Soil organic matter effects on US maize production and crop insurance payouts under drought

Daniel Kane[1*], Mark A. Bradford[1], Emma Fuller[2], Emily E. Oldfield[1], Stephen A. Wood[1,3]

[1]School of the Environment, Yale University, New Haven, CT 06511, USA

[2]Granular, Inc., San Francisco, CA 94103, USA

[3]The Nature Conservancy, Arlington, VA 22203, USA

*Corresponding author: daniel.kane@yale.edu

Climate change is predicted to increase the incidence and severity of droughts and floods[1,2], thereby increasing the risk of crop failures and yield losses[3]. Conservative estimates for maize suggest yields could drop between 20-80% in the US under plausible future climate scenarios[4–6]. Higher levels of soil organic matter may help mitigate yield losses from drought because there are positive relationships between soil organic matter and soil water retention[7,8]. Yet recent work suggests the actual effect of soil organic matter on plant available water is modest[9], and it is unclear whether these effects on water retention are great enough to reduce drought-induced yield losses. This lack of evidence is particularly pronounced at the scales at which many agricultural financing programs operate, meaning agricultural finance incentives are decoupled from soil-based, climate-adaptation strategies. Here, we evaluate relationships among soil organic matter, maize (*Zea mays* L.) yield, and drought between 2000-2016 at the county-level in the United States of America. We show that for 12,376 county-years in US corn-growing counties, those with higher soil organic matter were associated with greater yields and lower yield losses under drought, and this effect was most pronounced under severe droughts. An increase of 1% soil organic matter was associated with a yield increase of 2.2 Mg ha^{-1} under the most severe drought conditions. Further, higher soil organic matter is associated with lower rates of crop insurance payouts. An increase of 1% soil organic matter was associated with a reduction in the mean proportion of liabilities paid by 36% under the most severe drought. Our results suggest that soil organic matter content is an important predictor of the resilience of maize systems to drought at regional to continental scales and that data on soil organic matter should be used in agricultural policy and financial planning.

Rainfed agriculture, which made up 75% of global cropland use as of 2000[10], is highly susceptible to extreme weather conditions, such as heat and drought. Extreme heat accelerates plant development, effectively shortening growing season length and reducing harvest index, and in extreme instances, extreme heat causes plant reproductive failures, such as kernel abortion in maize, that drastically reduce yields[4,11,12]. Similarly, drought leads to elevated vapor pressure deficit which can lead to increased transpiration by plants, closing of stomata, and ultimately reduced rates of photosynthesis that slow plant growth and reduce grain yields[5,13].

Increasing soil organic matter can increase soil water holding capacity on similarly textured soils[9,14] and improve water infiltration[15–17] by supporting greater aggregate formation and, hence, a greater volume of pore spaces[18]. Researchers have argued that soils with higher organic matter can retain more water under vapor pressure deficit, protecting crops from losses induced by extreme heat and drought better than low organic matter soils[7,8,19]. Studies have demonstrated that higher soil organic matter is associated with lower long-term interannual yield variability at regional scales[20,21]. But lower variability is not necessarily indicative of greater resilience or protection against yield losses under adverse conditions. Some field-level studies have shown that practices known to increase soil organic matter can protect yields[22,23], but these do not explicitly test the relative effect of organic matter and do not provide information at county or regional scales which are arguably most relevant to policy initiatives.

In light of these evidence gaps, we quantified the impact of soil organic matter on agronomic risk to drought in the United States of America. We analyzed county-level maize (*Zea mays* L.) yield[24] and crop insurance payouts in the US[25] from 2000 - 2016 in combination with soil characterization data from the US Department of Agriculture's gSSURGO data product[26] and county-level Standardized Precipitation Evapotranspiration Index (SPEI) data from the Centers for Disease Control[27]. Data were gathered for 754 counties where maize production was predominantly rain-fed, representing a total of 12,376 county-years, 5945 of which experienced drought conditions over the summer growing season. We hypothesized that counties with higher levels of soil organic matter in surface soils (0-30 cm depth) where most of the fine-root biomass is found, would be less prone to yield losses in drought years given expected positive effects of soil organic matter on crop water availability, and that as a result, a lower proportion of crop insurance liabilities would be paid out in drought years.

We found that across all county-years and possible weather conditions, soil organic matter content was a strong positive predictor of yield. Soil organic matter had a standardized marginal effect of 0.84, meaning an increase of 1% soil organic matter was associated with an increase in yields of 0.84 Mg ha^{-1}. This observed relationship between soil organic matter and yield is consistent with other studies which show yield increases are associated with higher levels of soil organic matter[28,29]. Our data extend these observations by showing that as drought conditions became more severe, the marginal effect of soil organic matter on yields increased (Figure 1; Supplementary Table 1). For example, under moderate drought conditions ($-0.46 \leq$ SPEI < 0.12) an increase of 1% soil organic matter was associated with an increase in yields of 0.76 Mg ha^{-1}, and under very severe drought conditions (SPEI < -1.02), a 2.2 Mg ha^{-1} increase. Interaction effects of other soil properties with soil organic matter also emerged across different drought categories, but sensitivity analyses of these interaction effects revealed that the main effect of soil organic matter remained robust across typical variation in soil properties (Supplementary Information).

To more fully evaluate whether this greater relative yield advantage under drought is because soil organic matter protects against drought-induced yield losses, we also evaluated the relationship between soil organic matter and annual yield anomaly, the ratio of observed yield to expected yield estimated from long-term trends[30]. Under very severe drought conditions (SPEI < -1.01) an increase of 1% soil organic matter content was associated with a mitigation of yield losses of 12%, under severe drought conditions (-1.01 ≤ SPEI < -0.46) this effect decreased to a 5% mitigation, and then under moderate drought (-0.46 ≤ SPEI < 0.10) the effect was non-significant (Supplementary Table 2). Further examination of yield anomaly data revealed that counties with lower soil organic matter content may outperform expectations based on historical yield trends in favorable conditions, but are more likely to underperform expectations under adverse conditions. Whereas counties with higher soil organic matter consistently yield nearer expected yields based on historical trends, even under drought conditions. For example, in counties with greater than 2.5% soil organic matter content, the global mean of soil organic matter content in this study, the interquartile range for yield anomalies was 98% to 107% of expected yield under normal conditions and 91% to 104% under all drought conditions. By contrast, counties with lower than 2.5% soil organic matter ranged from 99% to 114% of expected yield under normal conditions and 82% to 104% under all drought conditions. This pattern is consistent with previous studies demonstrating that higher soil carbon is associated with lower long-term interannual yield variability[20,21]. But it offers additional insight by demonstrating that those reductions in interannual variability are partly explained by the association of higher soil organic matter and lower yield losses under drought conditions.

Given the decrease in yield risk associated with greater levels of soil organic matter, we expected that lower yield risk would be reflected in crop insurance payouts to farmers. Specifically, we expected that counties with higher soil organic matter would have lower loss cost[31], a metric based on the ratio of total indemnities to total liabilities. Our results support this expectation, showing that soil organic matter is associated with reduced loss cost under drought conditions and that the marginal effect of soil organic matter increases as drought severity increases (Figure 3). Under very severe drought conditions (SPEI $\leq$ -1.02), an increase in soil organic matter of 1% was associated with a 36% reduction in loss cost (Supplementary Table 3). Similar to yield anomaly, though, this effect decreases sharply as SPEI approaches normal. Soil organic matter was associated with an 8.4% reduction in loss cost under severe drought conditions (-1.02 < SPEI $\leq$ -0.46) and just a 4% reduction under moderate drought (-0.46 < SPEI $\leq$ 0.10). Nevertheless, given the expectation of increasing frequency of severe droughts[2,32,33], our results suggest that it would be strategic for rain-fed US agriculture to directly incorporate differences in soil properties into policy and insurance planning for yield resilience.

