submitted to **Journal of Space Weather and Space Climate**

Application Usability Levels:

A Framework for Tracking Project Product Progress

A. J. Halford[1], A. C. Kellerman[2], K. Garcia-Sage[3,4], J. Klenzing[4], B. A. Carter[5], R. M. McGranaghan[6,7], T. Guild[1], C. Cid[8], C. J. Henney[9], N. Yu. Ganushkina[10,11], A. G. Burrell[12], M. Terkildsen[13], D. T. Welling[14,10], S. A. Murray[15], K. D. Leka[16], J. P. McCollough[9], B. J. Thompson[4], A. Pulkkinen[4], S. F. Fung[4], S. Bingham[17], M. M. Bisi[18], M. W. Liemohn[10], B. M. Walsh[19], and S. K. Morley[20]

[1] Space Sciences Department, Aerospace Corporation, Chantilly, Virginia, USA. e-mail: Alexa.J.Halford@aero.org
[2] Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, USA.
[3] Catholic University of America, Washington D.C., USA.
[4] NASA GSFC, Heliophysics Science Division Greenbelt, MD 20771, USA.
[5] School of Science, RMIT University, Melbourne, Australia.
[6] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[7] University Corporation for Atmospheric Research, Boulder, Colorado, USA.
[8] Departamento de Física y Matemáticas, Universidad de Alcalá, Alcalá de Henares (Madrid), Spain.
[9] Space Vehicles Directorate, Air Force Research Laboratory, Kirtland AFB, New Mexico, USA.
[10] University of Michigan Climate and Space department, Ann Arbor, MI, USA.
[11] Finnish Meteorological Institute, Helsinki, Finland.
[12] Space Science Division, U.S. Naval Research Laboratory, Washington, DC 20375, USA.
[13] Space Weather Services, Bureau of Meteorology, Sydney, Australia.
[14] Physics Department, University of Texas at Arlington, Arlington, TX, USA.
[15] School of Physics, Trinity College Dublin, Ireland.
[16] NorthWest Research Associates, Boulder, Colorado, USA.
[17] Met Office, Fitzroy Road, Exeter, Devon, EX1 3PB, UK.
[18] RAL Space, Science & Technology Facilities Council - Rutherford Appleton Laboratory, Harwell Campus, Oxfordshire, OX11 0QX, UK.
[19] Center for Space Physics, Boston University, Boston, USA.
[20] Space Science and Applications, Los Alamos National Laboratory, Los Alamos, NM, USA.

ABSTRACT

The space physics community continues to grow and become both more interdisciplinary and more intertwined with commercial and government operations. This has created a need for a framework to easily identify what projects can be used for specific applications and how close the tool is to routine autonomous or on-demand implementation and operation. We propose the Application Usability Level (AUL) framework and publicizing AULs to help the community quantify the progress of successful applications, metrics, and validation efforts. This framework will also aid the scientific community by supplying the type of information needed to build off of previously published work and publicizing the applications and requirements needed by the user communities. In this paper, we define the AUL framework, outline the milestones required for progression to higher AULs, and provide example projects utilizing the AUL framework. This work has been completed as part of the activities of the Assessment of Understanding and Quantifying Progress working group which is part of the International Forum for Space Weather Capabilities Assessment.

Key words. Tracking Progress – Metrics and Validation – Applied Space Weather

1. Introduction

As a field, space physics has quickly evolved beyond science inquiries and pure research. We are currently at the point where new opportunities and a need for interdisciplinary and applied space weather research have notably increased. As such, research-to-research and research-to-operations communication frameworks have become important tools. These tools expedite both multidisciplinary research and the transition of research tools to applications. It is important that, as a community, we are able to identify which research applications are ready to be transitioned into technology and provide useful information for users. It is important that the scientific community is able to clearly communicate the progress of the transition process. It is equally important to provide a measure of the usability of a research project to a user-defined application. To effectively accomplish this task, researchers need clear communication of a user's requirements, needs, and metrics for successful use of an application. In this paper, we introduce a new framework to aid in communication and collaboration of space physics research applications.

1.1. Previous Tracking Frameworks

Our new framework was developed to address needs not met by existing tracking frameworks. The most well-known example of such a tracking framework is the Technology Readiness Level (TRL) system, which categorizes the maturity of a particular technology and its use for instrumentation (e.g., Mankins, 1995, 2009; Azizian et al., 2011; Olechowski et al., 2015, and references therein; European Space Agency (ESA) TRL definitions can be found at http://sci.esa.int/sci-ft/50124-technology-readiness-level/). The clear, consistent definitions of the TRLs allow the readiness of an instrument for a specific use to be determined independently and in comparison to currently available options.

NASA's Earth Sciences division's applied science program employs the Application Readiness Level (ARL) framework. The ARL framework is used to communicate how "ready" a given model or data analysis effort is for a particular utilization and industry partner. The framework's focus on research and development involves the industry partner at the project's start (Pulkkinen et al., 2017, see Figure 1, or https://www.nasa.gov/sites/default/files/files/ExpandedARLDefinitions4813.pdf). This includes forming the application requirements around the user needs. This framework has aided in the identification of obstacles. It is also used to assess programmatic health by comparing a project's progression through the framework against the distribution of funding across projects at different levels(e.g., see the Earth Sciences Division, Applied Science Programs Annual Reports: https://appliedsciences.nasa.gov/library-page).

Both ARLs and TRLs use single-digit level identifiers. They are used to communicate the advancement of applied products to the scientific, engineering, and funding communities. Though the two frameworks are designed for different types of products, the identifiers communicate similar information. The level identifiers enable researchers, funding agencies, and users to easily interact with each other and communicate progress towards the routine usage of these products.

1.2. A New Tracking Framework for Heliophysics

The existing frameworks each focus on tracking a particular type of product, and so do not fully meet the needs of the heliophysics community. Space physics products include observational data, derived indices, modeled outputs, and more. These products are often used together for different purposes. Each user will have different requirements for the application in terms of the type of product, robustness, and accuracy.

The unique needs of the space weather community led to the modification of existing research-to-application communication frameworks to create the Application Usability Level (AUL) framework. Applying AULs to model and data analysis efforts can benefit space physics research. These benefits include improving access to collaborators, project transparency, and communication of project results. As the requirements and user interests for each application are unique, the AUL framework uses specifically-tuned metrics. For instance, a research user interested in upper atmospheric coupling may want to know the flux and characteristic energy of precipitating electrons. Similarly, a satellite industry partner may want to predict satellite drag during a geomagnetic storm. A single research project may be able to provide both users with the products they need. However, the different outputs will require different metrics, implementation strategies, and time frames for implementation. Since the AUL framework is highly adaptable, it can help a single research project meet and track both of these user needs.

The AUL framework can bolster communication between researchers, users, funding bodies, and stakeholders. Using a standard framework provides a clear path for users and researchers to follow. This improves efficiency assuring that all components from the researchers' project to the user needs are considered. It enables communication about a proposal's development status, requirements for further progress, and achievable goals. Improved communication leads to better-targeted funding opportunities and proposals. The AUL framework can simplify comparisons of different projects working towards a specific application. This enables operational and funding services to select the most appropriate proposal for their requirements. It can also highlight gaps in knowledge, data, and technology, to aid the characterization of needs for new missions, instruments, and research or model development proposal calls.

In this article we introduce the AUL framework, a new heliophysics-focused research-to-application framework for communicating the usability and readiness of a product to its user. The AUL framework is constructed similarly to the existing TRL and ARL frameworks. The framework terminology is described in section 2. The details of the AUL framework are provided in section 3. Section 4 provides examples of AULs at different usability levels. The potential impact of using AULs is described in section 5. Full examples and tools to help aid in the adaptation of AULs are provided in the appendix and supplemental information.

2. Framework Terminology for Targeted Research

The AUL framework can be applied to any project where an expected outcome is the ongoing use of a product by another party (targeted research). In a very general sense, the framework can be applied as described in subsection 2.1. This paragraph contains italicized words that we have identified and defined in subsection 2.2. These definitions are included to avoid the confusion commonly encountered in interdisciplinary, multidisciplinary, and transdisciplinary projects. One of the strengths of AULs is its ability to enable communication between different groups. This includes researchers in the same field (unidisciplinary or interdisciplinary), scientists across disciplines (multidisciplinary or transdisciplinary), and with industry partners (transdisciplinary or applied).

2.1. AUL Framework General Use Example

We, a group of researchers, have a *project*. We believe that there is an *application* that it can be used for, and have determined that we will use the *AUL* framework and its *phases* to communicate the progress of this *project* towards a specific *application* to the specified *user*, as well as to the scientific community. First, we will identify and reach out to a potential *user* who might be interested in routinely using the *project*'s *product*. We then determine if the *project* is *viable* based on the *user*'s *requirements*. If it is, then we continue by defining *metrics* for *verifying* the *viability* and *feasibility* for the specific *application* with the *user*'s *requirements* in mind. If the project is not deemed *viable* or *feasible*, then the current *project* should be re-examined and potentially held off on until it is deemed both *viable* and *feasible*. As the *project* continues, dissemination of the progress, *metrics*, and *validation* efforts should be reported to both the *user* and the scientific community. Once the *product* is *validated* and demonstrated to work within the *relevant context*, it is *transitioned* over to regular on-demand use. *Validation* and *verification* efforts continue, now focusing on sustained usage in the *operating environment*. Each step in this development effort towards application readiness is given an *AUL* level number. For every step in this process, an *AUL* number is designated to denote and communicate the current state of the *project*.

2.2. AUL Terminology

Application - A specific use for a project, such as a data product from a mission, a service such as satellite hardware anomaly assessments, or a forecast of a specific quantity from a numerical model. Each application has its own unique requirements and metrics for validation. For instance, an application may be forecasts of surface charging events that have an 85% success rate and less than a 5% rate of false alarms.

Applied Research - Research pursued with a focus on providing a practical application or new technology, often in an operational environment.

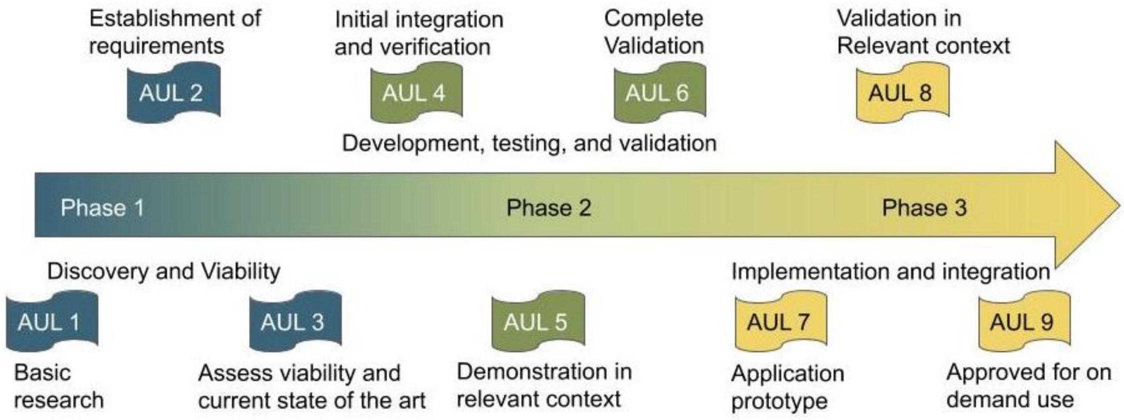

Fig. 1. Application Usability Level (AUL) diagram. The progress of a project towards an application moves from AUL 1 to AUL 9, passing through three main phases: discovery, development, and implementation. Each of these are described in the main text.

AUL - The Application Usability Levels (AULs) are the scale that tracks the progress of work on a given project for a specific application, as summarized in Figure 1 and Table 1. More details about the three phases and nine levels of the AUL framework are found in Section 3.

Feasibility - The ability to achieve success with the available resources.

Metric - A quantitative measure of project or application performance. When applied to project progress, this constitutes measures that define whether the project meets its goals and milestones. When applied to applications, metrics consist of quantities appropriate for measuring performance, such as accuracy, bias, or skill score.

Operations to Research (O2R) - Through the process of transitioning targeted or applied research to operations, new phenomena or discoveries can be found and inform subsequent research projects. This part of the feed back process is refereed to as Operations to Research (O2R)

Operational Environment - The conditions in which the application will be used. For example, a geophysical research model that will be delivered to the Community Coordinated Modeling Center (CCMC) for on-demand runs will define their operation environment as the computer system used by CCMC.

Phases - AULs are grouped into three phases: 1) discovery and viability; 2) development, testing, and validation; and 3) usability, final validation, and implementation. Each of these phases is discussed in more detail in the following sections and summarized in Figure 1 and Table 1.

Product - A project outcome that is routinely used to enhance the decision making process of a user or provide input to another research project or application.

Project - A research or development initiative designed to make progress towards a single application. Examples of projects include the development or modification of models, new uses for available data, using a data product to improve decision making, and using current knowledge to develop future projects.

Research to Research (R2R) - A targeted research project where both the 'researcher' and 'user' are researchers who may be in the same sub-field, or in completely different disciplines.

Research to Operation (R2O) - A targeted or applied research project which takes a research application and transitions it into the operational environment.

Requirements - The set of necessary conditions outlined by the user, which may include metrics, time frames, and operational environments that the project must meet for the resulting application to be considered successful.

Relevant context - The environment in which the project must be validated and verified (e.g., during geomagnetic storm periods or in interplanetary space).

Targeted Research - Investigations pursued with a specific objective.

Transition - The process or set of activities that take a product or service from a testing environment and move it to an operational environment.

User - The anticipated person or group who will make use of or operate the project's application. This may be another researcher, broker, or industry partner. Other common terms for user, appropriate for different fields, include 'end user', 'forecaster', 'customer', or even 'another collaborator'.

Validation - The determination of the skill of the project's outputs, quantified by identified metrics for the defined operating environment and relevant context.

Verification - The determination that the product conforms with the project requirements, as described in the relevant design documents.

Viability - The project's value or level of return for the researchers and users.

Table 1. A brief description of the AUL phases and levels

Phase	Phase definition	AUL	Level description
Phase 1	**Discovery and Viability**	1	Basic research
		2	Establishment of users and their requirements
		3	Assess viability and current state of the art
Phase 2	**Development, Testing, and Validation**	4	Initial integration and verification
		5	Demonstration in the relevant context
		6	Completed validation
Phase 3	**Implementation and Integration into Operation**	7	Application prototype
		8	Validation in relevant context
		9	Approved for on-demand use

Terms may have different definitions within different communities. For example, the terms *validation* and *verification* which can be particularly confusing. Their definitions differ between the operational and modeling communities. The operational community typically uses *verification* to mean that a system meets end-user requirements and uses *validation* to mean that the system is right for the project. The modeling community typically uses *verification* to mean that the code operates correctly and uses *validation* to mean that the results are accurate. For this paper, we have settled on the listed definitions based on commonalities between the two communities. Both identify *verification* as meeting basic requirements and *validation* as the appropriateness of the result within a given environment.

3. The Application Usability Level (AUL) Framework

The new AUL framework benefits research by providing a structured approach for tracking the progress of a project towards an application. Once the needs of a specific application have been defined, the same metrics may be used by the community to assess the progress of several projects towards the same application. This allows for easy comparison of projects and provides insight into a project's progress. In this section, we will define each of the three phases in the framework, the individual levels that make up each phase and step through the necessary milestones to achieve a given AUL. A summary figure of the AUL levels and phases can be seen in Figure 1 and is outlined in Table 1. A checklist for the three phases is provided in the supplemental information. More resources can be found through our team website at https://ccmc.gsfc.nasa.gov/assessment/topics/trackprogress.php.

3.1. Phase I: Discovery and Viability

Phase I is where fundamental research meets applied science. Not all research will or should progress beyond the very first AUL. However, if a potential user is identified (whether they are a fellow researcher or an industry partner), then the steps in this phase will determine whether the project should progress to phase II. Phase I provides researchers with justifiable confidence that their work is leading to a product for a specific application that will provide the user with the information they need. The first step in the progression of any project towards an application requires contacting a potential user to begin forming a partnership. The user's needs must be established through effective and clear communication about the requirements and metrics of the application. The researchers must determine if the project will be viable and feasible to satisfy the requirements for that specific application. It should be determined whether the current project represents the current standard or an improvement upon the state-of-the-art for that specific application.

3.1.1. AUL 1: Basic Research

This level is where the basic scientific concepts and projects are created and potential applications are identified. A project is considered to have an AUL 1 if the following milestones are achieved:

Milestones:

a) Basic research is documented and disseminated for the project, so that the usability may be assessed by way of the AUL method.

b) Ideas for how the project output(s) may enhance decision making or be applied to an end user application are generated.

c) Potential interested users are identified, but not necessarily contacted. This could occur, for example, through a literature search, conference attendance, or workshop participation.

3.1.2. AUL 2: Establishment of users and their requirements for a specific application

In this level, the application and project concept is formalized. An interested user is contacted, and their needs for a specific application are identified. This includes establishing requirements and defining metrics for measuring success. It should be noted that at this level, it is not necessary to show that the current research endeavors will result in the successful production of the identified application. A project is considered to be at AUL 2 if the following milestones are achieved:

Milestones:

a) Decide on the user(s), contact the user(s), and establish a reliable channel of communication that is used at a suitable frequency.

b) Formalization of the application and project concept.

c) Identification and formalization of the requirements and metrics necessary for successful application of the project for the user's needs.

