Preliminary hot spots and cold spots analysis of RMA drought-related losses

Prepared by Kelly Smith, Hasnat Aslam and Brian Fuchs, from the National Drought Mitigation Center, University of Nebraska-Lincoln, for the U.S. Department of Agriculture's Office of the Chief Economist

September 2023

Overarching questions

RQ1: How much does the USDM explain about drought-related losses?

RQ2: Can we identify areas that could be better served or that are particularly vulnerable?

From the original Statement of Work: Explore spatial patterns of drought indicators and RMA cause-of-loss data, including how "hot spots" may indicate fiscally riskier areas for crop insurance due to drought and how "cold spots" may indicate significant deficiencies in USDA programming outreach and technical assistance to communities in their uptake of important agricultural risk management programs.

Our study builds on previous work by Reyes and Elias (2019) that explored differences in patterns of loss over space and time, within the RMA cause of loss data. This study investigates the relationship between drought as measured by the U.S. Drought Monitor (Svoboda et al. 2002) and drought-related RMA claims. The U.S. Drought Monitor is not a trigger or criteria for RMA insurance.

Methods

Data

Study area

Our study area for this first stage of the project was the states in the Southwest and South Plains climate hubs: Nevada, Utah, Arizona, New Mexico, Texas, Oklahoma and Kansas, for the years 2000-2022. Differences in agriculture between the four westernmost states and the other three are readily apparent in maps. Crop cultivation is more widespread in Texas, Oklahoma and Kansas. We had a total of 530 counties in the study area.

Risk Management Agency Cause of Loss

We used Risk Management Agency Cause of Loss data to learn where producers filed claims to recoup drought-related insured losses. Drought was one of 43 causes of loss. After conferring with RMA experts, we combined losses from four categories -- Drought, Heat, Hot Wind and Failure of Irrigation -- as they are closely linked and would be difficult to separate. We also used certain program codes to pull out drought-related losses for forage, appearing as "ARPI" in this analysis.

Fields we used in our analysis were cause of loss, month and year of loss, county of loss, crop, and policies indemnified. We summarized the data for analysis by whether there was one or more policy indemnified by county, month, year, and crop.

Risk Management Agency Summary of Business

We used RMA's Summary of Business database to learn the years and counties in which different crops were insured. We inferred zeros – county-months without claims – if a crop was insured in a given county and year and if producers did not make a claim. "Missing values" appear on maps when no policies were sold for that crop in that county.

U.S. Drought Monitor

We used U.S. Drought Monitor county-level data. To convert from weekly to monthly data, we used values from the first map of each month. We assigned a category to each claim based on the worst category of drought affecting any part of the county, and calculated a yes/no "drought" variable, with moderate drought or worse being "drought," and none or abnormally dry being "no drought."

CropScape

We used CropScape, a remotely sensed dataset, to calculate estimates of how many acres of which crops are planted for each county and year. CropScape data begins in 2008. CropScape also reports crops that are cultivated twice in a year, so we regrouped CropScape data based on RMA crop descriptions.

Analyses

Data Visualization

Cropping patterns differ between states in the arid region of the Southwest Climate Hub and states in the Southern Plains Climate Hub. Fig. 1, Total Policies Earning Premium (2000-2022) shows many fewer crop insurance policies with claims filed in Nevada, Utah, Arizona and New Mexico, compared with Texas, Oklahoma and Kansas.

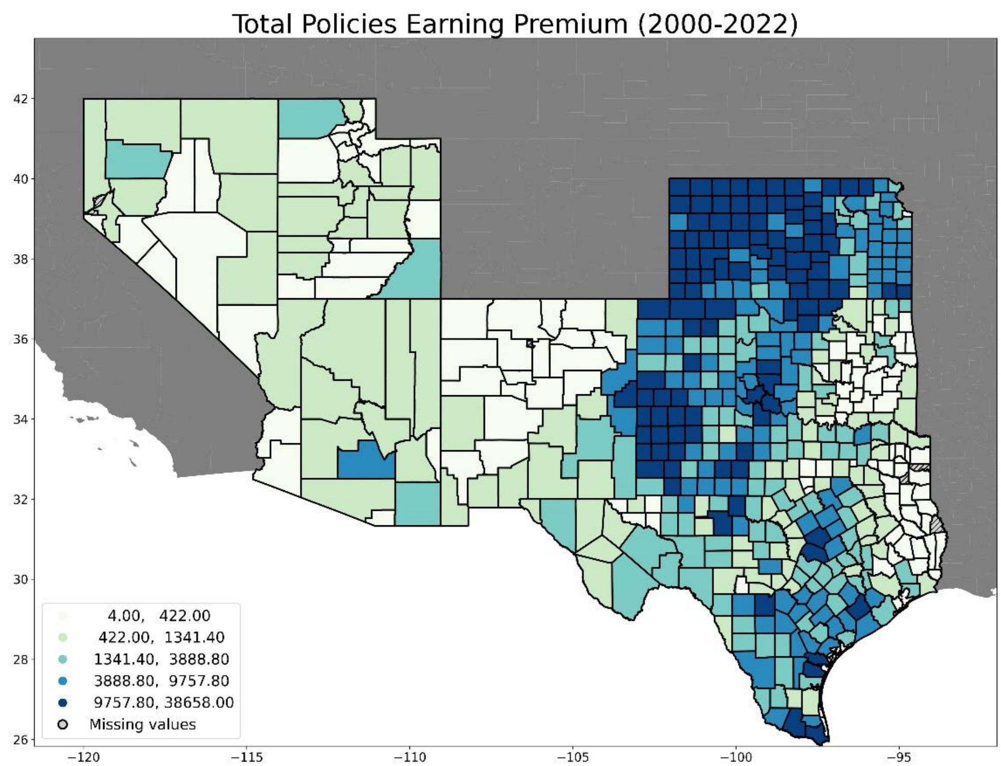

Fig. 1: The total number of claims filed by county, 2000-2022

Row crops such as wheat, corn and soy account for a higher proportion of agriculture in Kansas, Oklahoma and Texas, compared with Utah, Nevada, New Mexico and Arizona, where forage and specialty crops account for a greater share of acres planted. Fig. 2, Causes of Maximum Value of Liability normalized (adjusted for inflation) for 2000-2022 shows that loss codes related to forage and drought ("ARPI SCO …") accounted for the largest share of losses in nearly all of the counties in Nevada, Utah, Arizona and New Mexico. In contrast, drought-related losses to row crops accounted for the largest share of losses in Kansas, western Oklahoma and the Texas Panhandle.

