DRAFT of April 28, 2014

Proposal to 2014 LWS Bz Focused Science Topic

TITLE:
Robust Prediction of the Interplanetary Magnetic Field using Statistical and Physics-Based Model Approaches

Contents

1 Proposal Summary

Objectives: The primary objective of the proposed work is to derive the most robust and accurate prediction of the interplanetary magnetic field, nominally at least 24 hours in advance. Additionally, our approach will also provide estimates of solar wind speed, density, and temperature.

Methodology: Our approach is unique in that it focuses not primarily on understanding the physical processes that modulate the interplanetary magnetic field (although we anticipate that this will occur as an inevitable byproduct of the work), but with the single-minded objective of providing the most robust estimation of the interplanetary magnetic field, primarily at 1 AU. To achieve this, we will assemble all possibly viable techniques for predicting the interplanetary magnetic field, applying rigorous statistical methodologies to discriminate between, and combine them as appropriate. These include, but are not necessarily limited to: persistence (which provides the base prediction that any approach must exceed); historical pattern matching (using k-nearest neighbor, Euclidean distance, and dynamic time warping); neural networks; empirically based CME models (e.g., spheromak-cone model); and first-principle, physics-based models. *A priori*, we do not know which technique(s) will be superior; notably, though, we will pursue those that add the most value to the prediction, which may lead us in directions that we had not fully anticipated. Additionally, we will employ several data mining techniques (e.g., classification, clustering, support vector machines) to better understand the data-stream. Our team was carefully chosen to encapsulate the wide range of data collection, analysis, and both statistical and physics-based modeling expertise necessary to achieve our goals. Additionally, our team includes members from the the operations arena. We will also derive metrics for each of the in-situ measurements being predicted. We will quantify when, where, and to what extent they agree with the observations. Moreover, we will provide robust statistical measures of the uncertainty in the prediction. By so doing, we will provide necessary targets that other contemporaneous, or subsequent investigations must improve upon.

Proposed Contributions to the Focus Team Effort: Our proposed effort seeks to address all goals of this Focused Science Topic. Specifically, we will be able to provide estimates of the three components of the interplanetary magnetic field at least 24 hours in advance (and over longer periods up to 4-5 days within commensurately larger uncertainties). Additionally, the approach will allow us to provide predictions at other locations such as Mercury (Messenger) and Jupiter (Juno) and elsewhere (Solar Orbiter, Solar Probe+). Finally, our methodology lends itself to prediction of the remaining plasma parameters, including solar wind speed, density, and temperature.

Team Members: Pete Riley, Roberto Lionello, Jon A. Linker (PI, Co-I, and collaborator, respectively, PSI), Piet Martens and Rafal Angryk (Co-Is, Georgia State University), Chris Russell and Roger Ulrich (Co-Is, UCLA), Vic Pizzo, Curt de Koning, Alysha Reinard (Co-Is, NOAA), Todd Hoeksema and Yang Liu (Co-Is, Stanford), Tim Horbury (Collaborator, Imperial College, London), and Matt Owens (Collaborator, University of Reading).

2 Science, Technical Aspects, and Management

2.1 Scientific Background

2.1.1 Introduction

In this proposal, we outline a program to derive the most robust and accurate prediction of IMF B_z with up to 24 hours advance warning. In essence, we aim to build up a prediction of IMF B_z by combining several distinct elements that arise on different temporal and spatial scales: slowly-evolving stream structure; rapidly-changing large scale perturbations from ICMEs; and high-frequency fluctuations from waves and turbulence. To quantify our success, we will develop a set of metrics that will reliably track our progress during the course of the investigation. In this section, we review what is currently known about IMF B_z, what work has already been done to attempt to predict, and what lessons can be learned from the previous LWS Focused Science Topic (FST) addressing the same issue.

Conceptually, it is illuminating to consider the various processes that contribute to a non-zero z-component of the IMF. The large-scale quiescent heliospheric magnetic field has no net B_z. There is a semiannual variation in geomagnetic activity (the "Russell-McPherron" effect (Russell & McPherron, 1973)) that derives not from the true B_z component, but the projection of the IMF into the geomagnetic coordinate system. Waves and turbulence can be superposed on top of this large-scale picture (e.g., Horbury & Balogh (2001)), but in, and of themselves, they do not actively drive space weather. From a geo-effective viewpoint, large solar eruptions generating coherent flux rope structures that propagate relatively undisturbed to 1 AU represent the major source, particularly if the axis of the flux rope lies in, or near to, the ecliptic plane. In addition, however, fast CMEs drive fast-mode shocks ahead of them that compress the IMF, amplifying the wave/turbulent fluctuations. Additionally, draping of the large-scale field around the ejecta can result in large, sustained values of B_z (Gosling & McComas, 1987).

Predicting the consequences of interplanetary conditions on geomagnetic activity (at least as defined by such indices as D_{st} and Kp) has generally been easier than predicting interplanetary conditions. Burton et al. (1975) developed an empirical relationship between velocity, density, and the southward component of the IMF upstream of Earth and D_{st}. Later, feed-forward neural network algorithms with one hidden layer and back-propagation learning were developed to predict D_{st} one hour in advance (Lundstedt, 1996).

Predicting IMF B_z with any reliability, during disturbed time periods is a challenging and under-appreciated goal. It is important to distinguish between studies that aim to understand the magnetic-field properties of particular events retrospectively from those that attempt to predict them in advance. Consider, for example, the comparison of a global MHD simulation with observations shown in Figure 1. This could be interpreted as a validation that global MHD models are capable of reproducing the essential features of ICMEs. If so, it might further be inferred that such models should be a useful predictive tool. While we do not dispute that this will ultimately transpire, we caution that such comparisons necessarily amplify the positive aspects of the comparison and, indadvertedly, convey the idea that the model is more accurate and robust than it is in reality. To emphasize this point, consider that this event was chosen because: (1) it was extremely simple; and (2)

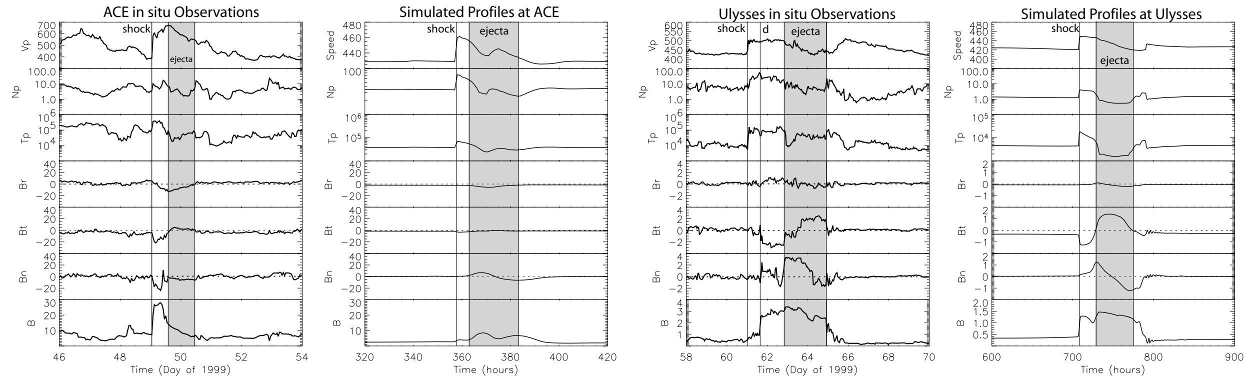

Figure 1: *Comparison of observed plasma and magnetic field parameters (a and c) with simulated parameters (b and d) for an ICME observed by the ACE at 1 AU and in the ecliptic plane (a and b), and 13 days later by the Ulysses spacecraft at 5 AU and 22°S heliographic latitude (c and d). Adapted from Riley et al. (2003).*

it matched the model results. The serendipity of identifying such a simple event that was observed by two heliospherically-separated in-situ spacecraft for which one of our limited number of simulations matched well cannot be overstated. In contrast, if we are seeking to predict an event, we do not have the luxury of hindsight to choose the model result that most matches the observations.

Currently, although a number of ideas and models have been proposed, there are no reliable predictions for IMF B_z. Perhaps the most defensible model at present is that of "persistence," that is, the value of B_z in the next hour, day, or even week is its current value. On average ($>$ 1hr, say) this value will be zero. During the passage of large, coherent magnetic clouds, on the other hand, it may have a sustained value that is substantially different from zero. It is not clear (to us, at least) which of the large number of possible ideas that could be applied to address this problem are most likely to be successful. As discussed below, various combinations of both statistical and physics-based models should be considered. Statistical techniques have met with only limited success within heliophysics, primarily, we believe, because they de-emphasize the physical underpinnings of the processes under study. Physics-based approaches range from purely empirical models, through semi-empirical, to "first principles" models. We anticipate that, at least initially, the simplest models will prevail over more complex formalisms, but, just as was the case in terrestrial weather forecasting, numerical models will eventually supersede them.

Figure 2 summarizes the essential components of our proposed effort. Each of these modules will be described in more detail in the sections that follow. Here, we remark on several points that we believe will ensure the success of our investigation. First, the overall framework is the most important aspect, not any specific model. Our team will develop a toolkit that can be accessed and used by all team members initially, and the community-at-large later, allowing them to objectively test various prediction approaches. Second, our approach is comprehensive. We cannot predict which approach will provide the optimum prediction. Indeed it is likely to be a combination of approaches which, when appropriately combined, lead to the best forecasts. Moreover, different techniques will likely mature and/or supersede others either during the lifetime of this study or later. Third, some

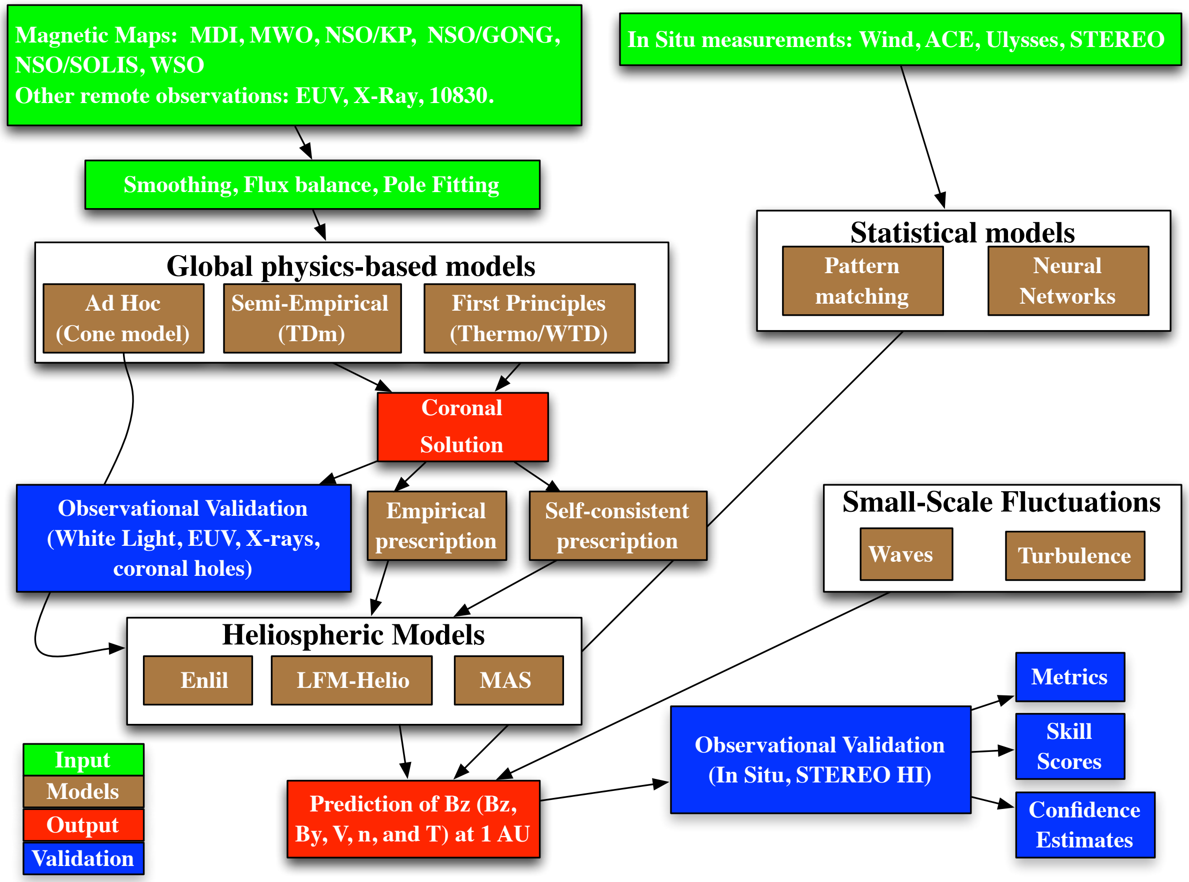

Figure 2: *Main elements of our proposed work flow. Extensive use will be made of both remote solar observations and in situ measurements (green), a range of physics-based and statistical models (brown), use of model output (red) for validation and prediction (blue). See text for more details.*