3.1.3. AUL 3: Assess viability of concept and current state of the art

To reach AUL 3, the feasibility and viability of achieving success for a specific application as defined in AUL 2, must be assessed by both the users and researchers. Building upon a proof-of-concept study, the requirements for project improvement and chosen metrics are re-examined and updated as needed. A demonstration that the project represents the state-of-the-art, an improvement, or value-add to the state-of-the-art is completed. A project is considered to have an AUL 3 if the following milestones are achieved:

Milestones:

a) Documentation and dissemination of the project's expected advancements from the current state-of-the-art used towards the identified application along with the proposed metrics for the specified application.

b) Perform the initial analysis of the individual project components, to determine the viability and feasibility of the entire project.

c) Complete a detailed characterization of the baseline performance and limitations with respect to the application.

d) Determine the viability and feasibility of the proposed project towards improving upon the state of the art for the identified application. If the project is deemed not viable or feasible, the project is put on hold until the identified roadblocks are removed.

3.2. Phase II: Development testing and validation

In Phase I, the current state of the art is identified, basic research into current limitations and expected areas for improvements is completed, initial communication with the end user is established, and a proof-of-concept and show of viability is made. Phase II focuses on finalizing development of the new state-of-the-art project integrating the resulting tools into the identified applications, demonstrating the feasibility of the new product and validating the new system.

3.2.1. AUL 4: Initial integration and verification

In this level, the basic prototype is completed and initial integration into the user application is started. To achieve AUL 4, it must be verified that all components work together. In addition, a project is considered to have AUL 4 if the following milestones are also achieved:

Milestones:

a) Integration of the individual components into the application.

b) Organizational challenges and human process issues (if applicable) are identified and managed.

3.2.2. AUL 5: Demonstration in the relevant context

In this level the viability of the project is determined for the specified relevant context (e.g., storm, substorm, or quiet time conditions). A project is considered to have AUL 5 if the following milestones are achieved:

Milestones:

a) The project team must articulate and disseminate the viability for the improvement upon the state of the art.

b) Application components integrated into a functioning application system for use during the given relevant context parameters.

3.2.3. AUL 6: Complete validation

While in AUL 5 the potential is articulated, in AUL 6 the potential is fully demonstrated, and this is stated as a major increase in the applications usability and ability to become the new standard for the user. Any application components already deployed in the user's operating environment are tested in their operational and/or decision making context. A project is considered to have AUL 6 if the following milestones are achieved:

Milestones:

a) Prototype application system beta-tested in a simulated operational environment.

b) Projected improvements in performance of the state-of-the-art and/or decision making activity demonstrated in simulated operational environment.

c) Documentation and dissemination of the specific application and associated metrics and the projects progress towards this application.

3.3. Phase III: Implementation and Integration into operational status

While Phase I and II focused on the development and initial validation of the new model/data analysis effort for a specific application, Phase III is where it becomes handed off and fully integrated into the end user's application. This also includes new validation efforts to determine how well the new model/data analysis effort performs in a "real world" setting. Validation and continued use in an operational environment drives discovery of new science questions, problems, and new applications. Although this is the final phase for the current application with its specific requirements and metrics, the search for the next new and improved application continues.

3.3.1. AUL 7: Application prototype

All portions of the new project are integrated into the user's application and the functionality has been established. A project is considered to have AUL 7 if the following milestones are achieved:

Milestones:

a) The system must be fully integrated into the operational environment specified by the user.

b) The system's functionality is tested and demonstrated in the user's specified relevant context.

c) Project team must demonstrate the functionality of the new system for the user's application and disseminate the results.

3.3.2. AUL 8: Validation in relevant "real world" environment

At AUL 8, the new project is fully integrated into the user application system and is initially validated by the user. The application is proven to work in its final form in the relevant context and operational environment either meeting or surpassing the initially identified requirements and metrics. In addition, user documentation including verification and validation/metric results, any limitations of the new project, training documentation, and maintenance documentation are completed. Ideas for future developments are documented. A project is considered to have AUL 8 if the following milestones are achieved:

Milestones:

a) The user must approve the addition of the new project to their operational application system.

b) Finalized application system tested, proven operational, and shown to operate within the specified requirements and metrics.

c) Applications qualified and approved by the user.

d) User documentation and training completed.

3.3.3. AUL 9: Approved for on-demand use towards stated application

At AUL 9, the project is the new state of the art and has been proven to work in a sustained manner. Continued validation efforts, completed by the user(s) and likely in concert with the researcher(s), are performed for the project's sustained use in the operational environment. A project is considered to have AUL 9 if the following milestones are achieved:

Milestones:

a) Sustained and repeated use of the application by the specified users.

b) The continued validation of the project in the operational environment.

c) Dissemination of the validation efforts, metrics, and new state of the art project to the relevant community for the specific application.

3.4. The next application

The AUL outline and process shows the path and progress from research to application. It does not as explicitly show the feedback from application or operations to research. As with all research, there are always new questions and better tools and methods which become available and allow for improvements, better forecasting, smaller error bars, etc. Once a project has reached AUL 9, there are often new areas for improvements which have been identified, new science questions which were uncovered during the process, and of course, new potential applications realized through working with the users. This framework is less of a line and more of a set of branching trees as shown in Figure 2 or a spiral as shown in Figure 3. In sections 4 and A, the AUL framework is applied to cutting edge research and R2O processes. In many of these examples, refinements of the metric criteria and completely new applications or users are identified. Thus, with each new application (along with specific metrics and operating criteria), the model or data product begins a new AUL branch.

3.5. Dissemination of Results

Throughout the AUL framework there are 6 milestones that require the dissemination of results. AUL 1 requires that the basic science has been documented. AUL 3 requires that the expected

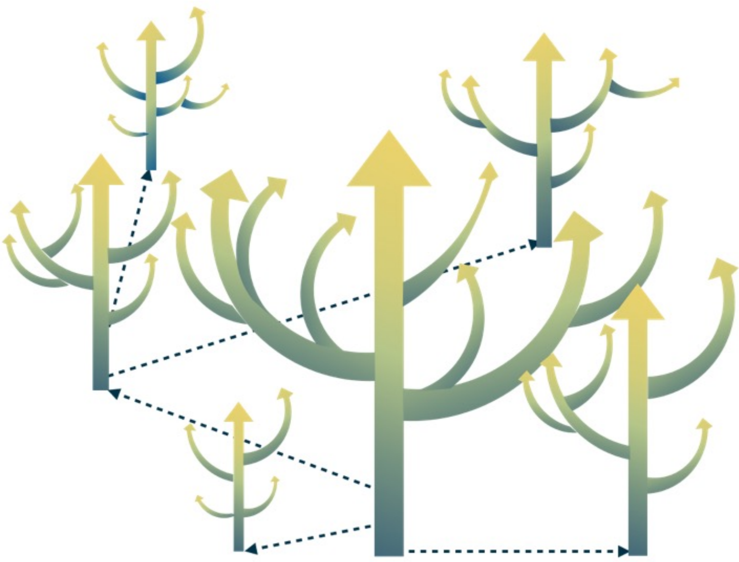

Fig. 2. Each AUL can spawn many branches through working with users and whole new trees which may remain connected through their common roots, the same basic research (like an aspen grove) as suggested in Sections 4.1 and 4.4. These new applications may be identified at any stage through the process and will have their own users, requirements, and metrics and will progress through the AUL framework at their own pace.

advancements are shared while AUL 5 requires documentation of the viability for improvement upon the state of the art. All three of these milestones help determine the viability and feasibility of the proposed project to address the needs of the user. The finalized metrics and requirements of the specific application are documented in AUL 6, and in AUL 7 the new functionality of the project in the user environment is shown. Continued validation, and how the project is the new operational state of the art is documented in AUL 9.

Peer Reviewed Papers: Although it may depend on who the researchers and users are, one natural method for dissemination of the project's progress is through peer-reviewed papers. For many projects, this would be the ideal method for completing the relevant milestones in AULs 1, 6, and 9. This has two primary benefits; 1) the assessment of the new AUL standing is reviewed by an independent assessor and 2) it advertises the application and the advancements of the project. In order to help facilitate the adaption of instrument-like or "AUL papers", we have provided both LaTeXand Microsoft word templates on our team website at the CCMC and as of this publication can be found at https://ccmc.gsfc.nasa.gov/assessment/topics/trackprogress.php. We expect that an AUL paper would be a concise piece, outlining the application, the researcher, the user, and the requirements along with the relevant advancements through the milestones. It should also include references to previous AUL papers or other research actives (papers, conference talks, ect.) to show previous progress to the current AUL as well as any new milestones met. These templates may also guide the introduction of AULs into more traditional paper formats.

The above paragraph assumes that the researcher would be writing the AUL papers, however, it has been proposed that the user may also be interested in writing such papers. One of the current hurdles in completing targeted, transdisciplinary, or applied research is identifying who may have

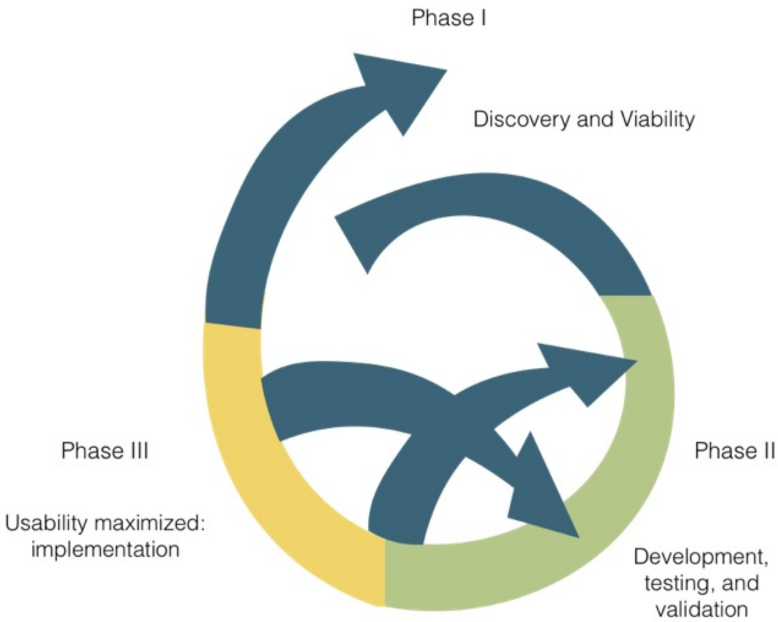

Fig. 3. While working through an AUL, roadblocks can appear which will temporarily lower the AUL of the project, similar to how a change in an instrument project can lower the TRL of a hardware project. An example of this would be a component of a large model may have changed and need to be validate, or a change to the processing of the data which is an input to the project. Anecdotal experience has shown how reaching an AUL 9 does not mean the end of a project. By the time an application reaches AUL 9, new requirements are often defined and a new application and project are created. This may be as simple as the same user has identified the need of new requirements and metrics for a new (perhaps an improvement on the same, but now identified as a new) application, leading to a spiraling of the AULs, as suggested and demonstrated in Sections 4.3 and 4.7

the required research and who may benefit from the research. It has been suggested at workshops on AULs that users may be interested in writing AUL 1 papers as a call for help finding interested researchers to address their needs.

Conference Talks and Posters: Presenting results and advancements at conferences is another method for disseminating results and advancements through the AUL framework. For many projects, this would be the ideal method for completing the relevant milestones in AULs 3, 5, and 7. Often it is expected that these advancements would be shared within the appropriate science sessions. However, the Assessment of Understanding and Quantifying Progress working group has convened sessions at the Geospace Environment Modeling (GEM) workshop, the American Geophysical Union Fall meeting, et cetra convened specific sessions that have focused on sharing examples of AUL progressions, and discussions of research to operations and transdisciplinary research. We encourage and support the idea of more sessions at conferences focused on reporting on new applications, advancements of projects, and lessons learned.

Websites and Online Documentation: Peer-reviewed papers and conference presentations are not the only methods for disseminating results, although they are perhaps the most familiar to researchers. The Assessment of Understanding and Quantifying Progress working group which is part of the International Forum for Space Weather Capabilities Assessment has been developing a website for the dissemination and tracking of applications and projects. The website will host applications with the specified end users and their requirements and the projects working towards the specific application. This will also allow for new researchers to find applications and their requirements which may be applicable for their research. It will also enable users to find applications relevant to their need and researchers who may be able to address their application needs.

3.6. Best Practices

The above framework was purposefully written in a general format so that it can be adapted to the needs of the researcher, the user, and the specific project. However, many projects will have a set of best practices (e.g., software (Burrell et al., 2018, e.g.,), data set production (Wilkinson et al., 2016, e.g.,)) and below we will discuss a few which will be important for many if not all projects and applications.

Training: In AUL8, milestone d) training must be completed. This is a very important step in making sure that the user understands the full capabilities of the product, and will use the product properly. It is necessary to train not just on the typical types of events or runs expected, but to also train for the less likely or extreme scenarios. This is especially important for applications related to space weather forecasting where training should occur on both real-time data, and the range of events where different decisions would need to be made.

Determination of availability of data and robustness of the system: When looking at the feasibility of a project, it is important to consider the impact of data outages. This can be mitigated by having redundancy in the system. It is important to test the robustness, the likelihood of an outage in order to determine if the project is feasible.

Version Control: As data, codes, and programming languages evolve, version control is vital. Version control ensures that stable releases are available, while further testing and development continue. As new data, models, ect. are developed, they will also go through the AUL framework to show their advancement to the application, and validate their usage.

Continued testing of operational availability: In phase 3, best practices dictate that project validation and the determination of a project's operational availability continue.

Standardized formats: There are likely to be either multiple projects working towards a single application (e.g., the CME arrival scoreboard, or for data management FAIR (Wilkinson et al., 2016, e.g.,)) or multiple projects which use similar data types. In order to make it easy for others to adopt the outputs of the projects, or use the data for inputs, it is important to use data and meta-data standardized practices. There are many groups working towards defining the appropriate set of standardizations including the Information Architecture group through the International Forum for Space Weather Capabilities and Assessment.

These are just a few of the best practices one needs to keep in mind (for example, software best practices are outlined in Burrell et al., 2018), but they largely fall under the AULs which are focused on determining either the feasibility and viability of the project, or in the transitioning of the project from the researcher to the user. It is important for both the researcher and user communities to continue to evaluate what best practices are needed and include them when completing the relevant milestones.

4. Summary of Example Projects using AULs in the Appendix

Here we will present a summary of examples that cover many different aspects of the heliosphere and how the AUL framework can be applied. Extended versions of these examples can be found in the appendix A. For each example in the appendix, we'll reflect on how the AUL framework could be used to assess the progress of each project towards the specified application. Many of the projects were initially developed without the AUL framework, however, they all exhibited best practices and hence we can use the AUL framework to measure their progress, and point out how and where it may be useful in each case. The Phase I Example 4.1, describes a research area which has identified multiple potential applications but has not yet identified a user. Example 4.1 also provides a look at how to start the AUL process and targeted research more generally. The Phase II examples include Example 4.2 has a research user whereas Example 4.4 which has a government agency as the user and Example 4.5 has an industry user. Example 4.3 shows how as a project moves through the AUL levels, new applications can be found by a change in the requirements of the user. The Phase III examples include Example 4.7 where one can see how an AUL 9 project often leads to the identification of a new application and thus a new project at AUL 1. Example 4.6 shows that there may be many different user communities interested in the product and application. And finally, Example 4.8 shows how a project may not be finished through a single funding opportunity, and in this specific example, each phase in the AUL framework was funded through a separate opportunity.

For the ease of the reader, we have provided a table of the examples, a brief summary of the project, phase, and user, which section they are in, where the longer version of the example can be found (Table 2).

4.1. Identifying a potential new application to track with the AUL framework: Phase I AUL 1 project

J. Klenzing and A. G. Burrell

The following is a summary of a phase 1 AUL 1 example from the ionospheric community. This example shows a data project, and how one can start using the AUL framework. A full version of this example can be found in Appendix A.1

The use of average solar extreme ultraviolet (EUV) as a potential driver of ionosphere-thermosphere (I-T) models is an example of an early-phase AUL 1 project. This radiation heats the thermosphere and creates the ionosphere through direct ionization. I-T models use proxies for this radiation, such as Sunspot Number or the $F_{10.7}$ index. These proxies are used in part because they are long-running history and continuity. However, observations during the recent solar minimum have suggested that their utility may not extend to periods of extremely low solar activity Emmert et al. (2010); Klenzing et al. (2011); Solomon et al. (2013). Additionally, their variability over the 27-day solar rotation cycle shows significant deviation Chamberlin et al. (2007).

