Forecasting Extreme Ultraviolet Thermospheric Drivers from Solar Imaging using Deep Learning

Edward J. E. Brown [1,2], Filip Svoboda [2], Nigel P. Meredith[1]

[1]British Antarctic Survey, Natural Environment Research Centre, Cambridge, UK
[2]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK

Key Points:

- EUV irradiance, 0.05nm to 121nm, is forecasted from 9.4, 33.5, and 160nm solar images using deep learning up to 4 days in advance
- The models outperform the persistence model at a 4 day horizon with a reduction of 36.45% in the mean absolute percentage error
- Models developed can provide forecasts of irradiance for use in thermospheric density models, important for vital spacecraft operations

Abstract

The thermosphere is heated by incident extreme ultraviolet (EUV) radiation from the Sun. Accurately forecasting this driver of thermospheric density is of paramount importance to spacecraft operations, such as collision avoidance, in an increasingly crowded low Earth orbit environment. This study uses deep learning techniques to forecast solar irradiance from three solar EUV/UV image channels (9.4nm, 33.5nm, and 160.0nm) taken by the Solar Dynamics Observatory up to four days in advance. The proposed model is able to forecast 23 wavelength bands from 0.05nm to 121nm (produced from the FISM2 empirical irradiance model (Chamberlin et al., 2020)) with a mean absolute percentage error of 6.0% and an average improvement in the mean absolute percentage error of 36.45% over the baseline persistence model at a 4 day time horizon. The study further validates the model and derives insights about its internal working by testing and investigating the importance of its core components.

Plain language summary

Vital low Earth orbit spacecraft operations, such as collision avoidance, rely on accurate modelling of the thermosphere. The thermosphere is heated, amongst other phenomena, by extreme ultraviolet (EUV) radiation from the Sun. This study uses deep learning-powered computer vision techniques to forecast irradiance from the Sun up to four days in advance using solar imaging. The results show a 6.0% mean absolute percentage error at four days which is a significant reduction of 36.45% in mean absolute percentage error over a baseline persistence model. This represents an opportunity to provide accurate solar EUV irradiance forecasts that will be useful for thermospheric modelling and therefore spacecraft operations.

1 Introduction

Recent high-profile satellite losses, such as the February 2022 SpaceX loss event, have underscored the importance of forecasting thermospheric drivers. The thermosphere is heated up, amongst other phenomena, by direct energy deposits from solar extreme ultraviolet (EUV) radiation (Berger et al., 2020). The density of the thermosphere therefore varies with the incident irradiance. Since satellite drag is proportional to the density of thermosphere, in which low-Earth orbit satellites operate, the ability to forecast solar EUV irradiance will therefore help forecast the density of the thermosphere and provide more accurate orbital propagation estimates. This ultimately aids spacecraft operations such as collision avoidance- important in an increasingly crowded low Earth orbit environment.

Features on the Sun such as active regions or coronal holes rotate in and out of view (from Earth's point of view) on an approximately 27 day cycle. This rotation period changes depending on the latitude on the solar disk, with the rotation period increasing at higher latitudes. If an active region is on the eastern limb of the solar disk that feature will soon be rotated out of view. This suggests that knowing what features are going to be soon disappearing from (and indeed appearing into) view will give insight into what the future irradiance conditions will be.

Solar images can be processed by machine learning models to make predictions. Indeed, previous works such as Upendran et al. (2020), Brown et al. (2022), Raju and Das (2021) have done just that with the solar wind speed, Bernoux et al. (2022) with the Kp index, and Hu et al. (2022) to forecast the Dst index. Sun et al. (2023) focus on solar active regions to forecast the chance of a flare occurring at short timescales from cropped images of active regions. In the space of modelling irradiance, Brown et al. (2021) use precomputed solar features from an autoencoder trained on solar images to forecast the F10.7 radio measurement - a solar irradiance proxy. Szenicer et al. (2019) create a virtual instrument to replace the MEGS-A instrument and nowcast (as opposed to forecast) various EUV irradiance emission lines using solar imaging. This presented study forecasts a multi-band EUV irradiance spectrum, selected for its relevance to the thermosphere, from full-disk solar images up to four days in advance.

2 Data

2.1 Solar Images

The solar images were obtained from the Atmospheric Imaging Assembly (AIA) on board the Solar Dynamics Observatory (SDO). SDO was launched on 11 February 2010 and began collecting science quality data on 1 May 2010 from a non-equatorial geosynchronous orbit (Pesnell et al., 2012).

The AIA takes images of the full Sun in two ultraviolet wavelength bands, centered at 160 and 170nm, seven extreme ultraviolet (EUV) wavelength bands, centred at 9.4, 13.1, 17.1, 19.3, 21.1, 30.4, and 33.5 nm, and one visible wavelength band, centered at 450nm (Lemen et al., 2012). The study uses the SDOML dataset, which is a specially curated set of AIA images that have been prepared for consumption by machine learning algorithms. To create this dataset, the AIA images were processed by performing various instrumental corrections, downsampled to useable spatial and temporal resolutions and synchronized both spatially and temporally (Galvez et al., 2019). The resulting data set, used in this study, contains 8 and a half years of images in each of the wavelength bands every 6 minutes from May 2010 to December 2018.

2.2 Solar Irradiance Spectra

The Flare Irradiance Spectral Model - Version 2 (FISM2) is an empirical model of the solar spectral irradiance from 0.01-190nm created to fill spectral and temporal gaps in actual measurements (Chamberlin et al., 2020). The model provides a "daily" and a "flare" product. The "flare" product is used in the study. This provides one spectrum every 60 seconds and includes variations due to the solar cycle, solar rotation and solar flares. Both products are available across the entire wavelength range in 0.1nm bins or binned into 23 'stan bands' covering the EUV from 0.01 to 121 nm (Solomon & Qian, 2005) for direct input into Earth atmosphere models such as the Whole Atmosphere Community Climate Model (WACCM) (Hurrell et al., 2013; Marsh et al., 2013). The "stan bands" are used as they are important for modelling the relationship of solar irradiance on thermospheric density, through the relevant atmosphere models. They are measured in the photons per cm^2 per second. Some of the bands have multiple categories representing weighted-averages from thresholds of nitrogen cross-sections (low, medium, and high).

2.3 Autocorrelation

The degree of autocorrelation in spectral irradiance is an important consideration when forecasting. Autocorrelation refers to the degree of correlation between a measured quantity and itself at a certain time lag. For example, if a particular spectral band is highly autocorrelated at a 1-day time lag, its measurements taken on successive days are likely

to be similar. Indeed, achieving a close fit between a forecast and the predicted feature is not always indicative of good performance if the feature is highly autocorrelated, because the current value, often referred to as the 'persistence value' is itself a good predictor, which one has to outperform to demonstrate any significant ability to forecast. Consequently, accounting for this will be an important component of the evaluation methodology. Indeed, the comparison to the persistence model (where the persistence value is used as the predictor) and persistence score (section 3.3) becomes central to the reporting of the model results.

3 Methodology

3.1 Model

The proposed model adapts and builds on the fixed-delay single-channel image-based Solar Wind speed prediction model proposed by Brown et al. (2022) by allowing for multiple input channels, and a variable forecast horizon. The original study forecasted a scalar solar wind speed at a 96 hour lag from a single 21.1nm image, using the OmniWeb (OmniWeb data archive, n.d.) solar wind and the SDOML Extreme UV image (Galvez et al., 2019) datasets. The here-proposed model uses 9.4nm, 33.5nm, and 160nm images to forecast a 23-entry vector of irradiance predictions for bands between 0.05 and 121.0 nm at the 24h, 48h, 72h, and 96h lags.

Figure 1 shows the forecasting setup over a sample period from the 16^{th} to the 24^{th} of July, 2018. Shown in the figure are three solar images at 9.4nm to highlight what is happening on the Sun. The expectation is that the proposed model can be trained to recognise that the active region will soon be rotated out of view, resulting in a drop in irradiance. The top graph shows the model forecast at 4 days for the 1.8-3.2nm band.

