



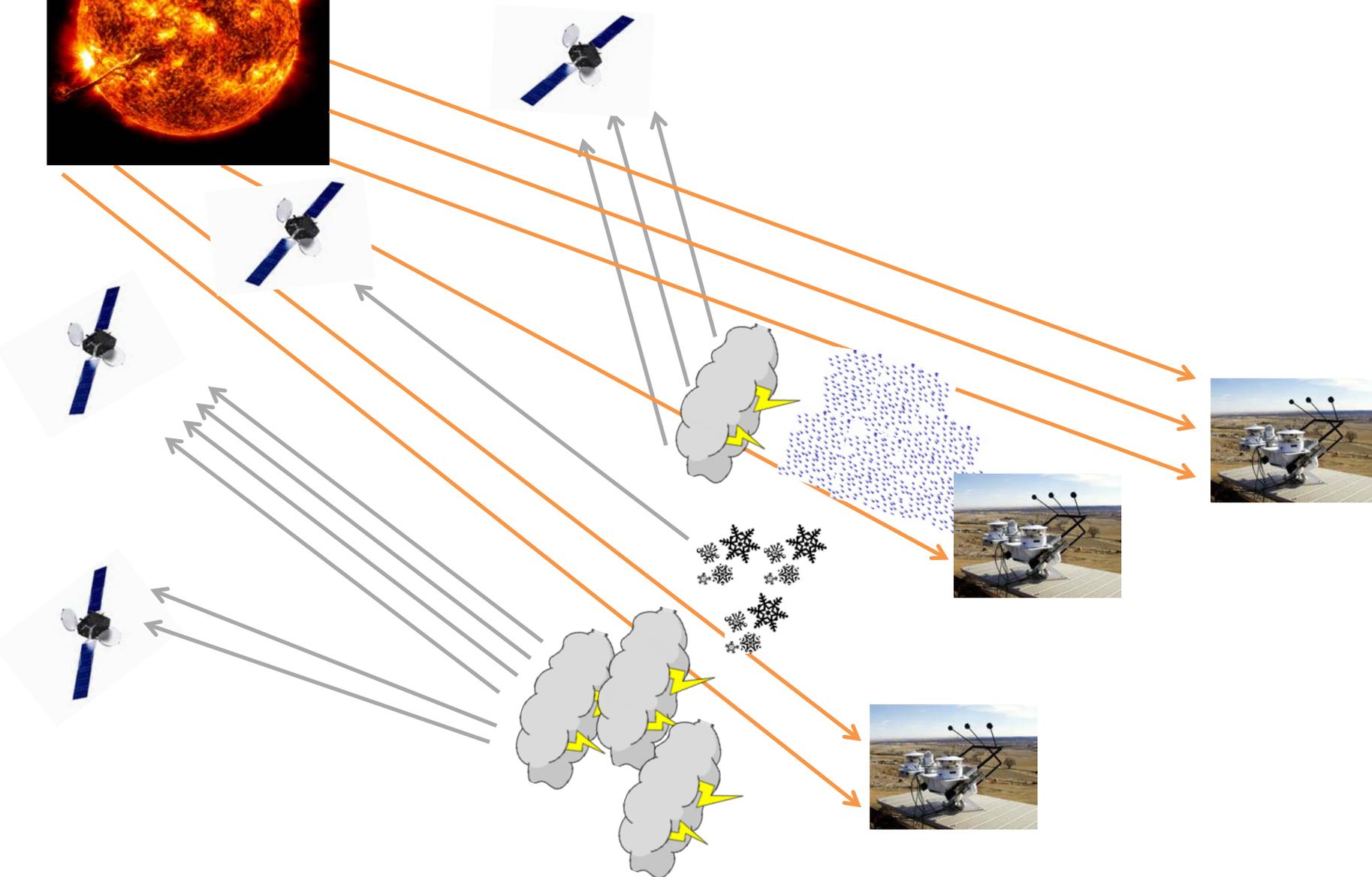
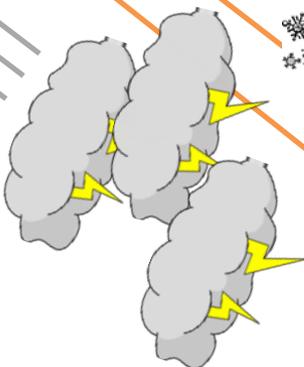
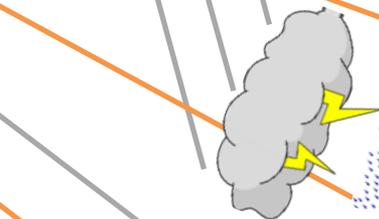
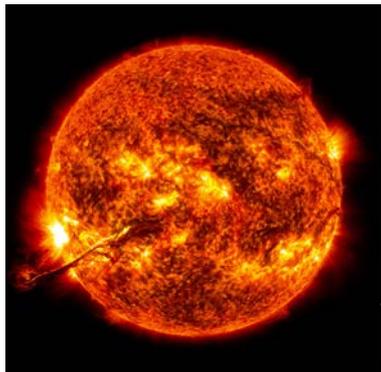
# **Multivariate linear regression technique for computing solar irradiance estimations using the SURFRAD and ISIS networks**

