Enabling data-driven modeling and real-time prediction of the dynamic solar atmosphere

Decadal Survey for Solar and Space Physics (Heliophysics) SSPH 2024-2033

Category/Topic	Basic Research/Solar Physics
Principal Author	Lucas A. Tarr ⓘ, National Solar Observatory, ltarr@nso.edu
Co-Authors	[2]Matthias Rempel ⓘ, [1,3]Andrey N. Afanasyev, [4]Georgios Chintzoglou ⓘ, [2]Masumi Dikpati, ⓘ, [2]Yuhong Fan ⓘ, [1,3]Maria Kazachenko ⓘ, [1]N. Dylan Kee ⓘ, [5]James Leake ⓘ, [6,7]KD Leka ⓘ, [2]Anna Malanushenko ⓘ, [1]Valentin Martinez-Pillet ⓘ, [8]Xudong Sun ⓘ, [9]Tibor Török ⓘ, [1,3]Benoit Tremblay ⓘ, [10]Brian Welsch ⓘ

[1]National Solar Observatory; [2]High Altitude Observatory, National Center for Atmospheric Research; [3]Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder; [4]Lockheed Martin Solar & Astrophysics Laboratory [5]Goddard Space Flight Center, National Aeronautics and Space Administration; [6]NorthWest Research Associates; [7]Nagoya University; [8]Institute for Astronomy, University of Hawaii; [9]Predictive Science Inc.; [10]University of Wisconsin - Green Bay

SYNOPSIS: Cutting edge research of the region between the solar photosphere and Alfvén surface, where the solar wind disconnects from the solar surface, increasingly relies on large-scale computational modeling. This is true not just for numerical experiments that only involve simulations, but also for interpreting the ever-more-intricate observations coming out of new facilities. A paradigm shift is currently underway in the field, away from *ad hoc* and *ab initio* models of the atmosphere and towards simulations that are directly driven by observations. While this topic is currently at the level of basic research into the techniques, the intent is for it to serve both basic research and operational space-weather needs. We must advance data driven simulations and their supporting infrastructure to the point where community members can use time dependent dynamic models on a regular basis as a tool to interpret observations and test physical theories.

Because of the complexity of the task, the immense computational resources required, and the required longevity of such a project, we argue that this effort should be strategically funded. Such an effort is required to make the most out of current and upcoming multi-mission, multi-instrument, heterogeneous observational data.

- Dedicated funding is required for collaborative model development that allows for continuous community use, support, and contribution, *separate from the application of models to specific scientific analyses and papers*

- Data and computational resources should be hosted within the same facility to provide tightly integrated dynamic modeling, forward modeling, and data analysis. This facility needs to combine easy access to simulation and observational data and easy access to open-source codes that can be run as-is or modified as-needed

- The broad scope of data driven methods require the combination of 4π steradian spectropolarimetric observations for self-consistent global models and high temporal/spatial resolution, multi-height observations for detailed studies of local physical processes
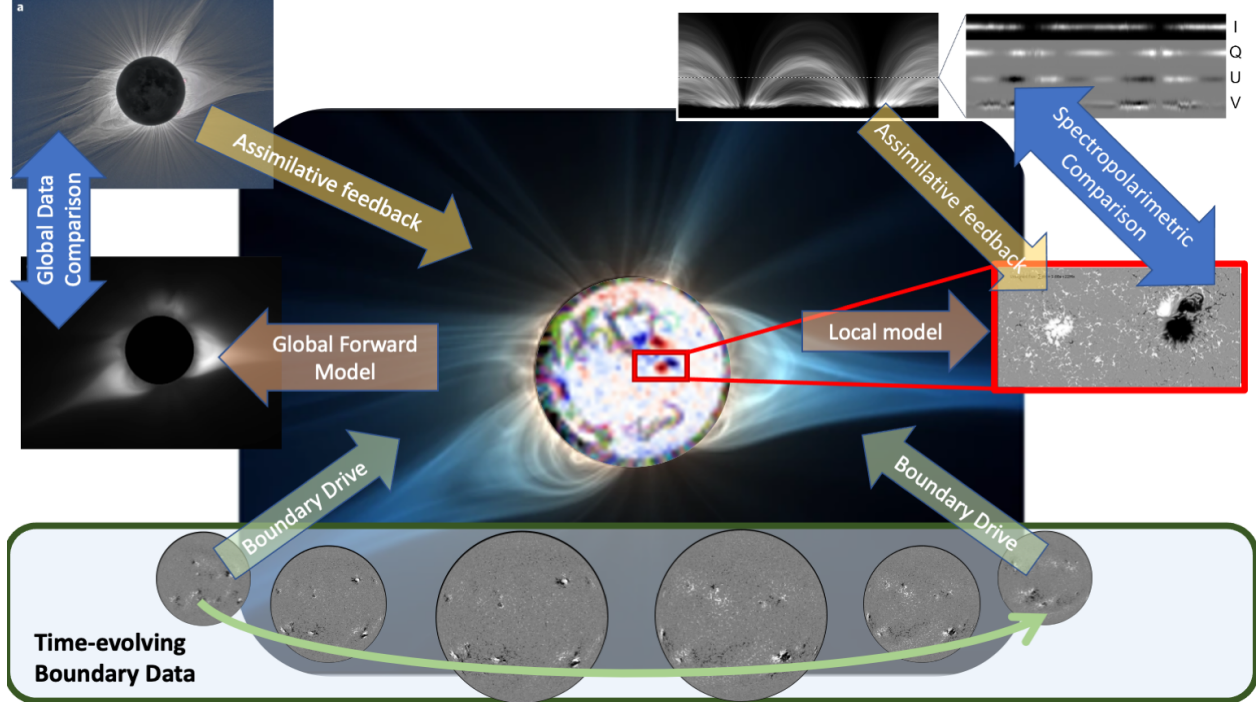
Figure 1: Cartoon overview of the data-driving and assimilative modeling framework. The lower level represents the continuous boundary driving data upon which both global (center,left) and detailed local (right) simulations are based. Additional images show recent examples of forward modeling through such simulations. Closing the feedback loops between modeling and observations is required both to interpret the latest observations via inversions and realize true, in-the-volume data assimilation. Data and modeling adapted from [1, 2, 3]; solar eclipse image in upper left © 2017 Miloslav Druckmüller, Peter Aniol, Shadia Habbal.