of the techniques we will consider offer little in the way of physical understanding of the processes driving variations in B_z. However, practically speaking, they may provide the best near-term forecasts. Fourth, the package we propose to develop will be based on standard, open source programming practices. This will allow a greater number of contributors to it and extend its life beyond the confines of a four-year program.

Previous studies that have assessed our ability to predict the bulk solar wind speed, an undoubtedly easier quantity to forecast have revealed that models can only modestly better "persistence." Persistence refers to the concept that a good estimate of what the solar wind will be doing in the next hour, day, or even week, can be found by looking at what it is doing now. Norquist & Meeks (2010) also demonstrated that, at least until recently, persistence remains the best technique for near-term forecasts. The persistence method is still applied by terrestrial weather forecasters, particularly in locations such as southern California, where conditions vary little from one day to the next. Additionally, we can define "recurrence" as a prediction based on observed values 27 days ago. However, such predictions are likely to be valid only under "all quiet" or "all clear" conditions. In an updated study of several numerical models, Norquist (2013) showed that recurrence had more skill than the models for three out of five years studied. Moreover, they make the most sense for parameters that have systematic variations even during quiescent conditions, such as speed, density, and temperature. Even the radial and azimuthal components of the IMF would be amenable to prediction using recurrence. However, during such times, although B_z will fluctuate rapidly due to both waves and turbulence – both of which are well-characterized, at least in a statistical sense, on average it would be predicted to remain zero. The periods of most interest, when B_z deviates from zero for substantial periods of time, occur during the passage of a coronal mass ejection (CME) and the sheath region that precedes a fast, and usually more geo-effective event. Although there is occasionally a recurrent aspect to such transient phenomena (such as when the same long-lived active region launches a sequence of ejecta), in general these are sporadic events.

The difficultly in predicting B_z at 1 AU during periods that are likely to be geo-effective cannot be over-emphasized. Currently, our best "nowcast" or near-term forecast of B_z, at least within the next 3-6 hours, is probably simple persistence. This is suggests that two quite diverse avenues can be pursued for improving on this. First, using a combination of observations and models to propagate what we see at the Sun to the vicinity of the Earth. We will outline a broad array of observational signatures and broad spectrum of models that can help us accomplish this. Second, extrapolating what we see in the vicinity of the Earth back toward the Sun to predict how B_z will change over the next 12-24 hours. The former approach will lead to predictions that can be made days in advance of the observations but, as we will discuss, is fraught with uncertainty. The second approach, will limited to near-term predictions will likely provide more accurate forecasts; at least for the duration of this proposed effort. However, neither is obviously better than than other. Our proposed investigation will create the necessary infrastructure to develop both paths, quantitatively estimating their accuracy and providing key confidence intervals. Additionally, by employing simple ensemble techniques, we can combine the various approaches yielding an optimized forecasting scheme that can incorporate new and novel approaches as they are uncovered.

It is worth noting that a cost benefit analysis suggests that the prediction of intervals of B_z with the highest likelihood of geomagnetic consequences should have the highest priority. Had the July 23, 2012 event observed *in situ* by the STEREO A spacecraft instead intercepted Earth, the massive field strengths (in excess of 100 nT), a good proportion of which was southward, would have produced unprecedented consequences to Earth's technological systems (Russell et al., 2013; Baker et al., 2013; Riley et al., 2013b).

The concepts and models that we propose to assess in this effort may at first seem overly-ambitious. However, our diverse team has expertise that overlaps with every technique we will implement. Moreover, our approach is one of utmost pragmatism. Ideas that work will continue to be pursued, but ideas that fail will be discounted (not necessarily discarded though). We believe it is important, however, to begin our effort with the most thorough and comprehensive mindset. Many ideas that have previously been explored have been through eyes of optimism. Techniques are developed and presented in the light most favorable to that technique. Rarely does a peer-reviewed study propose a novel technique only to dismiss it in the Discussion in favor of a competitor's approach. By developing a common framework for testing all ideas, our investigation can objectively and quantitatively assess them. Although team members are contributing their own models for assessment, the team as a whole will evaluate and guide the evolution of the effort over the four-year program.

Our investigation will also provide insight into the types of missions NASA might consider if the prediction of B_z warrants sufficient priority. For example, a spacecraft located at the L4 Lagrangian point could provide important Heliospheric Imaging data, together with key observations of the photospheric magnetic field that will, in ~ 7 days lie directly under the sub-solar point. On the other hand, that same spacecraft located at L1 might provide more relevant data on energetic particles accelerated by the CME-driven shock. Alternatively, perhaps a combined mission with limited and distinct instrumentation on each spacecraft would provide the optimum set of data.

2.1.2 Lessons Learned from the Previous FST

In 2007, the LWS program awarded grants to five groups (Vourlidas, Yurchyshyn, Lugaz, Jahn, and Bieber) to assemble a TR&T Focus Team on "Prediction of the Interplanetary Magnetic Field Vector Bz at L1." Although worthwhile studies were undoubtedly performed by each group, there are several reasons that the FST was not as successful as it could have been (Vourlidas, Personal Communication, 2014). First, the team members did not work well together. To a large degree, they did not "buy in" to the topic's goal. Second, the proposals were on very disparate topics, including muon measurements and global CME modeling. Our proposed effort resolves, or at least mitigates these issues. By pre-defining the team, we have identified the key people and their roles in achieving the single objective of the proposed work. We have "buy in" from the members because all agreed to participate as necessary to support the project's goal. Moreover, various subsets of the proposed team have worked successfully within other teams over many years.

Since the previous FST on this topic, seven years have elapsed. During this time, numerical models have improved, both remote and *in situ* observations have expanded, and our understanding of both the eruption process and evolution of CMEs through the solar wind is arguably better. In spite of a series of papers documenting these advances, it is not clear that any of them have improved our predictive capabilities. Reproducing the essential features of a particular set of observations is a distinct and considerably more tractable goal than predicting those same observations in advance. Moreover, whereas such studies necessarily emphasize the positive aspects of model/data comparisons (because little insight would be gained from poor comparisons), predictive studies should be more objective. In fact, such investigations will be encouraged to quantify the poor predictability of models in the early phases of the work, otherwise they will be unable to demonstrate improvement. By purposefully not advocating a particular approach to improving our prediction of B_z, we believe we can be more objective, and, additionally, identify and pursue those methodologies that are most promising. Moreover, we can provide a framework for other researchers in the field to test their techniques, subjecting them to the same metrics as are imposed on us.

2.2 Scientific Objectives

The primary scientific objective of this investigation is to apply a comprehensive suite of prediction algorithms, including new ones developed during the course of the investigation, to provide the most reliable and robust prediction of IMF B_z. Additionally, we will also compute the other magnetic field components (B_x and B_y) as well as the bulk solar wind speed, density, and plasma temperature. To accomplish this goal requires that we:

1. Develop an easy-to-use framework for testing B_z prediction algorithms;
2. Develop a rigorous set of metrics and skill scores that include estimates of uncertainty;
3. Test currently most promising statistical and numerical modeling techniques;
4. Develop a prioritized set of new predictive techniques;
5. Provide the completed framework as an open source resource to the scientific community.

2.3 Perceived Impact of the Proposed Work

The proposed data analysis and modeling techniques, together with the application of metrics and skill scores represent a comprehensive and unique approach to solving a specific applied problem. If successful, our work will substantially advance our ability to forecast potentially geo-effective conditions with 24 (and up to 4 days) advance notice. At the very least, we will be able to accurately quantify the current state of predictive capabilities, discriminate between approaches that are likely to be successful from those that are not, and provide a framework that future investigations can incorporate.

Although not objectives of the proposed effort, our work will naturally address a number of important questions. For example, where should spacecraft be positioned, and what quantities should they measure, to give us the optimum data to 'assimilate' into the model? Sensitivity studies exploring such options can be used to constrain the optimum vantage points.

2.4 Technical Approach and Methodology

2.4.1 Introduction

In this investigation, we propose to apply a range of statistical and physics-based modeling techniques to uncover the optimum combination of approaches that can provide the most accurate prediction of B_z. Our approach is unique, we believe, in that from the outset, our goals are defined not by the approaches that we propose to implement, but by the single objective we seek: The best prediction of B_z at 1 AU. Because of this, we envisage that some approaches will wain in importance during the course of the project, while others will prosper. Nevertheless, in the sections that follow, we briefly describe what we believe are the most promising initial avenues that we will pursue. Additionally, a crucial component of the proposed work will be the development of a comprehensive set of metrics for assessing the predictive capabilities, including the ongoing calculation of skill scores. Finally, to maximize the value of the proposed work, and leave a working legacy, all development will – from the outset – be undertaken within a publicly available SVN repository, allowing other researchers access to, and the ability to contribute to the tools developed.

2.4.2 Statistical Models for Predicting B_z

Over the last decade or so, data mining and knowledge discovery approaches to revealing and predicting a range of phenomena within datasets have blossomed (Han et al., 2006; Pang-Ning et al., 2006). Moreover, a variety of publicly available and scientifically vetted toolkits have been developed for applying these techniques (e.g., Weka, DBMiner, Rapidminer, Rattle, MCL++). These machine learning approaches have not been widely embraced in the solar/heliophysics community. In part, this may have been due to their limited ability to make accurate predictions as well as the hurdle of understanding new concepts. However, a more likely explanation is that, for the most part, such techniques do not directly address the question of *why* a particular relationship exists. In fact, the second layer in a typical neural network is known as the "hidden" layer. For our purposes, however, this understanding of the physical processes is of secondary importance to prediction. Additionally, recent advances and refinements, as well as the incorporation of these routines into a wide variety of scientific analysis packages (e.g., *RWeka* in R and *mlpy* in Python) allows them to be applied relatively easily.

