SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

HAIMOS Ensemble Forecasts for Intra-day and Day-Ahead GHI, DNI and Ramps

University of California San Diego

PI: Carlos F. M. Coimbra

Co-PI: Hugo Pedro (Presenter)

Project Team: UCSD and Clean Power Research (CPR)

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Mechanical and Aerospace Engineering

COIMBRA RESEARCH GROUP

CER
Center for Energy Research

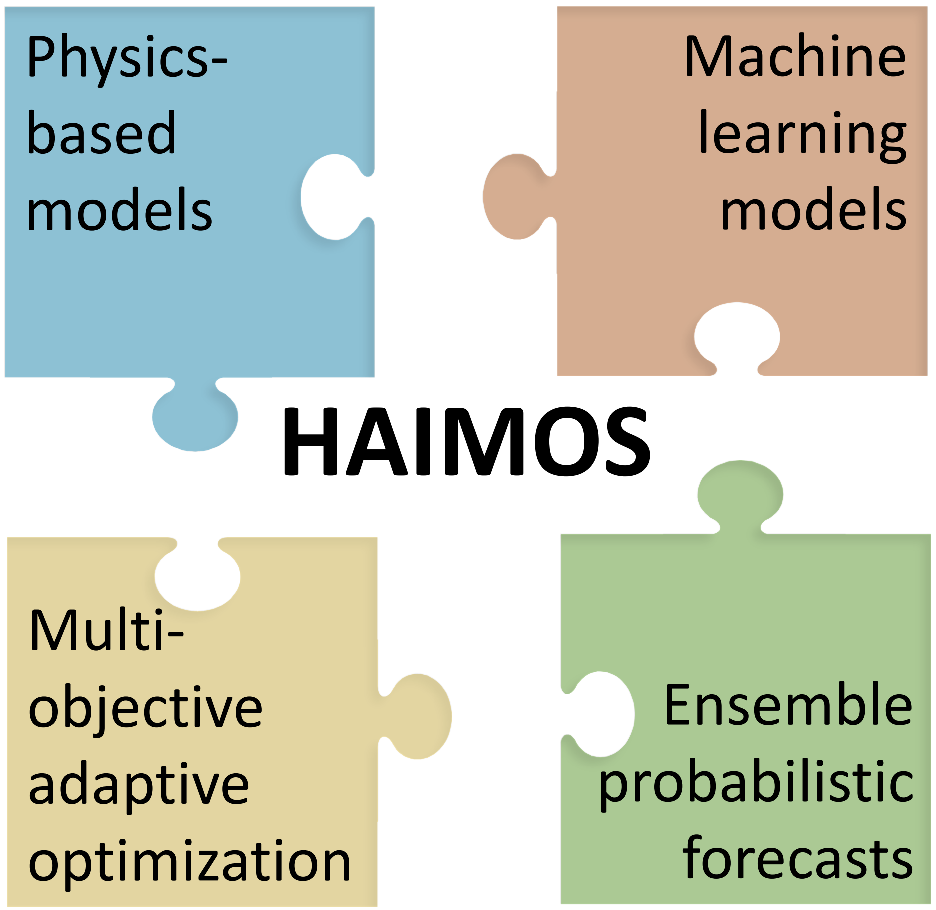
Project summary

Development of a Hybrid Adaptive Input Model Objective Selection (**HAIMOS**) ensemble model to improve solar irradiance (GHI and DNI) and cloud cover forecasts.

- Comprehensive optimization framework:
  - Every aspect of the model (data preprocessing, input selection, etc.) will be subject to adaptive optimization to reduce bulk error metrics, predict ramp onset, etc.
- Ingestion of new-generation cloud cover products
  - High-resolution rapid refresh satellite images, cloud cover modeling and forecasting using Large Eddy Simulation) to improve cloud optical depth forecast and irradiance forecasts

Goal: Increase the state of the art forecast skill from their present values of 10 to 30%. Aim to achieve the 50% forecast skill level **consistently** for both GHI and DNI.

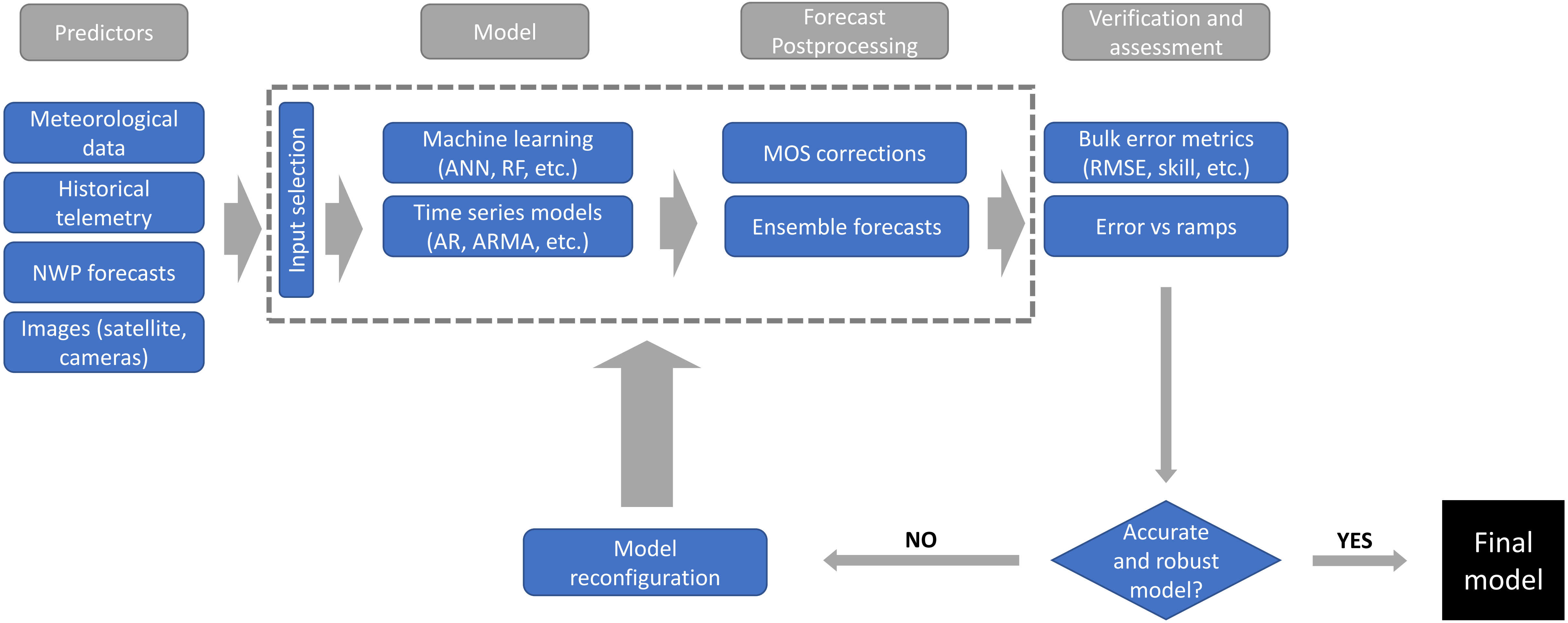
Technical Approach



- Combines innovations in **machine learning** algorithms (deep-learning, feature engineering, etc.) with detailed **physics-based models** for cloud cover and cloud optical depth forecasts.
- Integrates information derived from the **new GOES satellites sensors and cloud cover simulations from LES**.
- Spatial and temporal sensing/modeling of clouds at much higher resolutions than previously available.

