Data-Driven Solar Surface Flux Transport Modeling with Uncertainty Quantification

Katherine Keegan
Department of Mathematics, Emory University
400 Dowman Drive, Atlanta, GA, 30322, USA
katherine.emiri.keegan@emory.edu

Nina Bonaventura
Steward Observatory, University of Arizona
933 N. Cherry Ave., Tucson, AZ, USA
n.r.bonaventura@gmail.com

Plinio Guzman
Fused.io
pgzmnk@gmail.com

Nishu Karna
Center for Astrophysics | Harvard & Smithsonian
Cambridge, MA, 02138, USA
nishu.karna@cfa.harvard.edu

Bibhuti Kumar Jha
Southwest Research institute
1301 Walunt St., Boulder, CO 80302, USA
bibhuti.jha@swri.org

Shea Hess Webber
HEPL, Stanford University
452 Lomita Mall, Stanford, CA 94305
shessweb@stanford.edu

Spiridon Kasapis
Department of Astrophysical Sciences
171 Broadmead St., Princeton, NJ
skasapis@princeton.edu

Andrés Muñoz-Jaramillo
Southwest Research institute
1301 Walunt St., Boulder, CO 80302, USA
andres.munoz@swri.org

Abstract

Solar active regions (ARs) are areas of increased magnetic flux on the Sun's surface. Their downstream effects, including phenomena such as solar flares, Solar Energetic Particle events (SEPs) and coronal mass ejections (CMEs), can impair Earth- and space-based human infrastructure. To mitigate these dangers, the prediction of such events and their precursors –namely the ARs– are necessary, and uncertainty quantification (UQ) of these predictions are crucial for subsequent human decision-making. In this manuscript, we present Active Region Characterization and Analysis of Dynamics and Evolution (ARCADE), an ongoing effort to provide AI-driven and UQ-equipped forecasting pipeline for solar surface magnetic field evolution. ARCADE is trained by propagating a ResNet through a numerical simulation of the solar surface magnetic field. In this work, we validate our numerical integrator against the state-of-the-art Advective Flux Transport model (AFT) and demonstrate that ARCADE offers high-accuracy forecasts with consistent uncertainty estimations. Moreover, ARCADE is the first model of surface magnetic evolution to produce forecast of the emergence of flux several hours into the future. Finally, we provide a convenient user interface for forecasting with UQ, accessible upon request. This work is a promising first step for hybrid physics-ML methods for solar magnetic field forecasting with UQ, offering clear interpretability and usability for heliophysicists and stakeholders in space weather prediction.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: AI4Science.

1 Introduction

The growing interest in human spaceflight and the increased reliance on satellites and other space-based infrastructure have made reliable space weather forecasting a critical need. While progress has been made in predicting individual space weather events such as solar flares (5; 9), coronal mass ejections (13), and solar energetic particles (1; 10), forecasting their source regions (11; 10)—solar active regions (ARs)—still remains underdeveloped. Active regions are localized areas on the Sun's surface dominated by strong, concentrated magnetic fields. They appear as sunspots in visible light and are the origins of most major eruptive solar phenomena, making their early detection and characterization essential for space weather prediction. Detecting the growth or emergence of ARs before they are visible at the surface requires analyzing the rising magnetic flux, the dynamics of the Sun's interior (22; 20; 12) and most importantly the transport of solar surface magnetic flux.

The current state-of-the-art (SOTA) approaches to surface flux transport (SFT) are the Advective Flux Transport (AFT) model (20; 18) and the Air Force Data Assimilative Photospheric Flux Transport (ADAPT) (7) model. They have found widespread success and longevity in current space weather forecasting capabilities, and are focused mainly on modeling the evolution of the magnetic field around the solar far-side and at the poles, which are typically hidden from Earth's point of view. As a result, these models do not enable accurate short-term forecasts, which are critical for human operations dependent on space weather. Instead, they impose measurements of the magnetic field as seen from Earth, or estimations of the magnetic field based on far-side extreme ultraviolet observations (18). Due to their focus on long-term evolution, no SFT model is currently able to self-consistently grow or emerge new magnetic flux without assimilating observations. Additionally, no existing SFT model includes estimations of the surface velocity patterns that drive solar evolution, relying instead in a statistical approximation to these flows.

The objective of this work is to use AI to improve SFT modeling by enabling the simultaneous estimation of the small-scale flows that drive magnetic evolution and the simulated evolution of the solar surface magnetic field given the magnetic flux emergence and estimated velocity field. It is important to highlight that we are deliberately using small ResNets as part of this work (each with ~ 8K parameters). We aim to show that small models can produce competitive performance and consistent UQ estimates when acting in synergy with physical models.

1.1 Equation of magnetic flux transport

The mathematical model of SFT, guiding the time t evolution of the magnetic field $\vec{B} = \vec{B}(\theta, \varphi, t)$, arises from the ideal magnetohydrodynamics as the magnetic induction equation, with the colatitude $\theta \in [0, \pi]$ radians and the longitude $\varphi \in [0, 2\pi]$ radians:

$$\frac{\partial \vec{B}}{\partial t} = \eta \nabla^2 \vec{B} + \nabla \times \left[\vec{v} \times \vec{B}\right] + \vec{S}(\theta, \varphi, t), \tag{1}$$

where $\vec{v} = \vec{v}(\theta, \varphi, t)$ represents the velocity field governing the advection of the magnetic field strength and $\vec{S}(\theta, \varphi, t)$ represents the magnetic flux growth/emergence term, and η the solar turbulent diffusivity. If we deliberately restrict our model to only consider radial magnetic flux (a common SFT assumption), the equation simplifies to