The fact that we found that soil organic matter appeared to impart such effective protection against yield losses under severe drought appears inconsistent with results from recent studies using large soil databases into how soil organic matter influences the plant available water capacity in soils. Briefly, the ability of soils to provide water to plants is often estimated as 'available water capacity', which typically is the difference in water content of saturated soil samples dried on pressure plates at -1500 kPa and -33 kPa[26,34]. These analyses have suggested that the net impact of soil organic matter on available water capacity is relatively modest and contingent on soil texture[9,35]. To investigate potential discrepancies in conclusions between these past studies and our work, we performed a series of confirmatory path analyses[36,37] to investigate the extent to which soil organic matter associations with yields under drought and non-drought conditions are related to its influence on available water capacity and cation exchange capacity, used here as a proxy measure of soil fertility. We found that soil organic matter was strongly associated with cation exchange capacity but only weakly associated with available water capacity, and under both drought and non-drought conditions, cation exchange capacity and available water capacity were positively associated with yields. However, our confirmatory path analysis also suggested that soil organic matter had an independent, unmediated positive effect on yields under both drought and non-drought conditions (Figure 3; Supplementary Table 4).

These results confirm that soil organic matter has a positive influence on yields via its effects on available water capacity and soil fertility. But it also suggests that soil organic matter likely influences plant water availability and soil fertility in ways not captured by how those properties are commonly measured. Although our analyses cannot resolve these additional influences, we do know that soil organic matter affects other important soil biophysical properties, such as porosity, bulk density, and water infiltration[15,17,18,35]. Favorable changes in all of these properties may increase the soil volume from which plants can draw water and may effectively increase the supply of water to plants between rain events. In addition, soil organic matter is also an important source of key nutrients for plant growth. Under drought conditions, water transpiration and radiation efficiency in maize plants increase with increasing nitrogen fertilizer use[38] and nitrogen fertilizer can be important for maintaining key metabolic functions and increasing yield[39]. Further work is required to ascertain whether soil organic matter has similar, nutrient-mediated effects under drought conditions.

While our findings generally support the rationale of soil health initiatives to increase agriculture resilience by rebuilding soil organic matter[40], we acknowledge that our analyses are based on subcontinental-scale variation in soil organic matter and yield outcomes. As such, they cannot be used to argue directly that field-scale increases in soil organic matter, achieved through agricultural management, will lead to meaningful yield protection under drought. Additional evidence is required to understand whether increases in soil organic matter over time are associated with resilience to drought conditions at the farm scale. Nevertheless, our results do appear to have the potential to directly inform agricultural financing programs and policy. At present, knowledge of risk is incorporated into US Federal Crop Insurance Programs (FCIP) only indirectly. Premiums are based on the Actual Production History (APH) of a given area and farm, and current policy dictates that APH be calculated based on a 10-year trend excluding years in which yield losses were extreme[41]. While differences in soil organic matter and other biophysical limitations to resilience are arguably endogenous to these yield data, APH may become less predictive of risk under future climate scenarios where drought frequency is predicted to increase. It may therefore be more strategic for policy planning for agricultural resilience to explicitly consider differences in soil properties, such as organic matter levels, across counties. For instance, if maize yields in counties with low soil organic matter are particularly vulnerable to drought, it may make more sense to incentivize a transition to crops that are more appropriate to soil and predicted climate conditions in those counties, than to focus exclusively on economic protection through insurance.

Our analyses demonstrate that counties with higher mean soil organic matter content are associated with lower maize yield loss due to drought, that this relationship is strongest under severe drought conditions, and this increased yield protection translates to lower crop insurance payouts under drought conditions. Furthermore, we demonstrate that these impacts are not solely mediated through the impact of soil organic matter on conventional measures of plant available water, but likely occur through additional pathways that influence soil water supply and use by plants, which appear to collectively provide the yield protection benefits we document here. At least at the county level then for US rain-fed maize agriculture, soil organic matter content appears to be an important predictor of resilience to the type of drought conditions that are likely to occur more frequently under future climate scenarios. Further work should investigate whether similar benefits of soil organic matter for yield protection are afforded by agricultural managements that build organic matter in agricultural soils worldwide.

In the interim, our analyses highlight the potential value of integrating soil information into resilience planning as agricultural outcomes become more uncertain with the increasing incidence and severity of extreme weather events.

Figures

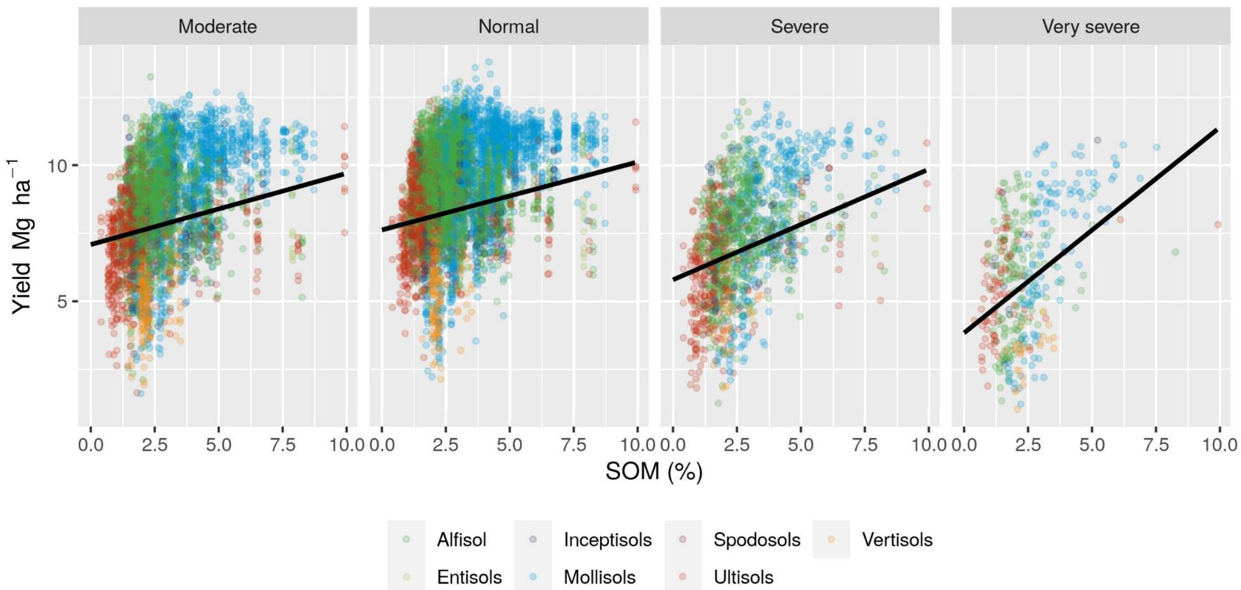

Figure 1. Maize yields increase with soil organic matter content (%), with the effect becoming more pronounced with increasing drought severity. Drought levels are based on the following ranges of Standardized Precipitation Evapotranspiration Index (SPEI): Very severe < -0.99; Severe $\geq$ -0.99 < -0.44; Moderate $\geq$ -0.44 < 0.12; Normal $\geq$ 0.12. Trendlines represent predicted yields based on the marginal effect of soil organic matter.