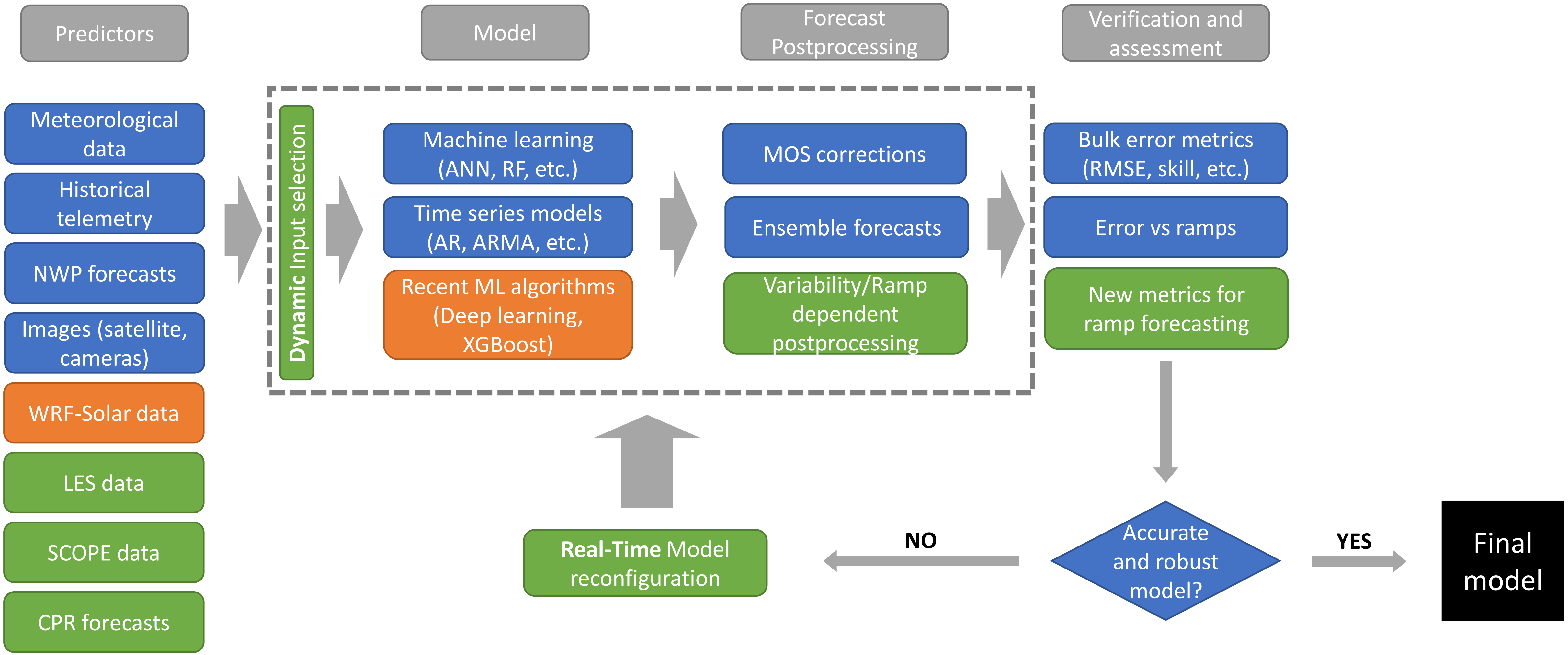
Technical Approach

- Typical approach to irradiance forecasts



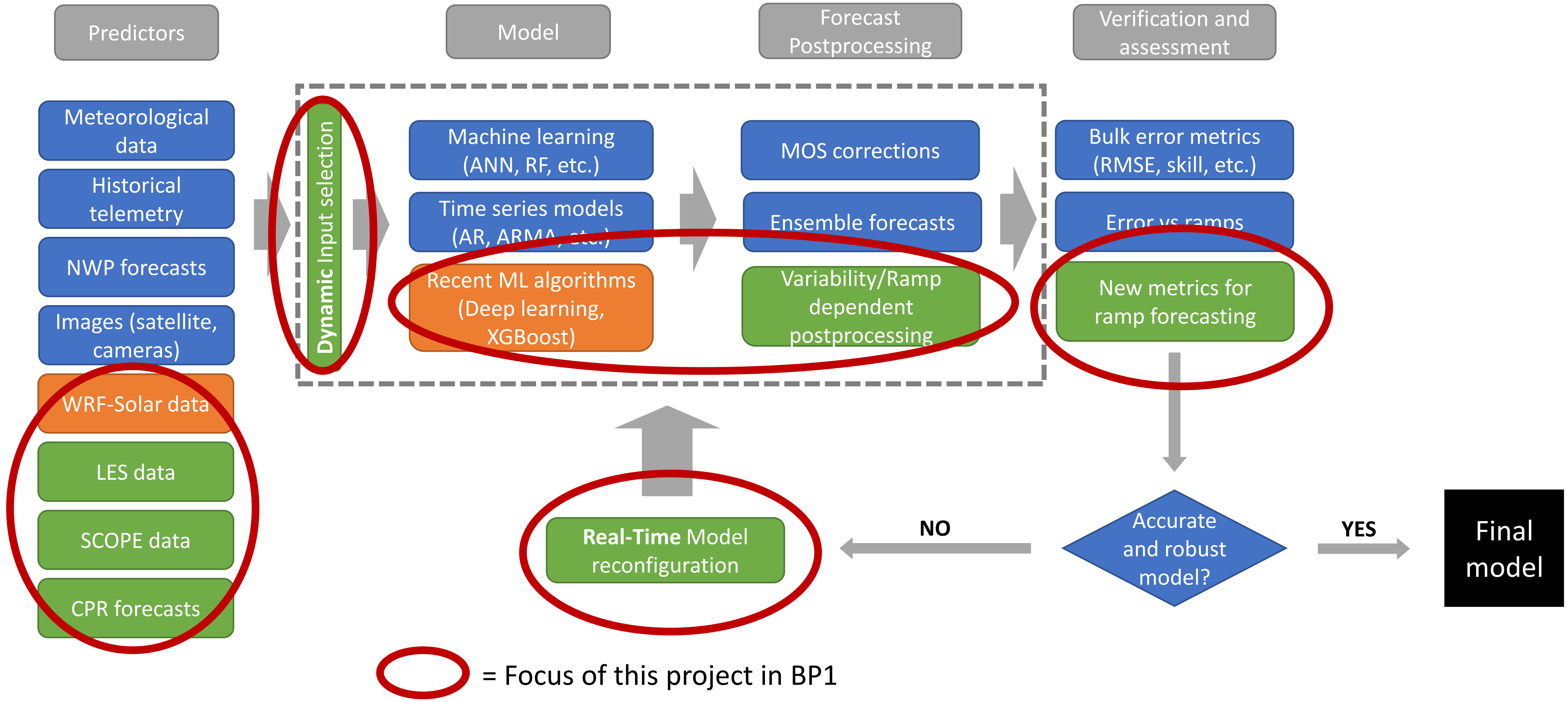
Technical Approach

- HAIMOS approach



Technical Approach

- HAIMOS approach



Task Summary for BP1

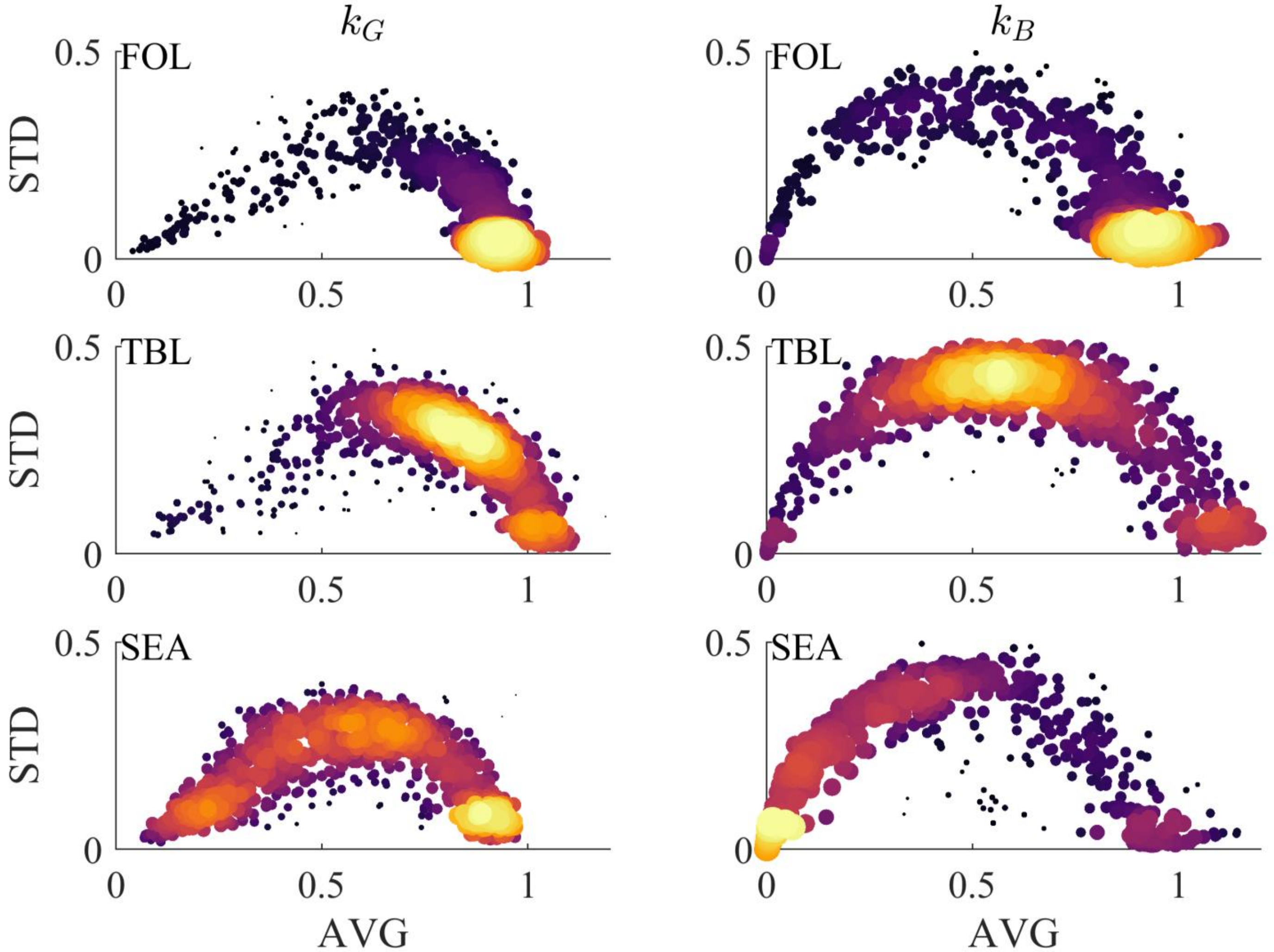
- Task 1: Selection of training sites and data collection
- Task 2: Commercial Irradiance Forecasts
  - Subtask 2.1: Augment Task 1 database with commercial irradiance forecasts
  - Subtask 2.2: Real-time access to commercial irradiance forecasts
- Task 3: Improvement of cloud cover forecast
  - Subtask 3.1: Improvement of cloud identification accuracy
  - Subtask 3.2: Improvement of cloud optical depth estimation
  - Subtask 3.3: WRF-Solar simulations
  - Subtask 3.4: Cloud fraction and optical depth forecast
- Task 4: Set up HAIMOS framework
  - Subtask 4.1: Input selection for HAIMOS
  - Subtask 4.2: Training and optimization of machine-learning models for HAIMOS
  - Subtask 4.3: Developing adaptive protocols for HAIMOS

Task 1 : Selection of training sites and data collection

- Technical approach
  1. Collected data from several candidate locations in the CONUS (SURFRAD, SOLRAD, UCSD)
  2. Data quality control
  3. Data normalization using a clear-sky model
  4. Computed several metrics to characterize data variability
     1. Daily average and standard deviation clear-sky index (AVG$(k_I)_d$, STD$(k_I)_d$)
     2. Ramp density
     3. Persistence model performance
  5. Selected 6 sites with diverse irradiance levels of variability
  6. Collected exogenous data for the selected sites

Task 1 : Selection of training sites and data collection

Daily average and standard deviation clear-sky index $(\mathrm{AVG}(k_I)_d, \mathrm{STD}(k_I)_d)$