$$\frac{\partial B_r}{\partial t} = \underbrace{\eta_h \nabla^2_{h,r} B_r}_{\text{Diffusion}} \underbrace{- \frac{1}{R_\odot \sin\theta} \frac{\partial}{\partial \varphi}[v_\varphi B_r] - \frac{1}{R_\odot \sin\theta} \frac{\partial}{\partial \theta}\left[v_\theta B_r \sin\theta\right]}_{\text{Advection}} + \underbrace{S_r(\theta, \varphi, t)}_{\text{Source}}, \tag{2}$$

where the radial magnetic field $B_r(\theta, \varphi) : \mathbb{R}^2 \to \mathbb{R}$ is now scalar-valued, $R_\odot$ is the solar radius (a scalar constant), and $v_\theta(\theta, \varphi, t)$ and $v_\varphi(\theta, \varphi, t)$ are the latitudinal and longitudinal components of the velocity flow.

Following standard practice, we split both $v_\theta(\theta, \varphi)$ and $v_\varphi(\theta, \varphi)$ into a global axisymmetric flow, and a small-scale turbulent flow. For colatitude we have

$$v_\theta(\theta, \varphi, t) = \texttt{MF}(\theta) + v'_\theta(\theta, \varphi, t), \quad \texttt{MF}(\theta) = -v_0 (\cos\theta)^p \sin\theta, \tag{3}$$

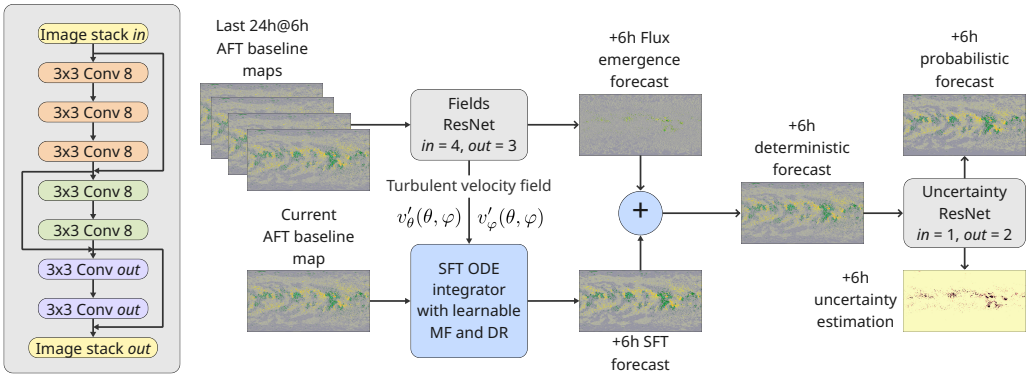

Figure 1: ARCADE ML pipeline: We use the AFT baseline (20) as our input and output data. We use small ResNets (left block; $\sim$ 7k parameters) to enable the generalization of our small-scale fields ($v'_\theta(\theta,\varphi)$ & $v'_\varphi(\theta,\varphi)$) and flux emergence ($S_r(\theta,\varphi,t)$) terms. We use `torchdiffeq` (2) to numerically integrate the SFT ODE (Eq. 2) and create the initial +6h forecast. Another small ResNet performs the uncertainty quantification.

where $\mathtt{MF}(\theta)$ represents the *meridional flow* (MF) (6), v_0 controls the peak pole-ward speed and p controls how sharply the flow peaks toward mid-latitudes. For longitude we have

$$v_\varphi(\theta,\varphi,t) = \Omega(\theta) + v'_\varphi(\theta,\varphi,t), \quad \Omega(\theta) = A + B\sin^2\theta + C\sin^4\theta, \tag{4}$$

where $\Omega(\theta)$ represents the *differential rotation* (DR) (8), and A, B, C are constants controlling the rotation rate and its slow down towards the poles.

Notably bereft of parametrization is the *source term*, $S_r(\theta,\varphi,t)$. This term accounts for the subsurface magnetic activity that ultimately gives rise to the observable magnetic activity at the surface, and can be thought of the source term for surface flux growth and new emergence. Due to inherent measurement limitations, one cannot fit the source term directly to data, but it may be learned in a semi-supervised manner concurrently with the remaining parameters in Equation 2. For simplicity, we will use “emergence" throughout this paper to indicate both AR growth and new emergence.

When combined, Equations 2-4 provide a collection of six learnable parameters, the unknown source flux emergence term $S_r(\theta,\varphi,t)$, and two turbulent velocity components $v_\theta(\theta,\varphi)$ and $v_\varphi(\theta,\varphi)$. The Active Region Characterization and Analysis of Dynamics and Evolution (ARCADE) project employs a **hybrid Ordinary Differential Equation-Neural Network (ODE-NN)** approach to simultaneously and robustly learn all nine terms on a **short time-horizon**, incorporates various methods for **uncertainty quantification (UQ)**, and offers **novel infrastructure for visualization and forecasting with new data** to support both heliophysics research and space-based operations communities.

2 Methodology

Our approach for SFT forecasts with learnable components takes inspiration from ClimODE (21), a neural ODE approach that respects underlying physics while integrating NNs for scalable parametrization of complex flows and fields.

Fig. 1 shows a summary chart of our hybrid physics-ML workflow. We explain how it connects together in the following sections: First, we describe our data and train/validation split (see Sec. 2.1). Second, we describe our numerical integrator and how it compares with a SOTA model (AFT, specifically; see Sec. 2.2). Third, we describe the combination of the ODE solver with our AI models to enable +6h forecast of the solar magnetic field in Sec. 2.3. Finally, we describe our UQ module in Sec 2.4.

