Climate-Driven Doubling of U.S. Maize Loss Probability: Interactive Simulation with Neural Network Monte Carlo

A. Samuel Pottinger
University of California, Berkeley

Lawson Connor
University of Arkansas, Fayetteville

Brookie Guzder-Williams
University of California, Berkeley

Maya Weltman-Fahs
University of California, Berkeley

Nick Gondek
University of California, Berkeley

Timothy Bowles
University of California, Berkeley

Abstract

Climate change not only threatens agricultural producers but also strains related public agencies and financial institutions. These important food system actors include government entities tasked with insuring grower livelihoods and supporting response to continued global warming. We examine future risk within the U.S. Corn Belt geographic region for one such crucial institution: the U.S. Federal Crop Insurance Program. Specifically, we predict the impacts of climate-driven crop loss at a policy-salient "risk unit" scale. Built through our presented neural network Monte Carlo method, simulations anticipate both more frequent and more severe losses that would result in a costly doubling in the annual probability of maize Yield Protection insurance claims at mid-century. We also provide a configurable open source pipeline and interactive visualization tools to further explore these results. Altogether, we fill an important gap in current understanding for climate adaptation by bridging existing historic yield estimation and climate projection to predict crop loss metrics at policy-relevant granularity.

1. Introduction

Public institutions such as government-supported crop insurance play an important role in agricultural stability (Mahul and Stutley 2010). To support climate adaptation efforts, we provide a neural network Monte Carlo method which we use to examine the U.S. Federal Crop Insurance Program inside the U.S. Corn Belt geographic region. Adding to existing work regarding global warming impacts to these essential food systems actors (Diffenbaugh et al. 2021), we build upon prior climate projections (Williams et al. 2024) and remote sensing yield estimations (Lobell et al. 2015) to predict future insurance indemnity claims at an institutionally-relevant spatial scale.

1.1. Background

Global warming threatens production of key staple crops, including maize (Rezaei et al. 2023). Climate variability already drives a substantial proportion of year-to-year crop yield variation (Ray et al. 2015) and continued climate change may reduce planet-wide maize yields by up to 24% by the end of this century (Jagermeyr et al. 2021). The growing frequency and severity of stressful weather conditions (Dai 2013) to which maize is increasingly susceptible (Lobell et al. 2020) pose not only a threat to farmers' revenue (Sajid et al. 2023) but also strain the institutions established to safeguard those producers (Hanrahan 2024). These important organizations are also often tasked with supporting food systems through evolving growing conditions and climate change (RMA 2022).

Within this context, the United States of America is the world's largest maize producer and exporter (Ates 2023). Its government-backed Federal Crop Insurance Program covers a large share of this growing risk (Tsiboe and Turner 2023). The costs of crop insurance in the U.S. have already increased by 500% since the early 2000s with annual indemnities reaching $19B in 2022 (Schechinger 2023). Furthermore, retrospective analysis attributes 19% of "national-level crop insurance losses" between 1991 and 2017 to climate warming, an estimate rising to 47% during the drought-stricken 2012 growing season (Diffenbaugh et al. 2021). Looking forward, Li et al. (2022) show progressively higher U.S. maize loss rates as warming elevates.

1.2. Prior work

Modeling possible changes in frequency and severity of crop loss events that trigger indemnity claims is an important step to prepare for the future impacts of global warming. Related studies have predicted changes in crop yields at broad scales such as the county-level (Leng and Hall 2020) and have estimated climate change impacts to U.S. maize within whole-sector or whole-economy analysis (Hsiang et al. 2017). These efforts include traditional statistical models (Lobell and Burke 2010) as well as an increasing body of work favoring machine learning approaches (Leng and Hall 2020). Finally, the literature also consider how practice-specific insurance subsidies intersect

with grower practices (Connor et al. 2022; Wang et al. 2021; Chemeris et al. 2022) and observed resilience (Renwick et al. 2021; Manski et al. 2024).

Despite these prior contributions, insurance instruments often use highly localized variables such as an individual farm's last ten years of yield for a specific crop (RMA 2008). Therefore, to inform policy, research must include more geographically comprehensive granular models than previous studies (Leng and Hall 2020) and, in addition to predicting yield (Lobell et al. 2015; Jagermeyr et al. 2021; Khaki and Wang 2019; Ma et al. 2024), need to simulate insurance instrument mechanics. Of particular interest, we fill a need for climate-aware simulations of loss probability and severity within a "risk" or "insured" unit, a geographic scale referring to a set of agricultural fields that are insured together (FCIC 2020).

1.3. Contribution

We address this need for granular future loss prediction through neural network Monte Carlo. We provide these projections at the policy-relevant risk unit scale, probabilistically forecasting institutional outcome metrics under climate change with a focus on the U.S. Corn Belt. This nine-state geographic region inside the United States is essential to the nation's maize crop (Green et al. 2018). We specifically model the Yield Protection plan, one of the options under the popular Multi-Peril Crop Insurance Program (RMA 2024). Furthermore, by contrasting predictions at approximately one and three decades to a "counterfactual" which does not include further climate warming, we quantitatively highlight insurer-relevant effects of climate change.

2. Methods

To predict the probability and severity of indemnity claims, we first model maize yield loss metrics using a neural network at an insurer-relevant spatial scale before simulating changes under different climate conditions with Monte Carlo.

2.1. Definitions

Insurers pay out based on the magnitude of a yield loss across the aggregation of all fields in an insured unit. This covered loss (l) is the difference between actual yield (y_{actual}) and a coverage level (c) percentage of an expected yield ($y_{expected}$) (RMA 2008):

$$l = \max(c * y_{expected} - y_{actual}, 0). \tag{1}$$

Typically, $y_{expected}$ is the average of the most recent ten years of yield for the insured crop (RMA 2008) as reported by the producer:

$$y_{expected} = \frac{y_{historic}[-d:]}{d}. \tag{2}$$

Next, we define p_l as probability of a loss that may incur a Yield Protection claim:

$$p_l = P\left(\frac{y_{actual} - y_{expected}}{y_{expected}} < c - 1\right) = P(y_{\Delta\%} < c - 1). \tag{3}$$

Generally, the severity (s) of a loss when it occurs defines the size of the claim:

$$s = \max(-1 * y_{\Delta\%} - (1 - c), 0). \tag{4}$$

We simulate the more common (FCIC 2023) 75% coverage limit ($c = 0.75$) but our interactive tools explore other coverage levels.

2.2. Data

As Yield Protection operates at the level of a risk unit, modeling these formulations requires highly local yield and climate information. Therefore, we use the Scalable Crop Yield Mapper (SCYM) from Lobell et al. (2015) which, from 1999 to 2022 at 30 meter resolution across the US Corn Belt, derives yield estimates from remote sensing and benefits from substantial validation efforts (Deines et al. 2021). Meanwhile, we use climate data from CHC-CMIP6 (Williams et al. 2024) which, at daily 0.05 degree or approximately five kilometer scale, offers both historic data on growing conditions from 1983 to 2016 as well as future projections. Specifically, CHC-CMIP offers both a 2030 and 2050 series which each contain multiple years[1]. This provides the following daily climate variables for modeling: precipitation, temperature (min and max), relative humidity (average, peak), heat index, wet bulb temperature, vapor pressure deficit, and saturation vapor pressure. Note that we prefer SCYM over recent alternatives (Ma et al. 2024) given temporal overlap with CHC-CMIP6.

Neighborhoods

We align these variables to a common grid in order to create discrete instances needed for model training and evaluation. More specifically, we create "neighborhoods" (Manski et al. 2024) of geographically proximate fields paired with climate data through four character geohashing[2] (Niemeyer 2008). After neural networks predict neighborhood yield distributions, we simulate risk units within each of these cells by sampling SCYM pixels within each neighborhood to approximate risk unit size and portfolio effects.

Yield deltas

Having created these spatial groups, we model against SCYM-observed deviations from yield expectations (($y_{actual} - y_{expected})/y_{expected}$) which can be used to calculate loss probability (l) and severity (s). Reflecting the mechanics of Yield Protection policies, this step converts to a distribution of changes or "yield deltas" relative to the average production histories.

---

[1]In choosing from its two available scenarios, we prefer the "intermediate" SSP245 within CHC-CMIP6 over the extreme SSP585 per the advice of Hausfather and Peters (2020). Note that, due to the available timeseries, simulations can confirm if observed effects worsen under continued warming.

