Forecasting Valley Fever (Coccidioidomycosis) Incidence via Soil Moisture Conditions: Leveraging an Extended *In Situ* Soil Moisture Record

E.J. Coopersmith[1], M.H. Cosh[1]., J.E. Bell[2,3,4], and J. Shriber[4,5]

USDA-ARS-Hydrology and Remote Sensing Laboratory, Beltsville, MD
Cooperative Institute for Climate and Satellites, Asheville, NC
NOAA-National Climate Data Center, Asheville, NC
Center for Disease Control, Atlanta, GA
Rollins School of Public Health, Atlanta, GA

(http://www.cdc.gov/fungal/diseases/coccidioidomycosis/)

USDA-ARS HRSL Hydrology & Remote Sensing Lab Beltsville, Maryland, USA

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION U.S. DEPARTMENT OF COMMERCE

Valley Fever Incidence is soil moisture dependent…
…but this hypothesis could not be tested!

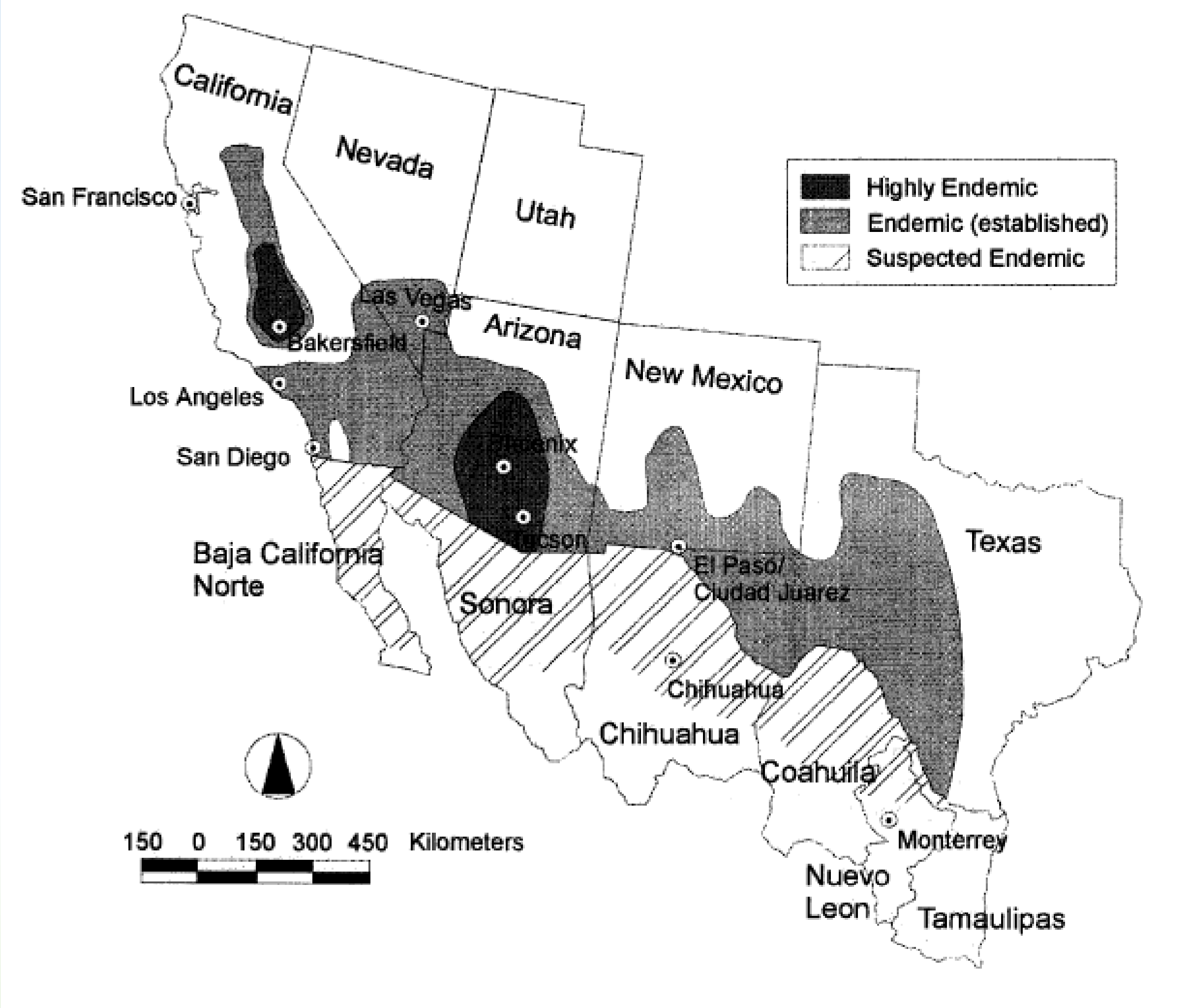
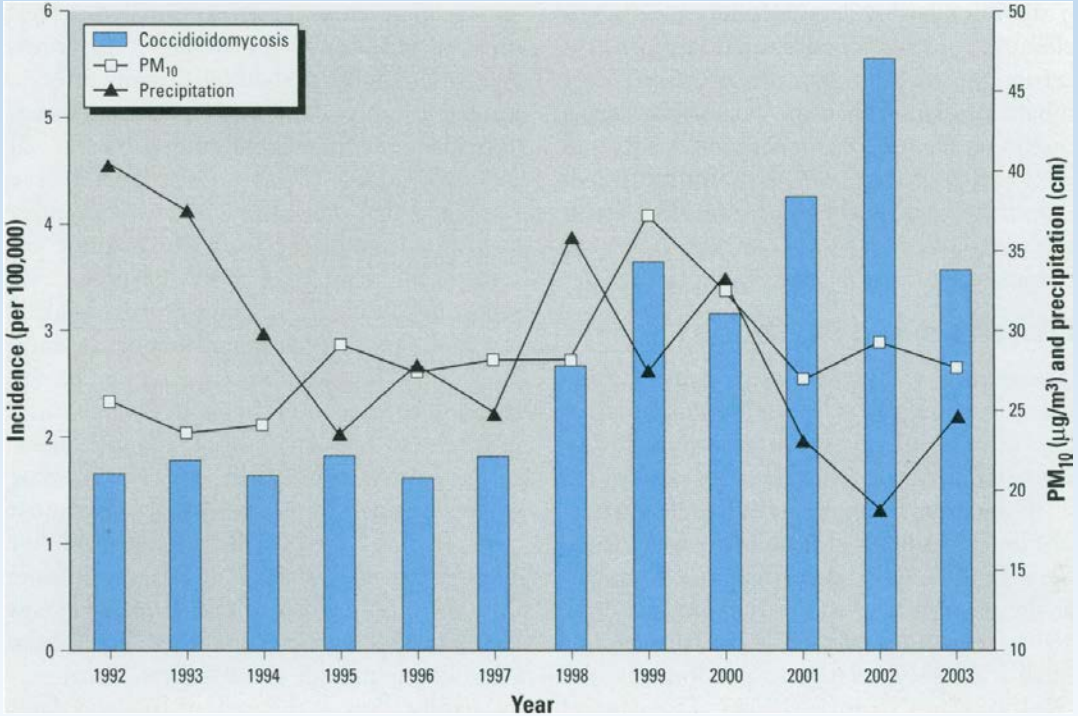