This project is classified at AUL 1 as it uses existing published scientific knowledge to present a new idea for improving a specific group of space weather applications. To advance to an AUL

Table 2. Table of examples given in the paper

Example	Phase	AUL	Research Sub-field	Primary User	longer example
4.1 Identifying a new application	1	1	Ionosphere	Researcher models	A.1
4.2 Application for another researcher	2	5	Ionosphere	The AMIE model	A.2
4.3 Branching applications	1, 2	2, 5	Magnetosphere	External Business	A.3
4.4 Transitioning a research model to a government user	2	6	Ionosphere	Australian Bureau of Meteorology	A.4
4.5 Validating in an operational environment	2	6	Magnetosphere	Industry/government	A.5
4.6 Identifying new transformative research by working with the user	3	8	Solar	Government/Air Force	A.6
4.7 Identifying new applications	1, 3	1, 9	GICs	Red Eléctrica de España, REE	A.7
4.8 Funding applications through the three phases	3	9	Magnetosphere	British Antarctic Survey	A.8

of 2 or higher, the project developers need to work with I-T model developers to determine if an improved solar EUV forcing index is viable and feasible for improving specific applications (for example, satellite drag calculations) and specify the level of improvement required for the application. Multiple AUL 2 applications could be identified from this AUL 1 project, Figure 2, each with different requirements.

4.2. An application for another researcher: Phase II AUL 5 project

R. M. McGranaghan

The following is a summary of a phase 2 AUL 5 example from the ionospheric community. This example shows a modeling project which has another research team as a user. A full version of this example can be found in Appendix A.2

In this example the user is a researcher using the Assimilative Mapping of Ionospheric Electrodynamics (AMIE) Richmond et al. (1988) which requires an ionospheric conductance model to infer global polar maps of electrodynamic variables on a roughly 1.5°×10° latitude×longitude grid at variable time resolution, where the time resolution is dependent on the cadence of input observations. As researchers in the ionospheric and magnetospheric communities need event specific outputs of electrodynamic fields run on demand, the operating environment is the researcher's local computer and the relevant context changes for the specific research application. The metrics needed for this application has been outlined in Cousins et al. (2015) and McGranaghan et al. (2016), in which the accuracy of the conductance model is determined by the extent to which it provides consistency between AMIE output using two different sets of input observations.

All milestones through AUL 5 have been completed. Because current efforts are underway to demonstrate the potential performance improvement to the AMIE model provided by the conductance model, requirements for AUL 6 have not yet been completed. Therefore, this application is currently at AUL 5.

4.3. Branching applications identified through continued communication with users during development: Phase I and II projects

A. C. Kellerman

The following is a summary of a phase 2 AUL 5 example from the magnetospheric community. This example shows a modeling project which has an industry user. It also serves to demonstrate that while a project may be at a higher AUL for one application, even a small change in the needs of the end user necessitates that the AUL be reassessed, and likely reverted back to an earlier level (see Figure 3). It is then necessary to work through each level again to ensure that the new needs can, and will be met in a systematic and robust manner. A full version of this example can be found in Appendix A.3.

A fast real-time data assimilative framework has been developed, to produce a hindcast/forecast of the Earth's radiation belt electron dynamics. In 2016, contact was made with an external business that was interested in utilizing the real-time hindcast data to provide a real-time tool for determining if recent spacecraft anomalies were a result of space weather. A personal meeting was set up and the needs/requirements were discussed with the interested party. In this initial interaction, the need was to provide a regular output of the hindcast and forecast electron PSD and to provide a full explanation of the content and format of the available files. Through these interactions, all milestones through AUL 5 would be satisfied. However, since complete validation has not been completed and disseminated to the community, this project would not have reached AUL 6.

After further discussions with the user, there was a new requirement that errors be included with the hindcast and forecast. Since this information was not readily available at the time, the project would be assessed at AUL 2, as the same user and communication channels exist, the project has a new formalized application, and there are new requirements that address the needs of the user. Note that a revision of the users' needs necessitates the definition of a new application, as the metrics/requirements of the project have changed, and one may no longer compare the project AUL for this new application with the previous one.

4.4. Transitioning science to a government user: Phase II AUL 6 project

B.A. Carter and M. Terkildsen

The following is a summary of a phase 2 AUL 6 example from the ionospheric community. This example shows a modeling project which a government user. This example also shows how one can work towards completing higher milestones while working to overcome others at lower levels. A full version of this example can be found in Appendix A.4.

A series of recent and ongoing studies into modeling the occurrence of EPBs using the Thermosphere Ionosphere Electrodynamics General Circulation Model (TIEGCM) – a global coupled 4-D model of the ionosphere-thermosphere system (Qian et al., 2014, and references therein) – are discussed here in the context of the AULs where the user is Australia's Bureau of Meteorology (BoM). The Bureau of Meteorology's Space Weather Services (BoM-SWS) is Australia's sole provider of space weather products and forecasts. Therefore, an ongoing collaboration with BoM-SWS has helped bridge the communications and knowledge gaps between researchers and potential users of scintillation forecasts. This project is currently at an AUL 6 and has completed all previous milestones.

The Raleigh-Taylor (R-T) growth rate threshold was chosen by eye to be $0.4 \times 10^{-3} s^{-1}$ for all six stations, which may not be appropriate for stations in different longitude sectors. One further complication is that the R-T growth rate is not a measurable quantity, so it cannot be directly verified against direct observations. As such, a more rigorous and systematic analysis that investigates the TIEGCM R-T growth rate threshold is needed before the Milestones of AUL 6 are achieved.

Initial work is being done towards achieving AUL 7. The scintillation forecasting scheme used by Carter et al. (2014b) in an operational environment, with the intention of providing 'beta' scintillation forecasts for key users. Proceeding into AUL 7 will help with the challenge of verifying the scintillation forecasts and advisories in terms of user experience. Proceeding with the development of a working prototype and delivering forecasts has the added benefit of informing/educating the users of potential vulnerabilities to their system(s).

4.5. Validating in an operational environment for multiple users, industry and government: Phase II AUL 6 project

T. Guild

The following is a summary of a phase 2 AUL 6 example from the magnetospheric community. This example shows a modeling project an industry and government user. It also provides an example of how a state of the art and "operational" tool may not be at an AUL 9. A full version of this example can be found in Appendix A.5.

The SEAES tool grew out of a need to quickly assess the likelihood of the space environment causing a satellite anomaly. It has been developed at The Aerospace Corporation Koons and Gorney (1988); O'Brien (2009). The SEAES algorithms produce a hazard quotient. This environment/anomaly likelihood relationship is derived from associating historical anomalies or their proxies to space environment measurements on the same satellites. A key user requirement for SEAES is speed: providing a hazard quotient, or likelihood that an anomaly is due to space weather, in near-real-time to influence decisions made during satellite anomaly investigations.

SEAES completed its Phase 1 milestones working closely with satellite operators during anomaly investigations where the space environment's role needed to be determined. The interaction with users informed the development of a prototype application, outlined the requirements, and culminated in a published description of the algorithms in O'Brien (2009). The SEAES algorithms have been implemented in an operational environment at NOAA/SWPC and are available to SWPC users at the following link: https://www.swpc.noaa.gov/products/seaesrt. This completes Phase 2, through AUL 6. However, the SEAES application has never been validated the user environment, nor have the results been disseminated, failing AUL 7 and above.

4.6. Transformative and translational research identified by the needs of the user: Phase III AUL 8 project

C. J. Henney

The following is a summary of a phase 3 AUL 8 example from the solar community. This example shows a modeling project for a government user. This example also shows how two projects when combined for a new application initially start at an AUL1 and must advance through all AULs with the new requirements. A full version of this example can be found in Appendix A.6.

The ADAPT (Air Force Data Assimilative Photospheric Flux Transport) project Arge et al. (2010, 2011); Hickmann et al. (2015) provides a sequence of best estimates of the instantaneous global spatial distribution of the solar photospheric magnetic field as a function of time. Initiated in 2008 and driven by community user interests, the objective of the ADAPT project combined two Phase II (AUL 5) projects Worden and Harvey (2000); Hickmann et al. (2015), to produce global magnetic maps with realistic estimates of the uncertainty (Milestones a-c of AUL 1). An essential element during development of ADAPT has been the vital feedback and collaboration with active users (Milestones a-c in AUL 2) to assess the viability of the global maps (Milestones a - d AUL 3). Another set of fundamental steps was to integrate the ADAPT software within a prototyping environment (Milestones a-b in AUL 4), iterate on map quality, and meta-data improvements (Milestones a-b AUL 5). The ADAPT model has been running at the National Solar Observatory for 5 years, generating public global magnetic maps for user validation (Milestones a-c AUL 6). The core functionality is installed at NOAA/SWPC (early stages of AUL 8 a-b) to be validated for driving WSA-Enlil, in collaboration with the CCMC.

4.7. Identifying new applications and research projects from previous targeted research: Phase III AUL 9 project

C. Cid

The following is a summary of a phase 3 AUL 9 example from the ground induced current community. This example shows a data project for a government user. This example also shows how through completion and continued validation of an AUL 9 project can lead to the funding of new targeted basic research and AUL 1 projects. A full version of this example can be found in Appendix A.7.

After the October 2003 Halloween Storm, affecting electric utilities in South Africa Kappenman (2005), low, and mid geomagnetic latitude countries were made aware that their power grids might be vulnerable to this hazard. The magnitude of the Spanish geomagnetic latitude is similar to South Africa. This fact was the impetus for a chain of projects that concluded with the development of a new index for nowcasting geomagnetic disturbances and the geomagnetically induced currents risk in Spain.

This project has fulfilled all the milestones through AUL 9. LDi and LCi products have been implemented through the ESA SSA Space Weather Service Network (http://swe.ssa.esa.int). The validation of the products continues while working in an operational environment. However, this is not the end of the story, new research goals were formed to answer the questions opened during the previous project. Now that we know that the new geomagnetic indices LDi and LCi are useful for nowcasting the disturbances at the ground level for local users, our aim is to forecast these indices from solar wind data. Revising the performance of the models that forecast ground magnetic disturbances from solar wind data, we discovered that they provide good results for smooth changes in local ground records, but not for fast changes, which are the most relevant for power grid users. To achieve this new goal, we need to understand the complex physics that solar wind-magnetosphere interactions rely on during transient phenomena. Some steps have already been taken Saiz et al. (2016) and a basic research project, funded by the Spanish government, is ongoing and now at AUL 1.

4.8. Funding an application's progress through the three phases: Phase III AUL 9 Project

N.Yu. Ganushkina

The following is a summary of a phase 3 AUL 9 example from the magnetospheric community. This example shows a modeling project for researcher and government users. This example also shows how projects may be funded as they go through different phases. A full version of this example can be found in Appendix A.8.

The Inner Magnetosphere Particle Transport and Acceleration Model (IMPTAM) was developed for purely scientific purposes (Ganushkina et al., 2000, 2001). IMPTAM moved to the AULs in Phase I when the project called SPACECAST was funded in 2011. The main goal of this project was protecting space assets from high energy particles. The identified users for IMPTAM were BAS (British Antarctic Survey, Cambridge, UK) and ONERA (Office National D'Etudes et de Recherches Aerospatiales, Toulouse, France). The first nowcast version of IMPTAM for < 100 keV electrons (Ganushkina et al., 2013, 2014) running online in real time (http://fp7spacecast.eu/) was developed. Validation of the IMPTAM output has been ongoing since the initial operation online (Ganushkina et al., 2015). Phase II continued during the next project, SPACESTORM (http://www.spacestorm.eu/) and at the completion of the milestones in AUL 6 and previous levels. This application of IMPTAM fully entered into Phase III as the integrated system was implemented at the user's system.

At present, IMPTAM is in another AUL spiral, Figure 3, as part of the on-going project PROGRESS funded by the European Union's Horizon 2020 research and innovation programme. In this project, IMPTAM is undergoing transformations to operate as a predictive tool.

5. Summary and Discussion

The proliferation and variety of models and data sets is a sign of the advancement and complexity of our field. As our field becomes more inter, multi, and transdisciplinary, researchers need to provide a wide array of information, data, and predictions. Our field now regularly encompasses collaborations including coupling and interactions across the heliosphere, data collection for Earth and astrophysical sciences, and applications to planetary and exoplanetary environments. Communication within our field and with related research areas is crucial to the effectiveness of our research outcomes. Similarly, there is a strong need for better communication with communities outside of academic research. One necessary step for R2R, R2O, and O2R is recognizing the needs of the user and how those needs inform requirements. This includes identifying the most useful metrics, accuracy, and format of information. As technology becomes more susceptible space weather, the challenge for researchers is to communicate clearly to users the capabilities of their research and how it can be beneficial for their needs (e.g., National Research Council, 2009; Caldwell et al., 2017; Cassak et al., 2017, and references therein).

To further enable communication of a project's progress towards defined outcomes, we have proposed the Application Usability Level framework. The AUL framework provides a step-by-step approach for tracking a project's progress towards a specific application. The framework we have presented here is intended for communication both within basic research fields and with industry users. As such, we have framed this work around two populations: *researchers* and *users*. The users may refer either to non-academic users or to other researchers. However, in the case of non-academic fields, users and researchers alike may benefit from a translator, i.e. a *broker*, who may help with the effective transition of research to operations or from one research field to another. In such cases, the expected users may not have the means or resources to fully explore the possibilities of a given model or data product for their application. Likewise, researchers might not be best positioned to appreciate the user's needs. Independent subject matter experts can be critical as brokers. They can ensure that the AULs are developed and tailored correctly for a given project. Brokers can include forecasters at government agencies (e.g., NOAA and the UK Met Office), government and government-funded scientists (including FFRDCs and government labs), academics or industry partners. Brokers should be sought out as needed early in the AUL process. In many cases, the brokers will become the user for many AUL pathways.

Within the AUL framework, the validation needs and subsequent definition of metrics are set early in the process (AUL 2). While the framework described in this paper applies to individual users with specific needs, the space physics community as a whole has a role to play in enabling the discovery and viability testing in Phase 1. The definition of a standard set of metrics for a given application, such as the CCMC's CME arrival scoreboard, can simplify the process in Phase 1. This can help ensure that each user is applying the right tool for the job. The validation and metric needs in the case of benchmarking involve the uniform application of metrics across different data or model frameworks which can measure improvement over time. The Community Coordinated Modeling Center (CCMC) has a unique role in helping to define and retain standards for use across the Heliophysics community. These efforts can be found through the work done in the International Forum for Space Weather Capabilities Assessment working groups (For more information on, and to get involved with these efforts see https://ccmc.gsfc.nasa.gov/assessment/forum-topics.php and other papers in this special issue). Community efforts such as Coupling, Energetics, and Dynamics of Atmospheric Regions (CEDAR), GEM, and Solar Heliospheric and INterplanetary Environment (SHINE) also play a role in testing these metrics. They provide an arena to test their ease of application, their usefulness for a constantly evolving scientific community, and in employing the standard metrics with cutting-edge models that may not yet be available at the CCMC.

As implied above, previous large scale community-driven efforts have focused on the validation of models and cross-calibration of instruments. These community efforts have been vital to the progression of our field but have often been centered around the needs of the researcher. The initial vision for the GEM workshop was to create multiple magnetospheric modules that would eventually be combined to produce a comprehensive model of the geospace environment and its interactions with the solar environment Roederer (1988). Efforts like the GEM Workshop (e.g., Raeder et al., 1998; Birn et al., 2001; Jordanova et al., 2006; Rastätter et al., 2013), the CCMC (e.g., Bellaire, 2006; National Research Council, 2003, 2013), and the Center for Integrated Space weather Modeling (CISM) (e.g., Spence et al., 2004; National Research Council, 2013) were instituted to enable coordination and intercommunication within and between codes. Specifically, these groups encourage the coupling of codes solving for different regions of space for the purpose of predicting the properties and variability of the space environment. The CCMC has played a crucial role in making these models available to the public. The CCMC has further helped in the validation of models and communication with users outside the space physics community. These and many other efforts continuing to make strides in improving and validating predictive models, defining metrics, and enabling communication within the field (e.g., Owens et al., 2008; Quinn et al., 2009; Shim et al., 2012; Honkonen et al., 2013; Pulkkinen et al., 2013; Rastätter et al., 2013; Gordeev et al., 2015; Glocer et al., 2016). The AUL framework can help with identifying and providing the data products for inputs into these models as shown in example 4.1. It can also help with the coordination of coupling models, as shown in examples 4.2, and 4.6. And finally, the AUL framework can inform industry users and forecasters as to the usability of the project as demonstrated in examples 4.4–4.7, and 4.8.

In this paper, we outlined the AUL framework and defined the different phases, levels, and the milestones necessary to reach each step for a project's AUL. We discussed potential methods of disseminating results as well as best practices. Several example summaries were provided, and full examples can be found in the appendix, A, which shows how this framework can be applied to current projects as well as how working with users can lead to scientific discoveries and future projects. Development of the AUL framework began at the first working meeting for the International Forum for Space Weather Capabilities Assessment by members of the Assessment of Understanding and Quantifying Progress working group. The aim of this working group is to develop a framework to aid in tracking the progress of our field and to provide a path for clear communication between researchers, funding agencies, and users.

Appendix A: Longer version of example projects using the AUL framework

Within the appendix we provide longer more explicit versions of the examples within the primary text of the paper.

A.1. Identifying a potential new application to track with the AUL framework: Phase I AUL 1 project

J. Klenzing and A. G. Burrell

The use of average solar extreme ultraviolet (EUV) as a potential driver of ionosphere-thermosphere (I-T) models is an example of an early-phase AUL 1 project. A brief outline follows for the start of a project set to use the AUL framework.