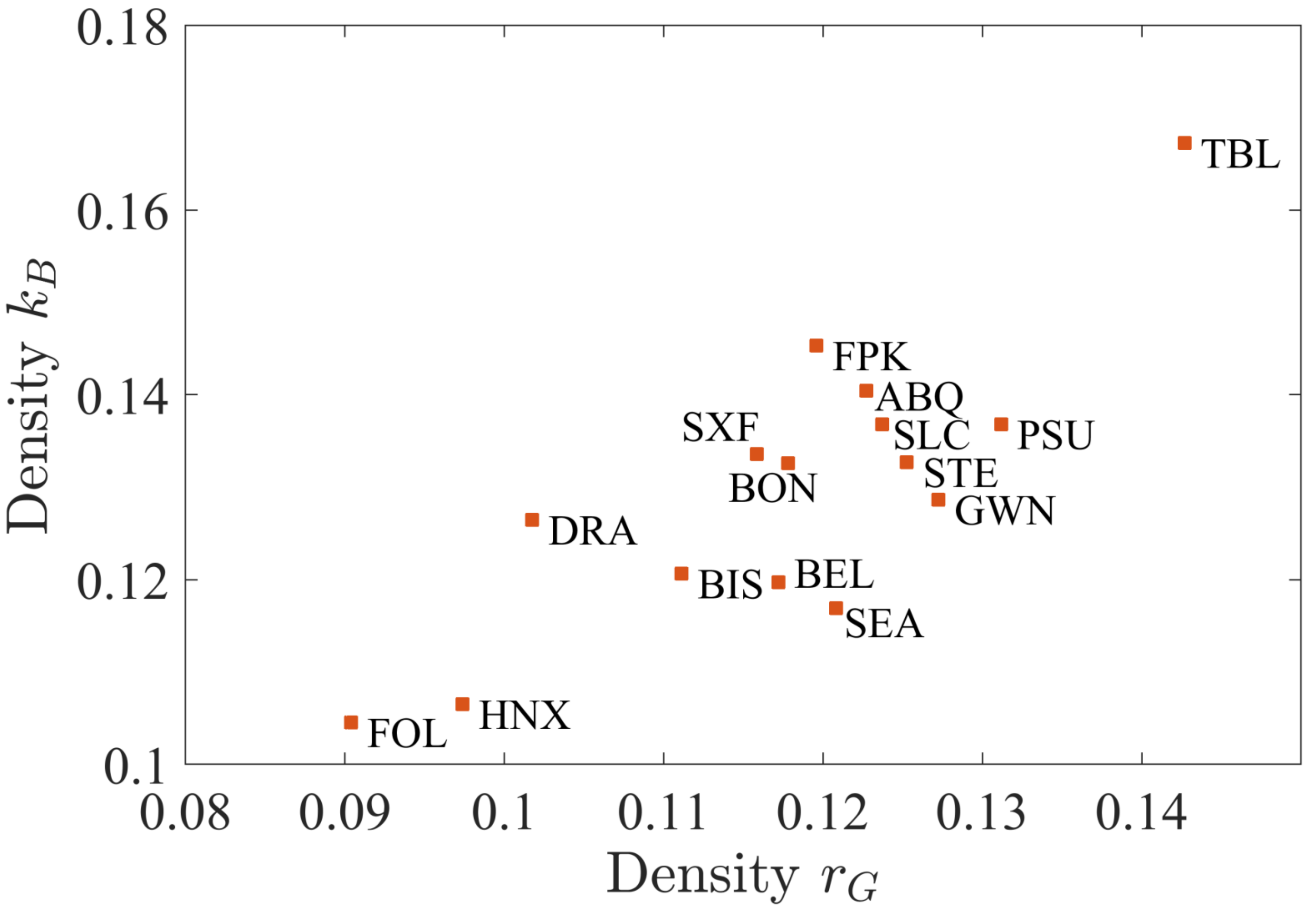


Plots of $(\mathrm{AVG}(k_I)_d, \mathrm{STD}(k_I)_d)$ for three locations in the initial dataset. From top to bottom: Folsom, CA (FOL), Table Mountain, CO (TBL), and Seattle, WA (SEA). The left column shows results for GHI and the right column the results for DNI. Lighter and larger dots indicate large concentrations of $(\mathrm{AVG}(k_I)_d, \mathrm{STD}(k_I)_d)$ pairs.

Task 1 : Selection of training sites and data collection

Ramp density: probability of the occurrence of large ramps

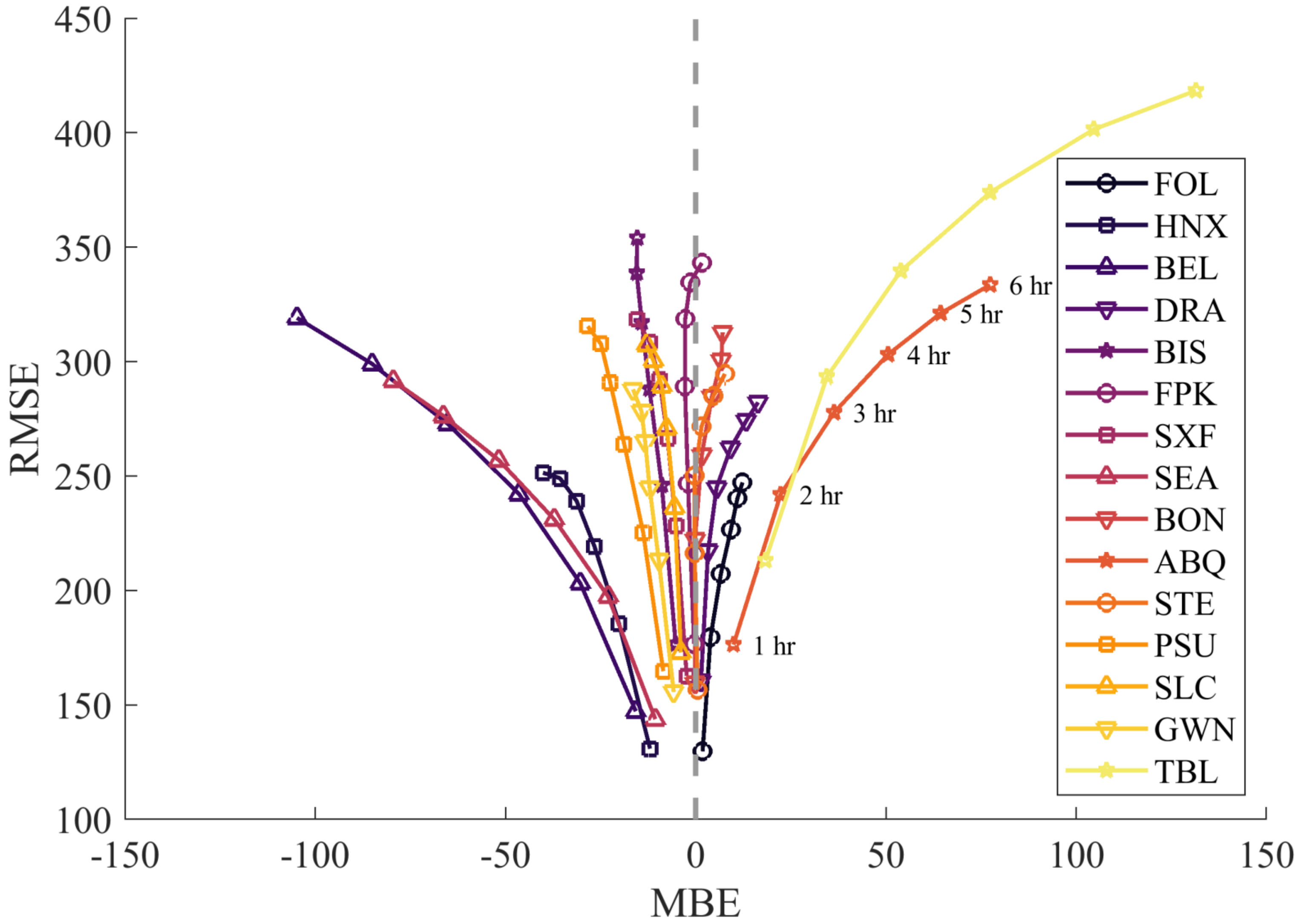
$$r_I(t) = \bar{k}_I(t + 1\text{hr}) - \bar{k}_I(t)$$



DNI ramp density versus GHI ramp density for all candidate locations.

Task 1 : Selection of training sites and data collection

Persistence model performance



RMSE vs MBE for hourly DNI forecasts 1 to 6 hours ahead of time for the 15 candidate sites.
The annotations in the right-most curve indicate the forecasting horizon.