C. T. M. Clack, A. Alexander and A. E. MacDonald

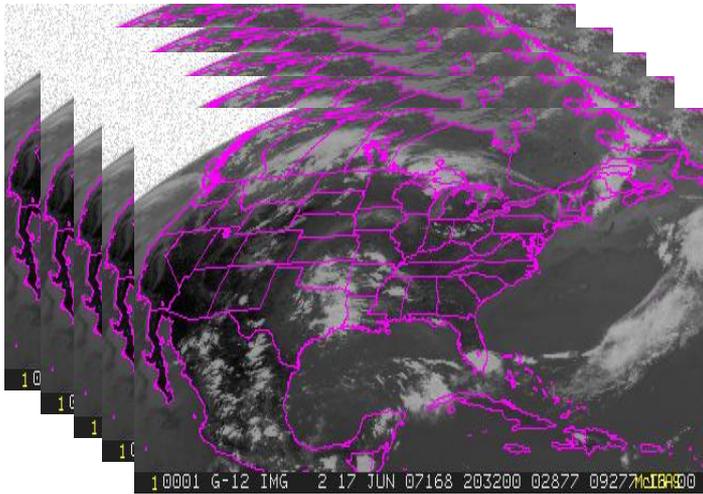
# Purpose

- The create *accurate* total, direct (normal and horizontal), and diffuse irradiance estimations.
- Leverage satellite, model hydrometeors, and high quality surface measurements to train the technique.
- Apply the technique over the CONUS domain to create an hourly data set of irradiance resource assessment (2006-2008 currently).

# Basic Technique I

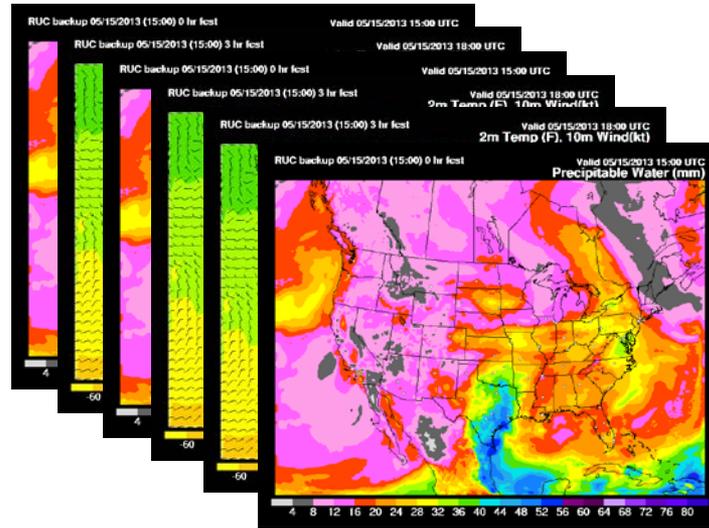


# Basic Technique II



5 Satellite Channels (where available)

+



6 NWP Hydrometeors (Rapid Update Cycle)

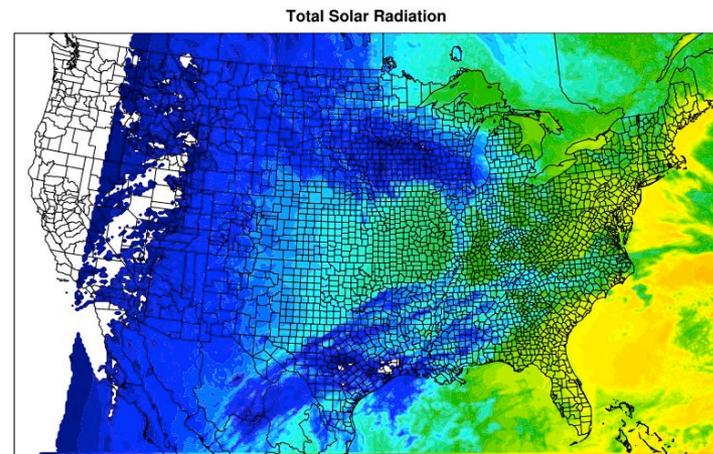
+

Calculated top of atmosphere  
Irradiance

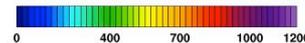
+

Calculated Zenith Angle

=



CONTOUR FROM 0 TO 1200 BY 25



# Linear Multivariate Multiple Regression

$$Y_{n \times p} = Z_{n \times (r+1)} \beta_{(r+1) \times p} + \epsilon_{n \times p},$$

$$E(\epsilon_{(i)}) = 0, \quad \text{Cov}(\epsilon_{(i)}, \epsilon_{(k)}) = \sigma_{ik} I, \quad i, k = 1, 2, \dots, p.$$

- We have  $p(=3)$  irradiance fields to calculate and  $n(=55258)$  observation of each field. The observations are taken from 10 sites (6 SURFRAD and 4 ISIS)
- The regressors ( $\beta$ ) are the satellite data (3 infrared channels, a visible channel, and a water vapor channel), the RUC Assimilation Model values for water within the column (snow, ice, etc...), the temperature from the model, the calculated top of atmosphere irradiance, and the zenith angle.
- The measurements are taken from 2006 – 2008, and averaged over the top of the hour (for 12 minutes) and matched up with the model data.
- The data is quality controlled, and all night-time measurements were removed.

