Temporal and Spatial Trends in Winter Wheat Yields Monitored Via Multiple Online Data Sources

Peter Claussen
Pete@gdmdata.com

Introduction

Recent studies have suggested that crop yields gains have slowed or stagnated in many regions globally [7, 6], while others shows patterns of increasing yield [8] . A simple analysis of winter wheat yield in the central United States shows that while yield gains have slowed for some (southern) regions, relative to mid-century gains, yield improvement have accelerated in other (northern) regions (Figure 1a). This implies a shift in winter wheat productions zones, perhaps related to climate change.

To explore this shift, data from several public databases were combined for covariate analysis to determine if changes in environmental and socio-economic variables correlate with geospatial yield trends. For simplicity, this analysis was centered on the Great Plains states plus adjacent states with testing stations cooperating in the Hard Winter Wheat Regional Nursery [1]. (Figure 1b). The period from 1984-2014 was chosen by inspection of the inflection points in (Figure 1a).

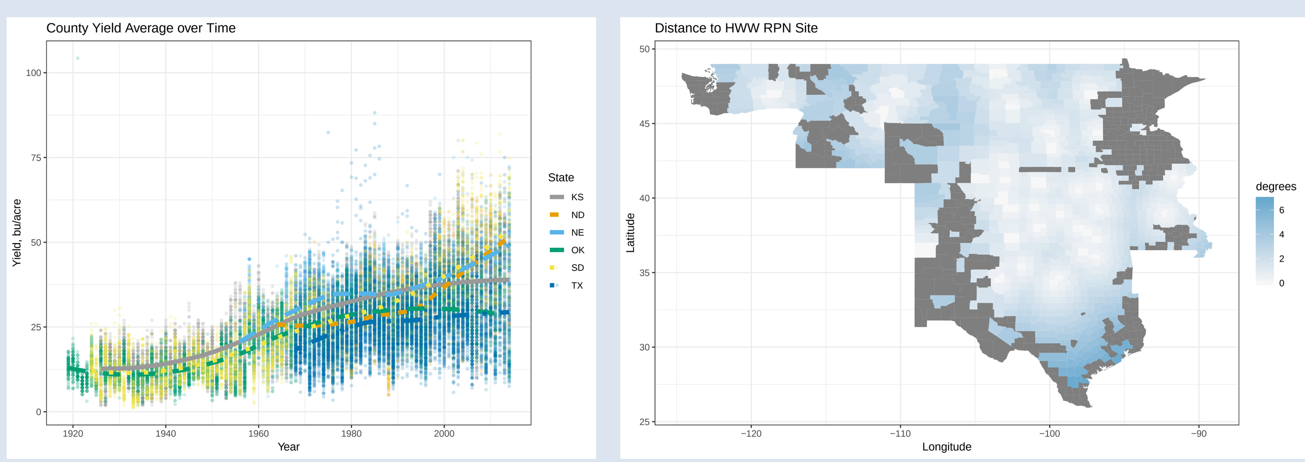

Figure 1: a. County average wheat yields (bu/acre) for six states in the Great Plains. b. Approximate distance (degrees latitude/longitude) from county center to nearest HWW RPN site.

Materials and Methods

Data sources and brief descriptions:

- USDA NASS [5] Yield, acreage, income.
- NuGIS [2] Soil N, P , K addition and withdrawal.
- CDC WONDER [3, 4] Air and surface temperatures, precipitation.
- HWW RPN [1] Testing of advanced breeding lines.

For all data, linear regression coefficients on year x_i and county n were computed in R using the model $y_{ni} = a_n + b_n x_i + e_{ni}$. Weighted means were computed by by $\bar{y}_n = \hat{a}_n + \hat{b}_n \times 1999$. Slopes were normalized to a relative rate as $\%\hat{b_n} = 100 \times \hat{b}_n / \bar{y}_n$.

Conclusion

This analysis demonstrates that multiple data sources can be integrated to provide insight into regional yield trends. However, difficulties managing heterogeneous data, too numerous to list here, present a challenge to updating this analysis as new data are available.

References

[1] Hard Winter Wheat Regional Nurseries www.ars.usda.gov/main/docs.htm?docid=11932.

[2] IPNI. 2012. A Nutrient Use Information System (NuGIS) for the U.S. Norcross, GA. www.ipni.net/nugis.

