# A Humidity-Driven Prediction System for Influenza Outbreaks

Heidar Thor Thrastarson
Joint Institute For Regional Earth System Science and Engineering, University of California Los Angeles
NASA Jet Propulsion Laboratory, California Institute of Technology

Joao Teixeira
NASA Jet Propulsion Laboratory, California Institute of Technology

Enoch Yeo
Harvey Mudd College

NASA Sounder Science Team Meeting, September, 2016

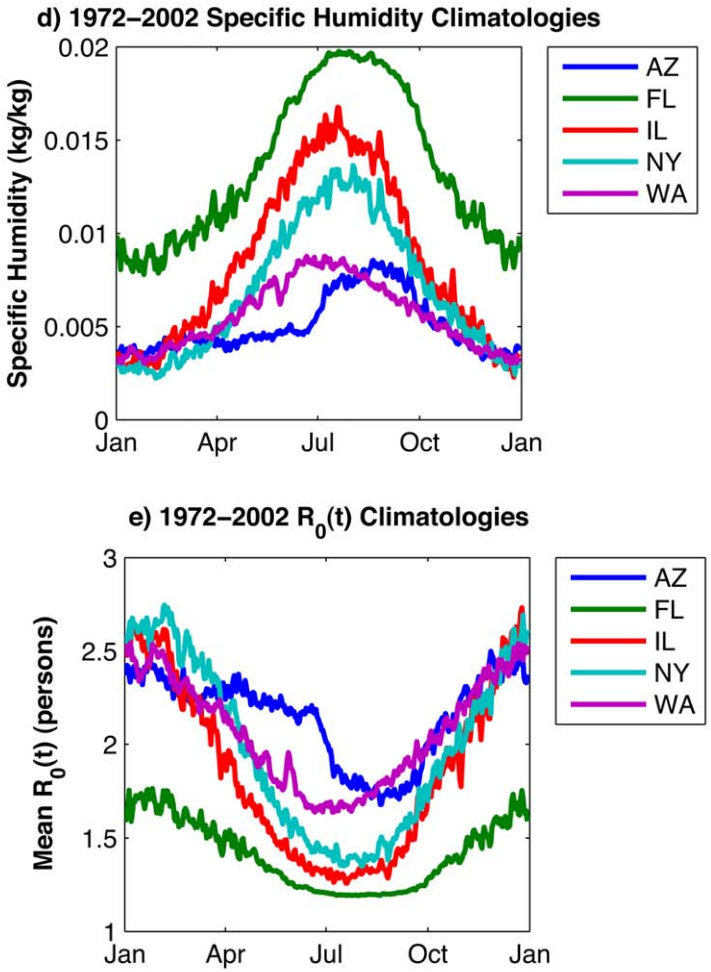
# Background - Motivation

- Seasonal influenza epidemics are a major public health concern, resulting in millions of cases of severe illness and 250,000 to 500,000 deaths worldwide each year as well as a large economic toll.
- In temperate regions influenza incidence generally has pronounced peaks in the winter, but the specific timing, magnitude and duration of individual local outbreaks in any given year are variable and not well explained.
- If the timing and intensity of seasonal influenza outbreaks can be forecast, this would be of great value for public health response efforts.
- It could guide both mitigation and response efforts, planning and stockpiling of vaccines and drugs, management of hospital resources, focusing of efforts to areas with more urgent need.
- Recent studies have highlighted a role of absolute (or specific) humidity conditions as a leading explanation for the seasonal behavior of influenza outbreaks.

© 2016. All rights reserved.

# Absolute Humidity as a Driver of Influenza Seasonality

- Lab experiments: Absolute humidity strongly modulates the airborne survival and transmission of the influenza virus.
- Climate & Influenza data records: Increased wintertime influenza-related mortality in the US associated with anomalously low absolute humidity levels
- Humidity-driven epidemiological models yielding promising results
- The reason for the humidity–influenza relationship is not well established but drying of mucous membranes and/or humidity effects on droplet sizes and travel range could play a role.