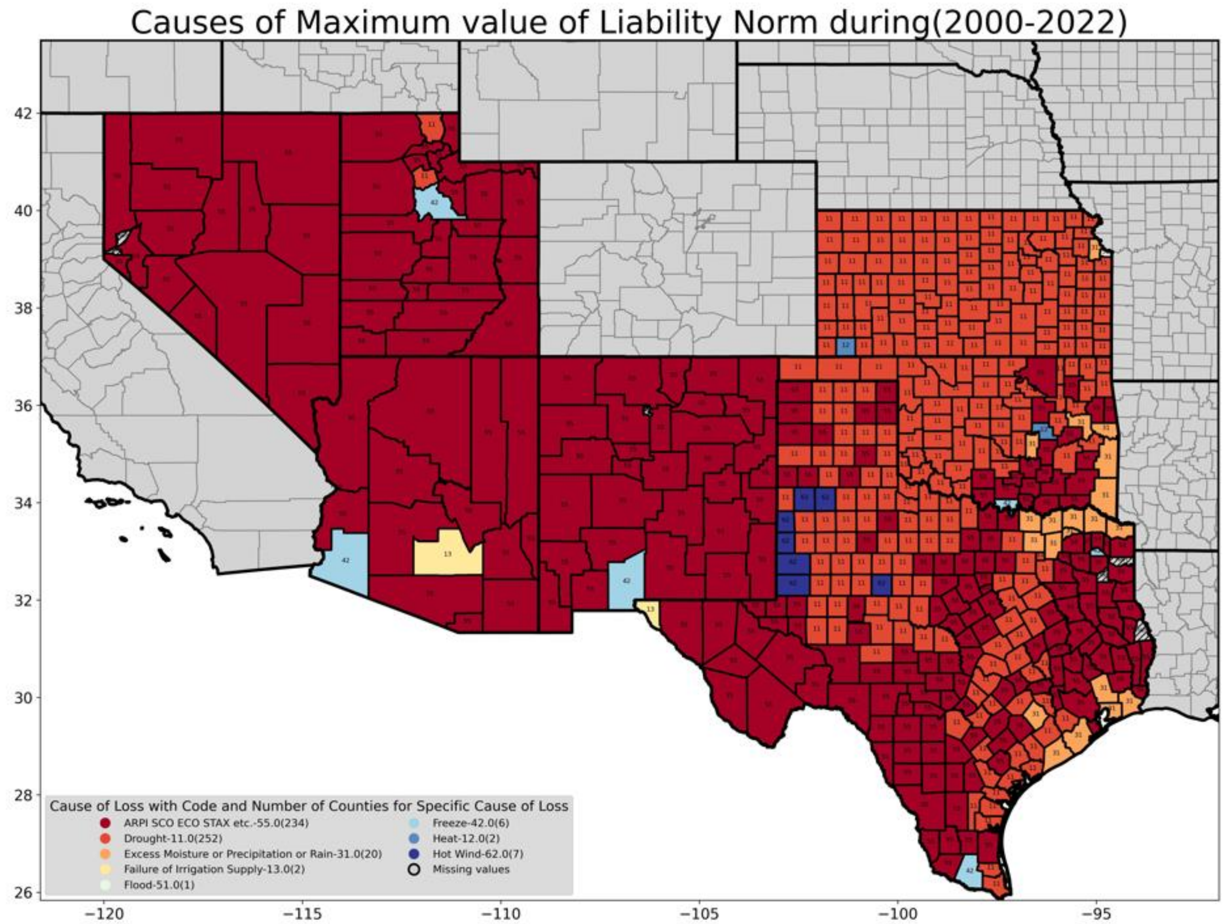

Fig. 2: Maximum causes of loss by county, 2000-2022, adjusted for inflation

These geographic differences affect what analyses are possible as well as interpretation of subsequent analyses. Generally speaking, lower numbers produce more statistical variability.

In the discussion below we focus on maps and statistics for wheat, the most widely insured crop across Texas, Oklahoma and Kansas, and for pasture, range and forage, the most widely insured crop across New Mexico, Arizona, Nevada and Utah. But we also produced maps and statistics for all of the crop data within our study area, and these are available for more detailed exploration.

Odds Ratio

To see whether the odds of drought-related claims being filed are greater during drought, we created a contingency table comparing county-months with and without drought and with and without drought-related claims (Table 1). Computing an odds ratio allowed us to characterize the odds that drought would result in claims for a given crop and county (Equation 1). A chi-square test (Equation 2) with one degree of freedom in all cases provided a measure of statistical significance as well as associated expected values and standardized residuals statistics. Expected values are calculated by multiplying the row and column totals for each cell and then dividing by the total number of cases in the table. Expected values tell us in this case how many claims there would be if drought had no effect. The chi-square test is a comparison of observed and expected values. For the chi-square test of independence to be valid, each of the expected values needs a value of at least five. We mapped odds ratios to detect spatial patterns and added p value symbology.

Table 1: Odds ratio table

	Claims		
		Yes	No
	Yes	a	b
Drought	No	c	d

a = claims during drought
b = no claims during drought
c = claims during no-drought
d = no claims during no drought

Equation 1:

$$odds\ ratio = (\frac{a}{b})/(\frac{c}{d})$$

Equation 2:

$$\chi^2 = \Sigma\frac{(O_i - E_i)^2}{E_i}$$

Where O_i is the observed value for in each cell and $E_i$is the expected value for each cell.

For example, the contingency table and chi-square values for wheat, with county counts summed for the entire state of Kansas, are shown in Table 2. The R package gmodel provides a helpful visualization of the table as well as odds ratio and chi-square statistics.

Table 2: Kansas wheat, all years, county counts aggregated to state

Cell contents:
Count
Expected values
Chi-square contribution
Standardized residuals

	Claims	No Claims	Row Total
Drought: D1-D4	10509 8610 419 20	10348 12247 294 -17	20857
Drought: None-D0	2451 4350 829	8085 6186 583	10536

	-29	24	
Column Total	12960	18433	31393

Odds ratio calculation: (10509/10348)/(2451/8185) = 1.02/0.3 = 3.4

County-level maps of the values in each cell are also informative. They provide insights on spatial patterns for counts of drought + claims, drought + no claims, no drought + claims, and no drought + no claims, as well as county-level expected values and residuals.

Comparison of Acres Cultivated and Acres Insured

CropScape accuracy varies across region, crop and year. To account for that, we looked at CropScape in five-year windows, the current year and the previous four, so our hotspots-cold spots analysis begins in 2012. For data comparability between CropScape and RMA Cause of Loss, some sub-categories in CropScape are reclassified into main crop names, i.e., hybrid corn seed and corn are combined as corn, and grain sorghum, hybrid sorghum seed and silage sorghum are combined as sorghum. This reduces the number of crops in CropScape from 41 to 31: Pistachios, Soybeans, Oranges, Canola, Sorghum, Oats, Barley, Beans, Rice, Apples, Peanuts, Grapes, Peppers, Cherries, Potatoes, Cabbage, Triticale, Cucumbers, Sunflower, Wheat, Onions, Corn, Pecans, Peaches, Sugarcane, Millet, Peas, Safflower, Alfalfa, Cotton and Rye.

We:

Calculated the maximum value of each crop area for each year and the previous four years, and joined CropScape data with RMA Summary of Business data based on year, crop and county.

Calculated the difference between insured acres and cultivated acres (maximum value) for each year, crop and county. We expressed this as a simple difference (Equation 3),

Equation 3:

$$Acres\ Difference = CropScape\ Acres - RMA\ Insured\ Acres$$

as a percentage (Equation 4),

Equation 4:

$$Percentage\ of\ Insured = \frac{RMA\ Insured\ Acres}{CropScape\ Acres} \times 100$$

and as a normalized insured index (Equation 5).

