



TRILLIUM USA

FDLXHELIO.ORG

A Scientific Cloud Computing Platform for Ingestion and Processing of SDO Data

DASH 2023

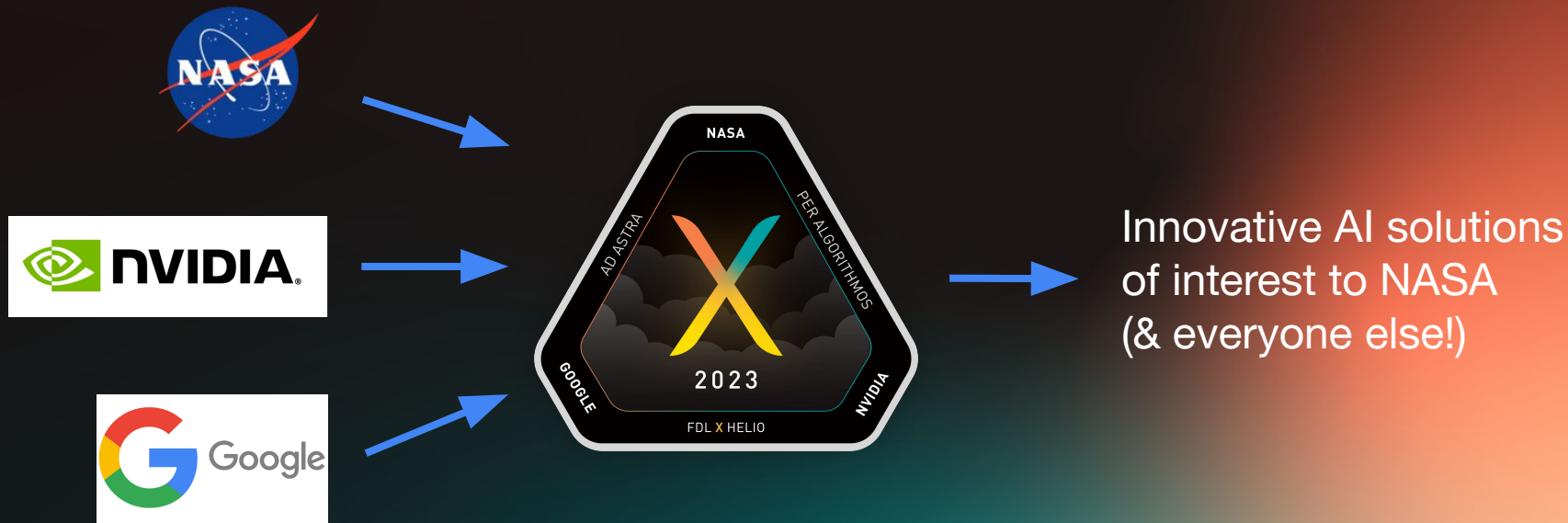
Monday 8th October

William Fawcett¹

Daniel Gass², Manuel Indaco³, Richard Galvez⁴,
Andrés Muñoz-Jaramillo⁵, Paul Wright⁶

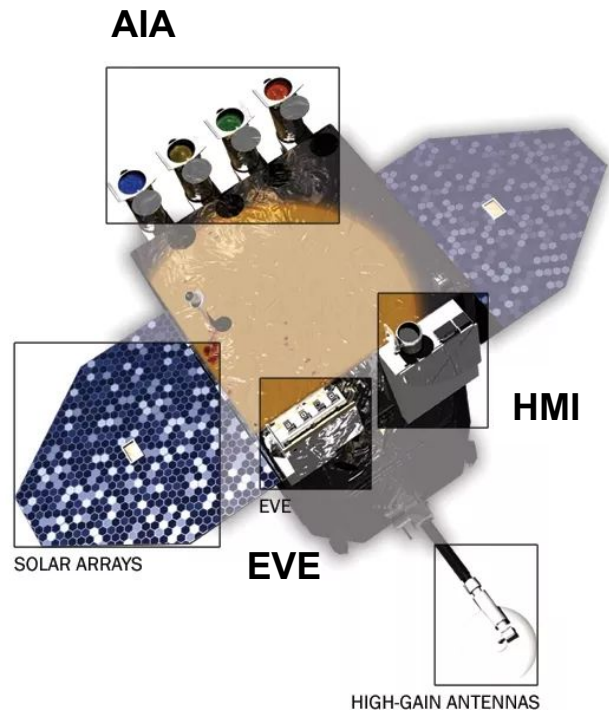
¹University of Cambridge, ²UCLAN, ³Auburn University, ⁴DataTalk AI,
⁵Southwest Research Institute, ⁶Dublin Institute for Advanced Studies





Innovative AI solutions
of interest to NASA
(& everyone else!)

The Solar Dynamics Observatory



The Atmospheric Imaging Assembly (AIA)

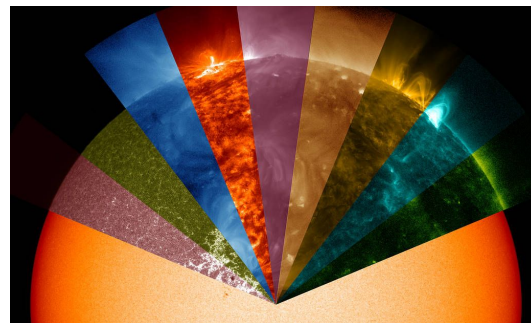
4096 x 4096 full-sun images in 10 channels

The Helioseismic and Magnetic Imager (HMI)

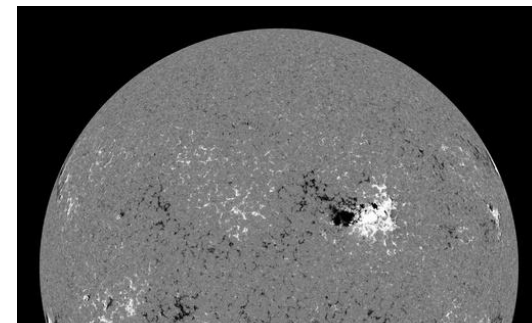
Effectively image sun's magnetic activity

