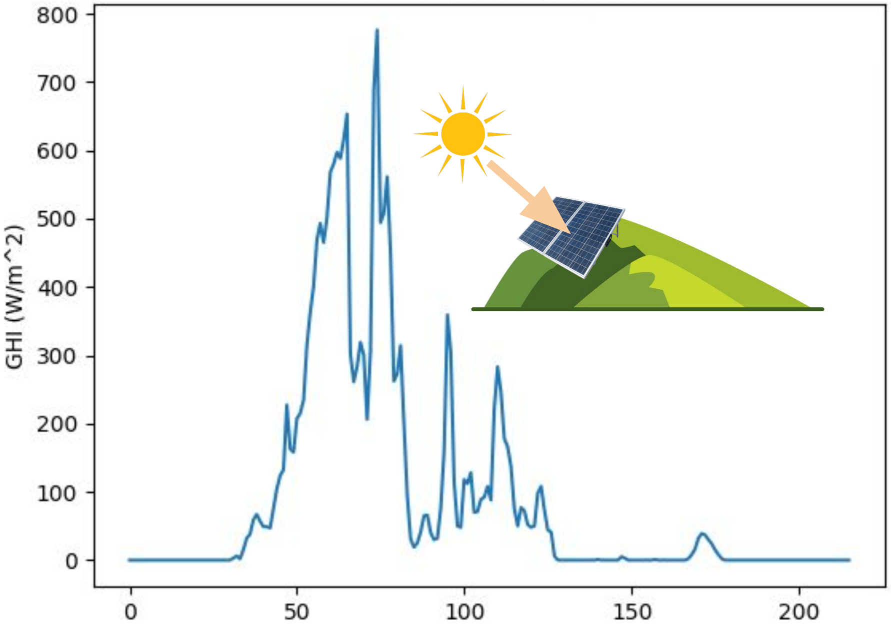
Short-Term Solar Irradiance Forecasting Using Calibrated Probabilistic Models

Eric Zelikman*, Sharon Zhou*, Jeremy Irvin*
Cooper Raterink, Hao Sheng, Anand Avati, Dr. Jack Kelly
Professor Ram Rajagopal, Professor Andrew Y. Ng†, Dr. David John Gagne†

Solar Energy

- Adopting solar in the electricity sector is essential to reducing GHG emissions[1]
- **Solar is highly volatile and intermittent**, so forecasting models are necessary for power system cost-effectiveness and security[2]
- Most are not probabilistic, but characterizing uncertainty can aid real-time grid integration of solar energy and help gauge when to deploy new storage[3,4]

A day of solar in Austin, TX



5 min intervals throughout the day

[1]A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering 2016*.
[2]Review of photovoltaic power forecasting. *Solar Energy 2016*.
[3]The use of probabilistic forecasts: Applying them in theory and practice. *IEEE Power and Energy Magazine 2019*.
[4]Energy storage sizing in presence of uncertainty. *PESGM 2019*

Solar Energy

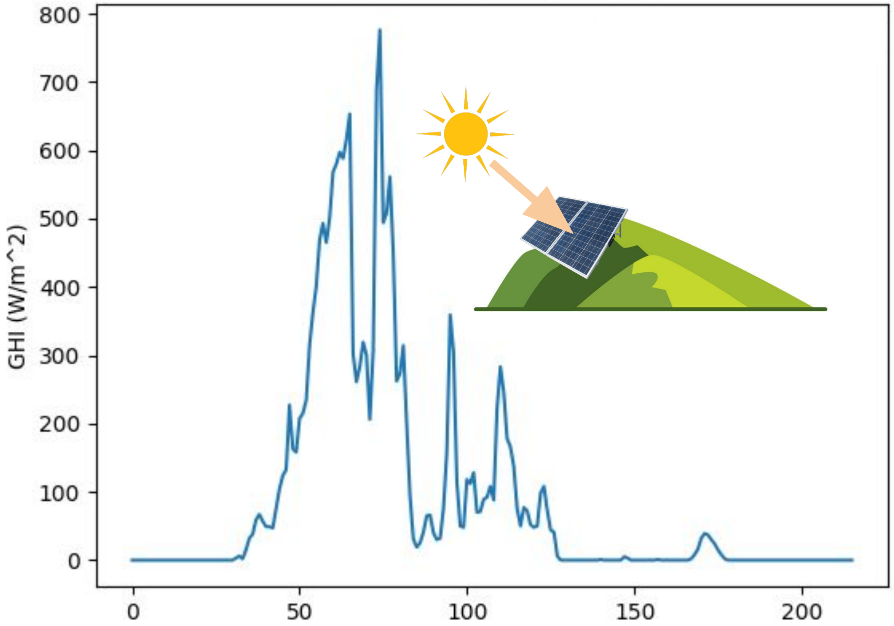
- Adopting solar in the electricity sector is essential to reducing GHG emissions[1]

