



Reconstruction of Solar Extreme-ultraviolet Irradiance Using Ca II K Images and SOHO/SEM Data with Bayesian Deep Learning and Uncertainty Quantification

Haodi Jiang^{1,2} , Qin Li^{1,3} , Jason T. L. Wang^{1,4} , Haimin Wang^{1,3,5} , and Serena Criscuoli⁶ ¹ Institute for Space Weather Sciences, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA; haodi.jiang@shsu.edu² Department of Computer Science, Sam Houston State University, Huntsville, TX 77341, USA³ Center for Solar-Terrestrial Research, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA⁴ Department of Computer Science, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA⁵ Big Bear Solar Observatory, New Jersey Institute of Technology, 40386 North Shore Lane, Big Bear City, CA 92314, USA⁶ National Solar Observatory, 3665 Discovery Drive, Boulder, CO 80303, USA

Received 2025 April 20; revised 2025 August 5; accepted 2025 August 9; published 2025 September 16

Abstract

Solar extreme-ultraviolet (EUV) irradiance plays a crucial role in heating the Earth's ionosphere, thermosphere, and mesosphere, affecting atmospheric dynamics over varying time scales. Although significant effort has been spent studying short-term EUV variations from solar transient events, there is little work to explore the long-term evolution of the EUV flux over multiple solar cycles. Continuous EUV flux measurements have only been available since 1995, leaving significant gaps in earlier data. In this study, we propose a Bayesian deep learning model, named SEMNet, to fill the gaps. We validate our approach by applying SEMNet to construct Solar and Heliospheric Observatory/Solar EUV Monitor EUV flux measurements in the period between 1998 and 2014 using Ca II K images from the Precision Solar Photometric Telescope. We then extend SEMNet through transfer learning to reconstruct solar EUV irradiance in the period between 1950 and 1960 using Ca II K images from the Kodaikanal Solar Observatory. Experimental results show that SEMNet provides reliable predictions along with uncertainty bounds, demonstrating the feasibility of Ca II K images as a robust proxy for long-term EUV fluxes. These findings contribute to a better understanding of solar influences on Earth's climate over extended periods.

Unified Astronomy Thesaurus concepts: [Solar atmosphere \(1477\)](#); [Solar extreme ultraviolet emission \(1493\)](#); [Ground-based astronomy \(686\)](#); [Neural networks \(1933\)](#)

1. Introduction

Solar extreme-ultraviolet (EUV) irradiance is a major source of heating and ionization in the upper atmosphere of the Earth (J. L. Lean et al. 2011; A. Vourlidis & S. Bruinsma 2018). The solar EUV spectrum covers the wavelength range from 1 to 120 nm, varying on timescales of minutes, days, months, to multiple years in a solar cycle. The variability of the EUV spectrum is an important driver of space weather (J. Liliensten et al. 2008; M. Haberreiter et al. 2014). When solar activity is elevated, enhanced solar EUV driving causes adverse space weather effects such as radio communication blackouts (A. Szenicer et al. 2019; M. Ishii et al. 2024).

The lack of continuous long-term observations in the EUV is a long-standing problem in solar irradiance studies. Consistent EUV flux measurements have been achieved since 1995. The longest coverage is offered by the Solar EUV Monitor (SEM) on board the Solar and Heliospheric Observatory (SOHO), which provides three EUV bands of measurements, covering from 1995 to the present (D. L. Judge et al. 2002). Later, the Solar EUV Experiment on board the Thermosphere Ionosphere Mesosphere Energetics and Dynamics satellite measures solar spectral irradiance from 0.1 to 194 nm in 1 nm intervals that are closely related to the energetics in the highly variable layers of the Earth's atmosphere above 60 km (T. N. Woods et al. 2005). More recently, the EUV Variability Experiment (EVE) on board the

Solar Dynamics Observatory (SDO) has measured solar EUV irradiance from 0.1 to 105 nm with unprecedented spectral resolution (0.1 nm) and temporal cadence (10 s) since 2010 (T. N. Woods et al. 2012). In addition, the Extreme Ultraviolet Sensor of the Geostationary Operational Environmental Satellite (GOES) on board the GOES-R series satellites has measured solar irradiance on specific spectral lines from 5 to 127 nm since 2017 (F. G. Eparvier et al. 2009).

Ca II K images provide proxies for the spatial distribution of solar emission (S. Criscuoli & I. Ermolli 2008). The relations between indices extracted from Ca II K imagery or from disk-integrated Ca II K emission, and solar irradiance, especially in the UV, have been pointed out in several studies (e.g., H. P. Warren et al. 1996; R. Kariyappa 2000; M. F. Naqvi et al. 2010; J. L. Linsky 2017; G. A. Chapman et al. 2024). Such relations are used to develop models of solar irradiance (J. M. Fontenla et al. 2011; S. Criscuoli et al. 2018; F. Berrilli et al. 2020; T. Chatzistergos et al. 2024; V. Penza et al. 2024), thus allowing one to construct long and uniform time series of irradiance variability necessary for Earth climate studies (L. J. Gray et al. 2010; J. H. Jungclaus et al. 2017). Scaling relations between Ca II K and UV/EUV emission are also found in other stars, and are of fundamental importance because UV/EUV measurements are quite rare for stars and strongly hampered by interstellar medium absorption (A. G. Sreejith et al. 2020). Here, we propose using observed Ca II K images to reconstruct solar EUV irradiance, specifically SEM EUV fluxes. Our work is motivated by the research of P. Foukal (1998), who used digitized Ca II K spectroheliograms obtained at the Mount Wilson Observatory between 1905 and 1984, and measured the area variations of

plages and enhanced networks. The author then calibrated these area variations against the F10.7 index between 1947 and 1984 and constructed a proxy of the F10.7 index, extending it back to 1905.

As mentioned above, SEM EUV flux measurements are collected in multiple channels, including first-order flux and central-order flux. The first-order flux measurement focuses on the ultraviolet range between 26 and 34 nm. It is derived by averaging the count rates from two channels (CH1 and CH3), which are designed to measure solar emissions in this narrow UV range. This data is crucial for studying the impacts of ionizing radiation on the Earth's atmosphere, especially the ionosphere and thermosphere. The data helps to understand solar variability and its influence on space weather. The central-order flux measurement covers a broader wavelength range from 0.1 to 50 nm, capturing more of the solar spectrum's high-energy photons. It is derived from a single channel (CH2) that measures a wider range of solar emissions. This broader range of flux is used for more comprehensive solar monitoring, needed for studying the solar corona and the overall energy output of the Sun. The data helps to understand solar radiation's broader impacts on the Earth's environment, including climate and atmospheric chemistry.

D. Judge et al. (2000) reported that the relative uncertainty of the SEM clone over a 3 yr period is only $\pm 5.5\%$ (1σ), with an absolute uncertainty of 8.5% when transferring the absolute flux calibration to the SOHO Charge, Element, and Isotope Analysis System (CELIAS)/SEM instrument. Our work is concerned with a long-term trend and variation in solar EUV irradiance. With the inspiration of the work of P. Foukal (1998), and the long-term availability of SEM EUV flux measurements, we developed a Bayesian deep learning model, named SEMNet, for solar EUV reconstruction and validated our model by using Ca II K images from the Precision Solar Photometric Telescope (PSPT) to construct SOHO/SEM EUV flux measurements in the period between 1998 and 2014, while minimizing the influence of solar transient events such as solar flares and coronal mass ejections. We then extended SEMNet through transfer learning to reconstruct solar EUV irradiance in the period between 1950 and 1960 using Ca II K images from the Kodaikanal Solar Observatory (KSO). It is worth noting that we use full-disk Ca II K images rather than K line fluxes in our work because these images contain significantly more information than a single flux value. The images preserve spatial and structural variations across the solar disk, allowing SEMNet to learn features such as the size, brightness, and distribution of plages and active regions, some of which contribute strongly to EUV radiation. The spatial and structural richness enables a more accurate and physically grounded mapping from chromospheric emission to the EUV flux. Our deep learning model is capable of capturing such a mapping, as the convolutional architectures of the model can learn hierarchical features hidden in subtle regional patterns (A. S. Razavian et al. 2014). In contrast, scalar K line fluxes lack spatial context and restrict input variability, making a machine learning model more prone to overfitting and less capable of distinguishing spatial configurations crucial for EUV emission.

Deep learning has been used for image-to-image and image-to-index transformations in solar physics (T. Chatzistergos et al. 2019b; A. Pineci et al. 2021; J. Son et al. 2021; H.-J. Jeong et al. 2022; H. Jiang et al. 2023). For example,

T. Chatzistergos et al. (2019b) reconstructed unsigned magnetic fields from Ca II K images. A. Pineci et al. (2021) evaluated He I 1083 nm images for extreme EUV emission. The authors applied a convolutional neural network (CNN) model to examine the nonlinear relationships in the data. They trained and validated the CNN model using historical and contemporaneous solar disk images. A. Szenicer et al. (2019) presented another CNN model to map SDO/Atmospheric Imaging Assembly (AIA) observations to spectral irradiance measurements. In contrast to the above work, we aim to use SEMNet and Ca II K images to reconstruct SEM flux measurements with uncertainty quantification.

The remainder of this paper is organized as follows. Section 2 describes the PSPT Ca II K images and the SOHO/SEM EUV flux measurements used in our study. Section 3 presents the workflow and architecture of SEMNet. Section 4 reports the experimental results in which we validate SEMNet by using it to construct the SEM EUV flux measurements in the period between 1998 and 2014. This section also reports the results in which we extend SEMNet through transfer learning to reconstruct solar EUV irradiance in the period between 1950 and 1960 using the KSO Ca II K images. Section 5 presents a discussion and concludes the paper.

2. Observations and Data Preparation

We utilized a combination of observational data from different solar instruments to train our SEMNet model for reconstructing EUV irradiance from ground-based Ca II K images. The observational data include full-disk Ca II K images from the PSPT operated at the Mauna Loa Solar Observatory in Hawaii (R. L. Coulter & J. R. Kuhn 1994; O. R. White et al. 2000), along with continuous measurements of the solar broadband EUV flux in the 0.1–50 nm and 26–34 nm wavelength ranges obtained from SOHO/SEM.