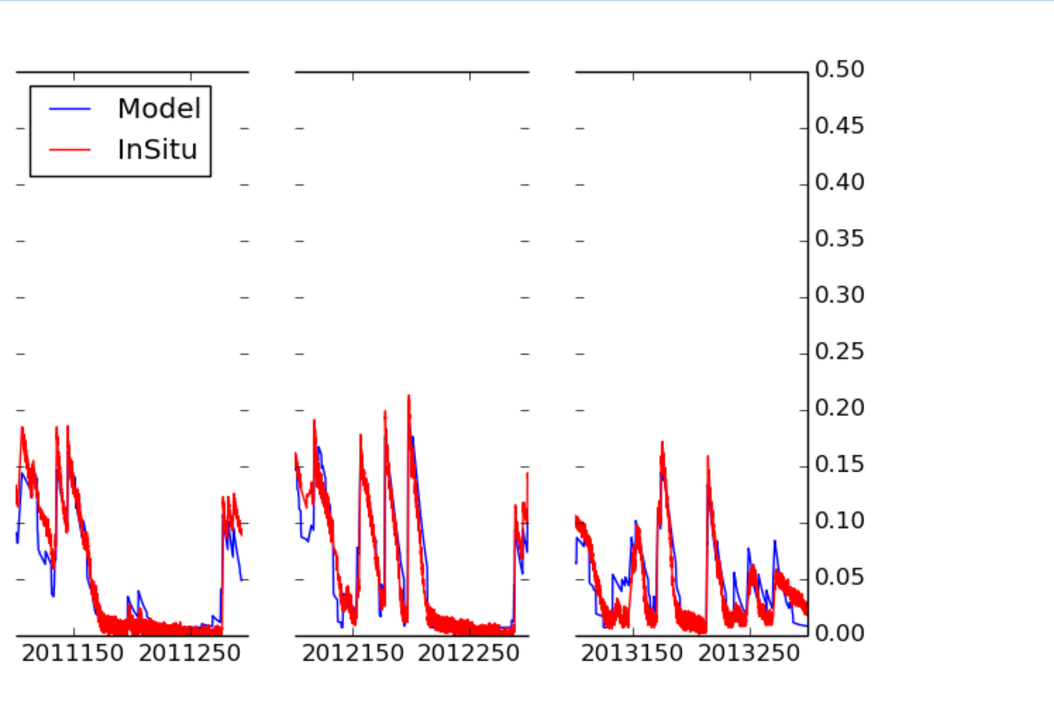
Fig. 1 Areas of the United States and northern Mexico that are considered endemic for valley fever. (Adapted from Kirkland and Fierer 1996)



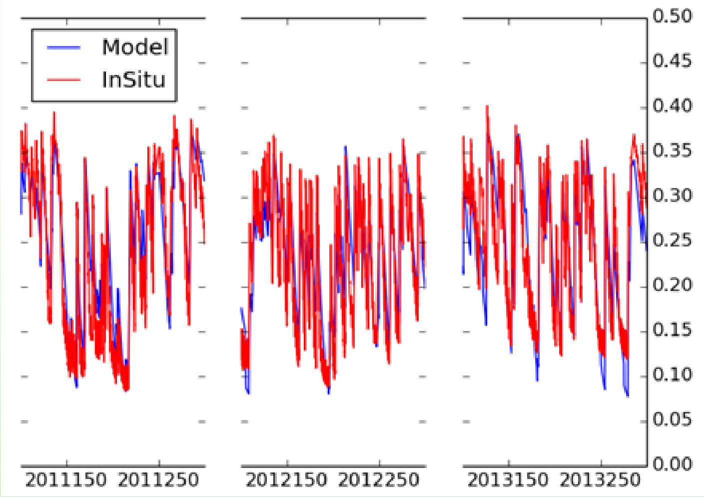
Kolivras et al., (2003, left) and Comrie et al. (2005, above) compared the numbers of reported cases to precipitation and temperature…

…but now soil moisture data are available…

USCRN records begin between 2009 and 2012, but SCAN sensors have been installed for longer.

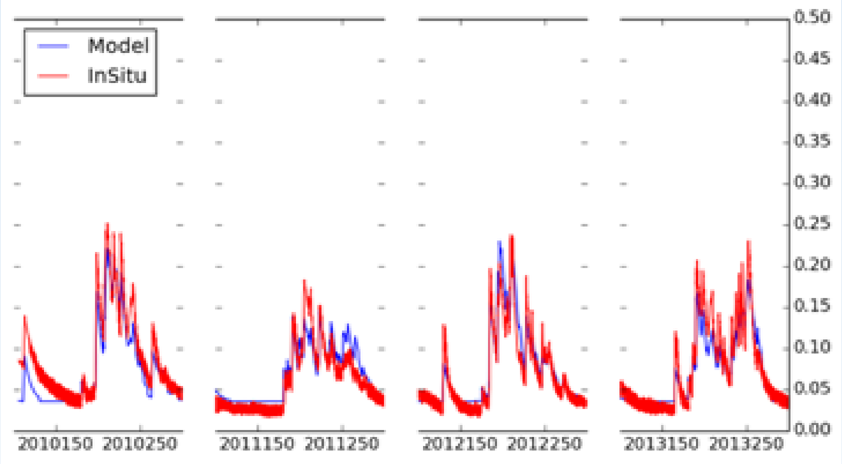


Top: SCAN #2021, Washington. Calibration (2011-2013) and Validation (2007-2010)
Bottom: SCAN #2039, Virginia. Calibration (2011-2013) and Validation (2003-2010)
(Coopersmith et al, 2015)



Analysis from Coopersmith et al (2015) verified that one can calibrate a model in recent years, and apply it *retroactively*.