1. State of the art

Multi-wavelength observations in the solar photosphere, chromosphere and lower solar corona provide the inner dynamic boundary condition for the heliosphere. Understanding the 3D dynamic state of the magnetic field and plasma parameters requires models that combine individual observations (typically constraining only a small fraction of the volume in a thin radial layer) with models that provide a volume filling global context. The main open questions in the physics of the solar atmosphere — how the atmosphere above the temperature minimum region is heated, how and where the various types of solar wind arise, and what mechanisms trigger individual eruptions as flares and coronal mass ejections (CMEs) — all require knowledge of the 3D structure of the magnetic field and plasma properties above the photosphere. At present, that knowledge is severely limited. As just one demonstration, the current status of solar eruption forecasting makes clear that there is a massive overlap in the parameter spaces of eruptive and non-eruptive active regions [4, 5]; this implies that we fundamentally do not know what the 3D magnetic or thermodynamic structure of the corona is and therefore have limited ability to assess its stability.

There are primarily two groups of models that simulate evolving coronal magnetic fields in 3D: quasi-static (or time-independent) and dynamic (time-dependent). In the quasi-static group, the potential, linear, and non-linear force-free field (NLFFF) extrapolations have been developed. These models apply the vacuum-limit assumption, which assumes that magnetic pressure dominates the gas pressure (low-β regime). Another group of models, that use fewer approximations, are time-dependent models. The time-dependent models come in a variety of setups. Smaller-domain but high-resolution fluid models, such as Bifrost [6, 7] and MURaM [8, 9, 10], have the

most comprehensive physics and allow for forward modeling of observables from visible to EUV and soft X-ray wavelengths, but are computationally expensive. Large-scale models, including global fluid models [1], have lower resolution and simplified physics, but allow for the modeling of the evolution of individual eruptions originating within the global solar corona out to large radial distances from the Sun [11]. The ingestion of observations in all models is critical for the modeling of actual solar events. The research focus is currently moving rapidly from the domain of data-inspired models (where the initial setup mimics certain properties observed on the Sun) to data-driven models that rely on routine vector magnetic field observations from HMI/SDO and use new techniques to derive electric fields or plasma velocities. How to incorporate the increasing number of multi-height observations is at this point an open research topic. The future requires a focused community plan to make the most of the deluge of heterogeneous data coming from a number of sources, notably the NSF's new flagship ground-based visible and infrared Inouye Solar Telescope [12], the corona-specific CoMP [13] and COSMO [14] facilities, and radio observations from, e.g., ALMA [15] and EOVSA [16] (see also the FASR white paper [17]).

