Supplemental File S5. Temperature-Humidity analysis

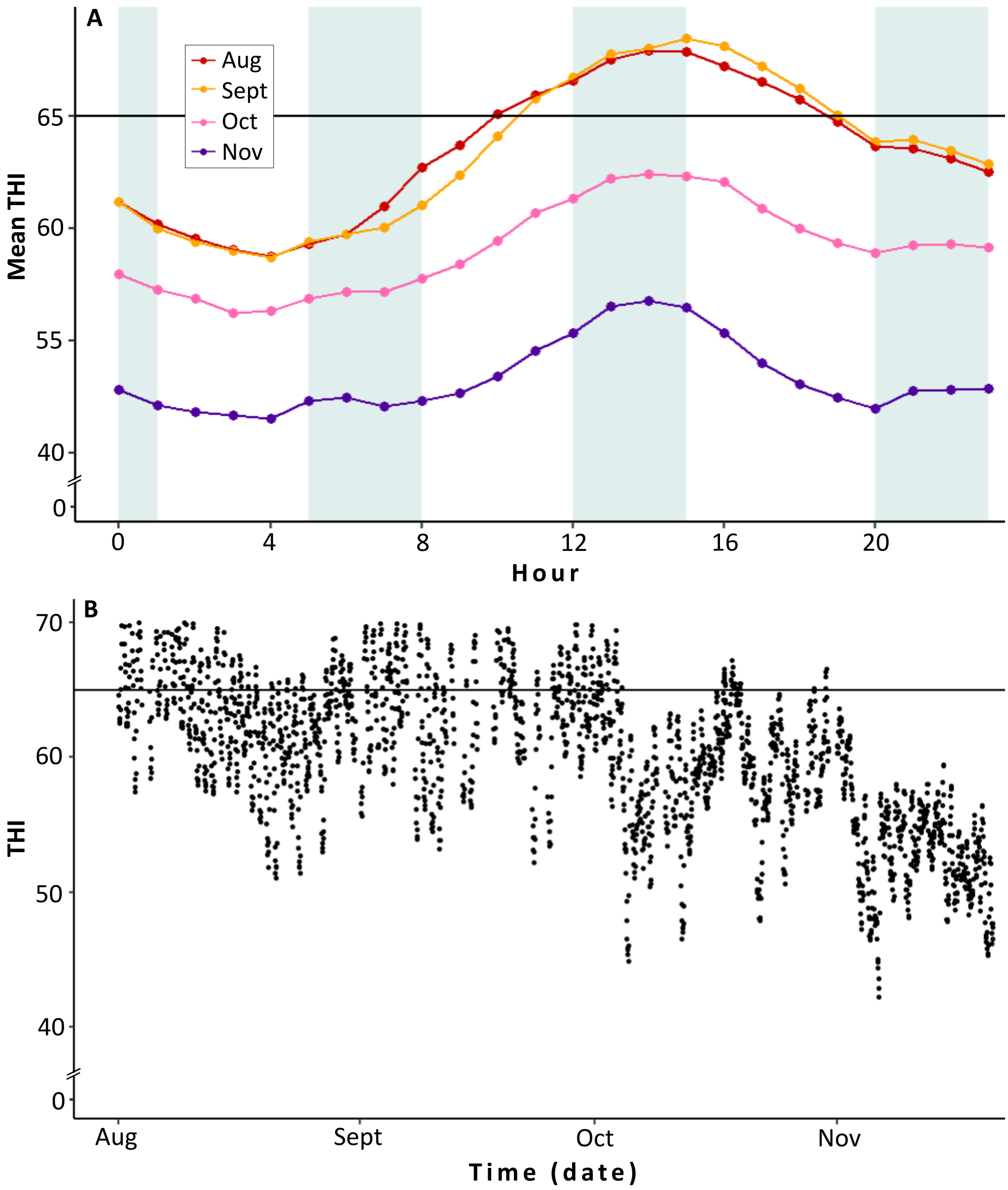
Previous studies have used both temperature and the Temperature-Humidity Index (**THI**) to measure the impacts of heat stress on animals. THI considers both the effects of temperature and relative humidity associated with thermal stress, and was initially developed to gauge human heat discomfort (Thom, 1959), although numerous modifications have been created for use on cattle e.g., (Berry, 1964; Mader et al., 2006; Morton et al., 2007). THI is relatively simple to measure, including on farms with large herds (Li et al., 2009), although most studies do not usually account for localized 'microclimates' (fine-scale spatiotemporal variations in temperature and humidity) within indoor environments (Bouraoui et al., 2002; Morton et al., 2007; McDonald et al., 2020).

