Climate, weather, space weather: model development in an operational context

Doris Folini

J. Space Weather Space Clim. 2018, 8, A32

https://doi.org/10.1051/swsc/2018021

JSWSC

Available online at:
www.swsc-journal.org

Developing New Space Weather Tools: Transitioning fundamental science to operational prediction systems

RESEARCH ARTICLE OPEN ACCESS

Climate, weather, space weather: model development in an operational context

Doris Folini[1,2,*]

[1] Institute for Atmospheric and Climate Science, ETH Zurich, 8092 Zurich, Switzerland
[2] École Normal Supérieur, Lyon, CRAL, UMR CNRS 5574, Université de Lyon, 69364 Lyon Cedex 07, France

Received 13 June 2017 / Accepted 30 April 2018

Abstract – Aspects of operational modeling for climate, weather, and space weather forecasts are contrasted, with a particular focus on the somewhat conflicting demands of "operational stability" versus "dynamic development" of the involved models. Some common key elements are identified, indicating potential for fruitful exchange across communities. Operational model development is compelling, driven by factors that broadly fall into four categories: model skill, basic physics, advances in computer architecture, and new aspects to be covered, from costumer needs over physics to observational data. Evaluation of model skill as part of the operational chain goes beyond an automated skill score. Permanent interaction between "pure research" and "operational forecast" people is beneficial to both sides. This includes joint model development projects, although ultimate responsibility for the operational code remains with the forecast provider. The pace of model development reflects operational lead times. The points are illustrated with selected examples, many of which reflect the author's background and personal contacts, notably with the Swiss Weather Service and the Max Planck Institute for Meteorology, Hamburg, Germany. In view of current and future challenges, large collaborations covering a range of expertise are a must – within and across climate, weather, and space weather. To profit from and cope with the rapid progress of computer architectures, supercompute centers must form part of the team.

Keywords: space weather / forecasting / methodology / strategy / meteorology

1 Introduction

Model development in the context of an operational chain or forecast service implies that one has to deal with the somewhat conflicting demands of "operational stability" versus "dynamic development". On the one hand, one does not want to touch the operational model, whose compliance with specifications has been ascertained – a state loosely referred to as "operational stability" in the following. Specifications typically depend on the forecast service and cover aspects from false alarm rate over time-criticality of model run time to compatibility with the entire operational chain. On the other hand, model development is compelling for a number of reasons, to be detailed later. Such "dynamic development" typically breaks the aforementioned compliance with specifications, which ultimately has to be re-established to regain "operational stability". Although a topic of discussions and meetings, the subject gets rather little coverage in the literature (Steenburgh et al., 2014).

The present paper takes a cross-community view on the question, from the perspective of climate, weather, and space weather. Earlier studies demonstrated the potential for mutual learning across these three communities (Siscoe, 2007). All three of them aim at predicting the future physical state of the "Sun–Earth" system in a likelihood sense and rely on the same, basic building blocks for their models: physical insight and associated equations – from empirical dependencies to basic physical laws – that characterize the time evolution of the system, numerical simulations to integrate the physical equations in time, and measurement data for model initialization (often via advanced data assimilation techniques) and validation (Tsagouri et al., 2013; Bauer et al., 2015; Palmer, 2016). To translate the output of the numerical model into customer specific products, expert knowledge (forecasters) and dedicated back end models are essential. For example, in the context of emergency preparedness in case of a nuclear power plant accident, a Lagrangian particle dispersion model may be used as back end model to translate the numerical weather prediction (NWP) into a fall-out prediction (Szintai et al., 2009). All is time critical, the prediction must be available before the real event. Differences among the communities include the envisaged lead time, from minutes

*Corresponding author: Doris.Folini@env.ethz.ch

to centuries, the availability of measurement data for model initialization and validation, or the interest given to extreme events.

Customer needs – more accurate and comprehensive forecasts – may be seen as the ultimate driver behind the further development of operational forecast models. Specific drivers of development are suggested to fall into four broad categories: improvement of model skill, insight from basic science, exploitation of new computer architectures, and coverage of entirely new aspects, e.g., an additional physical model, new observational data, or customer needs. Clearly, these four categories are not completely independent of each other. For example, covering additional physics and transitioning to a more powerful computer architecture may go hand in hand. The different drivers reflect in the emerging projects in the form of project size, duration, or composition of the development team. Their relative importance depends on the community. Regular examination of driving factors ideally forms part of the operational context or even, in the case of model skill, of the forecast operational chain. Permanent exchange with basic research outside the operational service is an asset in this process: it provides another perspective, potentially taps complementary expertise, and enables co-development projects that translate driving factors into new (operational) code. The ultimate responsibility for the operational code (coding rules, documentation, verification, validation, etc.) resides, however, with the operational service institution. It is argued that such exchange is beneficial to all parties involved. On the side of research, benefits lie with the use of the operational code as a clean, well documented and well tested starting point for research or a showcase application demonstrating the power of new technologies. The paper aims at illustrating these points for each community and at carving out the potential for mutual learning across communities.

Section 2 takes a community specific point of view, likely with some bias to the author's own background in climate, weather, and astrophysics, especially when it comes to concrete examples, which are often from the Swiss Weather Service (MeteoSwiss) and the Max Planck Institute for Meteorology, Hamburg, Germany. Some aspects of the operational chain and service are addressed, thereby embedding the focal point of the paper: how different communities cope with the challenge of "operational stability" versus "dynamic development". Section 3 deals with similarities across communities and what the different communities might learn from each other. Conclusions are presented in Section 4.

2 Operational predictions and development in different communities

The objective of this section is to provide some characterization of the three communities of interest – NWP in Section 2.1, climate in Section 2.2, and space weather prediction (SWP) in Section 2.3. Each of these three sections is structured around roughly the same basic points. The goals of each community are sketched, including the targeted temporal and spatial scales. Aspects of associated operational modeling are given, from institutional issues to assessment of model skill. Model development within this context is then illustrated, form factors triggering development projects, over how such projects work, to integration of the new development into the operational model. Some examples are given.

2.1 Weather prediction

2.1.1 Objectives

Operational NWP deals with lead times of hours to days and seasons (with decadal predictions emerging) on global to regional to local scales (Meehl et al., 2014). Specific back-end codes that run after the NWP model are used to meet demands from a wide range of customers, from governments (e.g., flight safety, hurricanes) over business (e.g., tourism, water management) to individuals (e.g., pollen forecast). Interests are with both, "ordinary weather" and extreme events.

2.1.2 Modeling aspects

The range of involved scales, from global to local, translates into running a hierarchy of models: a global scale model (few 10 km grid cell size) provides the large scale dynamics and "long term perspective", which are used as boundary conditions for a more finely resolved (kilometer scale) regional model or a hierarchy thereof. Coupling across scales is one way (from coarse to fine) or, more recently, two-way (from coarse to fine and from fine to coarse) (e.g., Reinert et al., 2017). It may involve the same model at different resolutions, or different models (see below).