Currently, magnetohydrodynamic (MHD) models of the coupled photosphere-corona atmosphere, such as those that study CME/flare events, typically derive boundary conditions for the magnetic field from the observed photospheric magnetograms and produce the pre-eruptive configuration using (1) some form of boundary driving of magnetic flux, such as flux emergence, shear flows, and helicity condensation [18, 19, 20, 21], (2) nonlinear force-free field (NLFFF) extrapolations [22], or (3) analytical flux-rope models that are inserted into the source region of the eruption [11]. In many cases, the lower boundary driving and/or the construction of the force-free field and the inserted flux ropes are still largely *ad hoc* and not well constrained by either the observations or the physics of the driving layer. As a result, generally only qualitative agreement is obtained between the modeled magnetic field evolution and the observed event. Truly quantitative models of flare and CME events that are well constrained by observations are yet to be developed.

Another challenge is bridging the gap between physically required and numerically feasible resolution and cadence. Lately, two types of data-driven models have been developed for this purpose: magneto-friction (MF) [18] and MHD models [24]. The MF model assumes that the plasma velocity in the induction equation is proportional to the local Lorentz force; the subsequent plasma evolution leads to a relaxation of a magnetic configuration toward a

Figure 2: Left: Synthetic emission proxy from a data-driven radiative zero-beta MHD simulation of active region (AR) 11158 derived with the Coronal Global Evolutionary Model (CGEM). The run used electric fields inverted from HMI observations as a photospheric driver. Right: Observed emission of AR 11158 in AIA 131A high-temperature channel before the X2.2 flare. Adapted from [23].

force-free state. MF is more computationally efficient than MHD and is suitable for description of the slow quiescent evolution of active regions (ARs), but not for modeling of the eruptions. The MHD models explicitly solve a full set of MHD equations including the plasma properties. The MHD approach is suitable for modeling the rapid evolution of ARs during eruptions but is currently too computationally expensive to model their long-term quiescent evolution. A hybrid
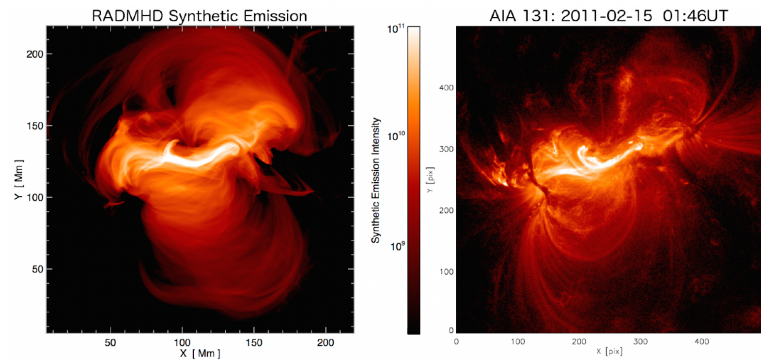
framework, where the MF model is used to model quiescent periods of AR evolution and the MHD model is used to model flaring periods of AR evolution, has been recently developed within the Coronal Global Evolutionary model (CGEM [23], see Figure 1). Other global approaches use locally concentrated or adaptively refined grids (e.g. the SWMF [25]). The future lies in data-driven models that can be implemented through: (1) boundary driving, the use of temporal sequences of photospheric electric fields (derived from vector magnetograms to represent the realistic flux transport) at the lower boundary for a time-dependent coronal field model (e.g. [26]); and (2) data assimilation, the use of temporal sequences of observations for updating the physical state of a model through statistical methods such as Ensemble-Kalman filters (EnKF).

Data Assimilation (DA) is widely used and well established in the Earth atmospheric community (e.g. to forecast the weather); however, the use in solar physics is currently limited to applications of solar-cycle forecasting [27, 28]. The full implementation of DA (through EnKF) in MHD models is a step beyond boundary driving which provides the following advantages: (1) evolution of an ensemble model that can account for observation uncertainties and calculate model errors; (2) correction of the full model state in response to new observations; (3) the ability to assimilate a wide variety of observations, including remote sensing and in-situ observations, not limited to just the lower boundary of the system. However, EnKF DA entails substantial computing costs due to requiring (1) adequate ensemble runs and (2) forward modeling of observables from physical model-outputs to compare the model with real observations at every assimilation-step.

