BIOSYSTEMS ENGINEERING 226 (2023) 238–251

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Determining the onset of heat stress in a dairy herd based on automated behaviour recognition

Hang Shu [a,b,c], Jérôme Bindelle [c], Leifeng Guo [a,*], Xianhong Gu [b,**]

[a] Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100086, China
[b] State Key Laboratory of Animal Nutrition, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, Beijing 100193, China
[c] AgroBioChem/TERRA, Precision Livestock and Nutrition Unit, Gembloux Agro-Bio Tech, University of Liège, 5030 Gembloux, Belgium

ARTICLE INFO

Article history:
Received 1 September 2022
Received in revised form
15 January 2023
Accepted 16 January 2023

Keywords:
Smart livestock farming
Animal welfare
Thermal comfort
Group measurement
Behavioural index

Dairy cows have various strategies for dealing with heat stress, including a change in behaviour. The aim of this study was to propose a deep learning-based model for recognising cow behaviours and to determine critical thresholds for the onset of heat stress at the herd level. A total of 1000 herd behaviour images taken in a free-stall pen were allocated with labels of five behaviours that are known to be influenced by the thermal environment. Three YOLOv5 architectures were trained by the transfer learning method. The results show the superiority of YOLOv5s with a mean average precision of 0.985 and an inference speed of 73 frames per second on the testing set. Further validation demonstrates excellent agreement in herd-level behavioural parameters between automated measurement and manual observation (intraclass correlation coefficient = 0.97). The analysis of automated behavioural measurements during a 10-day experiment with no to moderate heat stress reveals that lying and standing indices were most responding to heat stress and the test dairy herd began to change their behaviour at the earliest ambient temperature of 23.8 °C or temperature-humidity index of 68.5. Time effects were observed to alter the behavioural indicators values rather than their corresponding environmental thresholds. The proposed method enables a low-cost herd-level heat stress alert without imposing any burden on dairy cows.

1. Introduction

Homeotherms, including dairy cows, constantly maintain thermal equilibrium with their environments through thermoregulation (Kadzere, Murphy, Silanikove, & Maltz, 2002). Heat stress is defined as the demand made by the environment for heat dissipation (Silanikove, 2000). It is triggered when the thermal environment exceeds the upper critical threshold of the thermoneutral zone, inducing a variety of physiological and behavioural responses to reduce heat production and increase heat dissipation (Becker, Collier, & Stone, 2020). Due to the lack of real-time, large-scale, and automated measurement of animal-based indicators, heat

* Corresponding author.
** Corresponding author.
E-mail addresses: guoleifeng@caas.cn (L. Guo), guxianhong@vip.sina.com (X. Gu).
https://doi.org/10.1016/j.biosystemseng.2023.01.009
1537-5110/

Nomenclature

AIC	Akaike information criterion
AP	Average precision
CI	Comfort index (%)
CSI	Cow stress index (%)
drinking%	Percentage of cows drinking
eating%	Percentage of cows eating
FN	False negative
FP	False positive
FPS	Frames per second
IoU	Intersection over Union
lying%	Percentage of cows lying
mAP	Mean average precision
Pr	Precision
R	Recall
RH	Relative humidity (%)
SUI	Stall-use index (%)
Ta	Ambient temperature (°C)
THI	Temperature-humidity index
t_N	Total inference time (s)
TP	True positive
YOLO	You Only Look Once

mitigation in practice has long depended on environmental indicators and their critical thresholds (Shu, Wang, Guo, & Bindelle, 2021). Environmental indicators, however, do not reflect the actual response of the animal, making it difficult to evaluate the effectiveness of cooling measures (Collier, Dahl, & VanBaale, 2006).

Dairy cows take a series of behavioural changes to cope with heat stress. As environmental temperature increases, cows will spend more time standing to increase surface area for better heat dissipation (Cook, Mentink, Bennett, & Burgi, 2007). However, this may result in a significant reduction in sleeping time, posing a potential risk to cow welfare (Becker et al., 2020). Cows will drink more frequently under heat stress but with less water each time (Galán et al., 2018). Cows will also reduce feed intake and subsequent rumination to reduce metabolic heat production (Collier, Renquist, & Xiao, 2017). Therefore, recognising changes in behavioural patterns ascribable to heat stress can help quantify the true response of cows.

Recently, deep learning-based methods have allowed the automated recognition of basic cow behaviours such as lying, standing, and drinking with an accuracy of up to 0.976 (Fuentes, Yoon, Park, & Park, 2020; Wu et al., 2021). Further quantification of the results with association to animal growth, health, and welfare, and the extent to which they can be used to improve decision making have been highlighted for future work (Chen, Zhu, & Norton, 2021). Tsai, Hsu, Ding, Rustia, and Lin (2020) analysed how drinking time and frequency were affected by heat stress after detecting drinking behaviour with a convolutional neural network. Still, further application of deep learning techniques is required so that the detection of multiple behaviours enables a more comprehensive analysis of when heat stress is triggered.

Although progress has been made in computer vision-based individual identification in free-stall barns (Xiao, Liu, Wang, & Si, 2022), issues such as lack of colour pattern and occlusion still lead to poor identification. On the other hand, detections from deep learning methods provide an opportunity to calculate herd-level behavioural indices which have been commonly used for evaluating cow comfort. For example, cow lying index, which is defined as the number of cows lying in the stall divided by the total number of cows, has been calculated automatically with a computer vision-based system (Porto, Arcidiacono, Anguzza, & Cascone, 2013). Other indices related to free-stall usage and cow comfort would require knowing whether the cow is standing on the stall bed and whether the cow is eating or drinking. Therefore, a detailed behavioural recognition method that addresses the above questions is still required to compute these indices in an automated way.

Scan sampling, as a common method in animal research, is often used to record herd-level behavioural indices at predetermined intervals. Traditionally, scan sampling requires manual checks through direct observations or video recording, both of which are time- and labour-consuming. With the help of automated behaviour recognition, video frames can be processed in real time. However, continuous processing and storing of data would be a waste of time and memory. Scan sampling is still of great value, especially for behaviours that basically follow continuous and diurnal patterns, such as lying and standing. The sampling interval should always be determined in accordance with specific purposes. For example, studies or regular checks aimed at ascertaining standing and lying patterns may use sampling intervals of 30 or even 60 min (Mitloehner, Morrow-Tesch, Wilson, Dailey, & McGlone, 2001).

By combining deep learning and scan sampling methods, the onset of heat stress is promising to be determined at the herd level with the memory and power of local devices greatly saved. Therefore, the aim of this study was (1) to train and validate a deep learning-based model to recognise cow behaviours and further calculate herd-level behavioural indicators, and (2) to develop critical thresholds of heat stress at herd level based on behavioural indicators.

2. Materials and methods

All protocols involving animals were approved by the Experimental Animal Care and Committee of Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (approval number IAS2021-220).

2.1. *Experimental design*

The experiment was conducted on a free-stall dairy farm in May 2021 in Shandong, China, which has a temperate continental monsoon climate with hot and humid summers. The experiment consisted of two periods, in which the first three days in early May were designed to collect data for training behaviour recognition models while the other ten days during mid-May when the environment got warmer were designed to explore how the onset of heat stress can be determined through automated behaviour recognition.

2.1.1. Housing, animals, and management

The barn had four pens with a 4-row head-to-head design and was covered by a double-pitched roof (gradient 15%). The experimental pen (11 m × 96 m, oriented along the north–south axis) housed 79 lactating Holstein-Friesian dairy cows with 128 stalls and 128 headlocks. The pen could be evenly divided into eight areas, with each having a feeding zone of 16 headlocks and a parallel resting zone of 10–20 stalls with or without a water trough (see Supplementary Material Fig. S1). At the beginning of the experiment, the cows had a mean ± SD milk yield of 30.4 ± 11.8 kg day^{-1}, parity of 2.8 ± 1.4, and days in milk of 273.1 ± 117.3. The cows were milked three times daily at 08:30, 16:30, and 00:00 h in a parlour that was about 20 m away from the barn. All cows were observed to return to the pen within 1 h of departure. A total mixed ration was delivered three times daily after milking. Clean drinking water was delivered in five troughs. Both feed and water were provided ad libitum to all cows. Electronic fans (1.1 m in diameter; capacity: 25000 m^3 h^{-1} each; see Fig. S1) were turned on when the indoor temperature reached 20 °C whereas sprinklers remained closed during the experiment. Stalls were sand-bedded to a depth of about 150 mm and raked once daily while the cows were away for morning milking.

2.1.2. Behavioural and environmental measurement

Cow behaviour was recorded using eight closed-circuit video cameras (DS-2CD3T86FWDV2; Hikvision, Hangzhou, China) which were evenly spaced and placed opposite the pen's longitudinal axis at a height of about 6 m and an angle of about 45° downward (see Fig. S1). Each camera was able to capture a feeding zone with 16 headlocks and a parallel resting zone, allowing the complete side view of the pen to be captured. The eight cameras were linked, synchronised, and controlled using an eight-channel video recording system (DS-7808N-K2; Hikvision, Hangzhou, China). It is recommended that video recordings taken between 08:00 and 15:00 h is best for representing daily behavioural pattern in summer (Uzal Seyfi, 2013). Besides, behavioural assessment should be performed at least 1 h after cows return from morning milking to avoid being affected by intensive feeding (Overton, Sischo, Temple, & Moore, 2002). Accordingly, video recording was performed from 10:30 to 15:00 h on each test day as adapting the previous recommendations to the actual schedule.

