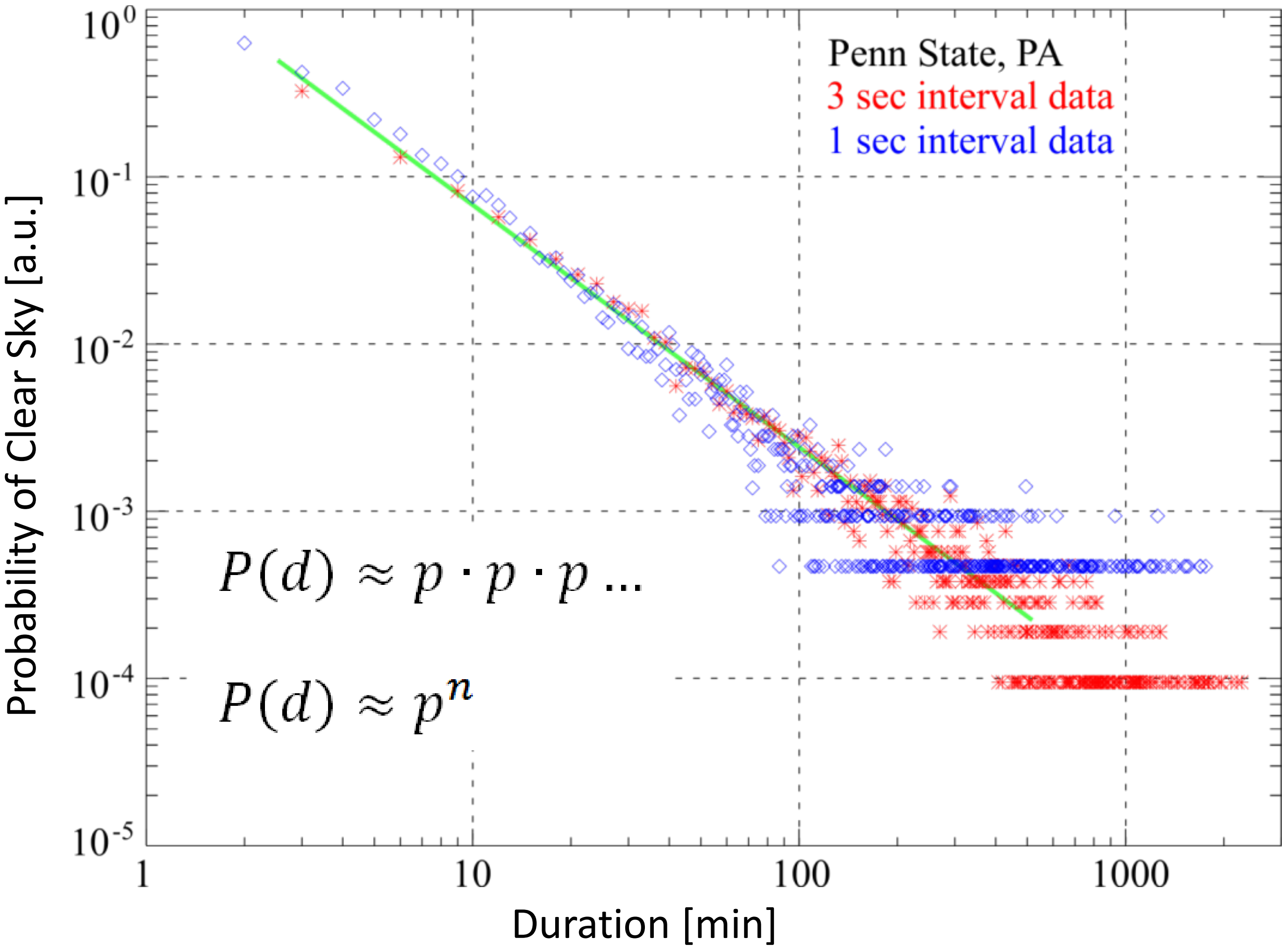
Machine-learning based enhancements for renewable energy forecasting:

From Research to Applications

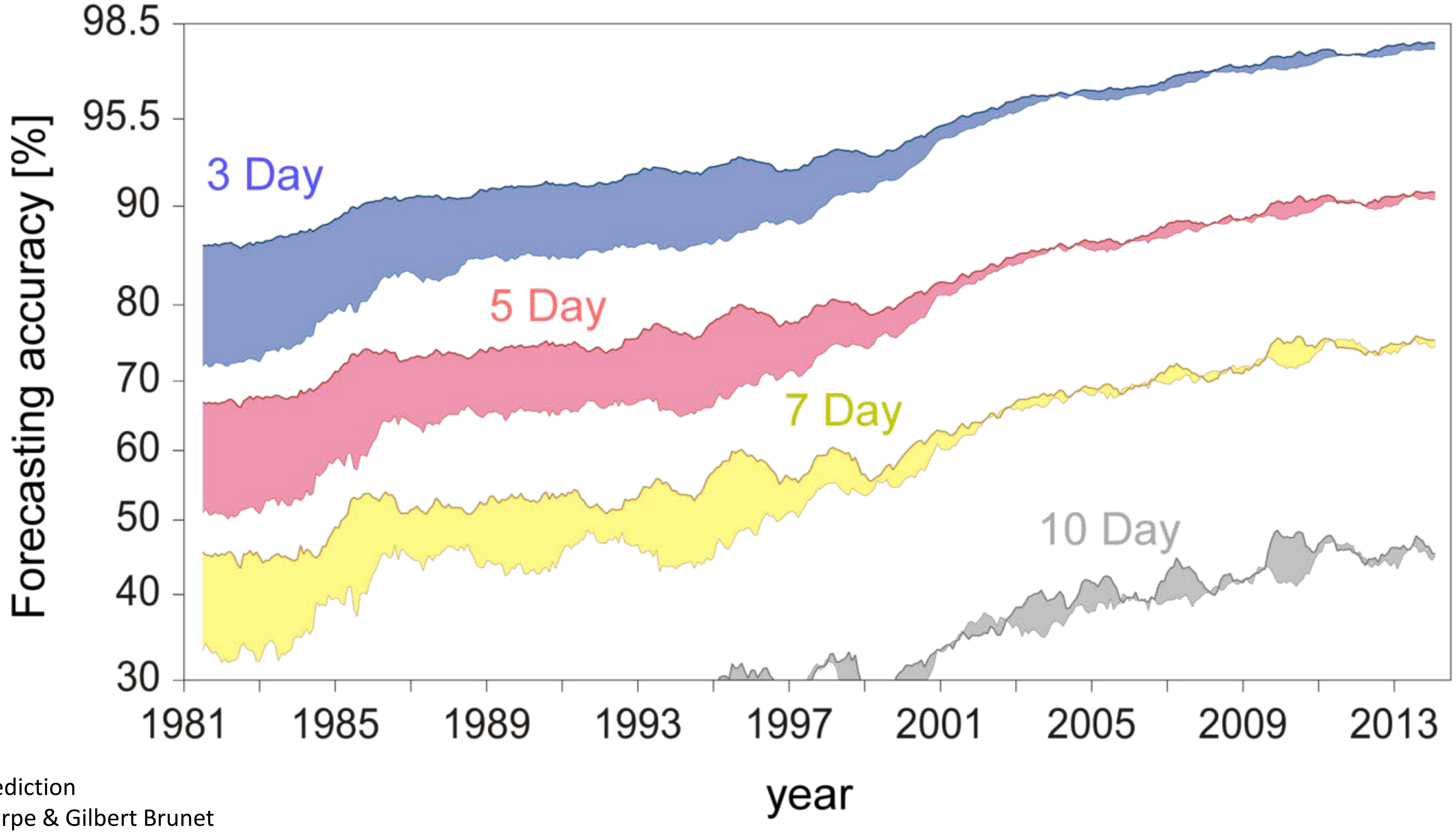
The challenge is that the *probability of a clear sky* for an extended period is (actually) very low*



Probability for longer periods of clear sky

- Becomes exponentially lower
- Follows a power law over 4 orders of magnitudes
- Similar behavior at other locations