Here we explain in further detail how humidity and temperature were combined and assessed in relation to the observed bunching metrics. Although humidity was not recorded within the barn, so we were unable to assess micro-variation in humidity, macro-level changes in humidity should be reflected from recordings taken from a local weather station even though these are not a direct measure of the indoor environment. Moreover, similarly weak relative humidity data has been used in previous studies assessing the impacts of heat stress on cattle e.g., (McDonald et al., 2020). We therefore obtained hourly relative humidity readings Harold Hill weather station (Weather Underground, 2014a), approximately 11km from the study location. We also considered data from Stanford Rivers (approximately 7km from the study site), but we did not use these data because of frequent non-sensical recordings, with the maximum of 100% humidity frequently recorded for an unrealistic number of consecutive time points (Weather Underground, 2014b). We thus calculated Temperature Humidity Index (THI) using the equation from (Kibler, 1964) (Equation 1):

Equation 1. $1.8 \times T - (1 - RH)(T - 14.3) + 32,$

where BT = barn temperature (°C), and RH = relative humidity.

Supplemental Figure S5.1A shows that THI using humidity from Harold Hill peaked during the afternoon. The mean hourly THI reaches above the hypothesized breakpoint of THI = 65 during the afternoon in August and September, but not during October and November (Supplemental Figure S5.1A). Supplemental Figure S5.2B shows THI per hour across time, showing that most of the values above the hypothesized breakpoint of THI = 65 lie in August and September (~88.52%, n = 455 of 514 hours).



Supplemental Figure S5.1. (A) Mean 24-hourly Temperature-Humidity Index (THI) averaged over all days per month, and (B) mean THI recorded during the entire study period (August to November 2014). In A, months shown are August = red, September = orange, October = pink, November = purple. The data along the x axis shaded in light grey correspond to milking periods (05:00-07:59, 12:00-14:59, 20:00-22:59) and periods during which the sensor reset (23:00 to 00:59). In B, the black data points are included in the analysis ("analysed"). The remaining data points, in grey, were removed from the analysis due to milking or because they correspond to when the system reset ("removed"). Instances where THI $\geq$ 65 are above the black horizontal line (n = 451 of 1125 hours).

Analysis was conducted to assess the relationship between the THI values from Harold Hill weather station and each of the bunching metrics (core range = CR, full range (FR), intercow distance (ICD) and nearest neighbor distance = NND), to determine whether the relationships vastly differ from those using temperature alone (Table 2 in main paper). For this, we used a breakpoint of THI = 65, the typical THI in UK barns according to six months of data collected from six farms in 2021 (Chamberlain et al., 2022), roughly equivalent to 20°C at a relative humidity of 60% (the mean hourly value over the study duration).

Using the hypothesized breakpoint of THI = 65, the breakpoint is significant for all the metrics except ICD considering all the data, although this also becomes non-significant when excluding outliers for FR and NND (n = 5 data points) to satisfy model assumptions (Supplemental Table S5.1-2; Supplemental Figure S5.3). The relation is (significantly) negative above this breakpoint, before and after excluding the outliers (n = 5 data points), except for with NND after excluding outliers (Supplemental Table S5.1-2; Supplemental Figure S5.3). It is not unexpected that the negative pattern does not hold for NND; the positive relationship with THI is negligible and non-significant (e = 0.001, $P = 0.51$; Supplemental Table S5.1), like the non-significant weak relationship found between NND and barn temperature (e = -0.01, $P = 0.19$; Table 2 in main paper).

