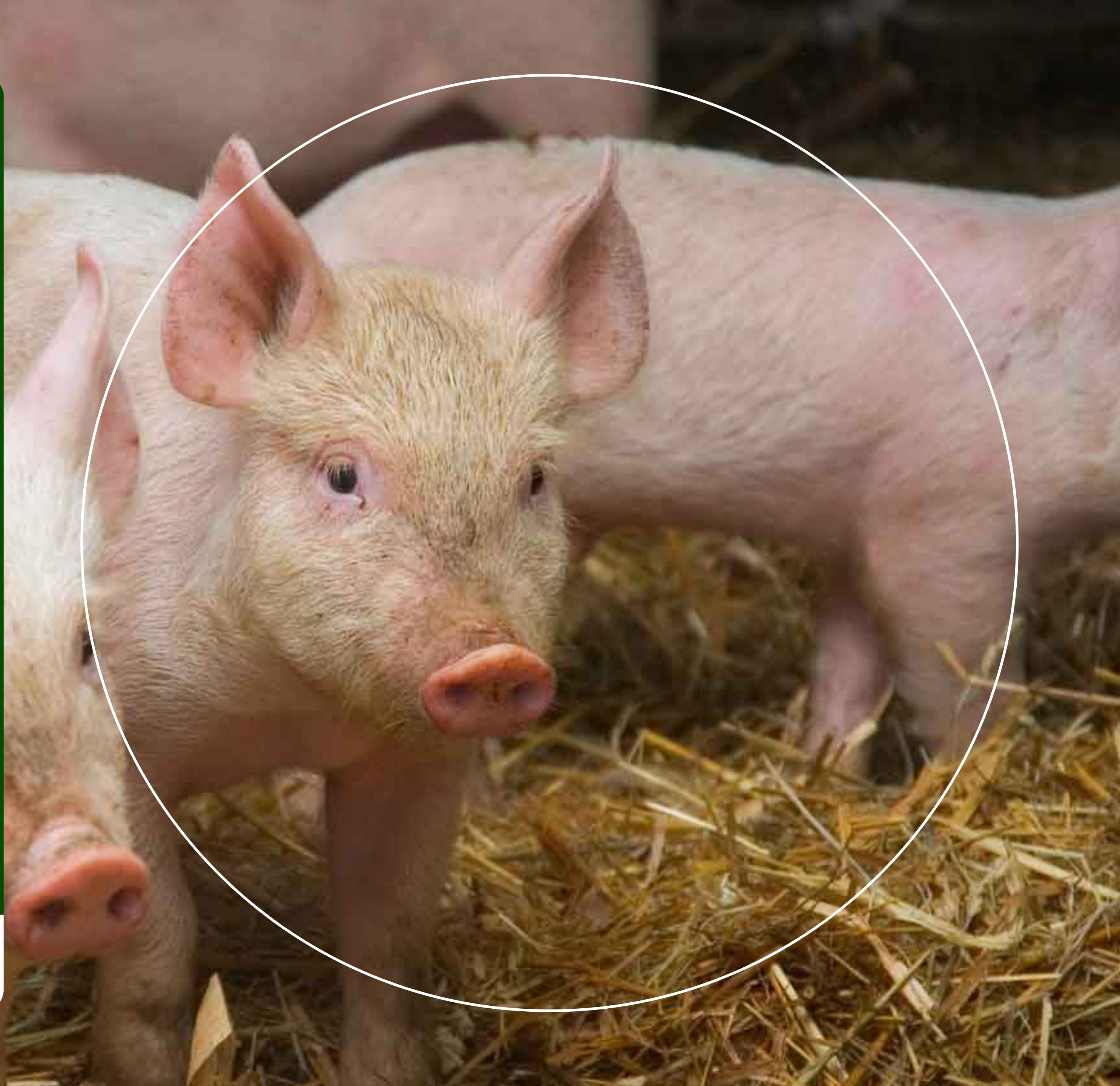
Improved detection of pig–enrichment interactions using time-windowed machine learning

Hadi Taghvafard, Bing Han, Fleur Veldkamp, Rudi M. de Mol, Ingrid C. de Jong, Dirkjan Schokker, Michel J. Counotte

Wageningen Bioveterinary Research
Wageningen Livestock Research

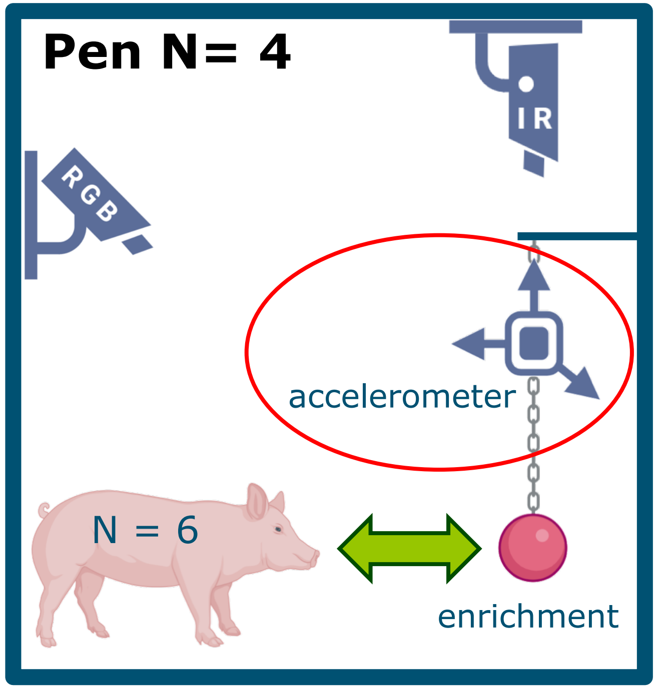


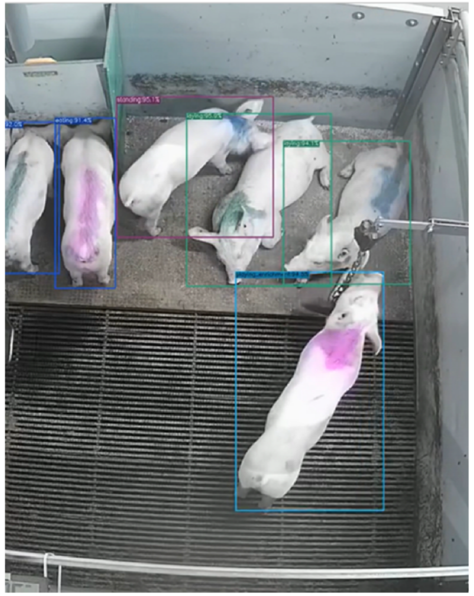
Background

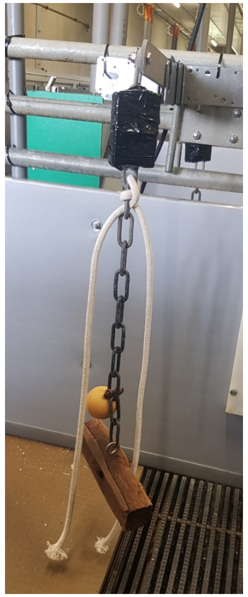
Why monitor interaction with enrichment?