The challenge is that historically NWP* model accuracy improvements have been (only) ~6% per** decade



* Numerical weather prediction

** Peter Bauer, Alan Thorpe & Gilbert Brunet

doi:10.1038/nature14956

IBM decided to leverage "big" data analytics to improve accuracy of forecasting

NWP improvements have been averaging 6% per decade*

This approach: Complementing NWP with machine-learning and **big data** analytics

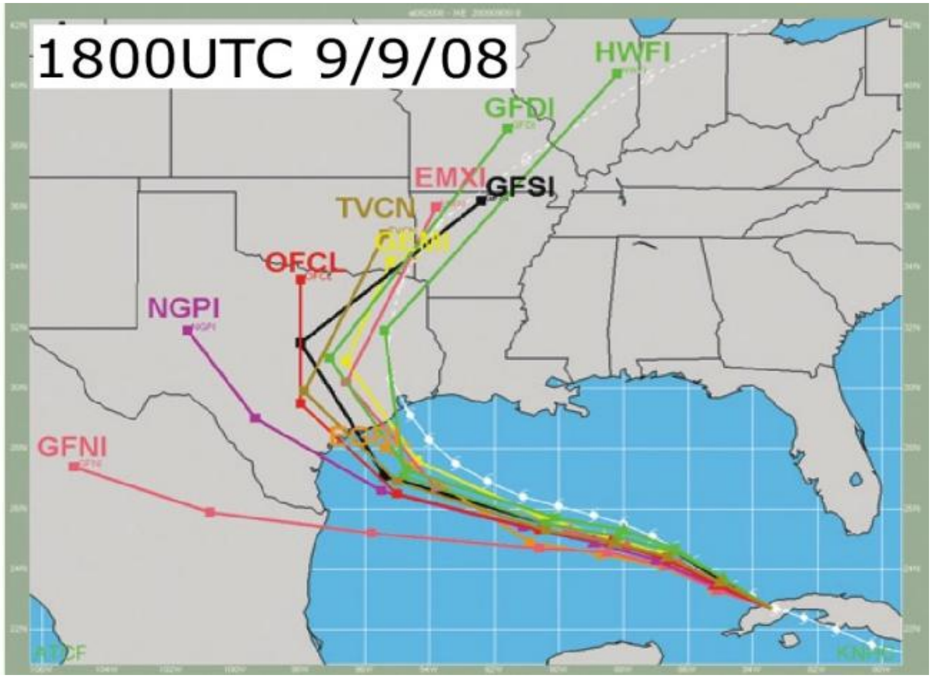
Everybody talks about big data but what is big data really?

- Big data is too big to be “moved”
- Forecasting models are becoming bigger
  - Global Forecast System (GFS): 140GB/day, increasing to 1.5TB/day
  - Global Ensemble Forecast System: 302GB/day, increasing to 3TB/day
  - Generally forecasts are not being stored
- E.g., what does 3TB mean?
  - It takes 10 hours to load 3TB from the disk to the memory
- Efficient processing requires
  - abandon file-based systems (grib, hdf, netCDF)
  - indexing of raw data
  - processing it in a massively distributed system

Key idea: Use **historical forecasts** and weather data to learn which model is better, when, where and under what situation

- Different forecasting models provide varying accuracies depending on weather situation etc.
- Apply deep machine learning / "adaptive mixture of experts" to learn *from historical data* which model is better when, where and under what situation
- Obtain dynamically optimal blending coefficients for different models to create **a super forecast**
- Adaptive mixture of expert approach has been successfully applied to:
  - Jeopardy! Challenge
  - Speech recognition
  - Medical diagnosis
  - ….

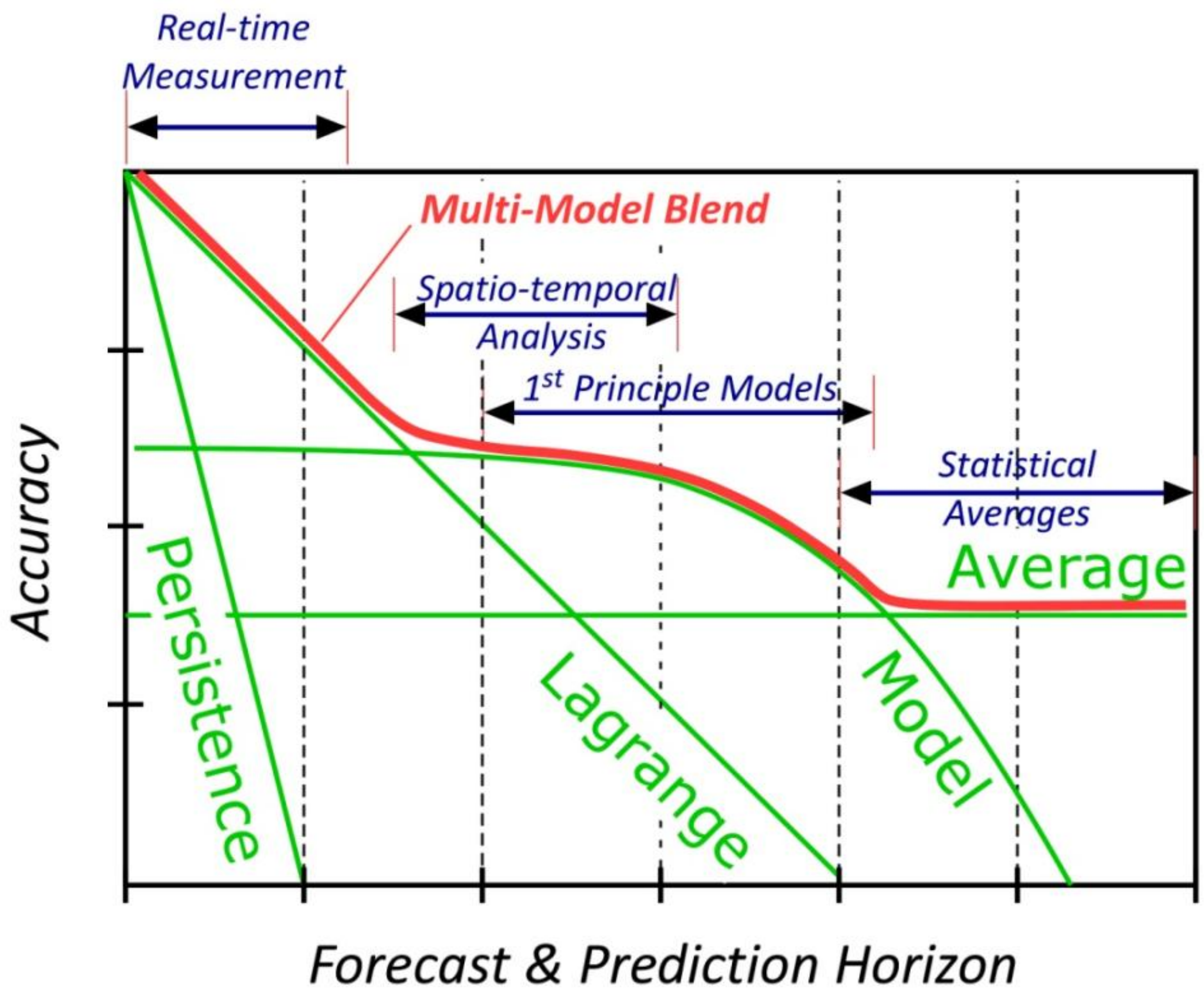
Hurricane Ike path forecasts from 9 different weather models*