Shaman et al. (2009), Shaman et al. (2010)

# Model of Influenza Outbreaks

- We have developed and implemented a SIRS (Susceptible-Infectious-Recovered-Susceptible) type numerical prediction system that is driven by specific humidity to predict influenza outbreaks.

- Two coupled first-order Ordinary Differential Equations numerically solved for the number of susceptible and infected/infectious people in a given population.

- Rate of infections and recoveries parameterized in terms of average length of immunity and mean infectious period and assumed to have a simple dependence on input specific humidity.

# SIRS Model

$$\frac{dS}{dt} = \frac{N - I - S}{L} - \frac{\beta I S}{N} - \alpha$$

$$\frac{dI}{dt} = \frac{\beta I S}{N} - \frac{I}{D} + \alpha$$

Specific humidity

$$\beta = \frac{R_0}{D} = \frac{1}{D}[R_{0min} + (R_{0max} - R_{0min})\mathrm{e}^{aq}]$$

| | |
|---|---|
| $N$ | Population size |
| $S$ | Susceptible persons |
| $I$ | Infectious (= infected) persons |
| $L$ | Average immunity duration |
| $D$ | Mean infectious period |
| $\alpha$ | Rate of (travel-related) import of virus into model domain |
| $\beta$ | Contact rate |
| $R_0$ | (Daily) basic reproductive number |
| $a$ | (Negative) coefficient in contact rate exponential |
| $q$ | Specific humidity |