- Behavioural changes can be early indicators of health and welfare in pigs.
- Interaction with enrichment material can reflect interest in the environment and may be relevant for tail-biting risk.
- Manual observation is informative, but non-continuous and bias.
- Non-invasive sensors can provide continuous and objective behavioural monitoring.

Original sensor study







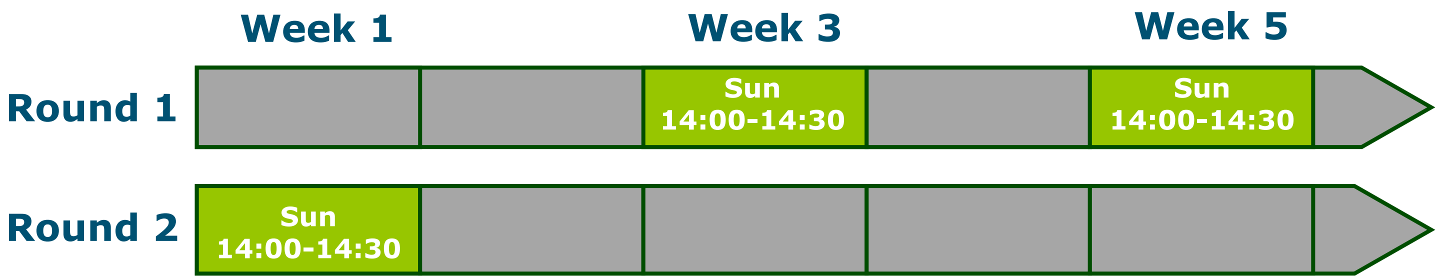
Contents lists available at ScienceDirect
Applied Animal Behaviour Science
journal homepage: www.elsevier.com/locate/applanim

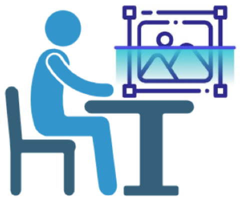
Validation of non-invasive sensor technologies to measure interaction with enrichment material in weaned fattening pigs

Fleur Veldkamp[a,b,*], Tomas Izquierdo Garcia-Faria[a], Vivian L. Witjes[c], Johanna M.J. Rebel[b], Ingrid C. de Jong[a]

[a] Wageningen Livestock Research, Wageningen University and Research, De Elst 1, NL-6708 WD Wageningen, the Netherlands
[b] Adaptation Physiology Group, Wageningen University and Research, De Elst 1, NL-6708 WD Wageningen, the Netherlands
[c] Farm Animal Health, Department Population Health Sciences, Veterinary Medicine, Utrecht University, Yalelaan 7, NL-3584 CL Utrecht, the Netherlands

- Semi-commercial farm, weaned pigs.
- Accelerometer: X,Y, Z values > threshold = interaction (per second).
- Manual video scoring (per second) = gold standard.





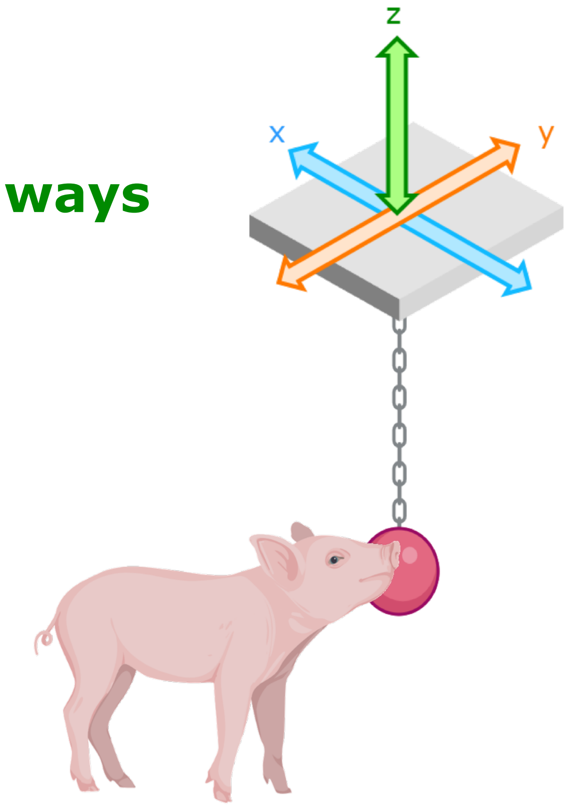
Shake
Carry
Nose
Bite
Chew
Root
Lie
Body
No action

- Intentional interactions
- All interactions

Original analysis and limitation

The accelerometer detected movement, but movement was not always the interaction

- Performance for intentional interaction was moderate
- Main problem: chain movement can come from true interaction, accidental contact or residual swinging.