For the time scales of interest, initial conditions play a crucial role. Consequently, much effort has gone in model initialization techniques that furnish initial conditions close to observations/reality, yet compatible with the physics covered by the (imperfect) model (e.g., 4D Var, Courtier et al., 1994; Trémolet, 2006; Zhang & Pu, 2010; Shaw & Daescu, 2017). To address the issue of weather being a chaotic system (Lorenz, 1963), modern weather forecasts provide likelihoods. This means that not one simulation but an ensemble of simulations differing in their initial conditions must be run in a time-critical fashion – the prediction must be available way ahead of the predicted event. More recently, ensembles are also being used to explicitly sample model uncertainty, for example due to the choice of numerical values in sub-grid-scale parameterizations (e.g., Leutbecher et al., 2017). Ensemble predictions provide a framework to extend the forecast lead time in a meaningful way (Buizza & Leutbecher, 2015). But they are also a significant cpu cost factor (e.g., Leutbecher et al., 2017). Similarly, finer grid resolution increases the cpu costs, yet is highly desirable to better capture relevant small scale features, like mountains or coast-lines, and to resolve physical processes that otherwise have to be included via sub-grid scale parameterization, e.g., convection and associated precipitation (e.g., Langhans et al., 2012; Ban et al., 2014).

The above demands – hierarchy of models, observation based model initialization, ensemble prediction, all time-critical – necessitate NWP and associated model development to be an overall highly collaborative effort. This although individual weather services typically run an operational model/chain themselves. An impression can be obtained from the project web-pages of some regional scale consortia (COnsortium for Small-scale Modeling, COSMO, http://www.cosmo-model.org/; High Resolution Limited

Area Model, http://hirlam.org/; the Weather Research Forecasting model, http://www.wrf-model.org). From these consortia it also becomes obvious that hierarchies often go not only across models but also across institutions (e.g., global from the European Center for Medium-Range Weather Forecasts, regional from COSMO), making clean interfaces a must. The latter enable a comparatively easy hand shake between different models or also models and observations or models and costumers. The concrete meaning of the term interface is correspondingly broad, ranging from definitions for observations (what, when, how), over unit conventions and coordinate system definitions to data formats (netcdf and grib) and detailed documentations of data and model. Some of these standards have been put forward by the World Meteorological Organization (WMO, http://www.wmo.int). Others may be seen more as the result of co-evolution among different stake-holders. With regard to the model as such, the large consortia and the strict rules (e.g., coding but also licensing issues) that have to go with them add to the basic challenge of "operational stability" versus "dynamic development".

2.1.3 Model development

A main development driver in NWP is the permanent evaluation of forecast skill. This is possible because of the short lead times, because interest is also with ordinary weather, i.e., each forecast and not only extreme events, and because there is ample observational data at hand on Earth. Comparison of forecast and reality is typically done via an institution specific skill metric. Other measures may be added on top. For example, as part of the operational chain at MeteoSwiss forecasters meet after each shift, while memory is still fresh and as part of the operational chain, with people from modeling to provide their impression of the model's performance for the last forecasting period (personal communication P. Steiner, MeteoSwiss). Several times a year this information is analyzed for model deficiencies that escaped the automated skill score.

Model improvement can then come via an internal project at the weather service or via a joint project with universities (e.g., improved terrain following coordinates, Schär et al., 2002). On the university side, a motivation for such a joint project is the free use of most or all of the operational NWP codes, tools, and data for their research. For the weather service, this form of exchange allows to carry out also larger exploratory projects with a close link to basic atmospheric science, which might be difficult to realize otherwise, e.g., for lack of expert knowledge or funding possibilities. Such projects may last anywhere from months to years and may cover anything from small adaptations to adding new capabilities/components to the model. If they lead to improved predictions within the research context, they are professionally implemented and tested (verification, validation, calibration/tuning, effect on customer specific back-ends) by the weather service, following strict rules and protocols. If all tests are passed, the development enters the operational code. Corresponding releases take place several times per year at MeteoSwiss. Note that this implies that forecasters and customers must cope with frequent – albeit typically small – changes of the product they get.

Another development driver may be summarized as insight from basic research. Into this category falls the awareness that ensemble simulations can be used to translate imperfections in initial conditions and model formulation into meaningful probabilistic forecasts. Associated development projects tend to be rather long and complex.

Yet another development driver is the advance of computer architectures. The reward is, roughly speaking, more computation in shorter time, potentially even for less energy. In terms of the above examples, it enables finer resolution and larger ensembles. This potential has spurred interest in porting (operational) NWP to modern computer architectures, in particular, graphical processing units/GPUs (Michalakes & Vachharajani, 2008; Shimokawabe et al., 2014; Vanderbauwhede & Takemi, 2016; Deconinck et al., 2017), but there are also attempts toward using cloud computing (Molthan et al., 2015; Siuta et al., 2016; Blaylock et al., 2017; Chen et al., 2017). Associated challenges comprise scalability (the model should run on thousands of cores), fault awareness and tolerance (the model should tolerate failure of a thread or core), advanced data compression techniques and I/O, or also concepts for portability (computer architectures may differ among consortia members) and composability (relevant physics may differ among users, e.g., mountain snow pack or storm surges) or modularity (Bauer et al., 2015; Palmer, 2015; Düben & Dawson, 2017). Tackling these challenges requires – besides domain scientists – highly specialized knowledge on computing related topics, ideally in close collaboration with a large supercompute center, several years time, and likely some shift of paradigms. Among the latter are, for example, the question of programming language (replacement of fortran, at least in parts), the requirement of bit-reproducibility, partial use of reduced precision arithmetic (Düben et al., 2014), online analysis/re-calculation to reduce I/O, or whether domain scientists are willing to give up control over parts of the model (e.g., solution of a Laplacian) in favor of domain specific languages or computer architecture specific libraries.

On the positive side, the range of stake-holders involved can also open additional funding opportunities for development projects. Recent example of this kind are the project on Energy-efficient Scalable Algorithms for Weather Prediction at Exascale (http://www.hpc-escape.eu/) or also the Center of Excellence in Simulation of Weather and Climate in Europe (https://www.esiwace.eu/). Both initiatives explicitly address both, NWP and climate. In addition, there are many smaller initiatives e.g., within the context of Partnership for Advanced Computing in Europe (http://www.prace-ri.euor). So far, to the author's knowledge, only one operational NWP model has been successfully ported to GPUs, namely that of MeteoSwiss (e.g., Fuhrer et al., 2014; Gysi et al., 2015; Leutwyler et al., 2016; Prein et al., 2017). The corresponding development branch of the code was largely decoupled from the operational version for years. The effort turned out to be a win-win situation for all partners: a more powerful yet less energy intense code for operational forecasts and research, as well as a showcase application that highlights the possibilities offered by new computer architectures at the Swiss National Supercompute Center.