The Extreme ultraviolet Variability Experiment (EVE)

Measures EUV irradiance in select ion ranges

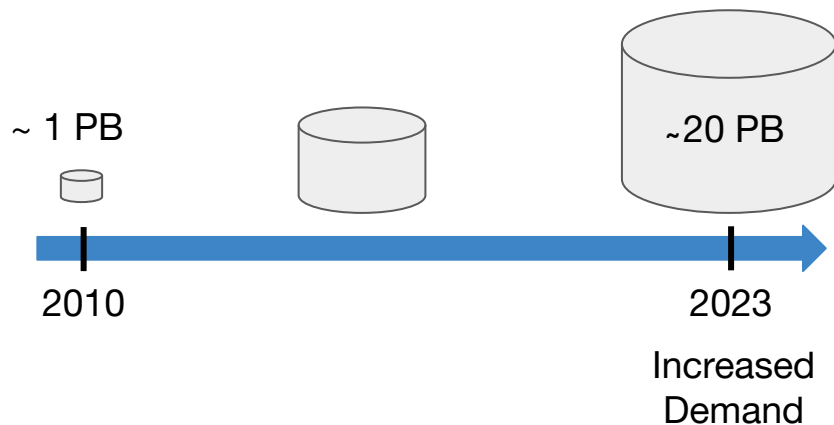


AIA

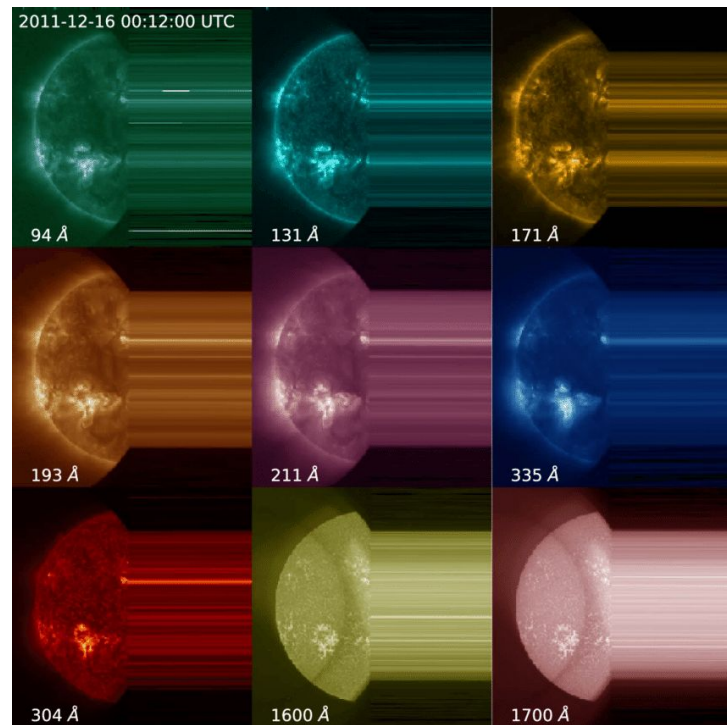


HMI

SDO Data Infrastructure Issues



- Data infrastructure designed decades ago, SDO now has massive data volumes (~20 PB).
- Scientists need access to curated data and the compute resources to perform large-scale analysis.
- Clear need for an **automated** pipeline to prepare and serve data.



**Can we make SDO data
more accessible for
everyone?**

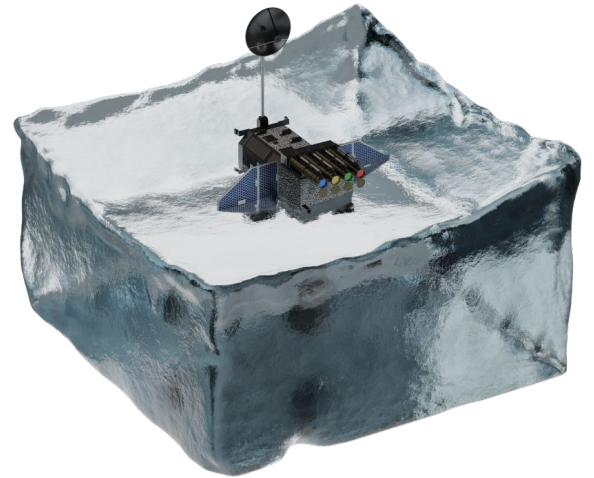
It can be difficult to get SDO data

- Egress limitations from JSOC
- Many different types of data product to choose from
- Data types at different calibrations / cadences / resolutions
- No compute resources available