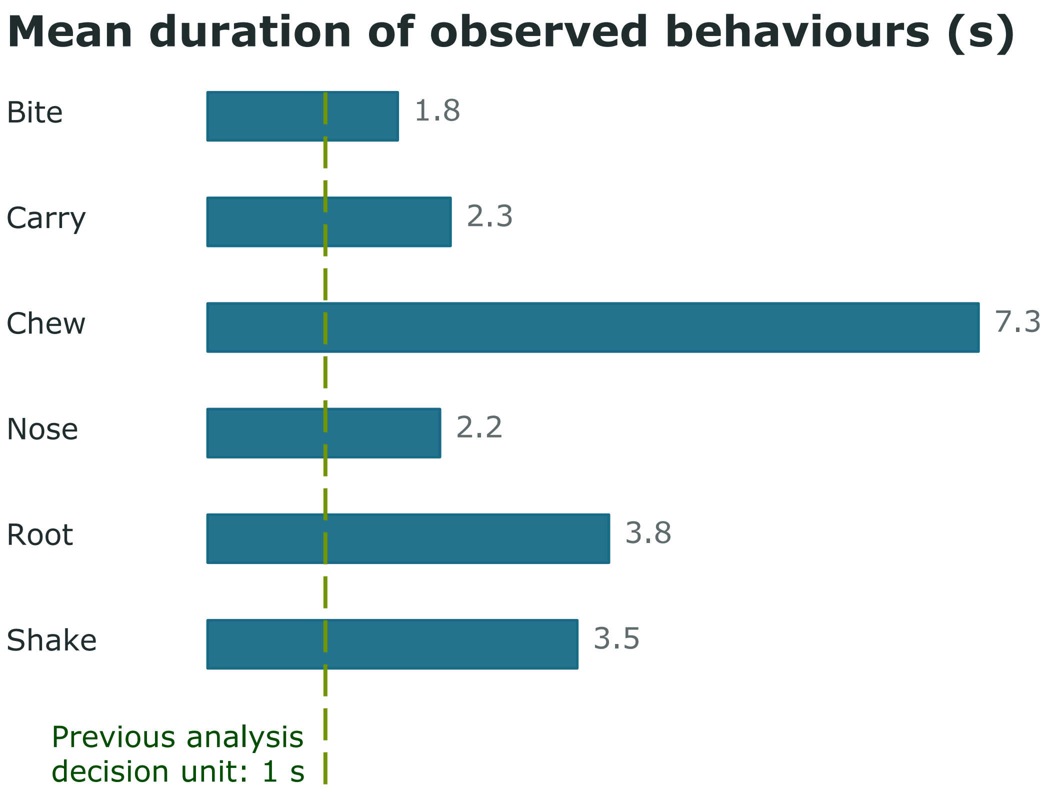
	F1	MCC
X-axis	0.482	0.345
Y-axis	0.524	0.401
Z-axis	0.465	0.320
XYZ	0.474	0.333

One sensor signal can mean different things

- True interaction → chain moves → Accelerometer signal
- Accidental touch → chain moves → Accelerometer signal
- Residual swing → chain moves → Accelerometer signal

Behaviour is not a single isolated second

- Pig–enrichment interaction unfold over short periods.
- But the approach judged each second separately.
- A single second may miss the pattern of an interaction.
- Time windows can better match the duration of real interaction.



Hypothesis and aim

Can the same sensor become more useful with better data processing?

Hypothesis:

Time-window-based signal features capture interaction dynamics better than point-in-time accelerometer values.

Aim:

Improve detection of true pig–enrichment interaction from enrichment-mounted accelerometer data.

Single-second value
limited context

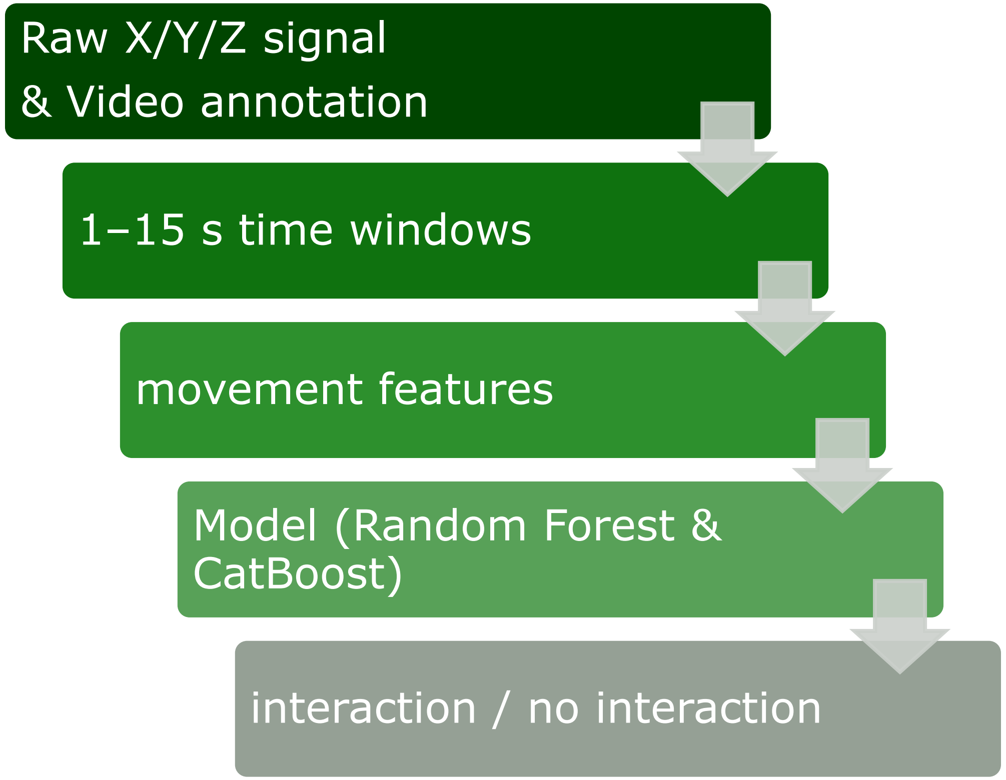
→

Short time window
movement pattern over time

Analytical approach

From sensor data to labelled time windows

- Re-used accelerometer data and manual video annotations.
- Synchronise accelerometer data with manual video annotation.
- Divided X/Y/Z signals into 1–15 s time windows.
- Extracted signal-level features from each window.
- Model trained and evaluated.
  - Tree-based ensemble models performed best in initial testing
  - Final comparison focused on Random Forest and CatBoost



Feature extraction

What did we extract from each time window (1 – 15s)?

We transformed raw X/Y/Z accelerometer values into movement of the enrichment, not yet which behaviour the pig performed.

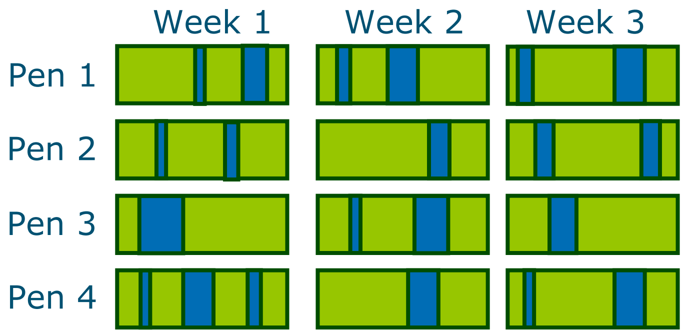
Movement descriptions used by the model

Question	Example features
How much did the enrichment move?	vector magnitude, max_vm, energy, SMA
How irregular was the movement?	sd_x, sd_y, sd_z, sd_vm
How did the chain move in space?	roll, pitch, yaw
Was the movement repeated (rhythm)?	frequency features, entropy, autocorrelation

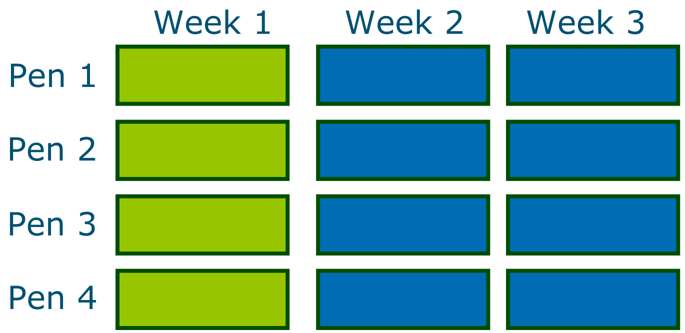
Training-testing scenarios

How the model could be trained in a realistic animal study?

