Extreme Weather Impacts on US Dairy Production: Evidence from 8 Million Cows

Ziyi Lin, Cornell University, zl349@cornell.edu
Ariel Ortiz-Bobea, Cornell University, ao332@cornell.edu
Christopher Wolf, Cornell University, caw364@cornell.edu
Kristan Reed, Cornell University, kfr3@cornell.edu

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

The Impact of Extreme Heat on Milk Production: Evidence from 8 Million Cows

Ziyi Lin[1]*, Kristan Reed[2], Christopher Wolf[1] and Ariel Ortiz-Bobea[3]

Abstract

We examine the effect of extreme heat on US milk production. We rely on a detailed panel dataset comprised of more than 150 million daily cow lactation records (2006-2016) from about 8 millions cows, that we link with gridded weather information. Our findings indicate that heat has a nonlinear effect on milk yield. Milk yields remain mostly insensitive to changes in temperature up to a threshold beyond which yields decrease precipitously. We also find that heat has a substantial negative effect on milk quality. Unlike yields, the detrimental of heat on fat and protein content are gradual, indicating that quality losses occur at moderate temperatures well below when temperature effects on yield manifest themselves. Additionally, we find considerable heterogeneity in the effect of heat across several dimensions, including farm size and location, cow breed and age.

JEL Codes: Q54, Q12, Q15.

Keywords: Extreme temperature, heat, dairy production, milk yields, adaptation, climate change.

[1] Charles H. Dyson School of Applied Economics and Management, Cornell University.

[2] Department of Animal Science, Cornell Unviersity.

[3] Jeb E. Brooks School of Public Policy, Cornell University.

* Corresponding author (Lin): zl349@cornell.edu.

1. Introduction

With the threat of long-term warming trends, heat stress in dairy production has become a serious problem in the US dairy industry. However, the impact on the milk component and the heterogeneity analysis remains unclear.

Currently, productivity-related studies mostly focus on milk yield loss. St-Pierre, Cobanov and Schnitkey (2003) estimate that heat stress cause ~ $900 million loss in the dairy industry. Mauger *et al.* (2015) predicts that with the global warming trend, the total economic loss will increase to ~ $2 billion at the end of this century.

However, compared to milk yield, dairy products are more relevant to milk content. For example, the quality of cheese and butter depends on the fat percentage, while yogurt and ice cream have standards for protein content. But there are few articles discussing the climate impact on milk quality (i.e., fat and protein) with large-scale datasets. The existing study mainly set experiments and compares the milk quality of the trail and compare groups in different temperature chambers ((Cowley *et al.*, 2015); (Gao *et al.*, 2017)). But those experiment methods are often confined to small data or short periods. Conversely, our study used large-scale and daily-level milk component records to study the impact of extreme heat.

Additionally, to give an unbiased estimation of the climate impact, we need to take the role of individual and farm information into consideration. Berry and Cromie (2009) found that Jersey cattle tend to have better heat tolerance. Aguilar, Misztal and Tsuruta (2009) found that the younger cattle will less influence by heat exposure. The larger farms also tend to have more advanced management strategies to mitigate the adverse impacts; since the larger farms may have access to imply better cooling systems (e.g., sprinklers, misters, and fans), housing infrastructure (e.g. barns and shading), and nutrition management (Bishop-Williams *et al*., 2015).

However, the current research doesn't yet have clear evidence discussing micro-level factors. Most existing studies rely on aggregate data such as state-level surveys (Gisbert-Queral et al., 2021). Such aggregate analyses conceal important heterogeneity that may help better understand adaptation mechanisms. Few studies analyze farm-level data (e.g. Key and Sneeringer, 2014), but those are generally confined to very limited annual farm surveys and don't provide temporally disaggregated data necessary to better understand the effects of weather shocks occurring over brief periods of time.

Conversely, this study relies on the granular dataset, which is at a fine temporal scale and the cow level for a large population and across regions, to understand the role of the heterogeneity differences. By estimating milk production and shedding light on farm management, this paper aims to provide insights for improving the climate resilience of US dairy.

In this paper, we first improve the estimation of the effect of heat stress on milk production, allowing for non-linearities. We then seek to better understand the underlying heterogeneity of this response function. Specifically, we explore how factors such as cow breed, management, and acclimation modulate the effect of heat on milk yield. We ultimately seek to quantify the potential impacts of extreme weather on US dairy production, allowing for various types of adaptation.

The paper is organized as follows: Section 2 is an overview and descriptive analysis of the data. Section 3 presents the methodology of the models used in this study. Section 4 provides the empirical result and analysis. Section 5 is a summary. The additional supportive materials, including graphs and tables, are reported in the Appendix.

2. Data

2.1 Data sources

2.1.1 Dairy data: Cattle Test-day (TD) Lactation Records

We use cattle lactation records from the US Council on Dairy Cattle Breeding (CDCB) (https://www.uscdcb.com). These lactation records are typically used to inform cattle breeding decisions. Every year, the CDCB database tracks hundreds of thousands of lactating cows across the US. For each cow lactation, the dataset provides detailed information in the form of 10 to 20 daily snapshots or test days (TD). Each record also provides information about the entire lactation.

The summary of each lactation includes the information of the individual, the lactation, and the herd. The individual information includes ID number, breed, sex, and age. The lactation information includes the length, the starting date, the actual milk yield, and the standardized 305-day milk yield. The herd information includes herd ID numbers and county-level geographical locations.

Besides records for the whole lactation, CDCB also collected individual milk test results on a monthly basis. During the typical 10-month lactation period, each lactation records have 10-20 test day records. The test-day (TD) milk yield record provides a daily snapshot of the milk production of a specific cow on a given date. It offers progress (days in milk) and milk production value (yield, protein, and fat percentages). Note that the farmers suspend the TDs record if they find the cattle are not performing well. To avoid selection bias, we only keep the records with at least 10 TDs records (Li *et al.*, 2022).

As a result, our dataset contains 150 million test day (TD) records from 12.8 million lactations from the year 2006 to the year 2016. Those records come from a total of 8 million individuals located in 47 states, 21,056 herds. For reference, the total dairy cow population in the United States is approximately 9 million. 75% of the data came from 5 states: Wisconsin (25%), Pennsylvania (13.3%), New York (12.9%), Minnesota (9.18%), and Michigan (6.92%). The dataset includes 22 distinct breeds, with Holstein cows comprising 91.8% of the sample population.

2.1.2 Weather data: Temperature Humidity Index (THI)

To describe the heat stress on dairy cattle, we use the temperature humidity index (THI). We rely on Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset to obtain the weather outcomes. The PRISM dataset is developed at Oregon State University, which offers standardized gridded climate data with daily temperature and humidity information at a 4km resolution (2.5 $\times$ 2.5 $mile$ grid cell) across the contiguous US since 1981.

The temperature-humidity index (THI) is a widely-used measurement for heat stress in dairy research since the year 1960(Bianca, 1962). Heat stress is a multifaceted phenomenon that results from a complex interplay of various environmental factors such as temperature, relative humidity, solar radiation, air movement, and precipitation (Bohmanova, Misztal and Cole, 2007). Thus the combined measurement of multiple environment outcomes more precisely describe the cattle's feeling about the environment than the sole effect (Council, 1971). Also, the abundance of data for temperature and humidity makes THI an ideal indicator for the empirical estimation (Correa-Calderon *et al.*, 2004).

Over the years of development, there are multiple formulas exist for computing the THI using varying weights and measurements of temperature and humidity. Given that 95% of our dataset corresponds to Holstein, we have selected the formula that best characterizes heat stress in this breed, as described in Key and Sneeringer (2014) and St-Pierre, Cobanov and Schnitkey (2003). This methodology was initially introduced in 1985 (Yousef and others, 1985) through Equation 1.

$$THI = (\text{dry bulb temperature } ^{\circ}C) + (0.36 \times \text{dew point temperature } ^{\circ}C) + 41.2 \quad (1)$$

where the dry bulb temperature measures the temperature, it is approximated by ambient air temperature; dew point temperature, a constant through one day, measures the humidity. It is the temperature at which the air must be cooled to achieve saturation, i.e., the temperature at which the relative humidity is 100%. Generally, dairy cows experience heat stress when the THI calculated by this formula is above 72 °F.

2.2 Key variables and summary tables

2.2.1 Time trend in milk yield

The assumption of our paper following existing research is that the cow will produce less milk when experiencing heat stress. In order to study the climate impact, we need to control the time trend. We found that the milk yield follows two kinds of time trends: a cyclical mode during each lactation process and an increasing trend over the years.

In Figure 1, we plot the lactation curves. During each lactation, the milk yield follows a first increasing then decreasing trend. Each lactation has an average length of 305 days, with an ascending phase leading up to the peak at around 200 days and a descending phase following up the peak. In the appendix, we show the formula used Khandekar model to prove this result with our data.

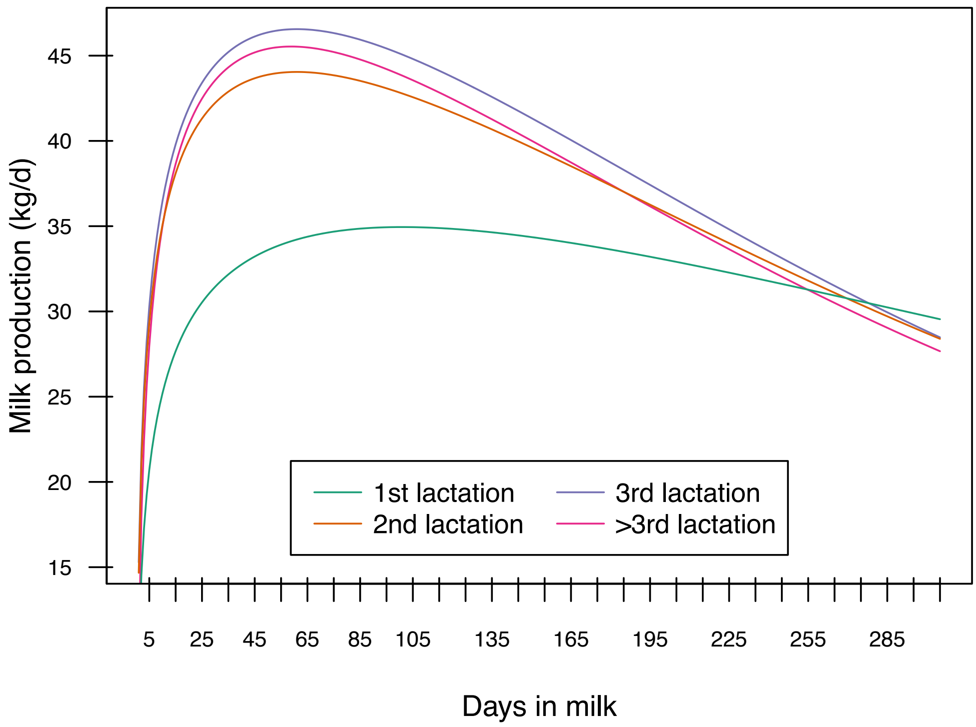


Figure 1: Lactation curves for each lactation

In Figure 2, we plot the annual mean value of the daily milk yield record using the total lactation record, the test-day record, and the USDA record. The three lines share a similar increasing time trend after the first year. This may be due to the farmers applying advanced agricultural technology and undertaking management actions to cope with climate risks such as fans, sprinklers, ventilation systems, and barns (Dairy, 2014).

