A HISTORY OF GLOBAL NUMERICAL WEATHER PREDICTION. The spectacular development of computational resources in the past decades has had a profound impact on the field of numerical weather and climate modeling. It has facilitated significant improvements in the description of key physical processes such as radiative transfer and has led to more accurate flow solvers. In addition, it has enabled sophisticated data-assimilation and ensemble-prediction schemes, both of which have turned out to be vital for improved prediction skill. Parallel to these developments, the increased computational power has led to a gradual but steady refinement of the computational grid. This has allowed models to resolve an increasingly large portion of the atmospheric scales of motion visualized in Fig. 1. The unresolved scales need to be approximated in a statistical way through statistical parameterizations, inevitably involving uncertain assumptions. The parametric representation of unresolved processes, especially processes related to clouds and convection, is considered to form a major source of uncertainty in weather (Slingo and Palmer 2011) and climate models (Dufresne and Bony 2008).

The historic evolution of computational grids is illustrated in the top panel of Fig. 2, which shows how the spatial scales treated by operational global numerical weather prediction (NWP) models have evolved in time. The range of resolved scales is visualized by a horizontal bar, with the largest scale (domain size) at the right and the smallest scale (resolution) at the left. The width of the bar is therefore a key measure of computational cost. Due to ever-increasing computational resources, operational NWP models have undergone an exponential increase in horizontal resolution. This growth started in 1974 with the model of the National Meteorological Center (NMC).

**Fig. 1.** Illustration of the atmospheric scales of motion, following Smagorinsky (1974). Common atmospheric phenomena are categorized according to their typical length scale (blue shades roughly indicate where the phenomena’s energy is most often concentrated). Important length scales are \( L_{PBL} \) and \( L_{Trop} \), the typical height of the planetary boundary layer and that of the troposphere, respectively.
at 300-km resolution (denoted N74; see Shuman 1989) and continued up to the resolution of 16 km that is now used by the latest version of the European Centre for Medium-Range Weather Forecasts model (E79–E10; see e.g., Simmons et al. 1989; European Centre for Medium-Range Weather Forecasts 2014). The red bars illustrate the computational breakthroughs by Miura et al. (M06; 2007), who simulated the global weather for one week at 3.5-km resolution, and by Miyamoto et al. (M13; 2013), who simulated 12 h at 0.87-km resolution some years later. While such exceptional cases cannot be performed on a regular basis, they illustrate the very limit of what is currently possible.

Operational global NWP models are presently on the verge of using resolutions finer than the depth of the troposphere, $L_{Trop}$ (10 km) (see Fig. 1). This implies that they are beginning to resolve the vertical convective overturning by cumulus clouds, but still need its partial parameterized representation. This obstacle, known as the “gray zone” or "Terra Incognita" (Wyngaard 2004) is like the proverbial “chasm” that cannot be crossed in small steps. Ideally the representation of convective overturning at these resolutions should be distributed smoothly (i.e., as a function of resolution), between the subgrid parameterizations on the one hand and explicit simulation on the other (see, e.g., Molinari and Dudek 1992; Wyngaard 2004; Arakawa et al. 2011). This can be achieved by making parameterizations “scale aware,” but a general framework for such an approach is presently lacking within the context of NWP models (e.g., particularly for convection).

If resolutions of 100 m are used, the vertical overturning of cumulus convection is resolved, but one could argue that a next gray zone of three-dimensional turbulence is entered: whereas the largest turbulent eddies, having the size of the depth of the planetary boundary layer $L_{PBL}$ (1 km), are resolved, the smaller turbulent eddies are still unresolved. Fortunately, the self-similar nature of inertial-range turbulence (with scales of 100 m or less), visualized in Fig. 1, can be used to express the transport of unresolved eddies in terms of the resolved eddies in a truly scale-aware fashion (see next section). It would therefore be a cornerstone achievement to perform global NWP at a turbulence-resolving grid resolution of 100 m. A naive extrapolation of the historical development displayed in the top panel of Fig. 2 suggests that this could be feasible on a daily basis around 2060.