2.2 Climate projections

2.2.1 Objectives

Climate projections aim at lead times of decades to centuries, at global to regional spatial scales. The term "operational" is hardly used in the context of climate

modeling. Yet the model data entering the International Panel for Climate Change (IPCC) reports (IPCC, 2013), for example the global climate model (GCM) data from the Coupled Model Intercomparison Project Phase 5 (CMIP5), fulfill some operational key characteristics: the data are to some degree customer driven (United Nations Environment Program, WMO; governments or business for selected back-end products), have to be available in time (for the IPCC report), and respect a number of specifications. The specifications arise from the desire to be able to compare and re-use the model data submitted to IPCC and are defined by working groups from the climate modeling community as such (for CMIP, Taylor et al., 2012; Eyring et al., 2016a). Broadly speaking, they detail the setup of some common simulations and of some aspects of the model. An example for the former is the demand to perform a simulation covering the years 1850–2005 with predefined input data, like annual mean greenhouse gas concentrations. Model specifications include that a certain amount of physical components must be covered (e.g., atmosphere, ocean) and that standards enabling data exchange be respected (e.g., file format, unit conventions, coordinate system definitions, variable names). From this perspective, climate projections also face the issue of "operational stability" versus "dynamic development" and accompanying effects on the development process (Jakob, 2014).

The term "operational" refers, however, to a quite different product than in NWP, in terms of lead time and spatio-temporal scales, but also concerning product details – rather customer tailored in the case of NWP, closer to the model output as such in the case of climate. Associated are differences in terms of operational model (e.g., physical components covered), embedding into an operational chain (more elaborate in NWP), bodies behind the operational product (larger in NWP), and model development. The points are further illustrated in the following, with focus on GCMs/Earth system models (ESMs) and CMIP5. Similar considerations apply to regional climate models.

2.2.2 Modeling aspects

The long lead times are a distinguishing feature of (operational) climate modeling. They have several consequences for the design, operation, and further development of such models. First, they require ESMs to take into account additional system components besides the planetary atmosphere, for example oceans, sea-ice, or vegetation. The exchange between corresponding model components (or models, for short) is mostly two-way, i.e., information (e.g., an energy flux) is passed from one model to another and vice versa. Developments on one model thereby tend to impact others. Model components (e.g., atmosphere or ocean) are typically developed and brought to operational stability on a stand alone basis, before being coupled and finally adjusted across components (see e.g., Hourdin et al., 2017). Second, the long lead times allow for and also demand for longer model run times, from about a week to several months for one simulation. Third, data assimilation for initialization is less of an issue than in NWP, as either the system memory is much shorter than the lead time (e.g., for the atmosphere) or because comparatively little observational data for assimilation is available (e.g., for the deep ocean). To nevertheless arrive at a controlled initial state, an ESM is typically relaxed by running it for several thousand years with fixed setup, e.g., conditions as of the year 1850. Fourth, to arrive at a controlled initial state that is compatible with observed climate variables (e.g., a global mean temperature remaining around 13.7 °C in 1850), numerical values in sub-grid scale parameterizations have to be adjusted: the model has to be calibrated or tuned (e.g., Hourdin et al., 2017; Schmidt et al., 2017). Because of the long time scales involved (run time and system memory) this process takes from several months to over a year and consumes a significant amount of cpu. The calibration and relaxation process, as well as the generally long integration times for actual operational production, reflect in comparatively long (years) intervals between individual operational code versions. The production phase of one operational version typically coincides with the development phase of the subsequent model version and resources (people, cpu) have to be split between the two.

Operational climate modeling and associated model development is essentially shouldered by individual research institutions (around thirty in the case of the last IPCC report/CMIP5, often in a stable team with some supercompute center). This is in contrast to NWP, where large, often transnational consortia play an important role (see Sect. 2.1). The institution based approach has the advantage that the exchange between development and operation is typically easier. Operational and research codes are closely related, the latter just being (research purpose adapted) branches of the former. In this way, research can rely on the operational code as a solid basis that continues to be well tested by research and operational activities. Basic research developments thus are made already within the context of the operational code, although often not in "operational coding quality". On the downside, an institution shouldering both, research and operational modeling, must split its resources, like people or available cpu time, between the two. At times, this may result in operational goals (like IPCC) being prioritized over "pure research", one reason being that the former tend to be more time critical than the latter. Corresponding concerns were raised, for example, during a recent workshop on Earth System Modeling (https://www.4icesm.eu).

Rules and regulations associated with the operational model are mostly institution specific as well, except for specifications regarding data exchange (see above). They are a must given typical code sizes, and range from coding rules over version control (e.g., svn, git) to validation/model skill (see below). Yet they tend to be less strict and comprehensive than in NWP, possibly reflecting the different sizes of the communities behind a single model. NWP likely also requires more regulations as the operational model forms part of a complex operational chain, from data assimilation to user specific back end models and products (see Sect. 2.1), while climate projections are more stand alone. They rely on comparatively simple input data (no data assimilation) and provide essentially the model output as such (no user specific back end products or models).

2.2.3 Model development

With climate research and operational climate modeling coming out of essentially one hand, much model development

simply results from research projects – except for the operational implementation itself, which is often done by a few, specialized people.

A main development driver is again model skill, now typically with regard to observation based climatological mean quantities, possibly their variance, like (global mean) temperature, the Indian monsoon, or El Niño. Corresponding skill metrics are typically institution specific. A common skill metric that would allow to compare model skill across institutions is suggested for CMIP6 (Eyring et al., 2016a). Comparison of observed and modeled historical climate, roughly from 1850 till 2000, is more of a final (yet crucial) test for an ESM than a primary driver of its development (Hourdin et al., 2017). Translating a lack of model skill into a concrete, targeted development project tends, however, to be even more challenging in an ESM (with atmosphere, ocean, sea ice, vegetation, etc.) than in NWP.

An emerging field of climate model development is the use of an ESM in NWP mode and testing its skill correspondingly (e.g., Simmons et al., 2016). The rational behind is that climate may be seen as the long-term statistical average of weather, thus a good climate model should also have good NWP properties. As the atmospheric component of the ESM is tested against more detailed observational data than the above mentioned climatological mean quantities, it is put on more firm physical ground. The approach is also a step toward bridging the lead time gap between NWP and climate, thus toward seamless predictions (Palmer et al., 2008; Simmons et al., 2016).

A lasting driver of model development in climate is the coverage of additional system aspects, like for example replacing prescribed plants and carbon sources/sinks with a carbon cycle model comprising interactive vegetation (e.g., trees growing or dying), carbon storage in oceans, etc. Associated stand alone models may originate from pure research. Once such a model seems mature enough, in terms of science and efficiency of execution/cpu requirements, one may try to couple this model as yet another component to the ESM. Keeping ESMs modular such that they allow for easy coupling of new stand alone models is an endeavor. It is also a challenge with regard to coding, including the potentially conflicting demands of modularity on the one hand and optimization for execution speed on the other hand.