The model in figure 1 is trained to evaluate observed solar conditions and extrapolate into the future by accounting for the rotational and the likely surface-condition developments. Indeed, on the on the 16^{th} (green-coded image, left) the model considers the solar conditions and predicts that by the 20^{th} (red-coded image, central), the active region will be partially obscured as it will be on the edge of the disk, and so its predicted irradiance is decreasing. Similarly, on the 20^{th}, based on the red image (center) the model predicts that by the 24^{th} the active region will have rotated out of view. Consequently, in predicts a significant drop in the forecasted irradiance, relative to the value on the 20^{th}. The model's prediction is vindicated by the observation on the 24^{th}, the blue-coded image(right) which shows that the active region has disappeared.

The internal mechanics of the predictive model are demonstrated on the example of July the 20^{th} at 00.00.00. The conditions at the moment of the prediction are illustrated by the central image in the top row, highlighted in red. The model consumes 9.4nm, 33.5nm, and 160.0nm EUV images, these are shown directly below.

The model consumes image patches derived from the 9.4nm, 33.5nm, and 160.0nm images. These input images are stacked as three channels, normalized, scaled, and then split into 196 patches, each 14 pixels wide, 14 pixels high and 3 channels deep. Shown in the diagram are three representatives of these patches, highlighted in violet, purple, and orange. The 196 patches are then passed through individually-applied linear projections,

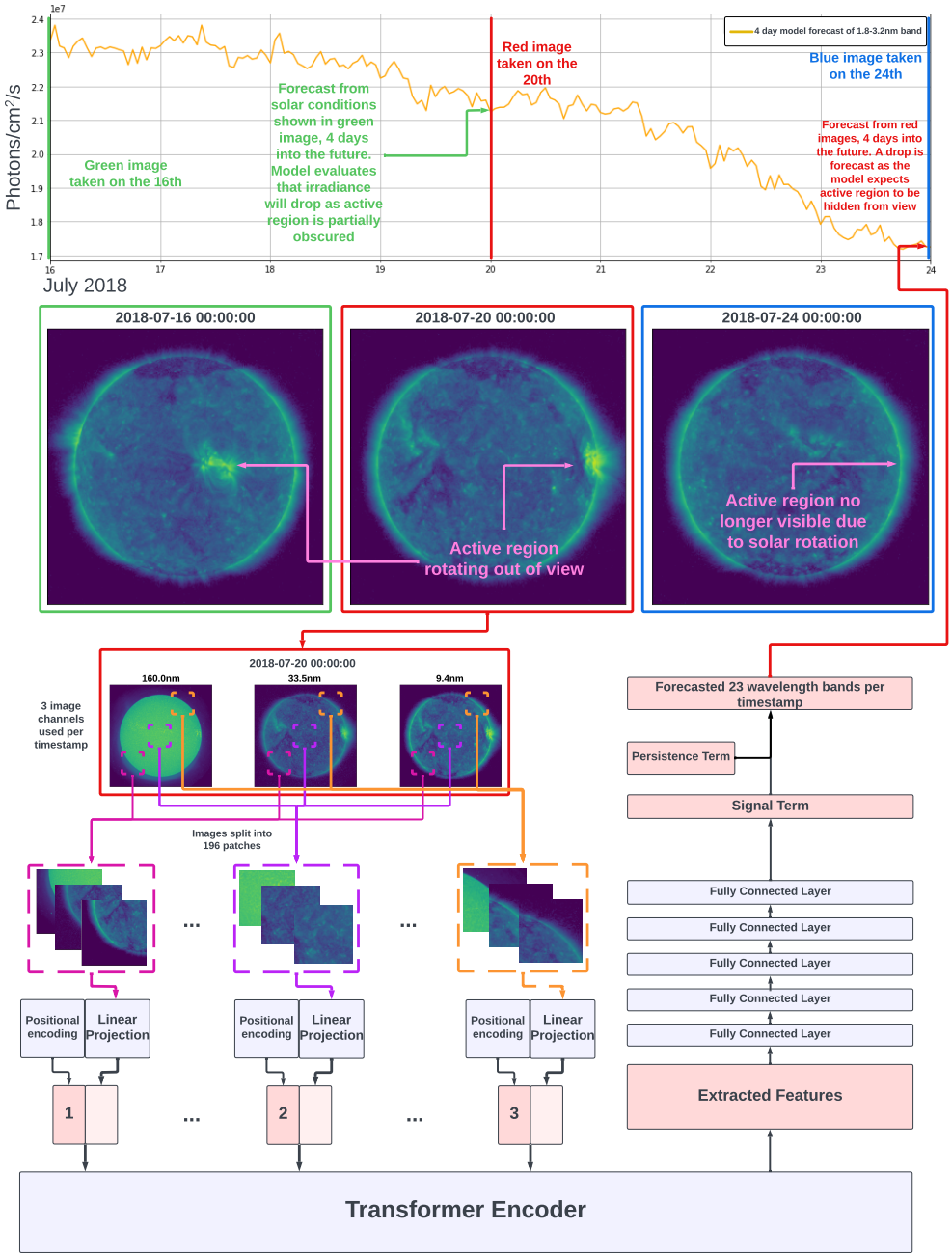


Figure 1: Forecasting irradiance procedure from solar images using the proposed model. The top graph shows the 4-day forecast over a period from the 16^{th} to the 24^{th} of July, 2018 for a disappearing active region. The three central images are 9.4nm images of the solar disk on the 16^{th}, 20^{th}, and 24^{th} respectively (color-coded green, red, and blue). Below the images is a diagram showing the model architecture for producing the forecasts.

to condense their informational content. Then they are augmented with additive positional encoding, to preserve relative position information. The resulting set of 196 condensed image slices is then fed into the vision transformer (Dosovitskiy et al., 2020) encoder.

The vision transformer is a highly complex model that replaces the fixed set of parameters learned by other machine learning models, with a trainable sub-routine that learns to generate weights for the input image's analysis. This generation is performed from scratch and separately for each input image. As the sub-routine for generating these weights uses the input image itself in the weight generation, this is often referred to as self-attention.

The solar irradiance forecast is calculated by the additive combination of a signal term and a persistence term. The signal term is derived from the vision transformer's output. The features extracted by the transformer, collected in a 768-long vector, are processed by a five-layer fully connected neural network classifier. The persistence term is the solar irradiance value at the time when the input images were taken.

The selection of the 9.4nm, 33.5nm, and 160.0nm images is desirable for two reasons. Firstly, using three channeled input works well with the three-channel-based pre-trained transformer feature extractor. Replicating the transformer backbone, or training a version from scratch could allow for increasing the channel count. This, however, would come at a considerable resource and data cost. Using the three-layer setup was judged to be a suitable compromise between cost, feasibility, and performance. Secondly, the three wavelengths were chosen such that they cover as much of the "stan" band wavelength range as possible. The 9.4nm and 33.5nm images cover the widest wavelength range within the SDO EUV (as opposed to UV or visible) image dataset. The wavelength of the upper bound in the the 23 FISM2 'stan bands' is at 121.0nm. Therefore the next closest channel, in terms of absolute wavelength, to this value is the 160.0nm image channel. Consequently, these three image channels cover the forecasted spectrum to the fullest extent available from the SDOML image dataset, within the constraint of three total image channels.

3.2 Mean Absolute Percentage Error (MAPE)

As the number of photons per cm^2 per second varies very significantly between irradiance bands, a unit-less error metric is necessary to compare the accuracy of predictions corresponding to different bands. Accordingly, this work reports the mean absolute percentage error (MAPE):

$$\frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - x_i}{x_i}\right|$$

where y_i and x_i are the predicted and true photons per cm^2 per second for the i^{th} data point respectively. For example, if the true value of the photons per cm^2/s for a particular wavelength band is 10, but the model outputs 8, then the contribution of that particular data point to the mean absolute percentage error is 20%. The MAPE reports the average contribution across all datapoints.