Full-disk photographs of the Sun in the resonance K line of singly ionized calcium, Ca II K, at 3933.67 Å were first obtained in the early 1890s by G. E. Hale (1893) with the spectroheliographs. Since then, regular observations with similar instruments have been made at various sites around the world. The potential of the full-disk Ca II K observations to serve as a proxy of the magnetic field makes them very valuable for studies of the evolution of solar activity (A. Ortiz & M. Rast 2005; M. Loukitcheva et al. 2009; A. A. Pevtsov et al. 2016; G. Petrie et al. 2021; J. Singh et al. 2023). Ca II K images provide direct information on plage and network regions on the Sun. Through the connection to solar surface magnetic fields, they have provided an excellent source to study solar magnetism for more than a century. The Ca II K data obtained from PSPT are characterized by their high spatial resolution of $2'' \text{ pixel}^{-1}$ scale and high photometric accuracy, with an unprecedented 0.1% pixel-to-pixel relative photometric precision (I. Ermolli et al. 2022). In our study, Ca II K (393.415 nm) with a full width half-maximum (FWHM) of 0.25 nm was used (I. Ermolli et al. 1998, 2022). These data are unique in bridging historical and modern full-disk solar observations, particularly at the Ca II K wavelength.

The SEM (D. L. Judge et al. 1998), part of CELIAS (D. Hovestadt et al. 1995) on board the SOHO mission, is a highly stable grating spectrometer that has provided nearly continuous absolute EUV irradiance measurements since 1995. SEM operates in two key bandpasses: a broad range

from 0.1 to 50 nm and a narrowband channel spanning 26–34 nm, the latter centered on the He II 30.4 nm emission line. The instrument utilizes a free-standing transmission grating and silicon photodiode detectors designed for long-term stability (H. S. Ogawa et al. 1992). Initial calibrations were performed at the National Institute of Standards and Technology/Synchrotron Ultraviolet Radiation Facility prior to launch (D. R. McMullin et al. 2002). To maintain calibration, an ongoing series of sounding rocket measurements has been conducted using a clone of the SEM instrument, which is calibrated before and after each flight, alongside a neon Rare Gas Ionization Cell absolute detector. These calibration flights have indicated some sensitivity loss in the SEM instrument, attributed to the buildup and polymerization of hydrocarbon contaminants by solar UV radiation, which has been mitigated through a time- and wavelength-dependent degradation model. Refinements to the SEM data processing algorithm (S. Wieman et al. 2014), especially during the SOHO and SDO overlapping operational period, have further improved the data accuracy, bringing the SEM irradiance measurements to a better agreement with those of the SDO/EVE instrument (T. N. Woods et al. 2012).

To ensure a consistent construction period, we selected the time frame between 1998 and 2014, during which both Ca II K images and SOHO SEM flux data (in the 0.1–50 nm and 26–34 nm wavelength ranges) are available. Specifically, we collected PSPT Ca II K images from the Laboratory for Atmospheric and Space Physics (LASP)⁷ at the University of Colorado Boulder, between 1998 March and 2014 December. We downloaded the SOHO/SEM daily average flux measurements from the LASP Interactive Solar Irradiance Datacenter (LISIRD).⁸ On each day, there is one or more PSPT Ca II K images corresponding to a pair of SOHO SEM flux readings in the two wavelength ranges. Furthermore, since the Ca II K images are appropriate for studying the long-term (more than 27 days) evolution of the least energetic part of the EUV spectrum but are less effective for studying short-term variations (T. Dudok de Wit et al. 2008), we excluded data samples having anomalous SEM flux measurements from our study, where the anomalies may have been caused by solar transient events. The reason for doing so is that including those anomalies, totaling 16 with flux values significantly higher than their neighboring flux values, may cause our model to pay attention to the anomalies (rare short-term variations), which would weaken the model’s ability to capture common long-term patterns in the data (C. C. Aggarwal 2016).

In training our model here, we built training samples by selecting Ca II K images from PSPT taken at different timestamps on the same day, along with their corresponding SEM flux measurement in an EUV band on that day, which was used as the ground-truth label for the multiple PSPT Ca II K images on that day. During testing, we randomly selected a single PSPT Ca II K image on each day and used it as a test sample. The training set contains samples from January to June in the years between 2000 and 2013, with a total of 12,343 training samples. The test set contains samples from the years 1998, 1999, and 2014, as well as from July to December between 2000 and 2013, totaling 1690 test samples. The validation set contains 500 samples randomly chosen from the same periods as the test set but with different

timestamps. The samples in the training, test, and validation sets span a wide range of solar activity from solar minima to solar maxima between 2000 and 2013. Note that the training set and the test set are disjoint, and hence our SEMNet model can make inferences on the data that it has never seen during training. Original PSPT Ca II K images, with a resolution of 2048×2048 pixels, were resized to 256×256 pixels to improve training time, although this reduction in size did not affect model performance. Due to different scales in the data, we applied normalization to each PSPT Ca II K image by dividing each pixel value in the image by 1000 and to each SEM flux measurement by multiplying the measurement by 1000.

3. Methodology

3.1. Overview of Our Approach

SEMNet is designed to capture the hidden relationships between Ca II K images and SEM EUV fluxes. Inspired by the Monte Carlo dropout framework (MC dropout) proposed by Y. Gal & Z. Ghahramani (2016) and used in H. Jiang et al. (2021), we develop a ResNet-like architecture (K. He et al. 2016) with MC dropout to predict SEM flux measurements with uncertainty quantification. Specifically, stochastic dropouts are applied, and the model output can be approximately treated as a random sample generated from the posterior predictive distribution. Consequently, the uncertainty of the model and the uncertainty of the data can be estimated by calculating the variance of the predictions of the model over multiple repetitions.

During training, SEMNet takes as input pairs of Ca II K images and their corresponding SEM flux measurements in the 0.1–50 nm and 26–34 nm wavelength ranges. The SEM flux measurements are used as ground-truth labels to guide the learning process in which the weights of the neurons in SEMNet are optimized. During testing/inference, the trained SEMNet model takes individual 256×256 full-disk Ca II K images as input and produces as output SEM flux measurements in the two wavelength ranges with uncertainties. Figure 1 explains how the trained model works during testing.

3.2. The SEMNet Model

Figure 2 presents the architecture of SEMNet, which is a ResNet-like model (K. He et al. 2016) with two types of residual blocks. These residual blocks are designed to learn residual functions with respect to the input layer, which mitigates the problem of vanishing gradients and accelerates convergence. Residual block 1 consists of two convolutional layers. The first convolutional layer has a kernel size of 3×3 with a stride of 1. This layer is followed by batch normalization and a LeakyReLU activation with a slope of 0.2, which maintains nonlinearity while reducing the risk of dying neurons. The second convolutional layer also employs a kernel size of 3×3 with a stride of 1, ensuring dimensional consistency across the feature maps. Batch normalization is applied postconvolution. A shortcut connection is used to add the input of this residual block to the output of the second convolutional layer. The final output is activated using LeakyReLU. Residual block 2 contains two convolutional layers similar to those of residual block 1. However, unlike residual block 1, residual block 2 has a stride of 2 in the first convolutional layer to enable downsampling. A shortcut connection is used, which undergoes a 1×1 convolutional

⁷ https://lasp.colorado.edu/pspt_access/.

⁸ https://lasp.colorado.edu/lisird/data/soho_sem_PID.

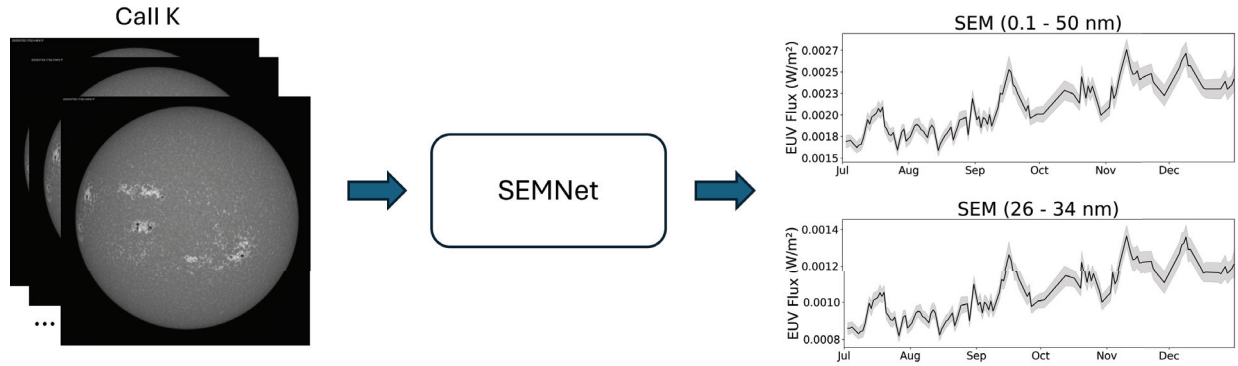


Figure 1. Illustration of the testing of SEMNet. The trained SEMNet model takes as input individual 256×256 full-disk Ca II K images and produces as output corresponding SEM flux measurements in the 0.1–50 nm and 26–34 nm wavelength ranges, where the shaded areas represent estimated uncertainties.