[1]A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering 2016.*

Solar Energy

- Adopting solar in the electricity sector is essential to reducing GHG emissions[1]
- **Solar is highly volatile and intermittent**, so forecasting models have become necessary[2]

A day of solar in Austin, TX



5 min intervals throughout the day

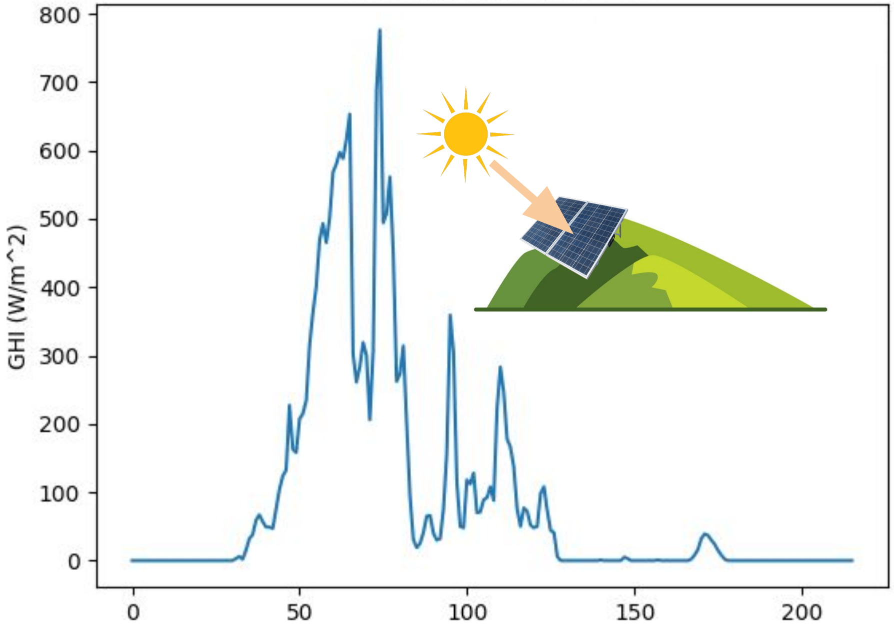
[1]A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering 2016.*

[2]Review of photovoltaic power forecasting. *Solar Energy 2016*.

Solar Energy

- Adopting solar in the electricity sector is essential to reducing GHG emissions[1]
- **Solar is highly volatile and intermittent**, so forecasting models have become necessary[2]
- Most are not probabilistic, but characterizing uncertainty is very useful[3,4]

A day of solar in Austin, TX



5 min intervals throughout the day

[1]A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering 2016.*

[2]Review of photovoltaic power forecasting. *Solar Energy 2016*.

[3]The use of probabilistic forecasts: Applying them in theory and practice. *IEEE Power and Energy Magazine 2019*.

[4]Energy storage sizing in presence of uncertainty. *PESGM 2019*

Probabilistic Solar Forecasting: Current Problems

- Numerical weather prediction (NWP) models
  - Cannot be used on short timescales
  - Computational inefficiency
- ML models
  - Generally rely on traditional models
  - Perform substantially worse than NWP where comparable
- Probabilistic smart persistence
  - Can be defined in several ways
  - Some remarkably good baselines
  - Consistently worse than NWP and machine learning