In the year 2006 (the first year of our dataset), the total lactation record and USDA record have the lowest value, while the test day records have the highest value. This is because our data begins from the year 2006 and the test day records are mostly from the beginning of each lactation. The average days in milk of TD records are 200 days, while in the year 2006, the average days in milk are only 110 days. According to the lactation curves, the milk yield from the early stage of each lactation is higher than those from the later stage. Thus, we get an increasing milk yield over years and the abnormal increasing in the first year is due to the nature of our dataset.

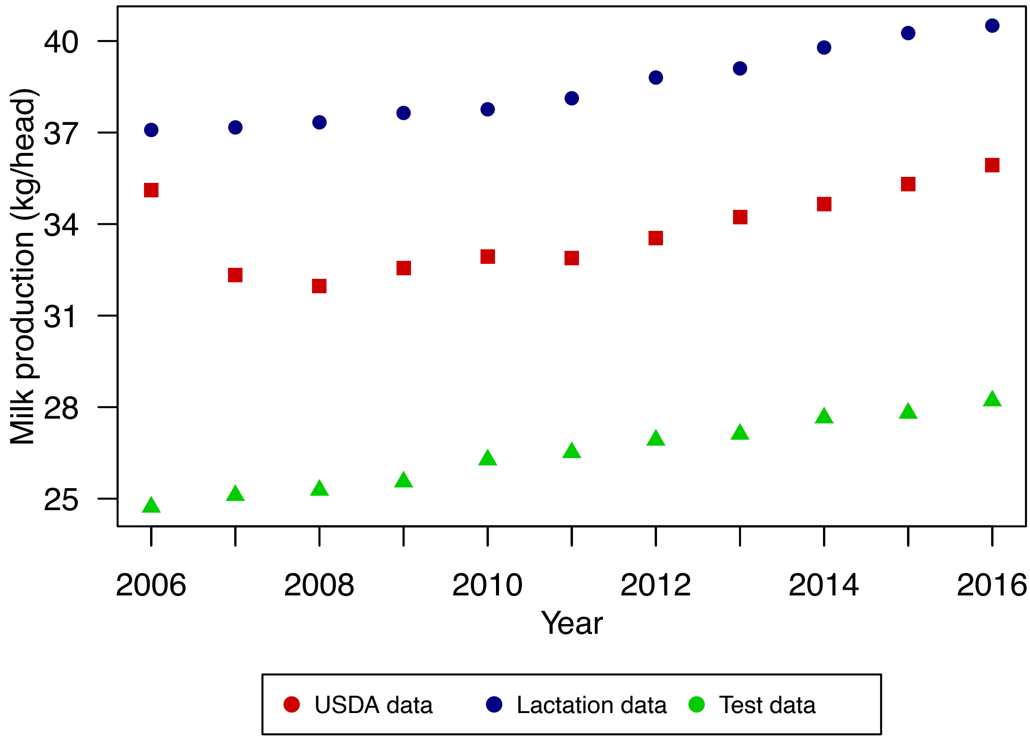


Figure 2. Milk yield trend compared between USDA and CDCB dataset

2.2.2 Herd size

In our study, we include a total of 27,000 herds. To estimate the ability to mitigate extreme heat, we take the size of a herd as the observable variable for management power.

According to Figure 3, the herd size has continuous increase over time. And the larger herds have higher milk productivity. Figure 4 is a Sankey diagram that shows the flow of herd size between different sets. The purple bars representing the smallest herds (with sizes less than 30) exhibit the highest values in the first three years, followed by a gradual decline over time. In contrast, the yellow bars representing the largest herds (with more than 500 heads) show an increasing trend over the years. Also, all levels of the herds have a tendency to flow into larger herd levels. For example, the herds that fall into the level of 30 – 50 head more likely to flow into the 50 - 100 head or higher level, while only a few flow into the < 30 head level.

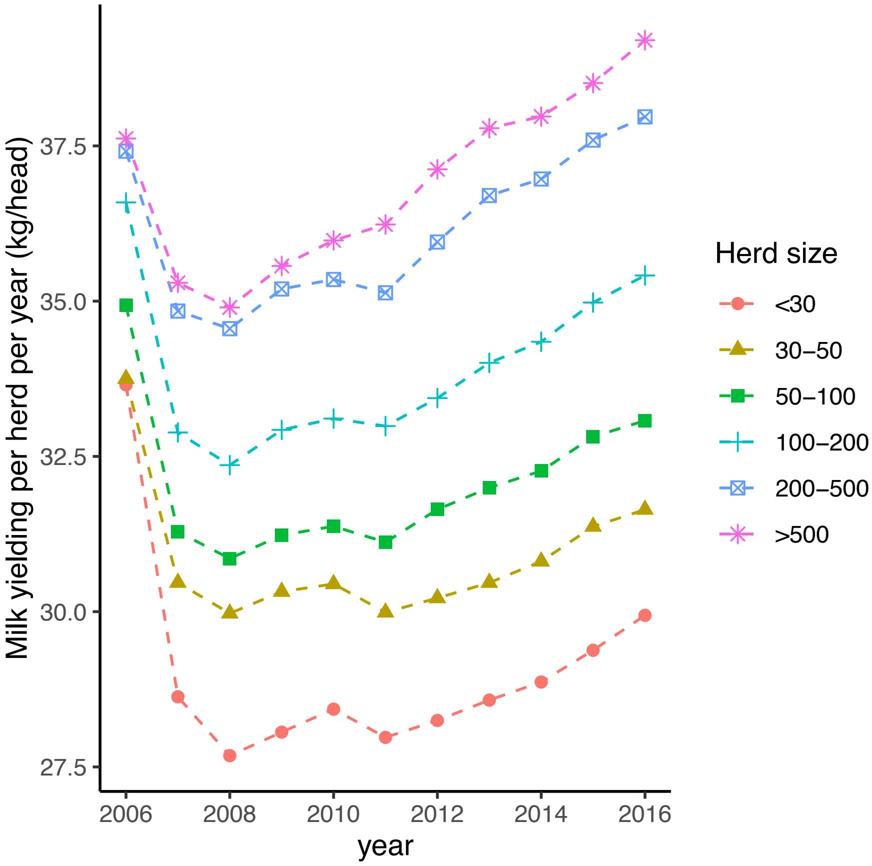


Figure 3. Milk yield in different herd size

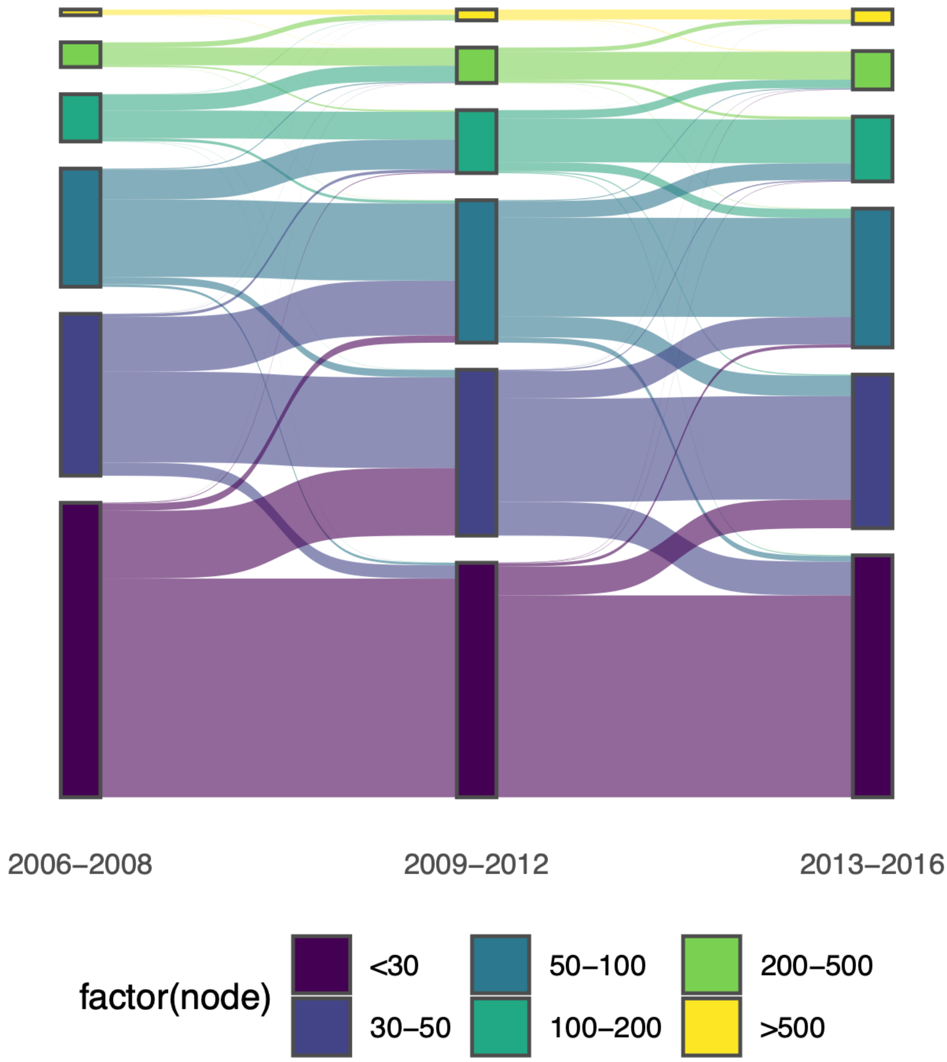


Figure 4. Sankey plot for herd size

2.2.3 THI bins

Our overall approach to modeling the effects of heat stress on milk yield exploits presumably random fluctuations in weather conditions in the days preceding a milk test. In other words, we're essentially capturing the effect of unexpectedly hot or cold episodes preceding a pre-scheduled test day in milk data.

When considering the heat stress in a longer time period, a delayed response of test-day milk yield is expected after the onset of the heat stress. A core measure in existing research regarding US dairy is using the average value of THI in the previous week of each test day (Hill and Wall, 2015).

However, the weather parameter summarized across a week's timescale from the test day leads to three biases: 1) The temporal averaging conceals the exposure to the extreme value, which biases our analysis of heat stress on milk production; 2) The spatial averaging includes irrelevant units for dairy production; 3) The cool night temperatures may counteract the effects of high values during the day and offer the relief of heat stress (Carabaño *et al.*, 2014).

To better account for the effect of recent heat stress, our strategy is to portray the distribution of THI over the week preceding each test day. That is, we build a measure of time spent at various levels of THI over the week before each observation. To do this from daily minimum and maximum temperatures, we derive a THI "time path" based on a double sine curve interpolated every 15 minutes for each spatial grid cell. We then recover the amount of time spent each week at various THI intervals ranging from -10 to 95 and aggregate to the county level.

Step 1: Daily THI calculation with PRISM data

Within a period of 24 hours, the temperature-humidity index was assumed to follow a perfect sine function (St-Pierre, Cobanov and Schnitkey, 2003). This naturally requires the data for the daily minimum temperature, maximum temperature, and dew point temperature.