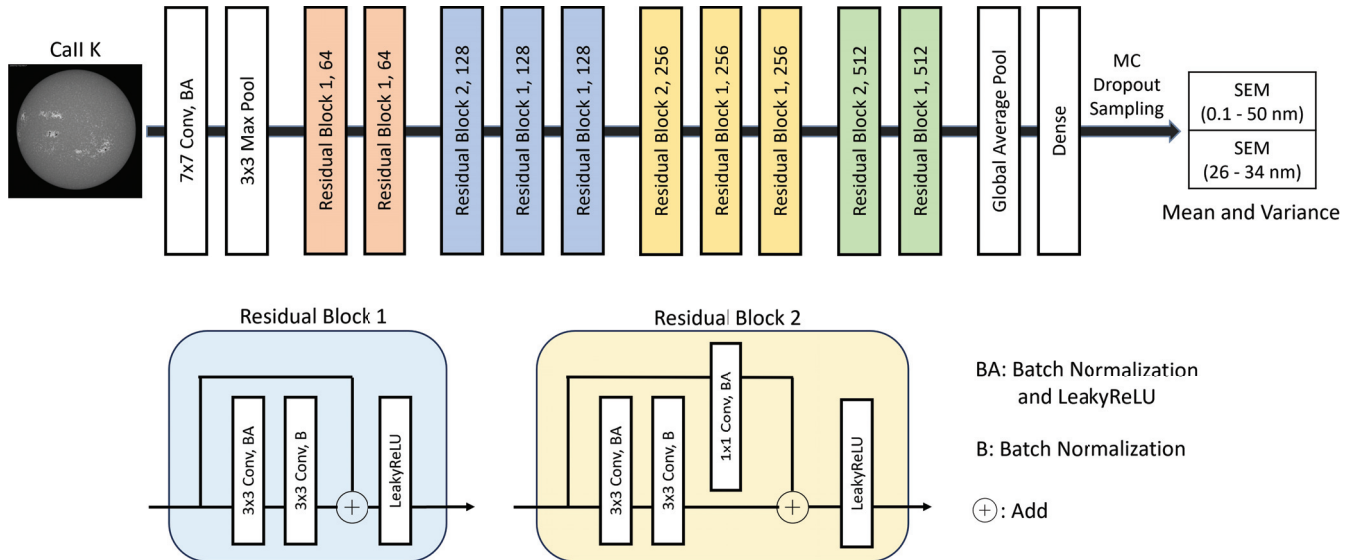


Figure 2. Illustration of the architecture of SEMNet, which is a ResNet-like model with uncertainty quantification.

transformation, to add the input of this residual block to the output of the second convolutional layer. This 1×1 convolution in residual block 2 ensures a dimension match during addition. Like residual block 1, the final output of residual block 2 is activated using LeakyReLU. Our SEMNet model alternates between the two types of residual blocks: residual block 1 (without downsampling) and residual block 2 (with downsampling), as in ResNet (K. He et al. 2016). This design allows the model to retain high spatial dimensions in the earlier layers, which is important for capturing fine-grained features in Ca II K images, while progressively reducing spatial dimensions and increasing the depth of features in the later (deeper) layers to encode higher-level representations. The use of shortcut connections throughout the model helps stabilize training and mitigate vanishing gradient issues.

We feed the SEMNet model with a 256×256 Ca II K image. The model produces as output a SEM flux measurement in the 0.1–50 nm and 26–34 nm wavelength range, respectively, along with uncertainty estimates. The input image first passes through a convolutional layer with 64 filters, each having a kernel size of 7×7 , with a stride of 2. This initial convolution is followed by batch normalization and LeakyReLU with a slope of 0.2. A 3×3 max-pooling layer with a stride of 2 further downsamples the feature maps, capturing key spatial structures from the input. The

network then goes through four residual block stages, with 64, 128, 256, and 512 filters, respectively. Following the residual stages, a global average pooling layer is used to reduce the spatial dimensions of the feature maps. This layer significantly reduces the number of parameters, enhancing generalization and providing translation invariance to the network. To avoid overfitting and enable the Monte Carlo (MC) dropout sampling technique (H. Jiang et al. 2021), a dropout layer with a rate of 0.2 is applied, ensuring regularization by randomly deactivating neurons during training and inference. The final layer is a fully connected dense layer with two output units designed for the regression task at hand. A linear activation function produces the output for the SEM flux in the 0.1–50 nm and 26–34 nm wavelength range, respectively.

We use the Adam optimizer (I. Goodfellow et al. 2016) for model training, which minimizes the Huber loss (J. T. Barron 2019), defined as

$$L_{\delta}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta \\ \delta \times (|y - \hat{y}| - \frac{1}{2}\delta) & \text{otherwise} \end{cases} \quad (1)$$

where y is the observed SEM flux value, \hat{y} is the predicted SEM flux value, and δ is the threshold, with a default value of

0.5, that determines the transition between the quadratic loss and the linear loss. This loss function reduces the sensitivity to outliers compared to the traditional MSE loss (I. Goodfellow et al. 2016), which makes it well suited to handle the inherent variability of the data.

3.3. Uncertainty Quantification

A Bayesian neural network combines a neural network with a probabilistic framework for uncertainty quantification provided by Bayesian statistics. By placing a prior to the network parameters W , we aim to find the posterior distribution of W given the training data D , i.e., pairs of Ca II K images and their corresponding SEM EUV flux measurements. This process is referred to as posterior inference. In theory, achieving exact posterior inference in a deep network is computationally intractable (Y. Gal & Z. Ghahramani 2016), and modifying its architecture for a full Bayesian treatment can be impractical. As a result, approximate methods that preserve the original model structure are often preferred in practice. Inspired by the approach suggested in the literature (Y. Gal & Z. Ghahramani 2016; A. Mobiny et al. 2021), we adopt the MC dropout sampling technique to approximate the uncertainty of the model. The algorithm works as follows: given an input Ca II K image x , the neural network output is computed with stochastic dropouts applied to the network. Specifically, each hidden unit is randomly dropped with a fixed probability p . This stochastic feed-forward process is repeated T times, producing a set of outputs $\{\hat{y}_1, \dots, \hat{y}_T\}$. These samples approximate the draws from the posterior predictive distribution (Y. Gal & Z. Ghahramani 2016). The variance σ_{model}^2 , which quantifies the uncertainty of the model, is then approximated using the sampling variance:

$$\sigma_{\text{model}}^2 = \text{Var}(f^{\text{SEMNet}}(x)) = \frac{1}{T} \sum_{t=1}^T (\hat{y}_t - \bar{\hat{y}})^2, \quad (2)$$

where $\hat{y} = f^{\text{SEMNet}}(x)$ is the model prediction and $\bar{\hat{y}} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t$ is the predicted mean (Y. Gal & Z. Ghahramani 2016). (In our study, T is set to 50 and the dropout probability p is set to 0.2.) Furthermore, following the adaptive approach proposed in L. Zhu & N. Laptev (2017), we let $\hat{f}^{\text{SEMNet}}(\cdot)$ be an unbiased estimator of the network, and $X' = \{x'_1, \dots, x'_V\}$, $Y' = \{y'_1, \dots, y'_V\}$ be the validation data where x'_i , $1 \leq i \leq V$, is a Ca II K image and y'_i is its corresponding SEM flux measurement in the validation set, and V is the cardinality of the validation set. Then, we estimate the noise level during the validation phase, known as the inherent noise, σ_{noise}^2 , via:

$$\sigma_{\text{noise}}^2 = \frac{1}{V} \sum_{v=1}^V (y'_v - \hat{f}^{\text{SEMNet}}(x'_v))^2. \quad (3)$$

The final prediction uncertainty, σ_{pred}^2 , is defined as $\sigma_{\text{model}}^2 + \sigma_{\text{noise}}^2$. Once the prediction uncertainty is determined, we can construct an approximate α -level prediction interval, $(\bar{\hat{y}} - z_{\alpha/2} \times \sigma, \bar{\hat{y}} + z_{\alpha/2} \times \sigma)$, for the model prediction where σ is $\sqrt{\sigma_{\text{pred}}^2}$ and $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of a standard normal distribution. For example, when $z_{\alpha/2} = 2$, it approximates a confidence interval of 95%, which is widely used to assess the reliability of a prediction (A. Gelman et al. 2013).

4. Experiments and Results

4.1. Evaluation Metrics

We adopt three metrics, namely the root mean square error (RMSE; T. Chai & R. R. Draxler 2014), the mean relative error (MRE; A. Szenicer et al. 2019), and the coefficient of determination (R^2 ; D. Chicco et al. 2021) to quantitatively evaluate the performance of SEMNet. Furthermore, we use empirical coverage (EC; J. Breidenbach et al. 2016) to evaluate the uncertainty estimation performance of SEMNet. The first metric is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{\hat{y}}_i - y_i)^2}, \quad (4)$$

where n is the number of samples in the test set that contains one Ca II K image/sample per day, $\bar{\hat{y}}_i$ is the predicted mean, and y_i is the observed value for the i th test sample. This metric calculates the square root of the average squared difference between the predicted mean value and the observed value. A smaller RMSE signifies a better fit of a model to the data with better model performance.

The second metric is defined as

$$\text{MRE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\bar{\hat{y}}_i - y_i}{y_i} \right| \times 100\%, \quad (5)$$

which calculates the average relative difference between the predicted mean $\bar{\hat{y}}_i$ and the observed value y_i . A smaller MRE indicates better model performance.

The third metric is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^n (\bar{\hat{y}}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (6)$$

which measures the strength of the relationship between the predicted mean $\bar{\hat{y}}_i$ and the observed value y_i where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. The metric value ranges from $-\infty$ to 1, with a higher value indicating better model performance.

The fourth metric refers to the proportion of observed values that fall within a specified prediction interval. It is used to evaluate the reliability of a model's uncertainty estimates, which is defined as

$$\text{EC} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(y_i \in [\bar{\hat{y}}_i - z_{\alpha/2} \times \sigma_i, \bar{\hat{y}}_i + z_{\alpha/2} \times \sigma_i]) \times 100\%, \quad (7)$$

where σ_i is the predicted standard deviation for the i th test sample and $\mathbb{I}(\cdot)$ is the indicator function, which is equal to 1 if the inside condition is true and 0 otherwise. For example, a confidence level of 95%, with $z_{\alpha/2} = 2$, corresponds to a prediction interval of $[\bar{\hat{y}}_i - 2\sigma_i, \bar{\hat{y}}_i + 2\sigma_i]$. Intuitively, EC represents the coverage rate. For instance, a coverage rate of 99% means that 99% of observed/true values fall within the prediction interval.

It should be pointed out that RMSE provides an overall assessment of the magnitude of prediction errors (W. Zhang et al. 2022), while MRE offers a normalized measure of prediction errors relative to true EUV fluxes (A. Szenicer et al. 2019). The two metrics are highly correlated and complementary to each other: RMSE reflects the absolute difference,

Table 1
Performance Metric Values of SEMNet and Related Methods

Wavelength	Metric	SEMNet	ANet3	EfficientNetB0	ViT
0.1–50 nm	RMSE	0.000200	0.000268	0.000534	0.000890
	MRE	6.66%	8.23%	17.5%	28.7%
	R ²	0.926	0.867	0.471	−0.466
	EC($\mu \pm 2\sigma$)	99.53%
26–34 nm	RMSE	0.000097	0.000129	0.000219	0.000400
	MRE	6.17%	7.94%	14.7%	26.04%
	R ²	0.924	0.866	0.612	−0.283
	EC($\mu \pm 2\sigma$)	99.82%

whereas MRE shows the relative difference between predicted and true EUV fluxes. R² indicates how well a model explains the variance in true EUV fluxes (L. Ye et al. 2025). EC quantifies how well the specified prediction interval covers true EUV fluxes (J. Breidenbach et al. 2016). Together, these four metrics help us to better understand how accurate a model is from an absolute and relative point of view, how well the model fits the data, and how reliable the model’s uncertainty estimates are.