DEVELOPMENT OF LES MODELS. Pioneered by Lilly (1962) and Deardorff (1972), Large-Eddy Simulation (LES) models employ subgrid transport parameterizations that exploit the self-similarity of inertial-range turbulence to represent atmospheric turbulence, instead of resolving motions down to the millimeter scale. Therefore, resolutions on the order of 100 m can suffice to simulate turbulent transport in the absence of external complexities, like detailed terrain features, although higher resolutions and therefore higher effective Reynolds numbers are desirable (see, e.g., Sullivan and Patton 2011). The historic development of computational grids for this type of model is shown...
in the bottom panel of Fig. 2. Initially, efforts were concentrated on improving the model characteristics and including more model components (D72, S76, B81; Deardorff 1972; Sommeria 1976; Bougeault 1981), but since the mid-nineties, there has been interest in increasing the domain size (C92, R04; Chlond 1992; de Roode et al. 2004). An important subsequent development were the simulations by Khairoutdinov and Randall (K06; 2006) and Khairoutdinov et al. (K09; 2009), who extended the domain size beyond 100 km. In that respect, it is tempting to also extrapolate the trend in the lower panel and imagine when the scale of the Earth will be reached. Interestingly, such thought experiments demonstrate that the final result of both approaches is the same: a global turbulence-resolving model. The refinement of global models might seem the most natural approach to this challenge, as these models have all of the important components already in place, having been running in an operational setting for decades. The gray zone is not easily crossed, however, especially for the traditional hydrostatic models. Promising work is currently being undertaken to develop nonhydrostatic global models that resolve convection when the employed resolution and parameterizations allow it (e.g., the MPAS-model, Skamarock et al. 2012).

An alternative approach to the gray-zone problem may be to use LES modeling as a starting point and develop a framework to enlarge the employed domain. The idea of using such an LES-based enlargement approach was clearly expressed in a recently funded massive German initiative called HD(CP)2 (Rauser 2014), which formulated the aim to simulate the weather over Germany with a resolution of 100 m. However, the realization of such an ambition is a major challenge.

While focusing on the domain expansion of “classical” LES models and the grid refinement of NWP models, we entirely neglected an important body of intermediate-scale models (i.e., limited-area models, or mesoscale models). Mesoscale models (such as WRF) have been developed that close the gap between NWP and LES models (e.g., Skamarock et al. 2008; Mirocha et al. 2013), or allow detailed representation of a specific subdomain through a nested LES (Moeng et al. 2007). Scale-aware turbulence models as developed by, for example, Bou-Zeid et al. (2005), Hatlee and Wyngaard (2007), and Perot and Gadebusch (2009) may allow such models to flexibly scale toward typical LES resolutions. For the sake of simplicity, however, we focus here on the two “ends” of the spectrum, highlighting the contrast between domain size and resolution.

**WEATHER SIMULATION WITH LES?** Here we perform an exploratory exercise into the LES-enlargement approach by employing the newest computer architectures—particularly graphical processing units (GPUs). GPUs have gone through tremendous development in the past decade, aided by the large financial stimulus from commercial gaming. In contrast to the central processing unit (CPU), GPUs were explicitly designed for parallel computing. Recently, GPUs evolved into general-purpose devices for massively parallel computations. In this form, they can provide an enormous computational boost, provided the programs are specifically adapted to the GPU design. A short explanation of GPU modeling is given in the sidebar.

Because relatively simple parameterizations facilitate quick adaptation to a new computer architecture, we were able to create GALES, a GPU-resident Atmospheric LES (Schalkwijk et al. 2012; Heus et al. 2010). As a result, simulations that normally require roughly 50 processors can be performed on a single GPU with comparable speeds. Yet, to enable the massive simula-

**Fig. 3.** The simulated cloud field by GALES on (a) 28 May 2006 (cloud streets) and (b) 28 Jun 2009 (thunderstorm).
tions presented below, GALES was further extended to exploit the multi-GPU hardware as featured by the latest “accelerated” supercomputers. The hardware in this study was provided by the French CURIE Hybrid supercomputer, allowing for simulations using as many as 256 GPUs simultaneously. This enabled us to simulate a domain that spans 400 × 400 km² with 100-m resolution, enough to cover the Netherlands.