3.3 Persistence Score (PS)

Given the high degree of autocorrelation in the irradiance data, model performance needs to be judged in the broader context of the task difficulty. As one always has the option to revert back to a simple persistence model, i.e. one using the current value without change as the prediction, it is important to compare the performance of forecasting models not only to the ground truth, but also to the performance of the baseline persistence model. Consequently, this study also reports the Persistence Score (PS):

$$(1 - \frac{M_m}{M_p}) * 100$$

where M_m and M_p are the MAPE for the model and the persistence model respectively. The score ensures that the models are actually adding value by re-expressing MAPE in terms of performance compared to the baseline persistence model. Its value potentially ranges from $-\infty$ to 100, with higher scores corresponding to better performance, and 0 being parity between the two models. To further contextualise the metric: if a model has a MAPE of 30% and the persistence model has a MAPE of 50%. Then the persistence score is $(1-\frac{30}{50})\times 100 = 40$. This shows that the improvement of the model MAPE compared to the persistence model MAPE has resulted in a reduction in the MAPE of 40% (as a percentage of the persistence model MAPE).

3.4 Dataset Preparation, training details, and validation setup

Details concerning the scaling and pre-processing of the irradiance and image data, as well as the training and validation procedures are included in the Appendix.

4 Experimental Results

First, the importance of joint inclusion of at least three input image channels is established, as opposed to a single channel alone (4.1). Then, the method is shown to be a robust and accurate forecaster of solar irradiance (4.2). Finally, the value of the persistence term is then highlighted and evaluated (4.3).

4.1 Relative and joint importance of the EUV/UV images

We first turn to the question of the relative and the joint importance of the EUV/UV images at 9.4nm, 33.5nm and 160nm wavelengths in modelling the fuller spectra. The model is first trained to nowcast the irradiance bands, that is, to calculate the same-period irradiance values from the contemporary images. The results of models trained on single image channels are compared to a model trained using all three images as an input. The setup is similar to that in Szenicer et al. (2019), however, there, multiple irradiance lines, not bands, were nowcast from EUV solar images in order to prove the potential for extracting irradiance information from EUV images. Fundamentally, the objective is to demonstrate how much information the chosen model can gain from using multiple image channels compared to each channel individually, testing the choice of multiple image channels in the forecasting model. The model is the same as the forecasting model described in section 3.1, without the persistence term, and used to nowcast current irradiance values from current images, rather than forecasting future irradiance values from current images.

Figure 2 shows the results of nowcasting experiment with 9.4nm, 33.5nm, 160.0nm, and a model trained on all three.

The top panel of figure 2 shows that the shorter wavelength bands are broadly reconstructed with a much higher MAPE. In other words, they are harder to reconstruct. The underlying cause of this may be their instability, or generally higher noise content. However, the most likely, and the leading, cause in the performance deterioration is the insufficient range of the solar images. The shortest wavelength represented in the image channels is 9.4nm which is significantly longer than the shortest 0.05-0.4nm band.

The middle table of figure 2 presents the raw data provided in the top graph. The raw data shows that the three-input model is very capable of reconstructing the bands from images, especially for longer wavelength bands with a MAPE as low as 0.38% for the 65.0-79.8nm high band. The entries are colour-coded so that within each irradiance band the darker the colour is, the lower is the relative ranking of its entry. White entries correspond to the best model for the given band, while dark grey entries correspond to the worst-performing model. Two key observations can be made here. First, the combina-

tion of all three models is a clear top performer of the comparison as it ranks first in all but 2 of the 23 bands. This is to be expected as adding image channels increases relevant information about the irradiance spectra. Second, for shorter wavelength bands, it is generally the case that 9.4nm images give better reconstruction than 33.5nm, which are better than 160.0nm. Beyond 91.3nm irradiance, however, the ordering switches, with 160.0nm no longer being the worst performer.

The bottom panel of figure 2 reports the MAPE difference, expressed as a percentage, of the single channel models, the one based on 9.4nm (blue), on 33.5nm (red), and on 160.0nm (yellow), compared to the three-channel model. Generally, it shows that there is a significant performance improvement for using all three channels compared to just using a single channel (with two minor magnitude exceptions at the 7.0-15.5nm and 15.5-22.4nm bands). For the very short first and second wavelength bands at 0.05-0.4nm and 0.4-0.8nm, there is a vastly significant increase in the MAPE for 33.5nm and 160.0nm. With error increases above 50% for both the 33.5nm and 160.0nm bands. However, the error increase is still significant at 10.5% and 20.5 when using the 9.4nm channel. This particular result is striking as it implies that even though the 0.05-0.4nm and 0.4-0.8nm bands are more than an order of magnitude shorter than the wavelength of the shortest image channel used, the combination of using all the image channels still significantly improves performance compared to the shortest wavelength channel alone.

In general, the observed behaviour suggests that shorter wavelength EUV images contain more information relevant for reconstructing shorter wavelength irradiance, mid to longer wavelength EUV images contain more information relevant to mid to long wavelength irradiance bands, and that the proposed feature extractor, when used with all three channels at the same time, can appropriately link inputs with the relevant parts of its output vector resulting in superior modelling than using single channels alone.

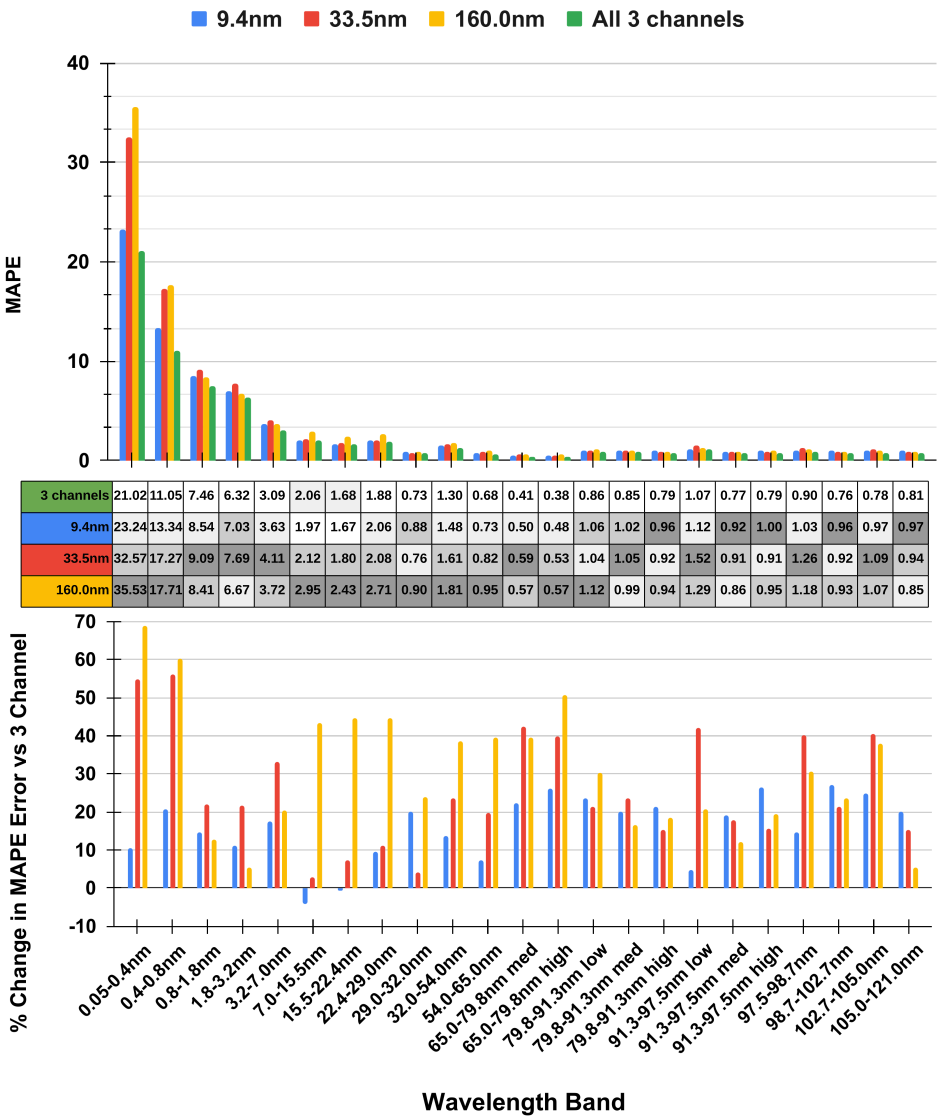


Figure 2: FISM2 replication using a model trained on 9.4nm, 33.5nm, 160nm, and all three images at once. The top graph shows the overall MAPE for each band for each model. The central table shows the raw MAPE data but colour-coded to show positional order of each model (darker represents a higher MAPE in relation to the others). The bottom graph shows the percentage reduction in MAPE of the three channel model versus the three single channels alone.