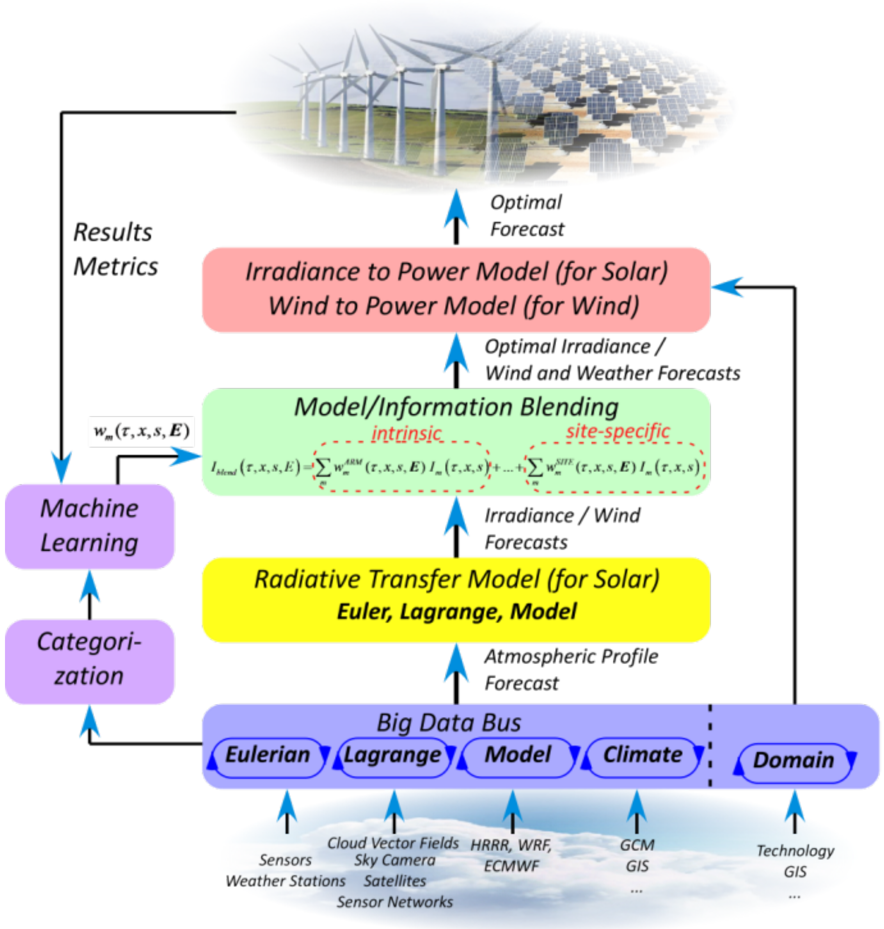
*M.J. Brennan, S.J. Majumdar, Weather and Forecasting 26, 848 (2011) An Examination of Model Track Forecast Errors for Hurricane Ike (2008) in the Gulf of Mexico

Multiple information and data sources are being fused to create a super forecast

- **Persistence:**
  - Real-time power data
  - Weather station data
- **Lagrangrian Forecast Models:**
  - Sky camera model
  - Satellite-based (GOES), advection models
  - Time-series models
- **Weather Forecast Models:**
  - Rapid Refresh (RAP)
  - Hi-Resolution Rapid Refresh (HRRR)
  - Short-Range Ensemble Forecast (SREF)
  - North American Mesoscale Forecast (NAM
  - Global Forecast System (GFS)
  - European Center for Medium range Weatl (ECMWF)
- **Climate Models:**
  - Climate Forecasting System (CFS)



An general platform for accurate and adaptive forecasting



Key technologies:

- uses **big data** information processing (hadoop, hbase,) technologies
- applies **!Jeopardy like machine learning approaches** to blend outputs from multiple models and to enhance system intelligence, adaptability and scalability.

System includes:

- Big Data Bus*
- Radiative Transfer Model (for Solar)
- Model / Information Blending
- Irradiance to Power Model (for Solar)
- Wind to Power Model (for Wind)
- Machine / Learning and Categorization

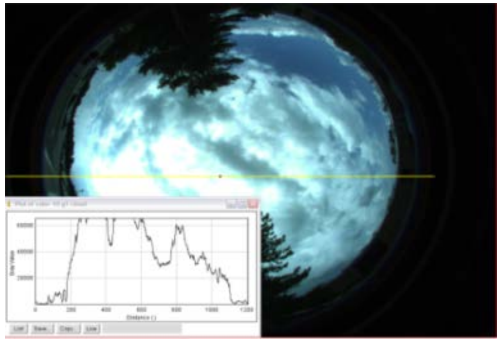
- **System provides a *platform* to optimally leverage current and future forecasting capabilities and models.**
- **Adapts autonomously to different metrics and applications.**

Lagrangrian forecasts using sky cameras

IBM Cloud Imaging System
without mechanical shutter

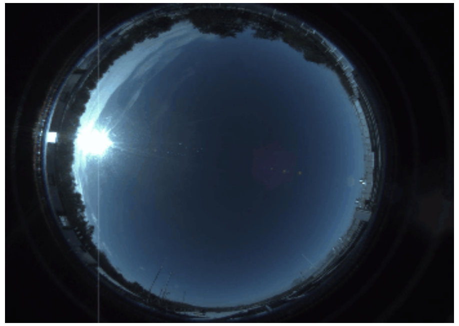
"Fish eye" lens

24 bit camera with several gain stages

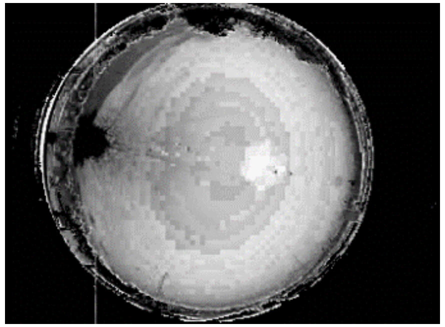


- Sky camera with fish eye lenses detects arrival incoming clouds
  - Field of view ~ 2 miles, no mechanical parts
- Multiple sky cameras increases prediction horizons and allow cloud height detection

SkyCam Image

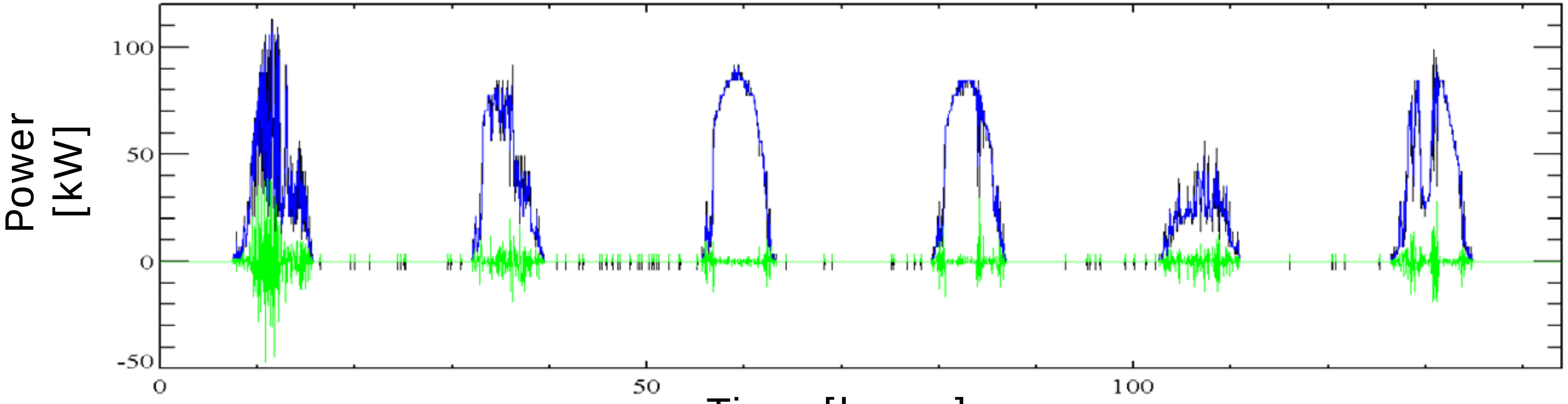


Sky Transparency



6 min optical flow based forecasts for 6 consecutive days

Measured
Forecast
Error



Short-term optical flow based forecasting with Navier-Stokes Modeling using GOES Satellites

Conventional Cloud Propagation:

