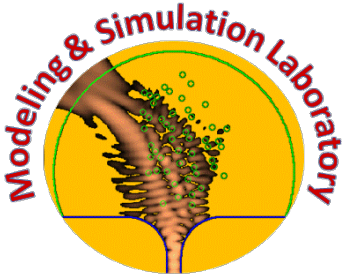


Develop Space Weather Forecast Strategy with SWFT

Chunming Wang

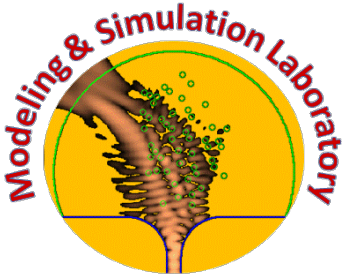
Modeling & Simulation Laboratory
Department of Mathematics
University of Southern California



Learning to Forecast Space Weather with Existing Models and Data

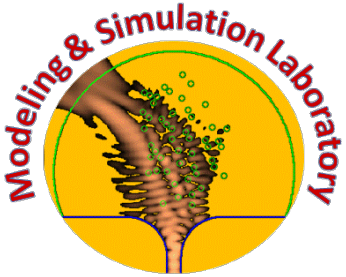
- Capability for forecast can be developed with imperfect model and incomplete observation.
  - Measures of success in forecasting can provide insight for model improvements.
- Focus on forecast draws attention to a different set of issues than model development.
  - What, when and how to forecast, how to measure success?
- Success in forecast provides validation for understanding.

Development of Forecast Strategy and Model Progress Concurrently



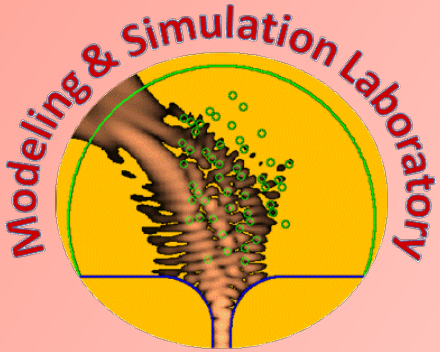
Physical Principles and Empirical Evidence are Crucial for Forecasting

- Laws of physics provide links between solar and interplanetary magnetic field (IMF) events to thermosphere and ionosphere variations.
  - Available data are insufficient for complete model calibration.
  - Models are most reliable for quiet and regular conditions.
- Historical evidence exists to connect solar and IMF anomalies to ionosphere disturbances.
- Successful forecast must leverage all available information.
  - Select features of ionosphere most susceptible to solar and IMF.
  - Combine empirical data and models to forecast extreme events.

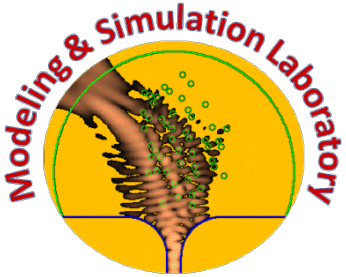


A Test-bed Can be an Incubator for Space Weather Forecasting Strategies

- A database of high quality historical space weather observations enables training and validation of forecast strategies.
- A collection of empirical and first principle physics models, as well as, data analysis tools facilitates development of new forecast approaches.
- A powerful computational and data storage platform provides capability to analyze a large scale of historical data and leverage previously computed intermediate results.



Forecasting Solar Wind



Forecasting Solar-Wind Using Past Observations

- Solar wind velocity is a key parameter used by thermosphere-ionosphere models.
- When these models are used to forecast conditions in the ionosphere, forecasted values of solar wind speed are required.
- Multi-linear regression method can be used to derive a forecast value for the solar wind by examining the correlation between past values of space weather observation to future one.