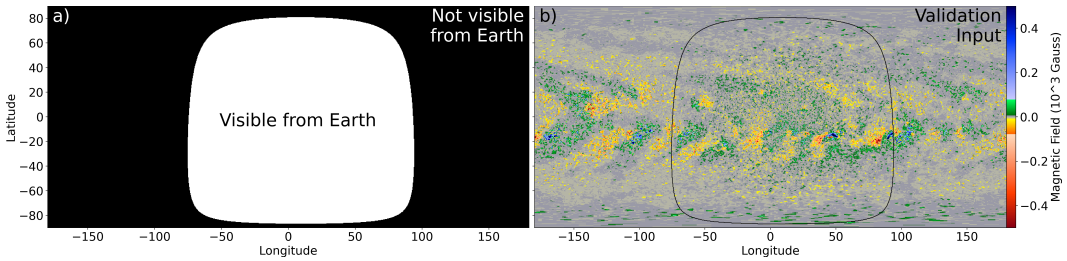

Figure 2: Hybrid observation/simulation AFT baseline map. **a)** Only a portion of the full solar surface is visible from Earth. AFT assimilates observations of the visible part of the Sun into its maps and uses SFT simulations to advance time in the parts of the solar surface that are not visible from Earth. **b)** AFT baseline map from Jan-01-2014. We used it as input for our SFT model validation (see Fig. 3). The assimilated visible part of the Sun is marked using a solid black line.

2.1 Data: The AFT Baseline

The AFT model (20; 19) has been used to create a baseline of magnetic maps covering the full solar surface during the Solar Dynamics Observatory (SDO; 14) era (2010-present) at a 6 hour cadence[1]. This corresponds to 22,364 maps at a resolution of 512 pixels in latitude and 1024 pixels in longitude. AFT maps are created by integrating the SFT equation (Eq. 2) using a prescribed set of global and convective flows, while assimilating near-side solar magnetograms observed by the SDO Helioseismic & Magnetic Imager (HMI) (16; 17). Fig. 2-b shows the AFT baseline map for Jan-01-2014. Fig. 2-a shows which part of the solar surface was visible from the Earth that day. By continuously assimilating observations from the near side, AFT maps provide an ideal dataset to develop, validate, and test our AI pipeline. On one hand, AFT maps contain the real observations of the near-side. On the other hand, the part of the Sun that is not visible from Earth is evolved by integrating the well-understood SFT equation, excepting the source flux term.

2.1.1 Data Splits

There is a significant amount of auto-correlation in solar data at the 6h-timescales we aim to forecast. Because of this, it is crucial to split the data in a way that minimizes the possibility of information leakage across training-validation-test splits. At the same time, it is important to retain a good representation of long-term solar cycle (quasi-decadal) variability. We follow the standard practice in heliophysics where data is split equally every year such that part of each year is present in all three splits. Our split, designed to ensure that there are flux emergence events in every split goes as follows (this split is applied for every year between 2010 and 2025):

Training: December, January, February, March, April, & May.

Buffers: June & November.

Validation: July & October.

Test: August & September.

This split places a full solar rotation of buffer between the Training and Validation splits, as well as two full rotations of buffer (including validation) between the Training and Test splits.

2.2 Numerical Integration

The SFT equation (Eq. 2) is an ODE with both temporal and spatial derivatives. For the spatial derivatives, we implement both first and second order derivatives using a convolutional layer with fixed weights, no biases, and with gradient tracking disabled. Throughout the model we use the 8th order centered differencing stencils.[2] (4)

[1]The AFT baseline can be downloaded from `https://data.boulder.swri.edu/lisa/AFT_Baseline/`

[2]Finite differencing stencils of different orders can be found at `https://www.ams.org/journals/mcom/1988-51-184/S0025-5718-1988-0935077-0/S0025-5718-1988-0935077-0.pdf`

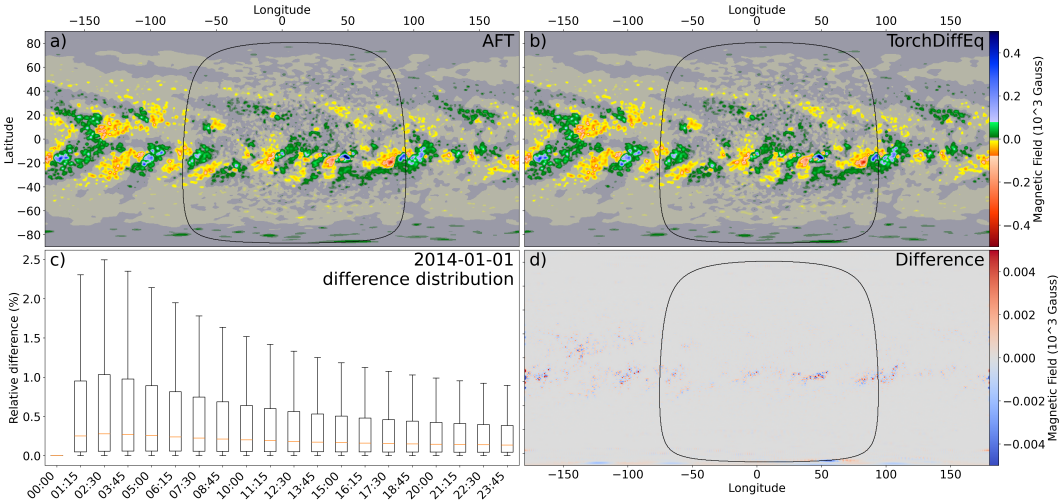

Figure 3: Validation against the SOTA AFT model. **a)** AFT simulation 24 hours into the future starting from the Jan-01-2014 AFT baseline map shown in Fig. 2-b. **b)** Output of our implementation, evolved 24 hours into the future using the same initial condition. **c)** Boxplot showing the distribution of relative absolute difference between the AFT evolution and our time-integrated approach. The results are shown relative to the AFT value and are shown as percents. The boxes indicate the first and third quartiles (i.e. 25% and 75% of the differences). The center notch represents the median error (less than 0.5% for all times). The whiskers indicate the Inter-Quartile Range. **d)** Difference between the 24h AFT and our SFT forecasts. Note the difference of two orders of magnitude between the two colorbar scales.