Since dew point temperature is a constant value through a day, minimum THI (THI_{min}) is calculated using the minimum temperature, whereas maximum THI (THI_{max}) is calculated using the maximum temperature (the detail is in Equation 2) (Key and Sneeringer, 2014). We then create a sine function passing through the minimum and maximum value of THI with Equation 2.

$$THI = \frac{THI_{max} - THI_{min}}{2} sin\left(\frac{2\pi}{24}x - \frac{\pi}{2}\right) + \frac{THI_{max} + THI_{min}}{2} \quad (2)$$

where,

$$THI_{max} = T_{\max} + 0.36 \times t_d + 41.2$$
$$THI_{\min} = T_{\min} + 0.36 \times t_d + 41.2$$

Step 2: Calculate time spent in each THI interval in the previous week

Based on the daily THI path, the next step is to create a consecutive path for the previous seven days and count the exposure bins as how much time each gridded cell is exposed to each interval in the previous week. This process is simulated by a double-sine function with the 15-minute' interpolation. Take the one-day cycle as an example, the first quarter of the phase is using the minimum value from the previous day and the maximum value from today; the second and third quarters are using the minimum and maximum value from today; and the last quarter is using the minimum value from today and maximum value from tomorrow. The time path for the previous seven days, which contains 7 cycles, can be found in Figure 5.

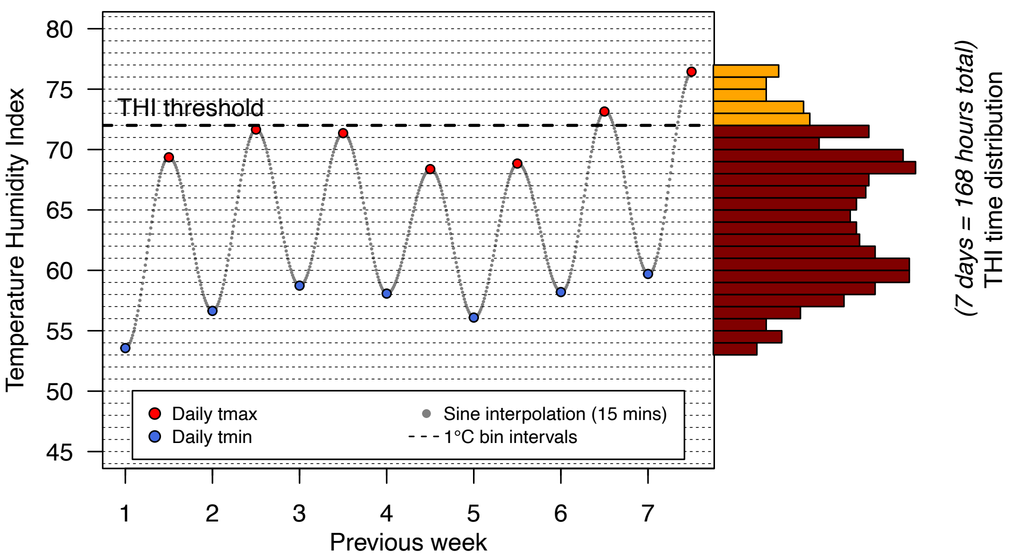


Figure 5. Weekly THI distribution with a randomly selected pixel

Figure 5 shows the THI distribution of a randomly selected pixel. Each point in the graph represents a 15-minute interpolation, which means the cow spent 15 minutes in the corresponding THI interval. The time distribution of THI intervals is shown on the right of the graph. 72 is the threshold for milk production. The bars in yellow color represent the time that cows are spent beyond the threshold. Similar to the degree days in crop science, the time spent beyond the threshold is defined as THI degree days.

Step 3: Aggregate grid data to county-level data

Because we ignore the precise location of cows within a county, based on the calculation of each grid cell, we aggregate the bins to the county level for analysis. To eliminate the cells with no agricultural activity, we perform this aggregation based on the National Land Cover Database (NLCD), by assigning aggregation weights proportional to agricultural land cover.

This process yields a panel dataset with daily county-level observations corresponding to the THI exposure "bins" over the previous week.

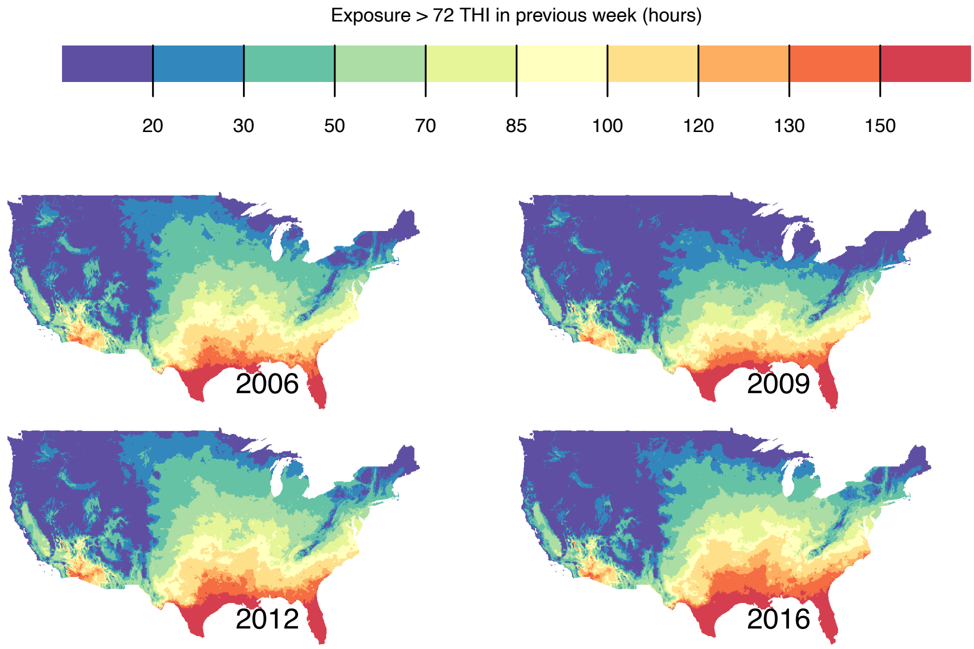


Fig 6. Annual exposure to THI above 72 across the US for specific years

In nation wide, there are heterogeneity differences in the THI exposure time. We focus on the number of THI degree days for analysis. In Figure 6, we take the year 2006, 2009, 2012, and 2016 for example, and show the hours spent in THI above 72 in the previous week. Overall speaking, southeastern regions (e.g., Florida, Georgia, Alabama, and Texas) have the most exposure to high THI while northern regions (e.g. New York, Wisconsin, and Minnesota) are mostly exposed to the low THI environment.

In Table 1, we make the summary table for the exposure time in each 10-bin level. We analyze summer months and all year separately. The exposure time has the most distribution in the level 65 - 75 THI bins for both two settings. However, the summer months have a distribution with a smaller standard deviation.

Table 1 Descriptive statistics for the exposure time (hours) in each THI intervals

Levels	Full year				Summer			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
< -5	0	0.1	0	34.40	0	0	0	0
(-5, 5]	0.1	1.7	0	108.0	0	0	0	0
(5, 15]	0.9	6.1	0	148.0	0	0.2	0	48.90
(15, 25]	3.5	13.0	0	145.1	0.1	1.6	0	101.7
(25, 35]	11.4	24.4	0	167.6	2.4	9.8	0	151.8
(35, 45]	27.1	35.6	0	168.0	14.7	25.9	0	167.6
(45, 55]	33.2	33.6	0	168.0	30.5	31.5	0	167.8
(55, 65]	37.3	33.6	0	168.0	42.6	34.6	0	167.9
(65, 75]	39.1	41	0	168.0	73.0	32.6	0	168.0
(75, 85]	15.2	28.6	0	168.0	34.4	35.4	0	168.0
>85	0.3	2.5	0	115.6	30.8	15.6	1	115.6

3. Regression models

We assume that the temperature-humidity index has a cumulative information and nonlinear marginal effects on milk yielding. Following (Gisbert-Queral et al., 2021) and (Schlenker and Roberts, 2009), the regression model is defined as Equation 3.

$$log(M_{it}) = \int_{\underline{THI}}^{\overline{THI}} g(THI)\phi_{it}(THI)d(THI) + z_{it}\delta_{it} + c_i + \epsilon_{it} \tag{3}$$

where $log(M_{it})$ is the logarithm of the milk yielding of individual i in test day t. g(THI) is the marginal effect of THI, it depends nonlinearity on THI. $\phi_{it}(THI)$ is the distribution of THI for cattle i and in the previous week of day t. δ_{it} analyzed the cow's attribute. It includes four control factors. The management and technology factors are denoted by herd id (h_{it}) and year (y_{it}). The milking progress is controlled by lactation times (l_{it}) and days in milk (dim_{it}). A time-invariant individual fixed effect c_i is to control for heterogeneity such as breed and genes. And the error term is clustered at herd by year level. Figure 16 in the Appendix reports that the setting of fixed effect and clustering form aligned well with our data. The full model is presented in the form in Equation 4.

$$log(M_{it}) = \int_{\underline{THI}}^{\overline{THI}} g(THI)\phi_{it}(THI)d(THI) + h_{it} + y_{it} + l_{it} + dim_{it} + c_i + \epsilon_{it} \quad (4)$$

This continuous representation is theoretical and is not tractable for estimation. To solve that, we estimate the integral with 1-unit THI interval in Equation 5. Note that to avoid the noisy estimation by having too little exposure at the tails of the THI distribution, the paper bottom and top the exposure data and the bins ranged from 5 to 85.

$$log(M_{it}) = \sum_{THI=5}^{85} g(THI)\,[\phi_{it}(THI+1) - \phi_{it}(THI)] + z_{it}\delta_{it} + c_i + \epsilon_{it} \quad (5)$$

However, estimating the effect of more than 100 individual THI bins on milk yield would lead to collinearity. If we assume that the effect of neighboring bins has similar effects, we can reduce the dimensionality of the estimation problem.