This allows a USCRN soil moisture record (installed in Arizona in 2010, for example)…



…to be extended back to precipitation sensor's original installation, sometime in 2002.

This record can now be compared to reported cases of valley fever in the 21st century, normalized by county population.

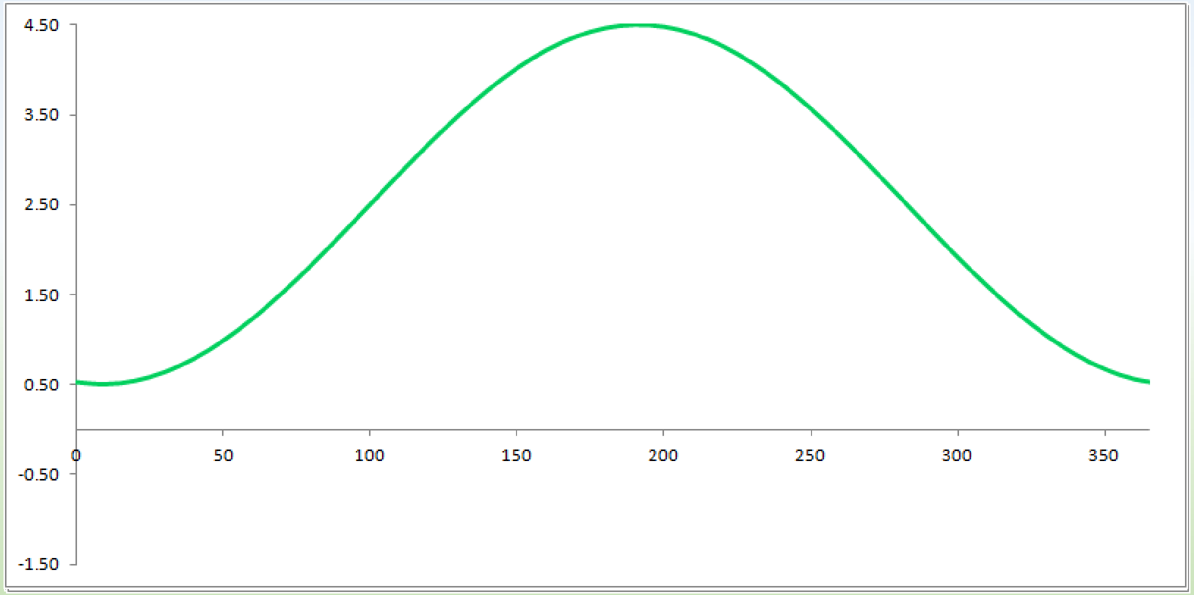
Brief Technical Description of the Model
(The "eyes glaze over" part of the presentation)

The diagnostic soil moisture equation:
(Pan et al, 2003; Pan, 2012)

$$\theta_{est} = \theta_{re} + (\phi_e - \theta_{re})(1 - e^{-c_4\beta})$$

$$\beta = \sum_{i=2}^{i=n-1}\left[\frac{P_i}{\eta_i}\left(1 - e^{-\frac{\eta_i}{z}}\right)e^{-\sum_{j=1}^{j=i-1}\left(\frac{\eta_j}{z}\right)}\right] + \frac{P_1}{\eta_1}\left(1 - e^{-\frac{\eta_1}{z}}\right)$$