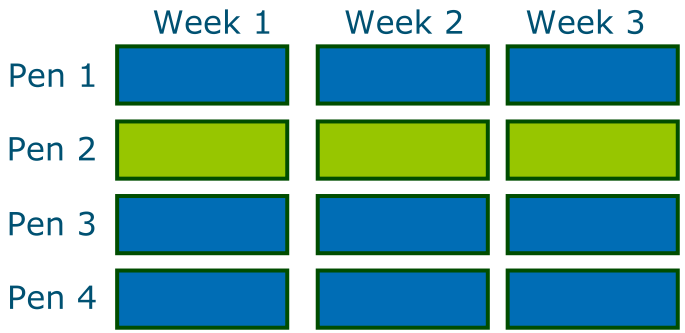
Scenario 1 (Random 70% → 30%)
- Internal benchmark performance



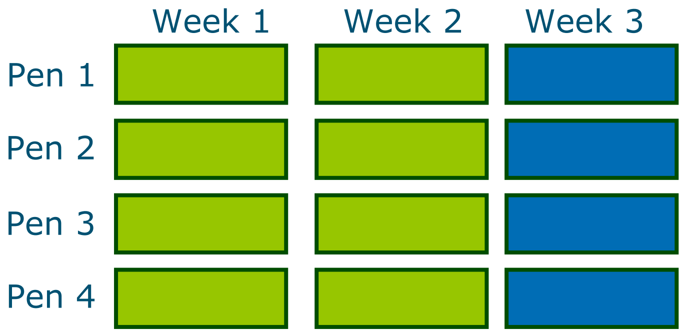
Scenario 2 (wk1 → wk2,3)
- Can one early annotated dataset predict later observations?



Scenario 3 (one pen → rest pens)
- Can one representative pen train the model?



Scenario 4 (wk1,2 → wk3)
- Does more early annotated data improve later prediction?



train → test

How did we evaluate model performance?

- The models tried to classify each sensor data time window as either true intentional interaction (shake, carry, nose, bite, chew or root) or no interaction.

Metric	Meaning	How to read it
Accuracy	How often the model is correct overall	Higher = better (can be misleading when “no interaction” is common)
Sensitivity	Shows whether the model misses real pig–enrichment interactions	High sensitivity = fewer missed interactions
Specificity	Shows whether the model avoids false alarms	High specificity = fewer false interaction
PPV	Shows how trustworthy a positive prediction is	High PPV = predicted interactions are more reliable
F1	Can the model detect real interactions, while avoiding too many false interaction. (Sensitivity & PPV)	Higher F1 = better balance for detecting the positive class (Useful when interaction events are less frequent than no-interaction events)
MCC	Shows how well the model separates interaction from no-interaction, considering both missed interactions and false alarms.	Higher MCC = model detects interactions while also correctly ignoring no-interaction periods

Results

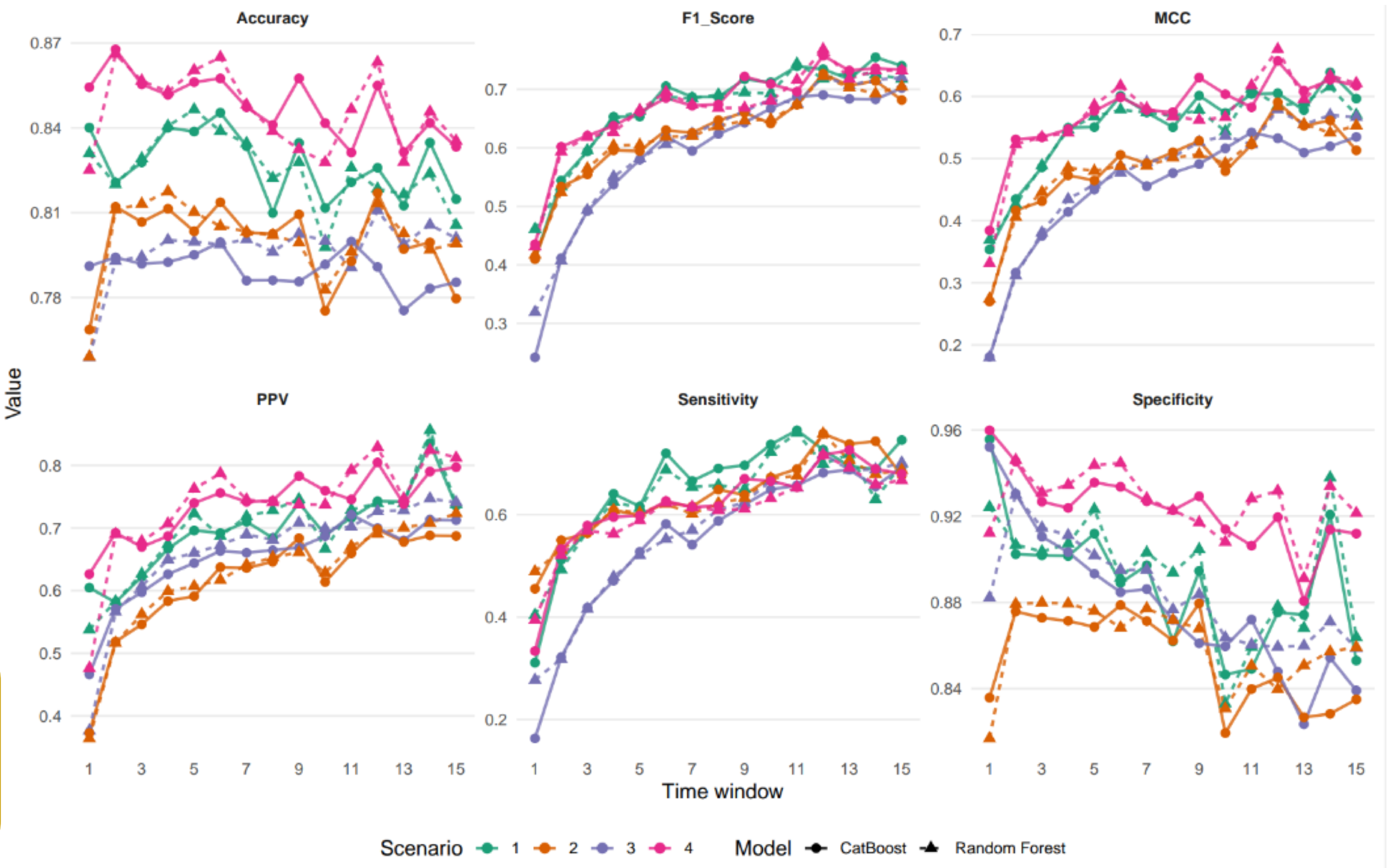
Time-window length and training scenario both affected performance