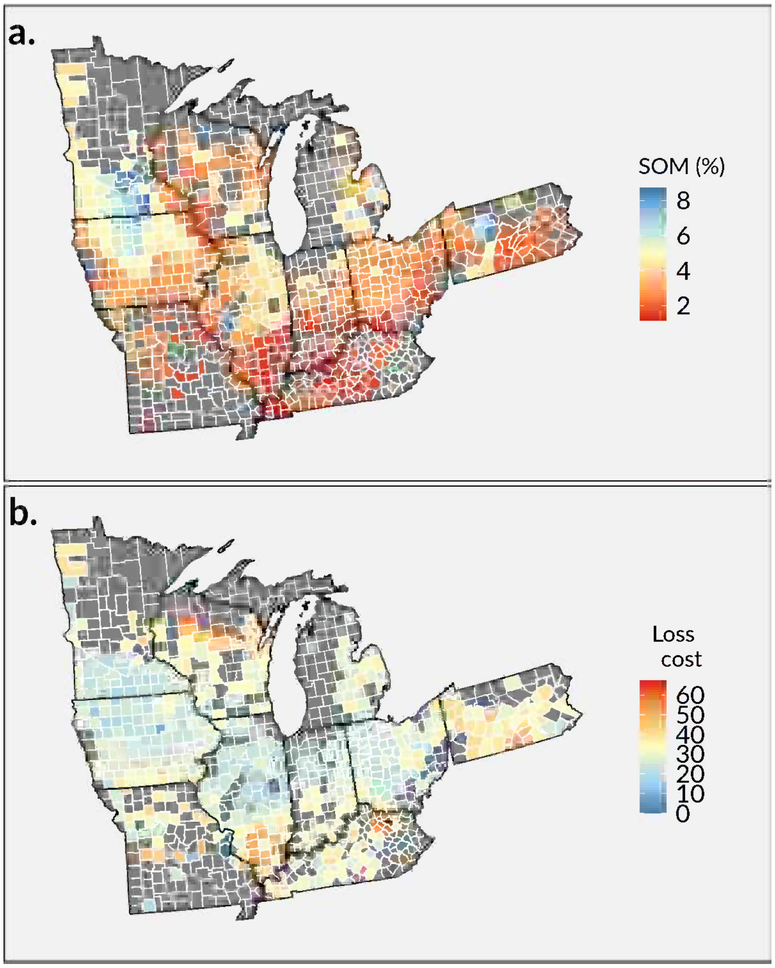

Figure 2. Higher levels of soil organic matter were associated with lower loss cost under drought conditions. Figure 2a is a map of mean soil organic matter content across counties (n=730) within a subset of states (PA, WI, MI, MN, IL, IN, IA, OH, MO, and KY) included in this study. Figure 2b is a map of mean loss cost under all drought conditions in those same counties. The marginal effect of soil organic matter on loss cost under all drought conditions is -6.47 (σ=0.84, p<0.001).

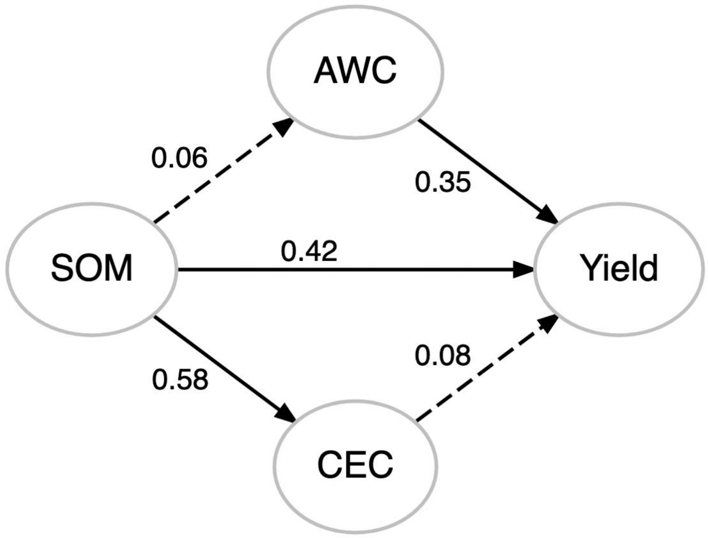

Figure 3. Increases in yield under drought associated with increasing soil organic matter appear only partially mediated by increases to soil available water capacity induced by higher soil organic matter. Figure 3 is a diagram of a structural equation model of best fit developed during a series of confirmatory pathway analyses. Numbers on each arrow represent standardized coefficients of the relationship between the variables connected, and the direction of arrows indicate variable dependence. Solid lines indicate statistically significant ($p \leq 0.05$) relationships and dotted lines indicate non-significant relationships.

METHODS:

Data collection

Maize yield data

We collected mean maize (*Zea mays* L.) yield (Mg ha^{-1}) data for all US counties between the years 2000 to 2016 for which maize yield data were available from the USDA National Agricultural Statistics Service (NASS)[24] accessed via the *rnassqs* package[42] in R. All work in R was conducted in R version 3.6.3[43]. Data were limited to the years 2000-2016 to limit the confounding effect improvements in maize genetics may have on yield data, and other data used to estimate drought incidence in each county detailed below were only available beginning in the year 2000. We removed data from counties in which corn was not grown for at least 16 of the 17 years in the study period. Yield data were then detrended on a county basis per the method detailed in Lu et al. 2017[30]. Briefly, we fit a locally weighted regression (LOWESS) model wherein yield was the dependent variable and year was the independent variable. Models were fit using the R package *caret*[44] using a 10-fold cross validation approach wherein the span parameter was constrained to a range of 5-10 years and the degree parameter was allowed to be either 1 (linear) or 2 (polynomial). We then added the yearly residuals of these models to the long-term county average yield to estimate detrended yield for each county-year in the study period, and we also divided the observed yield by the predicted yield of these models for each county-year to estimate yield anomaly.

Because soil organic matter might protect yields by improving soil water dynamics, irrigation could mask the effects of soil organic matter on agricultural resilience. As such, we restricted the analysis to primarily non-irrigated acres. We also retrieved data from USDA NASS[24] via the *rnassqs*[42] on the total number of corn acres harvested in each county-year and the total number of irrigated corn acres harvested in each county-year for the agricultural census years of 1997, 2002, 2007, and 2012. We then calculated the percent of maize-growing acres that were irrigated in each county for each census year, averaged those figures across all four census years, and filtered yield data to those counties in which corn-growing acres were on average <u><</u>5% irrigated during the study period and in which average acres irrigated had a standard deviation of <u><</u>1% across the four census years for which data were retrieved.

Drought data

We retrieved county-level Standardized Precipitation-Evapotranspiration Index (SPEI) figures from the Center for Disease Control[27]. SPEI is a multi-scalar drought index based on the similar Standardized Precipitation Index[45,46]. Differences between cumulative monthly precipitation and potential evapotranspiration are calculated for a chosen time scale (i.e. 1 month, 3 months, etc.). These figures are then standardized based on a log-logistic distribution so that they are comparable across locations. The CDC dataset we used is based on a 1-month timescale. We summarized SPEI data in each county-year by calculating the mean monthly SPEI for the summer growing season of May to August.

In addition to SPEI, we also retrieved drought data from The National Drought Mitigation Center[47], which reports on a daily basis the percent area of each county that is at each level of their drought index: D0 (no drought) – D4 (severe drought). This drought index is a categorization based on the Palmer Drought Severity Index, Standardized Precipitation Index, soil and streamflow models, and local expert assessment/verification by USDA field agents[48]. We then converted these coverage statistics to a daily Drought Severity Classification Index (DSCI), per the instructions of the US Drought Monitor website[49]. Briefly, DSCI is a weighted sum of the percent area of each county under each drought index level. We then subset daily DSCI data to the months of May to August, the months that are most crucial to maize growth and yield, and averaged them for each county-year across the study period. DSCI data were not used in our primary analyses but were instead used in a set of sensitivity analyses (Supplementary Information) to determine the relative importance of how drought is quantified on estimating the mitigating effect of soil organic matter on yields under drought.

Crop insurance data

The USDA Risk Management Agency collates a variety of data on the US crop insurance market on an annual basis, including total liabilities, total indemnities, and cause of loss. These data are available at the county-level via the USDA Southwest Climate Hub's AgRisk Viewer[25]. We retrieved data on total liabilities and total indemnities in USD for maize due to loss by drought for the same set of counties for which we retrieved yield data. We then used these data to calculate loss cost[31], the ratio of total indemnities to total liabilities in a given county-year. Loss cost implicitly accounts for differences in USD figures across years due to inflation, as well as differences across counties due to differences in total output and coverage patterns.