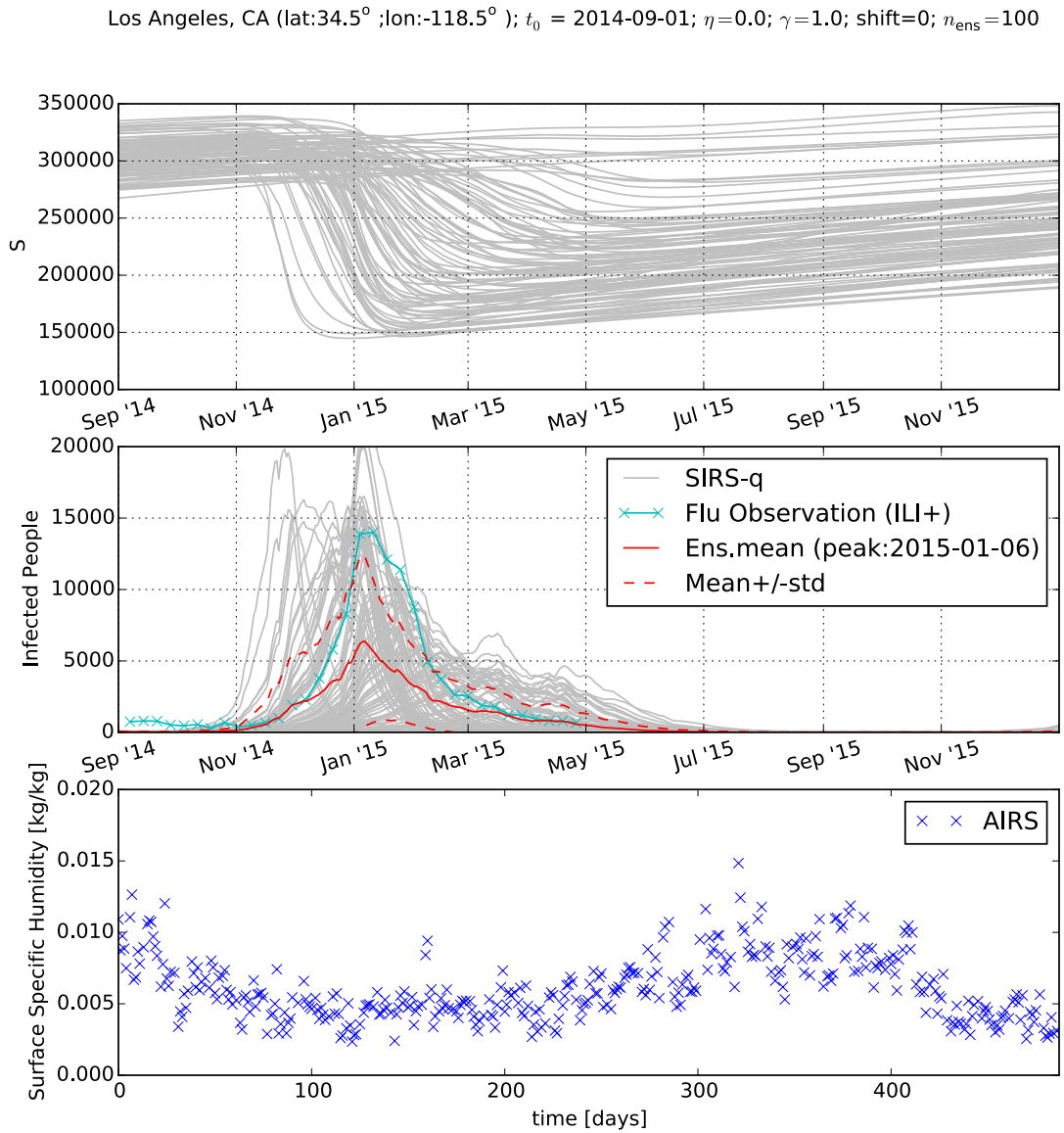
# Humidity Data

- For the humidity, we use satellite data from the NASA AIRS instrument.
- The model is integrated for a specified time period, with humidity data updated each day (when data is available) and the ODE's solved in the intervals between, on a specified time grid.
- We use surface water mass mixing ratio (close to specific humidity), data for the ascending mode of the satellite, so daytime values, level 3 data (version 6), averaged onto 1x1 degree latitude/longitude uniform grid cells.
- We model several US cities and for each of them we take the single grid cell that contains the largest part of that city (also considered state/ regional averages).