Task 1 : Selection of training sites and data collection

Selected sites

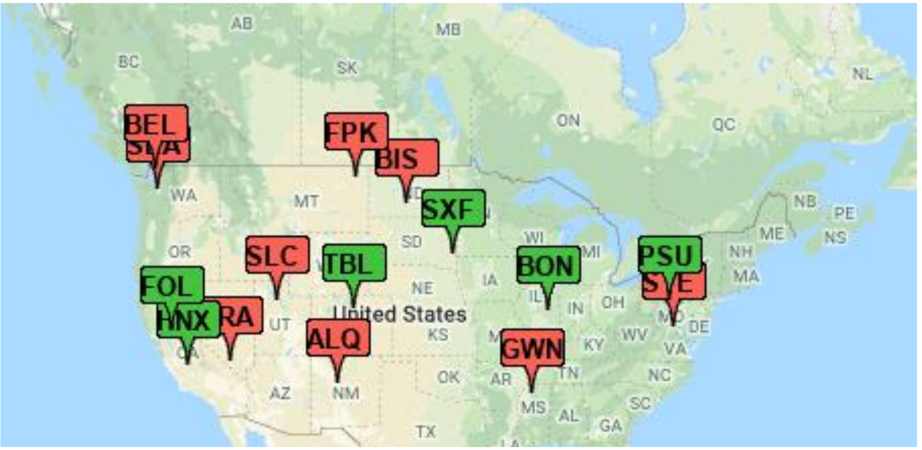
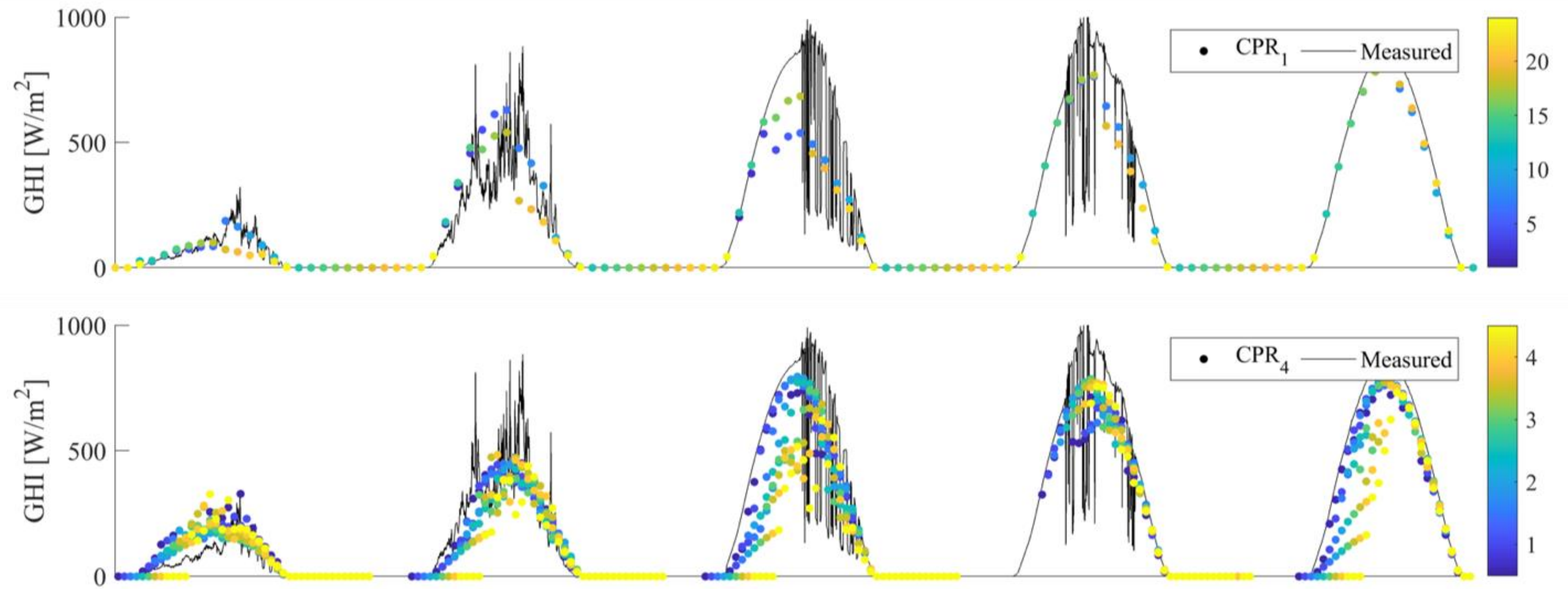


Table 2: List of selected sites and the respective selection metrics. All values are unitless except for MBE_{Δ} and $RMSE_{\Delta}$ which are given in Wm^{-2}.

Variability	Site ID	$AVG(k_I)_d$		$STD(k_I)_d$		ε_I		MBE_{Δ}		$RMSE_{\Delta}$	
		GHI	DNI	GHI	DNI	GHI	DNI	GHI	DNI	GHI	DNI
Large	TBL	0.78	0.60	0.23	0.29	0.14	0.17	[3, 29]	[18, 131]	[123, 227]	[213, 418]
Large	PSU	0.61	0.37	0.21	0.23	0.13	0.14	[-47, -7]	[-28, -9]	[104, 205]	[165, 315]
Medium	BON	0.66	0.46	0.20	0.22	0.12	0.13	[-37, -6]	[-0, 7]	[93, 189]	[160, 313]
Medium	SXF	0.71	0.49	0.18	0.22	0.12	0.13	[-39, -5]	[-15, -2]	[88, 193]	[163, 318]
Low	FOL	0.78	0.65	0.13	0.16	0.09	0.10	[-20, -2]	[2, 12]	[64, 132]	[130, 247]
Low	HNX	0.80	0.66	0.15	0.18	0.10	0.11	[-53, -8]	[-40, -12]	[66, 158]	[131, 251]

Task 2: Commercial Irradiance Forecasts

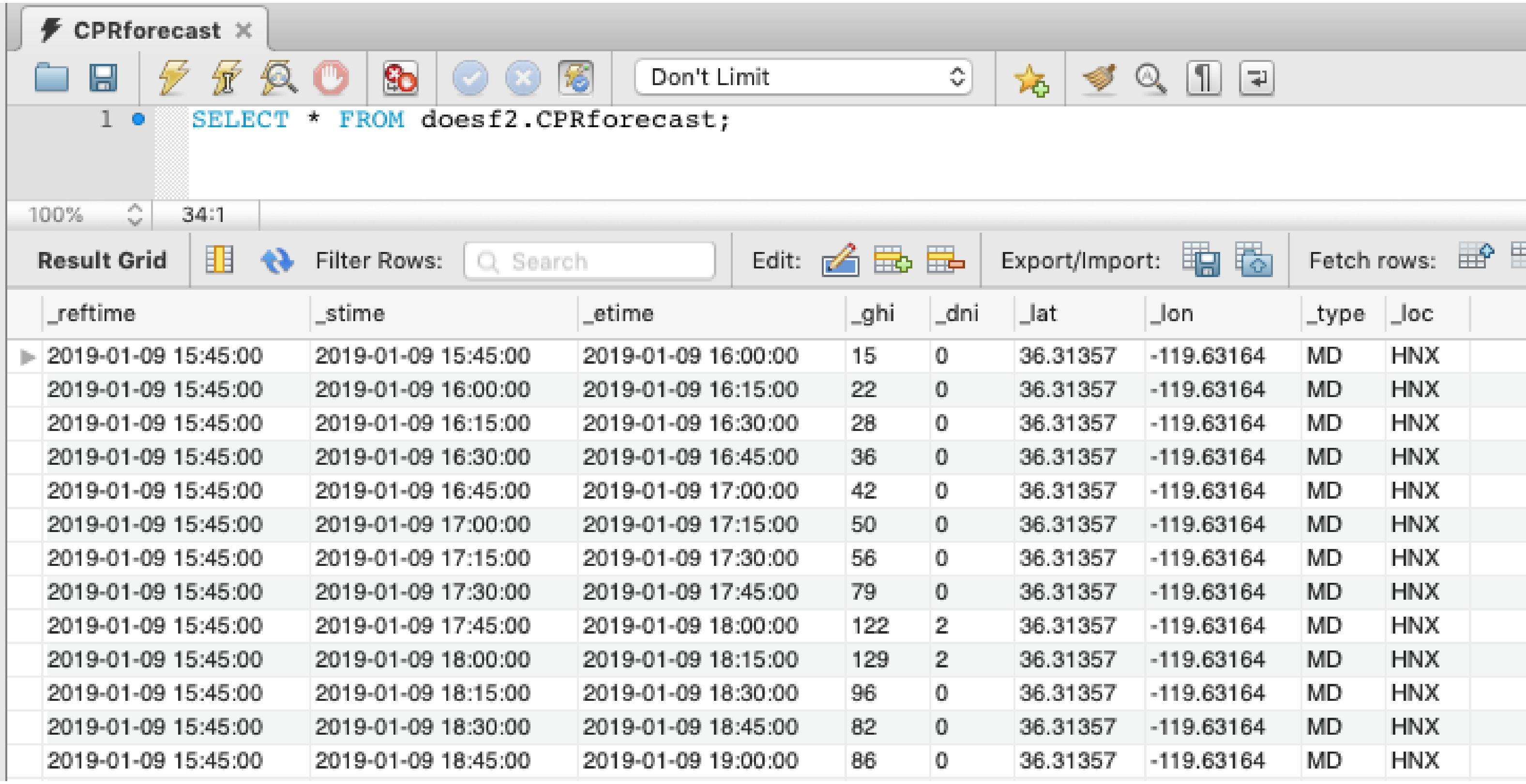
- Benchmark HAIMOS against state-of-the-art commercially available irradiance forecasts.
- Use CIF as inputs to HAIMOS
- CPR provided historical forecasts (backcasting) that will be used to augment the database from Task 1.



CIF forecasts for GHI for a five-day period (September 8 to September 13, 2018) in Bondville (BON). Measured data is plotted in black.

Task 2: Commercial Irradiance Forecasts

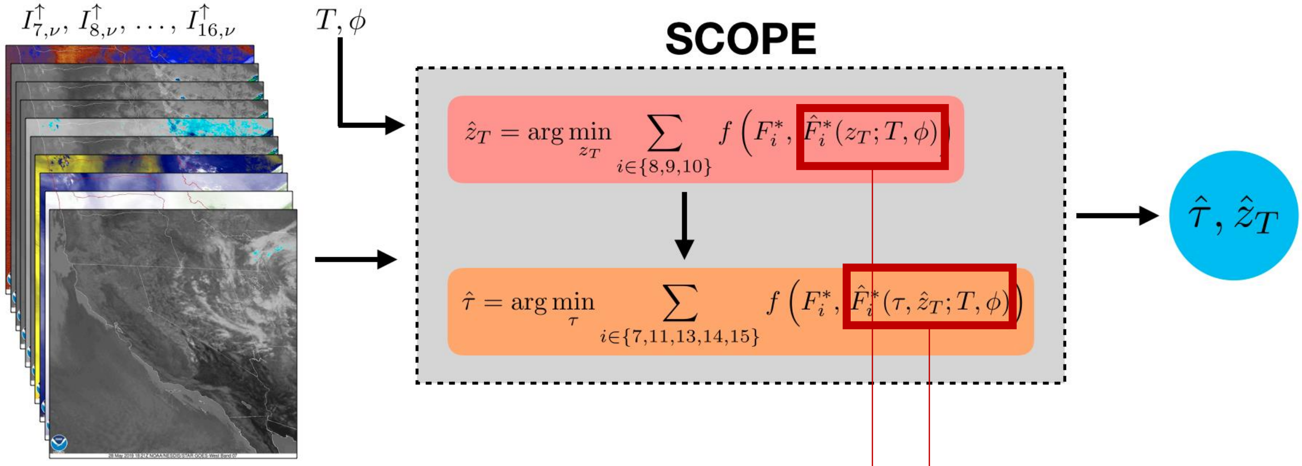
- UCSD uses a time-based job scheduler to run a python script that uses CPR's API to retrieve forecasts in real-time.
- The forecasts include GHI and DNI with a 15-minutes temporal resolution out to 2 days ahead of time.
- The data is then stored into a MySQL database.



