, the algorithm does introduce an arbitrary long-range transport of mass since the same amount of mass is added or subtracted everywhere. Several more sophisticated and more local algorithms have been developed. There is, however, a degree of arbitrariness in the way these "mass-fixing" algorithms repeatedly correct the total mass without ensuring the fulfillment of the continuity equation for individual grid cells. In other words, the mass-fixing algorithms ensure global but not local mass conservation.

While several schemes have been developed with the seven desirable properties suggested by Rasch and Williamson (1990) in simple test settings, their application in complete model systems may deteriorate some of the desirable properties such as conservation and monotonicity. For example, a so-called inherently-mass-conservative (IMC) advection scheme may not conserve mass when applied in typical model settings where pressure is the vertical coordinate. The nature of the problem is described in detail in the next section and is surprisingly not mentioned frequently in the literature.

The mass-wind inconsistency Consider the problem of advecting a tracer. In other words, the continuity equation for the tracer in question must be solved. As mass conservation is important, a so-called IMC numerical method is used for the continuity equation. Most atmospheric models use a pressure-based vertical coordinate. Thus the prognostic variable of the IMC transport scheme is usually $q\,\Delta p$, where Δp is the pressure level thickness (the horizontal area element has been omitted) and q is the mixing ratio. In order to solve the tracer transport equation the wind field and the pressure level thicknesses must be specified. If these are given at every time step and on the same grid as used for the tracer advection, the advection is performed *online*. This is typically the situation when performing tracer advection in an atmospheric GCM. If the wind and pressure data are not given at every time step and maybe on another grid, so that both interpolation in time and space is needed, then the advection is performed *offline*. This is typically the situation in a chemical transport model (CTM). In either situation the consistency between the wind and mass fields is not necessarily guaranteed. What is the reason for this ?

If $q = 1$ everywhere, the tracer transport scheme effectively solve the continuity equation for air as a whole. Hereby the pressure levels implied by the advection scheme are obtained. The problem

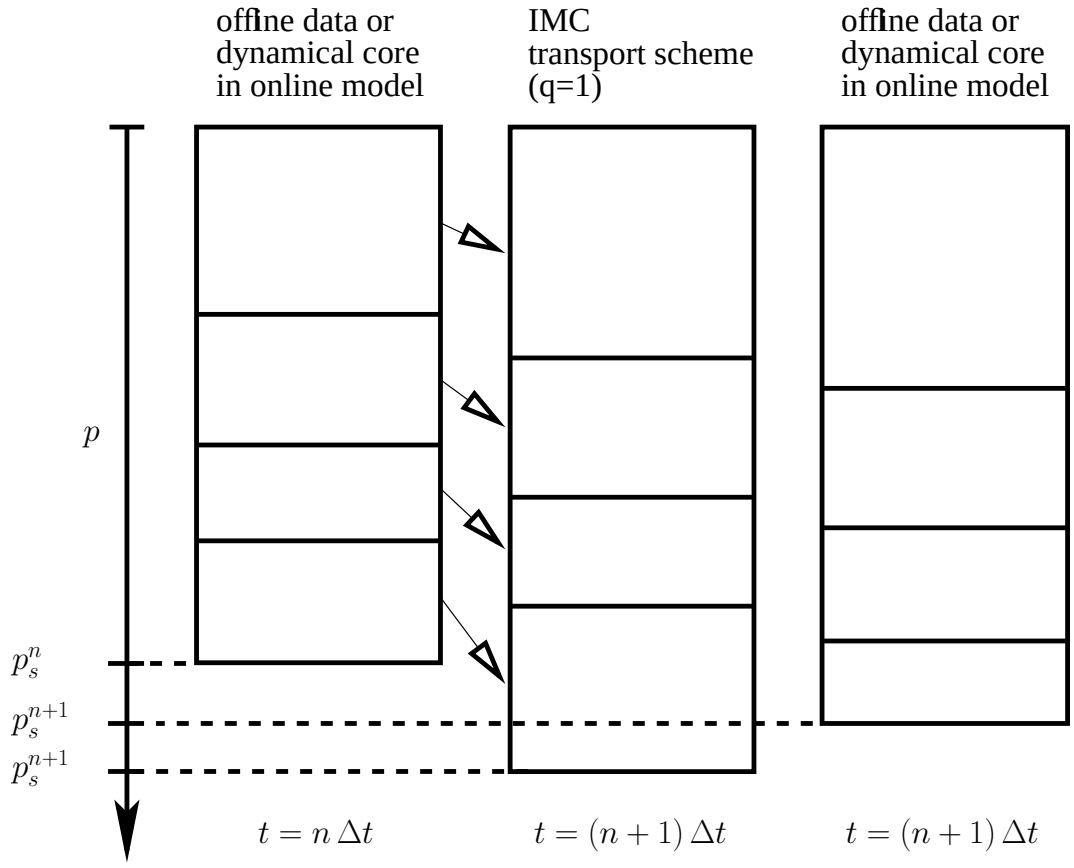


Fig. 1: Graphical illustration of the mass-wind inconsistency. The figure shows the location of pressure levels at the beginning of a time step $t = n\,\Delta t$ (left), after one time step $t = (n+1)\,\Delta t$ using an IMC transport scheme (middle) and given by *offline* data or predicted by the continuity equation of the dynamical core (right), respectively. If the vertical levels implied by the transport scheme and the dynamical core or *offline* data do not coincide an inconsistency between the mass and wind fields exists and affects the accuracy of the tracer advection.

is that the pressure levels implied by the *offline* or *online* data do not necessarily coincide with the vertical levels implied by the tracer transport scheme when setting $q = 1$ everywhere (see Fig. 1). In other words, there is a lack of consistency between the mass and wind fields. In an *offline* setting assimilated analysis or re-analysis data is often used to provide the horizontal wind $\mathbf{v}$ and Δp to the transport scheme. The data must be spatially and temporally interpolated to accommodate the grid and time step used by the transport scheme. In such a situation there is no consistency between the mass and wind fields unless *a posteriori* consistency correction methods are applied to the *offline* data. Performing the transport *online* increases the internal consistency, but it does not necessarily guarantee it. A typical setting is that a spectral dynamical core provides $\mathbf{v}$ and Δp to a so-called IMC transport scheme. However, in this situation the mass-wind inconsistency is present as well. When the continuity equation is solved using the spectral method it provides a solution different

from the one computed by the transport scheme with $q = 1$. Hence the pressure levels implied by the spectral dynamical core are different from the ones implied by the IMC transport scheme. In addition, if the transport scheme uses a time step different from the dynamical core (and maybe a different grid), a temporal (and spatial) interpolation is needed. These interpolations may also contribute to the lack of consistency between the mass and wind fields.

The consequences of the mass-wind inconsistency in long simulations can be severe. An IMC transport scheme predicts $\Delta p\, q$, but often the mixing ratio q is needed (for example for parameterizations). During the conversion from tracer mass $\Delta p\, q$ to tracer mixing-ratio q, mass conservation is lost if Δp provided by the *offline* or *online* data is used. Jöckel et al. (2001) ran a low resolution transport model using a so-called IMC advection scheme. For passive tracers initialized at different locations in the atmosphere the variations in the total mass were up to 70 % in a one year simulation. The inconsistency especially affects the vertical transport around the tropopause. If no inherent consistency exists one can attempt to restore the mass by altering the mixing ratios *a posteriori*. Jöckel et al. (2001), who discuss the mass-wind inconsistency in detail, investigated the effects of various mass-fixer algorithms. But all fixers have severe disadvantages such as violation of shape preservation or introduction of non-physical transport components. Alternatively, one could divide by Δp forecasted by the tracer advection scheme with $q = 1$, instead of Δp provided by the *offline* or *online* data. But in this case the fields start to develop independently; $\Delta p\, q$ develops according to the advection scheme and is never "synchronized" with Δp from the *offline* or *online* data, which is consistent with the horizontal wind field used by the advection scheme. The error therefore accumulates.

Instead of altering mixing ratios one can adjust the velocity field such that the tracer advection equation is consistent with the mass field, i.e. the winds are corrected so that the vertical integrated divergence of mass matches the surface pressure tendency (Cameron-Smith et al. 2002). This restoration algorithm is referred to as a pressure fixer. This approach is not completely satisfactory either since "true" wind data are modified to provide mass-wind consistency. It can, however, be used to indicate how severe the mass-wind inconsistency problem is, i.e. by running a model with and without a pressure fixer. Horowitz et al. (2004) have run the global Model of Ozone Research version 2 (MOZART-2) with and without a pressure fixer. Near the tropopause (where the vertical gradient of the ozone-mixing ratio is large) the difference between the two runs was approximately 187 Tg yr^{-1}. Assuming that the pressure fixer is perfect, it can be estimated that a spurious source of ozone of 187 Tg yr^{-1} is caused by the mass-wind inconsistency problem. This is not a negligible amount. For example, the spurious source of ozone is equal to the estimated amount of influx to the troposphere from the stratosphere in the northern extratropics. This is, of course, only an indication of the magnitude of the problem. In order to estimate the systematic spurious sources and sinks a fully consistent model must be run but the estimates provided by Horowitz et al. (2004) suggest the mass-wind inconsistency can introduce significant errors. Inconsistent advection schemes that fail to conserve mass locally introduce spurious instantaneous sources and sinks which erroneously alter the mixing ratios. This is probably a serious problem since mixing ratios are important for chemical reactions.

As discussed above, the problem of performing accurate tracer advection in a model using a pressure based vertical coordinate is not limited to the tracer transport scheme, but also to the consistency between the scheme and the wind and pressure data supplied to it. In an *offline* setting there is little choice but to use some kind of *a posteriori* correction method. But in an *online* transport model the consistency can be guaranteed if the same numerical method is used for the continuity equation of the dynamical core as for tracer advection. This is, however, much easier said than done in the majority of model settings since their dynamical cores are based on non-IMC schemes (e.g., traditional semi-Lagrangian models such as in the IFS[1], HIRLAM[2] and ECHAM5[3]). Changing the numerical method for the continuity equation for air as a whole impacts all other prognostic equations used in a model. Therefore all discretization in the model as a whole must be carefully rethought if inherent mass conservation is to be introduced.

1.1.2 *Semi-Lagrangian models*

Numerical models can be divided into two categories: Eulerian and semi-Lagrangian. Most Eulerian models solve the continuity equation for air using an IMC method and can therefore perform consistent *online* transport of tracers. The mass of air and atmospheric constituents is conserved. However, many Eulerian models have high numerical dispersion (this includes spectral models when the flow is non-linear) and as a consequence non-local features develop. On the contrary, the traditional semi-Lagrangian models have small dispersion errors, but are based on non-conservative discretizations and therefore do not conserve the mass. When applying an IMC advection scheme for tracer transport, the mass and wind fields provided to the transport scheme by the semi-Lagrangian model are not consistent with the IMC advection scheme. Despite this trade-off, semi-Lagrangian models have become very common operational models since they run several times faster than their Eulerian counterparts while retaining similar accuracy. The recognition of the mass-wind inconsistency problem as being a potential source of serious errors, and that semi-Lagrangian models are desirable but not IMC, motivated this thesis work. An elaboration of the latter motivation constitutes the following paragraphs.

During the last 20 years semi-Lagrangian transport (SLT) schemes have become widely adopted in atmospheric models because they offer certain advantages over other advection schemes (Staniforth and Côté 1991). Semi-Lagrangian schemes are not limited by the Courant-Friedrichs-Levy (CFL) time-step condition. They can therefore be used in high-resolution models where gravity and sound waves, if included, are treated semi-implicitly without the need for a very short time step. The long time steps allow models to include sub-grid-scale processes, which are computationally expensive, without depleting computational efficiency. Part of the cost of SLT schemes is the calculation of upstream trajectories. As this is done only once at every time step, the SLT schemes become relatively cheaper for a large number of tracers. In addition to the long time step

[1] White (2001)
[2] McDonald (1994)
[3] Roeckner et al. (2003)

advantage, SLT schemes minimize computational dispersion, can handle sharp discontinuities, and the vertical advection is formally more accurate than finite difference operators.

As any other numerical method the semi-Lagrangian approach is not perfect. The principal problem is that the SLT schemes currently used do not conserve mass, which can result in a significant drift in the global mass fields. As discussed previously, *ad hoc a-posteriori* algorithms may be employed to restore global mass conservation, but there is a degree of arbitrariness in the way these "mass-fixing" algorithms repeatedly correct the total mass without ensuring the fulfillment of the continuity equation for individual grid cells. It is therefore not only mathematically more rigorous, but also desirable to use IMC methods.

The application of traditional semi-Lagrangian methods in shallow water and baroclinic models have shown that for high Courant number flow over orography semi-Lagrangian models encounter problems; a spurious numerical resonance may develop. The problem has been studied in detail and sought reduced by using spatial decentering. The decentering reduces the noise to an acceptable level, but it does also seem to reduce the accuracy of the semi-Lagrangian method. It can easily be verified that the formal accuracy of the semi-Lagrangian method is reduced by using decentering. Also in simulations it has been shown that the forecast skill is reduced by applying decentering (e.g. Lauritzen et al. 2005). Clearly, modelers would rather do without decentering and use other methods to control noise and help reduce the orographic resonance problem.

In traditional semi-Lagrangian models the equations of motion contain terms which are discretized in an Eulerian fashion. For example, the discretized continuity equation is based on a formula containing Eulerian advection terms $\mathbf{v}\nabla \ln p_s$ (where p_s is the surface pressure). This is somewhat inconsistent with the semi-Lagrangian method, which is based on the discretization of the total derivative $d/dt = \partial/\partial t + \mathbf{v}\nabla$. Similarly, the approximation of the conversion term in the thermodynamic equation ω/p, where ω is the vertical velocity in pressure coordinates, is somewhat inconsistent; a diagnostic expression for ω/p is obtained by use of the continuity equation discretized in an Eulerian fashion. However, in a semi-Lagrangian model it would be more consistent to base the discretized vertical velocity on $\omega = dp/dt$ rather than the Eulerian formula. Similarly, the continuity equation can be discretized on a different form so that Eulerian advection terms in the discretized formulas can be avoided. The reason for discretizing terms in the equations of motion in an Eulerian fashion is probably to keep some similarity to Eulerian models.

Since semi-Lagrangian methods are widely used and recognized as accurate and very efficient, it would be worth a significant research effort to add inherent mass conservation to the long list of desirable properties these methods already possess. While reformulating the numerical scheme it would be opportune to introduce a more semi-Lagrangian discretization of the equations of motion as discussed in the preceding paragraph.

1.2 Research questions

A number of IMC SLT schemes, also referred to as cell-integrated semi-Lagrangian (CISL) schemes, have recently been developed for two-dimensional transport on the sphere (e.g., Lin and Rood 1996; Nair and Machenhauer 2002; Nair et al. 2002; Zerroukat et al. 2004b). These schemes do not only conserve mass locally but are also transportive, efficient, accurate, and have shape-preserving options. So far, however, little work has been done to extend these methods to the full dynamical equations of a global model.

To exploit the efficiency of semi-Lagrangian methods it is important to use long time steps, which requires a semi-implicit treatment of the fast waves. To date the CISL schemes have not been coupled with a semi-implicit time stepping in a full model. This is a necessary thing to do, otherwise the resulting semi-Lagrangian model will not be computationally competitive.

It is the goal of this thesis to design a baroclinic IMC semi-implicit semi-Lagrangian (SISL) model that can perform consistent *online* tracer transport. This requires the use of the same IMC SLT scheme for the continuity equation for the air as well as for the tracers, and the development of a semi-implicit time-stepping method for the CISL scheme. As mentioned above, when introducing a new scheme for the continuity equation of the dynamical core the discretizations of all model equations must be carefully rethought. It is this process which makes up the thesis.

The research has been performed within the framework of HIRLAM and therefore some similarity to the HIRLAM system is kept in the new models. In particular, the testing of the new CISL models is limited to non-global domains. The long-term goal, however, is to develop a global model and the extension of the limited area models to a global domain is considered on a theoretical level. The implementation of such a model is, however, beyond the scope of this thesis.

The main research questions addressed in this thesis are all the problems resulting from introducing CISL schemes in full models. They are briefly stated as:

- The CISL transport schemes have been developed and tested on the sphere, and found efficient and accurate. Is it possible to derive a semi-implicit CISL scheme for the continuity equation of the dynamical core in atmospheric models ?
- Traditional semi-Lagrangian grid-point methods are known to produce noise for high Courant number flow over mountains. To alleviate noise problems a decentering can be used at the cost of reduced accuracy. It would be desirable to develop a semi-Lagrangian method which is not susceptible to the noise problems. Are semi-implicit CISL models susceptible to the orographic resonance problem ?
- It appears inconsistent that traditional semi-Lagrangian models are designed only to differ from Eulerian ones in the advection part. Thereby some terms in the equations of motion are treated in an Eulerian fashion: in particular, the conversion term in the thermodynamic equa-

tion, the vertical part of the trajectory scheme and the right-hand side of the continuity equation. It would be desirable to discretize the equations of motion in a more semi-Lagrangian fashion. Is this possible ?

- Is it worthwhile in terms of accuracy and computational efficiency to introduce inherent mass conservation in semi-Lagrangian dynamical cores ?

1.3 Overview of the thesis

Chapter II The basis building block of the models derived here is the semi-Lagrangian method, which has been thoroughly reviewed in Staniforth and Côté (1991). Since the last review, however, many exciting new developments have been published in the meteorological literature and have not been reviewed yet[4]. In Chapter II an overview of the semi-Lagrangian method is given: various trajectory schemes are discussed before considering semi-Lagrangian solutions to the advection problem. Both newer traditional and the recently developed CISL schemes are discussed in detail. The application of the semi-Lagrangian method to the full system of equations is discussed in connection with the baroclinic model in Chapter IV.

Chapter III Shallow water models are the simplest possible models that capture the essential features of large-scale geophysical flows and the numerical difficulties encountered here are similar to those associated with the horizontal part of three-dimensional applications. Hence a first step toward the goal of developing an IMC semi-Lagrangian semi-implicit dynamical core for weather and climate models is to derive a shallow water model with the desired properties. In Chapter III a semi-implicit CISL shallow water model is described. The research is performed within the framework of HIRLAM and thus the shallow water model is implemented for a limited area domain. The accuracy of the new model is assessed using standard test cases adapted to a limited area on the sphere.

Chapter IV The next step in the development is the generalization of the limited area CISL shallow water model presented in Chapter III to three dimensions. In Chapter IV a hydrostatic semi-implicit CISL model is derived. Changing the scheme used for the continuity equation of the dynamical core does, of course, also affect how the other equations are discretized. A discretization which more consistently enforces semi-Lagrangian forms is derived. The new dynamical cores are tested using the recently developed Jablonowski-Williamson baroclinic wave test case.

[4] a review is currently under way (Machenhauer et al. 2005)

Chapter V Chapter V provides a summary of the scientific merit and major accomplishments of this thesis. Future research directions are proposed based on the advances presented here.

Chapter II

Overview of the semi-Lagrangian method

The equations of motion for the weather and climate system can be derived in either a Lagrangian form or an Eulerian form. In an Eulerian system the observer describes the evolution of the flow from a fixed point in the coordinate system, whereas Lagrangian equations describe the evolution of the flow that would be observed following the motion of each individual fluid parcel. The Eulerian method has the advantageous property of having a regular mesh throughout the integration, but it often suffers from overly restrictive time step limitations. The Lagrangian method is normally not subject to strict time step limitations, but introduces another problem: an initially regular mesh will quickly evolve into an irregular mesh with high concentration of mesh points in convergent areas and low concentration of mesh points in divergent areas. Thus the accuracy of the method will be lower in divergent areas compared to convergent areas of the domain. In order to avoid the problems of non-uniform meshes and at the same time allow long time steps, one can periodically map the distribution from the irregular (Lagrangian) mesh back to the regular (Eulerian) grid and then start all over. Instead of following the same set of fluid parcels during the entire integration, a new set of fluid parcels are chosen after each time step. Hereby the distribution of parcels can be kept quasi uniform throughout the integration. This method is called the *semi-Lagrangian* method.

2.1 Trajectory determination

The first step of the semi-Lagrangian method is to determine the parcel trajectories. Mathematically this corresponds to the integration of the first-order ordinary differential equation

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}(\mathbf{r}, t), \tag{1}$$

where $\mathbf{v}$ is the velocity vector $\mathbf{v} = (u, v)$ and $\mathbf{r}$ is the displacement vector. The kinematic relation (1) is among the prognostic equations of a semi-Lagrangian model. The accuracy of the method used for computing trajectories is crucial for the overall accuracy of the model. In particular, the use of long time steps emphasizes the need for reducing time-truncation errors associated with the scheme.

The discussion here is restricted to two-time-level schemes and only backward trajectories are discussed, i.e. consider parcels which depart from time level (n) and arrive at grid points at time level $(n + 1)$. For simplicity assume Cartesian geometry. The spatial location of the departure point is specified with subscript $*$ so $\mathbf{r}_*^n$ denotes the location of the departure point. The departure point is given by

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \int_{n\,\Delta t}^{(n+1)\,\Delta t} \mathbf{v}(\mathbf{r}, t)\, dt. \tag{2}$$

The integral on the right-hand side of (2) is approximated using velocities at time levels (n) and $(n - 1)$.

The simplest algorithm for finding the backward trajectory is Euler's method

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \mathbf{v}^n(\mathbf{r}^{n+1})\, \Delta t, \tag{3}$$

where $\mathbf{v}^n(\mathbf{r}^{n+1})$ refers to the velocity at time level (n) evaluated at the arrival point. Euler's method is first-order accurate in time, which can lead to large truncation errors especially when used with large time steps. (Robert 1981) found that the time truncation error should be no worse than $\mathcal{O}(\Delta t^2)$ to keep the overall errors at an acceptable level.

A variety of $\mathcal{O}(\Delta t^2)$ schemes have been proposed in the literature. A very popular scheme is the second-order implicit midpoint method where

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \tilde{\mathbf{v}}^{n+1/2}\left(\frac{\mathbf{r}^{n+1} + \mathbf{r}_*^n}{2}\right)\Delta t, \tag{4}$$

is iterated with (3) as a first guess (McDonald and Bates 1987, Temperton and Staniforth 1987). The velocity field is first extrapolated to time level $(n+1/2)$, $\tilde{\mathbf{v}}^{n+1/2}$, and thereafter interpolated to the approximate midpoint of the trajectory, $(\mathbf{r}^{n+1} + \mathbf{r}_*^n)/2$. Linear interpolation in space is found to be sufficient and typically a few iterations of (4) are needed (e.g. Staniforth and Côté 1991).

The extrapolation in time for obtaining the middle point of the semi-Lagrangian trajectory has, however, been identified as a potential source of instability. Non-meteorological noise has been observed in forecasts using two-time-level semi-Lagrangian models at a number of meteorological centers, e.g., in the Aire Limité Adaptation dynamique Développement InterNational/Limited Area Central European (ALADIN/LACE) model (Gospodinov et al. 2001), HIRLAM (McDonald 1999) and IFS model (Hortal 2002). This has forced modelers to rethink the design of 2TLSL models. Here we focus on the trajectory computations.

McDonald (1999) successfully removed the noise in forecasts made with HIRLAM by using an alternative approximation to the extrapolated velocity on the right-hand side of (4). The velocity extrapolated to time level $(n+1/2)$, $\tilde{\mathbf{v}}^{(n+1/2)}$, can be written as a linear combination of the known velocities at time levels (n) and $(n-1)$ evaluated at the arrival point $\mathbf{r}^{n+1}$, departure point $\mathbf{r}_*^n$ and $\mathbf{r}^{n+1} - 2\,\Delta t\,\mathbf{v}_*^n$, respectively:

$$\tilde{\mathbf{v}}^{(n+1/2)} = \sum_{\mu=1}^{6} \mathbf{w}_\mu\,\chi_\mu, \tag{5}$$

where

$$\begin{aligned}
\chi_1 &= \mathbf{v}^n(\mathbf{r}^{n+1}),\\
\chi_2 &= \mathbf{v}^{n-1}(\mathbf{r}^{n+1}),\\
\chi_3 &= \mathbf{v}^n(\mathbf{r}_*^n),\\
\chi_4 &= \mathbf{v}^{n-1}(\mathbf{r}_*^n),\\
\chi_5 &= \mathbf{v}^n\left[\mathbf{r}^{n+1} - 2\,\Delta t\mathbf{v}^n(\mathbf{r}^{(n+1)})\right],\\
\chi_6 &= \mathbf{v}^{n-1}\left[\mathbf{r}^{n+1} - 2\,\Delta t\mathbf{v}^n(\mathbf{r}^{(n+1)})\right]
\end{aligned}$$

and $\mathbf{w}_i$ are weights. Requiring second-order temporal accuracy imposes constraints on the weights $\mathbf{w}_i$. Three free parameters result. Several schemes proposed in the literature belong to this family

of schemes, for example the *extrapolation along the trajectory* scheme[5] (equation 42 in Temperton and Staniforth 1987) and the more economical scheme described in Hortal (1998)[6]. By running the forecast model and measuring the level of noise McDonald found that the first *extrapolation along the trajectory* scheme produces almost noise-free forecasts, while the Hortal scheme only reduces the noise. The noise could be reduced even further by an "optimal" choice of the free parameters. The optimal scheme, measured in terms of noise, used χ_1, χ_3, χ_4, χ_5 and χ_6 for the approximation of the trajectory.

By using more values for the approximation of $\tilde{\mathbf{v}}^{n+1/2}$ without increasing the formal accuracy, the approximation for $\tilde{\mathbf{v}}^{n+1}$ is somewhat smoothed. McDonald's scheme seems to remove the noise because of the smoothing. On the other hand a smoothing could lead to a decrease in accuracy. McDonald considered only noise levels in his study and did not report the forecast accuracy in terms of verification scores. So it is an open question how the overall accuracy of the forecast is affected by this trajectory scheme that produces practically noise-free forecasts. For example, Hortal (2002) reported no noise problems when using the *extrapolation along the trajectory* scheme, but got much worse verification scores.

All departure point algorithms mentioned so far assume straight line trajectories in a "space-time diagram", i.e. they do not take the acceleration into account. When using long time steps it would be desirable to include the accelearation in the trajecory estimation. Several trajectory schemes have been proposed in the literature.

The location of the departure point $\mathbf{r}_*^n$ can be written in terms of a Taylor series expansion about the departure point

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \sum_{\nu=1}^{N} \frac{(\Delta t)^\nu}{\nu!} \frac{d^\nu \mathbf{r}_*^n}{dt^\nu}, \tag{6}$$

or about the arrival point

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} + \sum_{\nu=1}^{N} \frac{(-\Delta t)^\nu}{\nu!} \frac{d^\nu \mathbf{r}^{n+1}}{dt^\nu}, \tag{7}$$

where d/dt is the material derivative (also called total or Lagrangian derivative)

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y}. \tag{8}$$

By including more terms in the Taylor series expansion the trajectory is no longer a straight line in a "space-time diagram" where time is plotted on the y axis and departure point distance on the x axis. The question is how to approximate the derivatives $\frac{d^\nu \mathbf{r}}{dt^\nu}$, $\nu = 2, .., N$.

Motivated by noise problems in the operational two-time-level semi-Lagrangian (2TLSL) model at the European Center for Medium-Range Weather Forecasts (ECMWF), Hortal (2002) derived

[5] $\mathbf{w}_1 = 0$, $\mathbf{w}_2 = 0$, $\mathbf{w}_3 = \frac{3}{2}$, $\mathbf{w}_4 = 0$, $\mathbf{w}_5 = 0$ and $\mathbf{w}_6 = -\frac{1}{2}$.
[6] $\mathbf{w}_1 = \frac{1}{2}$, $\mathbf{w}_2 = 0$, $\mathbf{w}_3 = 1$, $\mathbf{w}_4 = -\frac{1}{2}$, $\mathbf{w}_5 = 0$ and $\mathbf{w}_6 = 0$.

an improved trajectory scheme called 'Stable Extrapolation Two-Time-Level Scheme' (SETTLS). (6) was chosen as the basis for the trajectory scheme with

$$\frac{d\mathbf{r}_*^n}{dt} = \mathbf{v}_*^n,$$
$$\frac{d^2\mathbf{r}_*^n}{dt} = \frac{\mathbf{v}^n(\mathbf{r}^{n+1}) - \mathbf{v}_*^{n-1}}{\Delta t}.$$

The choice for the approximation for the acceleration was chosen after exploring many possibilities and so that the scheme could also be used for the other prognostic equations without complicating the elliptic equation associated with the semi-implicit system. The formal temporal accuracy of the scheme is second order and the computational cost of the scheme is comparable to that of the classical mid-point method (4). The treatment of the right-hand side of the trajectory equation was also used for the non-linear terms on the right-hand side of the remaining equations of motion. The forecast skill was not effected negatively when switching from the mid-point method to the SETTLS and noise problems were reduced. During sudden stratospheric warmings, however, Hortal (2004) report "new" noise problems.

McGregor (1993) proposed to discard the Eulerian velocity change in the approximation of the derivative

$$\frac{d}{dt} \approx \mathbf{v} \cdot \nabla, \tag{9}$$

and to use the formula in which the Taylor series expansion is about the arrival point (7). By using (7) for the approximation the need for spatial interpolation is eliminated. Using McGregor's scheme the location of the departure point is given by

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \Delta t\, \tilde{\mathbf{v}}^{n+1} + \sum_{\nu=1}^{N-1} \frac{(-\Delta t)^{\nu+1}}{(\nu+1)!} \frac{d^\nu}{dt^\nu} \left(\tilde{\mathbf{v}}^{n+1}\right), \tag{10}$$

where

$$\frac{d}{dt} \approx \tilde{\mathbf{v}}^{n+1/2} \cdot \nabla, \tag{11}$$
$$\widetilde{(\cdot)}^{n+1} = 2\,\widetilde{(\cdot)}^{n} - \widetilde{(\cdot)}^{n-1}, \tag{12}$$

and the higher-order derivatives are defined recursively

$$\frac{d^\nu \mathbf{v}}{dt^\nu} = \frac{d}{dt}\left(\frac{d^{\nu-1}\mathbf{v}}{dt^{\nu-1}}\right) \quad \nu = 2, 3, .., N-1. \tag{13}$$

Note that the scheme does not involve iterations. McGregor's formulation is general in the sense that high-order terms can easily be included since they are defined recursively. The order of accuracy can therefore easily be increased. Note that the formal order of accuracy of iterative methods does not increase when increasing the number of iterations. Nair et al. (2003) reported better results with McGregor's trajectory algorithm compared to a Runge-Kutta scheme for advection

experiments in Cartesian and spherical geometry. The scheme has also been extended to spherical geodesic grids, where better results were obtained compared to the iterative midpoint method (Giraldo 1999). It is, however, not know if it alleviates noise problems in baroclinic models. A critic of the scheme is that it assumes $\partial \mathbf{v}/\partial t = 0$. That assumption is also used in the midpoint method and other iterative methods. The SETTLS scheme, however, does not assume that $\frac{\partial \mathbf{v}}{\partial t}$ is zero.

Instead of using higher-order trajectory schemes one could alternatively split the trajectory into k segments so that each sub trajectory takes $\Delta t/k$ (D.L. Williamson 2004, personal communication). For each segment a low order efficient scheme could be used. As far as the author is aware the method has not been tested in baroclinic models so far.

The noise problems encountered in 2TLSL models have emphasized the fact that the computation of the trajectories can not be studied as an isolated problem. The differential equation for the trajectories is part of the prognostic equations of the full model and it is important that it "interacts well" with the remaining model equations. For example, the SETTLS scheme was designed so that the right-hand side of the equations of motion was treated in exactly the same way as the right-hand side of the trajectory equation. There is a need to incorporate trajectories in a more consistent manner in semi-Lagrangian models instead of treating trajectory determination and the solution to the remaining equations of motion as two separate tasks. Trajectories are always computed using explicit schemes. In a SISL model one should ideally solve the trajectory equation using the same semi-implicit scheme,i.e. the divergence should be averaged along the trajectory. It does, however, seem difficult to design a method where the trajectory algorithm is included in the elliptic system of a SISL model. The issue of incorporating trajectories more into the dynamics and studying the effect of trajectories on model properties other than stability, has been very little discussed in the literature. Exceptions are (Staniforth et al. 2003) and Cordero et al. (2005) who studied the impact of trajectories on the dynamical equivalence between momentum and angular momentum formulations of the equations of motion, and the impact on trajectories on the vertical models in a non-hydrostatic column model, respectively. Clearly, more research is desirable in this area.

Having discussed trajectory algorithms published in the meteorological literature and their use in baroclinic models, the next step in the semi-Lagrangian method is the solution to the equations of motion. In the following the focus is on the continuity equation, but the treatment of the the other equations of motion is similar and is discussed in Chapter III and IV in connection with the shallow water equations and the primitive equations, respectively.

2.2 The form of the continuous equations

Consider the continuity equation on flux form and advective form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \, \mathbf{v}) = 0, \tag{14}$$

$$\frac{d \rho}{d t} + \rho \, \nabla \cdot \mathbf{v} = 0, \tag{15}$$

respectively, where $\mathbf{v}$ is the velocity field, and ρ is some mass-specific quantity. The continuity equation is alternatively called the transport or advection equation. The numerical solution to this type of equation is probably the most studied problem in computational fluid dynamics since it represents a fundamental property of fluid flow. The flux-form equation for the transport of a passive scalar is closely associated to (14)

$$\frac{\partial}{\partial t} (\rho \, q) + \nabla \cdot (\rho \, q \, \mathbf{v}) = 0, \tag{16}$$

where q is the concentration of the tracer per unit mass (also referred to as mixing ratio). The two flux-form equations (14) and (16) imply an advective form of the tracer transport equation

$$\frac{d \, q}{d t} = 0, \tag{17}$$

which simply states that q is conserved along characteristics of the flow. Both the flux form and advective form of the continuous equations are differential forms. Alternatively, the equations of motions can be written on integro-differential form which is the partial differential equation integrated over an infinitesimal volume. If the volume is stationary the form is Eulerian, while a Lagrangian form results when the volume moves with the flow. Integrating the flux-form continuity (14) and tracer advection equation (16) over a stationary volume δV, the Eulerian integro-differential form results

$$\frac{\partial}{\partial t} \left(\overline{\psi} \, \delta V \right) + \oint \oint \oint_{\partial V} \psi \cdot \mathbf{n} \, dS = 0, \quad \psi = \rho, \, \rho \, q, \tag{18}$$

where

$$\overline{\psi} = \frac{1}{\delta V} \iiint_{\delta V} \psi \, dV \tag{19}$$

is the average of ψ over δV, ∂V is the boundary of δV and $\mathbf{n}$ the normal vector to ∂V. The divergence theorem, also known as Gauss's theorem, has been used to convert the volume integral of the divergence term to a surface integral. In a Lagrangian method, where the volume moves with the flow $\delta V = \delta V(t)$, there is no flux through ∂V and hence the Lagrangian version of (18) is given by

$$\frac{d}{d t} \left(\overline{\psi} \, \delta V \right) = 0, \tag{20}$$

which simply states that $\overline{\psi}$ is conserved for a volume moving with the flow.

Numerical methods based on the integral of a conservation law over a volume are called *finite volume* (FV) methods. An alternative and perhaps more descriptive name also used in the literature is *cell-integrated* methods. Eulerian and Lagrangian FV discretizations of the continuity equation are based on (18) and (20), respectively. Since FV or *cell-integrated* methods are based on tracking the integral of ψ, in this case mass, they are locally conservative. In Eulerian FV methods the fluxes through the boundaries of cells are tracked and mass flowing out of a cell wall is gained in the neighboring cell, thus guaranteeing that no mass is lost or added. In a Lagrangian FV method the mass is tracked as it moves with the flow, and as long as all the volumes span the entire domain the mass is conserved both globally and locally. Here only semi-Lagrangian methods are reviewed. An updated review of Eulerian, as well as semi-Lagrangian, finite-volume methods is currently under way (Machenhauer et al. 2005 in prep.).

2.3 Requirements for transport schemes

Before discussing the many finite-volume schemes used in the atmospheric sciences, it is important to realize which properties a transport scheme ideally should posses. The equation subject to the toughest requirements is probably the continuity equation for noisy tracers such as moisture. Rasch and Williamson (1990) have defined seven widely accepted desirable properties for transport schemes: accurate, stable, computationally reasonable, transportive, local, conservative and shape-preserving.

Accurate The high accuracy property is, of course, the primary aim for any numerical method and all but the efficiency requirement are part of the overall accuracy. Note that for a flow with shocks or sharp gradients the formal order of accuracy in terms of Taylor series expansions does not necessarily guarantee a high level of accuracy. Part of the accuracy is also the rate of convergence of the numerical algorithm.

Widely used measures of accuracy in the meteorological community for idealized test cases, are the standard error measures l_1, l_2 and l_∞ (e.g., Williamson et al. 1992):

$$l_1(\psi) = \frac{I(|\psi - \psi_E|)}{I(|\psi_T|)}, \tag{21}$$

$$l_2(\psi) = \frac{\left\{I\left[(\psi - \psi_E)^2\right]\right\}^{1/2}}{\left\{I\left[(\psi_T)^2\right]\right\}^{1/2}}, \tag{22}$$

$$l_\infty(\psi) = \frac{\max\left[I(|\psi - \psi_E|)\right]}{\max\left[I(|\psi_T|)\right]}, \tag{23}$$

where $I(\cdot)$ denotes the integral over the entire domain, ψ is the numerical solution and ψ_E is the exact solution if it exists. In case an exact solution does not exist ψ_E is a high resolution

reference solution. l_1 and l_2 are measures for the global "distance" between ψ and ψ_E, and l_∞ is the maximum deviation of ψ from ψ_E over the entire domain. In addition to the measures l_1, l_2 and l_∞, the normalized maximum and minimum values are also used to indicate errors related to overshooting and undershooting.

Stable The stability property ensures that the solution does not "blow up" during the time of integration. Usually the stability of Eulerian methods is governed by the Courant-Friedrichs-Levy (CFL) condition, which in one dimension is given by

$$\max \left| \frac{u \, \Delta t}{\Delta x} \right| \leq 1, \tag{24}$$

where u is the velocity, Δt the time step and Δx the grid interval. Hence a fluid parcel may not travel more than one grid interval during one time step. This overly restrictive time step limitation is usually alleviated in Lagrangian methods and can be replaced by the less severe Lipschitz criterion for stability

$$\left| \frac{\partial u}{\partial x} \right| \, \Delta t < 1, \tag{25}$$

(Smolarkiewicz and Pudykiewicz 1992), which guarantees that parcel trajectories do not cross during one time step. Hence in semi-Lagrangian models the time step can be chosen for accuracy and not for stability.

Note that for global models the efficiency and stability of the schemes are often challenged by the convergence of the meridians near the poles. If conventional latitude-longitude grids are used, special care must be taken in the vicinity of the poles. The problem can also be tackled by using other types of grids that do not have these singularities or at least reduce the effect of them, for example the icosahedral-hexagonal grid used operationally by the Deutscher Wetterdienst (e.g., Sadourny et al. 1968; Williamson 1968; Thuburn 1997; Majewski et al. 2002) and the cubed sphere approach (Sadourny 1972). These grids are more isotropic than conventional latitude-longitude grids, i.e. all cells have nearly the same size contrary to latitude-longitude grids where the areas decrease as aspect ratios increase toward the poles. This effect can be alleviated by using a Gaussian reduced grid in which the number of longitudes decrease toward the poles. It is an important part of accuracy that the advection schemes can transport distributions across the poles without distorting them and without imposing severe time-step limitations.

Computationally reasonable Computing resources are not unlimited and, given the complexity of geophysical fluid dynamics, the algorithms should be computationally efficient in order to allow for high resolution runs and/or a large number of prognostic variables. Efficiency is, however, hard to measure objectively. One measure for the efficiency of an algorithm is the number of elementary mathematical operations or the total number of floating-point operations per second (FLOPS) used by the algorithm. The advantage of counting FLOPS is that it can be done without turning

the computer on and is therefore a machine independent measure. But the number of FLOPS only captures one of several dimensions of the efficiency issue. The actual program execution involves subscripting, memory traffic and countless other overheads. In addition different computer architectures favor different kinds of algorithms and compilers optimize code differently. Measuring efficiency in terms of the execution time on a specific platform can be misleading for a user on another computer platform. Weather prediction and climate models are executed on massively parallel computersm wherefore the efficiency is partly determined by the amount of communication between the nodes. Hence the parallel programmer is concerned about algorithms being local thus minimizing the need for communication between the nodes. Nevertheless the most important measure of efficiency is probably the level of simplicity of the algorithm.

Transportive, local, conservative and shape-preserving The transportive and local property guarantee that information is transported with the characteristics and that only adjacent grid values affect the forecast at a given point. Integral invariants of the corresponding continuous problem as well as the shape of the distribution for non-divergent flows should ideally be preserved in the numerical solution. If the velocity field is divergent the shape of the distribution may be altered in the form of new extrema. So in the divergent case the numerical scheme should reproduce the physical extrema without creating spurious numerical extrema.

Integral invariants should, of course, be conserved for any kind of flow. For long simulations the conservation properties become increasingly important as numerical sources and sinks can degrade the accuracy significantly over time (e.g., Moorthi et al. 1995). Hence for climate models the finite-volume methods are very attractive given their inherent conservation properties.

In addition to the seven desirable properties of Rasch and Williamson (1990) even more desirable properties have emerged in the literature.

Consistency The consistency property is less frequently discussed in the literature. Notable exceptions are Jöckel et al. (2001) and Byun (1999). If $q = 1$ in the tracer transport equation (16), it mathematically degenerates to the continuity equation (14). This should ideally be the case numerically as well. If the two equations are solved using the same numerical method on the same grid and using the same time step, the consistency is, of course, guaranteed. However, in a realistic and practical setting found in many atmospheric models, the consistency is harder to achieve. The consistency property, or rather the lack of it, has been discussed in detail in the Introduction.

Preservation of constancy Another desirable property is the ability of the scheme to preserve a constant tracer field for a non-divergent flow.

2.4 Grid-point semi-Lagrangian transport schemes

Grid-point semi-Lagrangian schemes (also referred to as traditional semi-Lagrangian schemes) have been thoroughly reviewed in Staniforth and Côté (1991) and hence the method will only be discussed briefly here with updates.

The traditional semi-Lagrangian scheme is based on the discretization of (17)

$$\psi^{n+1} = \psi^n_*, \tag{26}$$

where the superscript refers to the time level and the subscript $*$ refers to the evaluation of ψ at the upstream departure point. The generic variable ψ is used (which for the continuity equation for a tracer is q) since the method also applies to the other equations of motion. Since the departure point does not necessarily coincide with a grid point, some kind of interpolation must be invoked in order to approximate ψ^n_*. In principle, any kind of interpolation may be used.

Probably the most widely used interpolator for two-dimensional problems is bicubic Lagrange interpolation or quasi biparabolic interpolation. It is a good compromise between accuracy and computational efficiency (Bates and McDonald 1982). Since it is such a well established and widely used method, it will be used as a reference throughout this study. Another popular interpolation method in semi-Lagrangian schemes is cubic splines (e.g, Purnell 1976;Riishøjgaard et al. 1998). To reduce the computational cost of fully two-dimensional methods, cascade methods have been introduced (Purser and Leslie 1991). Here the problem is reduced to two one-dimensional problems. This is done by defining a Lagrangian mesh (the regular mesh advected one time step) and then performing the interpolation along the Lagrangian longitudes and latitudes. Several schemes published in the meteorological literature are based on this method (e.g., Nair et al. 1999; Laprise and Plante 1995; Sun and Yeh 1997), which provides an economical and accurate alternative to fully two-dimensional methods.

2.4.1 *Enforcing monotonicity*

Monotonic filter The interpolators mentioned so far do not preserve the shape of the distribution unless special measures are taken. Using the ideas of flux-corrected transport (FCT; Zalesak (1979)), undershots and overshoots can be eliminated by using a simple and efficient filter. The resulting scheme is called *quasi-monotone semi-Lagrangian* scheme (QMSL) (Bermejo and Staniforth 1992). The filter works as follows. Whenever the upstream interpolated value is greater/(smaller) than the surrounding grid point values, the value is set equal to the maximum/(minimum) of the surrounding grid-point values. So the scheme reduces to low order whenever the high-order method overshoots or undershoots. The algorithm is currently used operationally in the IFS. The monotonicity is, however, enforced at the expense of accuracy. E.g. the mass is increased and there is a significant decrease in accuracy in terms of standard error measures (see table 1). The QMSL filter clips local extrema that in some situations should be retained (e.g.,

Method	$\int \psi / \int \psi_0$	$\int \psi^2 / \int \psi_0^2$	$Max(\psi)$	$Min(\psi)$	E_{DISS}	E_{DISP}
SL	1.00	0.90	4.51	-0.62	2.32(-3)	5.71(-2)
QMSL	1.01	0.82	3.99	0.00	0.96(-2)	0.95(-1)
CQMSL	1.00	0.81	3.99	0.00	1.00(-4)	6.46(-2)

Table 1: Error measures after six revolutions of a rotating slotted cylinder (see e.g., Bermejo and Staniforth 1992 for details) using cubic-spline interpolation with no "correction", with a quasi-monotone semi-Lagrangian (QMSL) filter (Bermejo and Staniforth 1992) and with the quasi-monotone semi-Lagrangian (CQMSL) algorithm of Priestley (1993), respectively. The error measures are (left to right) the mass, square of the mass, minimum value, maximum value and the dissipation and dispersion error measures introduced by Takacs (1985), respectively. All values are from the publications referred above.

see Fig. 2c in Nair et al. 1999).

More advanced filters in two dimensions are computationally expensive, but in one dimension the filter can be improved to deal with situations where the local extrema are physical, and then be applied to each one-dimensional sub-problem in cascade schemes (Sun and Yeh 1997; Nair et al. 1999).

Splines Alternative approaches for doing shape-preserving semi-Lagrangian advection is to use splines. The methods can be efficiently extended to two and three dimensions using a tensor product approach. Monotonocity is enforced by constraining derivative estimates (e.g., Williamson and Rasch 1989; Holnicki 1995). The unmodified splines tend to have better conservation properties than schemes using conventional interpolation. When modifying the derivative estimates in order to ensure monotonicity the scheme conserves mass less well as the peaks of small-scale structures are clipped. Hence, as for the schemes using Lagrange interpolation, also splines enforce monotone at the expense of conservation and vice versa.

2.4.2 *Mass fixers*

For general flows the traditional schemes do not conserve mass[7]. In order to restore mass one must periodically add or remove mass to the system. The question is how to do that without altering the shape of the distribution and without degrading the local property. The simplest procedure for restoring mass conservation is to add mass gained or lost to the average. Thereby a long-range transport of mass is introduced into the system and the numerical method is no longer local. This

[7] for divergence-free fbws the cubic spline method does conserve mass (Bermejo 1990)

method is therefore not desirable although widely used.

Priestley (1993), Bermejo and Conde (2002) and Sun and Sun (2004) have developed *ad hoc* mass-restoration algorithms which attempt to restore mass conservation while not violating monotonicity and locality requirements. Although the papers solve the problems differently, the basic idea is the same. In general the high order approximation to ψ_*^n can be written as a weighted sum of the surrounding grid-point values. Mass conservation can be regained by carefully choosing the weights. In Bermejo and Conde (2002) a new set of weights are computed by a minimization process. The monotone and conservative solution ψ is sought by minimizing the squared difference between ψ and the known non conservative but monotone solution with global mass conservation as a constraint. Mathematically it is a classical minimization problem that can be solved using Lagrange multipliers. A similar approach was taken by Sun and Sun (2004). In Priestley (1993) ψ_*^n is written as a linear combination of a high-order and low-order solution. The two solutions are weighted so that mass is conserved.

All the mass-fixing methods do, however, not guarantee entirely local mass conservation. The scheme of Priestley (1993) is less accurate with respect to the mean square error and the dispersion error compared to the unaltered solution, but slightly more accurate with respect to the dissipation error (see table 1). The scheme is, however, from 25% to 76% more expensive than the unaltered scheme, depending on which computer architecture is used (Table 4 in Priestley 1993). Similar conclusions hold for the scheme of Bermejo and Conde (2002). As already discussed in detail in the Introduction, all mass-fixing algorithms are non-local.