Environmental parameters including ambient temperature (Ta, °C) and relative humidity (RH, %) were measured at an interval of 10 min using a total of six Kestrel 5000 environment meters and Kestrel 5400 heat stress trackers (accuracy: ±0.4 °C Ta, ±1% RH; Nielsen-Kellerman, Boothwyn, PA, USA; see Fig. S1). These sensors were evenly distributed in the pen and were fixed at a height of 2.2 m. The measurements from all sensors were averaged for representing the global environment inside the pen. The temperature-humidity index (THI) was calculated according to Eq. (1) (NRC, 1971).

$$THI = (1.8 \times Ta + 32) - (0.55 - 0.005 \times RH) \times (1.8 \times Ta - 26) \quad (1)$$

2.2. Development of behaviour recognition model

2.2.1. Data preparation

Video frames from the first three-day experiment were extracted at an interval of 6 min by using a self-written program in Python. This interval was set for increasing the heterogeneity of the training data since the cows changed their behaviour less frequently. The extracted frames were in JPG format and were further corrected for distortion, cropped, and resized using OpenCV. The final images for training and evaluation had a resolution of 1920 by 1080 pixels. Finally, a total of 1000 images were chosen after eliminating low-quality frames (e.g., lens covered by flies).

To maintain the good quality of training data, all cows presented in all images were annotated as per the definition presented in Table 1. Five target behaviours known to be influenced by the thermal environment were carefully defined to be exclusive, meaning that the cows were not able to perform two behaviours simultaneously. “Standing-in” was previously subdivided into “perching” and “standing” based on whether all four feet or just two feet touched a stall (Cook, Bennett, & Nordlund, 2005). However, this was abandoned due to insufficient data in each subgroup. The annotation tool LabelImg (https://github.com/tzutalin/labelImg) was used to allocate the appropriate class to each cow per image with a bounding box. A total of 1000 annotated images were randomly split into a training set (60%), a validation set (20%), and a testing set (20%). A detailed description of the behaviour recognition dataset with regards to the number of labels per class is given in Table 2.

2.2.2. Deep learning algorithm and transfer learning

YOLO (You Only Look Once) is a popular one-stage framework for object detection. Unlike two-stage methods (e.g., Faster R-CNN), one-stage methods skip the region proposal stage, and regard object detection as a regression task with class probability and coordinates of the bounding box as the outcome. As a result, one-stage methods have much higher inference speeds and relatively lower accuracy.

YOLOv5 is a recent popular version of the YOLO-series of algorithms. Adapting from YOLOv4, YOLOv5 uses Focus structure with CSPdarknet53 as the backbone and introduces Spatial Pyramid Pooling method, mosaic training, self-

Table 1 – Definition of the target cow behaviours.

Behaviour	Definition
Drinking	Standing by a water trough with mouth in the trough
Eating	Standing with neck in a feeding rack
Lying	Lying in total lateral or sternal recumbency within a stall
Standing-in	Standing with two or more feet touching a stall bed
Standing-out	Standing or walking outside the stall but not eating

Table 2 – Overview of the behaviour recognition dataset. Images were extracted using 6-min scan sampling from 10:30 to 15:00 h during the three-day experiment.

Dataset	Number of cow labels					
	Drinking	Eating	Lying	Standing-in	Standing-out	Total
Training (600 images)	164	269	3326	509	2682	6950
Validation (200 images)	49	73	1237	170	836	2365
Testing (200 images)	40	103	1152	186	769	2250
Total	253	445	5715	865	4287	11565

adversary training, and multi-channel feature. It has been reported that YOLOv5 has a much higher inference speed and smaller size compared with its previous versions without sacrificing accuracy (Yang et al., 2020). Thus, YOLOv5 was chosen to recognise cow behaviour in the present study. Specifically, three architectures were trained, with size increasing from YOLOv5s, YOLOv5m, to YOLOv5l.

Deep learning methods require a lot of time and memory to train on a large number of images. Transfer learning, which involves transferring previously learned knowledge from a related task, is expected to accelerate the training process and usually produce better results than training from scratch. The pre-trained weights used in this study were provided by the authors of YOLOv5 based on the COCO dataset which is a benchmark object detection dataset published by Microsoft (Lin et al., 2014).

The training of the different YOLOv5 architectures was performed in Python 3.8 language with Pytorch 1.8.0. on a 64-bit version Windows 11 laptop with NVIDIA GeForce RTX 3060 GPU and 6 GB video memory. The batch size was set to 8 and the epoch was set to 100. Hyperparameters were set as default.

2.2.3. *Performance evaluation*

In object detection tasks, an Intersection over Union (*IoU*) threshold has to be given first to determine whether a predicted bounding box should be classified as positive or negative. The *IoU* is the overlap of the predicted and ground-truth bounding boxes divided by their union, as expressed in Eq. (2):

$$IoU = \frac{Prediction \cap Ground\ truth}{Prediction \cup Ground\ truth} \tag{2}$$

In this study, an *IoU* threshold of 0.5 was adopted by convention. Thus, a detection with an *IoU* $\geq$0.5 was classified as true positive (*TP*), a detection with an *IoU* <0.5 was classified as false positive (*FP*), and a ground truth presented but failed to be detected is classified as false negative (*FN*).

Afterward, the precision (*Pr*), recall (*R*), and average precision (*AP*) were calculated according to Eqs. (3)–(5). The *Pr* indicates the proportion of the predicted bounding boxes being correctly detected whereas the *R* indicates the proportion of the ground-truth bounding boxes being correctly detected. The ideal object detector should have high *Pr* and *R* at the same time. However, there is a trade-off between the two metrics depending on the confidence threshold. Confidence represents the probability (0–1) of a bounding box containing an object and the predictions with class probabilities lower than a given confidence threshold will be removed. A very high confidence threshold will discourage the model from making positive predictions, thus increasing *Pr* and decreasing *R*, and vice versa. The *AP* is a commonly recommended metric since it summarises the *Pr* along with the *R* at all possible confidence thresholds. This was done by default setting the confidence threshold to 0.001.

$$Pr = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$AP = \int_0^1 Pr_{(R)} dR \tag{5}$$

In such a multi-class task, the mean average precision (*mAP*) was used to evaluate the overall performance, as expressed in Eq. (6):

$$mAP = \frac{1}{C} \sum_{i=1}^{C} AP_i \tag{6}$$

where C takes 5, indicating 5 classes (i.e., “Drinking”, “Eating”, “Lying”, “Standing-in”, “Standing-out”). In addition, frames per second (*FPS*) was used to indicate the inference speed, as expressed in Eq. (7):

$$FPS = \frac{N}{t_N} \tag{7}$$

where t_N is the total inference time (s) on *N* images.

Unlike *AP* calculations, where all potential confidence thresholds are required, the confidence threshold for actual inference must be tuned and specified for better detection. The confidence threshold for inference on the testing set was determined by checking the global maximum on the F1 confidence curve. The F1 score is the harmonic mean of Pr and R, as expressed in Eq. (8):

$$F1\ score = 2 \times \frac{Pr \times R}{Pr + R} \tag{8}$$

2.3. *Behavioural indicator calculation and heat stress determination*

2.3.1. *Behaviour recognition*

The video frames were further extracted using 30-min scan sampling from 10:30 to 15:00 h since this method has been validated to be effective and efficient for analysing cow behaviour (Mattachini, Riva, & Provolo, 2011; Uzal Seyfi, 2013). Consequently, 10 scan samples were obtained for each of the 13 test days, each containing eight images from eight cameras. The proposed behaviour recognition model was applied to all images per scan sample.

2.3.2. Detection filtering

The eight cameras captured all 128 headlocks in the feeding zone, but partially overlapped at the far end of the view (i.e., resting zone) due to the fact that faraway objects naturally appear smaller than closer ones. However, it is important to count each cow only once when calculating herd-level behavioural indicators. Thus, processing had to be done to filter each detection per image per camera to ensure that only the detections located within the area of interest were counted. As shown in Fig. 1, the area of interest was predetermined for each camera by using a polygon that covered the entire floor of the feeding zone and its parallel resting zone. The filtering was based on Jordan Curve Theorem (Hales, 2007). For each predicted bounding box, a ray was first drawn horizontally to the right of the centre of its lower boundary line. This was to ensure that the cows stood or lied within the area. If the number of intersections was odd then the centre was inside the polygon and the predicted bounding box was kept. This filtering was based on the coordinates of the polygon and those of the lower centres of the bounding boxes using a self-written program in Python. The filtered number of detections per class per image was therefore obtained.