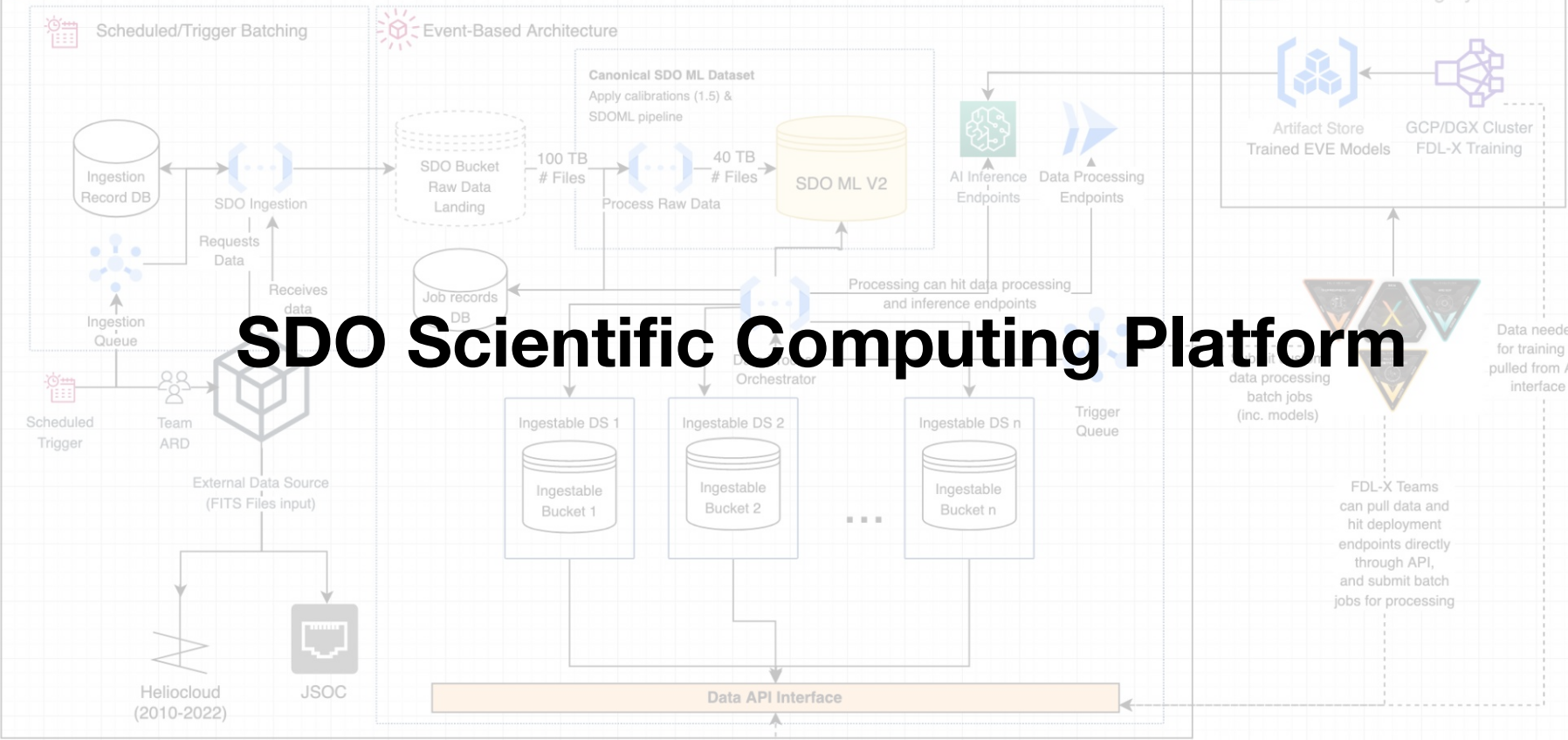


The SDO Machine Learning dataset (SDOML)

- The SDOML dataset was introduced in FDL2018 and has had subsequent improvements over the years
- Some key points are:
 - AIA & HMI images, and ion irradiance from EVE
 - Images corrected for instrument degradation
 - Calibration to “level 1.5”
 - Solar disk position & size harmonized across images
 - 512x512 resolution (easy for ML)
- The dataset is a curated, **machine-learning ready** dataset
- SDOMLv2 **only had SDO data up to 2020**
- **Addition of new data not automated**