Basic physics in the sense of switching to more physically sound sub-grid-scale parameterization is another important driver. However, within the operational context this driver has the downside that physically improved parameterization may first result in reduced model skill, as errors that used to compensate no longer do so (Jakob, 2014; Hourdin et al., 2017). Additional development time is needed, which is potentially (too) costly for an individual institution. Similar considerations apply with regard to exploitation of new computer architectures. Promises lie, for example, with better spatial resolution to address regional projections and the role of clouds. Challenges include lack of man power, large codes, code portability, and uncertainty about the longevity of different accelerator technologies – thus about whether portation would pay off at all.

The situation may change as climate and weather models approach and synergies can be exploited in larger teams. An example in this direction is the ICOsahedral Non-hydrostatic model (ICON) used by the German Weather Service as operational global NWP model and by the Max Planck Institute for Meteorology, Hamburg, Germany as global climate model. Together with other stake-holders from NWP, supercompute centers, and universities, several joint projects are under way to explore portation of ICON to new computer architectures (e.g., within the Platform for Advanced Scientific Computing, http://www.pasc-ch.org/).

In summary, developments in an operational ESM context go at a much slower pace than NWP development just because of the physical and lead time scales involved in the problem. However, the number of people involved in model development likely also plays a role. Whether larger collaborations would make the development process more efficient and faster is a matter of debate (Palmer, 2016). One danger to ever growing collaborations is a loss of diversity.

2.3 Space weather prediction

2.3.1 Objectives

SWP deals with perturbations originating at the Sun – eruptive perturbations, notably X-ray flares, coronal mass ejections (CMEs), and solar energetic particle events (SEPs), as well as more persistent features, notably coronal holes and associated wind streams – and their propagation toward and effects at or near the Earth. Customer specific interests include satellite safety, high frequency communication black outs, or damages to electrical power grids from geomagnetically induced currents (e.g., Sibley et al., 2012; Riley et al., 2018). Lead times depend on the type of event and typically range from minutes to days, although longer time scales are also of interest (e.g., recurrence of coronal hole associated with the solar rotation period of 27 days; the solar cycle of about 11 and 22 years Watermann et al., 2009; Singh et al., 2010). Operational products range from publicly available activity indices to customer tailored quantities (see e.g., Araujo-Pradere, 2009; Tsagouri et al., 2013; Steenburgh et al., 2014; Schrijver et al., 2015; Bonadonna et al., 2017).

2.3.2 Modeling aspects

Operational space weather providers include consortia (e.g., the teams participating in the ESA Space Situational Awareness (SSA) Programme's Space Weather Service Network, http://swe.ssa.esa.int/) and organizations engaged in NWP (e.g., the SWP Center (SWPC), of the National Oceanic and Atmospheric Administration, http://www.swpc.noaa.gov). New developments in space weather services are taking place in a number of countries, e.g., the United Kingdom, Belgium, Poland, Sweden, Austria, Australia, Brazil, Mexico, Canada, Korea, Japan, China, Indonesia, India, or also South Africa. A collaborative network of space weather service-providing organizations around the globe is provided by the International Space Environment Service (ISES). The ISES mission is to improve, to coordinate, and to deliver operational space weather services through a network of Regional Warning Centers (http://www.spaceweather.org/).

The different types of perturbations (X-ray flares, SEPs, CMEs, coronal holes) find their correspondence in rather separated modeling communities (Zhao & Dryer, 2014; Luhmann et al., 2015; Barnes et al., 2016; Reiss et al.,

2016; Cranmer et al., 2017; Murray et al., 2017). Further splitting of modeling activity occurs for regions closer to Earth (magnetosphere, ionosphere/thermosphere, Earth atmosphere and surface) because of traditional scientific domains, specific customer needs, as well as the physical processes involved (Lathuillère et al., 2002). Models range from empirical to semi-empirical to physics based. An impression of the emerging, rather fragmented modeling landscape may be obtained from the SWPC web-page. A number of projects aim at combining this existing expertise to arrive at more comprehensive space weather models. This includes coupling of different models, i.e., using the output of one model as input / initialization of another model. The coupling is typically one way, with information being passed from Sun to Earth. Concrete initiatives include the Space Weather Modeling Framework at the University of Michigan (http://csem.engin.umich.edu/ Tóth et al., 2005, 2012) and the Virtual Space Weather Modeling Center (VSWMC, https://esa-vswmc.eu/). A complementary, more integrative approach that stresses the critical linking of multiple scales at shocks, interfaces, and reconnection sites, is taken by the Space Weather Integrated Forecasting Framework (http://www.swiff.eu/ Lapenta et al., 2013).

Model initialization relies on satellite or other observational data and uses a range of data assimilation techniques (e.g., Ensemble Kalman Filtering or Ensemble Optimal Interpolation methods Hickmann et al., 2015; Murray et al., 2015). Data assimilation for model initialization is, however, not as widely spread in SWP as in NWP. Several reasons for this difference are identified by Lang et al. (2017), a major one being data availability. Translating the uncertainty from initialization into probabilistic forecasts using different ensemble techniques (Schunk et al., 2014; Elvidge et al., 2016; Knipp, 2016; Owens et al., 2017) is getting standard. Dependence on initial conditions can be chaotic (as in NWP, e.g., magnetosphere/ionosphere/thermosphere Horton et al., 2001; Mannucci et al., 2016; Wang et al., 2016) or non-chaotic (e.g., CME propagation toward Earth Cash et al., 2015; Lee et al., 2013, 2015; Pizzo et al., 2015). In the later case, the accuracy of a prediction will strongly depend on the quality of the underlying initial data, for example the parameters of a CME and the characterization of the solar wind to be passed. Depending on how model skill is evaluated it may then rather be an evaluation of "initial condition skill" than of actual model skill. Also, one model may outperform another because it was designed to take the imperfection of the initial data into account.

2.3.3 Model development

Regarding model development, the relative importance of (observation based) model initialization as well as the many existing, largely stand alone models allow for development of individual models without affecting others. Together with the short lead times, thus short run times, this potentially allows for overall short development cycles. An interesting account of a concrete development project in the context of operational SWP is given by Steenburgh et al. (2014). It addresses a wider scope than the present paper, touching for example also on visualization tools for the model results. With the rich and detailed presentation of real world issues that have to be dealt with to make a model operational, the paper makes an ideal, complementary reading to the present paper. Like the present paper, it stresses the importance of close exchange between research and operational people.