4.2. Performance Assessment and Comparison

We conducted a series of experiments to evaluate the performance of SEMNet and compare it with related methods, including ANet3 (A. Szenicer et al. 2019), EfficientNetB0 (M. Tan & Q. Le 2019), and the Vision Transformer (ViT; A. Dosovitskiy et al. 2021). ANet3 was used in previous work to map AIA images to EUV fluxes (A. Szenicer et al. 2019). EfficientNetB0 and ViT are well-known models for image classification. To adapt the related models to our work, we replaced the final classification layer of each model with a dense layer comprising two neurons, activated by a linear activation function. Unlike SEMNet, which produces a predicted mean and predicted variance, the three related methods predict a single EUV flux measurement for each wavelength range, given an input PSPT Ca II K image. We chose the Adam optimizer for all three methods to achieve the best performance and the fastest convergence. All methods used the same training, validation, and test sets as described in Section 2.

Table 1 presents the experimental results based on the samples in the test set, where the best values are highlighted in bold. The table includes three performance metrics, RMSE, MRE, R², and an uncertainty metric, EC($\mu \pm 2\sigma$), in two wavelength ranges, 0.1–50 nm and 26–34 nm, where μ denotes the predicted mean and σ denotes the predicted standard deviation. It can be seen from Table 1 that SEMNet outperforms the other methods in terms of both RMSE and MRE, achieving the lowest MRE of 6.66% for the 0.1–50 nm wavelength range and 6.17% for the 26–34 nm wavelength range. SEMNet also achieves the highest R² values, with 0.926 for the 0.1–50 nm wavelength range and 0.924 for the 26–34 nm wavelength range, showing a strong correlation between the predicted and observed EUV flux measurements. ViT has the worst performance, with the highest MRE values of 28.7% for the 0.1–50 nm wavelength range and 26.04% for the 26–34 nm wavelength range, and negative R² values. ViT is designed to capture complex spatial relationships in an image, which are lacking in this application. Another reason is the limited variability of the images, which leads to poor ViT

performance (A. Szenicer et al. 2019; J. Jang & D. Hwang 2022). We also performed additional experiments using other deep learning models, e.g., ResNet50 (K. He et al. 2016), which produced worse results than SEMNet.

The uncertainty metric EC($\mu \pm 2\sigma$) provides valuable information on the reliability of the predictions made by SEMNet. This metric assesses how well prediction intervals capture true/observed values, offering a practical assessment of SEMNet’s uncertainty estimates. Table 1 shows that EC($\mu \pm 2\sigma$) is 99.53% for the 0.1–50 nm wavelength range and 99.82% for the 26–34 nm wavelength range, indicating that almost all true/observed values fall within the prediction interval of $[\mu - 2\sigma, \mu + 2\sigma]$. SEMNet is conservative in the sense that the confidence level of the model is 95%, i.e., the model expects that 95% of the true/observed values fall within the prediction interval, but the coverage rate, EC, is actually higher than 95%. Conservative uncertainty estimates are suitable for long-term EUV reconstruction, where identifying a trend is key. In summary, SEMNet effectively balances accuracy (in terms of RMSE, MRE, and R²) and reliability (in terms of the EC metric), making it a feasible tool for applications that require high reconstruction accuracy and quantification of uncertainty.

Figure 3 shows the long-term evolution of observed and predicted EUV flux measurements in the period between 1998 and 2014 based on the samples in the test set. In the top panel of Figure 3, the predicted SEM flux for the 0.1–50 nm wavelength range closely matches the observed SEM flux, achieving an MRE of 6.66%. The shaded region represents the prediction interval of $[\mu - 2\sigma, \mu + 2\sigma]$. Approximately 99.53% of the true/observed values fall within this interval, with a confidence level of 95%. The bottom panel of Figure 3 shows the predicted SEM flux for the 26–34 nm wavelength range. The predicted flux aligns well with the observed data, with an MRE of 6.17%. Approximately 99.82% of the true/observed values fall within the prediction interval, with a confidence level of 95%.

Figure 4 presents 2D histograms of the observed and predicted EUV flux measurements based on the samples in the test set, where the x-axis (y-axis) in a histogram represents the observed measurements (predicted measurements). In the diagrams, the closer the points are scattered around the diagonal line, the better the predicted measurements are compared to the observed measurements. It can be seen from Figure 4 that when the EUV flux is small, the two measurements tend to agree closely with each other. However, when the EUV flux is large, the SEMNet predictions become less accurate. This happens probably because there are relatively few training samples in periods of strong solar

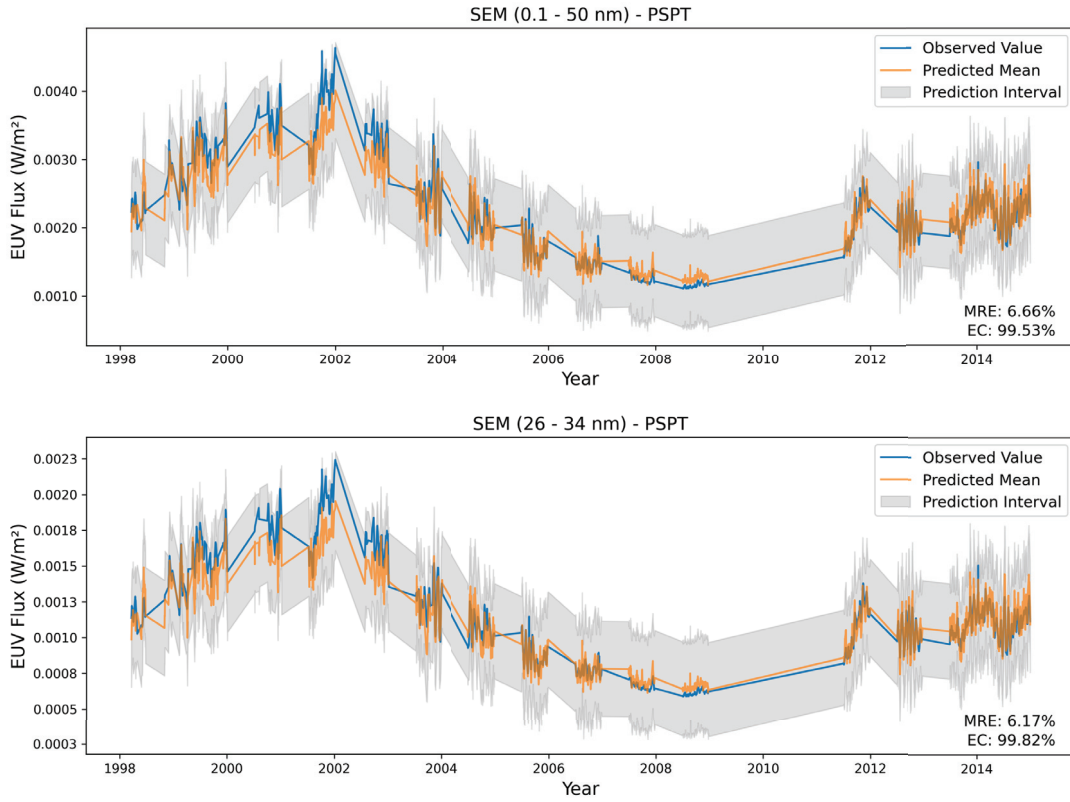


Figure 3. Observed and SEMNet-predicted EUV flux measurements in the period between 1998 and 2014 based on the PSPT Ca II K images in the test set. A shaded region represents the prediction interval of $[\mu - 2\sigma, \mu + 2\sigma]$. Top panel: results for the 01–50 nm wavelength range. Bottom panel: results for the 26–34 nm wavelength range.

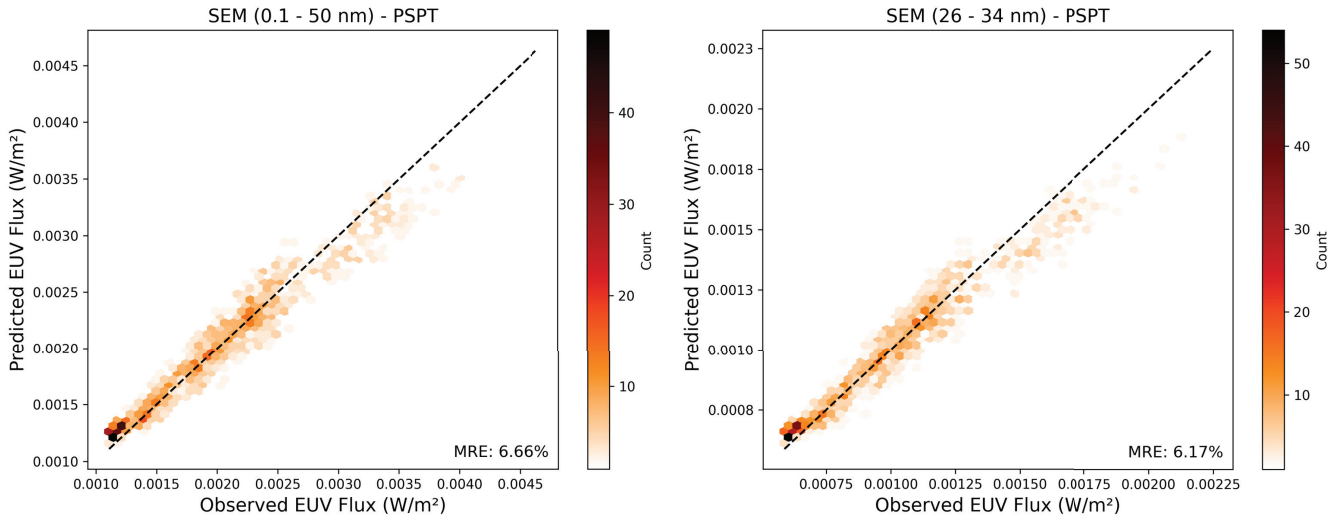


Figure 4. 2D histograms of observed and SEMNet-predicted EUV flux measurements based on the PSPT Ca II K images in the test set. Left panel: histogram for the 0.1–50 nm wavelength range. Right panel: histogram for the 26–34 nm wavelength range.

activity. Moreover, Figure 4 shows that our model performs better in the 26–34 nm wavelength range with an MRE of 6.17% than in the broader 0.1–50 nm wavelength range with an MRE of 6.66%. This happens probably because the 26–34 nm wavelength range, dominated by the He II 30.4 nm line, is produced in the high chromosphere or transition region, which are layers more directly linked to Ca II K emission. In contrast, the 0.1–50 nm wavelength range includes a

substantial coronal contribution, which is less directly linked to Ca II K emission.