4.2 Forecasting Results

Having established the extent to which the model is able to recreate the FISM2 'stan' bands from images and that using multiple image channels improves upon single-image models, the study looks at the ability for the model described in 3.1 to forecast the spectra at a 24 to 96 hour horizon.

Table 1 reports the Mean Absolute Percentage Error (MAPE) and the Persistence Score (PS) of the 24, 48, 72, and 96 hour (1-4 days) forecasts of the 23 irradiance bands, ranging from 0.05nm to 121.0nm. The proposed method is shown to be a robust and accurate forecaster of solar irradiance both in absolute terms (MAPE) and when compared relative to the persistence baseline (PS).

Regarding the MAPE scores, the study makes the distinction between results for wavelengths below and above 7.0nm, as shown in the 3rd and 2nd last rows in Table 1. This distinction is made because, within bands below 7.0nm, the wavelengths are outside of the range reflected in the images, i.e. 9.4nm to 160.0nm. The 7.0-121.0nm subcategory achieves incredibly low errors at 0.88, 1.19, 1.40, and 1.72 for 1-4 days respectively. The 0.05-7.0nm range achieves 13.13, 17.60, 18.83, and 21.43, which is much higher than that of the longer bands. This performance disparity makes physical sense as the shorter wavelengths' forecasts are extrapolating the short wavelength irradiance from image data which does not cover those bands within their range. The MAPE across the 23 bands is 3.54, 4.76, 5.19, 6.00 for 1-4 days respectively.

The persistence score averaged across all bands is 16.88, 26.67, 35.27, 36.45 for 24, 48, 72, and 96 hours, respectively. These values represent a significant improvement over the baseline persistence model, with the strongest over-performance seen for the longer forecasting horizons. The magnitude of the over-performance is very significant. Indeed, the mean persistence score of 36.45 at 96 hours means that the model's MAPE is 36.45% lower compared to MAPE of the persistence model.

The individual bands' results show an improvement over the baseline persistence model in the vast majority of bands (23) and time horizons (4), with 91 out of 92 band/time horizon categories returning a positive, meaning outperforming, persistence score. The extent by which the model outperforms the persistence model increases the further into the future the forecast is made, shown by the increasing persistence score with time horizon, with very few exceptions. Notably, the shortest wavelength has very high MAPE, as high as 52.02% at 96 hours. This band is particularly difficult to forecast due to dramatic swings in its absolute value over short timescales. However, the high persistence

Table 1: Solar irradiance forecasting results for models trained with 24, 48, 72, and 96 hours. The Mean Absolute Percentage Error (MAPE) and the Persistence Score (PS) are reported for each band individually, the two subcategories 0.05nm-7.0nm and 7.0-121.0nm, and the total.

Band	24 hours		48 hours		72 hours		96 hours	
	MAPE	PS	MAPE	PS	MAPE	PS	MAPE	PS
0.05-0.4nm	38.46	33.93	47.27	33.68	47.6	40.77	52.02	44.81
0.4-0.8nm	16.82	25.15	22.37	23.66	23.68	32.04	26.94	34.36
0.8-1.8nm	4.94	9.47	8.61	8.14	10.58	16.96	12.88	17.84
1.8-3.2nm	3.46	11.2	6.3	9.89	7.97	17.46	9.87	17.21
3.2-7.0nm	1.99	11.98	3.43	17.09	4.32	25.46	5.45	24.21
7.0-15.5nm	1.44	12.18	2.21	21.0	2.75	26.77	3.4	25.64
15.5-22.4nm	1.15	12.9	1.82	23.69	2.28	30.14	2.86	29.02
22.4-29.0nm	1.25	25.69	2.07	32.3	2.61	39.24	3.36	38.21
29.0-32.0nm	0.61	29.12	0.91	40.78	1.13	47.61	1.51	45.31
32.0-54.0nm	1.02	21.31	1.51	32.51	1.88	39.42	2.39	39.11
54.0-65.0nm	0.78	9.59	0.96	26.55	1.11	36.0	1.36	36.59
65.0-79.8nm med	0.53	9.05	0.64	21.45	0.68	32.89	0.8	35.62
65.0-79.8nm high	0.52	12.84	0.62	25.65	0.69	35.68	0.82	37.04
79.8-91.3nm low	0.9	18.91	1.17	29.93	1.28	41.2	1.54	43.33
79.8-91.3nm med	0.82	19.89	1.04	35.55	1.2	44.66	1.49	45.65
79.8-91.3nm high	0.79	20.06	1.01	32.9	1.15	42.86	1.38	45.21
91.3-97.5nm low	1.08	-0.31	1.22	21.14	1.45	27.48	1.64	33.06
91.3-97.5nm med	0.63	24.27	0.85	37.8	1.0	46.34	1.27	46.17
91.3-97.5nm high	0.71	26.78	0.98	39.51	1.17	47.98	1.48	48.05
97.5-98.7nm	1.32	3.49	1.57	5.63	1.57	18.86	1.7	22.91
98.7-102.7nm	0.72	22.52	0.94	36.78	1.11	44.88	1.38	45.59
102.7-105.0nm	1.07	12.43	1.27	23.3	1.37	33.46	1.57	36.8
105.0-121.0nm	0.52	15.86	0.71	34.45	0.85	43.02	1.0	46.56
0.05-7.0nm mean	13.13	18.35	17.60	18.49	18.83	26.54	21.43	27.69
7.0-121.0nm mean	0.88	16.48	1.19	28.94	1.40	37.69	1.72	38.88
Total mean	3.54	16.88	4.76	26.67	5.19	35.27	6.00	36.45

score of 44.81 at 4 days, really highlights that the model adds value compared to the persistence value alone.

In conclusion, the methods in this study represent an opportunity to provide accurate forecasts of thermospheric drivers using solar imaging. These results also open a question of how important is the inclusion of the persistence term in the model, given these significant performance increases over the persistence model.

4.3 Persistence term evaluation

Figure 3 reports the percentage changes in MAPE, across the 23 irradiance bands and for the 4 forecasting horizons, between the previously described full model, based on the persistence-signal structure, and a model that outputs the forecasted irradiance directly without using the persistence term as an input.

Adding the persistence term significantly aids the model across the predicted irradiance bands at shorter timescales. At the 24h horizon, in some cases the model's MAPE improves by up to 54.5% when the persistence term is added to it. This change is, however, not uniform across bands. Furthermore, the benefit diminishes at longer timescales. Indeed, the benefit is mostly single digit persistence term benefit at the 96h horizon, even occasionally resulting in worse performance for some bands.

These results point to the utility of the persistence term at shorter timescales. It is highly beneficial to the 24h forecast, however this effect diminishes as the time horizon is extended. More generally, the importance of using the persistence term in the modelling points to a broader idea of combining solar images and time series of data in forecasting phenomena at Earth, especially at shorter time scales.