Regarding drivers of model development, model skill is again a prominent driver, despite the complicating aspects mentioned above. For the many empirical models used in and specifically designed for SWP, model skill may be the single most important development driver. Evaluation of model skill comes in the form of automated, model specific skill metrics, but also in at least some institutions in a "soft variant" via regular meetings between modelers and forecasters (Steenburgh et al., 2014). Different actors in (operational) SWP tend to use different measures of model skill, there is no wide acceptance of a best approach. There are, however, initiatives toward more unified and thus comparable model skill evaluation. The ISES network as well as the ESA SSA Space Weather Service Network, both mentioned already earlier, engage in this direction. Efforts by the Community Coordinated Modeling Center (CCMC, https://ccmc.gsfc.nasa.gov/) build on simulations of the same, real events and use of the same skill metric for all comparable models. A corresponding platform for the comparison of real-time forecasts, the predictions being uploaded to CCMC by different providers, is operational for CMEs, under implementation for flares, and planned for SEPs. The approach is interesting as it allows to identify systematic deficiencies across models. In practical terms this also means assistance of community wide model validation efforts (e.g., Pulkkinen et al., 2013; Rastätter et al., 2014; Glocer et al., 2016; Welling et al., 2017). SWP here is a trend setter. In climate, a prescribed, common skill metric is only suggested for the next inter-comparison, CMIP6 (Eyring et al., 2016b). In NWP, such a common skill metric is less straightforward and maybe less appropriate, as relevant weather characteristics are rather regional and very different in, say, Central Europe and Southern India. Common are, however, the concepts and ideas behind the skill metric. All three communities thus may profit from associated guidelines and recommendations put forward by the WMO Joint Working Group on Forecast Verification Research (JWGFVR).

Basic physics appears a key driver of operational model development for those models that are physics based and used in both research and operational SWP (e.g., WSA-Enlil, see Mays et al., 2015, for an overview). Much like in climate and NWP, the use of the same model potentially exploits synergies (man power, expertise, tools, etc.) and inspires improvements also of the operational code version. The use of a common code facilitates transfer from research developments into operation. As such a code is potentially used in widely different environments, e.g., in terms of computer architecture, operating system or file system, possibly in serial and parallel mode, this promotes overall robustness and portability of the code (Steenburgh et al., 2014).

Advances in the form of more, better, and different observational data are a definite driver of model development in operational SWP. This applies especially but not exclusively for empirical models, which rely most heavily on observations for model design and application. Changes in computer architecture, by contrast, seem less of a driver (see e.g., Feng et al., 2013). Potential reasons could be lack of free resources or that advances are expected rather from progress in physical

understanding and improved models (empirical and physics based) than from more cpu.

The relative importance of (observation based) model initialization allows for development of individual models without affecting others. Together with the short lead times, thus short run times, this potentially allows for overall short development cycles.

3 Operational modeling and development – synthesis across communities

The previous section outlined some characteristics of modeling and model development in an operational context in individual communities, notably NWP, climate, and SWP. What can be synthesized?

A first impression is that exchange between "basic research" and "operational forecasting" ideally is a two-way road with added value for both sides. From research to the operational side, there is the asset of enabling exploitation of new scientific developments in an operational context. Also, there is the research view on what a model may or may not be able to realistically capture. The opposite direction – what research might benefit from operational modeling – is less frequently highlighted. Yet it is common practice, at least in NWP and climate, that operational codes and even entire modeling chains or parts thereof are used for basic research. This allows the latter to build on a code that runs reliably, is well written and tested (verified and validated), well optimized, and comes with a version history and a wealth of useful tools, ranging from code development over validation and visualization to well defined data formats and interfaces. Also, over time the permanent operational use exposes potential weaknesses of the model much better than few dedicated tests in a pure science project. Turning again to the direction from research to operation, use of the operational code in research potentially facilitates transfer of improved functionality back to the operational code.

A second aspect concerns model skill. Skill metrics and associated scores are an ubiquitous concept. They are a driver of model development, although typically a lack of skill not easily translates into a concrete cause thereof, a concrete development project. They are attractive because they are quantitative and reproducible, thus comparable, and transparent. However, they measure only what they are designed for, and the concrete design of the metric – what quantity is evaluated, against which benchmark, and what norm is used, e.g., root mean square or other – often varies with model and/or institution. This also because models are always imperfect, and a decision has to be taken as to which aspect of the model should weigh heavier in the skill score. To compensate for the first flaw (rigid design), a "soft model skill assessment" (exchange between modelers and forecasters) forms part of the operational chain in some NWP and SWP institutions. Climate modeling is less subject to this flaw in the first place, as research and operation are much closer, on a personal level, including exchange about model performance beyond predefined skill metrics. The second flaw (model / institution specific score) is typically addressed via dedicated model intercomparison projects, e.g., CMIP mentioned in Section 2.2 or CCMC (see Sect. 2.3). Validation of models against the same skill metrics and observational data is most valuable as it allows to identify systematic deficiencies across models. For regional NWP, model inter-comparison is less practical as models are calibrated for good performance within their operational region.

Third, there is the interplay between drivers, lead time, and time scales of model development. Put simply, short lead times reflect in short model run times and frequent comparison of "prediction" and "reality". The latter can be an essential driver of model development in an operational context, provided that enough data for model evaluation are available. For NWP this is indeed the case, resulting in frequent (possibly) small changes to the operational model version (up to several times a year in NWP, see Sect. 2.1). Whether an equally high update frequency of the operational model is desirable in SWP is less clear. The short run times would allow for fast development cycles. The comparatively large number of stand alone models involved seems of little consequence as they are largely independent of each other (see Sect. 2.3). An obstacle may be observational data coverage to evaluate (long term) model performance, especially when it comes to (rare) extreme events. Also, frequent updates of the operational model imply that customers have to be prepared for slight but frequent changes of their products, similar as in the case of NWP (see Sect. 2.1). Climate modeling has much longer development cycles. There is no short time development driver in the form of daily comparison of forecast and reality. And changes in one of the sub-model components (ocean, atmosphere) can necessitate adjustments in another component, if only via "re-tuning" the model (see Sect. 2.2). A consequence of the longer development cycle is that customer products (e.g., for IPCC) tend to differ substantially between subsequent operational versions.

Finally, there are common challenges ahead, a major one being advances in computer architectures. On the positive side, these are an important factor enabling better models, better predictions. However, to take advantage of this progress, one has to cope with the fact that computer architectures, chips, and disks, follow their own evolution. It seems highly unlikely that research or operational services can exert any substantial influence on chip or disk manufacturers. Slight influence on a concrete machine procurement might be taken via the benchmark applications normally run by a supercompute center before buying a new machine. How much adaptation then is compelling or rather a matter of choice – Is a GPU code a must? Is a commercial cloud an alternative? – is a subject of much debate. Opinions also diverge on adaptation strategies, one idea being to separate codes into machine specific back ends and science code front ends. The coding and maintenance of such back ends beyond show case applications is, however, an open questions. Whether sufficiently many actors will embark on this approach and could agree to common back ends, thereby distributing associated costs, remains to be seen.