Multi-linear Regression Requires Computation of Covariance

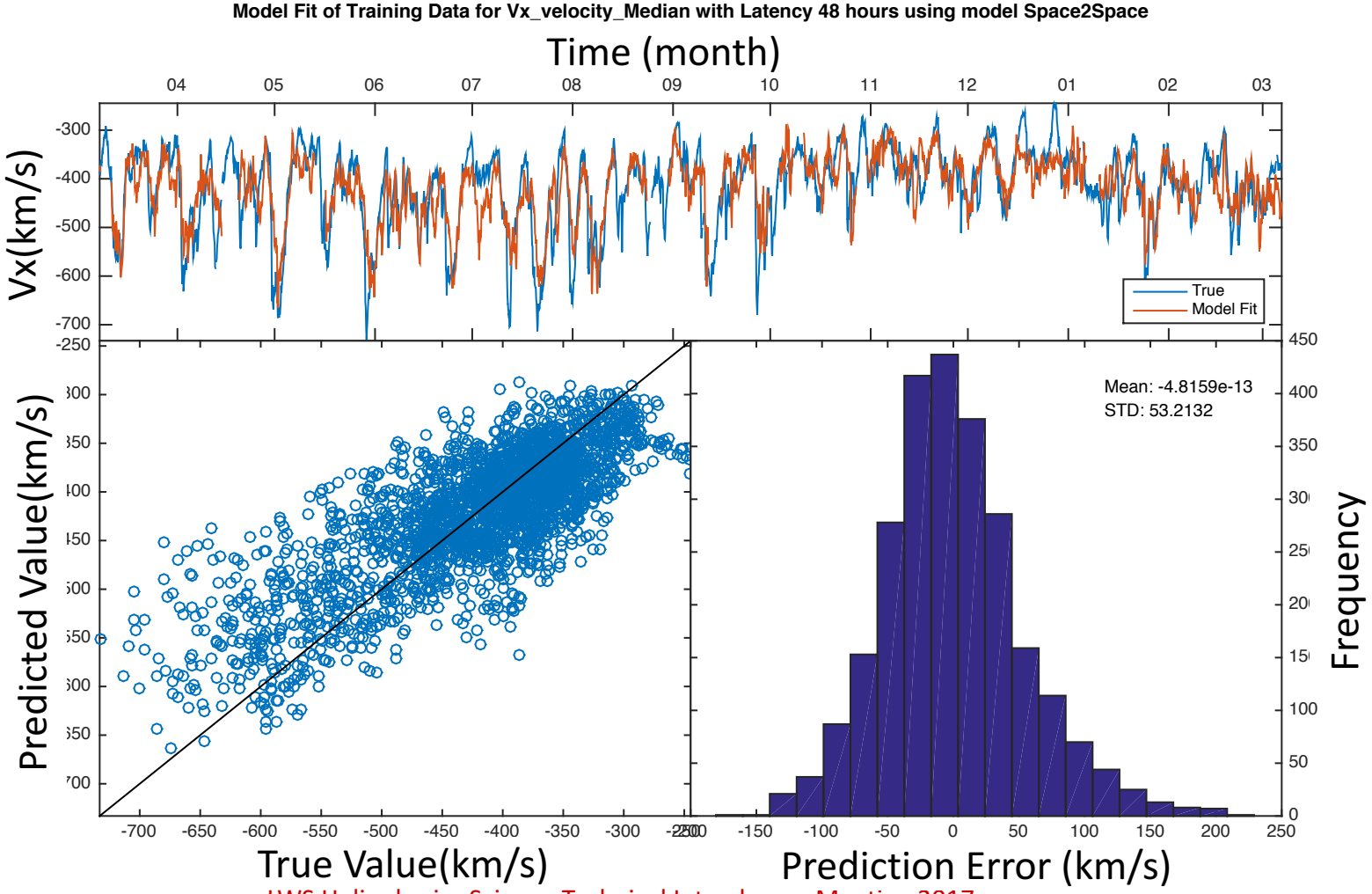
- Regression consists of finding optimal coefficiets for a linear model of the form

$$v_X(t) = T\begin{pmatrix} \vec{u}(t-d) \\ \vdots \\ \vec{u}(t-d-m) \end{pmatrix}, \vec{u}(t) = \begin{pmatrix} v_X(t) \\ \vdots \\ B_X(t) \\ \vdots \\ n_{SunSPot}(t) \end{pmatrix}.$$