# Linear Multivariate Multiple Regression

- Method relies on high quality ground measurements to train the regression procedure.



Validation Sites

$$\hat{\beta}_{(i)} = (Z'Z)^{-1}Z'Y_{(i)}$$

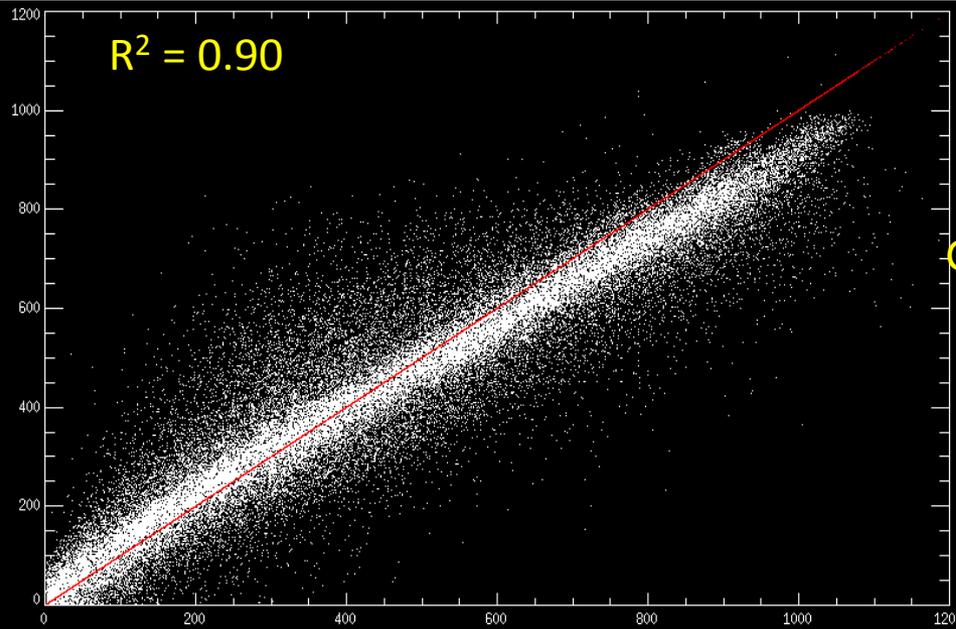
Can use numerous mathematical techniques to compute the coefficients. We do not go into that here... (I used SVD).

# Results I

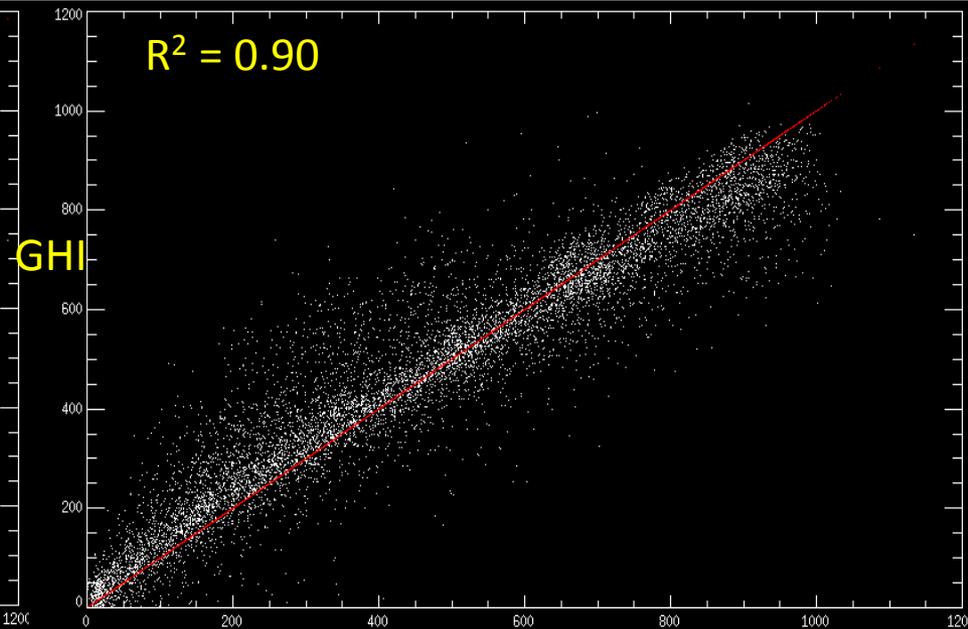
- The regression had differing success with total, direct, and diffuse

(Regression Correlation)	Satellite & Model	Satellite	Model
GHI	<b>0.94378774</b>	0.93182838	0.91139053
DNI	<b>0.78874645</b>	0.72949635	0.54405547
DHI	<b>0.83337251</b>	0.81288982	0.69247811