For each magnetogram, B_r^n, serving as an initial condition, we use the `torchdiffeq` (2) library to evolve B_r^n according to Equation 2 from time $t = n$ to $t = n + 1$. In particular, we use a fourth order Runge-Kutta solver with fixed $\Delta t = 15$ minutes. Solvers in `torchdiffeq` are fully implemented in PyTorch, enabling backpropagation through the numerical integration.

2.2.1 Numerical Grid and Boundary Conditions

For this work, we use a uniform grid with 1024 points in longitude and 512 points in colatitude. This is the resolution of the AFT baseline maps (see Sec. 2.1). Our boundary conditions are standard for the SFT community: Periodic boundary conditions in longitude and zero magnetic field at the poles. We implement our boundary conditions by padding our tensors prior to the application of the finite differencing scheme and later removing the padding once the integration step is done.

2.2.2 Validation against AFT

Before we combine the `torchdiffeq` numerical integration of the SFT equation with the AI model learnable component presented in Fig. 1, we validated the results of our numerical integration against AFT. To do this we start with an AFT baseline map and advance the simulation forward 24 hours to compare the results. Fig. 3 shows a comparison between AFT (Fig. 3-a) and `torchdiffeq` (Fig. 3-b). Fig. 3-c uses boxplots to show the evolution of the relative difference between AFT and `torchdiffeq`. In spite of some significant differences between the integration schemes (like a non-uniform grid in AFT), the difference is very small (a median of 0.2%). This difference is slightly larger in the first steps as both simulations deal slightly differently with noise, but at the end of the 24h simulation the difference has converged (see Fig. 3-d).

2.3 Hybrid ODE-NN Forecasts

In order to enable the short (6 hour) forecast of the solar surface magnetic field, it is necessary to have a way to estimate the small-scale flows that transport the magnetic field, as well as estimating how much additional flux will appear in the photosphere in the next 6 hours. To this date, there has not been a way to enable either of these tasks. The closest method for estimating velocity flows

is the imposition of a pre-calculated turbulent velocity field such as in the AFT model, while flux emergence uses the assimilation of observations with the SFT output on the visible side. To overcome these limitations, we train an AI model to estimate both velocity fields and flux emergence from past magnetic field evolution (see Fig. 1). In particular we use the last 24 hours worth of AFT baseline maps (4 maps) to estimate flows and fields. The flows are added to the global flows (MF and DR) and used by the `torchdiffeq` integrator to advance time by 6 hours. After this, the estimated flux emergence is added to produce the final forecast.

2.3.1 Neural Network Model

We use small ResNets with convolutional layers to parametrize all unknown functions in our pipeline. As shown in Fig. 1, our ResNets are built using 7 convolutional layers with three skip connections. Each convolutional layer has 8 channels. We use this architecture for both our fields and noise networks. These networks do not employ batch normalization due to small batch sizes in our experiments, and we employ LeakyReLU activations throughout all architectures. The LeakyReLU activation was chosen because the solar magnetic field has both positive and negative values and to reduce the chance of dead connections. Each ResNet has approximately 8k parameters.

As mentioned above, the choice of using small models is made deliberately to evaluate the potential of an ODE-NN hybrid in learning specific heliophysical quantities (the velocity fields and flux emergence terms). This design choice reduces the parameter space dimensionality significantly while ensuring interpretability, as all model-learned representations directly correspond to known physical processes governing solar magnetic field evolution, and ultimately lead to the competitive preliminary results we present here with validation against AFT simulation.

2.4 Uncertainty Quantification

It is well-understood that space-based observations, like all measurement-based data, carry irreducible uncertainties and noise due to instrument limitations, un-accounted-for probabilistic physical variables, and other forms of interference. To model this inherent stochasticity as well as the uncertainty in a model-generated output, uncertainty quantification is crucial for trustworthy adoption of AI-driven results. To this end, we construct a neural network (which we will refer to as a *noise network*) on top of the deterministic forecasts generated in Section 2.3, which returns a Gaussian mean and standard deviation over the deterministic forecast $\hat{B}_r^{n+1}$. This matches the approach to UQ modeling in ClimODE. We also include Fourier positional encoding as input to the noise network for interpretability of all input features. Moreover, we impose a baseline known level of data uncertainty, $\xi \in \mathbb{R}$, and train the noise network to learn any additional uncertainty. To train this noise network we use the following slightly modified *negative log-likelihood* (NLL) loss function on the predicted mean $\hat{\mu}^{n+1}$ and variance $\hat{\sigma}^{n+1}$:

$$\mathcal{L}_{\text{NLL}}\big(\hat{B}_r^{n+1} + \hat{\mu}^{n+1},\, \hat{\sigma}^{n+1},\, B_r^{n+1},\, \xi\big) = \text{GaussianNLL}(\tilde{B}_r^{n+1}, \tilde{\sigma}_i, B_r^{n+1}) \tag{5}$$

where $\tilde{B}_r^{n+1} = \hat{B}_r^{n+1} + \hat{\mu}^{n+1}$ is the magnetogram prediction adjusted for the learned forecast mean and $\tilde{\sigma} = \left|\hat{\sigma}^{n+1}\right| + \xi$ ensures that we are only able to learn a nonnegative addition to the baseline aleatoric uncertainty ξ. The $\hat{\cdot}$ notation denotes a raw model output. This NLL loss is added to the MSE loss on the raw predictions so that both models are trained simultaneously, allowing for *uncertainty-aware* training.