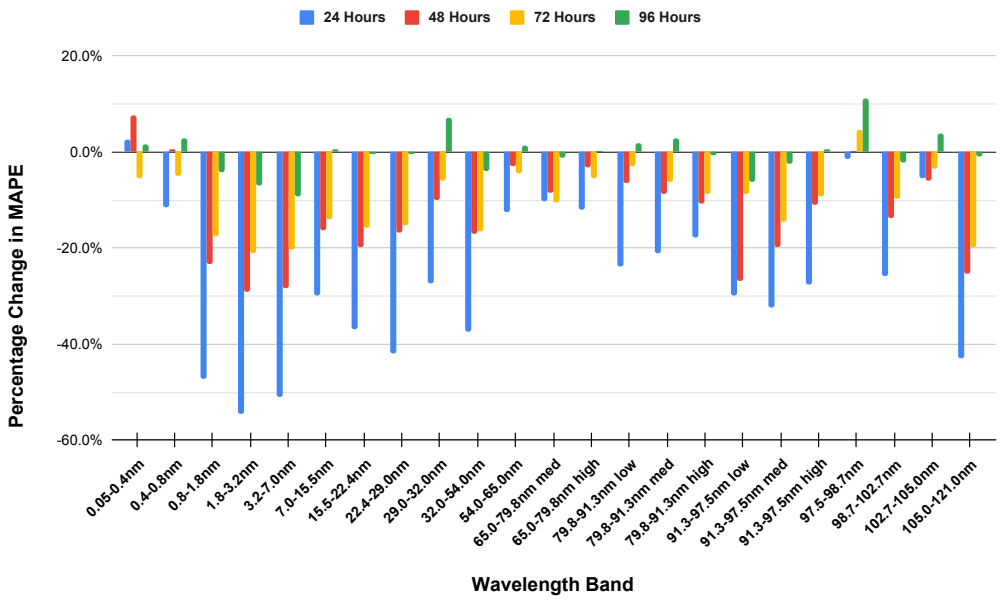


Figure 3: Percentage reduction in MAPE by using the persistence value as an input to the model versus without

5 Discussion and Extensions

This study can be furthered in at least three distinct directions. Firstly, it was limited by the use of three image channels, the 9.4nm, 33.5nm, and 160.0nm, even though the SDO datasets offers many more wavelengths. It is likely that using more image wavelengths would improve performance further, and so future studies may look to use all 12 image channels available in the SDO/SDOML dataset.

Secondly, only one architecture was tested in this work. This choice was driven by a previous work that established a very similar architecture as a top performing feature extractor for solar images (Brown et al., 2022). However, future methodological developments, especially in the space of data efficiency, i.e. the amount of data needed to train a given model, have the potential to improve on the current results even allowing training an all SDO wavelengths from scratch on a custom architecture.

Finally, the study's forecast horizon was set at 1-4 days, i.e. 24 to 96 hours. The choice was based on practical resource considerations, and limited by the amount of compute power available to the authors. The chosen horizons are believed to be sufficiently diverse to demonstrate the key forces behind the model's performance, in particular the changes in the relative importance of the signal and the persistence terms. However, finer sampling and longer time horizons might be of interest to some readers.

6 Conclusions

Vital spacecraft operations, such as collision avoidance, rely on accurate modelling of the thermosphere. Solar extreme ultraviolet radiation is a main driver of thermospheric density. Forecasting the density of the thermosphere therefore relies on accurate forecasts of the incident radiation. This study uses deep learning computer vision techniques to forecast the solar irradiance from solar images in 23 wavelength bands from 0.05nm to 121.0nm. The principle results are:

- Using multiple image channels improves EUV irradiance modelling, with significant improvements at shorter wavelength bands, even at wavelength bands significantly shorter than the shortest wavelength image channel used.
- The model is able to forecast 23 wavelength bands from 0.05nm to 121nm (produced from the FISM2 empirical irradiance model (Chamberlin et al., 2020)) up to four days in advance with a MAPE of 6.0% and an average improvement in the mean absolute percentage error of 36.45% over the baseline persistence model.
- Including the persistence term (the current value of the irradiance being forecast) as an input to the model results significantly increases model performance at shorter timescales, with diminishing value at the 4 day mark.

Future extensions of this study might seek to expand its scope by improving the efficiency of data processing and consumption through the use of any potential future developments in the field of deep learning vision. Alternatively, the study can be used as a proof of concept for the sub-domain of irradiance/thermospheric forecasting and a springboard for further research.

7 Data Availability Statement

The SDOML image data is available here: `https://purl.stanford.edu/vk217bh4910`

The FISM2 irradiance data is available at: `https://lasp.colorado.edu/lisird/`

The full codebase for training and evaluating models is available here: `https://github.com/eddbrown/irradiance-forecast`

Acknowledgements

For this study, we acknowledge extensive use of the SDOML (Galvez et al., 2019) dataset. These images are processed versions of images taken by the Atmospheric Imaging Assembly (Lemen et al., 2012) instrument aboard the Solar Dynamics Observatory (Pesnell et al., 2012).

For experiment tracking and general project organisation, we acknowledge extensive use of the software provided by Weights and Biases (Biewald, 2020).

For ease of using pretrained deep learning models, we acknowledge the use of the pytorch image models library (Wightman, 2019).

Many thanks are given to the British Antarctic Survey IT team for their computing services, enabling this work.

We would like to acknowledge funding from the Natural Environment Research Council grants NE/V00249X/1 (Sat-Risk) and NE/R016038/1

Lastly, the empirical model that is FISM2 (Chamberlin et al., 2020) has been an incredible resource for exploring irradiance modelling and forecasting.

8 Appendix

8.1 Data preparation

Each of the three input images to the model is subjected to the same pre-processing, and this is very similar to that first proposed in (Upendran et al., 2020). First, the images are clipped between two pixel values. Next the natural logarithm of that clipped value is taken. Then, they are linearly re-scaled so the minimum value is 0 and the maximum value is 1. Finally, the images are resized to 224 by 224 pixels before use in the model. The clipping values are chosen by taking the 20th and 99.99th percentile pixel values taken from 100 images randomly sampled from the dataset. The former percentile is justified by the fact that roughly 20% of the image is off-disk background. The latter by the desire to remove the extreme positive outliers from skewing the empirical data distribution, hence the 99.99^{th} percentile. In the here-presented dataset, a sample of 100 images randomly selected from the dataset had the minimum and maximum pixel values 0.52 and 781.95 for 9.4nm, 2.1 and 946.49 for 33.5nm, 5.04 and 1367.08 for 160nm. In order to improve information extraction from the dataset, data augmentation techniques first proposed in Brown et al. (2022) were employed. That is, transformed images were used in training alongside the untempered, original, data. Specifically, images were randomly flipped north to south. It is not claimed that these Suns are 'valid' scientifically. However, the features present in these images are valid enough to contain features relevant to training, while at the same time being distinct enough to be useful data augmentation techniques. Empirically, this treatment aids in model convergence through regularization.

As was the case with images, irradiance was logged and scaled. Each datum at each band was individually natural-logged and then scaled to between 0 and 1 using the minimum and maximum of the full training set per band. Outliers in the irradiance data are particularly severe. In particular, over the 2010-2018 period when powerful flares happened, the FISM2 model had outliers with wavelength bands many orders of magnitude higher than normal. For example, on 2012-03-23 at 00:30:00, the 105-121nm band had 3.2×10^{19} photons per cm^2 per second, which is nine orders of magnitude greater than its median of 3.04×10^{10} photons per cm^2 per second. Observations with irradiance values more 1,000 times over the median were excluded from the dataset on the grounds of being outliers. In total, only 31 observations out of the approximately 120,000 were dropped on the basis of this rule.