The specific EUV radiation that is used as a fundamental driver of I-T models is the spectra from 0.05–105.0 nm. This radiation heats the thermosphere and creates the ionosphere through direct ionization. Historically, I-T models have used proxies for this radiation, such as Sunspot Number (SSN) or the $F_{10.7}$ index. These proxies are used in part because of the long-running history and data continuity. However, observations during the recent solar minimum have suggested that the utility of these proxies may not be extended to periods of extremely low solar activity Emmert et al. (2010); Klenzing et al. (2011); Solomon et al. (2013). Additionally, the variability of these proxies over the 27-day solar rotation cycle shows significant deviation Chamberlin et al. (2007). Because the solar atmosphere can have different transmission properties for wavelengths of nm (EUV) vs cm ($F_{10.7}$), the day-to-day variability of these parameters can match quite well at times while being substantially different at other times.

This project is classified at AUL 1, since it uses existing published scientific knowledge to present a new idea for improving a specific group of space weather applications. To advance to an AUL of 2 or higher, the project developers would need to work with I-T model developers to determine if an improved solar EUV forcing index is viable and feasible for improving specific applications (for example, satellite drag calculation and collision avoidance due to thermospheric heating) and specify the level of improvement required for the application. Potentially, multiple AUL 2 applications could be identified from this AUL 1 project, as shown in Figure 2, each with different requirements.

A.2. Example of a Phase II project - R. M. McGranaghan

A.2.1. Application: Conductance models to calculate high-latitude ionospheric electrodynamic fields

In this example the end user is the Assimilative Mapping of Ionospheric Electrodynamics (AMIE) Richmond et al. (1988) which requires an ionospheric conductance model to infer global polar maps of electrodynamic variables (electric and magnetic fields and horizontal and field-aligned currents) on a roughly 1.5 *x* 10° latitude *x* longitude grid at variable time resolution, where the time resolution is dependent on the cadence of input observations. As researchers in the ionospheric and magnetospheric communities need event specific outputs of electrodynamic fields run on demand, the operating environment is the researcher's local computer and the relevant context changes for the specific research application.

Metrics to validate the conductance model are challenging due to the fact that conductance cannot be directly measured. However, the metric needed for this application has been outlined in Cousins et al. (2015) and McGranaghan et al. (2016), in which the accuracy of the conductance model is determined by the extent to which it provides consistency between AMIE output using two different sets of input observations (i.e., space-based magnetic perturbation observations from AMPERE and ground-based ionospheric convection observations from SuperDARN). The specific bounds in the error of the AMIE procedure due to the conductance model will depend on the goals of the specific research application. There are a number of efforts currently working on providing the necessary inputs to the AMIE model which are described below.

A.2.2. GLobal airglOW (GLOW) model Soloman et al 1988

The conductance model makes use of the GLOW electron transport and upper atmospheric chemistry model with specification of the auroral particle precipitation by Defense Meteorological Satellite Program (DMSP) satellites observations to calculate in-track conductance estimates, and then performs an assimilation of these in-track data to obtain global high-latitude conductance distributions McGranaghan et al. (2015, 2016).

For this application one set of users are research modelers using the AMIE procedure introduced above. The GLOW model is currently being validated for various geomagnetic conditions which cover the environmental conditions necessary for this specific application. However, the assimilative

GLOW model has already been examined for a characteristic event containing several periods of both quiet and geomagnetic storm and shown to provide greater accuracy than conductance models currently in wide use McGranaghan et al. (2016). The GLOW model will be able to provide the conductances necessary to run AMIE in the required environments for the end user and thus this application has passed the AUL 1 benchmarks.

This application has been deemed feasible for the research application by both the end user researchers as well as the developers. The assimilative conductance model is capable of providing the necessary conductances at the spatial and temporal resolutions required by the AMIE model and can be run on demand in a timely manner. The model has been tested and validated McGranaghan et al. (2016) and the detailed characterization of the baseline performance and limitations have been completed. Thus, Phase I development is completed by meeting the AUL 3 along with all previous Milestones and Levels.

Phase II is focused on the development, testing, and validation of the conductance model for the application of providing on-demand conductances for input into the AMIE model for research purposes. The model is running in the operational environment and has been demonstrated to be able to be run on demand for the end user needs. The organizational challenges have been managed.

These activities satisfy the AUL level 4 milestones. Future efforts to validate the global conductance patterns from the assimilative GLOW model will involve systematic testing across different relevant contexts. Additionally, ongoing efforts are comparing GLOW model output to other conductance models in various forms (e.g., Grubbs et al., 2018) and may lead to new metrics for this application. To complete the AUL 5 Milestones, we have articulated the potential improvement upon the state of the art McGranaghan et al. (2016) and created the capability to run the model during the relevant context conditions necessary for the research (quiet and storm time). During both quiet and storm conditions the model can meet the requirements needed for the AMIE procedure, and, thus, Milestones and Levels through AUL 5 have been fulfilled. Finally, because current efforts are underway to demonstrate the potential performance improvement to the AMIE model provided by the conductance model, requirements for AUL 6 have not yet been completed. Therefore, this application is currently AUL 5.

A.3. Example of a Phase III project: AUL 9 - A. C. Kellerman

A.3.1. Hindcasting and Forecast Radiation Belt Electron Fluxes

The Versatile Electron Radiation Belt (VERB) code Subbotin and Shprits (2009) has been recently combined with a Kalman filter Kalman (1960), data from the Van Allen Probe MagEIS Blake et al. (2013) and REPT Baker et al. (2013) instruments, and data from the GOES MAGED and EPEAD instruments in order to develop a data-assimilative code Shprits et al. (2013); Kellerman et al. (2014). The computational requirements for a full three dimensional Kalman filter may be quite large in the domain required for radiation belt simulations an alternative *split-operator* approach was introduced Shprits et al. (2013). The data-assimilative code was applied to study the March 1991 superstorm, leading to the discovery of a 4-zone structure in the Earth's radiation belts, identification of local acceleration events during a historical geomagnetic superstorm, and the development of the first data-assimilative radiation belt forecast model Kellerman et al. (2014). The forecast model runs at UCLA has been in operation online since February 2015, and has recently been adapted to provide output for users, outside of the research community.

The on-line forecast model is largely a research model, adapted to run automatically every two hours, producing a hindcast and a forecast. The hindcast assimilates available spacecraft observations, in this case real-time GOES primary and secondary data, and the real-time Van Allen Probes MagEIS and REPT data. The forecast utilizes the VERB code and forecast Kp to predict the change in electron phase space density (PSD) across multiple values of the three adiabatic invariants Roederer (1970); Schulz and Lanzerotti (1974).

A.3.2. Phase I

The first step required to take the project forward was to identify how this tool may be used for decision making or a particular application. The forecast model provides a recent hindcast of the state of the Earth's radiation belts, and a forecast of the state up to two days in the future. Several recorded failures of spacecraft electrical systems have been reported in the past as a result of geomagnetic activity. One such example is the failure of the attitude control system on the Galaxy 4 spacecraft in 1998 Baker et al. (1998). In order to determine whether a recent failure is due to geomagnetic activity it may be useful to have a real-time monitor of the radiation environment, which provides information to operators. The identification of this application *satisfies Milestone a) in Phase 1, AUL 1.*

The development of the VERB code 2 has been documented in the literature Subbotin and Shprits (2009), the code has been tested for numerical accuracy Subbotin and Shprits (2009), and validated against spacecraft observations Subbotin et al. (2011); Kim et al. (2012). The data assimilative model was tested for sensitivity to the datasets included in Kellerman et al. (2014). Current documentation of the code is available to all users, and there is an extensive set of examples which provide ease of access to the code. Both of these products are available on request, and are maintained on a dedicated Gitlab server. The data assimilative aspect of the code has been tested and published in the literature Shprits et al. (2013); Kellerman et al. (2014), and example scripts to load the reanalysis and conduct investigations are available also via a Gitlab server. *Together these items satisfy Milestone b) in Phase 1, AUL 1.*

Over the past few years, contact has been made between researchers and users who may be interested in the real-time operational forecast model output. Most of the contact occurred at national and international conferences, and through email exchange. *This list of potential users addresses Milestone c) in Phase 1, AUL 1, and hence the project should be rated AUL 1*

Contact was made with business who was interested in utilizing the real-time data to determine potential risks to spacecraft as a result of deep dielectric charging Meulenberg (1976). A personal meeting was set up and the needs/requirements were discussed with the interested party. In this initial interaction, the need was to provide regular output of the hindcast and forecast electron PSD, and to describe the file format and any other information associated with the data contained in the files. Email and telephone methods for communication were set up, and a schedule of activities/contact times were agreed upon. *These conditions address Milestones a) and b), Phase 1, AUL 2*

At this stage, the application has been identified - An industry user who wants to develop risk-assessment software, and at this stage just wants to begin using the data for testing. It was also decided at this stage that it would not be feasible (or necessary) to implement the model remotely, and so the implementation would remain at UCLA. Therefore, the only metric for success were to ensure a sufficient storage device, 24/7 access, and some file format changes for brevity. *With the metrics for success identified, Milestone c), Phase 1, AUL 2 is complete, and the project should be rated at AUL 2*.

At the initiation of the project, no other data-assimilative radiation belt forecast model was in existence, and hence the project can be considered state of the art. In the given application, the research model may be used to inform decision making for the satellite industry, and there will be a user who will regularly look at the output from the model. Both of these are beneficial for the research direction, as feedback will be received that may help to inform model development in the future. As mentioned previously, there was only a few modifications necessary to achieve success, in terms of AUL ascension, for this project, which likely would provide great feedback to research and development, hence the project is considered viable and implementation will be feasible. *Milestones a)-c), AUL 3, Phase 1 is complete*. The project should be rated at AUL 3.

A.3.3. Phase II

The user requires a model that can provide information for determining whether deep dielectric charging may have occurred for a recent anomaly. The forecast version of the data-assimilative VERB code, was developed largely from the published version Kellerman et al. (2014), and implements the validated and tested VERB code Subbotin and Shprits (2009); Kim et al. (2012); Drozdov et al. (2015); Aseev et al. (2016) to provide model-matrices for the assimilation framework. The code was implemented into an operational framework, capable of producing electron PSD nowcasts and forecasts. All organizational challenges were overcome, and the project has been *integrated into the environment required by the application, completing Milestones a) and b), AUL 4, Phase 2*. The project should be rated at AUL 4.

The model was set up to run automatically, producing documented output files required by the user every 2 hours. These files are currently being loaded remotely by the user on a regular basis. The application has been integrated into a functioning application system, providing state-of-the-art estimates of electron PSD in real time. The project should be rated at AUL 5.

The application has been tested first in a simulated environment, and now in an operational environment. Documentation and further testing is still underway for this project, and hence it can not be rated at AUL6.

A.3.4. Back to Phase I

Since implementing the project, there has been further discussions with the user. There is a new requirement that errors be included with the hindcast and forecast electron PSD. A data product with known errors requires further investigation, and development of the model. For this new application the project should be rated at AUL 2, as the milestones for AUL 3 require some further work before they can be considered complete. Note that a revision of the users needs necessitates that one treats the application as a new application, as the metrics/requirements have changed, and one can no longer compare the project AUL for this new application with the older one.

This example serves to demonstrate that while a project may be at AUL 5 for one application, a small change in the needs of the end user necessitates that the AUL be reassessed, and likely reverted back to an earlier level. It is then necessary to work though the levels again to ensure that the new needs can, and will be met in a systematic and robust manner.

A.4. Example of a Phase II project: AUL 6 - B.A. Carter and M. Terkildsen

A.4.1. Ionospheric scintillation prediction

Equatorial Plasma Bubbles (EPBs) are low plasma density structures that rise up into the high plasma density in the Earth's ionosphere during the nighttime hours (e.g., Kelley, Michael C. and Makela, Jonathan J. and de La Beaujardière, Odile and Retterer, John, 2011). EPBs, also known as Convective Ionospheric Storms, generate a spectrum of plasma waves/irregularities that cause random fluctuations (i.e., "scintillations") in the amplitude and phase of radio waves that propagate through them; e.g., those used for Positioning, Navigation and Timing (PNT). The amplitude and phase scintillation can cause Global Navigation Satellite System (GNSS) receivers to loose lock with one or more satellites, which can adversely impact the PNT results. The impact of ionospheric disturbances was highlighted in the 2015 multi-agency Space Weather Action Plan (National Science & Technology Council, 2015 (Lanzerotti, and references within), and NASA's Heliophysics Living with a Star Program identified "Physics-based Scintillation Forecasting Capability" as one of seven Strategic Science Areas in their recent decadal plan NASA (2015). Therefore, a current focus of the ionospheric research community has been to understand the driving mechanisms of the growth of EPBs, with the end-goal of developing an accurate EPB forecasting capability.

EPBs are known to be caused by the Generalized Rayleigh-Taylor (R-T) plasma instability (e.g., Sultan, 1996), in which a sharp vertical gradient in the plasma density in the bottom of the F layer, coupled with an upward plasma drift, creates instability in the ionospheric plasma. The plasma perturbations generated in the bottom-side of the F layer (approx. 150-200 km altitude) undergo rapid nonlinear growth into large 'bubbles' of low-density plasma that rise towards the topside of the F layer.

Empirical and phase screen propagation models have been shown to be very useful in not only capturing the EPB occurrence climatology, but also reproducing typical scintillation levels that are observed on the ground; the Wideband ionospheric scintillation model (WBMOD) Secan et al. (1995) and the Global Ionospheric Scintillation Propagation Model (GISM) Béniguel, Yannick and Hamel, Pierrick (2011) are notable examples. On the other hand, the development of a physics-based prediction capability for EPBs has two primary challenges. Firstly, physics-based predictions must have the ability to predict the occurrence of EPBs by simulating the background ionospheric conditions; particularly the daily changes in the upward plasma drift near sunset. The next challenge is to then numerically model the formation and small-scale structure(s) of the EPBs themselves (see review by Yokoyama, 2017). In achieving both of these challenges, an EPB prediction capability would be able to forecast the occurrence of EPBs, their spatial extent and their impact on radio waves. The AUL example discussed here deals with the first of these challenges; i.e., daily EPB occurrence.

A series of recent and ongoing studies into modeling the occurrence of EPBs using the Thermosphere Ionosphere Electrodynamics General Circulation Model (TIEGCM) – a global coupled 4-D model of the ionosphere-thermosphere system (Qian et al., 2014, and references therein) – are discussed here in the context of the AULs.

A.4.2. Phase I

AUL 1: Basic Research

While GNSS and Satellite Communications users are now widely recognized as the primary end-users, EPB prediction has been a topic of significant research effort since before the Global Positioning System was deployed. As such, the community benefits from decades of basic research into what began as a pure scientific curiosity into "Equatorial Spread F" Booker and Wells (1938). Basic research into EPBs is an ongoing topic, but the physical mechanism that drives the generation of EPBs is well-understood to be the R-T plasma instability.

The field has been working with the rationale that scintillation event forecasts would be useful for both Satellite Communications and GNSS users. This AUL framework would help enable communication with potential end users and help publicly track the progress of this project towards meeting their needs/requirements. Therefore, the AUL 1 Milestones have been achieved.

AUL 2: Establishment of users and their requirements for a specific application

Currently, the global community of GNSS users spans across many key industries, and one of those is aviation, with which weather forecasting agencies, such as Australia's Bureau of Meteorology (BoM), have existing communications channels. Further, the Bureau of Meteorology's Space Weather Services (BoM-SWS) is Australia's sole provider of space weather products and forecasts. Therefore, an ongoing collaboration with BoM-SWS has helped bridge the communications and knowledge gaps between researchers and potential end users of scintillation forecasts.

From this collaboration, we've learned that both GNSS-based positioning and surveillance and the use of Satellite Communications in the aviation sector are growing rapidly in line with the move to Performance Based Navigation. Amplitude scintillation can have a significant impact on aircraft using GNSS for Required Navigation Performance-based flight navigation. Further, ionospheric scintillation may disrupt satellite communications-based technology used by aircraft, with the potential to impact both communications and surveillance. The aviation sector, through the International Civil Aviation Organization (ICAO), have identified ionospheric scintillation (both current and forecast conditions) as one of a number of space weather information requirements for aviation users. These requirements will be formalized in ICAO Standards and Recommended Practices, as a recommendation for ICAO-nominated space weather centres to provide ionospheric scintillation advisories and forecasts to aviation users at not more than 6 hourly intervals. With these requirements in mind, the Milestones for AUL 2 have been achieved.

AUL 3: Assess viability of concept and current state of the art

Currently, scintillation products are built around recent or current observations from ground-based GNSS receivers, or based on climatological models (e.g., Secan et al., 1995; Béniguel, Yannick and Hamel, Pierrick, 2011). Very few scintillation forecast products are currently openly available, and to the authors' knowledge, those that do exist are largely built around climatology or extrapolating recent conditions. Consequently they tend to capture the seasonal climatology, but not necessarily the day-to-day variability in ionospheric scintillation, which is necessary for end users to implement effective mitigation strategies.