2.3.3. Behavioural parameters

The number of detections for each behavioural class per scan was determined after merging the results into each scan sample. The herd-level behaviour distribution of each scan was naturally calculated by dividing the number of detections for each behavioural class by the total number of detections. Except for the percentage of cows lying, several herd-level behavioural indices have often been used for characterising lying and standing behaviours. With the automated measurements, comfort index (*CI*), stall-use index (*SUI*), and cow stress index (*CSI*) were calculated according to Eqs. (9 - 11) (Mattachini et al., 2011; Overton et al., 2002):

$$CI = lying\,/\,(lying + standing{-}in) \tag{9}$$

$$SUI = lying\,/\,(lying + standing{-}in + standing{-}out) \tag{10}$$

$$CSI = (standing{-}in + standing{-}out)\,/\,n \tag{11}$$

where n denotes the total number of cows.

To further evaluate the proposed behaviour recognition model, the scan samples from the first three-day recording were manually observed to count the frequency of each behaviour. The manual results of behavioural parameters including the percentage of all five target behaviours as well as advanced lying and standing indices (i.e., *CI*, *SUI*, and *CSI*) were then calculated. The intraclass correlation coefficient was computed to assess the agreement between manual and automated methods using the icc function from the "irr" package (R version 3.4.4; https://R-project.org).

2.3.4. Threshold development

Since the proposed model worked well in recognising herd-level behavioural parameters, the remaining 10-day data were further used to explore the herd-level behavioural pattern with respect to the onset of heat stress. A total of six herd-level behavioural parameters, including the percentages of cows drinking (drinking%), eating (eating%), and lying (lying %), as well as *CI*, *SUI*, and *CSI*, were used as animal-based indicators, whereas *Ta* and *THI* were used as environmental indicators. Data from 11:00 to 14:30 h were used for further analyses since their behavioural results show a clear association with environmental indicators.

All statistical analyses in this section were performed using R software. Spearman's rank correlation analysis was performed using the cor function to explore how these indicators were associated with each other. Piecewise regression models were used to fit the response of animal-based indicators to environmental indicators and locate the breakpoint at which this response changed in trend. The piecewise models were built with the "segmented" package which works by updating the existing models with two or more breakpoints based on Davies test (Muggeo, 2008). Therefore, basic piecewise models were built by updating simple linear regression models fitted with the lm function. The models are written as Eq. (12):

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 (X_i - X_{bp}) X_k + \varepsilon_i, X_k = \begin{cases} 0 \text{ if } X \le X_{bp} \\ 1 \text{ if } X > X_{bp} \end{cases} \tag{12}$$

where Y_i is the animal-based indicators, β_0 is the population intercept, β_1 is the left slope, X_i is the environmental indicators, β_2 is the difference between the right slope and left slope, X_{bp} is the breakpoint, X_k is the dummy variable, and ε_i is the random residual for the i-th observation.

If the basic piecewise models converged, advanced piecewise models with the random effect of time of day were built to separate the profiles of different hours. Otherwise, the results would be shown with simple linear regression models. Advanced piecewise models were built by updating linear mixed models fitted with the lme function included in the "nlme" package. The random effect of time of day was

Fig. 1 – Schematic of detection filtering taking one camera for example. The predefined red polygon represents the area of interest and only the bounding boxes with their lower centres in the area are kept.

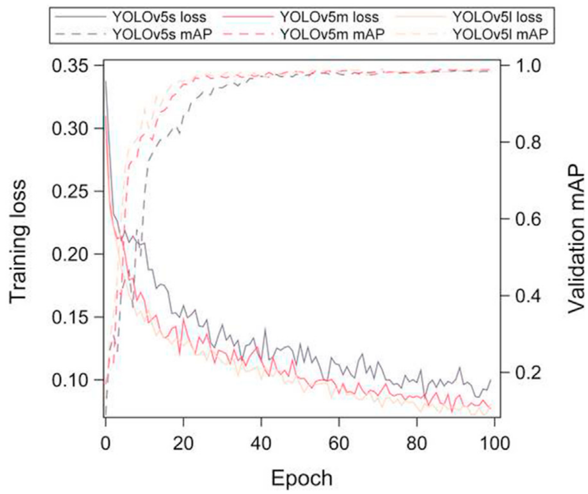

Fig. 2 – Comparison of training loss and validation mean average precision (*mAP*) for three YOLOv5 architectures.

included for every model parameter, including intercept, slope difference, and breakpoint. For j-th time of day, the model can be written as Eq. (13):

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \beta_{2j}(X_{ij} - X_{bpj})X_k + \varepsilon_{ij}, X_k = \begin{cases} 0 \; if \; X \le X_{bpj} \\ 1 \; if \; X > X_{bpj} \end{cases} \quad (13)$$

where each parameter is given by the sum of fixed and random effects (e.g., β_{0j} is the sum of fixed term β_0 and random term d_j).

3. Results and discussion

3.1. *Performance of behaviour recognition*

Figure 2 showing the training process of three YOLOv5 architectures indicates a faster convergence always in the validation *mAP* than the training loss. As expected, the training converged earlier as the network deepens in size. As shown in Table 3, the *mAP* on the testing set was roughly close among YOLOv5 models, in which "Drinking" was consistently the most difficult to detect, probably due to the most limited data to train. This is somehow inevitable when using such an imbalanced dataset with "Standing-out" being almost 16 times more represented than "Drinking". In the case of FPS, models with smaller sizes show much higher inference speed, with YOLOv5s increasing by 32.7% compared with YOLOv5m, and by 65.9% compared with YOLOv5l. Other algorithms commonly used for object detection were not compared since the YOLOv5 algorithms have already shown extremely good performance. In fact, YOLOv5 has been reported to have a dramatic increase in inference speed compared with YOLOv3 and YOLOv4 (Lv et al., 2022). Our results show that all three YOLOv5 architectures should be effective and capable of being deployed on mobile terminals. Anyway, we simply chose the YOLOv5s model for further application due to its ability to further compress the weight size while maintaining accuracy.

The F1 confidence curve indicates that the F1 score of all behavioural classes peaked at a confidence threshold of 0.696 (Fig. 3(a)). Thus, further evaluation and application were performed with confidence threshold set to 0.696. The confusion matrix shown in Fig. 3(b) indicates that the major misclassification was marked between drinking and standing-out behaviours and between standing-in and lying behaviours, with 1 out of 40 "Drinking" labels being misclassified as "Standing-out" and 4 out of 186 "Standing-in" labels being misclassified as "Lying", respectively. The example results shown in Fig. 3(c) demonstrate a good ability in dealing with occlusion caused by facilities or other cows.

When calculating herd-level behavioural parameters for the 30 scan samples from the first three-day recording, the intraclass correlation coefficient between manual and automated methods was 0.97. Moreover, the overall linear relationship shows that only 11 out of 240 observations were classified to be outliers by the 95% prediction limits (Fig. 4). Collectively, these results demonstrate an excellent agreement between manual and automated methods in obtaining herd-level behavioural parameters. Of note, given that the view provided by our cameras might be difficult to show the relationship between a cow's head and the trough when there was a strong occlusion, both the ground truth and detections could have underestimated the incidence of drinking, thereby affecting the calculated behavioural parameters. For example, the drinking% and the SUI would be underestimated whereas the *CSI* would be overestimated. This bias should cause negligible effects on the current results since such occlusion happened in only five scan samples (a total of eight occluded cows) during the entire 13 test days and drinking played a limited role in the equation compared with standing and lying. However, it may have a stronger impact on behavioural parameter calculation when facing a higher farming density and more occlusions. A top-view camera, as presented by Tsai et al. (2020), could be a solution to eliminating this measurement bias. Another possible way is to introduce new labels describing cow behaviour in the drinking area. This is of interest since it can be very crowded in drinking areas during hot seasons and cows would even compete for troughs (McDonald, von Keyserlingk, & Weary, 2020).

Table 3 – Performance comparison of three YOLOv5 architectures.

Model	Average precision					*mAP*	FPS	Size (M)
	Drinking	Eating	Lying	Standing-in	Standing-out			
YOLOv5s	0.944	0.995	0.995	0.995	0.994	0.985	73	13.7
YOLOv5m	0.963	0.995	0.995	0.995	0.995	0.988	55	40.2
YOLOv5l	0.956	0.995	0.995	0.994	0.995	0.987	44	88.5

mAP = mean average precision; *FPS* = frames per second.