Time-windowed effect

- Sensitivity, PPV, F1 and MCC showed better overall performance with longer windows (~5–6 s)
- Specificity decreased with longer windows
- Accuracy fluctuated; no clear time-window benefit

Training-scenario effect

- S1: useful benchmark, less realistic for deployment
- S2: early-only training insufficient (but different production round played a role?)
- S3: cross-pen transfer challenging
- S4: strongest practical calibration-start strategy



- Time-windowed features improved behaviour detection
- Early calibration (annotation) within the same study context is promising

Results

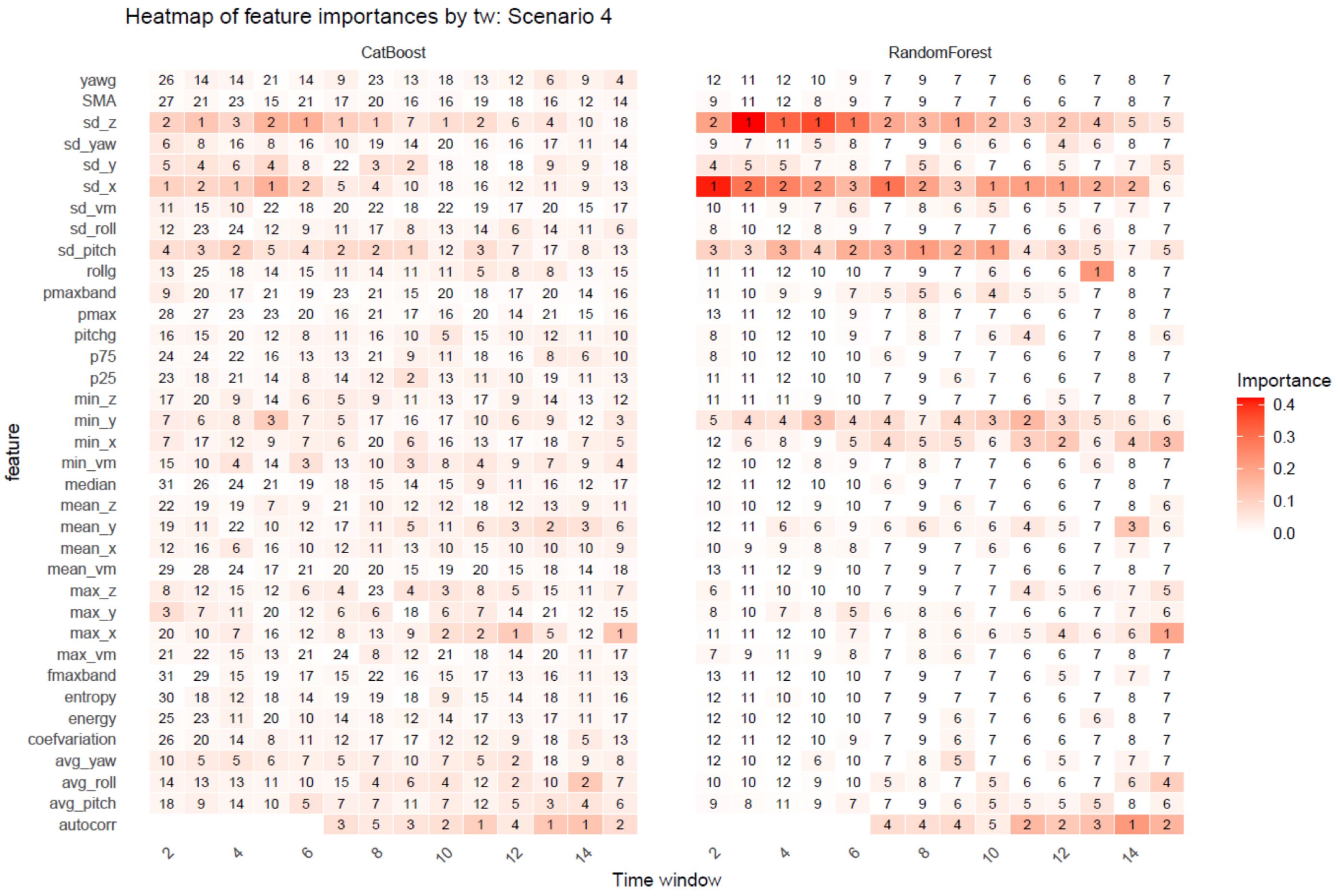
Which movement features helped detect interaction?

Permutation-based feature importance

- One feature shuffled at a time
- Importance = drop in model accuracy

Feature importance

- Different key features between models
- Variability features consistently important: sd_x, sd_z
- max_x more important with longer time windows
- CatBoost used broader feature set than Random Forest (tree depth setting).



True interaction is better captured by movement patterns than by one raw acceleration value

Conclusion

Time-windowed signal features improved detection of pig–enrichment interaction compared with point-in-time analysis.

- Very short windows miss behavioural structure.
- Longer windows improve sensitivity but can include residual movement (swinging).
- Some early group-specific annotation may be needed.
- Future work should improve detection of brief interactions and test long-term robustness.

The sensor is only as useful as the data-processing workflow behind it.

Acknowledgement

Wageningen Livestock Research

Fleur Veldkamp

Rudi M. de Mol

Ingrid C. de Jong

Wageningen Bioveterinary Research

Hadi Taghvafard

Bing Han

Dirkjan Schokker

Michel J. Counotte

Funding

- The Dutch ministry of agriculture fishery and nature (LVVN) KB-56-001-016 (AI DRIVEN INNOVATIONS)
- LVVN under the Wageningen University & Research Knowledge Base Programme on Future Food Systems KB-53-000-001 (AI tools to enable high-throughput phenotyping for breeding and health monitoring).