2.4.3 *Summary*

The traditional semi-Lagrangian schemes have several advantages over other numerical methods. They are economical in terms of CPU time and memory usage, are unconditionally stable and have small amplitude and phase errors for smooth flows. The efficiency is increased even further if a large number of tracers are transported simultaneously.

Common to all the traditional semi-Lagrangian schemes is that they are all cast in non-conservative form, and consequently have intrinsic difficulties in conserving mass. This can result in a significant drift in the global mass fields. To restore global mass conservation *ad hoc a-posteriori* algorithms must be employed. There is a degree of arbitrariness in these "mass-fixing" algorithms and they may degrade preexisting desirable properties. It is therefore not only mathematically more rigorous, but also desirable to use inherently conserving methods. During the last decade it has been demonstrated that it is possible to design semi-Lagrangian schemes which are inherently mass conservative. These schemes are the subject of the next section.

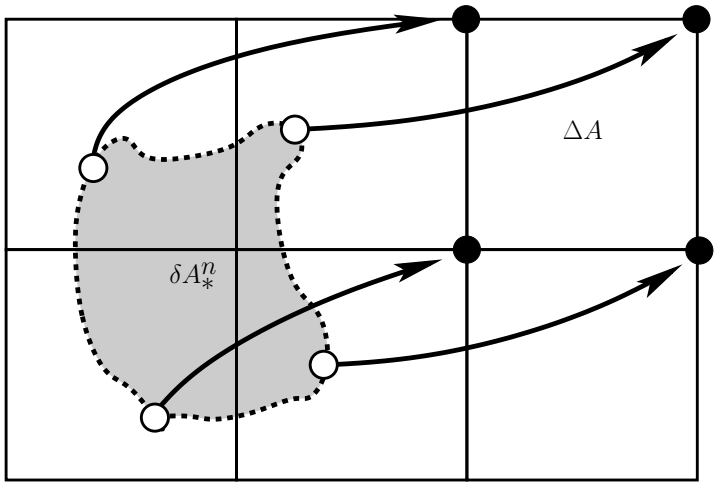


Fig. 2: The regular arrival cell with area ΔA and the irregular departure cell (shaded region) with area δA_*^n in the continuous case. The arrows are the parcel trajectories from the departure points (open circles) which arrive at the regular cell vertices (filled circles).

2.5 Cell-integrated semi-Lagrangian transport schemes

The semi-Lagrangian scheme can either be based on backward or forward trajectories, i.e. by considering parcels arriving or departing from a regular grid, respectively. The majority of semi-Lagrangian schemes are based on backward trajectories because it is usually simpler to remap from a regular to a distorted mesh. The deformed grid resulting from tracking the parcels moving with the flow is referred to as the *Lagrangian grid* while the stationary and regular grid is referred to as the *Eulerian grid*. The curve resulting from tracking a set of points along a latitude is referred to as a *Lagrangian latitude*. Similarly for *Lagrangian longitudes*.

Using backward trajectories the two-dimensional discretization of (20) leads to the CISL scheme

$$\overline{\psi}_{exp}^{n+1} \Delta A = \overline{\psi}_*^n \, \delta A_*^n \tag{27}$$

where

$$\overline{\psi}_*^n = \frac{1}{\delta A_*^n} \iint_{\delta A_*^n} \psi^n \, dA \tag{28}$$

is the integral mean value of ψ over the irregular departure cell area δA_*^n and $\overline{\psi}_{exp}^{n+1}$ is the mean value of ψ over the regular arrival cell area ΔA (see Fig. 2). The approximation of the integral on the right-hand side of (27) employs two steps. Firstly, defining the geometry of the departure cell. Secondly, performing the remapping, i.e. computing the integral over the departure cell using some reconstruction of the sub-grid distribution at the previous time step. The geometrical definition of the departure cell and the complexity of the sub-grid-scale distribution are crucial for the efficiency

and accuracy of the scheme. Before describing how the upstream integral can be approximated, the reconstruction of the sub-grid distribution is discussed.

2.5.1 *Sub-grid representation*

One-dimensional reconstructions Several one-dimensional methods for reconstructing the sub-grid distribution have been published in the literature. The simplest sub-grid representation is a piecewise constant function followed, in complexity, by a piecewise linear representation (van Leer 1977). Both methods are computationally cheap, monotonic and positive definite, but on the other hand excessively damping and therefore not suited for long runs. To reduce the dissipation to a tolerable level, the sub-grid-cell representation must be polynomials of at least second degree. Requirements of computational efficiency puts an upper limit to the order of the polynomials used, which explains why the predominant choice is second order.

The coefficients of the polynomials are determined by imposing constraints. Apart from the basic requirement of mass conservation within each grid cell, the choice of constraints is not trivial. Probably the simplest parabolic fit is obtained by requiring that the polynomial

$$p_i(x) = (a_0)_i + (a_1)_i\, x + (a_2)_i\, x^2, \quad x \in \left[x_{i-1/2}, x_{i+1/2}\right[\tag{29}$$

not only conserves mass in the ith grid cell

$$\int_{x_{i-1/2}}^{x_{i+1/2}} p_i(x) = \Delta x_i\, \overline{\psi}_i \tag{30}$$

but also in the two adjacent cells:

$$\int_{x_{i+1/2}}^{x_{i+3/2}} p_i(x) = \Delta x_{i+1}\, \overline{\psi}_{i+1}, \tag{31}$$

and

$$\int_{x_{i-3/2}}^{x_{i-1/2}} p_i(x) = \Delta x_{i-1}\, \overline{\psi}_{i-1}, \tag{32}$$

(Laprise and Plante 1995). Here $x_{i+1/2}$ and $x_{i-1/2}$ refer to the position of the ith cell border. Substituting (29) into (30), (31) and (32), and evaluate the analytic integrals, result in a linear system that can easily be solved for the three unknown coefficients $(a_0)_i$,$(a_1)_i$ and $(a_2)_i$. Performing this operation for all cells, a global piecewise-parabolic representation is obtained. Note that the method is only locally of second order since it is not necessarily continuous across cell borders. This method is referred to as the piecewise parabolic method 1 (PPM1).

An alternative way of constructing the parabolas, which ensures a globally continuous distribution, is the piecewise-parabolic method of Colella and Woodward (1984) (hereafter referred to as PPM2)

which has been reviewed in the context of meteorological modeling in Carpenter et al. (1990). It is convenient to use the cell average, $\overline{\psi}_i$ and the value of p_i at the left and right cell border, $(a_L)_i = p_i(x_{i-1/2})$ and $(a_R)_i = p_i(x_{i+1/2})$, respectively, instead of using $(a_0)_i$, $(a_1)_i$ and $(a_2)_i$ to define the parabolas. The equivalent formula for $p_i(x)$ is given by

$$p_i(x) = \overline{\psi}_i + (\delta a)_i \, x + (a_6)_i \, \left(\tfrac{1}{12} - x^2\right), \tag{33}$$

where $(\delta a)_i$ is the mean slope of p_i

$$(\delta a)_i = (a_R)_i - (a_L)_i,$$

and $(a_6)_i$ is the "curvature" of the parabola

$$(a_6)_i = 6\,\overline{\psi}_i - 3\,\left[(a_L)_i + (a_R)_i\right].$$

The first constraint is, of course, mass conservation (30). Secondly, the value of p_i at the left cell border, $(a_L)_i$, is fitted with a cubic polynomial using surrounding cell average values. Similarly, for the right border value, $(a_R)_i$. The result is

$$\begin{aligned}(a_L)_i &= \tfrac{1}{2}(\overline{\psi}_{i+1} + \overline{\psi}_i) + \tfrac{1}{6}\left(\Delta a_i + \Delta a_{i-1}\right), \\ (a_R)_i &= (a_L)_{i+1}\end{aligned}$$

for an equidistant grid (for a non-equidistant grid see Colella and Woodward 1984). This uniquely defines the parabolas and guarantees that the global sub-grid distribution is continuous across cell borders. Zerroukat et al. (2002) found in passive advection tests with their scheme that, when using the PPM2 for the sub-grid cell reconstructions (where the parabolas were continuous across cell borders), more accurate solutions were obtained compared to PPM1 (in which the distribution is not necessarily continuous across cell borders).

Instead of using the PPM1/2 Zerroukat et al. (2002) used a cubic generalization of the piecewise parabolic method for the reconstruction of the sub-grid-cell distributions. Of course any kind of reconstruction which is mass conserving can be used, for example rational functions as used in the transport scheme of Xiao et al. (2002) or parabolic splines. At present the most wide-spread sub-grid cell reconstruction method is PPM2.

Note that without further constraining the coefficients of the parabolas they do not guarantee monotonicity or positive definiteness. Standard filters (or limiters) that accommodate these requirements are given in Colella and Woodward (1984) and Lin and Rood (1996). The monotonic filter proposed in Colella and Woodward (1984) is very damping since it reduces the sub-grid distribution to a constant when the polynomial has monotonicity violating variation (that is when $p_i(x)$ takes values outside the range of a_L and a_R). Hence there are two situations in which the sub-grid scale distribution is modified: when $\overline{\psi}_i$ is a local extremum and when $\overline{\psi}_i$ is in between a_R and a_L, but sufficiently close to one of the values so that the parabola takes values outside the range of the edge values. The clipping proposed by Colella and Woodward (1984) can significantly reduce the accuracy in idealized advection tests (see Table 2). An alternative monotonicity filter has been

designed by Zerroukat et al. (2004a). It is formulated for cubic polynomials but can be adapted to the piecewise parabolic method (Dr. M. Zerroukat 2004, personal communication). It first detects if the monotonicity violating behavior is for an extremum or not. If $\overline{\psi}_i$ is an extremum then the high-order parabola is retained otherwise the order of the fitting polynomial is consecutively reduced until the variation is monotone. Then the severe clipping of "peaks" is eliminated and grid-scale noise is removed but without excessive damping. Note that the PPM2 (which unmodified is globally continuous) will be rendered discontinuous at some cell borders after the application of filters.

Fully two-dimensional methods The PPM in one dimension can be directly extended to two dimensions as has been done by Rančić (1992). This fully biparabolic fit involves the computation of nine coefficients, which makes the method computationally expensive. The computational cost can be reduced significantly by using a quasi-biparabolic sub-grid cell representation. Contrary to fully biparabolic fits, the quasi-biparabolic representation does not include the "diagonal" terms and simply consists of the sum of two one-dimensional parabolas, one in each coordinate direction. Using the form (33) for the parabolas, the quasi-biparabolic sub-grid-cell representation is given by

$$p_{ij}(x,y) = \overline{\psi}_{ij} + a^x_{ij}\,x + b^x_{ij}\,\left(\frac{1}{12} - x^2\right) + a^y_{ij}\,x + b^y_{ij}\,\left(\frac{1}{12} - y^2\right),$$

where a^x,b^y and a^y, b^y are the coefficients of the parabolic functions in each coordinate direction (Machenhauer and Olk 1998). This representation significantly reduces the computational cost of the sub-grid-cell reconstruction but, of course, does not include variation along the diagonals of the cells.

By using one-dimensional filters that prevent undershoots and overshoots to the parabolas in each coordinate direction, monotonicity violating behavior can be reduced but not strictly eliminated. In case of negative values at the cell boundaries of both unfiltered one-dimensional parabolic representations, even larger negative values may be present in one or more of the cell corners when the 1D representations are added. The monotone and positive definite filters eliminate only the negative values at the boundaries and not the larger negative corner values.

2.5.2 *Remapping*

For realistic flows the upstream cells deform into non-rectangular and possibly locally concave shapes (see Fig. 2). The question is how to integrate efficiently over a complex area. Several approaches have been suggested in the literature.

Fully two-dimensional schemes In Fig. 3 the departure cells of four different CISL schemes are shown. The simplest departure cell approximation in terms of geometry is the configuration of Laprise and Plante (1995); the departure cell is defined as a rectangle where the edges have the

same orientation as the arrival cell (see Fig. 3a). This is done by tracing the traverse motion of cell edges A, B, C and D, and not the cell vertices. Hereby the upstream cell retain orthogonality, which simplifies the upstream integral. On the other hand mass conservation is lost since the cells can overlap and have gaps between them. The cell geometry must be more advanced in order to obtain mass conservation, which is the primary design goal of CISL schemes.

Rančić (1992) defines the departure cell as a quadrilateral by tracking backward the cell vertices E, F, G, and H, and connecting them with straight lines (Fig. 3b). The vertices are not necessarily aligned with the coordinate axis which leads to some algorithmic complexity. In addition, the sub-grid-scale distribution used by Rančić was a piecewise biparabolic representation which in itself is quite expensive to compute. The combination of the complex geometry of the departure cell and the fully two-dimensional sub-grid-cell representation makes the scheme approx. 2.5 times less efficient than the traditional semi-Lagrangian advection scheme. This has hindered the scheme for use in full models. In order to speed up the remapping process, Machenhauer and Olk (1998) simplified both the geometry of the departure cell and the sub-grid-scale distribution. The departure cell is defined as a polygon with sides parallel to the coordinate axis (Fig. 3c). The sides parallel to the x-axis are at the y-values of the departure points, and the sides parallel to the y-axis pass through I, J, K, and L, located halfway between the departure points. With this form of departure cell and by using the pseudo-biparabolic sub-grid-scale distribution, the integral over the departure cell can be computed much more efficiently compared to the approach taken by Rančić (1992). For advection in Cartesian geometry NM02 reported a 10% overhead compared to the traditional semi-Lagrangian scheme. The NM02 scheme is hereafter referred to as the pseudo-biparabolic CISL (PB-CISL) scheme.

Cascade schemes The remapping can also be performed by splitting it into two one-dimensional remapping steps using the so-called cascade approach (Purser and Leslie 1991; Leslie and Purser 1995; Rančić 1995; Nair et al. 1999).

One such approach is the conservative cascade CISL scheme (CC-CISL) of Nair et al. (2002). In this scheme the departure cells are also defined as polygons with sides parallel to the coordinate axis and in each one-dimensional cascade step the PPM2 is used. Compared to the PB-CISL scheme the departure cell geometry in the CC-CISL scheme is defined somewhat differently (see Fig. 3d). Two of the sides parallel to the y-axis, $x = x(E)$ and $x = x(G)$, are defined as in the PB-CISL scheme and the remaining two sides are at the Eulerian longitude $x = x_i$. The sides parallel to the x-axis are determined from the intermediate Lagrangian grid points M, N, O, P, Q, and R defined as $y = 1/2\,\{y(\mathrm{M}) + y(\mathrm{N})\}$, $y = 1/2\,\{y(\mathrm{O}) + y(\mathrm{P})\}$, $y = 1/2\,\{y(\mathrm{P}) + y(\mathrm{Q})\}$, and $y = 1/2\,\{y(\mathrm{R}) + y(\mathrm{M})\}$, respectively. The y-values of the intermediate points are determined by cubic Lagrange interpolation between the y-values of four adjacent departure points along the *Lagrangian latitude* (dashed line on Fig. 3d). The upstream integral is computed by a remapping in the north-south direction from the Eulerian cells to the intermediate cells (crosshatched rectangular regions on Fig. 3d) followed by a remapping along the *Lagrangian latitudes* from the intermediate cells to the departure cells. Since the two remappings are one dimensional the scheme is more than

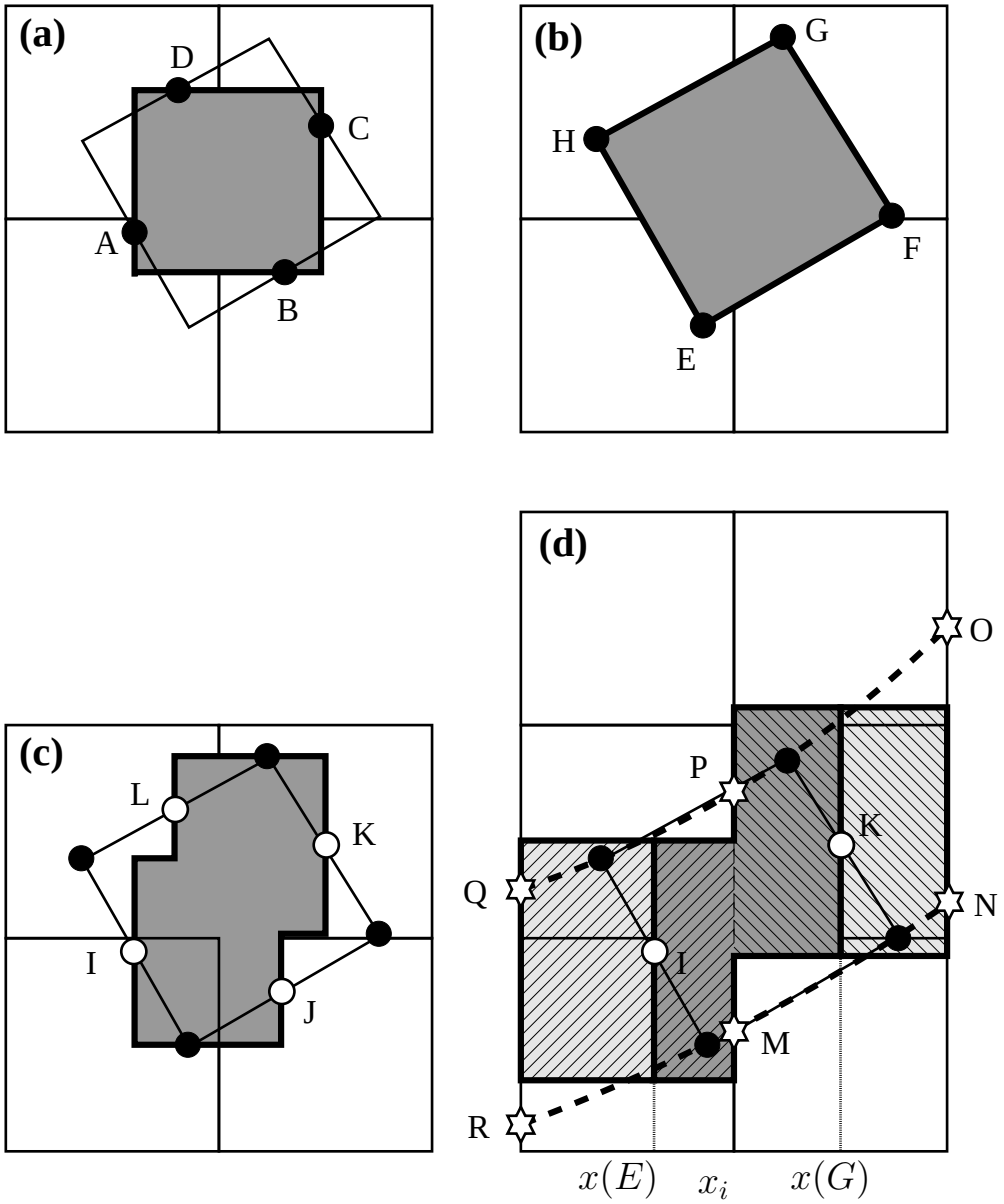


Fig. 3: The departure cells (shaded area) when using the (a) scheme of Laprise and Plante (1995), (b) scheme of Rančić (1992), (c) PB-CISL scheme and (d) CC-CISL scheme, respectively. The filled circles are the departure points, open circles the midpoints between the departure points, asterisks are the intermediate grid points which are used to define the intermediate cells in the cascade scheme (crosshatched area).

twice as efficient as the PB-CISL scheme (Nair et al. 2002).

Note that the departure areas in Fig. 3b, c and d completely cover the entire integration area without overlaps or cracks. Hence the total mass is conserved exactly. The size of the PB-CISL area

Scheme	l_1	l_2	l_∞	Min	Max
PB-CISL	0.075	0.051	0.083	-0.0088	-0.083
CC-CISL	0.051	0.039	0.076	-0.0070	-0.076
PB-CISL-M	0.077	0.089	0.18	-0.0038	-0.18
CC-CISL-M	0.070	0.086	0.186	0.0	-0.186

Table 2: Standard non-dimensional error measures for solid body rotation of a cosine bell for the PB-CISL scheme and the CC-CISL, respectively. The angle between the axis of solid body rotation and the polar axis is 30°, the resolution is 2.8125° and one revolution is completed in 256 time steps. The letter M denote the monotonic option (Colella and Woodward 1984). Additional details are in Chapter III.

(Fig. 3c) is equal to that of the Rančić area (Fig. 3a) whereas the CC-CISL area (Fig. 3d) only approximately so. Therefore, accepting the Rančić definition of the departure area as the most accurate, the PB-CISL area is the most accurate one of the two other schemes. The "jump" in the north and south walls in the CC-CISL scheme are not necessarily midway between the departure points, and hence the CC-CISL scheme may conserve mass locally less accurately than the PB-CISL scheme. On the other hand the CC-CISL scheme may in certain cases obtain a more accurate sub-grid-scale representation, namely when significant variations are along the *Lagrangian latitudes* and these are sloping toward north-east and south-east. In such situations the PB-CISL sub-grid-scale representation becomes less accurate due to the missing "diagonal" terms (see Table 2). Generally the most accurate sub-grid-scale representation of the three schemes is obtained with the full biparabolic representation of Rančić (1992).

The schemes discussed so far are based on integrating over areas which, given the departure cell topology for complex flows, can lead to some algorithmic complexity. Unlike the "area approach" Laprise and Plante (1995), Rančić (1995) and Zerroukat et al. (2002) have developed conservative cascade schemes which do not make explicit reference to departure areas during the remapping procedure. Hence they are not cell-integrated schemes in the sense considered so far where the "actual" integral over an area is computed.

The scheme of Laprise and Plante (1995) is not based on the finite-volume approach, but mass conservation is obtained globally anyway by differentiating a cumulative mass function. Rančić (1995) assigned mass to nodes (or "mass points") and used a cascade scheme similar to the one of Purser and Leslie (1991) based on forward trajectories. Instead of remapping along Lagrangian latitude and longitudes using Lagrange interpolation, a finite-volume method with PPM2 was used for the remapping. However, the application of the scheme for non-uniform grids and the extension to spherical geometry is not obvious.

A scheme that has also been extended to spherical geometry is the one of Zerroukat et al. (2002) called the SLICE (Semi-Lagrangian Inherently Conserving and Efficient) scheme. The scheme

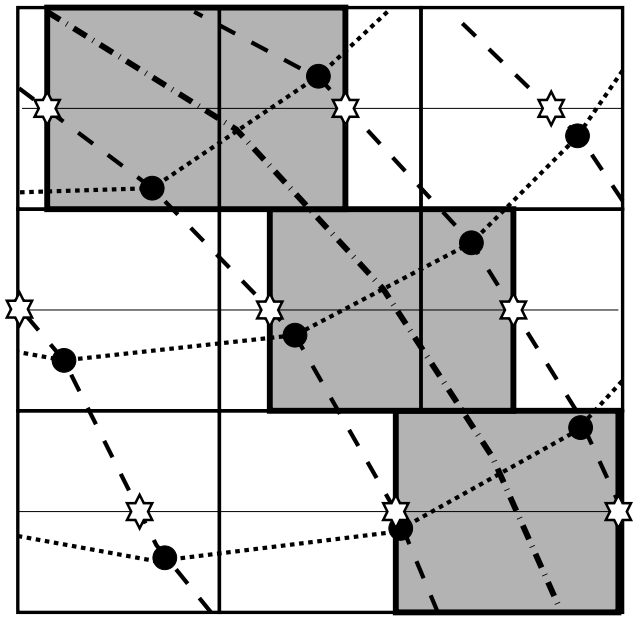


Fig. 4: Graphical illustration of the remappings in the SLICE scheme. The filled circles are the departure points corresponding to the cell vertices. The dotted/dashed lines are the Lagrangian latitudes/longitudes defined by connecting the departure points which arrive along the same latitude/longitude with straight line segments. The shaded areas are the intermediate Eulerian areas which are defined by the crossings between the Lagrangian longitudes and the Eulerian latitudes passing through the center of the Eulerian cells (thin lines). The crossings are marked with asterisk. The dash-dotted line is the line along which the cumulative distance function is defined and is used for the second remapping.

does not make explicit reference to areas during the remapping but is, however, similar to the CC-CISL scheme. As in the CC-CISL scheme the problem is divided into two one-dimensional remappings; one along an Eulerian coordinate direction and then one along a Lagrangian curve corresponding to the translation by the winds of one of the coordinate isolines. The remapping procedure is graphically illustrated on Fig. 4. As in the PB-CISL and CC-CISL schemes the cell vertices are tracked backward. The corresponding departure points are connected with straight lines to define *Lagrangian longitudes* and *latitudes*. Regular intermediate cells are defined by the intersections between the *Lagrangian longitudes* and the *Eulerian latitudes* that pass through the center of the cells. Similarly to the CC-CISL scheme the cell averages are mapped from the Eulerian cells to the regular intermediate cells defined by the intersections. The remap from the intermediate cells to the departure cells is quite different from the CC-CISL scheme. A cumulative distance function is defined in order to define the north-south cell borders along the *Lagrangian longitudes* passing through the center of the cells. The second remapping is performed along the *Lagrangian longitude* and by using the cumulative distance function to indicate the location of the cell "walls" along the *Lagrangian longitudes*. Hence the method makes no explicit reference to areas during the remapping procedure.

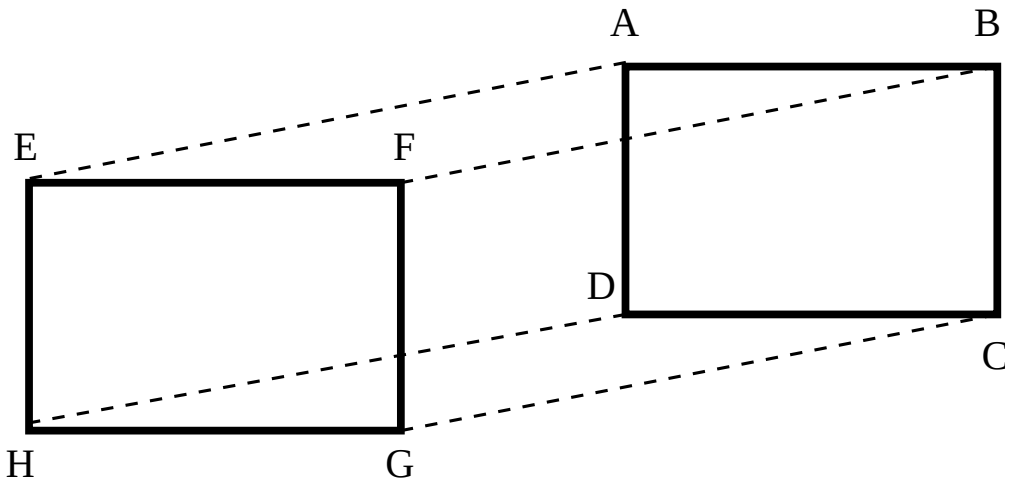


Fig. 5: The figure illustrates the equivalence between Eulerian flux-form and CISL schemes. See text for details.

A great potential of cascade schemes is that they may be extended to three dimensions without excessive computational cost and algorithmic complexity. For example, three-dimensional Lagrangian interpolation requires $\mathcal{O}(p^3)$ operations, where p is the formal order of accuracy of the interpolator, while cascade schemes require $\mathcal{O}(p)$ operations (e.g., Purser and Leslie 1991). Fully higher-dimensional cell-integrated schemes, on the other hand, increase rapidly in complexity as the number of dimensions is increased.

2.5.3 *A flux-form semi-Lagrangian scheme*

Lin and Rood (1996) and Leonard et al. (1996) have developed a scheme referred to as a finite-volume semi-Lagrangian scheme although it is quite different from the CISL schemes discussed so far. The approximation of the upstream integral can also be formulated in terms of fluxes instead of an integral over the departure area. The equivalence is illustrated on Fig. 5 for a uniform velocity field. The regular arrival cell is the rectangle ABCD and the corresponding departure cell is EFGH. The trajectories are the dashed lines connecting vertices of the departure and arrival cell. In the continuous case the integral over the departure cell is identical to the sum of the fluxes through the arrival cell walls, i.e. by adding the integrals over ABFE, DCGH, BCGF and ADHE computed with a positive sign for inflow and negative for outflow. For a "geometrical" CISL scheme using long time-steps, the integrals approximating the fluxes will stretch over several Eulerian cells and it is therefore more economical to integrate over the departure cell.

The fluxes can, however, be approximated using a less conceptional but more economical approach as done in Lin and Rood (1996) and Leonard et al. (1996). The respective schemes are identical, except for implementation details. The schemes approximate the fluxes using a combination of

one-dimensional flux-form and advective form operators. Thereby not only the transport components along coordinate axis are included as in conventional operator splitting schemes, but also the traverse motion is included. In fact for a uniform advecting velocity the method is equivalent to a cell-based semi-Lagrangian scheme in terms of the equivalence described above (Leonard et al. 1996). Contrary to the CISL scheme the flux-form schemes of Lin and Rood (1996) and Leonard et al. (1996) preserve a constant for non-divergent flow fields. They are extended to long time-steps by using integer advection and are therefore not semi-Lagrangian schemes in the traditional sense. The schemes are widely used in operational models for tracer advection. It is, however, not clear how the schemes could be incorporated in the dynamical core of existing semi-Lagrangian models since they have not been formulated with semi-implicit time stepping. For explicit time-stepping a dynamical core has been designed based on the Lin and Rood (1996) scheme (Lin 2004).

2.5.4 *Extensions to spherical geometry*

The singularities on the sphere are one of the main challenges for transport schemes. The number of schemes developed in Cartesian geometry is significantly larger than the number of schemes formulated for spherical geometry and that can transport quantities accurately and without excessive time-step limitations over the poles. For example the CISL schemes of Rančić (1992), Rančić (1995) and Laprise and Plante (1995) discussed in the previous sections have not been extended to spherical geometry.

Most algorithms require a certain amount of "engineering" to tackle the pole problem that often reduces the efficiency and simplicity of the algorithms. In Cartesian geometry the most accurate approximation to a departure cell, given the departure points, is the quadrilateral resulting from connecting the departure points with straight lines. Similarly, in spherical geometry the cells defined by connecting the departure points with great circle arcs are the optimal choice. But as in Cartesian geometry integrating along the optimal curves leads to complicated and computationally expensive algorithms. Therefore as in the Cartesian case the area approximation must be simplified.

The PB-CISL and CC-CISL schemes are extended to spherical geometry by using the μ grid, i.e. a latitude-longitude grid in which the latitude θ is replaced by $\mu = \sin\theta$. This transformation is invariant in the sense that the departure cells and corresponding upstream integrals take exactly the same form as in Cartesian geometry. In the vicinity of the poles, however, the cells on the μ-grid are a poor representation of the cells on the spherical latitude-longitude grid and some "engineering" is needed. In the PB-CISL scheme local area-preserving tangent planes at the poles are introduced. The areas in which the tangent planes are used is referred to as the *polar cap*. It is assumed that the east and west walls of the cells are straight lines in the tangent coordinate system. Using that assumption, the remapping procedure on the *polar cap* is formally equivalent to the Cartesian case. However, the east and west sides have a strong curvature when viewed in the μ-coordinate system. Hence extra latitudes are introduced in the *polar cap* so that the assumption of straight east and west walls in the tangent plane does not significantly degrade the accuracy. In the Lagrangian belt

Scheme	l_1	l_2	l_∞
PB-CISL	0.063	0.046	0.048
CC-CISL	0.054	0.042	0.065
SLICE-S	0.079	0.049	0.042

Table 3: Standard non-dimensional error measures for solid body rotation of a cosine bell over the pole for the PB-CISL scheme, the CC-CISL and SLICE-S scheme, respectively. The resolution is 2.8125° and one revolution is completed in 256 time steps. For additional details see e.g. NM02.

containing the pole point (referred to as the *singular belt*) the algorithm breaks down since the Lagrangian cell containing the Eulerian pole is not well defined. The total mass inside the *singular belt* can, however, easily be computed. The mass inside the *singular belt* is then distributed among the cells using Lagrange weights.

In cascade schemes the pole problem manifest itself by the fact that some Lagrangian latitudes may not cross a given Eulerian longitude. The CC-CISL scheme was first formulated for meridional Courant numbers less than unity, $C_\theta < 1$ (Nair et al. 2002), and later extended to longer time steps (Nair 2004). The schemes were extended to spherical geometry using the μ-grid which, as for PB-CISL, means that, away from the polar caps, the algorithm takes the same form as in the Cartesian case. In the general CC-CISL scheme the two-dimensional PB-CISL scheme was used over the polar caps and the unmodified CC-CISL scheme elsewhere (Nair 2004). Hereby, most of the efficiency of the CC-CISL scheme was retained while allowing for long time-steps over the *polar caps*.

The SLICE-S scheme uses a regular latitude-longitude grid (Zerroukat et al. 2004b). The cascade approach is based on the crossings of Lagrangian longitudes and Eulerian latitudes. Near the poles it can not be guaranteed that all Lagrangian longitudes cross a specific Eulerian latitude. In such cases the mass is simply distributed to the closest crossings based on distance-dependent weights. In principle the algorithm as described so far handles the pole problem, but the grid distortion near the poles can make the cascade axis nearly parallel instead of nearly orthogonal, making the remapping inaccurate. Therefore the mass is redistributed with a "post fix" procedure based on Lagrange weights similarly to the procedure in the PB-CISL scheme.

Of the three cascade schemes described here the SLICE-S scheme is the simplest regarding algorithmic complexity. It does not use high resolution *polar cap* belts and tangent planes. Since the CC-CISL and PB-CISL increase the resolution over the polar belts the accuracy of these schemes tends to be higher for transport over the poles compared to the SLICE-S scheme, but at an extra computational cost (see Table 3).

2.5.5 *Summary*

A review of the CISL schemes published in the literature has been given. These schemes can be divided into two categories: Fully two-dimensional methods and cascade schemes which split the problem into two one-dimensional ones. In general the CISL schemes perform much better than the traditional non-conservative schemes: they conserve mass exactly, permit long time steps, have the option of being monotone or simply positive definite, and have small phase and amplitude errors. For passive advection the two-dimensional schemes have a small computational overhead compared to traditional schemes, while cascade schemes are slightly more efficient than traditional methods. The accuracy of cascade and fully two-dimensional schemes are comparable in idealized advection tests. As discussed in this review of CISL schemes, the cascade approach may have an advantage when the flow is along cell diagonals. In this case the remapping is performed along the characteristics of the flow and thereby "catch" more of the diagonal variation.

The main criticism of CISL schemes is the amount of "engineering" needed near the poles. This leads to some algorithmic complexity. Another potential problem with CISL schemes is their extension to three dimensions. A fully three-dimensional extension of the fully two-dimensional scheme would be very expensive and complex. The cascade approach, on the other hand, seems straight forward to extend to three dimensions at little extra computational cost.

So far CISL schemes have only been applied to advection of passive tracers and have not been applied to full model systems. As discussed in the Introduction, accurate tracer advection requires that the mass fields and wind fields supplied to the advection scheme are consistent. The only way to ensure this (without *ad hoc* "fixing" algorithms) is to use the CISL scheme for the continuity equation of the dynamical core as for tracer advection.

To exploit the efficiency of semi-Lagrangian methods in full models it is, however, important to use long time steps which requires a special treatment of the equation terms involved in gravity wave motions, for example treating the fast waves semi-implicitly. So far CISL schemes have not been applied to full models. The flux-form semi-Lagrangian scheme of Lin and Rood (1996), which is an Eulerian scheme extended to long time steps by integer advection, has been applied in an hydrostatic dynamical core using explicit time stepping (e.g., Lin 2004). For the application of CISL schemes in complete models they must be coupled with semi-implicit time stepping in order to remain efficient.

Chapter III

Inherently mass-conservative SISL shallow water model

The first step toward the long term goal of developing a dynamical core for a global, baroclinic, quasi-hydrostatic and SISL model that is locally as well as globally mass conserving is taken in this Chapter: a mass-conserving SISL shallow water system is set up and tested for a limited area in spherical geometry. The limited area models (LAMs) cover a rectangular area around the equator of a rotated spherical latitude-longitude grid that does not include the pole. Hence the pole problem is not explicitly treated here, but possible extensions to a global domain are discussed. The continuity equation is solved using a CISL method on the μ grid, i.e. a latitude-longitude grid in which the latitude θ is replaced by $\mu = \sin(\theta)$ (Nair and Machenhauer 2002, hereafter referred to as NM02), whereas the momentum equations are solved using the traditional grid-point semi-Lagrangian method on a rotated spherical latitude-longitude grid. The semi-implicit continuity equation is formulated such that the resulting elliptic equation is in the same form as for the traditional SISL grid-point system. The results of test cases with the mass-conservative model are compared to results obtained with a shallow water version of the traditional non-conserving SISL HIRLAM (McDonald and Haugen 1992; McDonald and Haugen 1993).

The main novelty of the present study is the coupling of the CISL scheme with a semi-implicit time stepping. A preliminary study was made by Machenhauer and Olk (1997), who derived a one-dimensional SISL shallow water model based on CISL advection schemes conserving either mass and total energy or mass and angular momentum. The model is yet to be extended to two dimensions and spherical geometry. In one dimension the derivation of the semi-implicit scheme is rather straight forward, but the two-dimensional extension is not obvious. The reason is that the semi-implicit scheme must be consistent with the cell-integrated method and in two dimensions this geometric approach leads to a very complicated elliptic system. Instead a more simple approach suggested by Kaas et al. (2005) for the formulation of a semi-implicit CISL scheme, which leads to a simpler elliptic system, is followed.

The formulation of the traditional semi-Lagrangian method used for the momentum equations and the formulation of the CISL method used for the continuity equation are presented in section 1. The accuracy of the model has been assessed using selected problems from the standard test suit of Williamson et al. (1992, hereafter referred to as W92) . The adaptation to a LAM setup and the results from numerical integrations of selected test cases are presented in section 2. Hereafter possible extensions to a global domain are discussed (section 3).

3.1 The model

3.1.1 *Governing equations*

The equations used here are the shallow water continuity equation

$$\frac{d\Phi}{dt} = -\Phi \nabla \cdot \mathbf{v}, \tag{34}$$

and the momentum equations on component form (e.g., McDonald and Bates 1989)

$$\frac{d\,u}{dt} = -\frac{1}{a\,\cos\theta}\frac{\partial \Phi}{\partial \lambda} + f\,v, \tag{35}$$

$$\frac{d\,v}{dt} = -\frac{1}{a}\frac{\partial \Phi}{\partial \theta} - f\,u, \tag{36}$$

where Φ is the geopotential height, f the Coriolis parameter, $\mathbf{v}$ is the velocity vector $\mathbf{v} = (u, v)$, d/dt the horizontal total derivative, ∇ the horizontal gradient operator, a the radius of the earth and λ the longitude and the θ latitude. In a semi-Lagrangian model the trajectory equation

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \tag{37}$$

is one of the prognostic equations as well. Equations (34), (35), (36) and (37) are the basis for the discretizations in traditional semi-Lagrangian models.

In order to formulate a CISL continuity equation (34) is integrated over an infinitesimal area δa moving with the fluid. For consistency the divergence $\nabla \cdot \mathbf{v}$ in (34) must be expressed on the Lagrangian form

$$\mathbb{D} = \frac{1}{\delta a}\frac{d}{dt}\left(\delta a\right). \tag{38}$$

Using (38) and integration by parts, (34) becomes the CISL continuity equation

$$\frac{d}{dt}\left(\Phi\,\delta a\right) = 0. \tag{39}$$

The discretizations in the model presented here are based on (35), (36), (37) and (39).

3.1.2 *Trajectory algorithm*

The first step in the semi-Lagrangian algorithm is the computation of the trajectories. The trajectory algorithm used here works directly in (λ, θ) coordinates so $\mathbf{r} = (\lambda, \theta)$. To accommodate the semi-implicit formulation of the continuity equation, an iterative departure point algorithm in which the trajectory is split into a part involving only velocities extrapolated to time level $(n+1)$ and a part involving only velocities at time level (n) is used. This is done by splitting the trajectory into two segments (see Fig. 23):

1. The first contribution is the trajectory from the departure point $\mathbf{r}_*^n$ to the trajectory midpoint $\mathbf{r}_{*/2}^{n+1/2}$. Approximating this half trajectory with a Taylor series expansion about the departure point, only velocities at time level (n) are used. The Nth order approximation is given by

$$\mathbf{r}_{*/2}^{n+1/2} = \mathbf{r}_*^n + \mathcal{C}_1, \tag{40}$$

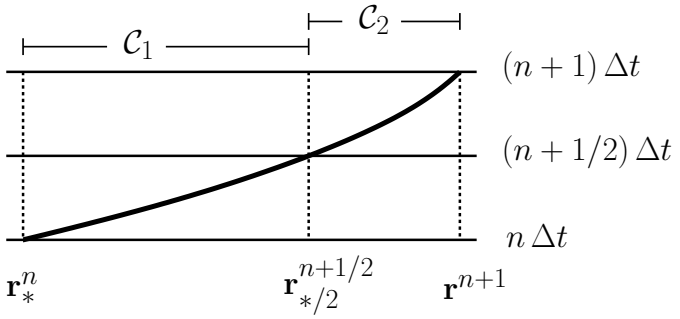


Fig. 6: A "space-time diagram" (time is plotted on the y axis and departure point distance on the x axis) of a trajectory. $\mathbf{r}^{n+1}$ is the arrival point, $\mathbf{r}_*^n$ is the departure point and $\mathbf{r}_*^{n+1/2}$ denotes the trajectory midpoint. $\mathcal{C}_1$ and $\mathcal{C}_2$ are the two contributions to the trajectory. See text for details.

where

$$\mathcal{C}_1 = \frac{\Delta t}{2}\,\mathbf{v}_*^n + \sum_{\nu=1}^{N-1} \frac{1}{(\nu+1)!}\left(\frac{\Delta t}{2}\right)^{\nu+1}\left(\frac{d^\nu \mathbf{v}}{dt^\nu}\right)_*^n. \tag{41}$$

The first guess departure point is the arrival point $\mathbf{r}^{n+1}$. The total derivative is approximated as in McGregor (1993), i.e. by discarding the Eulerian velocity change

$$\frac{d\,\mathbf{v}}{dt} \approx \mathbf{v}\cdot\nabla\mathbf{v}, \tag{42}$$

where the operator ∇ is expressed in spherical coordinates. Higher-order derivatives are defined recursively

$$\frac{d^\nu \mathbf{v}}{dt^\nu} = \frac{d}{dt}\left(\frac{d^{\nu-1}\mathbf{v}}{dt^{\nu-1}}\right) \quad \nu = 2,3,..,N-1. \tag{43}$$

2. The second contribution is the trajectory from the midpoint $\mathbf{r}_{*/2}^{n+1/2}$ to the arrival point $\mathbf{r}^{n+1}$. A Taylor series expansion about the arrival point involves only extrapolated velocities $\tilde{\mathbf{v}}^{n+1}$, where the operator $\widetilde{(\cdot)}^{n+1}$ is defined by

$$\widetilde{(\cdot)}^{n+1} = 2\,(\cdot)^n - (\cdot)^{n-1}. \tag{44}$$

The "half-trajectory" is given by

$$\mathbf{r}_{*/2}^{n+1/2} = \mathbf{r}^{n+1} - \mathcal{C}_2, \tag{45}$$

where

$$\mathcal{C}_2 = \frac{\Delta t}{2}\,\tilde{\mathbf{v}}^{n+1} - \sum_{\nu=1}^{N-1} \frac{1}{(\nu+1)!}\left(-\frac{\Delta t}{2}\right)^{\nu+1}\frac{d^\nu}{dt^\nu}\left(\tilde{\mathbf{v}}^{n+1}\right). \tag{46}$$

Combining (40) and (45) the departure point is given in terms of the sum of two contributions

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - (\mathcal{C}_1 + \mathcal{C}_2)\,. \tag{47}$$

To increase accuracy $\mathcal{C}_1$ is iterated. Note that the second contribution $\mathcal{C}_2$ is based entirely on grid-point values of the velocity field and is not iterated. For computational efficiency the first order ($N = 1$) approximation to the departure point

$$\mathbf{r}_*^n = \mathbf{r}^{n+1} - \frac{\Delta t}{2}\left(\mathbf{v}_*^n + \tilde{\mathbf{v}}^{n+1}\right), \tag{48}$$

is used for the Lagrangian divergence computations (see below). For all other terms the acceleration ($N = 2$) is included in the trajectory computations and three iterations are used. For efficiency bilinear, biquadratic and bicubic Lagrange interpolation is used for the first, second and third iteration, respectively. In the present study it was not investigated if fewer iterations affect the accuracy of the model. This will be the scope of future work.

3.1.3 *The traditional semi-implicit semi-Lagrangian discretization*

The shallow water equations (34), (35) and (36) can be written as

$$\frac{d\psi}{dt} = N_\psi + L_\psi \qquad \psi = u, v, \Phi, \tag{49}$$

where L_ψ and N_ψ are the linear and non-linear parts of the right-hand side of the equation for ψ, respectively. The explicit two-time-level traditional semi-Lagrangian discretization of (49) is given by

$$\psi_{exp}^{n+1} = \psi_*^n + \Delta t\,\left\{(L_\psi)_{*/2}^{n+1/2} + (N_\psi)_{*/2}^{n+1/2}\right\}, \tag{50}$$

where the known terms have been collected on the right-hand side. Values at time level $(n + 1/2)$ are obtained by linear extrapolation in time. The classical semi-implicit technique introduced by Robert et al. (1972) treats the non-linear terms explicitly and the linear part semi-implicitly in terms of a temporal average. The traditional SISL scheme is given by

$$\psi^{n+1} = \psi_*^n + \Delta t\,(N_\psi)_{*/2}^{n+1/2} + \frac{\Delta t}{2}\left\{(L_\psi)^{n+1} + (L_\psi)_*^n\right\}, \tag{51}$$

(e.g., Temperton and Staniforth 1987). It can easily be shown that (51) can be written as

$$\psi^{n+1} = \psi_{exp}^{n+1} + \frac{\Delta t}{2}\left\{(L_\psi)^{n+1} - \widetilde{(L_\psi)}^{n+1}\right\}, \tag{52}$$

by adding and subtracting $(L_\psi)_{*/2}^{n+1/2}$, using the approximation

$$(L_\psi)_{*/2}^{n+1/2} = \frac{1}{2}\left\{(L_\psi)_*^n + \widetilde{(L_\psi)}^{n+1}\right\}, \tag{53}$$

and (50). Motivated by noise and stability problems several other ways to evaluate the non-linear terms have been suggested in the literature. For example, the spatially averaged approach where $(N_\psi)^{n+1/2}_{*/2}$ in (51) is replaced by

$$\frac{1}{2}\left\{(N_\psi)^{n+1/2} + (N_\psi)^{n+1/2}_{*}\right\}, \tag{54}$$

(Kaas 1987) and the stable extrapolation two-time-level scheme (SETTLS) used at the ECMWF

$$\frac{1}{2}\left\{2\,(N_\psi)^{n}_{*} - (N_\psi)^{n-1}_{*} + (N_\psi)^{n}\right\}, \tag{55}$$

(Hortal 2002). The stability of these schemes, and others, is discussed in Durran and Reinecke (2004). Noise problems associated with orographic forcing can also be alleviated by using decentering (also referred to as uncentering or offcentering in the literature) in which a small offcentering in the time averaged terms is applied (Rivest et al. 1994).

For the momentum equations the traditional semi-Lagrangian scheme (52) on an Arakawa C-grid (Arakawa and Lamb 1977) with the centered approximation to the non-linear terms (54) is used. Treating the continuity equation in the same way, the configuration corresponds to a shallow water version of the semi-Lagrangian HIRLAM (hereafter referred to as SW-HIRLAM), which will be used for comparison throughout this study. The linear and non-linear terms are given in Appendix D.

3.1.4 *The discrete CISL continuity equation*

The explicit continuity equation is solved in spherical geometry on a μ grid using the discrete form of (39)

$$\overline{\Phi}^{n+1}_{exp}\,\Delta A = \overline{\Phi}^{n}_{*}\,\delta A^{n}_{*} \tag{56}$$

where

$$\overline{\Phi}^{n}_{*} = \frac{1}{\delta A^{n}_{*}} \iint_{\delta A^{n}_{*}} \Phi^{n}(\lambda,\mu)\,dA \tag{57}$$

is the integral mean value of the geopotential over the irregular departure cell area δA^{n}_{*} and $\overline{\Phi}^{n+1}_{exp}$ is the mean value of the geopotential over the regular arrival cell area ΔA. The departure area is defined by tracking backward the trajectories from the corner points of the arrival cell (see e.g., NM02 and Nair et al. 2002). Under the assumption that the time step is chosen such that the trajectories do not cross, the departure cell is "well defined" and the method is guaranteed to be conservative; if cells are not "well defined" it is detected immediately by the lack of mass conservation. The computation of the integral on the right-hand side of (56) employs two steps. Firstly, defining the geometry of the departure cell. Secondly, performing the remapping, i.e. computing the integral over the departure cell using some reconstruction of the sub-grid-cell distribution at the previous time step. As discussed in detail in section **2.5**, the geometrical definition of the departure cell and the complexity of the sub-grid-scale distribution are crucial for the efficiency of the scheme.