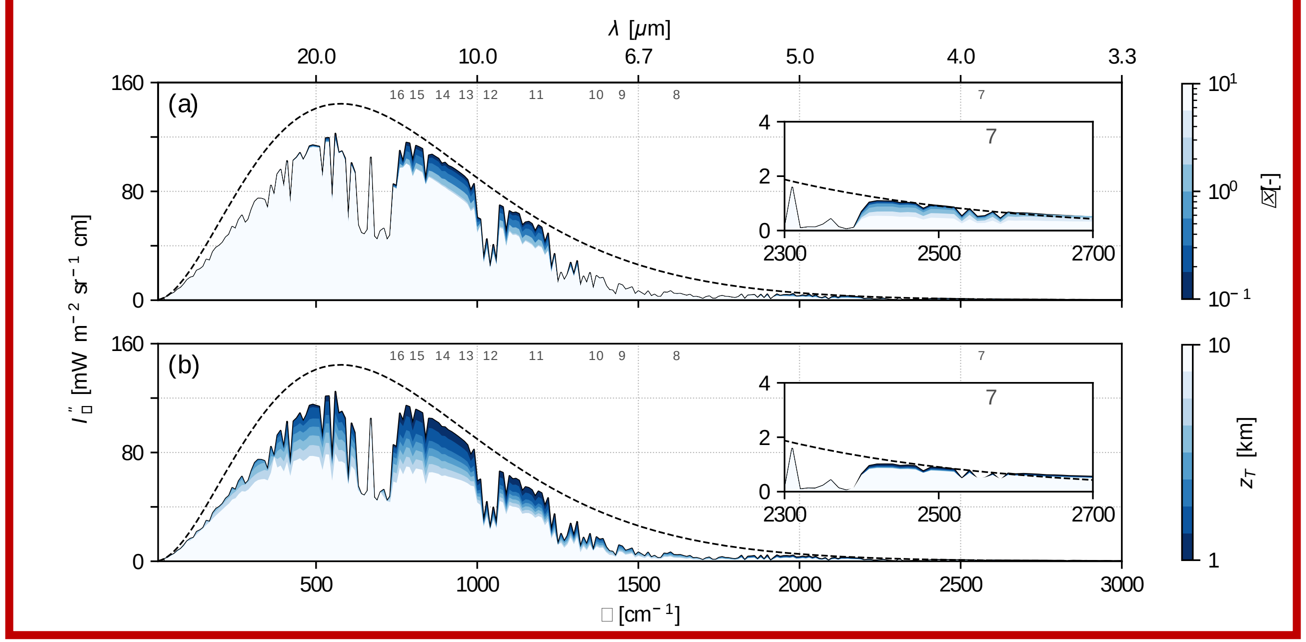
_reftime	_stime	_etime	_ghi	_dni	_lat	_lon	_type	_loc
2019-01-09 15:45:00	2019-01-09 15:45:00	2019-01-09 16:00:00	15	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 16:00:00	2019-01-09 16:15:00	22	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 16:15:00	2019-01-09 16:30:00	28	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 16:30:00	2019-01-09 16:45:00	36	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 16:45:00	2019-01-09 17:00:00	42	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 17:00:00	2019-01-09 17:15:00	50	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 17:15:00	2019-01-09 17:30:00	56	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 17:30:00	2019-01-09 17:45:00	79	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 17:45:00	2019-01-09 18:00:00	122	2	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 18:00:00	2019-01-09 18:15:00	129	2	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 18:15:00	2019-01-09 18:30:00	96	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 18:30:00	2019-01-09 18:45:00	82	0	36.31357	-119.63164	MD	HNX
2019-01-09 15:45:00	2019-01-09 18:45:00	2019-01-09 19:00:00	86	0	36.31357	-119.63164	MD	HNX

Task 3: Improvement of cloud cover forecast

- Spectral Cloud Optical Property Estimation (SCOPE)
  - couple radiative modeling with high-resolution spectral satellite imagery
  - real-time, accurate estimation of cloud optical properties
- Approach: compare outgoing longwave radiation (OLR) at the top of the atmosphere (TOA) from remote sensing and radiative modeling.
- Radiative model (Li et al. (2018)*)
  - spectrally-resolved and computationally efficient radiative model



RTM model variation with cloud optical depth (Top) and cloud top height (Bottom)

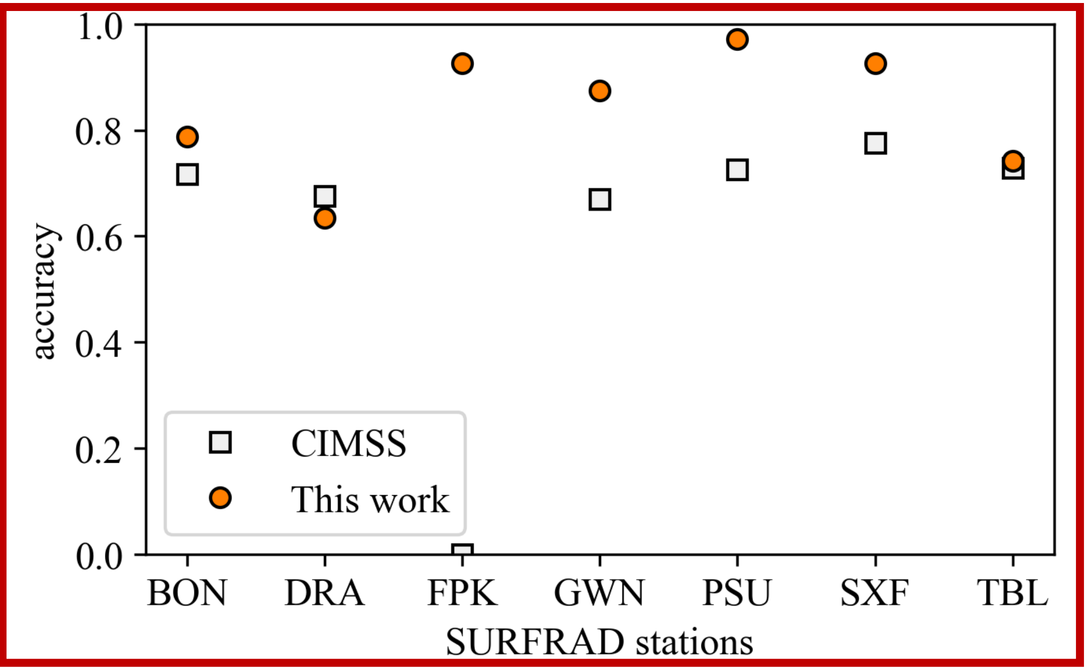


*Li, Liao and Coimbra (2018) "Spectral model for clear sky atmospheric longwave radiation"

Task 3: Improvement of cloud cover forecast

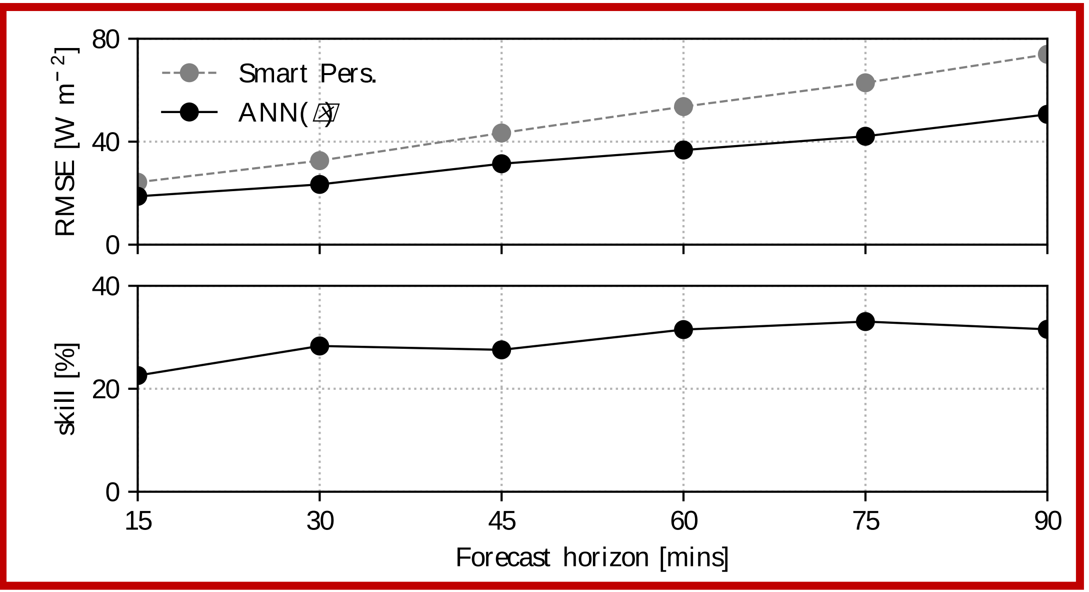
SCOPE's validation

1. Data: 1 complete year (2018) at 5-minute resolution for seven stations (SURFRAD).
   Validate against DLW measured at each site.
2. Clear-sky identification for SURFRAD sites. Compared against CIMSS data



SCOPE's data in forecasting

- GHI forecasting with estimated cloud optical properties from SCOPE
- GHI forecast performance for BON (testing set) for the early-morning hours
- Forecasts produced before sunrise.