Soil data

To aggregate soil data for each county in our analysis we first identified those areas in each county that are primarily used for maize production. To do so, we used the “Corn Frequency” raster available at USDA NASS’s CropScape portal, which reports how many years between 2008 – 2017 each pixel was used for maize production[50]. We then subset this raster to only those pixels in which maize was produced for 2 or more years to generate a masking layer which we then used in subsequent steps.

Soil data were collated from USDA’s Gridded Soil Survey Geographic (gSSURGO) Database, a spatial geodatabase based on the extensive field survey and laboratory work of the USDA National Cooperative Soil Survey[26]. The gSSURGO database is spatially organized as a series of discrete polygons referred to as map units that are composed of different component soil series. Associated with each soil series is characterization data organized by pedological horizon, including soil texture, soil organic matter, and measures of soil biophysical characteristics. As such, we first used the *aqp*[51] package in R to convert characterization data for each component soil series to a fixed depth increment of 0-30 cm to represent the typical rooting zone of maize. We then calculated a representative map unit value for each soil characteristic soil organic matter (%), clay content (%), H3O+ concentration (mol), cation exchange capacity (meq 100 g soil^{-1}), and available water capacity (%) by calculating a map unit mean weighted by the relative proportions of each component soil series in a map unit. We then converted the data to a raster format and used the masking layer described previously to remove soil data in each county from areas where maize is not consistently grown. Finally, we used these masked rasters to calculate county-level means for all soil properties we then used in our analyses.

Data analysis

Initial data analysis demonstrated that the yield response to SPEI begins to saturate above SPEI values of 0, indicating that when the balance of precipitation and evapotranspiration is negative, yields drop below the typical yield potential of a given area. Additionally, we found that when SPEI decreased (i.e. drought conditions became more severe) the impact of soil organic matter was greater on maize yield (Mg ha^{-1}), yield anomaly, and loss cost (Supplementary Information). As such, we chose to subset our data into different levels of drought severity based on SPEI and then analyze each subset to understand how the effect of soil organic matter on each outcome variable changed as drought severity changed. Drought severity thresholds were calculated based on the global mean and standard deviation of SPEI. Very severe drought was defined as greater than two standard deviations from the global mean (SPEI $\leq$ -1.02); severe drought as between one and two standard deviations from the mean (-1.02 < SPEI $\leq$ -0.46); moderate drought as between one standard deviation from the mean and the mean (-0.46 < SPEI $\leq$ 0.10); and normal conditions as greater than or equal to the mean (SPEI > 0.10).

Within each of these drought subsets we then fit a series of models wherein the dependent variable in these models was either yield (Mg ha^{-1}), yield anomaly, or loss cost. Independent variables across all models included soil organic matter, soil clay content, and soil H_3O^+

concentration. These variables were chosen by fitting a multivariate linear model with multiple potential independent variables then assessing variance inflation factors to eliminate spurious, highly collinear variables (Supplementary Information). We also included a random effect of State to account for possible impacts of geographic differences in management and production environment on model outcomes not accounted for in the data we collected. For yield and yield anomaly, linear mixed effects models including all possible interaction effects were fit using a restricted maximum likelihood approach in the *lme4*[52] package in R. For loss cost, a mixed effects Tobit regression model was fit using a Newton-Raphson maximization approach in the *censReg*[53] package in R to account for the fact that loss cost was left censored at a value of 0. Prior to model fitting, all observations of each independent variable were standardized so that coefficient estimates would also be standardized. Data standardization was done by subtracting the mean of a given variable from each observation and then dividing that value by 2x the standard deviation of that variable[54].

Finally, we conducted a series of confirmatory path analyses to evaluate to what extent the impacts of soil organic matter on maize yields were mediated by their impacts on soil available water capacity and cation exchange capacity, a proxy variable for soil fertility, under both drought and normal conditions. First, we split data into those observations from normal SPEI years (SPEI > 0.10) and drought SPEI years (SPEI $\leq$ 0.10), and then calculated the mean yield (Mg ha^{-1}) for each county under either set of conditions. We then employed a piecewise structural equation modeling approach using the *piecewiseSEM*[55] package in R to fit models in which the effects of soil organic matter were either partially mediated or fully mediated by its effects on available water capacity and cation exchange capacity. Briefly, the fully mediated SEM was such that available water capacity and cation exchange capacity were modeled as functions of soil organic matter and yield was modeled as a function of available water capacity and cation exchange capacity. Whereas, in the partially mediated SEM, available water capacity and cation exchange capacity were modeled the same way, but soil organic matter was included as an additional independent variable for modeling yield. To determine the most parsimonious model, we compared models via an analysis of variance and on the basis of AIC/BIC scores (Supplementary Information). Coefficients were extracted from the final model and standardized to then assess whether or not effects of organic matter on yields under drought were mediated by its impacts on available water capacity and cation exchange capacity. An initial analysis with all data in either SPEI category indicated a negative relationship between cation exchange capacity and yields. When we manually inspected data we found that this result was the consequence of outlying cation exchange capacity values, defined as being two standard deviations greater than the mean (CEC $\geq$ 31.9), with extremely high clay content (Supplementary Information). As such, we removed these outliers for the final path analysis to better estimate effects on typical soils, but include results from a path analysis including these observations in Supplementary Information.

Acknowledgements

This research was conducted as part of the Managing Soil Carbon working group with the Science For Nature and People Partnership. DK was supported by a Yale School of the Environment Fellowship, and a 2017 Soil Health Institute Literature Review grant from the Soil Health Institute was important in supporting early work that led to the development of this article.

Author contributions

DK, EF, and SW conceived the original project idea. DK and EF developed initial datasets and data processing code. DK and SW developed and verified the analytical methods. All authors contributed to analysis and interpretation of results and development of the final manuscript.

Additional information

Supplementary information is available for this paper. Correspondence and requests for materials should be addressed to DK.

References

1. Luber, G. & McGeehin, M. Climate Change and Extreme Heat Events. *Am. J. Prev. Med.* **35**, 429–435 (2008).
2. Meehl, G. A. *et al.* Global climate projections. *Clim. Change* **3495**, 747–845 (2007).
3. Urban, D., Roberts, M. J., Schlenker, W. & Lobell, D. B. Projected temperature changes indicate significant increase in interannual variability of U.S. maize yields. *Clim. Change* **112**, 525–533 (2012).
4. Hsiao, J., Swann, A. L. S. & Kim, S.-H. Maize yield under a changing climate: The hidden role of vapor pressure deficit. *Agric. For. Meteorol.* **279**, 107692 (2019).
5. Lobell, D. B. *et al.* The critical role of extreme heat for maize production in the United States. *Nat. Clim. Change* **3**, 497–501 (2013).
6. Schlenker, W. & Roberts, M. J. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci.* **106**, 15594–15598 (2009).
7. Bot, A. & Benites, J. The importance of soil organic matter: Key to drought-resistant soil and sustained food production. (Food and Agriculture Organization of the United Nations, 2005).
8. Iizumi, T. & Wagai, R. Leveraging drought risk reduction for sustainable food, soil and climate via soil organic carbon sequestration. *Sci. Rep.* **9**, 1–8 (2019).
9. Minasny, B. & McBratney, A. B. Limited effect of organic matter on soil available water capacity. *Eur. J. Soil Sci.* **69**, 39–47 (2018).
10. Portmann, F. T., Siebert, S. & Döll, P. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and

hydrological modeling. *Glob. Biogeochem. Cycles* **24**, (2010).