8.2 Training and evaluation

In order to have data from across the solar cycle represented equally in training, validation, and the test sets a targeted dataset partition method was employed. The desired relative 30-5-5 train-validation-test split was achieved by the repeated application of a data assignment pattern on data queried at 30 minute intervals between May 1, 2010 and December 31, 2018. Its dual role was to ensure both the equal sampling across the solar cycle and the prevention of train-validation and validation-test set cross-contamination. The latter was a very real possibility due to the data's high degree of autocorrelation at 48 hours. The sampling method is very similar to Upendran et al. (2020) and Brown et al. (2022). In a single application of the data-splitting pattern, the first 30 days of data were assigned to the train set, followed by 2 days of buffer data that were discarded. Next 5 days of data were earmarked for validation, and followed by further 2 days of buffer. Finally, a 5 day section was assigned to the test set, and again followed by a 2 day buffer. The final dataset size consisted of approximately 90,000, 15,000, and 15,000 images for the training, validation, and test set, respectively.

The model's transformer backbone was initialised using a set of pre-trained general-purpose object recognition weights. Its other layers were initialised using LeCun Uniform initializer (Paszke et al., 2019). All model parameters, including those of the transformer backbone, were trained.

The model was trained using the Adam optimizer (Kingma & Ba, 2014) (batch size 32) in PyTorch (Paszke et al., 2019) for 20 epochs, using mean squared error (MSE) loss function. It was validated on the full validation set after each epoch. The weights of the model are saved after every epoch. The set of weights that performed best over the validation set are used as the final set of weights for the model. The performance of this model is then evaluated on the test set and reported.

References

Berger, T., Holzinger, M., Sutton, E., & Thayer, J. (2020). Flying through uncertainty. *Space Weather*, *18*(1), e2019SW002373.

Bernoux, G., Brunet, A., Buchlin, É., Janvier, M., & Sicard, A. (2022). Forecasting the geomagnetic activity several days in advance using neural networks driven by solar euv imaging. *Journal of Geophysical Research: Space Physics*, *127*(10), e2022JA030868.

Biewald, L. (2020). *Experiment tracking with weights and biases.* (Software available from wandb.com)

Brown, E. J., Bonasera, S., Benson, B., Pérez-Hernández, J. A., Acciarini, G., Baydin, A. G., ... Jah, M. K. (2021). Learning the solar latent space: sigma-variational autoencoders for multiple channel solar imaging. *NeuRIPS: Machine Learning for Physical Sciences*.

Brown, E. J., Svoboda, F., Meredith, N. P., Lane, N., & Horne, R. B. (2022). Attention-based machine vision models and techniques for solar wind speed forecasting using solar euv images. *Space Weather*, *20*(3), e2021SW002976.

Chamberlin, P. C., Eparvier, F. G., Knoer, V., Leise, H., Pankratz, A., Snow, M., ... Woods, T. N. (2020). The flare irradiance spectral model-version 2 (fism2). *Space Weather*, *18*(12), e2020SW002588. doi: https://doi.org/10.1029/2020SW002588

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, *abs/2010.11929*.

Galvez, R., Fouhey, D. F., Jin, M., Szenicer, A., Muñoz-Jaramillo, A., Cheung, M. C. M., ... Thomas, R. (2019, may). A machine-learning data set prepared from the NASA solar dynamics observatory mission. *The Astrophysical Journal Supplement Series*, *242*(1), 7. doi: 10.3847/1538-4365/ab1005

Hu, A., Shneider, C., Tiwari, A., & Camporeale, E. (2022). Probabilistic prediction of dst storms one-day-ahead using full-disk soho images. *Space Weather*, *20*(8), e2022SW003064.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., ... others (2013). The community earth system model: a framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lemen, J. R., Title, A. M., Akin, D. J., Boerner, P. F., Chou, C., Drake, J. F., ...

others (2012). The atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo). *Solar Physics*, *275*, 17–40.

Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J.-F., Calvo, N., & Polvani, L. M. (2013). Climate change from 1850 to 2005 simulated in cesm1 (waccm). *Journal of climate*, *26*(19), 7372–7391.

OmniWeb. (n.d.). *Omniweb data archive.* `https://omniweb.gsfc.nasa.gov/form/dx1.html`. (Accessed: 2021)

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc.

Pesnell, W. D., Thompson, B. J., & Chamberlin, P. (2012). The solar dynamics observatory (sdo).

Raju, H., & Das, S. (2021). Cnn-based deep learning model for solar wind forecasting. *Sol Phys*. doi: https://doi.org/10.1007/s11207-021-01874-6

Solomon, S. C., & Qian, L. (2005). Solar extreme-ultraviolet irradiance for general circulation models. *Journal of Geophysical Research: Space Physics*, *110*(A10).

Sun, D., Huang, X., Zhao, Z., & Xu, L. (2023). Deep learning–based solar flare forecasting model. iii. extracting precursors from euv images. *The Astrophysical Journal Supplement Series*, *266*(1), 8.

Szenicer, A., Fouhey, D. F., Munoz-Jaramillo, A., Wright, P. J., Thomas, R., Galvez, R., ... Cheung, M. C. (2019). A deep learning virtual instrument for monitoring extreme uv solar spectral irradiance. *Science advances*, *5*(10), eaaw6548.

Upendran, V., Cheung, M. C. M., Hanasoge, S., & Krishnamurthi, G. (2020). Solar wind prediction using deep learning. *Space Weather*, *18*(9), e2020SW002478. (e2020SW002478 10.1029/2020SW002478) doi: https://doi.org/10.1029/2020SW002478

Wightman, R. (2019). *Pytorch image models.* `https://github.com/rwightman/pytorch-image-models`. GitHub. doi: 10.5281/zenodo.4414861

Figure 1.