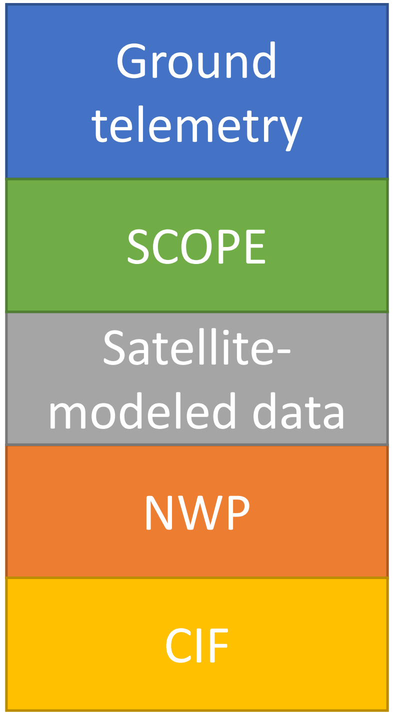


Task 4: Set up HAIMOS framework

- Motivation: Improving the forecast skill and especially the prediction during large variability periods depends on the selection of the proper inputs available in the search space and machine-learning (ML) algorithm
- Technical approach:
  1. Input selection
     - Feature engineering.
     - Selection of the proper inputs available in the search space.
     - Input selection based on techniques such as clustering analysis and time series correlation analysis.
  2. Training and optimization of ML algorithms
     - Identification of the best type of nonlinear approximator to map the input data into the target irradiance data
     - ANNs, kNNs, SVR, non-linear least-squares, Gradient boosting
  3. Developing adaptive protocols for HAIMOS
     - Continuous training
     - Analog training

Task 4: Set up HAIMOS framework

- Input selection
  - Currently we have ~200 input variables available
  - All irradiance data normalized using a clear-sky model



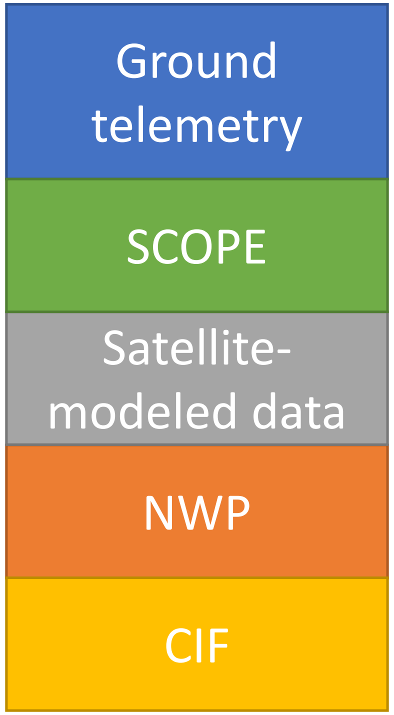
Task 4: Set up HAIMOS framework

- Input selection
  - Currently we have ~200 input variables available
  - All irradiance data normalized using a clear-sky model

Ground telemetry	GHI and DNI data Additional features (averages, STD, …)
SCOPE	COD, cloud height, cloud mask
Satellite-modeled data	Satellite-derived GHI and DNI from a domain surrounding target locations (49 nodes)
NWP	NAM, HRRR, WRF-Solar, …
CIF	GHI and DNI forecasts from CPR

Task 4: Set up HAIMOS framework

- Input selection
  - Currently we have ~200 input variables available
  - All irradiance data normalized using a clear-sky model



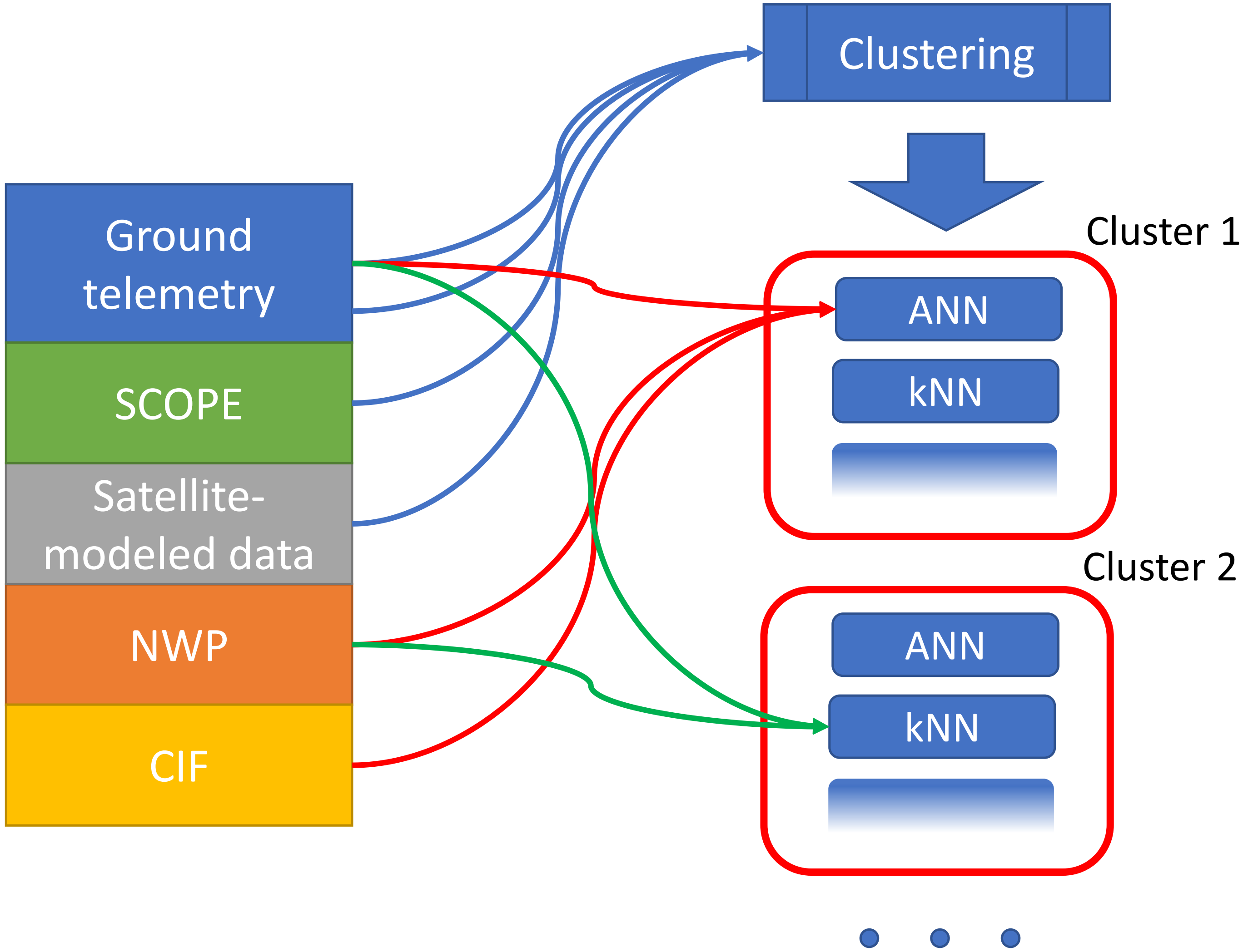
- Greedy algorithm
- Test all inputs in terms of bias-variance metrics
- Selects input that ranks highest
- Iterates over unselected inputs until no improvements are observed.

Task 4: Set up HAIMOS framework

- Input and machine-learning approximator selection is done using a genetic algorithm (GA)

Task 4: Set up HAIMOS framework

- Input and machine-learning approximator selection is done using a genetic algorithm (GA)