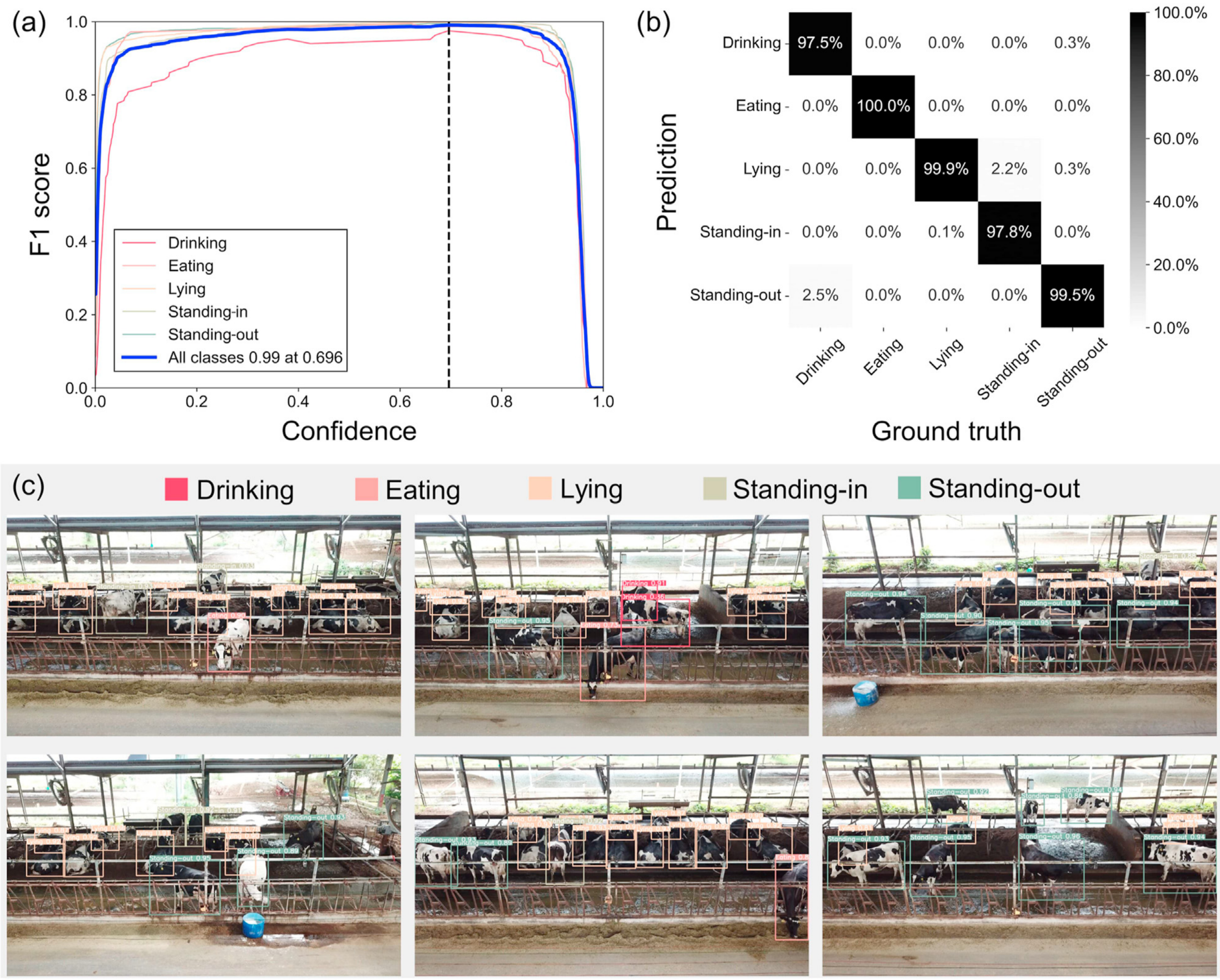

Fig. 3 – Detailed performance of the YOLOv5s model. (a) F1 confidence curve with the black vertical dashed line indicating the best F1 score at a confidence threshold of 0.696. (b) Confusion matrix (normalised by column) with confidence threshold set to 0.696. (c) Example results of behavioural recognition with confidence threshold set to 0.696.

3.2. *Behavioural pattern under heat stress conditions*

As shown in Fig. 5, the 10-day experiment conducted in mid-May well captured the beginning of heat stress with daily mean *Ta* rising from 14.7 to 25.8 °C and daily mean *THI* rising from 58.4 to 74.4. The heat stress threshold for high-producing dairy cows (≥35 kg day^{-1}) has been updated to a daily mean *THI* of 68 by Collier, Laun, Rungruang, and Zimbleman (2012). According to their revised *THI* thresholds, the study herd first stayed within the thermoneutral zone and then experienced mild to moderate heat stress during the 10-day experiment.

The 100% bar chart presents the herd-level behavioural pattern from 10:30 to 15:00 h (Fig. 6). The complete statistics used for plotting can be found in Table S1 (supplementary material). As expected, lying consistently occupied the largest proportion. Lying has long been used for indicating cow comfort and welfare, and cows should spend most idle time lying (Tucker, Jensen, de Passillé, Hänninen, & Rushen, 2021). Besides, cows had the lowest lying% and a relatively high eating% at 10:30 h compared with other scan samples, indicating that the effect of intensive feeding was still lasting 2 h after leaving for the morning milking which was scheduled at 08:30 h. Afterward, the lying% raised dramatically until it peaked at 11:00 h when the majority of cows were resting on their stall beds. The lying% then followed a decreasing trend from 11:00 to 14:30 h, which was right opposite to the rising *Ta* and *THI*. These patterns are consistent with Overton et al. (2002) and Mattachini et al. (2011), who found that a herd needs 2–3 h after leaving for milking to finish feeding and return to rest, and increasing environmental temperature will decrease the lying% during idle time. Therefore, the eight scan samples from 11:00 to 14:30 h were used for further analyses exploring the effect of heat stress on cow behaviour due to a clear relationship between cow behaviour and thermal environment during this period.

The spaghetti plot shows the temporal pattern of six herd-level behavioural indicators with regard to two environmental indicators (Fig. 7). The complete statistics used for plotting can be found in Table S1 (supplementary material). Consistent with lying%, indices describing lying behaviour (i.e., *CI* and

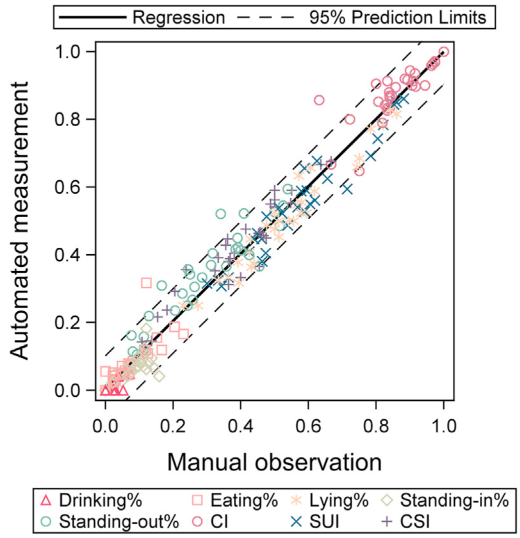

Fig. 4 – Comparison of herd-level behavioural parameters measured half-hourly by manual and automated methods from 10:30 to 15:00 h during the three-day experiment. Drinking% = percentage of cows drinking; Eating % = percentage of cows eating; Lying% = percentage of cows lying; Standing-in% = percentage of cows standing-in; Standing-out% = percentage of cows standing-out; *CI* = comfort index; *SUI* = stall-use index; *CSI* = cow stress index.

SUI) increased to a peak at 11:00 h and followed an overall decreasing trend until 14:30 h. The only difference between lying%, *CI*, and *SUI* was the calculation of the denominator. Neither *CI* nor *SUI* takes eating into account for calculation. An increasing eating% would therefore lower the denominator and finally increase their results. This leads to fluctuations in *CI* and *SUI* from 11:30 to 12:00 h and from 13:30 to 14:00 h. Moreover, *CSI*, as an index describing idle standing behaviour, showed an almost horizontally symmetrical trajectory to *SUI*.

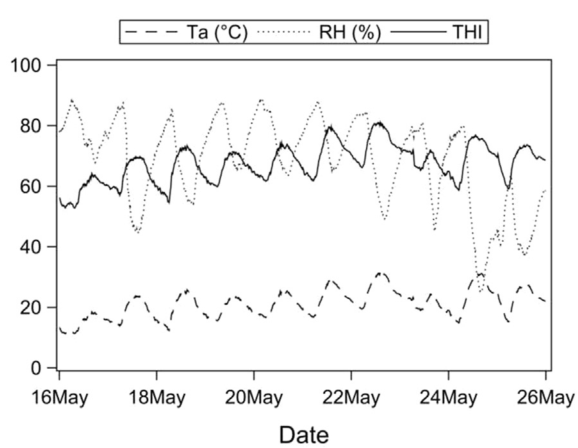

Fig. 5 – Overall variation of indoor ambient temperature (*Ta*, °C), relative humidity (*RH*, %), and temperature-humidity index (*THI*) during the 10-day experiment with a measurement interval of 10 min.