Regression Data Points

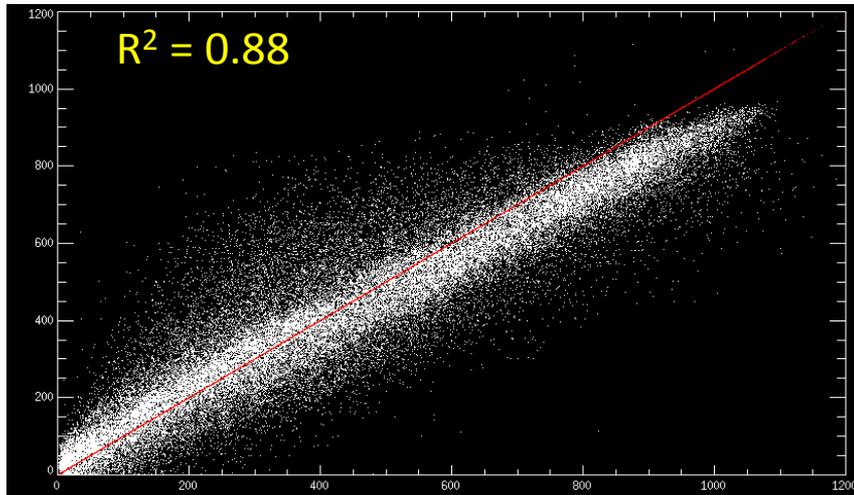


Verification Data Points

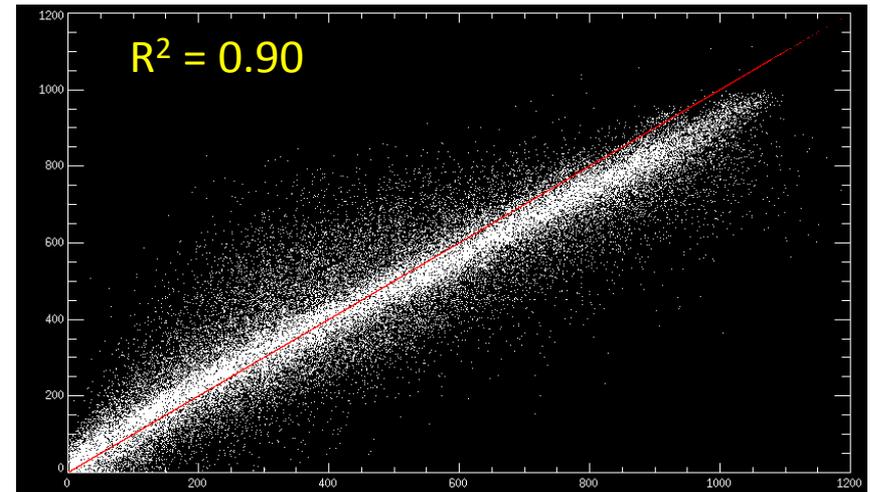


# Results II - GHI

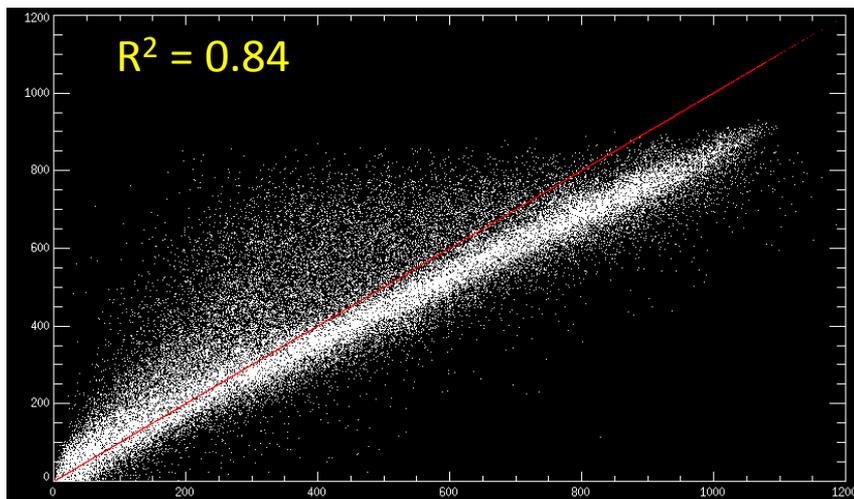
Satellite Data Only



Both Data Together

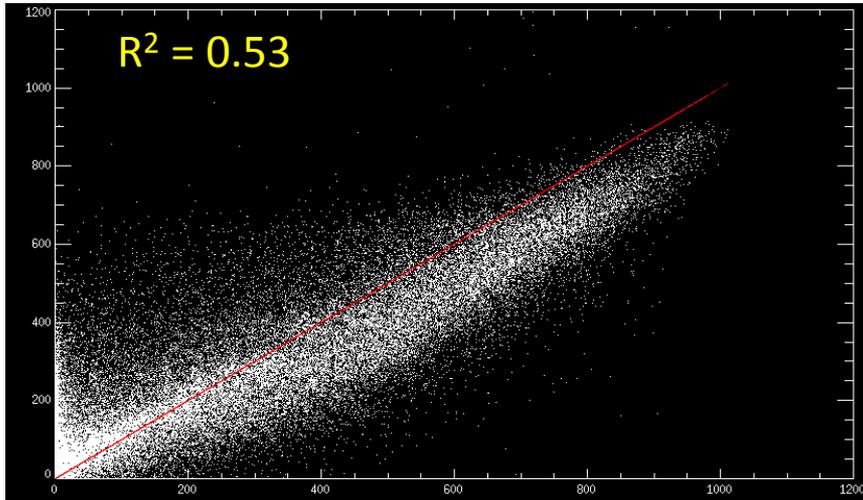


Hydrometeor Data Only

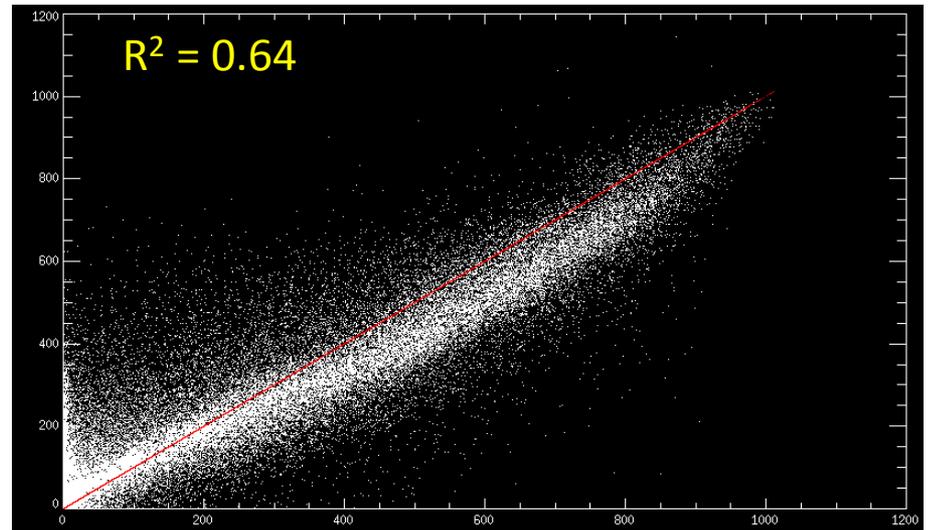


# Results III - DNI

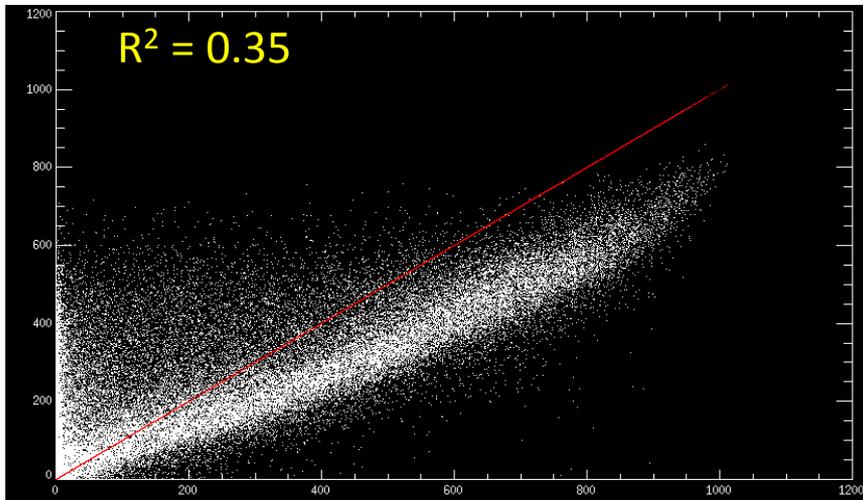