[2]This algorithm creates hierarchical grid cells where each point is assigned a unique string through hashing. The first four characters identifies a grid cell (approx 28 by 20 km) which contains all points with the same first four characters of their geohash. Our interactive tools evaluate alternative neighborhood sizes (number of geohash characters).

2.3. Regression

We next build predictive models for distributions of yield deltas.

Input vector

To predict yield delta distributions per year ahead of Monte Carlo simulations, we describe each of the nine CHC-CMIP6 variables as min, max, mean, and standard deviation of each month's daily values. We also input year and baseline variability in the form of neighborhood historic absolute yield mean ($y_{\mu-historic}$) and standard deviation ($y_{\sigma-historic}$). See interactive tools for further exploration.

Response vector

Prior work suggests that yields follow a beta distribution (Nelson 1990) but the expected shape of changes to yield (yield deltas) is unknown. Therefore, our open source pipeline can predict shape parameters for either a normal distribution or beta distribution. We choose the appropriate shape by calculating the skew and kurtosis of the observed yield deltas, using the normal distribution if meeting approximate normality per Kim (2013) or beta distribution otherwise. We predict two parameters for normal (mean, std) and four for beta (center, scale, a, b) (**SciPy** 2024). This use of summary statistics helps ensure appropriate dimensionality for the dataset size (Alwosheel et al. 2018).

Task

Our regressors (f) use neighborhood-level climate variables (C) and historic yield information to predict future yield changes ($y_{\Delta\%}$) per year. We preprocess these inputs using z-score (Kim et al. 2024):

$$f(C_z, y_{\mu-historic-z}, y_{\sigma-historic-z}) \mathrel{\hat{=}} y_{\Delta\%}(x) = \frac{y_{actual} - y_{expected}}{y_{expected}}. \tag{5}$$

Neural network

Machine learning performs well in prior studies on related problems (Leng and Hall 2020; van Klompenburg et al. 2020). Among those options, we use feed forward neural networks as they support multi-variable output within a single model (Brownlee 2020a). This approach may also offer better out-of-sample estimation (Mwiti 2023). With that in mind, we "grid search" (Joseph 2018) in order to find a suitable combination of model hyper-parameters, trying permutations from Table 1 before retraining on all available data ahead of simulations. We use AdamW (Kingma and Ba 2014; Loshchilov and Hutter 2017) and non-output neurons use Leaky ReLU activation per Maas et al. (2013). We instance weight by the neighborhood maize growing acreage. Finally, we contrast our results to analogous sweeps of Gaussian Process (Pedregosa et al. 2011) and LSTM (Hochreiter and Schmidhuber 1997) as described in the appendix.

2.4. Evaluation

Across feed forward, LSTM, and Gaussian Process sweeps, we choose our model using each candidate's ability to predict into future years, a task representative of the Monte

Table 1: Parameters permuted to find an optimal configuration.

Param	Options	Description	Purpose
Layers	1 − 6	Num feed forward layers to include. Two layers include 32 and then 8 nodes. Layer sizes are {512, 256, 128, 64, 32, 8}.	More layers might allow networks to learn more sophisticated behaviors but also might overfit to input data.
Dropout	0.00, 0.01, 0.05, 0.10, 0.50	This dropout rate applies across all hidden layers.	Random disabling of neurons may address overfitting.
L2	0.00, 0.05, 0.10, 0.15, 0.20	The L2 regularization strength to apply across all hidden layer neuron connections.	Penalizing networks with edges that are "very strong" may confront overfitting without changing the structure of the network itself.
Attr Drop	9	Retraining where the sweep individually drops each of the input distributions or year or keeps all inputs.	Removing attributes helps determine if an input may be unhelpful.

Table 2: Post-hoc trials after model selection.

Trial	Evaluate	Train	Test
Random	Ability to predict generally.	Random 75% of year / geohash combinations.	The remaining 25% of year / region combinations.
Temporal	Ability to predict into future years.	All data from 1999 to 2013 inclusive.	All data 2014 to 2016 inclusive.
Spatial	Ability to predict into unseen geographic areas.	All four character geohashes in a randomly chosen 75% of three character regions.	Remaining 25% of regions.
Climatic	Ability to predict into out of sample growing conditions.	All years but 2012.	2012 (unusually dry / hot)

Carlo simulations. Specifically, we **train** on data between 1999 to 2012 inclusive, **validate** on 2014 and 2016 to compare candidates from grid search, and **test** on 2013 and 2015 which serve as a fully hidden set estimating future performance (Brownlee 2020b).

Having performed model selection, we further evaluate our chosen regressor through tests in Table 2 which estimate performance in different tasks one may consider using this method. These post-hoc trials use only training and test sets as we fully retrain models using unchanging sweep-chosen hyper-parameters as described in Table 1. Note that some of these tests use "regions" which we define as all geohashes sharing the

same first three characters, creating a grid of 109 x 156 km cells (Haugen 2020) each including all neighborhoods (four character geohashes) found within. Finally, lacking a prior study with an identical response variable, we compare these results to recent literature performing absolute yield estimation (Khaki and Wang 2019).

2.5. Simulation

Neural network predictions of future yield delta distributions feed into Monte Carlo simulations (Metropolis 1987) which estimate probabilities and severity of losses at the risk unit scale. Figure 1 depicts this operation executed for each of the 17 years within the 2030 and 2050 CHC-CMIP6 series (Williams et al. 2024).
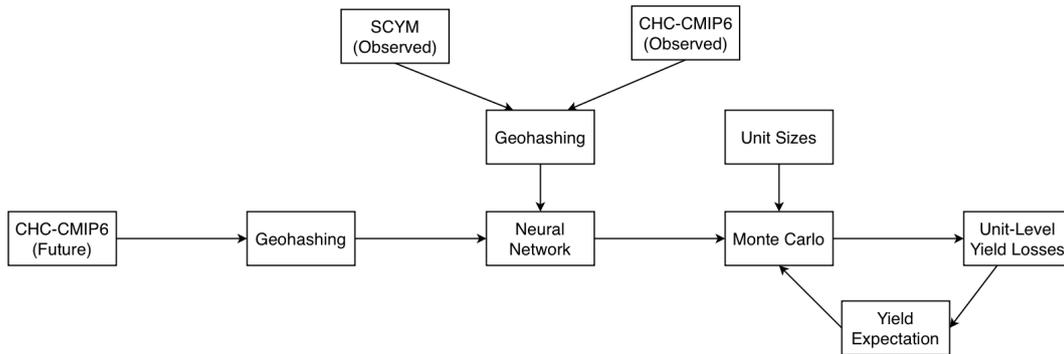

Figure 1: Model pipeline overview diagram. Code released as open source.

Altogether, these trials simulate risk unit yield deltas to consider a distribution of future outcomes and systems-wide institution-relevant metrics like claims rate (p_l). Note that CHC-CMIP6 provides multiple years per series but not specific future years like 2035. Therefore, we report results as distributions of outcomes per the 2030 and 2050 series.

Trials

Each Monte Carlo trial involves multiple sampling operations. First, we sample climate variables and model error residuals to propagate uncertainty (Yanai et al. 2010). Next, we draw yield multiple times to approximate the size of a risk unit with its portfolio effects. Note that the size but not the specific location of insured units is publicly disclosed. Therefore, our Monte Carlo draws the geographic size of each insured unit randomly from historic data (RMA 2024). In determining number of samples to take for statistical tests, we conservatively assume one kilometer resolution when sampling given SCYM use of Daymet (Thornton et al. 2014).

Statistical tests

Comparing predicted annual yield delta distributions from SSP245 to the no further warming "counterfactual" scenario per neighborhood, we determine significance ($p < 0.05/n$) by Mann Whitney U (Mann and Whitney 1947), a test appropriate for the heterogenous variance observed within our dataset (McDonald 2014). We also apply Bonferroni-correction (Bonferroni 1935) to control family-wise error in a large number of tests (one per neighborhood per year). All that said, our interactive tools allow for exploration of other test configurations.

3. Results

Our simulations anticipate loss probabilities (p_l) to double at mid-century relative to the no further warming counterfactual while the size of those losses (s) also increase. For insurers, this may translate to both more numerous and larger indemnity claims.

3.1. Aggregation outcomes

Historic data spanning 1999 to 2016 include a median of 83k field-level SCYM yield estimations per neighborhood represented within annual neighborhood-level yield distributions. While yield is often not normally distributed, 97% of neighborhoood yield *delta* distributions exhibit approximate normality per Kim (2013), representing almost all maize acreage. Therefore, we report outputs assuming normally distributed yield deltas. However, our appendix provides further statistics and beta distribution results.