Supplemental Table S5.1. Linear regression model outputs below and above the hypothesized breakpoint of Temperature-Humidity Index (THI) = 65 (*n* = 998 and 283 data points, respectively), using humidity from a local weather station, Harold Hill. Chow test outputs testing the significance of the hypothesized breakpoint are also shown. Each test is conducted for each bunching metric, **with** outliers[1]

Metric	**Chow-test**	**THI < 65** *[Shapiro-Wilks; NVC]*	**THI ≥ 65** *[Shapiro-Wilks; NVC]*
CR	F = 12.87, *p* = 2.94e-6*	e = -0.13, SE = 0.06, t-value = -2.38, *P* = 0.02* *[W = 0.98, P = 5.54e-9*; Chi-sq = 0.97, P = 0.33]*	e = -1.44, SE = 0.18, t-value = -7.86, *P* = 8.12e-14* *[W = 0.99, P = 0.006*; Chi-sq = 0.68, P = 0.41]*
FR	F = 8.56, *p* = 0.0002*	e = -0.20, SE = 0.11, t-value = -1.85, *P* = 0.06 *[W = 0.99, P = 1.61e-7*; Chi-sq = 0.48, P = 0.49]*	e = -1.77, SE = 0.47, t-value = -3.81, *P* = 0.0002* *[W = 0.98, P = 0.0001*; Chi-sq = 5.44, P = 0.02*]*
ICD	F = 1.98, *p* = 0.14	e = -0.01, SE = 0.005, t-value = -2.78, *P* = 0.005* *[W = 0.94, P < 2.2e-16*; Chi-sq = 61.71, P = 3.99e-15*]*	e = -0.24, SE = 0.03, t-value = -6.97, *p* = 2.28e-11* *[W = 0.91, P = 6.84e-12*; Chi-sq = 0.73, P = 0.39]*
NND	F = 7.94, *p* = 0.0004*	e = 0.01, SE = 0.001, t-value = 8.19, *P* = 7.88e-16* *[W = 0.99, P = 0.0003*; Chi-sq = 24.46, P =7.59e-7*]*	e = -0.002, SE = 0.007, t-value = -0.23, *P* = 0.82 *[W = 0.98, P = 0.0001*; Chi-sq = 0.96, P = 0.33]*

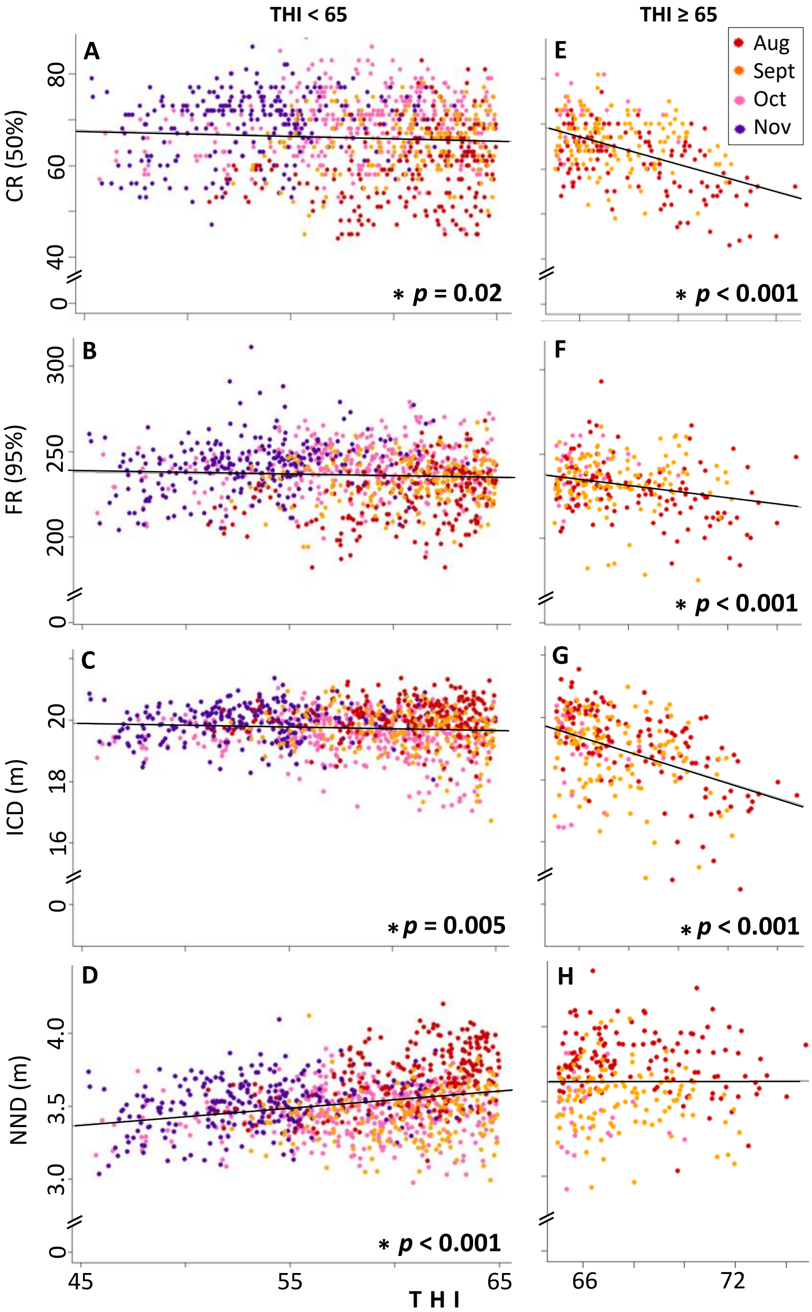
[1]Bunching metrics: CR = core range, FR = full range, ICD = intercow distance, NND = nearest-neighbor distance.