11. Craufurd, P. Q. & Wheeler, T. R. Climate change and the flowering time of annual crops. *J. Exp. Bot.* **60**, 2529–2539 (2009).
12. Sage, R. F. & Kubien, D. S. The temperature response of C3 and C4 photosynthesis. *Plant Cell Environ.* **30**, 1086–1106 (2007).
13. Lobell, D. B. *et al.* Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. *Science* **344**, 516–519 (2014).
14. Hudson, B. D. Soil organic matter and available water capacity. *J. Soil Water Conserv.* **49**, 189–194 (1994).
15. Boyle, M., Frankenberger, W. T. & Stolzy, L. H. The Influence of Organic Matter on Soil Aggregation and Water Infiltration. *J. Prod. Agric.* **2**, 290–299 (1989).
16. Elliott, J. A. & Efetha, A. A. Influence of tillage and cropping system on soil organic matter, structure and infiltration in a rolling landscape. *Can. J. Soil Sci.* **79**, 457–463 (1999).
17. Franzluebbers, A. J. Water infiltration and soil structure related to organic matter and its stratification with depth. *Soil Tillage Res.* **66**, 197–205 (2002).
18. Lado, M., Paz, A. & Ben-Hur, M. Organic Matter and Aggregate Size Interactions in Infiltration, Seal Formation, and Soil Loss. *Soil Sci. Soc. Am. J.* **68**, 935–942 (2004).
19. Carminati, A. & Javaux, M. Soil Rather Than Xylem Vulnerability Controls Stomatal Response to Drought. *Trends Plant Sci.* S1360138520301187 (2020).
20. Williams, A. *et al.* Soil Water Holding Capacity Mitigates Downside Risk and Volatility in US Rainfed Maize: Time to Invest in Soil Organic Matter? *PLOS ONE* **11**, e0160974 (2016).
21. Pan, G., Smith, P. & Pan, W. The role of soil organic matter in maintaining the productivity

and yield stability of cereals in China. *Agric. Ecosyst. Environ.* **129**, 344–348 (2009).

22. Bowles, T. M. *et al.* Long-Term Evidence Shows that Crop-Rotation Diversification Increases Agricultural Resilience to Adverse Growing Conditions in North America. *One Earth* **2**, 284–293 (2020).

23. Gaudin, A. C. M. *et al.* Increasing Crop Diversity Mitigates Weather Variations and Improves Yield Stability. *PLoS ONE* **10**, e0113261 (2015).

24. USDA National Agricultural Statistics Service. Data and Statistics. https://www.nass.usda.gov/Data_and_Statistics/index.php (2020).

25. USDA Southwest Climate Hub. AgRisk Viewer. https://swclimatehub.info/rma/rma-data-viewer.html (2020).

26. Soil Survey Staff. Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations served by the USDA-NRCS. United States Department of Agriculture, Natural Resources Conservation Service. *Geospatial Data Gateway* https://gdg.sc.egov.usda.gov/ (2015).

27. National Environmental Public Health Tracking Network. Standardized Precipitation Evapotranspiration Index, 1895-2016 | Data | Centers for Disease Control and Prevention. *Data.CDC.gov* https://data.cdc.gov/Environmental-Health-Toxicology/Standardized-Precipitation-Evapotranspiration-Inde/6nbv-ifib/data (2020).

28. Oldfield, E. E., Bradford, M. A. & Wood, S. A. Global meta-analysis of the relationship between soil organic matter and crop yields. *Soil* **5**, 15–32 (2019).

29. Oldfield, E. E., Wood, S. A. & Bradford, M. A. Direct evidence using a controlled greenhouse study for threshold effects of soil organic matter on crop growth. *Ecol. Appl.* **n/a**, e02073.

30. Lu, J., Carbone, G. J. & Gao, P. Detrending crop yield data for spatial visualization of drought impacts in the United States, 1895–2014. *Agric. For. Meteorol.* **237–238**, 196–208 (2017).

31. Reyes, J. J. & Elias, E. Spatio-temporal variation of crop loss in the United States from 2001 to 2016. *Environ. Res. Lett.* **14**, 074017 (2019).

32. Hayhoe, K., VanDorn, J., Croley, T., Schlegal, N. & Wuebbles, D. Regional climate change projections for Chicago and the US Great Lakes. *J. Gt. Lakes Res.* **36**, 7–21 (2010).

33. Adams, R. M. *et al.* Global climate change and US agriculture. *Nature* **345**, 219–224 (1990).

34. Cassel, D. K. & Nielsen, D. R. Field Capacity and Available Water Capacity. in *Methods of Soil Analysis* 901–926 (John Wiley & Sons, Ltd, 2018).

35. Libohova, Z. *et al.* Reevaluating the effects of soil organic matter and other properties on available water-holding capacity using the National Cooperative Soil Survey Characterization Database. *J. Soil Water Conserv.* **73**, 411–421 (2018).

36. Lefcheck, J. *Structural Equation Modeling in R for Ecology and Evolution*. https://jslefche.github.io/sem_book/index.html (2019).

37. Shipley, B. Confirmatory path analysis in a generalized multilevel context. *Ecology* **90**, 363–368 (2009).

38. Teixeira, E. I. *et al.* The impact of water and nitrogen limitation on maize biomass and resource-use efficiencies for radiation, water and nitrogen. *Field Crops Res.* **168**, 109–118 (2014).

39. Zhang, L.-X., Li, S.-X., Zhang, H. & Liang, Z.-S. Nitrogen Rates and Water Stress Effects on Production, Lipid Peroxidation and Antioxidative Enzyme Activities in Two Maize (Zea mays

L.) Genotypes. *J. Agron. Crop Sci.* **193**, 387–397 (2007).

40. Bradford, M. A. *et al.* Soil carbon science for policy and practice. *Nat. Sustain.* **2**, 1070–1072 (2019).

41. Bryant, L. & O'Connor, C. Creating Incentives to Improve Soil Health Through the Federal Crop Insurance Program. in *Global Soil Security* (eds. Field, D. J., Morgan, C. L. S. & McBratney, A. B.) 403–409 (Springer International Publishing, 2017).

42. Potter, N. *et al. rnassqs: Access the NASS 'Quick Stats' API*. R package v. 0.5.0. https://CRAN.R-project.org/package=rnassqs (2019).

43. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2020).

44. Kuhn, M. *et al. caret: Classification and Regression Training*. R package v. 6.0-86. https://CRAN.R-project.org/package=caret (2020).

45. Vicente-Serrano, S. M. *et al.* Performance of Drought Indices for Ecological, Agricultural, and Hydrological Applications. *Earth Interact.* **16**, 1–27 (2012).

46. Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *J. Clim.* **23**, 1696–1718 (2009).

47. The National Drought Mitigation Center. U.S. Drought Monitor. https://droughtmonitor.unl.edu/ (2019).

48. The National Drought Mitigation Center. Statistics Explanation. https://droughtmonitor.unl.edu/AboutUSDM/AbouttheData/StatisticsExplanation.aspx (2019).

49. The National Drought Mitigation Center. Drought Severity and Coverage Index. https://droughtmonitor.unl.edu/AboutUSDM/AbouttheData/DSCI.aspx (2019).

50. Han, W., Yang, Z., Di, L. & Yue, P. A Geospatial Web Service Approach for Creating on-Demand Cropland Data Layer Thematic Maps. *Trans. Asabe* **57**, 239–247 (2014).

51. Beaudette, D., Roudier, P. & Brown, A. *aqp: Algorithms for Quantitative Pedology*. R package v. 1.19. https://CRAN.R-project.org/package=aqp (2020).

52. Bates, D. *et al. lme4: Linear Mixed-Effects Models using 'Eigen' and S4*. R package v. 1.1-23. https://CRAN.R-project.org/package=lme4 (2020).

53. Henningsen, A. *censReg: Censored Regression (Tobit) Models*. R package v. 0.5-32. https://CRAN.R-project.org/package=censReg (2020).