2.4.2.1 Artificial Neural Networks: Artificial Neural networks (ANNs) are models inspired by neurons in the brain. Conceptually, ANNs consists of three layers: an input layer, one or more "hidden" layers, and an output layer. Figure 3 illustrates the connectivity of nodes (i.e., "neurons") between these layers. Biological connections between neurons have been simulated in ANNs using links between the neurons (i.e., graph nodes) that have adaptive weights, and an activation function that can cause the neuron to fire with a value dependent on the type of function and the neuron's activation threshold. Intuitively, ANNs (particularly the multi-layer feed-forward techniques) can be thought of as a nonlinear generalization of a linear filter. In fact, this picture is illustrated within the field of Space Weather, where ANNs have found modest success in the prediction of geomagnetic indices based on solar wind parameters (Lundstedt, 1996). Of more relevance, Wintoft & Lundstedt (1999) attempted to predict the daily solar wind speed at 1 AU using a neutral network based on flux tube expansion factor (computed from a potential field source surface (PFSS) model). They were able to obtain correlation coefficients as high as 0.57 for monthly averages, which dropped to 0.38 for daily averages. These disappointing results are likely more due to the inputs than the technique, as Riley & Luhmann (2012); Riley et al. (2014) have shown that expansion factor alone is a relatively poor predictor of solar wind speed, the distance from the coronal hole boundary providing much higher correlations (~ 0.75).

Liu et al. (2011) applied a particular type of ANN technique, known as support vector regression (SVR) to forecast solar wind velocity. They were able to predict speeds with an accuracy of over 90%, however, this was limited to a 3-hour prediction, and did not perform well during CME intervals. The authors suggested that the inclusion of expansion factor, f_s, would include the prediction of the velocity pattern. We suggest that while this may modestly improve prediction, there are other solar-related parameters, such as the distance from the coronal hole boundary (DCHB), which have been shown to produce measurably better correlations with *in-situ* measurements (Riley & Luhmann, 2012; Riley et al., 2014).

Qahwaji et al. (2008) applied ANN (the Cascade Correlation Neural Network algorithm) as well as Support Vector Machines (SVM) techniques to predict CMEs based on the proper-

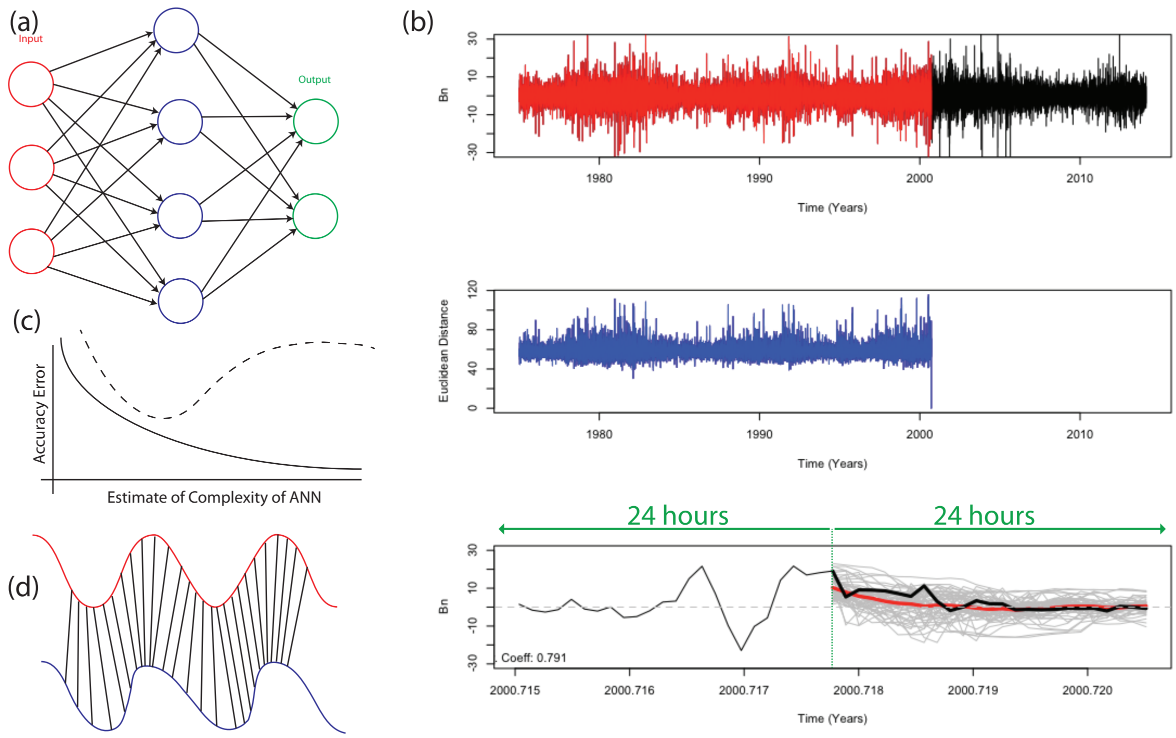

Figure 3: *(a) Illustration of three-layer neural network. (b) Example of pattern matching through in-situ measurements, predicting B_z for next 24 hours. The top panel shows B_n, with the data used for the prediction highlighted in red. The middle panel shows the Euclidean distance for the pattern shown in the left-hand side of the bottom panel. The right-hand-side of the bottom panel shows predictions based on the top-20 patterns identified from the Euclidean distance, together with the actual data (black) and the ensemble average prediction (red), based on these realizations (correlation coefficient ~ 0.8). (c) Schematic of accuracy error when the ANN is overfit (solid line) or not (dashed). The estimate of complexity could be the number of iterations, neurons, or hidden layers. (d) Illustration of dynamic time warping for B_z time series.*

ties of flares (intensity, duration, duration of decline/growth). While promising, the approach suffers from several issues: False positives; use of only M/X-class flares; and the neglect of erupting filaments/prominences. More fundamentally, for our purposes, it is not obvious how this connection between flares and CMEs can improve our ability to predict B_z. On the other hand, within the framework of an ensemble approach, it straightforward to include the prediction as an input, allowing a statistical model to tune the weights.

In recent years, ANNs have been significantly refined and applied to a number of fields, being adopted into the field of “Machine Learning.” They remain, however, simple to implement and the algorithms lend themselves easily to being parallelized. Although they have been criticized for obfuscating the underlying physics (although see Fu (2003) for a different perspective), since the primary goal of our proposed work is prediction, and not understanding, this limitation should not penalize the approach.

The most common error in using ANNs is the lack of respect for the complexity of the models they are capable of generating (whilst still being only visible as a “black box” to the user). To understand this more clearly, consider the analogy of the curve-fitting problem: Intuitively, we invoke "Occam's razor", that is, the principle of parsimony, and choose the lowest-order polynomial that still fits our data. Thus the main challenge with the complexity of ANNs is that the nonlinear models that the NN built to reflect our data are so complex that they must be treated as a "black box." This often leads to the creation of ANNs that show excellent results during the training phase (interpolation), but which perform poorly when applied to the prediction of future, or unseen data (extrapolation). The greater the number of hidden layers, or even nodes in the ANN, the higher the risk. This problem is well known in fields outside Solar Physics (Astion et al., 1993; Alman & Ningfang, 2002; Tetko et al., 1995), leading researchers to validate the quality of their ANN using the entire error of the training charts and not simply a single error measure (Figure 3). In our investigation, we will monitor our ANNs for such errors and minimize their impact.

2.4.2.2 Pattern Matching Approaches: A precursor to the pattern matching approaches that we plan to incorporate into analysis was developed by Chen et al. (1996). Their insight relied on identification of well-defined and coherent magnetic cloud structures. They suggested, and presented examples whereby certain classes of long-duration events could be identified at 1 AU by *in-situ* spacecraft relatively early through the rotation of the large-scale magnetic field and the trailing portion (up to 80%) could be predicted by extrapolating a sinusoidal pattern forward in time. Particularly for flux ropes whose axes is located in the ecliptic plane and points along the azimuth direction, the variation in B_z approximates a ±sine function. If the helicity of the event is such that a northward excursion of B_z occurs first, a reliable prediction of both the timing and magnitude of the all-important southward excursion can be made. Although this may appear to limit the applicability of this technique to a certain class of events, it is precisely this signature that results in the largest geomagnetic consequences. We outline several refinements to this basic idea, including flux-rope fits to real-time data and machine learning concepts, such as pattern recognition in the proposed work. A third novel idea is to combine cone-model simulations that predict the dynamic evolution of the ejecta in which the flux rope is embedded with the simple force-free flux rope extrapolation (e.g., Russell et al. (1990)); the deformation of the plasma being used to non-linearly stretch the force-free flux rope predicted from the small section of observed time series. Not only will these approaches provide more robust predictions, they will provide key estimates of the uncertainty in the prediction, which was not addressed in the pioneering study by Chen et al. (1996).

Our illustration here has been limited to a simple one-to-one mapping of the profiles. Given the variability in scale size of transient phenomena, we anticipate incorporating dynamic time warping (DTW) to uncover similarities in time series where the structures are either moving at different speeds or have expanded in a non-linear way (Keogh & Pazzani, 1998; Ding et al., 2008; Keogh et al., 2001, 2004). The two time series are 'warped' non-linearly along the time axis. Packages exist in *R* for applying DTW to time series. This approach has been successfully applied to a wide range of topics, including speech recognition, climatology, and aviation (Rakthanmanon et al., 2013).

Figure 3(b) illustrates how simple pattern matching techniques can be employed to predict, at least near term the evolution of B_z, particularly when there are coherent variations, such as during the passage of a magnetic cloud. The prediction was made by taking the previous 24 hours of B_z data from a point shortly after a large ICME was detected by Wind and sliding this swath of data along the entire time series, computing the Euclidean distance (essentially an estimate of χ^2) between it and the full dataset. We then chose the best 20 matches and used the data during the following 24 hours as a prediction for the current time. These are shown by the grey traces. The ensemble average of these traces is shown in red and compared with the actual data (thick black line). The correlation was ~ 0.8. No attempt was made to "tune" or tweak this comparison. We believe it demonstrates that these techniques, and important refinements to them may be a useful addition to our prediction tool chest. One noteworthy improvement involves the implementation of a dynamic time warping algorithm (Figure 3 (c)). CMEs in the solar wind, whether or not they are formally "self similar" occur on a range of spatial (and hence temporal) scales from a few hours (Moldwin et al., 2000) to several days. Yet the magnetic structure contained within them is often similar, being some approximation to a force-free flux rope. The speed at which the ICME travels through the ambient solar wind, as well as the structure of the ambient solar wind further modulates these features.

Solar feature detection in SDO observations have also yielded some surprises, challenging previously established "rules of thumb." Martens et al. (2014), for example showed that the hemispheric chirality preference for filaments, established by Pevtsov et al. (2003) appears to disappear during parts of the decline of cycle 23 and is less well established during the onset of cycle 24. To the extent that filaments and sigmoids can be used as proxy indicators of ICME properties, and particularly for any inference on flux rope axis orientation, these quantitative trends can be incorporated into a comprehensive model.

2.4.3 Empirically-based Models of ICMEs

The distinction between "empirical" and "physics-based" models for ICMEs and/or their associated shocks is somewhat subjective and arbitrary. However, tracking features through various remote solar and *in-situ* measurements surely qualifies as an empirical model. DeForest et al. (2013) outlined a comprehensive analysis of the December 12, 2008 CME from its origin in the low corona through its interception with near-Earth spacecraft. While the study identified the origin of various features of the ICME, from it predictive standpoint, it also illustrated the unique capabilities of heliographic imagers to provide real-time global information on the structure as it propagates from the Sun to 1 AU. With a goal of providing 24-hour predictions, these observations can provide either direct information that can be incorporated into a prediction of plasma conditions that will occur at 1 AU, or, as constraints to global MHD models, which themselves may contain the important magnetic field information. At the least, such observations can provide key timing information for both the CME-driven shock, the following sheath region, and the flux rope/ejecta itself.