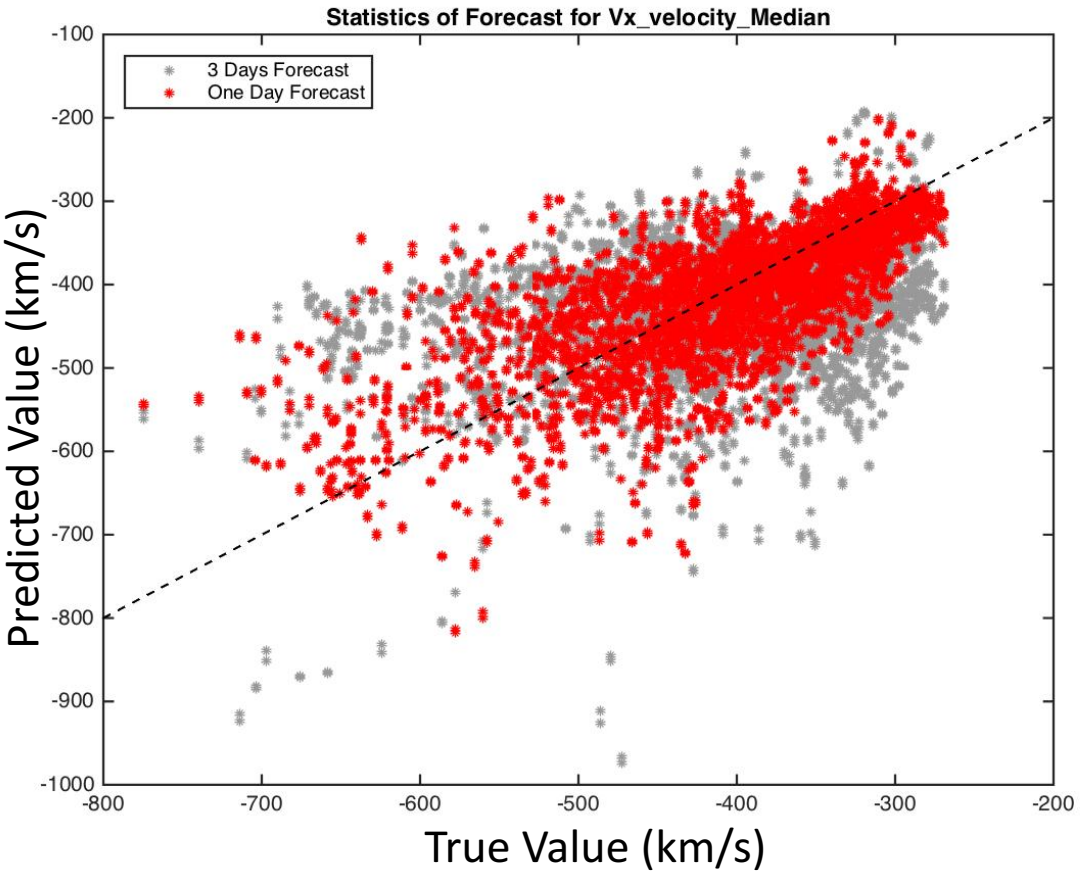
- Optimization for T consists of solving the least square minimization problem using historic data

$$\min_T \sum_{k=1}^{N} \left| v_X(k) - T\begin{pmatrix} \vec{u}(k-d) \\ \vdots \\ \vec{u}(k-d-m) \end{pmatrix} \right|^2 .$$

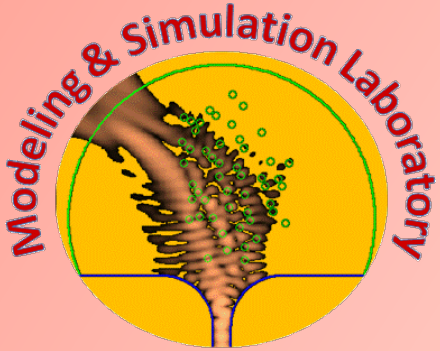
Data for Model Training, Verification and Validation Can be Reviewed