Except for our inclusion of known baseline uncertainty and using fixed Fourier positional encoding instead of a trainable positional encoding network, our approach to UQ is nearly identical to that in ClimODE, where it is termed as an emission head. We note that for full model UQ over the probabilistic forecasts, we would require a secondary, epistemic (or model-based) UQ technique, such as Laplace Approximation (3), to be used on the noise network itself. Although this remains future work, our current framework provides a strong probabilistic foundation for uncertainty-aware space weather model training.

3 Results

Our model is the first implementation of a hybrid ODE-NN in heliophysics. It is also the first time that an SFT model has been used to enable *short-term* solar magnetic evolution forecast including

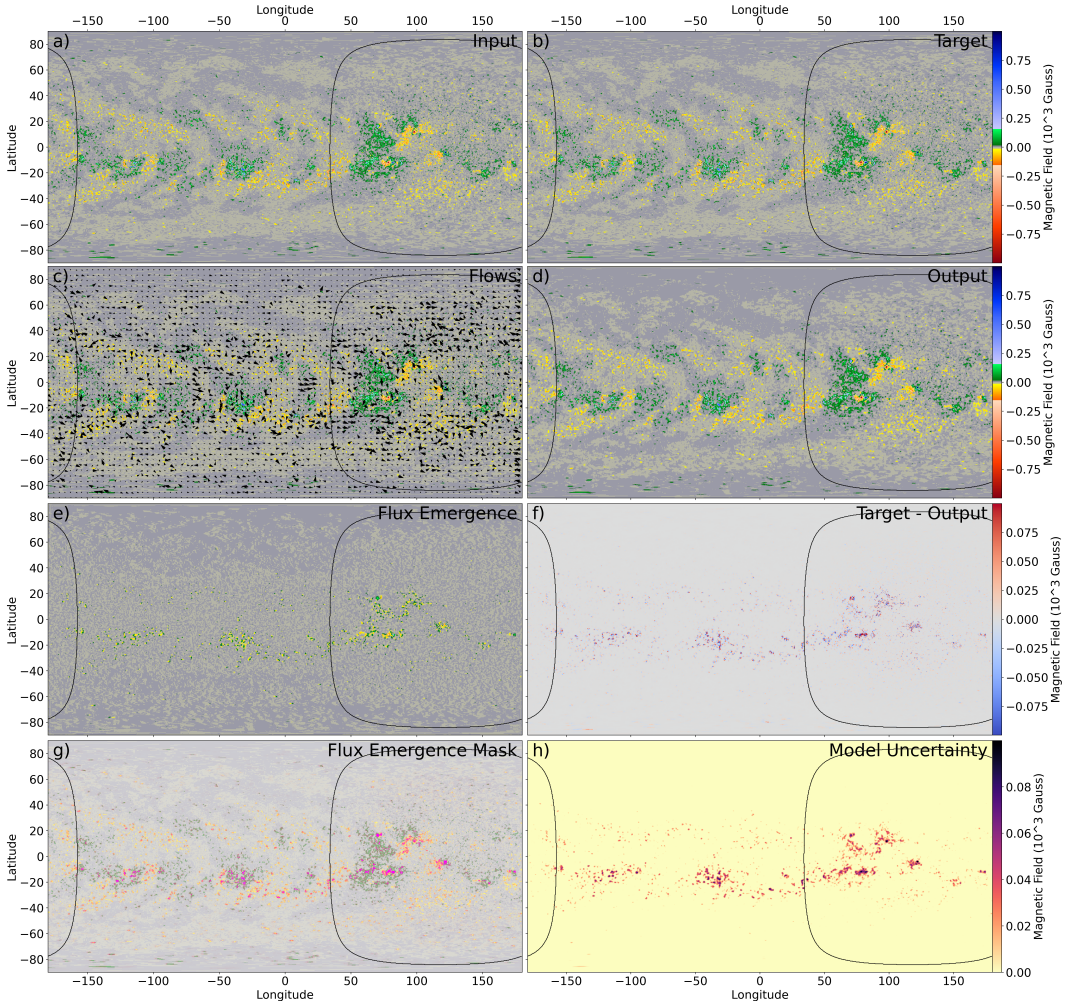

Figure 4: A 6h forecast example for Jun-06-2014 for the entire pipeline shown in Fig. 1. **a)** Input AFT baseline map. Visible part of the Sun is shown using a solid black line. **b)** Target AFT baseline map. **c)** Small-scale flows estimated by the fields ResNet. **f)** +6h forecast error. Note that the color-scale is one order of magnitude less than the one used for Input, Target, and Output. **g)** Flux emergence mask highlighting in magenta the regions for which our model forecasts flux emergence in the next 6 hours. **h)** Model uncertainty produced by the noise ResNet. Note that the colorscale is also one order of magnitude smaller than the Input, Target, and Output magnetic field.

effective forecasting of emerging flux. Fig. 4 shows a variety of visualizations of our results for Jul-6-2014: magnetic AR growth in our validation set. We would like to highlight the following:

Test Mean Square Error: The MSE calculated on our test set is 52.6 G^2. The square root of this value is 7.25 G. This is an encouraging result because it is lower than the instrumental uncertainty of SDO/HMI which is 15 G.