2. Current roadblocks and how to clear them

In the next decade we must support the development of time-dependent data-driven simulations using physically consistent boundary conditions and in-the-volume assimilation that treat both the magnetic and plasma variables. Data-driven MHD following a general radiative fluid or multi-fluid approach is necessary because: **(i) Static extrapolations lack sufficient physics and constraints to be correct** in the sense that an inferred coronal state cannot confidently be stated to be near the true state. Such extrapolations cannot distinguish between, for instance, proposed theories for solar eruptions. Further, time series of extrapolations are not causally connected and therefore allow changes in the 3D state from one time to the next that are disallowed by any physical process. **(ii) Magneto-frictional methods also lack correct physics**, i.e. they omit the material plasma and include non-physical evolution, even though they do provide a major improvement on static models via built-in hysteresis. **(iii) Spectropolarimetric data require sophisticated inversions to infer the 3D magnetic and plasma structure** in both the lower atmosphere [29] and the corona [30]. Currently, inversions almost exclusively assume independent 1D radiative transfer problems without reference to or knowledge of dynamically consistent 3D models. While recent efforts try to include spatial coupling [31] or infer plasma parameters on a spatial rather than optical depth grid [32, 33], in practice, performing an inversion will always be a highly model-dependent problem from two different perspectives: first, it depends upon physical models that allow inversion of the combined diagnostics from broadband emission and polarized spectra to recover the physical characteristics of the emitting plasma; and second, computational extrapolation models must be used to fill in gaps in the observational data: spatially, temporally, and in terms of sensitivity to each plasma property (magnetic field, density, temperature, ionization degree, etc.). *Data-driven simulations, which are dynamically constrained by the underlying equations of motion, are a powerful tool to drastically reduce the uncertainties in both types of model dependencies.* **(iv) Currently,**

the inner boundary conditions and initial state of both global forecasting models and smaller-scale detailed modeling of individual events are mostly *ad hoc*. For example, CMEs inserted into a background solar wind to simulate space weather events are either initially unstable or are *ad hoc* driven until they erupt [11]. Instead, data-driving provides a scaffolding on which to hang multi-scale studies which must accurately capture and exchange the effects of physical processes occurring at each of the scales in the multi-scale simulation suite, e.g., providing the first layer of the linked series of models in the white paper by Allred et al. [34]. Physically consistent boundary conditions capture physics at larger scales outside each subset of the simulation suite, and data assimilation captures physics at sub-grid scales. Having dynamical, data constrained models will provide realistic inner boundary conditions for the heliosphere and initial conditions for more detailed local models.

To move beyond the limitations just described, time dependent data-driven models are required, but only possible given the confluence of 3 critical elements:

- The availability of data (both remote sensing and in-situ) with a stable quality, time-duration, spatial coverage, and cadence, that constrain enough of the MHD state vector;

- MHD models with a sufficient sophistication in terms of physics and their ability to simulate a solar-like setup in terms of domain extent and time-scales;

- The ability to ingest these data into models and to continuously update and correct the model state to reflect the observed conditions on the Sun.

3. Critical elements

3.1. *Data Requirements, Observing Facility requirements*

We need to have data that are capable of both feeding simulations and assessing their performance. The recent progress in data-driven models has relied critically on routine vector magnetic field observations from the HMI/SDO and new approaches to derive electric fields (or plasma velocities) from these observations. The need for such observations will only increase once full data assimilation approaches are incorporated into MHD models, and new methods must also include better constraints on the thermodynamic variables; uncertainties, in particular systematics in current and future observations, will need to be better quantified. Future routine observations of vector magnetic fields in the chromosphere and transition region (e.g., ngGONG [35]) can provide the much needed additional constraints within the volume, particularly of the pre-eruption chromospheric and coronal magnetic field [36]. Global operational models require continuous observations of the vector magnetic field and at least one thermodynamic variable over the whole 4π surface of the Sun. Future space missions need to focus on multi-spacecraft constellations mapping out a larger area of the heliosphere [37].

3.2. *Model requirements*

- Improving models such that large-scale simulations of the solar atmosphere allow for a detailed modeling of processes and the forward modeling of observables through a combination of implemented physics and numerical resolution. Forward modeling of spectropolarimetric

observables is often done as a post processing step that is poorly integrated with MHD models and is a large bottleneck for individual researchers. This bottleneck includes both easy access to extensive MHD datasets for a broad swath of the community as well as the computational resources required for the most expensive and diagnostic-rich forward models. The FORWARD [38] project was an early example of building the forward modeling side of this endeavor, but isn't directly coupled with a broad variety of simulations.

- The full implementation of data assimilation to ingest a wide range of remote sensing and in-situ observations from heliospheric observatories. Unlike boundary driven models, data assimilation can ingest a multitude of simultaneous observations within the entire simulation volume and account for observation uncertainties and calculate model errors. While data assimilation is widely used in models of Earth's atmosphere, it has rarely been explored and applied in the context of the solar atmosphere.