User Defined New Forecast Model Can be Tested on SWFT

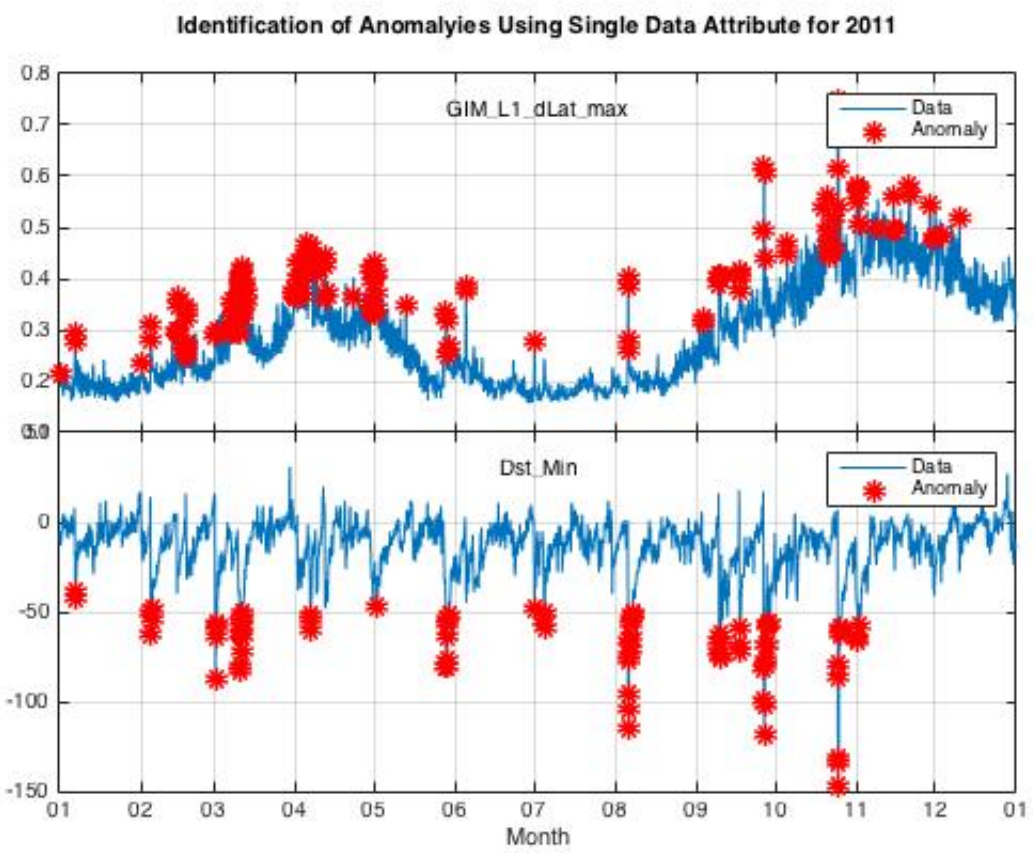


- SWFT script can be used to evaluate a new forecast strategy using extended time periods for evaluation.
- Scoring of the new forecast strategy can be made using repositories of model output.