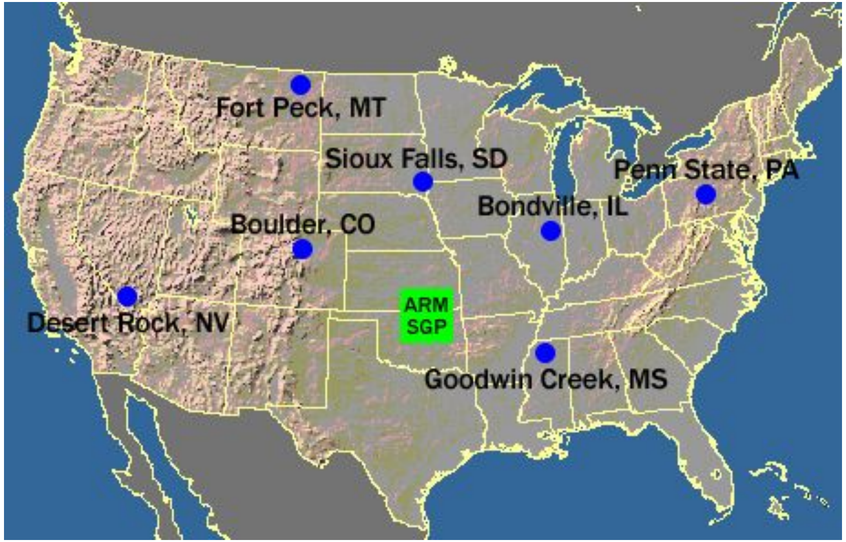
Solar forecasting methods for renewable energy integration. *Progress in Energy and Combustion Science 2013.*

Modern probabilistic ML can substantially improve solar forecasting

Methods

Data: SURFRAD Network

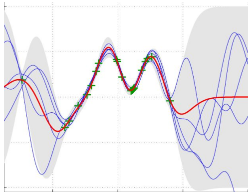
- NOAA's Surface Radiation (SURFRAD) Network[5]
- Seven stations throughout U.S.
- Measure *solar irradiance* (GHI) at 5min resolution
- Meteorological inputs



[5]SURFRAD (Surface Radiation Budget) Network. *Global Monitoring Laboratory.*

Probabilistic Models

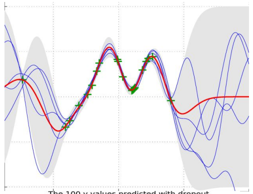
Probabilistic Models

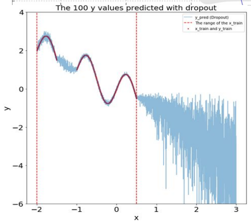


- Gaussian Process[6]

[6]Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances. *IEEE Signal Processing Magazine 2013.*

Probabilistic Models





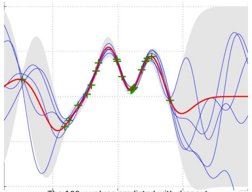
- Gaussian Process[6]

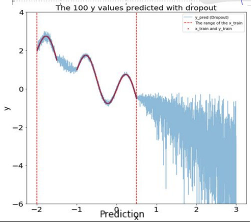
- Dropout Neural Network[7]

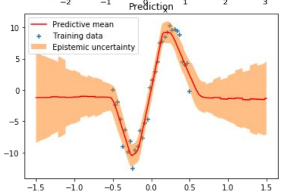
[6] Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances. *IEEE Signal Processing Magazine 2013.*

[7] Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *ICML 2016*.

Probabilistic Models



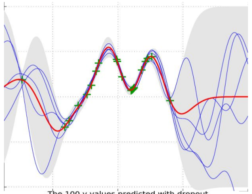


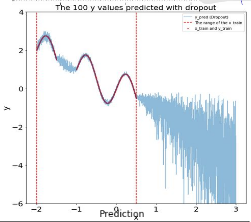


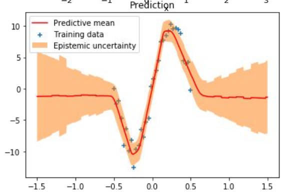
- Gaussian Process[6]
- Dropout Neural Network[7]
- Variational Neural Network[8]

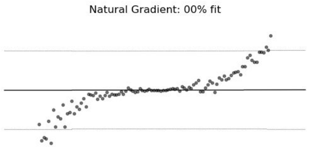
[6]Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances. *IEEE Signal Processing Magazine 2013.*
[7]Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *ICML 2016*.
[8]What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? *NeurIPS 2017*.