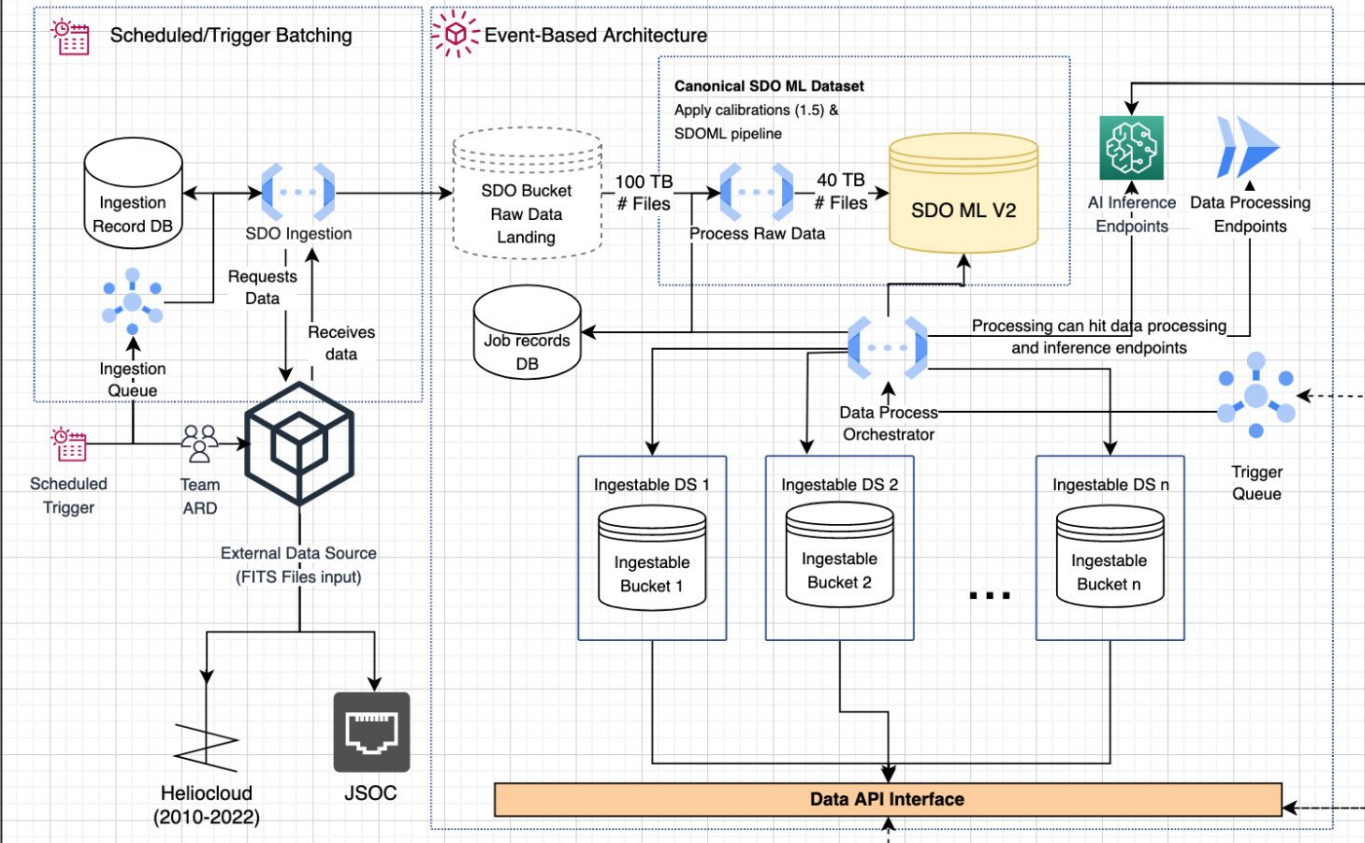


SDO Scientific Computing Platform

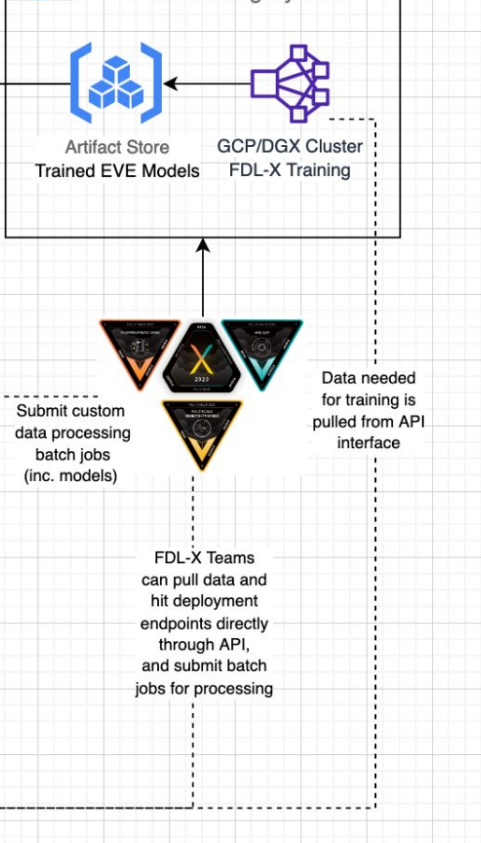




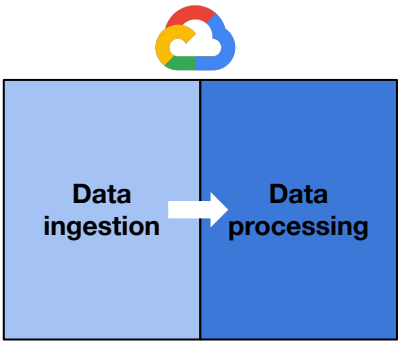
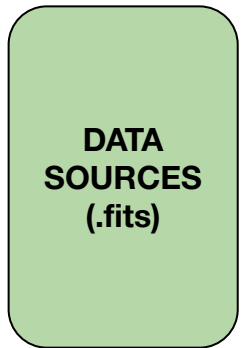
GCP SDO Data Cloud Architecture



GCP / NVIDIA FDL-X Training System



Data from new sources must be processed and aligned with existing SDOMLv2

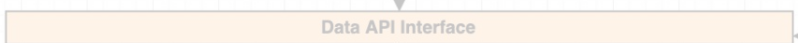
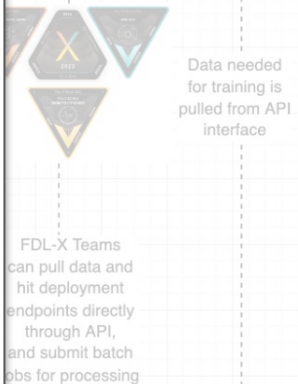
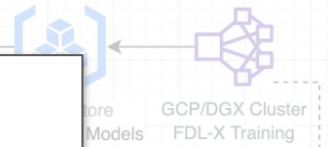
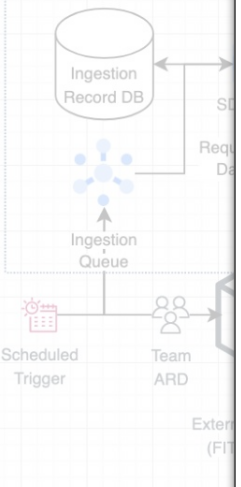


512x512 resolution
Level 1.5
6 minute cadence (AIA)
12 minute cadence (HMI)