Thank you!

bing.han@wur.nl

Additional slides

Definition of different behaviours

Behaviour	Description
Shake	While holding the object in its mouth, the animal energetically moves the enrichment material from side to side using its neck and head.
Carry	Animal securely holds the enrichment material in its mouth, while moving in a forward/backward/sideward direction.
Nose	Animal moves snout along or close to enrichment material, without holding the enrichment material in its mouth.
Bite	Animal bites the enrichment material once, without keeping the enrichment material in its mouth.
Chew	Animal chews on the enrichment material, without moving the enrichment material forward/backward/sideward.
Root	Animal nudges or lifts enrichment material with movement of the snout, without keeping the enrichment material in its mouth.
Lie	Physical contact with the enrichment material other than mouth or snout (limbs, body, etc.) while lying down.
Body	Physical contact with the enrichment material other than mouth or snout (limbs, body, etc.) while standing/walking/running.
No action	No physical contact with the enrichment material.

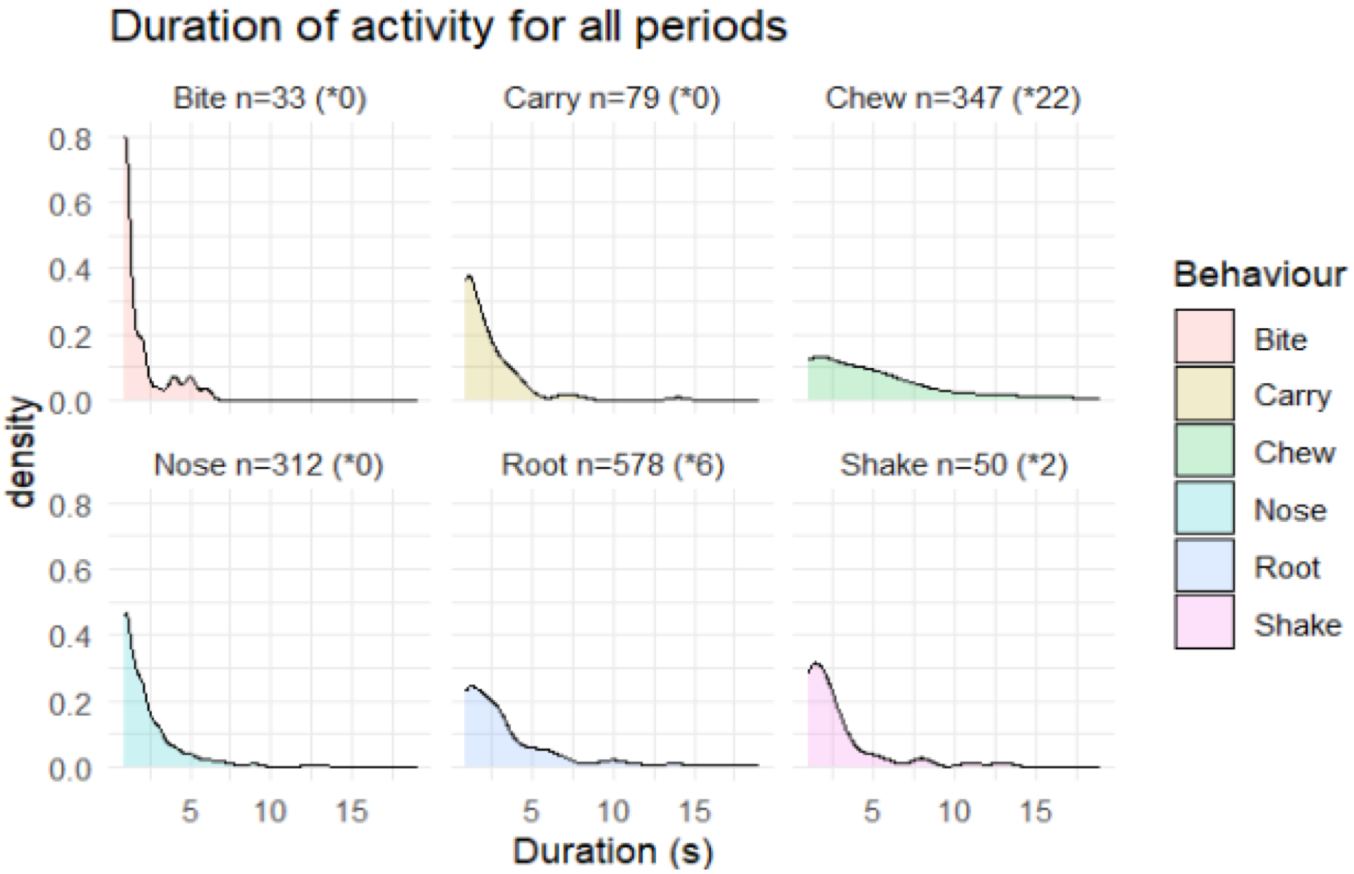
Number of observations per outcome when considering different time-window lengths (time-window).

Time-window (s)	Time-windows with ‘ true interaction’	Total time-windows	Proportion time-windows with ‘true interaction’
1	4030	21600	0.19
2	2216	10800	0.21
3	1601	7200	0.22
4	1279	5400	0.24
5	1081	4320	0.25
6	945	3600	0.26
7	838	3084	0.27
8	771	2700	0.29
9	706	2400	0.29
10	663	2160	0.31
11	620	1956	0.32
12	589	1800	0.33
13	558	1656	0.34
14	524	1536	0.34
15	506	1440	0.35

Density plots of the duration of the different behaviours.

- The plots are truncated at a duration of 20 seconds.
- n = total number of instances.
- The number after the asterisk between brackets = number of instances > 20 seconds.

Behaviour	Count	Total (s)	Median (s)	Mean (s)	SD (s)	Max (s)
Bite	33	59	1	1.8	1.4	6
Carry	79	179	2	2.3	2.0	14
Chew	347	2519	4	7.3	10.7	109
Nose	312	700	2	2.2	1.9	13
Root	578	2171	3	3.8	3.9	34
Shake	50	176	2	3.5	4.6	22