- The use of accelerator technologies such as GPUs and physics parameterization utilizing machine learning to boost computation speed. This will allow models to run faster than real time in order to enable research on a large number of observed events and allow for operational space weather modeling. These developments are only possible if the field widely adopts the latest computing technologies (such as the use of accelerators in form of GPUs) and stays on the forefront of new developments. While some models have been refactored for GPU use (e.g. MAS, [39]) or are in the process of refactoring (e.g. MURaM, [40]), the field of solar physics overall is already 1 decade behind the curve in adopting GPU computing and other emerging exa-scale technologies. We need targeted support for this transition in a way that allows continued in-community development of new codes through a culture change in training software engineers and domain scientists in the use of exa-scale technology.

3.3. *The need for community coordination*

Building a simulation framework like we propose is too great an effort for a single-PI, university-scale research program. There is broad interest in these tasks within the community, as evidenced by a number of recent workshops and conference sessions : "Model-Coupling workshop" (Boulder 2018); "Data-Driven Models of the Solar Progenitors of Space Weather and Space Climate" (Nagoya 2018); "Data-driven 3D Modeling of Evolving and Eruptive Solar Active Region Coronae" (ISSI 2022); and numerous sessions at the TESS, SPD, AGU, and COSPAR meetings. However, progress is currently somewhat piecemeal. What is lacking is a concerted, community-informed effort to decide the best way to implement the tight integration of data driven modeling with observational data analysis. **We therefore argue that this capability should be coordinated and supported at the national level**.

- Solar observations are incredibly sophisticated. Modern analyses increasingly rely on multi-mission, multi-instrument, heterogeneous datasets that combine imaging, spectroscopy, polarimetry, and interferometry between radio and x-ray wavelengths [41, 42]. The type of modeling framework we propose is necessary to interpret the latest observations coming out of recent and upcoming solar-capable facilities (DKIST, COSMO, UCoMP, ALMA, EOVSA, SST, GREGOR, PSP, SO, MUSE, EUVST, . . .).

- The primary development of many commonly used codes in solar physics is based in Europe (e.g. MHD codes such as BiFROST [6, 7, 43], MURAM [8, 9, 40, 44], MPI-AMRVAC [45], MANCHA3D [46], CO5BOLD [47], and inversion codes such as STiC [48], NICOLE [49], HAZEL [50], DeSIRe [51]). Therefore the European community is in a better position to fully take advantage of the latest observations, *especially the copious amount of spectropolarimetric data just starting to come on line from DKIST [12]*. If the US wants to stay at the forefront of solar research in the coming decade, studying coupled dynamics of the solar atmosphere, including the chromosphere, is crucial. This requires sophisticated multi-fluid, multi-species codes with non-equilibrium ionization coupled to global scales. In order to stay competitive the US must support the *development* of data driven codes, not just their use in specific analyses.

- The massive amounts of observational and simulation data, each of which require large computational resources to produce and analyze, need to be stored at the same location. As an example, during a recent campaign, several hours of observations from the DKIST generated roughly 20Tb of data. Handling this amount of data efficiently, and cross-comparing to a similar amount of simulation data, is an extraordinarily challenging problem beyond what can or should be required of individual researchers: models need to be tied to the evolving requirements of facilities that are supported at the national level.

- The "joined at the hip funding" of computational modeling and observations was a massive success for the IRIS/BiFROST spacecraft and simulation effort. Later, forward modeling through numerical models (BIFROST, MURaM, RADYN [52, 53]), was explicitly used in the design phase of the recently selected MUSE mission [54, 55]. Radiation MHD simulations were used in the design phase of DKIST instrumentation to simulate instrument performance. Similar capabilities should be specifically supported for new facilities in order to fully realize their individual *and collective* potential.

4. Final Statement

The developments listed in §3.3 are critical to enable a comprehensive investigation of physical processes in the coupled solar atmosphere. They are required to make the most of combined observations from multiple observatories and instruments in a consistent dynamical context given by data-driven MHD models. Self-consistently driven models will answer fundamental questions about the solar atmosphere: what is its 3D structure, how did that structure arise, and is it stable? What does its evolution tell us about the Sun's internal dynamo and how it couples outward to the solar wind? Ultimately, these models will allow the science-to-operations transition and enable us to make the jump from flare and CME forecasting based on empirical relations to those derived directly from ensemble modeling around a dynamically constrained initial condition, thus giving the physical basis for eruption probability and timing, estimation of CME strength, speed, and magnetic field orientation— the vital elements of space weather prediction that will allow humanity's continued exploration of the heliosphere.

References

[1] Mikić *et al.* Predicting the corona for the 21 August 2017 total solar eclipse. *Nature Astronomy* **2**, 913–921 (2018). doi: `10.1038/s41550−018−0562−5` .