Several techniques have been developed to compute the arrival time of IP shocks. The Shock Time Of Arrival (STOA) model (Dryer & Smart, 1984), for example, rely on the time, location, drift speed, and duration of radio burst measurements. Kadinsky-Cade et al. (1998) used IMP-8 and Wind data to evaluate this model during the interval 1991-1997 and found that the average deviation between the predicted and actual arrival of the shock was ± 36 hours. Riley & McComas (2009) developed a set of fluid conservations to infer near-Sun properties of CMEs from their in-situ measurements. The technique was validated using numerical model results and a case study (November 22, 2001 CME) and can be used to estimate the transit time for ICMEs. Other approaches relying on white-light observations and effective acceleration profiles (Gopalswamy et al., 2001) and low-frequency radio measurements (Reiner et al., 2007) have also been proposed. Gopalswamy et al. (2005) estimated typical errors of ~ 10.7 hours in the predicted arrival time of the shock.

Bieber et al. (2013) presented a technique for inferring B_z from neutron monitor data, based on the idea that cosmic rays hitting the Earth have already passed through and interacted with the IMF upstream of the Earth. They found that the correlation between the predicted and measured B_z for 161 ICMEs ranged from $\sim 85\%$ to $\sim 60\%$ for predictions 1-3 hours into the future.

Numerous attempts have been made to relate magnetic structures observed in connection with a solar eruption and the subsequent properties of the magnetic flux rope observed at 1 AU. These have met with mixed success. In an early study, Hoeksema & Zhao (1992) correlated the source regions of five strong $-B_z$ events observed at 1AU, inferring that the computed field orientation at the point in the corona were the acceleration of the ejecta ceased matched with *in-situ* measurements. In another study, Leamon et al. (2004) related the magnetic properties of magnetic clouds to their associated active region sources. In particular, they looked at the helicity, finding that the twist within the cloud was typically an order of magnitude greater from the twist measured in the active region. Bothmer & Schwenn (1998) found a good correlation between MCs and large quiescent filament disappearances. Specifically, the magnetic orientation of the prominence axis, the polarity of the overlying field lines, and the helicity pattern of the filaments "agreed well" with the properties of the MCs. While qualitative inferences of the degree of match cannot be used in a quantitative prediction scheme, the results suggest that a statistically-based approach, such as a neural network or regression technique, which incorporates all of the potentially relevant features, may yield a predictive capability.

Fry et al. (2003) compared predictions of shock arrival times with the HakamadaAkasofuFry version 2 (HAFv.2) model finding that the model performance was as good, but no better than the simpler STOA model.

Vandegriff et al. (2005) described a technique for predicting the arrival time of interplanetary shocks at Earth, not based on the forward propagation of signatures discerned near the Sun but based on energetic particle signatures observed at 1 AU prior to the shock's arrival. They found that the predicted arrival time of the shock was accurate to within 8.9 hours 24 hours in advance and 4.6 hours when the shock was 12 h from the spacecraft. Although these uncertainties are not obviously better than NOAA/SWPI's estimates of ± 6 h based on a prediction made several days in advance, they offer a unique extra piece of information – that the event has generated a shock that will intercept Earth. Cone-model-based predictions do not necessarily have the ability to predict whether the shock will necessarily impact the Earth, or whether it will be a glancing collision. Moreover, in principle, analysis of the energetic particle signatures can provide insight into the strength of the following shock, which in turn, provides albeit indirect information about the properties of the driver ejecta.

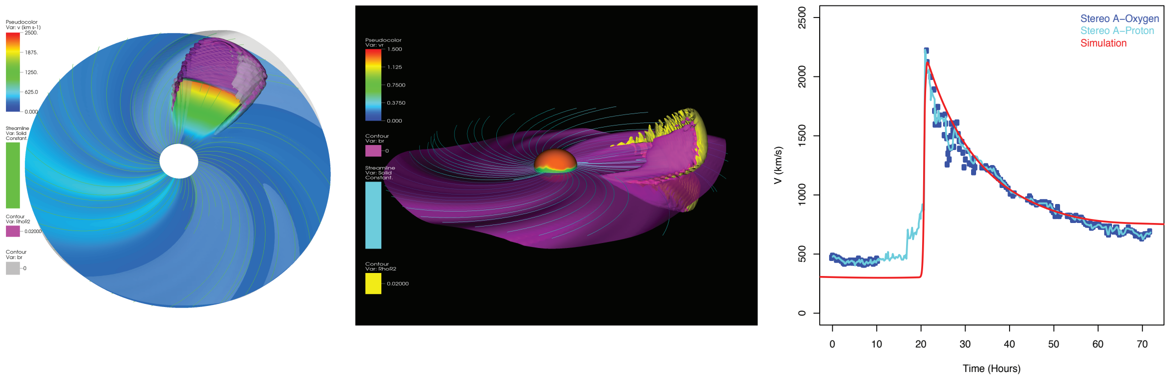

Figure 4: *(a) and (b): Two 3-D views of ICME (simulation 15) as it approaches 1 AU. The legends to the left of each panel indicate what parameters are being displayed. A selection of interplanetary magnetic field lines are also shown. (c): Comparison a hypothetical observer (simulation 15) at 1 AU with STEREO A velocity measurements. This ICME was launched with an initial speed of 2500 km/s, consistent with remote solar observations.*

2.4.3.1 Cone-Model ICMEs A subset of PSI's modeling suite, CORHEL, consists of the WSA-Enlil model, with a cone-model CME generator, which recently became the National Weather Service's first space weather model, running real-time at NOAA's Space Weather Prediction Center (SWPC) (Farrell, 2011). Although Enlil is a global MHD code, it currently treats CMEs as plasma ejecta, specifying the location of 'launch' (on a sphere of height $30R_S$), direction, speed, density, and duration. We have developed a similarly capable version of the cone model (Riley et al., ????) using a recently developed massively parallel interplanetary version of our global MHD algorithm (Lionello et al., 2013). In either case, however, the lack of magnetic structure within the ejecta precludes it from giving meaningful predictions of the flux rope field. It can, however, provide information on the field in the sheath region, which, at least for fast events, has been draped over the ejecta. We have also begun investigating techniques for incorporating magnetic fields within the ejecta. One idea that appears promising was motivated by more theoretical results from TDm simulation results, suggesting that substantial reconnection could occur under a flux rope leading to a spheromak-like structure. Although we found that the *in-situ* profiles from such an event can mimic the observed rotations within observed ejecta (e.g., Figure 5(c)), it remains to be seen whether this can lead to a robust estimate of flux rope fields. An alternative idea, which would allow one to retain the simplicity of the cone-model approach would be to "kinematically" incorporate, say, a force-free type flux rope close to the Sun (e.g., Owens et al. (2006)), using the evolving velocity field to distort the flux rope as it is propagated through the inner heliosphere (Riley & Crooker, 2004). This would require knowledge of the orientation of the flux rope axis, which could be inferred from several remote solar observations, although not without substantial caveats.

Krall & Cyr (2006) developed a simple technique for parameterizing a three-dimensional flux rope based on white-light observations. The synthetic images produced capture many of the features observed in the eruption of simple limb events, and, although they did not develop it, in principle, the technique can be used to estimate the orientation of the flux-rope axis. This, however, would require observations at quadrature to predict B_z at Earth, or an *in situ* spacecraft located at quadrature to test the efficacy of the model to predict B_z orientations. As with many of these techniques, the concept is simple, its implementation is straightforward, and it's ability to predict B_z has not been ascertained. Thus, it represents an ideal method for testing.

2.4.3.2 Predicting B_z within Sheath Regions of ICMEs Echer et al. (2008) estimated that the sheath regions of ICMEs are responsible for $\sim 25\%$ of all intense ($Dst < -100$nT) geomagnetic storms. Thus, an accurate estimate of these intervals alone is an important component of any prediction scheme. Although current cone models do not contain intrinsic magnetic fields within the ejecta, they should, in principle, be able to reproduce the essential large-scale features of CME sheaths. We envisage at least two processes that could contribute to producing a non-zero B_z within sheath regions. The first, as already noted is due to the process of draping. As ambient solar wind is swept up, field lines are caught up and dragged through the solar wind by the ejecta. A second, and more subtle effect, however, concerns the effects of the shock itself. It is well known, based on the Rankine-Hugoniot (RH) relations, that the magnetic field changes direction across a shock (e.g., Priest (1982)). Specifically, the magnitude of the tangential (relative to the shock normal) component is not conserved. Across a fast-mode shock this translates into a compression, with the resulting effect that the magnetic field vector is deflected away from the shock normal. Thus, with prior knowledge of the likely properties of the shock as passes Earth, the amplification of the tangential component can be estimated and converted into a B_z component (Horbury, Personal Communication, 2014). In principle, this effect should already be directly included within global cone-model simulations such as Enlil. However, this information can also be derived independently from Heliospheric Imager observations when available, which will provide estimates of the shock/sheath orientation and strength, coupled with more reliable estimates of the modeled ambient solar wind at Earth. Moreover, global MHD models of CME ejecta themselves are not capable of reproducing the small-scale (but large amplitude) fluctuations produced by waves and turbulence, opening up the possibility of using the RH approach to estimate the amplification of this contribution to B_z. In the proposed investigation, we would use various combinations of these inputs, as well as other heuristic approaches to estimating the characteristics of CME-sheath regions (e.g., Russell & Mulligan (2002)), testing the predictions with a large sample of previously identified CME-driven shocks (e.g., Pulupa et al. (2010)).

2.4.4 Physics-based Models of ICMEs

The modeling group at PSI have developed state-of-the-art global models of CME initiation and evolution for many years (e.g., Linker & Mikić (1995); Linker et al. (1996, 2001); Riley et al. (2003); Linker et al. (2003); Riley et al. (2007); Riley et al. (2008); Riley et al. (2008)). However, they are – at least at present – time consuming; both in terms of computer time and personnel. Thus, we developed an interim solution that bridges the gap between "empirically-based" models and these "first principles" models. The current version of this practical CME generator relies on a modified description of the TItov-Démoulin flux rope model. The essential difference between these types of models and either empirically-based models or statistical models is that we anticipate deriving a greater physical understanding of the processes that give rise either to the eruption and/or evolutionary properties of CMEs and their interplanetary counterparts. For this specific investigation, however, this promised understanding is of value only if allows for a better prediction. We further anticipate that, at least initially, these types of models will not provide forecasts that are as accurate as the simpler approaches. However, it is likely that this will change in the foreseeable future. Moreover, by including them at the outset, we provide quantitative skill scores that these or any comparable model must exceed to be considered better from a predictive standpoint.

2.4.4.1 Pre-Existing Flux Rope Eruptions: The modified Titov-Démoulln (TDm) model is an analytic model of a force-free coronal flux rope embedded within the solar corona The eruption can be initiated by either allowing the initially out-of-equilibrium flux rope to erupt, or by first establishing an equilibrium for the flux rope, then introducing a perturbation (such as canceling flows near the polarity inversion line), which then produces an eruption. Because the initiation process is not treated self-consistently, we cannot address fundamental questions about how, and when the flux rope formed. On the other hand, for the proposed work, answers to such questions are only of value if they qualitatively effect the resulting magnetic structure of the ejecta. We have found that the more self-consistent the model, the more idealized, or generic the solutions tend to be. By constraining the boundary conditions with a limited number of easily specifiable parameters, however, we are able to better match actual observations. Additionally, we believe it is possible to use available observations to constrain the initial characteristics of the flux rope (location, orientation, axial field strength), providing the key ingredients for a predictive type of model, in much the same way that cone-model simulations are routinely driven by parameters derived from white-light observations (Odstrcil & Pizzo, 2009).