Equation 5:

$$Normalized\ Insured\ Index = \frac{CropScape\ Acres - RMA\ Insured\ Acres}{CropScape\ Acres + RMA\ Insured\ Acres}$$

To summarize the 2012-2022 data:

1. Statistics of difference in acres are calculated:
    - Maximum values

  - Minimum Values
  - First quantile (Q1 25th percentile)
  - Second quantile (Q2 50th percentile)
  - Third quantile (Q3 75th percentile)
  - (Fourth quantile would be the same as maximum values)
2. Statistics of percentage of insured acres are calculated:
  - Maximum values
  - Minimum values
  - Mean values
  - First quantile (Q1 25th percentile)
  - Second quantile (Q2 50th percentile)
  - Third quantile (Q3 75th percentile)
3. Statistics of normalized insurance index are calculated.
  - Maximum values
  - Minimum values
  - Mean values
  - First quantile (Q1 25th percentile)
  - Second quantile (Q2 50th percentile)
  - Third quantile (Q3 75th percentile)

Results

Odds Ratio for all crops

We found statistically significant relationships between months with drought according to the USDM and months when producers filed drought-related crop insurance claims, especially in Kansas, Oklahoma and Texas, where more crops insurance policies are sold. We calculated odds ratios for each crop individually and for all crops combined. We also produced a suite of maps and associated tables for each crop and for all crops combined. Fig. 3 shows the odds ratio and probabilities based on chi-square tests for all crops and forage.

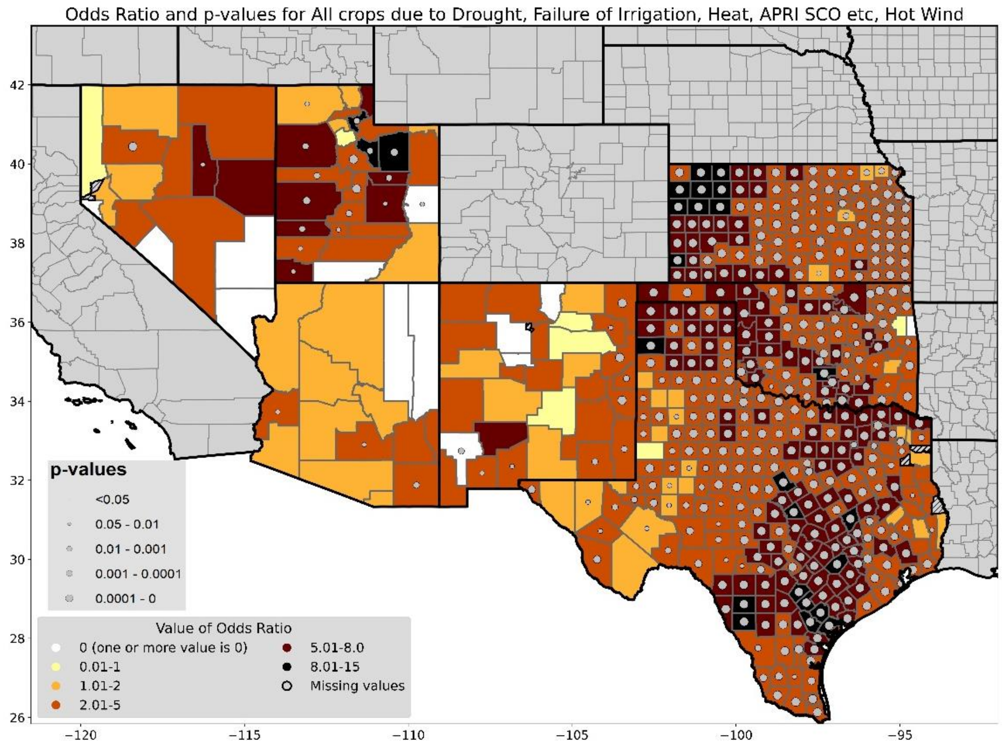

Fig. 3: Odds ratios and p-values for all crops and forage, 2000-2022

When we aggregated county-months by state for all crops, including pasture and forage, we found that odds of claims during drought were at least 3:1 in Oklahoma, Kansas, Texas and Utah, over 2:1 in Nevada, and less than 2:1 in Arizona and New Mexico. All of the state odds ratios were statistically significant. See Table 3.

Table 3: Odds ratio table aggregated to state for all crops including pasture and forage

State	Odds Ratio	Chi-square Statistic	Probability	a	b	c	d	resid.a	resid.b	resid.c	resid.d	conditions
Oklahoma	4.52	3300.172	0	16652	2951	6516	5224	17.96	-30.23	-23.21	39.07	true
Kansas	4.23	4760.196	0	45876	11455	5232	5523	13.70	-23.76	-31.62	54.86	true
Texas	3.91	8838.016	0	61649	10448	22215	14712	26.29	-47.99	-36.73	67.06	true
Utah	3.04	189.928	0	1942	179	5126	1434	5.18	-10.84	-2.94	6.16	true
Nevada	2.07	35.936	0	804	83	3290	703	2.19	-5.01	-1.03	2.36	true
Arizona	1.91	27.969	0	1153	87	2841	410	1.51	-4.29	-0.93	2.65	true
New Mexico	1.64	59.222	0	3258	362	5844	1066	2.30	-5.82	-1.67	4.21	true

Notes on table: The chi-square statistic is computed with one degree of freedom.
a = Drought + Claims
b = No Drought + Claims
c = Drought + No Claims
d = No Drought + No Claims

Conditions, TRUE = expected values are >= 5, meaning the chi-square test is statistically valid
Residuals are a relative description of how much each cell contributed to the chi-square statistic.

Wheat odds ratio

Producers insured wheat in 357 counties in the study area, mainly in Kansas, Oklahoma and Texas. Values for 290 counties met the conditions for the chi-square test of statistical significance (each expected value >=5). Seven counties in Texas, with no claims during non-drought, showed infinite odds of drought affecting claims. For five counties in Texas and one in Kansas, the odds of drought having an effect on claims were greater than 10:1. For 190 counties, the odds of drought affecting claims were between 2:1 and 10:1, and 167 of those had chi-square statistics over 3.840, the threshold for statistical significance at the 0.05 level. For the 87 counties with odds ratios of 2:1 or less, 75 had chi-square statistics under 3.840, not considered statistically significant. Odds ratios and probabilities based on the chi-square test are mapped for wheat by county in Fig. 4.