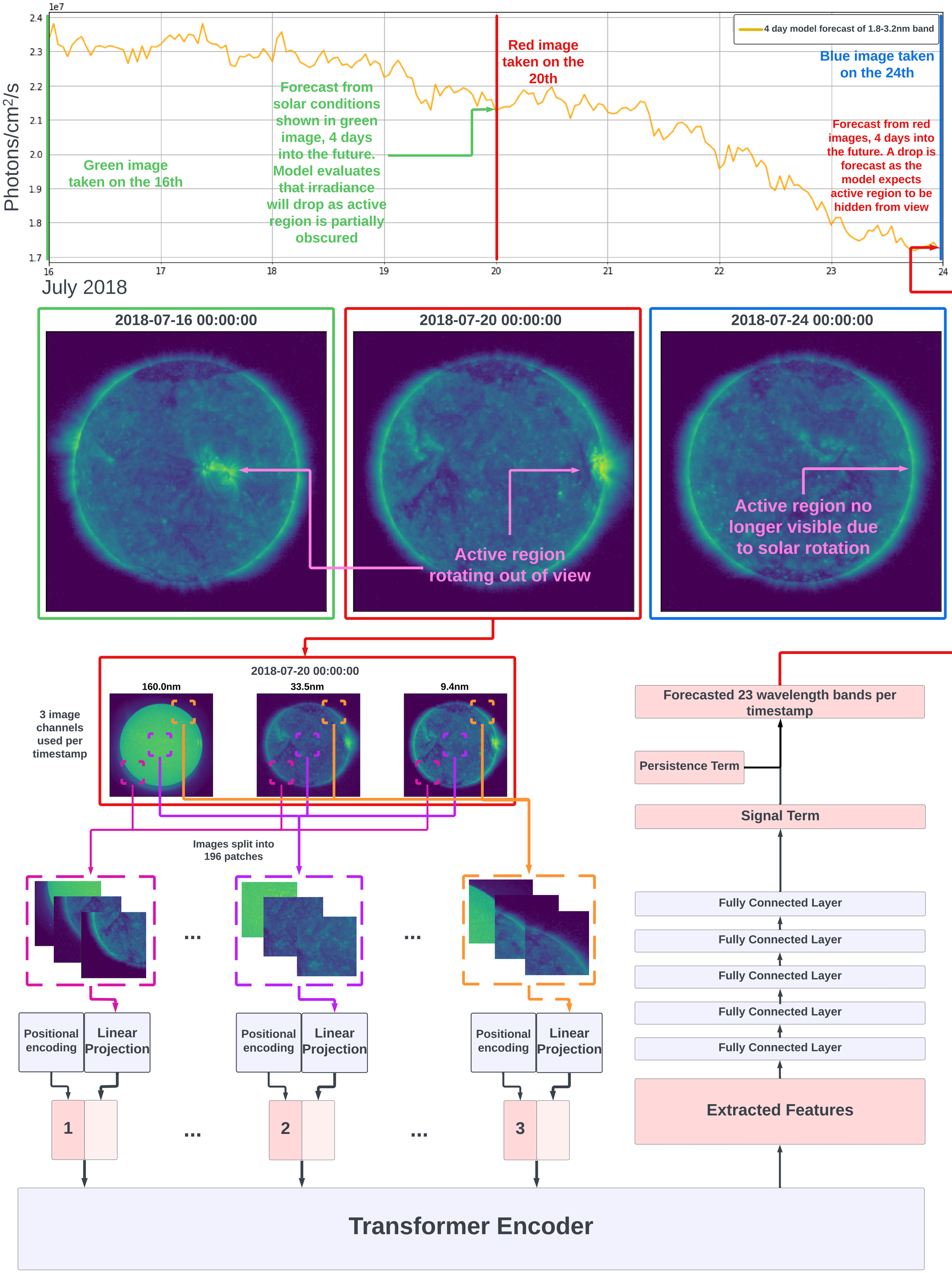
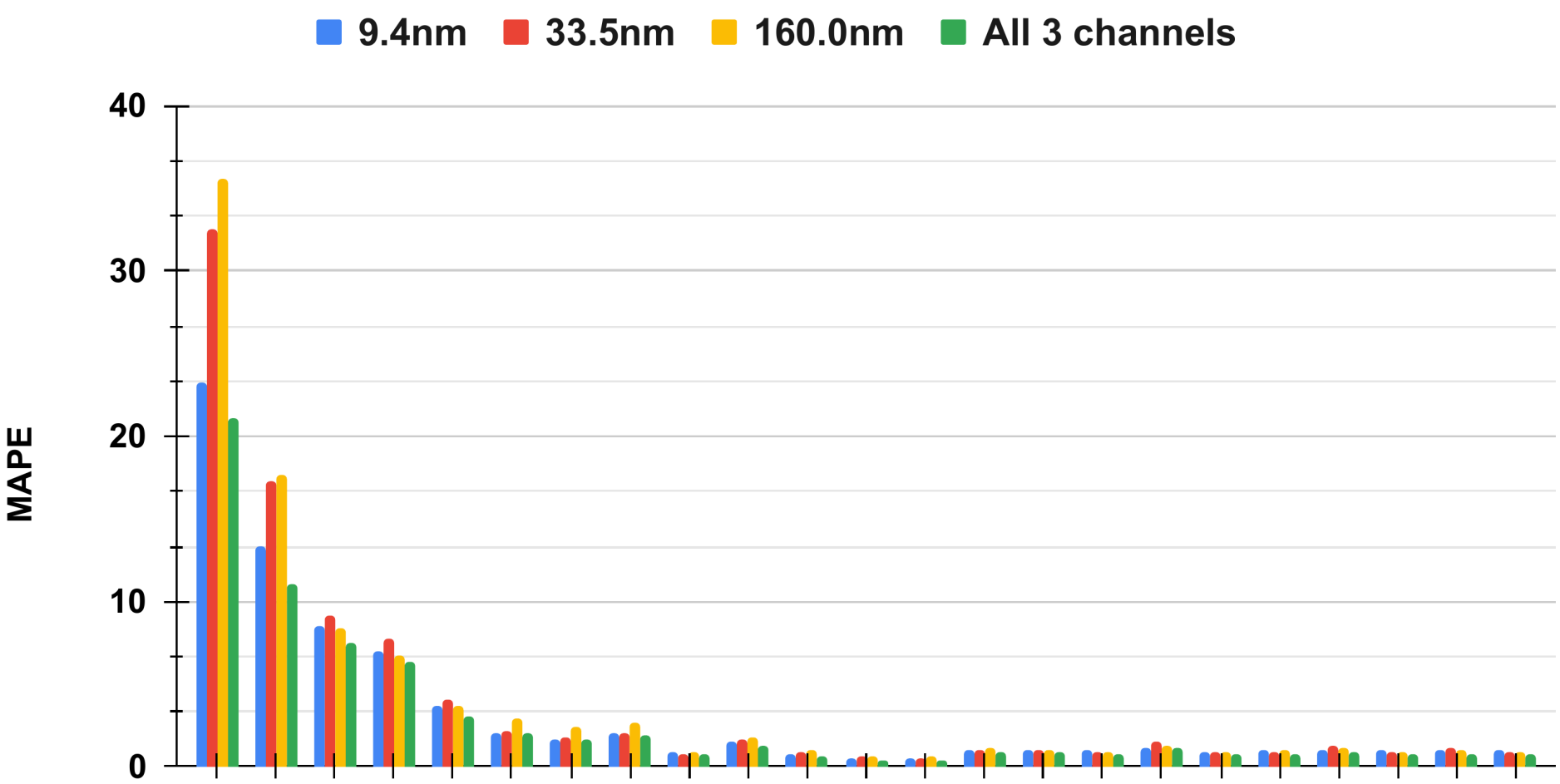

1e7
Photons/cm²/s
4 day model forecast of 1.8-3.2nm band
Green image taken on the 16th
Forecast from solar conditions shown in green image, 4 days into the future. Model evaluates that irradiance will drop as active region is partially obscured
Red image taken on the 20th
Blue image taken on the 24th
Forecast from red images, 4 days into the future. A drop is forecast as the model expects active region to be hidden from view
July 2018
2018-07-16 00:00:00
2018-07-20 00:00:00
2018-07-24 00:00:00
Active region rotating out of view
Active region no longer visible due to solar rotation
2018-07-20 00:00:00
160.0nm
33.5nm
9.4nm
3 image channels used per timestamp
Images split into 196 patches
Positional encoding
Linear Projection
1
2
3
Transformer Encoder
Forecasted 23 wavelength bands per timestamp
Persistence Term
Signal Term
Fully Connected Layer
Fully Connected Layer
Fully Connected Layer
Fully Connected Layer
Fully Connected Layer
Extracted Features

Figure 2.



3 channels	21.02	11.05	7.46	6.32	3.09	2.06	1.68	1.88	0.73	1.30	0.68	0.41	0.38	0.86	0.85	0.79	1.07	0.77	0.79	0.90	0.76	0.78	0.81
9.4nm	23.24	13.34	8.54	7.03	3.63	1.97	1.67	2.06	0.88	1.48	0.73	0.50	0.48	1.06	1.02	0.96	1.12	0.92	1.00	1.03	0.96	0.97	0.97
33.5nm	32.57	17.27	9.09	7.69	4.11	2.12	1.80	2.08	0.76	1.61	0.82	0.59	0.53	1.04	1.05	0.92	1.52	0.91	0.91	1.26	0.92	1.09	0.94
160.0nm	35.53	17.71	8.41	6.67	3.72	2.95	2.43	2.71	0.90	1.81	0.95	0.57	0.57	1.12	0.99	0.94	1.29	0.86	0.95	1.18	0.93	1.07	0.85