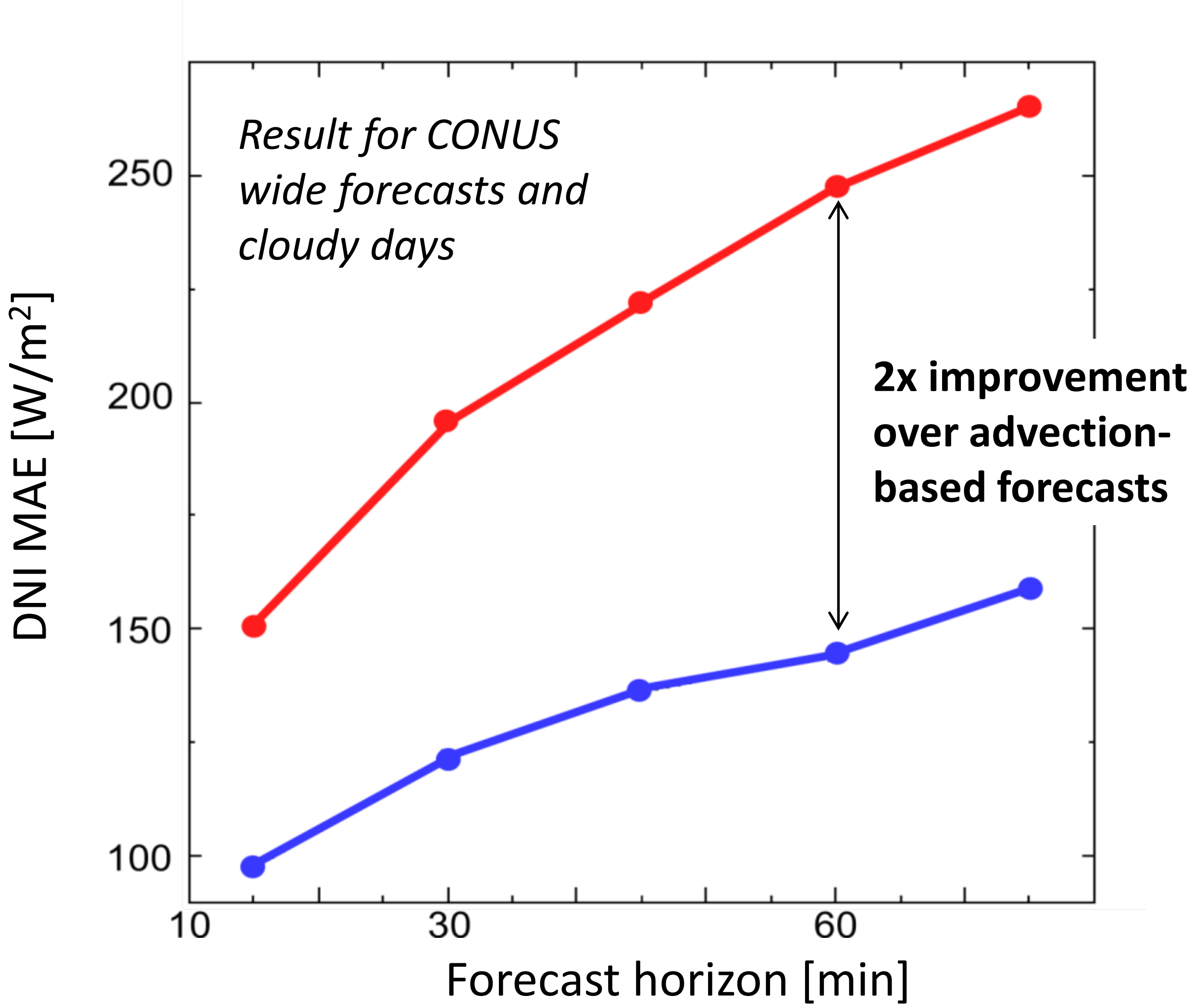
- Using (filtered) NWP wind field
  Inaccurate wind (error in cloud height estimate)
- Using wind field from optical flow
  Neglecting wind dynamics in hour-ahead.

***New method** keeps accurate wind field determined by optical flow, but captures basic wind dynamics.*

1. Optical flow estimates wind field using two consecutive cloud images (optical depth)

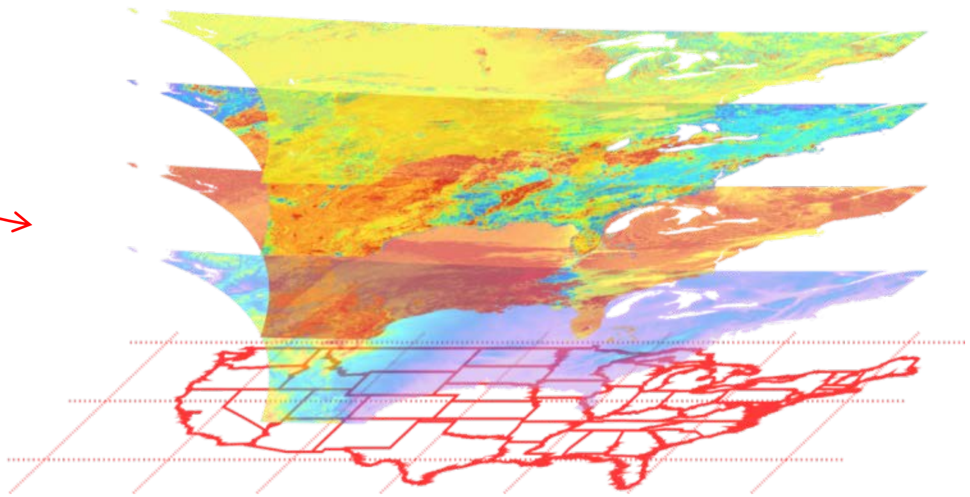
Fit an initial condition of 2D Navier-Stokes Equation to two consecutive optical flows

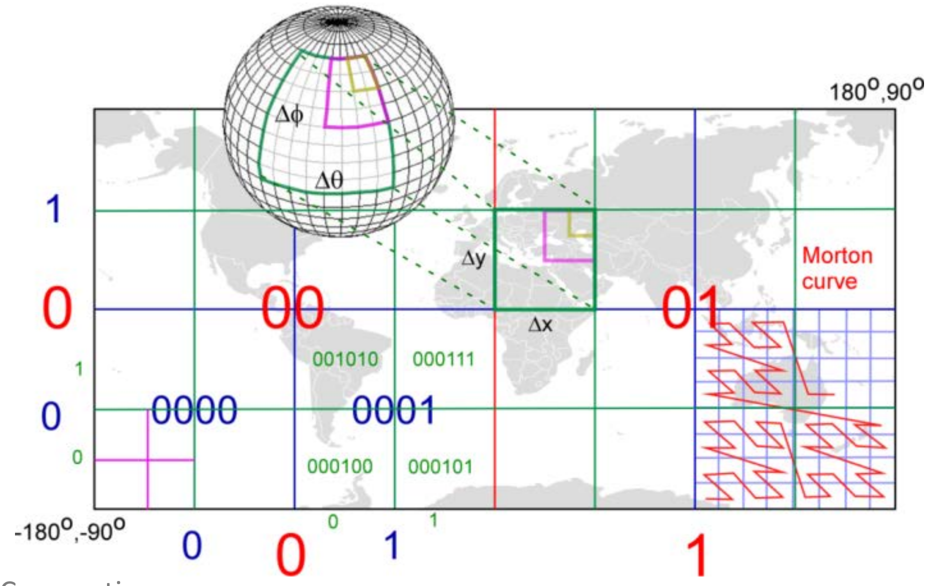
Forecast optical flow and use it to predict cloud images