- Also have a version with NCEP humidity forecasts incorporated (but still always using latest observed AIRS values to initialize)

# Influenza Incidence Data

- Near real-time estimates of influenza infection rates based on online-search queries are available from Google Flu Trends (GFT)
- For HHS (Health and Human Services) regions, the Center for Disease Control (CDC) provides weekly surveillance records for the proportion of doctor's visits for influenza-like illness (ILI)
- They also provide lab virology results for the percentage of influenza positive samples
- We combine these to get an estimate of influenza incidence to compare with and incorporate in the model (ILI+)

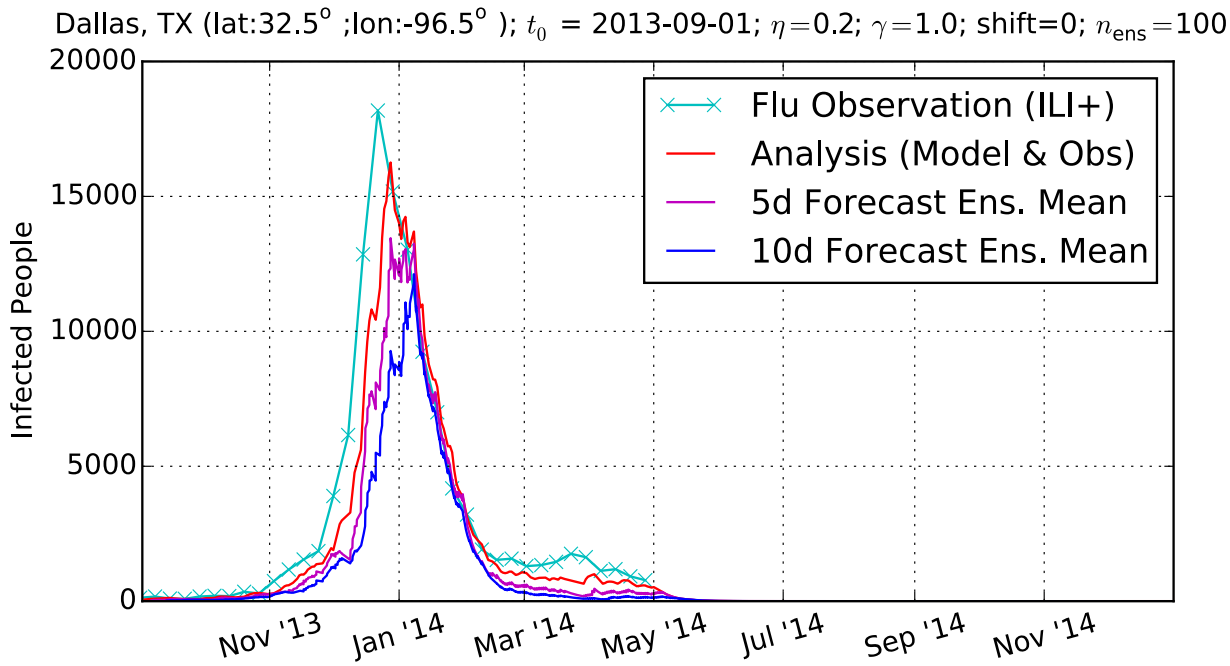
# Illustrative Example – Retrospective Simulation