Probabilistic Models





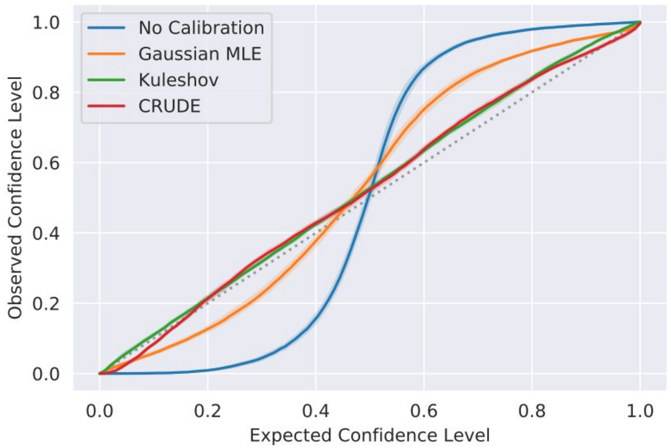




- Gaussian Process[6]
- Dropout Neural Network[7]
- Variational Neural Network[8]
- NGBoost[9]

[6] Gaussian Processes for Nonlinear Signal Processing: An Overview of Recent Advances. *IEEE Signal Processing Magazine 2013.*
[7] Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *ICML 2016*.
[8] What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? *NeurIPS 2017*.
[9] NGBoost: Natural Gradient Boosting for Probabilistic Prediction. *ICML 2020*.

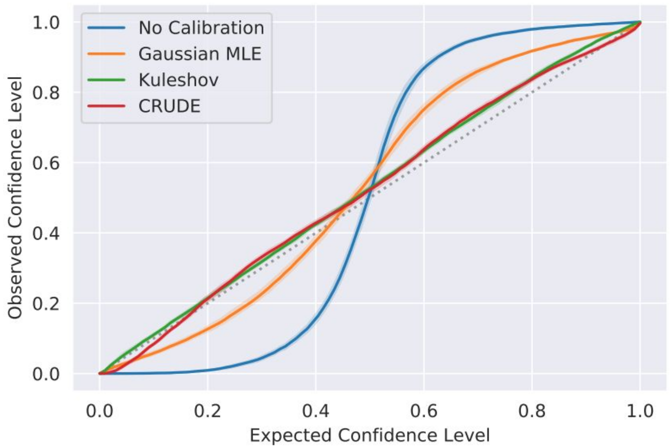
Sharpness Subject to Calibration



Calibration curve for a Gaussian process regression model forecasting in Penn State, PA

- What defines a good probabilistic forecast?

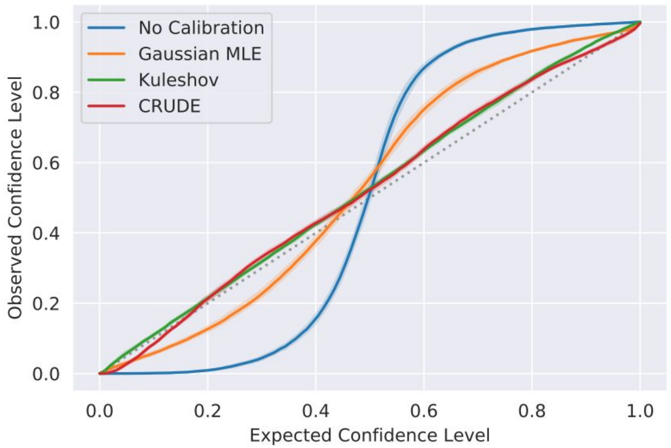
Sharpness Subject to Calibration



Calibration curve for a Gaussian process regression model forecasting in Penn State, PA

- What defines a good probabilistic forecast?
- *Calibration*
  - Are the probabilistic forecasts consistent with the observations?
  - Measures whether predicted distributions correctly capture confidence levels.

Sharpness Subject to Calibration



Calibration curve for a Gaussian process regression model forecasting in Penn State, PA

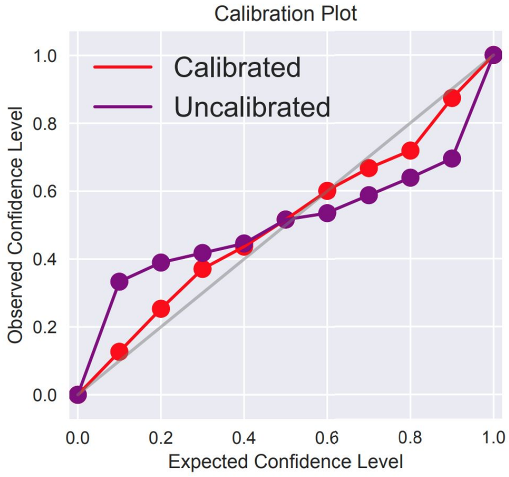
- What defines a good probabilistic forecast?
- *Calibration*
  - Are the probabilistic forecasts consistent with the observations?
  - Measures whether predicted distributions correctly capture confidence levels.
- *Sharpness*
  - Is the probability distribution tight?
  - Sharper models are better, subject to calibration.