[3] Moderate Resolution Imaging Spectroradiometer (MODIS) Daily Land Surface Temperature (LST), years 2003-2008 on CDC WONDER Online Database, released 2012. wonder.cdc.gov/nasa-lst.html.

[4] North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index, years 1979-2011 on CDC WONDER Online Database, released 2013. wonder.cdc.gov/NASA-NLDAS.html.

[5] United States Department of Agriculture, National Agricultural Statistics Service quickstats.nass.usda.gov.

[6] R Graybosch, HE Bockelman, KA Garland-Campbell, DF Garvin, and T Regassa. *Yield Gains in Major U.S. Field Crops*, chapter Wheat. ASA, CSSA and SSSA, 5585 Guilford Rd. Madison, WI 53711-5801, USA, September 2014.

[7] M Lin and P Huybers. Reckoning wheat yield trends. *Environ Res Lett*, 7, July 2012.

[8] DK Ray, N Ramankutty, ND Mueller, PC West, and JA Foley. Recent patterns of crop yield growth and stagnation. *Nature communications*, 3:1293, December 2012.

Spatial Plots, Winter Wheat Yield and Covariates

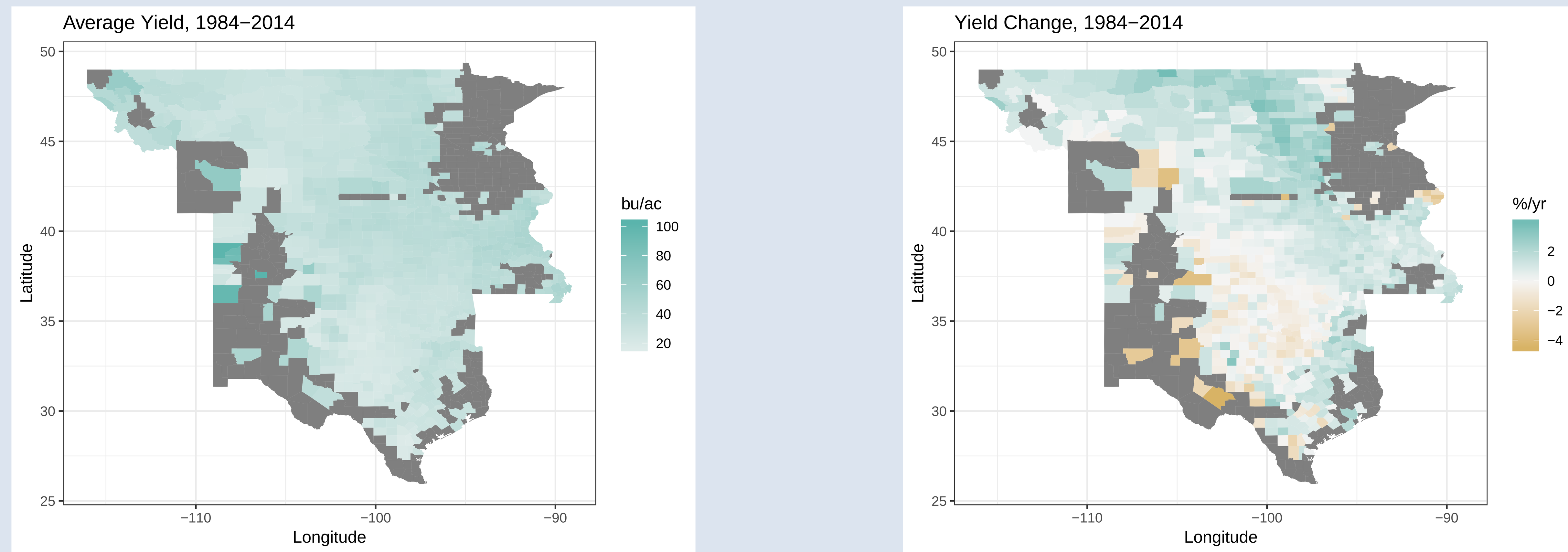

Figure 2: a. Average winter wheat yields by county. b. Yearly percent change in county winter wheat yields