- Results from the SIRS model for Los Angeles 2014-15.
- 'Raw' simulations (without data assimilation), but with hindsight knowledge of observed humidity
- Ensembles with parameters from normal distributions
- The gray lines show the curves for Susceptible and Infected people from a 100 member ensemble, with the ensemble mean and stdev in red. The ILI+ data is in cyan.
- Humidity time series for the corresponding periods are shown (blue crosses) in the bottom panel.

# Data Assimilation and Forecast Ranges

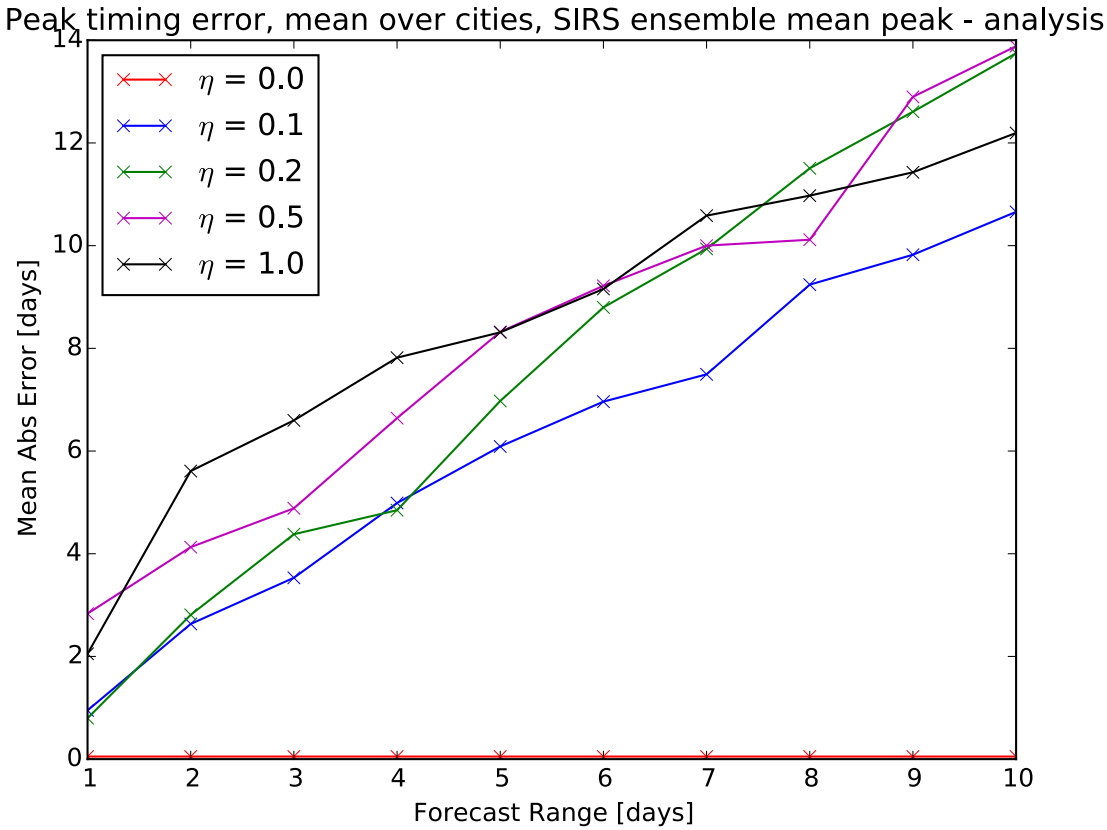
- Simple data assimilation with weighted sum of flu data (ILI+) and previous model results used to re-initialize model
- Pseudo-forecasts (in hindsight with observed humidity) run with different ranges.