Here the schemes of NM02 and Nair et al. (2002) are used and hereafter referred to as the pseudo biparabolic CISL (PB-CISL) scheme and the conservative cascade CISL scheme (CC-CISL), respectively. As the PB-CISL and CC-CISL scheme use the piecewise parabolic method of Colella and Woodward (1984) for reconstructing the sub-grid-cell distributions the monotonic (M), positive-definite (P) and semi-monotonic (SM) constraints as described in Lin and Rood (1996) can optionally be used.

3.1.5 *Derivation of the semi-implicit system*

In order to formulate the semi-implicit continuity equation, the Lagrangian divergence (38) is discretized by a centered approximation

$$\mathbb{D}^{n+1/2} = \frac{1}{\Delta A}\frac{\Delta A - \delta A_*^n}{\Delta t}. \tag{58}$$

Note that the discretized divergence is determined entirely by the trajectories and the geometrical definition of the departure area. Since the trajectory is split into a part using only extrapolated velocities and velocities at time level (n), the departure area can be split into two parts. The Lagrangian divergence is defined in terms of areas hence $\mathbb{D}$ can be split into two parts as well

$$\mathbb{D}^{n+1/2} = \frac{1}{2}\left[\mathbb{D}(\tilde{\mathbf{v}}^{n+1}) + \mathbb{D}(\mathbf{v}^n)\right]. \tag{59}$$

The first term on the right-hand side of (59) is given by

$$\mathbb{D}(\tilde{\mathbf{v}}^{n+1}) = \frac{1}{\Delta A}\frac{\Delta A - \delta A(\tilde{\mathbf{v}}^{n+1})}{\Delta t}, \tag{60}$$

where the departure area $\delta A(\tilde{\mathbf{v}}^{n+1})$ is computed using only extrapolated winds. As mentioned in section 2b only first-order trajectories (48) are used for the Lagrangian divergence computations, i.e. the area is computed using the departure points located at

$$\mathbf{r}^{n+1} - \frac{1}{2}\,\tilde{\mathbf{v}}^{n+1}\,\Delta t. \tag{61}$$

For all other terms the acceleration is included in the trajectory computations.

The semi-implicit continuity equation is derived as in the one-dimensional shallow water system of Machenhauer and Olk (1997). Isolating δA_*^n in (58) and thereafter substituting the resulting expression into the explicit continuity equation (56), the explicit prediction can be written as

$$\overline{\Phi}_{exp}^{n+1}\Delta A = \overline{\Phi}_*^n\,\Delta A - \Delta t\,\mathbb{D}^{n+1/2}\Phi_{00}\,\Delta A - \Delta t\,\mathbb{D}^{n+1/2}\,\overline{\Phi'}_*^n\,\Delta A, \tag{62}$$

where the geopotential $\overline{\Phi}_*^n$ in the divergence terms has been split into a constant geopotential Φ_{00} and a perturbation $\overline{\Phi'}_*$. Treating the linear term implicitly, the semi-implicit continuity equation becomes similar to that of the traditional system:

$$\overline{\Phi}^{n+1} = \overline{\Phi}_{exp}^{n+1} - \frac{\Delta t}{2}\,\Phi_{00}\,\left[\mathbb{D}^{n+1} - \mathbb{D}(\tilde{\mathbf{v}}^{n+1})\right], \tag{63}$$

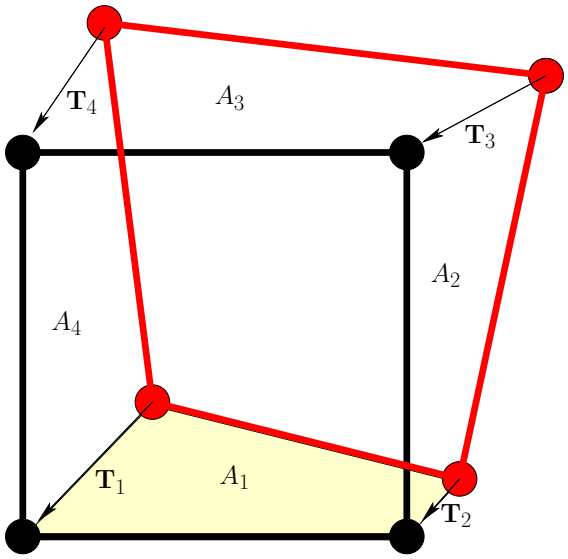


Fig. 7: Graphical illustration of how the divergence $\mathbb{D}^{n+1}$ is computed. The black rectangle is the arrival cell with area ΔA. The red quadrilateral is the departure cell with area δA_*^n. The corner points of the departure cell follow straight line trajectories as indicated by the displacement vectors $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$ and $\mathbf{T}_4$. A_1 (yellow) is the area enclosed by $\mathbf{T}_1$, $\mathbf{T}_2$ and the south side of the arrival and departure cell. A_2, A_3 and A_4 are defined similarly.

Since the arrival and departure areas span the entire domain without overlaps or cracks, (63) conserves mass when integrated over the domain, and assuming there is no mass flux through the domain boundaries. Derivation of the full semi-implicit system using (63) and (52) with $\psi = u, v$, requires a formula for the Lagrangian divergence as a function of wind components. The derivation of such a formula is the topic of the next section.

3.1.6 *The "ideal" semi-implicit system*

The "ideal" semi-implicit system turns out to be quite complicated and is therefore not used for the model presented here. For completeness, however, the derivation is discussed in some detail.

To simplify the derivation assume Cartesian geometry and a grid in which the velocity components are known at cell vertices (e.g., Arakawa A-grid). The derivation of the elliptic equation requires an expression for the divergence $\mathbb{D}$ at the new time level $(n+1)$ as a function of velocity components. In other words, Given $\mathbf{v}^{n+1}$ find an expression for $\mathbb{D}^{n+1}$.

The cell divergence at time level $(n+1)$ is approximated using the following one-sided formula

$$\mathbb{D}^{n+1} = \frac{1}{\Delta A}\frac{\Delta A - \delta A_*^n}{\Delta t}, \tag{64}$$

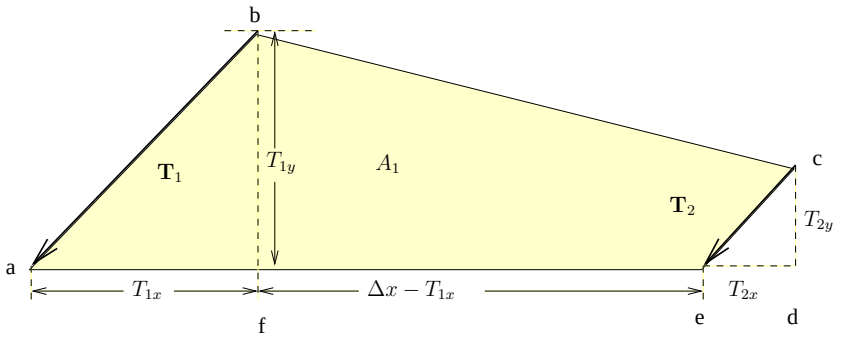


Fig. 8: Graphical illustration of how A_1 (yellow) is calculated. A_1 can be written as a sum of the area of the two triangles 'abf' and 'ecd', and the rhomboidal 'fbce'.

where δA_*^n is the area of the cell which vertices are found by using backward straight line trajectories based on $\mathbf{v}^{n+1}$ (Fig. 7). Define the displacement vectors $\mathbf{T}_i = (T_{ix}, T_{iy})$, $i = 1, .., 4$, as the straight line trajectories from the corner points of the departure cell to the associated arrival cell vertices. $\mathbf{T}_i$ is given by

$$\begin{aligned}
\mathbf{T}_1 &= \Delta t\ (u_{ij}, v_{ij})^{n+1}\,, \\
\mathbf{T}_2 &= \Delta t\ (u_{i+1j}, v_{i+1j})^{n+1}\,, \\
\mathbf{T}_3 &= \Delta t\ (u_{i+1j+1}, v_{i+1j+1})^{n+1}\,, \\
\mathbf{T}_4 &= \Delta t\ (u_{ij+1}, v_{ij+1})^{n+1}\,.
\end{aligned}$$

The displacement vectors $\mathbf{T}_1$ and $\mathbf{T}_2$ and the south sides of the departure and arrival cell define a quadrilateral which area is denoted A_1. Similarly define areas A_2, A_3 and A_4 (see Fig. 7). The area of the departure cell δA_*^n can be written in terms of A_i and ΔA

$$\delta A_*^n = \Delta A + A_2 + A_3 - A_1 - A_4. \tag{65}$$

A_i can be calculated by writing A_i in terms of a rhomboidal and two triangles (see Fig.8). The result is

$$\begin{aligned}
A_1 &= -\tfrac{1}{2}\,(T_{1y} + T_{2y})\ (-T_{2x} - \Delta x + T_{1x}) + \tfrac{1}{2}\,T_{1x}\,T_{1y} - \tfrac{1}{2}\,T_{2x}\,T_{2y}, \\
A_2 &= -\tfrac{1}{2}\,(T_{2x} + T_{3x})\ (-T_{3y} - \Delta y + T_{2y}) + \tfrac{1}{2}\,T_{2x}\,T_{2y} - \tfrac{1}{2}\,T_{3x}\,T_{3y}, \\
A_3 &= -\tfrac{1}{2}\,(T_{3y} + T_{4y})\ (-T_{3x} - \Delta x + T_{4x}) + \tfrac{1}{2}\,T_{4x}\,T_{4y} - \tfrac{1}{2}\,T_{3x}\,T_{3y}, \\
A_4 &= -\tfrac{1}{2}\,(T_{1x} + T_{4x})\ (-T_{4y} - \Delta y + T_{1y}) + \tfrac{1}{2}\,T_{1x}\,T_{1y} - \tfrac{1}{2}\,T_{4x}\,T_{4y}.
\end{aligned}$$

Using (64), (65) and the formulas for A_i and $\mathbf{T}_i$, the divergence $\mathbb{D}^{n+1}$ can, after some algebra, be written as

$$\begin{aligned}
\mathbb{D}^{n+1} = \frac{1}{2}\,(\delta_x u_{ij} + \delta_y v_{ij} + \delta_x u_{ij+1} + \delta_y v_{i+1j}) \\
+ \frac{1}{2}\frac{\Delta t}{\Delta x\,\Delta y}\,[(\delta_{\nwarrow} v_{ij})\ (\delta_{\nearrow} u_{ij}) + (\delta_{\nearrow} v_{ij})\ (\delta_{\nwarrow} u_{ij})]
\end{aligned} \tag{66}$$

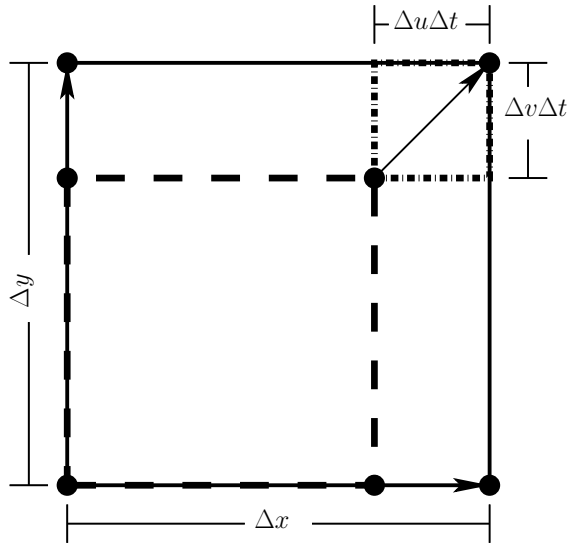


Fig. 9: A simple example illustrating the divergence computed using finite differences and the divergence computed using areas. The solid line rectangle is the arrival cell and the dashed rectangle is the departure cell. When computing the divergence using the Eulerian formula it differs from the Lagrangian one by the area enclosed by the dashed-dotted line.

where the finite difference operators are defined by

$$\begin{aligned}
\delta_x Q_{ij} &= \frac{Q_{i+1j} - Q_{ij}}{\Delta x}, \\
\delta_y Q_{ij} &= \frac{Q_{ij+1} - Q_{ij}}{\Delta y}, \\
\delta_{\nearrow} Q_{ij} &= Q_{i+1j+1} - Q_{i,j}, \\
\delta_{\nwarrow} Q_{ij} &= Q_{ij+1} - Q_{i+1j}.
\end{aligned}$$

The latter two operators are associated with the derivatives along the diagonal. The expression enclosed in the first set of brackets on the right-hand side of (66) corresponds to the Eulerian approximation of the divergence at the cell center. Computing the divergence using the Eulerian formula corresponds to the cell configuration of Laprise and Plante (1995) in which cells may overlap or have cracks between them. The terms enclosed in the square brackets on the right-hand side of (66) are the diagonal terms. These terms become increasingly important if the departure cell walls are not aligned with the coordinate axis. A simple example illustrates the order of magnitude of the terms.

Consider a velocity field which is a function of the geometric distance from the center of the coordinate system. After one time step a cell placed with its south-west corner at the center will have expanded as depicted on Fig. 9. If the divergence is computed using the Eulerian formula it

is given by

$$D = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y}, \tag{67}$$

where Δu and Δv are the differences between the velocities at the respective departure and arrival points. In terms of areas, (67) corresponds to the divergence computed using the Lagrangian formula but where the small dashed-dotted rectangle on Fig. 9 has been included twice. $\mathbb{D}$ is given by

$$\mathbb{D} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta y} - \Delta t \, \frac{\Delta u \, \Delta v}{\Delta x \, \Delta y}. \tag{68}$$

Although the difference between the Eulerian and Lagrangian divergence is second order, $\mathbb{D}$ can not be replaced by D in the cell-integrated model. If the areas used for divergence computations do not span the entire domain without cracks or overlaps, mass conservation is lost. Hence the last term on the right-hand side of (68) can not be discarded although it is small compared to the other terms.

The derivation of the elliptic equation associated with the semi-implicit system proceeds by substituting for $\mathbf{v}^{n+1}$ in (66) using the momentum equations. The diagonal terms in (66), however, involve products that complicate the elliptic equation. Extension to spherical geometry makes matters worse. The most obvious way to define the areas in spherical geometry is probably to connect the cell vertices with great circle arcs. The formula for the area of the spherical polygon is, however, complicated and would therefore result in a complex elliptic equation (see Appendix E). Alternatively, the μ grid could be used. In that case (66) would involve the sine function which would severely complicate the elliptic system as well. Unless the expression for the Lagrangian divergence can be simplified in a consistent manner, i.e. area-preserving manner, this approach seems to result in an elliptic equation which, if solvable, would be computationally very expensive to solve. In addition, the derivation has been performed under the assumption of having defined the velocities at cell vertices. If that is not the case some kind of interpolation or averaging must be employed to get $\mathbf{v}$ at the corner points. Consequently, the complexity of the formula for $\mathbb{D}$ would increase by using, for example, an Arakawa C-grid.

3.1.7 *A simpler semi-implicit system*

As discussed the "ideal" semi-implicit system results in an elliptic equation more complicated than the elliptic equation associated with the traditional system. Rather than pursuing that avenue, a method which results in an elliptic equation in the same form as the one associated with the traditional system is used. Thereby the method is more easily implemented in existing systems. The algorithm is a spherical extension of the method presented in Kaas et al. (2005).

By replacing $\mathbb{D}^{n+1}$ with the discretized Eulerian divergence D^{n+1} in (63) an elliptic equation in the same form as the traditional system results. In this context though the discretized Eulerian divergence is inconsistent since it corresponds to a discretized Lagrangian divergence with a departure cell area different from that defined for the CISL scheme in question. The "Eulerian" departure cell

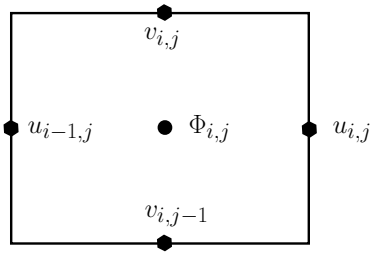


Fig. 10: The placement of the Eulerian cell with respect to the Arakawa C-grid.

areas, which correspond to the cell configuration of Laprise and Plante (1995), generally overlap and do not span the entire domain. Consequently, the mass conservation is lost. However, the mass conservation property is recovered if the inconsistency introduced by using the Eulerian divergence is corrected for in the next time step. The SISL cell-integrated continuity equation becomes

$$\overline{\Phi}^{n+1} = \overline{\Phi}^{n+1}_{exp} - \frac{\Delta t}{2}\,\Phi_{00}\,\left[D^{n+1} - \mathbb{D}(\tilde{\mathbf{v}}^{n+1})\right] + \frac{\Delta t}{2}\,\Phi_{00}\,\overline{\left[D^{n} - \mathbb{D}(\mathbf{v}^{n})\right]}\frac{\delta A^n_*}{\Delta A}. \tag{69}$$

Note that the correction term (last term on the right-hand side of (69)) is integrated over the departure area. Note that we use first-order trajectories for the cell divergence so that the Lagrangian divergence will be close to the Eulerian one and thus minimizing the magnitude of the correction term.

The scheme conserves mass both locally and globally. As discussed in connection with the ideal semi-implicit continuity equation (63), mass is conserved globally since the Lagrangian divergence terms give no contribution when integrated over the entire domain, assuming no mass flux through the boundaries of the domain. With the same assumption and given that a conservative method is used, the Eulerian divergences D^n and D^{n+1} integrated over the domain yield both zero.

The inconsistent semi-implicit term is corrected explicitly, which could be a potential source of instability. Numerically it was found that by choosing Φ_{00} sufficiently large, the scheme is stable for large Courant numbers. A similar condition for stability was derived by Simmons and Temperton (1997) as a result of a stability analysis performed on a two-time-level SISL baroclinic model. They found that the reference surface pressure should be higher than the actual surface pressure in order to have stable integrations. The phase speed of the gravity waves in the model is increased as Φ_{00} is increased, but that is normally of no concern in meteorological applications.

In SW-HIRLAM, as in many atmospheric models, a staggered Arakawa C-grid is used. That is, the mass point is situated in the center of a grid cell and the velocity components at the cell sides as indicated on Fig. 10. The principal argument for this choice is that, compared to a non-staggered A-grid with the same resolution, the truncation error of the centered second order finite difference approximations for the pressure gradient and the divergence are minimized. To keep the model as close as possible to SW-HIRLAM also here an Arakawa C-grid for the cell-integrated versions is used. Consequently, the discretization of the momentum equations and the Eulerian divergence

used in the correction procedure are as in SW-HIRLAM. In addition the elliptic equation is in the same form as the one in SW-HIRLAM when a C-grid is used.

The algorithm for the cell-integrated models is as follows:

- Interpolate the velocities to cell corner points (using cubic Lagrange interpolation) and compute the trajectories for cell vertices,
- add the correction term
$$\frac{\Delta t}{2}\,\Phi_{00}\,\left[D^n - \mathbb{D}(\mathbf{v}^n)\right], \tag{70}$$
to $\overline{\Phi}^n$,
- compute the upstream integral of
$$\overline{\Phi}^n + \frac{\Delta t}{2}\,\Phi_{00}\,\left[D^n - \mathbb{D}(\mathbf{v}^n)\right], \tag{71}$$
- compute $\mathbb{D}(\tilde{\mathbf{v}}^{n+1})$,
- now all known terms on the right-hand side of (69) are computed and the algorithm continues exactly as for the traditional system (e.g., Staniforth and Côté 1991).

Since a "half-implicit-Coriolis scheme" is used (see Appendix D), the elliptic equation is a Helmholtz equation. It is solved using the solver of HIRLAM, which applies a Fourier transform in the zonal direction (see Appendix D). The divergence is assumed zero on the boundaries.

When using the scheme of NM02 for the continuity equation the version of the new model is refered to as the PB-CISL shallow water model (PB-CISL-SWM) and the conservative cascade shallow water model (CC-CISL-SWM) when using the CC-CISL scheme. Note that the Lagrangian divergence is computed according to the departure cell geometry associated with the respective schemes.

3.2 Results of some tests

The widely accepted test suit proposed by W92 has been used to assess the accuracy of the model. The W92 test cases, however, are formulated for a global domain. Here the tests are run on a limited area domain, which, of course, requires frequent specification of boundary fields. The analytic solution or a reference solution is used to provide values at the boundaries. But in W92 the reference solutions were only provided daily, which is too infrequent for updating boundary fields in a LAM. Hence the high-resolution spectral model used in W92 has been rerun to provide

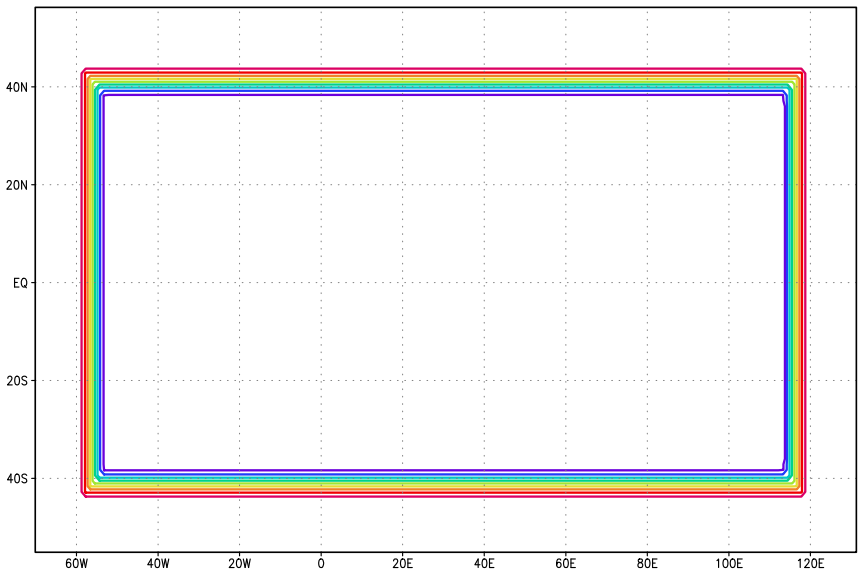


Fig. 11: The boundary relaxation function. The halo zone, where the fields are entirely specified by the reference solution, surrounds the 9° wide relaxation zone. The contours are from 0 to 0.9 and the contour interval is 0.1. The relaxation function increases monotonically from the inner domain toward the halo zone.

boundary data more frequently. A detailed description of the reference solutions and the parameters used for the reference model runs can be found in Jakob et al. (1993).

The LAM uses a rotated spherical latitude-longitude grid. Subsequent reference to the LAM grid refers to the rotated grid and not necessarily to a geographical latitude-longitude grid. The active grid points, i.e. grid points at which the solution is computed by the LAM, extends 180° in the longitudinal direction and from 45°S to 45°N on the LAM grid unless stated explicitly otherwise. Around the active domain there is a halo zone where the reference solution is prescribed. In a 9° wide zone adjacent to the boundaries of the active domain the solution is relaxed toward the boundary field using a relaxation coefficient, which is decreasing from one to zero over the relaxation zone (Davies 1976) (see Fig.11). The resolution for the runs is 1.125° unless stated explicitly otherwise. To minimize the errors introduced by the boundary relaxation, the boundary fields are updated with the reference solution at every time step. The error measures are computed every hour. It would be unfair to compare the performance of the LAM to the performance of global models because of the boundary relaxation. Instead a reference shallow water version of the semi-Lagrangian HIRLAM has been run. SW-HIRLAM uses decentering, filtering of the non-linear terms in time and the noise-reducing trajectory scheme of McDonald (1999) (for additional details see section 2c and Undén 2002). Since not all test cases require decentering for stable and noise-free integrations, the SW-HIRLAM has been run without decentering as well. The filtering of the non-linear terms is used in the semi-Lagrangian HIRLAM for stability reasons, but in the tests run here the accuracy was not affected by turning the filter on or off. Only when noise was

present the filter slightly reduced it.

For the SW-HIRLAM a constant geopotential Φ_{00} which is equal to the maximum geopotential height of the initial field, Φ^0_{max}, is used. For the cell-integrated models a constant geopotential which is 50% larger than Φ^0_{max} is used (see discussion in section 2d).

3.2.1 *Advection experiments (test case 1)*

Test case one consists of solid body rotation on the sphere of a cosine bell. For this test case a non-rotated (geographical) grid is used. The wind field is given by

$$u = u_0 \ (\cos\theta \ \cos\alpha + \cos\lambda \ \sin\theta \ \sin\alpha) \tag{72}$$

$$v = -u_0 \ \sin\lambda \ \sin\alpha \tag{73}$$

where α is the angle between the axis of solid body rotation and the polar axis and u_0 is a constant given by

$$u_0 = \frac{2\,\pi\,a}{(12 \text{ days})}, \tag{74}$$

where $a = 6.37122 \times 10^6$ m is the radius of the earth. Williamson and collaborators proposed advection along the geographical equator ($\alpha = 0.0$) and over the pole ($\alpha = \frac{\pi}{2}$;$\alpha = \frac{\pi}{2} - 0.05$). These orientations favor the transport scheme since the flow is approximately parallel to the coordinate axis. To complement the results published for the PB-CISL and the CC-CISL schemes, the advection test has been run for $\alpha = 30°$. The flow results in deformed departure cells which makes the problem more challenging than if the departure cells were rectangles as is the case for $\alpha = 0.0$. The cell deformations affect the upstream integral and divergence computations. For this experiment a global domain in the east-west direction and a resolution of $2.8125° \times 2.8125°$ is used.

Standard non-dimensional error measures (see e.g., W92) for the advection test for different schemes are listed in Table 4. As expected, errors have increased compared to advection for $\alpha = 0.0$ (compare to Table 1 in Nair et al. 2002). In general the CC-CISL outperforms the PB-CISL scheme. This is because the sub-grid-scale representation in the PB-CISL scheme eliminates the cross terms and therefore only includes variation along Eulerian latitudes and longitudes. The CC-CISL scheme may in certain cases include some of the diagonal variation since the second remapping is along the Lagrangian latitudes. In the present case the northern part of the cosine bell is compressed in the (λ, μ) coordinate system as it approaches the most northern position and vice versa for the most southern position. Obviously, this asymmetry is better resolved along sloping Lagrangian latitudes than along the Eulerian latitudes. When enforcing the P and SM filters the accuracy is improved. On the other hand the M filter severely damps and decreases the accuracy. It may be possible to improve the monotonicity filter by adapting the filter of Zerroukat et al. (2004, manuscript submitted to *Q. J. R. Meteorol. Soc.*) to the piecewise parabolic method. Note also that 1D filters that prevent undershoots and overshoots in one dimension do not necessarily guarantee

Scheme	l_1	l_2	l_∞	Min	Max
PB-CISL	0.075	0.051	0.083	-0.0088	-0.083
CC-CISL	0.051	0.039	0.076	-0.0070	-0.076
PB-CISL-P	0.043	0.040	0.082	-0.00079	-0.082
CC-CISL-P	0.033	0.034	0.077	0.0	-0.077
PB-CISL-M	0.077	0.089	0.18	-0.0038	-0.18
CC-CISL-M	0.070	0.086	0.186	0.0	-0.186
PB-CISL-SM	0.047	0.041	0.082	-0.0032	-0.082
CC-CISL-SM	0.035	0.034	0.076	0.0	-0.076
TRAD	0.25	0.15	0.15	-0.017	-0.151
TRAD-M	0.25	0.19	0.22	0.0	-0.217

Table 4: Standard non-dimensional error measures for solid body rotation of a cosine bell for the PB-CISL scheme, the CC-CISL and the traditional semi-Lagrangian (TRAD) scheme using bicubic Lagrange interpolation, respectively. The angle between the axis of solid body rotation and the polar axis is $30°$, the resolution is $2.8125°$ and one revolution is completed in 256 time steps. The letters P, M and SM denote positive definite, monotonic and semi-monotonic options, respectively (Lin and Rood 1996). The monotonic filter for TRAD is described in Bermejo and Staniforth (1992). Exact trajectories were used.

these properties when applied in each coordinate direction to the pseudo biparabolic sub-grid-cell representation of the PB-CISL scheme. In the case of negative values at the boundaries of both unfiltered 1D parabolic representations of a cell an even larger negative value may be present in one or more of the cell corners when the 1D representations are added. The 1D M, P and SM filters eliminate only the negative values at the boundaries and not the larger negative corner values. The same has been observed with the scheme of Lin and Rood (1996) and by NM02. The cascade scheme does not produce negative values after the application of the M, P or SM filter since the two 1D parabolic representations are applied successively and are never added. Using either the PB-CISL or CC-CISL scheme the accuracy is greatly improved compared to the traditional semi-Lagrangian scheme.

For a non-divergent flow the advection of a constant, e.g. Φ_{00}, should ideally yield Φ_{00} at all times. It is well known that CISL schemes do not have this property due to the approximation of the departure cell which fails to preserve the area exactly (e.g., Laprise and Plante 1995). Consequently the Lagrangian divergence computations are not strictly correct, whereas the Eulerian formula for divergence yields zero at the machine precision. To investigate the severity of the problem the constant field $\Phi_{00} = 50000$ m^2 s^{-2} has been advected one time step in the solid body rotation flow of test case one. The results for $\Delta t = 4050$ s are shown in Fig. 12 and 13 for exact trajectories and when using different departure point algorithms, respectively. The errors introduced by the inaccuracies in the departure cell approximations are of the order of 1 m^2 s^{-2} for both CISL schemes. Compared to Φ_{00} these errors are considered small. Note, however, that the departure

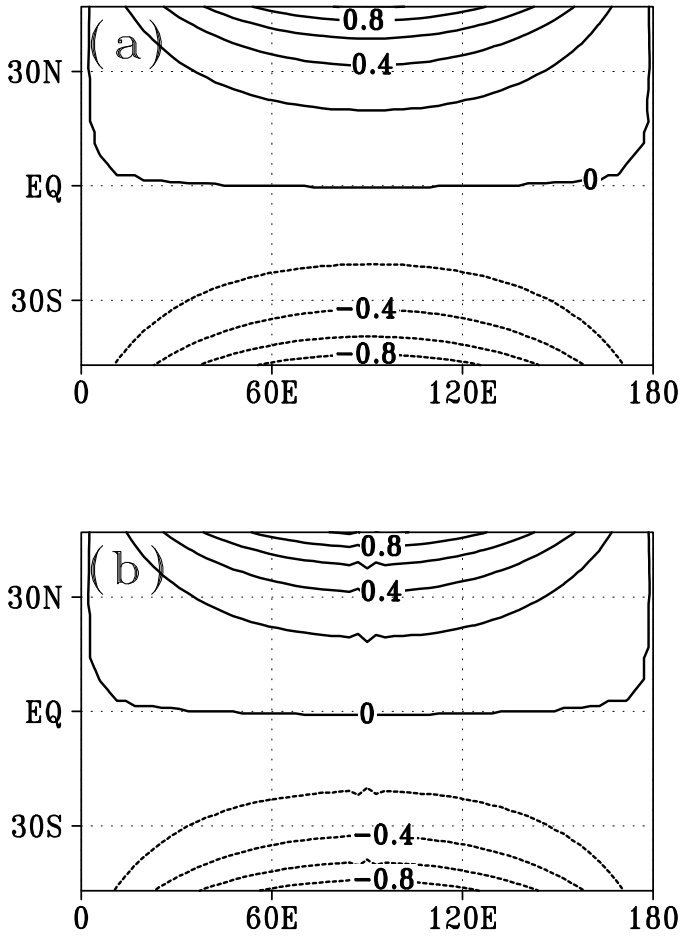


Fig. 12: The constant geopotential $\Phi_{00} = 50000$ m^2 s^{-2} has been advected one time step ($\Delta t = 4050$ s) using the solid body rotation flow of test case one and exact trajectories. The figure shows the deviation from Φ_{00} when using the (a) PB-CISL and (b) CC-CISL scheme, respectively. The contour interval is 0.2 m^2 s^{-2}.

point algorithm of McDonald (1999), which reduces noise in the baroclinic HIRLAM, negatively affects the accuracy of the trajectories.

3.2.2 *Non-linear geostrophic flow (test case 2)*

The flow field is a geostrophically balanced flow given by (140), (141) and

$$\Phi = g\,h_0 - \left(a\,\Omega\,u_0 + \frac{u_0^2}{2}\right) \times \left(-\cos\lambda\,\cos\theta\,\sin\alpha + \sin\theta\,\cos\alpha\right)^2, \tag{75}$$

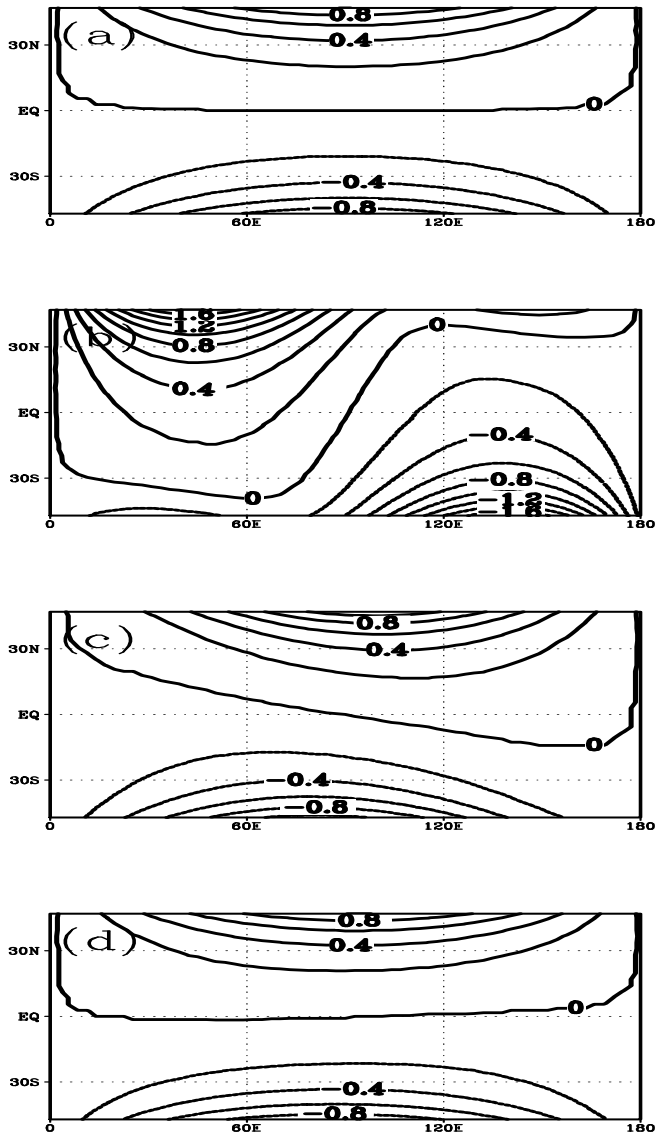


Fig. 13: Same as Fig. 12 but for approximated trajectories and only for the cell geometry used in the PB-CISL scheme. The departure point algorithms used are the (a) implicit-midpoint method, (b) method of McDonald (1999), (c) SETTLS and (d) new departure point algorithm, respectively. For all algorithms three iterations have been used with bilinear, biquadratic and bicubic interpolation for the first, second and third iteration, respectively.

model	resolution (°)	$l_1(h)$	$l_2(h)$	$l_\infty(h)$
SW-HIRLAM	0.5625	3.973E-6	5.931E-6	1.753E-5
SW-HIRLAM no decentering	0.5625	2.126E-6	2.645E-6	6.640E-6
PB-CISL-SWM	0.5625	1.588E-6	2.022E-6	4.293E-6
CC-CISL-SWM	0.5625	1.636E-6	2.037E-6	3.858E-6
SW-HIRLAM	1.125	1.231E-5	1.768E-5	5.077E-5
SW-HIRLAM no decentering	1.125	6.880E-6	1.036E-5	3.863E-5
PB-CISL-SWM	1.125	5.551E-6	7.185E-6	1.826E-5
CC-CISL-SWM	1.125	5.286E-6	6.695E-6	1.469E-5
SW-HIRLAM	2.25	4.829E-5	7.243E-5	2.195E-4
SW-HIRLAM no decentering	2.25	5.607E-5	5.607E-5	2.099E-4
PB-CISL-SWM	2.25	2.887E-5	3.773E-5	9.297E-5
CC-CISL-SWM	2.25	2.791E-5	3.715E-5	9.505E-5

Table 5: Standard non-dimensional error measures for the geopotential height field after 10 days of simulation for test case two. See text or Appendix C for definition of the acronyms. The maximum Courant number is approx. 3.5 for all resolutions, which corresponds to a time step of 900 s, 1800 s and 3600 s for the high, medium and low resolution runs, respectively.

where g is the acceleration due to gravity, $g\,h_0$ is a constant set to $29400\ \mathrm{m^2\ s^{-2}}$ and Ω is the rotation rate of the earth. Ideally the numerical solution should not change in time. Again a geographical grid is used and $\alpha = 30°$. Since an analytic solution exists the problem is well suited for verifying the convergence of the numerical algorithm. In the context of the PB-CISL scheme the test case is quite useful. Firstly the PB-CISL algorithm favors flow which is either along longitudes or latitudes which is not the case in this experiment, and secondly it serves to investigate if the errors in the Lagrangian divergence computations cause problems for the semi-implicit scheme.

Standard non-dimensional error measures for test case two are listed in Table 5 for different model configurations and resolutions. As expected, all models converge. The accuracy of the mass-conservative models is significantly better than the accuracy of the SW-HIRLAM. The difference between the error measures of the CISL models is small and probably insignificant, given that only a limited area is considered.

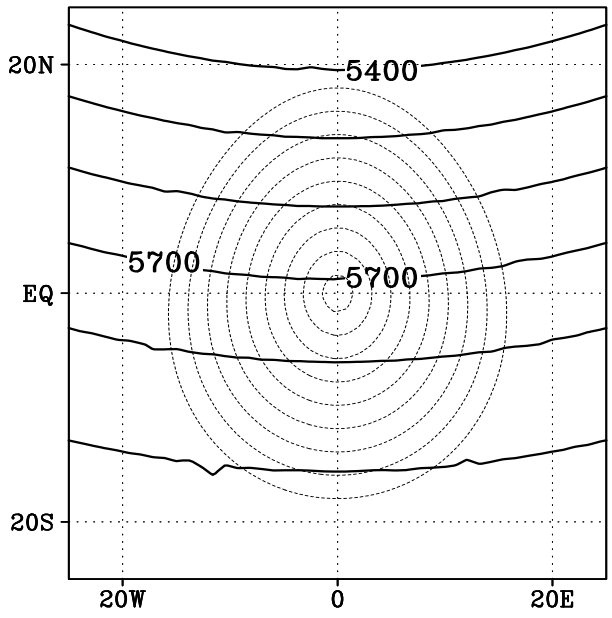


Fig. 14: (solid contours) The initial height field constructed on the LAM grid using the spectral coefficients of the T213 reference solution for test case five. The LAM grid is a rotated grid and the axis labels denote location on the LAM grid. Only part of the integration domain is shown. The solid line contour interval is 100 m. (dashed contours) The geopotential height of the orography, Φ_s. The dashed line contours start at 2000 m^2 s^{-2} and increase monotonically to 20 000 m^2 s^{-2} with an increment of 2000 m^2 s^{-2}. Note the noise in the height field near the rim of the mountain. The maximal deviation from the analytical initial condition is approximately 100 m^2 s^{-2}.

3.2.3 *Flow over an isolated mountain (test case 5)*

Test case five consists of the fields of test case two with $u_0 = 20$ m s^{-1} and $h_0 = 5960$ m and a surface orography given by

$$h_s = h_{s0} \left(1 - \frac{r}{R}\right), \tag{76}$$

where $h_{s0} = 2000\, m$, $R = \pi/9$, and

$$r^2 = \min\left[R^2, (\lambda - \lambda_c)^2 + (\theta - \theta_c)^2\right], \tag{77}$$

where $(\lambda_c, \theta_c) = (3\,\pi/2, \pi/6)$. The rotated south pole is located at $60°$S, $270°$W so that the peak of the mountain is at the origin of the LAM grid. The initial condition on the LAM grid is shown in Fig. 14. There is no analytic solution to this problem, so the high resolution spectral model is used to produce a reference solution. The representation of the mountain, however, is a challenge for the spectral scheme due to Gibbs's phenomena. As shown in Fig. 14 small-scale noise is present near the rim of the mountain already in the initial state. The maximum difference in the geopotential height field between the analytic initial condition and the initial condition constructed

using spectral coefficients is approx. 100 m^2 s^{-2}. In addition to the inaccuracies introduced by Gibbs's phenomena, the flow will initially experience an abrupt change due to the presence of the mountain, which suddenly is raised in the balanced flow. This creates an initial imbalance that excites gravity waves. The gravity waves excited by the imbalance in the initial condition will complete one revolution in approx. 44 hours. The phase and amplitude of the gravity waves in the global reference solution and in the LAM numerical solutions are not identical due to the boundary relaxation and the models representation of the gravity waves. Therefore extrema in the error measures appear approximately every second day when gravity waves enter the domain is seen (see Fig. 15). Since Φ_{00} is large in the CISL models, they perform worse than SW-HIRLAM as the gravity waves sweep through the domain. Otherwise the cell-integrated models perform much better than than SW-HIRLAM, while the cell-integrated models perform equally well. When not using decentering in the SW-HIRLAM the accuracy increases significantly but, as discussed in the next section, decentering in the SW-HIRLAM is needed to alleviate noise problems for general flows at high Courant numbers.

3.2.4 *Modified test case 5*

Semi-Lagrangian methods are known to produce spurious numerical resonance for flow over orography at high Currant number unless special measures are taken (e.g., Coiffier et al. 1987; Kaas 1987; Tanguay et al. 1992; Rivest et al. 1994; Lindberg and Alexeev 2000). Linear one-dimensional theory predicts noise problems when the flow satisfies

$$\frac{2\,u\,\Delta t}{L_m} = \pm 1, \pm 3, \pm 5, .., \tag{78}$$

where L_m is the wavelength of the mountain (Lindberg and Alexeev 2000). For the parameters of test case five and a time-step of 30 minutes the wavelength of the mountain should be less than 100 km in order to trigger resonance when using the traditional semi-Lagrangian scheme. Such a mountain cannot be resolved with the 1.125° resolution used in test case five. To stimulate the orographic noise set $u_0 = 50\,\mathrm{m\,s^{-1}}$, $\Delta t = 40$ minutes, and elongate the mountain in the north-south direction so that it extends throughout the domain (see Fig. 16). More precisely the mountain is given by (76) with (77) replaced by

$$r^2 = \min\left[R^2, (\lambda - \lambda_c)^2\right]. \tag{79}$$

The boundary field is not updated with data from a new high-resolution run, but is held fixed at the original initial condition since the focus is on the noise problem and not on accuracy.

In Fig. 38 the simulation using the SW-HIRLAM with and without decentering is shown as well as the results from the CISL models. The spatially averaging of the non-linear terms (54) and the filtering of the non-linear terms in time alleviate the noise problem for the traditional scheme, but do not eliminate it. When using decentering the noise is reduced to an acceptable level. On the contrary, the CISL models produce noise-free forecasts without using any decentering or filtering.

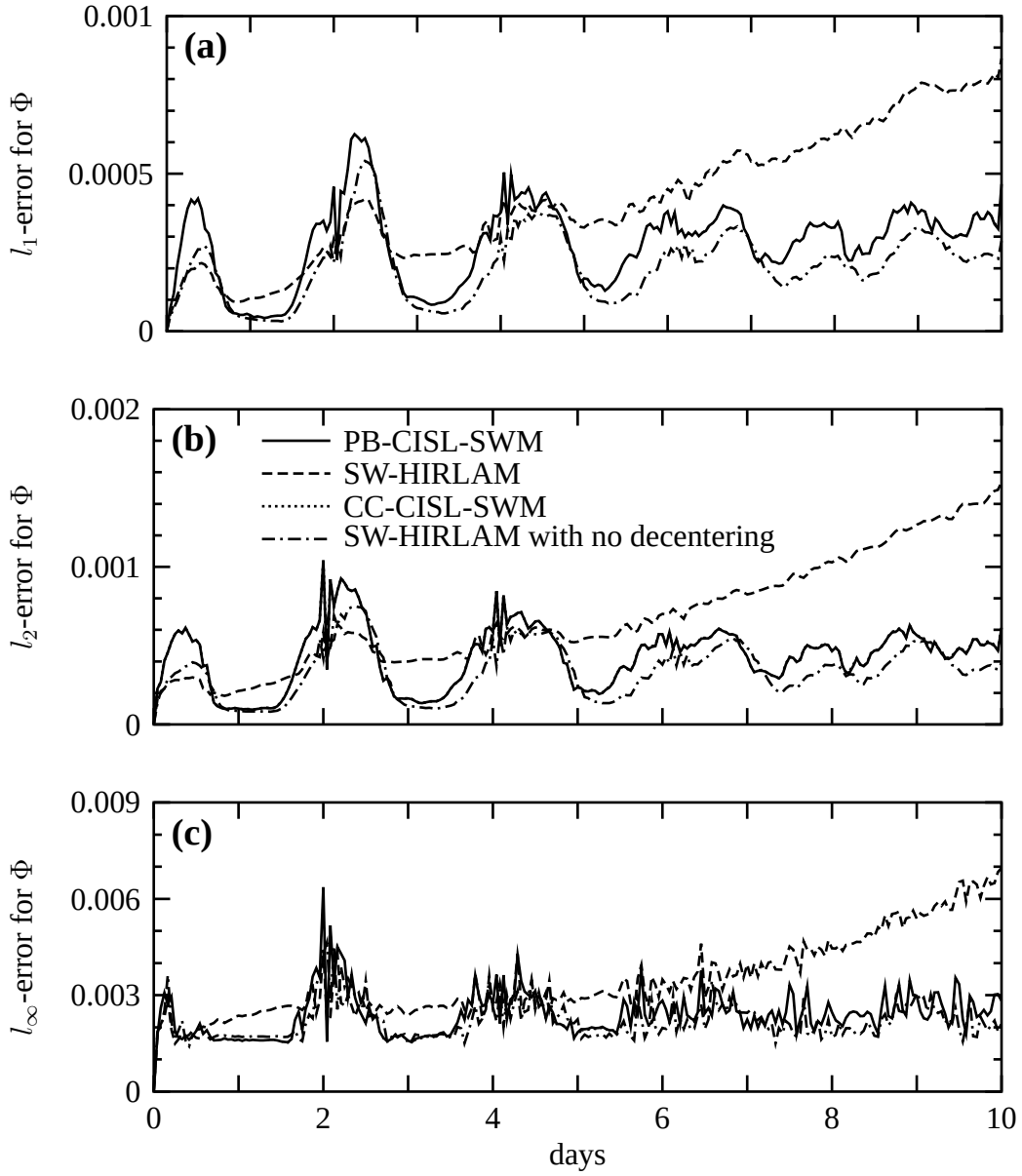


Fig. 15: Standard non-dimensional error measure (a) l_1, (b) l_2 and (c) l_∞ for the geopotential height as a function of time for test case five when using the (dash line) SW-HIRLAM, (dash-dot line) SW-HIRLAM without decentering, (solid line) PB-CISL-SWM and (dot line) CC-CISL-SWM, respectively. The x-axis are the same for all plots. The time step is 30 minutes, which corresponds to a maximal Courant number of approx. five. The boundary field is updated at every time step using the reference solution. The error measures for the CISL models are indistinguishable on the figure.