54. Gelman, A. Scaling regression inputs by dividing by two standard deviations. *Stat. Med.* **27**, 2865–2873 (2008).

55. Lefcheck, J., Byrnes, J. & Grace, J. *piecewiseSEM: Piecewise Structural Equation Modeling*. (2019).

Soil organic matter effects on US maize production and crop insurance payouts under drought: Supplementary Information

Supplementary methods:

Summary of data

Filtering to counties that had $\geq$ 15 yield observations and in which acres planted to maize were $\leq$5% irrigated reduced the dataset from an initial 2286 counties to 748 counties. The majority of these counties are located in areas typically identified as the US Corn Belt[1] with the notable exception of some counties in NE, KS, ND, and SD where irrigated production predominates. The overall mean of detrended maize yields was 8.59 Mg ha^{-1}, and yields varied across counties from 1.03 to 13.79 Mg ha^{-1}. Average soil organic matter content per the gSSURGO soil database across the studied counties was 2.9% and ranged from 0.41% to 9.9%. Soil orders included Alfisols, Ultisols, Mollisols, Inceptisols, Entisols, Spodosols and Vertisols, and Alfisols were the predominant order across all counties.

Drought data reflected significant, recorded drought events from within the study period. For example, several counties included in the study indicated extreme values on drought metrics in 2012 when a widespread drought affected a substantial portion of the U.S. corn belt and damaged yields[2]. Other less widespread and more acute drought events also appeared in these data, such as in 2002 when drought impacted several counties in OH, IL, and IN but others in IA were unaffected. Of the 12,376 county-years included in this study, 5945 experienced drought conditions, defined as SPEI below the global mean SPEI of 0.10 for the dataset.

Selection of independent variables

Since many soil properties predictive of crop yields are likely to be collinear with soil organic matter, inclusion of these properties as independent variables may reduce the accuracy of coefficient estimates for soil organic matter. As such, we performed an initial model specification analysis to determine a set of independent variables that would reduce collinearity. We fit a multivariate linear regression with yield (Mg ha^{-1}) as the dependent variable and no interaction effects. Independent variables included soil organic matter (%), soil clay content (%), soil H_3O^+ concentration (mol), and cation exchange capacity (meq 100g $soil^{-1}$). We then evaluated variance inflation factors to determine which independent variables could be removed to reduce multicollinearity. Cation exchange capacity had the highest variance inflation factor with a value of 4.6. We removed cation exchange capacity then fit the model again with the remaining three variables. Variance inflation factors remained within tolerance (1-2), so they were kept for use in further analyses.

Initial analyses and drought index sensitivity analysis

Prior to subsetting data into different ranges of drought severity, we fit an initial model to assess if soil organic matter interacted at all with the effect of drought conditions to affect yields. The dependent variable in these models was yield (Mg ha^{-1}) and the independent variables included soil organic matter (%), soil clay content (%), soil H_3O^+ concentration (mol), and mean standardized precipitation evapotranspiration index (SPEI) for the months of May to August. We also included a random effect of State to account for possible impacts of geographic differences in management and production environment on model outcomes not accounted for in the data we collected. In addition, we also developed a set of parallel models using the Drought Severity Classification Index (DSCI) to perform a sensitivity analysis to evaluate whether or not the interaction effects of soil properties with yield changed if drought was quantified in a different manner. Models including all possible interaction effects were fit using a restricted maximum likelihood approach in the *lme4*[3] package in R. Prior to model fitting, all observations of each independent variable were standardized so that coefficient estimates would also be standardized. Data standardization was done by subtracting the mean of a given variable from each observation and then dividing that value by 2x the standard deviation of that variable.

Results from the models using SPEI indicated that when SPEI decreased (i.e. drought conditions became more severe) the impact of soil organic matter was greater on maize yield (Mg ha-1) (Supplementary Table 5). Sensitivity analyses confirmed that when DSCI was substituted for SPEI, directionality, significance, and relative magnitude of estimated coefficients remained largely the same (Supplementary Table 6), as DSCI increased (drought conditions became more severe), the effect of soil organic matter on yields became more positive. To better understand the nature of this interaction effect, we then subset data to different ranges of drought severity for further analysis as described in the main text.

Analyses of interaction effects

While our primary interest was determining the main effect of soil organic matter on yields under variable drought conditions, several interaction effects with soil organic matter emerged for models in which yield or loss cost were the dependent variable. Here we document them and conduct a series of sensitivity analyses to understand why and how they emerged, and to demonstrate that the positive effects of soil organic matter on yields were robust across drought conditions and typical variation in soil properties.

For models in which yield (Mg ha^{-1}) was the dependent variable, which interaction effects emerged as statistically significant ($p < 0.05$) varied by drought category. For very severe drought, no interaction effects emerged; under severe and moderate drought two-way interaction effects emerged between soil organic matter and H_3O^+ concentration, soil organic matter and clay, as well as a three-way interaction effect between soil organic matter, clay, and H_3O^+ concentration; and under normal conditions, a two-way interaction between soil organic matter and clay emerged (Supplementary Table 1). Matching significant interaction effects had the same direction and similar magnitude across drought levels.

Analysis of model residuals indicated that the interaction effect of soil organic matter and soil clay content is such that the positive effect of soil organic matter on yields diminishes as clay content diminishes. Multiple sensitivity analyses to test the sensitivity of this effect to observations on either end of the distribution of clay content did not substantially change coefficients or their *p* values. These sensitivity analyses indicate that this is a valid interaction effect and that organic matter has a more limited positive effect on yields in counties where clay content is limited.

The negative interaction effect of soil organic matter and H_3O^+ concentration is such that the positive effect of soil organic matter on yields is diminished as H_3O^+ concentration increases. This interaction effect appears to be primarily driven by observations from 6 counties with very high H_3O^+ concentration that are also high soil organic matter. When observations from these counties are removed, the interaction effect is rendered non-significant (p=0.63 at moderate drought). Removing these observations also rendered the three-way interaction effect of soil organic matter, clay, and H_3O^+ concentration non-significant (p=0.76 at moderate drought), as these same observations were also very high clay content.

For models in which loss cost was the dependent variable, negative two-way interaction effects emerged between soil organic matter and clay content, as well as between soil organic matter and H_3O^+ concentration at the severe and moderate drought levels; at the very severe drought level a negative three-way interaction effect between soil organic matter, clay, and H_3O^+ concentration (Supplementary Table 3). The interaction effects between soil organic matter and clay content reflected the corresponding effects for yield discussed above. The mitigating (i.e. negative) effect of soil organic matter on loss cost diminished as clay content decreased, and sensitivity analyses indicate that this effect is robust against outliers.

The interaction effects between soil organic matter and H_3O^+ concentration did not reflect the corresponding effects from models in which yield was the dependent variable. As H_3O^+ concentration increases, the negative effect of soil organic matter on loss cost is made more strongly negative. Exploration of the residuals and the raw data indicated that this effect was due to a limited number of observations that had high H_3O^+ concentrations and high soil organic matter. At the severe drought level a very small number (n=5) of these observations also had a loss cost value of 0. When these observations were removed, the interaction effect was rendered non-significant (p=0.72). Similarly, at the moderate drought level a small number of observations (n=29) with both low clay and soil organic matter content that also had a loss cost of 0 appear to have been responsible for this interaction effect. Removing these observations rendered the effect non-significant (p=0.93). Given that this interaction effect was not robust to sensitivity analyses and was caused by opposite outlier cases at either drought level, it does not appear to be robust but is instead the result of anomalous observations.

Finally, the three-way interaction at the severe drought level is such that the negative effect of soil organic matter on loss cost becomes more strongly negative when soil clay content and H_3O^+ concentration increase. Evaluation of residuals and the data itself seem to indicate that this effect is largely the result of a small number of counties in Iowa that had normal yields, and thus low loss cost, despite bad weather conditions. These counties had low clay content and H_3O^+ concentrations in the lower quartile, producing the interaction effect we observed. Removing these observations (n=18) rendered the effect non-significant (p=0.13).