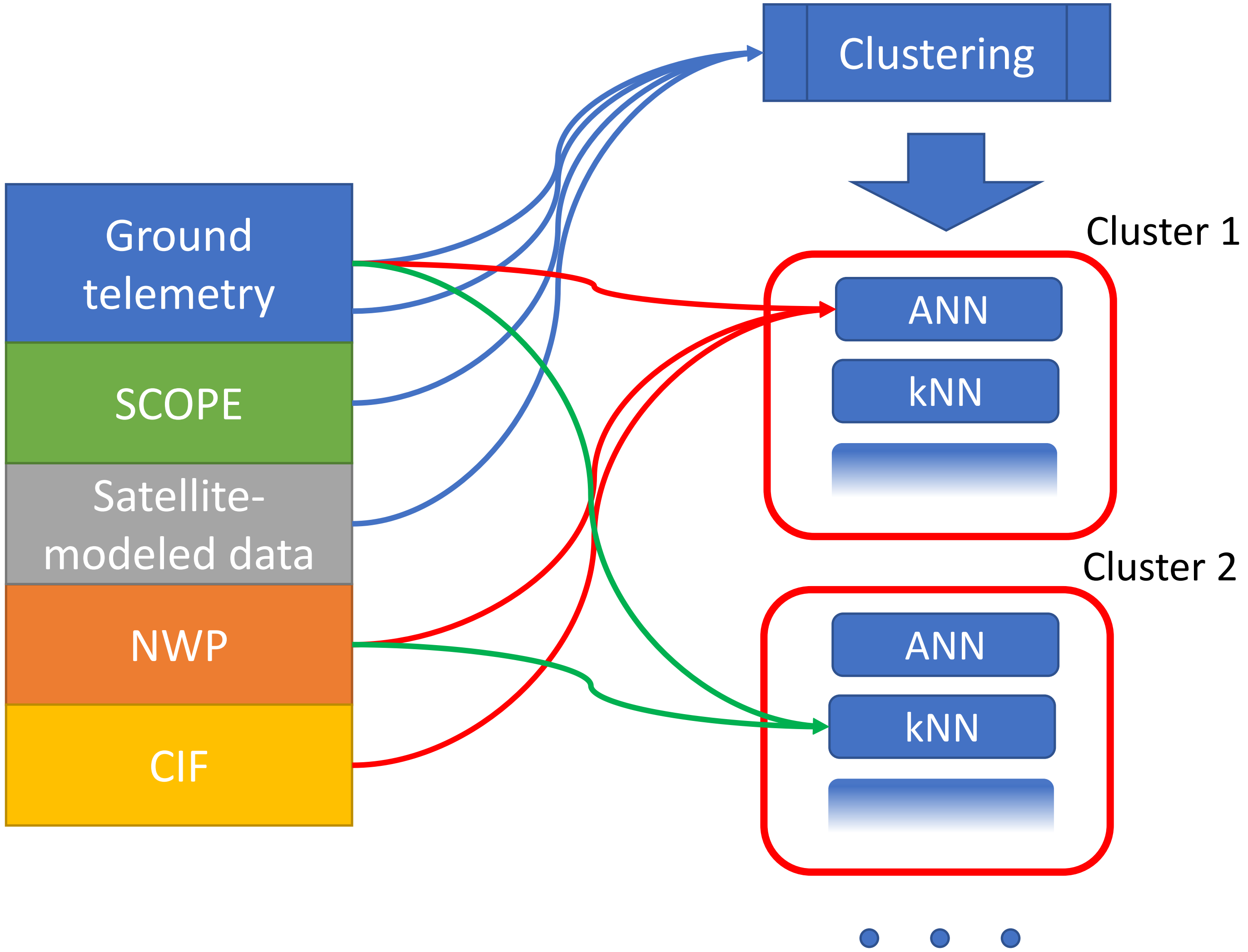
GA optimization

1. Number of clusters
2. Inputs used in clustering
3. Selection of ML algorithm

Greedy input selection optimization

1. Inputs for ML algorithm

Task 4: Set up HAIMOS framework



Once GA converges forecasts for new data (validation) are created

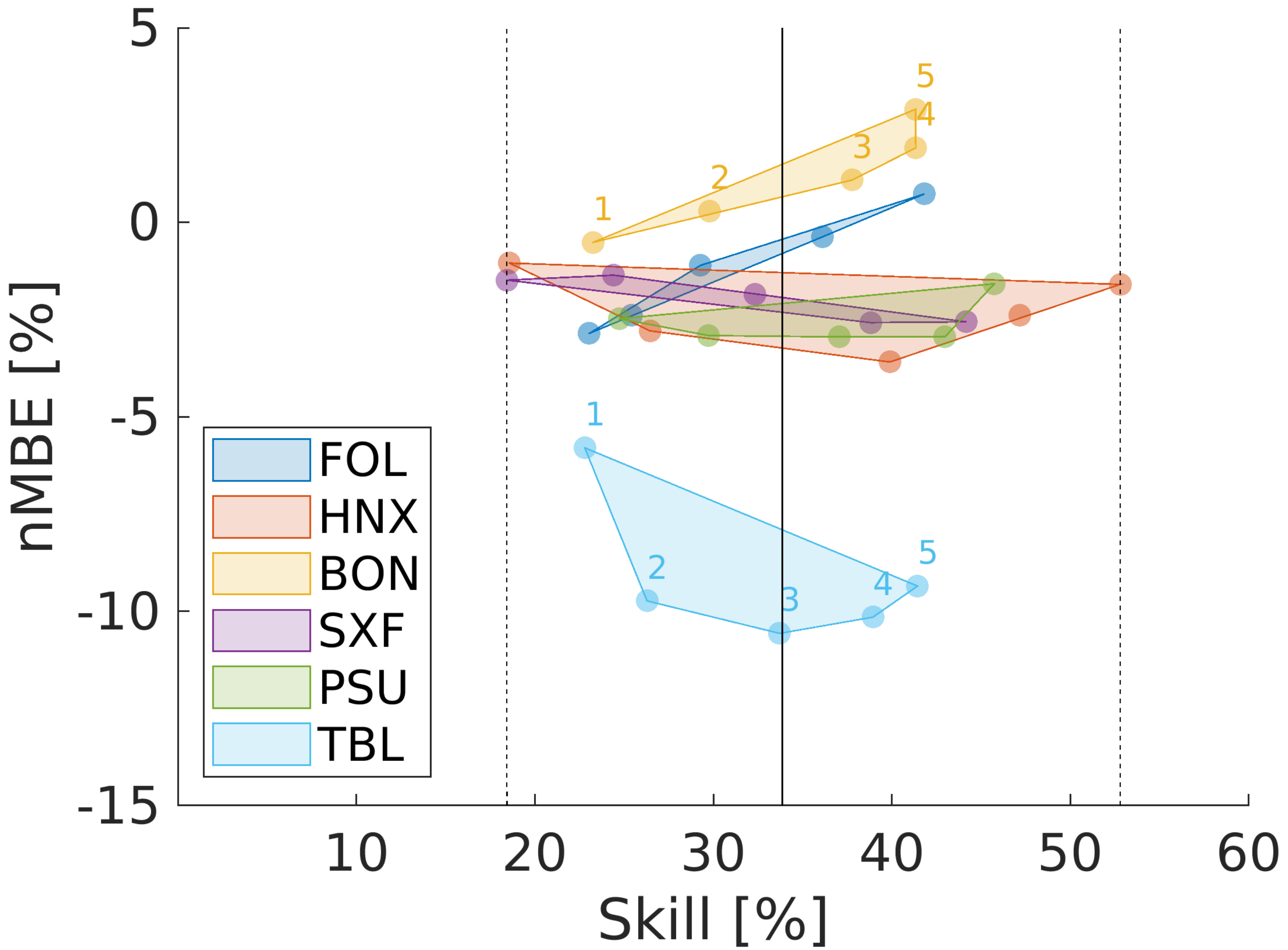
1. Assign new data to one of the clusters
2. Selected the input variables and model associated with the cluster
3. Produce forecast

Task 4: Set up HAIMOS framework

- Forecast validation
  - 6 selected locations
  - July 2017 to Dec 2018 (training from Jan 2016 to June 2017)
  - Error metrics
    - $nMBE$ (%)
    - Forecast skill: $s = \left(1 - \frac{\mathrm{RMSE(HAIMOS)}}{\mathrm{RMSE(Persist.)}}\right) \times 100\ (\%)$
  - Daytime only ($\theta_s < 85°$)
  - Results for periods of larger variability ($k_t < 0.9$)