3.2. Neural network outcomes

With bias towards performance in mean prediction, we select six hidden layers using 0.05 dropout and 0.05 L2 from our sweep with all data attributes included. Within the validation set, this leading feed forward neural network outperforms the leading Gaussian Process and LSTM result as shown in Table 3.

Table 3: Mean absolute validation set error in yield delta percentage points ($|(y_{actual} - y_{expected})/y_{expected} - y_{\Delta\%-Predicted}|$) for top models per model type.

Model	MAE for Mean Prediction	MAE for Std Prediction
Gaussian Process	14.4%	4.2%
Feed Forward	9.4%	3.2%
LSTM	18.0%	5.4%

Table 4 details that, when predicting neighborhood hidden test set yield delta distributions ($y_{\Delta\%}$), this selected feed forward model sees 6.2% MAE when predicting mean and 2.0% for standard deviation after retraining with train and validation together.

Table 4: Results of model training and selection.

Set	MAE for Mean Prediction	MAE for Std Prediction
Train	6.1%	2.0%
Validation	9.4%	3.2%
Test with retrain	6.2%	2.0%

In addition to the MAEs reported in Table 5 for varied test sets, we also find a 0.74 test set correlation coefficient. We observe that Khaki and Wang (2019) recently report a similar 0.73 in a separate but related absolute yield prediction task when using different but analogous data[3]. However, while Khaki and Wang (2019) focus on site-specific

---

[3]Khaki and Wang (2019) predict absolute yield. This study predicts percent risk unit yield change.

prediction in a specific future year so further improve upon their 0.73 by incorporating genotypical and soil data not available at our broad geographic scale, we contribute the mechanics required for insurance simulation such as predicting variance, incorporating error measures, and simulating risk units to calculate institutional-level outcomes.

Table 5: Results of tests after model selection across varied test set assignments.

Task	Test Mean Pred MAE	Test Std Pred MAE	% Units in Test Set
Random	5.0%	1.6%	15.4%
Temporal	8.3%	2.1%	17.0%
Spatial	4.7%	1.7%	24.8%
Climatic	5.2%	1.9%	5.2%

3.3. Simulation outcomes

Figure 2: Monte Carlo simulation results comparing SSP245 versus no warming counterfactual for (A) loss probability, (B) loss severity, and (C) change in average yields.

After retraining the sweep-selected configuration on all available data, Monte Carlo simulates risk units from which we derive overall system metrics like claims rate. To capture insurance mechanics, these trials track changes to average yields over time at the neighborhood and approximated risk unit level. Additionally, we also sample test set model residuals to account for error. Despite the conservative nature of the Bonferroni correction (McDonald 2014), 95.3% of maize acreage in SSP245 falls within a neighborhood with significant changes to claim probability ($p < 0.05/n$) at some point during the 2050 series simulations. From an insurance perspective, average covered loss increases over time while the claims rate elevates from 13% to 22% in the SSP245 2030 series before reaching 29% in the 2050 timeframe relative to the no further warming counterfactual. Simulation re-executions confirm result stability (see appendix).

4. Discussion

We observe policy-relevant dynamics when simulating insurance under climate change.

4.1. Geographic bias

Neighborhoods with significant results ($p < 0.05/n$) may be more common in some areas as shown in Figure 3. This spatial pattern may partially reflect that a number of neighborhoods have less land dedicated to maize, so simulations have smaller sample sizes and fail to reach significance. However, this may also mirror geographical bias in altered growing conditions. Reflecting empirical studies that document the negative impacts of heat stress and water deficits on maize yields (Sinsawat et al. 2004; Marouf et al. 2013), we note that spatial distribution of anticipated combined warmer and drier conditions partially mirror areas of lower yield predictions, possibly highlighting analogous stresses to 2012 and its historically poor maize production (Westcott and Jewison 2013).
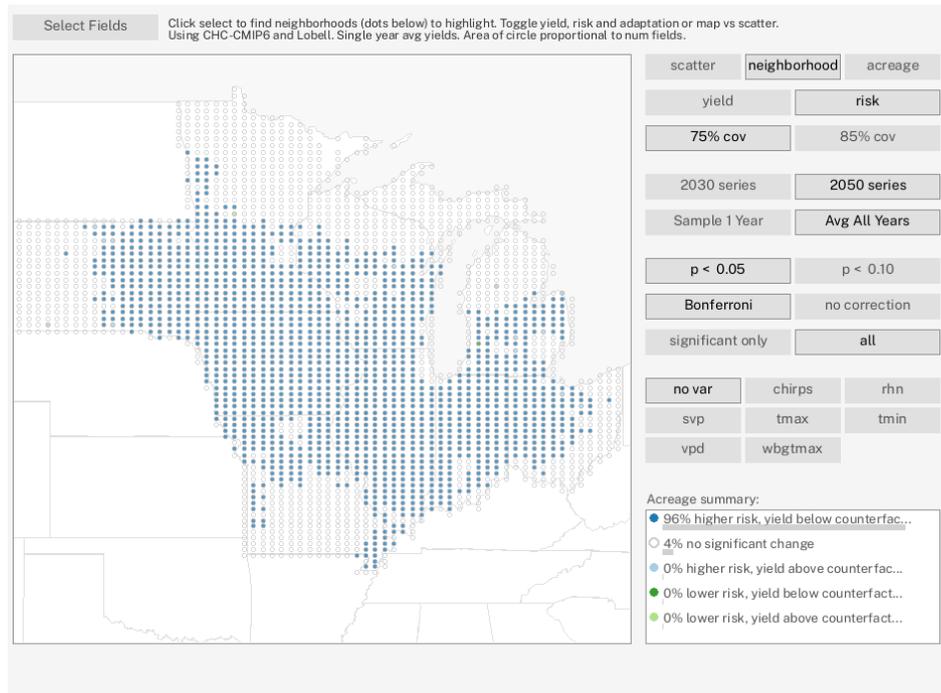

Figure 3: Interactive visualization showing 2050 series risk. Gray circle outlines show sparser neighborhoods with the 4% of maize growing acreage not reporting significant changes. Filled in dark blue circles show neighborhoods containing the 96% of acreage where loss risk increased ($p < 0.05/n$). Possible bias in Iowa, Illinois, and Indiana.

4.2. Yield expectations

Figure 4 reveals possible challenges with using simple averages in crop insurance products. While current instruments use $y_{expected}$ to capture changes to risk, our simulations anticipate higher yield volatility to skew yield delta distributions such that risk units see higher claims rates despite changing $y_{expected}$ values. Indeed, as described in our
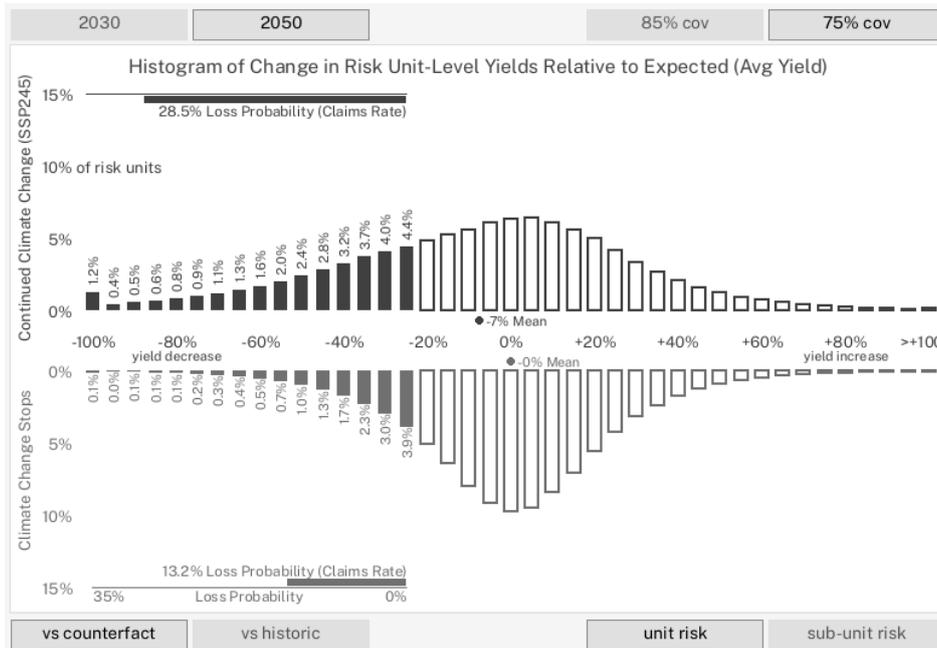
Figure 4: Interactive tool screenshot showing 2050 yield deltas distribution with climate change on the top and without further climate change (counterfactual) on bottom.

appendix, 12.7% of neighborhoods and 9.8% of counties under SSP245 in the 2050 series report both increased claims rates and increased average yields. In other words, yield volatility could elevate loss probability without necessarily decreasing overall mean yields enough to reduce claims rates through $y_{expected}$.