It is worth noting, however, that our investigation will not be restricted to applying the models described here. Should we find that another semi-empirical approach is likely to yield better predictions, we will employ that approach. The group at the University of Michigan, for example, pioneered the use of an out-of-equilibrium model for CMEs by superimposing the TD flux rope onto a coronal solution (Manchester IV et al., 2008; Lugaz et al., 2011). This approach has yielded valuable insight into the evolution of flux ropes through the corona and into the solar wind, but also suffered from notable limitations, including the introduction of artificial forces into the system as well as an overestimate of the likely energy available to the flux rope. The approach is straightforward to implement, however, and may lead to better predictions of B_z provided the inputs can be tuned and/or estimated from available observations. Similarly, Kunkel & Chen (2010) applied the erupting flux rope model (Chen, 1996) to reproduce, among other parameters, the magnetic field within the interplanetary flux rope. If this can be applied in a predictive manner, it is simple to implement and apply, thus lending itself to testing within this framework.

paragraph"First Principles" Models of CMEs

Our group at PSI has, and continues to investigate fundamental issues related to the initiation and evolution of CMEs in the corona and solar wind (e.g., Mikić et al. (2013)) and to improve the physics describing the origin of the slow and fast components of the solar wind (Lionello et al., 2014) through the incorporation of waves and turbulence. As noted, these types of simulations are extremely time consuming and could not, at least currently,

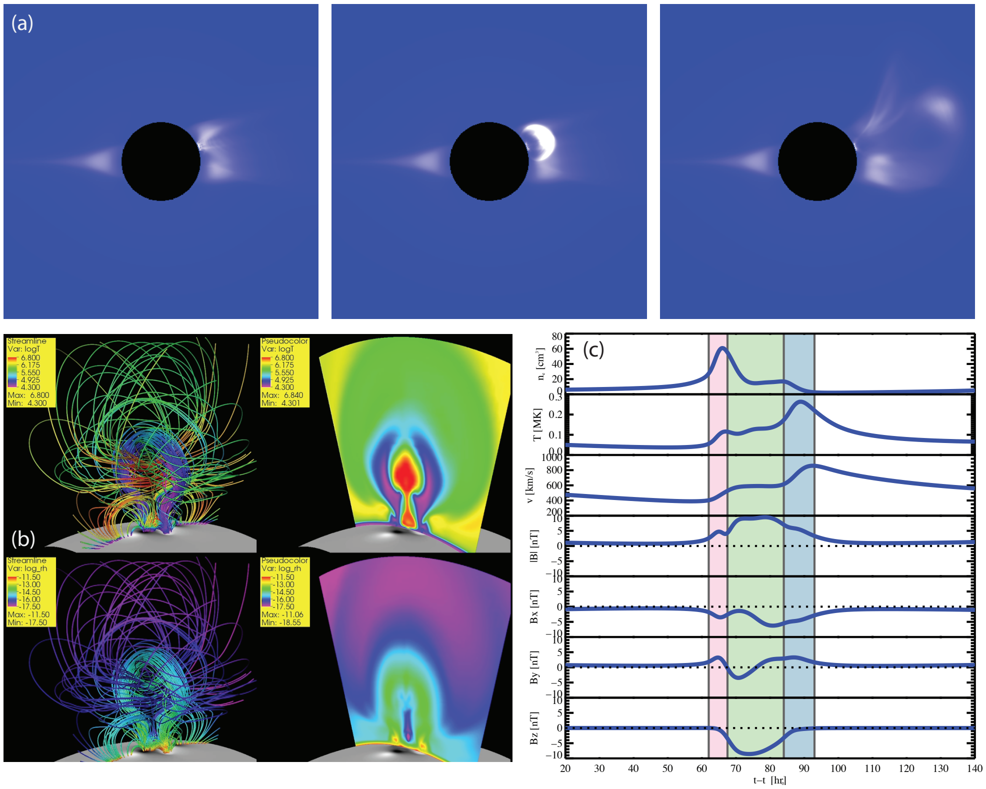

Figure 5: *(a) White-light images (polarization brightness) for the simulated CME eruption in (b). (b) Eruption of the TD flux rope after canceling flows at the photosphere are introduced. (b) Left: Magnetic field lines colored by the plasma temperature and; Right: plasma temperature and density in a plane passing through and parallel to the flux rope. (c) Time history of local plasma variables at 1AU for a point situated in the path of the ICME. The magnetic field vectors have been converted to Geocentric Solar Ecliptic (GSE) coordinates. The red highlighted region indicates the compressed solar wind downstream of the ICME. The green region indicates the flux system of the ICME. The blue region indicates the trailing plasma, which was compressed by a high speed stream behind it.*

be incorporated into a predictive scheme for forecasting B_z. However, we recognize that as these models improve and more computational resources are provided for their use, it may become feasible for them to find a place within the modeling framework. Thus, it is necessary to include them, to the extent possible, within any comprehensive framework. The development of this code is being funded as part of other investigations, thus, we propose to leverage this work here, principally, by analyzing model results for specific events, identifying where these model results are better or worse than the other approaches and identifying how such models can play a predictive role in the future. The questions we are asking using first-principle models may also provide insight into this proposed investigation. For example, why is there such a disparity between the twist observed in the source region of the CME and the measured twist of the ICME? (Leamon et al., 2004). Why are only a fraction of ICMEs observed to be, or at least contain flux ropes? (Riley & Richardson, 2013). How are ICMEs magnetically connected back to the Sun? (Riley et al., 2004).

Another exciting possibility exists through the use of MHD model results to map solar sources to their locations at 1 AU (Riley et al., 2008). Although the process of relating coronal signatures with their interplanetary counterparts would not initially, or perhaps definitively provide a one-to-one correspondence, such analyses could be used to identify the key solar signatures, which, *if present*, would predict a corresponding feature at 1 AU. This could, in turn, be combined with the statistical approaches defined above: Data mining techniques, such as those developed to recognize specific features in SDO remote solar observations (Martens et al., 2012; Banda & Angryk, 2009) could be used to define a necessary grouping of features with certain characteristics that correspond to a particular profile at 1 AU. This would be guided and/or confirmed through the physics-based simulations.

2.4.4.2 "Hybrid" Predictions: An intriguing possibility lies in combining the predictions from different types of empirical and physics-based CME models. For example, one could combine either white-light or HI observations with magnetic field observations of the flux rope observed by Messenger at Mercury. Although this would only work under certain conditions and geometries, it could provide a unique prediction of what the combined magnetic/plasma structure of the CME was near to the Sun, but sufficiently far that it could be evolved either kinematically (e.g., Riley & Crooker (2004)) or using a a (magneto)hydrodynamic algorithm. Alternatively, one could use the cone model, which has arguably the best capabilities of propagating the ejecta and associated shock to 1 AU, to specify the large scale non-magnetic properties of the event at 1 AU. Into this, a force-free, TDm, or even first-principles model could be embedded. The rationale is that for the force free model, there is no information about its propagation through the heliosphere, while for the latter two, the initial eruption speed may not be accurately produced by the simulation leading to inaccuracies in it's evolution between the Sun and 1 AU.

Typically, our global model solutions are constructed from coupled coronal and heliospheric models, where the output from the coronal solution is used to drive the heliospheric simulation. However, for the purposes of predicting solar wind structure only at one point in the heliosphere, e.g., at Earth, it is not clear that a global heliospheric model is required. In fact, we have shown that a simple mapping procedure that takes into account the dynamic evolution of solar wind streams can evolve solar wind structure from near the Sun to 1 AU (Riley et al., 2011). It has not yet been established whether the same technique could be successfully applied to ICMEs; however, given its simplicity and speed it should be included as a component in our comprehensive modeling approach for evolving structure near the Sun to 1 AU.

2.4.5 The Ambient Solar Wind

While prediction of the ambient solar wind in the absence of any transient phenomena is likely not a high priority from the perspective of geo-effective consequences (although all-clear forecasts are important), the role of the quiet-time solar wind is crucial in the evolution and hence prediction of large excursions in B_z produced by ICMEs. Fast CMEs, for example, are more significantly distorted and decelerated when slow and dense ambient wind lies ahead. Moreover, the properties of the wind into which the CME propagates will be modified by its passage: Large-scale deflections across the CME-driven shock will alter the perpendicular components of the IMF. Additionally, the presence of large-amplitude Alfvén waves in fast solar wind will result in the large and temporally compressed B_z fluctuations.

We have developed robust and relatively accurate models of the ambient solar wind over the last decade, making direct comparisons with observations (Riley et al., 2001, 2002, 2003; Riley, 2007b, 2010; Riley et al., 2010, 2011, 2012,?). We have found that the modeled results are capable of reproducing *in-situ* measurements with a typical accuracy of 0.75 (Riley et al., 2014), during relatively stable conditions (declining phase and solar minimum) and in the absence of obvious transient activity (Riley et al., 2014). These models, however, are demonstrably quite sensitive to the input magnetograms (Riley, 2007a; Riley et al., 2012; Riley et al., 2013a). To substantially improve our predictions, requires a thorough analysis of these inputs, which should include the best reconstruction of polar fields and, optimally, incorporate temporal evolution. As part of other funded studies, we are investigating which approaches most accurately reproduce *in-situ* measurements.

2.4.5.1 Persistence and Probabilistic Forecasts Although our goal will be to provide real-time, event-driven predictions, at a minimum, we can use combinations of persistence and probabilistic forecasts to make crude estimates for B_z in the absence of any other reliable information. Persistence refers to the assumption that the Sun today is exactly as it was one rotation earlier. If so, we can use those data to predict the next 24 hours or even 27 days, obviously with much larger uncertainties. Probabilistic forecasting summarizes what is known about likely future events by assigning a probability to a range of outcomes. This may offer some value during intervals surrounding stream interfaces (McPherron & Siscoe, 2004). Such techniques were employed in terrestrial meteorology prior to the development of sophisticated global circulation models. It is not yet clear whether space meteorology has yet crossed the threshold whereby models outperform these statistical approaches. Thus, it seems prudent to include them in our arsenal.

2.4.5.2 Waves and Turbulence Our discussion so far has focused on the large-scale variations in the IMF. However, a substantial amount of power is contained within higher frequency waves and turbulence. From a prediction point of view, these are the most challenging to attempt to forecast. However, from a geo-oeffective standpoint, their phase information is not as important as their statistical properties. Thus, we can superpose a contribution from waves and turbulence, based on properties of the predicted large-scale field at that time which contains contributions up to the highest frequencies, and which is statistically indistinguishable from observations. The properties of the turbulent solar wind are well established both as a function of solar wind (e.g., slow/fast) and phase of the solar cycle. The magnetosphere's response to them too will likely not depend on their relative phase (Merkin et al., 2007), at least beyond some threshold frequency.