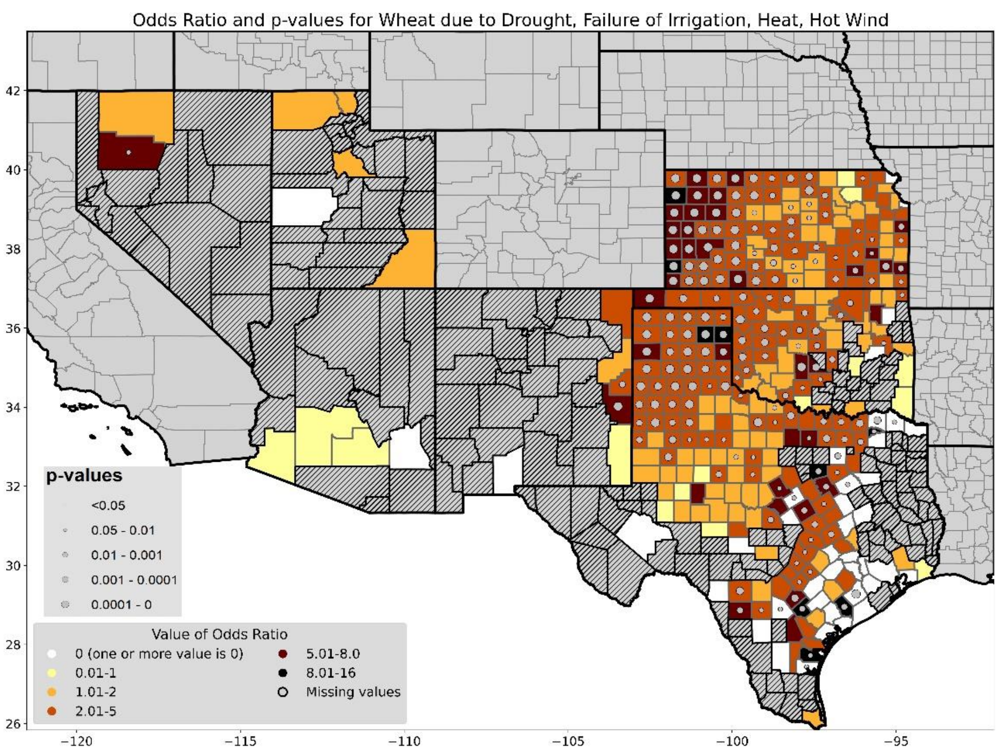

Fig. 4: Odds ratios and p-values for wheat, by county, 2000-2022

Computing odds ratios by state, we find statistically significant relationships between drought and claims for wheat in Kansas, Oklahoma, Texas and New Mexico, but not in Utah or Arizona (Table 4).

Table 4: Odds ratio aggregated by state for wheat

State	Odds Ratio	Chi-square Statistic	Probability	a	b	c	d	resid.a	resid.b	resid.c	resid.d	conditions
Nevada	4.20	5.472	0.0193	55	3	249	57	0.94	-2.12	-0.41	0.92	true
Kansas	3.35	2123.253	0.0000	10509	2451	10348	8085	20.46	-28.79	-17.16	24.14	true
Oklahoma	2.51	559.599	0.0000	4593	1109	6021	3647	10.44	-15.60	-8.02	11.98	true
Texas	2.41	1098.679	0.0000	10502	1983	23713	10800	14.82	-24.24	-8.91	14.58	true
New Mexico	1.99	21.941	0.0000	489	68	934	259	1.70	-3.54	-1.16	2.42	true
Utah	1.41	3.304	0.0691	289	46	671	151	0.66	-1.46	-0.42	0.93	true
Arizona	0.58	1.474	0.2247	43	9	887	108	-0.47	1.32	0.11	-0.30	true

Mapping the antecedent values – the number of county-months with and without drought and with and without claims – also provides some insight, particularly where values are lower and it is not possible to compute a statistically meaningful odds ratio (Fig. 5). For example, the contingency table maps for wheat show counties in Arizona with relatively high numbers of county-months with no claims filed despite the occurrence of drought.

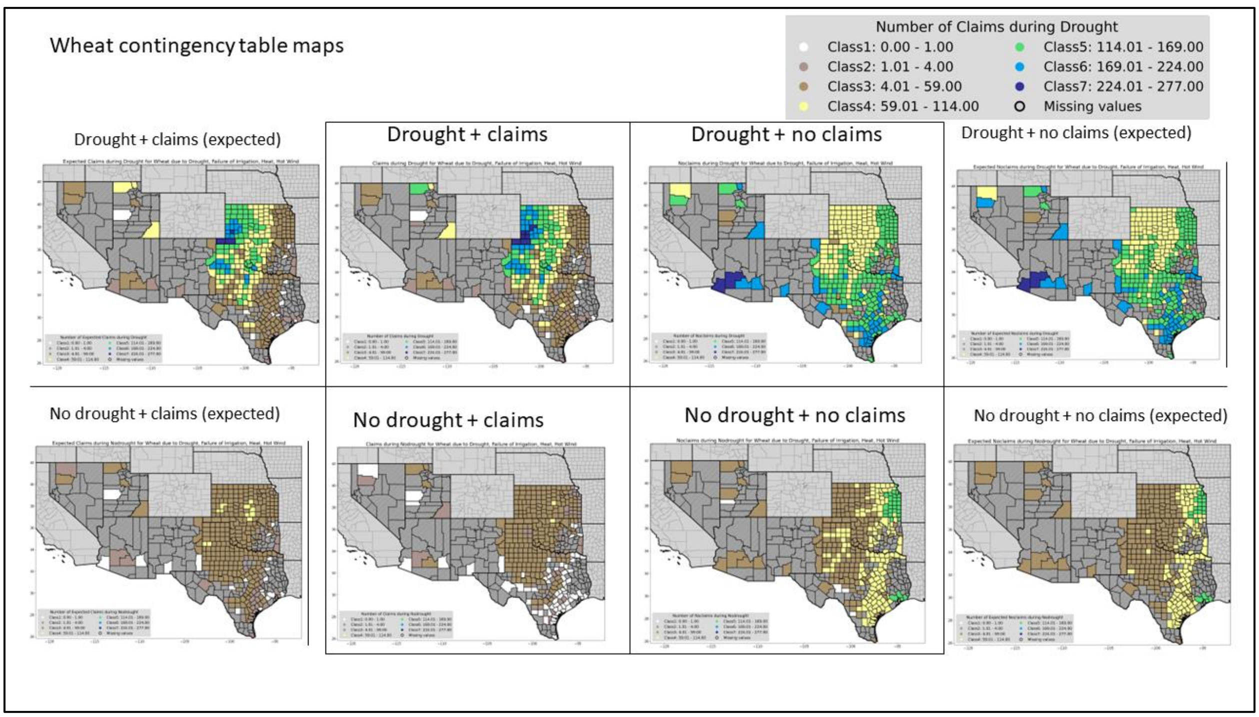

Fig. 5: Maps showing ranges of contingency table values by county for wheat, 2000-2022

Those counties are also places where, based on RMA Summary of Business data, it appears to be the norm to insure only a small portion of the acres of wheat that are planted. (Fig. 6)