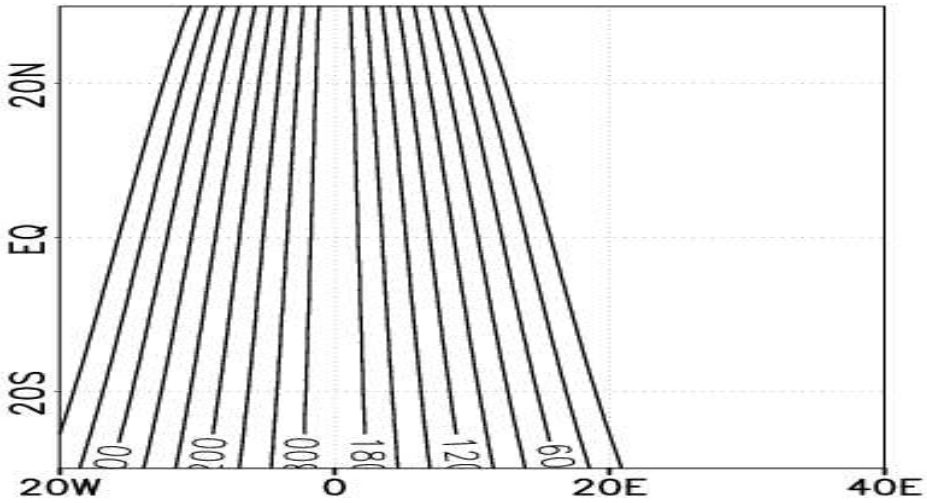


Fig. 16: Height of the mountain used in the modified test case 5. Countour interval is 200 m.

It could be speculated that the CISL models are susceptible to the noise problems at other flow regimes. The modified test case five was run with maximum Courant numbers ranging from approximately one to eight without encountering noise problems. The conservative CISL schemes are based on cell averages and it is suggested that consequently the resonance does not develop since it is *averaged out*.

3.2.5 *Rossby-Haurwitz wave (test case 6)*

The initial condition is a Rossby-Haurwitz wave with wavenumber 4. The problem has an analytic solution for the nonlinear barotropic vorticity equation on the sphere, but not for the nonlinear shallow water equations. It was previously believed that wavenumber 4 is stable (Hoskins 1973), but Thuburn and Li (2000) have shown that, if a perturbation is introduced, small-scale features develop in the vorticity field due to an instability of the flow. This, of course, questions the accuracy of the reference solution. Sufficient diffusion of the numerical method or explicitly added diffusion tends to suppress the instability. However, for the high-resolution models used in Thuburn and Li (2000), which include the spectral model used here, the dynamical instability first appears by day 10 of the simulation. Hence for the test runs here, which are less than 10 days, the instability may not degrade the accuracy of the reference solution.

It is difficult to argue that one particular region has more activity than others. Therefore the LAM grid is chosen to coincide with the geographical latitude-longitude grid.

The Rossby-Haurwitz wave propagates slowly eastward with small vacillations about the steady propagation wave structure. The error measures for the different models are shown in Fig. 18. The numerically simulated wave propagates with a phase speed close to the one predicted by the

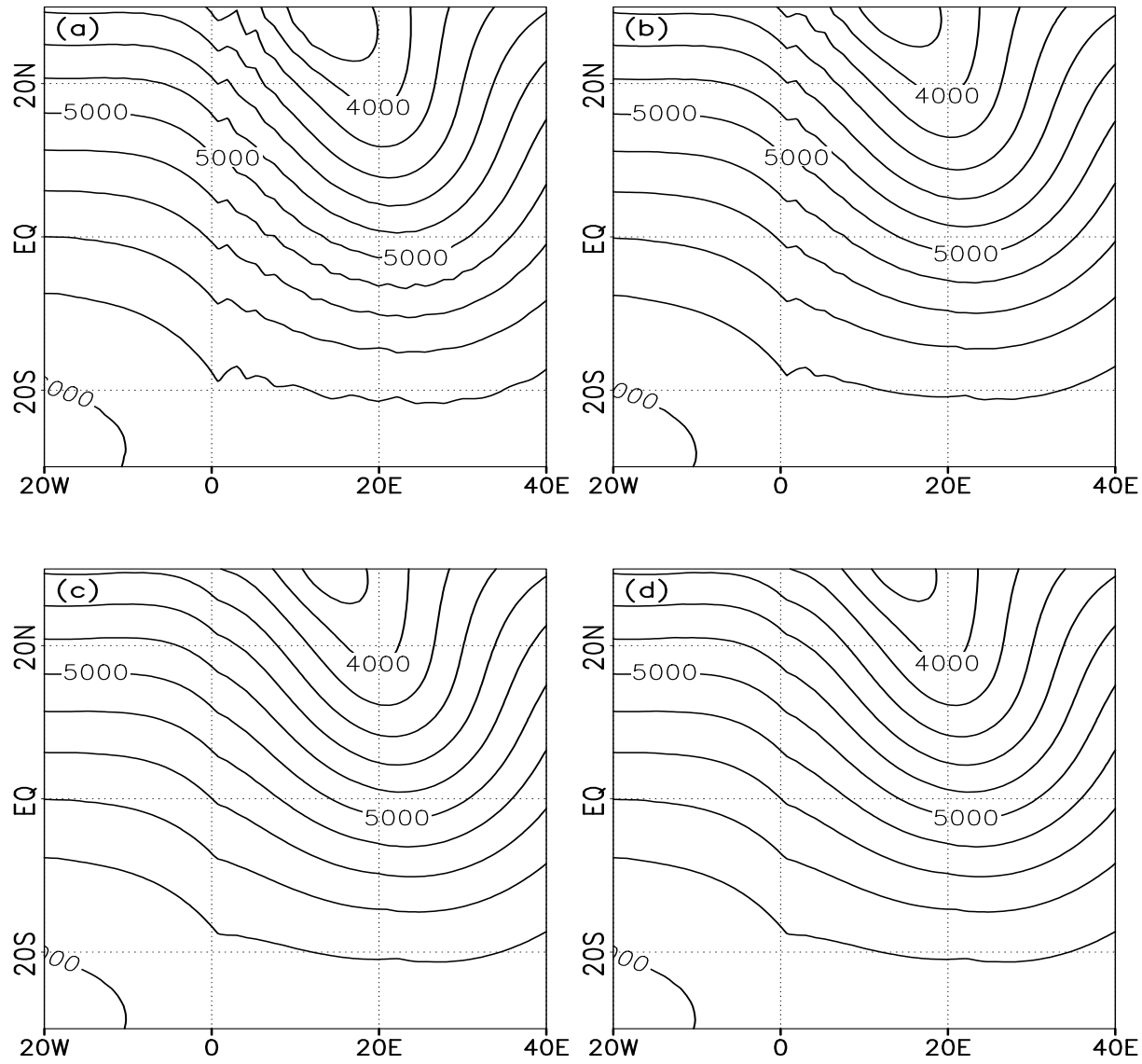


Fig. 17: The height of the free surface on the LAM grid at day one for the modified test case five when using the (a) SW-HIRLAM with no decentering, (b) SW-HIRLAM, (c) CISL-SWM and (d) CC-CISL-SWM, respectively. The axis labels denote location on the rotated grid and only part of the integration domain is shown. The contour interval is 200 m. Noise develops downstream of the mountain when using the traditional semi-Lagrangian method without decentering. The CISL methods use no decentering and show no sign of noise problems.

non-divergent barotropic model. The noise in the error measures is primarily due to the vacillations. The CISL models are as accurate as the SW-HIRLAM and perform equally well. The

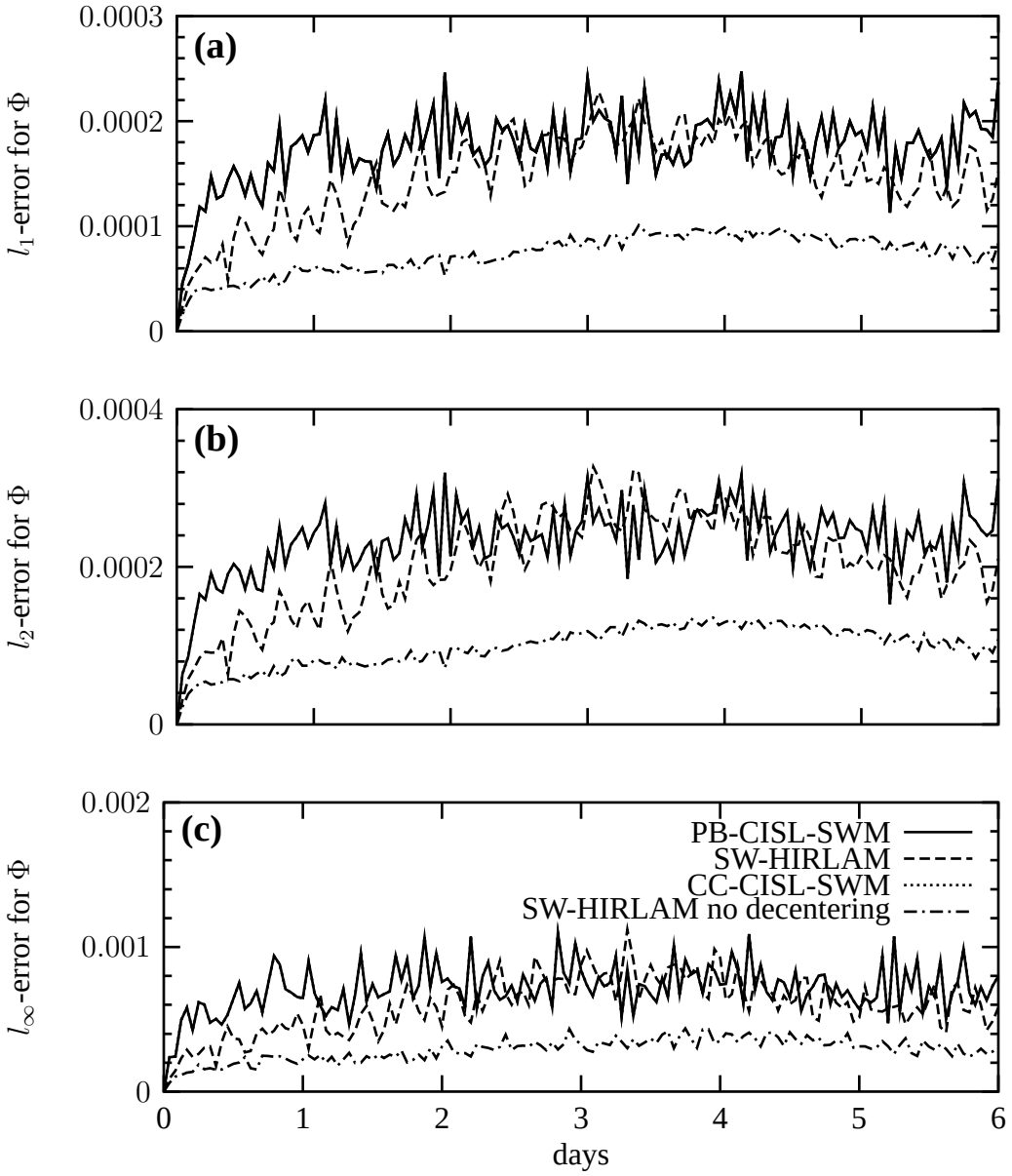


Fig. 18: Same as Fig. 15 but for test case six and a time step of 12 minutes. The error measures for the CISL models are indistinguishable on the figure.

SW-HIRLAM without decentering, however, performs significantly better than all other model configurations, but in general decentering is needed to alleviate noise problems.

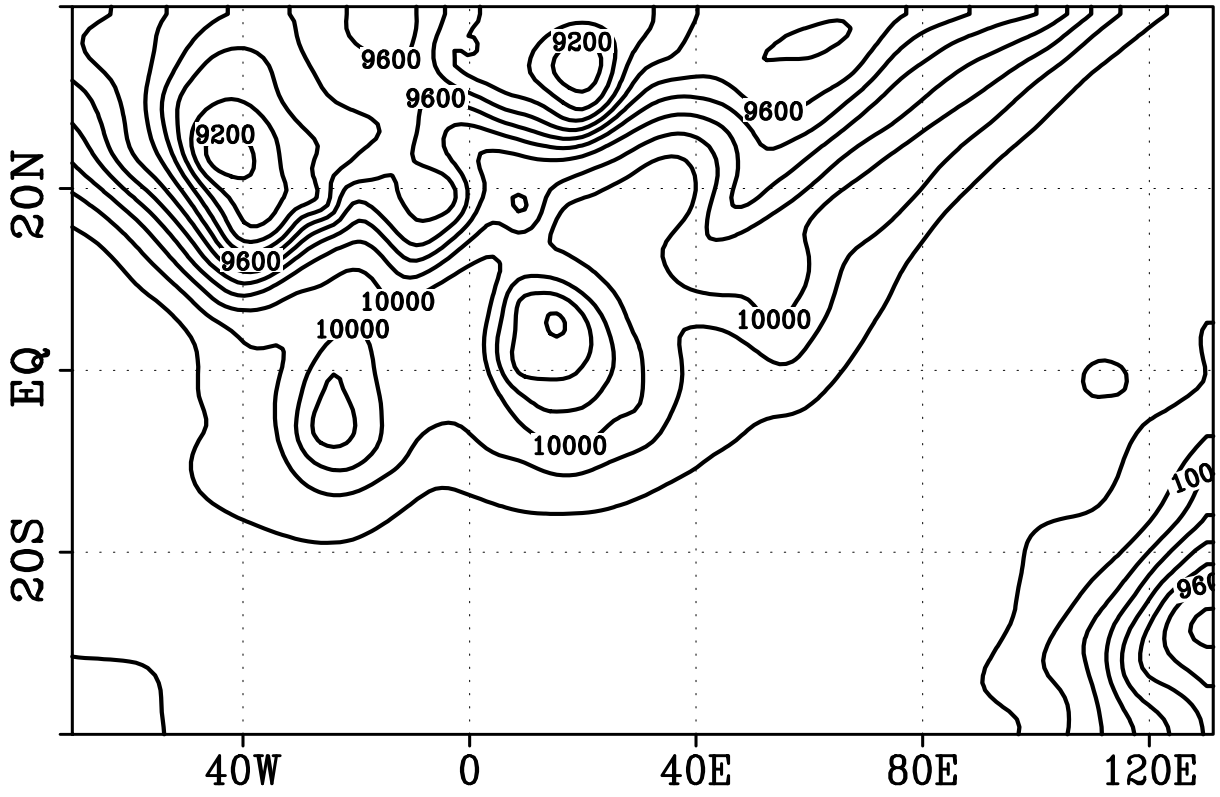


Fig. 19: The initial height field in meters on the LAM grid for test case seven. The rotated south pole is located at $40°$S, $0°$W and the axis labels denote position in the rotated grid. The contour interval is 100 m.

3.2.6 *Analyzed 500 mb height and wind field initial conditions (test case 7)*

Three different initial conditions for the 500 mb height and winds were proposed in W92. These initial conditions were chosen more or less randomly and in this context there is no particular reason for choosing one of the situations over the others. Here the initial data from GMT January 9, 1979, are used and the limited area is placed where a blocking develops (see Fig.19).

The error measures are shown in Fig. 20. The CISL models perform better than the traditional model with respect to l_1 and l_2. l_3 is slightly larger for the CISL models up to about day 4, after which the CISL models perform clearly better.

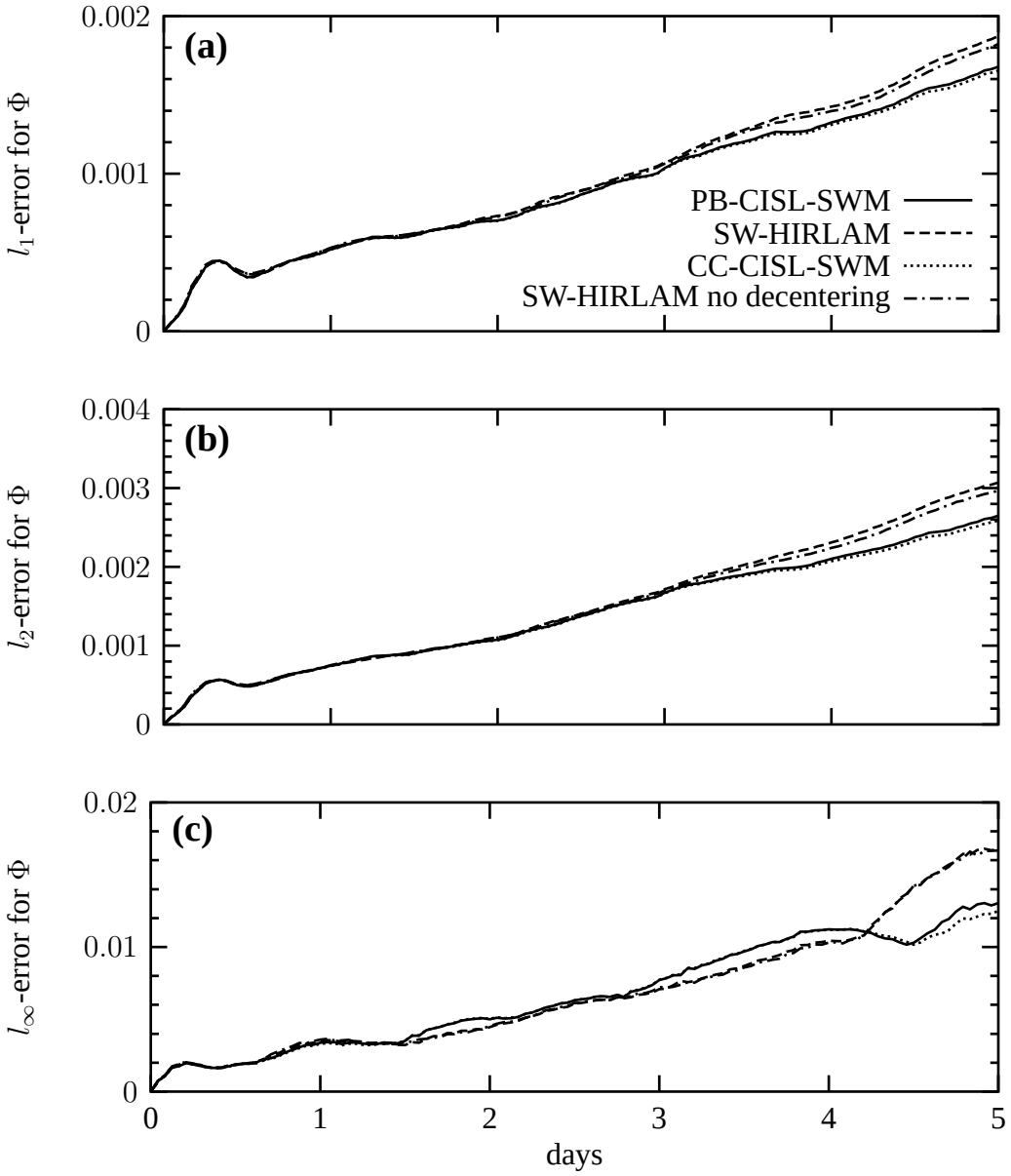


Fig. 20: Same as Fig. 15 but for test case seven and a time step of 24 minutes.

3.2.7 *Computational cost*

Ultimately the efficiency of using CISL schemes may only be discussed in full-blown-3D models implemented in multi-processing environments. However, the horizontal part of a 3D hydrostatic model is similar to a shallow water system, so it is important to discuss how the efficiency is

affected by going from passive CISL advection to a semi-implicit CISL shallow water model and to discuss the differences in the two CISL models tested here.

The CISL models have been implemented within the framework of the HIRLAM code without any attempts to optimize the code. The SW-HIRLAM code, however, is optimized and it would therefore make little sense to compare the actual CPU (Central Processing Unit) times for the two types of models. The computational cost for the momentum equations and the elliptic equation solver is identical for all model configurations. The explicit continuity equation is solved more efficiently using the CC-CISL scheme compared to the traditional method and less efficiently with the PB-CISL scheme. In general, the CISL schemes use more memory than conventional methods, but that would have little effect on execution time on massively parallel computers since the extra memory usage consists of local information only. Computation of the correction term is not needed for the traditional system. The correction term is computed very efficiently with the PB-CISL scheme since the departure cell area is defined explicitly in terms of the departure points. In the CC-CISL scheme the intersection between the Lagrangian latitudes and Eulerian longitudes must be computed in order to define the departure cell geometry, which reduces the computational advantage the CC-CISL scheme has over the PB-CISL scheme for explicit advection. Weightening these differences, a small computational overhead for the CISL models compared to the conventional model is expected for an efficient implementation. With the present coding it is not possible to decide about the relative efficiency of the two CISL models.

3.3 Possible extensions to a global domain

The model presented here is for a limited area and for wider applications an extension to a global domain is, of course, important. Two methods to extend the model to a global domain are suggested. One extension uses a global latitude-longitude grid, which requires a certain amount of "engineering" and approximations near the poles. Perhaps a more viable approach to avoid the pole problem is to use the so-called "Yin-Yang" grid for spherical geometry (Kageyama and Sato 2004).

The first extension suggested requires the momentum equations to be solved on vector form rather than on component form to avoid instability near the poles (e.g., Ritchie and Beaudoin 1994). In addition, the "half-implicit-Coriolis scheme" must be replaced by a fully implicit scheme and a vector form trajectory algorithm must be used. Then the semi-implicit CISL scheme for the continuity equation must be extended to a global domain.

Both CISL advection schemes used here have originally been formulated on the sphere and both use the same local approach for accurate transport over the polar regions, which allows for large meridional Courant numbers (NM02). In the vicinity of the poles a high-resolution tangent plane is introduced. Apart from the two cells which include the Eulerian poles, the departure cells are "well defined" on the tangent plane and the remapping is formally equivalent to the regular case. For the

polar singularity a “latitudinal” belt of Lagrangian cells which contain the Eulerian pole, referred to as the “singular belt”, is defined. The total mass inside the “singular belt” can be computed using the regular algorithm on the tangent plane. Thereafter the total mass is approximately distributed to each cell using the weights equal to the mass at the midpoints of each cell computed by the traditional semi-Lagrangian method .

To extend the treatment of the polar regions used in the CISL advection schemes to the semi-implicit ones, the Lagrangian divergence must be computed. Apart from the “singular belt”, the departure cells are “well defined” and the Lagrangian divergence is easily computed. For the “singular belt” the extension is also rather straight forward. An approach could be to advect a non-zero constant. Then the divergence will be readily available since the deviation from constancy is proportional to the change in area. Therefore, by using the same method of distribution of mass in the “singular belt” outlined above, also the divergence can be made available. Hence theoretically there should be no obvious obstacles in extending the limited are model to a global domain using this approach, but it does imply some algorithmic complexity.

Alternatively, the pole problem can be addressed by using the “Yin-Yang” grid recently developed for spherical geometry. This orthogonal grid is based on the overset grid methodology, i.e. two overlapping subgrids where each subgrid is a spherically rotated latitude-longitude grid. Thus, two LAMs are effectively coupled. Hence the model presented here can be applied without modifications on each subgrid. The pole problem is avoided at the expense of introducing an overlap region. This region is, however, static and mass in the interface region can be conservatively remapped using an approach such as the pseudo-piecewise parabolic method used here in one of the CISL model versions. A problem that must be solved is how to couple the elliptic systems of the two LAMs.

Chapter IV

Extension to a hydrostatic limited area model

The second step toward the long-term goal of designing an inherently mass-conservative SISL global model is presented in this chapter. Here the limited area shallow water model described in the previous chapter is extended to three dimensions. As in the shallow water model the CISL method is used for the continuity equation and the momentum equations are solved with the conventional semi-Lagrangian grid-point form. The thermodynamic equation is discretized on grid-point form as well, but several discretization options are considered.

An obvious question regarding the extension of the shallow water model to a baroclinic model is how to tackle the three-dimensional transport problem. Existing CISL methods are only two-dimensional and a fully three-dimensional upstream transport would be geometrically complex and computationally very expensive (unless special methods such as a cascade interpolation in the vertical is considered). Instead (Machenhauer and Olk 1998) suggested a hybrid trajectory method consisting of the upstream horizontal and downstream vertical transport of mass. They suggested the computation of upstream integrals of mass which arrive from each model layer to a particular vertical column of regular grid cells (see Fig. 21). The mass originating from a particular model layer was supposed to move with vertical walls along backward horizontal trajectories originally determined from winds in the layer in question, and its vertical displacement was supposed to be determined indirectly by the hydrostatic assumption. That is, the pressure at the top and the bottom of the transported mass should be computed in the arrival column as the weight of the mass above and below the transported mass. In particular, the surface pressure at the arrival grid column should be computed as the weight of the total mass transported to the column from all the model layers. Later Machenhauer (personal communication) suggested that this should be used to determine a mean vertical pressure velocity, $\omega = dp/dt$, along each trajectory, which could be used to determine the vertical part of the trajectory in the next iteration; or in other words the level to which the horizontal winds should be interpolated to. Furthermore, the mean vertical velocity in pressure coordinates could be used in the energy conversion term of the thermodynamic equation.

Thus, by this hybrid trajectory method the explicit solution to the continuity equation is completely consistent with the horizontal flow and the hydrostatic assumption. The hybrid trajectories are used to determine the semi-Lagrangian transport of the prognostic variables in the model. The parcel trajectories do, however, not necessarily arrive at model levels and therefore a vertical mapping of the prognostic variables back to model levels is needed after the completion of each time step. The problem is, however, one-dimensional.

For the models presented here the NM02 scheme is used, but in principle any two-dimensional CISL scheme formulated in spherical geometry can be applied. The baroclinic model has been developed within the framework of HIRLAM and hence some similarity to the HIRLAM system has been retained. For example, the Arakawa C-grid, the component form discretization of the momentum equations and the limited area domain are retained.

Several versions of the new inherently mass-conservative dynamical core are presented. The reference version is a “minimal modifications to the HIRLAM” version which use the discretizations of the momentum and thermodynamic equations of HIRLAM only modified for the hybrid trajec-

tory. This version is referred to as the reference CISL HIRLAM (R-CISL-HIRLAM). In HIRLAM and R-CISL-HIRLAM the energy conversion term in the thermodynamic equation is not consistent with the discretized continuity equation used in the respective models. A modified version of R-CISL-HIRLAM is presented, in which the discretized conversion term in the thermodynamic equation is consistent with the discretized semi-implicit CISL continuity equation. This model version will be referred to as the "consistent omega-p" CISL HIRLAM (COP-CISL-HIRLAM).

In section 1 the new R-CISL-HIRLAM is presented. First the continuous equations of motion are reformulated for the new dynamical core. Then the discretization grid is defined, in particular the placement of the cells with respect to the Arakawa C-grid. The discretization of the hydrostatic equation is very similar to traditional semi-Lagrangian models where the 'full-level' geopotential is defined in terms of a free parameter α (Simmons and Burridge 1981, hereafter referred to as SB1981). In traditional models the 'full-level' pressure is not explicitly needed, but in the new dynamical core it is. An approximation consistent with the choice of α is used. The new hybrid trajectory scheme is a three-dimensional extension of the algorithm used in the shallow water model (see chapter III). The hybrid approach involves the solution to the explicit CISL continuity equation and therefore the discretization of the continuity equation is presented in connection with the trajectory algorithm. Hereafter a derivation of the semi-implicit CISL continuity equation, which is a direct extension of the shallow water one, is given. The remaining equations of motion are discretized as in HIRLAM, but modified for the hybrid trajectory, i.e. the upstream integral/interpolation is only two-dimensional and at the end of each time step a vertical mapping from the grid implied by the semi-implicit CISL continuity equation back to the model levels is needed.

Section 2 considers the COP-CISL-HIRLAM where the discretization of the energy conversion term is consistent with the discretized semi-implicit CISL continuity equation. The linear part for the semi-implicit thermodynamic equation must be reformulated for the new discretization of the conversion term.

In section 3 the accuracy of the model is assessed using the idealized baroclinic wave test case recently developed by Jablonowski and Williamson (2005, hereafter referred to as JW2005). The new dynamical cores have been implemented within the framework of HIRLAM and thus the domain can not include the poles; neither can periodic boundary conditions be used in the east-west direction. Thus the boundary fields are held fixed at the initial condition. Since the reference solutions provided by JW2005 are for a global domain, the comparison with LAM runs can be precarious. It can not be ruled out that the boundaries significantly alter the LAM solution compared to a global model run. The new dynamical core can, however, be compared with HIRLAM runs for which the boundaries are treated in exactly the same way. Several versions of the new dynamical core are intercompared to demonstrate the influence of inherent mass conservation, consistent energy conversion term discretization, horizontal resolution and the number of iterations used in the trajectory algorithm.

At this point the conservation properties of the vertical discretization have not been considered. In

Eulerian models the discretization of the energy conversion term and the hydrostatic equation is such that the vertical discretization is not a spurious source or sink for total energy. The conservation properties of the vertical discretization in traditional semi-Lagrangian models and the new dynamical cores are discussed in section 4.

4.1 Reference CISL HIRLAM

4.1.1 *Reformulation of the continuous equations of motion*

Consider the equations of motion for a hydrostatic and moist-diabatic atmosphere using the vertical coordinate, $\eta(p, p_s)$, introduced by SB1981

$$\frac{d\mathbf{v}}{dt} = -\nabla\phi - \frac{R_d\, T_v}{p}\nabla p - f\,\mathbf{k}\times\mathbf{v} + \mathbf{P} + \mathbf{K}_v, \tag{80}$$

$$c_p\,\frac{dT}{dt} = R_d\, T_v\frac{\omega}{p} + P_T + K_T, \tag{81}$$

$$\frac{dq}{dt} = P_q + K_q, \tag{82}$$

$$\frac{d}{dt}\left(\ln\frac{\partial p}{\partial\eta}\right) = -D - \frac{\partial\dot{\eta}}{\partial\eta}, \tag{83}$$

$$\frac{\partial\Phi}{\partial\eta} = -\frac{R_d\, T_v}{p}\frac{\partial p}{\partial\eta}, \tag{84}$$

(e.g., White 2001) where

$$T_v = T\left[1 + \left(\tfrac{1}{\epsilon} - 1\right)\, q\right], \tag{85}$$

$$\epsilon = \frac{R_d}{R_v}, \tag{86}$$

$$\frac{d}{dt} = u\frac{\partial}{\partial t} + v\frac{\partial}{\partial t} + \dot{\eta}\frac{\partial}{\partial\eta}, \tag{87}$$

$$c_p = c_{pd}\,\left[1 + \left(\tfrac{c_{pv}}{c_{pd}} - 1\right)\, q\right] \tag{88}$$

and $\dot{\eta}$ is the η-coordinate vertical velocity. The terms K_T, K_q and $\mathbf{K}_v$ represent horizontal diffusion processes , and P_T, P_q and $\mathbf{P}_v$ represent physical processes. The symbols are conventional and are listed in Appendix A.

The continuity equation (83) is converted to cell-integrated form by replacing the right-hand side , which is the three-dimensional divergence, with its Lagrangian form

$$\frac{1}{\delta V}\frac{d\,\delta V}{dt}, \tag{89}$$

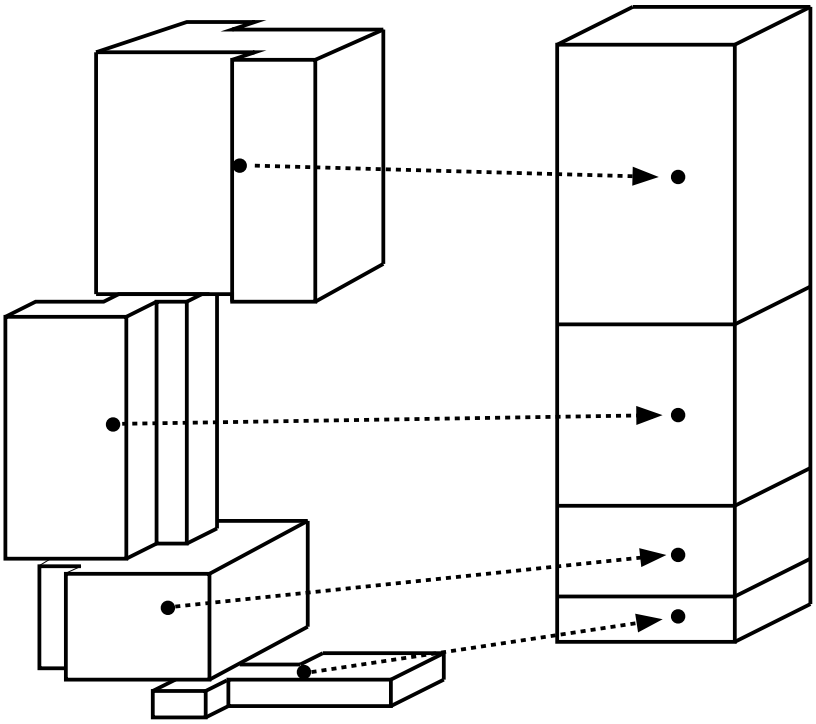


Fig. 21: Schematic illustration of the departure and arrival cells which make up the deformed column on the left and the regular column on the right, respectively (in the special case where all cells in each model level have the same pressure thickness; in general the departure cell thickness will be fitted with a piecewise quasi-biparabolic function). The cells move with vertical walls and the horizontal extension is a polygon as in the two-dimensional CISL scheme of Nair and Machenhauer (2002). The filled circles indicate the center of mass of the departure and arrival cells, respectively. Note that the vertical levels in the arrival column are the ones implied by the advection scheme and not the new model levels based on the new hydrostatically determined surface pressure, p_s^{n+1}, and the predefined hybrid coefficients (92).

where δV is an infinitesimal volume element in the (λ, θ, η)-coordinate system. Rearranging terms and making use of the chain rule for differentiation, it may be shown that the continuity equation can be written as

$$\frac{d}{dt}\left(\frac{\partial p}{\partial \eta}\,\delta V\right) = 0, \tag{90}$$

(Machenhauer 1992). The integro-differential form of the continuity equation (90) can be used to write the moisture equation, (82), in the following form

$$\frac{d}{dt}\left(q\,\frac{\partial p}{\partial \eta}\,\delta V\right) = \left(\frac{\partial p}{\partial \eta}\,\delta V\right)\,(P_q + K_q)\,. \tag{91}$$

Equations (80), (81), (84), (90) and (91) form the basis for the discretizations in the new dynamical cores.

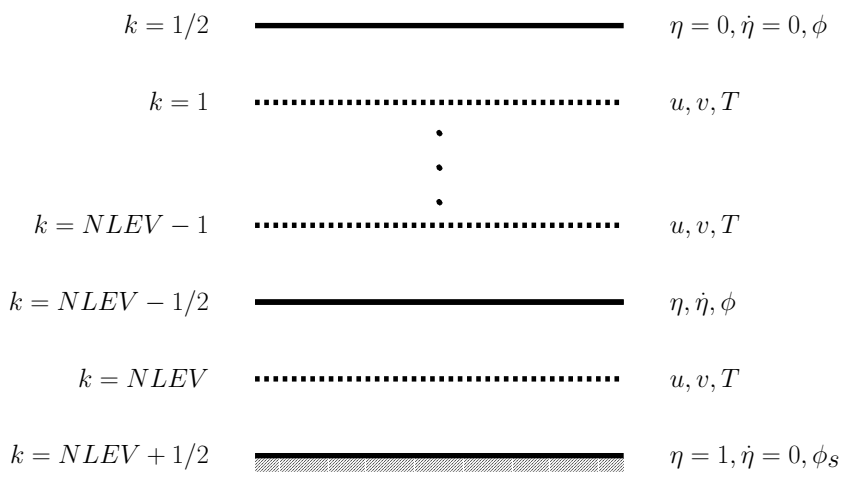


Fig. 22: The vertical staggering and indexing of the dependent variables. The atmosphere is divided into $NLEV$ layers and the indexes are decreasing with height from $k = NLEV + 1/2$ to $k = 1/2$. At the top and bottom of the atmosphere the vertical velocity is held at zero.

4.1.2 *Discretization grid*

In the horizontal an Arakawa C-grid is used (see Mesinger and Arakawa 1976 for the definition of grids A-E). The cells move with vertical walls (see Fig. 21). Horizontally the cells are defined such that the mass points are in their centers and the sides pass through the velocity points (see Fig. 10 in Chapter III). The indexes of the cells correspond to those of the mass point located in their center.

In the vertical the atmosphere is divided into $NLEV$ layers and the layer index is increasing toward the surface. The pressure at the interface between layer k and $k + 1$ is defined by

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2}\, p_s, \tag{92}$$

where $A_{k+1/2}$ and $B_{k+1/2}$ are predefined constants, p_s is the surface pressure and $0 \leq k \leq NLEV$. The vertical index $k + 1/2$ refer to a 'half level' and the pressure $p_{k+1/2}$ is the location of the 'bottom' of the cell with vertical index k. In the models presented here the top level pressure $p_{1/2}$ is zero. The pressure thickness of layer k is denoted $\Delta p_k = p_{k+1/2} - p_{k-1/2}$. In between the 'half levels' are the 'full levels' and their exact location is discussed in the next section. The vertical staggering of the dependent variables is a Lorenz-type staggering in which the temperature, u, v and q are at the 'full levels' and the geopotential, pressure and vertical velocity are at 'half levels' (Lorenz 1960). See Fig. 22 for additional details.

4.1.3 *Hydrostatic equation*

When using the Lorenz (1960) staggering of the variables the geopotential height is held at 'half levels'. In the momentum equations the 'full level' geopotential height Φ_k is needed since the pressure gradient force (PGF) involves the horizontal gradient of Φ_k. Hence the discretization of the PGF is related to the hydrostatic equation.

'Full level' geopotential height The finite difference analog of the hydrostatic equation (84) for a single layer and for the airmass from surface and up to 'half level' $k+1/2$ is given by

$$\Phi_{k+1/2} - \Phi_{k-1/2} = -R\,T_k\,\left(\Delta \ln p\right)_k\,, \tag{93}$$

$$\Phi_{k+1/2} = \Phi_s + R\sum_{\ell=k+1}^{NLEV} T_\ell\,\left(\Delta \ln p\right)_\ell\,, \tag{94}$$

respectively, where

$$\left(\Delta \ln p\right)_k = \ln\left(\frac{p_{k+1/2}}{p_{k-1/2}}\right). \tag{95}$$

To get Φ_k the almost universally adopted approach by SB1981 is used

$$\Phi_k = \phi_{k+1/2} + \alpha_k\,R_d\,(T_v)_k, \tag{96}$$

where the last term on the right-hand side of (96) is the approximation to

$$R\int_{p=p_k}^{p=p_{k+1/2}} (T_v)_k d\ln p. \tag{97}$$

Under the assumption that $(T_v)_k$ is constant over the half layer, $p = p_k$ to $p = p_{k+1/2}$, α_k is the approximation of the integral

$$\int_{p=p_k}^{p=p_{k+1/2}} d\ln p. \tag{98}$$

SB1981 approximates α_k with

$$\alpha_k = \begin{cases} \ln 2 & ,k=1, \\ 1 - \frac{p_{k-1/2}}{\Delta p_k}\,\left(\Delta \ln p\right)_k & ,k=2,..,NLEV, \end{cases} \tag{99}$$

(see SB1981 for details). This approximation shall be used in the models presented here as well.

For the discretization of the trajectory algorithm the pressure (which is held at 'half levels') is needed at 'full levels'. (98) can be used to define the 'full level' pressure consistent with the definition of α_k

$$p_k = p_{k+1/2}\,\exp\left(-\alpha_k\right), \tag{100}$$

where the integral constant has been set to zero so that $p_1 = \frac{1}{2} p_{3/2}$. Away from the upper boundary (100) yields

$$p_k = \frac{1}{2}\left(p_{k+1/2} + p_{k-1/2}\right) + \mathcal{O}\left[\left(\frac{\Delta p_k}{p_{k-1/2}}\right)^2\right]. \tag{101}$$

(101) can be derived as follows. Define $z = \Delta p_k / p_{k-1/2}$ for which α_k can be written as $\alpha_k = 1 - (1/z)\,\ln(1+z)$. Away from the upper boundary $z << 1$ and the logarithm in (99) can be expanded in terms of a Taylor series expansion about $z = 0$ so that α is given by

$$\alpha = 1 - \left(\frac{1}{z}\right)\,\ln(1+z), \tag{102}$$

$$= 1 - \left(\frac{1}{z}\right)\,\left[z - \frac{1}{2}\,z^2 + \mathcal{O}\left(z^3\right)\right], \tag{103}$$

$$= \frac{1}{2}\,z + \mathcal{O}\left(z^2\right). \tag{104}$$

A Taylor series expansion for $\exp(-\alpha_k)$ about $\alpha_k = 0$ is given by

$$\exp(-\alpha_k) = 1 - \alpha_k + \mathcal{O}\left(\alpha_k^2\right). \tag{105}$$

Substituting (104) into (105) completes the demonstration.

Linearized geopotential For the semi-implicit momentum equations the linearized geopotential is needed for the discretization of the PGF. Denote the linearized geopotential where constant terms have been discarded G_k; the constant terms are superfluous since the gradient of Φ is taken in the PGF.

Linearizing the geopotential

$$\Phi_k = \Phi_s + R \sum_{\ell=k+1}^{NLEV} T_\ell\ \left(\Delta \ln p\right)_\ell + \alpha_k\, R_d\ \left(T_v\right)_k, \tag{106}$$

about $p_s = p^{ref}$ and $T = T^{ref}$ yields

$$G_k = \phi_s + \left([\gamma]\,\mathbf{T}\right)_k + R_d\,\left(\frac{T}{p_s}\right)^{ref} p_s, \tag{107}$$

where $\mathbf{T}$ is a vector with the kth entry T_k. The matrix $[\gamma]$ is defined in Appendix F. The last term on the right-hand side of (107) is a linearizion of $R_{\rm d}\,T^{ref}\,\ln p_s$ using a Taylor series expansion, i.e. by using the equality

$$\ln p_s - \ln p_s^{ref} = \ln\left(1 + \frac{p_s - p_s^{ref}}{p_s^{ref}}\right). \tag{108}$$

and assuming that $\frac{p_s - p_s^{ref}}{p_s^{ref}}$ is sufficiently small, then the right-hand side of (108) can be approximated by a Taylor series expansion

$$\ln p_s = \ln p_s^{ref} + \frac{p_s}{p_s^{ref}} + \mathcal{O}\left[\left(\frac{p_s}{p_s^{ref}}\right)^2\right]. \tag{109}$$

4.1.4 *Trajectories*

To facilitate the computation of the upstream integral it is assumed that the cells move with rigid vertical walls (see Fig. 21) and a hybrid trajectory scheme is used. The trajectories are hybrid in the sense that the horizontal part is backward and the vertical part is forward i.e. in the horizontal cell vertices arrive at grid points and in the vertical the cells depart from model levels (Machenhauer and Olk 1998). The hybrid trajectory is graphically illustrated on Fig. 23. Since the vertical extension of the departure cell coincides with a model level and the walls are always vertical, the computation of the upstream integral is a two-dimensional problem. The methods used to approximate the upstream integral in NM02 and other 2D CISL schemes are therefore directly applicable. The arrival cell, however, does not necessarily coincide with a model level.

In a quasi-hydrostatic model the vertical velocities are diagnostic and computed from the continuity equation. In a traditional model, that is

$$\dot{\eta}_{k+1/2} = -B_{k+1/2}\, p_s \left(\frac{\partial \ln p_s}{\partial t}\right) - \sum_{\ell=1}^{k} \nabla \cdot \{\mathbf{v}_\ell\, \Delta p_\ell\}\,, \tag{110}$$

which is derived by partially summing the continuity equation on grid-point form and the surface pressure tendency equation

$$\frac{\partial \ln p_s}{\partial t} = -\frac{1}{p_s} \sum_{\ell=1}^{N} \nabla \cdot \{\mathbf{v}_\ell\, \Delta p_\ell\}\,. \tag{111}$$

also derived by vertical integration of the continuity equation on grid-point form. This is somewhat inconsistent with the semi-Lagrangian approach since Eulerian forms are used in the expression for the vertical velocity (110).

Here the vertical problem is solved using a vertical velocity in pressure coordinates (or equivalently a vertical grid) which is consistent with the explicit CISL semi-Lagrangian scheme used for the continuity equation. Hence the discretization of the explicit CISL continuity equation is discussed before deriving the trajectory algorithm.

Explicit continuity equation Assume that the hybrid trajectories are given. The explicit continuity equation is solved in spherical geometry on a μ grid using the hybrid trajectory scheme. The

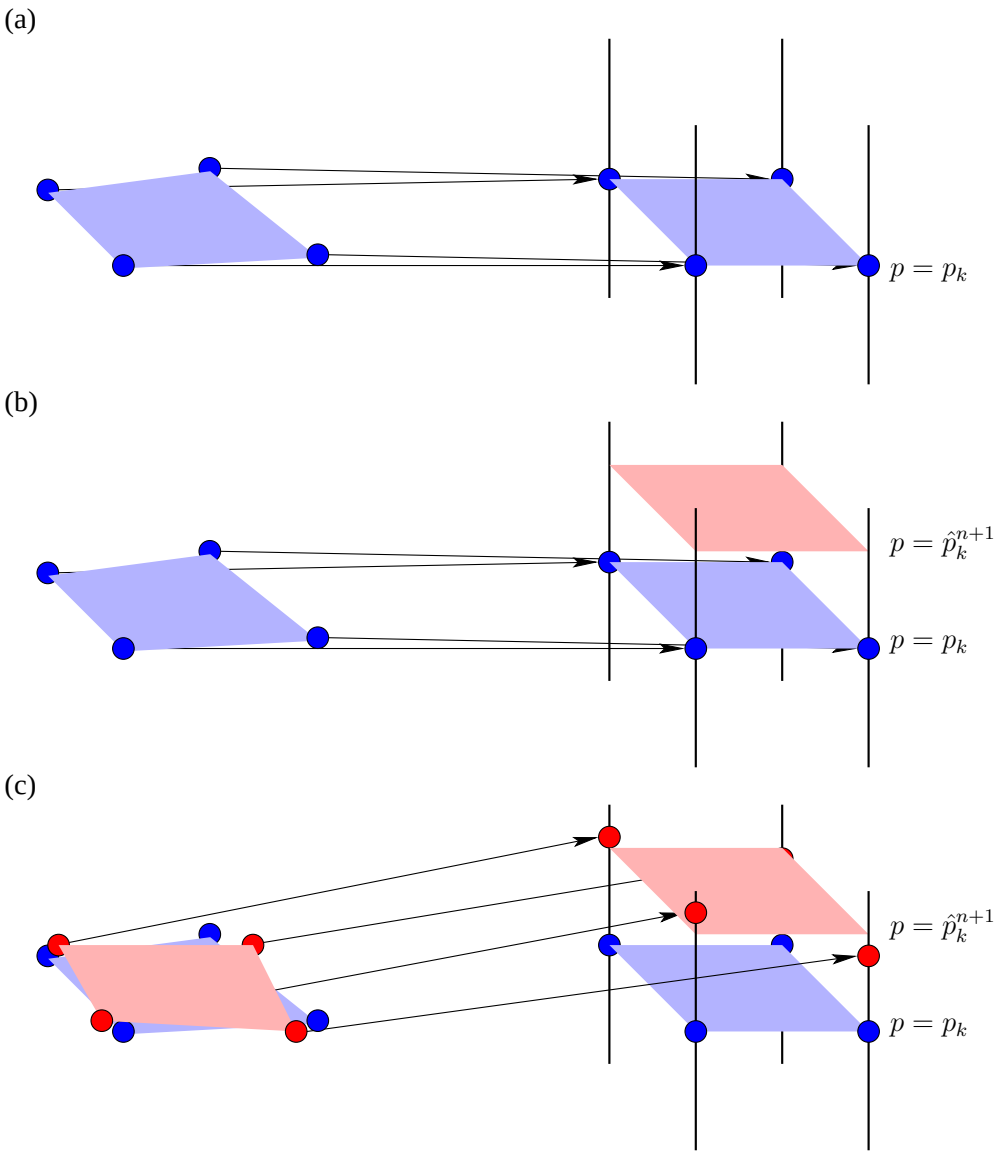


Fig. 23: A graphical illustration of the first iteration of the departure point algorithm consistent with the explicit CISL conitnuity equation. On the figures the regular/(irregular) arrival/(departure) cell is located to the right/(left) and the circles are the horizontal cell corner points. (a) Using the winds $\mathbf{v}_k^{n+1/2}$ construct backward trajectories of the cell corner points. All computations are performed in the model level, $p = p_k$. (b) Perform the upstream integral of mass to find the pressure levels implied by the continuity equation $p = \hat{p}_k^{n+1}$. (c) Interpolate the extrapolated part of the trajectories to $p = \hat{p}_k^{n+1}$ to find new (horizontal) trajectories. Note that the p coordinate of the arrival point, $\hat{p}_k^{n+1}$, does not necessarily coincide with a model level $p = p_k$. The new set of arrival points are not located exactly at the model level since the pressure is interpolated to cell vertices.

discretization of (90) can be written as

$$\overline{(\delta p_k)}^{n+1}_{exp}\,\Delta A = \overline{(\Delta p_k)}^{n}_{*}\,(\delta A_k)^n_* \tag{112}$$

where

$$\overline{(\Delta p_k)}^{n}_{*} = \frac{1}{(\delta A_k)^n_*}\iint_{(\delta A_k)^n_*} \Delta p^n_k(\lambda,\mu)\,dA \tag{113}$$

is the integral mean value of the pressure level thickness over the irregular departure cell area $(\delta A_k)^n_*$ and $\overline{(\delta p_k)}^{n+1}_{exp}$ is the mean value of the pressure level thickness over the regular arrival cell area ΔA. Here, and in the following, subscript $*$ and $*/2$ refers to the evaluation at the departure point or integral mean value over the departure cell and evaluation at the trajectory midpoint or the integral mean value over the cell at time level $(n+1/2)$, respectively.