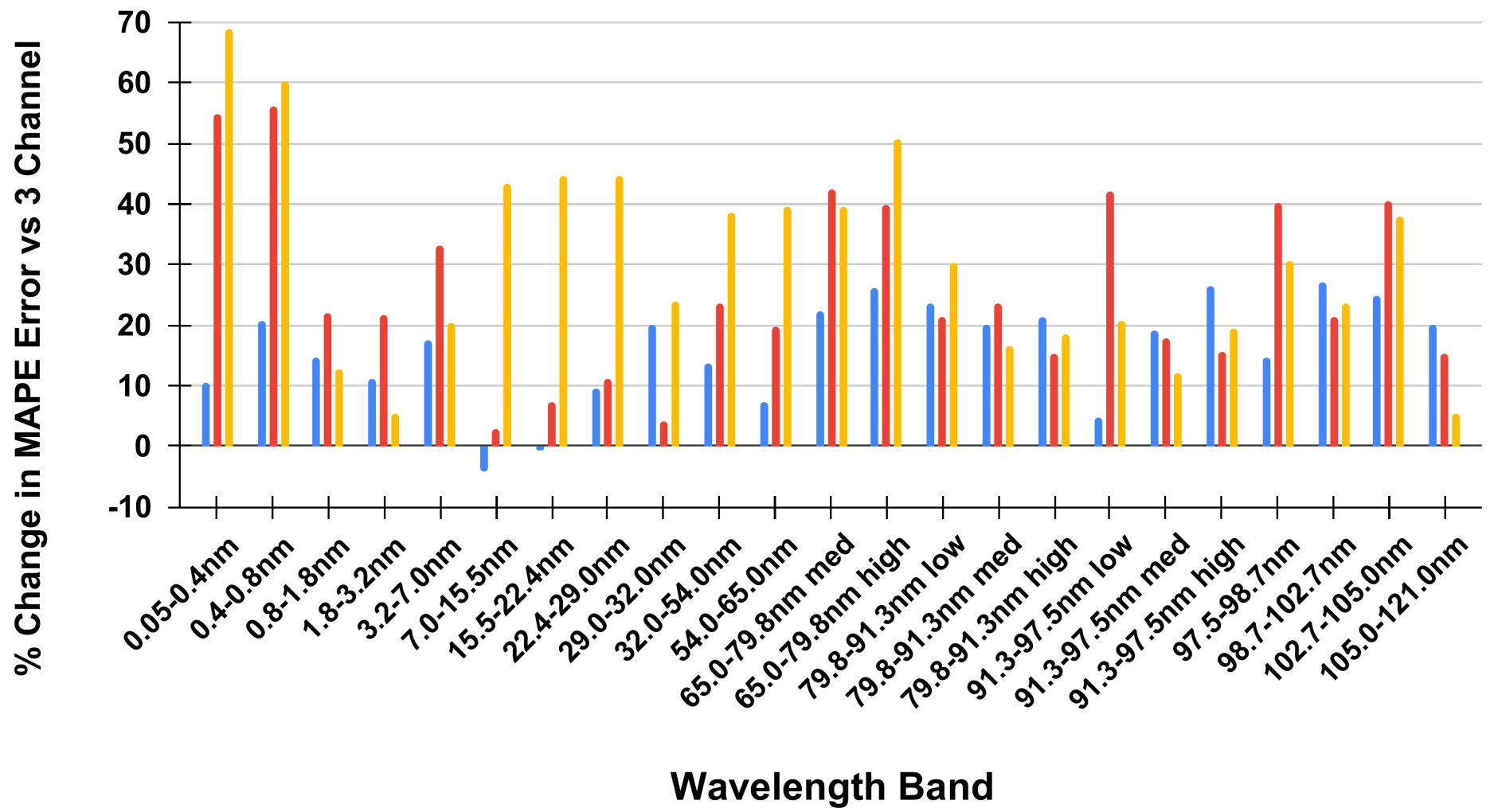
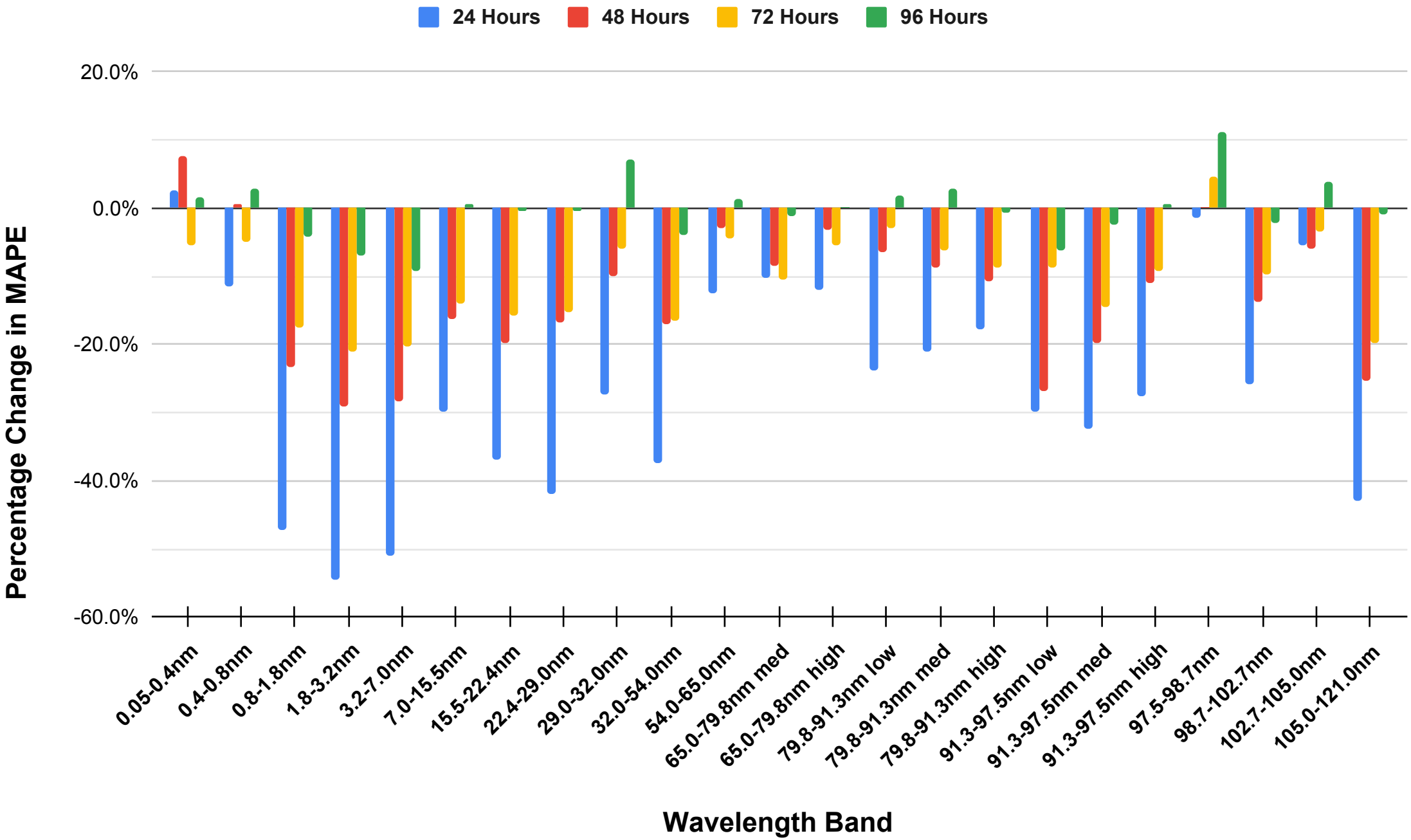


Figure 3.


24 Hours
48 Hours
72 Hours
96 Hours
20.0%
0.0%
-20.0%
-40.0%
-60.0%
Percentage Change in MAPE
0.05-0.4nm
0.4-0.8nm
0.8-1.8nm
1.8-3.2nm
3.2-7.0nm
7.0-15.5nm
15.5-22.4nm
22.4-29.0nm
29.0-32.0nm
32.0-54.0nm
54.0-65.0nm
65.0-79.8nm med
65.0-79.8nm high
79.8-91.3nm low
79.8-91.3nm med
79.8-91.3nm high
91.3-97.5nm low
91.3-97.5nm med
91.3-97.5nm high
97.5-98.7nm
98.7-102.7nm
102.7-105.0nm
105.0-121.0nm
Wavelength Band