*Significant *P*-values.

Supplemental Table S5.2. Linear regression model outputs below and above the hypothesized breakpoint of Temperature-Humidity Index (THI) = 65 (*n* = 998 and 278 data points, respectively. Chow test outputs testing the significance of the hypothesized breakpoint are also shown. Each test is conducted for each bunching metric, **without** outliers (*n* = 5 data points), based on model assumption testing[1]

Metric	**Chow-test**	**THI < 65** *[Shapiro-Wilks; NVC]*	**THI > 65** *[Shapiro-Wilks; NVC]*
CR	F = 5.84, *p* = 0.003	e = -0.11, SE = 0.05, t-value = -2.07, *P* = 0.04* *[W = 0.98, P = 5.54e-9*; Chi-sq = 0.97, P = 0.33]*	e = -1.52, SE = 0.17, t-value = -8.76, *P* < 2.e-16* *[W = 0.99, P = 0.06; Chi-sq = 2.60, P = 0.11]*
FR	F = 1.84, *p* = 0.16	e = -0.20, SE = 0.11, t-value = -1.85, *P* = 0.06 *[W = 0.99, P = 1.61e-7*; Chi-sq = 0.48, P = 0.49]*	e = -0.19, SE = 0.04, t-value = -4.98, *P* = 7.21e-7* *[W = 0.99, P = 0.06; Chi-sq = 5.78, P = 0.02]*
ICD	F = 1.87, *p* = 0.15	e = -0.01, SE = 0.005, t-value = -2.45, *P* = 0.01* *[W = 0.94, P < 2.2e-16*; Chi-sq = 61.71, P = 3.99e-15*]*	e = -0.25, SE = 0.03, t-value = -7.98, *P* = 3.84e-14* *[W = 0.95, P = 2.27e-8*; Chi-sq = 1.49, P = 0.22]*
NND	F = 0.14, *p* = 0.87	e = 0.01, SE = 0.001, t-value = 8.53, *P* < 2e-16* *[W = 0.99, P = 0.0003*; Chi-sq = 24.46, P = 7.59e-7*]*	e = 0.001, SE = 0.01, t-value = 0.14, *p* = 0.89 *[W = 1.00, P = 0.51; Chi-sq = 0.07, P = 0.79]*

[1]Bunching metrics: CR = core range, FR = full range, ICD = intercow distance, NND = nearest-neighbor distance.
*Significant *P*-values.



Supplemental Figure S5.2. Linear regression plot outputs below and above the hypothesized breakpoint of Temperature-Humidity Index (THI) = 65 (> THI = 65: *n* = 170 data points). Plots are shown for each bunching metric: core range (CR; 50%; measured in virtual cells of 2.25m[2]), full range (FR; 95%; measured in virtual cells of 2.25m[2]), intercow distance (ICD), nearest neighbor distance (NND), excluding outliers (*n* = 5 data points), based on model assumption testing. A single point represents an average per hour and points are colored by month: August = red, September = orange, October = pink and November = purple. Significant relationships are marked with an asterisk (*).