Figure 5 presents scatter plots between observed EUV flux measurements and predicted EUV flux measurements, respectively, based on the samples in the test set. The figure shows that the two EUV channels have similar trends. Specifically, there is a linear relationship, $y = 2.08x$, between the two EUV channels in the observed EUV flux measurements, in part

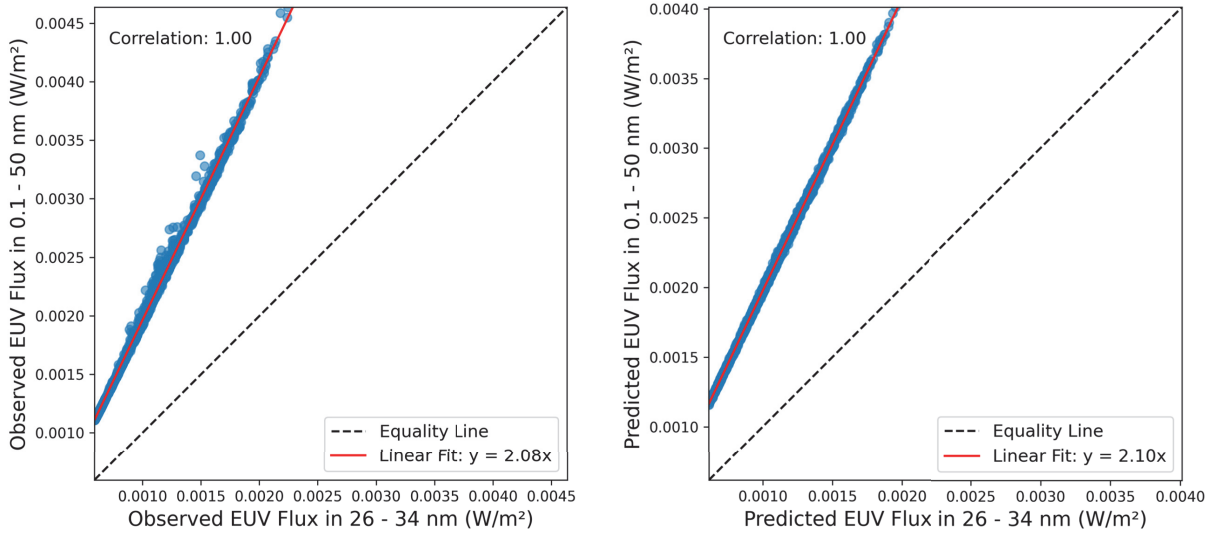


Figure 5. Scatter plots showing that the two EUV channels have similar trends in observed and SEMNet-predicted EUV flux measurements based on the PSPT Ca II K images in the test set. Left panel: scatter plot for the observed EUV flux measurements showing that there is a linear relationship, $y = 2.08x$, between the two EUV channels. Right panel: scatter plot for the predicted EUV flux measurements showing that there is also a linear relationship, $y = 2.10x$, between the two EUV channels. The similar trends in the data present another level of validation of SEMNet.

because there is an overlapped range between the two EUV channels. Furthermore, there is also a linear relationship, $y = 2.10x$, between the two EUV channels in the predicted EUV flux measurements. Similar trends in the data further validate the reliability of the predictions made by SEMNet, demonstrating the feasibility of Ca II K images as a robust proxy for the long-term EUV flux, a finding consistent with that in the literature (T. Dudok de Wit et al. 2008).

4.3. Case Studies on the Solar Maxima in Solar Cycles 23 and 24

To demonstrate how solar activity affects the predictions made by SEMNet, we considered the solar maximum (2001) in solar cycle 23 and the solar maximum (2014) in solar cycle 24. The Sun was very active in 2001 and relatively quieter in 2014. Figure 6 presents the results during the solar maxima. It can be seen from the figure that our SEMNet model performs reasonably well in these periods. Specifically, in 2001, the model achieved an MRE of 8.65% for the 0.1–50 nm wavelength range and 8.77% for the 26–34 nm wavelength range, as shown in the top two panels of Figure 6. In 2014, the model achieved an MRE of 5.21% for the 0.1–50 nm wavelength range and 4.86% for the 26–34 nm wavelength range, as shown in the bottom two panels of Figure 6. In addition, the coverage rate, EC, provides information on the reliability of the uncertainty estimates of the model. During 2001, which is a highly active period, $EC(\mu \pm 2\sigma)$ reached 95.29% for the 0.1–50 nm wavelength range and 98.82% for the 26–34 nm wavelength range, showing that the uncertainty intervals capture most of the variability in the data. During the relatively quieter period of 2014, $EC(\mu \pm 2\sigma)$ reached 100% coverage.

The larger deviations from observed values during the very active period (2001) of the Sun are probably due to fewer training samples for intense solar activity, as there may exist large variations in solar activity when the Sun is very active, while solar activity is more uniform in quieter periods. These

results underscore the importance of uncertainty quantification, which produces interval predictions, as opposed to point predictions, across varying solar conditions.

4.4. Application of SEMNet to Kodaikanal Digitized Ca II K Images via Transfer Learning

So far, we have used PSPT Ca II K images to predict SEM EUV flux measurements. Here, we use SEMNet for the digitized Ca II K data obtained from the KSO⁹ in India to show the applicability of our method to different types of Ca II K images. Historical Ca II K observations at Kodaikanal were performed using a spectroheliograph with a spectral width of 0.5 nm (M. Priyal et al. 2014). KSO uses photographic plates recorded through a telescope having a 30 cm objective with an f -ratio of $f/21$. The effective spatial resolution was about $2'' \text{ pixel}^{-1}$ for the majority of the documentation time. Recently, 16 bit digitization has been performed on these plates using a CCD sensor (with a pixel size of $15 \mu\text{m}$ cooled at -100°C) to generate 4096×4096 raw images (T. Chatzistergos et al. 2018, 2019a, 2021). We collected and preprocessed all available digitized Ca II K images from KSO, in the period from 1996 to 2007, where the images were corrected for center-to-limb variations. Specifically, to reduce cross-instrument domain discrepancies, we applied preprocessing steps such as intensity normalization and histogram matching to bring the KSO images to the same dynamic range and style as PSPT. These images were matched with the corresponding SEM flux readings in the 0.1–50 nm and 26–34 nm wavelength ranges, ensuring temporal alignment between solar images and spectral flux measurements.

We divided the total of 370 KSO Ca II K images, preprocessed to 256×256 pixels, together with their corresponding SEM flux measurements, into training, validation, and test sets. The training set contains 241 paired samples from 2000 to 2007. The validation set contains 50 randomly selected paired samples from 1996 to 2007. The test set

⁹ <https://kso.iap.res.in/data>

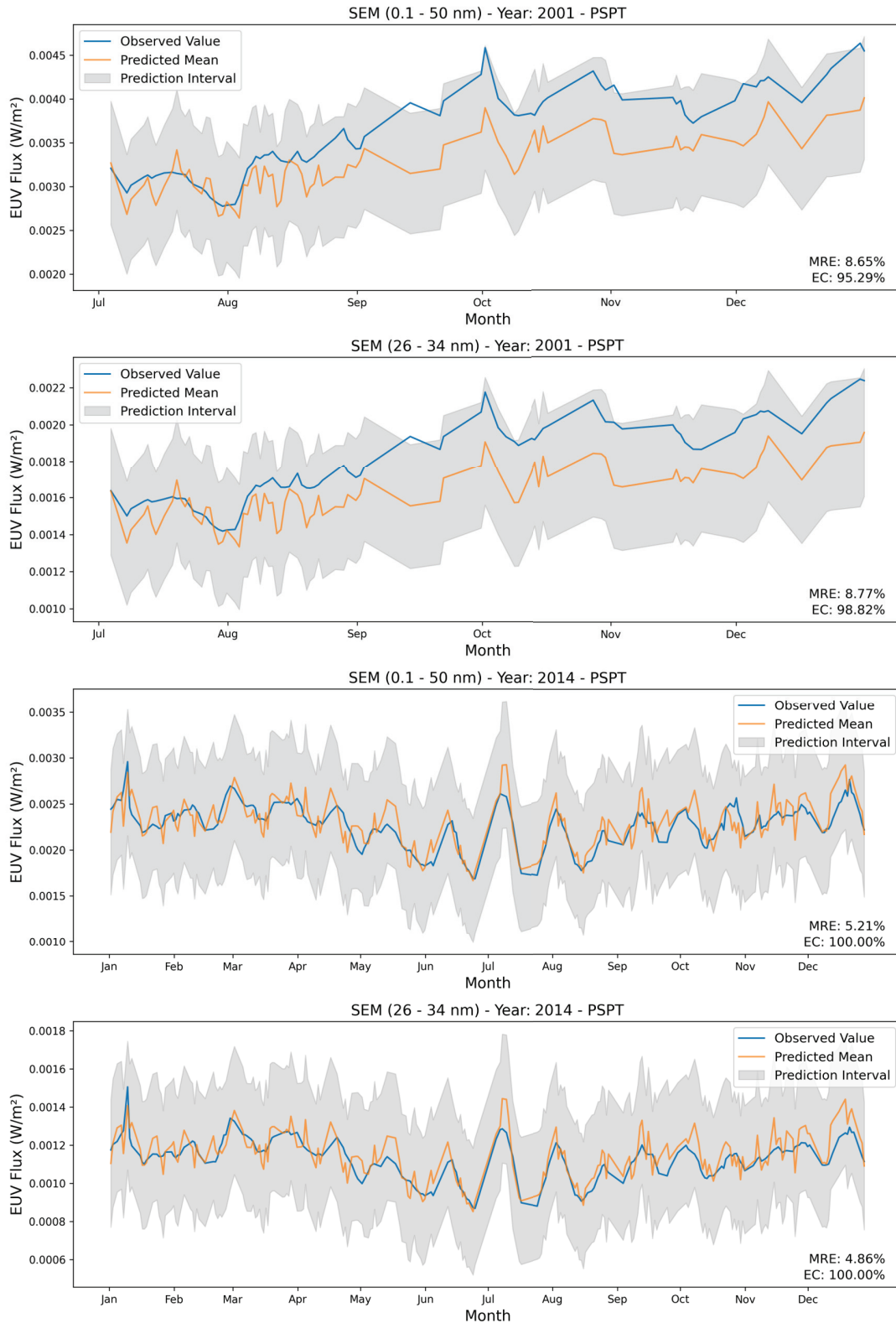


Figure 6. Observed and SEMNet-predicted EUV flux measurements during the solar maxima in solar cycles 23 and 24 based on the PSPT Ca II K images in the test set. A shaded region represents the prediction interval of $[\mu - 2\sigma, \mu + 2\sigma]$. Top two panels: results for the two wavelength ranges under consideration in 2001. Bottom two panels: results for the two wavelength ranges in 2014.