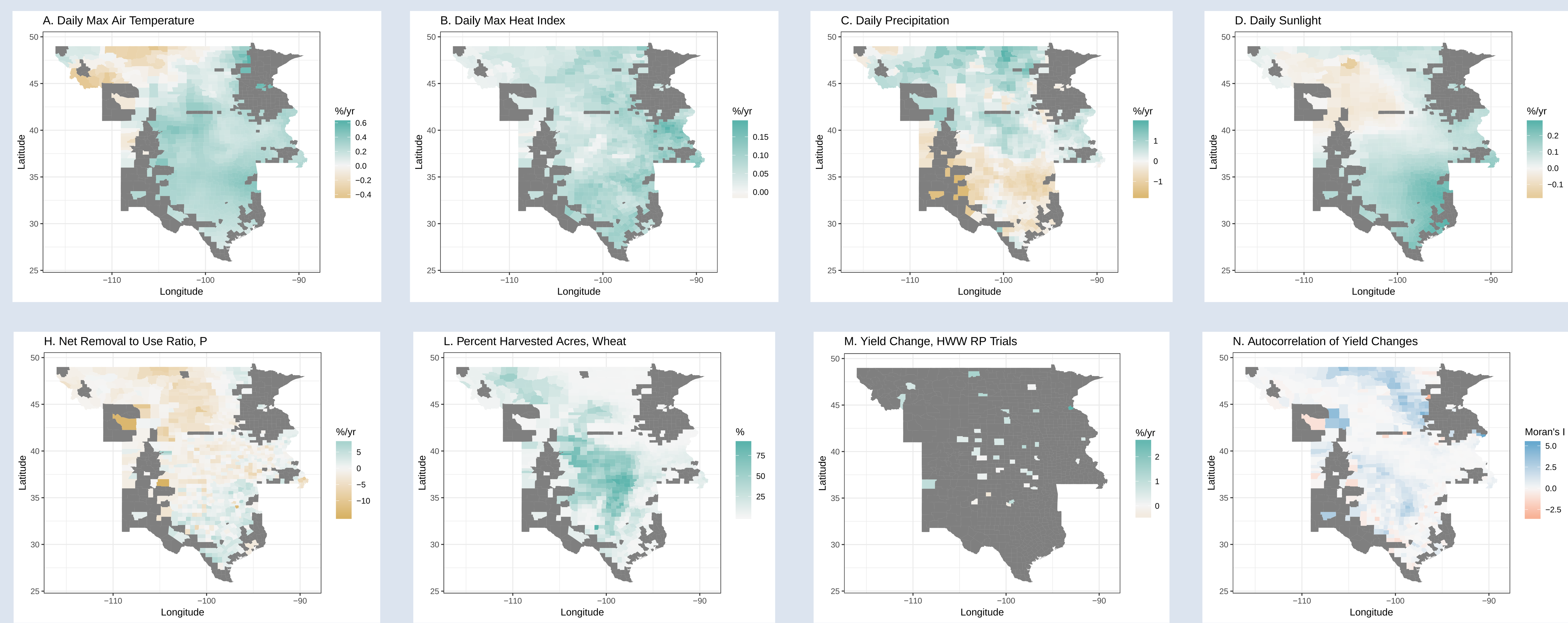

Figure 3: Spatial distribution of changing covariates correlated with yield changes. Alphabetic identifiers correspond to Table 1. Figures A,B,C,D,H and L represent county level changes over from 1984-2014. Figure M shows change in the average yields for entries in the HWW RPN. Figure N plots a spatial autocorrelation measure (Moran's I), and shows two areas where yield changes cluster. M and N are not included in the regression shown in Table 1.

Regression Coefficients

The linear model relating change in yield (Y_b) to relative change in covariates, $Y_b = \%\hat{b_1} + \%\hat{b_2} + \ldots + \%\hat{b_n}$, was fit in R using the `lm` function. An optimal linear model was found using the `step` function. Table 1 shows the result of this function. Coefficients listed are in % units, with the exception of Percent Harvested Acres, Wheat, which was computed as the percent wheat acres harvested of total crop acres harvested.

Covariate	Coefficient	
A. Daily Max Air Temperature [4]	-1.2558	***
B. Daily Max Heat Index [4]	2.8354	*
C. Daily Precipitation [4]	0.2645	***
D. Daily Sunlight [4]	0.7585	
E. Day Land Surface Temperature [3]	-0.0792	*
F. Farm Inputs, Tons P [2]	-0.0009	
G. Nutrient Removal by Crops, N [2]	-0.0114	
H. Net Removal to Use Ratio, P_2O_5 [2]	-0.0728	***
I. Sum of 21 Crop Acres Planted [2]	-0.1951	*
J. Sum of 21 Crop Acres Harvested [2]	0.3774	***
K. Total Cropland [2]	-0.1531	***
L. Percent Harvested Acres, Wheat	-0.0122	***

Table 1: Summary of best linear model coefficients. Coefficients of regression are significant at 0.1% (***) 1% (**) 5% (*) and 10%(.)