[2] Zhu, X. & Wiegelmann, T. Testing magnetohydrostatic extrapolation with radiative MHD simulation of a solar flare. *A&A* **631**, A162 (2019). doi: `10.1051/0004−6361/201936433` .

[3] Schad, T. & Dima, G. Forward Synthesis of Polarized Emission in Target DKIST Coronal Lines Applied to 3D MURaM Coronal Simulations. *SoPh* **295** (7), 98 (2020). doi: `10.1007/s11207−020−01669−1` .

[4] Leka, K. D., Barnes, G. & Wagner, E. L. The NWRA Classification Infrastructure: Description and Extension to the Discriminant Analysis Flare Forecasting System (DAFFS). *Journal of Space Weather and Space Climate* **8** (27), A25 (2018). doi: `10.1051/swsc/2018004` .

[5] Leka, K. D. *et al.* A Comparison of Flare Forecasting Methods. II. Benchmarks, Metrics, and Performance Results for Operational Solar Flare Forecasting Systems. *ApJS* **243** (2), 36 (2019). doi: `10.3847/1538−4365/ab2e12` .

[6] Gudiksen, B. V. *et al.* The stellar atmosphere simulation code Bifrost. Code description and validation. *A&A* **531**, A154 (2011). doi: `10.1051/0004−6361/201116520` .

[7] Martínez-Sykora, J. *et al.* Ion-neutral Interactions and Nonequilibrium Ionization in the Solar Chromosphere. *ApJ* **889** (2), 95 (2020). doi: `10.3847/1538−4357/ab643f` .

[8] Vögler, A. *et al.* Simulations of magneto-convection in the solar photosphere. Equations, methods, and results of the MURaM code. *A&A* **429**, 335–351 (2005). doi: `10.1051/0004−6361:20041507` .

[9] Rempel, M. Extension of the MURaM Radiative MHD Code for Coronal Simulations. *ApJ* **834** (1), 10 (2017). doi: `10.3847/1538−4357/834/1/10` .

[10] Cheung, M. C. M. *et al.* A comprehensive three-dimensional radiative magnetohydrodynamic simulation of a solar flare. *Nature Astronomy* **3**, 160–166 (2019). doi: `10.1038/s41550−018−0629−3` .

[11] Török, T. *et al.* Sun-to-Earth MHD Simulation of the 2000 July 14 "Bastille Day" Eruption. *ApJ* **856** (1), 75 (2018). doi: `10.3847/1538−4357/aab36d` .

[12] Rimmele, T. R. *et al.* The Daniel K. Inouye Solar Telescope - Observatory Overview. *SoPh* **295** (12), 172 (2020). doi: `10.1007/s11207−020−01736−7` .

[13] Tomczyk, S. *et al.* Alfvén Waves in the Solar Corona. *Science* **317** (5842), 1192 (2007). doi: `10.1126/science.1143304` .

[14] Tomczyk, S. *et al.* Scientific objectives and capabilities of the Coronal Solar Magnetism Observatory. *Journal of Geophysical Research (Space Physics)* **121** (8), 7470–7487 (2016). doi: `10.1002/2016JA022871` .

[15] Wootten, A. & Thompson, A. R. The Atacama Large Millimeter/Submillimeter Array. *IEEE Proceedings* **97**, 1463–1471 (2009). doi: `10.1109/JPROC.2009.2020572` .

[16] Gary, D. E. *et al.* Microwave and Hard X-Ray Observations of the 2017 September 10 Solar Limb Flare. *ApJ* **863** (1), 83 (2018). doi: `10.3847/1538−4357/aad0ef` .

[17] Gary et al. Particle Acceleration and Transport, New Perspectives from Radio, X-ray, and Gamma-Ray Observations. *SSPH 2024-2033 Decadal White Paper* (2022) .

[18] Cheung, M. C. M. & DeRosa, M. L. A Method for Data-driven Simulations of Evolving Solar

Active Regions. *ApJ* **757** (2), 147 (2012). doi: `10.1088/0004-637X/757/2/147` .

[19] Jiang, C., Wu, S. T., Feng, X. & Hu, Q. Data-driven magnetohydrodynamic modelling of a flux-emerging active region leading to solar eruption. *Nature Communications* **7**, 11522 (2016). doi: `10.1038/ncomms11522` .

[20] Fan, Y. Modeling the Initiation of the 2006 December 13 Coronal Mass Ejection in AR 10930: The Structure and Dynamics of the Erupting Flux Rope. *ApJ* **824** (2), 93 (2016). doi: `10.3847/0004-637X/824/2/93` .