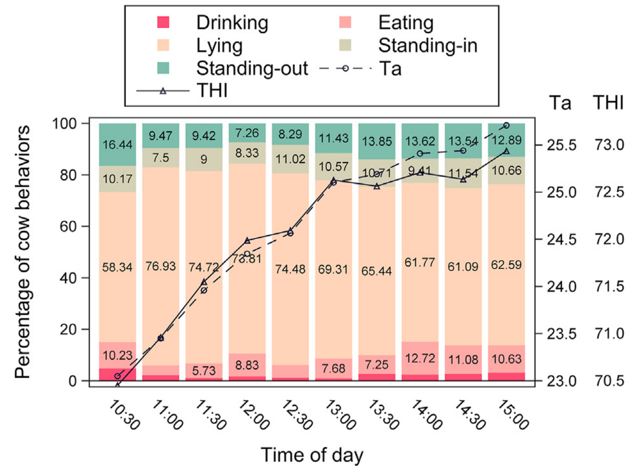

Fig. 6 – Herd-level behaviour distribution measured half-hourly by the proposed automated method, as well as the corresponding ambient temperature (*Ta*, °C) and temperature-humidity index (*THI*), averaged during the 10-day experiment.

These trends of lying%, *SUI*, and *CSI* are comparable with Mattachini et al. (2011) who manually observed cow behaviour through video recording with a 60-min scan sampling. Once again, this demonstrates the effectiveness of our proposed automated method for behaviour recognition.

Generally, lying and standing indices (i.e., lying%, *CI*, *SUI*, and *CSI*) and drinking% had strong correlations with *Ta* and *THI* (all $P < 0.01$), whereas eating% did not appear to have any correlation (both $P > 0.05$) (Fig. 8). Of note, *Ta* had a stronger correlation with cow behaviour than *THI* with exception of drinking% and eating%. A better correlation of *Ta* with animal-based indicators was also observed in our previous study when physiological indicators were used (Shu et al., 2022). To some extent, *Ta* seems to better describe environmental stress than *THI* in this specific environmental condition, probably because the effect of *RH* not yet playing an important role at the very beginning of summer.

Among lying and standing indices, *CI* correlated the weakest with *Ta* ($r = -0.360$, $P < 0.001$) and *THI* ($r = -0.322$, $P = 0.002$), whereas SUI correlated the strongest with *Ta* ($r = -0.744$, $P < 0.001$) and *THI* ($r = -0.673$, $P < 0.001$). Similarly, *CI* was found less susceptible to environmental temperature compared with lying% and *SUI* by Overton, Moore, and Sischo (2003), and the strongest correlation coefficient was found between *SUI* and *Ta* (–0.762) by Mattachini et al. (2011). It is well known that cows will change their drinking and feeding patterns to better coping with heat stress (Kadzere et al., 2002). The positive correlation of drinking% with *Ta* ($r = 0.357$, $P < 0.001$) and *THI* ($r = 0.393$, $P < 0.001$) observed in this study might be attributed to increased visit to the trough and increased time per visit, as previously identified by Tsai et al. (2020) and McDonald et al. (2020). Moreover, eating patterns can change through different strategies by cows with different production stages (Eslamizad, Lamp, Derno, & Kuhla, 2015) or social ranks (Olofsson, 1999). Thus, the small to no effect of heat stress on eating% is probably due to heterogeneous strategies taken among the herd. Indeed, some cows may

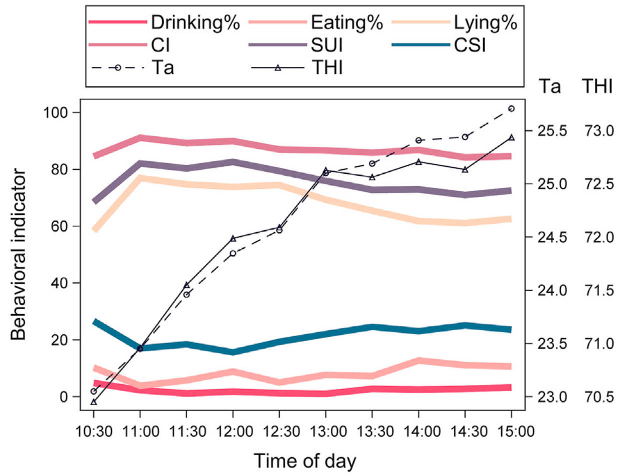

Fig. 7 – Herd-level behavioural indicators measured half-hourly by the proposed automated method, as well as the corresponding ambient temperature (*Ta*, °C) and temperature-humidity index (*THI*), averaged during the 10-day experiment. Drinking% = percentage of cows drinking; Eating% = percentage of cows eating; Lying% = percentage of cows lying; *CI* = comfort index; *SUI* = stall-use index; *CSI* = cow stress index.

reduce their feed intake with a longer eating time (Eslamizad et al., 2015; Herbut et al., 2021). Anyway, eating% alone appears to be insufficient to quantify eating behaviour, and more

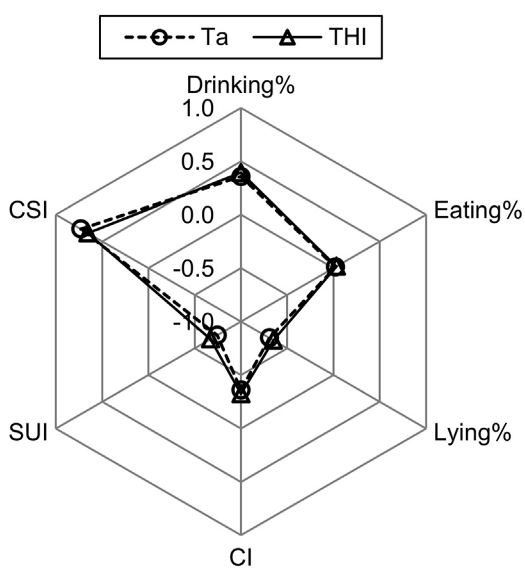

Fig. 8 – Spearman's rank correlation coefficients between herd-level behavioural indicators measured half-hourly by the proposed automated method from 10:30 to 15:00 h during the 10-day experiment and ambient temperature (*Ta*, °C) and temperature-humidity index (*THI*). Drinking % = percentage of cows drinking; Eating% = percentage of cows eating; Lying% = percentage of cows lying; *CI* = comfort index; *SUI* = stall-use index; *CSI* = cow stress index.

data from longer duration or smaller sampling intervals are required to evaluate its ability as a herd-level heat stress indicator.

3.3. *Critical threshold for determining the onset of heat stress*

The basic linear and piecewise models visualised in Fig. 9 show the benefit of using lying%, *SUI*, and *CSI* for indicating the onset of heat stress. Their statistical results presented in Table 4 indicate that the lowest upper critical *Ta* was associated with *SUI* and *CSI* (both 23.8 °C), and the lowest upper critical *THI* was associated with lying% (68.5). The slope differences were always higher with *Ta* than with *THI*, once again suggesting the superiority of *Ta* in representing the stress imposed by the current environment. As for the behavioural indicators that were not converged in the piecewise regression, drinking% and *CI* were positively and negatively related to environmental indicators, respectively, whereas eating% appeared to be independent of *Ta* and *THI*. These results are consistent with those from the correlation analysis.

Several studies have compared cow behaviour in different environmental classes (e.g., *THI* classes), and the upper limit of the class at which significance occurred was stated as a critical point. For example, a *THI* of 68 was determined since the *THI* class <68 had a significantly lower percentage of cows standing compared with other predetermined *THI* classes (Allen, Hall, Collier, & Smith, 2015). However, this method may be arbitrary and can lose information. Moreover, piecewise or segmented models, are used for developing critical thresholds due to their indicative parameters (i.e., slope and breakpoint). For example, Heinicke, Hoffmann, Ammon, Amon, and Amon (2018) determined a *THI* threshold (67) for total lying/standing time, number of lying/standing bouts, and lying bout duration. Our results, however, cannot be compared directly with theirs since *THI* and behavioural indicators were summarised as daily averages in their study.

The advanced piecewise models with *THI* as the predictor all failed to converge. The profiles of the advanced piecewise models with *Ta* as the predictor and lying%, *SUI*, and *CSI* as the outcomes are shown in Figs. 10–12. For each behavioural indicator, *Ta* breakpoints were roughly the same at different times of day, with differences only being observed since the ten thousandth place. The rounded *Ta* breakpoints among times of day were 23.88, 23.70, and 23.65 °C, for lying%, *SUI*, and *CSI*, respectively. The behavioural pattern can be found on the Y-axis with standing increasing and lying decreasing from 11:00 to 14:30 h. According to these findings, accumulated heat load over the observed period of time did not make cows more sensitive to heat stress at a particular time point. In a recent chamber study, the critical *Ta* threshold of respiration rate was found to be lower in the afternoon than in the morning, indicating that cows were more sensitive after longer exposure to heat stress (Zhou, Aarnink, Huynh, van Dixhoorn, & Groot Koerkamp, 2022). Anyway, more data is required to confirm our results since the subgroup sample size used in the analysis might not support a precise localisation of breakpoints. To sum up, any attempt to integrate behavioural indicators for heat stress evaluation should carefully consider the temporal pattern.