In a hydrostatic model atmosphere the average pressure at the interface between cell k and $(k+1)$ is simply the weight of the air above it

$$\overline{\left(\hat{p}_{k+1/2}\right)}^{n+1} = \sum_{\ell=1}^{k}\overline{(\delta p_\ell)}^{n+1}, \tag{114}$$

and by using (100) the 'full level' pressure is given by

$$\overline{(\hat{p}_k)}^{n+1} = \exp(-\alpha_k^{n+1})\,\overline{\left(\hat{p}_{k+1/2}\right)}^{n+1}, \tag{115}$$

where α_k^{n+1} is computed by substituting (114) into the formula for α_k (99). Note that the pressure levels implied by the advection scheme ($\hat{p}_{k+1/2}$) do not necessarily coincide with the model levels ($p = p_{k+1/2}$) defined in terms of p_s^{n+1} and the hybrid coefficients (92).

Extension of new iterative algorithm to the hybrid approach To accommodate the semi-implicit system a new departure point algorithm was developed for the semi-implicit shallow water model presented in Chapter III. The trajectory was divided into two segments, $\mathcal{C}_1$ and $\mathcal{C}_2$; the displacement from the departure point to the trajectory midpoint and from the trajectory midpoint to the arrival point, respectively. $\mathcal{C}_1$ was based on winds only at time level (n) while $\mathcal{C}_2$ was based on winds extrapolated to time level $(n+1)$. The vectors $\mathcal{C}_1$ and $\mathcal{C}_2$ are vectors in the (λ,θ)-coordinate system. This setup is directly extended to the hybrid trajectory method.

The horizontal segments $\mathcal{C}_1$ and $\mathcal{C}_2$ are given by

$$\mathcal{C}_1 = \frac{\Delta t}{2}\,\mathbf{v}^n_* + \sum_{\ell=1}^{N-1}\frac{1}{(\ell+1)!}\left(\frac{\Delta t}{2}\right)^{\ell+1}\left(\frac{d^\ell \mathbf{v}}{dt^\ell}\right)^n_*, \tag{116}$$

$$\mathcal{C}_2 = \frac{\Delta t}{2}\,\tilde{\mathbf{v}}^{n+1} - \sum_{\ell=1}^{N-1}\frac{1}{(\ell+1)!}\left(-\frac{\Delta t}{2}\right)^{\ell+1}\frac{d^\ell}{dt^\ell}\left(\tilde{\mathbf{v}}^{n+1}\right), \tag{117}$$

where $\widetilde{(\cdot)}$ is the linear extrapolation operator

$$\widetilde{(\cdot)}^{n+1} = 2\,(\cdot)^n - (\cdot)^{n-1}. \tag{118}$$

The total derivative is approximated with

$$\frac{d\,\mathbf{v}}{dt} \approx \mathbf{v} \cdot \nabla \mathbf{v}, \tag{119}$$

where $\mathbf{v}$ is the horizontal velocity vector (u, v) and ∇ is the horizontal gradient operator expressed in spherical coordinates. Higher-order derivatives are defined recursively

$$\frac{d^\ell \mathbf{v}}{dt^\ell} = \frac{d}{dt}\left(\frac{d^{\ell-1}\mathbf{v}}{dt^{\ell-1}}\right) \quad \ell = 2, 3, .., N-1, \tag{120}$$

as in (McGregor 1993).

The trajectories are computed for cell vertices. The displacements for velocity and mass points are obtained by quasi-bicubic horizontal interpolation. The algorithm proceeds as follows.

First guess (initialization)

1. Interpolate the velocity components u and v horizontally to cell Eulerian corner points using one-dimensional cubic Lagrange interpolation.
2. Compute the components of $(\mathcal{C}_1)_k$ and $(\mathcal{C}_2)_k$ at cell vertices. No interpolation is needed for that. $(\mathcal{C}_1)_k$ and $(\mathcal{C}_2)_k$ are now known at 'full level' cell vertices of the model grid.
3. A first guess departure point is then given by

$$\mathbf{r}_k - \left[(\mathcal{C}_1)_k + (\mathcal{C}_2)_k\right], \tag{121}$$

   where $\mathbf{r}_k$ is the location of the south-west corner of the Eulerian cell with vertical index k.

Iterations

1. Solve the continuity equation, i.e. perform the upstream integral of mass. Hereby the vertical grid implied by the advection scheme $\overline{\hat{p}}_{k+1/2}$ is obtained. The bar refers to the fact that the vertical grid is given in terms of cell averages and the hat refers to the fact that the pressure levels do not necessarily coincide with model levels, $p = p_{k+1/2}$.
2. The second contribution to the trajectory C_2 (which is known at cell vertices and 'full levels' on the model grid) is interpolated in the vertical to the 'full levels' of the grid implied by the advection scheme:

- Since $(\mathcal{C}_2)_k$ is known at cell vertices and $\overline{\hat{p}}_{k+1/2}$ is a cell average, some kind of horizontal interpolation of $\overline{\hat{p}}_{k+1/2}$ to cell vertices is needed. Here quasi-bicubic interpolation is used. Denote $\overline{\hat{p}}_{k+1/2}$ interpolated to the south-west cell vertex $\left(\hat{p}_{k+1/2}\right)_{SW}$
- The vertical interpolation of $(\mathcal{C}_2)_k$ from $p = (p_k)_{SW}$ to

$$\left(\hat{p}_k^{n+1}\right)_{SW} = \left[\exp\left(-\alpha_k^{n+1}\right)\hat{p}_{k+1/2}^{n+1}\right]_{SW}, \tag{122}$$

  is done by using one-dimensional third-order Lagrange interpolation. Near the surface extrapolation is used and near the upper boundary the order of interpolation is reduced so that no extrapolation is needed. Denote the interpolated/extrapolated value $[(\mathcal{C}_2)_k]_{V*}$.

3. Interpolate $(\mathcal{C}_1)_k$ in the horizontal to

$$\mathbf{r}_k - \left\{(\mathcal{C}_1)_k + [(\mathcal{C}_2)_k]_{V*}\right\}. \tag{123}$$

   Denote the interpolated value $[(\mathcal{C}_1)_k]_{H*}$

4. At this iteration the departure point is located at

$$\mathbf{r}_k - \left\{[(\mathcal{C}_1)_k]_{H*} + [(\mathcal{C}_2)_k]_{V*}\right\}. \tag{124}$$

5. If needed repeat steps 1 to 4.

Note that this departure point algorithm does not require three-dimensional interpolation. Only one-dimensional and two-dimensional interpolation are used.

In a semi-implicit model the trajectories should ideally be part of the semi-implicit system and not be computed explicitly. When using explicit trajectories the divergence is not averaged along the trajectory as done in the semi-implicit continuity equation. However, it seems to be difficult if not impossible to do that. Therefore, as all other conventional semi-Lagrangian models, there is a slight inconsistency between the way the trajectory and other model equations are treated. The vertical velocity in the new model is, contrarily to conventional semi-Lagrangian models, consistent with the explicit CISL continuity equation.

4.1.5 *Discretization of the equations of motion*

In the following the equations of motion are solved on the hybrid trajectory. When the full semi-implicit system has been solved, the vertical remapping of all prognostic variables back to the model levels is performed.

The semi-implicit CISL continuity equation The discretization of the explicit continuity equation was discussed in connection with the trajectory algorithm. The derivation of the semi-implicit continuity equation is a direct extension of the derivation in Chapter III for the CISL shallow water model.

Using the continuity equation (112) and the discretized Lagrangian divergence

$$\mathbb{D}_k^{n+1/2} = \frac{1}{\Delta A}\frac{\Delta A - (\delta A_k)_*^n}{\Delta t}, \tag{125}$$

the explicit CISL continuity equation can be written as

$$\frac{1}{\Delta t}\left[\overline{(\delta p_k)}_{exp}^{n+1} - \overline{(\delta p_k)}_*^n\right] = -\mathbb{D}_k^{n+1/2}\,\overline{(\delta p_k)}_*^n. \tag{126}$$

Linearizing the right-hand side about p_s^{ref}, treating the linear term as a temporal average, the "ideal" semi-implicit continuity equation results

$$\overline{\delta p}_k^{n+1} = \overline{(\delta p_k)}_{exp}^{n+1} - \frac{\Delta t}{2}\left[\mathbb{D}_k^{n+1} - \mathbb{D}\left(\widetilde{\mathbf{v}}_k^{n+1}\right)\right]\,\Delta p_k^{ref}, \tag{127}$$

(see Chapter III for details). The last term on the right-hand side of (127) is given by

$$\mathbb{D}(\tilde{\mathbf{v}}^{n+1}) = \frac{1}{\Delta A}\frac{\Delta A - \delta A(\tilde{\mathbf{v}}^{n+1})}{\Delta t}, \tag{128}$$

where the departure area $\delta A(\tilde{\mathbf{v}}^{n+1})$ is computed using only extrapolated winds. As discussed in Chapter III, this semi-implicit equation leads to a complicated elliptic equation. Hence the "predictor-corrector" approach of Kaas et al. (2005) is used:

$$\overline{\delta p}_k^{n+1} = \overline{(\delta p_k)}_{exp}^{n+1} - \frac{\Delta t}{2}\,\Delta p_k^{ref}\,\left[D_k^{n+1} - \mathbb{D}\left(\widetilde{\mathbf{v}}_k^{n+1}\right)\right] + \frac{\Delta t}{2}\,\Delta p_k^{ref}\,\overline{\left[D_k^n - \mathbb{D}\left(\mathbf{v}_k^n\right)\right]}\,\left(\frac{\delta A_*^n}{\Delta A}\right)_k. \tag{129}$$

where D is the Eulerian divergence based on a finite difference discretization of

$$D = \frac{1}{a\,\cos\theta}\left[\frac{\partial}{\partial\lambda}\left(u\,\cos\theta\right) + \frac{\partial v}{\partial\theta}\right]. \tag{130}$$

The last term on the right-hand side of (129), the so-called *correction term*, is integrated over the departure area.

The prediction of the surface pressure can be written as

$$\overline{(p_s)}^{n+1} + \frac{\Delta t}{2}\,\left([\nu]\,\mathbf{D}\right)^{n+1} = \left(\mathbf{R_p}\right)_k, \tag{131}$$

where

$$\left(\mathbf{R_p}\right)_k = \overline{(p_s)}_{exp}^{n+1} - \frac{\Delta t}{2}\,\left([\nu]\,\mathbf{A}_p\right), \tag{132}$$

and $[\nu]$ is a matrix defined in Appendix F and the known terms are given by

$$(\mathbf{A}_p)_\ell = \mathbb{D}\left(\widetilde{\mathbf{v}}_\ell^{n+1}\right) + \overline{\left[D_\ell^n - \mathbb{D}\left(\mathbf{v}_\ell^n\right)\right]_*}^n \left(\frac{\delta A_*^n}{\Delta A}\right)_\ell , \tag{133}$$

$$\overline{(p_s)}_{exp}^{n+1} = \sum_{\ell=1}^{NLEV} \overline{(\delta p_\ell)}_{exp}^{n+1} . \tag{134}$$

Grid-point SISL thermodynamic and momentum equations The thermodynamic and momentum equations are discretized as in HIRLAM but must be adapted to the hybrid trajectory. Using the notation introduced in Chapter III the SISL discretizations for u, v and T can be written as

$$\frac{\hat{\psi}_k^{n+1} - (\psi_k^n)_{H*}}{\Delta t} = \frac{1}{2}\left\{\left[(N_\psi)_k\right]_{V*}^{n+1/2} + \left[(N_\psi)_k\right]_{H*}^{n+1/2}\right\} + \frac{1}{2}\left\{(L_\psi)_k^{n+1} + \left[(L_\psi)_k\right]_{H*}^{n}\right\}, \quad \psi = u, v, T, \tag{135}$$

respectively. The subscript $H*$ refers to the interpolation to the departure point, which only requires horizontal interpolation. The subscript $V*$ refers to the vertical interpolation from model levels to the grid implied by the explicit CISL continuity scheme. The predicted value $\hat{\psi}_k^{n+1}$ is not located at a model level, but at the vertical grid implied by the semi-implicit CISL continuity equation. A slight inconsistency between $(\psi)_{V*}$ and $(\hat{\psi})$ is present since the vertical grid implied by the explicit and semi-implicit CISL continuity equation are not exactly identical. Note, however, that the vertical part of the trajectory algorithm in a traditional model is not consistent with either the explicit or semi-implicit continuity equation and thus more inconsistent with respect to the vertical part than the model presented here.

The discretized adiabatic SISL thermodynamic equation is given by

$$(T_k)^{n+1} + \frac{\Delta t}{2}\left(\left[\tau^A\right]\mathbf{D}\right)_k^{n+1} = \left(\mathbf{R}_T^A\right)_k , \tag{136}$$

where

$$\left(\mathbf{R}_T^A\right)_k = \left(\frac{\Delta t}{2}\right)\left(\hat{N}_T\right)_k^{n+1/2} + \left[T_k^n - \frac{\Delta t}{2}\left(\left[\tau^A\right]\mathbf{D}\right)_k^n + \frac{\Delta t}{2}\left(N_T\right)^{n+1/2}\right]_{H*} . \tag{137}$$

The non-linear term $(N_T)_k$ is given by

$$(N_T)_k = \left(\kappa\, T\, \frac{\omega}{p}\right)_k - \left(\left[\tau^A\right]\mathbf{D}\right)_k , \tag{138}$$

where

$$\left[\kappa\, T\, \frac{\omega}{p}\right]_k = \frac{\kappa\, T_k}{\Delta p_k}\left\{\Delta \ln p_k \left[p_s \frac{\partial \ln p_s}{\partial t} + \sum_{\ell=k+1}^{NLEV} \nabla\cdot(\mathbf{v}\,\Delta p)_\ell\right] + \left(\Delta \ln p_k - \alpha_k\right)\nabla\cdot(\mathbf{v}\,\Delta p)_\ell\right\}, \tag{139}$$

(see Undén (2002) for discretization details). Note that the energy conversion term is discretized in an Eulerian fashion.

The momentum equations are discretized on component form and using a 'half-implicit'-Coriolis scheme as in HIRLAM. The adiabatic SISL momentum equations are given by

$$\left[\hat{u} + \frac{\Delta t}{2}\left(\frac{1}{a\,\cos\theta}\,\frac{\partial G}{\partial \lambda} - f_0\,\hat{v}\right)\right]_k^{n+1} = (\mathbf{R}_u)_k \tag{140}$$

$$\left[\hat{v} + \frac{\Delta t}{2}\left(\frac{1}{a}\,\frac{\partial G}{\partial \theta} + f_0\,\hat{u}\right)\right]_k^{n+1} = (\mathbf{R}_v)_k \tag{141}$$

$\mathbf{R}_u$ and $\mathbf{R}_v$ are the explicit parts of the equation for u and v, respectively,

$$(\mathbf{R}_u)_k = \left[u_k^n - \frac{\Delta t}{2}\left(\frac{1}{a\,\cos\theta}\,\frac{\partial G_k^n}{\partial \lambda} - f\,v_k^n - (N_u)_k^{n+1/2}\right)\right]_{H*} + \frac{\Delta t}{2}\left[(N_u)_k^{n+1/2}\right]_{V*}, \tag{142}$$

$$(\mathbf{R}_v)_k = \left[v_k^n - \frac{\Delta t}{2}\left(\frac{1}{a}\,\frac{\partial G_k^n}{\partial \theta} + f\,u_k^n - (N_v)_k^{n+1/2}\right)\right]_{H*} + \frac{\Delta t}{2}\left[(N_v)_k^{n+1/2}\right]_{V*}, \tag{143}$$

where the non-linear terms are given by

$$N_u = -\frac{1}{a\,\cos\theta}\,\frac{\partial}{\partial \lambda}\,(\Phi - G) - \frac{R_d\,T_v}{a\,\cos\theta}\,\frac{\partial}{\partial \lambda}\,(\ln p) + \frac{u\,v}{a}\,\tan\theta + f'\,v, \tag{144}$$

$$N_v = -\frac{1}{a}\,\frac{\partial}{\partial \theta}\,(\Phi - G) - \frac{R_d\,T_v}{a}\,\frac{\partial}{\partial \theta}\,(\ln p) + \frac{u^2}{a}\,\tan\theta - f'\,u. \tag{145}$$

The linearized geopotential is given by (107). For the interpolation of the non-linear terms from model levels to the levels implied by the explicit CISL continuity equation, the pressure levels at velocity points are needed. The pressure levels, however, are known as cell-averages. 2D quasi-bicubic interpolation is used to get the pressures at velocity points.

The elliptic system The derivation of the elliptic system closely follows HIRLAM (see p.6-10 in Undén 2002). First an expression for the linearized geopotential G which combines the continuity and thermodynamic equations is derived. Then the resulting expression is substituted into the momentum equations, which subsequently are manipulated to form an Helmholtz equation.

By using the SISL continuity and thermodynamic equations the linearized geopotential at time level $(n+1)$ can be written as

$$\mathbf{G}^{n+1} + \frac{\Delta t}{2}\,\left[M^A\right]\mathbf{D}^{n+1} = \mathbf{H}^A \tag{146}$$

$\left[M^A\right]$ is the vertical structure matrix and is defined in Appendix F. The kth element in the vector $\mathbf{H}^A$ is given by

$$\left(\mathbf{H}^A\right)_k = \Phi_s + \left([\gamma]\,\mathbf{R}_T^{n+1}\right)_k + R_d\left(\frac{T}{p_s}\right)^{ref}(p_s)_{exp}^{n+1} - \frac{\Delta t}{2}\left(\left[M^B\right]\,\mathbf{A}_p\right)_k. \tag{147}$$

Substitute v from (141) in (140), and u from (140) in (141), and replace G using (146). The results are

$$\left[\hat{u} + b\left(\frac{\Delta t}{2}\right)^2 \left(\frac{1}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{F_+}{a}\frac{\partial}{\partial\theta}\right)[M]\,\mathbf{D}\right]_k^{n+1} = (\mathbf{Z}_u)_k, \tag{148}$$

$$\left[\hat{v} + b\left(\frac{\Delta t}{2}\right)^2 \left(-\frac{F_+}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{1}{a}\frac{\partial}{\partial\theta}\right)[M]\,\mathbf{D}\right]_k^{n+1} = (\mathbf{Z}_v)_k, \tag{149}$$

where

$$\mathbf{Z}_u = b\left\{\left[\mathbf{R}_u - \left(\frac{\Delta t}{2}\right)\frac{1}{a\,\cos\theta}\frac{\partial\mathbf{H}}{\partial\lambda}\right] + F_+\left[\mathbf{R}_v - \left(\frac{\Delta t}{2}\right)\frac{1}{a}\frac{\partial\mathbf{H}}{\partial\theta}\right]\right\}, \tag{150}$$

$$\mathbf{Z}_v = b\left\{\left[\mathbf{R}_v - \left(\frac{\Delta t}{2}\right)\frac{1}{a}\frac{\partial\mathbf{H}}{\partial\theta}\right] - F_+\left[\mathbf{R}_u - \left(\frac{\Delta t}{2}\right)\frac{1}{a\,\cos\theta}\frac{\partial\mathbf{H}}{\partial\lambda}\right]\right\}, \tag{151}$$

where $b = 1/(1 + F_+^2)$ and $F_+ = f_0\,\Delta t/2$. A vertically coupled system of Helmholtz equations results when applying the horizontal gradient operator to (148) and (149)

$$\left[1 - b\left(\frac{\Delta t}{2}\right)^2 [M]\,\nabla^2\right](\mathbf{D})_k^{n+1} = \nabla\cdot(\mathbf{Z})_k \qquad k = 1, .., NLEV, \tag{152}$$

where $(\mathbf{Z})_k$ is a two-dimensional vector with components $[(\mathbf{Z}_u)_k, (\mathbf{Z}_v)_k]$ and ∇^2 is the Laplacian

$$\nabla^2 = \frac{1}{a\,\cos\theta}\left[\frac{\partial}{\partial\lambda}\frac{1}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{\partial}{\partial\theta}\frac{\cos\theta}{a}\frac{\partial}{\partial\theta}\right]. \tag{153}$$

The system (152) is decoupled by diagonalizing the matrix $[M]$. Let $[E]$ be the matrix that diagonalizes matrix $[M]$, i.e.

$$[E]^{-1}[M][E] = \lambda, \tag{154}$$

then (152) can be written as a vertically decoupled set of Helmholtz equations

$$\left[1 - b\left(\frac{\Delta t}{2}\right)^2 \lambda_k\,\nabla^2\right]\left([E]^{-1}\,\mathbf{D}\right)_k^{n+1} = \nabla\cdot\left([E]^{-1}\,\mathbf{Z}\right)_k \qquad k = 1, .., NLEV. \tag{155}$$

Each Helmholtz equation is solved using the solver of HIRLAM which applies a Fourier transform in the zonal direction (Källén 1996, p. 2.14) and assumes that the divergence is zero on the boundaries.

Vertical remapping The vertical grid implied by the semi-implicit CISL continuity equation will, in general, differ from the models vertical grid. Therefore all prognostic variables must be remapped or interpolated back to the model levels at the end of each time step. It is important that this process is accurate since the vertical gradients can be large.

u, v and T are interpolated from the 'full level' pressure implied by the semi-implicit CISL continuity equation to the 'full levels' of the model grid. As in the trajectory algorithm, the interpolation is performed in pressure coordinates and the 'full level' pressure levels are computed from the 'half level' pressures using (100). Cubic Lagrange interpolation is used (as in the trajectory algorithm).

4.2 Consistent "omega-p" CISL HIRLAM

In the derivation of the R-CISL-HIRLAM mass conservation was enforced with minimum modifications to the existing HIRLAM dynamical core. However, the discretization of the thermodynamic equation is not completely satisfactory. In particular, the energy conversion term is not consistent with the discretized continuity equation (this is also the case for HIRLAM).

4.2.1 *Consistent ω/p discretizaton*

For the approximation of the vertical velocity in pressure coordinates the definition $\omega = dp/dt$ is used. Using the CISL approach a mean value of ω over $(\delta A_k)^{n+1/2}_{*/2}$ can be approximated by

$$\left[\overline{(\omega_k)}^{n+1/2}_{*/2}\right]_{exp} = \frac{1}{\Delta t}\left[\overline{(\hat{p}_k)}^{n+1}_{exp} - \overline{(p_k)}^{n}_{*}\right], \tag{156}$$

where $\overline{(\hat{p}_k)}^{n+1}_{exp}$ and $\overline{(p_k)}^{n}_{*}$ is the pressure averaged over the arrival and the departure cell, respectively. $\overline{(\hat{p}_k)}^{n+1}_{exp}$ is simply the sum of the weight of air above the arrival cell in level k, and is directly available by (115) after having solved the explicit CISL continuity equation. $\overline{(p_k)}^{n}_{*}$ is obtained by integrating $\overline{(p_k)}^{n}$, given by

$$p^n_k = \exp\left(-\alpha^n_k\right) \sum_{\ell=1}^{k} \Delta p^n_\ell, \tag{157}$$

over $(\delta A_k)^n_*$. That is

$$\overline{(p_k)}^{n}_{*} = \frac{1}{(\delta A_k)^n_*} \iint_{(\delta A_k)^n_*} p^n_k \, dA. \tag{158}$$

The explicit approximation to ω/p is given by

$$\frac{\left[\overline{(\omega_k)}^{n+1/2}_{*/2}\right]_{exp}}{\frac{1}{2}\left[\overline{(\hat{p}_k)}^{n+1}_{exp} + \overline{(p_k)}^{n}_{*}\right]}. \tag{159}$$

where $\left[\overline{(\omega_k)}^{n+1/2}_{*/2}\right]_{exp}$ is given by (156). Since all terms are computed using pressures predicted by the explicit continuity equation, the discretization of the energy conversion term is consistent

with the explicit CISL continuity equation. Note that, contrarily to conventional semi-Lagrangian models, no explicit extrapolation has been used. Extrapolation has only been used implicitly in terms of the extrapolation in the trajectory algorithm. Most importantly, the energy conversion term is discretized in a purely Lagrangian fashion.

For the semi-implicit discretization of the thermodynamic equation isolate $\overline{(\delta p)}^{n+1}_{exp}$ in (126), and using this to substitute for $\overline{(\hat{p})}^{n+1}_{exp}$ in (156). Then (156) can be written as

$$\overline{(\omega_k)}^{n+1/2}_{*/2} = \frac{1}{\Delta t}\left\{\exp\left(-\alpha_k^{n+1}\right)\sum_{\ell=1}^{k}\left[-\Delta t\,\mathbb{D}_\ell^{n+1/2}\,\overline{(\delta p_\ell)}^{n}_{*} + \overline{(\delta p_\ell)}^{n}_{*}\right] - \overline{(p_k)}^{n}_{*}\right\}. \tag{160}$$

Using (160) divided by $p_k = \exp(-\alpha_k)\,p_{k+1/2}$ for the energy conversion term, the linearization of the right-hand side of the thermodynamic equation about T^{ref} and p_s^{ref} yields

$$-\frac{R_d}{c_{pd}}\left(\frac{T}{p_{k+1/2}}\right)^{ref}\sum_{\ell=1}^{k}\Delta p_\ell^{ref}\,\mathbb{D}_\ell^{n+1/2}. \tag{161}$$

Modifying R-CISL-HIRLAM such that the ω/p term in the thermodynamic equation is consistent with the semi-implicit CISL continuity equation results in a system, COP-CISL-HIRLAM, identical to R-CISL-HIRLAM, but with the matrix $\left[\tau^A\right]$ replaced by $\left[\tau^B\right]$. Of course, instead of using (139) for the explicit energy conversion term (156) must be used.

As a result of a different linearization of the right-hand side of the thermodynamic equation the vertical structure matrix is different. The eigenvalues of the vertical structure matrix for the two new dynamical core versions are shown on Fig. 24

4.3 Preliminary tests

4.3.1 *Idealized dynamical core test cases*

Atmospheric models can generally be divided into two components or modules: firstly, the numerical solution to the dry, adiabatic, primitive equations also referred to as the *dynamical core*; secondly, the physics package dealing with processes not resolved on the numerical grid and which must therefore be represented using so-called parameterizations. The interaction between the dynamical core and the physics package is strongly nonlinear and it is therefore difficult to diagnose cause and effect associated with numerical approximations in a full model. It is therefore desirable to study model components on their own.

For the shallow water equations a standard test suite exists and is widely used by the modeling community. For the dynamical core of baroclinic models such a standard test suite does not exist.

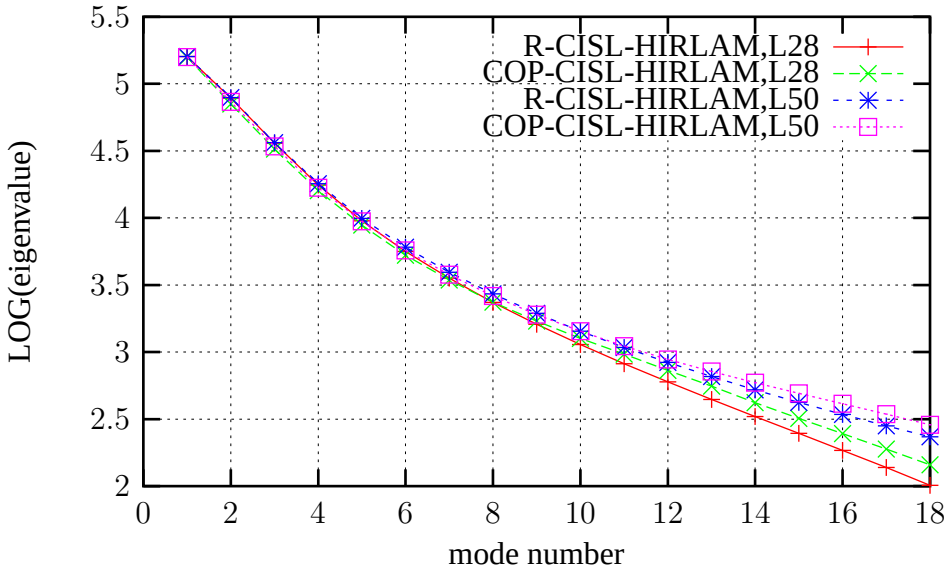


Fig. 24: The first 18 eigenvalues of the vertical structure matrix for model versions R-CISL-HIRLAM and COP-CISL-HIRLAM at two vertical resolutions as a function of eigenvector number. The vertical levels are defined as in JW2005 but with an extra level at the top of the atmosphere so that the upper most level pressure is zero (the levels for the 28-level version are defined in Appendix G). The reference temperature is 400 K and the reference pressure is 1000 hPa.

Several baroclinic test case have, however, been developed. The most widely spread 3D test is the one of Held and Suarez (1994) for general circulation models. The model is initialized with artificial data, for example isothermal atmosphere (e.g. 300K), zero winds and constant surface pressure (e.g. 1000 hPa) and no orography. Often the initial conditions are perturbed with a small random perturbation to break the symmetry. The physics package is replaced with idealized forcings given in terms of a Newtonian relaxation of the temperature and a Rayleigh damping of the wind at lower levels. These forcings represent sub-grid-scale physical processes and boundary layer friction, respectively. The test requires the model to be run for 1200 days until an equilibrium is reached and the subsequent analysis of the mean zonally averaged temperatures and mean zonally averaged zonal velocities. Since temporally and spatially averaged quantities are the proxies for model performance, the test aims at the long-term statistics of the computed circulation and not at small scale flow properties[8]. A similar test case was designed by Boer and Denis (1997), which also aims at long-term statistical properties of the model. These test cases are, however, incomplete in the sense that small scale features and forecast skill can not be measured. In addition, the dynamical core is not tested in isolation but with a simple forcing package which, as discussed above, makes it harder to diagnose cause and effect of model behavior.

Recently Polvani et al. (2004) and JW2005 have developed test cases aiming at forecasting skill of

[8]hence the model may also be initialized with "real" data

the dynamical core. The initial conditions are quasi realistic and the integration is only in the order of a few weeks. Both test cases are similar, but the Polvani et al. (2004) test case is with a second order diffusion term, which to some extent controls the solution. The JW2005 test case (also referred to as the Jablonowski-Willamson test case), on the other hand, does not include explicit diffusion terms and the initial condition is specified entirely by analytic expressions. Here the Jablonowski-Williamson test case has been used for the preliminary testing of the new dynamical cores.

4.3.2 *Jablonowski-Williamson test case*

The Jablonowski-Williamson tests consist of two test cases. In the first case the initial condition is given by analytic expressions which are a steady-state solution to the hydrostatic primitive equations. The ideal model should therefore not alter the initial conditions at any time during the integration. In the second test a perturbation is overlaid the steady state initial condition. Hereby a baroclinic wave is triggered (see Fig. 25). The wave train travels eastward and starts growing noticeably around day 4 and evolves rapidly thereafter with strong cyclogenesis at day 8. The initial conditions for the two test cases are given in Appendix H. There is no analytic solution to this test case and reference solutions from four different models are provided by the authors. A short summary of the characteristics of the simulations with the four dynamical cores is given next.

Characteristics of the global model runs JW2005 provided reference solutions to the baroclinic wave test case using four different dynamical cores based on very different numerical approaches, thereby providing independent estimates of the unknown true solution. The dynamical cores were the NASA/NCAR Finite Volume dynamics package (NASA/NCAR-FV), the NCAR spectral transform Eulerian and semi-Lagrangian dynamical cores of the Community Atmosphere Model CAM3 (EUL-CAM3 and SL-CAM3, respectively) and the icosahedral finite-difference model GME of the German Weather Service. NASA/NCAR-FV model and SL-CAM3 were run without explicit diffusion whereas fourth-order horizontal diffusion was applied in GME and EUL-CAM3. Note that the spectral models do, in addition, have a "cut off" of the short waves during the transform back and forth to spectral space. A mass fixer was used in SL-CAM3 and EUL-CAM3, and in SL-CAM3 also an energy fixer was applied. Neither GME or NASA/NCAR-FV uses mass or energy fixers.

Note that although SL-CAM3 and NASA/NCAR-FV were run without explicit diffusion, the models have numerical diffusion in terms of their interpolators. In the semi-Lagrangian model the upstream interpolation is a source of numerical diffusion. The finite-volume model was run with a monotonic constraint on the sub-grid-scale reconstructions and thereby providing numerical diffusion.

The models were run with 18, 26 and 49 vertical levels, respectively, and at horizontal resolutions ranging from about 5° to 0.25°. By comparing estimates of the reference solution produced by the

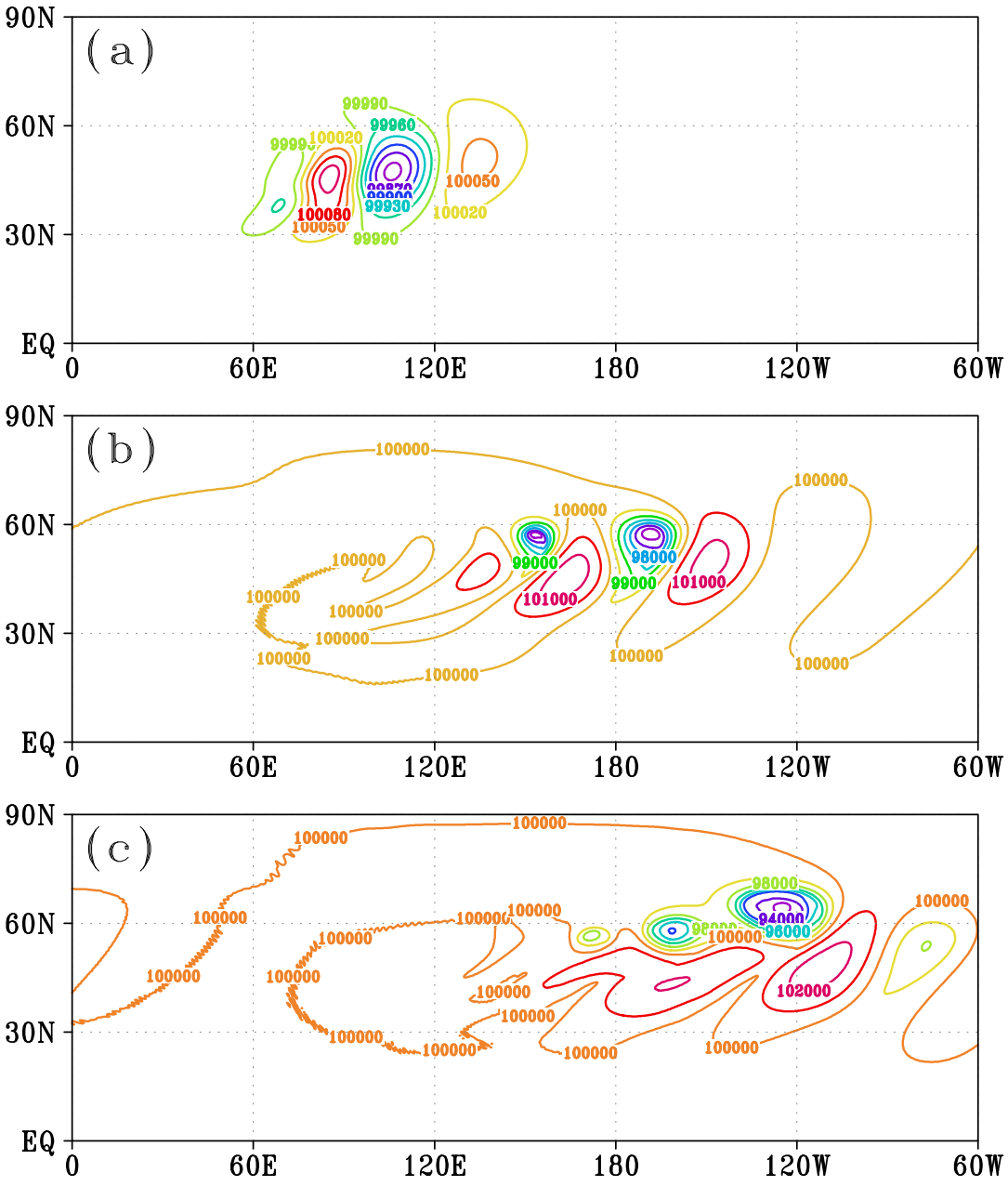


Fig. 25: The surface pressure (Pa) for the Jablonowski-Williamson baroclinic wave test case at day (a) 4, (b) 8, and (c) 10. The high-resolution reference solution has been computed with the semi-Lagrangian CAM3 model (T340 horizontal resolution and 49 levels in the vertical). Data courtesy of Dr. D. Williamson and Dr. C. Jablonowsky, NCAR.

highest resolution version of the same model it was found that the models converged for horizontal

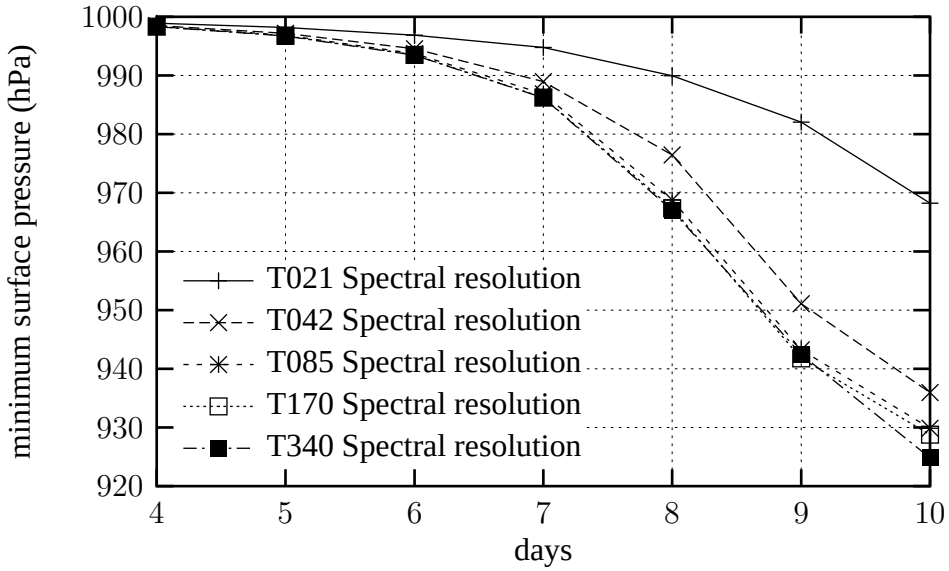


Fig. 26: The minimum surface pressure as a function of time for the Jablonowski-Williamson baroclinic wave. Solutions are computed with the NCAR spectral transform semi-Lagrangian dynamical core (SL-CAM3). Note that the higher the resolution the deeper the low pressure system. Data courtesy of Dr. C. Jablonowski, NCAR.

resolutions higher than approx. 1.5°. The difference between the next highest horizontal resolution run (about 0.5°) and the reference solution were well below the uncertainty limit.

The phase lag of the baroclinic wave was defined by the distance the surface pressure field must be shifted with respect to the high-resolution reference solution to minimize l_2 difference. Apart from the finite volume model, the wave propagated too far east at lower resolutions compared to the reference solution, and the phase lag decreased for increasing resolution (see Fig.10 of JW2005).

An indication of the amplitude errors can be obtained by considering the maximum and minimum surface pressures. Note, however, that the minimum and maximum surface pressure may not be representative for the amplitude error in the entire domain, but only for the area of strongest baroclinic activity. In general the minimum/maximum p_s decreases/increases for increasing resolution (see Fig. 26). So in that respect the more accurate the individual model the lower/higher the minimum/maximum p_s. An inter-model comparison of the minimum p_s as a function of time shows that all the high-resolution models lie within a range of 2 hPa (see Fig. 27). The NASA/NCAR FV has the highest minimum p_s. The flux-form advection scheme of Lin and Rood (1996), on which this model is based, uses a monotonic limiter which introduce extra numerical diffusion which might explain the higher minimum p_s.

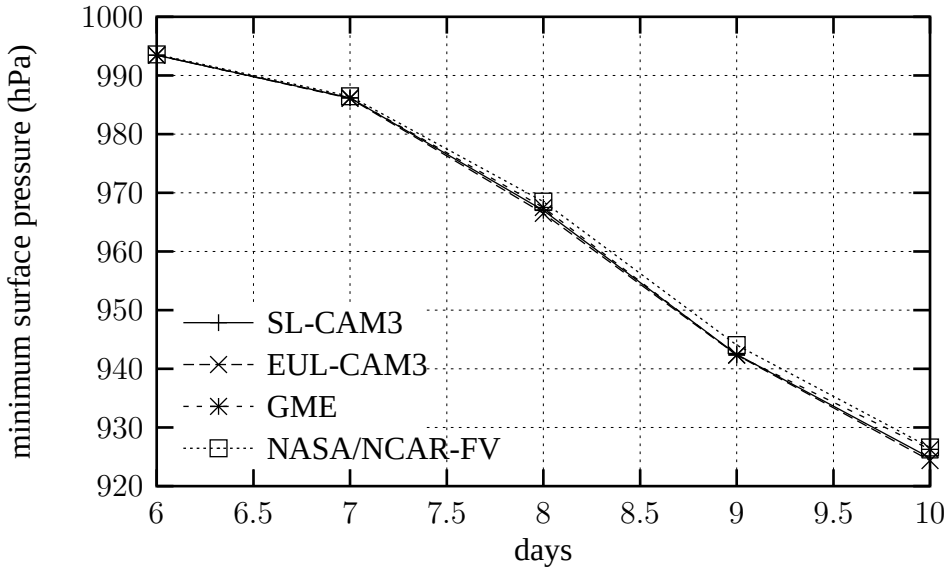


Fig. 27: The minimum surface pressure as a function of time for the Jablonowski-Williamson baroclinic wave in the four reference solutions provided by JW2005. At day 8 the minimum p_s are 967.0 hPa, 966.5 hPa, 967.5 hPa, and 968.5 hPa for the SL-CAM3, EUL-CAM3, GME, and NASA/NCAR-FV, respectively. Note that the NASA/NCAR-FV min(p_s) is slightly larger compared to the other models. Data courtesy of Dr. C. Jablonowski, NCAR.

4.3.3 *Experimental setup*

The implementation of the new dynamical cores has been performed within the framework of HIRLAM. Therefore the integration area can only cover a limited area on the sphere. It was found technically too extensive to set up a system similar to the shallow water model setup where the boundary values were provided by a high-resolution reference solution at every time step. The next best setup would be a global channel where the solution is relaxed toward the initial condition at the north and south boundaries while the domain would be periodic in the east-west direction. This is, however, not possible in HIRLAM since the elliptic solver for the SISL system assumes zero divergence on the boundaries; this would provide a spurious source of divergence in the overlap region where the domain is made periodic. Rather than programming a new elliptic solver these preliminary tests are run on a limited area domain large enough to contain the area of main activity for at least 8 days. At the boundaries the solution is held fixed at the initial condition and the divergence is assumed zero on the lateral boundaries (which is consistent with the initial condition which has zero divergence away from the perturbation).

The active domain is not rotated and extends from $-80°$S to $80°$N and in the longitudinal direction from 100°/(260°) west/(east) of the center of the perturbation located at $(20°E, 40°N)$. In an 8 grid-point wide zone around the active domain the fields are held fixed at the initial condition.

Inside the active domain the fields are relaxed toward the initial condition in an approximately 10° wide zone adjacent to the active domain boundary. A rotated grid was not chosen since the baroclinic wave would "hit" a lateral boundary after only a few days of integration.

Two horizontal resolutions are used. The standard resolution is $\Delta\lambda = 1.45^\circ$, $\Delta\theta = 1.15^\circ$, and the highest resolution is $\Delta\lambda = 0.74^\circ$, $\Delta\theta = 0.59^\circ$. In the vertical there are 27 levels and the placement of the levels is as in the JW2005, but with one more layer added at the top of the atmosphere such that the pressure at the upper boundary is zero as in HIRLAM. The time step for the standard and high-resolution runs is 30 and 15 minutes, respectively.

To keep solutions computed with the traditional HIRLAM (simply referred to as HIRLAM) stable and noise free it was necessary to use decentering, filtering of the non-linear terms in time and horizontal diffusion. By running only with decentering and not with horizontal diffusion or vice versa the solution was not noise free (not shown). The trajectory scheme used in HIRLAM is the one developed by McDonald (1999). Two iterations are used, and for interpolations in the trajectory scheme and for the interpolation of fields to the departure points, fully three-dimensional interpolation is used. For discretization details see Undén (2002).

The new dynamical core does not use decentering or filtering of the non-linear terms in time, but it was necessary to apply horizontal diffusion in order to avoid noise problems. Unless else stated explicitly all models were run with ∇^6 implicit horizontal diffusion on T, u and v (see p.12-13 in Undén 2002). The horizontal diffusion coefficients are listed in Appendix I. A monotone limiter is applied to the continuity equation of the dynamical core so that, if additional tracers are advected *online*, they are advected monotonically and consistently.

For the semi-implicit scheme a reference pressure of 1000 hPa is used. In HIRLAM the reference temperature is set to 300 K and in the new dynamical cores $T^{ref} = 400$ K. In the shallow water model the constant geopotential used for the semi-implicit scheme was chosen sufficiently large to avoid instabilities (see discussion in Chapter III). In a baroclinic model that corresponds to an increase in the reference temperature so that the equivalent depths are increased. As a starting point the reference temperature was chosen larger than what is normally used.

4.3.4 *High-resolution limited area simulations*

Figure 28 show the minimum and maximum surface pressure for different model configurations, and the surface pressure at days 6,7 and 8 for the high-resolution simulation of the baroclinic wave train with HIRLAM and COP-CISL-HIRLAM are shown on Fig. 29. With respect to the minimum/maximum surface pressure COP-CISL-HIRLAM has a stronger development compared to HIRLAM. It is during these days that the wave starts growing noticeably into strong cyclogenesis around day 8. At day 8 the minimum/maximum surface pressures are 963.06 hPa/1014.06 hPa and 964.56 hPa/1013.44 hPa for COP-CISL-HIRLAM and HIRLAM, respectively. The two lowest surface pressures are located further toward north-east in COP-CISL-HIRLAM compared to

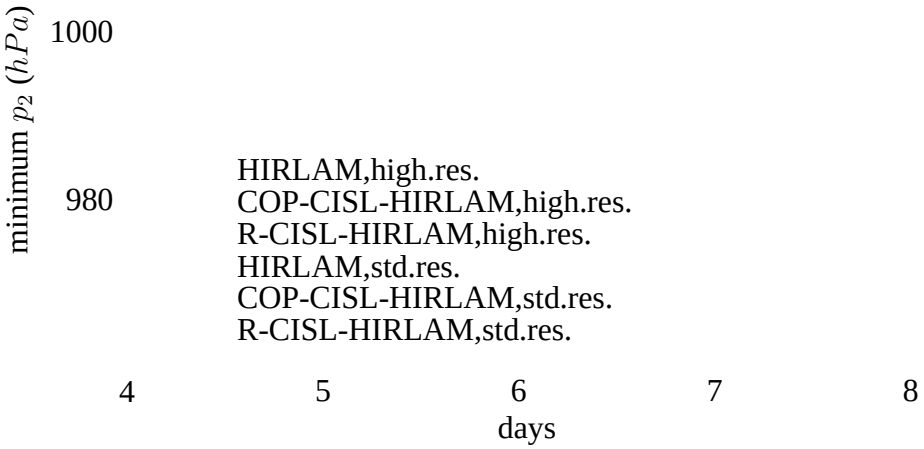


Fig. 28: The minimum and maximum surface pressure as a function of time for different model versions and resolutions. “std.res” and “high.res” refer to the standard- and high-resolution model runs, respectively.