- Results from the SIRS model for Dallas, 2013-2014.
- 5 day (magenta) and 10 day (blue) pseudo-forecast curves shown here, along with analysis (red) and ILI+ (cyan) data.

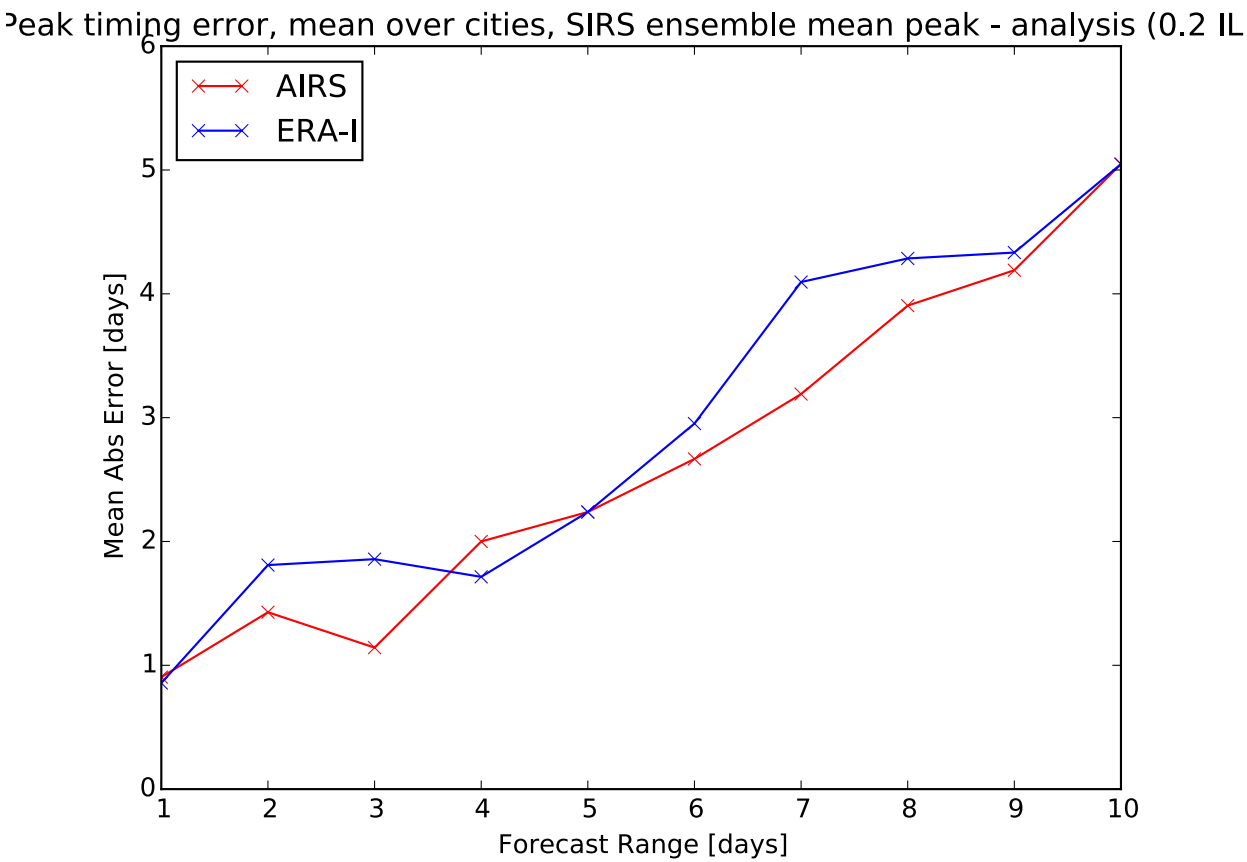
# Forecast Skill Measures

- Mean absolute peak timing error as a function of forecast range.
- Averaged over ensemble members and several locations (21 US cities) and seasons (2010-2014)
- Here different curves are for simulations with different data assimilation weights (eta=0 means flu observations ignored)



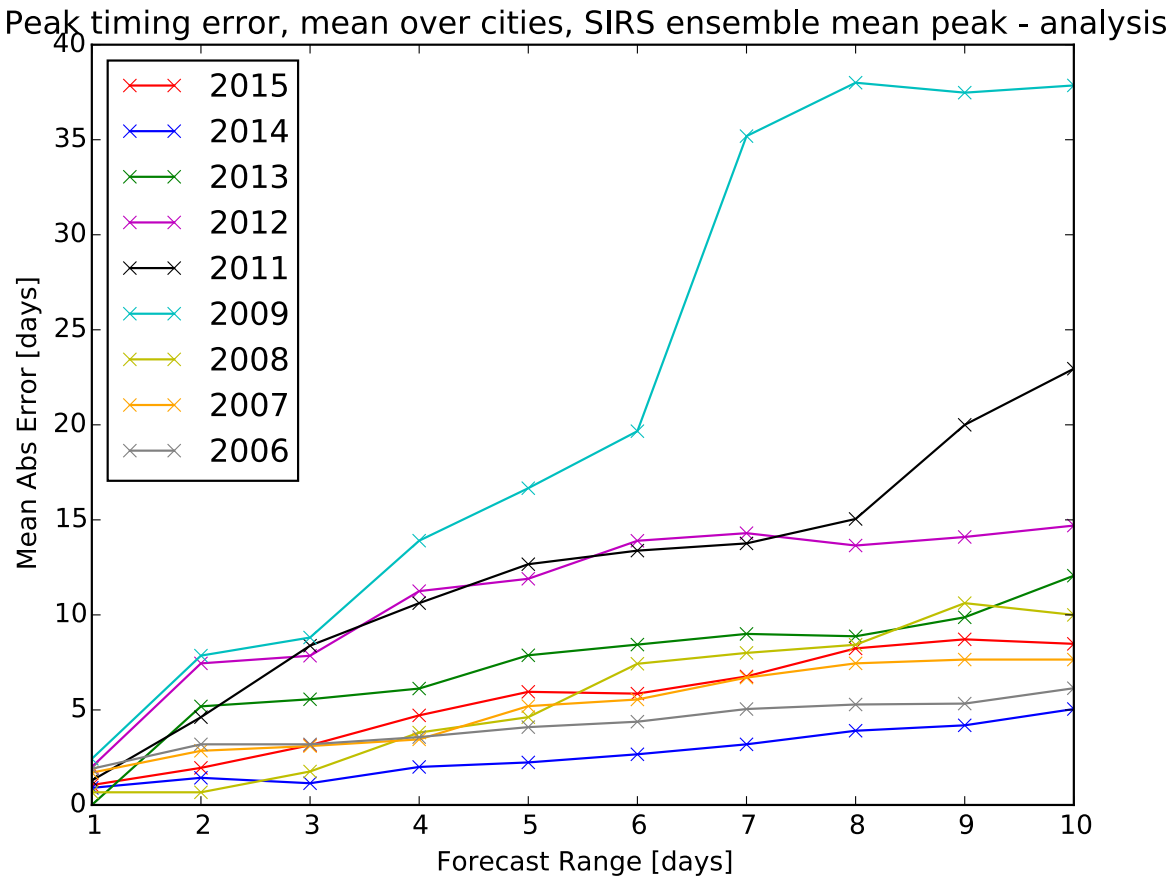
# Humidity Sources

- Comparing simulations driven with humidity from AIRS and ECMWF ERA-Interim (everything else equal)
- Averaged over several locations (21 US cities), here showing 2013-2014 season
- Source and details of surface specific humidity data do not seem to be a dominant source of uncertainty, but further characterization is needed