Small-scale velocity field: Fig. 4-c shows a vector plot over our input magnetic field. We find the average speed of the vectors in this quiver plot to be 14m/s. This value is lower than measured super-granular flows (15), but is likely a consequence of the fact that we use the same flow during the 6h integration step, which can also be seen as an average of time-varying cells.

Flux emergence term: Figs. 4-e & g show the signed and absolute value, respectively, of the flux emergence term. It is clear that the model is successful in capturing the growth of the AR, and its emerging magnetic flux successfully. However, considering that an AFT baseline map can only have flux emergence on the visible side (inside the solid black line), we note that the flux emergence term

is not pure, because the CNN is using the flux emergence term to facilitate some of the flux transport onto the far side.

Error and Uncertainty: Fig. 4-f shows the forecast error (Target-Output) and Fig. 4-h shows the estimated uncertainty. As can be seen in the scales of the colorbars, both are smaller than the forecasted magnetic field. Furthermore, error and uncertainty have similar orders of magnitude. This indicates that the uncertainty estimation module is producing sensible results.

4 Future Work and Conclusions

This study introduces ARCADE, the first machine learning framework designed to forecast solar surface magnetic flux emergence and evolution with quantified uncertainties. Our results demonstrate that ARCADE produces forecasts with errors lower than the instrumental uncertainties of SDO/HMI, which are also physically meaningful in their representation of magnetic field evolution. The model can successfully capture the growth of an AR, and predict the magnitude of the emerging magnetic flux.

Throughout this research, a variety of PDE solvers were used for evolving B_r^n to $\hat{B}_r^{n+1}$, with highly varying levels of training stability. In practice, either a simple Euler solver or the Runge-Kutta 4 solver were the most reliable. The objective of ARCADE is to produce short-term forecasts, so we intend to continue investigating advanced solvers in the future and possibly use adaptive time-stepping schemes that may offer improved training stability.

Since our priority is short-term forecasts to enhance flux emergence and AR modeling accuracy, this study focuses on 6-hour time horizons. Future work will focus on extending our results to 12- and 24-hour forecasts, which may necessitate further model design and numerical solver choices to ensure time stability at these slightly longer time scales. Moreover, all experiments were performed on a single GPU, which was found to be sufficient for AFT validation performance but memory-intensive. Expanding ARCADE to massive datasets of historical magnetograms will likely require distributed training techniques, such as data parallelism for multi-GPU training.

For robust end-to-end model UQ, we intend to add a Laplace approximation over the noise network to enhance our epistemic uncertainty understanding over the aleatoric uncertainty estimates that we already have learned. In particular, in ongoing and future work in UQ-driven and UQ-aware training, we are interested in constructing our training pipeline such that the summary statistics of the posterior over the weights generated by the Laplace approximation indicate confident aleatoric uncertainties. Such a pipeline would result in AI-driven forecasts that overcome existing issues in overconfident model predictions and that are ready for integration with scientific pipelines.

While our results achieve consistent uncertainty estimates and short-term forecasting with fixed Fourier positional encoding, trainable positional encoding may offer richer input features to the noise network. Future work may involve optimizing a spatial or temporal encoding to either enhance noise network predictions or provide additional model context throughout our training pipelines. Of particular interest to us is the construction of a temporal encoding for ARCADE that takes into account solar cycle variability. Existing models for Earth systems forecasting will often include spatiotemporal positional encoding that takes into account seasonal variability. Extending this technique for known repeated temporal patterns in solar activity would be novel and of interest to the heliophysics community.

The current ARCADE scheme uses magnetograms and AFT simulations of the surface magnetic flux. Incorporation of subsurface and solar atmospheric data would also contribute valuable information for prediction of flux emergence as well as improved flow field estimations. Future iterations of the ARCADE model will be able to ingest additional data from solar dopplergrams, which probe subsurface layers, as well as extreme ultraviolet images that capture atmospheric dynamics.

Overall, this paper implemented the first hybrid AI-physics model to forecast the evolution of the solar surface magnetic field. This was achieved with a relatively small and computationally inexpensive model, which proves the power of close integration between physics and AI. By reproducing SOTA results, ARCADE delivers on the promise that ML, as a method for solving the SFT problem, can contribute to improving predictions for heliophysics and space weather applications.

Acknowledgments and Disclosure of Funding

This work is a product of the Frontier Development Lab's Heliolab summer research incubator. The authors all acknowledge both financial and programmatic support from this program.

K.K. was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Department of Energy Computational Science Graduate Fellowship under Award Number DE-SC0023112.

References

[1] ALI, A., SADYKOV, V., KOSOVICHEV, A., KITIASHVILI, I. N., ORIA, V., NITA, G. M., ILLARIONOV, E., O'KEEFE, P. M., FRANCIS, F., CHONG, C.-J., KOSOVICH, P., AND MARROQUIN, R. D. Predicting Solar Proton Events of Solar Cycles 22-24 Using GOES Proton and Soft-X-Ray Flux Features. *Astrophys. J. Supp.270*, 1 (Jan. 2024), 15.

[2] CHEN, R. T. Q. torchdiffeq, 2018.

[3] DAXBERGER, E., KRISTIADI, A., IMMER, A., ESCHENHAGEN, R., BAUER, M., AND HENNIG, P. Laplace redux–effortless Bayesian deep learning. In *NeurIPS* (2021).

[4] FORNBERG, B. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation 51*, 184 (1988), 699–706.

[5] GEORGOULIS, M. K. Toward an efficient prediction of solar flares: Which parameters, and how? *Entropy 15*, 11 (2013), 5022–5052.

[6] GIZON, L., CAMERON, R. H., POURABDIAN, M., LIANG, Z.-C., FOURNIER, D., BIRCH, A. C., AND HANSON, C. S. Meridional flow in the sun's convection zone is a single cell in each hemisphere. *Science 368*, 6498 (2020), 1469–1472.