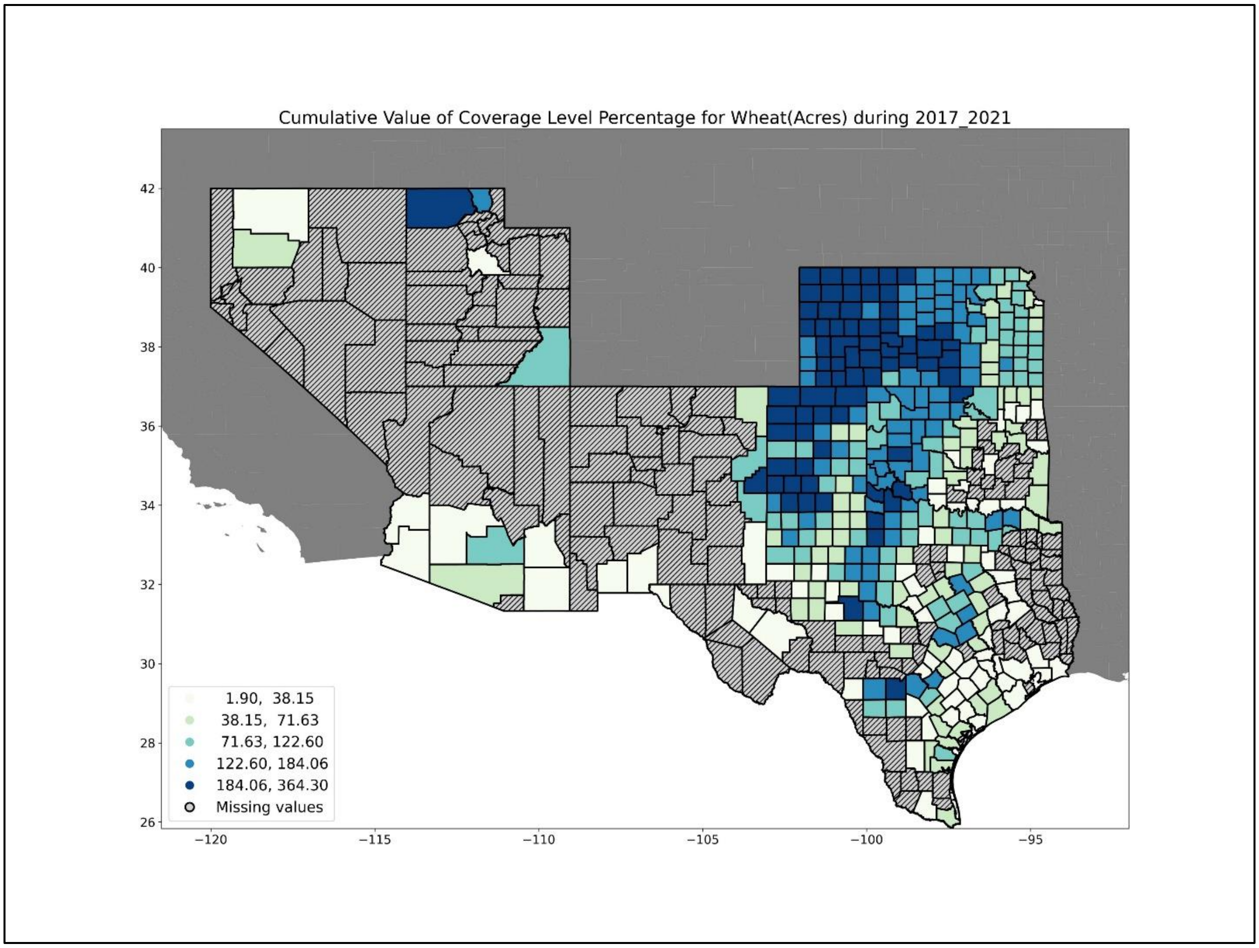

Fig. 6

Wheat: Comparisons of acres cultivated to acres insured

The comparison of acres planted to acres insured shows differences across regions and years when viewed as a percentage or index (Fig. 9), but shows a more consistent spatial pattern when viewed as the difference between acres planted and acres insured (Fig. 7). Percentage of Insured Cultivated Acres shows more variation between years, and the Normalized Insured Index, labeled as "Hotspot and Cold spot" maps, appears to be the most sensitive to changes across time (Fig. 9).

The map showing the difference in cultivated and insured acres (Fig.7) shows a consistent "cold spot" – acres that are planted but not insured – in central Oklahoma and in part of the Texas Panhandle. This spatial characterization is true for each of the 11 years from 2012 to 2022. The pattern in the map of the difference between cultivated and insured acres relates to the planted acreage of wheat (Fig. 8).

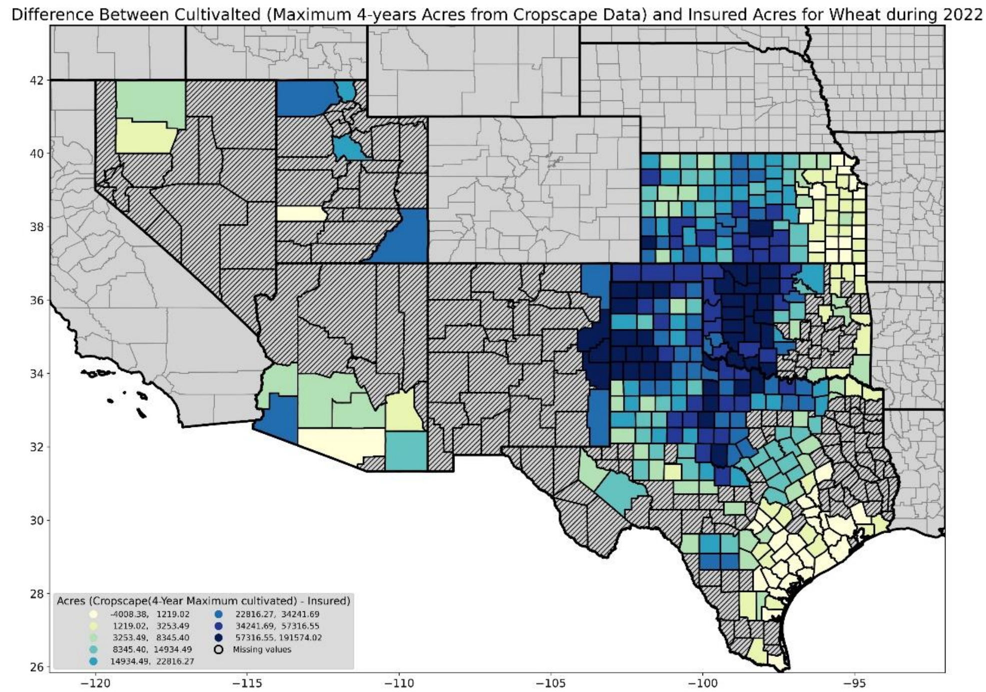

Fig. 7: Difference in cultivated and insured acres for wheat, 2022

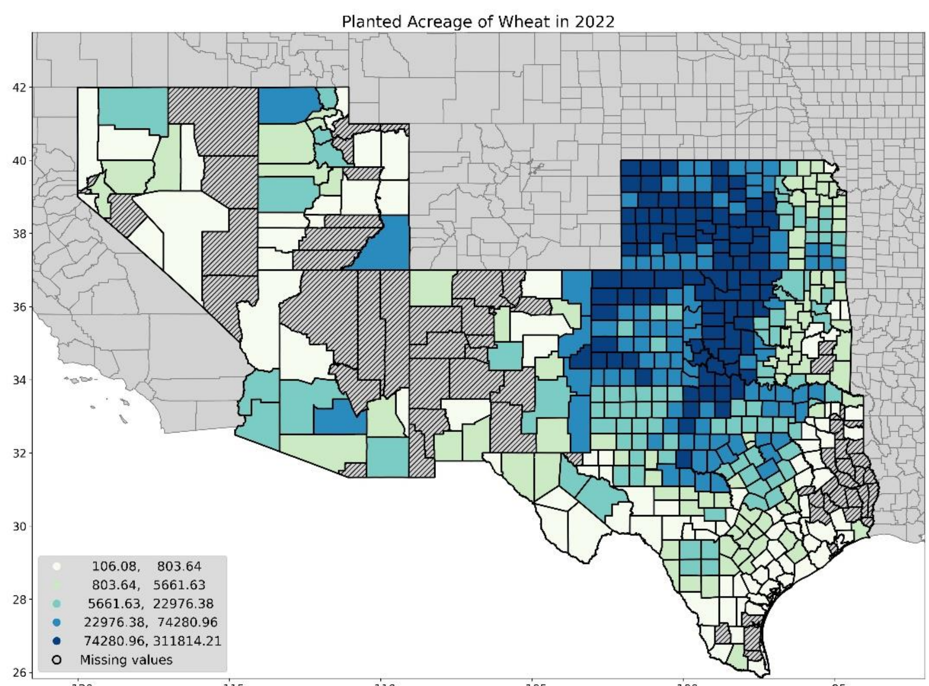

Fig. 8: Planted acreage of wheat in 2022, from RMA's Summary of Business data

Mapping the difference as a percentage or as the Normalized Insured Index (hotspots and cold spots) for wheat shows variation by region and year (Fig. 9). Maps of the Normalized Insured Index depict greater differences than maps of the percentage of planted acres, as shown in a comparison of select years. Although three more recent years in the study period – 2018, 2019 and 2022 – seem to show "cooler" values, a trend is not discernable.