In order to accomplish this task, a J-th Chebyshev polynomial ($T_{j()}$) was utilized to fit exposure data that had been aggregated over intervals of 1 °C, with the expression $g(THI) = \sum_{j=1}^{J} \gamma_j T_j THI$. The selection of polynomial degrees was conducted by estimating their increasing order until a stable relationship was observed, ultimately leading to the choice of degree 9 for the estimation. The analysis results are presented in the 15, and the complete function is displayed in Equation 6.

$$\begin{aligned} log(M_{it}) &= \sum_{THI=5}^{85} \sum_{j=1}^{9} \gamma_j T_j [\phi_{it}(THI+1) - \phi_{it}(THI)] + z_{it}\delta_{it} + c_i + \epsilon_{it} \\ &= \sum_{j=1}^{9} \gamma_j \sum_{THI=5}^{85} T_j [\phi_{it}(THI+1) - \phi_{it}(THI)] + z_{it}\delta_{it} + c_i + \epsilon_{it} \quad (6) \\ &= \sum_{j=1}^{9} x_{it}^{9} + z_{it}\delta_{it} + c_i + \epsilon_{it} \end{aligned}$$

4. Empirical Result

4.1 Climate impact on milk yield

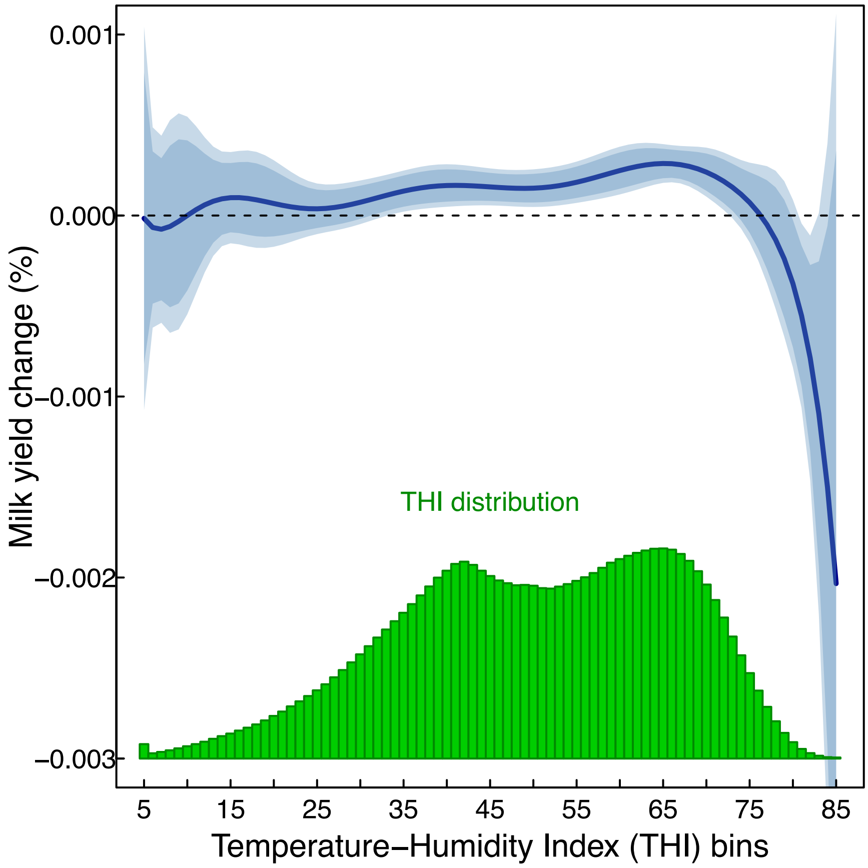


Figure 7. Nonlinear relation between THI and milk yield

Our results suggest a highly nonlinear relationship between THI and milk yield. The milk yield has an increasing trend with an optimal THI ranges around 72. Then the milk yield follows by a decline when THI exceeds 76. Figure 7 illustrates this relationship.

In Figure 7, The blue line shows a ninth-order Chebychev polynomial function of THI and the blue-colored bands around the response function correspond to 95% and 99% confidence intervals. The green histogram at the bottom represents the exposure to the temperature-humidity index for seven days before each test day. By construction, the distribution sums to 168 hours. The vertical axis represents the log of milk yield (kg per day) with the predicted yield normalized to zero. The horizontal axis is the temperature-humidity index.

To compare two points on the curve, a vertical difference of 0.01 represents a 1% difference in average yield for that day. For example, spending an additional hour in the THI interval at 84 – 85 rather than 44 – 45 in the previous week reduces hourly milk yield by 0.002 log points or 0.2%. To simplify the presentation of the numerical outcomes, we have translated the results from an hourly level to a daily level. Hence, an extension of time spent within the 84–85 THI range during the previous week, rather than the 44–45 range, results in decline in milk yield of 4.8% (24 × 0.2%).

The regression results for the polynomial model can be found in Appendix Table 2. The overall model exhibits statistical significance, as evidenced by the small p-values obtained from both the Wald and F tests, which are significantly lower than the conventional level of 0.05. All nine variables included in the polynomial function also exhibit statistical significance. Notably, the coefficients for the first and ninth degrees of freedom, ns1, and ns9, respectively, were significantly negative in both regressions. This suggests that cattle are susceptible to experiencing hot and cold stress when the temperature-humidity index (THI) reaches extreme values.

4.2 Climate Impact on milk component

The existing literature suggests that heat stress can significantly alter milk quality and components (Hammami *et al.*, 2013, 2013; Cowley *et al.*, 2015). Specifically, heat stress has been found to decrease the fat and protein content in milk, which can have significant implications for the dairy industry as a whole.

Our result aligns with the existing conclusion for the declining association between THI and milk components. As reported in Figure 8 and Figure 9, our results indicate that the response function curves for fat and protein percentage exhibit a continuous decreasing trend.

Compared to the milk yield, the heat stress had a more severe impact on fat and protein percentage. Specifically, when comparing spending one week in the THI interval of 84-85 versus 44-45, we observed that daily fat production decreased by 5.76% (with an hourly drop of 0.24%) and daily protein production decreased by 7.2% (with an hourly drop of 0.3%), while daily milk yield only decreased by 4.8%.

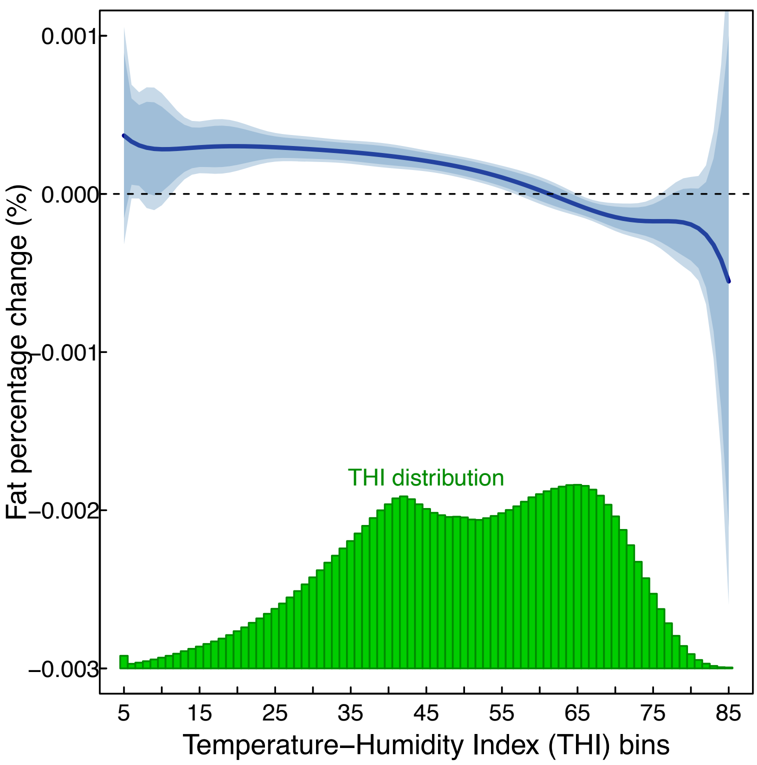


Figure 8. Nonlinear relation between THI and fat yield

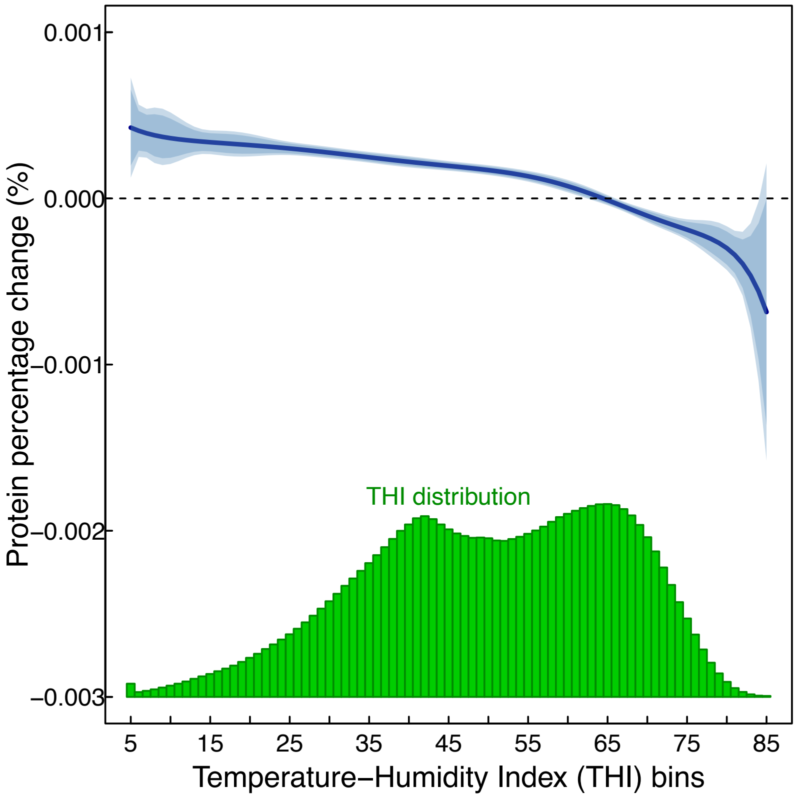


Figure 9. Nonlinear relation between THI and protein yield

The negative impacts of heat stress on fat and protein production also appear at an earlier stage than milk yield. Unlike the response function curve for milk yield, which exhibits an initial increase followed by a decrease, the response function curve for fat and protein shows a continuous decline. This indicates that while hot weather may initially boost milk yield, the milk quality is not improved by increasing THI.

To conclude, while previous research on the climate impact on the dairy industry has primarily focused on milk yields, our study highlights that milk components, specifically fat and protein, are also highly susceptible to extreme weather. Thus, policymakers and industry stakeholders should take into account the impact of heat stress on milk quality when considering strategies to mitigate the effects of extreme weather on the dairy industry.

4.3 Heterogeneity analysis

In this section, we do the heterogeneity analysis that estimated Equation 6 using the subsamples divided by different breeds, time frames, number of lactation times, places, and herd sizes.

The results are displayed in Figure 10. The horizontal axis depicts the percentage decline in daily milk yield for each 2-unit increase in THI. The y-axis on the right indicates the criteria used to divide the sub-samples, while the bars on the left y-axis represent the categories and their respective sample sizes.