Principal Components

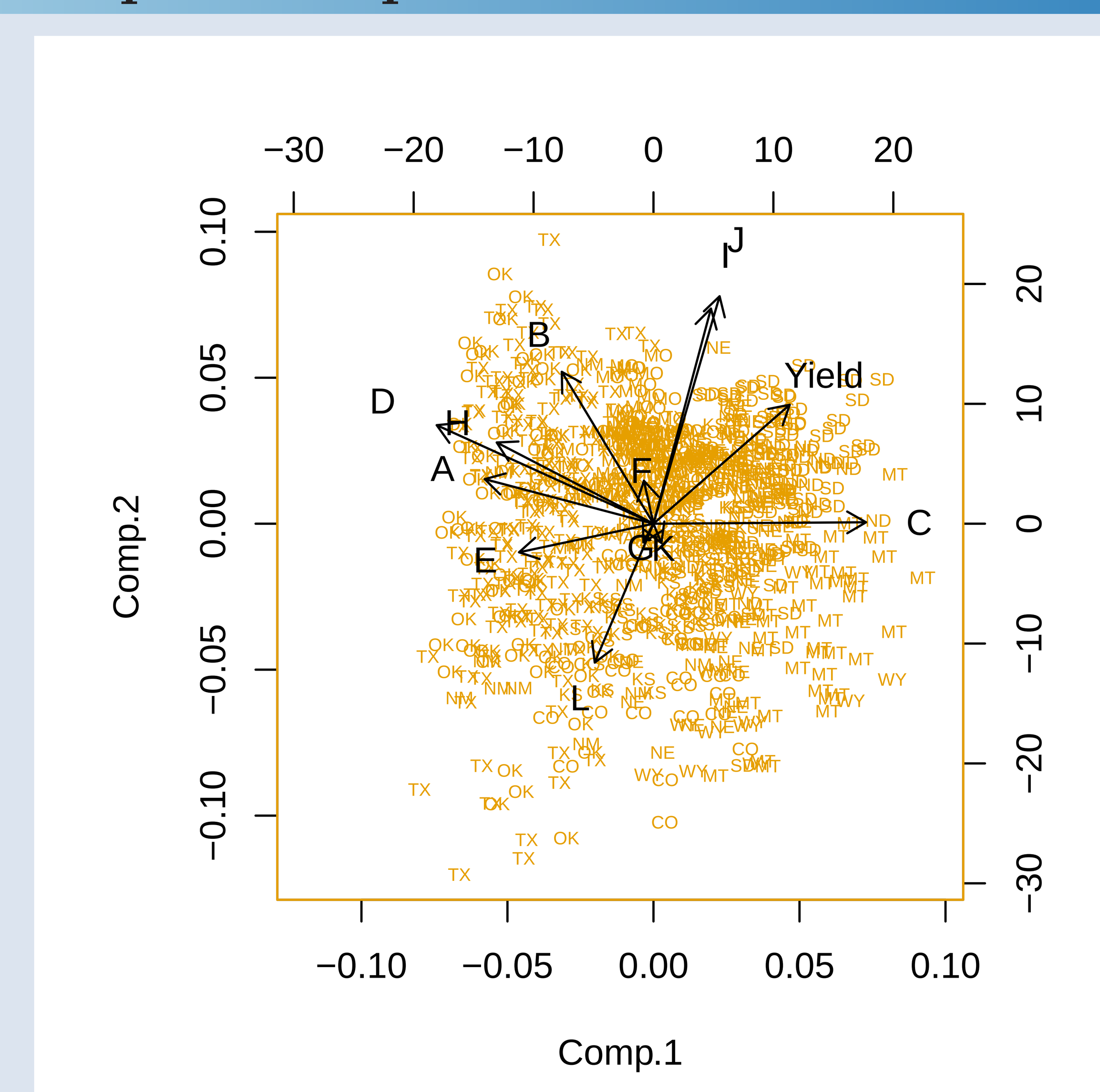

Figure 4: Principle components of covariates associated with yield changes. See Table 1 for large letter codes. Small letter codes represent data points for counties, denoted by state.