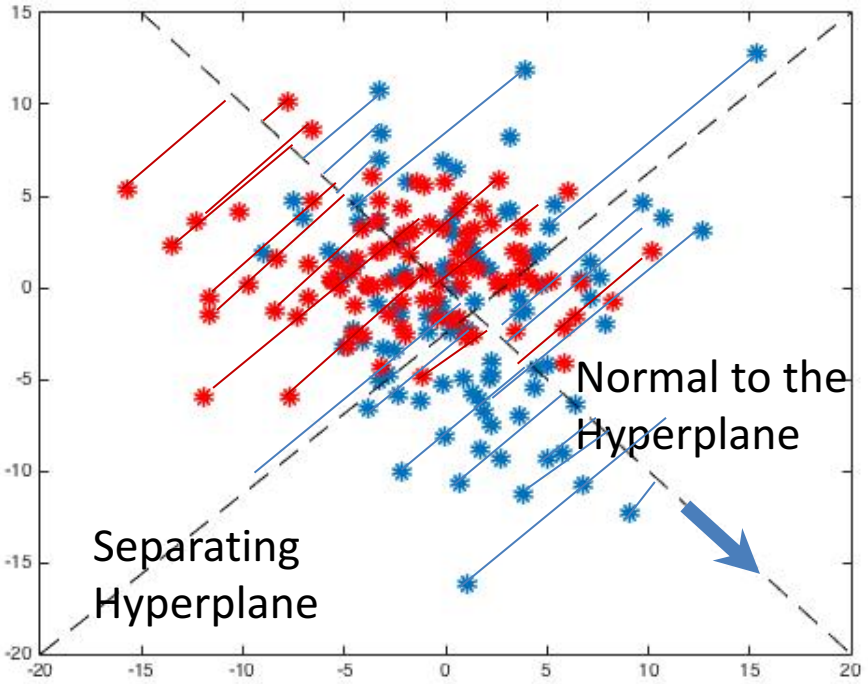
Forecasting Ionosphere Anomalies

Anomalies Can be Defined Using a Single or Multiple Metrics



- Instead of forecasting "observations", *anomalies* can be defined and forecast.
- Machine learning algorithms exist for making categorical forecasts.