Therefore, an EPB/scintillation prediction that is capable of capturing daily variability could significantly advance the state-of-the-art scintillation modeling/forecasting capabilities, and is the focus of the current project. The initial results of this analysis into the viability of using physics-based modeling for this purpose were published in Carter et al. (2014c) (discussed in further detail below), thus the AUL 3 Milestones have been completed.

A.4.3. Phase II

AUL 4: Initial integration and verification

Carter et al. (2014c) was the first to use the TIEGCM to directly calculate the flux-tube integrated R-T linear growth rate derived by Sultan (1996). In Carter et al. (2014c)'s work, a daily variation in the maximum R-T growth rate was revealed, and this variation showed a clear resemblance to the occurrence of amplitude scintillation, as measured using a ground-based GPS receiver. Further analysis in that study, and in a subsequent study Carter et al. (2014a), examined the source of the TIEGCM R-T growth rate daily variability, and found it to be caused by variations in the TIEGCM's magnetospheric input in the high-latitude region; i.e., the electric potential patterns that drive horizontal plasma drift. The high-latitude plasma flow variations were found to influence the thermospheric winds in the equatorial region hours later, and these changes were found to influence the strength of the R-T growth rate. These thermospheric wind variations that were modeled by the TIEGCM is understood to be the "disturbance dynamo" effect Blanc and Richmond (1980).

The analyses discussed above focused on the peak EPB season, when EPB occurrence is dictated by conditions that suppress, not enhance, EPB growth. As such, the TIEGCM's ability to show a decreased R-T growth rate on one day compared to the day prior represents an ability to model (and potentially forecast) daily variations in EPB activity. This analysis effectively completed the AUL 4 Milestone.

AUL 5: Demonstration in the relevant context

To demonstrate the feasibility of employing the TIEGCM R-T growth rate results in an operational EPB prediction environment, Carter et al. (2014b) used the Wing et al. (2005) forecast Kp index to drive the TIEGCM in a 5-month EPB prediction trial for six locations across Africa and Asia. In this analysis, a threshold R-T growth rate of $0.4{\times}10^{-3}s^{-1}$ was used to classify whether the day would be an EPB day or a non-EPB day. During peak EPB season, it was shown that the TIEGCM R-T growth rate predictions were successful in capturing non-EPB days, as measured by the ground-based GPS receivers.

Figure A.1 shows a 2-month subset of the results from the 5-month period analyzed by Carter et al. (2014b). In this analysis, the amplitude scintillation S4 index, which is measured each minute for each satellite-to-ground link, is used. Each hour, the 90th percentile of the S4 index from all satellite links 30° above the horizon, $GPS\,S4_{90}$, is taken to indicate the presence of elevated scintillation activity. The black solid curves in Figures A.1a-f are the daily maxima of the GPS $S4_{90}$ (GPS $S4_{90}$max) throughout March and April 2014 for three GPS station locations in Southeast Asia (Bangkok, Bangdun and Calcutta) and three locations in Africa (Kampala, Kisumu and Zanzibar). The blue curves indicate the daily maximum R-T growth rates that were calculated from the TIEGCM. The orange curves show the predicted GPS $S4_{90}$max from WBMOD Secan et al. (1995), for comparison with state-of-the-art. The horizontal dashed lines indicate the chosen GPS $S4_{90}$max=0.4 threshold (i.e., the difference between an EPB day and a non-EPB day) and the chosen R-T growth rate threshold of $0.4\times10^{-3}s^{-1}$. Based on these thresholds, the Heidke Skill Scores were calculated and are shown in the top-right of each panel, alongside the percentage of days with an accurate EPB occurrence forecast for each model. The black percentage indicates the success of the "persistence" forecast; i.e., what happened yesterday will happen today. Finally, Figure A.1g shows the nowcast Kp index (KpEst, black), the 1-hr Wing Kp forecast (Kp1Hr, green) Wing et al. (2005) and the $F_{10.7}$ solar flux (blue).

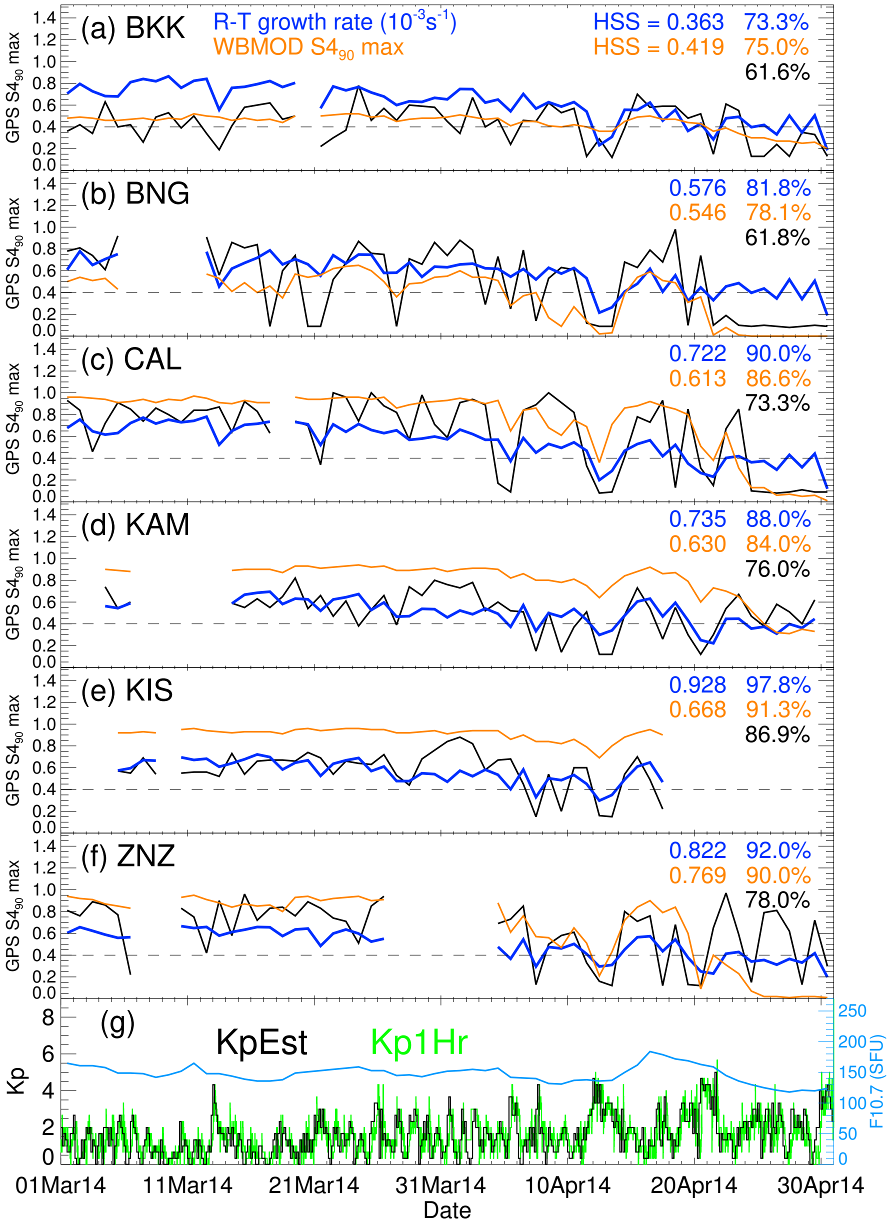

Fig. A.1. (a) to (f) The daily GPS $S4_{90}max$ observed by each GPS station throughout March and April of 2014 in black. The orange lines show the WBMOD predictions for GPS $S4_{90}max$ and the blue lines show the TIEGCM R-T growth rate. In the top-right of each panel is the corresponding Heidke Skill Score and the percentage of correct EPB/non-EPB days forecast. The black percentage indicates the “persistance” forecast result. The dashed horizontal line indicates the S4 and R-T growth rate thresholds. (g) The real-time observed Kp (KpEst, black), the 1-hour predicted Kp (Kp1Hr, green) and the $F_{10.7}$ solar flux (blue) throughout this period.

In this demonstration, it can be seen from both assessment metrics that the TIEGCM R-T growth rate is generally better at capturing the EPB daily variability than both the ARFL WBMOD and the persistence forecasts. The non-EPB days, e.g., April 13th for all stations, are characterized with

R-T growth rates of less than the $0.4 \times 10^{-3} s^{-1}$ threshold. It should be noted that the TIEGCM does not capture all non-EPB days; e.g., April 9th for both KAM and KIS stations.

Importantly, as the TIEGCM was driven using the 1-hr Wing Kp index forecast, this demonstration was designed to be comparable to an operational prediction environment that would or could be used by space weather forecasting agencies. With prediction accuracies generally consistently higher than those from WBMOD, and consistently higher than the persistence forecast, the TIEGCM was shown to be useful in the prediction of EPBs on a daily basis during peak EPB season, effectively completing the AUL 5 Milestones.

AUL 6: Complete validation

While this initial assessment is promising, there are some further questions that need answering prior to completely achieving AUL 6. In particular, the thresholds used for both the scintillation level and the R-T growth need to be further investigated.

The primary challenge with using a single S4 threshold is that the background electron density is proportional to the scintillation level Whalen (2009). Therefore, stations located under the equatorial anomaly trough (i.e., at the magnetic equator) are going to register lower S4 values compared to stations located under the anomaly crests. The levels across these different locations, which are also likely to change with time (e.g., season, solar activity, etc.) need to be quantified.

In the work discussed above, the R-T growth rate threshold was chosen by eye to be $0.4 \times 10^{-3} s^{-1}$ for all six stations, which may not be appropriate for all stations in different longitude sectors; Carter et al. (2014a)'s analysis uncovered notable differences between the optimal R-T growth rate thresholds between different longitude sectors. One further complication is that the R-T growth rate is not a measurable quantity, so it cannot be directly verified against direct observations. As such, a more rigorous and systematic analysis that investigates the TIEGCM R-T growth rate threshold is needed before the Milestones of AUL 6 are achieved.

Both of these aspects are part of ongoing work.

A.4.4. Current work and future plans

Firstly, current research is focused on quantifying the optimal/most reliable scintillation and R-T growth rate thresholds for given locations, and exploring the conditions under which these thresholds should be adapted. Verification of scintillation forecast products in terms of ground-based S4 estimates is straight forward and a good amount of high quality S4 data exists for this purpose. Many space weather agencies are also using proxies for ionospheric scintillation such as ROTI, and these show good correlation with scintillation indices.

Also, some initial work is being done towards achieving AUL 7 (i.e., Application prototype), in collaboration with the Australian BoM-SWS, and in consultation with the TIEGCM developers (National Center for Atmospheric Research, NCAR). The current goal is to set up the scintillation forecasting scheme used by Carter et al. (2014b) in an operational environment, with the intention of providing 'beta' scintillation forecasts for key end-users (such as aviation).

While it may seem premature to proceed into AUL 7 without having completed the validation in AUL 6, we expect that any findings related to the scintillation and modeled growth rate thresholds could be easily translated into an operational 'beta' scintillation forecasting system.

Proceeding into AUL 7 will help with the challenge of verifying the scintillation forecasts and advisories in terms of end user experience, primarily because end user experience varies with application, equipment, usage, tolerance, etc. Proceeding with the development of a working protype and delivering forecasts has the added benefit of informing/educating the end users of potential vulnerabilities to their system(s). This direct interaction with end users will also hopefully create a feedback loop that will allow for modifications to an operational scintillation forecasting system in order to make it more useful/informative for them. BoM have a regular program of engagement with the aviation community providing valuable feedback on pilot products and services. These interactions will thus help achieve AUL 8 (Validation in relevant "real world" environment, and eventually AUL 9 (Approved for on-demand use towards stated application).

Looking further forward, it is worth mentioning that research on global physics-based ionosphere-thermosphere modeling is continuing to advance; e.g., the most recent release of WACCM-X Liu et al. (2018). Further, data assimilation is being investigated as a tool for capturing daily R-T growth variability (e.g., Rajesh et al., 2017), and ground-to-topside modeling has been used to show that lower atmospheric forcing can be a significant source of daily variability in the R-T growth rate Shinagawa et al. (2018). Thus, research into using global ionosphere-thermosphere models for predicting EPB occurrence is expected to continue to adapt as these models and techniques continue to be expanded upon and improved.

A.5. Validating in an operational environment for multiple users, industry and government: Phase II AUL 6 project

T. Guild

The SEAES tool grew out of a need to quickly assess the likelihood of the space environment causing a satellite anomaly. It was originally developed at The Aerospace Corporation Koons and Gorney (1988) and modernized by O'Brien (2009). The SEAES algorithms produce a hazard quotient, which is the ratio of the instantaneous likelihood of an anomaly to its long-term mission-averaged likelihood of an anomaly. This environment / anomaly likelihood relationship is derived from associating historical anomalies or their proxies to space environment measurements on the same satellites, yielding a translation between environment and hazard. A key user requirement for SEAES is speed: providing a hazard quotient, or likelihood that an anomaly is due to space weather, in near-real-time to influence decisions made during satellite anomaly investigations.

SEAES completed its Phase 1 milestones during the early development at Aerospace, working closely with satellite operators during anomaly investigations where the space environment's role needed to be determined. The interaction with users informed the development of a prototype application, outlined the requirements, and culminated in a published description of the algorithms in O'Brien (2009). This satisfies all of the Phase 1 AUL milestones. This prototype application has been implemented in a relevant DOD computing network to facilitate delivering hazard quotients to users, and feedback to the development team. In addition, the SEAES algorithms have been implemented in an operational environment at NOAA/SWPC and are available to SWPC users at the following link: https://www.swpc.noaa.gov/products/seaesrt. This completes Phase 2, through AUL 6. However, the SEAES application has never been thoroughly validated with user-decided metrics, nor have the results been disseminated, failing AUL 7 and above. We can therefore claim SEAES should be ranked at an AUL 6.

A.6. Transformative and translational research identified by the needs of the user: Phase III AUL 8 project

C. J. Henney

Global solar magnetic maps are the primary input driver for most coronal and solar wind models, however, the assembling of such maps is challenging since the solar photospheric and chromospheric magnetic fields are currently only recorded for less than half of the solar surface at any given time. With a limited view of the sun, and the rotational period of the Sun as observed from Earth is approximately 27 days, global maps of the magnetic field include old data, ranging from 15 days at mid-latitudes to 6 months at the poles. The primary goal of the ADAPT (Air Force Data Assimilative Photospheric Flux Transport) project Arge et al. (2010, 2011); Hickmann et al. (2015) is to provide sequences of best estimates of the instantaneous global spatial distribution of the solar photospheric magnetic field as a function of time. Initiated in 2008 and driven by community user interests, the objective of the ADAPT project began by combining two Phase II (AUL 5) projects, photospheric magnetic flux transport model based on Worden and Harvey (2000) and rigorous data assimilation based on Kalman Filtering Hickmann et al. (2015), to produce global magnetic maps with realistic estimates of the uncertainty (completing Milestones a-c of AUL 1).

An essential element during the ADAPT model development has been the vital feedback and collaboration with active users (completing Milestones a-c in AUL 2) to assess the viability of the global maps within different scientific contexts (completing Milestones a - d AUL 3). For example, the ADAPT global maps have been used with time-dependent MHD simulations of the inner heliosphere Merkin et al. (2016), new techniques for driving non-potential solar coronal magnetic field modeling Weinzierl et al. (2016), ensemble modeling of the large CME during July 2012 Cash et al. (2015), scale-dependent data assimilation of solar photospheric magnetic fields Hickmann et al. (2016), and empirically driven time-dependent modeling of the solar wind Linker et al. (2016). Another fundamental step during the project development was to integrate the ADAPT software within a prototyping environment (completing Milestones a-b in AUL 4) and iterating on map quality, along with meta-data improvements, with various users (Milestones a-b AUL 5). For example, the ADAPT model has been running autonomously at the National Solar Observatory for the past 5 years, generating public global magnetic maps for user validation (completing Milestones a-c AUL 6). Integrating and running ADAPT autonomously within a prototype system, identifying and managing challenges, integrating the components, and prototype the system in a simulated operational environment, provided the critical real-world testing and feedback needed to ready the ADAPT software to be installed and run on demand at NOAA-SWPC. The core functionality of the ADAPT model, and this specific application is in the early stages Milestones a-b of AUL 8, now released and installed at NOAA/SWPC to be validated in the context of driving WSA-Enlil, in collaboration with the CCMC. For more background on ADAPT, and access to real-time ADAPT global solar magnetic maps, see www.nso.edu/data/nisp-data/adapt-maps/.

While searching for full-disk integrated metric parameters to validate the timing and amplitude of far-side flux evolution and emergence within ADAPT maps, a significant new application branched off the ADAPT development (Figure 2), the SIFT (Solar Indices Forecasting Tool) empirical models. A preliminary viability study of full-disk integrated parameters for global map feedback led to the discovery that flux transport modeling can be utilized to predict the observed $F_{10.7}$ (i.e., solar radio flux at 10.7 cm) values Henney et al. (2012) and bands within the VUV (vacuum ultraviolet, between 0.1 and 175 nm, which includes the XUV, EUV, and FUV) solar irradiance Henney et al. (2015). Solar $F_{10.7}$ and EUV are both key inputs to ionospheric and thermospheric models, and the ability to forecast these quantities more reliably allows for the possibility of advanced prediction of satellite drag and ionospheric structure, as proposed in Section 4.1. After completing the basic research][see][]Henn2015, iterating with users on the quality (i.e., showed improvement compared to models utilized by users) and final forecast product format (AUL 1-5), the SIFT model for $F_{10.7}$ has been user validated and is operating autonomously in a prototype mode (AUL 7) generating public predictions for users, along with providing real-time feedback on the ADAPT maps. The next step for the SIFT $F_{10.7}$ forecast model, to progress to AUL 8, is to validate the user application metrics were met within the time specifications. For more background on SIFT, and access to real-time SIFT forecasts, see www.nso.edu/data/nisp-data/sift-forecasts/.