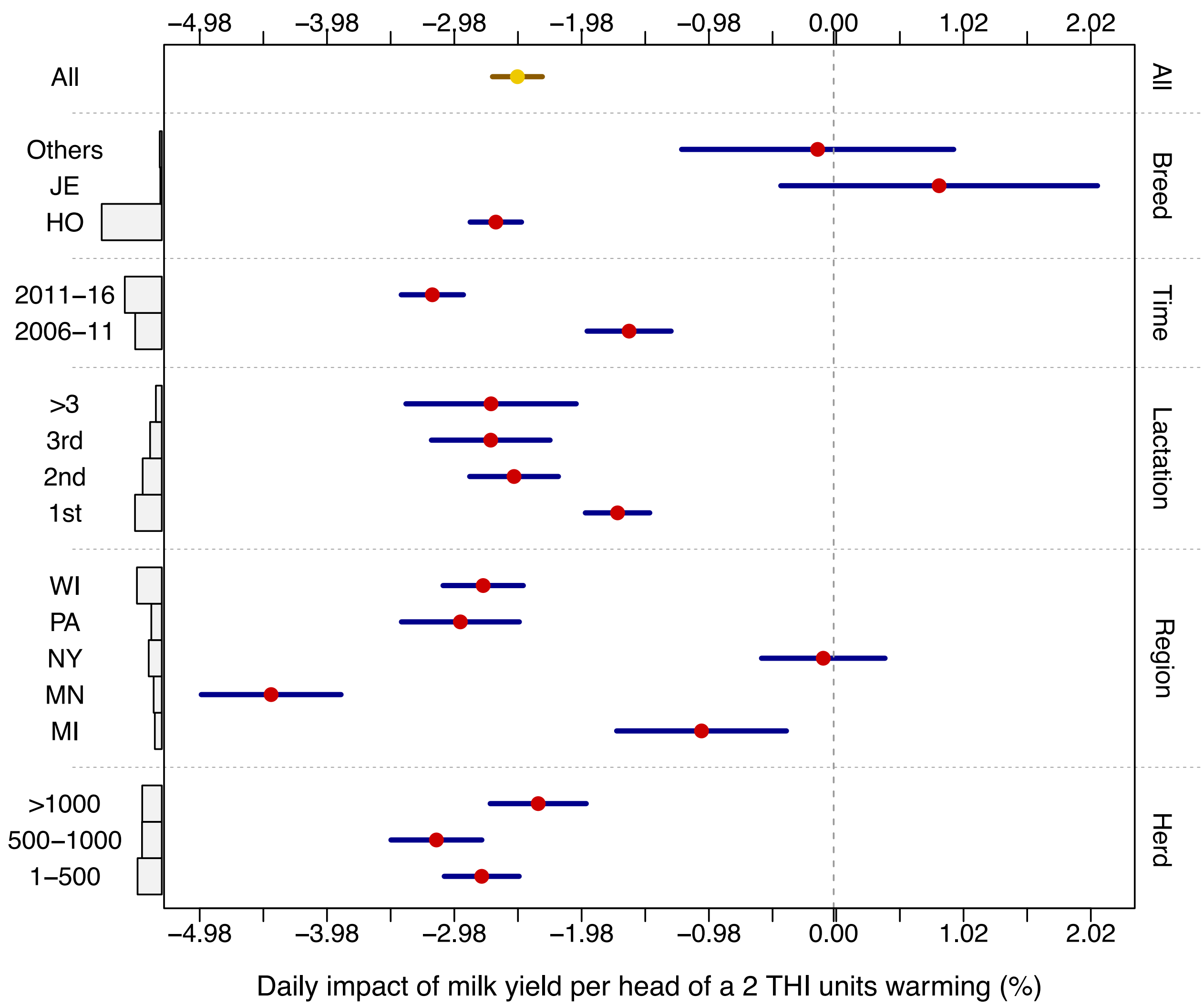


Figure 10. Heterogeneity analysis

4.3.1 Temporal factors

As evidenced by Figure 2, the dairy industry has witnessed a noteworthy escalation in milk production over the past decade. This substantial surge in milk production is widely attributed to effective genetic selection and advancements in farming practices (GROSU et al., 2013; Hazel, Heins and Hansen, 2020). If this assumption hold, than the cattle should have less sensitivity to heat stress as time goes by. To evaluate the heterogeneity, we compare the milk yield drop with two more units THI environment in different time periods.

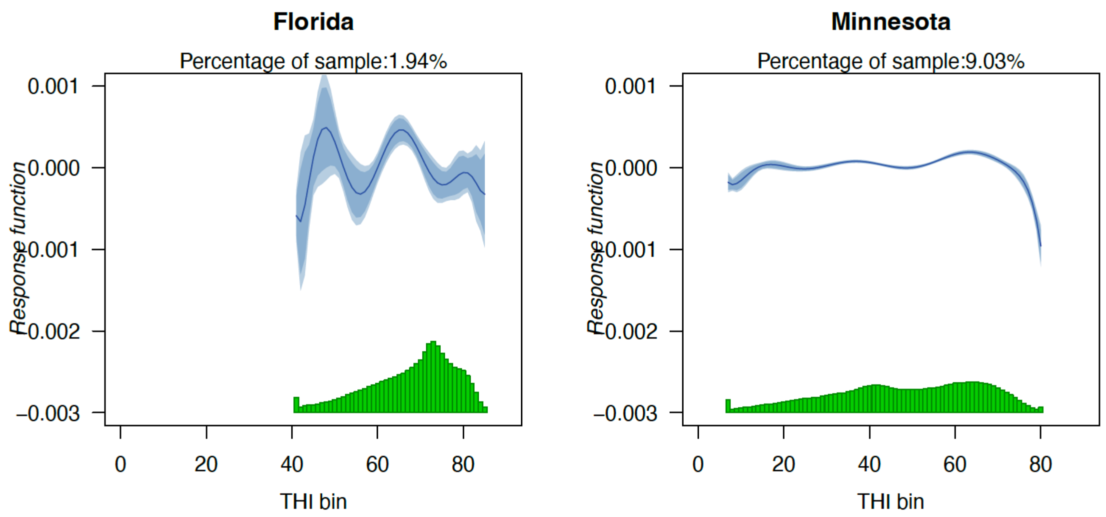
In the analysis, we divided our dataset into two distinct periods: the first half spanning from the year 2006 to the year 2011 and the second half spanning from the year 2011 to the year 2016. Our analysis reveals a significant differential in performance between the two periods. The latter half (i.e., the year 2011-2016) exhibits a superior capacity to withstand hot weather conditions in comparison to the former half (i.e., the year 2006-2011). Specifically, we observe a 2.45% decrease in daily milk yield for the latter half, whereas the former half experiences a more substantial drop of 4%. In the Appendix, we conducted the Chow test to ascertain the structural break in the relationship between the first and second halves of the sample.

Our findings offer a mechanism for explaining the recent increase in milk yield, indicating that the improvement in farming practices has led to an overall increase in milk yield by enhancing the ability to cope with hot weather and reducing sensitivity to THI.

4.3.2 Spatial factors

To assess the spatial heterogeneity, we estimated the response functions for each state. To avoid the noisy tail in THI bins distribution, we adjust the support for THI bins for each state by cumulating upward or bottom down.

In the initial calculation, we computed the total time spent in the preceding 7 days for each test day. Consequently, the cumulative time spent in the total bins should amount to 168 hours (7*24). Then, we cumulate upward or bottom down the bins that have accumulative times of less than 1 hour or exceeds 167 hours. Figure 11 displays a comparison between the hottest place, Florida, and the coldest place, Minnesota. As evident from the graph, Florida's THI bins are in the range of 40-85, whereas Minnesota has THI support ranging from 5-80.



(a) Response function in Florida (b) Response function in Minnesota

Figure 11. Response functions compared between the hottest and coldest states

Furthermore, we conducted a split-sample analysis of five key states - Wisconsin, Pennsylvania, New York, Minnesota, and Miami. Minnesota experience the most significant adverse effect. When exposed to 2-more units THI conditions, the dailu milk yield in Minnesota will drop approximately 3.48%. In contrast, New York demonstrated the most resilient performance. There is virtually no adverse impact for 2 more unit THI hotter environment in New York. The difference in performance may be attributed to differences in technical support. According to USDA, New York state has a more developed dairy industry than all other states.

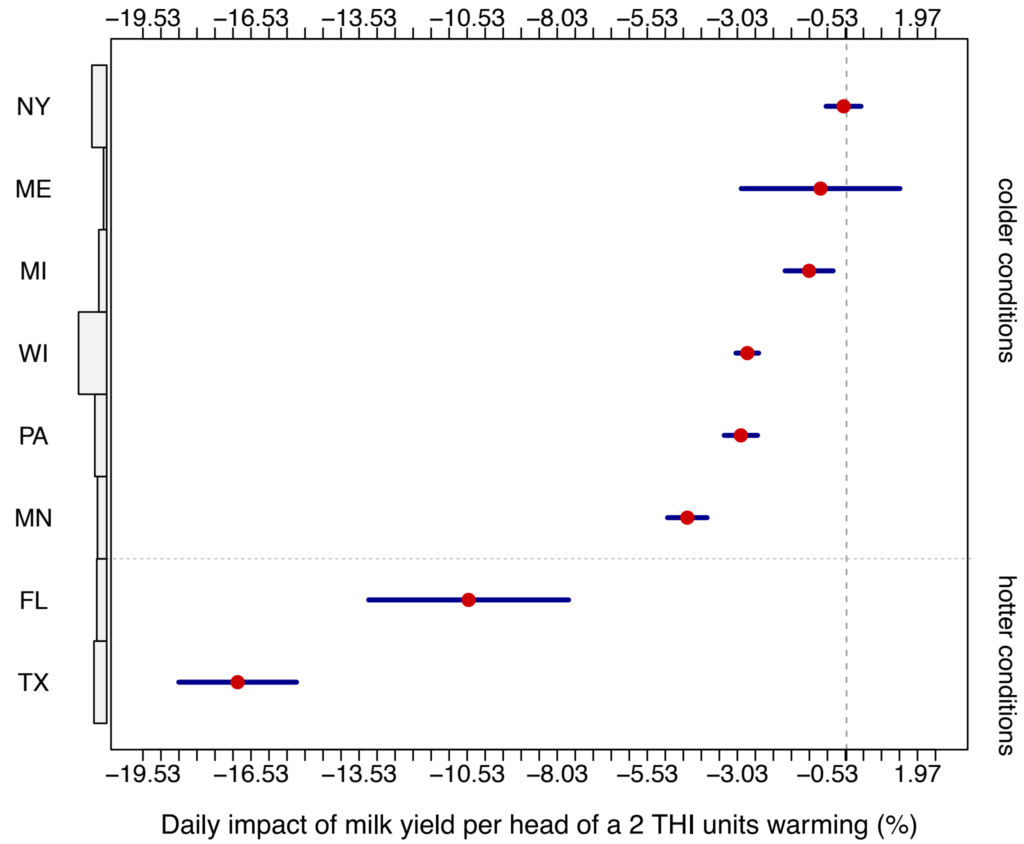


Figure 12. Heterogeneity analysis across hot and cold states

Another crucial factor in analyzing spatial heterogeneity is the local mean temperature. To test this assumption, we added Florida and Texas in our analysis. Those two states have a higher baseline temperature conditions than the five main states in the above analysis. The results are reported in Figure 12. The findings demonstrate that Texas and Florida are more likely to experience greater declines in milk production in environments with 2 more THI units. Specifically, daily milk yield in

Texas and Florida is estimated to drop by 17% and 10%, respectively. By comparison, the daily milk yield drop in states with colder conditions did not exceed 5%.

Therefore, our analysis indicates the existence of spatial heterogeneity between states. The states with higher baseline temperatures are likely to experience a more severe milk yield drop than cooler states.

4.3.3 **Management processes**

a. Breeds selection

The existing literature supports the superior thermal resilience of Jersey cattle relative to other breeds (Berry and Cromie, 2009). Our study corroborates this finding.