Post-hoc Calibration Methods

- Models are usually not well-calibrated by default
  - They’re often overconfident on unseen data
- Post-hoc calibration methods:
  - Gaussian MLE

Post-hoc Calibration Methods

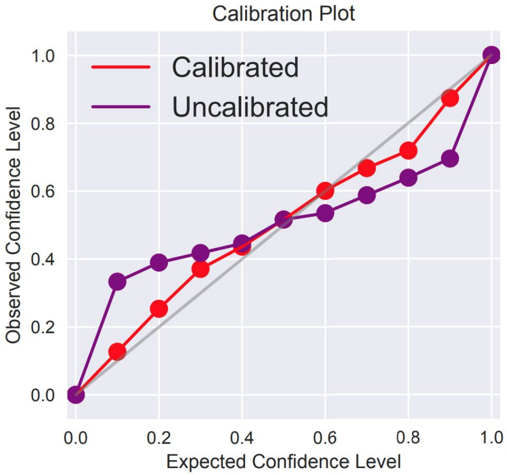
- Models are usually not well-calibrated by default
  - They’re often overconfident on unseen data
- Post-hoc calibration methods:
  - Gaussian MLE
  - Kuleshov: invert the calibration curve[10]

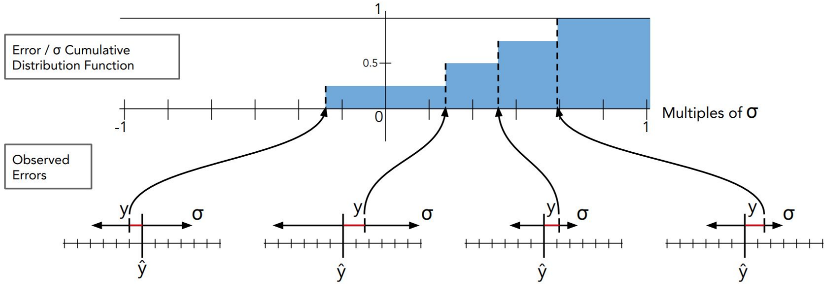


[10] Accurate Uncertainties for Deep Learning Using Calibrated Regression. *ICML 2018.*

Post-hoc Calibration Methods

- Models are usually not well-calibrated by default
  - They're often overconfident on unseen data
- Post-hoc calibration methods:
  - Gaussian MLE
  - Kuleshov: invert the calibration curve[10]
  - CRUDE: measure z-scores of observed errors[11]





[10]Accurate Uncertainties for Deep Learning Using Calibrated Regression. *ICML 2018.*

[11]CRUDE: Calibrating Regression Uncertainty Distributions Empirically. *ICML 2020 Workshop on Uncertainty & Robustness in Deep Learning.*

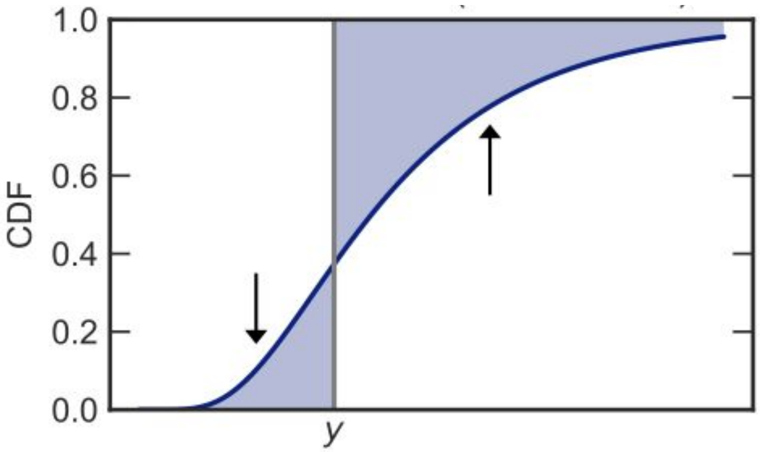
Performance Metric: CRPS

- Is there a metric which captures both calibration and sharpness?
- *Continuous Ranked Probability Score (CRPS)*
    - Area between the predicted CDF and a step function at the observed value