List of features derived from accelerometers

Parameter	Description	N
Basic statistics		
Min_x, min_y, min_z	Minimum of x, y, z	3
Max_x, max_y, max_z	Maximum of x, y, z	3
Mean_x, mean_y, mean_z	Mean of x, y, z	3
Sd_x, sd_y, sd_z	Standard deviation of x, y, z	3
Vector magnitude derived summaries (VM: v = (x^2 + y^2 + z^2)^1/2)		
mean_vm	Mean vector magnitude (VM) of acceleration	1*
sd_vm	Standard deviation of VM	1*
coefvariation	Coefficient of variation of VM	1*
min_vm	Minimum VM	1*
max_vm	Maximum VM	1*
25thp, median, 75thp	25th, 50th (median), and 75th percentiles of VM	3*
autocorr	Autocorrelation of VM with lag of 1 second	1*
Energy and absolute sum		
energy	energy of an accelerometer signal over a time-window (sum of x^2 + y^2 + z^2 across time-window)	1
SMA	sum of absolute values (x,y,z) over time-window	1
Rotational features and derived summaries (roll = tan^-1(y, z), pitch = tan^-1(x, z), and yaw = tan^-1 (y, x))		
avg_roll, avg_pitch, avg_yaw	Mean roll, pitch, and yaw angles	3
sd_roll, sd_pitch, sd_yaw	Standard deviation of roll, pitch, and yaw angles	3*
rollg, pitchg, yawg	Roll, pitch, and yaw angles of estimated gravity direction	3*
Fourier transformation of VM		
pmax	Maximum power in the power spectrum	1
fmaxband, pmaxband	Dominant frequency and power in the 0.3-3 Hz band	2
entropy	Spectral entropy of the frequency-domain signal	1

* features are not calculated for a 1 second time-window, since they require more than one observation.

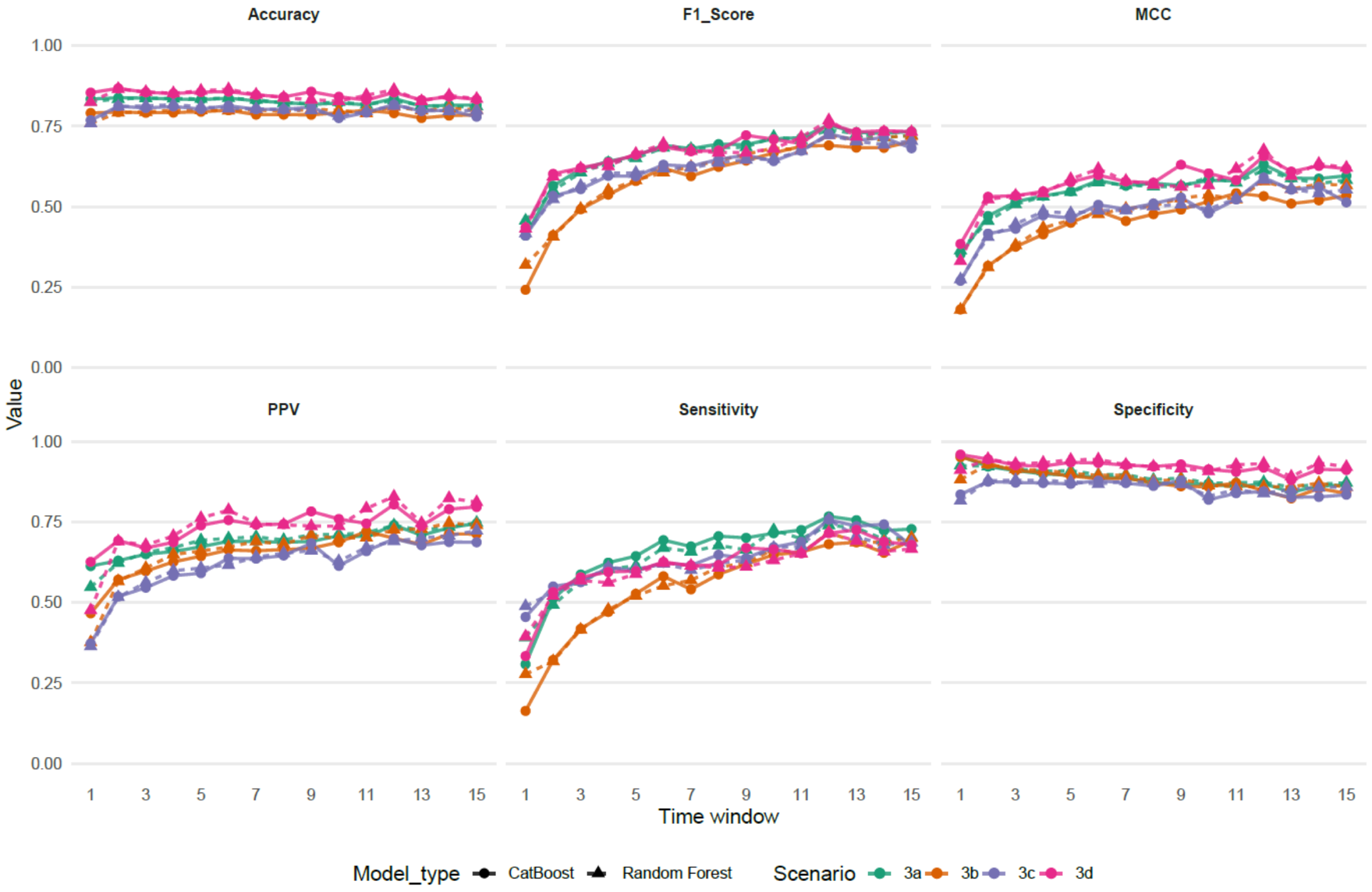
N=Number of features.