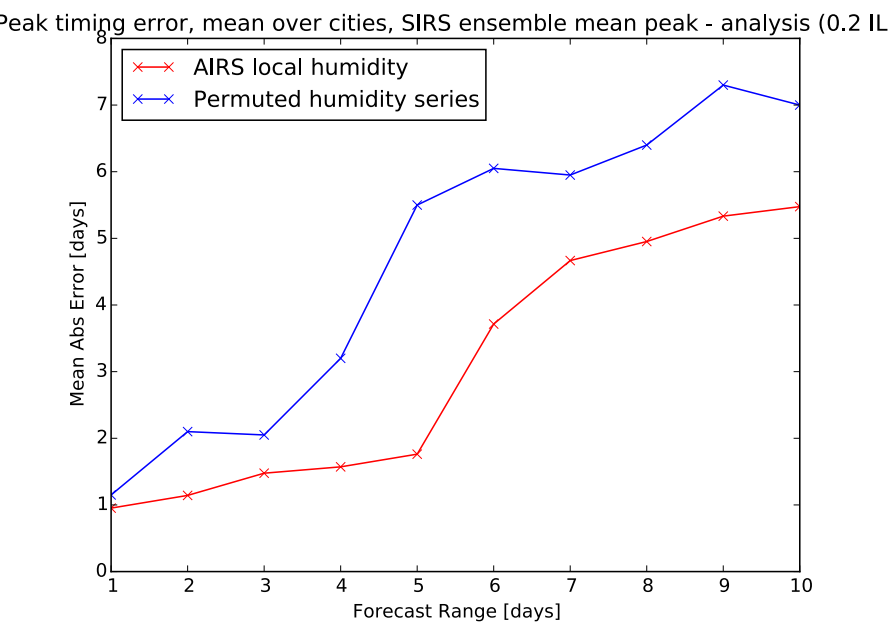
# 2005-2015 seasons

- Model performance can vary significantly with season
- ‘Out of season anomalies’ like 2009 swine-flu not expected to be captured well by this ‘universal’ model



# Testing if Local Humidity Conditions Really Drive Flu Seasonality

- The red line shows simulation results with model driven by actual local AIRS humidity observations
- The blue line shows results when using the AIRS datasets but permuting their location (so matching ILI+ data with model driven by wrong humidity time series on purpose) -- significantly worse results




- Results support the hypothesis that local humidity conditions are a significant driver for influenza seasonality

# Operational Forecasting System Incorporating Humidity Forecasts

- Having evaluated the model, studying past seasons with available humidity observations, we now incorporate NCEP forecasts for specific humidity

- Code modified to operationally produce forecasts on a regular basis

- Depending on data availability and lag time for reporting, we obtain and use a combination of influenza incidence data, humidity observations and forecasts and previous results from the SIRS model