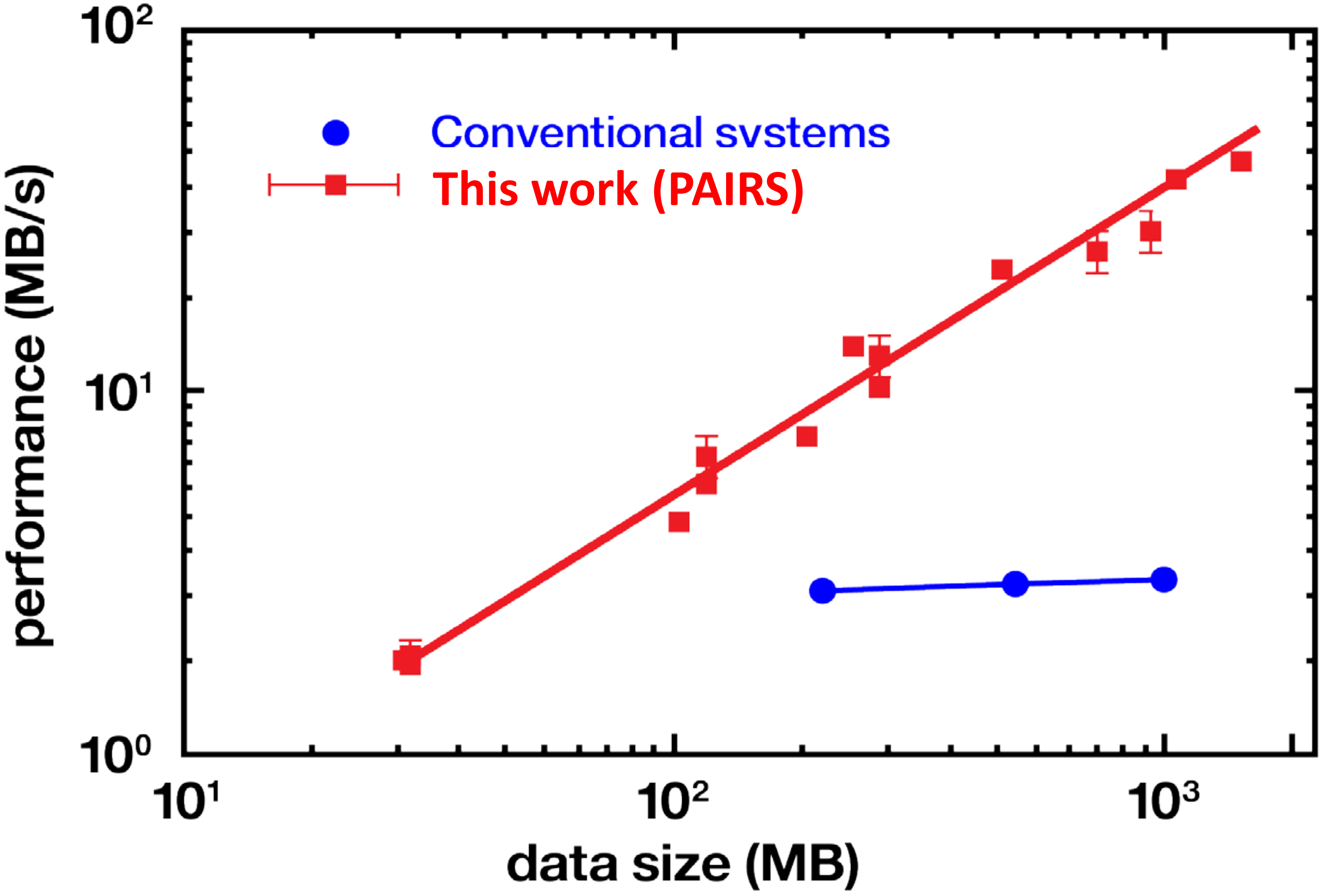
Big data bus indexes and aligns to a global spatio-temporal reference and indexing system

	Lagrangrian		*Model*						*Climate*
Models	**Sky Cam**	**GOES**	**RAP**	**HRRR**	**SREF**	**NAM**	**GFS**	**ECMWF**	**CFS**
Spatial Res & Coverage	Local 10 m	Global 4km	US 13km	US 3km	US 16km/ 40km	US 5km	Global 0.5 deg	Global 0.1 deg	Global 0.5 deg
Temporal Resolution	1 min	15 min	15 min 2D 1 hr 3D	15 min 2D 1hr 3D	1 hr for (40km) 3 hr for (16km)	1 hr	3 hr	1 hour	6 hr
Forecasting Horizon	10 min	4 hr	18 hr	15 hr	0-87 h	0-60 h	6-192 h	0 -60 h	6 months
Ensemble Forecast	No	No	No	No	CTL, P1, P2, P3, N1, N2, N3	No	No	NA	4 members





Scalable Machine-learning on big spatio-temporal data



- Machine learning performance of this technology is (almost) independent of data size
- Time to result is independent of how much data is processed
- Conventional systems require more time for larger data sizes

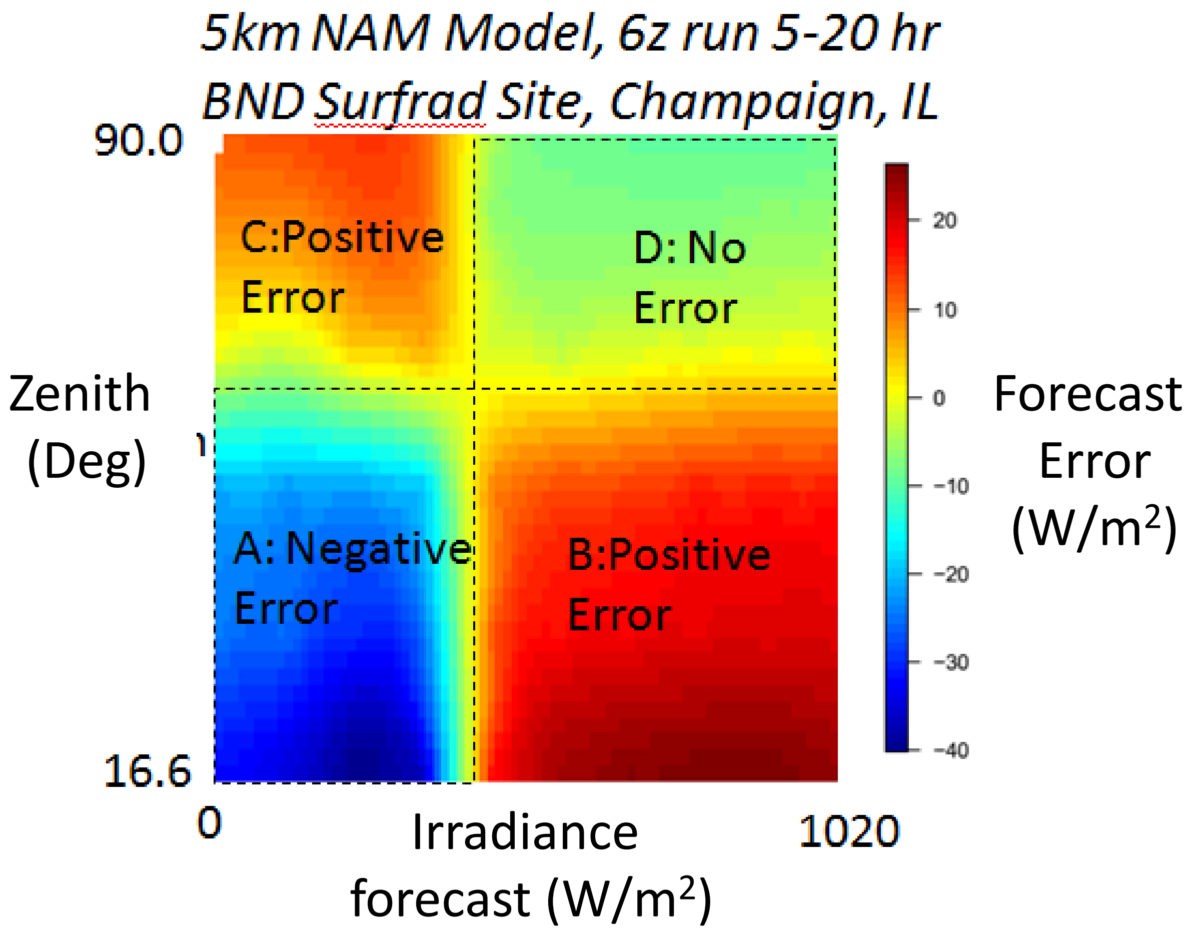
Improving accuracy using situation dependent, machine-learnt, multi-model blending

Question: **Which model is more accurate, when, where, under what weather situation?**

- Apply functional analysis of variance to understand 1st,2nd,3rd, ….order errors
- Model accuracy can depend strongly on **"weather situation"** category.
- "Weather situation" is determined using a set of parameters including forecasted ones on which model error depends on strongly.