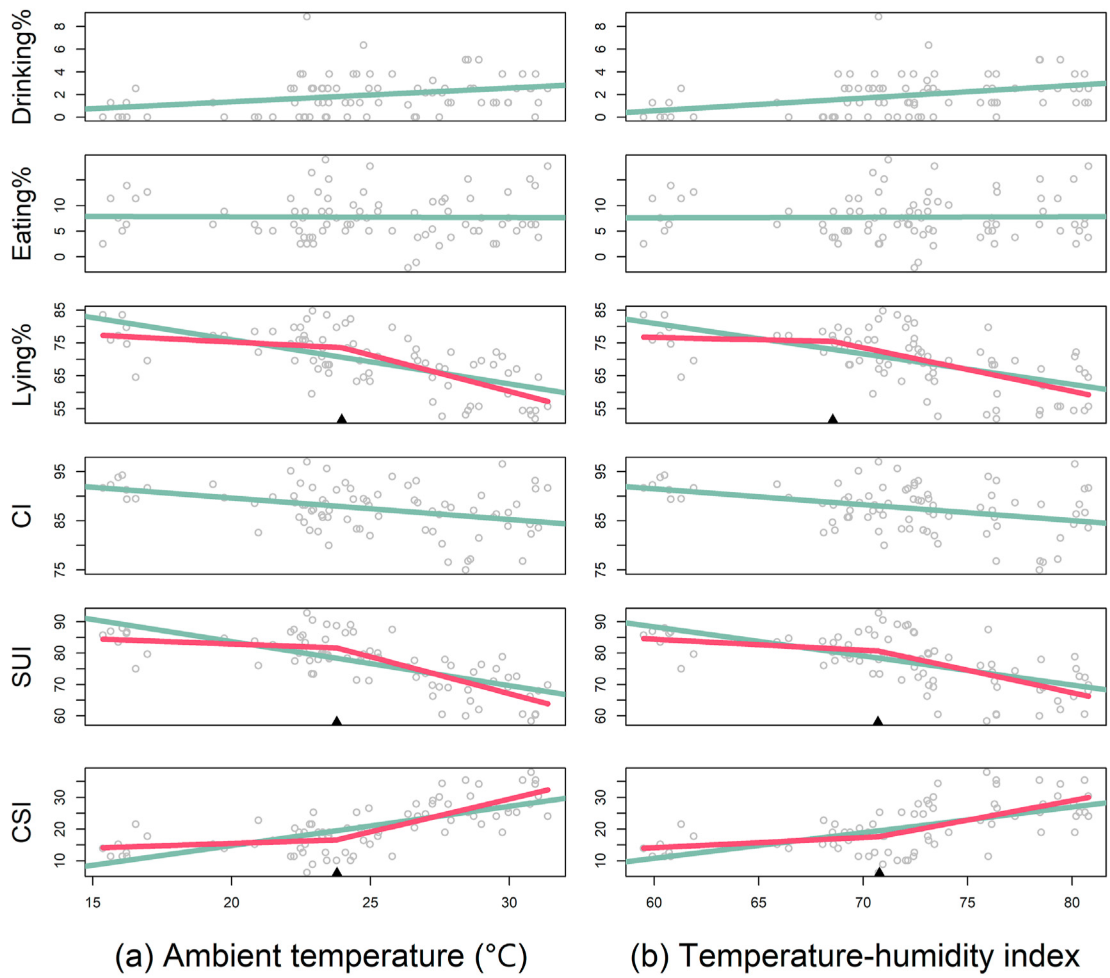

Fig. 9 – Automated measurements of herd-level behavioural indicators and their fitted profiles from linear regression (green) and piecewise regression (red) models with (a) ambient temperature and (b) temperature-humidity index as the predictor, respectively. Breakpoints are marked as a black triangle above the x-axis. Drinking% = percentage of cows drinking; Eating% = percentage of cows eating; Lying% = percentage of cows lying; *CI* = comfort index; *SUI* = stall-use index; *CSI* = cow stress index. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.4. *Strength and limitations*

To the best of our knowledge, this is the first study to determine heat stress based on herd-level behavioural recognition. Although individual measurements can help identify animals with the greatest risks and thereby customise heat abatement, this can be costly in free-stall barns since individual measurement and abatement require a lot of improvement on the existing facilities. In many cases, even if individual measurements have been done, the data have to be summarised to reflect the herd mean (Levit et al., 2021). Indeed, as long as heat abatement is implemented at the herd level, information on the individual level is not necessary for decision making (Winckler, 2019).

With the help of computer vision and scan sampling, our work offers a low-cost herd-level heat stress alert without imposing any burden on dairy cows. Besides, the effectiveness of heat abatement can be tracked by introducing more flexible scan sampling or even continuous measurement when necessary. It should be noted that cows would take advantage of nights when the temperature is thermally comfortable to relieve their accumulated heat load throughout the daytime (Allen et al., 2015). However, such nighttime behavioural data is missing in this study. Future works with reliable nighttime recording and transfer learning techniques are required to develop a behavioural recognition model that can work day and night as well as to customise heat stress thresholds for nighttime hours. By doing so, it is promising to develop a comprehensive protocol for heat stress detection, mitigation, and evaluation.

Our method is best suited for intensive farms where facilities and animals are highly standardised. From the facility perspective, our method for camera mounting can be directly applied in similar settings, but further evaluation in other designs (e.g., 4-row tail-to-tail) is required. In addition, our solution requires fewer cameras than the top-view method (Porto, Arcidiacono, Anguzza, & Cascone, 2015), because it includes more cattle in each view. Top-view cameras, as

Table 4 – Parameter estimates (mean ± SE) of basic piecewise regression models with ambient temperature (*Ta*, °C) and temperature-humidity index (*THI*) as the predictor, respectively. Behavioural indicators were measured half-hourly by the proposed automated method from 11:00 to 14:30 h during the 10-day experiment.[a]

Predictor	Outcome	Intercept	Breakpoint	Left slope	Right slope	AIC	
						Linear	Piecewise
Ta	Drinking%	−1.1 ± 1.1	N/A	0.12 ± 0.04	N/A	309.6	N/A
	Eating%	8.1 ± 2.9	N/A	−0.01 ± 0.12	N/A	464.9	N/A
	Lying%	84.0 ± 7.9	24.0 ± 1.3	−0.43 ± 0.37	−2.21 ± 0.42	534.0	527.8
	CI	98.4 ± 3.0	N/A	−0.44 ± 0.12	N/A	471.3	N/A
	SUI	89.5 ± 6.7	23.8 ± 1.0	−0.33 ± 0.32	−2.35 ± 0.36	516.0	502.2
	CSI	9.7 ± 6.2	23.8 ± 1.0	0.29 ± 0.29	2.07 ± 0.31	499.3	486.4
THI	Drinking%	−6.1 ± 2.4	N/A	0.11 ± 0.03	N/A	306.1	N/A
	Eating%	7.1 ± 6.5	N/A	0.01 ± 0.09	N/A	464.9	N/A
	Lying%	85.2 ± 35.6	68.5 ± 3.3	−0.14 ± 0.56	−1.32 ± 0.23	544.4	542.4
	CI	110.8 ± 6.8	N/A	−0.32 ± 0.09	N/A	472.5	N/A
	SUI	105.5 ± 20.8	70.7 ± 2.1	−0.35 ± 0.31	−1.43 ± 0.27	534.0	530.9
	CSI	−5.1 ± 18.9	70.7 ± 2.3	0.32 ± 0.28	1.23 ± 0.24	517.9	515.8

AIC = Akaike information criterion; Drinking% = percentage of cows drinking; Eating% = percentage of cows eating; Lying% = percentage of cows lying; *CI* = comfort index; *SUI* = stall-use index; *CSI* = cow stress index.

[a] N/A indicates that piecewise regression failed to converge. Intercept and left slope, in this case, represent the parameters of simple linear regressions.

stated before, can also be supplemented in specific areas (e.g. troughs) to provide more reliable recordings. From the animal perspective, cows raised on intensive farms are typically grouped by common influencing factors (e.g., productivity, lactation stage, and parity) and thus behave relatively homogeneously against heat stress. This will help to shrink within-herd variation, allowing abatement measures based on herd means to be useful for the majority of cows. This also allows precision management by customising herd-level heat stress thresholds with respect to different levels of heat sensitivity. Of note, any extrapolation and interpretation of the determined critical thresholds should carefully consider the impact of different management or facility conditions on cow behaviour, such as overstocking, milking frequency, and bedding materials (Hart, McBride, Duffield, & DeVries, 2013; Ito, Chapinal, Weary, & von Keyserlingk, 2014).