A.7. Example of a Phase III project: Developing new geomagnetic indices: LDi and LCi - C. Cid

Geomagnetically induced currents (GICs) are a ground level effect of solar activity, affecting electrically-conducting infrastructure (Pulkkinen et al., 2017, e.g.,). The large dependence of society on electric power makes GICs a natural hazard. Consequently, large efforts are being devoted to the assessment of the risk on electric grid at different levels. After the storm events of October 2003, affecting electric utilities in South Africa Kappenman (2005), low and mid geomagnetic latitude countries were made aware that their power grids might also be vulnerable to a hazard that it is usually seen as a high-latitude issue. The magnitude of the Spanish (northern) geomagnetic latitude is similar to South Africa (southern). This fact was the impetus for a chain of projects that concluded with the development of a new geomagnetic index for nowcasting geomagnetic disturbances and the GICs risk in Spain, which are discussed in terms of AULs.

A.7.1. Phase I: Discover and Viability

AUL 1: Basic Research

Basic research on geomagnetic indices and their relationship with the solar wind reaching the Earth is an active issue. In the context of a basic research project, Saiz et al. (2008) developed a warning procedure for large Dst with only one input: interplanetary magnetic field (IMF) B_z-component. There were two main concerns in this project: on one hand the severity of the storm (as measured by the peak Dst) was not well-related to the severity of ground space weather effects, which looked better associated with large values of Dst (at this point there were no contact to end-users); on the other, there was a lack of solar wind plasma data during the most severe events (when forecasting was more relevant), and only IMF was available. Moreover, this basic research project was trying to be useful for society, but, for that purpose an open question first needed to be solved: is the Dst a good proxy for severity of the ground effects? To answer this question potential interested users were identified but contacting them was not an easy task.

AUL 2: Establishment of users and their requirements for a specific application

In March 2011, Spanish Civil Protection organized a workshop on Space Weather inviting researchers and possible end-users. This workshop was an excellent forum for contacting a potential interested user in the project described above: the main Spanish power company (Red Eléctrica de España, REE). A first project was initiated for assessing the space weather risk on the Spanish power grid. The first objective of this project was to search if there was any relationship between failures recorded by the power company labeled as ‘unknown cause’ and some space weather proxies. Geomagnetic indices commonly used by the scientific community, like Kp or Dst, or even Dst were considered in this study, but no relationship was found.

In the framework of this project, Cid et al. (2014) perceived the large discrepancies among magnetic records during the main phase of some extreme storms from different observatories at similar latitude, but at different longitude. Differences between the observatories used to compute the Dst index were larger than 200 nT during the Quebec storm, but the main problem was not the magnitude of the discrepancies in the records, but when the disturbance was positive in some observatories while being negative in others, as happened in 29 October 2003 Cid et al. (2015). In that case, the main hazard for GICs was missed in global indices like the Dst (because of the average process). Then the analysis for assessing the space weather risk was focused in local magnetic records with temporal resolution of minutes. Although not being conclusive in the relationship between the problems in the network and space weather, this collaboration was the start of a new project consisting on the development of a new geomagnetic index useful to nowcast GICs in Spain.

AUL 3: Assess viability of concept and current state of the art

Regarding the viability and the feasibility of developing a local geomagnetic index with high temporal resolution and in real time, there were two main issues to consider. On the company side, measurements of GICs were needed to quantify the performance of the new index. REE agreed to set up the necessary equipment to continuously measure the current at the neutral of some transformers at specific locations carefully selected considering basic research. There were also local magnetic records continuously available from the San Pablo-Toledo geomagnetic observatory provided through the National Geographical Institute (IGN). On the researchers side, the new index should be obtained from local magnetic records. That seemed to be viable considering the experience of the research team, although due to the mid-latitude location of Spain, to remove daily variation from local magnetic records was clearly out of the state-of-the art. At this point the project has fulfilled all milestones of Phase 1.

A.7.2. Phase II: Development testing and validation

AUL 4: Initial integration and verification

Two new geomagnetic indices were introduced. To give a name to the indices, we used acronyms: LDi, for the Local Disturbance index, and LCi, for the Local Current index. The procedure to obtain the LDi consists on removing quiet variations from the horizontal component of the local magnetic field provided by the magnetometer. If these variations are properly removed, the result, i.e., the LDi, measures the magnetic disturbance. The LCi was obtained as the derivative of the one-minute LDi data, to consider the induction effects related to the temporal variation of the magnetic field. At this point, the final procedure to obtain two new geomagnetic indices was established and implemented to run in a server with a remote and restricted access section for the researchers and the company. The output was a real-time plot showing both indices in the last five days, which was updated automatically every minute.

AUL 5: Demonstration in the relevant context

The LDi and LCi were computed from historical data during the period 1998 to 2009 (a solar cycle). The indices seemed to work well during quiet time, and also during geomagnetic storm periods. The small deviation from zero value during quiet time allowed to estimate the uncertainty of the indices. Regarding the disturbed periods, the LDi peak value was -567 nT and was reached on 20 November 2003, matching with the most intense storm in the period analyzed, according to the Dst index. However, for this event, the LCi reached only 15 nT min^{-1}, a very small value when compared, for example, with the 60 nT min^{-1} reached on 29 October 2003 or the 68 nT min^{-1} on 31 March 2001.

GICs measurements were available only during minimum solar activity (the first record was measured on December 2013). The GICs records during the event on 23 May 2014, due to the arrival of a fast stream from a solar coronal hole, show by the first time that the development was going in the right way.

AUL 6: Complete validation

The event on 7-8 June 2014 provided the first data set with measurements clearly out of noise from the amperemeter at the neutral of one of the transformers to try a cross-correlation between the GICs measurements and the new geomagnetic indices. The results were successful as the LCi showed a linear relationship with those GICs records. The correlation coefficient was 0.7, which was a very good result considering the signal-to-noise ratio of the GICs measurements.

These results fully demonstrated the potential of LCi for the end user, but what about the LDi? Was this late index reduced only to the previous necessary step in the procedure of computing the LCi? The answer was definitively not. Even if is this index did not appear as having a potential for REE (but as an intermediate procedure), the measurement of the local geomagnetic disturbance should have at least the same potential as the Dst or the Kp geomagnetic indices. But for that purpose, the LDi needed first to be validated.

As the most sensible phase in the development of the LDi was to compute in real time the daily variation, the validation approach consisted on checking whether the quietest days according to the procedure used for LDi were the same as the Q-days considered by scientific community, i.e. the quietest days of each month deduced from the Kp indices [ref]. The validation results were successful and the procedure to obtain the index was documented and sent to the patent office for approval [Guerrero et al., 2016].

A.7.3. Phase III

AUL 7: Application prototype

The new indices appeared not only having potential for REE, but also for other power companies in Spain and neighbor countries. Also, companies affected by ground effects due to solar activity might be interested by these indices. Therefore, the prototype was integrated in SeNMEs, the Spanish Space Weather Service (www.senmes.es), not only for REE, but also for other potential users. The integration was made through two different products: real-time plots showing LDi and LCi for the last five days updated every minute, and also in the form of two color scale graphs or sentinels, one for each index, with a five-levels scale ranging from green (quiet) to red (highly disturbed). The sentinels were conceived as the best option to help decision makers in an operational environment.

The products were disseminated to other national potential end-users through a workshop organized by the Spanish Civil Protection and also to the scientific community during the European Space Weather Week.

AUL 8: Validation in relevant real world environment

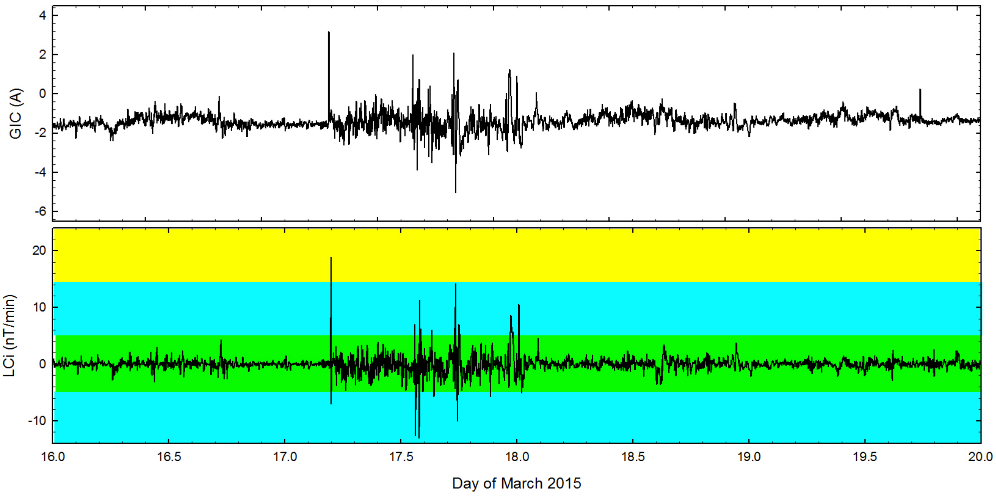

Fig. A.2. The geomagnetically induced current measured at a substation in the Northwest of Spain by REE during the period from 16 to 20 March 2015 (top panel) and the LCi geomagnetic index (bottom panel). Shadowed areas in bottom panel corresponds to the five levels scale introduced to help decision makers in an operational environment. Adapted from Cid et al. [2016]

.

The first severe geomagnetic storm of solar cycle 24, the St Patrick's Day storm, provided the opportunity to test the application developed and fully integrated in a real environment. Figure A.2 show (from top to bottom) the LCi index obtained from 17-19 March 2015 and the current measured in the neutral of the transformer of a Spanish capacity. As can be appreciated, LCi and the current recorded were very well related. Even more, the linear relationship was the same as that obtained during the previous period analyzed on 7-8 June 2014.

The original plot shown through the application have been colored in Figure A.2 to include the scales considered in the sentinels. As can be noticed, the LCi reached the yellow level according to our scales. Indeed, this storm was classified as C2.1. No consequences were reported by REE, although according to NOAA this was a G4 (severe) storm and the effects foreseen in power systems were the following: 'Possible widespread voltage control problems and some protective systems will mistakenly trip out key assets from the grid' (https://www.swpc.noaa.gov/noaa-scales-explanation). Moreover, the peak value reached by the LDi during this storm was comparable to the peak value of the Dst index (-223 nT).

After the St Patrick's Day storm, the application was considered to be fully operational and providing the specific requirements made by the power company REE. The user documentation and the training was completed at this stage.

AUL 9: Approved for on-demand use towards stated application

About 200,000 requests per week to the SeNMes server to check the products related to the new geomagnetic indices, not only from Spain, but also from world-wide locations, can be considered as a proof of interest from other end-users. In addition, these products have been implemented trough the ESA SSA Space Weather Service Network (http://swe.ssa.esa.int). The validation of the products is continued while working in an operational environment and users are being contacted through campaigns asking for feedback.

A.7.4. Back to Phase I

This is not the end of the story, but the beginning of new research goals to answer the questions opened during the previous project. Now we are aware that the geomagnetic indices LDi and LCi are useful for nowcast the disturbances at the ground level for local users. Our aim is now to forecast these indices from solar wind data.

The former projects reached a very important goal, as good nowcasting can be a good starting point for a good forecasting. Revising the performance of the models which forecast ground magnetic disturbances (non-local geomagnetic indices) from solar wind data, we discovered that they provide good results for smooth changes in ground local records, but not for fast (minute-scale) changes, which are the most relevant for power grid users. Therefore, to reach this goal, we need first to understand the complex physics that solar wind-magnetosphere interaction relies on during transient phenomena. Some steps have already been taken ahead Saiz et al. (2016) and a basic research project, funded by the Spanish government, is ongoing and now at AUL 1.

A.8. Example of a Phase III Project: Nowcast of keV electrons in the inner magnetosphere with IMPTAM - N.Yu. Ganushkina

The development of the Inner Magnetosphere Particle Transport and Acceleration Model (IMPTAM) was started as a tool to explain the observed features of ion dispersed structures in the inner magnetosphere seen at energy-time spectrograms from the CAMMICE/MICS instrument onboard the Polar spacecraft Ganushkina et al. (2000, 2001). One of the important results obtained from IMPTAM modeling was the ability of the model to reproduce the observed amount of ring current protons with energies > 80 keV during a storm recovery phase Ganushkina et al. (2005, 2006) by incorporating, in addition to the large-scale fields, transient fields associated with the dipolarization process in the magnetotail during substorm onset. The name IMPTAM appeared, actually, later in the study of the dependence of the modeled ring current on the representations of magnetic and electric fields and boundary conditions used in simulations Ganushkina et al. (2012). For this initial project, the model was used for purely scientific purposes, without any identification of potential users or specific applications.

IMPTAM moved to the AULs in Phase I when the project called SPACECAST was funded by the European Union Seventh Framework Programme (FP7/2007-2013) in 2011 (ended in February 2014). The main goal of this project was formulated as protecting space assets from high energy particles by developing European dynamic modeling and forecasting capabilities. The SPACECAST team consisted of leading experts from several EU countries providing their models for radiation environment for further development and inter-coupling inside the project. At that stage, the identified users for IMPTAM were BAS, British Antarctic Survey, (Cambridge, UK) and ONERA, Office National D'Études et de Recherches Aerospatiales, (Toulouse, France). British Antarctic Survey had their BAS radiation belts model Glauert and Horne (2005) and ONERA had their Salammbô global radiation belt model Beutier and Boscher (1995). Neither model included low energy electrons, the seed population of < 100 keV, which is critically important for radiation belt dynamics. The first nowcast version of IMPTAM for < 100 keV electrons Ganushkina et al.

(2013, 2014) running online in real time (http://fp7spacecast.eu/) was developed providing seed population for both BAS and Salammbô models Horne et al. (2013). All milestones in the three AULs were past in Phase I: basic scientific concepts and potential applications were identified in the beginning of the SPACECAST project (AUL 1); the users together with their requirements were identified (AUL 2); and IMPTAM was at the current state of the art Ganushkina et al. (2014) being able to reproduce the observed variations of keV electrons on the time scale of minutes. All the project's deliverables were successfully submitted in time, all the deadlines were met which were requirements from the European Commission.

IMPTAM moved to Phase II even during the SPACECAST project, since it went through development, testing, and validation. AUL 4 and AUL 5 were reached when IMPTAM was integrated into the functioning system of radiation belt models running at http://fp7-spacecast.eu/. Validation of the IMPTAM output (AUL 6) has been ongoing since initial operation online in real time in February 2013 Ganushkina et al. (2015). In a sense, IMPTAM was already in Phase III, since it has been operational for a full year when the SPACECAST project ended.

Phase II continued during the next project, SPACESTORM (http://www.spacestorm.eu/), funded by the European Union Seventh Framework Programme (FP7/2007-2013) in 2013 (ended in March 2017). The milestones in AULs 4–6 were completed and these efforts written up in the project's deliverables for the Commission. At the completion of the milestones in AUL 6 and previous levels, this application of IMPTAM fully entered into Phase III as the integrated system was implemented at the users system. That is, IMPTAM was implemented and integrated into operational status.

The SPACESTORM consortium consisted of five partners and the goal was to model severe space weather events and mitigate their effects on satellites by developing better mitigation guidelines, forecasting, and by experimental testing of new materials and methodologies to reduce vulnerability. During the SPACESTORM project, a "real-world" user was identified which was the group of project participants from ONERA's DESP, Space Environment Department. The presence of rapidly-varying low energy (<200 keV) electrons causes surface charging effects on satellites, changes in the satellite potential and deg-radiation of satellite surface materials. Therefore, the unique value of IMPTAM's ability to model the variations of keV electron fluxes at any satellite orbit in the inner magnetosphere was of the exceptional interest for users from ONERA. They identified a current need to determine the risks that extreme events present to critical spacecrafts in GEO and MEO (geosynchronous and medium Earth orbit, respectively). The special software called Spacecraft Plasma Interaction Software (SPIS) has been developed at ONERA under ESA (European Space Agency) and CNES (Centre National D'Études Spatiales, the French government space agency) funding. SPIS is used to assess surface charging levels of spacecraft immersed in severe GEO and MEO environments. The requirements set for IMPTAM were to provide locations and magnitudes of worst case electron fluxes (< 100 keV) at MEO by validating IMPTAM at GEO based on the database of surface charging events observed at LANL spacecraft Matéo-Vélez et al. (2018). This procedure follows all three AULs for Phase III: AUL 7 with application prototype in which the type and specifics of application was determined, AUL 8 with validation in relevant "real world" environment with the observed surface charging events, and AUL 9 with the on-demand use of IMPTAM by ONERA's SPIS software to compute surface charging for any event of interest.