y = **α**sin(x-**h**) + **v**



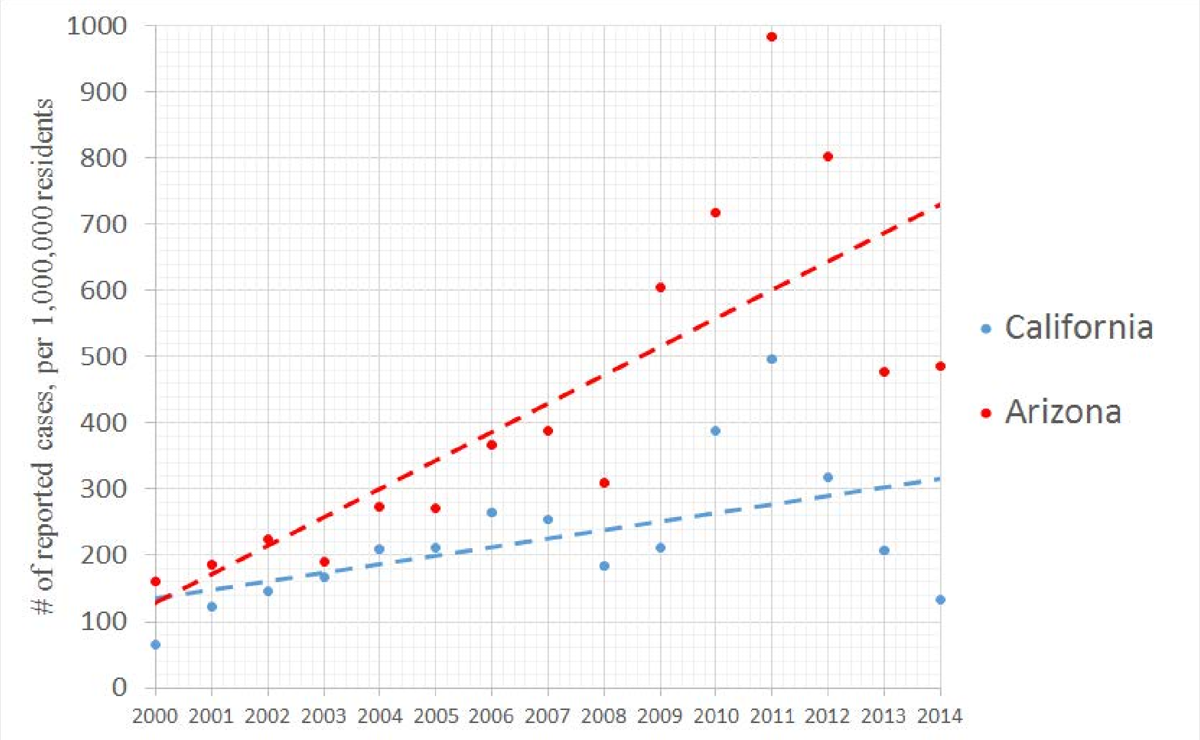
{**v**, **α**, **h**}

Framing the analysis:
Asking the appropriate questions

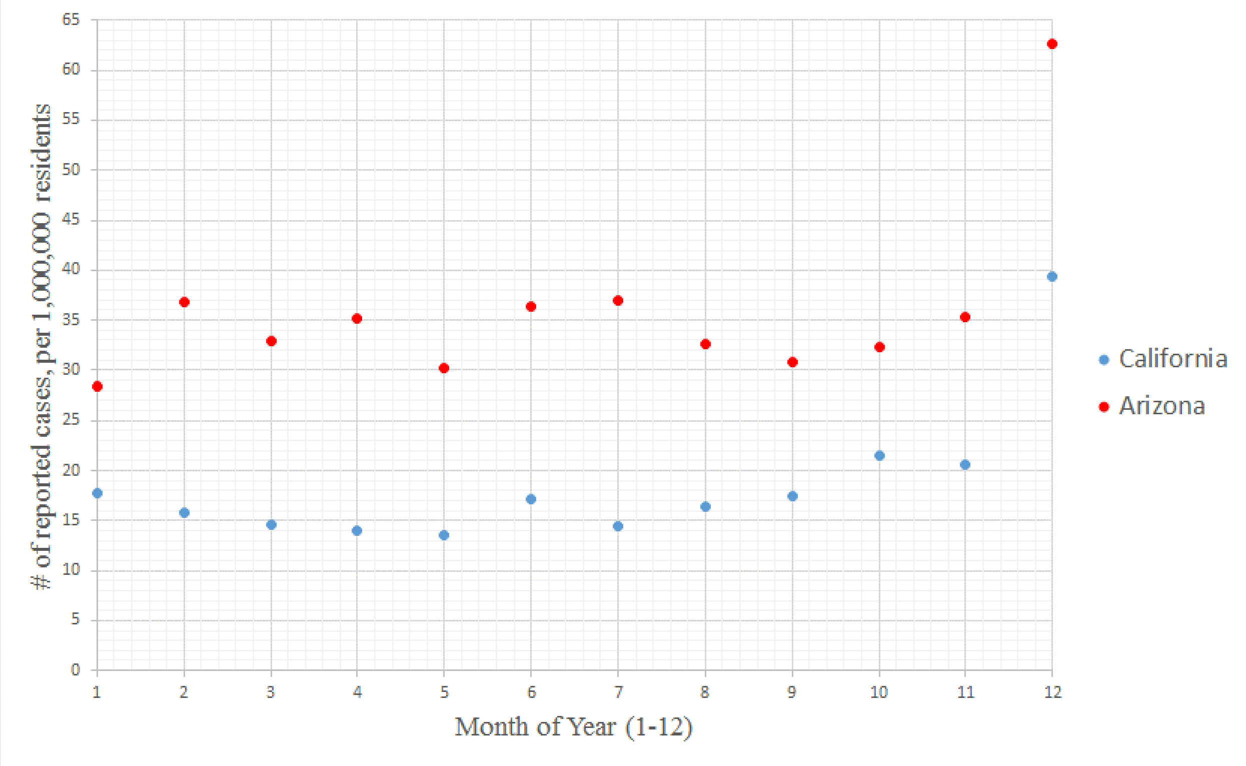
Is there an annual trend? Yes! So let us remove it from our data (by state):

Step 1 – Fit a simple linear regression: $normalized\_VF_y = \beta_0 + \beta_1 y$

Step 2 – Remove that linear, annual trend: $$detrended\_VF_{county,month,y} = \frac{normalized\_VF_{county,month,year}}{\beta_0 + \beta_1 y} * \frac{\sum_{i=2000}^{2013} normalized\_VF_i}{14}$$

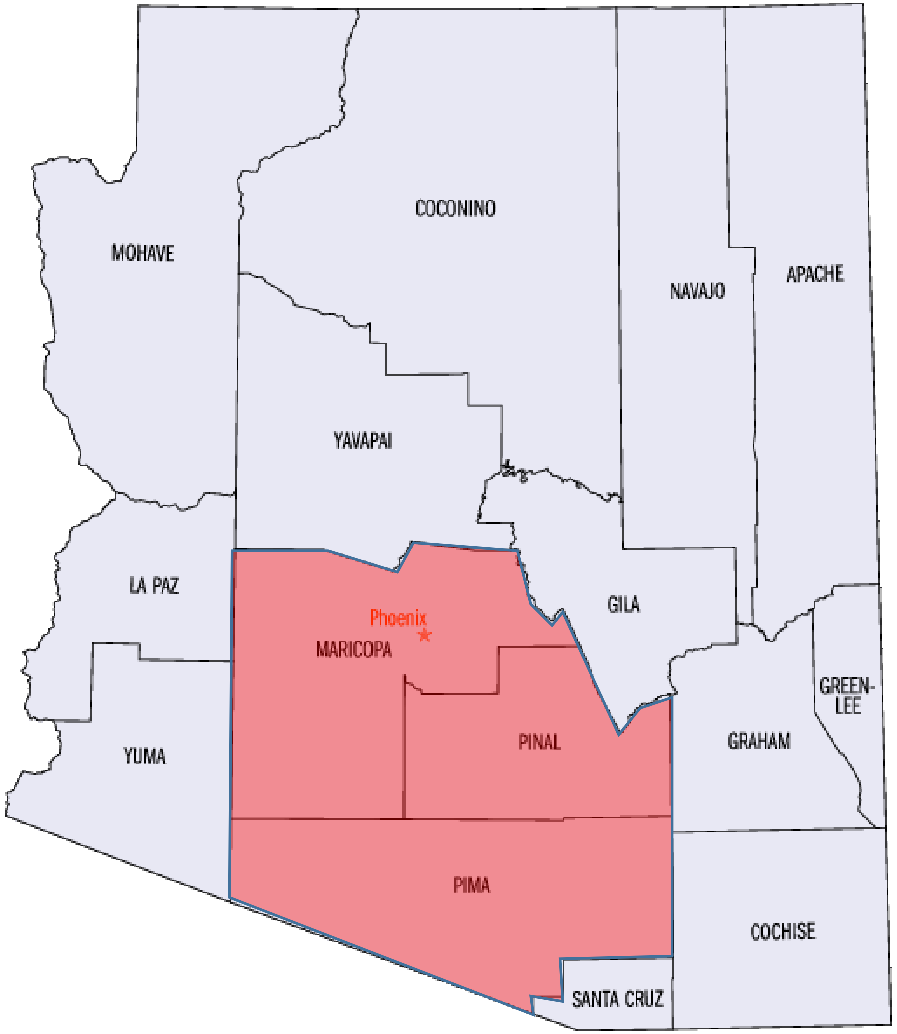


Is there a seasonal trend?