[7] HICKMANN, K. S., GODINEZ, H. C., HENNEY, C. J., AND ARGE, C. N. Data assimilation in the adapt photospheric flux transport model. *Solar Physics 290*, 4 (Apr 2015), 1105–1118.

[8] HOTTA, H., AND KUSANO, K. Solar differential rotation reproduced with high-resolution simulation. *Nature Astronomy 5*, 11 (Nov 2021), 1100–1102.

[9] JIAO, Z., SUN, H., WANG, X., MANCHESTER, W., GOMBOSI, T., HERO, A., AND CHEN, Y. Solar Flare Intensity Prediction With Machine Learning Models. *Space Weather 18*, 7 (July 2020), e02440.

[10] KASAPIS, S., KITIASHVILI, I. N., KOSOVICHEV, A. G., AND STEFAN, J. T. Prediction of intensity variations associated with emerging active regions using helioseismic power maps and machine learning. *The Astrophysical Journal Supplement Series* (2025).

[11] KASAPIS, S., KITIASHVILI, I. N., KOSOVICHEV, A. G., STEFAN, J. T., AND APTE, B. Predicting the emergence of solar active regions using machine learning. *Proceedings of the International Astronomical Union 19*, S365 (2023), 311–319.

[12] KOSOVICHEV, A. G., BASU, S., BEKKI, Y., BUITRAGO-CASAS, J. C., CHATZISTERGOS, T., CHEN, R., CHRISTENSEN-DALSGAARD, J., DONEA, A., FLECK, B., FOURNIER, D., ET AL. Structure and dynamics of the sun's interior revealed by the helioseismic and magnetic imager. *Solar Physics 300*, 5 (2025), 70.

[13] MAHARANA, A., CRAMER, W. D., SAMARA, E., SCOLINI, C., RAEDER, J., AND POEDTS, S. Employing the Coupled EUHFORIA-OpenGGCM Model to Predict CME Geoeffectiveness. *Space Weather 22*, 5 (May 2024), e2023SW003715.

[14] PESNELL, W. D., THOMPSON, B. J., AND CHAMBERLIN, P. C. The Solar Dynamics Observatory (SDO). *Solar Phys.275*, 1-2 (Jan. 2012), 3–15.

[15] RAST, M. P. The Scales of Granulation, Mesogranulation, and Supergranulation. *Astrophys. J.597*, 2 (Nov. 2003), 1200–1210.

[16] SCHERRER, P. H., SCHOU, J., BUSH, R. I., KOSOVICHEV, A. G., BOGART, R. S., HOEKSEMA, J. T., LIU, Y., DUVALL, T. L., ZHAO, J., TITLE, A. M., SCHRIJVER, C. J., TARBELL, T. D., AND TOMCZYK, S. The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO). *Solar Phys.275*, 1-2 (Jan. 2012), 207–227.

[17] SCHOU, J., SCHERRER, P. H., BUSH, R. I., WACHTER, R., COUVIDAT, S., RABELLO-SOARES, M. C., BOGART, R. S., HOEKSEMA, J. T., LIU, Y., DUVALL, T. L., AKIN, D. J., ALLARD, B. A., MILES, J. W., RAIRDEN, R., SHINE, R. A., TARBELL, T. D., TITLE, A. M., WOLFSON, C. J., ELMORE, D. F., NORTON, A. A., AND TOMCZYK, S. Design and Ground Calibration of the Helioseismic and Magnetic Imager (HMI) Instrument on the Solar Dynamics Observatory (SDO). *Solar Phys.275*, 1-2 (Jan. 2012), 229–259.

[18] UGARTE-URRA, I., UPTON, L., WARREN, H. P., AND HATHAWAY, D. H. Magnetic flux transport and the long-term evolution of solar active regions. *The Astrophysical Journal 815*, 2 (dec 2015), 90.

[19] UPTON, L., AND HATHAWAY, D. H. Effects of Meridional Flow Variations on Solar Cycles 23 and 24. *Astrophys. J.792*, 2 (Sept. 2014), 142.

[20] UPTON, L., AND HATHAWAY, D. H. Predicting the Sun's Polar Magnetic Fields with a Surface Flux Transport Model. *Astrophys. J.780*, 1 (Jan. 2014), 5.

[21] VERMA, Y., HEINONEN, M., AND GARG, V. Climode: Climate and weather forecasting with physics-informed neural odes, 2024.

[22] YEATES, A. R., CHEUNG, M. C., JIANG, J., PETROVAY, K., AND WANG, Y.-M. Surface flux transport on the sun. *Space Science Reviews 219*, 4 (2023), 31.