What seems certain is that complexity will grow and with it the need for expert knowledge to develop, optimize, and run corresponding well structured and modular models. In concert, there will be a growth in complexity of the hand shake between "research" and "operation", as well as between established categories, such as physics or domain science, scientific computing, and computational science. A partial answer to this increasing complexity must lie in education and community culture, promoting heterogeneous teams that cover all

necessary expertise and interact on an equal footing. An asset for such teams would be the possibility for science career tracks that are situated between or even alternate among established categories, have a long term perspective, and share the recognition of "pure science" tracks. Today, such cross-cutting careers are rather exceptional.

With the necessarily growing investments in codes and operational infrastructure, licensing issues potentially raise in relevance. Transnational collaborations and the associated mix of legal systems potentially add to the issue. Whether open source licensing could be an answer is a matter of debate.

4 Conclusion

This paper examined the somewhat conflicting demands of "operational stability" versus "dynamic development" in three different communities: NWP, SWP, and climate projections. Similarities and differences in how the three communities deal with this issue were identified. In all three communities, the lead is with the operational side, which is obliged or even legally bound to meet customer demands. The research community plays, however, an important part in an overall win-win situation. To what degree the latter is indeed recognized and appreciated by all partners is difficult to judge. Both sides have common goals – learn about dependencies in the system with the goal of making predictions and, ultimately, understand the underlying mechanism – and in pursuing them may benefit from each others strengths: clean, verified, and validated code with plenty of tools and data on the side of operational services and additional expertise, man power, time and funding for exploitative scientific studies on the research side. Mutual awareness of and respect for each other's strengths and limitations is essential for success. Leaving the community specific perspective while writing this paper, my conviction grew that mutual exchange among NWP, SWP, and climate may not always be easy but has much potential given the similarities in physics, equations, goals, and challanges ahead. Some of these challenges – notably the need for composability in view of modern computer architectures and, more generally, the growing complexity in operation and research – are not only common to all three communities but call for common action. Community building, education, but also career paths must account for these developments in order to ascertain long-term success.

Acknowledgements. I would like to thank Gianni Lapenta, KU Leuven, for inviting me to give a talk at the European Space Weather Week in 2016 on the topic of the present paper. Two anonymous referees, I would like to thank for their most constructive comments. To Christoph Schär and Martin Wild, Institute of Atmospheric and Climate Science, ETH Zurich, I am grateful for their permanent support, in this and other activities. Special thanks go to Philippe Steiner and Pirmin Kaufmann from MeteoSwiss for all the extensive discussions about NWP. Finally, I would like to thank Mickaël Melzani, Centre de Recherche Astrophysique de Lyon (CRAL), and Mark Dieckmann, Linköping University, for sharing with me their insight into plasma kinetics of magnetic reconnection and shocks. The editor thanks two anonymous referees for their assistance in evaluating this paper.

References

Araujo-Pradere EA. 2009. Transitioning space weather models into operations: the basic building blocks. *Space Weather* **7**: S10006. DOI:10.1029/2009SW000524.

Ban N, Schmidli J, Schär C. 2014. Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J Geophys Res* (*Atmos*) **119**: 7889–7907. DOI:10.1002/2014JD021478.

Barnes G, et al. 2016. A comparison of flare forecasting methods. I. Results from the all-clear workshop. *Astrophys J* **829**: 89. DOI:10.3847/0004-637X/829/2/89. 1608.06319.

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

Blaylock BK, Horel JD, Liston ST. 2017. Cloud archiving and data mining of high-resolution rapid refresh forecast model output. *Comput Geosci* **109**: 43–50. DOI:10.1016/j.cageo.2017.08.005.

Bonadonna M, Lanzerotti L, Stailey J. 2017. The National Space Weather Program: two decades of interagency partnership and accomplishments. *Space Weather* **15**: 14–25. DOI:10.1002/2016SW001523.

Buizza R, Leutbecher M. 2015. The forecast skill horizon. *Q J R Meteorolog Soc* **141**: 3366–3382. DOI:10.1002/qj.2619.

Cash MD, Biesecker DA, Pizzo V, de Koning CA, Millward G, Arge CN, Henney CJ, Odstrcil D. 2015. Ensemble modeling of the 23 July 2012 coronal mass ejection. *Space Weather* **13**: 611–625. DOI:10.1002/2015SW001232.

Chen X, Huang X, Jiao C, Flanner MG, Raeker T, Palen B. 2017. Running climate model on a commercial cloud computing environment. *Comput Geosci* **98**: 21–25. DOI:10.1016/j.cageo.2016.09.014.

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q J R Meteorolog Soc* **120**: 1367–1387. DOI:10.1002/qj.49712051912.

Cranmer SR, Gibson SE, Riley P. 2017. Origins of the ambient solar wind: implications for space weather. *Space Sci Rev* **212**: 1345–1384. DOI:10.1007/s11214-017-0416-y.

Deconinck W, et al. 2017. Atlas: a library for numerical weather prediction and climate modelling. *Comput Phys Commun* **220**: 188–204. 10.1016/j.cpc.2017.07.006.

Düben PD, Dawson A. 2017. An approach to secure weather and climate models against hardware faults. *J Adv Model Earth Syst* **9**: 501–513. DOI:10.1002/2016MS000816.

Düben PD, McNamara H, Palmer TN. 2014. The use of imprecise processing to improve accuracy in weather and climate prediction. *J Comput Phys* **271**: 2–18. DOI:10.1016/j.jcp.2013.10.042.

Elvidge S, Godinez HC, Angling MJ. 2016. Improved forecasting of thermospheric densities using multi-model ensembles. *Geosci Model Dev* **9**: 2279–2292. DOI:10.5194/gmd-9-2279-2016.

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE. 2016a. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci Model Dev* **9**: 1937–1958. DOI:10.5194/gmd-9-1937-2016.

Eyring V, et al. 2016b. ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geosci Model Dev* **9**: 1747–1802. DOI:10.5194/gmd-9-1747-2016.

Feng X, Zhong D, Xiang C, Zhang Y. 2013. GPU computing in space weather modeling. In: Pogorelov NV, Audit E, Zank GP, eds. Numerical Modeling of Space Plasma Flows (ASTRONUM2012),

vol. 474 of Astronomical Society of the Pacific Conference Series, pp. 131.

Fuhrer O, Osuna C, Lapillonne X, Gysi T, Cumming B, Bianco M, Arteaga A, Schulthess TC. 2014. Towards a performance portable, architecture agnostic implementation strategy for weather and climate models. *Supercomput Frontiers Innovat* **1**: 45–62. DOI:10.14529/jsfi140103.

Glocer A, et al. 2016. Community-wide validation of geospace model local K-index predictions to support model transition to operations. *Space Weather* **14**: 469–480. DOI:10.1002/2016SW001387.