Structural equation model fitting

To determine the best, most parsimonious structural equation models for our confirmatory path analysis, we fit two candidate models for data from either drought or non-drought conditions in which the effect of soil organic matter on yields was either fully or partially mediated by its impacts on available water capacity and cation exchange capacity. We then compared these models for best fit based on Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and an analysis of variance. In both cases, the model in which the effect of soil organic matter on yield was only partially mediated by its effects on available water capacity had lower AIC/BIC, and an analysis of variance indicated the partially mediated model was significantly different than the fully mediated model (Supplementary Table 7). This initial analysis also resulted in a negative effect of cation exchange capacity on yields under both drought and non-drought conditions. As this result runs counter to previous evidence suggesting cation exchange capacity is an important positive predictor of yields, we manually inspected the data to find that this result was the consequence of outlying cation exchange capacity values that correlate to extremely high clay content. To make our confirmatory path analysis more representative of conventional, maize-growing soil types, we removed these observations (n=20) and then refit the structural equation models. The partially mediated model was again the better fitting model, and coefficients for soil organic matter and available water capacity's influence on yields remained similar, but cation exchange capacity's coefficient changed to become positive (Supplementary Table 8).

Supplementary Tables:

Drought	Term	Coefficient	Std. error	p
Very severe	SOM	2.1976	0.3339	< 0.0001
	Clay	0.4999	0.3355	0.1371
	H_3O^+	0.5665	0.6013	0.3467
	SOM:Clay	0.6549	0.6312	0.3001
	SOM:H_3O^+	0.0426	0.9302	0.9635
	Clay:H_3O^+	1.0140	0.9089	0.2653
	SOM:Clay:H_3O^+	0.9120	1.6866	0.5890
Severe	SOM	1.1820	0.1270	< 0.0001
	Clay	-0.1776	0.1340	0.1852
	H_3O^+	-0.3637	0.1716	0.0342
	SOM:Clay	0.8457	0.2340	0.0003
	SOM:H_3O^+	-0.5317	0.2429	0.0288
	Clay:H_3O^+	-0.0274	0.2900	0.9249
	SOM:Clay:H_3O^+	-1.6099	0.5223	0.0021
Moderate	SOM	0.7616	0.0687	< 0.0001
	Clay	0.3132	0.0718	< 0.0001
	H_3O^+	-0.3794	0.0949	0.0001
	SOM:Clay	0.6793	0.1189	< 0.0001
	SOM:H_3O^+	-0.2837	0.1264	0.0249
	Clay:H_3O^+	-0.2869	0.1549	0.0640
	SOM:Clay:H_3O^+	-0.5372	0.2941	0.0678
Normal	SOM	0.7266	0.0514	< 0.0001
	Clay	0.2085	0.0577	0.0003
	H_3O^+	-0.5189	0.0856	< 0.0001
	SOM:Clay	0.5992	0.0985	< 0.0001
	SOM:H_3O^+	-0.1456	0.1219	0.2322
	Clay:H_3O^+	-0.2103	0.1445	0.1456
	SOM:Clay:H_3O^+	0.1430	0.2600	0.5822

Supplementary Table 1: Results of linear mixed effects models across multiple levels of drought severity relating yield (Mg ha^{-1}) to soil organic matter (SOM, %), clay (%), and H_3O^+ concentration (mol). Coefficients for terms are standardized. Sample numbers for each drought level are as follows: Very severe, n=410; Severe, n= 1537; Moderate, n=3998; Normal, n=6431.

Drought	Term	Coefficient	Std. error	*p*
Very severe	SOM	0.1154	0.0362	0.0015
	Clay	0.0203	0.0358	0.5700
	H_3O^+	-0.0095	0.0650	0.8834
	SOM:Clay	-0.0246	0.0680	0.7184
	SOM:H_3O^+	-0.1604	0.1011	0.1135
	Clay:H_3O^+	-0.0607	0.0984	0.5376
	SOM:Clay:H_3O^+	-0.2421	0.1827	0.1858
Severe	SOM	0.0478	0.0135	0.0004
	Clay	-0.0304	0.0138	0.0287
	H_3O^+	0.0018	0.0188	0.9249
	SOM:Clay	0.0075	0.0258	0.7714
	SOM:H_3O^+	-0.0172	0.0272	0.5266
	Clay:H_3O^+	0.0365	0.0324	0.2600
	SOM:Clay:H_3O^+	-0.0934	0.0590	0.1139
Moderate	SOM	0.0016	0.0068	0.8150
	Clay	0.0155	0.0068	0.0230
	H_3O^+	0.0079	0.0096	0.4081
	SOM:Clay	0.0050	0.0123	0.6846
	SOM:H_3O^+	0.0045	0.0133	0.7366
	Clay:H_3O^+	0.0033	0.0161	0.8353
	SOM:Clay:H_3O^+	0.0273	0.0312	0.3803
Normal	SOM	-0.0261	0.0048	< 0.0001
	Clay	0.0016	0.0053	0.7605
	H_3O^+	-0.0058	0.0080	0.4678
	SOM:Clay	-0.0088	0.0093	0.3449
	SOM:H_3O^+	-0.0072	0.0116	0.5345
	Clay:H_3O^+	-0.0185	0.0137	0.1778
	SOM:Clay:H_3O^+	0.0104	0.0249	0.6749

Supplementary Table 2: Results of linear mixed effects models across multiple levels of drought severity relating yield anomaly to soil organic matter (SOM, %), clay (%), and H_3O^+ concentration (mol). Coefficients for terms are standardized. Sample numbers for each drought level are as follows: Very severe, n=410; Severe, n= 1537; Moderate, n=3998; Normal, n=6431.

Drought	Term	Coefficient	Std. error	*p*
Very severe	SOM	-36.0966	4.7641	< 0.0001
	Clay	-17.8897	4.6736	0.0001
	H_3O^+	-32.0566	9.9484	0.0013
	SOM:Clay	-13.4321	9.5874	0.1612
	SOM:H_3O^+	-23.1025	15.2054	0.1287
	Clay:H_3O^+	-42.9249	13.4872	0.0015
	SOM:Clay:H_3O^+	-53.6069	24.4468	0.0283
Severe	SOM	-8.3689	1.4140	< 0.0001
	Clay	-1.3879	1.4720	0.3458
	H_3O^+	-5.5117	2.2093	0.0126
	SOM:Clay	-11.5867	2.7725	< 0.0001
	SOM:H_3O^+	-6.2995	3.0806	0.0409
	Clay:H_3O^+	-11.2259	3.7693	0.0029
	SOM:Clay:H_3O^+	0.9964	6.9334	0.8857
Moderate	SOM	-3.9536	0.7339	< 0.0001
	Clay	-0.7116	0.7060	0.3135
	H_3O^+	-4.3376	1.1342	0.0001
	SOM:Clay	-4.8935	1.3467	0.0003
	SOM:H_3O^+	0.9792	1.3513	0.4687
	Clay:H_3O^+	-7.4802	1.7291	< 0.0001
	SOM:Clay:H_3O^+	4.8528	3.6119	0.1791
Normal	SOM	-1.9504	1.3170	0.1386
	Clay	-0.4230	1.3886	0.7607
	H_3O^+	-8.7301	2.3950	0.0003
	SOM:Clay	3.4948	2.6662	0.1899
	SOM:H_3O^+	-0.9849	3.5804	0.7833
	Clay:H_3O^+	-7.4969	4.2564	0.0782
	SOM:Clay:H_3O^+	6.7002	8.1071	0.4086

Supplementary Table 3: Results of mixed effects Tobit regression models across multiple levels of drought severity relating loss cost to soil organic matter (SOM, %), clay (%), and H_3O^+ concentration (mol). Coefficients for terms are standardized. Sample numbers for each drought level are as follows: Very severe, n=410; Severe, n= 1537; Moderate, n=3998; Normal, n=6431.