Impacts to insurers

As $y_{expected}$ (FCIC 2020) may fail to offer sensitivity to these predicted changes, insurers may see higher risk. These possible "invisible" elevations in unit-level loss probability may highlight a need for instrument adaptation. For example, FCIP formulations may consider possibly including historic yield variability in establishing production histories and $y_{expected}$.

Impacts to growers

Some risk mitigating practices such as regenerative agriculture trade output for stability (Lobell et al. 2024), guarding against increased loss probability (Renwick et al. 2021) at the cost of slightly reduced average yield (Deines et al. 2023). Therefore, though insurance effects on grower behavior remain under investigation (Connor et al. 2022; Wang et al. 2021; Chemeris et al. 2022), our results may indicate how average-based expectations could possibly disincentivize growers from climate change preparation.

4.3. Comparison to recent actual claims rates

We generally predict a 13% claims probability in 2030 and 2050 "counterfactual" simulations which anticipate yields absent further climate change (future conditions similar to recent past). For comparison, insurance data in the years for which SCYM and

historic CHC-CMIP6 data are available (RMA 2024) report an annual claims rate around 14% (median). Despite similarity between our predictions and comparable recent actuals, a number of factors which are difficult to model would likely lead us to underestimate claims probability. First, field-level yield data and actual geographically specific risk unit structure are not currently public. While we sample units randomly based on expected size, growers likely optimize their unit structure when purchasing policies. Second, we do not have geographically specific unit locations for modeling trend adjustment and yield exclusion options[4]. Finally, while these limitations likely overall lead to a suppression of loss rates relative to actuals, policy changes over time could cause further fluctuations alongside growing condition variability. For example, 2014 saw a number of statutory changes to yield exclusions (ERS 2024). Altogether, the future may see substantial annual variation similar to the recent past even as our results still capture overall long term trends.

4.4. Future data

We acknowledge limitations arising from the currently available public datasets. First, though our interactive tools consider different spatial aggregations such as five character geohashes, future work may consider modeling with actual field-level yield data and the actual risk unit structure if later made public. Additionally, we focus on systematic changes in growing conditions impacting claims rates across a broad geographic scale so exclude highly localized effects like certain inclement weather which may require more granular climate predictions. This may be relevant to programs with smaller geographic portfolios. Next, our model shows signs that it is data constrained and additional years of training data may improve performance. Our pipeline should and can be re-run as future versions of CHC-CMIP6 and SCYM or similar are released. Furthermore, we recognize that the CHC-CMIP6 2030 and 2050 series make predictions for general timeframes and not individual specific years which may be valuable for future research. Finally, we acknowledge that SCYM and CHC-CMIP6 include limited uncertainty data.

4.5. Other programs

Outside of Yield Protection, future study could extend to the highly related Revenue Protection form of insurance. Indeed, the yield stresses that we describe in this model may also impact this other plan. On that note, we include historic yield as inputs into our neural network, allowing those data to "embed" adaptability measures (Hsiang et al. 2017) such as grower practices where, for example, some practices may reduce loss events or variability (Renwick et al. 2021). That said, we highlight that later studies looking at revenue may require additional economic information to serve a similar role.

4.6. Visualizations and software

In order to explore these simulations, we offer interactive open source web-based visualizations built alongside our experiments. These both aid us in constructing our own

---

[4]Under certain conditions, trend adjustment increases $y_{expected}$ above historic average (Plastina and Edwards 2014) to anticipate expected yield improvements while exclusions remove poor years from $y_{expected}$ (Schnitkey et al. 2015).

conclusions and allow readers to consider possibilities and analysis beyond our own narrative. Publicly available at `https://ag-adaptation-study.pub`, this software includes the ability to explore alternative statistical treatments and regressor configurations as well as generate additional geographic visualizations. Finally, we also offer an open source data science pipeline to build these models and run simulations.

5. Conclusion

We present Monte Carlo simulations on top of a neural network-based regressor for prediction of institution-relevant crop yield changes. We specifically simulate climate-driven system-wide impacts to maize growing conditions at a policy-relevant scale of granularity. Our results anticipate maize Yield Protection claim rates to double at mid-century for the U.S. Federal Crop Insurance Program (Multi-Peril Crop Insurance) within the U.S. Corn Belt relative to a no further warming counterfactual.

In addition to publishing our inputs and raw model outputs under a creative commons license, we explore the specific shape of these results from the perspective of insurance structures. First, we describe a possible agriculturally-relevant geographic bias in climate impacts. Second, we also highlight potential mathematical properties of interest including a predicted increase in volatility without fully offsetting average-based yield expectation measures ($y_{expected}$). These particular kinds of changes may pose specific threats to the current structure of existing insurance instruments.

Altogether, this study considers how this machine learning and interactive data science approach may understand existing food system policy structures in the context of climate projections. Towards that end, we release our software under permissive open source licenses and make interactive tools available publicly at `https://ag-adaptation-study.pub` to further interrogate these results. These visualizations also allow readers to explore alternatives to key analysis parameters. Aided by our open source pipeline, this work may inform agriculture policy response to continued climate change.

Acknowledgments

Data availability statement: Zenodo archives our open source code, data, and resources (Pottinger et al. 2024a,b). Public hosted software is available at `https://ag-adaptation-study.pub` and can investigate further metrics and alternative models.

References

Alwosheel, A., van Cranenburgh, S., and Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling*, 28:167–182, ISSN: 1755-5345, DOI: https://doi.org/10.1016/j.jocm.2018.07.002.

Ates, A. (2023). Feed grains sector at a glance. https://www.ers.usda.gov/topics/crops/corn-and-other-feed-grains/feed-grains-sector-at-a-glance/.

Bonferroni, C. (1935). Il calcolo delle assicurazioni su gruppi di teste. *Studi in Onore del Professore Salvatore Ortu Carboni*, https://api.semanticscholar.org/CorpusID:89994272.

Brazie, A. (2024). Designing the core gameplay loop: A beginner's guide. https://gamedesignskills.com/game-design/core-loops-in-gameplay/.

Brewer, C., Harrower, M., Sheesley, B., Woodruff, A., and Heyman, D. (2013). ColorBrewer 2.0. https://colorbrewer2.org.

Bridgwater, A. (2015). What are micro apps and why do they matter for mobile? https://www.forbes.com/sites/adrianbridgwater/2015/07/16/what-are-micro-apps-and-why-do-they-matter-for-mobile/.

Brown, M. (2024). The 100 games that taught me game design. https://www.youtube.com/watch?v=gWNXGfXOrro.

Brownlee, J. (2020a). Deep learning models for multi-output regression. https://machinelearningmastery.com/deep-learning-models-for-multi-output-regression/.

Brownlee, J. (2020b). What is the difference between test and validation datasets? https://machinelearningmastery.com/difference-test-validation-datasets.

Chemeris, A., Liu, Y., and Ker, A. P. (2022). Insurance subsidies, climate change, and innovation: Implications for crop yield resiliency. *Food Policy*, 108:102232, ISSN: 03069192, DOI: 10.1016/j.foodpol.2022.102232.

Connor, L., Rejesus, R. M., and Yasar, M. (2022). Crop insurance participation and cover crop use: Evidence from Indiana county-level data. *Applied Economic Perspectives and Policy*, 44(4):2181–2208, ISSN: 2040-5790, 2040-5804, DOI: 10.1002/aepp.13206.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, 3(1):52–58, ISSN: 1758-678X, 1758-6798, DOI: 10.1038/nclimate1633.

Deines, J. M., Guan, K., Lopez, B., Zhou, Q., White, C. S., Wang, S., and Lobell, D. B. (2023). Recent cover crop adoption is associated with small maize and soybean yield losses in the United States. *Global Change Biology*, 29(3):794–807, ISSN: 1354-1013, 1365-2486, DOI: 10.1111/gcb.16489.