No – There is not a (pronounced) seasonal pattern. Without *a priori* knowledge, this looks like noise. (Excluding December)

Constructing the dataset:
"What to leave in, what to leave out…"



Counties are included if and only if the average number of monthly cases exceeds 10…

…and we match up each county with its nearest USCRN or SCAN gauge to obtain soil moisture estimates…

…preparing a model estimate alongside the *in situ* estimates at each sensor location.

Framing the analysis:
With soil moisture data and normalized valley fever data, what can we compare?

3 Soil moisture series:
(*In situ*, model, and "merged")

7 Soil moisture metrics:
(ActualSM, Actual-ExpectedSM, hrs_above_5%, hrs_above_10%, hrs_above_15%, hrs_above_20%, hrs_above_25%)

6 Possible month ranges to aggregate soil moisture data:
(*ex: 1-month Jan., Feb., Mar., …*
2-month Jan-Feb, Feb-Mar, …
3-month Jan-Mar, Feb-Apr, …)

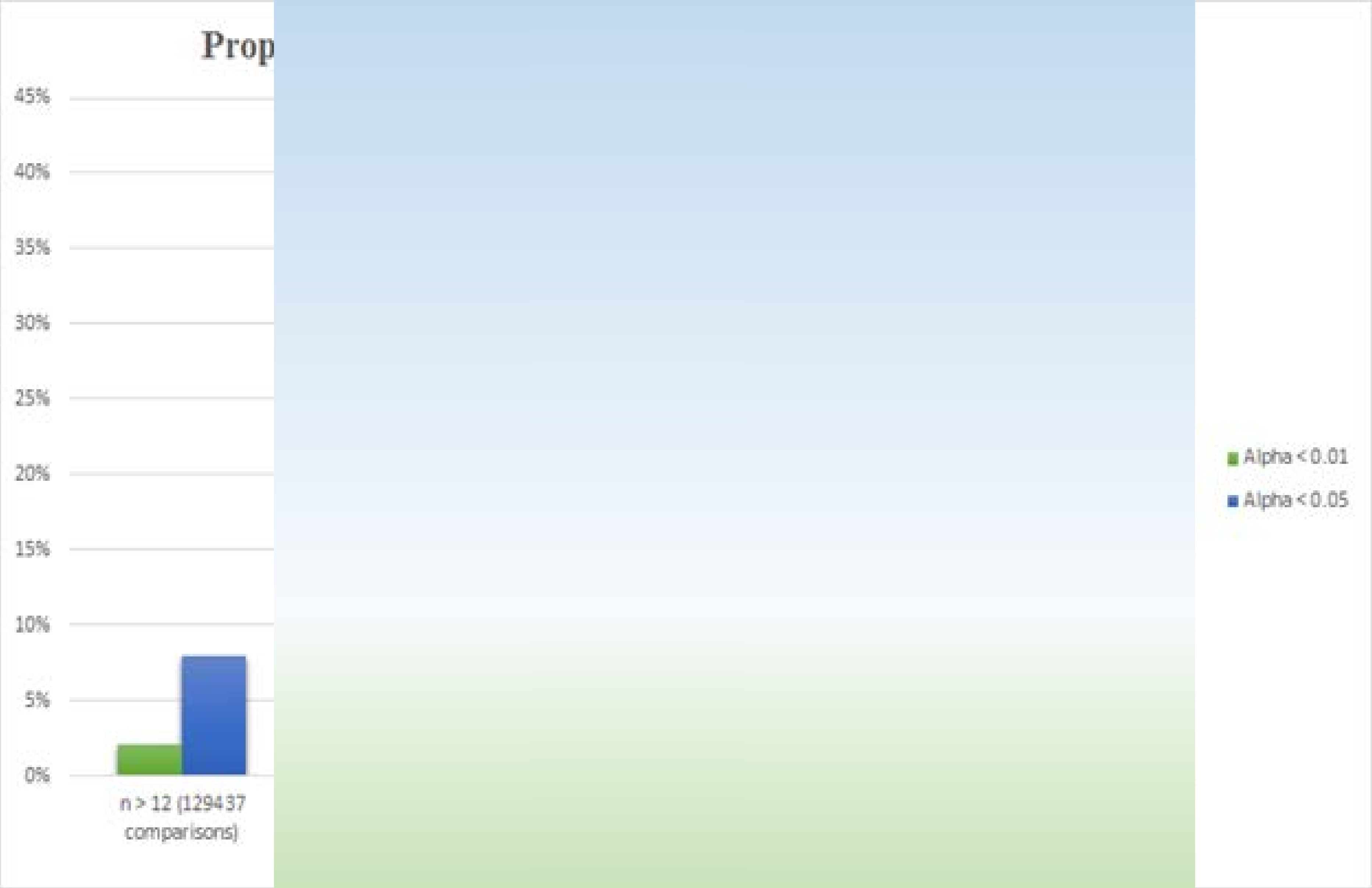
6 Possible month ranges to aggregate valley fever data:
(*ex: 1-month Jan., Feb., Mar., …*
2-month Jan-Feb, Feb-Mar, …
3-month Jan-Mar, Feb-Apr, …)