As evidenced by Figure 10, a 2-unit increase in the Temperature Humidity Index (THI) results in a mere 2.48% decrease in the daily milk yield for the overall sample. Since Holstein consists of 95% of our dataset, there is a negligible difference between Holstein cattle production and the overall sample. And the change in milk yield for all other breeds, apart from Jersey and Holstein, is close to 0 when exposed to a 2-unit increase in THI. However, Jersey cattle were even able to demonstrate an improvement of 1% in milk yield under the same conditions. These results provide additional evidence that although Holstein is the dominant breed in the US dairy industry, it may not perform optimally in a global warming environment compared to Jersey.

b. Lactation times

Lactation times (i.e. parity) in the dairy industry means the number of different times a female has had offspring or lactation. It has been widely anticipated that the initial lactation of cattle exhibits greater resilience to heat stress than subsequent lactation periods. One plausible explanation for this phenomenon is that initial lactation is characterized by lower milk yield and a relatively flatter yield curve (as confirmed by Figure 1), resulting in reduced metabolic and heat production (Kadzere *et al.*, 2002).

To testify to this assumption, we compare the influence of a hotter environment on the first, second, third, and higher than third lactations. Our analysis in Figure 10 demonstrates that every 2-unit elevation in THI exposure per interval corresponds to an approximate reduction of 2.04% in daily milk production. This effect is observed across most lactation periods, including the second, third, and all subsequent lactations. However, the first lactation records only a 1.6% decline in daily milk yield, signifying greater resilience to hot and humid weather conditions compared to later lactation rounds.

c. Herd size

Our result shows that the larger herds perform better than the smaller herds in heat stress. Herds with more than 1000 heads showed only a 2% daily milk yield drop, while small herds experienced a 2.5% drop. The Chow test in the appendix reveals a significant structural difference between small and large herds.

On one hand, the larger dairy farms tend to take confinement operations while the small farms are still having grazing systems. Thus, the cattle from the small farms are more likely to directly exposed in outdoor conditions. On the other hand, the larger farms have more advanced environmental control practices including sprinklers, fans, and shading (Hayakawa et al., 2016; McDonald, Law and Mosheim, 2020).

Thus, small herds remain the most vulnerable entities in the market when it comes to hot weather. And the policymakers should prioritize support for small herds to help them cope with hot weather conditions.

5. Conclusion

In conclusion, this study contributes to the literature on the impact of heat stress on the dairy industry by providing a unique dataset. Our granular data allows for the understanding of the weather impacts on milk components and a heterogeneity discussion in different conditions.

Our findings first confirm that heat stress has a significant negative impact on milk production. There is a clear threshold at a temperature-humidity index of 76. This result is consistent with previous studies that use aggregate data. Besides the discussion of milk yield, our study also highlights the influence of extreme heat on milk quality. We find the extreme heat leads to a more severe impact on fat and protein percentages than milk yield.

Our study proved that the dairy industry are having lower sensitivity to extreme weather in recent years. However, the risk associated with extreme weather varies among US dairy farms with different locations and management processes. Firstly, farms in hotter places tend to experience more significant decreases in milk production. Additionally, farms with less experience and technical support are more vulnerable in hot weather. Thirdly, small herds with grazing systems are more affected by heat stress than large herds with confinement operations. Finally, the heterogeneity between farms also exists in breed selection and parity. Farms with more cattle in Jersey breed and in their first lactation will be less influenced by heat stress.

Further research could expand the analysis to include the climate impact on milk quality and extend the study globally. The further discussion may also focus on the unbalanced development between small and large herds or hot and cold places.

References

Aguilar, I., Misztal, I. and Tsuruta, S. (2009) 'Genetic components of heat stress for dairy cattle with multiple lactations', *Journal of Dairy Science*, 92(11), pp. 5702–5711. Available at: https://doi.org/10.3168/jds.2008-1928.

Berry, D. and Cromie, A. (2009) 'Associations between age at first calving and subsequent performance in Irish spring calving Holstein–Friesian dairy cows', *Livestock Science*, 123(1), pp. 44–54.

Bishop-Williams, K.E. *et al*. (2015) 'Heat stress related dairy cow mortality during heat waves and control periods in rural Southern Ontario from 2010–2012', *BMC Veterinary Research*, 11(1), p. 291. Available at: https://doi.org/10.1186/s12917-015-0607-2.

Bohmanova, J., Misztal, I. and Cole, J.B. (2007) 'Temperature-humidity indices as indicators of milk production losses due to heat stress', *Journal of dairy science*, 90(4), pp. 1947–1956.

Carabaño, M. *et al*. (2014) 'Modeling heat stress effect on Holstein cows under hot and dry conditions: Selection tools', *Journal of Dairy Science*, 97(12), pp. 7889–7904.

Correa-Calderon, A. *et al*. (2004) 'Thermoregulatory responses of Holstein and Brown Swiss heat-stressed dairy cows to two different cooling systems', *International journal of biometeorology*, 48, pp. 142–148.

Council, N. (1971) 'A guide to environmental research on animals: National Academies'.

Cowley, F.C. *et al*. (2015) 'Immediate and residual effects of heat stress and restricted intake on milk protein and casein composition and energy metabolism', *Journal of Dairy Science*, 98(4), pp. 2356–2368. Available at: https://doi.org/10.3168/jds.2014-8442.

Dairy, U. (2014) *Dairy Cattle Management Practices in the United States*. USDA–APHIS–VS–CEAH–NAHMS.

Gao, S.T. *et al*. (2017) 'The effects of heat stress on protein metabolism in lactating Holstein cows', *Journal of dairy science*, 100(6), pp. 5040–5049.

Gisbert-Queral, M. *et al*. (2021) 'Climate impacts and adaptation in US dairy systems 1981–2018', *Nature Food*, 2(11), pp. 894–901.

GROSU, H. *et al*. (2013) 'HISTORY OF GENETIC EVALUATION METHODS IN DAIRY CATTLE I. DAUGHTER-DAM COMPARISONS', *SCIENTIFIC PAPERS*, p. 25.

Hammami, H. *et al*. (2013) 'Evaluation of heat stress effects on production traits and somatic cell score of Holsteins in a temperate environment', *Journal of dairy science*, 96(3), pp. 1844–1855.

Hazel, A., Heins, B. and Hansen, L. (2020) 'Fertility and 305-day production of Viking Red-, Montbéliarde-, and Holstein-sired crossbred cows compared with Holstein cows during their first 3 lactations in Minnesota dairy herds', *Journal of dairy science*, 103(9), pp. 8683–8697.

Hill, D.L. and Wall, E. (2015) 'Dairy cattle in a temperate climate: the effects of weather on milk yield and composition depend on management', *Animal*, 9(1), pp. 138–149. Available at: https://doi.org/10.1017/S1751731114002456.

Kadzere, C.T. *et al*. (2002) 'Heat stress in lactating dairy cows: a review', *Livestock production science*, 77(1), pp. 59–91.

Key, N. and Sneeringer, S. (2014) 'Potential effects of climate change on the productivity of US dairies', *American Journal of Agricultural Economics*, 96(4), pp. 1136–1156.

Li, M. *et al*. (2022) 'Investigating the effect of temporal, geographic, and management factors on US Holstein lactation curve parameters', *Journal of Dairy Science*, 105(9), pp. 7525–7538.

Mauger, G. *et al*. (2015) 'Impacts of climate change on milk production in the United States', *The Professional Geographer*, 67(1), pp. 121–131.

Schlenker, W. and Roberts, M.J. (2009) 'Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change', *Proceedings of the National Academy of Sciences*, 106(37), pp. 15594–15598. Available at: https://doi.org/10.1073/pnas.0906865106.

St-Pierre, N., Cobanov, B. and Schnitkey, G. (2003) 'Economic losses from heat stress by US livestock industries', *Journal of dairy science*, 86, pp. E52–E77.

Yousef, M.K. and others (1985) *Stress physiology in livestock. Volume I. Basic principles*. CRC press.

Appendix

A. Introduction Chapter of Appendix

A.1 Estimation of lactation curves

To specify this attribute of lactation curve, the paper uses the package 'lactcurves' in R to simulate 38 models for lactation curves and select the best model. In our data, 39 % of the records have the first lactation, 28% of the records have the second lactation, 17 % of the records have the third lactation, and rest 16% records have the lactation times range from 6 to 18. Thus, we calculate the lactation curve for the 1st, 2nd, 3rd lactation and lactation after 3rd time. After running a total of 38 models, we select Khandekar model as the best model for the selection criterion including R2, R2 adjustment, RSE, logL, AIC, AICC and BIC. The model is described in Equation 7.

$$nls(milk\ yielding) = a + b \times dim + c \times dim^2 + d \times dim^3 + f \times log(dim) \quad (7)$$

$$start = list(a = 15, b = -0.15, c = -0.00043, d = 0.0000005, f = 4.05)$$

A.2 Analysis of lactation status

In our dairy research analysis, we have observed that the year 2006 exhibits a higher milk yield than the subsequent years. However, this discrepancy in milk production is not indicative of superior dairy practices in that particular year. Instead, it can be attributed to the nature of our dataset construction. Since our dataset begins in 2006, the daily milk yield records for that year are predominantly derived from the initial stages of each lactation. These stages are characterized by relatively higher milk production compared to the later stages.

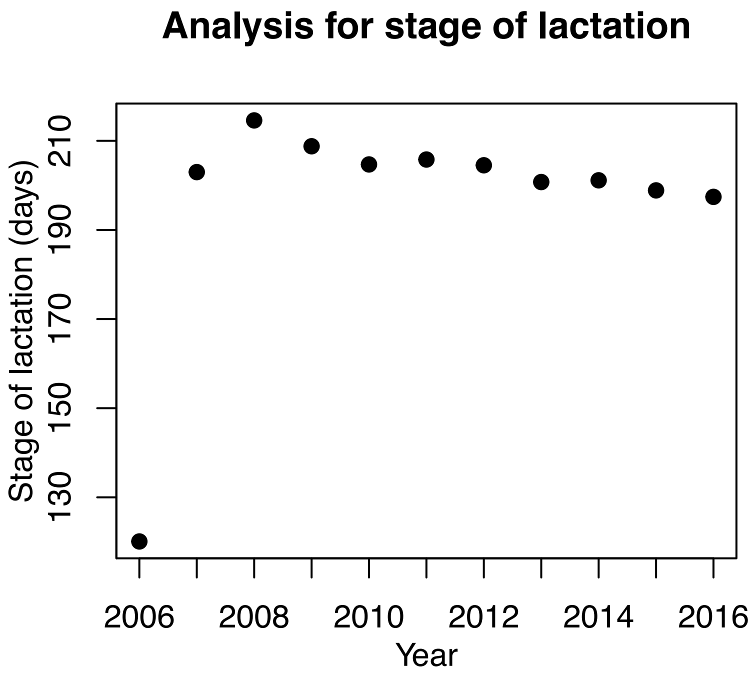


Figure 13. Stage of lactation

To substantiate this claim, we have plotted the number of days in milk for each test day in Figure 13. Days in milk is a term frequently used in dairy research to describe the number of days from the start of each lactation. As shown in the figure, the year 2006 has a significantly lower days in milk record (i.e., earlier lactation stage) compared to the later years. This finding confirms that the higher milk production observed in the year 2006 is a result of the dataset's construction rather than superior dairy practices.

A.3 Analysis of large herd size

Figure 14 analyzes the trend in the size of large herds over the past decade. As illustrated in the graph, there is a steady increase in the number of large herds, including the transformation of small herds into large ones in the last 10 years.