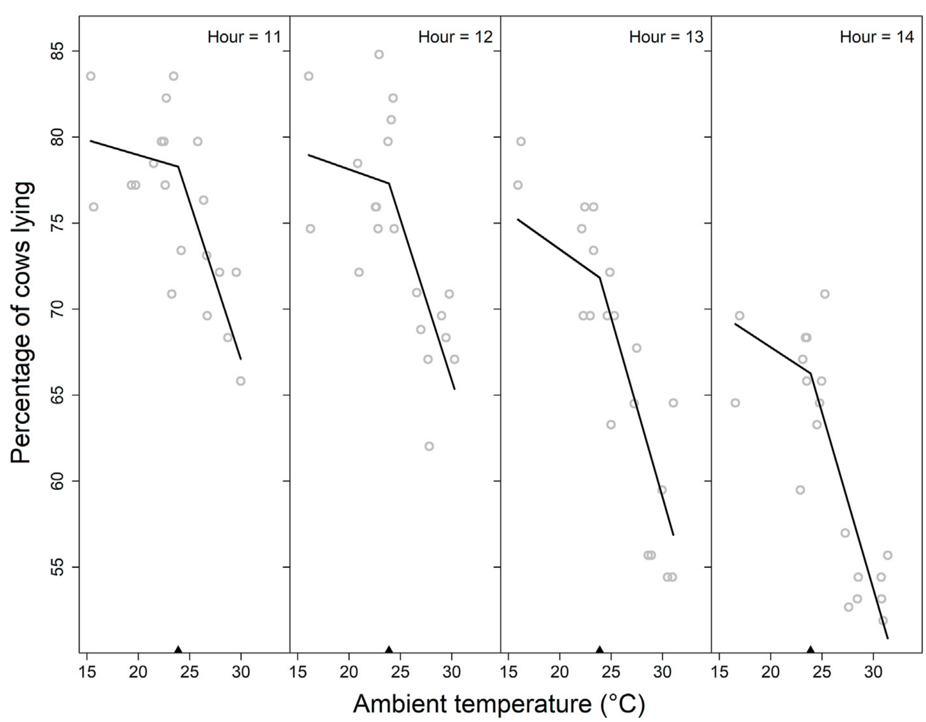

Fig. 10 – Automated measurements of the percentage of cows lying and their fitted profile from advanced piecewise regression with ambient temperature as the predictor. Random effects of time of day (h) were introduced for intercept, slope difference, and breakpoint. Breakpoints are marked as a black triangle above the x-axis.

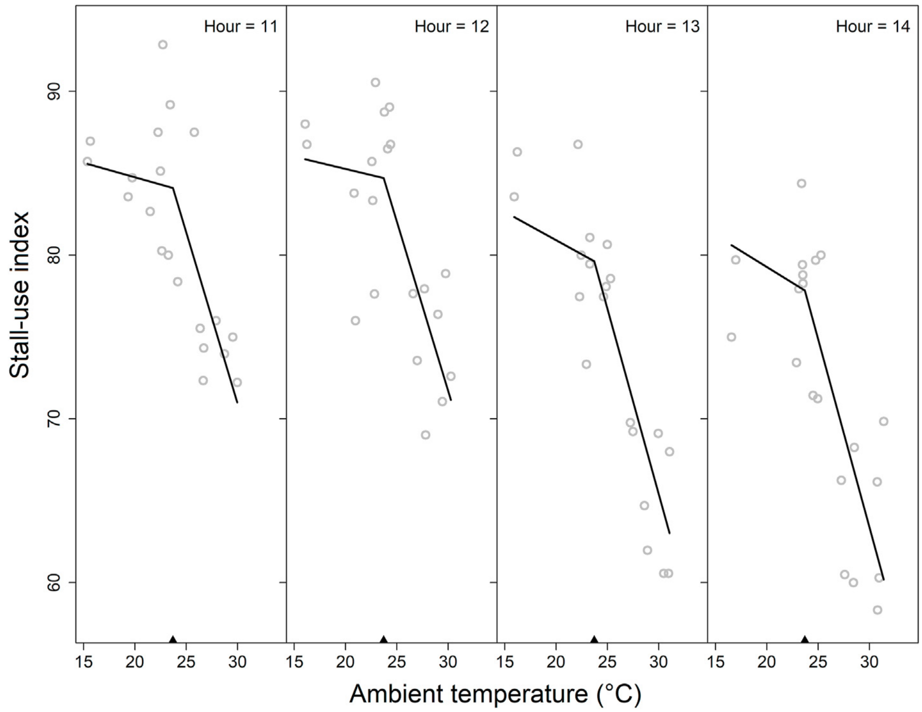

Fig. 11 – Automated measurements of stall-use index and their fitted profile from advanced piecewise regression with ambient temperature as the predictor. Random effects of time of day (h) were introduced for intercept, slope difference, and breakpoint. Breakpoints are marked as a black triangle above the X-axis.

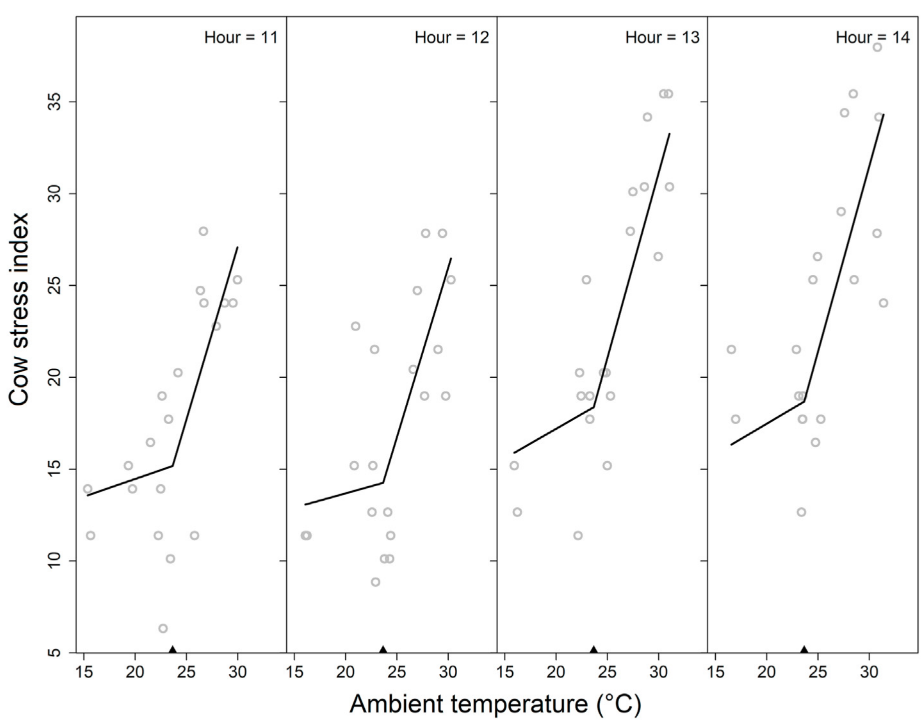

Fig. 12 – Automated measurements of cow stress index and their fitted profile from advanced piecewise regression with ambient temperature as the predictor. Random effects of time of day (h) were introduced for intercept, slope difference, and breakpoint. Breakpoints are marked as a black triangle above the X-axis.

4. Conclusions

This study has proposed a YOLOv5-based method for recognising cow behaviour with excellent *mAP* and inference speed. The ability of the proposed model in measuring herd-level behavioural indicators has been validated in comparison to manual observation. The automated measurements taken during the 10-day experiment reveal that lying and standing indices (i.e., lying%, *SUI*, and *CSI*) were most responding to heat stress and the test dairy herd began to change their behaviour at the earliest *Ta* of 23.8 °C or *THI* of 68.5. Collectively, the model and results presented in this paper have achieved a low-cost heat stress alert for the study herd without imposing any burden on dairy cows. Further study using multiple herds with varying characteristics is promising to customise herd-level heat stress thresholds.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Agricultural Science and Technology Innovation Program [ASTIP-IAS07], the Key Research and Development Project of Heibei Province [22326609D], the Major Science and Technology Program of Inner Mongolia Autonomous Region [2020ZD0004], and the Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences [CAAS-ASTIP-2016-AII]. The authors are grateful to Mr. Fuyu Sun and Mr. Xiaoyang Chen, as well as Yinxiang dairy farm, for their assistance in data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biosystemseng.2023.01.009.

REFERENCES

Allen, J. D., Hall, L. W., Collier, R. J., & Smith, J. F. (2015). Effect of core body temperature, time of day, and climate conditions on behavioral patterns of lactating dairy cows experiencing mild to moderate heat stress. *Journal of Dairy Science, 98*(1), 118–127. https://doi.org/10.3168/jds.2013-7704

Becker, C. A., Collier, R. J., & Stone, A. E. (2020). Invited review: Physiological and behavioral effects of heat stress in dairy cows. *Journal of Dairy Science, 103*(8), 6751–6770. https://doi.org/10.3168/jds.2019-17929

Chen, C., Zhu, W., & Norton, T. (2021). Behaviour recognition of pigs and cattle: Journey from computer vision to deep learning. *Computers and Electronics in Agriculture, 187*, Article 106255. https://doi.org/10.1016/j.compag.2021.106255

Collier, R. J., Dahl, G. E., & VanBaale, M. J. (2006). Major advances associated with environmental effects on dairy cattle. *Journal of Dairy Science, 89*(4), 1244–1253. https://doi.org/10.3168/jds.s0022-0302(06)72193-2

Collier, R. J., Laun, W. H., Rungruang, S., & Zimbleman, R. B. (2012). Quantifying heat stress and its impact on metabolism and performance. In *Paper presented at the Florida Ruminant Nutrition Symposium, Gainesville, FL, USA.*.