36 Possible lags between independent and dependent variables:
(*1-mo. lag, 2-mo. lag, …, 36-mo. lag*)

12 Possible months (or aggregations thereof):
(*We can start make the first predicted month Jan, Feb, Mar, …*)

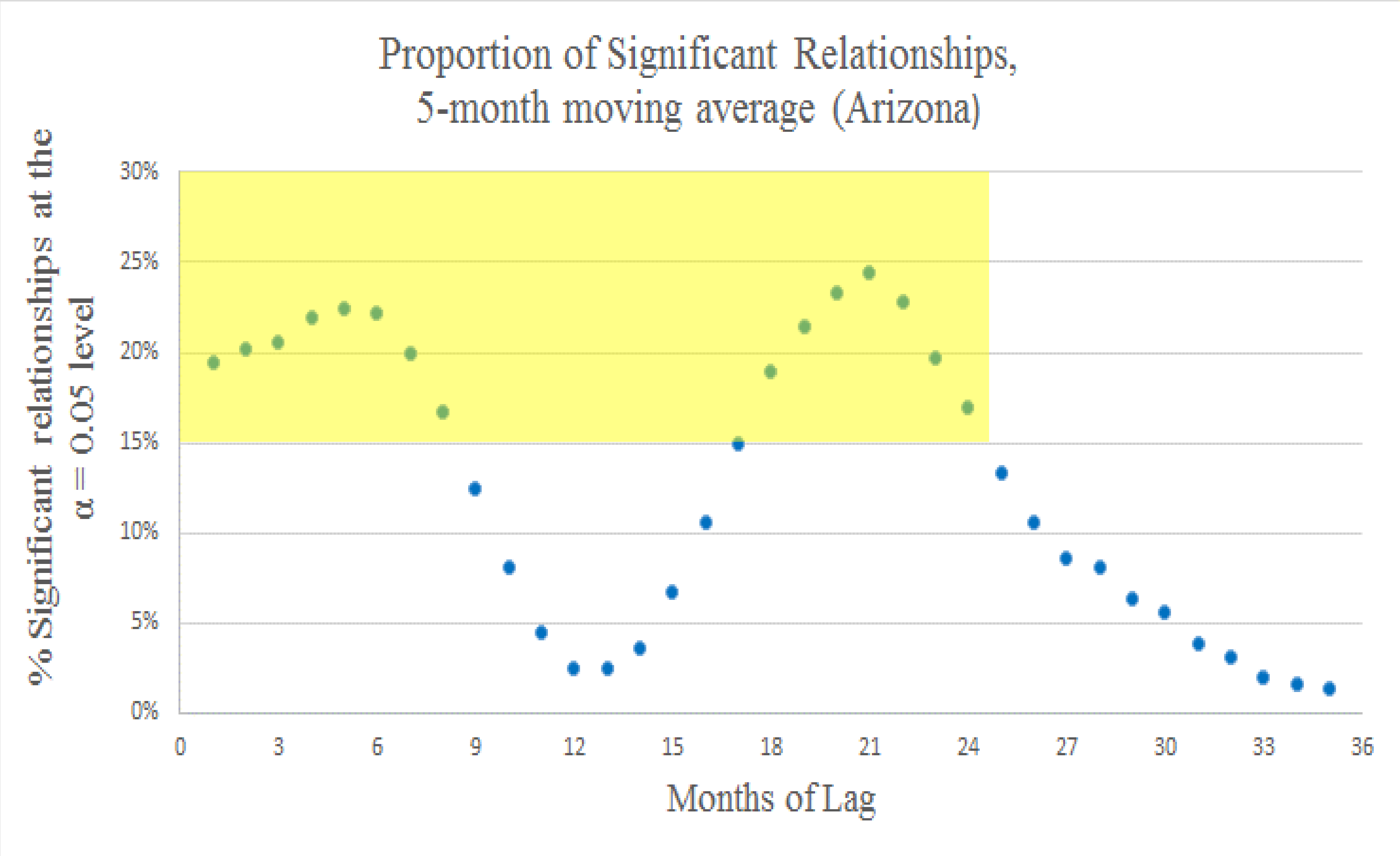
3 * 7 * 6 * 6 * 36 * 12 =
326,592 possible comparisons

Focusing the Lens:
Gaining insight in Arizona



Finally, we choose the number of months of ‘lag’ more intelligently (see Figure on next slide).

Choosing the Window:
Determining how far to look back in Arizona



We note that predictive power is greatest when lags are between 1 and 7 months or 17 and 24 months (not accidentally, these windows are staggered by one year)

Focusing the Lens:
Gaining insight in Arizona

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X-1
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X-2