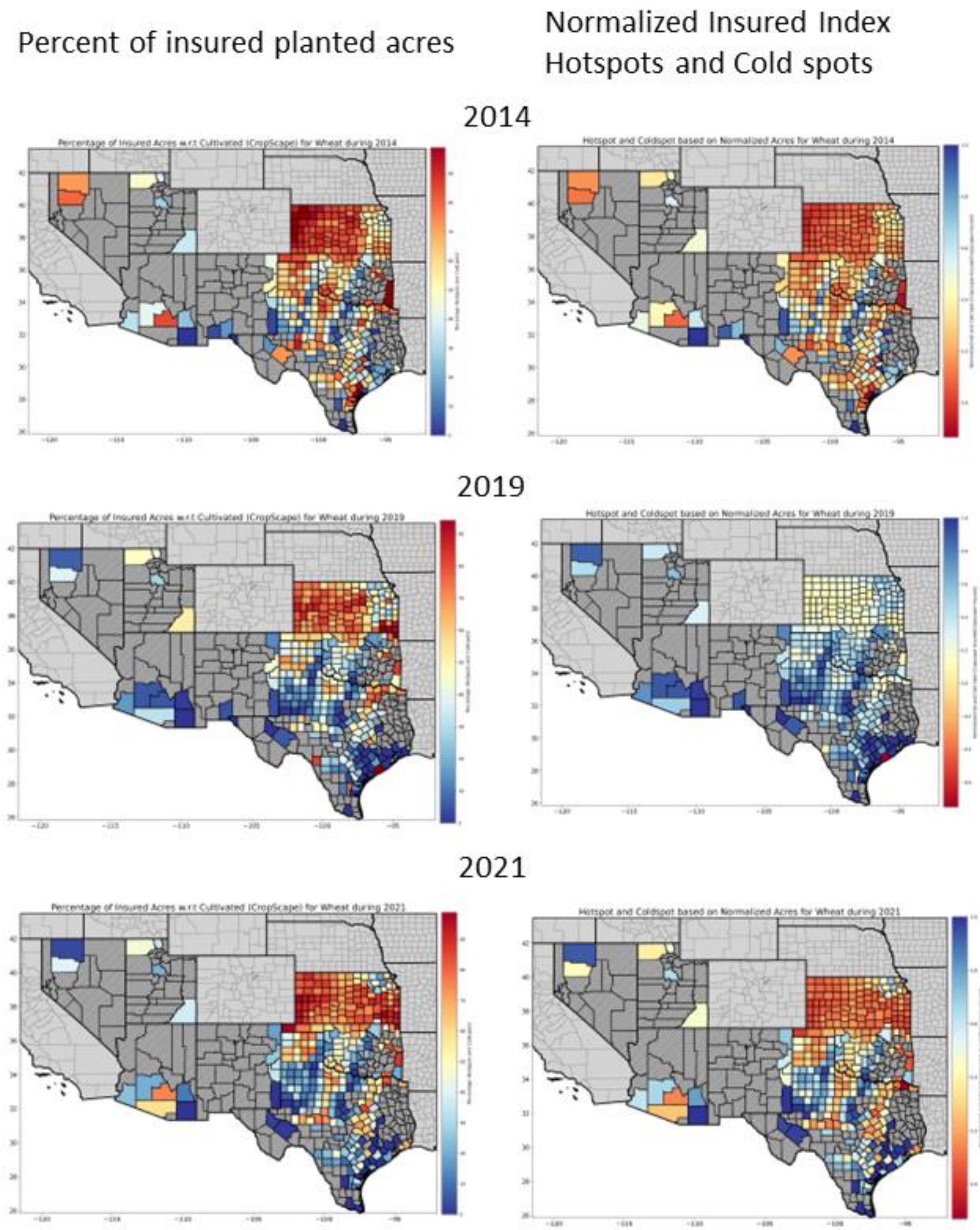

Fig. 9: Comparison of select years, percentage of insured cultivated acres with Normalized Insured Index, for wheat

Comparisons of maximum and minimum values for the percentage of insured planted acres and for the Normalized Insured Index also show greater variation in the latter (Fig. 10). Further research and analysis can determine whether these differences are an unintended product of data or methodology limitations, whether they are meaningful, and how they may be of use.

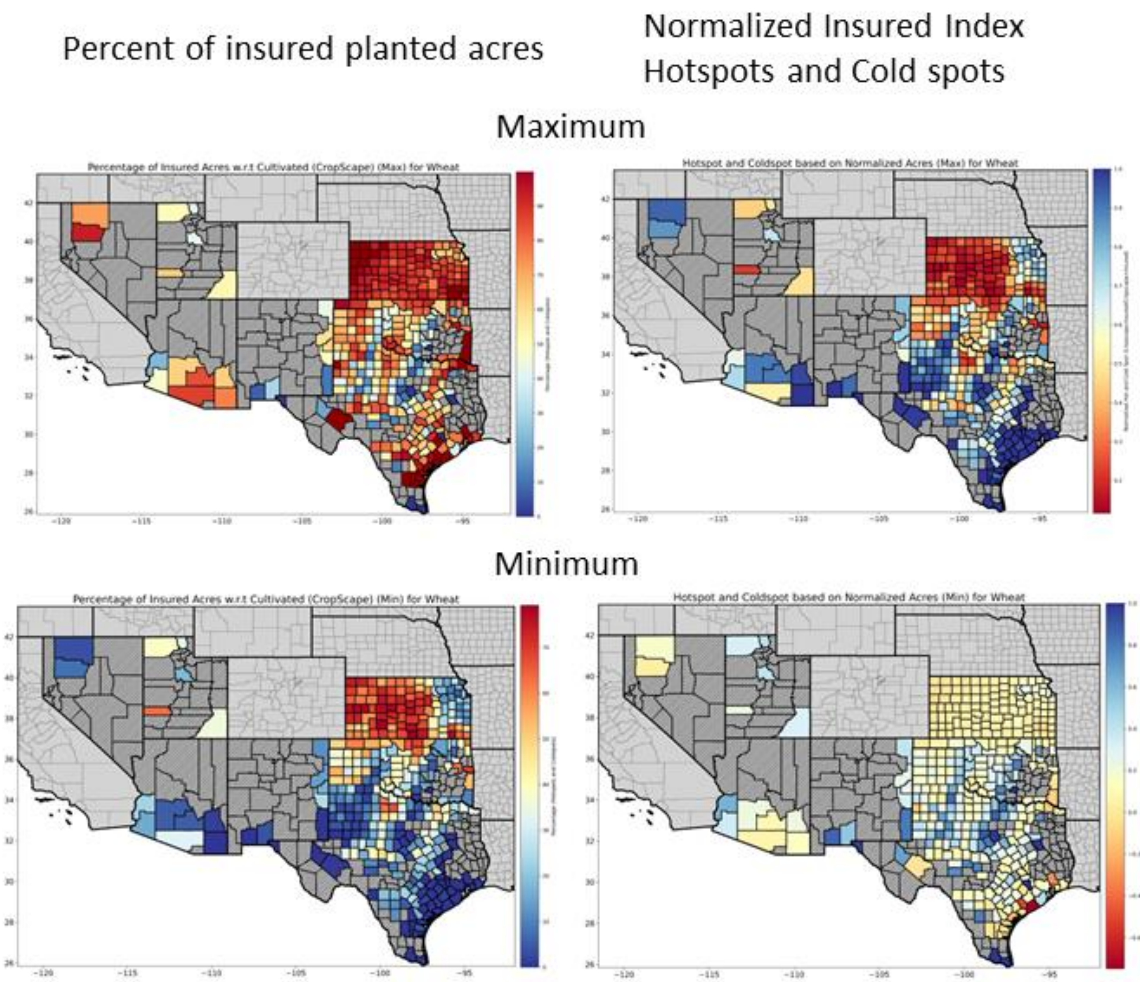

Fig. 10: Comparison of maximum and minimum values of percentage of insured cultivated acres with Normalized Insured Index for wheat