Example, NAM solar irradiance forecast

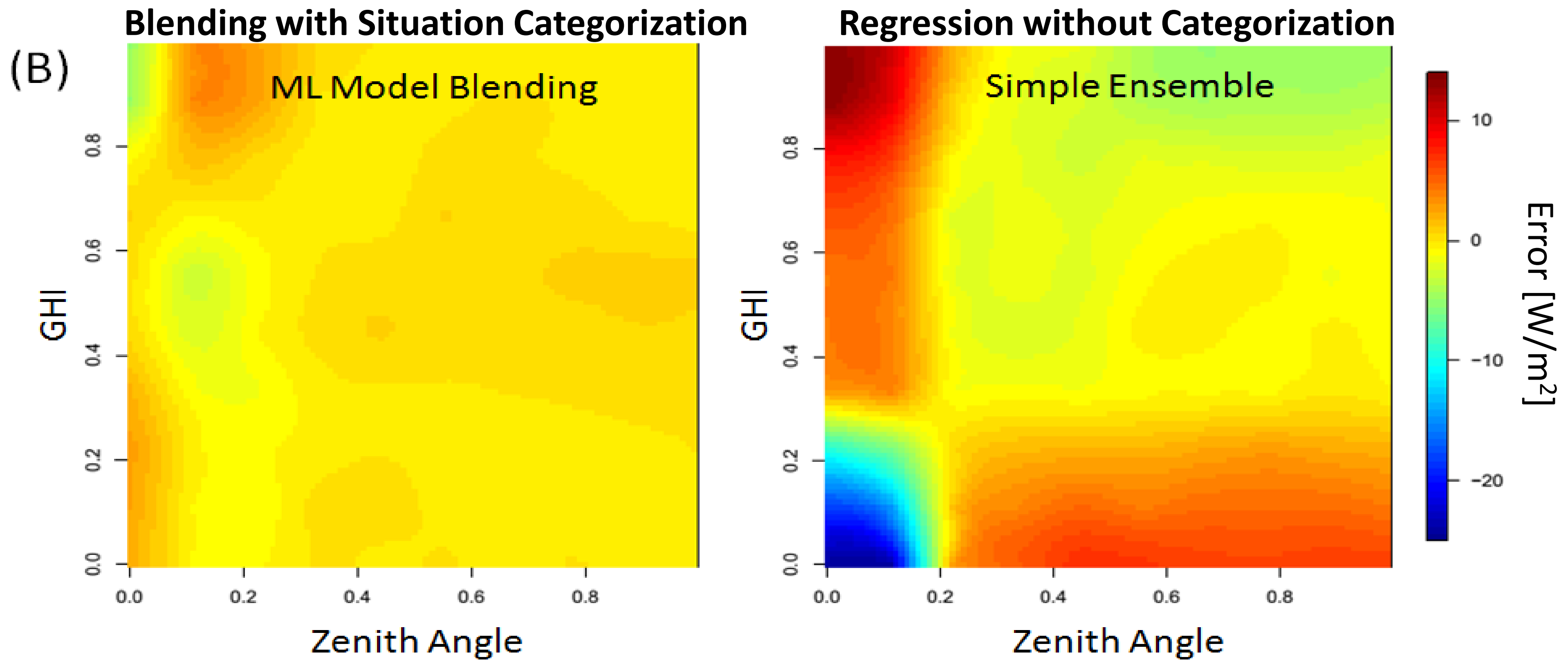
- Depends strongly GHI and solar zenith angle.
- The two parameters create four categories of situations below.



Reduction of forecast error using situation dependent machine learning based multi-model blending

Three models: RAP 11z (0-15hr), HRRR 11z (0-15hr), NAM 6z (5 to 20 hr ahead)

Average d for Seven Surfrad Stations.

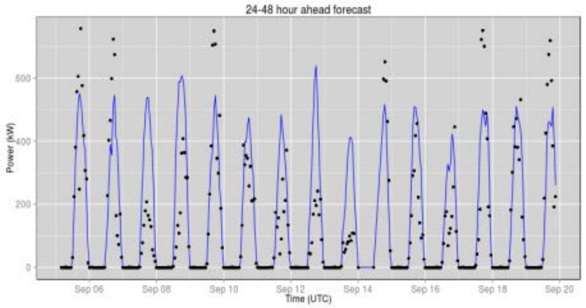


After model blending, "situation dependent " bias error is essentially eliminated.

Local, regional, and probabilistic forecasts

Single Plant – Fixed Systems

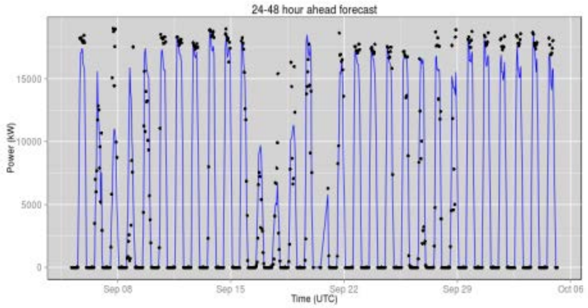
- Smyrna, TN
- 1MW Nameplate Capacity
- 24 to 48 hr ahead forecasts
- **MAPE 11% (2014-5-1 to 2014-10-31)**



Metrics	24-47 hr 2nd Yr
Correlation coef.	0.766
RMSE (MW)	0.15
NRMSE by capacity	0.155
MaxAE (MW)	0.48
MAE (MW)	0.111
MAPE by capacity	0.115
MBE (MW)	0.0211
KSIPer (%)	12.601
Stdev. (MW)	0.148
Skewness	0.296
Kurtosis	0.238
4RMQE (MW)	0.203
N4RMQE	0.21
95th percent (MW)	0.301

Single Plant – 1D Tracking System

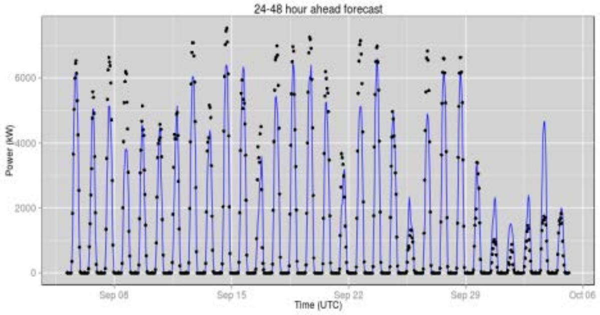
- TEP FRV Site, Marana, AZ
- 20MW AC Capacity
- 24 to 48 hr ahead forecasts
- **MAPE 11% (2014-5-1 to 2014-10-31)**



Metrics	24-47 hr ,2nd Yr
Correlation coef.	0.854
RMSE (MW)	3.4
NRMSE by capacity	0.17
MaxAE (MW)	18.5
MAE (MW)	2.23
MAPE by capacity	0.111
MBE (MW)	-0.194
KSIPer (%)	20.076
Stdev. (MW)	3.43
Skewness	1.41
Kurtosis	4.9
4RMQE (MW)	5.69
N4RMQE	0.282
95th percent(MW)	6.36

Regional Forecast for ISO-NE

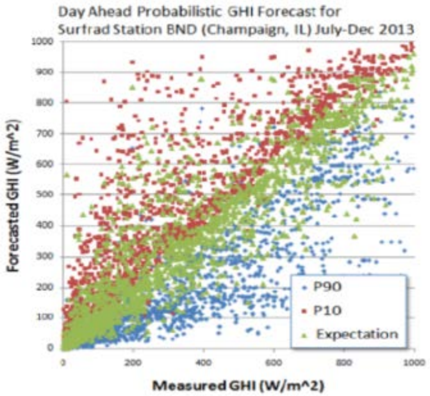
- South East Massachusetts Region
- 158 PV Plants, Total 10.4 MW
- 24-48 hr ahead forecasts at 3:30am EST daily
- **MAPE 5.0% (2014-5-1 to 2014-10-31)**



Metrics	24-48 hr, 2nd Yr
Correlation coef.	0.936
RMSE (MW)	79
NRMSE by capacity	0.0757
MaxAE (MW)	466
MAE (MW)	52.6
MAPE by capacity	0.0504
MBE (MW)	-7.2
KSIPer (%)	6.648
Std dev. (MW)	78.8
Skewness	0.539
Kurtosis	3.35
4RMQE (MW)	124
N4RMQE	0.119
95th percent (MW)	131

Probabilistic Forecasting

Built into the machine learning approach using "Weighted absolute deviations" type loss function as training target.