contains 79 paired samples that cover the years from 1996 to 1999. The training, validation, and test sets are disjoint. We adopt a transfer learning approach, where the SEMNet model,

initially trained using the PSPT data, is fine-tuned by being retrained on the 241 paired samples of KSO data in the training set here. Transfer learning provides an effective way to build a

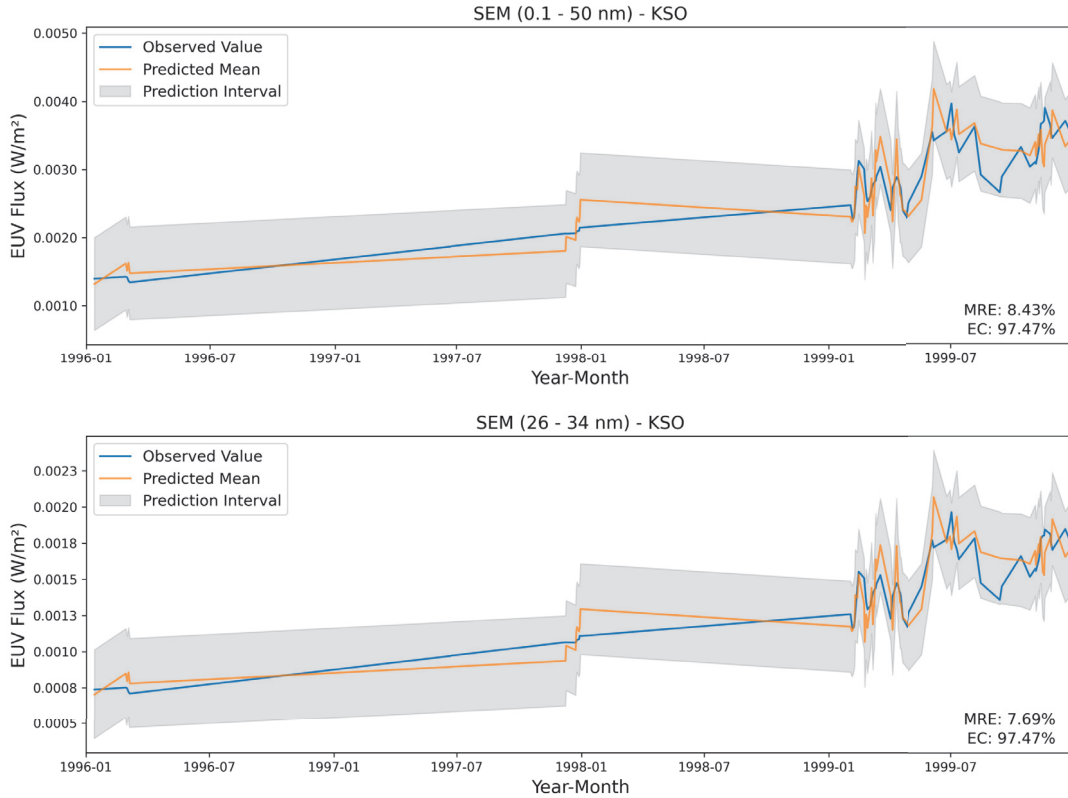


Figure 7. Observed and SEMNet-predicted EUV flux measurements in the period between 1996 and 1999 based on the KSO Ca II K images in the test set. Top panel: results for the 01–50 nm wavelength range. Bottom panel: results for the 26–34 nm wavelength range.

new model to specific needs without requiring substantial training data.

Figure 7 shows the observed and predicted EUV flux measurements in the period between 1996 and 1999 based on the KSO Ca II K images in the test set. The fine-tuned SEMNet model works reasonably well for the KSO data. Specifically, the model achieves an MRE and $EC(\mu \pm 2\sigma)$ of 8.43% and 97.47% (7.69% and 97.47%, respectively) for the wavelength range of 0.1–50 nm (26–34 nm, respectively). Compared to Figure 3, the results based on the PSPT data with an MRE of approximately 6% are slightly better than those based on the KSO data with an MRE of approximately 8%. This is understandable given that the PSPT data benefit from more advanced modern observation technologies compared to the digitized Ca II K images from KSO.

4.5. Reconstruction of Solar EUV Irradiance between 1950 and 1960 Using KSO Ca II K Images

Here, we collect the KSO Ca II K images in the period between 1950 and 1960 and attempt to use the fine-tuned SEMNet model to reconstruct solar EUV irradiance in this period. We considered this period because the F10.7 values, used as a reference here, were available after 1947. These F10.7 values were downloaded from LISIRD.¹⁰ Figure 8 presents the reconstruction results, which are consistent with solar activity in solar cycles 18 and 19. Specifically, the solar minimum (1954) has low EUV fluxes, while the solar maximum (1958) has high EUV fluxes (L. Svalgaard 2013;

I. G. Usoskin 2023). Because observed/true SEM EUV flux measurements are not available in this period, we present the F10.7 values as a reference in this period. F10.7 does not reflect EUV irradiance precisely due to an additional component in gyrosynchrotron emission (J. L. Lean et al. 2001), but would provide a general trend. Figure 8 shows that the F10.7 values are in good agreement with the reconstructed solar EUV irradiance with uncertainty estimates, especially in the solar minimum (L. Liu et al. 2006; Y. Chen et al. 2011).

5. Discussion and Conclusions

The SOHO mission has provided invaluable long-term data to advance our understanding of the Sun and its influence on space weather and Earth’s atmosphere. However, continuous monitoring of solar EUV irradiance remains a significant challenge. In this study, we develop a Bayesian deep learning model (SEMNet) to predict the SEM EUV flux using Ca II K images, focusing on the 0.1–50 nm and 26–34 nm wavelength ranges. SEMNet not only provides reasonably accurate predictions but also reliably quantifies uncertainty, an essential feature in scientific settings. Experimental results based on PSPT data show that SEMNet consistently achieves the lowest MREs, outperforming three related models (ANet3, Efficient-NetB0, and ViT).

After fine-tuning the SEMNet model via transfer learning, we then applied the fine-tuned model to predict EUV fluxes using KSO data in the period between 1996 and 1999, with reasonably good prediction results. Moreover, we used the fine-tuned model to reconstruct solar EUV irradiance in the

¹⁰ https://lasp.colorado.edu/lisird/data/noaa_radio_flux

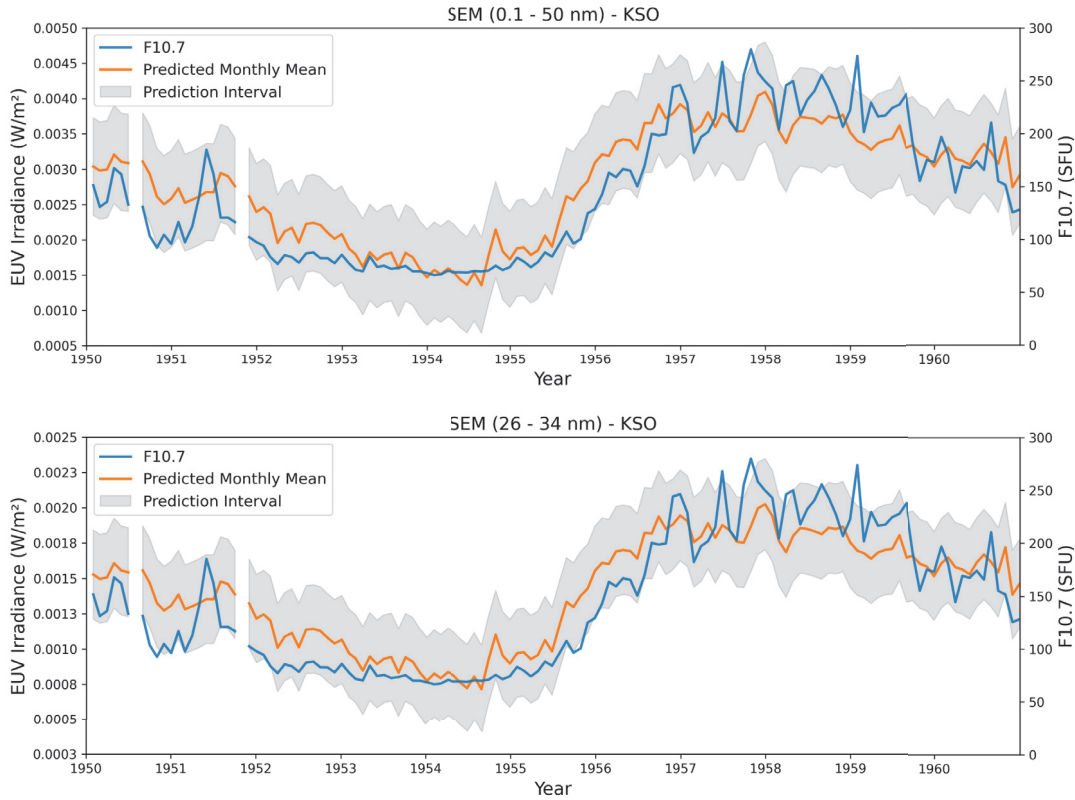


Figure 8. Reconstruction of solar EUV irradiance, more precisely SEM EUV fluxes, in the period between 1950 and 1960 based on the KSO Ca II K images in this period. Shown in the figure are F10.7 values, used as a reference, and SEMNet-predicted EUV flux measurements, averaged over each month for visibility, for the 0.1–50 nm wavelength range (top) and the 26–34 nm wavelength range (bottom). Gaps on the curves are due to missing F10.7 values or KSO images where the two types of data cannot be temporally aligned. The long-term evolution trend of F10.7 is similar to that of reconstructed SEM EUV fluxes.

period between 1950 and 1960, spanning solar cycles 18 and 19. The reconstruction results are consistent with solar activity in the two solar cycles.