Countdown Regression: Sharp and Calibrated Survival Predictions. *UAI 2019.*

Performance Metric: CRPS

- Is there a metric which captures both calibration and sharpness?
- *Continuous Ranked Probability Score (CRPS)*
    - Area between the predicted CDF and a step function at the observed value



Countdown Regression: Sharp and Calibrated Survival Predictions. *UAI 2019.*

Results

Comparison Between Our Models

Station	Gaussian Process				Dropout Neural Network				Variational Neural Net				NGBoost			
	None	MLE	C	Kul.	*None*	MLE	C	Kul.	*None*	MLE	C	Kul.	*None*	MLE	C	Kul.
Bondville, IL	101.3	53.2	48.5	48.6	48.5	46.0	43.6	44.0	42.0	42.0	41.8	41.9	**40.5**	**40.5**	40.6	40.6
Boulder, CO	110.9	61.7	56.4	56.5	59.3	55.8	53.3	53.9	48.6	48.9	48.3	48.6	**45.9**	46.1	46.0	46.2
Desert Rock, NV	96.6	44.3	35.4	35.7	37.2	40.8	36.1	36.2	31.4	32.5	30.0	30.3	27.9	30.1	**27.8**	28.2
Fort Peck, MT	97.5	50.7	43.6	43.4	41.6	41.9	38.9	39.0	37.9	46.8	37.5	37.6	**34.8**	35.2	35.0	34.9
Goodwin Creek, MS	119.2	59.8	54.7	54.9	57.9	53.3	51.6	51.5	46.9	46.9	46.7	46.9	**44.8**	45.0	**44.8**	45.1
Penn State, PA	111.6	58.8	53.9	53.3	56.5	51.2	49.5	48.0	47.4	47.4	47.3	47.0	**46.0**	46.6	46.1	**46.0**
Sioux Falls, SD	107.2	54.4	49.3	49.5	48.0	46.0	43.4	43.7	43.8	41.8	42.4	43.0	**37.9**	39.1	38.0	38.4

- NGBoost was consistently the best performing model
- Calibration had no substantial impact for short-term forecasting

Comparison To Prior Models

	CH-P	PeEn	MCM	NGB	%Δ
Bondville, IL	92.1	52.8	48.7	**40.5**	-16.8%
Boulder, CO	91.3	61.6	51.6	**45.9**	-11.0%
Desert Rock, NV	47.3	35.2	29.4	**27.9**	-5.1%
Fort Peck, MT	77.0	46.3	39.8	**34.8**	-12.6%
Goodwin Creek, MS	98.4	59.7	52.5	**44.8**	-14.7%
Penn State, PA	98.1	60.0	53.0	**46.0**	-13.2%
Sioux Falls, SD	86.8	47.8	41.0	**37.9**	-7.6%

Intra-hourly Performance

CH-P	GAU	NWP	NGB (+C)
78.1	52.7	**50.8**	53.1 (52.9)
75.7	64.2	64.6	**60.3 (60.4)**
37.7	42.5	39.2	**36.1 (35.8)**
64.8	49.9	48.0	**46.3 (46.2)**
82.3	58.3	**56.4**	56.9 (56.6)
83.4	**55.1**	57.4	58.8 (58.1)
74.3	50.6	**49.7**	58.6 (56.6)

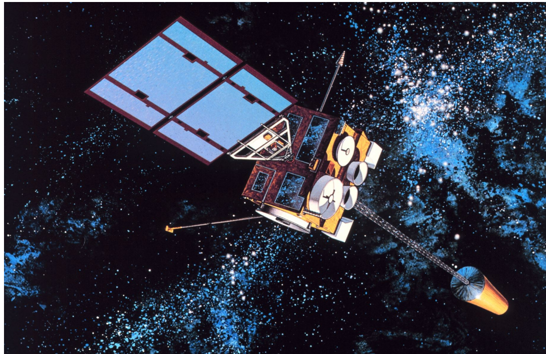
Hourly Resolution Performance

- NGBoost was consistently the best short-term forecasting model
- NGBoost with CRUDE calibration often outperformed NWP models

Visualization

Future Directions

- Incorporate satellite imagery to account for clouds
- An ablation study of various inputs would help
  - Can we predict irradiance accurately with only public data?
- Could the models perform better with better hyperparameters?



Thank you!

Thank you!