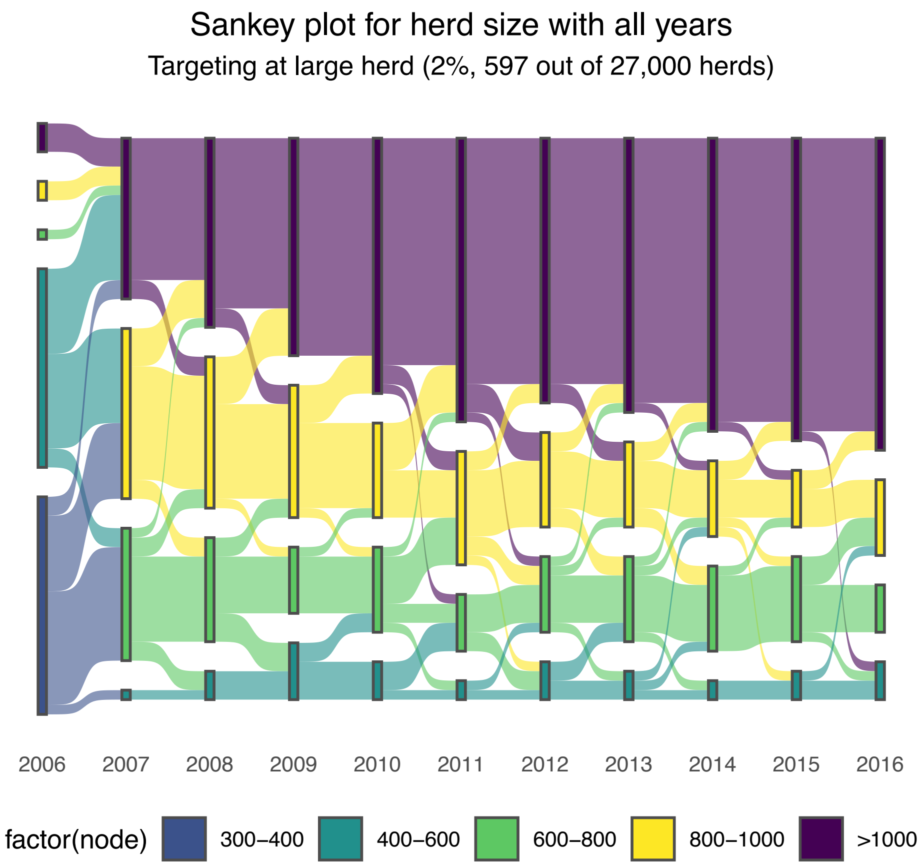


Figure 14. Large herds analysis

B. Methods Chapter of Appendix

B.1 Stability with higher degree of freedoms

In order to select the degrees of freedom, we estimated increasing order until the relationship between the milk yield and the THI appeared stable.

As shown in Figure 15, both Chebyshev and Spline functions are getting stable after the degrees of freedom as nine. Thus, in our analysis, the polynomial degrees for Chebyshev and Spline functions are all set as nine.

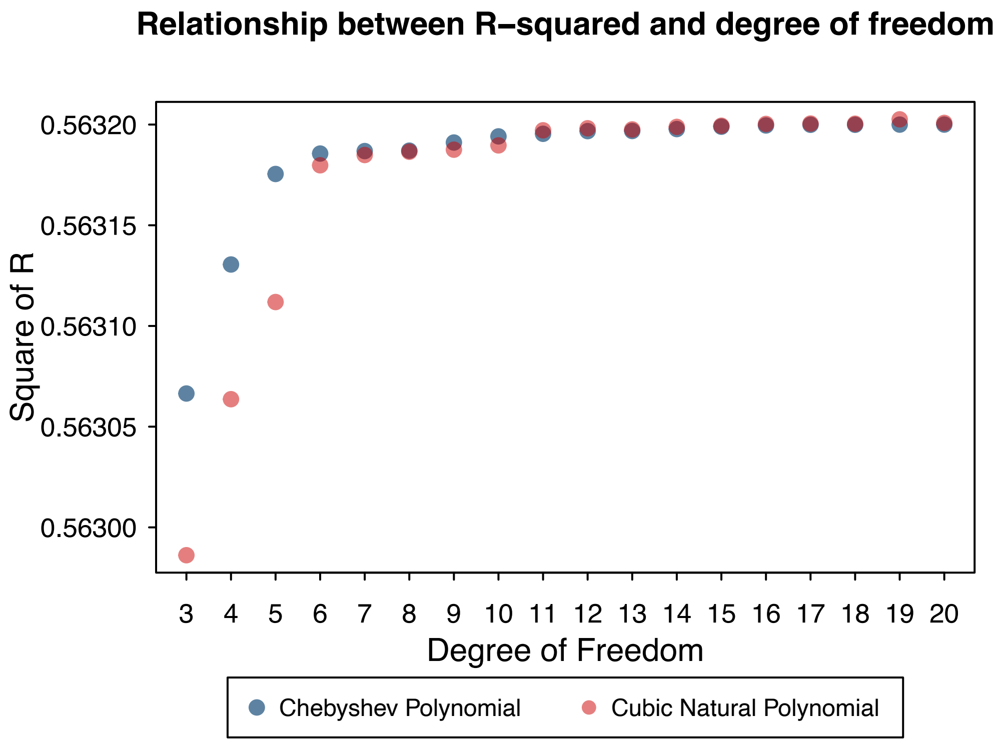


Figure 15. R-sqaured and degree of freedoms

B.2 Fixed effect and standard error clustering

To select the best form of fixed effect c and clustering form for standard error ϵ, we estimated multiple function settings, as illustrated in Figure 16.

Our analysis considered three different fixed effect settings (id, id by year, and id by month), three different time trends (NA, linear format, and quadratic format), and four types of clustering formats for the standard errors (NA format, herd, herd by year, and herd by month). For each model, we selected one type of time trend setting, fixed effect setting, and clustering setting, respectively. Consequently, our analysis estimated 14 models.

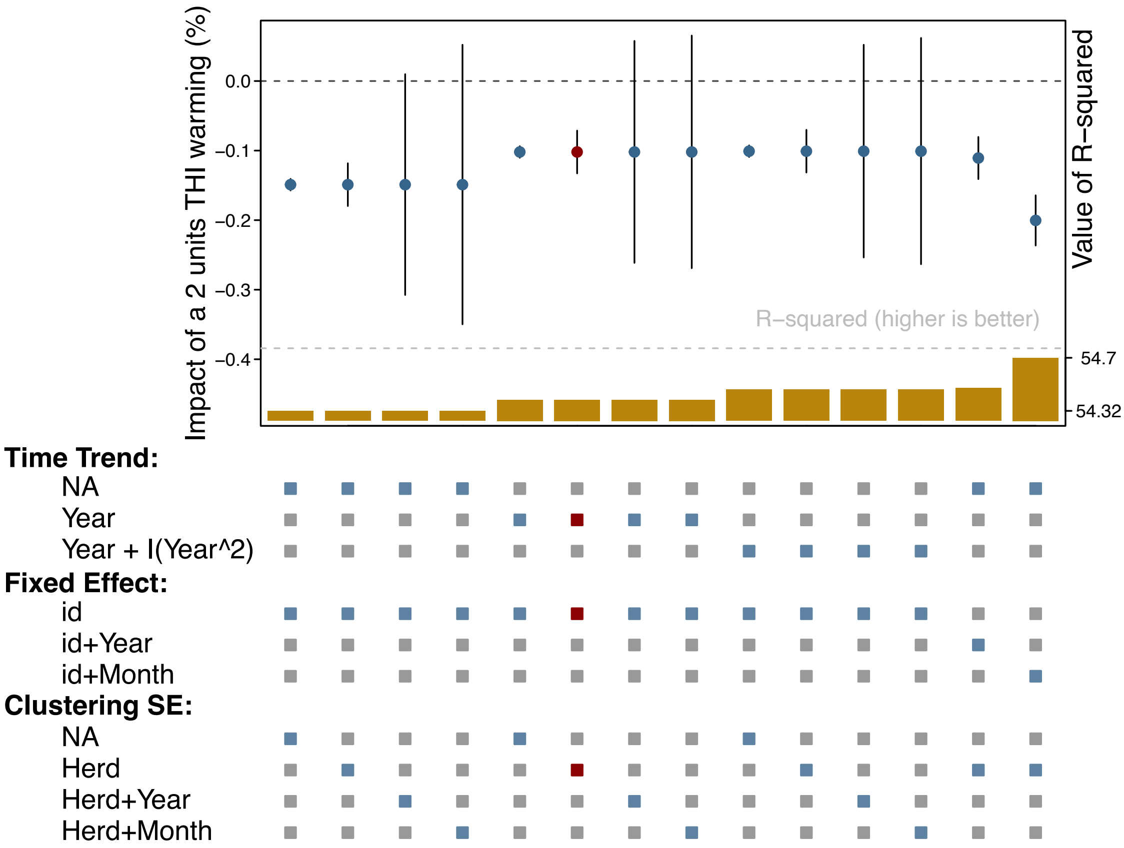


Figure 16. Analysis of multiple models

The corresponding grey squares at the bottom graph were highlighted as red for our model and blue for all other models. We then reported the estimated treatment effect on milk yield with an increase of 2 units of THI as Equation 8 on top of the graph. The y axis indicates the percentage decrease in milk production per hour for every 2 units of THI increase. Specifically, we isolate the effects of adjustments in exposure time for each THI intervals. This entails a uniform displacement of the entire Temperature-Humidity Index (THI) distribution by 2 units.

The blue points represented the estimated treatment effect, and the length of the line indicated the magnitude of variance. The yellow bars at the bottom of the graph showed the R-squared values.

$$\delta(log(M_{it})) = \int_{\underline{THI}}^{\overline{THI}} g(THI)\, \phi_{it}\delta(THI)d(THI) + h_{it} + y_{it} + l_{it} + dim_{it} + c_i + \epsilon_{it} \qquad (8)$$

From the graph, we infer that selecting the time trend as year, fixed effect as id, and clustering the error term on herd by year was an effective choice with a relatively high R^2 and moderate variance.

As a result, a time-invariant individual fixed effect c_i is to control for heterogeneity such as breed and genes. And the error term is clustered at herd by year level. The full model is presented in the form in Equation 4.