Logistic Regression Algorithms Helps Classify Data



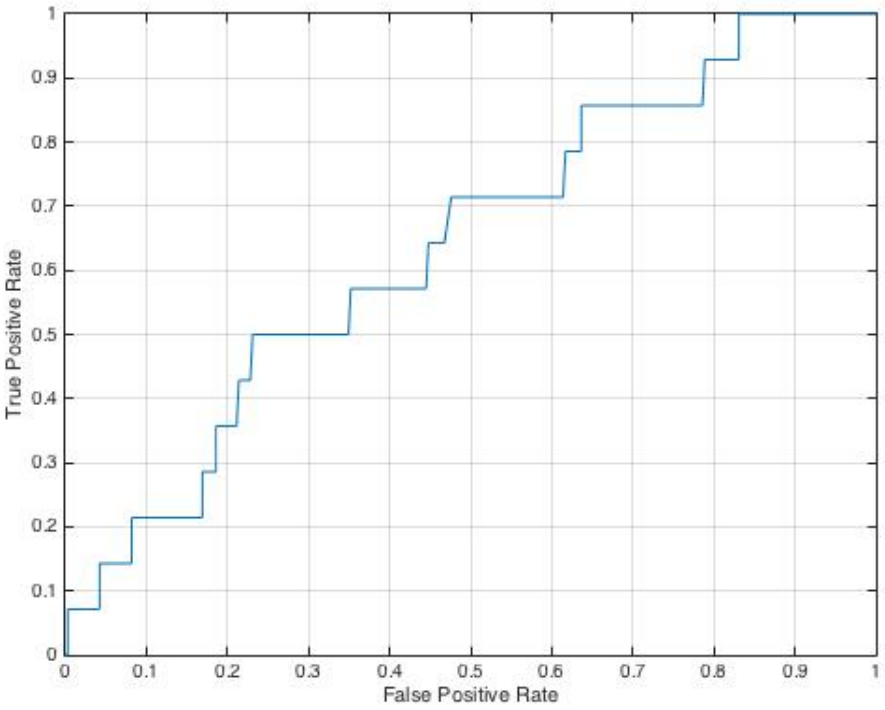
Data Score

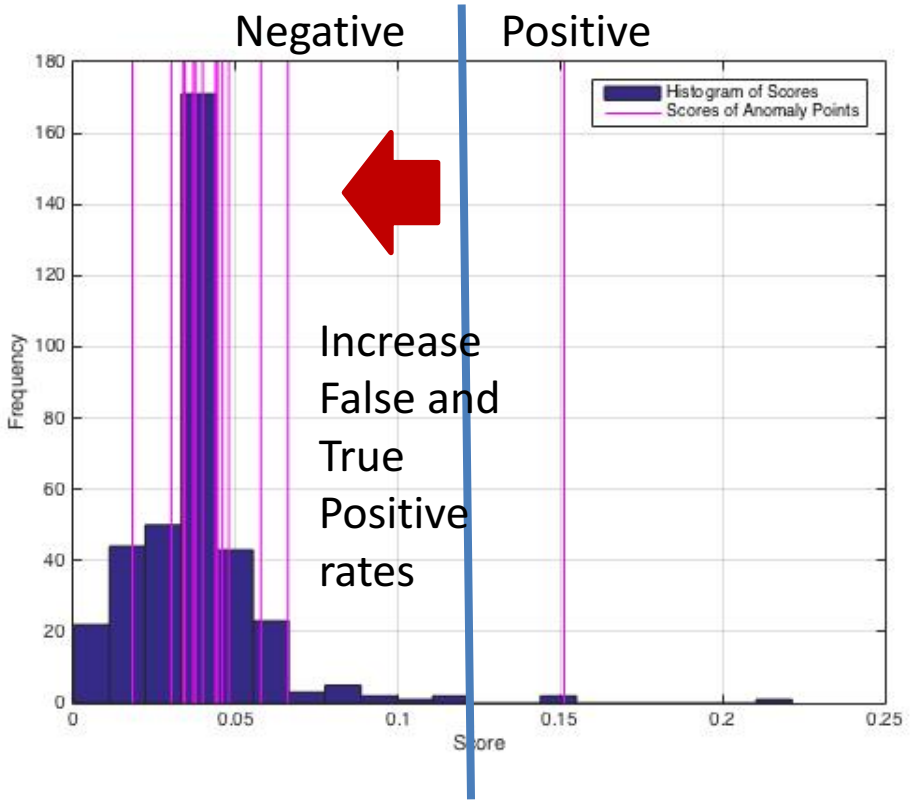
$$p(\vec{x}) = \frac{e^{\vec{a}^T\vec{x}+b}}{1+e^{\vec{a}^T\vec{x}+b}}$$

- Seeking hyperplane in multiple-dimensional space of forecast variables to best separate two categories of data.
- Find vector $\vec{a}$ and offset b to maximize the likelihood function

$$\prod_{k=1}^{N}\left(p(\vec{x}_k)\right)^{y_k}\left(1-p(\vec{x}_k)\right)^{1-y_k}$$

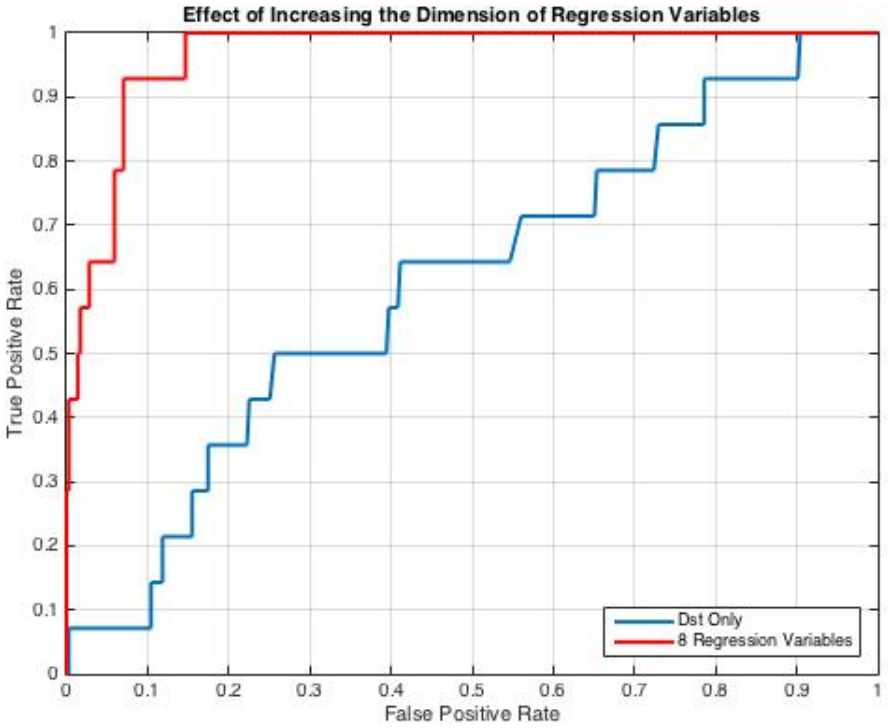
Success in Regression Training is Measured by True/False Positive Rates