Collier, R. J., Renquist, B. J., & Xiao, Y. (2017). A 100-Year Review: Stress physiology including heat stress. *Journal of Dairy Science, 100*(12), 10367–10380. https://doi.org/10.3168/jds.2017-13676

Cook, N. B., Bennett, T. B., & Nordlund, K. V. (2005). Monitoring indices of cow comfort in free-stall-housed dairy herds. *Journal of Dairy Science, 88*(11), 3876–3885. https://doi.org/10.3168/jds.S0022-0302(05)73073-3

Cook, N. B., Mentink, R. L., Bennett, T. B., & Burgi, K. (2007). The effect of heat stress and lameness on time budgets of lactating dairy cows. *Journal of Dairy Science, 90*(4), 1674–1682. https://doi.org/10.3168/jds.2006-634

Eslamizad, M., Lamp, O., Derno, M., & Kuhla, B. (2015). The control of short-term feed intake by metabolic oxidation in late-pregnant and early lactating dairy cows exposed to high ambient temperatures. *Physiology & Behavior, 145*, 64–70. https://doi.org/10.1016/j.physbeh.2015.03.044

Fuentes, A., Yoon, S., Park, J., & Park, D. S. (2020). Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. *Computers and Electronics in Agriculture, 177*, Article 105627. https://doi.org/10.1016/j.compag.2020.105627

Galán, E., Llonch, P., Villagrá, A., Levit, H., Pinto, S., & Del Prado, A. (2018). A systematic review of non-productivity-related animal-based indicators of heat stress resilience in dairy cattle. *PLoS One, 13*(11). https://doi.org/10.1371/journal.pone.0206520. e0206520-e0206520.

Hales, T. C. (2007). Jordan's proof of the Jordan curve theorem. *Studies in Logic, Grammar and Rhetoric, 10*(23), 45–60.

Hart, K. D., McBride, B. W., Duffield, T. F., & DeVries, T. J. (2013). Effect of milking frequency on the behavior and productivity of lactating dairy cows. *Journal of Dairy Science, 96*(11), 6973–6985. https://doi.org/10.3168/jds.2013-6764

Heinicke, J., Hoffmann, G., Ammon, C., Amon, B., & Amon, T. (2018). Effects of the daily heat load duration exceeding determined heat load thresholds on activity traits of lactating dairy cows. *Journal of Thermal Biology, 77*, 67–74. https://doi.org/10.1016/j.jtherbio.2018.08.012

Herbut, P., Hoffmann, G., Angrecka, S., Godyń, D., Vieira, F. M. C., Adamczyk, K., et al. (2021). The effects of heat stress on the behaviour of dairy cows – a review. *Annals of Animal Science, 21*(2), 385–402. https://doi.org/10.2478/aoas-2020-0116

Ito, K., Chapinal, N., Weary, D. M., & von Keyserlingk, M. A. G. (2014). Associations between herd-level factors and lying behavior of freestall-housed dairy cows. *Journal of Dairy Science, 97*(4), 2081–2089. https://doi.org/10.3168/jds.2013-6861

Kadzere, C. T., Murphy, M. R., Silanikove, N., & Maltz, E. (2002). Heat stress in lactating dairy cows: A review. *Livestock Production Science, 77*(1), 59–91. https://doi.org/10.1016/S0301-6226(01)00330-X

Levit, H., Pinto, S., Amon, T., Gershon, E., Kleinjan-Elazary, A., Bloch, V., et al. (2021). Dynamic cooling strategy based on individual animal response mitigated heat stress in dairy cows. *Animal, 15*(2), Article 100093. https://doi.org/10.1016/j.animal.2020.100093

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). Microsoft coco: Common objects in context. In *Paper presented at the European conference on computer vision, 8693*, 740–755.

Lv, J., Xu, H., Han, Y., Lu, W., Xu, L., Rong, H., et al (2022). A visual identification method for the apple growth forms in the

orchard. *Computers and Electronics in Agriculture, 197*, Article 106954. https://doi.org/10.1016/j.compag.2022.106954
Mattachini, G., Riva, E., & Provolo, G. (2011). The lying and standing activity indices of dairy cows in free-stall housing. *Applied Animal Behaviour Science, 129*(1), 18–27. https://doi.org/10.1016/j.applanim.2010.10.003
McDonald, P. V., von Keyserlingk, M. A. G., & Weary, D. M. (2020). Hot weather increases competition between dairy cows at the drinker. *Journal of Dairy Science, 103*(4), 3447–3458. https://doi.org/10.3168/jds.2019-17456
Mitloehner, F., Morrow-Tesch, J., Wilson, C., Dailey, J., & McGlone, J. (2001). Behavioral sampling techniques for feedlot cattle. *Journal of Animal Science, 79*, 1189–1193. https://doi.org/10.2527/2001.7951189x
Muggeo, V. M. (2008). Segmented: an R package to fit regression models with broken-line relationships. *R News, 8*(1), 20–25.
NRC. (1971). *A guide to environmental research on animals*. Retrieved from Washington, DC, USA.
Olofsson, J. (1999). Competition for total mixed diets fed for ad libitum intake using one or four cows per feeding station. *Journal of Dairy Science, 82*(1), 69–79. https://doi.org/10.3168/jds.S0022-0302(99)75210-0
Overton, M. W., Moore, D. A., & Sischo, W. M. (2003). Comparison of commonly used indices to evaluate dairy cattle lying behavior. In *Paper presented at the Fifth International Dairy Housing Conference for 2003*.
Overton, M. W., Sischo, W. M., Temple, G. D., & Moore, D. A. (2002). Using time-lapse video photography to assess dairy cattle lying behavior in a free-stall barn. *Journal of Dairy Science, 85*(9), 2407–2413. https://doi.org/10.3168/jds.S0022-0302(02)74323-3
Porto, S. M. C., Arcidiacono, C., Anguzza, U., & Cascone, G. (2013). A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. *Biosystems Engineering, 115*(2), 184–194. https://doi.org/10.1016/j.biosystemseng.2013.03.002
Porto, S. M. C., Arcidiacono, C., Anguzza, U., & Cascone, G. (2015). The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system. *Biosystems Engineering, 133*, 46–55. https://doi.org/10.1016/j.biosystemseng.2015.02.012
Shu, H., Guo, L., Bindelle, J., Fang, T., Xing, M., Sun, F., et al. (2022). Evaluation of environmental and physiological indicators in lactating dairy cows exposed to heat stress. *International Journal of Biometeorology, 66*(6), 1219–1232. https://doi.org/10.1007/s00484-022-02270-w
Shu, H., Wang, W., Guo, L., & Bindelle, J. (2021). Recent advances on early detection of heat strain in dairy cows using animal-based indicators: A review. *Animals, 11*(4), 980. https://doi.org/10.3390/ani11040980
Silanikove, N. (2000). Effects of heat stress on the welfare of extensively managed domestic ruminants. *Livestock Production Science, 67*(1), 1–18. https://doi.org/10.1016/S0301-6226(00)00162-7
Tsai, Y.-C., Hsu, J.-T., Ding, S.-T., Rustia, D. J. A., & Lin, T.-T. (2020). Assessment of dairy cow heat stress by monitoring drinking behaviour using an embedded imaging system. *Biosystems Engineering, 199*, 97–108. https://doi.org/10.1016/j.biosystemseng.2020.03.013
Tucker, C. B., Jensen, M. B., de Passillé, A. M., Hänninen, L., & Rushen, J. (2021). Invited review: Lying time and the welfare of dairy cows. *Journal of Dairy Science, 104*(1), 20–46. https://doi.org/10.3168/jds.2019-18074
Uzal Seyfi, S. (2013). Seasonal variation of the lying and standing behavior indexes of dairy cattle at different daily time periods in free-stall housing. *Animal Science Journal, 84*(10), 708–717. https://doi.org/10.1111/asj.12062
Winckler, C. (2019). Assessing animal welfare at the farm level: Do we care sufficiently about the individual? *Animal Welfare, 28*(1), 77–82. https://doi.org/10.7120/09627286.28.1.077
Wu, D., Wang, Y., Han, M., Song, L., Shang, Y., Zhang, X., et al. (2021). Using a CNN-LSTM for basic behaviors detection of a single dairy cow in a complex environment. *Computers and Electronics in Agriculture, 182*, Article 106016. https://doi.org/10.1016/j.compag.2021.106016
Xiao, J., Liu, G., Wang, K., & Si, Y. (2022). Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM. *Computers and Electronics in Agriculture, 194*, Article 106738. https://doi.org/10.1016/j.compag.2022.106738
Yang, G., Feng, W., Jin, J., Lei, Q., Li, X., Gui, G., et al. (2020). Face mask recognition system with YOLOV5 based on image recognition. In *Paper presented at the 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*.
Zhou, M., Aarnink, A. J. A., Huynh, T. T. T., van Dixhoorn, I. D. E., & Groot Koerkamp, P. W. G. (2022). Effects of increasing air temperature on physiological and productive responses of dairy cows at different relative humidity and air velocity levels. *Journal of Dairy Science, 105*(2), 1701–1716. https://doi.org/10.3168/jds.2021-21164