2.4.6 Relevant Data

Our proposed work will rely on the full range of datasets produced by NASA missions. Time series data of the IMF field vectors will of course be important, as well as speed, density, and temperature to validate model predictions. Additionally, remote sensing white light images will drive cone-model simulations and provide key clues about the three-dimensional

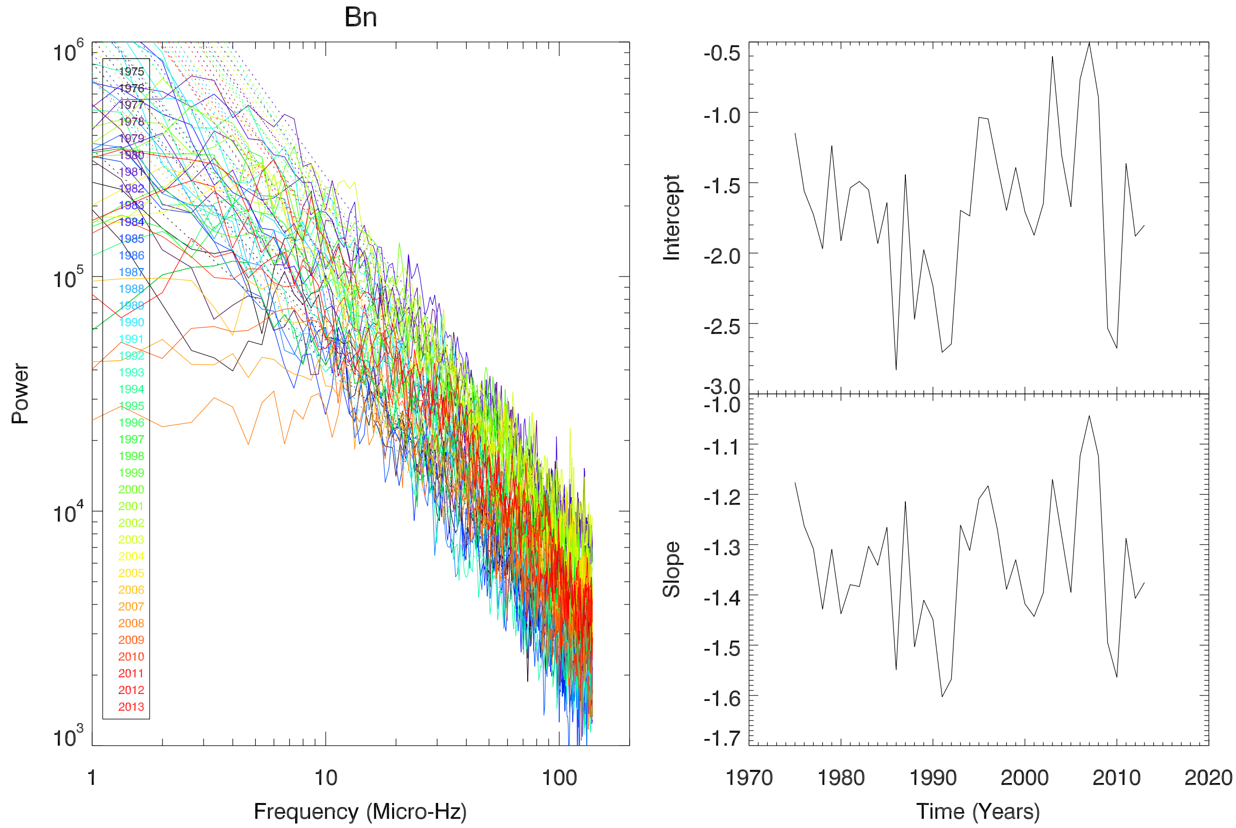

Figure 6: *Ensemble-averaged (21) power spectra of B_z for each year since 1975. (a) shows the spectra for each year. (b) and (c) show the intercept and slope, respectively, for a least-squares fit to the spectra above 20 micro-Hertz.*

structure of the ejecta, including, possibly, the orientation of the flux-rope axis. Heliospheric Imager data from STEREO will provide important information on the propagation and evolution of some ICMEs. EIT and X-ray images will provide important feedback on the quality of global model solutions and may provide "actionable" information for making predictions. Sigmoid structures, dimming regions, and EIT waves, for example, can be used to provide timing information during an eruption. On the other hand, while relevant, it is not *a priori* clear whether it will lead to useful information from a forecasting perspective. Similarly, flare data can be used as a measure of the speed of the ensuing ejecta, but it's relationship to the orientation of the ICME's magnetic field is likely tenuous. Photospheric magnetic field measurements, while providing a direct measurement of the parameter we are hoping to estimate at 1 AU must undergo significant transformation before reaching the Earth. Although several tentative steps have been taken to relate magnetic features in the photosphere to 1 AU observations of magnetic clouds (ref), the link remains weak. Finally, sub-surface measurements of flows...

We can leverage the work by the Montana State University group, who have developed a feature recognition algorithm for SDO observations (Schuh et al., 2014). Various combinations of different structures/phenomena can be incorporated into the pipeline we are proposing and tested for efficacy.

2.4.7 Ensemble Modeling

A crucial advance that propelled terrestrial modelers forward was the development of ensemble modeling techniques, and we can implement some of these approaches in our proposed investigation. Ensemble forecasting is defined as a method of prediction that relies on the use of a representative sample of possible future states to derive a prediction. One of the appealing aspects of such an approach is that it offers a rigorous method for computing confidence bounds of the solution by estimating the uncertainty in the ensemble (Wilks, 2006). Moreover, the mean of the ensemble of forecasts is, or should be more accurate than the forecast from any individual member; the reason being that the random, or unpredictable regions of the forecast tend to cancel one another, while the aspects of the forecast that the majority of the models agree on are not removed (Warner, 2010).

Ensemble modeling techniques have only recently and tentatively been applied to the heliophysics environment (e.g., Riley et al. (2012); Riley et al. (2013); Pizzo (2014)). However, they have been developed, tested, and rigorously applied within the terrestrial weather community for more than a decade, thus, providing a wealth of resources that can be adapted to space weather phenomena. For our specific objective, however, we can simplify matters, by defining the following ensemble model:

$$\hat{B}_z(t) = \sum_{i=1}^{K} \omega_i B_z(t), \tag{1}$$

where ω_i are the model weights and sum to one ($\sum_{i=1}^{K} \omega_i = 1$) (Wichard & Ogorzalek, 2004). Intuitively, equation (1) illustrates that those predictions that produce robust and reliable forecasts will also be associated with large weights. Less reliable models will be discounted and even removed from the ensemble if $\omega_i < \omega_c$.

2.4.8 Metrics, Probabilities, Skill Scores, and Confidence Intervals

Model validation refers to the task of identifying the strengths and weaknesses of a model through detailed comparisons of model output with observations. Similarly, metrics are used to measure the long-term trends in model improvement. Whereas model validation involves a comprehensive comparison of all available data, metrics target a small set of specific benchmarks. Our goal here is singularly focused: to obtain the best prediction of B_z with up to 24 advance warning. Although improvements in attaining this goal will undoubtedly help us predict other parameters of interest, attempting to optimize prediction of multiple parameters could adversely affect our ability to predict B_z.

We will define a number of metrics to track our progress. For example, we will employ both mean-square error (MSE) estimates, as well as event-based tests, such as the arrival time of sector boundary crossings. From these analyses, it is straightforward to define "skill scores," which track the improvements in the model over time. Once defined, metrics are simple to compute on a regular basis.

A confidence interval (CI) is used to indicate the reliability of an estimate. It is a measure of our confidence with a particular prediction, or alternatively the likely uncertainty, typically given at the 95% level. Thus, in addition to estimating retrospectively how well our

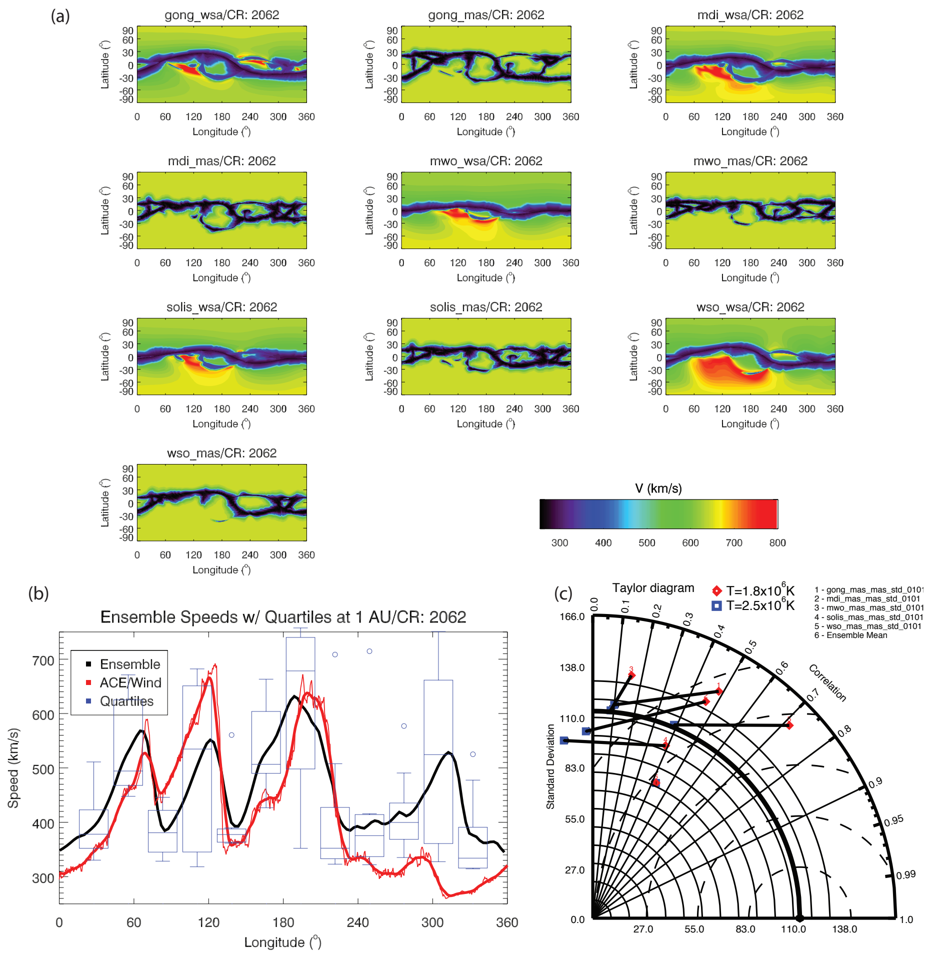

Figure 7: *(a) A subset of model realizations computed for CR 2062. (b) 'Whisker" plot of model realizations from (a) at* $30R_S$ *in the equatorial plane: Black line is the ensemble solution; red are in-situ measurements; and boxes with "whiskers" summarize the variability of the realizations. (c) A "Taylor diagram" for summarizing the performance of model runs. See Taylor (2001) for further details. The blue squares mark the new locations of the solutions when the base temperature is raised from* $1.8 \times 10^6 K$ *to* $2.5 \times 10^6 K$.

predictions performed, we will also provide prospective confidence intervals for all predictions made. CIs can be constructed in a number of ways. We could, for example, combine the output from different models (both statistical and mathematical) to derive an estimate of the CI. Not surprisingly, the CIs in a predicted time series of B_z would increase as the prediction stretched farther into the future. Importantly, they would provide a direct and quantitative indication of the quality of the prediction to the user. Moreover, as the prediction scheme improved, uncertainties would be seen to decrease, providing an additional metric to track.

The forecast effort is the difference between the actual measured value of B_z ($B_{z,t}$) and the forecast ($\hat{B}_{z,t}$):

$$E_t = B_{z,t} - \hat{B}_{z,t} \tag{2}$$

From this, we can develop a number of measures of the aggregate error, which describe the accuracy of the prediction over some interval $t = 1, N$. These include the mean absolute error (MAE), mean absolute percentage error (MAPE), mean absolute deviation (MAD), percent mean absolute deviation (PMAD), mean squared error (MSE), and root mean squared error (RMSE). For illustrative purposes, and because of their application in space weather model comparisons, MSE and RMSE are:

$$MSE = \frac{\sum_{t=1}^{N} E_t^2}{N} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N} E_t^2}{N}} \tag{4}$$

Additionally, the forecast skill score (SS) is defined by:

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{ref}} \tag{5}$$

where MSE_{ref} is the mean square error calculated from some standard, reference model, typically the simplest and least accurate in the suite. Thus, SS serves as a measure of how well (or poorly) model refinements perform.