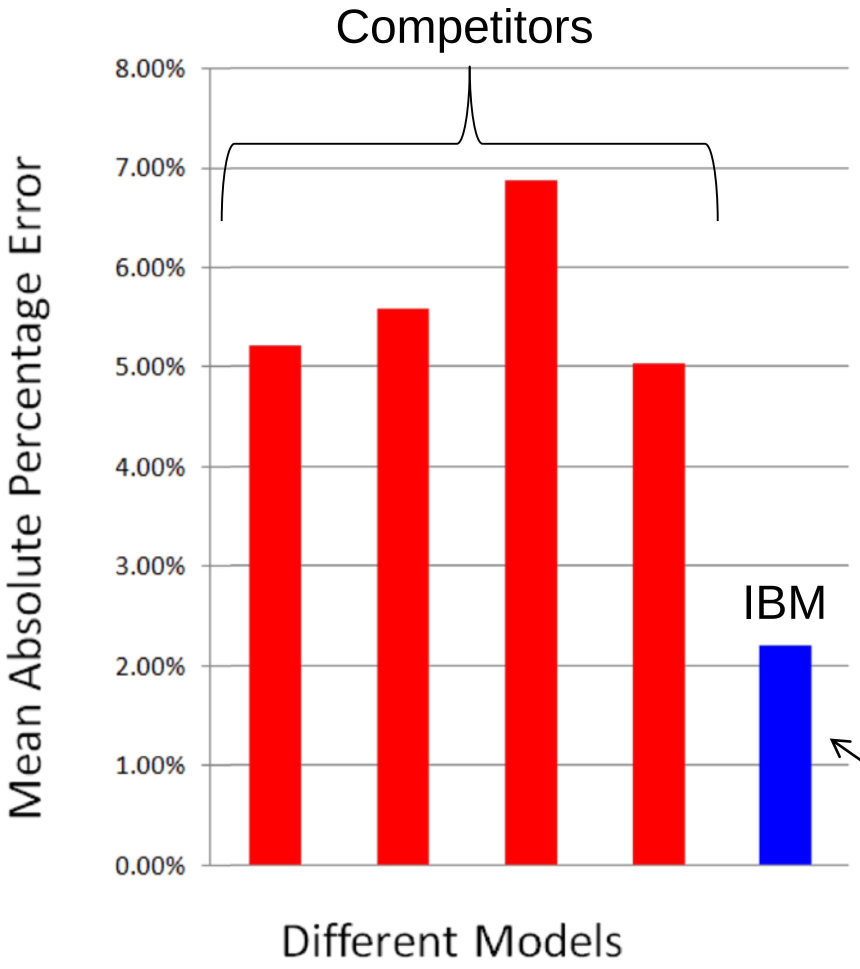


Quantile Reliability

Targeted Quantile	Actual Quantile
99%	99.4%
90%	92.2%
10%	9.9%
1%	0.8%

NOAA BND Surfrad Station
05/2013 to 01/2014

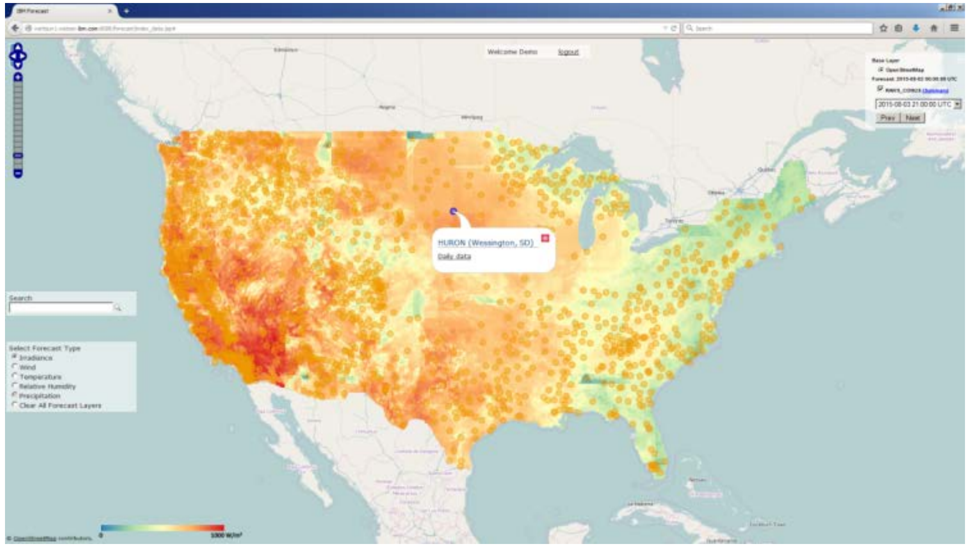
In Vendor trials we reduced forecast error by more than 30% over the next best forecasts



1.935 MW Fixed Array in Italy

30 % reduced error

We scaled the technology to continental wide forecasting and beyond.



- **> 35 % improved accuracy with respect to next best model at 1600 sites across the United States**
- **SMT provides gridded forecasts**
- **Continuously learns and improves**

Publically available web access to forecasts of 1600 sites across the US
http://server01.mmthub.com:9080/forecast/
User id: demo; Password: demo

Performance Metrics
2015-05-10 to 2015-05-24

IBM Blended GHI 0-48hr Ahead

SiteSet	# of Sites	MAE* (W/m^2)	MAPE(%)	RMSE (W/m^2)	NRMSE (%)	MBE (W/m^2)	MaxAE (W/m^2)	KSIPer (%)	pcoe
RAWS_CONUS	1641	135.54	13.55	192.99	19.29	19.84	668.06	16.62	0.78

NOAA NAM GHI 0-48hr Ahead

SiteSet	# of Sites	MAE* (W/m^2)	MAPE(%)	RMSE (W/m^2)	NRMSE (%)	MBE (W/m^2)	MaxAE (W/m^2)	KSIPer (%)	pcoe
RAWS_CONUS	1641	179.06	17.91	253.60	25.37	118.94	814.81	22.27	0.75

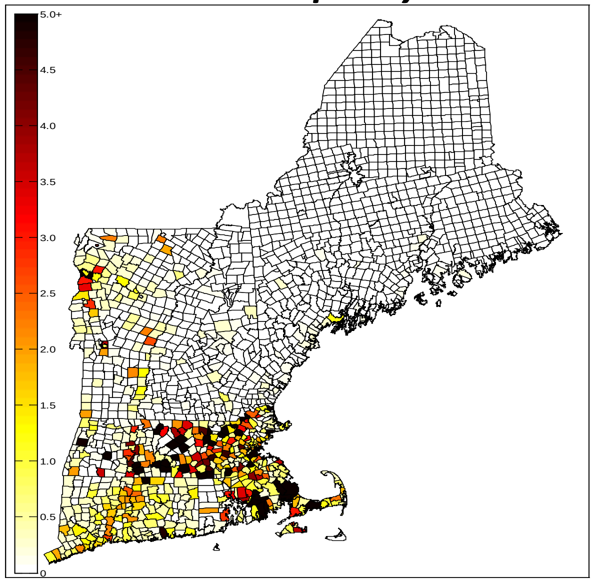
NOAA SREF GHI 0-48hr Ahead

SiteSet	# of Sites	MAE* (W/m^2)	MAPE(%)	RMSE (W/m^2)	NRMSE (%)	MBE (W/m^2)	MaxAE (W/m^2)	KSIPer (%)	pcoe
RAWS_CONUS	1641	188.44	18.84	262.97	26.30	129.56	837.83	23.64	0.74