HIRLAM. Behind the baroclinic wave train a second wave train is visible in the western part of the domain. This is a wave triggered at the boundaries and is discussed later.

On Fig. 30 the surface pressure of the high-resolution runs with COP-CISL-HIRLAM and R-CISL-HIRLAM are plotted. In terms of the extrema in the surface pressure the development is stronger with COP-CISL-HIRLAM compared to R-CISL-HIRLAM. At day 8 the minimum/maximum surface pressure for R-CISL-HIRLAM is 964.71 hPa/1013.72 hPa. At day 6 and 7 the phase of the wave in the cell-integrated models is very similar, but at day 8 the minima in p_s have propagated further south-east in R-CISL-HIRLAM compared to COP-CISL-HIRLAM. As expected, the effect of using a energy conversion term consistent with the discretized semi-implicit CISL continuity equation is a stronger baroclinic development.

4.3.5 *Standard-resolution LAM simulations*

Here the standard-resolution simulations ($\Delta\lambda = 1.45°$, $\Delta\theta = 1.15°$) are compared with the high-resolution runs ($\Delta\lambda = 0.74°$, $\Delta\theta = 0.59°$) with the same model. In the standard-resolution run with HIRLAM the minima in p_s are located further toward south-east compared to the high-resolution runs while the trailing highs propagate a little slower in the standard-resolution simulation (Fig. 31). Similarly, the wave propagates faster in the low-resolution run compared to the high-resolution run with COP-CISL-HIRLAM (Fig. 32) and R-CISL-HIRLAM (not shown).

With regard to the wave amplitude the cell-integrated models and traditional models behave differently. The lower resolution runs with the cell-integrated models have lower/higher minimum/maximum surface pressures compared to the high-resolution runs with the same model. On the contrary, HIRLAM behaves as the reference solutions provided by JW2005, i.e. the higher the resolution, the higher/lower the maximum/minimum surface pressures.

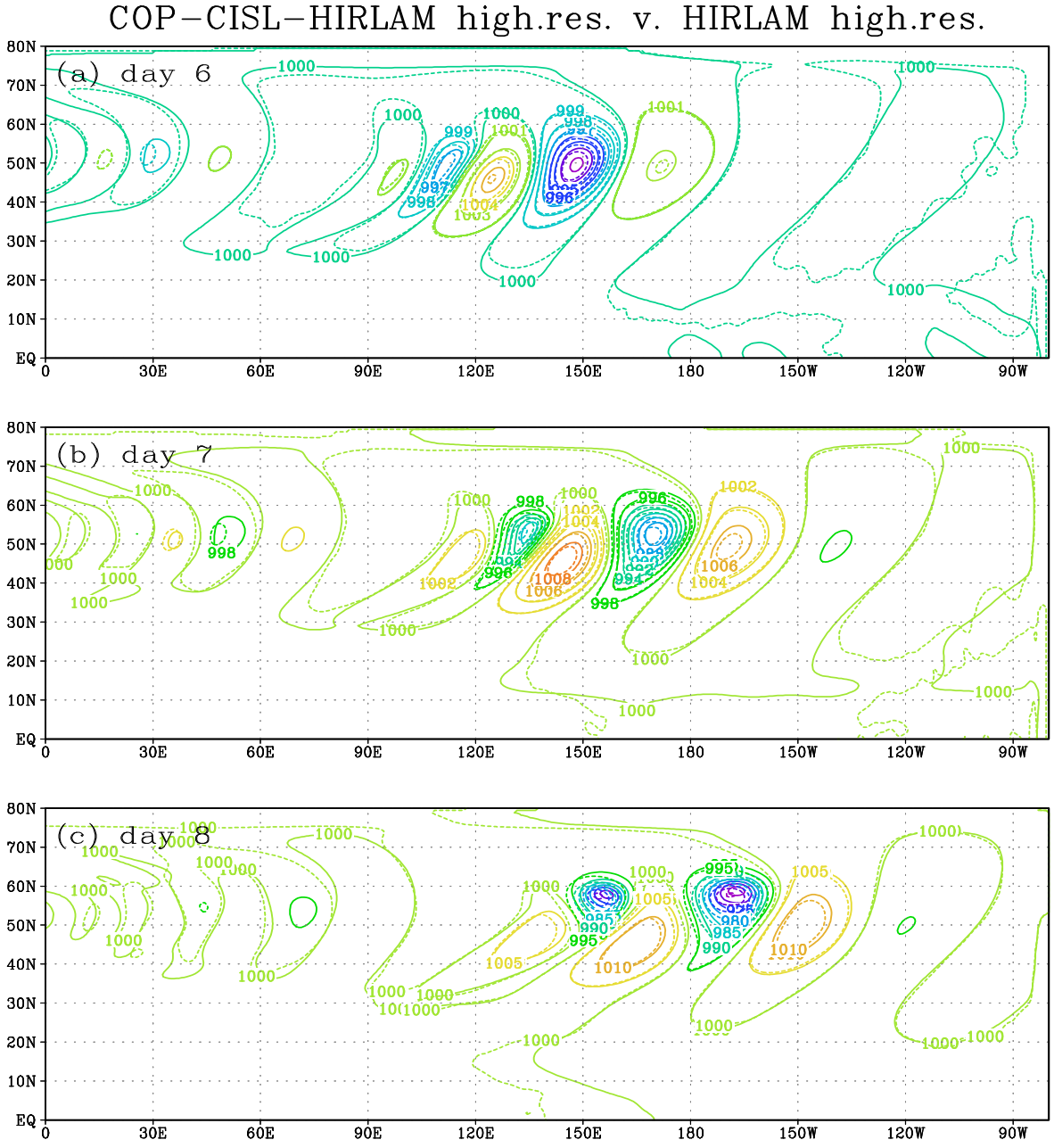


Fig. 29: Surface pressure (hPa) for the high-resolution runs with COP-CISL-HIRLAM (solid contours) and HIRLAM (dashed contours) at days (a) 6, (b) 7 and (c) 8, respectively. The solid and dashed contour colors represent the same contour levels. The contour interval is (a) 1 hPa, (b) 2 hPa, and (c) 5 hPa.

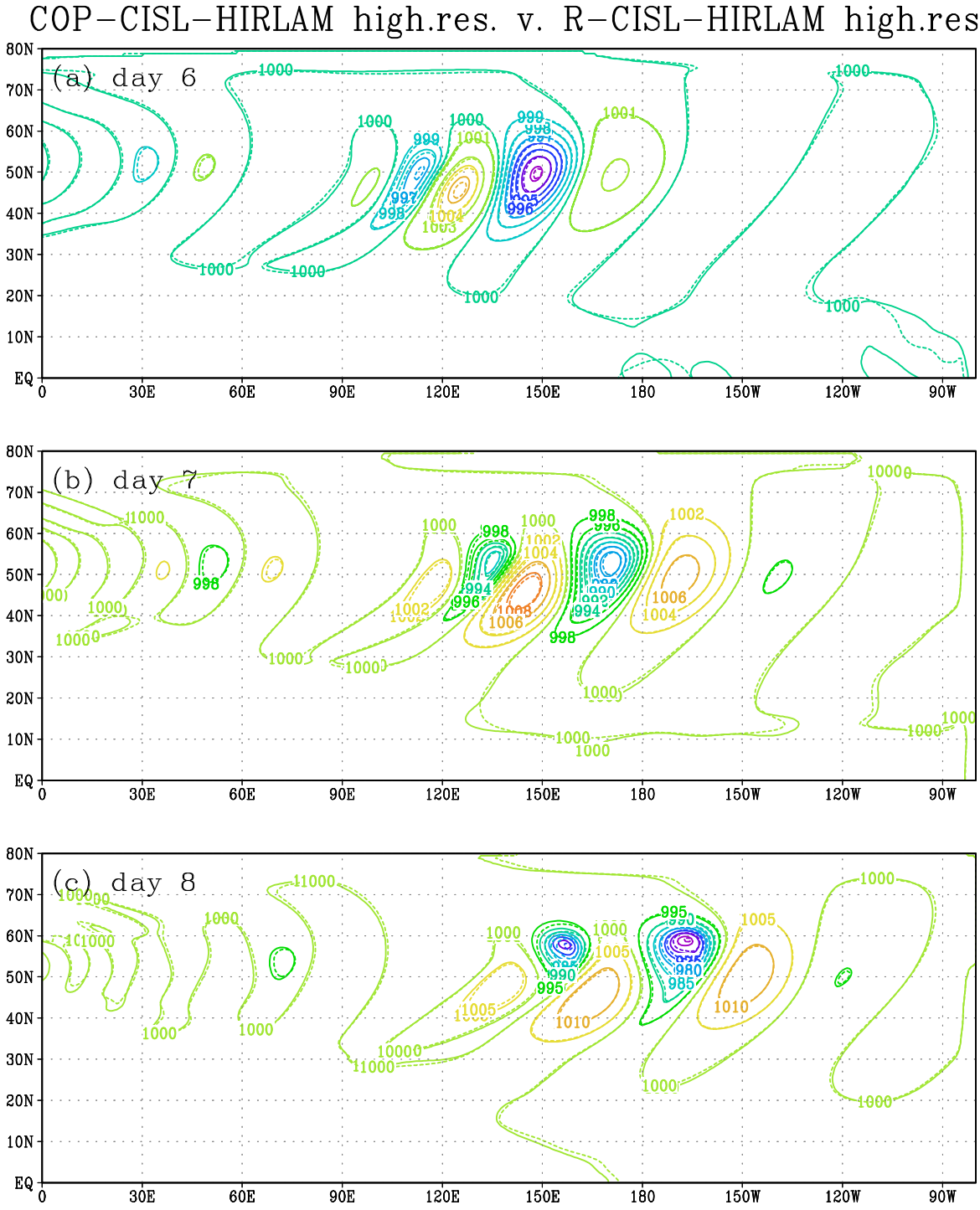


Fig. 30: Same as Fig. 29 but for COP-CISL-HIRLAM (solid contours) and R-CISL-HIRLAM (dashed contours).

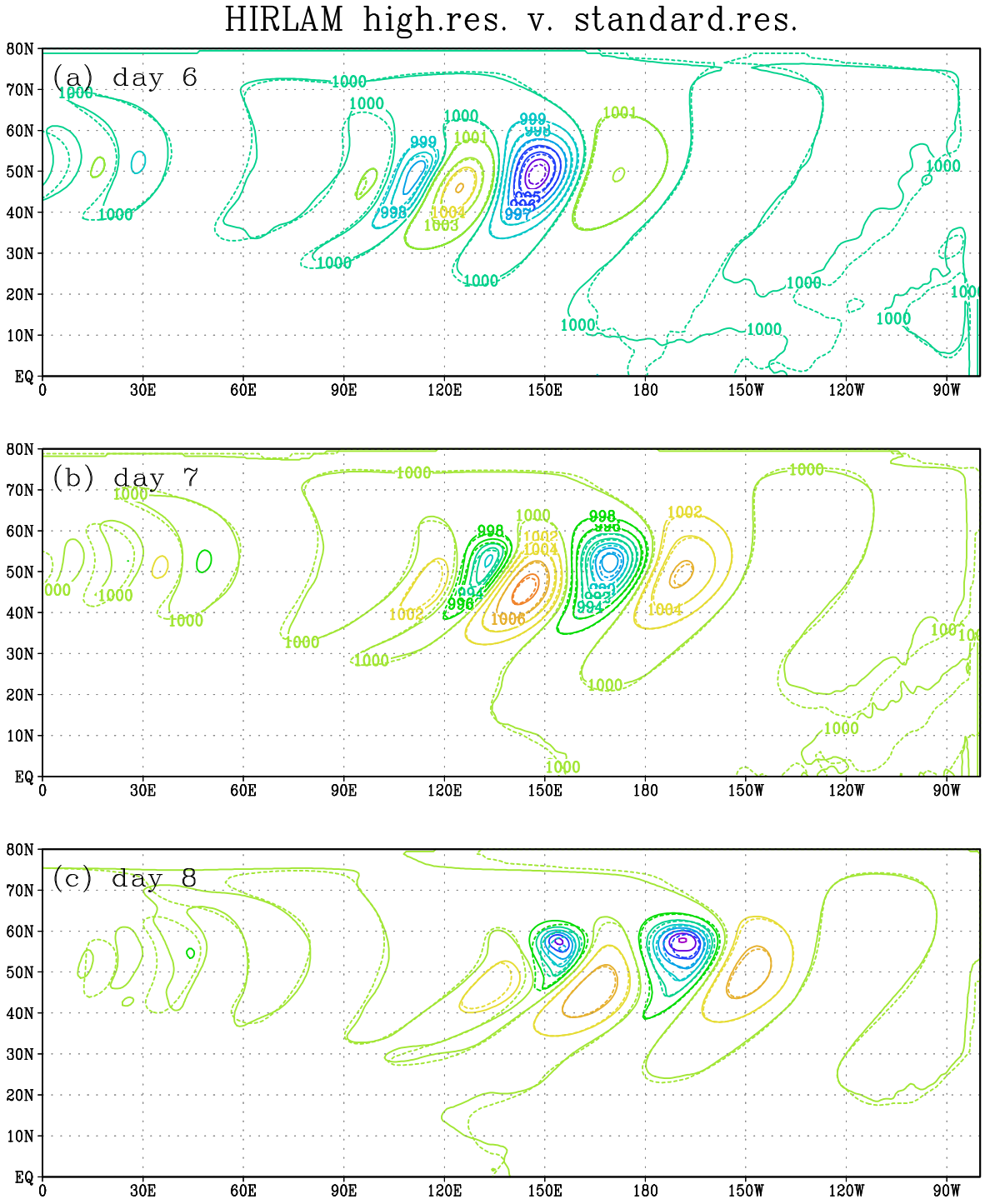


Fig. 31: Same as Fig. 29 but for the high-resolution (solid contours) and standard-resolution (dashed contours) runs with HIRLAM.

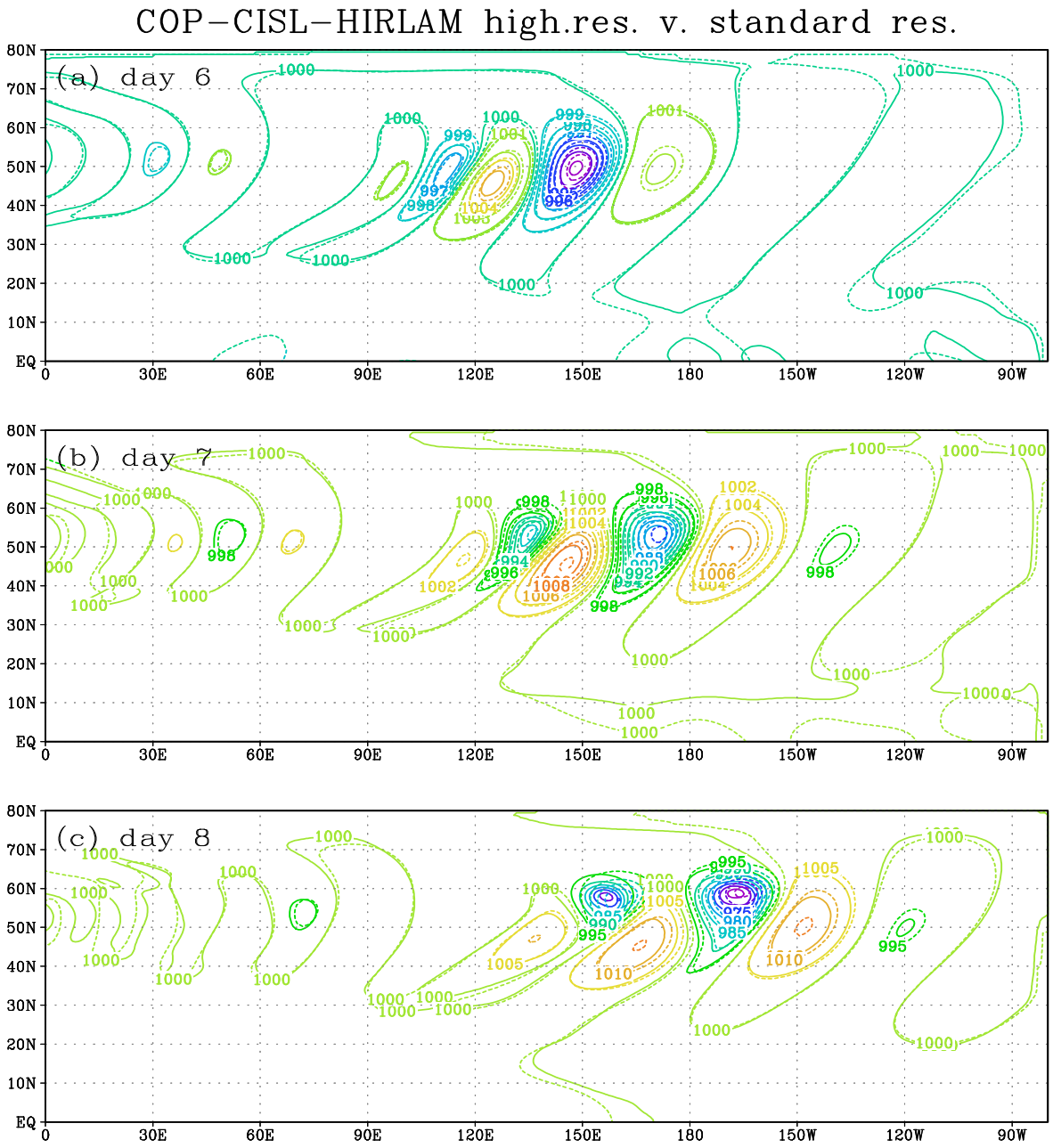


Fig. 32: Same as Fig. 29 but for the high-resolution (solid contours) and standard-resolution (dashed contours) runs with COP-CISL-HIRLAM.

Modified model runs For COP-CISL-HIRLAM different configurations have been examined to assess their influence on the forecast. The modifications tested are:

- **Number of iterations in the trajectory algorithm** COP-CISL-HIRLAM uses two iterations in the departure-point algorithm. Increasing the number of iterations to three has minimal effect on the forecast (not shown); the changes in the extrema in the surface pressure are on the order of 0.01 hPa at day 8. When using only one iteration the changes in the surface pressure field are clearly visible (see Fig. 33). The eastern part of the wave propagates faster compared to the model using two or three iterations, while the western part of the wave propagates more slowly. The extrema are weaker for the one-iteration model version. These tests suggest that the optimal number of iterations is two.

- **Constraints on the reconstruction of the sub-grid-scale distribution** In COP-CISL-HIRLAM a monotonic filter is used in the parabolic reconstruction of the sub-grid-scale distributions. Running the model with no limiter or a positive definite filter did not impact the forecast (not shown). This suggests that the limiters are not invoked very often.

- **Reference temperature** For stability the cell-integrated models use a reference temperature higher than the reference temperature used in traditional models. The higher the reference temperature, the deeper the equivalent depths, and consequently, the faster the gravity waves propagate in the semi-implicit model. COP-CISL-HIRLAM uses a reference temperature of 400 K. By decreasing the reference temperature to 350 K the model still ran stably and no clear differences from the run with $T^{ref} = 400$ K were observed (not shown). By decreasing the reference temperature even further to 300 K the forecast became noisy and after day 6 it became unstable.

- **Order of sub-grid-scale reconstruction** As expected, by reducing the order of the sub-grid-scale reconstructions to zeroth order, i.e. using cell averages for the reconstruction, the amplitude of the wave is reduced (see Fig. 34). The extrema in the surface pressure, however, are located toward the south-west compared to the standard run with the same model. In other words, the phase of the wave is closer to HIRLAM when increasing the numerical diffusion in the CISL model.

- **Decreasing the order of horizontal diffusion** By using fourth-order instead of sixth-order implicit horizontal diffusion, a less scale-selective damping is introduced, i.e. low-order diffusion damps the longer wave lengths more than the high-order diffusion (see e.g., Fig. 4 in McDonald 1994). The coefficients in the model runs are chosen such that the e-folding time for the 2-Δx wave (on the order of 30 minutes) is approximately the same, regardless of the order of the horizontal diffusion. As in the case of increased numerical diffusion, the low pressures are located further southward when decreasing the order of horizontal diffusion (Fig. 35). Also the phase speed of the wave train is changed. Exactly the same is observed for HIRLAM when going from ∇^6 to ∇^4 diffusion (not shown). Doubling the diffusion coefficient when using sixth-order horizontal diffusion has little effect on the simulation (not shown). These results indicate that the order of the horizontal diffusion to some extent

control the simulated solution. This was also recently observed by Polvani et al. (2004) in their baroclinic test case.

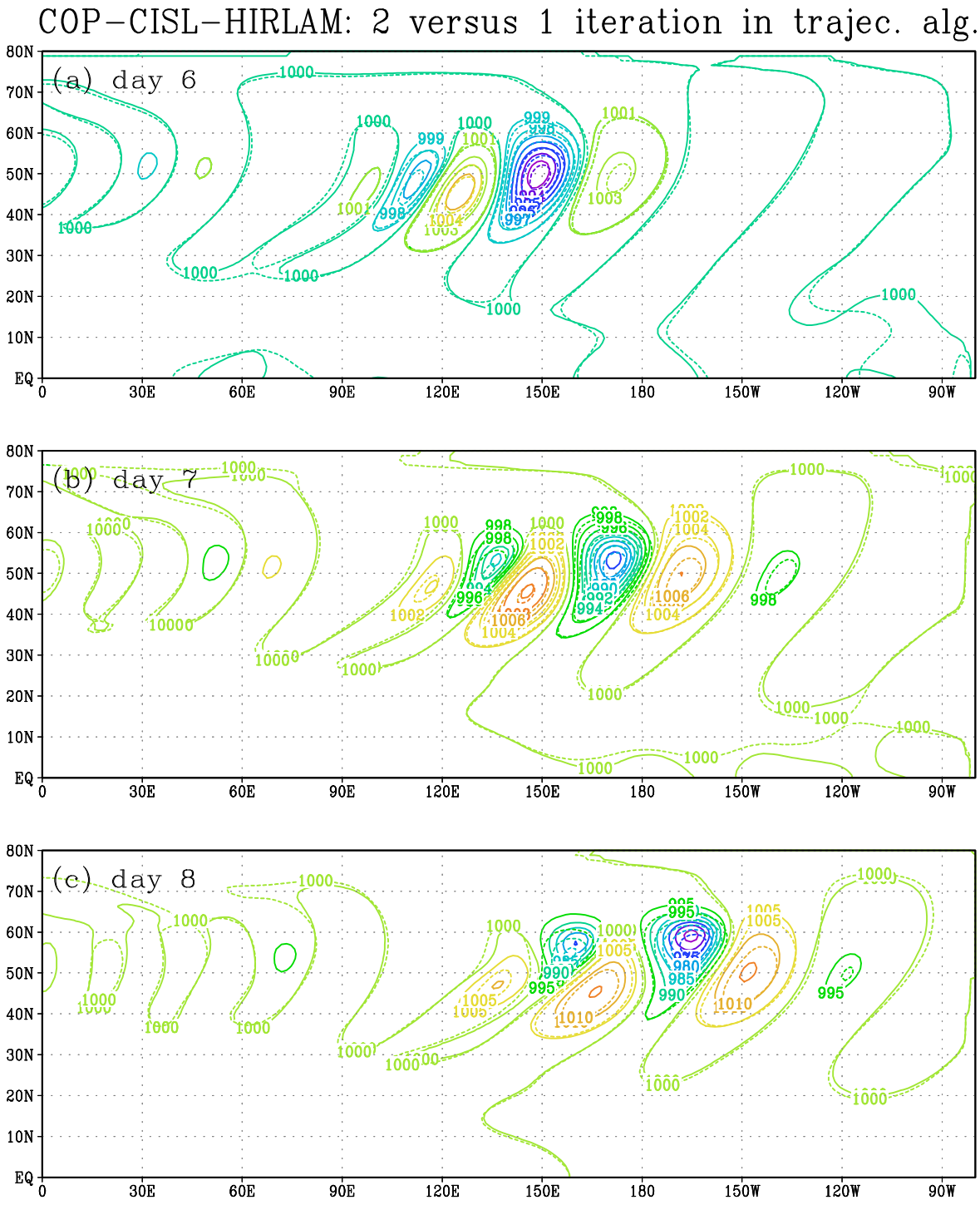


Fig. 33: Same as Fig. 29 but for COP-CISL-HIRLAM using one (solid contours) and two (dashed contours) iterations in the trajectory algorithm.

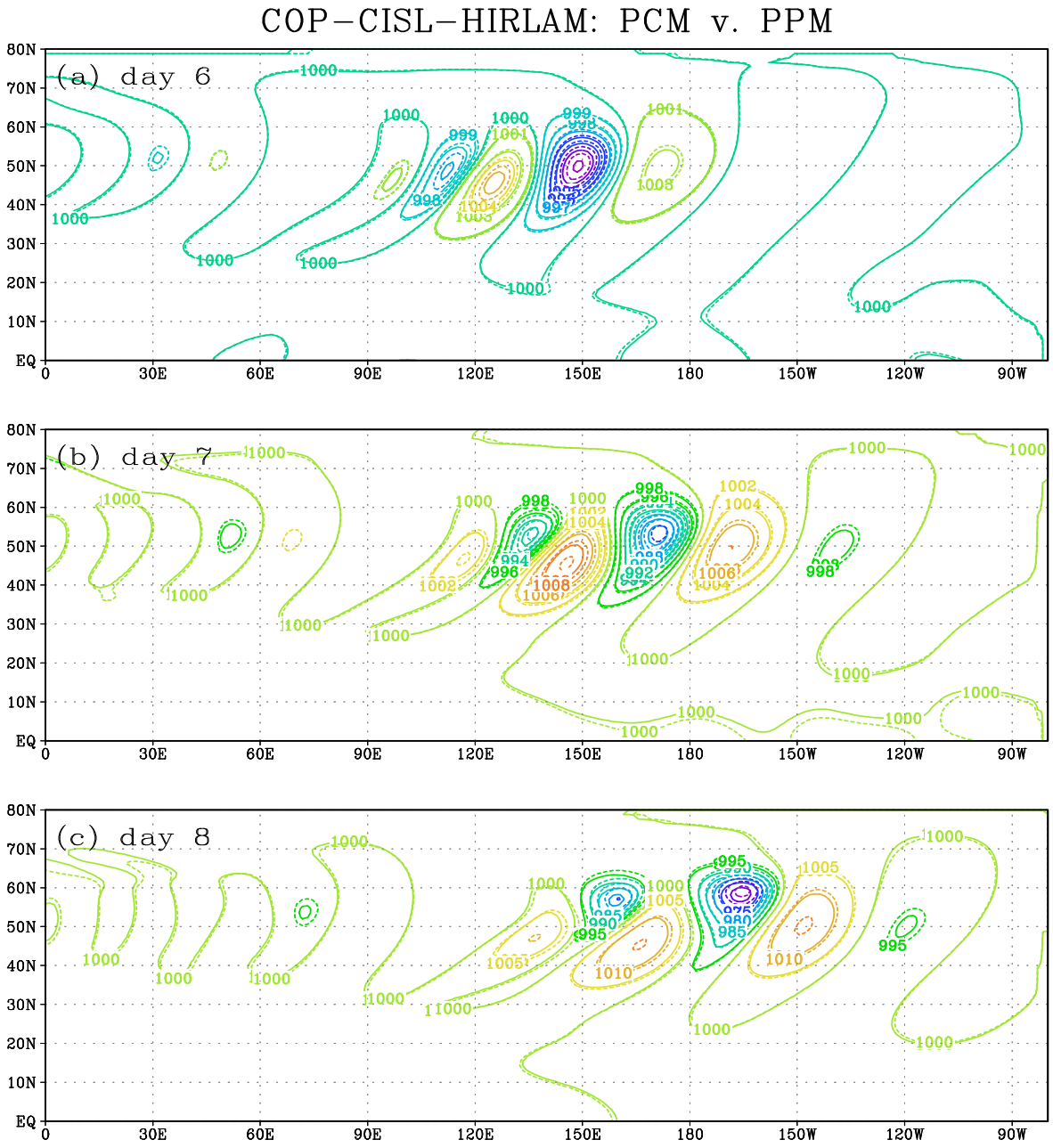


Fig. 34: Same as Fig. 29 but for standard COP-CISL-HIRLAM using the piecewise-parabolic method (solid contours) and the piecewise-constant method (dashed contours) for the sub-grid-scale reconstructions. The algorithm has not been modified regarding vertical interpolation.

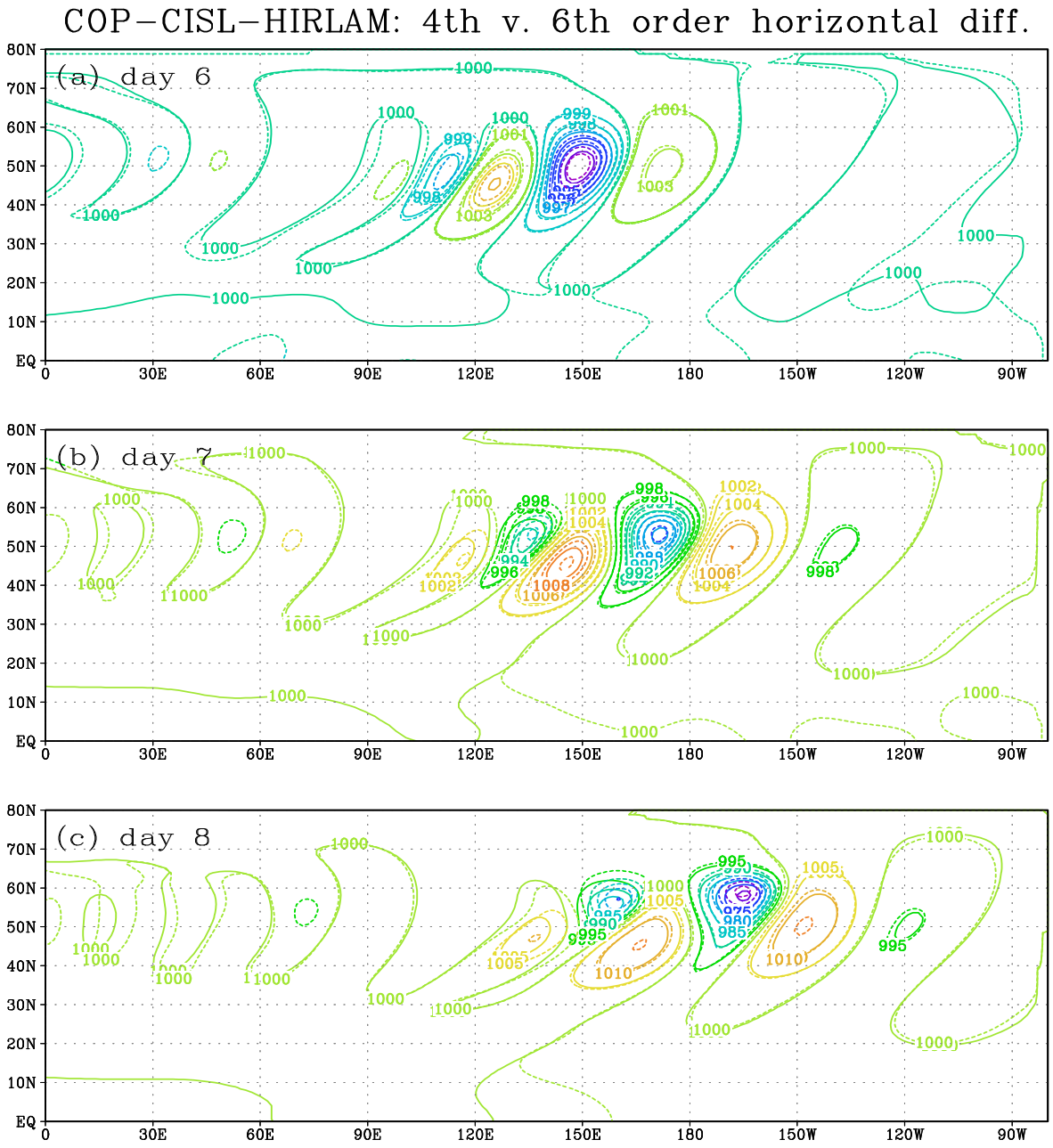


Fig. 35: Same as Fig. 29 but for COP-CISL-HIRLAM using fourth-order (solid contours) and sixth-order (dashed contours) implicit horizontal diffusion.

4.3.6 *Effects of the boundaries*

The effects of the boundaries are assessed in this subsection. In the HIRLAM setup the solution is relaxed toward the initial condition in the boundary region and the elliptic solver assumes zero divergence on the boundaries. The runs with the global models show very little activity in the Southern Hemisphere during at least the first 8 days of integration. Hence, by considering the solution in the Southern Hemisphere for the LAM, the effects introduced at the boundaries can be studied in isolation.

Fig. 36 shows the surface pressure at days 1 to 8 for the high-resolution run with R-CISL-HIRLAM. For the high-resolution global runs the departures from 1000 hPa are on the order of 0.01hPa for days 1 to 8 in the Southern Hemisphere. For the LAM runs a wave is triggered at the boundaries (hereafter referred to as the "boundary wave"), which deepens as it travels eastward. At day 8 the amplitude has reached several hPa. The "boundary wave" resembles the wave triggered by the velocity perturbation in the Northern Hemisphere, but the perturbations in the surface pressure are on a smaller scale.

A "boundary wave" in the Southern Hemisphere has a counterpart symmetric about the equator in the Northern Hemisphere. It is possible that the "boundary wave" in the Northern Hemisphere interacts with the baroclinic wave triggered by the velocity perturbation, and thereby render comparison with global runs precarious. In addition, the north and south boundaries may also influence the computed solution compared to the global run.

When running COP-CISL-HIRLAM with the center of the initial perturbation located 60° from the western boundary instead of 100°, the effects are clearly visible in the surface pressure field (see Fig. 37). The "boundary wave" is closer to the baroclinic wave train in the simulation where the east border is closer to the initial perturbation, and by day 6 this is visibly affecting the tail of the baroclinic wave. As a result the simulations evolve differently, emphasizing that the effects of the boundaries are affecting the model simulations. A doubling of the width of the boundary-relaxation zone does not visibly reduce the impact on the surface pressure of the baroclinic wave train (not shown).

In this connection it is also noted that, even though HIRLAM uses decentering and filtering of the non-linear terms in time, the solution is noisy in the less active parts of the domain (see Fig. 38). On the contrary, the cell-integrated models are smooth throughout the domain. In "real world" applications it is very important that the fields are not noisy. Hence it is an important property of the CISL approach that they show no sign of noise problems.

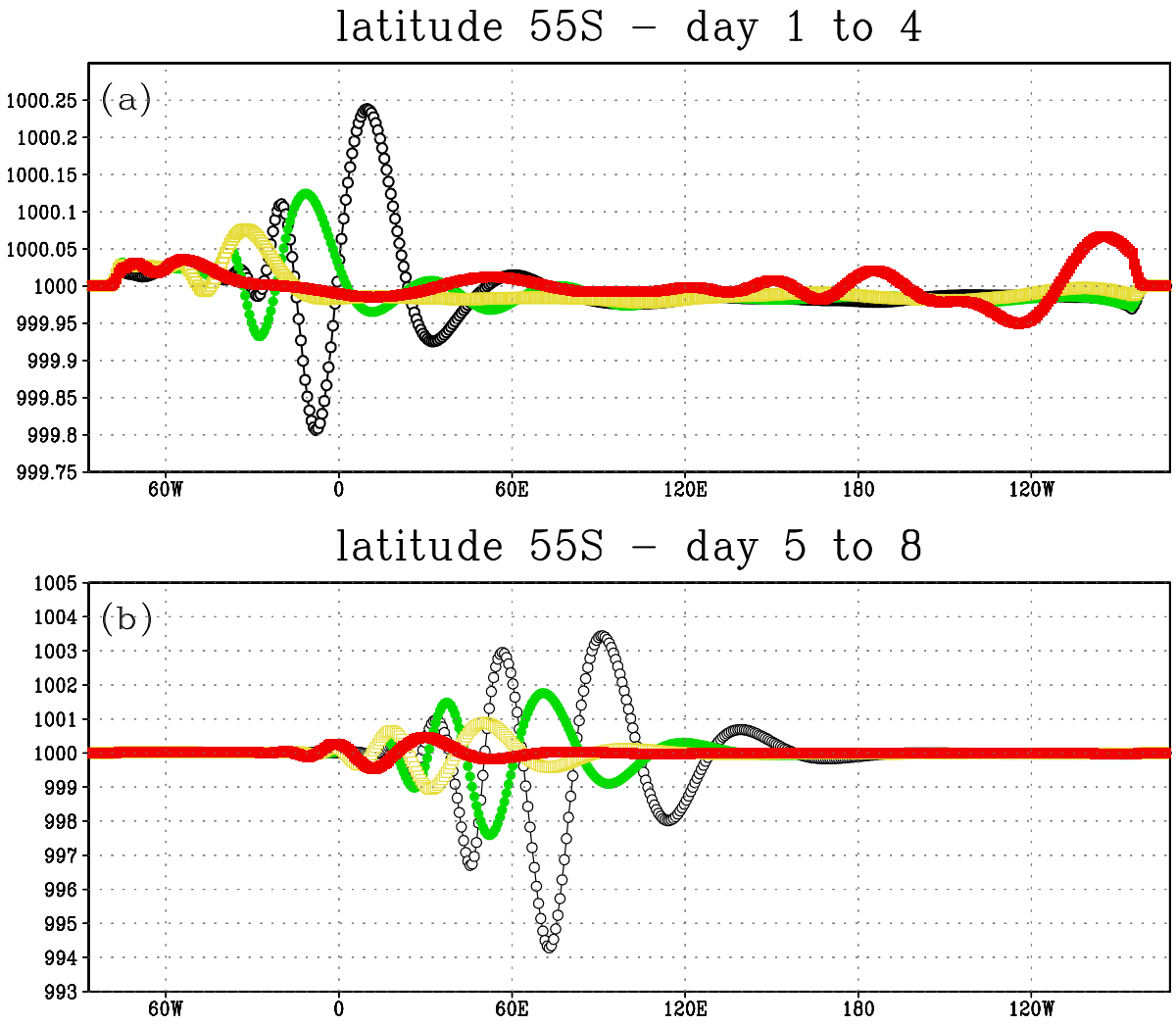


Fig. 36: The surface pressure (hPa) for the high-resolution run with R-CISL-HIRLAM along latitude 55S. (a)/(b) shows day 1/5 (red), 2/6 (yellow), 3/7 (green) and 4/8 (black). The plots show the entire longitudinal extent of the domain including relaxation zone and the passive boundary points where the solution is held fixed at the initial condition. The wave train on the figure is triggered at the boundaries and is referred to as the “boundary wave”.

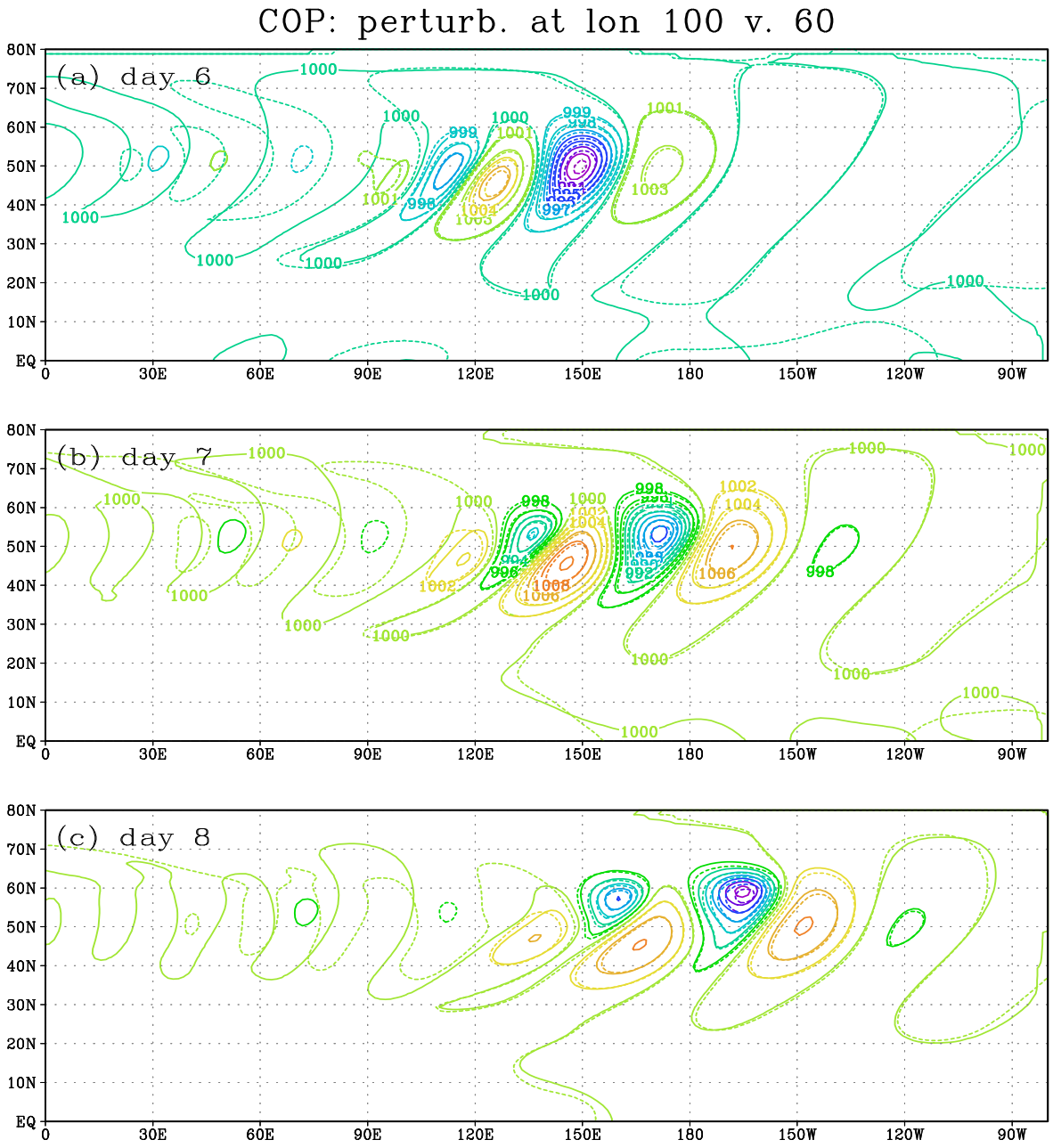


Fig. 37: Same as Fig. 29 but for standard-resolution run with COP-CISL-HIRLAM with the western border located 100° (solid contours) and 60° from the center of the initial perturbation.

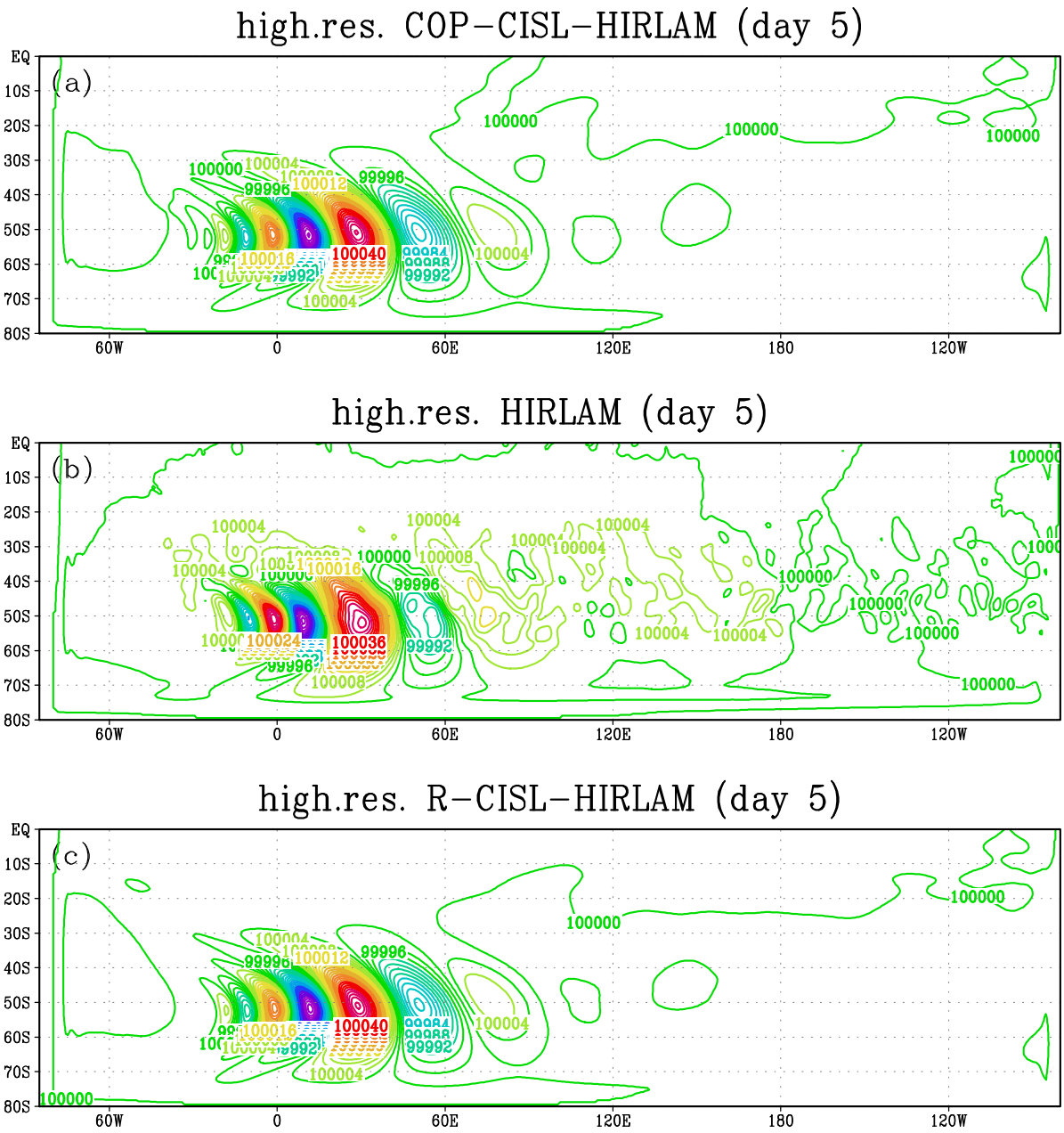


Fig. 38: The surface pressure (Pa) in the Southern Hemisphere at day 5 for the high-resolution runs with (a) COP-CISL-HIRLAM, (b) HIRLAM and (c) R-CISL-HIRLAM, respectively. The contour interval is 2 Pa. Note the noise in the simulation with HIRLAM.

4.3.7 *Comparison of LAM runs with global model simulations*

The LAMs use several approximations which can be inaccurate to use in a global model. These are:

- A 'half-implicit'-Coriolis scheme is used, i.e. f is divided into two parts (a constant f_0 and a derivation from the constant f') where the constant part is treated semi-implicitly and the non-constant part is treated explicitly. In a global model the departure from f_0 is large when approaching the polar areas.
- The departure point algorithm works directly in (λ, θ)-coordinates, i.e. the angular displacements over one time step (when not including the acceleration) are given by

$$\Delta\lambda = \Delta t \, \frac{d\,\lambda}{dt}(t) + \mathcal{O}(\Delta t^2), \tag{162}$$

$$\Delta\theta = \Delta t \, \frac{d\,\theta}{dt}(t) + \mathcal{O}(\Delta t^2), \tag{163}$$

  where

$$\frac{d\,\lambda}{dt}(t) = \frac{u\,[\lambda(t), \theta(t), t])}{a\,\cos\theta(t)}, \tag{164}$$

$$\frac{d\,\theta}{dt}(t) = \frac{v\,[\lambda(t), \theta(t), t])}{a}. \tag{165}$$

  These approximations become increasingly inaccurate the closer the departure point is to the poles. In global models the departure points are usually computed in three dimensions as suggested by Ritchie (1987,1988).