Thus far, we have assumed a goal of predicting B_z as a function of time. It is likely, however, that at least for some parties, a forecast probability would be more appropriate. For event predictions, such as the time of arrival of a shock, a high-speed stream, or a sector boundary crossing, this can be phrased simply by the probability that the event will, or will not occur within the next 24 hours, say. We might generalize this concept for B_z by estimating the probability that B_z exceeds some negative threshold, say, -10nT for some period of time, say, 6 hours, again within the next 24 hours, say. These values would be situation-dependent: Different customers would likely have different requirements if the forecast probability was to be useful to them. 'All clear' forecasts based on the absence of such events would also likely have value to a wide range of users.

2.5 Proposed Contributions to the Focus Team Effort

NASA's Living With a Star program (LWS) seeks to "improve our understanding of how and why the Sun varies, how the Earth and solar system respond, and how the variability and response affects humanity in space and on Earth." More generally, NASA's Heliophysics Research Program aims to "Understand the fundamental physical processes of the space environment from the Sun to Earth,..." and to "enabl[e] the capability to predict the extreme and dynamic conditions in Space. Our objectives directly address these science goals by investigating the properties of CMEs, whose impact with the Earth's magnetosphere can cause a range of adverse effects. Ultimately, these results may drive, or at least contribute to models with predictive capabilities. By providing the most robust and accurate forecasts of B_z our investigation directly addresses these goals. In addition to providing a clear practical benefit to NASA's stakeholders in the form of the predictions, our work will also advance our understanding of the processes that give rise to sustained intervals of negative B_z; specifically fast ICMEs with embedded flux ropes.

2.5.1 Relevance to the scientific objectives of the Focused Topic

Our proposed investigation directly addresses all objectives of the focused topic. Specifically, we will provide a continuous estimate of B_z (as well as B_x, B_y, bulk solar wind speed, density, and temperature) primarily in the vicinity of Earth, but also at other strategically important locations, such as at Mercury and, in principle, along the orbits of Solar Orbiter, Solar Probe Plus, Sunjammer, and even upstream of Jupiter in support of the Juno mission. Our team includes representatives from the operational community and thus is sensitive to user needs. Moreover, our approach is comprehensive, yet prioritizes the impact and importance of fast ICMEs containing magnetic flux ropes as the primary concern from an operational standpoint.

2.5.2 Contributions to the Focused Science Team's effort

The team we have assembled combines all the skills necessary to meet the objectives of this FST. We are observers, modelers, theorists, and operators. We have expertise in the analysis of remote solar observations (white-light and HI, EUV, X-ray, and photospheric magnetic field observations) and in-situ measurements. We represent the key developers of relevant models from simple *ad hoc* techniques to state-of-the-art global, time-dependent MHD simulations. We are a truly cross-disciplinary team, including experts from the field of computer science, magnetospheric physics, solar, and heliospheric physics.

2.5.3 Metrics and milestones for determining success of proposed research

Our goal is to produce a framework that is available to the scientific community by the end of the project. In Section 2.6 we summarize the main milestones we will achieve. Our metrics for success will literally be metrics defined to quantitatively assess our ability to predict B_z. We believe that to claim to be able to make an accurate real-time, continuous prediction of B_z on the timescale of four years is probably naive. We do, however, believe that we can objectively quantify our initial abilities and track the progress during the course of the investigation. Based on some of the promising ideas presented here, we are confident that our forecasting abilities will improve substantially during the project because of the comprehensive and goal-driven focus of our team.

2.6 Outline of the General Plan of Work

This proposal is for four years. Our principal goals are to: (1) define a set of robust metrics for quantitatively, reliably, and efficiently assessing a range of prediction techniques for estimating IMF B_z at 1 AU with up to 24 hours advance warning; and (2) develop techniques for predicting B_z – both in the short and long term – that have the most likelihood of success.

- First Year
  - Develop a B_z Prediction Framework (BPF) that can be downloaded and installed by all team members. We will use a combination of shell scripts, Fortran, C/++,and R/Python served from and Subversion (SVN) repository, providing a platform-agnostic toolkit;

  - Populate the BPF with a limited subset of the predictive algorithms outlined above. These will be prioritized based on several factors, including likelihood of success and ease of implementation;
  - Assemble the necessary datasets as well as any relevant meta-datasets for driving, testing, and validating the BPF. Some will be stored directly in the SVN repository while others will be either disseminated from PSI or from their original source location;
  - Complete initial set of validation studies, computing relevant metrics and skill scores. Focus initially on predicting ambient solar wind and simple CME events;

- Second Year
  - Incorporate more models into BPF;
  - Incorporate dynamic time warping into time series pattern recognition technique;
  - Compare various physics-based model predictions for a set of ICMEs that cover the range of observed events;
  - Write set of papers documenting the BPF as well as the initial results from its application and present results at scientific meetings;

- Third Year
  - Develop a statistical model (e.g., NN) that incorporates automatically recognized solar features (dimmings, sigmoids, flares, etc);
  - Initiate contact with NASA (CCMC) and NOAA (SWPC) and discuss possible avenues for delivering and/or operationalizing the BPF;
  - Perform parametric studies of all models and combinations thereof to optimize the predictive capability of the BPF;
  - Revise and refine validation studies, computing a comprehensive set of skill scores and model uncertainties for all model combinations. Publish results in peer-reviewed journal and present current status at selection of meetings and/or workshops;

- Third Year
  - Investigate efficacy of running first principle models for specific events as a predictive model;
  - Complete development of BPF, including full documentation, and make available to scientific community;
  - Deliver a BPF to NASA/CCMC and/or NOAA/SWPC;
  - Write final set of peer-reviewed manuscripts comprehensively describing the BPF and the the various modules developed for it. Attend selection of scientific meetings to promote the use and availability of the BPF.

Although our proposed work appears ambitious, it is necessary if we are to make meaningful advances in our ability to predict B_z. Additionally, we will leverage a wide range of past and existing studies spearheaded by co-investigators of the present proposal. This include: (1) A recently-funded Air Force project (Jon Linker, PI) to use global CME models to predict solar wind conditions, and B_z, in particular, at 1 AU; (2) an NSF-funded project to understand extreme space weather events all the way from the Sun to the Earth (co-Is Pete Riley and Jon Linker); (3)

Our team has, or is developing most of the models we are proposing to test in this effort. Thus, the emphasis is more on development of the framework to quantitatively test them as well as the insight on how to best combine them to improve our predictive capabilities, not on the development of the models themselves, which has or is being funded in complementary work. A crucial aspect of our approach concerns our biweekly web-based meetings and twice-yearly face-to-face meetings. We believe that through these largely conceptualizing interactions, our team of experts will be able to formulate novel and related ideas that can then be tested within the BPF.

2.7 Management Plan

The investigation outlined here requires a well-directed and coordinated management plan. In addition to the intrinsic value bought by individual team members, we believe that a key factor in our success lies in the members' abilities to think outside of their specific areas of expertise. It is crucial that we meet on a regular basis to develop and test new ideas and report back on progress made. We propose to hold monthly web based meetings (using Citrix's GoToMeeting) as well as two in-person meetings, one in Southern California and the other in Boulder, Colorado. Each team member is fully aware and has agreed to the specific and focused goals of the proposed work and will devote their effort to achieving them. We anticipate that as certain avenues of research prove more or less promising, our emphasis may fall into adjacent and even unforeseen areas.

Pete Riley (PI, Predictive Science, San Diego, CA) will manage, and be responsible for the completion of the proposed research. He will work with Co-I's Roberto Lionello, Chris Russell, Roger Ulrich, Vic Pizzo, Curt de Koning, Alysha Reinard, Todd Hoeksema, and Yang Liu, and collaborators Jon Linker, Tim Horbury and Matt Owens, to undertake the tasks outlined here. Dr. Lionello will..., etc. In addition, we currently support undergraduate interns on a variety of computer- and science-related projects, and we will include them in portions of this investigation, as appropriate. The study will be carried out in collaboration with Drs. Horbury and Owens at Imperial College and the University of Reading, respectively, who will provide expertise and insight during the planning stages of the investigation as well as interpretation of the results. This will complement projects they are currently managing at their institutions.

In addition to the specific areas of expertise outlined above, a crucial element in the selection of the team members lies in their breadth of study over the years. This FST's objective is unique in its specific objective, requiring that the team be able to adapt to initial results and possibly move in directions that they had not originally considered. For example, should it be found that sub-surface inputs are only marginally relevant, Dr. Reinard's expertise in analyzing and interpreting *in situ* measurements will assume a larger role. Similarly, Dr.

Pizzo's vast expertise in modeling both from first principles to empirically based approaches demonstrate his ability to adapt to even more obscure modeling approaches, including those encapsulated under the "machine learning" moniker. Dr. Riley has worked both with statistical and physics-based models, and has analyzed a wide array of both remote solar and *in situ* measurements. All team members are both aware of, and have agreed that their role and contribution to the success of this proposal may require them to undertake new and perhaps unanticipated studies in pursuit of the effort's goals.

References

Owens, M. J., Merkin, V. G., & Riley, P. 2006, J. Geophys. Res., 111, A03104

Alman, D. H., & Ningfang, L. 2002, Color Research & Application, 27, 122

Astion, M., Wener, M., Thomas, R., Hunder, G., & Bloch, D. 1993, Clinical chemistry, 39, 1998

Baker, D., Li, X., Pulkkinen, A., Ngwira, C., Mays, M., Galvin, A., & Simunac, K. 2013, Space Weather, 11, 585

Banda, J. M., & Angryk, R. A. 2009, in Fuzzy Systems, 2009. FUZZ-IEEE 2009. IEEE International Conference on, IEEE, 2019–2024

Bieber, J., Evenson, P., Kuwabara, T., & Pei, C. 2013, in AGU Fall Meeting Abstracts, Vol. 1, 2146

Bothmer, V., & Schwenn, R. 1998, Ann. Geophys., 16, 1

Burton, R. K., McPherron, R. L., & Russell, C. T. 1975, J. Geophys. Res., 80, 4204

Chen, J. 1996, J. Geophys. Res., 101, 27499

Chen, J., Cargill, P. J., & Palmadesso, P. J. 1996, Geophysical research letters, 23, 625

DeForest, C., Howard, T., & McComas, D. 2013, The Astrophysical Journal, 769, 43

Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., & Keogh, E. 2008, Proceedings of the VLDB Endowment, 1, 1542

Dryer, M., & Smart, D. 1984, Advances in Space Research, 4, 291

Echer, E., Gonzalez, W., Tsurutani, B., & Gonzalez, A. 2008, Journal of Geophysical Research: Space Physics (1978–2012), 113

Farrell, P. 2011, www.bu.edu/cas/news/press-releases/cism/

Fry, C., Dryer, M., Smith, Z., Sun, W., Deehr, C., & Akasofu, S.-I. 2003, Journal of Geophysical Research: Space Physics (1978–2012), 108

Fu, L.-M. 2003, Neural networks in computer intelligence (Tata McGraw-Hill Education)

Gopalswamy, N., Lara, A., Manoharan, P., & Howard, R. 2005, Advances in Space Research, 36, 2289

Gopalswamy, N., Lara, A., Yashiro, S., Kaiser, M. L., & Howard, R. A. 2001, J. Geophys. Res., 106, 29207

Gosling, J. T., & McComas, D. J. 1987, Geophys. Res. Lett., 14, 355

Han, J., Kamber, M., & Pei, J. 2006, Data mining: concepts and techniques (Morgan kaufmann)

Hoeksema, J. T., & Zhao, X. 1992, Journal of Geophysical Research: Space Physics (1978–2012), 97, 3151