To test the applicability of LES modeling to the daily weather, we have simulated three different “archetype” weather situations by nesting GALES in a weather model, as detailed in the appendix. These situations comprise fair weather cumulus clouds, cloud streets in high-wind conditions (Fig. 3a), and the development

**Fig. 4. Clouds over the Netherlands on 6 Jul 2004.** The image shows the simulated cloud field by GALES, with the corresponding satellite image as inset.

---

**GPU COMPUTATION**

The differences between the graphical processing unit (GPU) and the central processing unit (CPU) stem from their traditional roles in the computer. The CPU (the traditional “brains” of the computer) is responsible for the operating system (OS) and running programs. As a result, the CPU chip acquired a large cache (short-term, very fast memory for data in use) and a large scheduler (responsible for distributing the processing power over the various tasks and programs) to efficiently handle a relatively small number of unrelated and very diverse tasks. In contrast, GPUs typically perform a large number of nearly identical computations—they are responsible for the intensive vector/matrix computations associated with computer graphics—and therefore developed into massively parallel devices.

LES models perform calculations that are very similar to these matrix computations. First, the equation solved is the same for every grid node, and second, the data are ordered in a structured manner. This allows for a straightforward task division on the one hand, and efficient memory access on the other hand: each consecutive GPU core performs the same instruction, but on a consecutive data element.

GALES is written in C++ and CUDA: CPU instructions in C++ for program management and I/O tasks, CUDA for the GPU computations. The implementation for single-GPU purposes is described in Schalkwijk et al. (2012). To perform the large simulations shown in this manuscript, we have complemented the GPU parallelism with the more traditional method of message passing. The domain is divided into subdomains, which are distributed among the processors (CPUs). Each CPU offloads its subdomain to the GPU, which performs computations for
of a severe thunderstorm (Fig. 3b). For brevity, we will focus on the case of fair weather cumulus clouds (Fig. 4), which are the most abundant clouds on the globe (Rossow et al. 2005), but which are still poorly represented in conventional weather and climate models. Note that the simulations in this work require 4 h of computation for every simulated hour, and are therefore yet unfit as forecasts. Rather, they represent special cases of extreme computing and serve to illustrate what will become possible in the near future.

The cumulus case visualized in Fig. 4 is characterized by cumulus clouds that are typically 100 m–1 km wide and up to a few kilometers deep. A comparison with the satellite image at the same time (see inset in Fig. 4) shows that the simulation indeed forecasts the observed fair weather cumulus, which are notoriously

---

**Fig. 5.** The simulated cloud field that was shown in Fig. 4, but rendered from a different perspective.
difficult to represent in conventional models (Slingo and Palmer 2011). The high-resolution features are illustrated yet more clearly in Fig. 5, which shows a three-dimensional rendering of the same cloud field. The shadows underneath the clouds illustrate the possibility of representing the fine interaction between clouds, radiation, and the surface. This interaction directly affects the temperature and humidity in the atmospheric boundary layer, and thereby the daily weather.

In addition, the sea breeze effects, recognizable by clear skies in the coastal regions and intensified cloudiness further inland, are remarkably well represented. However, the present study should by no means be considered a comprehensive skill analysis. Our purpose is to present a proof of concept of predictive Large-Eddy Simulations and to illustrate its potential for operational forecasting.

In general, it is well accepted that many of the long-term biases in weather and climate models are due to an incorrect representation of the interactions between clouds and the large-scale circulation (Stevens and Bony 2013). Such biases include the wrong phasing of precipitation in the diurnal cycle (Betts and Jakob 2002), the misrepresentation of the Madden–Julian oscillation (Miura et al. 2007), and systematic errors in the precipitation patterns in the tropics. These biases are likely due to an incorrect interaction between the large-scale resolved dynamics and the parameterized cloud processes in present-day weather and climate models. Case studies on smaller domains have shown that many of these biases can be mitigated by using high-resolution simulations.