At present, IMPTAM is part of the on-going project PROGRESS (ssg.group.shef.ac.uk/progress/html) funded by the European Union's Horizon 2020 research and innovation programme (ends in July 2018). The overall aim of the PROGRESS project is to develop an accurate and reliable forecast of space weather hazards. In this project, IMPTAM is undergoing transformations to operate as a predictive tool (imptam.fmi.fi, https://ssg.group.shef.ac.uk/progress/html/imptam_results.phtml), not only as a near real time tool which it has been so far. To be a predictive tool, IMPTAM required the foretasted solar wind and IMF parameters and geomagnetic indices to drive it. IMPTAM can be considered predictive when reliable forecasts for its driving parameters become available within the PROGRESS project.

Acknowledgements. This work has been completed as part of the Assessment of Understanding and Quantifying Progress working group which is part of the International Forum for Space Weather Capabilities Assessment. A large thanks goes to the International Forum for Space Weather Capabilities Assessment, CCMC, Masha Kuznetsova, and Lawrence Friedl for the wonderful discussions during the workshop and since which resulted in this paper. The authors would like to thank overleaf which aided in our ability to collaborate among a large number of active authors. A.J.H. would like to thank all of the co-authors for their efforts and patience with putting together such a large collaboration, especially such an international one where co-authors were at times asked to stay up very late or get up way too early. All authors provided help throughout the entire paper but A.J.H. would like to point out where individual efforts were focused and greatly appreciated. Specifically, A.J.H. would like to thank, B.A.C, J.K, R.M.M, C.C., C.J.H., A.G.B., T.G., N.Y.G., and M.T. for their extensive work on providing examples for the paper and being Guinea Pigs for applying the AUL framework to their past and ongoing projects; M.W.L, B.M.W, K.D.L, S.F., M.M.B, and S.K.M., for their help writing and refining the intro and discussion sections and especially with help defining all the vocabulary; D.T.W. and A.G.B. for going above and beyond with help to state all of the communities ideas throughout the paper as succinctly as possible; A.P, and B.J.T for bringing their past experience using and working with ARLs to help refine the new AUL levels and making them more applicable and useful for the space weather community; K.G.-S. for helping with the entire paper and bringing a focus to past and present validation efforts from around the Heliophysics community; S.B., S.A.M, and J.P.M. for bringing the focus of the space weather user community to the paper; And last but definitely not least, A.C.K. - well, every first author needs a fantastic second author to help coordinate meeting efforts, providing an example project, and help outlining a paper as well as a cheerleader when things seem daunting. Along with all the co-authors, we would like to thank the larger Space Physics/Space Weather/Heliophysics community which has provided feedback and comments at many telecoms, conferences, and workshops. Along with the wider community, we'd like to extend a thank you to the reviewers for helping to make this paper more accessible, and encouraging us to provide tools and an infrastructure to help with easy implementation by the community. We hope through this large and active collaboration from across the community, we have developed and provided a new useful tool for applied and pure targeted space weather and space physics research.

The proposals that in part funded R. M. M's Ph.D. research which is used as an example in section 4.2 was the NSF Fellowship award DGE 1144083, NSF grant AGS-1025089, NASA grant NNX13AD64G, and the Vela Fellowship at the Los Alamos National Labs Space Weather Summer School. Portions of this research were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. RMM is currently supported by the NASA Living With a Star Jack Eddy Postdoctoral Fellowship Program, administered by the University Corporation for Atmospheric Research and coordinated through the Cooperative Programs for the Advancement of Earth System Science (CPAESS). S.A.M. is supported by the Irish Research Council Postdoctoral Fellowship Programme and the Air Force Office of Scientific Research award number FA9550-17-1-039. The projects leading to the results presented by N. Yu. G. have received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No 606716 SPACESTORM and from the European Union's Horizon 2020 research and innovation program under grant agreement No 637302

PROGRESS. Work of M.W.L. and N. Yu. G. in the US was conducted under NASA grants NNX17AI48G, NNX17AB87G, and 80NSSC17K0015, and NSF grant 1663770. B.A.C. and M.T.'s input was funded by an Australian Research Council Linkage Project grant (LP160100561). SKM acknowledges support from the US Department of Energy's Laboratory Directed Research and Development program (grant number 20170047DR). A.G.B.'s work is supported by the Office of Naval Research. AJH's work is supported in part by The Aerospace Corporation. This work did not provide any new datasets.

References

Arge, C. N., C. J. Henney, J. Koller, C. R. Compeau, S. Young, D. MacKenzie, A. Fay, and J. W. Harvey, 2010. Air Force Data Assimilative Photospheric Flux Transport (ADAPT) Model. *Twelfth International Solar Wind Conference*, **1216**, 343–346. 10.1063/1.3395870. 4.6, A.6

Arge, C. N., C. J. Henney, J. Koller, W. A. Toussaint, J. W. Harvey, and S. Young, 2011. Improving Data Drivers for Coronal and Solar Wind Models. In N. V. Pogorelov, E. Audit, and G. P. Zank, eds., 5th International Conference of Numerical Modeling of Space Plasma Flows (ASTRONUM 2010), vol. 444 of *Astronomical Society of the Pacific Conference Series*, 99. 4.6, A.6

Aseev, N. A., Y. Y. Shprits, A. Y. Drozdov, and A. C. Kellerman, 2016. Numerical applications of the advective−diffusive codes for the inner magnetosphere. *Space Weather*, **14**(11), 993–1010. 2016SW001484, 10.1002/2016SW001484. A.3.3

Azizian, N., T. Mazzuchi, S. Sarkani, and D. F. Rico, 2011. A framework for evaluating technology readiness, system quality, and program performance of U.S. DoD acquisitions. *Systems Engineering*, **14**(4), 410–426. 10.1002/sys.20186, https://onlinelibrary.wiley.com/doi/pdf/10.1002/sys.20186, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sys.20186. 1.1

Baker, D. N., J. H. Allen, S. G. Kanekal, and G. D. Reeves, 1998. Disturbed space environment may have been related to pager satellite failure. *EOS Trans.*, **79**, 477–477. 10.1029/98EO00359. A.3.2

Baker, D. N., S. G. Kanekal, V. C. Hoxie, S. Batiste, M. Bolton, et al., 2013. The Relativistic Electron-Proton Telescope (REPT) Instrument on Board the Radiation Belt Storm Probes (RBSP) Spacecraft: Characterization of Earth's Radiation Belt High-Energy Particle Populations. *Space Science Reviews*, **179**. 10.1007/s11214-012-9950-9. A.3.1

Bellaire, P., 2006. Community Coordinated Modeling Center 2005 Workshop Report. *Space Weather*, **4**(2). 10.1029/2005SW000206, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005SW000206. 5

Béniguel, Yannick and Hamel, Pierrick, 2011. A global ionosphere scintillation propagation model for equatorial regions. *J. Space Weather Space Clim.*, **1**(1), A04. 10.1051/swsc/2011004, URL https://doi.org/10.1051/swsc/2011004. A.4.1, A.4.2

Beutier, T., and D. Boscher, 1995. A three-dimensional analysis of the electron radiation belt by the Salammbô code. *Journal of Geophysical Research Space Physics*, **100**(A8), 14,853–14,862. 10.1029/94JA03066. A.8

Birn, J., J. F. Drake, M. A. Shay, B. N. Rogers, R. E. Denton, et al., 2001. Geospace Environmental Modeling (GEM) Magnetic Reconnection Challenge. *Journal of Geophysical Research: Space Physics*,

106(A3), 3715–3719. 10.1029/1999JA900449, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999JA900449. 5

Blake, J. B., P. A. Carranza, S. G. Claudepierre, J. H. Clemmons, W. R. Crain, et al., 2013. The Magnetic Electron Ion Spectrometer (MagEIS) Instruments Aboard the Radiation Belt Storm Probes (RBSP) Spacecraft. *Space Science Reviews*, **179**. 10.1007/s11214-013-9991-8. A.3.1

Blanc, M., and A. Richmond, 1980. The ionospheric disturbance dynamo. *Journal of Geophysical Research: Space Physics*, **85**(A4), 1669–1686. 10.1029/JA085iA04p01669, URL http://dx.doi.org/10.1029/JA085iA04p01669. A.4.3

Booker, H. G., and H. W. Wells, 1938. Scattering of radio waves by the F-region of the ionosphere. *Terrestrial Magnetism and Atmospheric Electricity*, **43**(3), 249–256. 10.1029/TE043i003p00249, URL http://dx.doi.org/10.1029/TE043i003p00249. A.4.2

Burrell, A. G., A. Halford, J. Klenzing, R. A. Stoneback, S. K. Morley, A. M. Annex, K. M. Laundal, A. C. Kellerman, D. Stansby, and J. Ma, 2018. Snakes on a SpaceshipAn Overview of Python in Heliophysics. *Journal of Geophysical Research: Space Physics*, **0**(0). 10.1029/2018JA025877, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2018JA025877, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JA025877. 3.6

Caldwell, B., E. McCarron, and S. Jonas, 2017. An abridged history of federal involvement in space weather forecasting. *Space Weather*, **15**(10), 1222–1237. 10.1002/2017SW001626, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017SW001626. 5

Carter, B. A., J. M. Retterer, E. Yizengaw, K. Groves, R. Caton, et al., 2014a. Geomagnetic control of equatorial plasma bubble activity modeled by the TIEGCM with Kp. *Geophysical Research Letters*, **41**(15), 5331–5339. 10.1002/2014GL060953, URL http://dx.doi.org/10.1002/2014GL060953. A.4.3, A.4.3

Carter, B. A., J. M. Retterer, E. Yizengaw, K. Wiens, S. Wing, et al., 2014b. Using solar wind data to predict daily GPS scintillation occurrence in the African and Asian low-latitude regions. *Geophysical Research Letters*, **41**(23), 8176–8184. 10.1002/2014GL062203. 4.4, A.4.3, A.4.3, A.4.4

Carter, B. A., E. Yizengaw, J. M. Retterer, M. Francis, M. Terkildsen, R. Marshall, R. Norman, and K. Zhang, 2014c. An analysis of the quiet time day-to-day variability in the formation of postsunset equatorial plasma bubbles in the Southeast Asian region. *Journal of Geophysical Research: Space Physics*, **119**(4), 3206–3223. 10.1002/2013JA019570, URL http://dx.doi.org/10.1002/2013JA019570. A.4.2, A.4.3

Cash, M. D., D. A. Biesecker, V. Pizzo, C. A. de Koning, G. Millward, C. N. Arge, C. J. Henney, and D. Odstrcil, 2015. Ensemble Modeling of the 23 July 2012 Coronal Mass Ejection. *Space Weather*, **13**, 611–625. 10.1002/2015SW001232. A.6

Cassak, P. A., A. G. Emslie, A. J. Halford, D. N. Baker, H. E. Spence, S. K. Avery, and L. A. Fisk, 2017. Space physics and policy for contemporary society. *Journal of Geophysical Research: Space Physics*, **122**(4), 4430–4435. 10.1002/2017JA024219, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JA024219. 5

Chamberlin, P., T. Woods, T. Woods, and F. G. Eparvier, 2007. Flare Irradiance Spectral Model (FISM): Daily component algorithms and results. *Space Weather: The International Journal of Research and Applications*, **5**(7), S07,005. 4.1, A.1

Cid, C., J. Palacios, E. Saiz, A. Guerrero, and Y. Cerrato, 2014. On extreme geomagnetic storms. *J. Space Weather Space Clim.*, **4**, A28. 10.1051/swsc/2014026, URL https://doi.org/10.1051/swsc/2014026. A.7.1

Cid, C., E. Saiz, A. Guerrero, J. Palacios, and Y. Cerrato, 2015. A Carrington-like geomagnetic storm observed in the 21st century. *J. Space Weather Space Clim.*, **5**, A16. 10.1051/swsc/2015017, URL https://doi.org/10.1051/swsc/2015017. A.7.1

Cousins, E. D. P., T. Matsuo, and A. D. Richmond, 2015. Mapping high-latitude ionospheric electrodynamics with SuperDARN and AMPERE. *Journal of Geophysical Research: Space Physics*, **120**(7), 2014JA020463. 10.1002/2014JA020463, URL http://dx.doi.org/10.1002/2014JA020463. 4.2, A.2.1

Drozdov, A. Y., Y. Y. Shprits, K. G. Orlova, A. C. Kellerman, D. A. Subbotin, D. N. Baker, H. E. Spence, and G. D. Reeves, 2015. Energetic, relativistic, and ultra relativistic electrons: Comparison of long-term VERB code simulations with Van Allen Probes measurements. *Journal of Geophysical Research: Space Physics*, **120**, 3574–3587. 2014JA020637, 10.1002/2014JA020637. A.3.3

Emmert, J. T., J. Lean, and J. M. Picone, 2010. Record-low thermospheric density during the 2008 solar minimum. *Geophysical Research Letters*. 4.1, A.1

Ganushkina, N. Y., O. A. Amariutei, Y. Y. Shprits, and M. W. Liemohn, 2013. Transport of the plasma sheet electrons to the geostationary distances. *Journal of Geophysical Research: Space Physics*, **118**(1), 82–98. 10.1029/2012JA017923. A.8

Ganushkina, N. Y., O. A. Amariutei, D. Welling, and D. Heynderickx, 2015. Nowcast model for low–energy electrons in the inner magnetosphere. *Space Weather*, **13**(1), 16–34. 10.1002/2014SW001098. A.8

Ganushkina, N. Y., M. W. Liemohn, O. A. Amariutei, and D. Pitchford, 2014. Low–energy electrons (5–50 keV) in the inner magnetosphere. *Journal of Geophysical Research: Space Physics*, **119**(1), 246–259. 10.1002/2013JA019304. A.8

Ganushkina, N. Y., M. W. Liemohn, and T. I. Pulkkinen, 2012. Storm-time ring current: model-dependent results. *Annales Geophysicae*, **30**(1), 177–202. 10.5194/angeo-30-177-2012, URL https://www.ann-geophys.net/30/177/2012/. A.8

Ganushkina, N. Y., T. I. Pulkkinen, V. F. Bashkirov, D. N. Baker, and X. Li, 2001. Formation of intense nose structures. *Geophysical Research Letters*, **28**(3), 491–494. 10.1029/2000GL011955. A.8

Ganushkina, N. Y., T. I. Pulkkinen, and T. Fritz, 2005. Role of substorm-associated impulsive electric fields in the ring current development during storms. *Annales Geophysicae*, **23**(2), 579–591. 10.5194/angeo-23-579-2005, URL https://www.ann-geophys.net/23/579/2005/. A.8

Ganushkina, N. Y., T. I. Pulkkinen, A. Milillo, and M. Liemohn, 2006. Evolution of the proton ring current energy distribution during 21–25 April 2001 storm. *Journal of Geophysical Research: Space Physics*, **111**(A11). 10.1029/2006JA011609. A.8

Ganushkina, N. Y., T. I. Pulkkinen, V. A. Sergeev, M. V. Kubyshkina, N. E. Baker, D. N. andTurner, et al., 2000. Entry of plasma sheet particles into the inner magnetosphere as observed by Polar/CAMMICE. *Journal of Geophysical Research Space Physics*, **105**(A11), 25,205–2521914,853. 10.1029/2000JA900062. A.8

Glauert, S. A., and R. B. Horne, 2005. Calculation of pitch angle and energy diffusion coefficients with the PADIE code. *Journal of Geophysical Research: Space Physics*, **110**(A4). 10.1029/2004JA010851. A.8

Glocer, A., L. Rastätter, M. Kuznetsova, A. Pulkkinen, H. J. Singer, et al., 2016. Community-wide validation of geospace model local K-index predictions to support model transition to operations. *Space Weather*, **14**(7), 469–480. 10.1002/2016SW001387. 5

Gordeev, E., V. Sergeev, I. Honkonen, M. Kuznetsova, L. Rastätter, M. Palmroth, P. Janhunen, G. Tóth, J. Lyon, and M. Wiltberger, 2015. Assessing the performance of community-available global MHD models using key system parameters and empirical relationships. *Space Weather*, **13**(12), 868–884. 10.1002/2015SW001307. 5

Grubbs, G., R. Michell, M. Samara, D. Hampton, J. Hecht, S. Solomon, and J.-M. Jahn, 2018. A Comparative Study of Spectral Auroral Intensity Predictions From Multiple Electron Transport Models. *Journal of Geophysical Research: Space Physics*, **123**(1), 2017JA025026. 10.1002/2017JA025026, URL http://dx.doi.org/10.1002/2017JA025026. A.2.2

Henney, C. J., R. A. Hock, A. K. Schooley, W. A. Toussaint, S. M. White, and C. N. Arge, 2015. Forecasting solar extreme and far ultraviolet irradiance. *Space Weather*, **13**, 141–153. 10.1002/2014SW001118. A.6

Henney, C. J., W. A. Toussaint, S. M. White, and C. N. Arge, 2012. Forecasting $F_{10.7}$ with solar magnetic flux transport modeling. *Space Weather*, **10**, S02011. 10.1029/2011SW000748. A.6