Gysi T, Osuna C, Fuhrer O, Bianco M, Schulthess TC. 2015. STELLA: a domain-specific tool for structured grid methods in weather and climate models. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '15. New York, NY, USA: ACM, pp. 41:1–41:12. DOI:10.1145/2807591.2807627, ISBN 978-1-4503-3723-6.

Hickmann KS, Godinez HC, Henney CJ, Arge CN. 2015. Data assimilation in the ADAPT photospheric flux transport model. *Solar Physics* **290**: 1105–1118. DOI:10.1007/s11207-015-0666-3.

Horton W, Weigel RS, Sprott JC. 2001. Chaos and the limits of predictability for the solar-wind-driven magnetosphere-ionosphere system. *Phys Plasmas* **8**: 2946–2952. DOI:10.1063/1.1371522.

Hourdin F, et al. 2017. The art and science of climate model tuning. *Bull Am Meteorolog Soc* **98**: 589. DOI:10.1175/BAMS-D-15-00135.1.

IPCC. 2013. Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment Report of the intergovernmental panel on climate change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, DOI:10.1017/CBO9781107415324, www.climatechange2013.org., ISBN 978-1-107-66182-0.

Jakob C. 2014. Going back to basics. *Nat Clim Change* **4**: 1042–1045. DOI:10.1038/nclimate2445.

Knipp DJ. 2016. Advances in space weather ensemble forecasting. *Space Weather* **14**: 52–53. DOI:10.1002/2016SW001366.

Lang, M, Browne P, van Leeuwen PJ, Owens M. 2017. Data assimilation in the solar wind: challenges and first results. *Space Weather* **15**: 1490–1510. DOI:10.1002/2017SW001681.

Langhans W, Schmidli J, Schär C. 2012. Bulk Convergence of cloud-resolving simulations of moist convection over complex terrain. *J Atmos Sci* **69**: 2207–2228. DOI:10.1175/JAS-D-11-0252.1.

Lapenta G, et al. 2013. SWIFF: space weather integrated forecasting framework. *J Space Weather Space Clim* **3**: A05. DOI:10.1051/swsc/2013027.

Lathuillère C, Menvielle M, Lilensten J, Amari T, Radicella SM. 2002. From the Sun's atmosphere to the Earth's atmosphere: an overview of scientific models available for space weather developments. *Ann Geophys* **20**: 1081–1104. DOI:10.5194/angeo-20-1081-2002.

Lee CO, Arge CN, Odstrčil D, Millward G, Pizzo V, Quinn JM, Henney CJ. 2013. Ensemble modeling of CME propagation. *Sol Phys* **285**: 349–368. DOI:10.1007/s11207-012-9980-1.

Lee CO, Arge CN, Odstrcil D, Millward G, Pizzo V, Lugaz N. 2015. Ensemble modeling of successive halo cmes: a case study. *Sol Phys* **290**: 1207–1229. DOI:10.1007/s11207-015-0667-2.

Leutbecher M, et al. 2017. Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Q J R Meteorolog Soc* **143**, DOI:10.1002/qj.3094.

Leutwyler D, Fuhrer O, Lapillonne X, Lüthi D, Schär C. 2016. Towards European-scale convection-resolving climate simulations with GPUs: a study with COSMO 4.19. *Geosci Model Dev* **9**: 3393–3412. DOI:10.5194/gmd-9-3393-2016.

Lorenz EN. 1963. Deterministic non-periodic flow. *J Atmos Sci* **20**: 130–148. DOI:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Luhmann J, Mays ML, Odstrcil D, Bain H, Li Y, Leske R, Cohen C. 2015. Challenges in Forecasting SEP Events. In: AAS/AGU Triennial Earth-Sun Summit, vol. 1 of AAS/AGU Triennial Earth-Sun Summit, pp. 112.01.

Mannucci AJ, Hagan ME, Vourlidas A, Huang CY, Verkhoglyadova OP, Deng Y. 2016. Scientific challenges in thermosphere-ionosphere forecasting – conclusions from the October 2014 NASA JPL community workshop. *J Space Weather Space Clim* **6**: E01. DOI:10.1051/swsc/2016030.

Mays ML, et al. 2015. Ensemble Modeling of CMEs Using the WSA-ENLIL+Cone Model. *Sol Phys* **290**: 1775–1814. DOI:10.1007/s11207-015-0692-1.

Meehl GA, et al. 2014. Decadal climate prediction: an update from the trenches. *Bull Am Meteorolog Soc* **95**: 243–267. DOI:10.1175/BAMS-D-12-00241.1.

Michalakes J, Vachharajani M. 2008. GPU Acceleration of numerical weather prediction. *Parallel Proc Lett* **18**: 531–548. DOI:10.1142/S0129626408003557.

Molthan AL, Case JL, Venner J, Schroeder R, Checchi MR, Zavodsky BT, Limaye A, O'Brien RG. 2015. Clouds in the cloud: weather forecasts and applications within cloud computing environments. *Bull Am Meteorolog Soc* **96**: 1369–1379. DOI:10.1175/BAMS-D-14-00013.1.

Murray SA, Bingham S, Sharpe M, Jackson DR. 2017. Flare forecasting at the Met Office Space Weather Operations Centre. *Space Weather* **15**: 577–588. DOI:10.1002/2016SW001579.

Murray SA, Henley EM, Jackson DR, Bruinsma SL. 2015. Assessing the performance of thermospheric modeling with data assimilation throughout solar cycles 23 and 24. *Space Weather* **13**: 220–232. DOI:10.1002/2015SW001163.

Owens MJ, Riley P, Horbury TS. 2017. Probabilistic solar wind and geomagnetic forecasting using an analogue ensemble or "similar day" approach. *Sol Phys* **292**: 69. DOI:10.1007/s11207-017-1090-7.

Palmer T. 2015. Modelling: build imprecise supercomputers. *Nature* **526**: 32–33. DOI:10.1038/526032a.

Palmer TN. 2016. A personal perspective on modelling the climate system. *Proc R Soc Lond Ser A* **472**: 20150772. DOI:10.1098/rspa.2015.0772.

Palmer TN, Doblas-Reyes FJ, Weisheimer A, Rodwell MJ. 2008. Toward seamless prediction: calibration of climate change projections using seasonal forecasts. *Bull Am Meteorolog Soc* **89**: 459. DOI:10.1175/BAMS-89-4-459.

Pizzo VJ, de Koning C, Cash M, Millward G, Biesecker DA, Puga L, Codrescu M, Odstrcil D. 2015. Theoretical basis for operational ensemble forecasting of coronal mass ejections. *Space Weather* **13**: 676–697. DOI:10.1002/2015SW001221.

Prein AF, Rasmussen R, Stephens G. 2017. Challenges and advances in convection-permitting climate modeling. *Bull Am Meteorolog Soc* **98**: 1027–1030. DOI:10.1175/BAMS-D-16-0263.1.