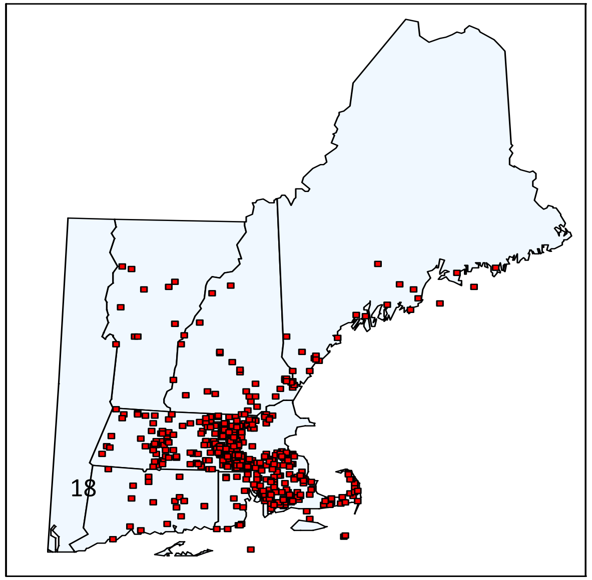
*Metrics Suites: Excluding night time values for solar. Hover mouse over the acronyms to see full definition.
IBM Blended GHI 0-48hr Ahead

Load forecasting for ISO-New England

Estimated PV capacity in ISO-NE



Measurement station in ISO-NE

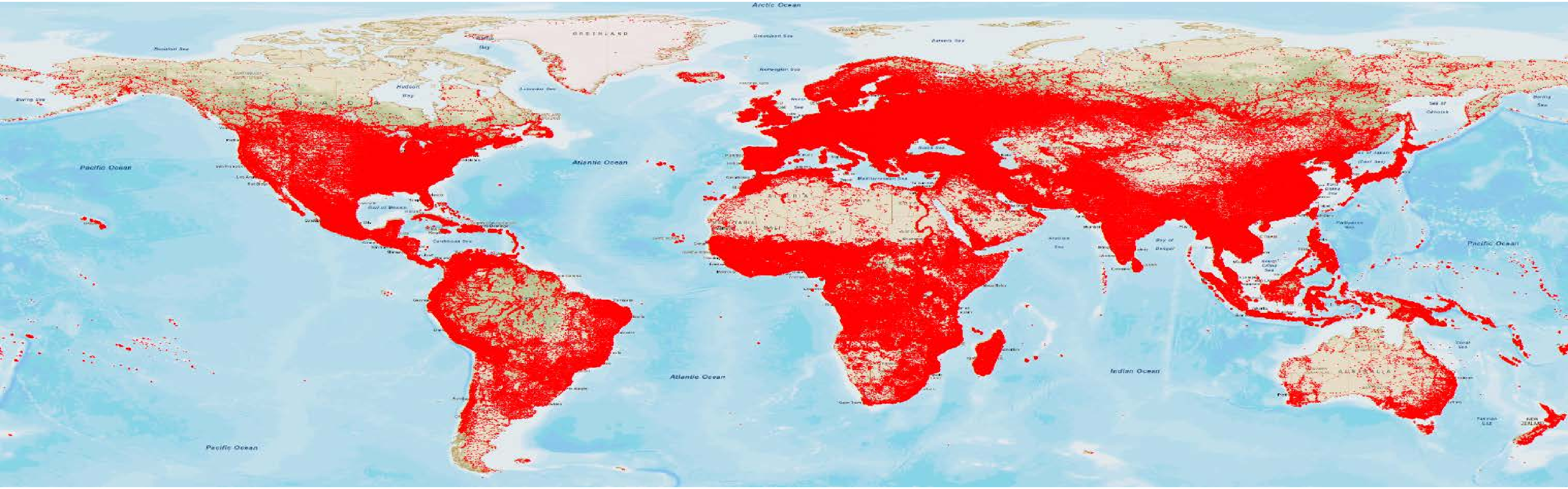


- IBM provides point forecasts for 665 sites in 9 dispatch zones
- Trains the model on 665 sites
- IBM scales forecast using estimated PV capacity for each dispatch zone
- ISO-NE feeds the forecasted data as an input into a neural network for load predictions

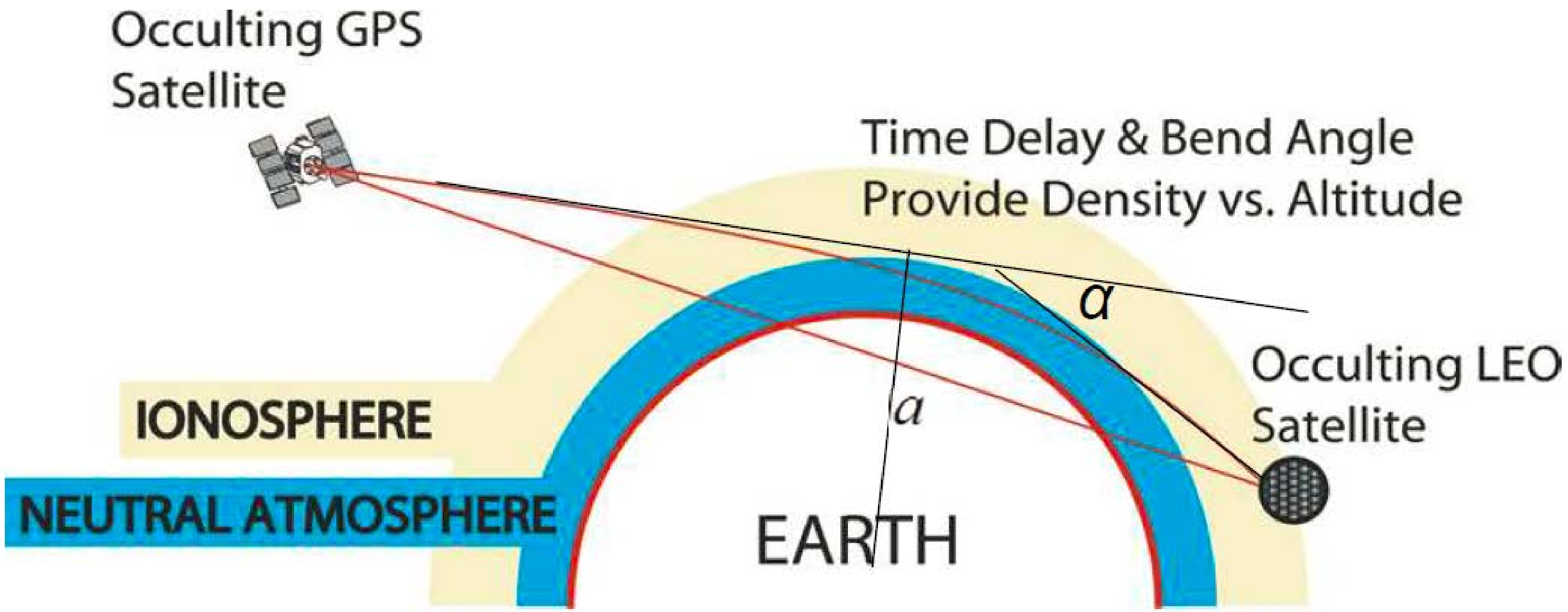
Commercialization of the technology?

- IBM acquired the world's largest private weather enterprise, the Weather Company (TWC), commonly known as the Weather Channel. TWC currently handles over 26 billion data requests per day, and push data to 40M+ cell phones
- TWC owns weather underground which will give unique access to new data
- Situation-dependent machine-learning model blending is the next generation technology to upgrade DiCast

IBM’s World Wide Weather Monitoring Network using Weather Underground



What is next? GPS-RO



- GPS Radio Occultation (GPS-RO) is an technique for measuring 3D weather variables (temperature, humidity etc) of the Earth’s atmosphere from space
- Explore opportunities to leverage GPS-RO for enhanced machine-learning
- Early results show drastic improvements

Summary

- The development of this technology improved solar forecasting accuracy by approximately 30%
- Technology is being commercialized by being integrated the IBM PAIRS geospatial big data platform
- Started to work with TWC
- The technology has been transferred to NREL to ensure is continue to serve the public good and the PV community