Deines, J. M., Patel, R., Liang, S.-Z., Dado, W., and Lobell, D. B. (2021). A million kernels of truth: Insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. *Remote Sensing of Environment*, 253:112174, ISSN: 00344257, DOI: 10.1016/j.rse.2020.112174.

DellaFave, R. (2014). Designing RPG mini-games (and getting them right). https://www.gamedeveloper.com/design/designing-rpg-mini-games-and-getting-them-right-.

Diffenbaugh, N. S., Davenport, F. V., and Burke, M. (2021). Historical warming has increased U.S. crop insurance losses. *Environmental Research Letters*, 16(8):084025, ISSN: 1748-9326, DOI: 10.1088/1748-9326/ac1223.

Economic Research Service (2024). 2014 farm bill. https://www.ers.usda.gov/topics/farm-bill/2014-farm-bill/.

Federal Crop Insurance Corporation (2020). Common crop insurance policy 21.1-BR. https://www.rma.usda.gov/policy-procedure/general-policies/basic-provisions-21-1-br.

Federal Crop Insurance Corporation (2023). *Crop Insurance Handbook: 2024 and Succeeding Crop Years*. https://www.rma.usda.gov/sites/default/files/handbooks/2024-18010-1-Crop-Insurance-Handbook.pdf.

Green, T. R., Kipka, H., David, O., and McMaster, G. S. (2018). Where is the USA Corn Belt, and how is it changing? *Science of the Total Environment*, 618:1613–1618, ISSN: 0048-9697, DOI: 10.1016/j.scitotenv.2017.09.325.

GSA (2024). Public sans. https://public-sans.digital.gov/.

Hanrahan, R. (2024). Crop insurance costs projected to jump 29%. https://farmpolicynews.illinois.edu/2024/02/crop-insurance-costs-projected-to-jump-29/.

Haugen, B. (2020). Geohash size variation by latitude. https://bhaugen.com/blog/geohash-sizes/.

Hausfather, Z. and Peters, G. P. (2020). Emissions — the 'business as usual' story is misleading. *Nature*, 577(7792):618–620, ISSN: 0028-0836, 1476-4687, DOI: 10.1038/d41586-020-00177-3.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780, ISSN: 0899-7667, DOI: 10.1162/neco.1997.9.8.1735.

Hsiang, S., Kopp, R., Jina, A., Rising, J., Delgado, M., Mohan, S., Rasmussen, D. J., Muir-Wood, R., Wilson, P., Oppenheimer, M., Larsen, K., and Houser, T. (2017). Estimating economic damage from climate change in the United States. *Science*, 356(6345):1362–1369, ISSN: 0036-8075, 1095-9203, DOI: 10.1126/science.aal4369.

Jagermeyr, J., Muller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J. A., Fuchs, K., Guarin, J. R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A. K., Kelly, D., Khabarov, N., Lange, S., Lin, T.-S., Liu, W., Mialyk, O., Minoli, S., Moyer, E. J., Okada, M., Phillips, M., Porter, C., Rabin, S. S., Scheer, C., Schneider, J. M., Schyns, J. F., Skalsky, R., Smerald, A., Stella, T., Stephens, H., Webber, H., Zabel, F., and Rosenzweig, C. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, 2(11):873–885, ISSN: 2662-1355, DOI: 10.1038/s43016-021-00400-y.

JM8 (2024). The secret to making any game satisfying | design delve. https://www.youtube.com/watch?v=ORsb7g_ioLs.

Joseph, R. (2018). Grid search for model tuning. https://medium.com/data-science/grid-search-for-model-tuning-3319b259367e.

Keras Contributors (2024). **Keras**. https://keras.io.

Khaki, S. and Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in Plant Science*, 10:621, DOI: 10.3389/fpls.2019.00621.

Kim, H.-Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, 38(1):52, ISSN: 2234-7658, 2234-7666, DOI: 10.5395/rde.2013.38.1.52.

Kim, Y.-S., Kim, M. K., Fu, N., Liu, J., Wang, J., and Srebric, J. (2024). Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustainable Cities and Society*, page 105570, ISSN: 22106707, DOI: 10.1016/j.scs.2024.105570.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. DOI: 10.48550/ARXIV.1412.6980.

Leng, G. and Hall, J. W. (2020). Predicting spatial and temporal variability in crop yields: An inter-comparison of machine learning, regression and process-based models. *Environmental Research Letters*, 15(4):044027, ISSN: 1748-9326, DOI: 10.1088/1748-9326/ab7b24.

Lewis, C. (1982). Using the "thinking-aloud" method in cognitive interface design. Research Report RC9265, IBM TJ Watson Research Center (1982).

Li, K., Pan, J., Xiong, W., Xie, W., and Ali, T. (2022). The impact of 1.5 °C and 2.0 °C global warming on global maize production and trade. *Scientific Reports*, 12(1):17268, ISSN: 2045-2322, DOI: 10.1038/s41598-022-22228-7.

Lobell, D., di Tommaso, S., Zhou, Q., Ma, Y., Specht, J., and Guan, K. (2024). The mixed effects of recent cover crop adoption on U.S. cropland productivity. *Research Square*, DOI: 10.21203/rs.3.rs-5146628/v1.

Lobell, D. B. and Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150(11):1443–1452, ISSN: 0168-1923, DOI: 10.1016/j.agrformet.2010.07.008.

Lobell, D. B., Deines, J. M., and Tommaso, S. D. (2020). Changes in the drought sensitivity of US maize yields. *Nature Food*, 1(11):729–735, ISSN: 2662-1355, DOI: 10.1038/s43016-020-00165-w.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., and Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, 164:324–333, ISSN: 00344257, DOI: 10.1016/j.rse.2015.04.021.

Loshchilov, I. and Hutter, F. (2017). Decoupled weight decay regularization. In *International Conference on Learning Representations*. https://api.semanticscholar.org/CorpusID:53592270.

Luigi Contributors, S. (2024). **Luigi**.

Ma, Y., Liang, S.-Z., Myers, D. B., Swatantran, A., and Lobell, D. B. (2024). Subfield-level crop yield mapping without ground truth data: A scale transfer framework. *Remote Sensing of Environment*, 315:114427, ISSN: 0034-4257, DOI: 10.1016/j.rse.2024.114427.

Maas, A. L., Hannun, A. Y., and Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28. JMLR. https://api.semanticscholar.org/CorpusID:16489696.

Mahul, O. and Stutley, C. (2010). Government support to agricultural insurance: Challenges and options for developing countries. *The World Bank Open Knowledge Repository*, https://openknowledge.worldbank.org/entities/publication/8a605230-3df5-5a4e-8996-5a6de07747e1.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1):50–60, ISSN: 0003-4851, DOI: 10.1214/aoms/1177730491.

Manski, S., Socolar, Y., Goldstein, B., Pizzo, G., Ahmed, Z., Connor, L., Cross, H., Fettes, K., McLauchlan, A., Pham, L., Viens, F., and Bowles, T. M. (2024). Diversified crop rotations mitigate agricultural losses from dry weather. *agriRxiv*, page 20240168962, ISSN: 2791-1969, DOI: 10.31220/agriRxiv.2024.00244.

Marouf, K., Naghavi, M., Pour-Aboughadareh, A., and Naseri rad, H. (2013). Effects of drought stress on yield and yield components in maize cultivars (Zea mays L.). *International Journal of Agronomy and Plant Production*, 4:809–812. https://api.semanticscholar.org/CorpusID:130960873.

McDonald, J. (2014). *Handbook of Biological Statistics*. Sparky House Publishing, 3rd edition, http://www.biostathandbook.com/#print.

Metropolis, N. (1987). The beginning of the Monte Carlo method. https://sgp.fas.org/othergov/doe/lanl/pubs/00326866.pdf.

Mwiti, D. (2023). Random forest regression: When does it fail and why? https://neptune.ai/blog/random-forest-regression-when-does-it-fail-and-why.

Nelson, C. H. (1990). The influence of distributional assumptions on the calculation of crop insurance premia. *North Central Journal of Agricultural Economics*, 12(1):71–78, ISSN: 01919016, http://www.jstor.org/stable/1349359.

Niemeyer, G. (2008). geohash.org is public! https://web.archive.org/web/20080305102941/http://blog.labix.org/2008/02/26/geohashorg-is-public/.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). **Scikit-learn**: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830. https://dl.acm.org/doi/10.5555/1953048.2078195.

Plastina, A. and Edwards, W. (2014). Trend-adjusted actual production history (APH). https://www.extension.iastate.edu/agdm/crops/html/a1-56.html.