Pulkkinen, A, et al. 2013. Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather* **11**: 369–385. DOI:10.1002/swe.20056.

Rastätter L, Tóth G, Kuznetsova MM, Pulkkinen AA. 2014. CalcDeltaB: an efficient postprocessing tool to calculate ground-level magnetic perturbations from global magnetosphere simulations. *Space Weather* **12**: 553–565. DOI:10.1002/2014SW001083.

Reinert D, Prill F, Frank H, Zängl G. 2017. ICON database reference manual. Tech. rep., Offenbach: Deutscher Wetterdienst, DOI:10.5676/DWD_pub/nwv/icon_1.1.12.

Reiss MA, Temmer M, Veronig AM, Nikolic L, Vennerstrom S, Schöngassner F, Hofmeister SJ. 2016. Verification of high-speed solar wind stream forecasts using operational solar wind models. *Space Weather* **14**: 495–510. DOI:10.1002/2016SW001390.

Riley P, Baker D, Liu YD, Verronen P, Singer H, Güdel M. 2018. Extreme space weather events: from cradle to grave. *Space Sci Rev* **214**: 21. DOI:10.1007/s11214-017-0456-3.

Schär C, Leuenberger D, Fuhrer O, Lüthi D, Girard C. 2002. A new terrain-following vertical coordinate formulation for atmospheric prediction models. *Mon Weather Rev* **130**: 2459. DOI:10.1175/1520-0493(2002)130<2459:ANTFVC>2.0.CO;2.

Schmidt GA, Bader D, Donner LJ, Elsaesser GS, Golaz J-C, Hannay C, Molod A, Neale RB, Saha S. 2017. Practice and philosophy of climate model tuning across six US modeling centers. *Geosci Model Dev* **10**: 3207–3223. DOI:10.5194/gmd-10-3207-2017.

Schrijver CJ, et al. 2015. Understanding space weather to shield society: a global road map for 2015–2025 commissioned by COSPAR and ILWS. *Adv Space Res* **55**: 2745–2807. DOI:10.1016/j.asr.2015.03.023.

Schunk RW, et al. 2014. Ensemble modeling with data assimilation models: a new strategy for space weather specifications, forecasts, and science. *Space Weather* **12**: 123–126. DOI:10.1002/2014SW001050.

Shaw JA, Daescu DN. 2017. Sensitivity of the model error parameter specification in weak-constraint four-dimensional variational data assimilation. *J Comput Phys* **343**: 115–129. DOI:10.1016/j.jcp.2017.04.050.

Shimokawabe T, Aoki T, Onodera N. 2014. High-productivity framework on gpu-rich supercomputers for operational weather prediction code ASUCA. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, SC '14. Piscataway, NJ, USA: IEEE Press, pp. 251–261. DOI:10.1109/SC.2014.26. ISBN 978-1-4799-5500-8.

Sibley A, Biesecker D, Millward G, Gibbs M. 2012. Space weather, impacts and forecasting: an overview. *Weather* **67**: 115–120. DOI:10.1002/wea.1915.

Simmons A, et al. 2016. Observation and integrated Earth-system science: a roadmap for 2016–2025. *Adv Space Res* **57**: 2037–2103. DOI:10.1016/j.asr.2016.03.008.

Singh AK, Siingh D, Singh RP. 2010. Space weather: physics, effects and predictability. Surv Geophys 31: 581–638. DOI:10.1007/s10712-010-9103-1.

Siscoe G. 2007. Space weather forecasting historically viewed through the lens of meteorology. In: Bothmer V, Daglis IA, eds. Space Weather- Physics and Effects. Springer, pp. 5–30. DOI:10.1007/978-3-540-34578-7_2. chap. 2.

Siuta D, West G, Modzelewski H, Schigas R, Stull R. 2016. Viability of cloud computing for real-time numerical weather prediction. *Weather Forecast* **31**: 1985–1996. DOI:10.1175/WAF-D-16-0075.1.

Steenburgh RA, Biesecker DA, Millward GH. 2014. From predicting solar activity to forecasting space weather: practical examples of research-to-operations and operations-to-research. *Sol Phys* **289**: 675–690. DOI:10.1007/s11207-013-0308-6.

Szintai B, Kaufmann P, Rotach MW. 2009. Deriving turbulence characteristics from the COSMO numerical weather prediction model for dispersion applications. *Adv Sci Res* **3**: 79–84. DOI:10.5194/asr-3-79-2009.

Taylor KE, Stouffer RJ, Meehl GA. 2012. An overview of CMIP5 and the experiment design. *Bull Am Meteorolog Soc* **93**: 485–498. DOI:10.1175/BAMS-D-11-00094.1.

Tóth G, et al. 2005. Space Weather Modeling Framework: a new tool for the space science community. *J Geophys Res* (*Space Phys*) **110**: A9. DOI:10.1029/2005JA011126.

Tóth G, et al. 2012. Adaptive numerical algorithms in space weather modeling. *J Comput Phys* **231**: 870–903. DOI:10.1016/j.jcp.2011.02.006.

Trémolet Y. 2006. Accounting for an imperfect model in 4D-Var. *Q J R Meteorolog Soc* **132**: 2483–2504. DOI:10.1256/qj.05.224.

Tsagouri I, et al. 2013. Progress in space weather modeling in an operational environment. *J Space Weather Space Clim* **3**: A17. DOI:10.1051/swsc/2013037.

Vanderbauwhede W, Takemi T. 2016. An analysis of the feasibility and benefits of GPU/multicore acceleration of the weather research and forecasting model. *Concurr Comput Pract Exp* **28**: 2052–2072. DOI:10.1002/cpe.3522.

Wang C, Rosen IG, Tsurutani BT, Verkhoglyadova OP, Meng X, Mannucci AJ. 2016. Statistical characterization of ionosphere anomalies and their relationship to space weather events. *J Space Weather Space Clim* **6**: A5. DOI:10.1051/swsc/2015046.

Watermann J, et al. 2009. Models of solar wind structures and their interaction with the Earth's space environment. *Space Sci Rev* **147**: 233–270. DOI:10.1007/s11214-009-9494-9.

Welling DT, Anderson BJ, Crowley G, Pulkkinen AA, Rastätter L. 2017. Exploring predictive performance: a reanalysis of the geospace model transition challenge. *Space Weather* **15**: 192–203. DOI:10.1002/2016SW001505.

Zhang H, Pu Z. 2010. Beating the uncertainties: ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction. *Adv Meteorol* **432160**: 10. DOI:10.1155/2010/432160.

Zhao X, Dryer M. 2014. Current status of CME/shock arrival time prediction. *Space Weather* **12**: 448–469. DOI:10.1002/2014SW001060.

Cite this article as: Folini D. 2018. Climate, weather, space weather: model development in an operational context. *J. Space Weather Space Clim.* **8**: A32