**CONCLUSION.** Bearing in mind our earlier extrapolation anticipating high-resolution operational NWP models in the year 2060, it is useful to estimate when a global 100-m simulation will become technically possible in a special case of extreme computing. For GALES we know exactly how additional GPUs improve the simulation speed and the domain size (see sidebar). This leads us to conclude that the world’s top GPU-accelerated supercomputer, Oak Ridge’s Titan, could run a simulation of 3,200 x 3,200 km² domain at 100-m resolution already today (for reference, the Earth’s diameter is roughly 6,400 km). This virtual breakthrough is illustrated with the red bar in the bottom panel of Fig. 2 (tagged T14). If, in addition, we extrapolate the historic supercomputing trend (Strohmaier et al. 2005), we find that a global turbulence-resolving simulation will be possible in less than 10 years.

Of course, this approach also faces challenges in its route to global simulations. In many cases, LES models require modifications to the pressure solver and even to the grid itself to handle the Earth’s curvature and orography, not to mention the assimilation of observational data. Although such challenges might entail significant additional computational cost, they do not conflict with the model design: many issues have already been solved in the fields of weather forecasting or, for example, fluid dynamics. The relatively low number of assumptions in LES models aids the quick adaptation to new computational strategies, and the models are well prepared to utilize the exponential increase of computational power. Therefore, we should not be surprised when the first global LES experiments appear within the next decade.

**ACKNOWLEDGMENTS.** We acknowledge PRACE for awarding us access to the CURIE hybrid nodes based in France at TGCC, owned by GENCI and operated by CEA. The authors gratefully thank the editor and three anonymous reviewers for their suggestions to improve the manuscript. Thanks also to Anton Beljaars, Peter Sullivan, Henk Dijkstra, and Bjorn Stevens for their helpful comments on an earlier draft of the paper.

**Fig. 6.** Domain nesting. The solid black line represents the domain of RACMO, spanning 126 x 120 points in the horizontal directions; the dashed line indicates the transition between inner domain and border region. The solid blue line depicts the domain of GALES. Colors indicate the surface level in meters.
**APPENDIX: SETUP.** The initial and lateral boundary conditions were provided by the Regional Atmospheric Climate Model (RACMO) that was developed at the KNMI (Royal Netherlands Meteorological Institute). RACMO is a hydrostatic limited-area model utilizing the physics parameterization package from the ECMWF model, extensively described in van Meijgaard et al. (2008). The domain of RACMO is illustrated in Fig. 6. The LES domain ranges approximately from 2.61° to 8.64° E and from 50.41° to 54.09° N (400 km x 400 km) and is marked in blue.

The original LES model is extensively described in Heus et al. (2010); only case-specific details are treated here. Following Böing et al. (2012), the Boussinesq approximation was replaced by an anelastic approximation to account for density variations in the vertical direction, and a simple, single-moment ice microphysics scheme is employed. These modifications facilitate the extension of the domain to the 15-km altitude used in this study.

The vertical resolution increases exponentially from \( \Delta z = 40 \text{ m} \) at the surface to 70 m at the top of the domain. Combined with a horizontal resolution of \( \Delta x = \Delta y = 100 \text{ m} \), the cumulus case, having boundary layer heights \( z \) between 3.5 km and 5 km, can be characterized by \( 45 < z / \Delta z < 65 \) (with \( \Delta z = \sqrt{\Delta x \Delta y / \Delta z} \); which corresponds to an effective Reynolds number \( \text{Re}_z = (z / \Delta z)^{1/2} \approx 200 \). At these values for \( z / \Delta z \), the first- and second-moment statistics start to converge (Sullivan and Patton 2011), but higher effective Reynolds numbers are desirable.

During the LES runs, we retained periodic boundary conditions to allow developed turbulent fluctuations to reenter on the opposite side. Within a border region around the domain, the average state is relaxed toward RACMO. This procedure, outlined as follows, effectively provides boundary layer conditions to the mean state while retaining turbulent fluctuations.