Drought	Model	Response	Predictor	Coefficient	Std. error	*p*
Drought	Partial	Yield	CEC	0.2007	0.0102	< 0.0001
		Yield	AWC	0.3253	0.0222	< 0.0001
		Yield	SOM	0.3979	0.0383	< 0.0001
		AWC	SOM	0.0557	0.0573	0.1332
		CEC	SOM	0.5848	0.1247	< 0.0001
	Full	Yield	CEC	0.4727	0.0087	< 0.0001
		Yield	AWC	0.2324	0.0233	< 0.0001
		AWC	SOM	0.0557	0.0573	0.1332
		CEC	SOM	0.5848	0.1247	< 0.0001
Normal	Partial	Yield	CEC	0.2044	0.0097	< 0.0001
		Yield	AWC	0.4085	0.0210	< 0.0001
		Yield	SOM	0.2337	0.0363	< 0.0001
		AWC	SOM	0.0557	0.0573	0.1332
		CEC	SOM	0.5848	0.1247	< 0.0001
	Full	Yield	CEC	0.3642	0.0078	< 0.0001
		Yield	AWC	0.3540	0.0208	< 0.0001
		AWC	SOM	0.0557	0.0573	0.1332
		CEC	SOM	0.5848	0.1247	< 0.0001

Supplementary Table 4: Results of partially and fully mediated piecewise structural equation models under drought and non-drought conditions. SEMs relate the effect of soil organic matter (SOM, %) to yield (Mg ha^{-1}) via its impacts on available water capacity (AWC, %) and cation exchange capacity (CEC, meq 100 g $soil^{-1}$). Coefficients for terms are standardized. Models for the drought category include observations at all levels of drought severity (i.e. very severe, severe, moderate). To more accurately model typical soil types, observations with a CEC value greater than 31.92 meq 100 g $soil^{-1}$ were excluded. Sample numbers for each drought level are as follows: Drought, n=5740; Normal, n=6307. Model inputs were based on the mean value for each term at the county level.

Term	**Coefficient**	**Std error**	***p***
SPEI	1.1507	0.0325	< 0.0001
SOM	0.8316	0.0404	< 0.0001
H_3O^+	-0.4497	0.0613	< 0.0001
Clay	0.1992	0.0440	< 0.0001
SPEI:SOM	-0.6661	0.0659	< 0.0001
SPEI:H_3O^+	-0.1508	0.1070	0.1588
SOM:H_3O^+	-0.3258	0.0852	0.0001
SPEI:Clay	-0.1686	0.0647	0.0092
SOM:Clay	0.6487	0.0745	< 0.0001
H_3O^+:Clay	-0.2630	0.1029	0.0106
SPEI:SOM:H_3O^+	0.2498	0.1599	0.1182
SPEI:SOM:Clay	0.2117	0.1276	0.0972
SPEI:H_3O^+:Clay	-0.1153	0.1934	0.5512
SOM:H_3O^+:Clay	-0.4649	0.1881	0.0134
SPEI:SOM:H_3O^+:Clay	0.5682	0.3571	0.1115

Supplementary Table 5: Results of linear mixed effects models relating yield (Mg ha^{-1}) to Standardized Precipitation Evapotranspiration Index (SPEI), soil organic matter (SOM, %), clay (%), and H_3O^+ concentration (mol). Coefficients for terms are standardized. Sample number was n=12,376.

Term	**Coefficient**	**Std error**	***p***
DSCI	-1.2245	0.0388	< 0.0001
SOM	0.8423	0.0407	< 0.0001
H_3O^+	-0.4235	0.0622	< 0.0001
Clay	0.2474	0.0443	< 0.0001
DSCI:SOM	0.4675	0.0760	< 0.0001
DSCI:H_3O^+	0.3370	0.1302	0.0096
SOM:H_3O^+	-0.1634	0.0857	0.0565
DSCI:Clay	0.2943	0.0663	< 0.0001
SOM:Clay	0.6290	0.0748	< 0.0001
H_3O^+:Clay	-0.1733	0.1029	0.0923
DSCI:SOM:H_3O^+	-0.2908	0.1942	0.1342
DSCI:SOM:Clay	-0.5208	0.1282	0.0001
DSCI:H_3O^+:Clay	0.1289	0.2005	0.5204
SOM:H_3O^+:Clay	-0.1827	0.1881	0.3316
DSCI:SOM:H_3O^+:Clay	-0.3559	0.3689	0.3347

Supplementary Table 6: *Supplementary Table 5*: Results of linear mixed effects models relating yield (Mg ha^{-1}) to Drought Severity Classification Index (DSCI), soil organic matter (SOM, %), clay (%), and H_3O^+ concentration (mol). Coefficients for terms are standardized. Sample number was n=12,376.

Drought	**Model**	**AIC**	**BIC**	**Fisher C**	***p***
Drought	Full	343.85	389.75	323.85	---
	vs. Partial	219.4	269.89	197.4	< 0.01
Normal	Full	260.89	306.8	240.89	---
	vs. Partial	219.4	269.89	197.4	< 0.01

Supplementary Table 7: Results of a fit comparison and analysis of variance for partially and fully mediated piecewise structural equation models under drought and non-drought conditions. SEMs relate the effect of soil organic matter (SOM, %) to yield (Mg ha^{-1}) via its impacts on available water capacity (AWC, %) and cation exchange capacity (CEC, meq 100 g $soil^{-1}$). Models for the drought category include observations at all levels of drought severity (i.e. very severe, severe, moderate). Sample numbers for each drought level are as follows: Drought, n=5945; Normal, n=6431. Model inputs were based on the mean value for each term at the county level.

Drought	Model	Response	Predictor	Coefficient	Std. error	*p*
Drought	Partial	Yield	CEC	-0.0752	0.0074	0.0155
		Yield	AWC	0.4465	0.0206	< 0.0001
		Yield	SOM	0.5282	0.0353	< 0.0001
		AWC	SOM	0.0677	0.0575	0.0642
		CEC	SOM	0.4522	0.1604	< 0.0001
	Full	Yield	CEC	0.1693	0.0078	< 0.0001
		Yield	AWC	0.4231	0.0244	< 0.0001
		AWC	SOM	0.0677	0.0575	0.0642
		CEC	SOM	0.4522	0.1604	< 0.0001
Normal	Partial	Yield	CEC	-0.1241	0.0071	0.0001
		Yield	AWC	0.5374	0.0198	< 0.0001
		Yield	SOM	0.3910	0.0340	< 0.0001
		AWC	SOM	0.0677	0.0575	0.0642
		CEC	SOM	0.4522	0.1604	< 0.0001
	Full	Yield	CEC	0.0569	0.0070	0.0745
		Yield	AWC	0.5201	0.0217	< 0.0001
		AWC	SOM	0.0677	0.0575	0.0642
		CEC	SOM	0.4522	0.1604	< 0.0001

Supplementary Table 8: Results of partially and fully mediated piecewise structural equation models under drought and non-drought conditions. SEMs relate the effect of soil organic matter (SOM, %) to yield (Mg ha^{-1}) via its impacts on available water capacity (AWC, %) and cation exchange capacity (CEC, meq 100 g $soil^{-1}$). Coefficients for terms are standardized. Models for the drought category include observations at all levels of drought severity (i.e. very severe, severe, moderate). Sample numbers for each drought level are as follows: Drought, n=5945; Normal, n=6431. Model inputs were based on the mean value for each term at the county level.

References

1. Green, T. R., Kipka, H., David, O. & McMaster, G. S. Where is the USA Corn Belt, and how is it changing? *Sci. Total Environ.* **618**, 1613–1618 (2018).
2. Rippey, B. R. The U.S. drought of 2012. *Weather Clim. Extrem.* **10**, 57–64 (2015).
3. Bates, D. *et al. lme4: Linear Mixed-Effects Models using 'Eigen' and S4*. R package v. 1.1-23. https://CRAN.R-project.org/package=lme4. (2020).