It should be noted that, although our model is trained on high-quality PSPT data, systematic errors are inevitable. There are three causes of systematic errors. First, the low-level fixed-pattern residuals and flat-field imperfections in the PSPT images may influence feature learning, particularly in low-contrast regions. Second, the SEM EUV flux values, used as ground-truth labels in model training, carry their own measurement uncertainties, especially during periods of elevated solar activity. These uncertainties propagate through the regression process; however, our prediction framework does not handle label-related uncertainties. Third, the PSPT and KSO data differ in many aspects, such as spatial resolution, photometric response, and dynamic range. When applying our model, initially trained on the PSPT data, to the KSO data through transfer learning, we applied preprocessing steps to reduce cross-instrument domain discrepancies, as described in Section 4.4. Although these preprocessing steps significantly reduce discrepancies, residual inconsistencies in noise texture and stray light remain a possible source of errors. Fine-tuning the model on the preprocessed KSO data further improves model generalization and accuracy. As revealed by our additional experimental study, the errors (in terms of MRE) caused by cross-instrument domain discrepancies are estimated to be 2.4% for the 0.1–50 nm wavelength range and 2% for the 26–34 nm wavelength range, respectively (see Figure 9 in the Appendix).

The PSPT Ca II K images have an FWHM bandpass of 0.25 nm, which includes both chromospheric emission and photospheric absorption. To assess the impact of FWHM on the scaling relation or mapping from Ca II K images to SEM EUV fluxes, we conducted an additional experiment. In this experiment, we downloaded PSPT Ca II K narrowband core (NBC) images with an FWHM of 0.1 nm from the LASP.¹¹ Our SEMNet model was trained on the PSPT Ca II K data (FWHM 0.25 nm) from January to June in the years between 2000 and 2013 (excluding the period of 2007 June 9–11), as described in Section 2, and then tested on three data sets, including the PSPT Ca II K images (FWHM 0.25 nm), NBC images (FWHM 0.1 nm), and preprocessed KSO images (FWHM 0.5 nm), respectively, in their overlapping period of 2007 June 9–11. It was observed that the model performs better on the NBC images (FWHM 0.1 nm) than on the preprocessed KSO images (FWHM 0.5 nm); see Figure 10 in the Appendix. This is understandable given that a narrower bandpass better isolates the chromospheric signal, which scales more directly with EUV emission. Another possible reason is that the KSO data has larger cross-instrument domain discrepancies than the NBC data. Moreover, it was observed that the model performs better on the PSPT Ca II K images (FWHM 0.25 nm) than on the NBC images (FWHM 0.1 nm). This happens probably because the model was trained on the PSPT Ca II K data, which might provide a favorable representation during inference.

¹¹ https://lasp.colorado.edu/lisird/data/pspt_cak_nbc_image_files

On the basis of these experimental results, we conclude that SEMNet is a feasible tool for studying solar activity and for reconstructing solar EUV irradiance. Ca II K images are available starting 1893, and images in this spectral range have been acquired and are still acquired at several observatories (T. Chatzistergos et al. 2019a, 2020, 2021, 2024). SEMNet therefore potentially allows one to extend EUV irradiance estimates back into the past.

Acknowledgments

The authors thank members of the Institute for Space Weather Sciences for fruitful discussions. The deep learning methods studied here were implemented in TensorFlow. H.J. acknowledges support from NASA grants 80NSSC24K0548 and 80NSSC25K7708. Q.L., J.W., and H.W. acknowledge support from NSF grants AGS-2149748, AGS-2228996, OAC-2320147, OAC-2504860, and NASA grants 80NSSC24K0548, 80NSSC24K0843, and 80NSSC24M0174.

Facilities: SOHO, HAO, KTT.

Appendix Supplementary Material

Figure 9 shows the prediction errors, measured by the mean relative error (MRE), obtained by applying the SEMNet model, initially trained on PSPT Ca II K data, to the PSPT Ca II K data set, the raw KSO Ca II K data set (before preprocessing), and the preprocessed KSO Ca II K data set with model fine-tuning, in their overlapping period from 1999 March to September, for wavelength ranges of 0.1–50 nm and 26–34 nm, respectively. Figure 10 compares the model predictions for varying FWHM. The SEMNet model, trained on PSPT Ca II K data, was applied separately to three data sets with different FWHM values, including the PSPT Ca II K images (FWHM 0.25 nm), NBC images (FWHM 0.1 nm), and preprocessed KSO images (FWHM 0.5 nm), during the period of 2007 June 9–11.

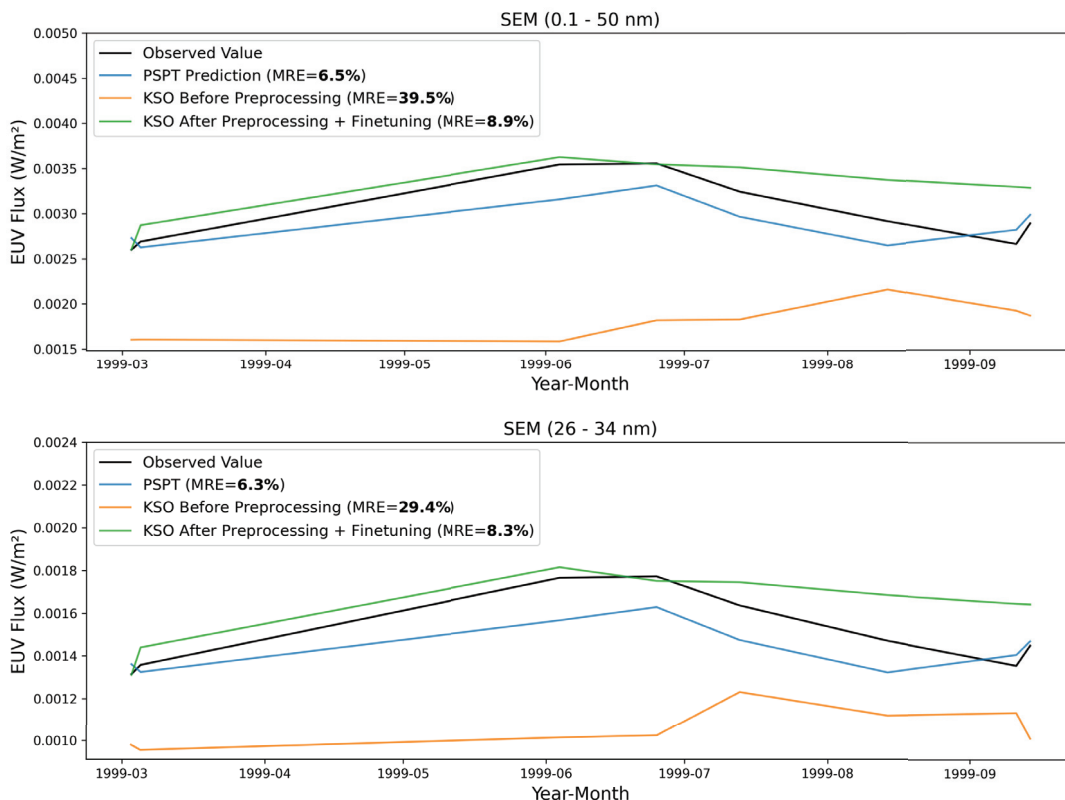


Figure 9. Comparison of model predictions on different data sources. Shown in the figure are observed and SEMNet-predicted EUV flux measurements on the PSPT Ca II K data set, the raw KSO Ca II K data set (before preprocessing), and the preprocessed KSO Ca II K data set with model fine-tuning, in their overlapping period from 1999 March to September. Top panel: results for the 0.1–50 nm wavelength range, with an estimated MRE of approximately 2.4% caused by cross-instrument domain discrepancies. Bottom panel: results for the 26–34 nm wavelength range, with an estimated MRE of 2% from cross-instrument effects.

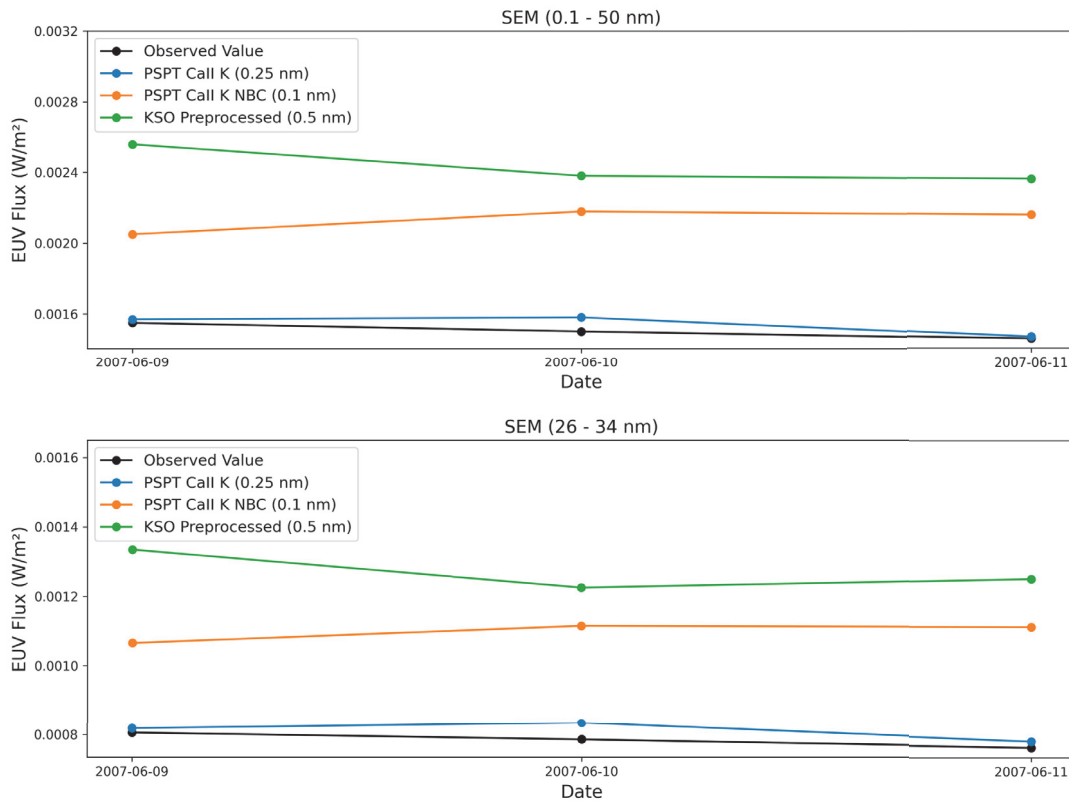


Figure 10. Comparison of model predictions for varying FWHM. Shown in the figure are observed and SEMNet-predicted EUV flux measurements on the PSPT Ca II K images (FWHM 0.25 nm), NBC images (FWHM 0.1 nm), and preprocessed KSO images (FWHM 0.5 nm), during the period of 2007 June 9–11. Top panel: results for the 0.1–50 nm wavelength range. Bottom panel: results for the 26–34 nm wavelength range.