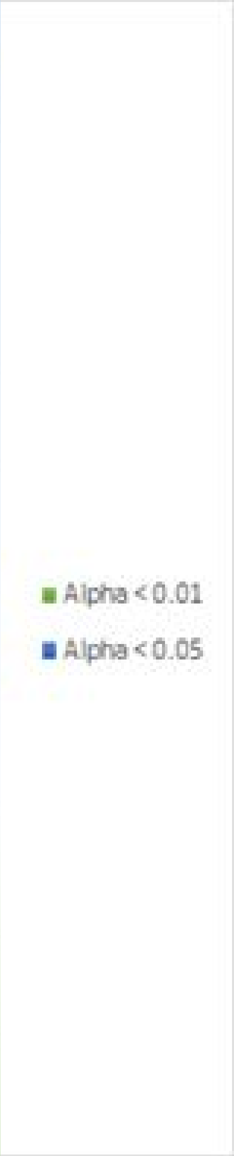
Cases of valley fever in year X…

…are inversely correlated with hours above $0.05 m^3/m^3$ in year X-1

…and directly correlated with hours above $0.05 m^3/m^3$ in year X-2

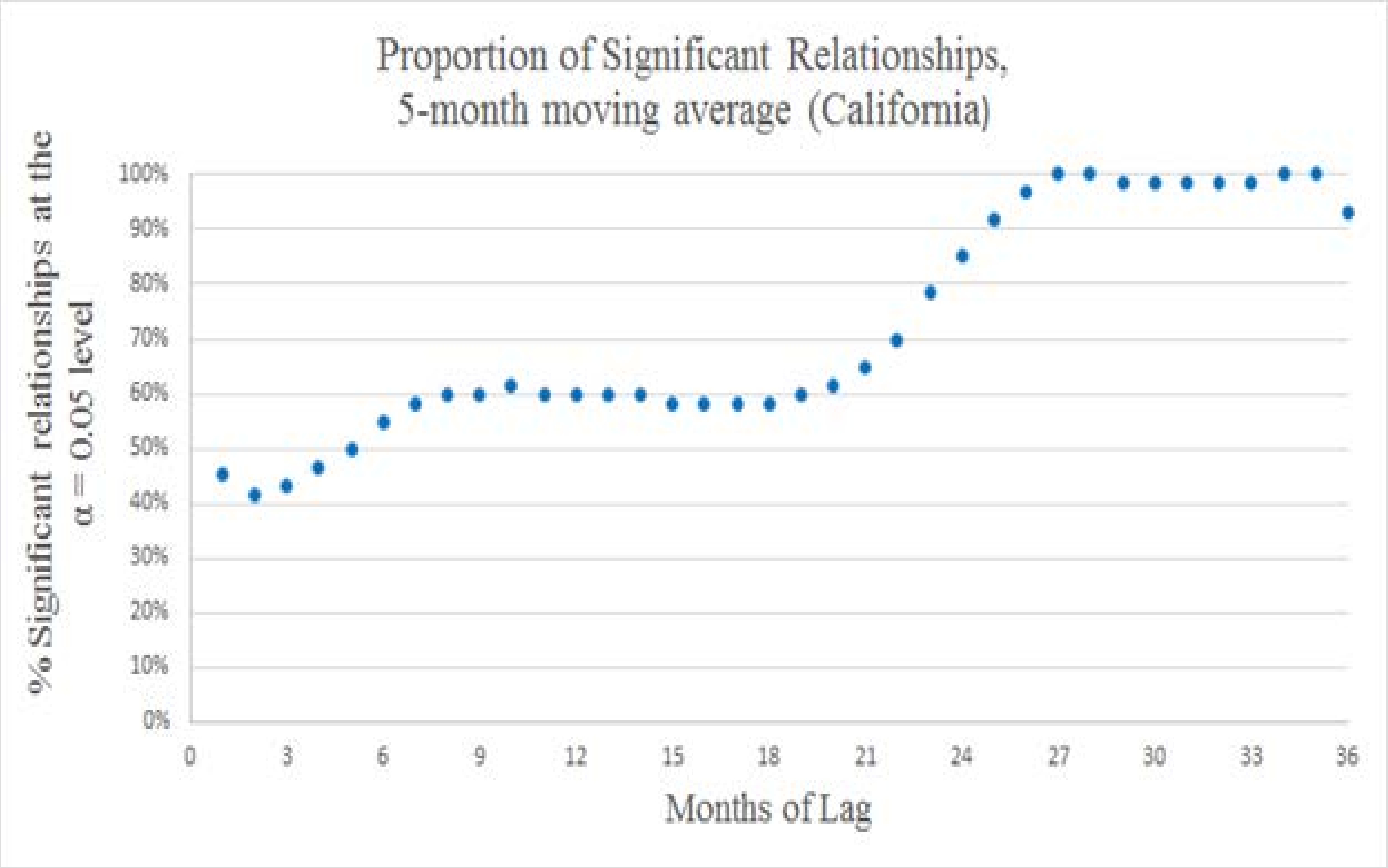
We analyze the remaining statistically significant relationships, and the following patterns emerge/recur.

Focusing the Lens:
Gaining insight in California



In this case, all potential “lags” are viable (see Figure on next slide).

Choosing the Window:
Determining how far to look back in California



Unlike Arizona, in California predictive power remains strong for the full three years examined.

Focusing the Lens:
Gaining insight in Arizona

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X-1
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X-2
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year X-3

Cases of valley fever in year X…

…are inversely correlated with hours above 0.10m^3/m^3 in year X-1

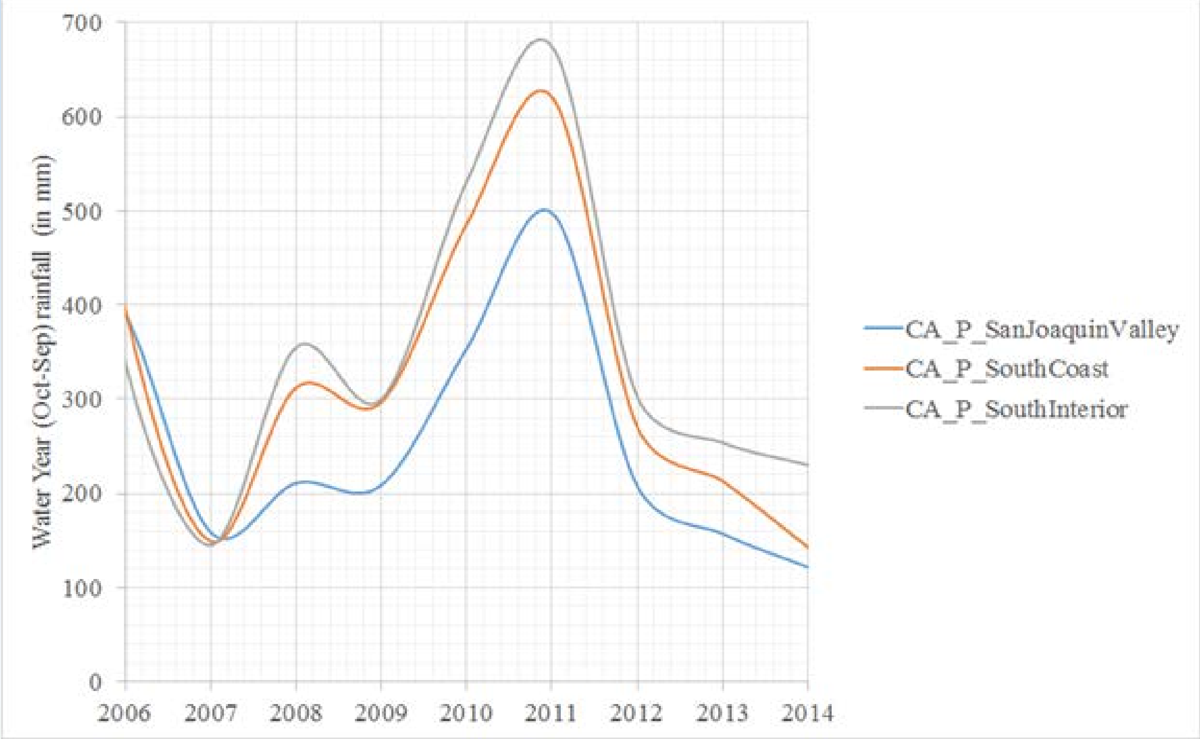
…and inversely correlated with hours above 0.10m^3/m^3 in year X-2