Pasture and forage odds ratio

Producers insured pasture and forage in 438 of the 530 counties in our study area (Fig. 11). Conditions for computing the chi-square test of statistical significance were met in 371 counties. The odds of drought affecting claims were infinite in Greeley County, Kansas, where no producers filed drought-related claims during non-drought months. For eight counties in Kansas and Oklahoma, odds were greater than 11:1 that drought affected claims. The odds were between 2:1 and 10:1 for 291 counties, with 226 having a chi-square statistic over 3.840, the threshold for 0.05 significance. None of the 71 counties with odds ratios less than 2:1 met the criteria for statistical significance.

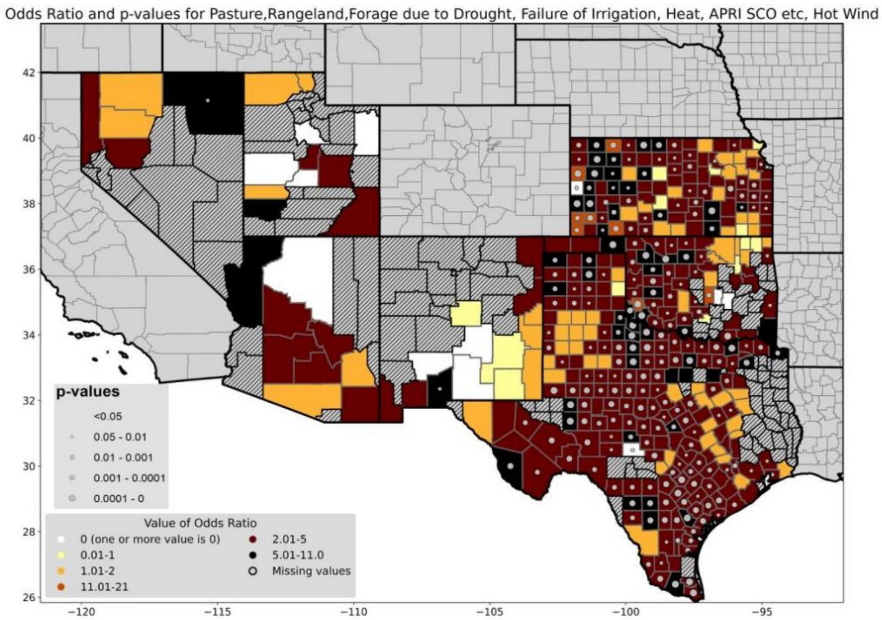

Fig. 11: Odds ratio and p-values for pasture and forage

Although numbers were too low in the four western states of our study area to perform a statistically valid odds ratio test for each county, aggregating by state and calculating odds ratios finds statistically significant relationships, ranging from 3.56:1 in Oklahoma to 1.89:1 in New Mexico (Table 5).

Table 5: Odds ratio for pasture and forage aggregated by state

State	Odds Ratio	Chi-square Statistic	Probability	a	b	c	d	resid.a	resid.b	resid.c	resid.d	conditions
Oklahoma	3.56	413.440	0.0000	2606	358	3000	1467	7.82	-13.71	-6.37	11.17	true
Texas	3.10	1658.806	0.0000	14947	2198	10851	4953	13.15	-24.97	-13.69	26.01	true
Kansas	2.94	572.704	0.0000	3642	756	4998	3048	10.65	-16.05	-7.87	11.86	true
Arizona	2.75	14.915	0.0001	428	24	305	47	0.78	-2.52	-0.89	2.85	true
Utah	2.61	10.604	0.0011	502	22	262	30	0.51	-1.97	-0.69	2.64	true
Nevada	2.47	9.367	0.0022	303	25	147	30	0.63	-1.79	-0.85	2.44	true
New Mexico	1.89	14.591	0.0001	770	67	652	107	0.89	-2.54	-0.93	2.67	true

Maps of the number of claims submitted and not submitted for forage during drought and non-drought months provide insight on spatial patterns (Fig. 12).

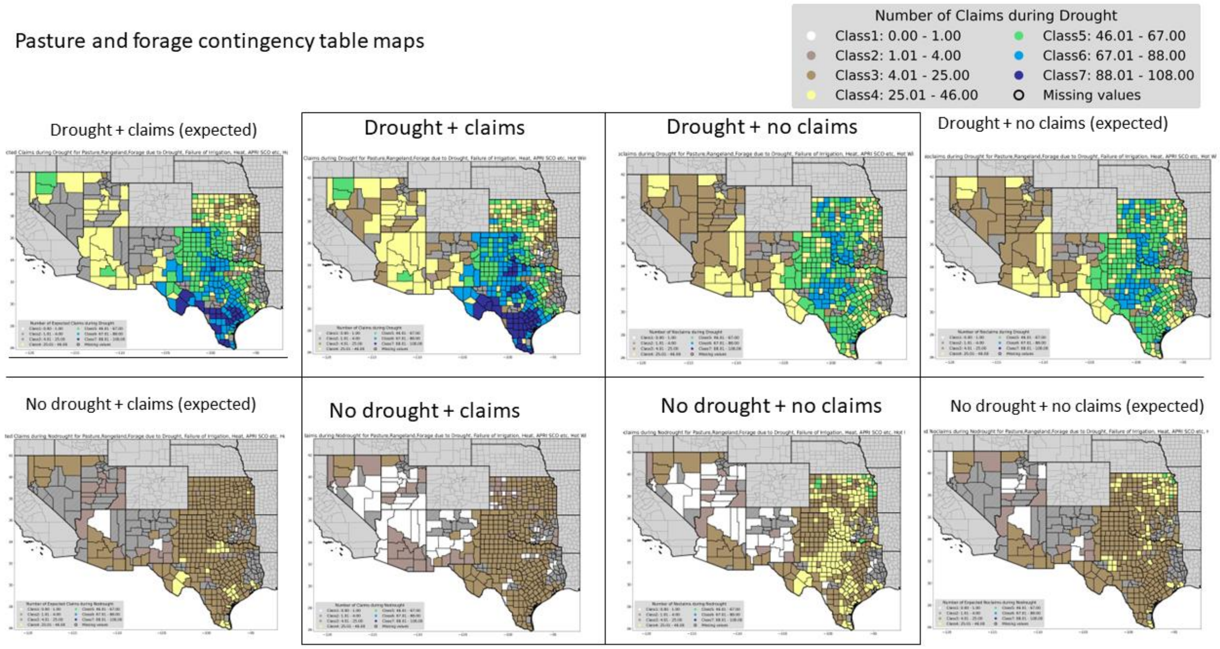

Fig. 12: Maps showing ranges of contingency table values for pasture and forage by county

Forage: Comparisons of acres cultivated to acres insured

The same analyses of acres planted and insured are not possible for pasture and forage with the data we used, i.e., we did not have comparable data from CropScape. The acres of pasture, rangeland and forage acres insured in 2022, according to RMA's Summary of Business data, was lower in the four western states of our study area, with the exception of eastern New Mexico (Fig. 13).