C. Analysis Chapter of Appendix

C.1 Polynomial regression result

Table 2. Regression result for milk yield

Dependent Variable:	log(milk)	
Model:	(1)	(2)
Variables		
ns1	-0.0003***	-0.0003***
	(4.97×10^{-5})	(5.02×10^{-5})
ns2	-0.0013***	-0.0013***
	(5.2×10^{-5})	(5.18×10^{-5})
ns3	-0.0010***	-0.0010***
	(5.42×10^{-5})	(5.41×10^{-5})
ns4	-0.0007***	-0.0006***
	(4.93×10^{-5})	(4.89×10^{-5})
ns5	-0.0002***	-0.0002***
	(4.73×10^{-5})	(4.69×10^{-5})
ns6	-4×10^{-5}	8.56×10^{-5}*
	(4.65×10^{-5})	(4.68×10^{-5})
ns7	0.0002***	0.0001**
	(5.61×10^{-5})	(5.65×10^{-5})
ns8	0.0001***	0.0002***
	(3.6×10^{-5})	(3.64×10^{-5})
ns9	-0.0001***	-0.0001***
	(4.62×10^{-5})	(4.68×10^{-5})
herd	-3.49×10^{-6}***	-3.66×10^{-6}***
	(1.3×10^{-6})	(1.28×10^{-6})
year	0.0443***	
	(0.0007)	
lac	0.0020*	-0.0009
	(0.0012)	(0.0011)
dim	-0.0015***	-0.0015***
	(8.6×10^{-6})	(8.59×10^{-6})
Fixed-effects		
id	Yes	Yes
year		Yes
Fit statistics		
Observations	157,556,011	157,556,011
R^2	0.55948	0.56033
RMSE	0.26427	0.26401
Wald (joint nullity), p-value	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$
F-test (projected), p-value	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$

Clustered (herd) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3. Regression result for fat and protein

Dependent Variables: Model:	fat percentage (1)	protein percentage (2)	log(milk) (3)
Variables			
ns1	-0.0059	-0.0050	-0.0003***
	(0.0001)	(5.78×10^{-5})	(5.02×10^{-5})
ns2	-0.0023	-0.0021	-0.0013***
	(0.0001)	(5.99×10^{-5})	(5.18×10^{-5})
ns3	0.0006	-0.0004	-0.0010***
	(0.0001)	(6.33×10^{-5})	(5.41×10^{-5})
ns4	0.0009	0.0003	-0.0006***
	(0.0001)	(5.92×10^{-5})	(4.89×10^{-5})
ns5	2.65×10^{-5}	0.0003	-0.0002***
	(0.0001)	(6.2×10^{-5})	(4.69×10^{-5})
ns6	-0.0002	-9.01×10^{-6}	8.57×10^{-5}*
	(0.0001)	(6.2×10^{-5})	(4.68×10^{-5})
ns7	-0.0003	-0.0001	0.0001**
	(0.0001)	(7.71×10^{-5})	(5.65×10^{-5})
ns8	4.69×10^{-5}	-3.38×10^{-5}	0.0002***
	(0.0001)	(4.63×10^{-5})	(3.64×10^{-5})
ns9	0.0001	9.36×10^{-5}	-0.0001***
	(0.0001)	(5.97×10^{-5})	(4.68×10^{-5})
lac	-0.0036	0.0021	-0.0010
	(0.0010)	(0.0004)	(0.0011)
dim	0.0010	0.0016	-0.0015***
	(1.06×10^{-5})	(4.94×10^{-6})	(8.59×10^{-6})
Fixed-effects			
id	Yes	Yes	Yes
Fit statistics			
RMSE	0.62522	0.23844	0.26401
R^2	0.44625	0.63280	0.5433
Wald (joint nullity), p-value	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$
F-test (projected), p-value	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$	$NaN \times 10^{-Inf}$

Clustered (herd) standard-errors in parentheses

C.2 Chow test for heterogeneity analysis

To analyze whether there is a structural break between the sub samples in our results, we have conducted a Chow test using Equation 9.

$$C = \frac{S_C - (S_1 + S2)/k}{(S_1+S_2)/(N_1+N_2-2k)} \quad (9)$$

H_0 : there is no structural break in the relationship between the sub-sample.

H_1 : there is a structural break in the relationship between the first and second half of the sample.

where C is the Chow test statistic, S_C, S_1 and S_2 are the sum of squared residuals from the combined data, sub-sample 1 and sub-sample 2 respectively. N_1 and N_2 are the sample size. k is the length of coefficients. The test statistic follows the F-distribution with k and $N_1 + N_2 - 2k$ degrees of freedom.

We did the Chow test for herd size and time frame. Our result shows that the Chow test statistic is 289,380 and 141,886.6, while the critical value for the F-test is $F(0.95, 12, 107752151) = 2.18$. Thus, we can reject the null hypothesis at 95% significance level, and accept that there is a structural break in the relationship between the first and second half of the sample and between the small and large herds.

D. Robustness Chapter of Appendix

D.1 Methods check

In this section, we introduce the step function, polynomial, and Chebychev functions with different degree of freedoms, respectively. Upon examination, we observed that these models exhibit comparable patterns to the primary model utilized in our prior analysis. As a result, the robustness check substantiates the robustness of our earlier model.

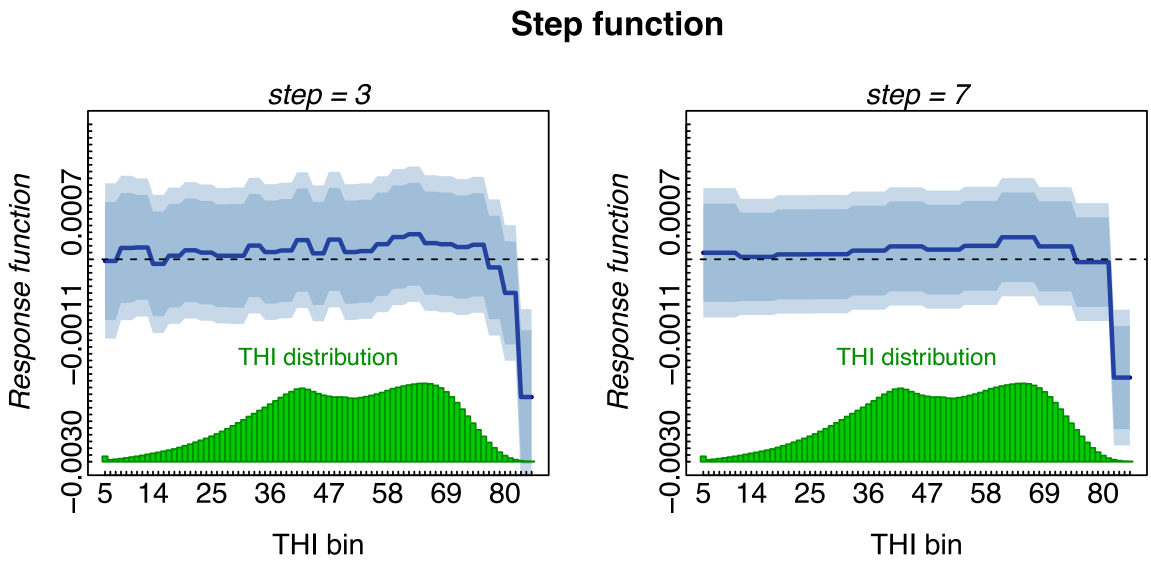


Figure 17. Nonlinear Relation Between THI and Milk Yield using Step Functions

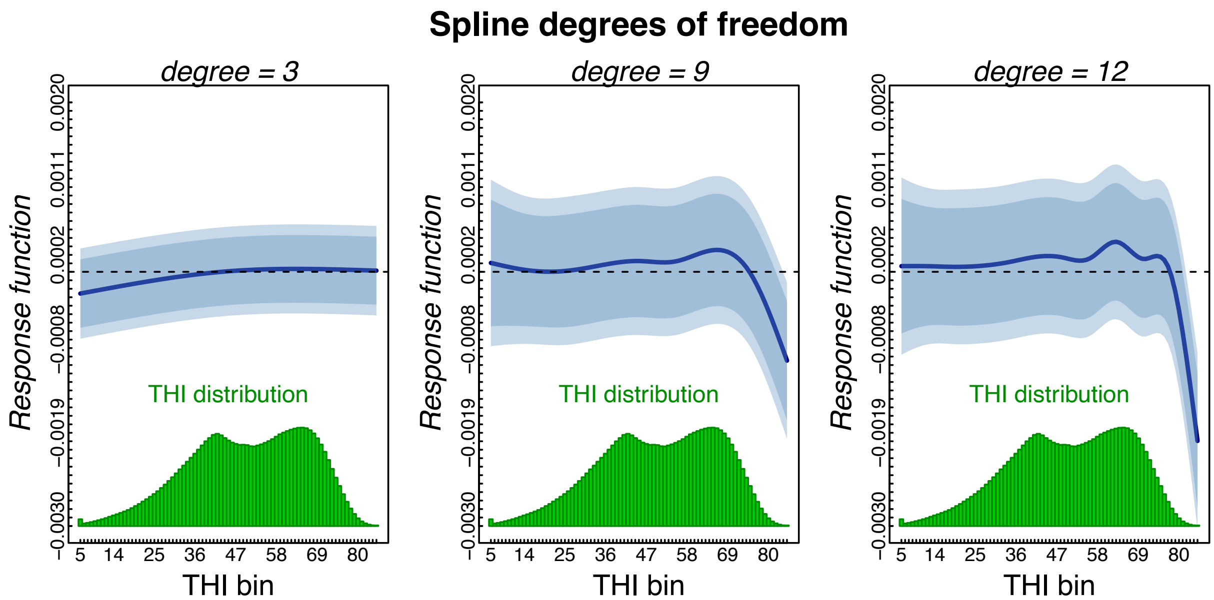


Figure 18. Nonlinear Relation Between THI and Milk Yield using Spline Functions

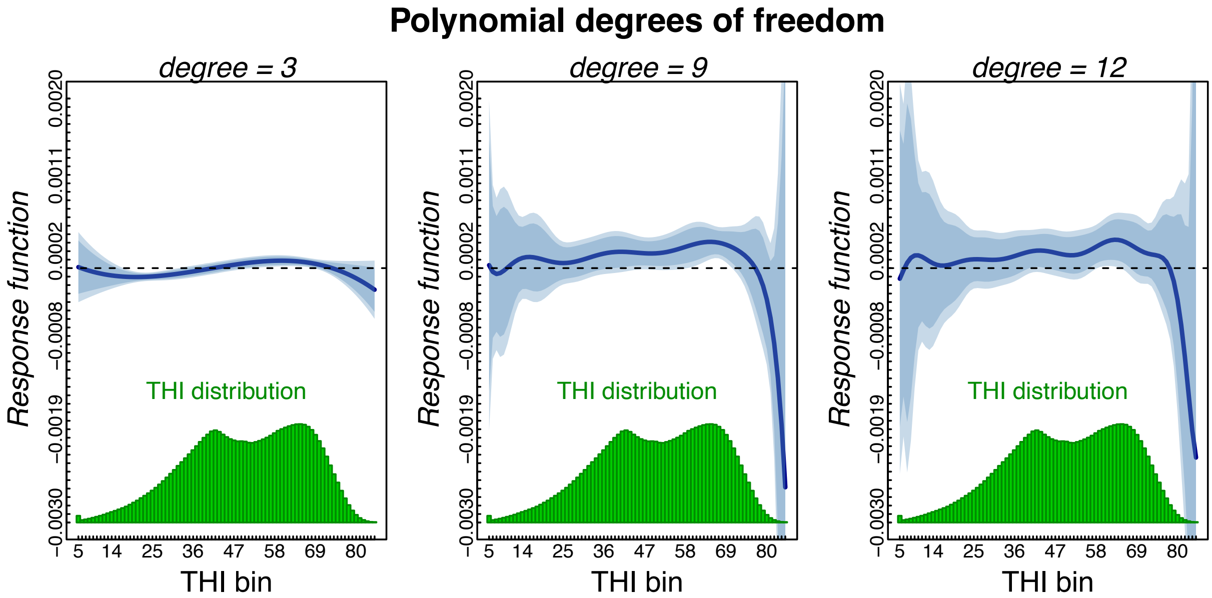


Figure 19. Nonlinear Relation Between THI and Milk Yield using Spline Functions