  Similarly, the upstream interpolation in the remaining model equations are performed directly in (λ, θ)-coordinates which becomes increasingly inaccurate the closer the interpolation is to the poles. The momentum equations are solved on component form in HIRLAM and thus the metric terms appear explicitly. Near the poles the metric terms become unstable. Therefore the momentum equations are normally discretized on vector form in global models (e.g., Ritchie 1988).

- When approaching the poles the CISL scheme must also be modified as described in Chapter II.

The approximations mentioned above and the effects introduced at the boundaries suggest that it can be precarious to compare global simulations with LAM runs. In these experiments the LAM approximations have been "stretched" to the limit since the meridional extension of the active domain is 160°. It can therefore not be ruled out that the LAM has introduced inaccuracies related to the spherical geometry. Here the global model run with SL-CAM3 (T340 horizontal resolution)

is compared to the high-resolution LAM runs. Note that the linear grid in the SL-CAM3 model has approximately double horizontal resolution compared to the high-resolution LAM resolution. In the vertical SL-CAM3 has twice as many levels as the high resolution LAMs.

The HIRLAM simulation has a stronger development in terms of extrema in the surface pressure compared to the global solution (see Fig. 39). At day 8 the minimum/maximum surface pressures are 967.02 hPa/1013.16 hPa and 964.56 hPa/1013.44 hPa for SL-CAM3 and HIRLAM, respectively. Especially the minimum p_s is significantly smaller in HIRLAM compared to SL-CAM3, although the models are based on the traditional semi-Lagrangian method. The location of the extrema in the two models also differs slightly. The high-resolution run with COP-CISL-HIRLAM differs even more from the global model run. At day 8 the minimum/maximum surface pressures are 967.02 hPa/1013.16 hPa and 963.06 hPa/1014.06 hPa for SL-CAM3 and COP-CISL-HIRLAM, respectively, and the extrema are located further toward north-east compared to the global model run. These results suggest that the boundaries and the approximations to spherical geometry in the LAM do affect the computed solution significantly.

The location of the minima in p_s in HIRLAM is closer to the minima in p_s in the global model run compared to the cell-integrated models. The question is why this is the case. One could suspect the "predictor-corrector" approach used in the semi-implicit CISL scheme to be the cause of the difference in phase speed of the baroclinic wave. To investigate if the explicit correction term was the cause of the differences in phase speed, the COP-CISL-HIRLAM was run with a time step of 150 s both explicitly and semi-implicitly (see Fig. 42). Although the two simulations were run with the western boundary located 100° and 60° from the western boundary, the differences in phase speed in these simulations are small compared to the difference between the location of the minimum p_s in HIRLAM and the cell-integrated models. Hence the correction term in the semi-implicit time-stepping scheme does not seem to introduce an erroneous phase speed. By comparing the COP-CISL-HIRLAM run using the standard time-step with the explicit version of COP-CISL-HIRLAM using a 150 s time-step, it is seen that the explicit run has a phase speed closer to HIRLAM. The explicit simulation, however, uses a time step that is 12 times smaller than the one used for the semi-implicit model (time steps larger that 150 s resulted in an unstable simulation in the explicit model). The explicit model does therefore introduce much more numerical diffusion due to the upstream remappings compared to the run with long time steps. It is therefore speculated that the slow phase speed when using $\Delta t = 150$ s compared to $\Delta t = 30$ minutes is caused by the extra numerical diffusion.

4.3.8 *Summary of the preliminary experiments*

The Jablonowski-Williamson baroclinic wave test case has been used to study the properties of the new dynamical cores compared to HIRLAM. The cell-integrated models could run stably without decentering and filtering of the non-linear terms in time. On the contrary, HIRLAM must use decentering and filtering of the non-linear terms in time in order to provide stable simulations. Both models have been run with the same sixth-order horizontal diffusion. Despite the filters in

HIRLAM the solutions were noisy in less active regions of the domain. The baroclinic development in COP-CISL-HIRLAM was stronger compared to HIRLAM. For R-CISL-HIRLAM the extrema in the surface pressures were similar to HIRLAM. The location of the extrema in the surface pressure field in HIRLAM and the cell-integrated models did not coincide exactly.

It was not possible to determine which model was the most accurate by comparing the LAM simulations with the reference solutions from global model runs. The lack of "adaptation" to the sphere and the boundaries in the LAMs influence the simulations to such an extent that judging the accuracy by comparing global and LAM runs is precarious. It is, however, speculated that the cell-integrated models are the most accurate for the following reasons:

- The cell-integrated models conserve mass locally and therefore the mass field is described more accurately in the inherently mass-conservative models.
- The CISL models discretize the continuity equation, and in one model version also the energy conversion term, in a Lagrangian and more consistent fashion.
- The departure point algorithm is consistent with the explicit CISL continuity equation, adding to the consistency of the model.
- In simulations of the baroclinic wave train it is seen that by increasing the numerical diffusion the solution "moves toward" the HIRLAM solution. For the location of the baroclinic development this also applies to the similarity with the global models. It is well known from idealized advection tests that traditional semi-Lagrangian methods are much more diffusive than CISL methods. Consequently, it can be speculated that the cell-integrated dynamical core is more accurate since it has less numerical diffusion. When replacing the sixth-order horizontal diffusion with fourth-order a similar change is observed.

However, to make a more comprehensive analysis of the new dynamical cores it is necessary to design and implement global model versions. The global solution could be compared with the JW2005 reference solutions, and other conservation properties, such as total energy, entropy, etc., could be evaluated.

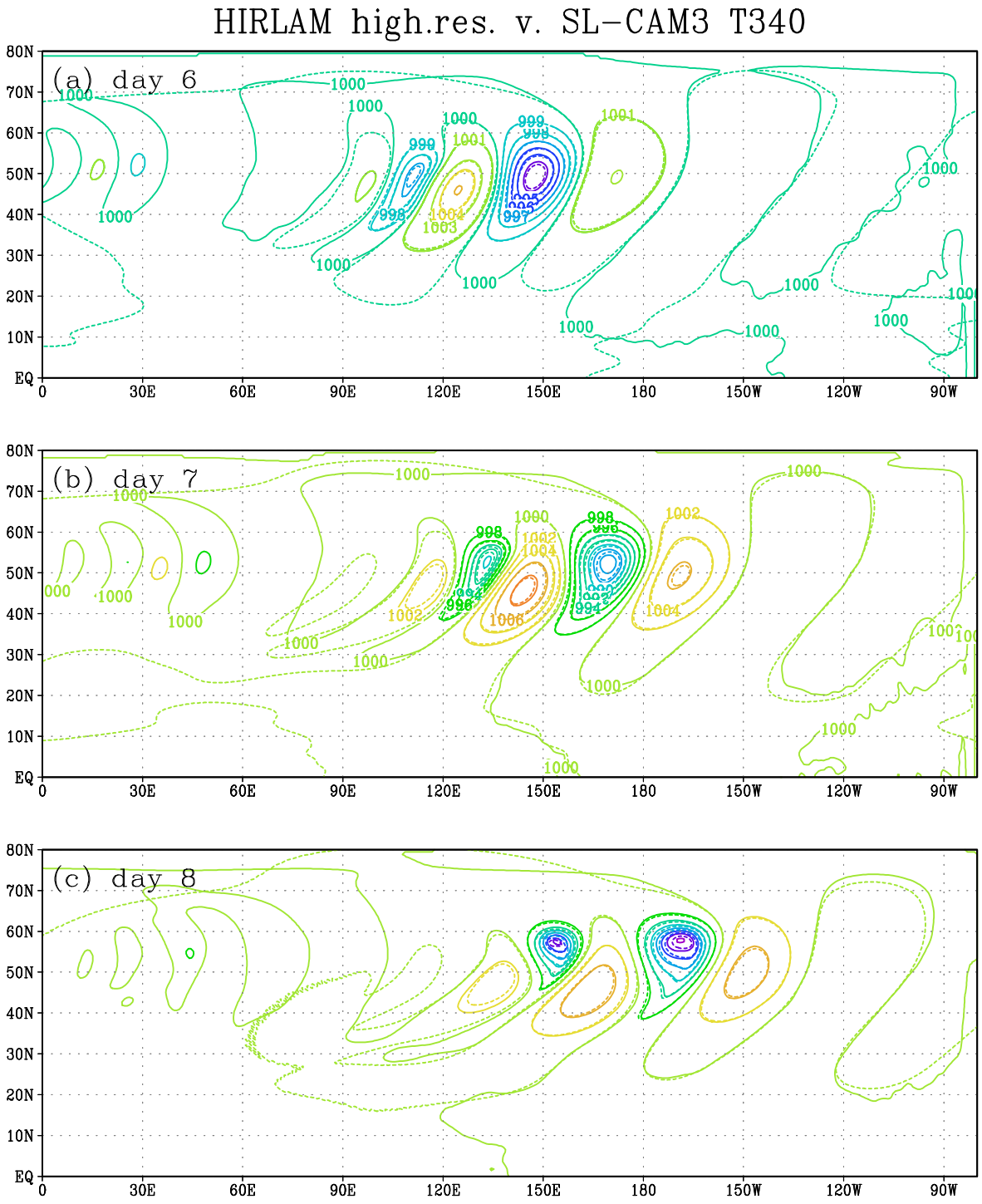


Fig. 39: Same as Fig. 29 but for high-resolution run with HIRLAM (solid contours) and SL-CAM3 using T340 spectral truncation (dashed contours).

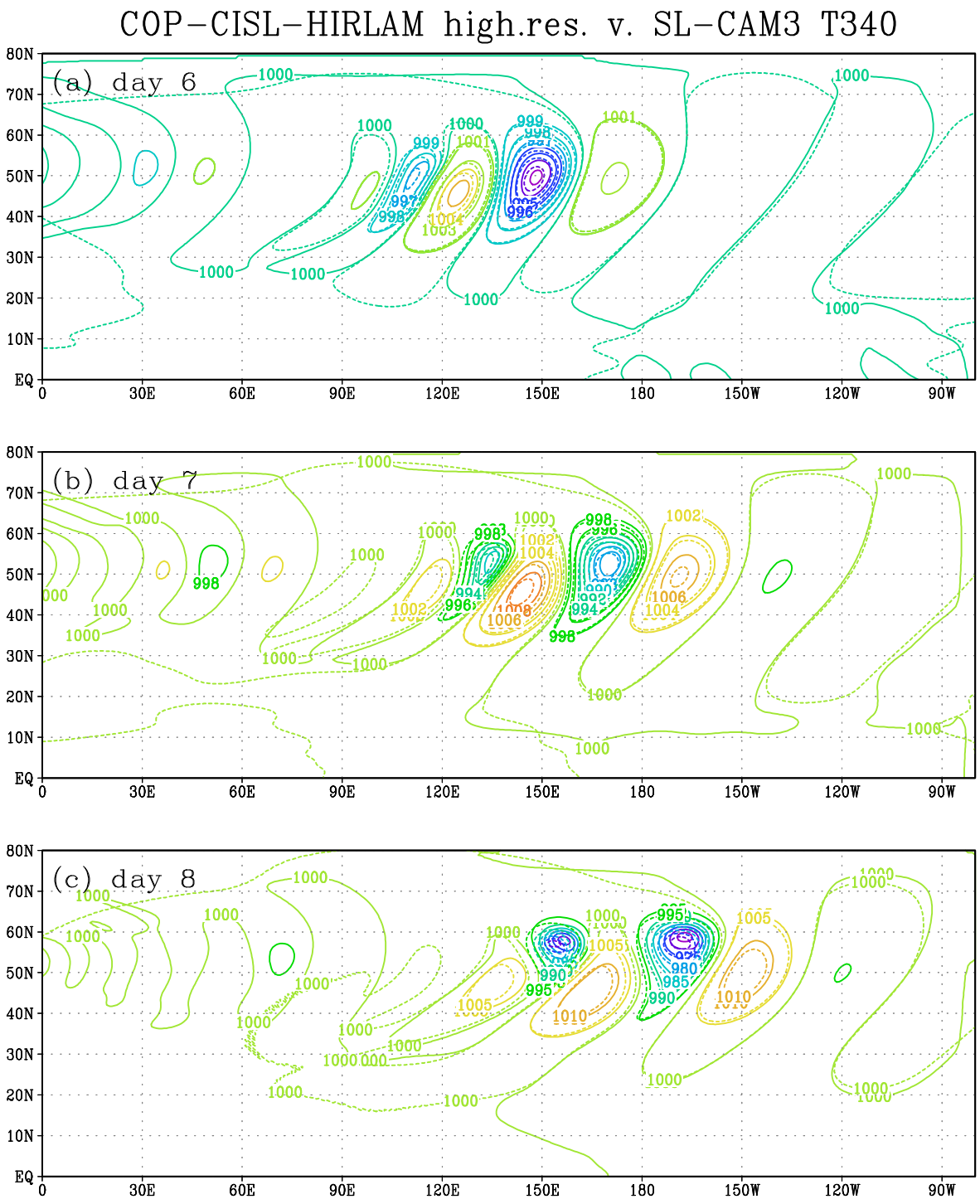


Fig. 40: Same as Fig. 29 but for high-resolution run with COP-CISL-HIRLAM (solid contours) and SL-CAM3 using T340 spectral truncation (dashed contours).

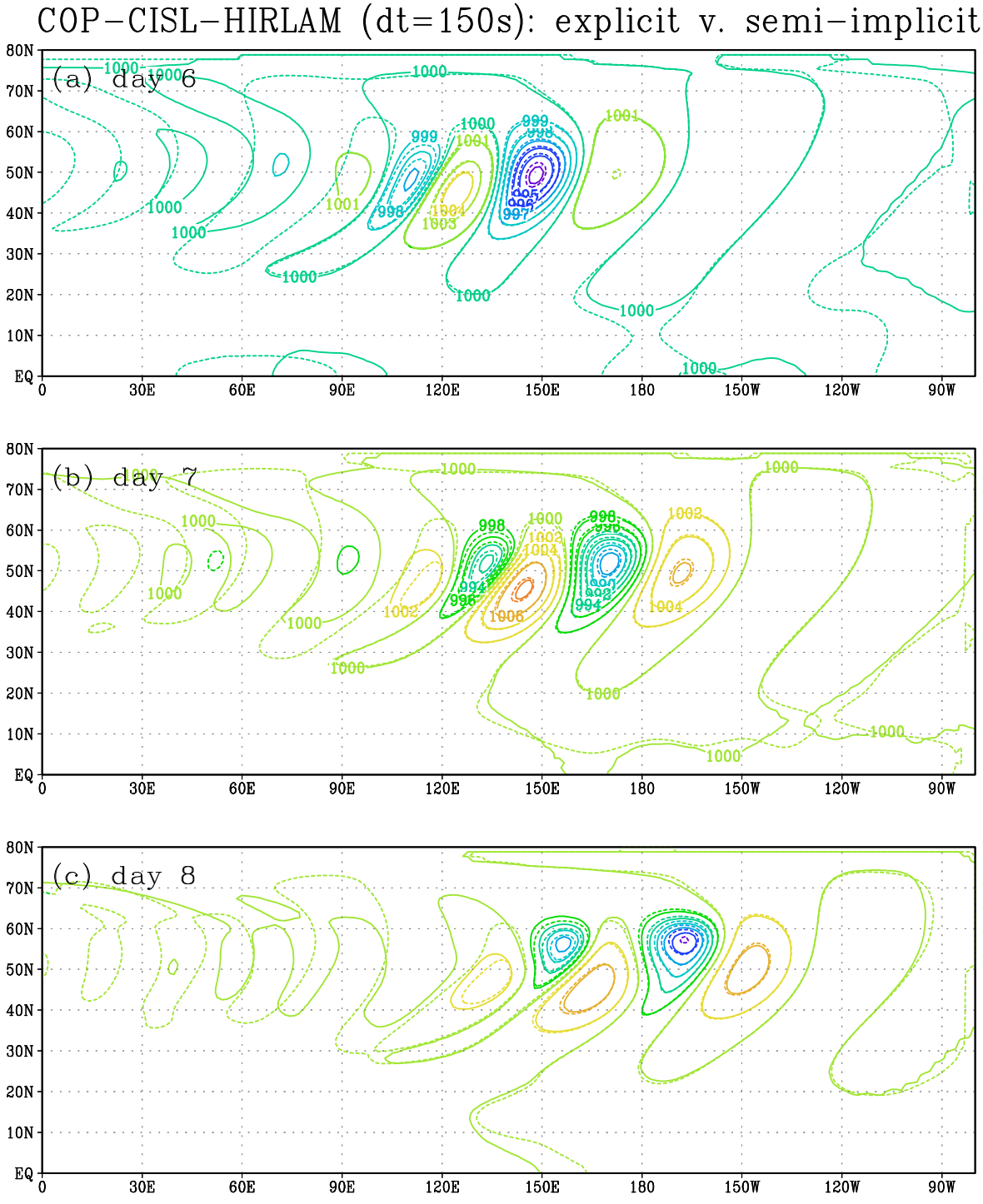


Fig. 41: Same as Fig. 29 but for standard resolution run with COP-CISL-HIRLAM using explicit (solid contours) and semi-implicit (dashed contours) time stepping. Both simulation are with the same time step of 150 s. In the explicit and semi-implicit run the west border of the domain is located 60° and 100° from the center of the initial perturbation, respectively.

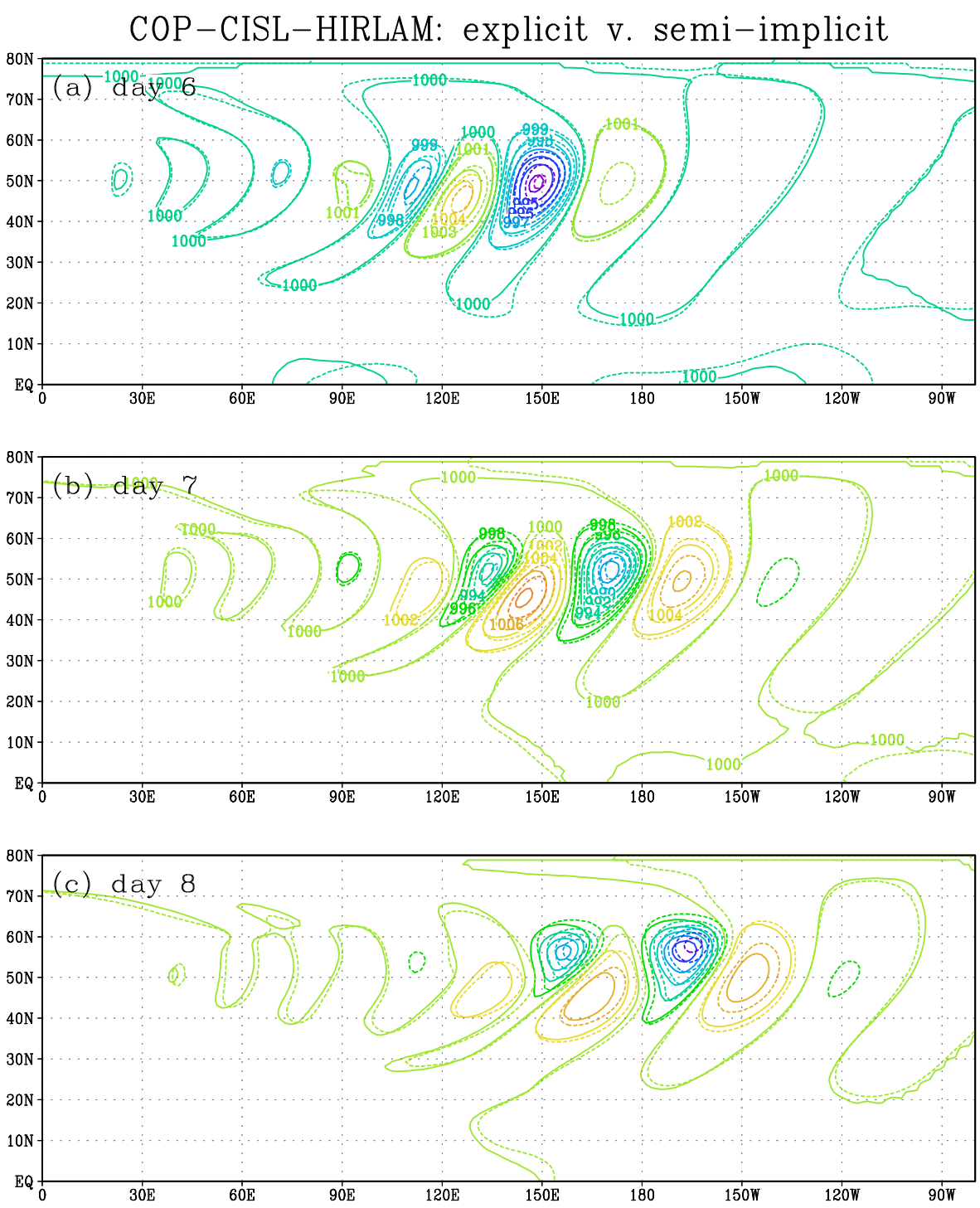


Fig. 42: Same as Fig. 29 but for standard resolution run with COP-CISL-HIRLAM using explicit (solid contours) and semi-implicit (dashed contours) time stepping. The semi-implicit simulation uses a 30 minuntes time step and the explicit run is with a 150 s time step. The west border of the domain is located 60° from the center of the initial perturbation, respectively.

4.4 Discussion on the conservation of the vertical discretization

Ideally the discretized baroclinic equations should maintain analogs of integral properties of the continuous equations such as local and global mass conservation, conservation of total energy, total potential enthalpy and total entropy (e.g., Arakawa and Lamb 1977). Most baroclinic models use some kind of staggering of the variables in the vertical, i.e. some variables are defined at so-called 'full levels' while others are defined in between 'full levels' at 'half levels'. Here a Lorenz type of staggering is used where the pressure and geopotential are defined at 'half levels' and the temperature and velocity components are held at 'full levels' (Lorenz 1960). In this situation there is some degree of freedom on how to approximate the PGF at 'full levels' since variables in the expression for the PGF are held at 'half levels'. Modelers generally attempt to use these degrees of freedom to ensure that the vertical discretization does not introduce spurious unphysical effects; for example by maintaining analogs of integral constraints. Consistency of the vertical differencing of the primitive equations has received some attention in the literature (Arakawa and Lamb 1977; Arakawa and Suarez 1983; Williamson 1988).

It is generally believed as very important that the vertical discretization of the PGF and the energy conversion term in the thermodynamic equation should not be a spurious source or sink for energy in the transformation from kinetic to potential energy. If the vertical discretization does not introduce spurious energy, it is refereed to as being *energy consistent*. In a hybrid vertical coordinate system the PGF involves the horizontal discretization of the geopotential. Consequently the discretization of the hydrostatic equation influences the *energy consistency*. Note that *energy-consistent* models do not necessarily conserve total energy but that are only guaranteed not to introduce spurious energy through their vertical discretization.

A condition for *energy consistency* can be derived by considering global integrals of the change in kinetic energy, $\frac{\partial K}{\partial t} = \mathbf{v}\,\frac{\partial \mathbf{v}}{\partial t}$, and the change in internal energy, $\frac{\partial I}{\partial t} = \frac{\partial}{\partial t}\,(c_p\,T)$. Substituting the equations of motion and hydrostatic balance into these expressions and performing algebraic manipulations, it is realized that, if the vertical discretization should be *energy consistent*, terms in the PGF should balance the energy conversion term in the thermodynamic equation

$$\int_0^1 \left\{ \nabla \cdot \left(\frac{\partial p}{\partial \eta}\,\mathbf{v} \right) \int_{p(\eta=0)}^{p(\eta=1)} R\,T_v\,d\ln p \right\} d\eta = \int_0^1 \left\{ R\,T_v\,\frac{\partial p}{\partial \eta}\,\frac{1}{p}\,(\mathbf{v}\,\nabla p - \omega) \right\}, \tag{166}$$

(see e.g., Collins et al. 2004 for details). In (166) it is assumed that horizontal integrals are taken and that divergence terms give no contribution.

In traditional semi-Lagrangian models the energy conversion term ω/p is discretized in an Eulerian fashion

$$\left(\frac{\omega}{p}\right)_k = -\sum_{\ell=1}^{NLEV} C_{k\ell}\,\left[\nabla \cdot (\mathbf{v}_\ell\,\Delta p_\ell)\,\right] + \mathbf{v}_k \cdot \nabla p_\ell, \tag{167}$$

where $C_{k\ell}$ is an approximation to $1/p_k$ for $\ell \leq k$ and zero elsewhere. Using (167) and the dis-

cretized inner integral on the left-hand side of (166), which is part of the hydrostatic balance,

$$\Phi_k = \Phi_s + R \sum_{\ell=\ell=k+1}^{NLEV} H_{k\ell} \, (T_v)_\ell \, , \tag{168}$$

where H is the hydrostatic matrix, it may be shown that the discretized analog of (166) holds column wise if

$$C_{k\ell} = \frac{H_{\ell k}}{\Delta p_k}. \tag{169}$$

(e.g., Collins et al. 2004). Note that the *energy consistency* is not enforced locally, but for each column of the model. This derivation assumes Eulerian finite differences, which give no global contribution by divergence terms. It is quite surprising that no matter how $H_{\ell k}$ is chosen the vertical discretization will not be a spurious source of energy as long as $C_{\ell k}$ satisfies (169). An almost universal choice of $H_{\ell k}$ is

$$H_{\ell k} = \begin{cases} 0 & , \ell < k, \\ \alpha_k & , \ell = k, \\ \Delta \ln p_\ell & , \ell > k, \end{cases} \tag{170}$$

(e.g., SB1981, Collins et al. 2004).

Similarly, requiring the vertical discretization not to be a spurious source or sink for total angular momentum, an expression for the PGF in terms of $H_{\ell k}$ results (SB1981). Using (170) for $H_{\ell k}$ results in a formula for the PGF which is identical to a finite difference discretization of

$$\nabla \ln p = \nabla \left[\frac{d}{dp} \left(p \, \ln p \right) \right]. \tag{171}$$

These vertical discretization choices will be referred to as the SB1981 discretizations.

Although these derivations assume Eulerian finite differences, the discretizations have been almost universally adopted in semi-Lagrangian models. That is, the Eulerian expression for the ω term and the PGF are used as grid-point values, and then interpolated along the trajectory according to the semi-Lagrangian scheme. Therefore in a traditional semi-Lagrangian model using the SB1981 expression for the PGF and energy conversion term, the conservation properties of the vertical discretization are only approximately maintained. Note also that current semi-Lagrangian models do not have intrinsic conservation properties. Therefore these models would not conserve total energy even if the vertical discretization was conservative.

Due to the similarity in the expressions for the energy conversion term based on the Eulerian discretization (167) and the relevant terms in the PGF, the condition for *energy consistency* becomes surprisingly simple for Eulerian models.

In a traditional semi-Lagrangian model there is *energy consistency* before the interpolation is applied and that is also the case for R-CISL-HIRLAM. For COP-CISL-HIRLAM the situation is

different. The Lagrangian discretization of the vertical velocity should balance Eulerian terms in order to have consistency (see (166)). A simple relationship between the coefficients of the hydrostatic equation and the energy conversion term in the thermodynamic equation can not be derived. Therefore the SB1981 discretizaton of the hydrostatic equation has been used although it does not guarantee *energy consistency*.

Chapter V

Summary and conclusions

5.1 Summary

A brief summary of the shallow water and baroclinic model derived in Chapter III and IV, respectively, is given.

5.1.1 *Shallow water model*

Two versions of a semi-implicit semi-Lagrangian shallow water model, which guarantees mass conservation by using cell-integrated methods for the continuity equation, have been developed. The two versions differ in the definition of their departure areas and in their sub-grid-scale representations. The version based on the NM02 scheme uses a pseudo-biparabolic representation which is the sum of two 1D piecewise parabolic functions along the Eulerian longitudes and latitudes. The version based on the advection scheme of Nair et al. (2002) uses a cascade approach with two successive remappings using 1D parabolic representations along Eulerian longitudes and Lagrangian latitudes, respectively. The mass is conserved locally in both versions and there is no need for ad hoc *a posteriori* mass-fixing algorithms as is the case for traditional semi-Lagrangian models. The cell-integrated models are efficient since they permit large time steps and their elliptic equation is in the same form as for the traditional semi-Lagrangian system and is relatively inexpensive to solve. In addition, desirable properties such as monotonicity and positive definiteness can easily be enforced.

The accuracy of the model has been assessed using selected test cases from the suite of test cases proposed by W92. For passive tracer advection the conservative transport schemes are much more accurate than traditional semi-Lagrangian methods. For shallow water model runs the accuracy and efficiency of the new models was estimated to be comparable to traditional semi-Lagrangian models. In all test cases the cell-integrated models performed better than the traditional semi-Lagrangian method with decentering. In a test case specially constructed to investigate potential noise problems, it was found that the traditional semi-Lagrangian method needed decentering for noise-free integrations. On the contrary, the cell-integrated models did not need decentering or other types of filters to maintain noise-free solutions. The shallow water tests performed revealed no clear difference in the accuracy of the two conservative model versions.

5.1.2 *Baroclinic model*

The shallow water model has been extended to three dimensions. The most apparent complication in the extension to three dimensions is probably the lack of three-dimensional CISL methods. Rather than developing a three-dimensional CISL scheme, hybrid trajectories in which the vertical part is forward and the horizontal part is backward have been used. Consequently, cells depart from model levels and the need for doing three-dimensional upstream integrals is no longer present. A price to pay is that the trajectories do not necessarily arrive at model levels and therefore a one-

dimensional remap of all prognostic variables from the arrival grid to the model grid is needed. The problem is, however, one-dimensional.

The vertical part of the trajectory scheme differs from traditional methods. Instead of approximating the vertical velocity by using formulas based on the continuity equation on grid-point form (as done in traditional semi-Lagrangian models), the vertical displacements are obtained by solving the explicit CISL continuity equation. That is, the explicit CISL continuity equation is solved at every iteration of the departure point algorithm to find the vertical grid implied by the advection scheme. This grid is used to interpolate the horizontal displacements in the vertical. Hereby the trajectory scheme is consistent with the explicit CISL continuity equation.

The momentum and thermodynamic equations have been discretized on grid-point form. Two different discretizations of the thermodynamic equation were considered. In the first version the thermodynamic equation is discretized as in HIRLAM (only adapted to the hybrid trajectory). In the second version the conversion term of the thermodynamic equation is consistent with the semi-implicit CISL continuity equation.

The new dynamical cores have been tested using the baroclinic wave test case recently developed by JW2005. As for shallow water model the implementation has been performed within the framework of HIRLAM and therefore only a limited-area setup has been available. It was found too comprehensive to specify the boundary fields using high-resolution reference solutions as was done in the shallow water model simulations. The boundary fields were held fixed at the initial condition throughout the integration. Consequently, it may be precarious to compare the limited area model runs with the global reference solutions provided by JW2005.

The simulations with the new models have been compared to runs with the traditional HIRLAM. As for the shallow water models the cell-integrated dynamical cores could run stably without needing decentering or filtering of the non-linear terms in time. All models, however, were run with sixth-order horizontal diffusion. The minimum/maximum surface pressure in the new dynamical core runs is lower/higher compared to HIRLAM. The location of the extrema in the surface pressure also differ. When increasing the numerical diffusion in the cell-integrated models the solution "moves toward" the HIRLAM solution both phase-wise and amplitude-wise. Therefore it is speculated that the inherently mass-conservative transport of mass, the more consistent discretizations (regarding trajectory algorithm and energy-conversion term) and the less numerical diffusion in the the cell-integrated models, make the baroclinic development more intense and alters the location of the extrema compared to HIRLAM. However, a global cell-integrated model is needed in order to make robust comparisons with simulations with the global reference solutions provided by JW2005 so that the many uncertainties caused by the limited area setup are eliminated.

Although the HIRLAM uses decentering, filtering of the non-linear terms in time and has more numerical diffusion in the continuity equation compared to the cell-integrated models, the solution was noisy in less active parts of the domain. On the contrary the cell-integrated models simulate smooth solutions in the entire domain. This important observation suggests that cell-integrated

models have less noise problems compared to traditional models which is important for a stable coupling with the physics, data-assimilation, and initialization.

5.2 Conclusions

5.2.1 *Major accomplishments*

For the first time an inherently mass-conservative semi-Lagrangian baroclinic model which uses a semi-implicit time stepping has been developed and tested. In addition to the local conservation of mass, this model can transport additional tracers consistently, i.e. there is no mass-wind inconsistency problem and mass is conserved locally also for the transport of atmospheric constituents. This adds an important property to the long list of desirable properties that semi-Lagrangian methods possess. In fact the lack of inherent mass-conservation has been the principal argument for not applying traditional semi-Lagrangian methods in many climate models. This obstacle has been overcome and could pave the road for more widespread application of semi-Lagrangian methods in climate modeling.

In addition to solving the problem of designing a practicable semi-implicit CISL scheme for the continuity equation, the equations of motion have been discretized in a more Lagrangian fashion compared to traditional semi-Lagrangian models. This increases the internal consistency of the model. The consistencies introduced are:

- The trajectories are computed consistently with the explicit CISL continuity equation.
- The energy conversion term in the thermodynamic equation is discretized in a Lagrangian fashion and consistent with the semi-implicit CISL continuity equation.
- The continuity equation has been discretized on cell-integrated form where there are no “forcing” terms on the right-hand side of the equation as there is on grid-point form. Thereby the continuity equation is discretized on an inherently mass-conservative form and in a more Lagrangian fashion compared to traditional semi-Lagrangian models.

In idealized test settings it has been demonstrated that the new dynamical core is stable for long time steps and produces forecasts which have a stronger baroclinic development compared to HIRLAM. The new dynamical cores did not need decentering and filtering of the non-linear terms in time to produce stable simulations as was the case in HIRLAM. Despite the filters and the more diffusive advection scheme used for the continuity equation, HIRLAM was noisy in less active parts of the domain whereas the cell-integrated models were smooth.

In addition, it has been demonstrated experimentally that the CISL shallow water models are not susceptible to the orographic noise problem commonly present in traditional semi-Lagrangian

models. In the baroclinic models it was also found that the cell-integrated method produces noise-free forecasts where as HIRLAM was noisy in less active parts of the domain. Hence the new dynamical cores produce smooth solutions without needing decentering or other types of filters which is important for practical applications such as a stable coupling with the physics, data-assimilation and initialization.

5.2.2 *Limitations of the work*

The most apparent limitations of this research are:

- The shallow water model could be tested using standard test cases where the boundary fields were prescribed and updated at every time step by a high resolution reference solution. For the baroclinic model it was too comprehensive to provide boundary data from a reference solution. Therefore the assessment of the accuracy of the new model was not optimal since the boundaries were held fixed at the initial condition.
- The efficiency of the new dynamical core could not be measured accurately since the implementation within the framework of HIRLAM was not optimized, but coded in an *ad-hoc* manner.
- In the limited area model simulations it is difficult to assess the conservation properties of total energy, entropy etc. because of the boundaries. Therefore the ability of the new models to conserve such quantities was not assessed. Since the new dynamical cores could not be formulated such that the vertical discretization is not a spurious source or sink for energy (see discussion in Chapter IV), it could not be diagnosed if it is a serious problem or not.

5.3 Future research directions

5.3.1 *Nonexistence of the orographic noise problem in CISL model*

During the experimental assessment of the accuracy of the CISL shallow water model it was found that the new model was not susceptible to the orographic resonance problem for high Currant flow over orography. Traditional semi-Lagrangian models do produce noise downstream of mountain ridges unless special measures are taken. A well known and widely applied method for reducing the noise is to apply decentering, which is a small offcentering in the time averaged terms (Rivest et al. 1994). As demonstrated in Chapter III, the decentering does alleviate the noise problem commonly present in traditional semi-Lagrangian models, but on the other hand it can have a quite strong negative impact on the accuracy.

Several articles analyze the orographic resonance problem of traditional semi-Lagrangian models in detail using linear analysis (e.g., Coiffier et al. 1987; Kaas 1987; Tanguay et al. 1992; Rivest et al. 1994; Lindberg and Alexeev 2000). They linearize the 1D shallow water equations on a f plane, keep space continuous, discretize in time, seek solutions and derive a condition for the orographic resonance. The result is grid independent.

It would be desirable to explain the "non-existence of the orographic resonance problem" for the CISL models with a similar analysis. As already mentioned in Chapter III, it was speculated that the resonance is averaged out by the cell-integrated method and therefore does not develop. Since space is kept continuous in the analysis published in connection with traditional semi-Lagrangian models, the proof also applies to CISL models; if space is kept continuous the CISL and traditional methods are identical. Hence a proof of the "non-existence of the orographic resonance problem" would require an analysis including the spatial discretization. This could be the scope of future work.

5.3.2 *Global CISL model*

The final step toward the long-term goal of developing a global inherently mass-conservative SISL dynamical core would be to extend the models presented in this thesis to a global domain. In Chapter III it was suggested how the limited area model could be extended to global domains. These suggestions also apply for the baroclinic model.

In a global setup several of the problems associated with the boundaries in HIRLAM would not be present. In addition, the ability of the models to conserve total energy, angular momentum and entropy are difficult to assess in a limited area setup. Therefore a global model is also needed for global benchmark tests and comparison with other dynamical cores.

Of course, dynamical cores can not be used for practical applications if they are not coupled to a physics package. Since the thermodynamic and momentum equations are discretized on grid-point form, it is possible that only small modifications are needed to existing physics packages used in traditional SISL models, when coupling the new dynamical cores to the parameterizations. The toughest test for a dynamical core is probably how well it interacts with the physics.

Other future research directions could be:

- *Alternative discretizations* In the new dynamical cores the momentum and thermodynamic equations were solved on grid-point form since the focus was on enforcing mass conservation. It would, however, be more consistent if these equations could be solved on CISL form as well instead of the hybrid approach where grid-point and cell-integrated discretizations are mixed.
- *Other prognostic variables* One could also use other prognostic variables such as total en-

ergy and entropy as suggested by Machenhauer (1992) where all model equations are cell-integrated and as a consequence the conservation of all the prognostic variables is guaranteed. It is, however, unknown if an efficient semi-implicit time-stepping scheme can be designed for such a model.

- *Discretization grid* In the new dynamical cores an Arakawa C-grid was used since such a grid is used in the HIRLAM. It would, however, be more natural to use a discretization grid in which the velocity components were held at cell vertices. The consequences of introducing such a grid were not considered in depth in this thesis, but one of the advantages of using such a grid would be that the many interpolations of the displacements from cell vertices to velocity points and mass points are avoided.

- *Non-hydrostatic extension* It was beyond the scope of this ph.d. thesis to extend the CISL ideas to non-hydrostatic models. For application of the CISL dynamical cores in meso-scale modeling and also in future climate models, a non-hydrostatic extension would be needed. It is, however, believed that global hydrostatic models still have applications for many years to come, although we are moving toward a more wide-spread use of global non-hydrostatic models.

Appendix

A List of symbols

a	radius of the earth
A_k and B_k	constants used the pressure p_k
c_{pd}	specific hear of dry air at constant pressure
c_{pv}	specific hear of water vapour at constant pressure
D	divergence calculated using the Eulerian formula
$\mathbb{D}$	divergence calculated using the Lagrangian formula
ΔA	regular arrival area
f	Coriolis parameter
g	acceleration due to gravity
G	linearized geopotential where constant terms are discarded
K_T	horizontal diffusion in thermodynamic equation
K_q	horizontal diffusion in continuity equation for water vapour
L	linear part of the right-hand side of a prognostic equation on the form $\frac{d\psi}{dt}$= ...
N	nonlinear part of the right-hand side of a prognostic equation on the form $\frac{d\psi}{dt}$= ...
n	time-level index
p	pressure
p^r	reference pressure
p_s	surface pressure
q	specific humidity
$\mathbf{P}$	contributions of parametrized physical processes in momentum equations
P_T	contributions of parametrized physical processes in thermodynamic equation
R_d	gas constant for dry air
R_v	gas constant for water vapour
T	temperature
T^{ref}	reference temperature
T_v	virtual temprerature defined by $T_v = T\ \left[1 + \left\{\frac{R_v}{R_d} - 1\right\}\ q\right]$
u	longitudinal/zonal velocity
v	latitudinal/meridional velocity
α	trajectory displacement along longitude
β	trajectory displacement along latitude
δ	$\delta = c_{p_{\text{vap}}}/c_{p_{\text{dry}}}$
η	vertical coordinate
$\dot{\eta}$	η-coordinate vertical velocity ($\dot{\eta} = d\eta/dt$)
ϕ	geopotential height
ϕ_s	surface geopotential
κ	$\kappa = R_{\text{dry}}/c_{p_{\text{dry}}}$
λ	longitude
θ	latitude

ω	vertical velocity in pressure coordinates ($\omega = dp/dt$)

B Notation

$(\cdot)_*^n$	quantity at the departure point
$(\cdot)_{*/2}^{n+1/2}$	quantity at the trajectory mid-point
$(\cdot)^{n+1}$	quantity at the arrival point
$(\cdot)_{H*}$	quantity interpolated to departure point (horizontal interpolation)
$(\cdot)_{V*}$	quantity interpolated to grid implied by advection scheme (vertical interpolation)

C List of Acronyms

ALADIN/LACE	Aire Limité Adaptation dynamique Développement InterNational /Limited Area Central European model (limited-area high resolution model used by Météo France and partners)
CAM3	NCAR Community Atmosphere Model version 3
CISL	Cell-Integrated Semi-Lagrangian
CFL	Courant-Friedrichs-Levy
CC-CISL	Conservative Cascade Cell-Integrated Semi-Lagrangian
CC-CISL-SWM	Conservative Cascade Cell-Integrated Semi-Lagrangian Shallow Water Model
CFL	Courant-Friedrichs-Levy
CTM	Chemical Transport Model
COP-CISL-HIRLAM	Consistent “omega-p” version of CISL HIRLAM
ECMWF	European Centre for Medium-Range Weather Forecasts
EUL-CAM3	NCAR spectral transform Eulerian dynamical core of the Community Atmosphere Model CAM3
FLOPS	FLoating-point Operations Per Second
GCMs	General Circulation Models
GME	icosahedral finite-difference model of the German Weather Service
HIRLAM	High Resolution Limited Area Model
FCT	Flux-Corrected Transport
FV	Finite-Volume
IFS	Integrated Forecast System (operational forecast model at ECMWF)
IMC	Inherently Mass-Conservative
LAM	Limited Area Model
M	Monotone
NASA/NCAR-FV	NASA/NCAR Finite-Volume dynamical core
NCAR	National Center of Atmospheric Research
P	Positive Definite
PB-CISL	Pseudo-Biparabolic Cell-Integrated Semi-Lagrangian
PB-CISL-SWM	Pseudo-Biparabolic Cell-Integrated Semi-Lagrangian Shallow Water Model
PGF	Pressure gradient force
PPM	Piecewise-Parabolic Method
PPM1	PPM with sub-grid scale representation of Laprise and Plante (1995)
PPM2	PPM with sub-grid scale representation of Colella and Woodward (1984)
QMSL	Quasi-Monotonic Semi-Lagrangian
R-CISL-HIRLAM	Reference version of CISL HIRLAM
SETTLS	Stable Extrapolation Two-Time-Level Scheme
SISL	Semi-Implicit Semi-Lagrangian
SL-CAM3	NCAR spectral transform semi-Lagrangian dynamical core of the Community Atmosphere Model CAM3
SLT	Semi-Lagrangian Transport

SLICE	Semi-Lagrangian Inherently Conserving and Efficient
SM	Semi-Monotone
SW-HIRLAM	Shallow Water version of the HIRLAM
TRAD	Traditional Semi-Lagrangian
2TLSL	2 Time-Level Semi-Lagrangian

D Code documentation for the shallow water models

A shallow water version of HIRLAM[9] has been coded i.e. the source code of the baroclinic forecast model has been changed to perform shallow water model runs. The implementation has been done such that minimal changes have been applied to the original code. All changes to the source code are marked with 'PHL'. Original source code not used has been out-commented with 'CPHL' and lines added to the code has been marked with '!ADD PHL'. Only few comments are present in the original source code. Many extra comments have been added to clarify the code. Each line with a comment added starts with 'CPHL'. Thereby it should be very easy to identify the changes to the original code.

The documentation is not exhaustive and the reader is referred to the code for further explanation. The equations provided here follow the code closely and it should be straight forward to identify the different formulas in the code by reading this documentation and the comments in the code.

The traditional semi-Lagrangian model

The prognostic equations can be written on the form

$$\frac{d\psi}{dt} = L_\psi + N_\psi, \tag{D-1}$$

where L and N are the linear and non-linear part of the right-hand side , respectively, and $\psi = u, v, \Phi$ (momentum equations and continuity equation, respectively). In spherical coordinates the linear and non-linear terms are given by

$$L_u = f_0\, v - \frac{1}{a\,\cos\theta}\,\frac{\partial}{\partial\lambda}(\Phi + \Phi_s), \tag{D-2}$$

$$L_v = -f_0\, u - \frac{1}{a}\,\frac{\partial}{\partial\theta}(\Phi + \Phi_s), \tag{D-3}$$

$$L_\Phi = -\Phi_{00}\, D, \tag{D-4}$$

$$N_u = f'\, v + \frac{u\, v}{a}\tan\theta, \tag{D-5}$$

$$N_v = -f'\, u - \frac{u^2}{a}\tan\theta, \tag{D-6}$$

$$N_\Phi = -\left(\Phi - \Phi_{00}\right)\, D, \tag{D-7}$$

where the "half-implicit Coriolis scheme" (McDonald and Haugen 1992) has been used, i.e. the Coriolis parameter is split into a constant part and the deviation from the constant, $f = f_0 + f'$. For discretization details see Undén (2002).

[9] the model is documented in Undén (2002)

(D-1) is discretized using the *centered Lagrangian explicit (2)* approximation of McDonald and Haugen (1992)

$$\frac{\psi^{n+1}-\psi^{n}_{*}}{\Delta t} = \tfrac{1}{2}\left[L^{n+1}_{\psi} + (L_{\psi})^{n}_{*}\right] + \tfrac{1}{2}\left[N^{n+1/2}_{\psi} + (N_{\psi})^{n+1/2}_{*}\right], \tag{D-8}$$

where the subscript $*$ refers to evaluation at the departure point. Collecting unknown terms on the left-hand side and known terms on the right-hand side (D-8) becomes

$$\psi^{n+1} - \tfrac{1}{2}\Delta t_{+}\, L^{n+1}_{\psi} = A_{\psi}, \tag{D-9}$$

where A_{ψ} is given by

$$\tfrac{1}{2}\Delta t_{+}\, N^{n+1/2}_{\psi} + \left[\tfrac{1}{2}\Delta t_{-}\, L^{n}_{\psi} + \tfrac{1}{2}\Delta t_{-}\, N^{n+1/2}_{\psi} + \psi^{n}\right]_{*}. \tag{D-10}$$

To control noise and help reduce the problem of orographic resonance first order spatial decentering is applied in (D-9) (Semazzi et al. (1995) and references herein). This is done by redefining Δt such that terms evaluated at the departure point and arrival point are decentered

$$\Delta t_{+} = (1+\epsilon_{g}), \tag{D-11}$$

$$\Delta t_{-} = (1-\epsilon_{g}), \tag{D-12}$$

$$\tag{D-13}$$

where ϵ is the decentering parameter. To enhance stability an additional filter is applied to the non-linear terms (see p.13 in Undén (2002) for details).

The prognostic variables are extrapolated to time-level $n+1/2$ in subroutine `COMPFX` where after the subroutine `SLDYN.f` calculates all the non-linear terms $N^{n+1/2}_{\psi}$ for $\psi = u, v, \Phi$. Then the subroutine `SLDYNM.f` interpolates the expression in the square brackets in (D-10) to the departure points and adds the remaining nonlinear term in (D-10) to complete the computation of A_{ψ}.