ORCID iDs

Haodi Jiang <https://orcid.org/0000-0001-6460-408X>
 Qin Li <https://orcid.org/0000-0002-3669-1830>
 Jason T. L. Wang <https://orcid.org/0000-0002-2486-1097>
 Haimin Wang <https://orcid.org/0000-0002-5233-565X>
 Serena Criscuoli <https://orcid.org/0000-0002-4525-9038>

References

- Aggarwal, C. C. 2016, *Outlier Analysis* (2nd ed.; Cham: Springer)
- Barron, J. T. 2019, in 2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 4326
- Berrilli, F., Criscuoli, S., Penza, V., & Lovric, M. 2020, *SoPh*, **295**, 38
- Breidenbach, J., McRoberts, R. E., & Astrup, R. 2016, *RSEnv*, **173**, 274
- Chai, T., & Draxler, R. R. 2014, *GMD*, **7**, 1247
- Chapman, G. A., Cookson, A. M., & Choudhary, D. P. 2024, *JSWSC*, **14**, 34
- Chatzistergos, T., Ermolli, I., Krivova, N. A., et al. 2020, *A&A*, **639**, A88
- Chatzistergos, T., Ermolli, I., Solanki, S. K., & Krivova, N. A. 2018, *A&A*, **609**, A92
- Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019a, *SoPh*, **294**, 145
- Chatzistergos, T., Ermolli, I., Solanki, S. K., et al. 2019b, *A&A*, **626**, A114
- Chatzistergos, T., Krivova, N. A., & Ermolli, I. 2024, *JSWSC*, **14**, 9
- Chatzistergos, T., Krivova, N. A., Ermolli, I., et al. 2021, *A&A*, **656**, A104
- Chen, Y., Liu, L., & Wan, W. 2011, *JGRA*, **116**, A04304
- Chicco, D., Warrens, M. J., & Jurman, G. 2021, *PeerJ Comput. Sci.*, **7**, e623
- Coulter, R. L., & Kuhn, J. R. 1994, in ASP Conf. Ser. 68, *Solar Active Region Evolution: Comparing Models with Observations*, ed. K. S. Balasubramaniam & G. W. Simon (San Francisco, CA: ASP), 37
- Criscuoli, S., & Ermolli, I. 2008, *A&A*, **484**, 591
- Criscuoli, S., Penza, V., Lovric, M., & Berrilli, F. 2018, *ApJ*, **865**, 22
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. 2021, 9th Int. Conf. on Learning Representations, ICLR 2021 (OpenReview) <https://openreview.net/forum?id=YicbFdNTTy>
- Dudok de Wit, T., Kretzschmar, M., Abouadarham, J., et al. 2008, *AdSpR*, **42**, 903
- Eparvier, F. G., Crotser, D., Jones, A. R., et al. 2009, *Proc. SPIE*, **7438**, 743804
- Ermolli, I., Fofi, M., Bernacchia, C., et al. 1998, *SoPh*, **177**, 1
- Ermolli, I., Giorgi, F., & Chatzistergos, T. 2022, *FrASS*, **9**, 352
- Fontenla, J. M., Harder, J., Livingston, W., Snow, M., & Woods, T. 2011, *JGRD*, **116**, D20108
- Foukal, P. 1998, *GeoRL*, **25**, 2909
- Gal, Y., & Ghahramani, Z. 2016, in Proc. of Machine Learning Research 48, Proc. of the 33rd Int. Conf. on Machine Learning, ed. M. F. Balcan & K. Q. Weinberger (New York: PMLR), 1050, <https://proceedings.mlr.press/v48/gal16.html>
- Gelman, A., Carlin, J. B., Stern, H. S., et al. 2013, *Bayesian Data Analysis* (3rd ed.; New York: Chapman and Hall/CRC)
- Goodfellow, I., Bengio, Y., & Courville, A. 2016, *Deep Learning* (Cambridge, MA: MIT Press)
- Gray, L. J., Beer, J., Geller, M., et al. 2010, *RvGeo*, **48**, RG4001
- Haberreiter, M., Delouille, V., Mampaey, B., et al. 2014, *JSWSC*, **4**, A30
- Hale, G. E. 1893, *MmSSI*, **21**, 68
- He, K., Zhang, X., Ren, S., & Sun, J. 2016, in 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE),
- Hovestadt, D., Hilchenbach, M., Bürgi, A., et al. 1995, *SoPh*, **162**, 441
- Ishii, M., Berdermann, J., Forte, B., et al. 2024, *AdSpR*, in press
- Jang, J., & Hwang, D. 2022, in 2022 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 20686
- Jeong, H.-J., Moon, Y.-J., Park, E., Lee, H., & Baek, J.-H. 2022, *ApJS*, **262**, 50
- Jiang, H., Jing, J., Wang, J., et al. 2021, *ApJS*, **256**, 20
- Jiang, H., Li, Q., Liu, N., et al. 2023, *SoPh*, **298**, 87
- Judge, D., Ogawa, H., McMullin, D., & Gangopadhyay, P. 2000, *PCEC*, **25**, 417
- Judge, D. L., Ogawa, H. S., McMullin, D. R., Gangopadhyay, P., & Pap, J. M. 2002, *AdSpR*, **29**, 1963
- Judge, D. L., McMullin, D. R., Ogawa, H. S., et al. 1998, *SoPh*, **177**, 161
- Jungclaus, J. H., Bard, E., Baroni, M., et al. 2017, *GMD*, **10**, 4005
- Kariyappa, R. 2000, *JApA*, **21**, 293

- Lean, J. L., White, O. R., Livingston, W. C., & Picone, J. M. 2001, *JGR*, **106**, 10645
- Lean, J. L., Woods, T. N., Eparvier, F. G., et al. 2011, *JGRA*, **116**, A01102
- Lilensten, J., Dudok de Wit, T., Kretzschmar, M., et al. 2008, *AnGeo*, **26**, 269
- Linsky, J. L. 2017, *ARA&A*, **55**, 159
- Liu, L., Wan, W., Ning, B., Pirog, O. M., & Kurkin, V. I. 2006, *JGRA*, **111**, A08304
- Loukitcheva, M., Solanki, S. K., & White, S. M. 2009, *A&A*, **497**, 273
- McMullin, D. R., Judge, D. L., Hilchenbach, M., et al. 2002, *ISSIR*, **2**, 135
- Mobiny, A., Yuan, P., Moulik, S. K., et al. 2021, *NatSR*, **11**, 5458
- Naqvi, M. F., Marquette, W. H., Tritschler, A., & Denker, C. 2010, *AN*, **331**, 696
- Ogawa, H. S., McMullin, D., Judge, D. L., & Korde, R. 1992, *Proc. SPIE*, **1745**, 123
- Ortiz, A., & Rast, M. 2005, *MmSAI*, **76**, 1018
- Penza, V., Bertello, L., Cantoresi, M., et al. 2024, *ApJ*, **976**, 11
- Petrie, G., Criscuoli, S., & Bertello, L. 2021, *Solar Magnetism and Radiation* (Washington, DC: American Geophysical Union), 83
- Pevtsov, A. A., Virtanen, I., Mursula, K., Tlatov, A., & Bertello, L. 2016, *A&A*, **585**, A40
- Pineci, A., Sadowski, P., Gaidos, E., & Sun, X. 2021, *ApJL*, **910**, L25
- Priyal, M., Singh, J., Ravindra, B., Priya, T. G., & Amareswari, K. 2014, *SoPh*, **289**, 137
- Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. 2014, in *Proc. of the 2014 IEEE Conf. Computer Vision and Pattern Recognition Workshops, CVPRW'14* (Piscataway, NJ: IEEE), 512
- Singh, J., Priyal, M., Ravindra, B., Bertello, L., & Pevtsov, A. 2023, *RAA*, **23**, 045016
- Son, J., Cha, J., Moon, Y.-J., et al. 2021, *ApJ*, **920**, 101
- Sreejith, A. G., Fossati, L., Youngblood, A., France, K., & Ambily, S. 2020, *A&A*, **644**, A67
- Svalgaard, L. 2013, *JSWSC*, **3**, A24
- Szenicer, A., Fouhey, D. F., Munoz-Jaramillo, A., et al. 2019, *SciA*, **5**, eaaw6548
- Tan, M., & Le, Q. 2019, in *Proc. of Machine Learning Research 97, Proc. 36th Int. Conf. Machine Learning*, ed. K. Chaudhuri & R. Salakhutdinov (New York: PMLR), 6105, <https://proceedings.mlr.press/v97/tan19a.html>
- Usoskin, I. G. 2023, *LRSP*, **20**, 2
- Vourlidas, A., & Bruinsma, S. 2018, *SpWea*, **16**, 5
- Warren, H. P., Mariska, J. T., Lean, J., Marquette, W., & Johannesson, A. 1996, *GeoRL*, **23**, 2207
- White, O. R., Fox, P. A., Meisner, R., et al. 2000, in *Solar Variability and Climate*, ed. E. Friis-Christensen et al., Vol. 11 (Berlin: Springer), 75, doi:10.1007/978-94-010-0888-4_7
- Wiemann, S., Didkovsky, L., & Judge, D. 2014, *SoPh*, **289**, 2907
- Woods, T. N., Eparvier, F. G., Bailey, S. M., et al. 2005, *JGRA*, **110**, A01312
- Woods, T. N., Eparvier, F. G., Hock, R., et al. 2012, *SoPh*, **275**, 115
- Ye, L., Liu, M., Fu, D., et al. 2025, *RemS*, **17**, 1720
- Zhang, W., Zhao, X., Feng, X., et al. 2022, *Univ*, **8**, 30
- Zhu, L., & Laptev, N. 2017, in *2017 IEEE Int. Conf. on Data Mining Workshops (ICDMW)* (Piscataway, NJ: IEEE), 103