Satellite Data Only



Both Data Together

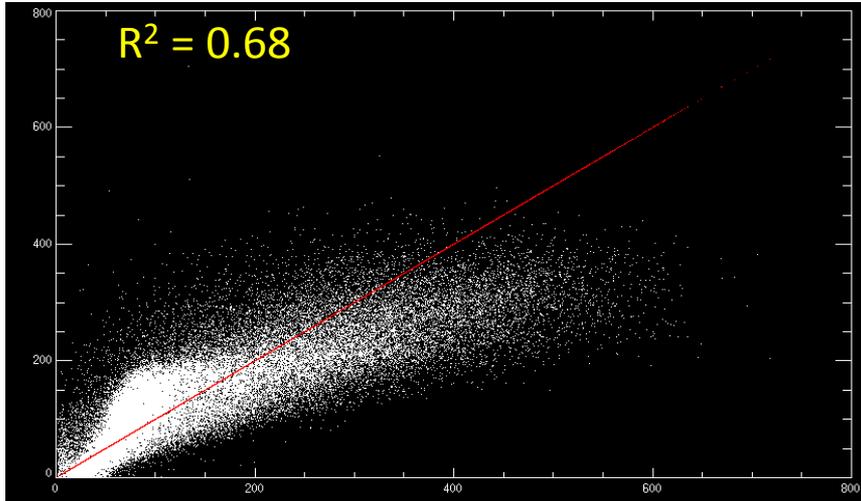


Hydrometeor Data Only

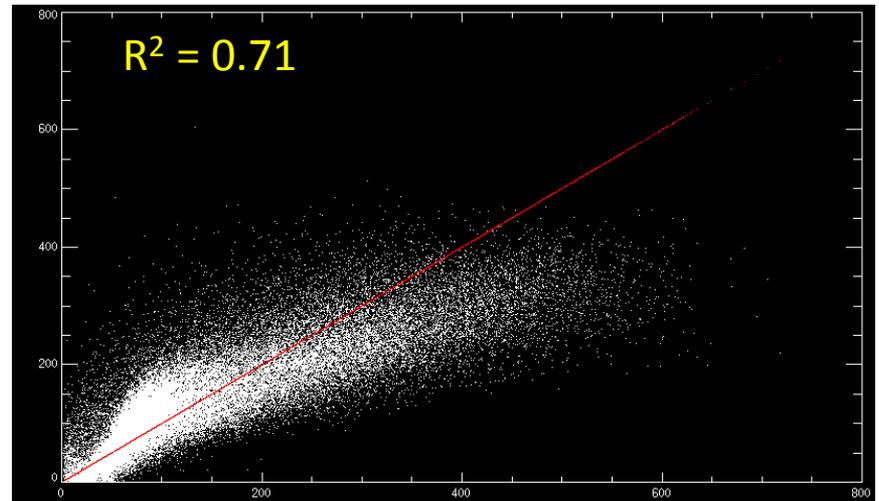


# Results III - DHI

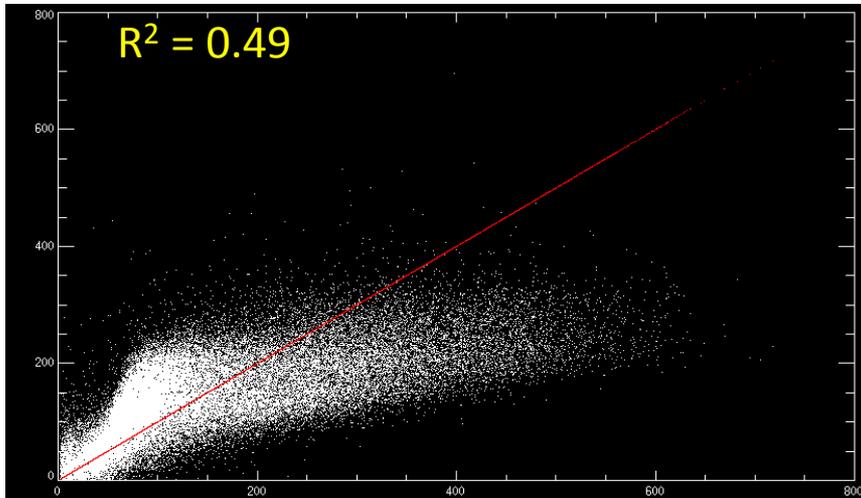
Satellite Data Only



Both Data Together

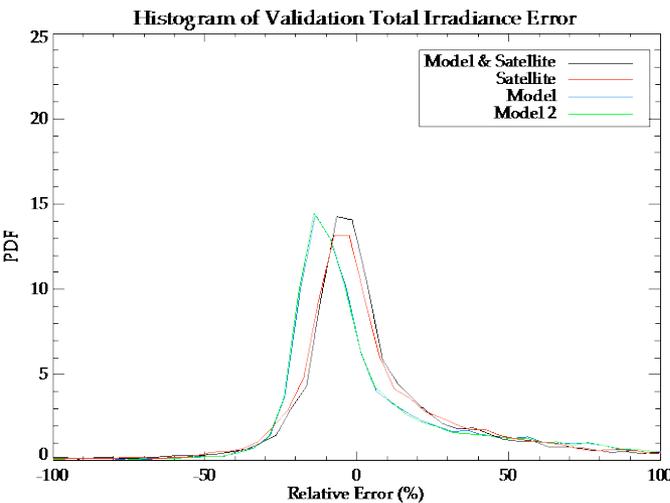
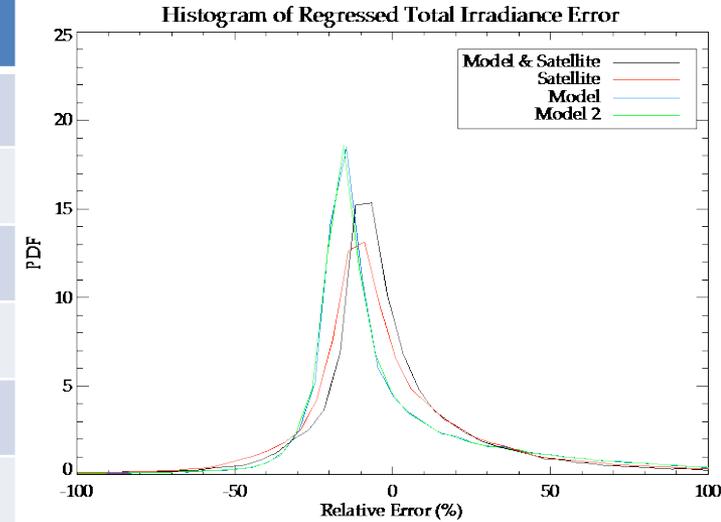


Hydrometeor Data Only



# Results IV – Metrics (GHI)

[GHI - Regression]	Satellite & Model	Satellite	Model
Bias (% & W/m <sup>2</sup> )	<b>-2.7% / -11.91</b>	-3.24% / -14.95	-4.20% / -18.53
RMSE (W/m <sup>2</sup> )	<b>89.37</b>	98.40	112.13
NRMSE (%)	<b>7.26</b>	7.99	9.11
CVRMSE (%)	<b>20.27</b>	22.32	25.44
STD (W/m <sup>2</sup> )	<b>88.57</b>	97.36	110.60
CV (%)	<b>20.09</b>	22.08	25.09