A Glossary

A.1 Notation

Term	Definition
$\theta \in [0, \pi]$	Colatitudinal coordinates in radians.
$\varphi \in [-\pi, \pi]$	Longitudinal coordinates in radians.
$\vec{B} = \vec{B}(\theta, \varphi, t) \in \mathbb{R}^3$	Vector-valued solar magnetic field .
$\vec{v} = \vec{v}(\theta, \varphi, t) \in \mathbb{R}^3$	Vector-valued magnetic velocity field.
$\vec{S}(\theta, \varphi, t) \in \mathbb{R}^3$	Vector-valued source (flux emergence) term.
$S_r(\theta, \varphi, t)$	Radial source (flux emergence) term.
$R_\odot \in \mathbb{R}$	Fixed solar radius constant.
η	Diffusivity parameter in vector-valued surface flux transport equation.
η_h	Diffusivity parameter in radial surface flux transport equation.
$B_r = B_r(\theta, \varphi, t) \in \mathbb{R}$	Radial magnetic field.
$v_\theta(\theta, \varphi, t) \in \mathbb{R}^2$	Latitudinal component of velocity flow.
$v_\varphi(\theta, \varphi, t) \in \mathbb{R}^2$	Longitudinal component of velocity flow.
$\Omega(\theta) \in \mathbb{R}^2$	Differential rotation.
$A, B, C \in \mathbb{R}$	Parameters of differential rotation.
$v_0 \in \mathbb{R}$	Peak pole-ward speed.
$p \in \mathbb{R}$	Parameter controlling how sharply the flow peaks toward mid-latitudes.
B_r^n	Radial magnetic field at time $t = n$ used as an initial condition for forward time evolution.
B_r^{n+1}	Radial magnetic field at time $t = n + 1$ used as a ground truth data sample to compare against evolved magnetogram.
$\hat{B}_r^{n+1}$	Predicted radial magnetogram.
$\xi \in \mathbb{R}$	User-chosen parameter controlling known baseline aleatoric uncertainty (e.g. from known instrument error).
$\hat{\mu}^{n+1}$	Predicted Gaussian mean from noise network corresponding to mean over probabilistic forecast.
$\hat{\sigma}^{n+1}$	Predicted Gaussian variance from noise network corresponding to variance over probabilistic forecast.
$\tilde{B}_r^{n+1}$	$\tilde{B}_r^{n+1} = \hat{B}_r^{n+1} + \hat{\mu}^{n+1}$; magnetogram forecast adjusted for Gaussian mean prediction from noise net.
$\tilde{\sigma}^{n+1}$	$\tilde{\sigma}^{n+1} = \lvert\hat{\sigma}^{n+1}\rvert + \xi$; Gaussian variance adjusted for known baseline aleatoric uncertainty.

A.2 Abbreviations

Term	Definition
AR	Active Region
SFT	Surface Flux Transport
CME	Coronal Mass Ejection
SEP	Solar Energetic Particle
AFT	Advective Flux Transport (existing approach to SFT)
ADAPT	Air Force Data Assimilation Photospheric Transport (existing approach to SFT)
ARCADE	Active Region Characterization and Analysis of Dynamics and Evolution
DR	Differential Rotation
MF	Meridional Flow
SDO	Solar Dynamics Observatory
HMI	Helioseismic & Magnetic Imager

B Computing

Development and training were conducted on an **NVIDIA A100 GPU** (40 GB memory) provisioned via Google Cloud virtual machines. The hardware provided sufficient memory bandwidth to execute a single training process.

To maximize efficiency and stability during training, the following strategies were employed:

- **Automatic Mixed Precision (AMP)** and half-precision operations were selectively applied to reduce memory usage while maintaining model accuracy.
- Training and inference were executed on **CUDA-enabled devices**, leveraging GPU acceleration for all deep learning and PDE-solving workloads.
- **Gradient accumulation** was used to lower memory costs.

C Dashboard UI to execute model and visualize results

The main functionality of the Dashboard UI includes:

- **Model Execution:** Users specify forecast parameters (day, month, year, horizon in hours) through a form. The UI backend issues an HTTP request to `Fused serverless`, which spins up a cloud function to run the model. The function returns a JSON payload containing run metadata and a signed URL pointing to the corresponding GCS bucket.
- **Visualization:** Model outputs, stored as temporally ordered `.png` files in a **Google Cloud Storage (GCS)** bucket, are read by the UI and animated using the `DeckGL` framework. The number of frames depends on cadence and horizon (e.g., a 15-minute cadence over 24 hours produces 96 frames). Typical resolutions include 512×1024 and 1024×2048, though the system supports arbitrary sizes.
- **Interactive Controls:** Users can zoom, pan, scroll through frames, and adjust contrast or color scales. The interface supports rendering one or two globes side-by-side, enabling comparisons such as:
  - Model outputs vs. simulation results,
  - Model outputs vs. uncertainty layers,
  - Different model versions,
  - Different data channels (e.g., magnetogram vs. AIA304).

These capabilities support users across multiple stages of interacting with the model:

1. **Develop and Debug:** Globe rendering highlights anomalies such as checkerboard artifacts from PDE derivatives or NaN emergence from solver step size.

2. **Forecasting:** Produces forward-looking predictions that can be explored interactively in real time.
3. **Presentations:** Forecasts can be exported into stop-motion animations using `Theatre.js`, an independent tool.
4. **Sharing:** Forecasts are distributed via public GCS bucket links, with permissions managed by the backend. These URLs can be embedded in Weights & Biases dashboards and MCP outputs for reproducibility and collaborative analysis.

D Experimental Setup

Configuration	Value	Notes
Optimizer	Adam	
Learning Rate	5×10^{-3}	Schedule (e.g., cosine, step decay)
Batch Size	26	
StepLR	$\gamma = 0.5$ frequency=10	step size = 1, interval=epoch
Epochs	50	

Table 1: Training configurations.

Configuration	Value	Notes
Function	`odeint_adjoint` (instead of `odeint`)	Mitigates memory costs
Solver	Runge-Kutta 4	`torchdiffeq`
Δt	15 minutes	
Data normalization	100 Gauss	
Time normalization	60 minutes	

Table 2: Numerical solver and data configurations.

Component	Hidden Channel List (L)	Residual Blocks / Stage (N)	# Trainable Parameters
Velocity Net	[8, 8]	[3, 2, 2]	6.1 K
Noise Net	[64, 32, 2]	[2, 2, 2]	182 K
PE (Fourier)	Fourier features ($B = 8$)	–	None

Table 3: Architecture configurations.