In order to derive the elliptic equation substitute u^{n+1} (equation (D-9) with $\psi = u$) into the equation for v^{n+1} and vice versa. The result is

$$\left[u + \frac{\Delta t}{2}\left(\frac{b}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{e}{a}\frac{\partial}{\partial\theta}\right)(\Phi+\Phi_{s})\right]^{n+1} = b\,A_{u} + e\,A_{v}, \tag{D-14}$$

$$\left[v + \frac{\Delta t}{2}\left(-\frac{e}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{b}{a}\frac{\partial}{\partial\theta}\right)(\Phi+\Phi_{s})\right]^{n+1} = -e\,A_{u} + b\,A_{v}, \tag{D-15}$$

where $b = 1/(1+F^{2}_{+})$, $e = F_{+}\,b$ and $F_{+} = \overline{f}\,\Delta t/2$. Finally substitute the equation for Φ^{n+1} into (D-14) and (D-15)

$$\left[u - \left(\frac{\Delta t}{2}\right)^{2}\overline{\Phi}\left(\frac{b}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{e}{a}\frac{\partial}{\partial\theta}\right) D\right]^{n+1} = Z_{u} \tag{D-16}$$

$$\left[v - \left(\frac{\Delta t}{2}\right)^{2}\overline{\Phi}\left(-\frac{e}{a\,\cos\theta}\frac{\partial}{\partial\lambda} + \frac{b}{a}\frac{\partial}{\partial\theta}\right) D\right]^{n+1} = Z_{v}, \tag{D-17}$$

where

$$Z_u = b\,(Z_1 + F_+\,Z_2) \tag{D-18}$$

$$Z_v = b\,(Z_2 - F_+\,Z_1) \tag{D-19}$$

and

$$Z_1 = A_u - \left(\frac{\Delta t}{2}\right)\,\frac{1}{a\,\cos\theta}\,\frac{\partial}{\partial\lambda}\,(A_\Phi + \Phi_s)\,, \tag{D-20}$$

$$Z_2 = A_v - \left(\frac{\Delta t}{2}\right)\,\frac{1}{a}\,\frac{\partial}{\partial\theta}\,(A_\Phi + \Phi_s)\,. \tag{D-21}$$

The elliptic equation results when taking the divergence of (D-16) and (D-17). The cross-derivative terms cancel out and an Helmholtz equation results

$$\left[1 - \overline{\Phi}\,\left(\frac{\Delta t}{2}\right)^2\,\nabla_H^2\right]\,D^{n+1} = \nabla_H\cdot\mathbf{Z} \tag{D-22}$$

where $\mathbf{Z} = (Z_u, Z_v)$, the divergence of $\mathbf{Z}$ is given by

$$\nabla_H\cdot\mathbf{Z} = \frac{1}{a\,\cos\theta}\left[\frac{\partial Z_u}{\partial\lambda} + \frac{\partial}{\partial\theta}\,(\cos\theta\,Z_v)\right], \tag{D-23}$$

and the Laplacian operator is given by

$$\nabla_H^2 = \frac{1}{a\,\cos\theta}\left[\frac{\partial}{\partial\lambda}\,\frac{1}{a\,\cos\theta}\,\frac{\partial}{\partial\lambda} + \frac{\partial}{\partial\theta}\,\frac{\cos\theta}{a}\,\frac{\partial}{\partial\theta}\right]. \tag{D-24}$$

The right-hand side of the Helmholtz equation is prepared in subroutine `SLEXPA.f` and the Helmholtz equation is solved in `HHSOLV`.

Once the Helmholtz equation is solved use the continuity equation to find the new geopotential height, and use (D-16) and (D-17) to find the velocity components. For efficiency note that

$$u^{n+1} = b\,\left(Z_1 - \frac{\Delta t}{2}\,\frac{1}{a\,\cos\theta}\,\frac{\partial Q}{\partial\lambda}\right) + e\,\left(Z_2 - \frac{\Delta t}{2}\,\frac{1}{a}\,\frac{\partial Q}{\partial\theta}\right) \tag{D-25}$$

$$v^{n+1} = -e\,\left(Z_1 - \frac{\Delta t}{2}\,\frac{1}{a\,\cos\theta}\,\frac{\partial Q}{\partial\lambda}\right) + b\,\left(Z_2 - \frac{\Delta t}{2}\,\frac{1}{a}\,\frac{\partial Q}{\partial\theta}\right), \tag{D-26}$$

where

$$Q = -\frac{\Delta t}{2}\,\overline{\Phi}\,D^{n+1}. \tag{D-27}$$

For efficiency the symmetry in (D-25) and (D-26) is explored in the code.

The CISL semi-Lagrangian model

A pseudo code for how the cell-integrated semi-implicit continuity equation is solved is given in Chapter III. By replacing A_Φ in (D-9) with the expression for the CISL scheme, the algorithm proceeds as for the traditional model.

Performing the upstream integral

The upstream integral is performed in `NMREMAP.f` and `CCSREMAPNS.f`/`CCSREMAPEW.f` for the fully two-dimensional CISL and the cascade CISL scheme, respectively. The location of the cell walls are computed in `CIBIXT.f` and the integral subroutines are called from here. In the cascade scheme the intermediate grid is computed using cubic or linear interpolation in subroutine `CCSDEPARTUREGRID` and the upstream integral is decomposed into two one-dimensional remappings performed first in the north-south direction (`CCSREMAPNS.f`) and then along the Lagrangian latitude (`CCSREMAPEW.f`). The one-dimensional remappings are performed as in the CISL scheme, which is described next.

The upstream integral in the CISL scheme is optimized by accumulating the coefficients for the parabolas along latitudes (`PREPREMAP.f`). By integrating the accumulated coefficients up to some Eulerian longitude, we get the mass from the reference line up to the Eulerian longitude enclosed between two adjacent Eulerian latitudes. This area is referred to as a stripe. The actual integral over each departure cell can be written as a sum of the accumulated integrals surrounding the cell scaled by their meridional extension (see Nair and Machenhauer 2002 for details). The integral over each stripe is computed in `NMACCMASSPPM.f`. See `NMREMAP.f` for additional details.

Helmholtz solver

The Helmholtz solver is the standard Helmholtz solver used in HIRLAM. The solver applies a Fourier transform in the λ-direction (Källén 1996, p. 2.14). The divergence is set to zero on the boundary.

Consider the general Helmholtz equation

$$\left(1 - b_0 \nabla^2\right) D = Q, \tag{D-28}$$

where b is given by

$$b_0 = \overline{\Phi} \left(\frac{\Delta t}{2}\right)^2, \tag{D-29}$$

The finite difference formula for the Laplacian (D-24) is

$$\nabla^2_{i,j} = \frac{1}{a^2}\left\{\left(1\frac{1}{\cos^2\theta_j}\,\delta^2_\lambda - \frac{\cos\theta_{j-1/2}+\cos\theta_{j+1/2}}{\cos\theta_j\ (\Delta\theta)^2}\right) + \frac{\cos\theta_{j-1/2}}{\cos\theta_j\ (\Delta\theta)^2} + \frac{\cos\theta_{j+1/2}}{\cos\theta_j\ (\Delta\theta)^2}\right\}, \tag{D-30}$$

where $\theta_j = j\Delta\theta$, $\theta_{j\pm1/2} = (j \pm 1/2)\,\Delta\theta$ and δ_λ is a zonal finite difference operator defined by

$$\delta_\lambda \phi_i j = \frac{\phi_{i+1,j} - \phi_{i-1,j}}{\Delta\lambda}, \qquad \delta^2_\lambda \phi_i j = \frac{\phi_{i+1,j} - 2\,\phi_{i,j} - \phi_{i-1,j}}{(\Delta\lambda)^2}. \tag{D-31}$$

Now apply a Fourier sine transformation in the λ-direction

$$D(\lambda, \theta = \theta_j) = \sum_\ell D_\ell\, e^{i\ell\lambda}, \tag{D-32}$$

where ℓ is the wave number and D_ℓ is the amplitude of the ℓth wave. $\delta^2_\lambda D$ is given by

$$\delta^2_\lambda D = \sum_{ell} D_\ell \frac{1}{(\Delta\lambda)^2}\left[e^{\imath\,\ell\,(\lambda+\Delta\lambda)} - 2\,e^{\imath\,\ell\,\lambda)} + e^{\imath\,\ell\,(\lambda-\Delta\lambda)}\right] \tag{D-33}$$

$$= -\left(\frac{2}{\Delta\lambda}\right)^2 \sum_\ell D_\ell\, \sin^2\left(\frac{\ell\,\Delta\lambda}{2}\right). \tag{D-34}$$

For each wave number ℓ the Fourier transformed Helmholz equation becomes

$$D_j + b_0\,\frac{1}{a^2}\left\{\left[\frac{1}{\cos^2\theta_j}\left(\frac{2}{\lambda}\right)^2 \sin^2\left(\frac{\ell\,\Delta\lambda}{2}\right) + \frac{\cos\theta_{j-1/2}+\cos\theta_{j+1/2}}{\cos\theta_j\ (\Delta\theta)^2}\right] D_j \right.$$
$$\left. - \frac{\cos\theta_{j-1/2}}{\cos\theta_j\ (\Delta\theta)^2}\,D_{j-1} - \frac{\cos\theta_{j+1/2}}{\cos\theta_j\ (\Delta\theta)^2}\,D_{j+1}\right\} = Q_j. \tag{D-35}$$

After rearranging terms (D-35) can be written as

$$A_j\,D_{j-1} + B_j\,D_j + C_j\,D_{j+1} = F_j\,Q_j, \qquad j = 1, \text{KLAT} \tag{D-36}$$

where

$$A_j = -\cos\theta_{j-1/2}, \tag{D-37}$$

$$B_j = E_j\,\cos\theta_j + \frac{\left[\frac{2}{\Delta\lambda}\,\sin\left(\frac{\ell\,\Delta\lambda}{2}\right)\,\Delta\theta\right]^2}{\cos\theta_j} + \cos\theta_{j-1/2} + \cos\theta_{j+1/2}, \tag{D-38}$$

$$C_j = -\cos\theta_{j+1/2}, \tag{D-39}$$

$$F_j = E_j\,\cos\theta_j, \tag{D-40}$$

$$E_j = \frac{a^2\,(\Delta\theta)^2}{b_0}, \tag{D-41}$$

and $KLAT$ is the number of latitudes. The system of equations (D-36) is a triangular system which can be solved efficiently using tridiagonal solvers (e.g, p.440-441 Durran 1999).

Departure point algorithms

The departure points are calculated by the subroutine CALPQR. The trajectory algorithm used in HIRLAM is a modified version of the usual iterative algorithm of Robert (1981) described in McDonald (1999). To accommodate the CISL scheme and in order to investigate the sensitivity of the semi-Lagrangian schemes on the accuracy of the trajectories, two other departure point algorithms have been implemented: the 'Stable Extrapolation Two-Time-Level Scheme' or SETTLS scheme by Hortal (2002) and the trajectory method described in this thesis (Chapter III).

E Area of a spherical polygon

The cell area in spherical geometry can be defined by connecting the four cell corner points, located at longitude λ_i and latitude θ_i ($i = 1, 4$), by great circle arcs. Then the cell takes the form of a spherical polygon which area is given by[10]

$$\text{Area} = \left(\sum_{i=1}^{4} \Theta_i - 2\,\pi \right)\, a^2, \tag{E-1}$$

where Θ_i are the interior angles of the polygon and also denote the vertices of the cell (see Fig.43). For computation of e.g. the interior angle Θ_1, consider the spherical triangle $\Theta_1\Theta_2\Theta_4$ as shown on Fig. 43. For simplicity it is assumed that Θ_4 coincides with the Earth's pole, so that the arc lengths of $\Theta_4\Theta_1$ and $\Theta_4\Theta_2$ are simply given by $\pi - \theta_1$ and $\pi - \theta_2$, respectively. The spherical cosine equation can be used to calculate Θ_4:

$$\cos(\Theta_4) = \frac{\cos(\lambda_2 - \lambda_1) - \cos(\pi - \theta_2)\,\cos(\pi - \theta_1)}{\sin(\pi - \theta_2)\,\sin(\pi - \theta_1)}. \tag{E-2}$$

Equation (E-2) can be substituted in the fourth term of the sum on the right-hand side of (E-1). Similarly for Θ_1, Θ_2 and Θ_3.

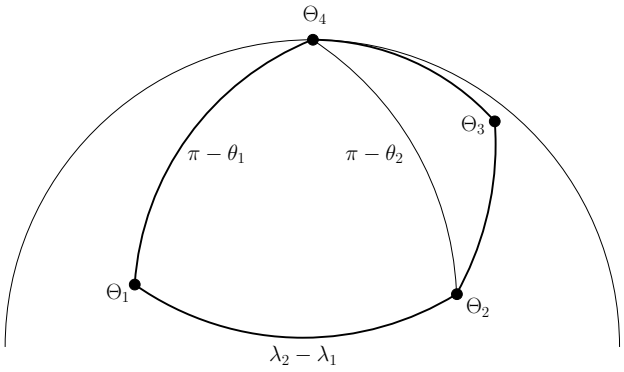


Fig. 43: The spherical polygon defining the cell in the special case where one of the vertices coincide with the Earth's north pole. Θ_i denote the cell corner points and interior angles of the vertices which are located at longitude λ_i and latitude θ_i. Along selected edges the arc lengths in radians are indicated.

[10] if the location of the cell corner points are known explicitly, the area of the spherical polygon can be computed effi ciently and accurately using the algorithm of Miller (1994)

F Definition of matrix operators

For the formulation of the baroclinic system it is convenient to introduce matrices $[\nu]$, $\left[\tau^A\right]$, $\left[\tau^B\right]$, $[\gamma]$, and $\left[M^A\right]$ and $\left[M^B\right]$ such that

$$[\nu]\,\mathbf{D} = \sum_{\ell=1}^{NLEV} \Delta p_\ell^{ref}\, D_\ell^{n+1}, \tag{F-1}$$

$$\left(\left[\tau^A\right]\mathbf{X}\right)_k = \frac{R_d\, T^{ref}}{c_{pd}} \left[\left(\frac{\Delta \ln p}{\Delta p}\right)_k^{ref} \sum_{\ell=1}^{k} \Delta p_\ell^{ref}\, X_\ell + \alpha_k^{ref}\, X_k\right] \qquad \mathbf{X} = \mathbf{D}, \mathbf{A}_p, \tag{F-2}$$

$$\left(\left[\tau^B\right]\mathbf{X}\right)_k = \frac{R_d\, T^{ref}}{c_{pd}} \left(\frac{1}{p_{k+1/2}}\right)^{ref} \sum_{\ell=1}^{k} \Delta p_\ell^{ref}\, X_\ell \qquad \mathbf{X} = \mathbf{D}, \mathbf{A}_p, \tag{F-3}$$

$$\left([\gamma]\,\mathbf{T}\right)_k = R_d \sum_{\ell=k+1}^{NLEV} \left(T\,\Delta \ln p^{ref}\right)_\ell + R_d \left(\alpha^{ref}\, T\right)_k, \tag{F-4}$$

$$\left(\left[M^X\right]\mathbf{D}\right)_k = \left(\left[[\gamma]\left[\tau^X\right] + R_d \left(\frac{T_k}{p_s}\right)^{ref} [\nu]\right]\mathbf{D}^{n+1}\right)_k \qquad X = A, B, \tag{F-5}$$

respectively. The reference 'half-level' pressures are obtained from (92) by chosing $p_s = p_s^{ref}$ and the coefficients α_k^{ref} are based on these reference values.

G Vertical η coordinate

The hybrid coordinate introduced by Simmons and Burridge (1981) are defined such that at the interfaces ('half-levels') the pressure is given by

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2}\, p_s, \tag{G-1}$$

where $A_{k+1/2}$ and $B_{k+1/2}$ are prescribed coefficients listed in table 9. The reference pressure p_s^{ref} is set to 10^4 Pa. The 'half-level' η-coordinate is defined by

$$\eta_{k+1/2} = \frac{A_{k+1/2}}{p_s} + B_{k+1/2}. \tag{G-2}$$

The 'full-level' η value is determined via the average $\eta_k = \frac{1}{2}(\eta_{k+1/2} + \eta_{k-1/2})$.

Index k	$A_{k+1/2}$	$B_{k+1/2}$	Index k	$A_{k+1/2}$	$B_{k+1/2}$
1	0.0000	0.00000000	15	6728.5740	0.14086370
2	219.4067	0.00000000	16	6410.5090	0.18077200
3	489.5209	0.00000000	17	6036.3220	0.22772200
4	988.2418	0.00000000	18	5596.1110	0.28295620
5	1805.2010	0.00000000	19	5078.2250	0.34793640
6	2983.7240	0.00000000	20	4468.9600	0.42438220
7	4462.3340	0.00000000	21	3752.1910	0.51431680
8	6160.5870	0.00000000	22	2908.9490	0.62012020
9	7851.2430	0.00000000	23	2084.7390	0.72353550
10	7731.27102	0.01505309	24	1334.4430	0.81767680
11	7590.1310	0.03276228	25	708.4990	0.89621530
12	7424.0860	0.05359622	26	252.1360	0.95347610
13	7228.7440	0.07810627	27	0.0000	0.98511220
14	6998.9330	0.10694110	28	0.0000	1.00000000

Table 9: Vertical coefficients $A_{k+1/2}$ and $B_{k+1/2}$ for the 27 level model. The coefficients are as the 26-layer model versions in Jablonowski and Williamson (2005, manuscript submitted to *Mon. Wea. Rev.*), but with one level added at the top of the atmosphere such that the upper boundary is at zero pressure.

H Jablonowski-Williamson baroclinic test case

Jablonowski and Williamson (2005, manuscript submitted to *Mon. Wea. Rev.*) have recently developed two test cases for dynamical cores. The initial conditions for the Jablonowski-Williamson test cases are given here.

Stationary steady state test case

The velocity field is given by

$$u(\lambda, \theta, \eta) = u_0 \, \cos^{3/2} \, \eta_v \, \sin^2(2\theta), \tag{H-1}$$

$$v(\lambda, \theta, \eta) = 0, \tag{H-2}$$

where the auxiliary variable η_v is given by

$$\eta_v = (\eta - \eta_0) \, \frac{\pi}{2}, \tag{H-3}$$

with $\eta_0 = 0.2$ and $u_0 = 35$ ms^{-1}. The flow field is comprised of two symmetric jets in the midlatitudes (see Fig. 44)

The horizontally averaged temperature is given by

$$\langle T(\eta) \rangle = T_0 \, \eta^{R_d \, \Gamma} g \quad , \quad (\text{for } \eta_s \geq \eta \geq \eta_t), \tag{H-4}$$

$$\langle T(\eta) \rangle = T_0 \, \eta^{R_d \, \Gamma} g + \Delta T \, (\eta_t - \eta)^5 \quad , \quad (\text{for } \eta_t > \eta), \tag{H-5}$$

where $\eta_t = 0.2$ is the tropopause level, $T_0 = 288$ K is the horizontal mean temperature at the surface, $\Gamma = 0.005$ K m^{-1} is the lapse rate, $\Delta t = 4.8 \times 10^5$ K, R_d the ideal gas constant and g the gravitational acceleration. The total temperature variation is the sum of the horizontal mean temperature and a horizontal variation at each level:

$$T(\lambda, \theta, \eta) = \langle T(\eta) \rangle + \frac{3}{4} \, \frac{\eta \, \pi \, u_0}{R_d} \, \sin(\eta_v) \, \cos^{\frac{1}{2}}(\eta_v) \, A_c \, 2 \, u_0 \, \cos^{\frac{3}{2}}(\eta_v) + a \, \Omega B_c, \tag{H-6}$$

where Ω is the Earth's angular velocity, a the radius of the earth, and

$$A_c = -2 \, \sin^6(\theta) \, \left[\cos^2(\theta) + \frac{1}{3} \right] + \frac{10}{63}, \tag{H-7}$$

$$B_c = \frac{8}{5} \, \cos^3(\theta) \, \left[\sin^2(\theta) + \frac{2}{3} \right] - \frac{\pi}{4}. \tag{H-8}$$

The temperature field is shown on Fig. 45.

The surface geopotential Φ_s balances the non-zero wind field at the surface and is determined by

$$\Phi_s(\lambda, \theta) = u_0 \, \cos^{\frac{3}{2}} \left[(\eta_s - \eta_0) \, \frac{\pi}{2} \right] \, A_c \, u_0 \, \cos^{\frac{3}{2}} \left[(\eta_s - \eta_0) \, \frac{\pi}{2} \right] + a \, \Omega \, B_c \tag{H-9}$$

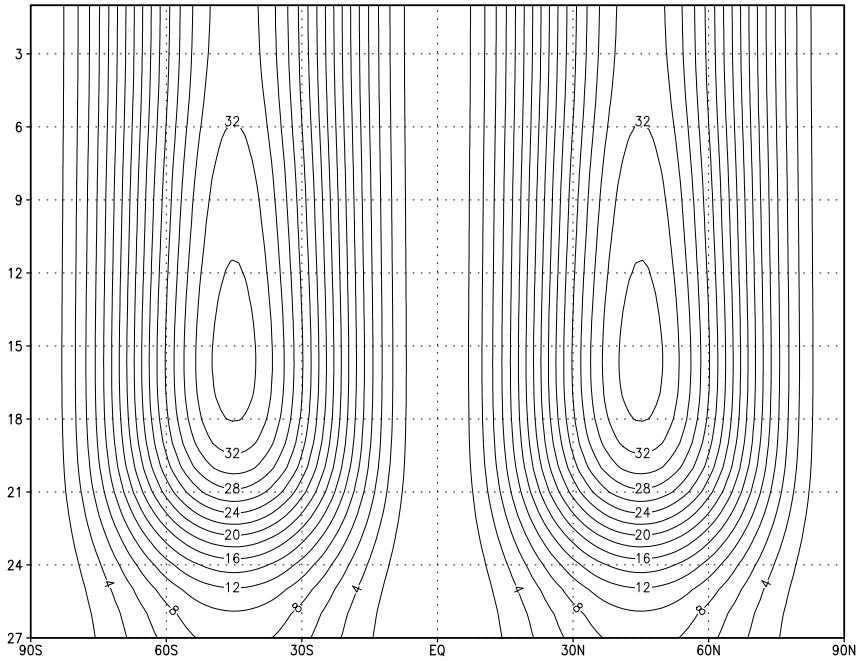


Fig. 44: Cross-section of the velocity field ($m\,s^{-1}$) plotted as a function of latitude and full levels (level one is the top of the atmosphere and level 27 the lowest level).

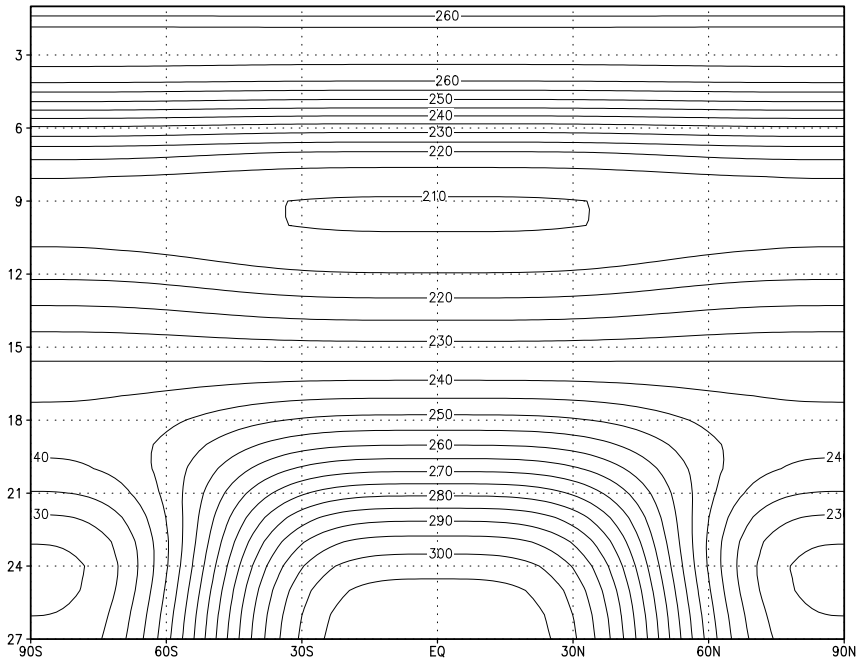


Fig. 45: Same as Fig. 44 but for the temperature field (K).

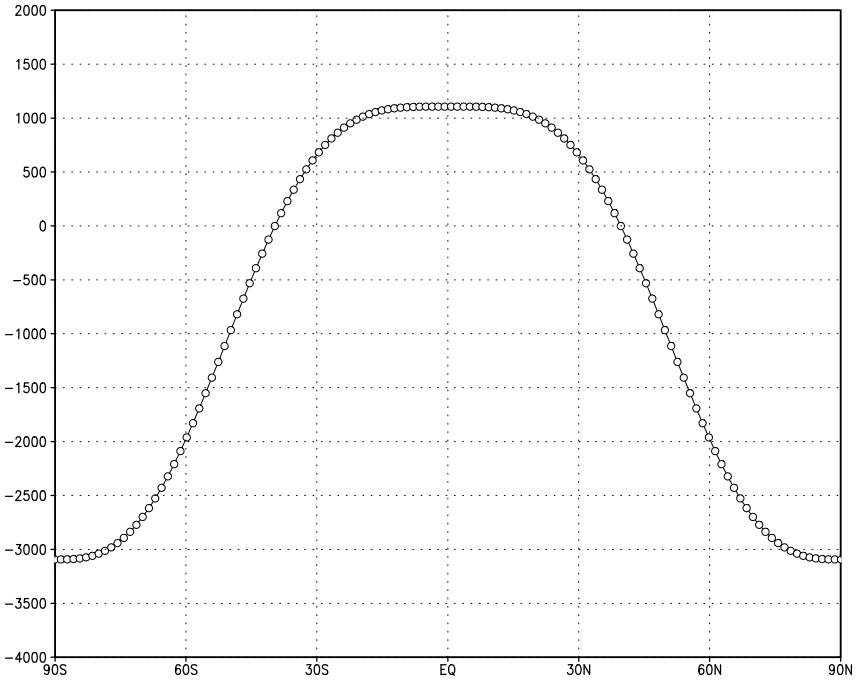


Fig. 46: The surface geopotential (m^2s^{-2}) along a longitude.

(see Fig. 46).

These analytic expressions are a solution to the adiabatic and frictionless primitive equations. A perfect model should preserve the initial conditions throughout the integration.

Baroclinic wave

To trigger a dynamically more interesting test case a perturbation in the velocity field is overlayed the steady-state initial conditions. The perturbation is placed at $(\lambda_c, \theta_c) = (\frac{2\pi}{9}, \frac{\pi}{9})$ and is given by

$$u'(\lambda, \theta, \eta) = u_p \, \exp\left[-\left(\frac{r}{R}\right)^2\right], \tag{H-10}$$

with radius $R = a/10$ and a maximum amplitude of $u_p = 1 \text{ m s}^{-1}$, and where r is the great circle distance given by

$$r = a \, \arccos\left[\sin(\theta_c) \, \sin(\theta) + \cos(\theta_c) \, \cos(\theta) \, \cos(\lambda - \lambda_c)\right]. \tag{H-11}$$

I Implicit horizontal diffusion

The zth order implicit horizontal diffusion is applied to u, v, and T as in HIRLAM. Following Undén (2002) the diffusion is applied at the end of each time-step[11] by solving the diffusion equation

$$\frac{\partial \psi}{\partial t} - K\left(\frac{\partial^z \psi}{\partial x} + \frac{\partial^z \psi}{\partial y}\right) = 0, \qquad \psi = u, v, T, \tag{I-1}$$

using a two-step split implicit approximation:

$$\frac{\psi^* - \psi^n}{\Delta t} - K\frac{\partial^z \psi^*}{\partial x} = 0, \tag{I-2}$$

$$\frac{\psi^{n+1} - \psi^*}{\Delta t} - K\frac{\partial^z \psi^{n+1}}{\partial x} = 0, \tag{I-3}$$

where $\partial\psi^z/\partial x$ and $\partial\psi^z/\partial y$ are computed using finite differences. For $z = 6$ that is

$$\frac{\partial^6 \psi}{\partial x^6} = \frac{1}{\Delta x^6}\left(\psi_{i-3,j} - 6\,\psi_{i-2,j} + 15\,\psi_{i-1,j} - 20\,\psi_{i,j} - 6\,\psi_{i+2,j} + \psi_{i+3,j}\right), \tag{I-4}$$

$$\frac{\partial^6 \psi}{\partial y^6} = \frac{1}{\Delta y^6}\left(\psi_{i,j-3} - 6\,\psi_{i,j-2} + 15\,\psi_{i,j-1} - 20\,\psi_{i,j} - 6\,\psi_{i,j+2} + \psi_{i,j+3}\right), \tag{I-5}$$

where $\Delta x = a\,\cos(\theta)\,\Delta\lambda$ and $\Delta y = a\,\Delta\theta$.

The diffusion coefficient is scaled such that the $2\,\Delta x$-wave has the same e-folding time regardless of resolution (McDonald 1998). In the 5 upper most levels the diffusion coefficient is increased. Levels with index higher than 6 use the same horizontal diffusion coefficient as level 6 uses.

$\Delta\lambda = 1.45°$		$\Delta\lambda = 0.74°$	
level index	k	level index	k
1	0.238E28	1	0.443E26
3	0.119E28	2	0.222E26
4	0.595E27	3	0.111E26
5	0.298E27	4	0.554E25
6	0.149E27	5	0.277E25

Table 10: ∇^6-diffusion coefficients for the 6 upper most layers. Below level 6 the diffusion coefficients are as in layer 6.

[11] after the boundary relaxation

References

Arakawa, A. and V. Lamb, 1977: Computational design and the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics*, **17**, 173–265.

Arakawa, A. and M. J. Suarez, 1983: Vertical differencing of the primitive equations in sigma coordinates. *Mon. Wea. Rev.*, **111**, 34–45.

Bates, J. and A. McDonald, 1982: Multiply-upstream, semi-Lagrangian advective schemes: Analysis and application to a multi-level primitive equation model. *Mon. Wea. Rev.*, **110**, 1831–1842.

Bermejo, R., 1990: Notes and correpondence on the equivalence of semi-Lagrangian schemes and particle-in-cell finite element methods. *Mon. Wea. Rev.*, **118**, 979–987.

Bermejo, R. and J. Conde, 2002: A conservative quasi-monotone semi-Lagrangian scheme. *Mon. Wea. Rev.*, **130**, 423–430.

Bermejo, R. and A. Staniforth, 1992: The conversion of semi-Lagrangian advection schemes to quasi-monotone schemes. *Mon. Wea. Rev.*, **120**, 2622–2632.

Boer, G. J. and B. Denis, 1997: Numerical convergence of the dynamics of a GCM. *Cliamte Dyn.*, **13**, 359–374.

Byun, D. W., 1999: Dynamically consistent formulations in meteorological and air quality models for multiscale atmospheric studies. Part II: Mass conservation issues. *J. Atmos. Sci.*, 3808–3820.

Cameron-Smith, P., P. Connell, and M. Prather, 2002: Pressure fixers: Conserving mass in chemical transport models by adjusting the winds, http://asd.llnl.gov/pfix/.

Carpenter, R., K. Droegemeier, P. Woodward, and C. Hane, 1990: Application of the piecewise parabolic method (ppm) to meteorological modeling. *Mon. Wea. Rev.*, **118**, 586–612.

Coiffier, J., P. Chapelet, and N. Marie, 1987: Study of various quasi-Lagrangian techniques for numerical models. *Proc. ECMWF Workshop on Techniques for Horizontal Discretization in Numerical Weather Prediction Models*, Shinfield Park, Reading, United Kingdom, European Centre for Medium-Range Weather Forecasts, 19–46.

Colella, P. and P. R. Woodward, 1984: The piecewise parabolic method (PPM) for gas-dynamical simulations. *J. Comput. Phys.*, **54**, 174–201.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, J. T. Kiehl, and B. Briegleb, 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Tech. Note, NCAR/TN-464+STR.

Cordero, E., N. Wood, and A. Staniforth, 2005: Impact of semi-Lagrangian trajectories on the discrete normal modes of a non-hydrostatic vertical-column model. *Q. J. R. Meteorol. Soc.*, **131**, 93–108.

Davies, H., 1976: A lateral boundary formulation for multi-level prediction models. *Q. J. R. Meteorol. Soc.*, **102**, 405–418.

Durran, D. R., 1999: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics*. Texts in Applied Mathematics, Vol. 32, Springer-Verlag.

Durran, D. R. and P. A. Reinecke, 2004: Instability in a class of explicit two-time-level semi-Lagrangian schemes. *Q. J. R. Meteorol. Soc.*, **130**, 365–369.

Eluszkiewicz, J., R. S. Hemler, J. D. Mahlman, L. Bruhwiler, and L. L. Takacs, 2000: Sensitivity of age-of-air calculations to the choice of advection scheme. *J. Atmos. Sci.*, **57**, 3185–3201.

Giraldo, F. X., 1999: Trajectory calculations for spherical geodesic grids in cartesian space. *Mon. Wea. Rev.*, **127**, 1651–1662.

Gospodinov, I., V. Spiridonov, and J. Geleyn, 2001: Second-order accuracy of two-time-level semi-Lagrangian schemes. *Q. J. R. Meteorol. Soc.*, **127**, 1017–1033.

Gravel, S. and A. Staniforth, 1994: A mass-conserving semi-lagrangian scheme for the shallow-water equations. *Mon. Wea. Rev.*, **122**, 243–248.

Held, I. and M. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteorol. Soc.*, **73**, 1825–1830.

Holnicki, P., 1995: A shape-preserving interpolation: Applications to semi-Lagrangian advection. *Mon. Wea. Rev.*, **123**, 862–870.

Horowitz, L. W., S. Walters, D. L. Mauzerall, L. K. Emmons, P. J. Rasch, C. Granier, X. Tie, J.-F. cois Lamarque, M. G. Schultz, G. S. Tyndall, J. J. Orlando, and G. P. Brasseur, 2004: A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2. *J. Geophys. Res.*, **108**, 4784, doi:10.1029/2002JD002853.

Hortal, M., 1998: Some recent advances at ecmwf. *LAM News*, **27**, 32–36.

— 2002: The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ecmwf forecast model. *Q. J. R. Meteorol. Soc.*, **128**, 1671–1687.

Hortal, M., 2004: Instabilities in the semi-lagrangian scheme in the stratosphere. *The 2004 Workshop on the solution of partial differential equations on the sphere*.

Hoskins, B. J., 1973: Stability of the Rossby-Haurwitz wave. *Q. J. R. Meteorol. Soc.*, **99**, 723–745.

Jablonowski, C. and D. L. Williamson, 2005: A baroclinic instability test case for atmospheric model dynamical cores. *Mon.Wea.Rev.*, submitted.

Jakob, R., J. Hack, and D. Williamson, 1993: Reference solutions to shallow water test set using the spectral transform method. Technical Report TN-388+STR, NCAR.

Jöckel, P., R. von Kuhlmann, M. G. Lawrence, B. Steil, C. Brenninkmeijer, P. J. Crutzen, P. J. Rasch, and B. Eaton, 2001: On a fundamental problem in implementing flux-form advection schemes for tracer transport in 3-dimensional general circulation and chemistry transport models. *Q.J.R.Meteorol.Soc.*, **127**, 1035–1052.

Kaas, E., 1987: *The construction of and tests with a multi-level, semi-Lagrangian and semi-implicit limited area model*. Master's thesis, Geophysics Institute, Copenhagen University, Geophysics Institute, University of Copenhagen, Juliane Maries vej 30, DK-2100 Copenhagen O, Denmark.

Kaas, E., P. H. Lauritzen, and B. Machenhauer, 2005: A shallow water model based on an accurate cell-integrated semi-Lagrangian and semi-implicit scheme, Danish Climate Center Report 05-02, Danish Meteorological Institute, Copenhagen, Denmark.

Kageyama, A. and T. Sato, 2004: The "Yin-Yang grid": An overset grid in spherical geometry. *Geochem. Geophys. Geosys.*, arXiv: physics/0403123.

Källén, E., 1996: *HIRLAM Documentation Manual*. [Available from SMHI, S-601 76 Norrköping, Sweden].

Laprise, J. and A. Plante, 1995: A class of semi-Lagrangian integrated-mass (SLIM) numerical transport algorithms. *Mon. Wea. Rev.*, **123**, 553–565.

Lauritzen, P. H., E. Kaas, and B. Machenhauer, 2005: A mass-conservative semi-implicit semi-Lagrangian limited area shallow water model on the sphere. *Mon.Wea.Rev.*, accepted for publication (electronic version available at http://www.gfy.ku.dk/˜peter/LauritzenEtAl.pdf).

Leonard, B. P., A. Lock, and M. MacVean, 1996: Conservative explicit unrestricted-time-step multidimensional constancy-preserving advection schemes. *Mon. Wea. Rev.*, **124**, 2588–2606.

Leslie, L. M. and J. R. Purser, 1995: Three-dimensional mass-conserving semi-Lagrangian scheme employing forward trajectories. *Mon. Wea. Rev.*, **123**, 2551–2566.

Lin, S. and R. Rood, 1996: Multidimensional flux-form semi-Lagrangian transport schemes. *Mon. Wea. Rev.*, **124**, 2046–2070.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307.

Lindberg, K. and V. A. Alexeev, 2000: A study of the spurious orographic resonance in semi-implicit semi-Lagrangian models. *Mon. Wea. Rev.*, **128**, 1982–1989.

Lorenz, E., 1960: Energy and numerical weather prediction. *Tellus.*, **12**, 364–373.

Machenhauer, B., 1992: A mass-, energy- and entrophy conserving semi-lagrangian and explicit integration scheme for the primitive meteorological equations. Technical Report 146, Max Planck Institute for Meteorology MPI, Bundesstrasse-55, D-20146, Hamburg, Germany.

Machenhauer, B., E. Kaas, and P. Lauritzen, 2005: *Special Volume on Computational Methods for the Ocean and Atmosphere*, Elsevier, chapter Finite Volume Techniques in Atmospheric Models. In prep.

Machenhauer, B. and M. Olk, 1997: The implementation of the semi-implicit scheme in cell-integrated semi-lagrangian models. *Numerical Methods in Atmospheric and Oceanic Modelling - The André J. Robert Memorial Volume*, C. Lin, R. Laprise, and H. Ritchie, eds., CMOS/NRC Research Press, 103–126.

Machenhauer, B. and M. Olk, 1998: Design of a semi-implicit cell-integrated semi-Lagrangian model. Max Planck Institute for Meteorology Tech. Rep. 265, Hamburg, Germany, 76-85.

Majewski, D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, 2002: The operational global icosahedral-hexagonal gridpoint model gme: Description and high-resolution tests. *Mon. Wea. Rev.*, **130**, 319–338.

McDonald, A., 1994: The HIRLAM two time level, three dimensional semi-Lagrangian, semi-implicit, limited area, grid point model of the primitive equations. Technical Report 17.

— 1998: Default horizontal diffusion coefficient in HIRLAM-4.1. HIRLAM newsletter No.31, electronic version available at http://hirlam.knmi.nl/.

— 1999: An examination of alternative extrapolations to find the departure point position in a "two-time-level" semi-Lagrangian integration. *Mon. Wea. Rev.*, **127**, 1985–1993.

McDonald, A. and J. Bates, 1987: Improving the estimate of the departure point position in a two-time level semi-Lagrangian and semi-implicit scheme. *Mon. Wea. Rev.*, **115**, 737–739.

— 1989: Semi-Lagrangian integration of a gridpoint shallow water model on the sphere. *Mon. Wea. Rev.*, **117**, 130–137.

McDonald, A. and J. Haugen, 1992: A two-time-level, three-dimensional semi-Lagrangian, semi-implicit, limited-area gridpoint model of the primitive equations. *Mon. Wea. Rev.*, **120**, 2603–2621.

McDonald, A. and J. E. Haugen, 1993: A two time-level, three-dimensional, semi-Lagrangian, semi-implicit, limited-area gridpoint model of the primitive equations. part ii: Extension to hybrid vertical coordinates. *Mon. Wea. Rev.*, **121**, 2077–2087.

McGregor, J. L., 1993: Economical determination of departure points for semi-Lagrangian models. *Mon. Wea. Rev.*, **121**, 221–230.

Mesinger, F. and A. Arakawa, 1976: Numerical methods used in atmospheric models. GARP Publication Series 17, WMO-ICSU Joint Organizing Committee, 64 pp.

Miller, R. D., 1994: Computing the area of a spherical polygon. *Graphics Gems IV*, P. S. Heckbert, ed., Academic Press Professional, Inc., 132–137.

Moorthi, S., R. Higgins, and J. Bates, 1995: A global multilevel atmospheric model using a vector semi-Lagrangian finite-difference scheme. part ii: Version with physics. *Mon. Wea. Rev.*, **123**, 1523–1541.

Nair, R., J. Côté, and A. Staniforth, 1999: Monotonic cascade interpolation for semi-Lagrangian advection. *Q. J. R. Meteorol. Soc.*, **125**, 197–212.

Nair, R. D., 2004: Extension of a conservative cascade scheme on the sphere to large courant numbers. *Mon. Wea. Rev.*, **132**, 390–395.

Nair, R. D. and B. Machenhauer, 2002: The mass-conservative cell-integrated semi-Lagrangian advection scheme on the sphere. *Mon. Wea. Rev.*, **130**, 649–667.

Nair, R. D., J. S. Scroggs, and F. H. M. Semazzi, 2002: Efficient conservative global transport schemes for climate and atmospheric chemistry models. *Mon. Wea. Rev.*, **130**, 2059–2073.

— 2003: A forward-trajectory global semi-Lagrangian transport scheme. *J. Comput. Phys.*, in press.

Polvani, M., R. K. Scott, and S. J. Thomas, 2004: Numerically converged solutions of the global primitive equations for testing the dynamical core of atmospheric gcms. *Mon.Wea.Rev.*, **132**, 2539–2552.

Priestley, A., 1993: A quasi-conservative version of the semi-Lagrangian advection scheme. *Mon. Wea. Rev.*, **121**, 621–632.

Purnell, D., 1976: Solution of the advective equation by upstream interpolation with cubic splines. *Mon. Wea. Rev.*, **104**, 42–48.

Purser, R. and L. Leslie, 1991: An efficient interpolation procedure for high-order three-dimensional semi-Lagrangian models. *Mon. Wea. Rev.*, **119**, 2492–2498.

Rančić, M., 1992: Semi-Lagrangian piecewise biparabolic scheme for two-dimensional horizontal advection of a passive scalar. *Mon. Wea. Rev.*, **120**, 1394–1405.

— 1995: An efficient, conservative, monotonic remapping for semi-Lagrangian transport algorithms. *Mon. Wea. Rev.*, **123**, 1213–1217.

Rasch, P. and D. Williamson, 1990: Computational aspects of moisture transport in global models of the atmosphere. *Q. J. R. Meteorol. Soc.*, **116**, 1071–1090.

Riishøjgaard, L., S. Cohn, and R. Menard, 1998: The use of spline interpolation in semi-Lagrangian transport models. *Mon. Wea. Rev.*, **126**, 2008–2016.

Ritchie, H., 1987: Semi-lagrangian advection on a Gaussion grid. *Mon. Wea. Rev.*, **115**, 608–619.

— 1988: Application of the semi-Lagrangian method to a spectral model of the shallow water equations. *Mon. Wea. Rev.*, **116**, 1587–1598.

Ritchie, H. and C. Beaudoin, 1994: Approximations and sensitivity experiments with a baroclinic semi-Lagrangian spectral model. *Mon. Wea. Rev.*, **122**, 2391–2399.

Rivest, C., A. Staniforth, and A. Robert, 1994: Spurious resonant response of semi-Lagrangian discretizations to orographic forcing: Diagnosis and solution. *Mon. Wea. Rev.*, **122**, 366–376.

Robert, A., 1981: A stable numerical integration scheme for the primitive meteorological equations. *Atmos.-Ocean*, **19**, 35–46.

Robert, A., J. Henderson, and C. Turnbull, 1972: An implicit time integration scheme for baroclinic models of the atmosphere. *Mon. Wea. Rev.*, **100**, 329–335.

Roeckner, E., G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. G. S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, A. Rhodin, U. Schlese, U. Schulzweida, and A. Tompkins, 2003: Echam5 model description. Technical Report 349, Max Planck Institute for Meteorology.

Sadourny, R., 1972: Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids. *Mon. Wea. Rev.*, **100**, 136–144.

Sadourny, R., A. Arakawa, and Y. Mintz, 1968: Integration of the nondivergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere. *Mon. Wea. Rev.*, **96**, 351–356.

Semazzi, F. H., J.-H. Qian, and J. S. Scroggs, 1995: A global nonhydrostatic semi-Lagrangian atmospheric model without orography. *Mon. Wea. Rev.*, **123**, 2534–2550.

Simmons, A. and D. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758–766.

Simmons, A. J. and C. Temperton, 1997: Stability of a two-time-level semi-implicit integration scheme for gravity wave motion. *Mon. Wea. Rev.*, **125**, 600–615.

Smolarkiewicz, P. K. and J. A. Pudykiewicz, 1992: A class of semi-Lagrangian approximations for fluids. *J. Atmos. Sci.*, **49**, 2082–2096.

Staniforth, A. and J. Côté, 1991: Semi-Lagrangian schemes for atmospheric models-a review. *Mon. Wea. Rev.*, **119**, 2206–2223.

Staniforth, A., A. White, and N. Wood, 2003: Analysis of semi-Lagrangian trajectory computations. *Q. J. R. Meteorol. Soc.*, **129**, 2065–2085.

Sun, W.-Y. and M.-T. Sun, 2004: Mass correction applied to semi-Lagrangian advection scheme. *Mon.Wea.Rev.*, **132**, 975–984.

Sun, W.-Y. and K.-S. Yeh, 1997: A general semi-Lagrangian advection scheme employing forward trajectories. *Q. J. R. Meteorol. Soc.*, **544**, 2463–2476.

Takacs, L., 1985: A two-step scheme for the advection equation with minimized dissipation and dipsersion errors. *Mon. Wea. Rev.*, **113**, 1050–1065.

Tanguay, M., E. Yakimiw, H. Ritchie, and A. Robert, 1992: Advantages of spatial averaging in semi-implicit semi-Lagrangian schemes. *Mon. Wea. Rev.*, **120**, 113–123.

Temperton, C. and A. Staniforth, 1987: An efficient two-time-level semi-Lagrangian semi-implicit integration scheme. *Q. J. R. Meteorol. Soc.*, **113**, 1025–1039.

Thuburn, J., 1997: A PV-based shallow-water model on a hexagonal-icosahedral grid. *Mon. Wea. Rev.*, **125**, 2328–2350.

Thuburn, J. and Y. Li, 2000: Numerical simulations of Rossby-Haurwitz waves. *Tellus*, **52A**, 181–189.

Undén, P., 2002: *HIRLAM-5 Scientic Documentation*. [Available from SMHI, S-601 76 Norrköping, Sweden].

van Leer, B., 1977: Towards the ultimate conservative difference scheme. IV: A new approach to numerical convection. *J. Comput. Phys.*, **23**, 276–299.

Watterson, I. and M. Dix, 2005: Effective sensitivity and heat capacity in the response of climate models to greenhouse gas and aerosol forcings. *Q. J. R. Meteorol. Soc.*, **131**, 259–279.

White, E. P., 2001: Part III: Dynamics and numerical procedures. *Integrated Forecasting System (IFS) documentation, Cycle CY23r4, ECMWF*, [Available online at www.ecmwf.int.].

Williamson, D., 1968: Integration of the barotropic vorticity equations on a spherical geodesic grid. *Tellus*, **20**, 642–653.

Williamson, D., J. Drake, J. Hack, R. Jakob, and P. Swarztrauber, 1992: A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.*, **102**, 211–224.

Williamson, D. and P. Rasch, 1989: Two-dimensional semi-Lagrangian transport with shape-preserving interpolation. *Mon. Wea. Rev.*, **117**, 102–129.

Williamson, D. L., 1988: The effect of vertical finite difference approximations on simulations with the NCAR Community Climate Model. *J. Climate*, **1**, 40–58.

Xiao, F., T. Yabe, X. Peng, and H. Kobayashi, 2002: Conservative and oscillation-less atmospheric transport schemes based on rational functions. *J. Geophys. Res.*, **107**, 4609, doi:10.1029/2001JD001532.

Zalesak, S. T., 1979: Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.*, **31**, 335–362.

Zerroukat, M., N. Wood, and A. Staniforth, 2002: SLICE: A semi-Lagrangian inherently conserving and efficient scheme for transport problems. *Q. J. R. Meteorol. Soc.*, **128**, 2801–2820.

— 2004a: A monotonic and positive-definite filter for a semi-Lagrangian inherently conserving and efficient (SLICE) scheme. *Q. J. R. Meteorol. Soc.*, submitted.

— 2004b: SLICE-S: A semi-Lagrangian inherently conserving and efficient scheme for transport problems on the sphere. *Q. J. R. Meteorol. Soc.*, **130**, 2649–2664.