Scheduled/Trigger Batching

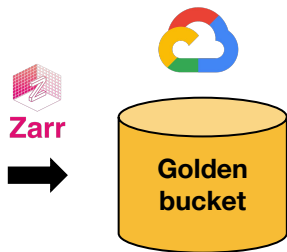
Event-Based Architecture

Canonical SDO ML Dataset



Pipeline in Action: data processing

Data processing



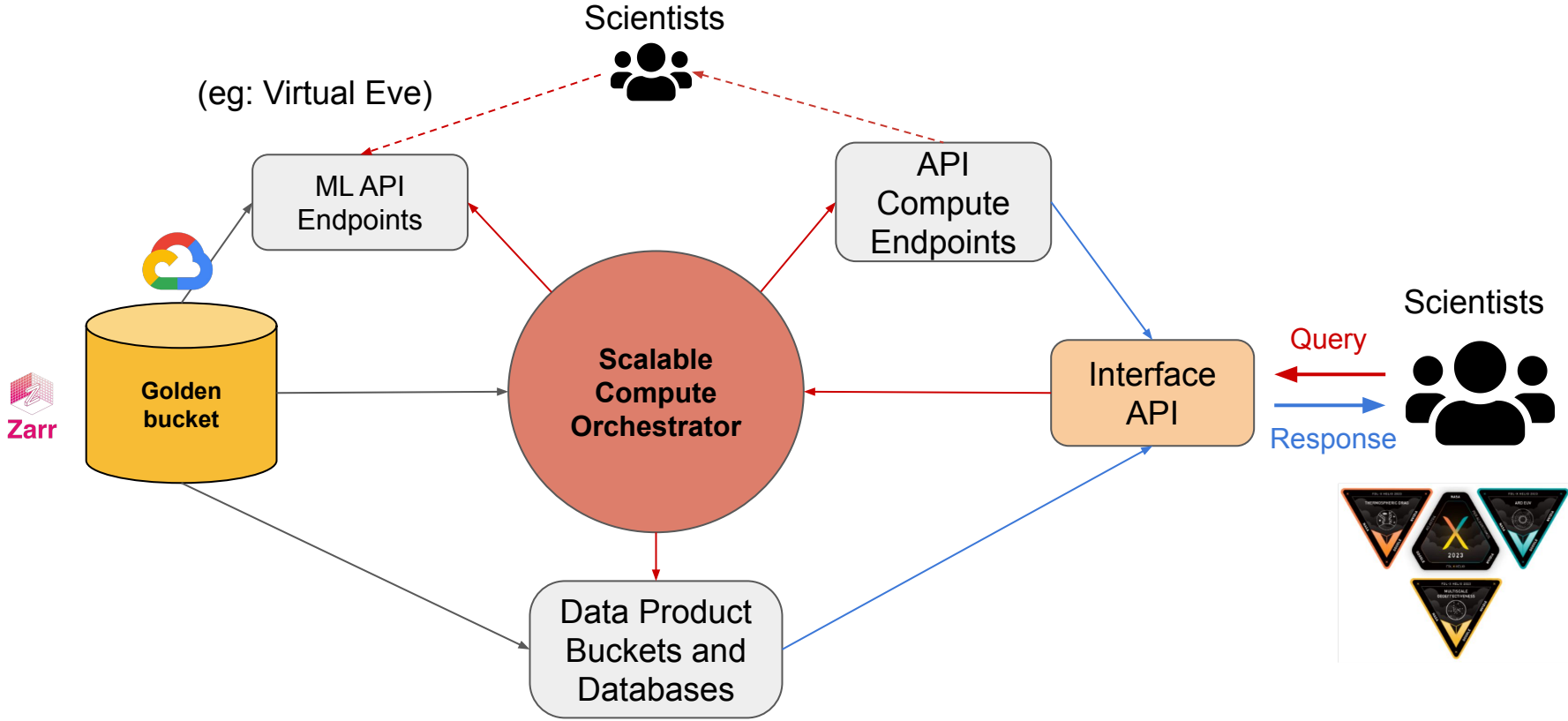
512x512 resolution
Level 1.5
6 minute cadence (AIA)
12 minute cadence (HMI)

1 year of HMI data
10 minutes to process



6000 concurrent machines

Serverless Computational Platform





Welcome to the FDL-X SDO Live Virtual EVE Demo. This app is designed to showcase the capabilities of the FDL-X SDO computational platform.