Forced breakpoints are outputted as approximately THI = 54 to 68, and the relations above these breakpoints remain negative, and are significant for all metrics except for NND (Supplemental Table S5.3).

Supplemental Table S5.3. The relation between Temperature-Humidity Index (THI) and each bunching metric. THI breakpoints were found using a data-driven approach. Outliers have been excluded, based on linear model assumption testing (n = 5 data points)[1]

Metric	BP	Chow-test	< BP	≥ BP
CR	68.27	F = 5.20, *P* = 0.01	e = -0.07, SE = 0.04, t-value = -1.68, *P* = 0.09	e = -2.36, SE = 0.47, t-value = -5.02, *P* = 2.29e-6*
FR	54.39	F = 2.34, *P* = 0.10	e = 0.90, SE = 0.43, t-value = 2.11, *P* = 0.04*	e = -0.58, SE = 0.11, t-value = -5.22, *P* = 2.19e-7*
ICD	65.52	F = 2.43, *P* = 0.09	e = -0.02, SE = 0.005, t-value = -3.45, *P* = 0.001*	e = -0.27, SE = 0.03, t-value = -7.92, *P* = 8.5e-14*
NND	68.47	F = 1.05, *P* = 0.35	e = 0.01, SE = 0.001, t-value = 10.23, *P* < 2e-16*	e = -0.01, SE = 0.02, t-value = -0.80, *P* = 0.42

[1]Bunching metrics: CR = core range, FR = full range, ICD = intercow distance, NND = nearest-neighbor distance.
*Significant *P*-values.

REFERENCES

Berry, I. L., M. D. Shanklin, and H. D. Johnson. 1964. Dairy shelter design based on milk production decline as affected by temperature and humidity. Trans. ASAE 7:0329–0331. https: //doi .org/10.13031/2013.40772.

Bouraoui, R., M. Lahmar, A. Majdoub, M. Djemali, and R. Belyea. 2002. The relationship of temperature-humidity index with milk production of dairy cows in a Mediterranean climate. Anim. Res. 51:479–491. https://doi.org/10.1051/animres:2002036.

Chamberlain, A. T., C. D. Powell, E. Arcier, and N. Aldenhoven. 2022. The relationship between on-farm environmental conditions inside and outside cow sheds during the summer in England: Can temperature humidity index be predicted from outside conditions?

Kibler, H. H. 1964. Environmental physiology and shelter engineering with special reference to domestic animals. LXVII, Thermal effects of various temperature-humidity combinations on Holstein cattle as measured by eight physiological responses. Mizzou 1–40.

Li, S., K. G. Gebremedhin, C. Lee, and R. Collier. 2009. Evaluation of thermal stress indices for cattle. Page 1, American Society of Agricultural and Biological Engineers Annual International Meeting. American Society of Agricultural and Biological Engineers, Reno, NV.

McDonald, P. V., M. A. G. von Keyserlingk, and D. M. Weary. 2020. Hot weather increases competition between dairy cows at the drinker. J. Dairy Sci. 103:3447–3458. https://doi.org/10.3168/jds.2019-17456.

Mader, T. L., M. S. Davis, and T. Brown-Brandl. 2006. Environmental factors influencing heat stress in feedlot cattle. J. Anim. Sci.84:712–719. https://doi.org/ 10.2527/2006.843712x.

Morton, J. M., W. P. Tranter, D. G. Mayer, and N. N. Jonsson. 2007. Effects of environmental heat on conception rates in lactating dairy cows: Critical periods of exposure. J. Dairy Sci. 90:2271–2278. https://doi.org/10.3168/jds.2006-574.

Thom, E. C. 1959. The discomfort index. Weatherwise 12:57–61. https://doi.org/10.1080/00431672.1959.9926960.

Weather Underground. 2014a. Harold Hill, Forecast for Romford, GB. Accessed Sep. 22, 2022. https://www.wunderground.com/history/daily/gb/romford.

Weather Underground. 2014b. Stanford Rivers, Forecast for Stanford Rivers, GB. Accessed Jul. 11, 2022. https://www.wunderground.com/dashboard/pws/ ISTANFOR2.