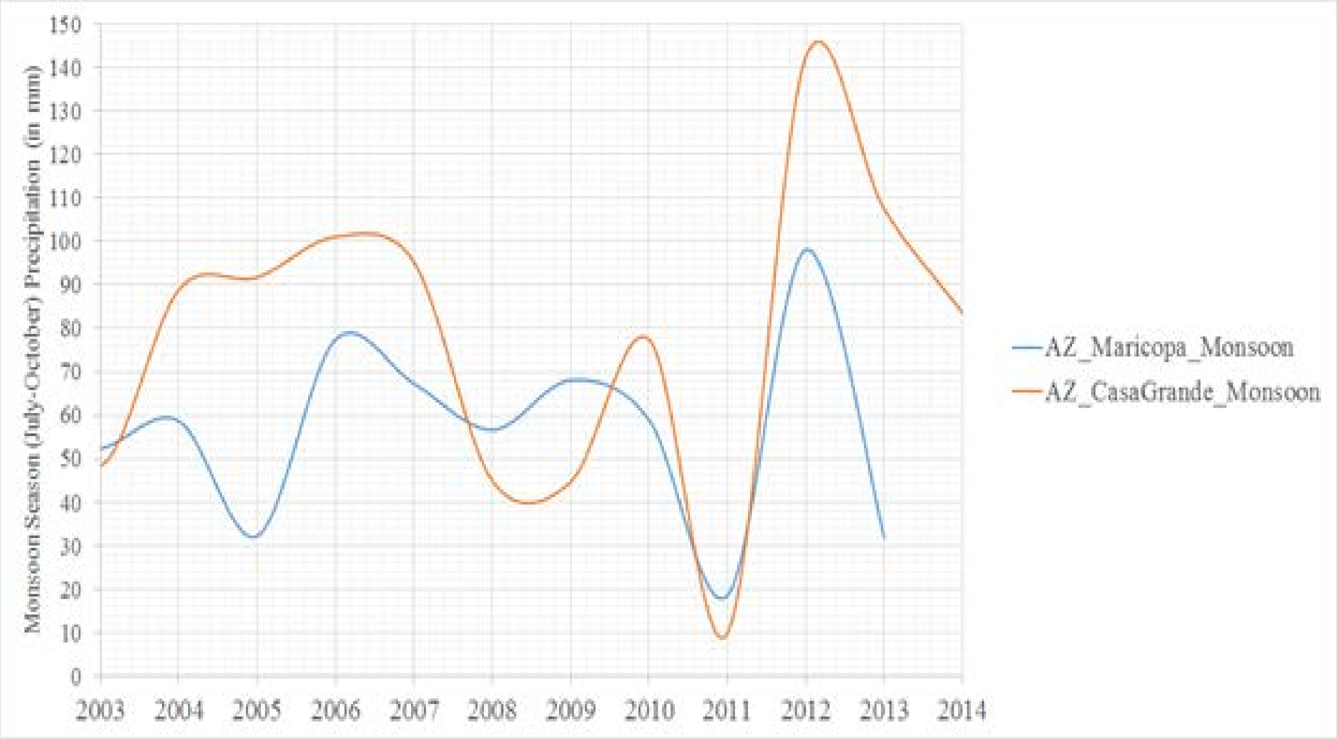
…and inversely correlated with hours above 0.10m^3/m^3 in year X-3

We analyze the remaining statistically significant relationships, and the following patterns emerge/recur.

Discussion – What did we learn?
(A brief hydro-climatological tangent)

In both states, dry soil during the summer portends higher incidence of valley fever in the coming year…
…is this an idiosyncratic property of a small number of years? (Did we get "lucky" with our data?)





California experienced a *wet* 2011 water year, which would entail a drop-off in 2012…

…but Arizona experienced a very *dry* 2011 summer, which would imply a jump in 2012.

Precipitation does not tell the full story, *soil moisture* data can help us better understand how and when valley fever spores form.

Works Cited

Bell, J. E., M. A. Palecki, C. B. Baker, W. G. Collins, J. H. Lawrimore, R. D. Leeper, M. E. Hall, J. Kochendorfer, T. P. Meyers, T. Wilson, and H. J. Diamond. 2013: U.S. Climate Reference Network soil moisture and temperature observations. J. Hydrometeorol., 14, 977-988. doi: 10.1175/JHM-D-12-0146.1

Comrie, A.C. “Climate Factors Influencing Coccidioidmycosis Seasonality and Outbreaks.” 2005. Environmental Health Perspectives. 113 (6), 688-692. doi: 10.1289/ehp.7786

Coopersmith, E.J, Cosh, M.H., Bell, J.E., 2015. Extending the soil moisture data record of the Climate Reference Network (CRN) and Soil Climate Analysis Network (SCAN). Advances in Water Res. 79, 80-90. doi:10.1016/j.advwatres.2015.02.006

Diamond, H.J., Karl, T.R., Palecki, M.A., Baker, C.B., Bell, J.E., Leeper, R.D., Easterling, D.R., Lawrimore, J.H., Meyers,T.P., Helfert, M.R., Goodge, G., Thorne, P.W. 2013. U.S. Climate Reference Network after one decade of observations: Status and assessment. Bull. Am. Met. Soc. Doi: 10.1175/BAMDS-D-12-00170.1

Kolivras, K., Comrie, A.C. “Modeling Valley Fever Incidence Based on Climate Conditions.” 2003. International Journal of Biometeorology. 47, 87-101. doi: 10.1007/s00484-002-0155-x

Pan F, Peters-Lidard CD, Sale MJ. An analytical method for predicting surface soil moisture from rainfall observations.” WRR. 2003; 39(11)

Pan F. Estimating daily surface soil moisture using a daily diagnostic soil moisture equation. Journal of Irrigation and Drainage Engineering. 2012; 138(7):625-631

Schaefer GL, Cosh MH, Jackson TJ. The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). Jour. of Oceanic and Atmospheric Tech. 2007; 24:2073-2077

Thank You for Listening
Any Questions?