The variables \( \{u,v,\theta,q\} \); longitudinal velocity, latitudinal velocity, liquid water potential temperature, and specific humidity, respectively, are relaxed toward the RACMO state in the border regions. First, the average of a variable \( \phi \in \{u,v,\theta,q\} \) is calculated within the border region:

\[
\langle \phi \rangle_b = \frac{\iint_S \phi \, dx \, dy}{\iint_S f \, dx \, dy}
\]

where \( S \) is a single subdomain of 1 GPU (25 km x 25 km), and \( f \) is a factor that creates a smooth transition between border region and inner domain: \( f = 1 \) at the boundary and \( f = 0 \) well into the inner domain. The tendency due to the boundary conditions (bc) is given by

\[
\left[ \frac{\partial \phi}{\partial t} \right]_{bc} = -\frac{f}{\tau} \left( \langle \phi \rangle_b - \phi^s \right)
\]

where the superscript \( R \) denotes RACMO’s value of the variable. This formulation preserves turbulence while allowing for heterogeneous boundaries. The border region comprises a 12-km-wide edge around the square domain, and the relaxation time scale \( \tau \) is 300 s, roughly the time scale at which high winds move through the border region. Another 12 km is removed from the simulation results to allow turbulence to adapt to the new mean state. Alternatively, one might provide in/outflow boundary conditions, but also here the supply of turbulent fluctuations remains a complex problem (e.g., Mirocha et al. 2013).

The conservation equation for an LES filtered variable \( \phi \) becomes:

\[
\frac{\partial \phi}{\partial t} = -\frac{1}{\rho_0} \nabla \cdot \rho_0 \bar{u} \phi - \frac{1}{\rho_0} \nabla \cdot \rho_0 \bar{u} \phi^s - w^b \frac{\partial \phi}{\partial z} - \frac{f}{\tau} \left( \langle \phi \rangle_b - \phi^s \right) + S_{bc}^\phi
\]

where \( \bar{u} = (u,v,w) \) is the velocity vector.

The first two terms on the right-hand side represent the advection of resolved and subfilter-scale turbulence, respectively. Subfilter-scale motions in GALES, denoted with the superscript \( s \), are treated through eddy viscosity/diffusivity fluxes, modeled as a function of the subfilter-scale turbulent kinetic energy; \( \rho_0 = \rho_0(z) \) is a base density profile that is dependent on height only.

The third and fourth term are meant to represent the effects of large-scale vertical and horizontal transport, respectively. The third term models subsidence, which acts as a slow downward (or upward) force. By using RACMO’s state for the lateral boundary conditions (fourth term), the LES is effectively provided with large-scale horizontal advection into the LES domain.

For \( \phi \in \{\theta,q\} \), the source term \( S_{bc}^\phi \) includes effects from precipitation and radiation. Since no GPU-accelerated radiation module was available to GALES at the time of the simulation, we employed RACMO’s radiative tendencies for the latter.
The momentum equations are essentially equal to Eq. (A1) if \( S \) is taken to include tendencies due to pressure fluctuations (Heus et al. 2010) and the ageostrophic acceleration \( 2\Omega x (\hat{a} - \hat{v}) \), with \( \hat{u} \) the geostrophic wind from RACMO.

Last, GALE'S surface model was modified to handle a heterogeneous roughness map at 100-m resolution, as well as a high-resolution land–sea map (converted from Farr et al. 2007). Based on the latter, the surface model switches between a prescribed sea surface temperature formulation and a land surface model that accepts the net radiative surface flux as input.

**FOR FURTHER READING**


European Centre for Medium-Range Weather Forecasts, 2014: Evolution of the ECMWF analysis and forecasting system. [Available online at www.ecmwf.int/products/data/operational_system/evolution.]


Radar and Atmospheric Science: A Collection of Essays in Honor of David Atlas
Edited by Roger M. Wakimoto and Ramesh Srivastava

This monograph pays tribute to one of the leading scientists in meteorology, Dr. David Atlas. In addition to profiling the life and work of the acknowledged “Father of Radar Meteorology,” this collection highlights many of the unique contributions he made to the understanding of the forcing and organization of convective systems, observation and modeling of atmospheric turbulence and waves, and cloud microphysical properties, among many other topics. It is hoped that this text will inspire the next generation of radar meteorologists, provide an excellent resource for scientists and educators, and serve as a historical record of the gathering of scholarly contributions honoring one of the most important meteorologists of our time.

Radar and Atmospheric Science:
A Collection of Essays in Honor of David Atlas


Price $75.00 list/$55.00 member

To place an order point your Web browser to www.ametsoc.org/amsbookstore

AMS BOOKS

RESEARCH APPLICATIONS HISTORY