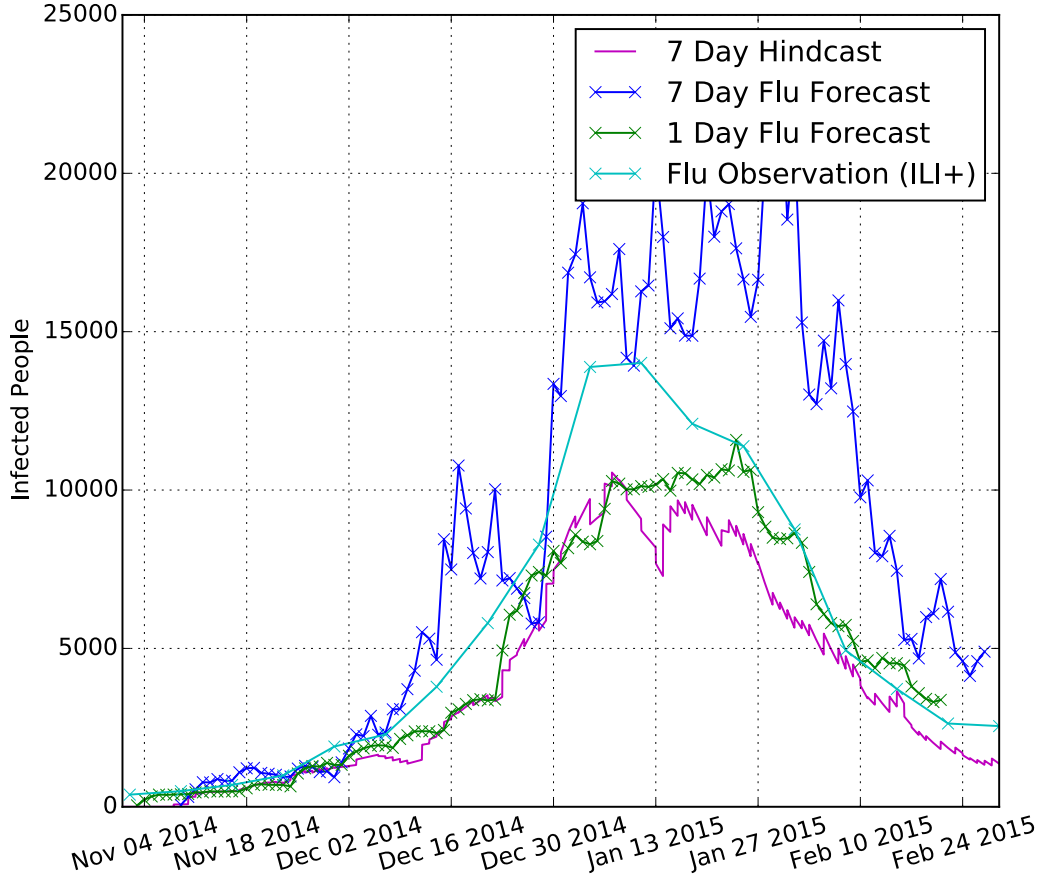
- On a given day, for a given location, we acquire the latest influenza incidence reported online and feed into the model, check for the most recent humidity observations (AIRS) to initialize model calculations, and automatically acquire NCEP humidity forecasts for all the days needed within a specified forecast range (like 1 or 2 weeks). All these datasets are acquired and processed to a conformed format in an automated way. Output data from simulations is collected, and ensemble statistics performed.

- A version of this system is now ready and working.

- Still need to make it more general/versatile to process and use different data sources, time/space resolution, etc., and validate.

# Operational Forecasting System Preliminary Results

Los Angeles, 2014-215 season



# Summary

- We have developed a humidity driven (by AIRS near-surface humidity) influenza model operating on a city, state or regional scale
- It is a promising tool to inform about influenza seasonality, peak timing, onset and relative intensity of outbreaks
- Simulations of past seasons 2005-2015 performed for validation, using observed (AIRS) humidity
- Results provide support for the hypothesis that local humidity conditions are a significant driver for the seasonality of influenza
- Humidity forecasts incorporated along with AIRS humidity into forecasting system, getting ready to run operationally

# Future Work

- Further establishing/verifying the environment-influenza link with AIRS and other data sources (humidity, temperature, anomalies vs. absolute values, pollution?)
- Further validation of the prediction system
- Obtaining and incorporating more specific influenza incidence data from medical networks and authorities and internet sources.
- Developing confidence and uncertainty measures (including effects of AIRS humidity data uncertainty)
- Better constraining of parameters and methods that account for their variability and uncertainty
- Applying the system and comparing to other regions, outside US
- Implementing and assessing longer term seasonal predictions (using climatology, maybe longer term humidity predictions in the future)
- Exploring different types of models, virus subtypes, population age structure, geographical spread