Horbury, T. S., & Balogh, A. 2001, J. Geophys. Res., 106, 15929

Kadinsky-Cade, K., Quigley, S., & Ginet, G. 1998, Trans. Am. Geophys. Un, 79, F712

Keogh, E., Chakrabarti, K., Pazzani, M., & Mehrotra, S. 2001, Knowledge and information Systems, 3, 263

Keogh, E., Chu, S., Hart, D., & Pazzani, M. 2004, Data mining in time series databases, 57, 1

Keogh, E. J., & Pazzani, M. J. 1998, in KDD, Vol. 98, 239–243

Krall, J., & Cyr, O. S. 2006, The Astrophysical Journal, 652, 1740

Kunkel, V., & Chen, J. 2010, The Astrophysical Journal Letters, 715, L80

Leamon, R. J., Canfield, R. C., Jones, S. L., Lambkin, K., Lundberg, B. J., & Pevtsov, A. A. 2004, J. Geophys. Res., 109, 5106

Linker, J. A., Lionello, R., Mikić, Z., & Amari, T. 2001, J. Geophys. Res., 106, 25165

Linker, J. A., & Mikić, Z. 1995, Astrophys. J. Lett., 438, L45

Linker, J. A., Mikić, Z., Lionello, R., Riley, P., Amari, T., & Odstrcil, D. 2003, Phys. Plasmas, 10, 1971

Linker, J. A., Mikić, Z., & Schnack, D. D. 1996, in A.S.P., ed. K. S. Balasubramaniam, S. L. Keil, & R. N. Smartt, Vol. 95, 208

Lionello, R., Downs, C., Linker, J. A., Török, T., Riley, P., & Mikić, Z. 2013, Ap. J., 777, 76

Lionello, R., Velli, M., Downs, C., Linker, J. A., Mikić, Z., & Verdini, A. 2014, Ap. J., 784, 120

Liu, D., Huang, C., Lu, J., & Wang, J. 2011, Monthly Notices of the Royal Astronomical Society, 413, 2877

Lugaz, N., Downs, C., Shibata, K., Roussev, I., Asai, A., & Gombosi, T. 2011, The Astrophysical Journal, 738, 127

Lundstedt, H. 1996, Surveys in geophysics, 17, 561

Manchester IV, W. B., Toth, G., Sokolov, I., Zurbuchen, T. H., & Kota, J. 2008, in in Particle Acceleration and Transport in the Heliosphere and Beyond, Proceedings of the 7th Annual International AstrophysicsConference, ed. G. Li, Q. Hu, O. Verkhogladova, G. Zank, R. P. Lin, & J. Luhmann, Vol. 1039 (Kauai, Hawaii: AIP, Woodbury, New York), 273

Martens, P. C., Yeates, A. R., & Pillai, K. G. 2014, in IAU Symposium, Vol. 300, IAU Symposium, 135–138

Martens, P. C. H., Attrill, G. D. R., Davey, A. R., Engell, A., Farid, S., Grigis, P. C., Kasper, J., Korreck, K., Saar, S. H., Savcheva, A., Su, Y., Testa, P., Wills-Davey, M., Bernasconi, P. N., Raouafi, N.-E., Delouille, V. A., Hochedez, J. F., Cirtain, J. W., Deforest, C. E., Angryk, R. A., de Moortel, I., Wiegelmann, T., Georgoulis, M. K., McAteer, R. T. J., & Timmons, R. P. 2012, Solar Phys., 275, 79

McPherron, R. L., & Siscoe, G. 2004, Space Weather, 2

Merkin, V., Owens, M., Spence, H. E., Hughes, W., & Quinn, J. 2007, Space Weather, 5

Mikić, Z., Török, T., Titov, V., Linker, J. A., Lionello, R., Downs, C., & Riley, P. 2013, in American Institute of Physics Conference Series, Vol. 1539, American Institute of Physics Conference Series, ed. G. P. Zank, J. Borovsky, R. Bruno, J. Cirtain, S. Cranmer, H. Elliott, J. Giacalone, W. Gonzalez, G. Li, E. Marsch, E. Moebius, N. Pogorelov, J. Spann, & O. Verkhoglyadova, 42–45

Moldwin, M., Ford, S., Lepping, R., Slavin, J., & Szabo, A. 2000, Geophysical research letters, 27, 57

Norquist, D. C. 2013, Space Weather, 11, 17

Norquist, D. C., & Meeks, W. C. 2010, Space Weather, 8

Odstrcil, D., & Pizzo, V. J. 2009, Solar Phys., 148

Pang-Ning, T., Steinbach, M., Kumar, V., et al. 2006, WP Co

Pevtsov, A. A., Fisher, G. H., Acton, L. W., Longcope, D. W., Johns-Krull, C. M., Kankelborg, C. C., & Metcalf, T. R. 2003, Astrophys. J., 598, 1387

Pizzo, V. 2014, Submitted to Space Weather

Priest, E. R. 1982, Solar magneto-hydrodynamics (Dordrecht, Holland ; Boston : D. Reidel Pub. Co. ; Hingham), 469 p.

Pulupa, M., Bale, S., & Kasper, J. 2010, Journal of Geophysical Research: Space Physics (1978–2012), 115

Qahwaji, R., Colak, T., Al-Omari, M., & Ipson, S. 2008, Solar Physics, 248, 471

Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., Zakaria, J., & Keogh, E. 2013, ACM Trans. Knowl. Discov. Data, 7, 10:1

Reiner, M. J., Kaiser, M. L., & Bougeret, J.-L. 2007, Astrophys. J., 663, 1369

Riley, P., , Ben-Nun, M., Lionello, R., Downs, C., Torok, T., & Linker, J. A. ????

Riley, P., , & Linker, J. A. 2014, Submitted to Space Weather

Riley, P. 2007a, Astrophys. J. Lett., 667, L97

—. 2007b, Journal of Atmospheric and Solar-Terrestrial Physics, 69, 32

Riley, P. 2010, in Twelfth International Solar Wind Conference, Am. Inst. Phys. Conf. Proc., ed. M. Maksimovic, K. Issautier, N. Meyer-Vernet, M. Moncuquet, & F. Pantellini, Vol. 1216, 323

Riley, P., Ben-Nun, M., Linker, J., Mikic, Z., Svalgaard, L., Harvey, J., Bertello, L., Hoeksema, T., Liu, Y., & Ulrich, R. 2013a, Solar Physics, 1

Riley, P., Ben-Nun, M., Lionello, R., Downs, C., Torok, T., Linker, J., Mikic, Z., Schwadron, N., & Gorby, M. 2013b, in AGU Fall Meeting Abstracts

Riley, P., & Crooker, N. U. 2004, Astrophys. J., 600, 1035

Riley, P., Gosling, J. T., & Crooker, N. U. 2004, Astrophys. J., 608, 1100

Riley, P., Linker, J., Mikic, Z., & Lionello, R. 2008, in in Particle Acceleration and Transport in the Heliosphere and Beyond, Proceedings of the 7th Annual International AstrophysicsConference, ed. G. Li, Q. Hu, O. Verkhogladova, G. Zank, R. P. Lin, & J. Luhmann, Vol. 1039 (Kauai, Hawaii: AIP, Woodbury, New York), 279

Riley, P., Linker, J. A., Lionello, R., & Mikic, Z. 2012, J. Atmos. Solar-Terr. Phys., 83, 1

Riley, P., Linker, J. A., & Mikić, Z. 2001, J. Geophys. Res., 106, 15889

—. 2002, J. Geophys. Res., 107, DOI 10.1029/2001JA000299

Riley, P., Linker, J. A., & Mikic, Z. 2012, in Solar Wind 13, ed. G. P. Zank (AIP)

Riley, P., Linker, J. A., & Mikić, Z. 2013, J. Geophys. Res., 118, 600

Riley, P., Linker, J. A., Mikić, Z., Odstrcil, D., Zurbuchen, T. H., Lario, D., & Lepping, R. P. 2003, J. Geophys. Res., 108, 1272

Riley, P., Lionello, R., Linker, J. A., Mikic, Z., Luhmann, J., & Wijaya, J. 2011, Solar Phys., 145

Riley, P., Lionello, R., Mikić, Z., & Linker, J. 2008, Astrophys. J., 672, 1221

Riley, P., Lionello, R., Mikić, Z., Linker, J., Clark, E., Lin, J., & Ko, Y.-K. 2007, Astrophys. J., 655, 591

Riley, P., Luhmann, J., Opitz, A., Linker, J. A., & Mikic, Z. 2010, J. Geophys. Res. (Space Physics), 115, 11104

Riley, P., & Luhmann, J. G. 2012, Solar Phys., 277, 355

Riley, P., & McComas, D. J. 2009, J. Geophys. Res., 114, 9102

Riley, P., Mikić, Z., & Linker, J. A. 2003, Ann. Geophys., 21, 1347

Riley, P., & Richardson, I. G. 2013, Solar Phys., 284, 217

Riley, P., Stevens, M., Linker, J. A., Lionello, R., Mikic, Z., & Luhmann, J. G. 2012, in American Institute of Physics Conference Series, Vol. 1436, American Institute of Physics Conference Series, ed. J. Heerikhuisen, G. Li, N. Pogorelov, & G. Zank, 337–343

Russell, C. T., & McPherron, R. L. 1973, J. Geophys. Res., 78, 92

Russell, C. T., Mewaldt, R. A., Luhmann, J. G., Mason, G. M., von Rosenvinge, T. T., Cohen, C. M. S., Leske, R. A., Gomez-Herrero, R., Klassen, A., Galvin, A. B., & Simunac, K. D. C. 2013, Ap. J., 770, 38

Russell, C. T., & Mulligan, T. 2002, Planet. Space Sci., 50, 527

Russell, C. T., Priest, E. R., & Lee, L.-C. 1990, Physics of magnetic flux ropes, Vol. 58 (American Geophysical Union)

Schuh, M. A., Banda, J. M., Bernasconi, P. N., Angryk, R. A., & Martens, P. C. H. 2014, Sol. Phys., 289, 2503

Taylor, K. 2001, J. Geophys. Res., 106, 7183

Tetko, I. V., Livingstone, D. J., & Luik, A. I. 1995, Journal of Chemical Information and Computer Sciences, 35, 826

Vandegriff, J., Wagstaff, K., Ho, G., & Plauger, J. 2005, Advances in Space Research, 36, 2323

Warner, T. 2010, Numerical Weather and Climate Prediction (Cambridge University Press)

Wichard, J., & Ogorzalek, M. 2004, in Proceedings of International Joint Conference on Neural Networks (IJCNN 2004), Citeseer, 1625–1629

Wilks, D. 2006, Statistical Methods in the Atmospheric Sciences, International Geophysics Series (Academic Press)

Wintoft, P., & Lundstedt, H. 1999, Journal of Geophysical Research: Space Physics (1978–2012), 104, 6729

A Facilities and Equipment

At Predictive Science Incorporated (PSI), San Diego, California, we maintain a set of Macintosh workstations that are more than adequate for meeting the data-processing, storage, and analysis requirements of the proposed investigation. Any necessary large-scale computations of CME initiation and evolution will be performed on on massively parallel computers at NASA (Pleiades) and NSF (Stampede) through other contracts, for which we currently maintain, and envisage receiving further allocations sufficient to perform the proposed work. We maintain a Subversion (SVN) repository for all of team-oriented code development projects (e.g., CORHEL) and will develop a complete SVN package to serve the algorithms and data for the proposed work. This will be used by all team members and, in due course, opened up to the scientific community.

B Curriculum Vitae

C Current and Pending Support

D Budget Justification

Narrative and Details

Direct Labor Summary

Travel

Other Direct Costs

Indirect Rates

Detailed Yearly Budgets