Pottinger, A., Connor, L., Guzder-Williams, B., Weltman-Fahs, M., and Bowles, T. (2024a). Data files for climate-based maize loss rate simulations. DOI: 10.5281/ZENODO.13356980.

Pottinger, A., Connor, L., Guzder-Williams, B., Weltman-Fahs, M., and Bowles, T. (2024b). Data pipeline and tool source code for climate-based maize loss rate simulations. DOI: 10.5281/zenodo.13356710.

Pottinger, A. S. (2024). **Sketchingpy**. https://sketchingpy.org.

Pottinger, A. S., Biyani, N., Geyer, R., McCauley, D. J., de Bruyn, M., Morse, M. R., Nathan, N., Koy, K., and Martinez, C. (2023). Combining game design and data visualization to inform plastics policy: Fostering collaboration between science, decision-makers, and artificial intelligence. DOI: 10.48550/ARXIV.2312.11359.

Pottinger, A. S. and Zarpellon, G. (2023). Pyafscgap.org: Open source multi-modal Python-based tools for NOAA AFSC RACE GAP. *Journal of Open Source Software*, 8(86):5593, ISSN: 2475-9066, DOI: 10.21105/joss.05593.

Python Software Foundation (2024). Python language reference. https://www.python.org/.

Ray, D. K., Gerber, J. S., MacDonald, G. K., and West, P. C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications*, 6(1):5989, ISSN: 2041-1723, DOI: 10.1038/ncomms6989.

Renwick, L. L. R., Deen, W., Silva, L., Gilbert, M. E., Maxwell, T., Bowles, T. M., and Gaudin, A. C. M. (2021). Long-term crop rotation diversification enhances maize drought resistance through soil organic matter. *Environmental Research Letters*, 16(8):084067, ISSN: 1748-9326, DOI: 10.1088/1748-9326/ac1468.

Rezaei, E. E., Webber, H., Asseng, S., Boote, K., Durand, J. L., Ewert, F., Martre, P., and MacCarthy, D. S. (2023). Climate change impacts on crop yields. *Nature Reviews Earth & Environment*, 4(12):831–846, ISSN: 2662-138X, DOI: 10.1038/s43017-023-00491-0.

Risk Management Agency (2008). Crop insurance options for vegetable growers. `https://catalog.hathitrust.org/Record/007425920`.

Risk Management Agency (2022). *Climate Adaptation Plan.* `https://www.usda.gov/sites/default/files/documents/3_FPAC_RMA_ClimateAdaptationPlan_2022.pdf`.

Risk Management Agency (2024). State/county/crop summary of business. `https://old.rma.usda.gov/en/Information-Tools/Summary-of-Business/State-County-Crop-Summary-of-Business`.

Sajid, O., Ortiz-Bobea, A., Ifft, J., and Gauthier, V. (2023). Extreme heat and Kansas farm income. `https://farmdocdaily.illinois.edu/2023/07/extreme-heat-and-kansas-farm-income.html`.

Schechinger, A. (2023). Crop insurance costs soar over time, reaching a record high in 2022. `https://www.ewg.org/research/crop-insurance-costs-soar-over-time-reaching-record-high-2022`.

Schnitkey, G., Sherrick, B., and Coppess, J. (2015). Yield exclusion: Description and guidance. `https://origin.farmdocdaily.illinois.edu/2015/01/yield-exclusion-description-and-guidance.html`.

SciPy (2024). `scipy.stats.beta` — **scipy** v1.41.1 manual. `https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.beta.html`.

Shneiderman, B. and Plaisant, C. (2006). Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, pages 1–7. ACM, DOI: `10.1145/1168149.1168158`.

Sinsawat, V., Leipner, J., Stamp, P., and Fracheboud, Y. (2004). Effect of heat stress on the photosynthetic apparatus in maize (Zea mays L.) grown at control or high temperature. *Environmental and Experimental Botany*, 52(2):123–129, ISSN: 00988472, DOI: `10.1016/j.envexpbot.2004.01.010`.

Thornton, P., Thornton, M., Mayer, B., Wilhelmi, N., Wei, Y., Devarakonda, R., and Cook, R. (2014). Daymet: Daily surface weather data on a 1-km grid for North America, version 2. DOI: `10.3334/ORNLDAAC/1219`.

Tsiboe, F. and Turner, D. (2023). Crop insurance at a glance. `https://www.ers.usda.gov/topics/farm-practices-management/risk-management/crop-insurance-at-a-glance/`.

Unwin, A. (2020). Why is data visualization important? What is important in data visualization? *Harvard Data Science Review*, DOI: `10.1162/99608f92.8ae4d525`.

van Klompenburg, T., Kassahun, A., and Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177:105709, ISSN: 0168-1699, DOI: `10.1016/j.compag.2020.105709`.

Victor, B. (2011). Explorable explanations. http://worrydream.com/Explorable Explanations/.

Wang, R., Rejesus, R. M., and Aglasan, S. (2021). Warming temperatures, yield risk and crop insurance participation. *European Review of Agricultural Economics*, 48(5):1109–1131, ISSN: 0165-1587, 1464-3618, DOI: 10.1093/erae/jbab034.

Westcott, P. C. and Jewison, M. (2013). Weather effects on expected corn and soybean yield. https://www.ers.usda.gov/publications/pub-details?pubid=36652.

Williams, E., Funk, C., Peterson, P., and Tuholske, C. (2024). High resolution climate change observations and projections for the evaluation of heat-related extremes. *Scientific Data*, 11(1):261, ISSN: 2052-4463, DOI: 10.1038/s41597-024-03074-w.

Yanai, R. D., Battles, J. J., Richardson, A. D., Blodgett, C. A., Wood, D. M., and Rastetter, E. B. (2010). Estimating uncertainty in ecosystem budget calculations. *Ecosystems*, 13(2):239–248, ISSN: 1432-9840, 1435-0629, DOI: 10.1007/s10021-010-9315-8.

A. Simulation Results

We first provide further detailed simulation results in Table 6.

Table 6: Details of Monte Carlo simulation results. Counterfactual is a future without continued warming. "2010 series" label used for consistency with 2030 and 2050 from CHC-CMIP6 though that "2010" language does not explicitly appear in their data model. Further results are available in Zenodo (Pottinger et al. 2024a).

Scenario	Series	Unit mean yield change	Unit loss robability	Avg covered loss severity
Historic	2010	18.6%	7.3%	13.8%
Counterfactual	2030	0.0%	13.3%	14.7%
SSP245	2030	-4.5%	22.3%	17.5%
Counterfactual	2050	-0.0%	13.2%	14.5%
SSP245	2050	-7.4%	28.5%	18.9%
		$y_{\Delta\mu}$	$p_{l-\mu}$	s_μ

When re-executing simulations 100 times to understand variability for system-wide metrics in Table 6, the range of all standard deviations of each metric's distribution is under 0.1% and the range under 1%, likely reflecting the high degree of aggregation in system-wide metrics. However, lacking confidence measures from SCYM and CHC-CMIP6, this post-hoc experiment cannot account for input data uncertainty. Finally, within these results, Table 7 shows that simulations report 13% of neighborhoods and 10% of counties seeing both increased average yields and increased claims rates together[5], likely reflecting increased year to year volatility.

Table 7: Frequency with which average yield and probability of claim both increase. Counterfactual assumes no further warming.

Series	Condition	Neighborhoods	Counties
2030	Counterfactual	3.6%	2.0%
2050	Counterfactual	3.7%	1.9%
2030	SSP245	1.5%	1.5%
2050	SSP245	12.7%	9.8%

For context, without yield exclusion, a year with claims for a risk unit would generally decrease the subsequent $y_{expected}$ for that risk unit. Therefore, one may expect generally few neighborhoods and counties to see both increased average yields and increased probability of claims when both are calculated over a multi-year period. However, the skew for the *multi-year distributions* of yield deltas (as opposed to any single set of

---

[5]Calculated across the entire SSP245 2050 series. We use geohash center to determine county (FCC 2024). To avoid noise, we consider increases in average yield and increases in claims rates of less than 2% as essentially unchanged for this specific post-hoc experiment. However, the gap persists between 2050 SSP245 and 2050 counterfactual frequencies even if this 2% noise filter is removed.

annual yield deltas) grows over SSP245 as reflected visually in our interactive tools: 2030 looks more like a normal distribution than 2050.

B. Methods and Input Data

This appendix further details specific implementation of the methods and input data.