Please Select Date Range

Start Date: 2010/02/01

Begin Time: 00:00

End Date: 2020/12/31

End Time: 23:59

Wavelength: 131A, 1600A, 1700A, 171A, 193A, 304A, 211A, 335A, 94A

Statistics: mean

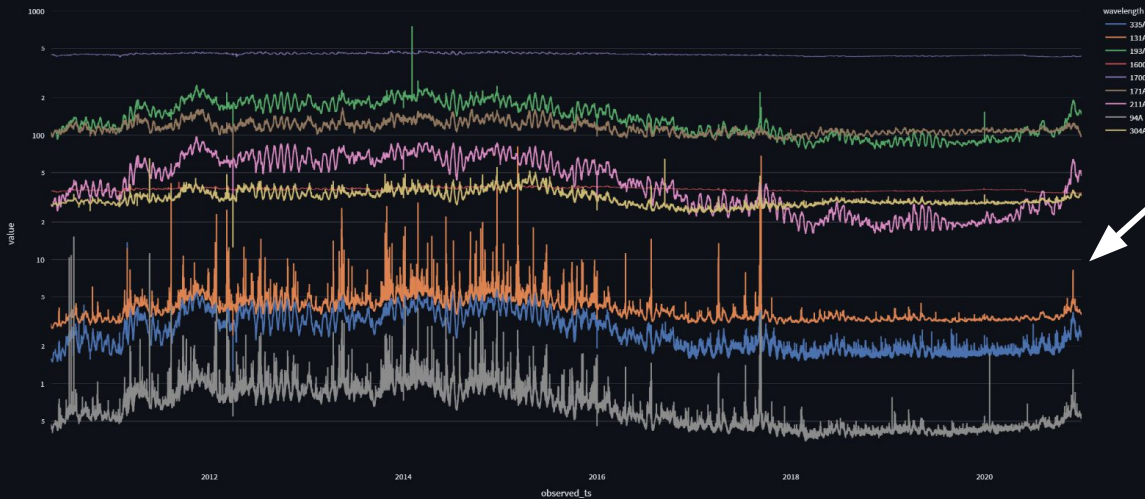
Log Y Axis

Analyze Data

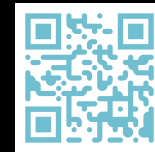
FDL-X SDO Computational Platform Demo

Select Page

AIA Statistics



sdomldemo.org/

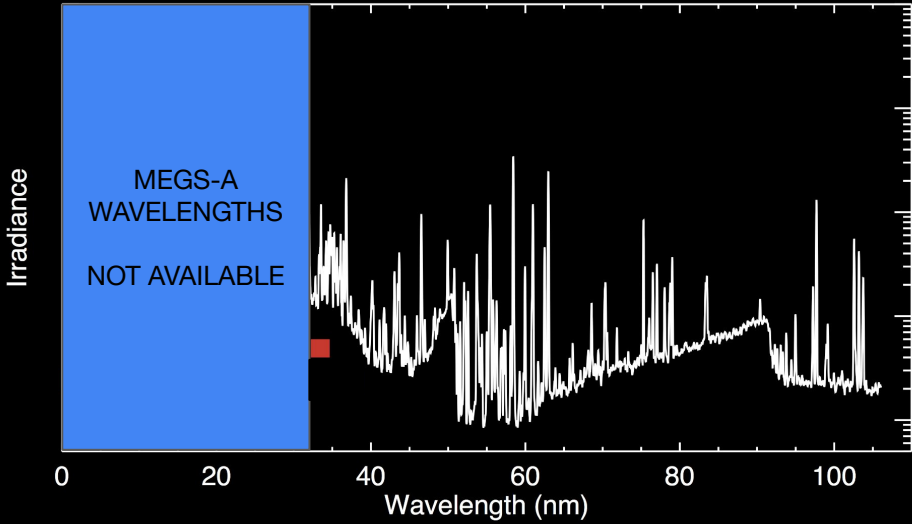


AIA summary statistics



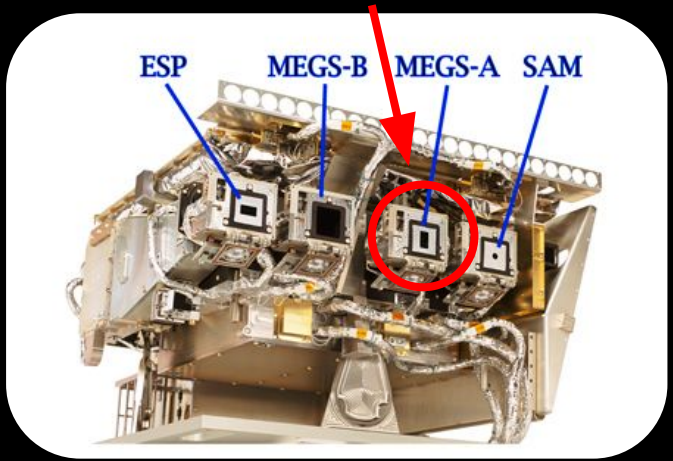
Case Study: EVE Instrument Virtualization

EVE: How it's going

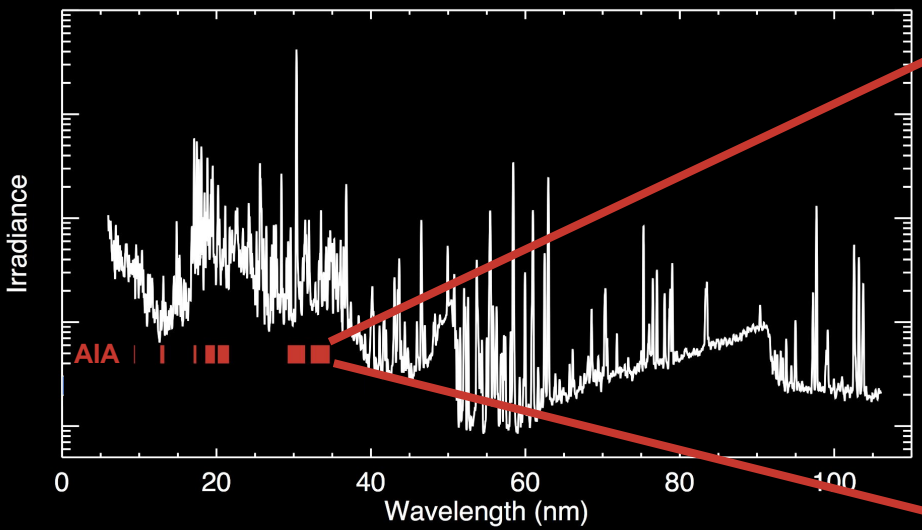


EVE Instrument

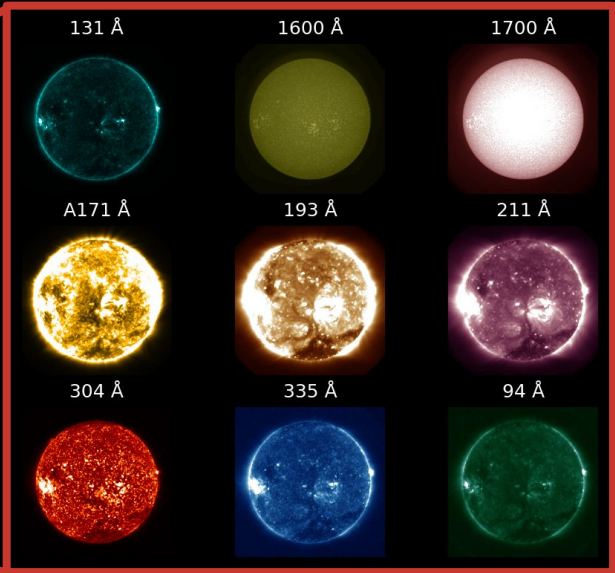
Partly broken since 2014
(Capacitor short in MEGS-A)