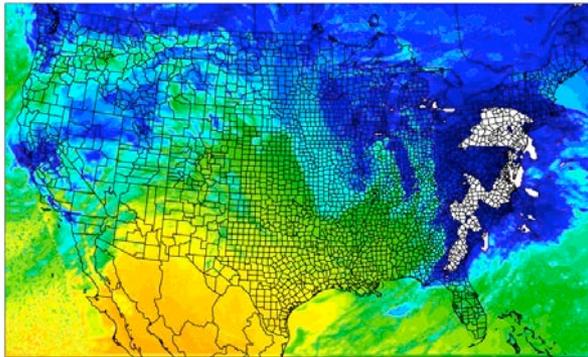


[GHI - Verification]	Satellite & Model	Satellite	Model
Bias (% & W/m <sup>2</sup> )	<b>+2.46% / +11.18</b>	+2.70% / +12.27	<b>+1.21% / +5.50</b>
RMSE (W/m <sup>2</sup> )	<b>87.42</b>	92.12	109.62
NRMSE (%)	<b>7.77</b>	8.14	9.69
CVRMSE (%)	<b>19.23</b>	20.26	24.11
STD (W/m <sup>2</sup> )	<b>86.71</b>	91.30	109.49
CV (%)	<b>19.07</b>	20.08	24.08

# Application I

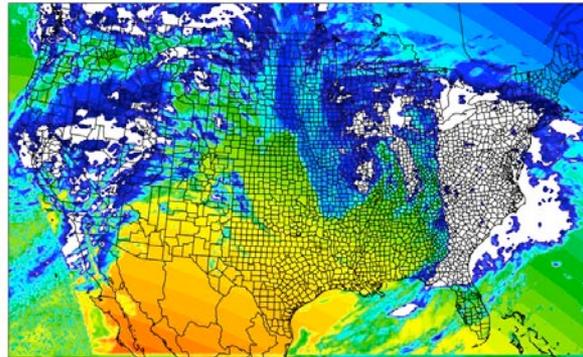
- Once the regression has been performed, and we have verified that it is robust, we apply the procedure throughout the Contiguous USA domain provided by the RUC and GOES data for 2006-2008.
- The GHI, DNI, and DHI are assimilation values (not forecasts).
- The irradiance values are utilized to provide an estimated PV output for each hour of 2006-2008 at each RUC cell over the CONUS. For this we do 96 regressions!

Total Solar Radiation



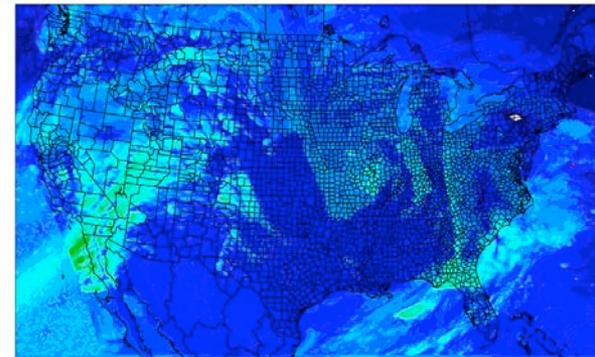
CONTOUR FROM 0 TO 1200 BY 25

Direct Normal Irradiance (DNI)



CONTOUR FROM 0 TO 1200 BY 25

Diffuse Horizontal Irradiance (DHI)