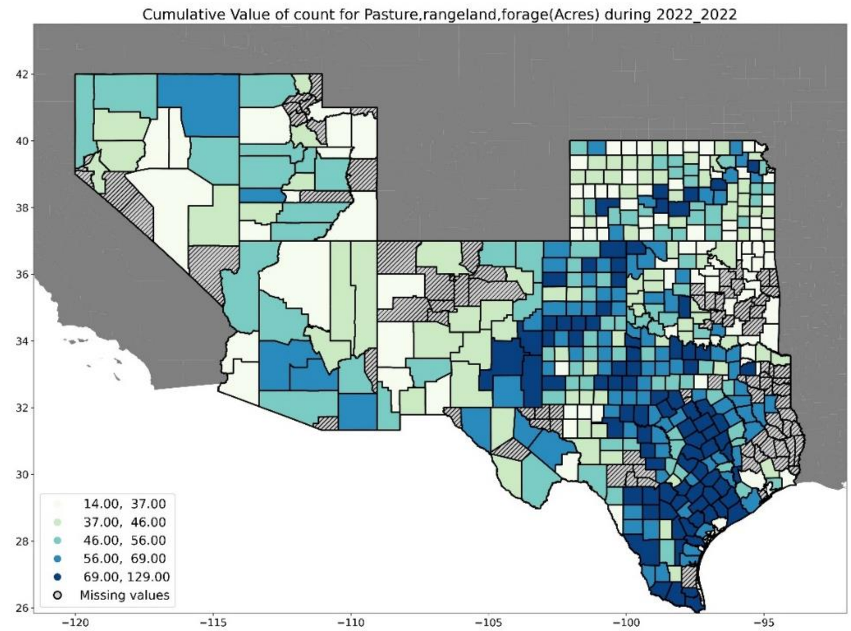

Fig. 13: Pasture acres insured in 2022, from RMA's Summary of Business data

Discussion

Wheat

Both the odds ratio analysis (Fig. 4) and the comparison of acres planted to acres insured (Fig. 9 and 10) suggest that much of western Kansas may be a "hotspot" for wheat, with the majority of planted acres insured, and with higher odds that drought-related claims will be filed in months when there is drought, according to the U.S. Drought Monitor.

Interpreting cold spots is more challenging because they seem to occur in areas where total planted acreage is less, and lower numbers can produce more variability. Areas such as the southern part of the Texas Panhandle and eastern New Mexico may be cold spots for wheat, consistently showing a lower percentage of planted acres that are also insured. Total acres planted are lower along the Texas Gulf Coast and in southern Arizona, which may contribute to higher variability. Those areas show both high and low proportions of insured acres over the years, in contrast to Kansas, which ranges from high to middling proportions of insured acres, with none on the low end.

Pasture

South Texas may be a hotspot for pasture and forage, based on the acres insured (Fig. 13), despite smaller county areas, and on the increased number of claims during drought, in contrast to other regions. The odds ratio map (Fig. 11) suggests a larger hotspot for pasture and forage insurance uptake and use, in areas of Texas, Oklahoma and Kansas.

Considerations, Limits and Caveats

It is easier to identify hot spots – areas with more insurance uptake and more frequent claims during drought – than cold spots – areas "that may indicate significant deficiencies in USDA programming outreach and technical assistance to communities in their uptake of important agricultural risk management programs."

Producers clearly buy less insurance in Nevada, Utah, Arizona and New Mexico than in Texas, Oklahoma and Kansas. Whether that represents a "deficiency in USDA programming and technical assistance" or rational economically driven decision-making in an arid climate is beyond the scope of this analysis. The RMA experts we consulted also mentioned that in some instances, producer preferences for percentage of coverage (deductible), vary by region.

NDMC staff note that the claims resolution process for pasture, rangeland and forage differs from that for row crops. Rather than sending a claims adjuster out to verify a loss, forage claims are confirmed using Climate Prediction Center precipitation data, which has some caveats related to interpolating point data. Conversations with livestock producers in the Southwest suggest that many do not believe this process is effective.

The Normalized Insurance Index or Hotspot-Cold spot analysis shows spatial and interannual variation in the proportion of planted acres that are insured, although further research is needed to determine the cause. A few possibilities include conditions when producers are making insurance purchase decisions such as commodity prices, whether it's already a dry year, and programmatic incentives or initiatives.

As a preliminary stage of research, our unit of analysis was county-months, converting data to presence-absence – whether or not a claim was "present" in a given county during the month – to avoid the additional complexity of summarizing quantified monthly data, which are typically summarized as annual data. Using the monthly information in RMA data provided a sufficiently fine temporal resolution to compare it with U.S. Drought Monitor data, which are weekly. However, an analysis of quantified losses may provide richer and more detailed insights. For example, it is possible that a claim early in the year accounted for a producer's entire acreage, leaving none to be claimed in later months, and leading to more instances of no claims despite drought occurring.

It is also possible that indicators other than the U.S. Drought Monitor would correlate more closely to conditions that producers experience in various regions.

Next Steps

Expansion of first stage: Expand odds ratio and hot spots-cold spots analysis to other parts of the country

Explore combining crop types to produce a comparison of acres planted to acres insured for pasture and forage

Quantitative analysis:

- Use RMA Summary of Business data to get total liabilities for the year for the crop and county.
- Can infer zeros if there were liabilities without indemnities.
- Exercise caution (or aggregate to annual) in comparing monthly indemnity data with annual liability data. A producer may have more than one indemnity against a total liability.

Qualitative: Interview producers, insurers and others to understand producers' decisions on insurance. Engage FSA staff and others to help ground-truth and interpret maps.

Demographic: Compare demographic data with insurance uptake, if data exists to go beyond the RMA's earlier analysis. An RMA Report to Congress, "Adequate Coverage for States and Underserved Producers," mandated by the 2018 Farm Bill, defines adequate coverage in relative terms, as a state having 50% or more of the national participation level, and an underserved producer as a beginning farmer or rancher, a veteran farmer or rancher, or a socially disadvantaged farmer or rancher, including members of Indian Tribes. The same report found that beginning farmers and veterans had the highest rates of insurance uptake of any of the groups. Women, Black, Asian, Hispanic, Native American/Indian and Native Hawaiian and other Pacific Islanders all had lower rates of uptake than beginning or veteran farmers.

References

Reyes, J. J., & Elias, E. (2019). Spatio-temporal variation of crop loss in the United States from 2001 to 2016. Environmental Research Letters, 14(7), 074017. https://doi.org/10.1088/1748-9326/ab1ac9

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D., & Stephens, S. (2002). The Drought Monitor. Bulletin of the American Meteorological Society, 83(8), 1181–1190. https://doi.org/10.1175/1520-0477-83.8.1181