[21] Mackay, D. H., DeVore, C. R., Antiochos, S. K. & Yeates, A. R. Magnetic Helicity Condensation and the Solar Cycle. *ApJ* **869** (1), 62 (2018). doi: `10.3847/1538-4357/aaec7c` .

[22] Guo, Y., Xia, C., Keppens, R., Ding, M. D. & Chen, P. F. Solar Magnetic Flux Rope Eruption Simulated by a Data-driven Magnetohydrodynamic Model. *ApJL* **870** (2), L21 (2019). doi: `10.3847/2041-8213/aafabf` .

[23] Hoeksema, J. T. *et al.* The Coronal Global Evolutionary Model: Using HMI Vector Magnetogram and Doppler Data to Determine Coronal Magnetic Field Evolution. *ApJS* **250** (2), 28 (2020). doi: `10.3847/1538-4365/abb3fb` .

[24] Hayashi, K., Feng, X., Xiong, M. & Jiang, C. An MHD Simulation of Solar Active Region 11158 Driven with a Time-dependent Electric Field Determined from HMI Vector Magnetic Field Measurement Data. *ApJ* **855** (1), 11 (2018). doi: `10.3847/1538-4357/aaacd8` .

[25] Tóth, G. *et al.* Space Weather Modeling Framework: A new tool for the space science community. *Journal of Geophysical Research (Space Physics)* **110** (A12), A12226 (2005). doi: `10.1029/2005JA011126` .

[26] Fisher, G. H. *et al.* The Coronal Global Evolutionary Model: Using HMI Vector Magnetogram and Doppler Data to Model the Buildup of Free Magnetic Energy in the Solar Corona. *Space Weather* **13** (6), 369–373 (2015). doi: `10.1002/2015SW001191` .

[27] Dikpati, M., Mitra, D. & Anderson, J. L. Role of response time of a Babcock-Leighton solar dynamo model in meridional flow-speed reconstruction by EnKF data assimilation. *Advances in Space Research* **58** (8), 1589–1595 (2016). doi: `10.1016/j.asr.2016.08.004` .

[28] Kitiashvili, I. N. Data Assimilation Approach for Forecast of Solar Activity Cycles. *ApJ* **831** (1), 15 (2016). doi: `10.3847/0004-637X/831/1/15` .

[29] del Toro Iniesta, J. C. & Ruiz Cobo, B. Inversion of the radiative transfer equation for polarized light. *Living Reviews in Solar Physics* **13** (1), 4 (2016). doi: `10.1007/s41116-016-0005-2` .

[30] Kramar, M., Lin, H. & Tomczyk, S. Direct Observation of Solar Coronal Magnetic Fields by Vector Tomography of the Coronal Emission Line Polarizations. *ApJL* **819** (2), L36 (2016). doi: `10.3847/2041-8205/819/2/L36` .

[31] Asensio Ramos, A. & Díaz Baso, C. J. Stokes inversion based on convolutional neural networks. *A&A* **626**, A102 (2019). doi: `10.1051/0004-6361/201935628` .

[32] Pastor Yabar, A., Borrero, J. M. & Ruiz Cobo, B. FIRTEZ-dz. A forward and inverse solver of the polarized radiative transfer equation under Zeeman regime in geometrical scale. *A&A* **629**, A24 (2019). doi: `10.1051/0004-6361/201935692` .

[33] Löptien, B., Lagg, A., van Noort, M. & Solanki, S. K. Connecting the Wilson depression to the magnetic field of sunspots. *A&A* **635**, A202 (2020). doi: `10.1051/0004-6361/201936975` .

[34] Allred, J.C. et al. Next-Generation Comprehensive Data-Driven Models of Solar Eruptive

Events. *SSPH 2024-2033 Decadal White Paper* (2022) .

[35] Hill, F. *et al. ngGONG: The Next Generation GONG - A New Solar Synoptic Observational Network*, Vol. 51, 74 (2019).

[36] Yadav, R. *et al.* Three-dimensional magnetic field structure of a flux-emerging region in the solar atmosphere. *A&A* **632**, A112 (2019). doi: `10.1051/0004-6361/201936790` .

[37] Raouafi et al. FIREFLY: Exploring the Heliosphere from the Solar Interior to the Solar Wind. *SSPH 2024-2033 Decadal White Paper* (2022) .

[38] Gibson, S. E. *et al.* Magnetic Nulls and Super-radial Expansion in the Solar Corona. *ApJL* **840** (2), L13 (2017). doi: `10.3847/2041-8213/aa6fac` .