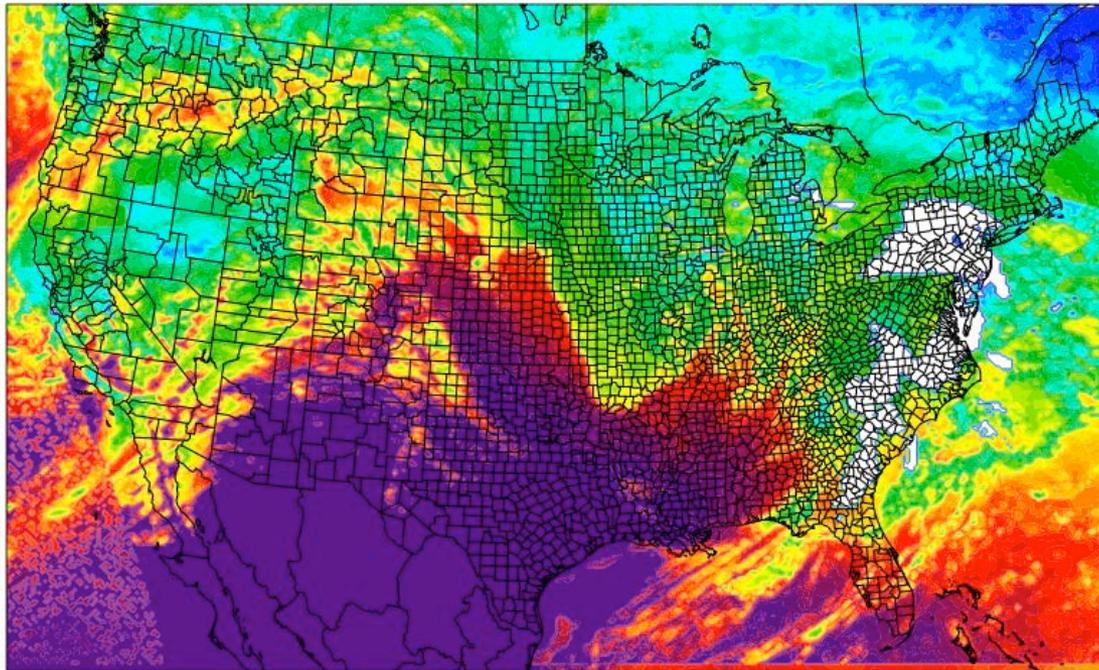
CONTOUR FROM 0 TO 1200 BY 25



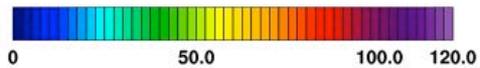
Time of Image: Jan 2 2000 UT 2006

# Application II

PV Generation



CONTOUR FROM 0 TO 1.2 BY .025

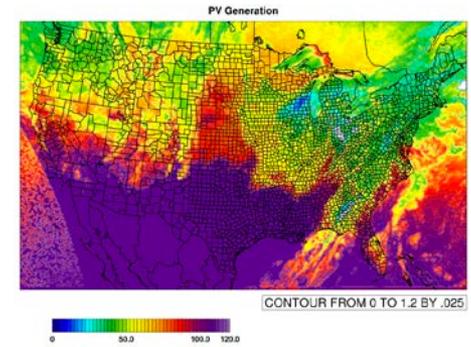
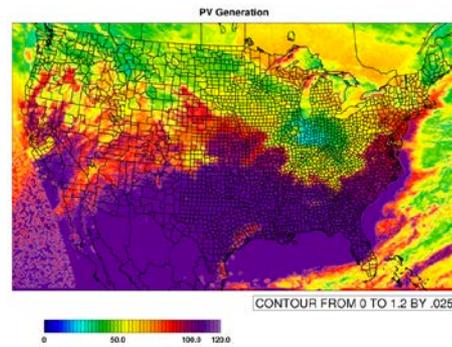
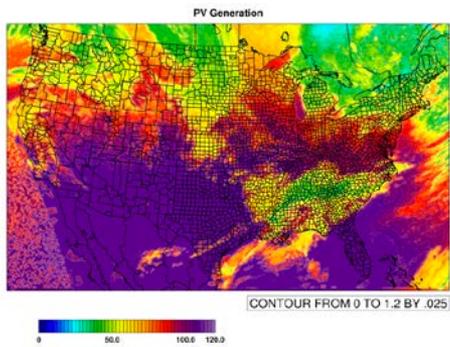
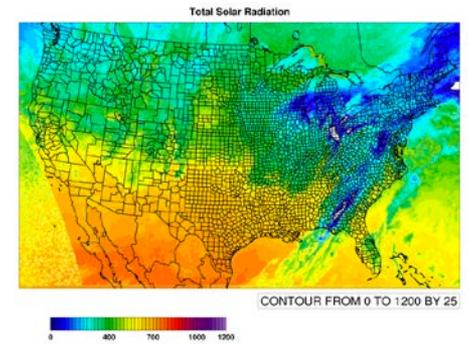
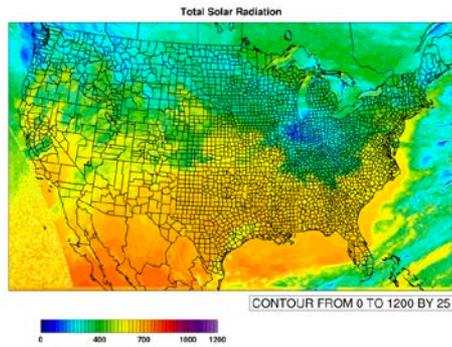
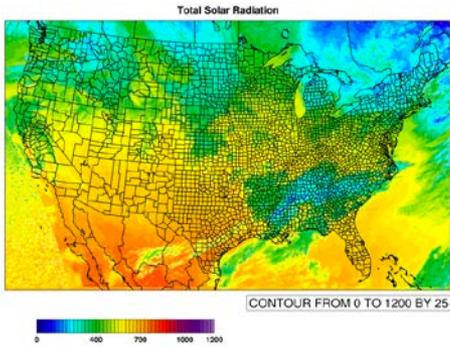


Time of Image: Jan 2 2000 UT 2006

# Application III - Future

- The results are promising, even though an older, lower resolution model was utilized for the regression.
- In the near future, we are going to do more analysis of the results at the individual sites, try to validate at more diverse locations, and improve the technique.
- After that, we envisage applying the procedure to the HRRR model (2012), with more high quality ground measurements.
- The procedure can be extended to forecast hours (utilizing a training set of approximately a year), and we hope to do this alongside Kathy Lantz and Joseph Michalsky.
- It is a technique developed to create a high quality hourly dataset of PV power output for use within the ESRL Optimization Code.

# Application IV - Samples



2006, February 6, 1900 UT

2007, February 6, 1900 UT

2008, February 6, 1900 UT

THANK YOU! ANY QUESTIONS?