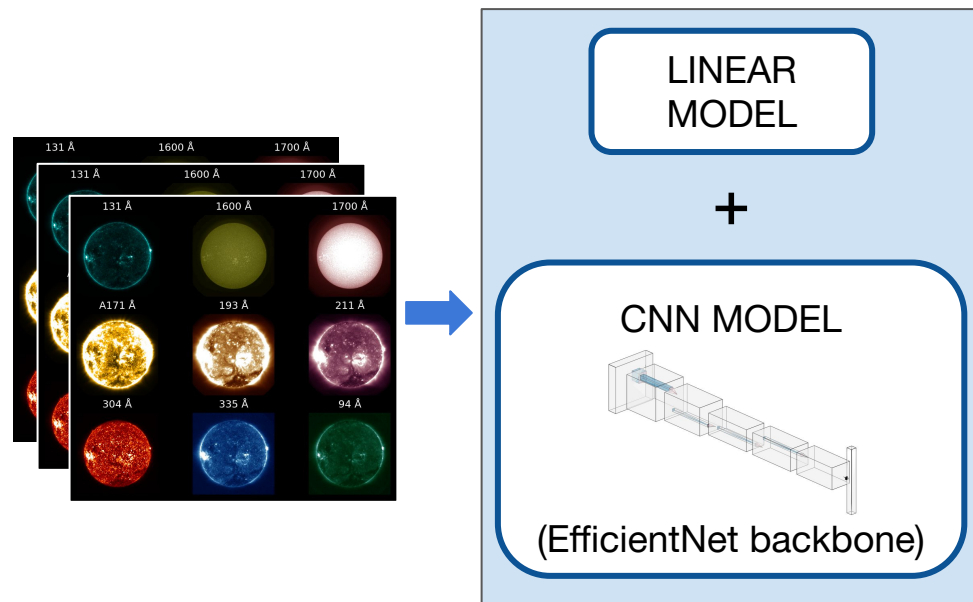
EVE: AIA to the Rescue



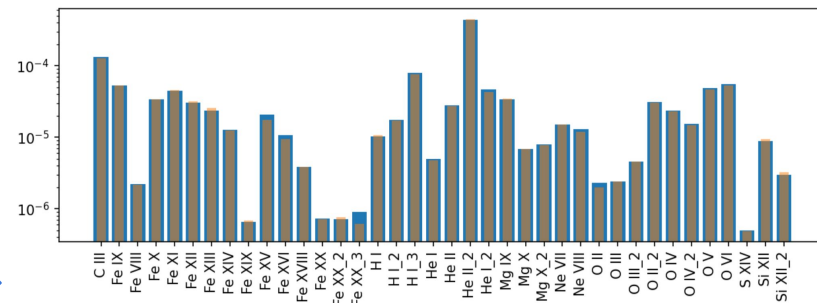
Our dataset



Virtual EVE: Hybrid Model



Irradiance Prediction



	Previous work (FDL 2018-2022)	This work (FDL 2023)
# ions	14	<u>38</u>
# wavelength	8	<u>9</u>
prediction	historical	<u>live inference</u>



Welcome to the FDL-X SDO Live Virtual EVE Demo. This app is designed to showcase the capabilities of the FDL-X SDO computational platform.

Please Select Date Range

Start Date

2017/09/26

Begin Time:

00:00

End Date

2017/09/27

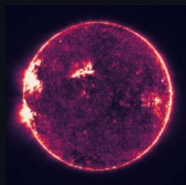
End Time:

23:59

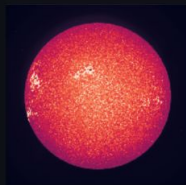
Virtual Eve

AIA Data

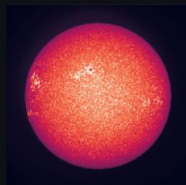
131A



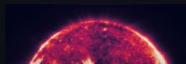
1600A



1700A



171A



193A



211A



Virtual EVE Irradiance





FDL-X SDO Computational Platform Demo

Welcome to the FDL-X SDO Live Virtual EVE Demo. This app is designed to showcase the capabilities of the FDL-X SDO computational platform.

Please Select Date Range

Start Date:

2010/02/01

Begin Time:

00:00

End Date:

2020/01/31

End Time:

23:59

Wavelength:

131A x

1700A x

304A x

94A x

Statistics:

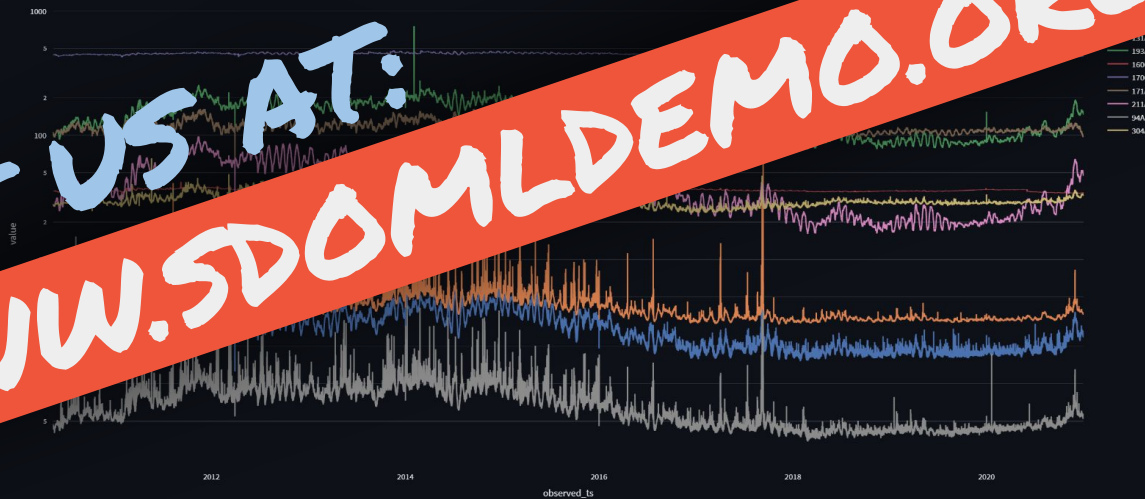
mean x

Log Y Axis

Analyze Data

Select Page

AIA Statistics



Results

SDO Computational Platform

A cloud-based large-scale scalable data ingestion, processing, and ML platform.

SDO ML V2

Analysis ready data for *all* SDO AIA, HMI, and EVE data

Virtual EVE

A deep learning model based on AIA input data providing live proxy EVE solar irradiance measurements to the community; a blueprint for future virtual instruments like it.

Impact

We've built the computational data analysis cloud-based platform that we believe SDO has always needed.

We produced a standardized and easy-to-access suite of SDO data ready for science and machine learning that live-updates.

We've built the necessary tools that bridge the divide between data and compute which will help further unlock the scientific value of the SDO mission (and others).