[39] Caplan, R. M. *et al. GPU Acceleration of an Established Solar MHD Code using OpenACC*, Vol. 1225 of *Journal of Physics Conference Series*, 012012 (2019). `1811.02605`.

[40] Wright, E. *et al. Refactoring the mps/university of chicago radiative mhd (muram) model for gpu/cpu performance portability using openacc directives*, PASC '21 (Association for Computing Machinery, New York, NY, USA, 2021). URL `https://doi.org/10.1145/3468267.3470576`.

[41] Anan, T. *et al.* Measurements of Photospheric and Chromospheric Magnetic Field Structures Associated with Chromospheric Heating over a Solar Plage Region. *ApJ* **921** (1), 39 (2021). doi: `10.3847/1538-4357/ac1b9c` .

[42] Kobelski, A. R. *et al.* A Publicly Available Multiobservatory Data Set of an Enhanced Network Patch from the Photosphere to the Corona. *ApJS* **261** (2), 15 (2022). doi: `10.3847/1538-4365/ac6b3b` .

[43] Nóbrega-Siverio, D., Martínez-Sykora, J., Moreno-Insertis, F. & Carlsson, M. Ambipolar diffusion in the Bifrost code. *A&A* **638**, A79 (2020). doi: `10.1051/0004-6361/202037809` .

[44] Przybylski, D. *et al.* Chromospheric extension of the MURaM code. *A&A* **664**, A91 (2022). doi: `10.1051/0004-6361/202141230` .

[45] Keppens, R., Teunissen, J., Xia, C. & Porth, O. Mpi-amrvac: A parallel, grid-adaptive pde toolkit. *Computers & Mathematics with Applications* **81**, 316–333 (2021). URL `https://www.sciencedirect.com/science/article/pii/S0898122120301279`. doi: `https://doi.org/10.1016/j.camwa.2020.03.023`, development and Application of Open-source Software for Problems with Numerical PDEs .

[46] Khomenko, E., Vitas, N., Collados, M. & de Vicente, A. Three-dimensional simulations of solar magneto-convection including effects of partial ionization. *A&A* **618**, A87 (2018). doi: `10.1051/0004-6361/201833048` .

[47] Freytag, B. *et al.* Simulations of stellar convection with CO5BOLD. *Journal of Computational Physics* **231** (3), 919–959 (2012). doi: `10.1016/j.jcp.2011.09.026` .

[48] de la Cruz Rodríguez, J., Leenaarts, J., Danilovic, S. & Uitenbroek, H. STiC: A multiatom non-LTE PRD inversion code for full-Stokes solar observations. *A&A* **623**, A74 (2019). doi: `10.1051/0004-6361/201834464` .

[49] Socas-Navarro, H., de la Cruz Rodríguez, J., Asensio Ramos, A., Trujillo Bueno, J. & Ruiz Cobo, B. An open-source, massively parallel code for non-LTE synthesis and inversion of spectral lines and Zeeman-induced Stokes profiles. *A&A* **577**, A7 (2015). doi: `10.1051/0004-6361/201424860` .

[50] Asensio Ramos, A., Trujillo Bueno, J. & Landi Degl'Innocenti, E. Advanced Forward Model-

ing and Inversion of Stokes Profiles Resulting from the Joint Action of the Hanle and Zeeman Effects. *ApJ* **683** (1), 542–565 (2008). doi: `10.1086/589433`.

[51] Ruiz Cobo, B. *et al.* DeSIRe: Departure coefficient aided Stokes Inversion based on Response functions. *A&A* **660**, A37 (2022). doi: `10.1051/0004-6361/202140877`.

[52] Carlsson, M. & Stein, R. F. Non-LTE Radiating Acoustic Shocks and CA II K2V Bright Points. *ApJL* **397**, L59 (1992). doi: `10.1086/186544`.

[53] Allred, J. C., Alaoui, M., Kowalski, A. F. & Kerr, G. S. Modeling the Transport of Non-thermal Particles in Flares Using Fokker-Planck Kinetic Theory. *ApJ* **902** (1), 16 (2020). doi: `10.3847/1538-4357/abb239`.

[54] De Pontieu, B. *et al.* Probing the Physics of the Solar Atmosphere with the Multi-slit Solar Explorer (MUSE). I. Coronal Heating. *ApJ* **926** (1), 52 (2022). doi: `10.3847/1538-4357/ac4222`.

[55] Cheung, M. C. M. *et al.* Probing the Physics of the Solar Atmosphere with the Multi-slit Solar Explorer (MUSE). II. Flares and Eruptions. *ApJ* **926** (1), 53 (2022). doi: `10.3847/1538-4357/ac4223`.