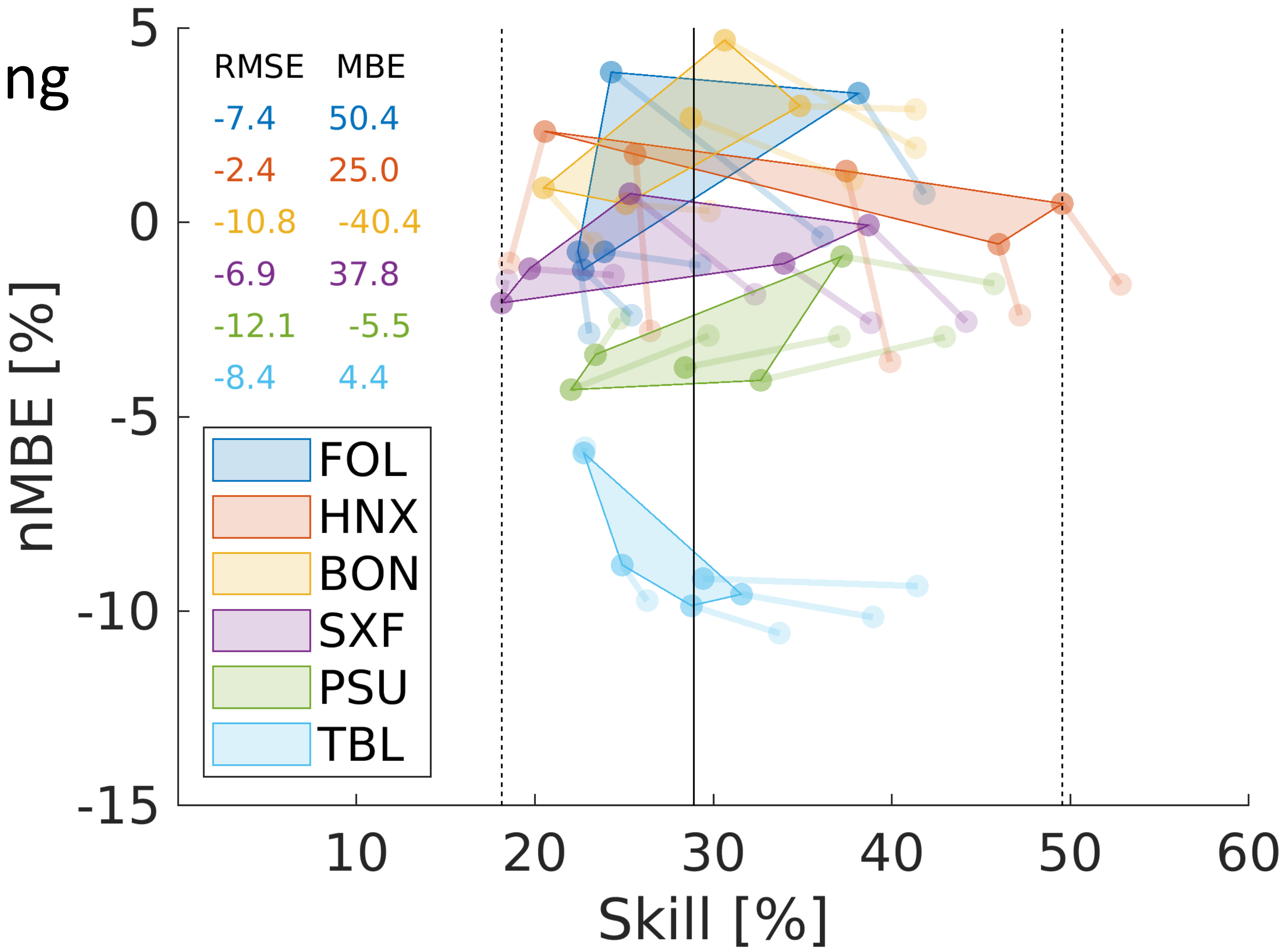
Task 4: Set up HAIMOS framework

- GHI forecasts
  - Hourly values
  - 1 to 5 hours ahead of time



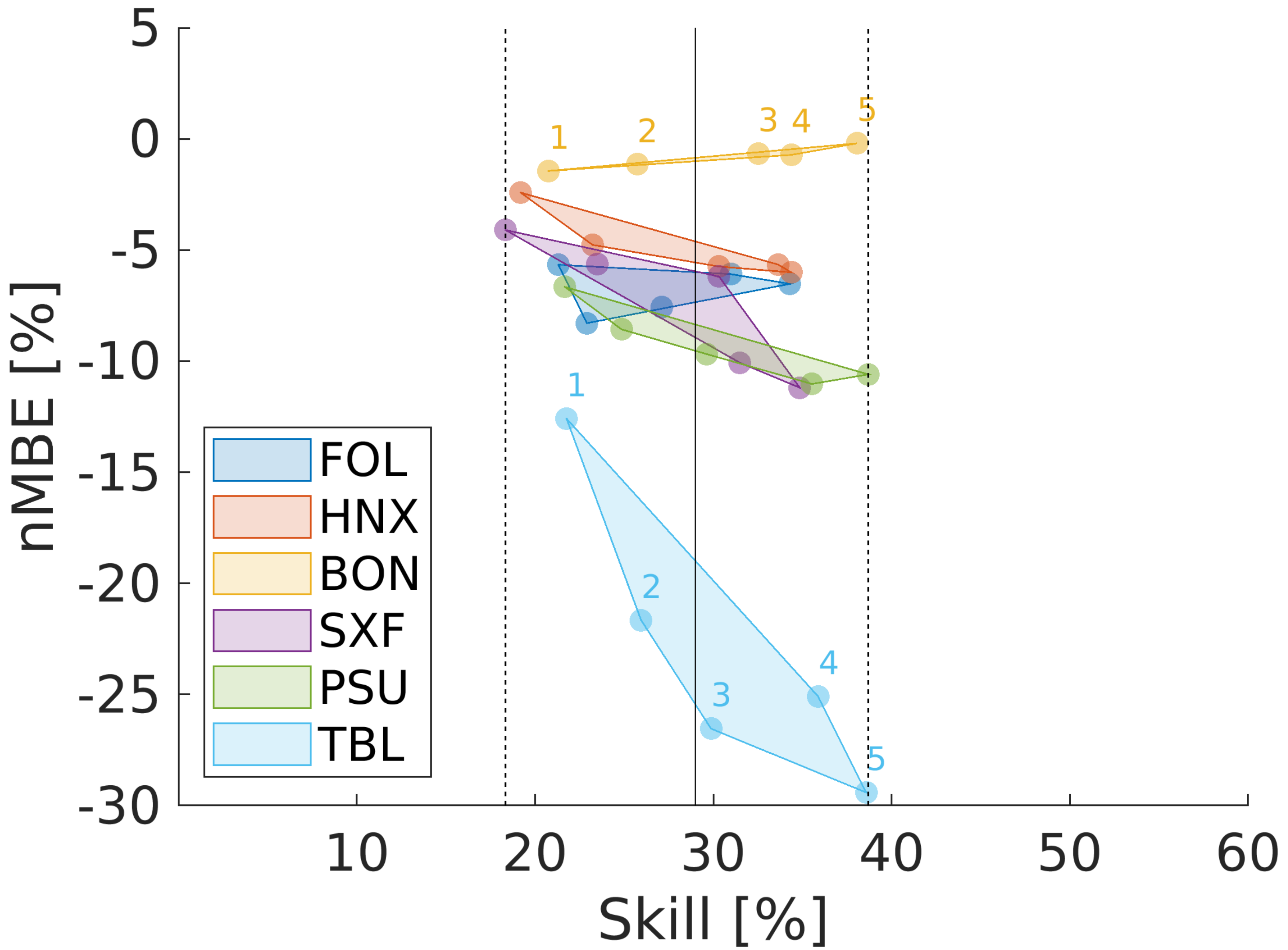
Task 4: Set up HAIMOS framework

- Adaptive HAIMOS
  - Continuous training
  - Analog training



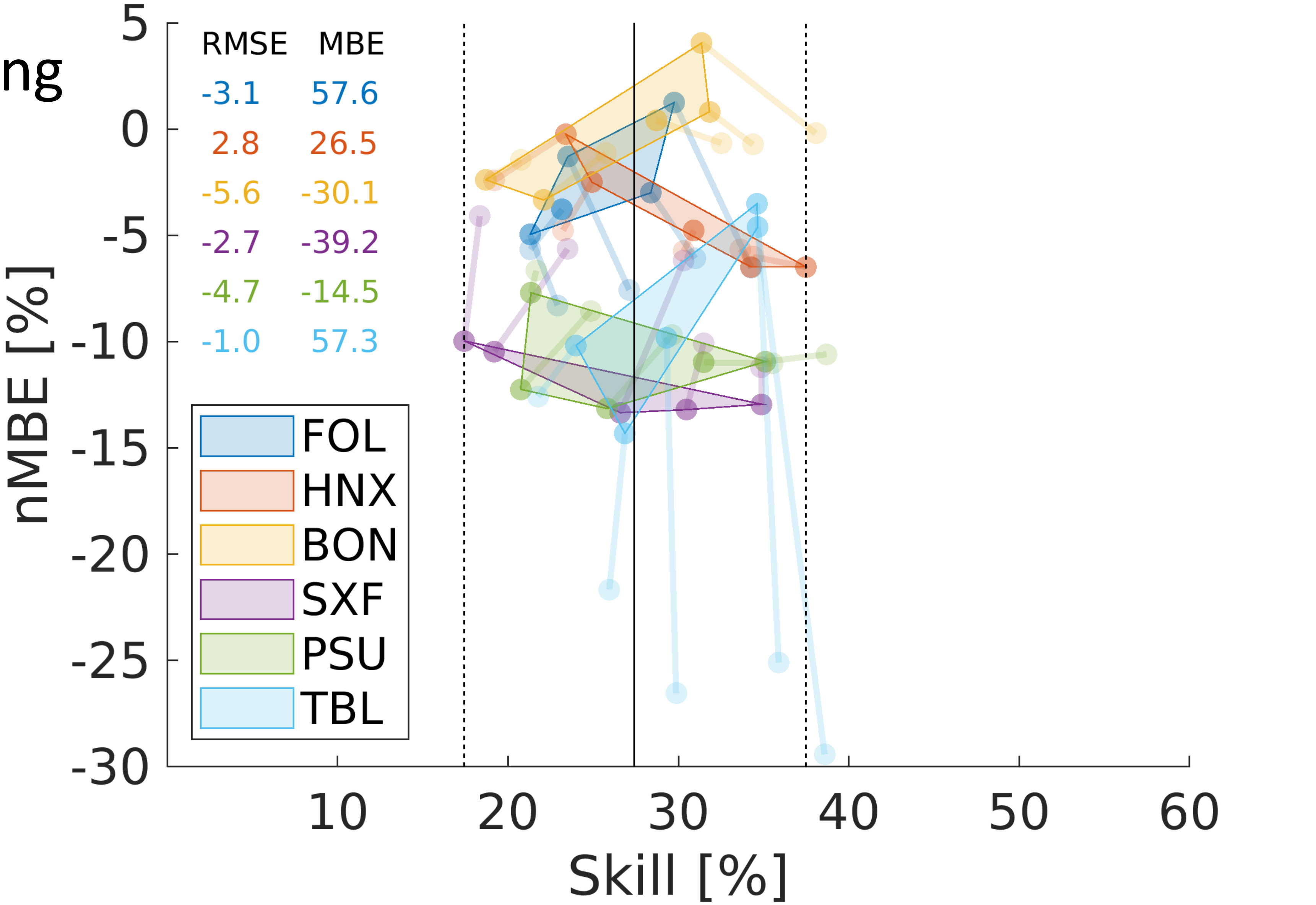
Task 4: Set up HAIMOS framework

- DNI forecasts
  - Hourly values
  - 1 to 5 hours ahead of time



Task 4: Set up HAIMOS framework

- Adaptive HAIMOS
  - Continuous training
  - Analog training



Next steps

- Optimization of HAIMOS framework
  - optimal objective function
  - optimal ensembling
  - probabilistic forecast and ramp onset forecast
- New-generation cloud forecasting tools
  - GOES-16,17 high-resolution cloud cover identification
  - LES cloud cover forecasting

More information

- Data, reports and additional information available at http://coimbra-server3.dynamic.ucsd.edu/doesf2

PROJECT PROFILE: University of California San Diego (Solar Forecasting 2)

Project Name: HAIMOS Ensemble Forecasts for Intra-day and Day-Ahead GHI, DNI and Ramps

Funding Opportunity: Solar Forecasting 2

SETO Subprogram: Systems Integration

Location: San Diego, CA

SETO Award Amount: $1,316,203

Awardee Cost Share: $162,500

Principal Investigator: Carlos F. M. Coimbra

This project will develop the Hybrid Adaptive Input Model Objective Selection (HAIMOS) ensemble model for solar irradiance forecasting. HAIMOS is a hybrid physics-based/data-driven model that forecasts direct normal and global horizontal irradiance (DNI and GHI) for horizons ranging from 1 to 72 hours. The project leverages the experience of the research team in solar forecasting to address key gaps in these technologies, namely: the lack of accurate solar forecasts for DNI and inaccurate forecasting of large irradiance ramps. The project aims to achieve a 50% forecast skill (improvement relative to a baseline persistence forecast) across a wide range of horizons for both GHI and DNI.

APPROACH

The proposed model (HAIMOS) combines Numerical Weather Prediction (NWP) forecasts, determinist

UC San Diego

U.S. DEPARTMENT OF ENERGY

Energy Efficiency & Renewable Energy

Clean Power Research

OUTCOMES

Project outcomes will be listed here for each reporting period (quarterly).

- Budjet Period 1 - Quarter 1
- Budjet Period 1 - Quarter 2

DELIVERABLES

Project deliverables will be listed here for each budget period.

Questions?