B.1. Insured risk unit data

Figure 5 visualizes anonymized information about risk structure (RMA 2024) where the USDA only indicates the county (not precise location) for each unit.
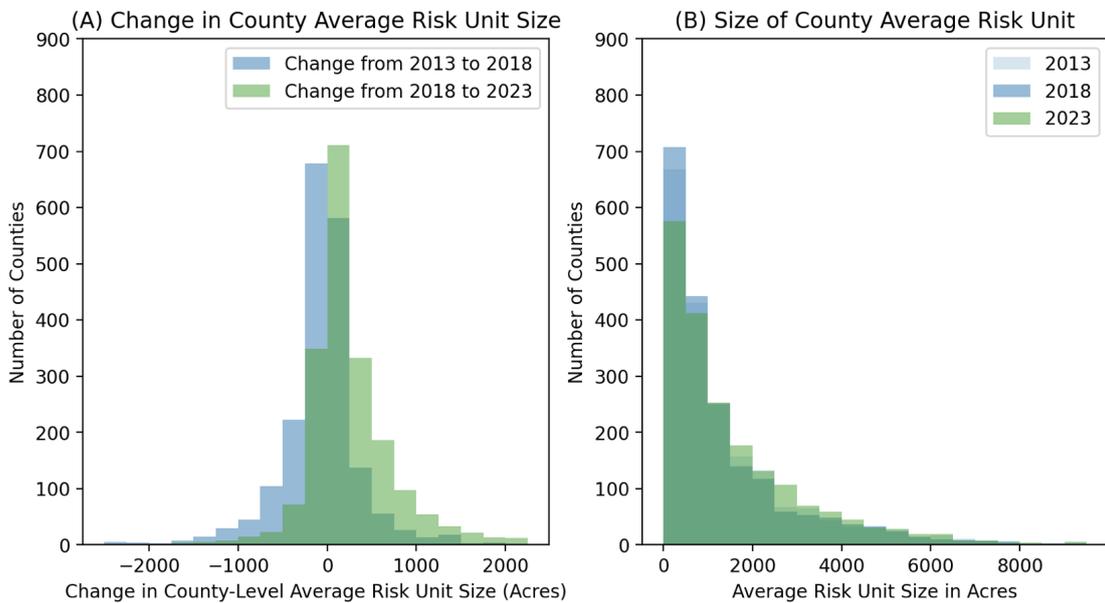

Figure 5: Examination of risk unit size in years 2013, 2018, and 2023. First, this figure shows how risk unit size changed between each year examined (A) to highlight that the structures do evolve substantially between years. However, these results also indicate that the overall distribution of risk unit sizes is relatively stable (B) when considered system-wide. Some extreme outliers not shown to preseve detail.

Year to year instability at the county level in unit size may reflect growers reconfiguring their risk structure to optimize rates as yield profiles change over time. All this in mind, sampling the risk unit size at the county level likely represents over-confidence (overfitting) to previous configurations. Instead, we observe that the system-wide risk unit size distribution remains relatively stable. This may suggest that, even as more local changes to risk unit structure may be more substantial between years, overall expectations for the size of risk units are less fluid. Therefore, we use that larger system-wide distribution to sample risk unit sizes within our Monte Carlo simulation instead of the county-level distributions. This also has the effect of propogating risk unit size uncertainty into results through the mechanics of Monte Carlo.

B.2. Yield distributions

To avoid a large increase in data requirements to cope with increased dimensionality (Alwosheel et al. 2018) and to perform fair comparisons to Gaussian Process trials, our neural network requires a distributional shape assumption to maintain a smaller output vector size. We decide the shape to predict based on observed skew and kurtosis of yield deltas. To that end, our open source pipeline can be run with beta or normal distribution assumptions. The former has precedent in the literature (Nelson 1990) but 97% of neighborhoods and maize growing acreage are approximately normal per Kim (2013). Table 8 shows that using beta distributions in our neural networks results in similar median absolute errors but elevated mean absolute errors. A minority population of neighborhoods causes this swing where small changes in beta distribution parameters can infrequently cause large error. Guided by stronger performance and the frequency of normality, we assume normal yield deltas in our main text. For memory efficiency, we sample 1,000 yield values per neighborhood per year from which risk units are further sampled. This behavior can be disabled if an a priori distribution shape assumption is made.

Table 8: Test set performance after retraining for predicting distribution location (mean or center) for both a normal distribution and beta distribution assumption.

Shape	Mean Absolute Error	Median Absolute Error
Normal	6.2%	5.9%
Beta	16.9%	7.1%

B.3. Neural network configuration

We next offer additional information about the specific neural network configuration chosen. Table 9 provides mean absolute error for the selected model from the sweep. A drop in error observed from validation to test with retrain[6] performance may be explained by the increased training set size. This may indicate that the model is data constrained by the number of years available for training. Our open source data pipeline can rerun analysis as input datasets are updated in the future.

Table 9: Residuals for the main training task with and without retraining.

Set	MAE for Mean Prediction	MAE for Std Prediction
Train	6.1%	2.0%
Validation	9.4%	3.2%
Test with retrain	6.2%	2.0%
Test without retrain	11.1%	2.4%

---

[6]Test with retrain specifically refers to retraining a model from scratch using the model configuration selected from our hyper-parameter sweep. This training spans across both training and validation data together. In both the "with retrain" and "without retrain" cases, the test set remains fully hidden.

Test set residuals are sampled during Monte Carlo to propogate uncertainty. That said, a relatively small sub-population of large percentage changes may skew results, causing mean and median error to diverge as shown in post-hoc tasks in Table 10.

Table 10: Results of tests after model selection with MAE and MdAE.

Task	Mean Absolute Error		Median Absolute Error	
	Test Mean Pred	Test Std Pred	Test Mean Pred	Test Std Pred
Random	5.0%	1.6%	5.1%	1.7%
Temporal	8.3%	2.1%	7.2%	2.2%
Spatial	4.7%	1.7%	5.0%	1.7%
Climatic	5.2%	1.9%	5.2%	1.8%

Even so, the overall error remains acceptable. In general, increased model size shows diminishing returns (four vs five neural network layers changes mean prediction MAE by less than one point). Our final chosen model has the following layer sizes: 512 neurons, 256 neurons, 128 neurons, 64 neurons, 32 neurons, 8 neurons. Empirically leading to generally better performance, we allow the model to use the count of growing condition estimations. This may serve as a possible measure of uncertainty. We also allow inclusion of the year. However, as can be executed in our open source pipeline, we find that including absolute year generally increases overfitting. Therefore, we use a relative measure (years since the start of the series within the simulations). Our simulations run for 17 relative years for each series.

B.4. Historic yield averages

Our simulations anticipate $y_{expected}$ to change over time as documented in our Zenodo record (Pottinger et al. 2024a). We sample ten years of historic yields per neighborhood per year per trial and we offset the yield deltas produced by the neural network accordingly as the simulated timeseries progresses: predictions for 2030 claims rate samples the 2010 (historic) series and 2050 samples the 2030 series. To prevent discontinuity in the data due to unknown bias, the 2010 deltas are retroactively predicted. Model error residuals are always sampled. Note that many growers will engage in at least simple crop rotations (Manski et al. 2024) which may change the locations in which maize is grown. SCYM implicitly handles this complexity but the set of geohashes present in results may vary from one year to the next in part due to this behavior. All that said, historic locations of growth from the past are sampled in the future simulations. These mechanics could operate for LSTM and Gaussian Process as well.

B.5. Coverage levels

We observe that there may be geographic bias in coverage levels. This may include some areas with different policy availability, possibly including geographically-biased supplemental policy usage. This results both from grower and institutional behavior and may prove important in specific prediction of future claims. However, lacking public data on coverage levels chosen with geographic specificity, we respond to this limitation by allowing for investigation of different coverage levels within our interactive

Table 11: Overview of explorable explanations.

Simulator	Question	Loop	JG
Rates	What factors influence the price and subsidy of a policy?	Iteratively change variables to increase subsidy.	Improving on previous hypotheses.
Hyper-Parameter	How do hyper-parameters impact regressor performance?	Iteratively change neural network hyper-parameters to see influence on validation set performance.	Improving on previous hyper-parameter hypotheses.
Distributional	How do overall simulation results change under different simulation parameters?	Iterative manipulation of parameters (geohash size, event threshold, year) to change loss probability and severity.	Deviating from the study's main results.
Neighborhood	How do simulation results change across geography and climate conditions?	Inner loop changing simulation parameters to see changes in neighborhood outcomes. Outer loop of observing changes across different views.	Identifying neighborhood clusters of concern.
Claims	How do different regulatory choices influence grower behavior?	Iteratively change production history to see which years result in claims under different regulatory schemes.	Redefining policy to improve yield stability.

tool. Though we do not believe this to impact our predictions of general claims probability and severity changes, this aspect may impact research making specific annual predictions. Therefore, we encourage future work on further investigation of coverage level selection and its intersection with climate change.