Comparison of model performance for scenario 3

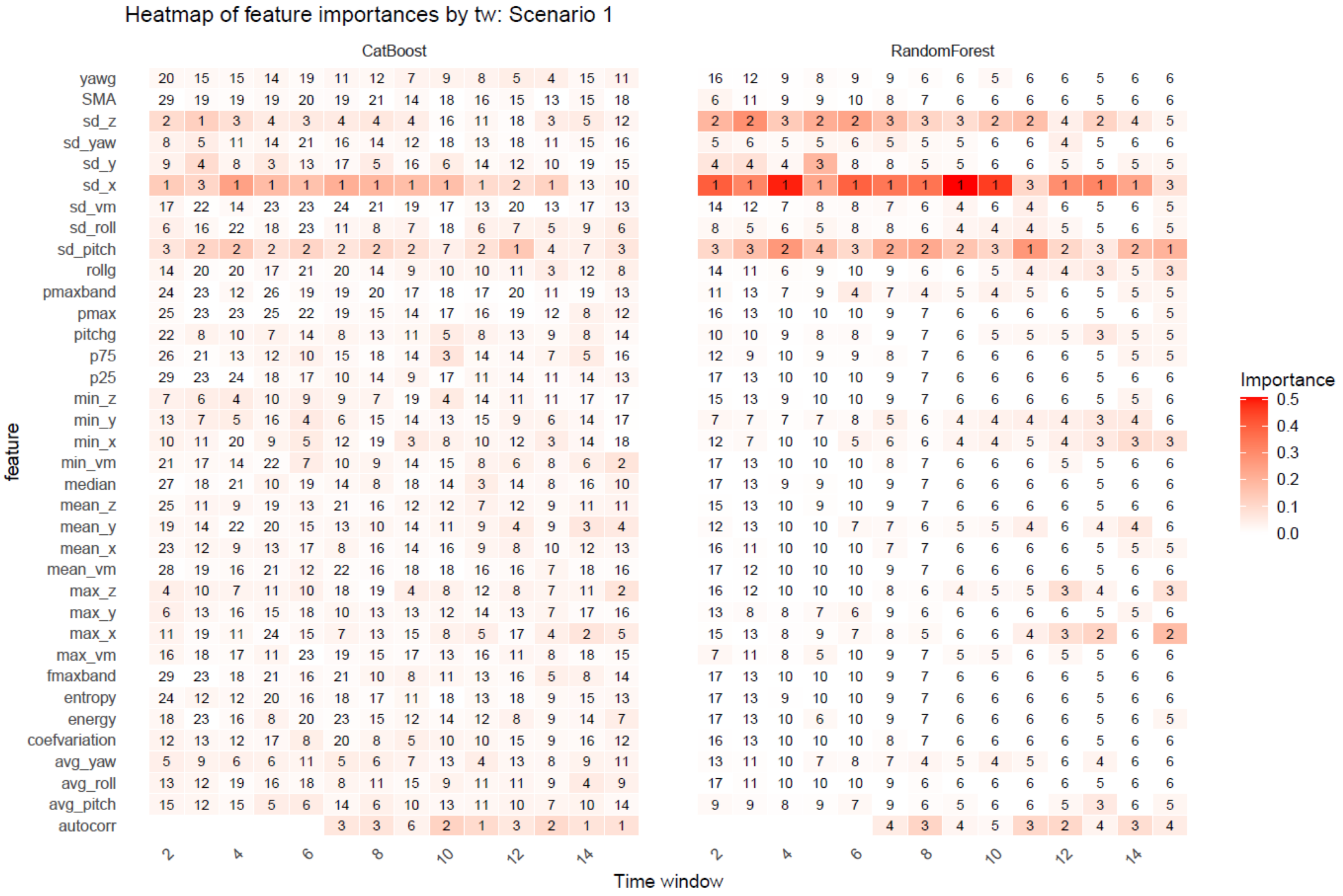
There was moderate variation comparing the selection of other pens as training set

Scenario 3 (one pen → rest pens)

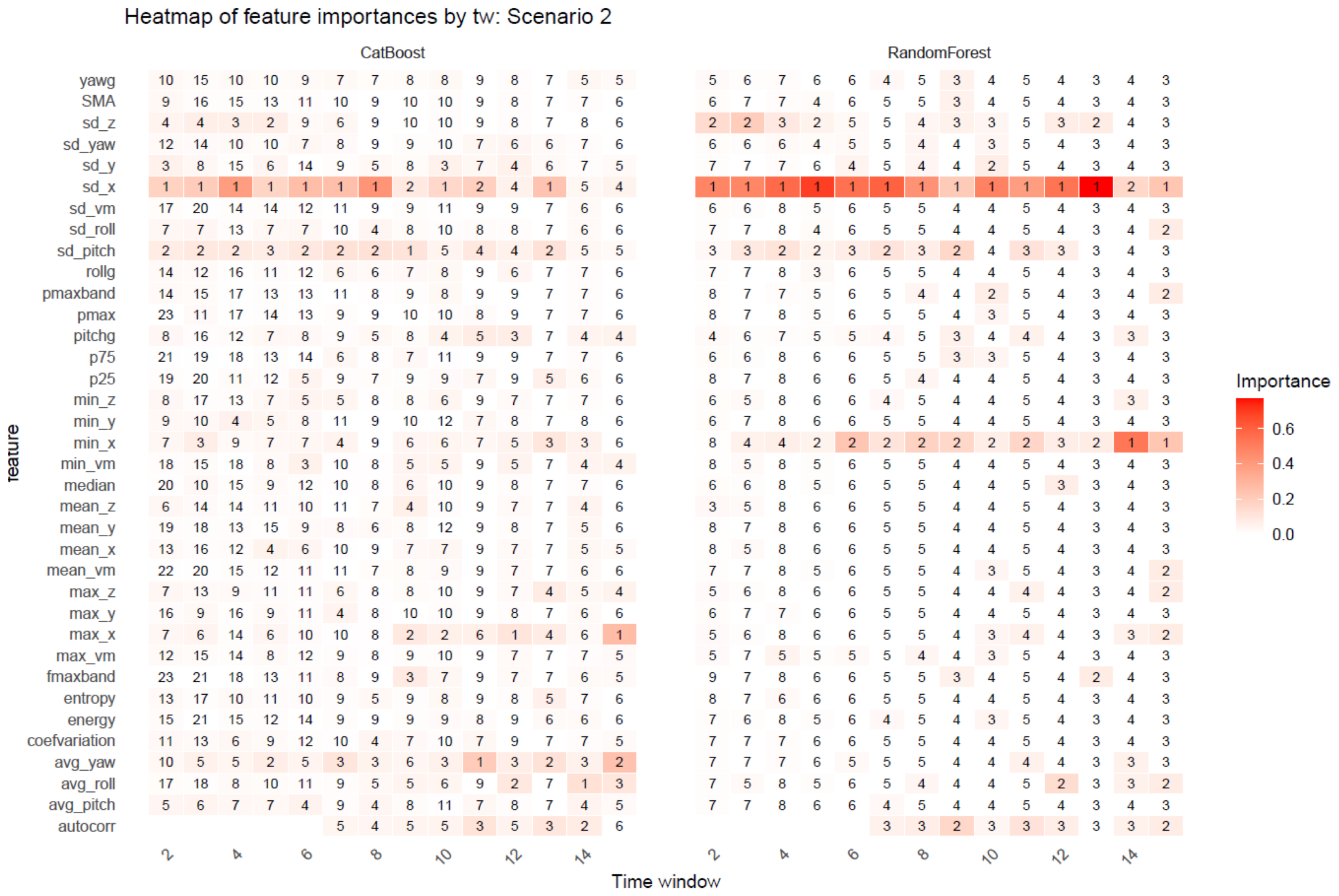
- 3a: training using pen 1 across all weeks;
- 3b: training using pen 2 across all weeks;
- 3c: training using pen 3 across all weeks;
- 3d: training using pen 4 across all weeks.



Feature importance scenario 1



Feature importance scenario 2



Feature importance scenario 3