Hickmann, K. S., H. C. Godinez, C. J. Henney, and C. N. Arge, 2015. Data Assimilation in the ADAPT Photospheric Flux Transport Model. *Sol. Phys.*, **290**, 1105–1118. 10.1007/s11207-015-0666-3, 1410.6185. 4.6, A.6

Hickmann, K. S., H. C. Godinez, C. J. Henney, and C. N. Arge, 2016. Scale-Dependent Data Assimilation of Solar Photospheric Magnetic Field. *10th IFAC Symposium on Nonlinear Control Systems*, **49**, 193–198. 10.1016/j.ifacol.2016.10.162. A.6

Honkonen, I., L. Rastätter, A. Grocott, A. Pulkkinen, M. Palmroth, J. Raeder, A. J. Ridley, and M. Wiltberger, 2013. On the performance of global magnetohydrodynamic models in the Earth's magnetosphere. *Space Weather*, **11**(5), 313–326. 10.1002/swe.20055. 5

Horne, R. B., S. A. Glauert, N. P. Meredith, H. Koskinen, R. Vainio, et al., 2013. Forecasting the Earth's radiation belts and modelling solar energetic particle events: Recent results from SPACECAST. *J. Space Weather Space Clim.*, **3**, A20. 10.1051/swsc/2013042, URL https://doi.org/10.1051/swsc/2013042. A.8

Jordanova, V. K., Y. S. Miyoshi, S. Zaharia, M. F. Thomsen, G. D. Reeves, D. S. Evans, C. G. Mouikis, and J. F. Fennell, 2006. Kinetic simulations of ring current evolution during the Geospace Environment Modeling challenge events. *Journal of Geophysical Research: Space Physics*, **111**(A11). 10.1029/2006JA011644, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JA011644. 5

Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering*, **82**(Series D), 35–45. A.3.1

Kappenman, J. G., 2005. An overview of the impulsive geomagnetic field disturbances and power grid impacts associated with the violent Sun-Earth connection events of 29-31 October 2003 and a comparative evaluation with other contemporary storms. *Space Weather*, **3**(8). 10.1029/2004SW000128. 4.7, A.7

Kellerman, A. C., Y. Y. Shprits, D. Kondrashov, D. Subbotin, R. A. Makarevich, E. Donovan, and T. Nagai, 2014. Three-dimensional data assimilation and reanalysis of radiation belt electrons: Observations of a four-zone structure using five spacecraft and the VERB code. *Journal of Geophysical Research: Space Physics*, **119**(11), 8764–8783. 10.1002/2014JA020171, URL http://dx.doi.org/10.1002/2014JA020171. A.3.1, A.3.2, A.3.3

Kelley, Michael C. and Makela, Jonathan J. and de La Beaujardière, Odile and Retterer, John, 2011. CONVECTIVE IONOSPHERIC STORMS: A REVIEW. *Reviews of Geophysics*, **49**(2). RG2003, 10.1029/2010RG000340, URL http://dx.doi.org/10.1029/2010RG000340. A.4.1

Kim, K., Y. Shprits, D. Subbotin, and B. Ni, 2012. Relativistic radiation belt electron responses to GEM magnetic storms: Comparison of CRRES observations with 3−D VERB simulations. *Journal of Geophysical Research: Space Physics*, **117**(A8). 10.1029/2011JA017460, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JA017460. A.3.2, A.3.3

Klenzing, J., F. Simões, S. Ivanov, R. A. Heelis, D. Bilitza, R. F. Pfaff, and D. E. Rowland, 2011. Topside equatorial ionospheric density and composition during and after extreme solar minimum. *Journal of Geophysical Research: Space Physics (1978–2012)*, **116**(A12), A12,330. 4.1, A.1

Koons, H. C., and D. J. Gorney, 1988. Spacecraft Environmental Anomalies Expert System: A Status Report. *Aerosp. Rep. ATR88(9562)1, Aerosp. Corp., El Segundo, Calif.* 4.5, A.5

Lanzerotti, L. J., ???? Space Weather Strategy and Action Plan: The National Program Is Rolled Out. *Space Weather*, **13**(12), 824–825. 10.1002/2015SW001334, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015SW001334, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015SW001334. A.4.1

Linker, J. A., R. M. Caplan, C. Downs, R. Lionello, P. Riley, Z. Mikic, C. J. Henney, C. N. Arge, T. Kim, and N. Pogorelov, 2016. An Empirically Driven Time-Dependent Model of the Solar Wind. In Journal of Physics Conference Series, vol. 719 of *Journal of Physics Conference Series*, 012012. 10.1088/1742-6596/719/1/012012. A.6

Liu, H., C. G. Bardeen, B. T. Foster, P. Lauritzen, J. Liu, et al., 2018. Development and Validation of the Whole Atmosphere Community Climate Model With Thermosphere and Ionosphere Extension (WACCM-X 2.0). *Journal of Advances in Modeling Earth Systems*, **10**(2), 381–402. 10.1002/2017MS001232. A.4.4

Mankins, J. C., 1995. Technology readiness levels. *White Paper, April*, **6**, 1995. 1.1

Mankins, J. C., 2009. Technology readiness assessments: A retrospective. *Acta Astronautica*, **65**(9), 1216 – 1223. Https://doi.org/10.1016/j.actaastro.2009.03.058. 1.1

Matéo-Vélez, J., A. Sicard, D. Payan, N. Ganushkina, N. P. Meredith, and I. Sillanpää, 2018. Spacecraft surface charging induced by severe environments at geosynchronous orbit. *Space Weather*, **16**(1), 89–106. 10.1002/2017SW001689. A.8

McGranaghan, R., D. J. Knipp, T. Matsuo, and E. Cousins, 2016. Optimal interpolation analysis of high-latitude ionospheric Hall and Pedersen conductivities: Application to assimilative ionospheric electrodynamics reconstruction. *Journal of Geophysical Research: Space Physics*, **121**(5), 2016JA022486. 10.1002/2016JA022486, URL http://dx.doi.org/10.1002/2016JA022486. 4.2, A.2.1, A.2.2

McGranaghan, R., D. J. Knipp, T. Matsuo, H. Godinez, R. J. Redmon, S. C. Solomon, and S. K. Morley, 2015. Modes of high-latitude auroral conductance variability derived from DMSP energetic electron precipitation observations: Empirical orthogonal function analysis. *Journal of Geophysical Research: Space Physics*, **120**(12), 2015JA021828. 10.1002/2015JA021828, URL http://dx.doi.org/10.1002/2015JA021828. A.2.2

Merkin, V. G., J. G. Lyon, D. Lario, C. N. Arge, and C. J. Henney, 2016. Time-dependent magnetohydrodynamic simulations of the inner heliosphere. *Journal of Geophysical Research (Space Physics)*, **121**, 2866–2890. 10.1002/2015JA022200. A.6

Meulenberg, A., Jr., 1976. Evidence for a new discharge mechanism for dielectrics in a plasma. *Studies in Condensed Matter Physics*, 237–246. A.3.2

NASA, 2015. Heliophyics Living With a Star Program, 10-Year Vision Beyond 2015, https://lwstrt.gsfc.nasa.gov/images/pdf/LWS_10YrVision_Oct2015_Final.pdf. *website*. A.4.1

National Research Council, 2003. The Sun to the Earth and Beyond: A Decadal Research Strategy in Solar and Space Physics. The National Academies Press, Washington, DC. ISBN 978-0-309-08509-0. 10.17226/10477, URL https://www.nap.edu/catalog/10477/the-sun-to-the-earth-and-beyond-a-decadal-research. 5

National Research Council, 2009. Severe Space Weather Events-Understanding Societal and Economic Impacts: A Workshop Report: Extended Summary. The National Academies Press, Washington, DC. ISBN 978-0-309-13811-6. 10.17226/12643, URL https://www.nap.edu/catalog/12643/severe-space-weather-events-understanding-societal-and-economic-impacts-a. 5

National Research Council, 2013. Solar and Space Physics: A Science for a Technological Society. The National Academies Press, Washington, DC. ISBN 978-0-309-16428-3. 10.17226/13060, URL https://www.nap.edu/catalog/13060/solar-and-space-physics-a-science-for-a-technological-society. 5

O'Brien, T. P., 2009. SEAES−GEO: A spacecraft environmental anomalies expert system for geosynchronous orbit. *Space Weather*, **7**(9). 10.1029/2009SW000473, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009SW000473, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009SW000473. 4.5, A.5

Olechowski, A., S. D. Eppinger, and N. Joglekar, 2015. Technology readiness levels at 40: A study of state-of-the-art use, challenges, and opportunities. In 2015 Portland international conference on management of engineering and technology (PICMET), 2084–2094. IEEE. 1.1

Owens, M. J., H. E. Spence, S. McGregor, W. J. Hughes, J. M. Quinn, C. N. Arge, P. Riley, J. Linker, and D. Odstrcil, 2008. Metrics for solar wind prediction models comparison of empirical, hybrid, and physics-based schemes with 8 years of L1 observations. *Space Weather*, **6**(8). 10.1029/2007SW000380. 5

Pulkkinen, A., E. Bernabeu, A. Thomson, A. Viljanen, R. Pirjola, et al., 2017. Geomagnetically induced currents: Science, engineering, and applications readiness. *Space Weather*, **15**(7), 828–856. 2016SW001501, 10.1002/2016SW001501, URL http://dx.doi.org/10.1002/2016SW001501. 1.1, A.7

Pulkkinen, A., L. Rastätter, M. Kuznetsova, H. Singer, C. Balch, et al., 2013. Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather*, **11**(6), 369–385. 10.1002/swe.20056. 5

Qian, L., A. G. Burns, B. A. Emery, B. Foster, G. Lu, A. Maute, A. D. Richmond, R. G. Roble, S. C. Solomon, and W. Wang. The NCAR TIE-GCM, in Modeling the Ionosphere–Thermosphere System (eds J. Huba, R. Schunk and G. Khazanov), 73–83. John Wiley & Sons, Ltd, 2014. ISBN 9781118704417. 10.1002/9781118704417.ch7, URL http://dx.doi.org/10.1002/9781118704417.ch7. 4.4, A.4.1

Quinn, J., J. Hughes, D. N. Baker, J. Linker, J. Lyon, S. C. Solomon, and M. Wiltberger, 2009. Building and using coupled models for the space weather system: Lessons learned. *Space Weather*, **7**(5). 10.1029/2009SW000462. 5

Raeder, J., J. Berchem, and M. AshourAbdalla, 1998. The Geospace Environment Modeling Grand Challenge: Results from a Global Geospace Circulation Model. *Journal of Geophysical Research: Space Physics*, **103**(A7), 14,787–14,797. 10.1029/98JA00014, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JA00014. 5

Rajesh, P. K., C. H. Lin, C. H. Chen, J. T. Lin, T. Matsuo, M. Y. Chou, W. H. Chen, M. T. Chang, and C. F. You, 2017. Equatorial plasma bubble generation/inhibition during 2015 St. Patrick's Day storm. *Space Weather*, **15**(9), 1141–1150. 10.1002/2017SW001641. A.4.4

Rastätter, L., M. M. Kuznetsova, A. Glocer, D. Welling, X. Meng, et al., 2013. Geospace environment modeling 2008-2009 challenge: Dst index. *Space Weather*, **11**(4), 187–205. 10.1002/swe.20036. 5

Richmond, A. D., Y. Kamide, B. H. Ahn, S. I. Akasofu, D. Alcaydé, et al., 1988. Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Combined incoherent-scatter radar and magnetometer measurements for January 18-19, 1984. *Journal of Geophysical Research: Space Physics*, **93**(A6), 5760–5776. 10.1029/JA093iA06p05760, URL http://dx.doi.org/10.1029/JA093iA06p05760. 4.2, A.2.1

Roederer, J. G., 1970. Dynamics of geomagnetically trapped radiation. Springer Verlag, New York. A.3.1

Roederer, J. G., 1988. GEM: Geospace Environment Modeling. *Eos, Transactions American Geophysical Union*, **69**(33), 786–787. 10.1029/88EO01064, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/88EO01064, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/88EO01064. 5

Saiz, E., C. Cid, and Y. Cerrato, 2008. Forecasting intense geomagnetic activity using interplanetary magnetic field data. *Annales Geophysicae*, **26**(12), 3989–3998. 10.5194/angeo-26-3989-2008, URL https://www.ann-geophys.net/26/3989/2008/. A.7.1

Saiz, E., A. Guerrero, C. Cid, J. Palacios, and Y. Cerrato, 2016. Searching for Carrington-like events and their signatures and triggers. *J. Space Weather Space Clim.*, **6**, A6. 10.1051/swsc/2016001, URL https://doi.org/10.1051/swsc/2016001. 4.7, A.7.4

Schulz, M., and L. J. Lanzerotti, 1974. Particle diffusion in the radiation belts. Springer, Berlin. A.3.1

Secan, J. A., R. M. Bussey, E. J. Fremouw, and S. Basu, 1995. An improved model of equatorial scintillation. *Radio Science*, **30**(3), 607–617. 10.1029/94RS03172. A.4.1, A.4.2, A.4.3

Shim, J. S., M. Kuznetsova, L. Rasttter, D. Bilitza, M. Butala, et al., 2012. CEDAR electrodynamics thermosphere ionosphere (ETI) challenge for systematic assessment of ionosphere/thermosphere models: Electron density, neutral density, NmF2, and hmF2 using space based observations. *Space Weather*, **10**(10). 10.1029/2012SW000851. 5

Shinagawa, H., H. Jin, Y. Miyoshi, H. Fujiwara, T. Yokoyama, and Y. Otsuka, 2018. Daily and seasonal variations in the linear growth rate of the Rayleigh-Taylor instability in the ionosphere obtained with GAIA. *Progress in Earth and Planetary Science*, **5**(1), 16. 10.1186/s40645-018-0175-8. A.4.4

Shprits, Y., A. Kellerman, D. Kondrashov, and D. Subbotin, 2013. Application of a new data operator-splitting data assimilation technique to the 3-D VERB diffusion code and CRRES measurements. *Geophysical Research Letters*, **40**(19), 4998–5002. 10.1002/grl.50969, URL http://dx.doi.org/10.1002/grl.50969. A.3.1, A.3.2

Solomon, S. C., L. Qian, and A. G. Burns, 2013. The anomalous ionosphere between solar cycles 23 and 24. *Journal of Geophysical Research: Space Physics*, **118**(10), 6524–6535. 4.1, A.1

Spence, H., D. Baker, A. Burns, T. Guild, C.-L. Huang, G. Siscoe, and R. Weigel, 2004. Center for integrated space weather modeling metrics plan and initial model validation results. *Journal of Atmospheric and Solar-Terrestrial Physics*, **66**(15), 1499 – 1507. Towards an Integrated Model of the Space Weather System, https://doi.org/10.1016/j.jastp.2004.03.029, URL http://www.sciencedirect.com/science/article/pii/S1364682604001567. 5

Subbotin, D. A., and Y. Y. Shprits, 2009. Three-dimensional modeling of the radiation belts using the Versatile Electron Radiation Belt (VERB) code. *Space Weather*, **7**, S10001. 10.1029/2008SW000452. A.3.1, A.3.2, A.3.3

Subbotin, D. A., Y. Y. Shprits, and B. Ni, 2011. Long-term radiation belt simulation with the VERB 3−D code: Comparison with CRRES observations. *Journal of Geophysical Research: Space Physics*, **116**, A12210. 10.1029/2011JA017019. A.3.2

Sultan, P. J., 1996. Linear theory and modeling of the Rayleigh-Taylor instability leading to the occurrence of equatorial spread F. *Journal of Geophysical Research: Space Physics*, **101**(A12), 26,875–26,891. 10.1029/96JA00682, URL http://dx.doi.org/10.1029/96JA00682. A.4.1, A.4.3

Weinzierl, M., A. R. Yeates, D. H. Mackay, C. J. Henney, and C. N. Arge, 2016. A New Technique for the Photospheric Driving of Non-potential Solar Coronal Magnetic Field Simulations. *Astrophys. J.*, **823**, 55. 10.3847/0004-637X/823/1/55. A.6

Whalen, J. A., 2009. The linear dependence of GHz scintillation on electron density observed in the equatorial anomaly. *Annales Geophysicae*, **27**(4), 1755–1761. 10.5194/angeo-27-1755-2009, URL https://www.ann-geophys.net/27/1755/2009/. A.4.3

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, et al., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, **3**, 160,018 EP –. URL https://doi.org/10.1038/sdata.2016.18. 3.6

Wing, S., J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman, K. Costello, M. Balikhin, and K. Takahashi, 2005. Kp forecast models. *Journal of Geophysical Research: Space Physics*, **110**(A4). A04203, 10.1029/2004JA010500, URL http://dx.doi.org/10.1029/2004JA010500. A.4.3, A.4.3

Worden, J., and J. Harvey, 2000. An Evolving Synoptic Magnetic Flux map and Implications for the Distribution of Photospheric Magnetic Flux. *Sol. Phys.*, **195**, 247–268. 10.1023/A:1005272502885. 4.6, A.6

Yokoyama, T., 2017. A review on the numerical simulation of equatorial plasma bubbles toward scintillation evaluation and forecasting. *Progress in Earth and Planetary Science*, **4**(1), 37. 10.1186/s40645-017-0153-6. A.4.1