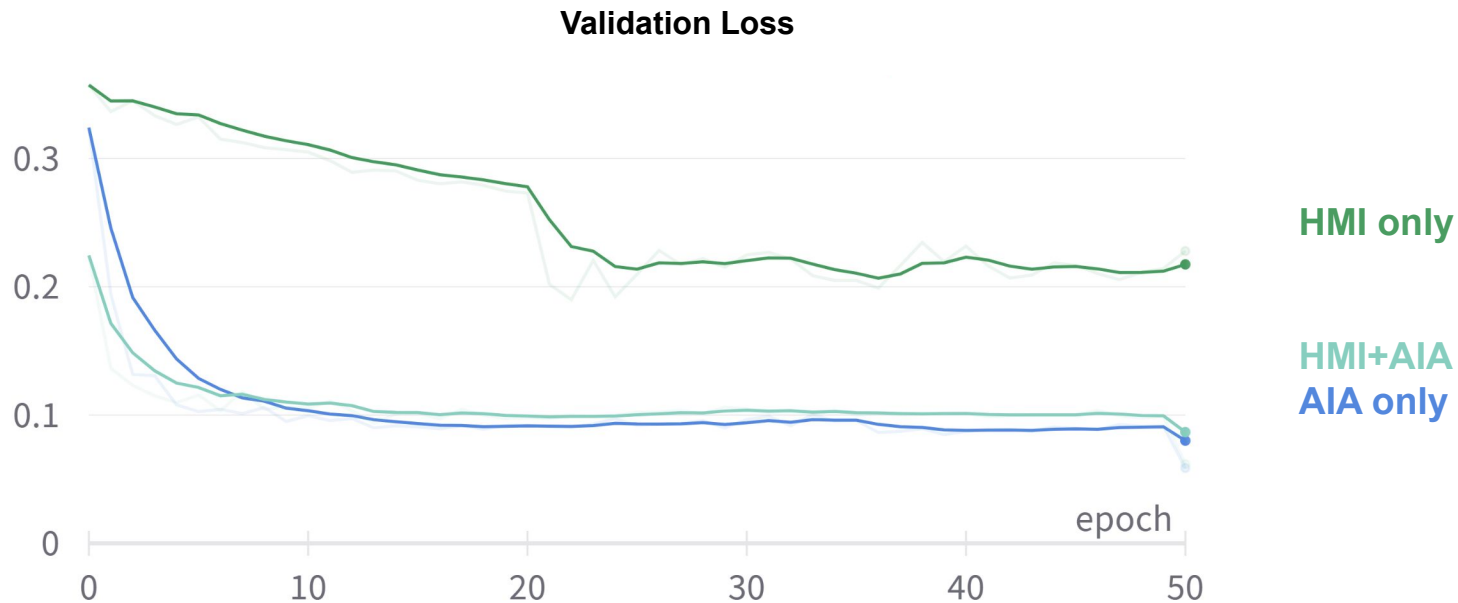


Thank you

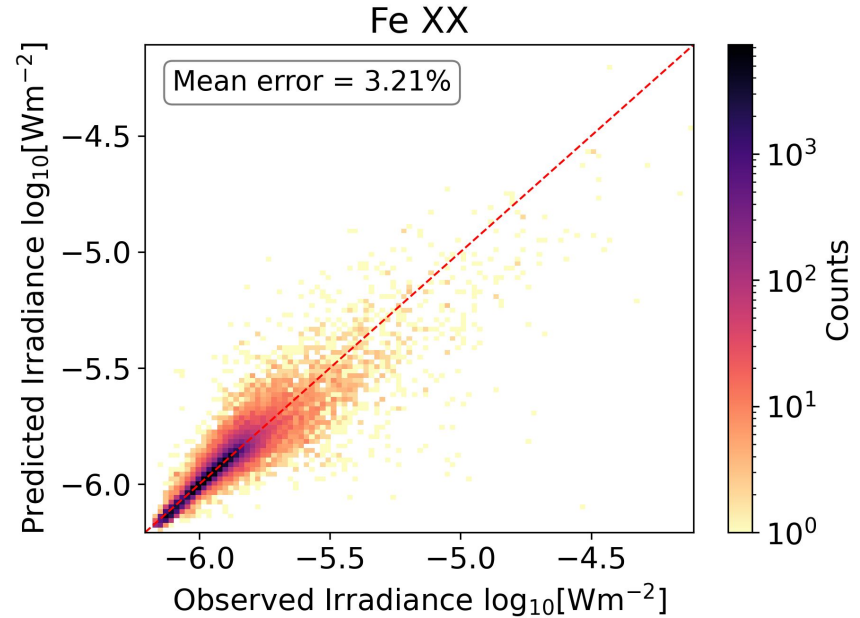
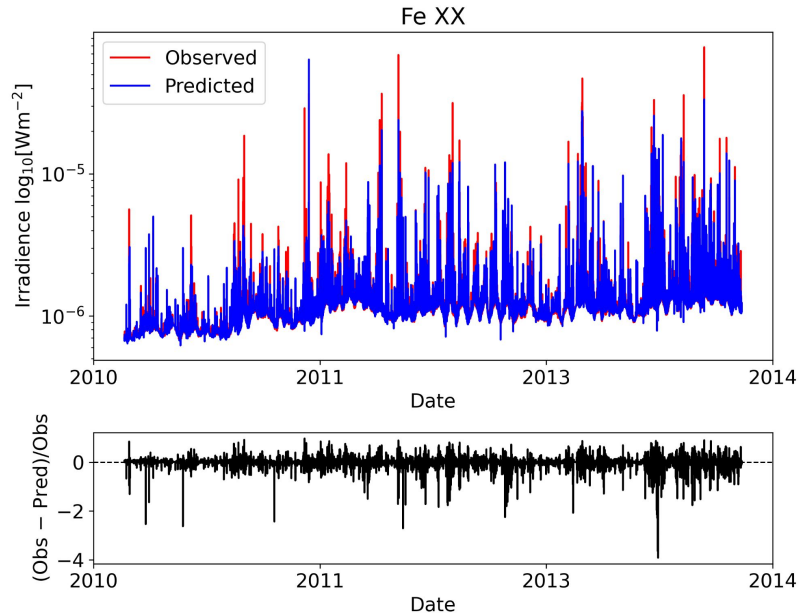


Virtual EVE Results: AIA is All You Need

HMI line-of-sight data **do not** improve irradiance prediction quality



Virtual EVE Results: Irradiance predictions



Virtual EVE Results: Irradiance predictions 2020

- Results show irradiance prediction vs observation (MEGS-B), 6 years after training data ends
- Shape and trend looking good, but there's a systematic offset
- Degradation corrections have been applied, but perhaps the degradation correction is slightly off?