C. Interactive Tools

Our interactive often tools operate as "explorable explanations" (Victor 2011). Listed in Table 11, we draw analogies to micro-apps (Bridgwater 2015) or mini-games (DellaFave 2014) in which the user encounters a series of small experiences that, each with distinct interaction and objectives, can only provide minimal instruction (Brown 2024). As these very brief visualization experiences cannot take advantage of design techniques like Hayashida-style tutorials (Pottinger et al. 2023), they rely on simple "loops" (Brazie 2024) for immediate "juxtaposition gratification" (JG) (JM8 2024), showing fast progression after minimal input. Following Unwin (2020), our custom tools first serve as internal exploratory graphics enabling the insights detailed in our results before acting as a medium for sharing our work. Figures and interactive tools use Color Brewer

(Brewer et al. 2013) and Public Sans (GSA 2024). Our pipeline is constructed using **Luigi** (**Luigi** Contributors 2024) in Python (**Python** Software Foundation 2024) with **Keras** (**Keras** Contributors 2024) while our tools use **Sketchingpy** (Pottinger 2024). For complete open source listing, see Pottinger et al. (2024b).

C.1. Internal use

First built during our own internal exploration of data, Table 12 outlines specific observations we attribute to our use of these tools.

Altogether, these tools serve to support our exploration of our modeling such as different loss thresholds for other insurance products, finding relationships of outcomes to different climate variables, understanding interaction with insurance mechanisms, answering geographically specific questions, and modification of machine learning parameters to understand performance.

Table 12: Observations we made from our own tools in the "exploratory" graphic context of Unwin (2020).

Simulator	Observation
Distributional	Dichotomy of changes to yield and changes to loss risk.
Claims	Issues of using average for FCIC (2020).
Neighborhood	Geographic bias of impact and model output relationships with broader climate factors.
Hyper-parameter	Model resilience to removing individual inputs.

C.2. Workshops

In addition to supporting our finding of our own conclusions, we release this software publicly at `https://ag-adaptation-study.pub`. For example, possible use of these tools may include workshop activity. To support use of these tools as supplement to this paper, we made the following changes[7]:

- We elect to alternate between presentation and interaction similar to Pottinger and Zarpellon (2023). However, we added the rates simulator to further improve presentation of the rate setting process due to the complexities of crop insurance, dynamics previously explained in static diagrams.

- Our single loop (Brazie 2024) designs may be better suited to the limited time-frame of a workshop. Therefore, we now let facilitators hold the longer two loop neighborhood simulator till the end by default.

---

[7]These were implemented in response to our work's participation in a "real-world" nine person workshop session encompassing scientists and engineers which was intended to improve these tools specifically through active co-exploration limited to these study results. We collect information about the tool only and not generalizable knowledge about users or these patterns, falling under "quality assurance" activity. IRB questionnaire on file. This was *not* a public workshop or a formalized academic conference presentation. We thank those who provided advice and who gave preference to be named in acknowledgements.

- The meta-parameter visualization specifically relies heavily on memory, recalling the association of graphical visuals to the configuration producing that graphic. Therefore, we now offer a "sweep" button for facilitators to show all results at once.

Later work may more broadly explore this design space through controlled experimentation (Lewis 1982) or diary studies (Shneiderman and Plaisant 2006).

D. Alternative Models

We contextualize our feed forward neural network by comparing to alternative machine learning approaches. However, as described further in this appendix, we still conclude that the presented feed forward neural network performs the best when tested in a temporally displaced hidden validation set.

D.1. Sweeps

Beyond the feed forward neural network, we consider two alternative types of models.

Gaussian process

First, using Pedregosa et al. (2011), we sweep Gaussian Process with varying kernels: Matern and RBF. For those sweep conditions in which Matern is used, we also try varying nu values: 1.0, 1.5, 2.0, 2.5. This sweep also tries optional addition of the White Kernel for Matern such that each Matern model is tried both with and without the White Kernel.

LSTM

Second, we also attempt LSTM with stacking. This uses the same L2 and dropout options as the feed forward network. However, this sweep varies numbers of stacked LSTM layers (8, 32, 128).

Memory constraints

Note that the LSTM sizes and Gaussian Process sampling were chosen as to fit within memory constraints of the distributed sweep workers (17.1 GB of RAM per model). These same constraints are imposed on the feed forward neural network.

D.2. Outcome

Sweeps chose Gaussian Process with Matern (nu = 0.25) and White Kernel while sweeps chose three layers LSTM with 0.1 L2 no dropout. However, all of the Gaussian Process and LSTM models underperformed relative to the feed forward neural network.

D.3. Discussion of alternative models

We note that it is possible that the LSTM underperformed the feed forward neural network due to increased data demands and the relatively small number of years for

which yield information is available in this problem[8]. However, we also observe significant year to year yield variability such that knowing the yield of a prior year may not necessarily provide additional advantage in predicting yield deltas within this specific task (with a limited year range) relative to knowing the number of years since the last yield was reported for an insured unit, a value already given to the model. We hope that highlighting this possibility may be useful to future work.

E. Alternative definitions

We provide equivalent mathematical definitions from the main text to support future work. First, covered loss is defined as actual yields dropping below coveragev level:

$$l = \max(c * y_{expected} - y_{actual}, 0). \tag{6}$$

This can be described as a percentage of that covered yield within some contexts where helpful:

$$l_\% = \max\left(\frac{y_{expected} - y_{actual}}{y_{expected}} - c, 0\right). \tag{7}$$

Furthermore, note that $y_{expected}$ is technically defined as the last ten years of yield for a crop. However, in practice, this may not be calendar years due to factors like crop rotations or due to farms with insufficient yield history:

$$y_{expected} = \frac{y_{historic}[-d:]}{d}, \tag{8}$$

$$y_{expected} = \frac{y_{historic}[-\min(10, |y_{historic}|):]}{\min(10, |y_{historic}|)}. \tag{9}$$

Next, the probability of experiencing a loss that may incur a Yield Protection claim (p_l) may be defined a few different ways depending on data available under different conditions:

$$p_l = P(l > 0) = P(c * y_{expected} - y_{actual} > 0); \tag{10}$$

$$p_l = P\left(\frac{y_{actual} - y_{expected}}{y_{expected}} < c - 1\right); \tag{11}$$

$$p_l = P(y_{\Delta\%} < c - 1). \tag{12}$$

Finally, the severity (s) of a loss may also take multiple forms:

$$s = \frac{l}{y_{expected}}; \tag{13}$$

$$s = \max\left(c - \frac{y_{actual}}{y_{expected}}, 0\right); \tag{14}$$

$$s = \max(-1 * y_{\Delta\%} - (1 - c), 0). \tag{15}$$

Our interactive tools further explain these formulations and how they fit together to define preimums and claims.

---

[8]Note that only one yield estimation is available per year.

Affiliation:

A. Samuel Pottinger
Eric and Wendy Schmidt Center for Data Science and Environment
University of California, Berkeley
201 Wellman Hall, Berkeley 94720, CA, USA
E-mail: sam.pottinger@berkeley.edu
URL: https://gleap.org

Lawson Connor
Department of Agricultural Economics and Agribusiness
University of Arkansas, Fayetteville
URL: https://policy.uark.edu/about/faculty-list/uid/lconnor/name/
Lawson+Connor/

Brookie Guzder-Williams
Eric and Wendy Schmidt Center for Data Science and Environment
University of California, Berkeley
URL: https://dse.berkeley.edu/people/brookie-guzder-williams

Maya Weltman-Fahs
Eric and Wendy Schmidt Center for Data Science and Environment
University of California, Berkeley
URL: https://dse.berkeley.edu/people/maya-weltman-fahs

Nick Gondek
Eric and Wendy Schmidt Center for Data Science and Environment
University of California, Berkeley
URL: https://dse.berkeley.edu/people/nick-gondek

Timothy Bowles
Department of Environmental Science, Policy & Management
University of California, Berkeley
URL: https://nature.berkeley.edu/agroecologylab/lab-members/timothy-
bowles/