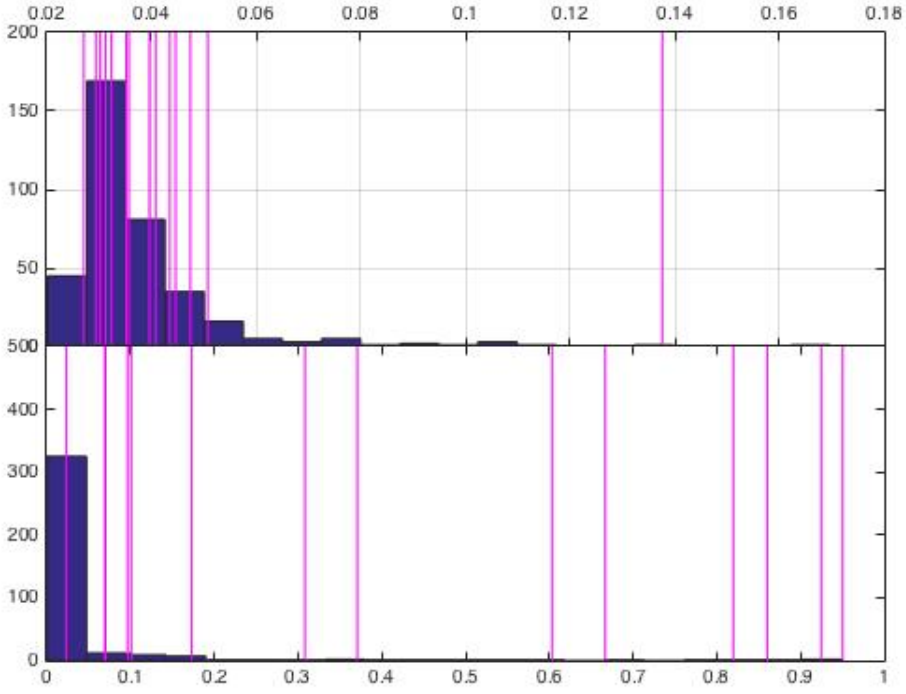




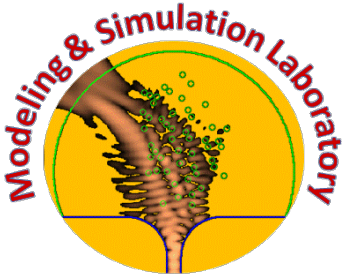
- ROC curve represents the trade-off between true and false positive rates.

Increasing the Dimension of Regression Variables Improves Fit





- Adding regression variables in general improves classification of training data but not necessarily forecast.



Conclusion

- SWFT provides a community platform for the development of forecast strategies.
  - SWFT allows user to define anomalies when quantitative forecasts are challenging
  - Acquire new data, indices, proxies and use them to test forecast strategies
  - Incorporate repositories of model output
  - Experiment with different machine learning approaches.
- Machine learning as an effective way to lower the barrier to entry for space weather forecast.
  - Two current mathematics Ph.D students are involved in developing machine learning based forecast approaches.