Classification of animal behaviour

– with focus on 3d accelerometers

Heidi Rautiainen

Outline

- (animal fitted) sensors beyond location
- 3D accelerometers – applied perspective
- Behaioural classification
  - Signal processing
  - Segementation
  - Feature extraction and selection
  - predictive modelling

Individually fitted sensors beyond positional/locational sensors (e.g., GPS)

BIOLOGGING

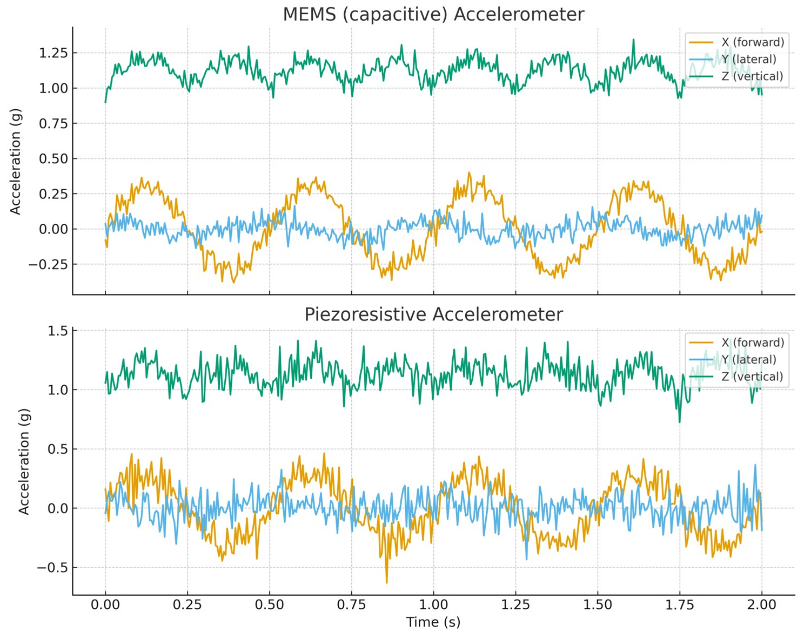
Activity, behaviour and energetics
- **Accelerometers**
- Gyroscopes
- Magnetometers
- IMU = acc + gyro (+ magnetometer)

Depth, light pressure sensors

Physiological data
- Heart rate
- Body temperature
- Sound

Measures static (gravity) and dynamic acceleration

(changes in velocity)



Animal-attached 3D-accelerometers (X, Y, Z)

- Most common type of accelerometer (for behavioural classification)
- Changes in velocity (m/s^2) of the body over time
- To remotely monitor fine-scale behaviours and body posture
- To estimate energy expenditure
- Unlimited by:
  - observer bias
  - disturbing animals
  - visibility
  - scale of space (and time)
- Possible to montior behaviour over months/years
- Attached on collars, leg bracelets, harnessess, glue-on tags etc.

Application, examples

- Acceleration-based (behavioural) proxy for welfare assessment, e.g.,
  - stress in shelter dogs (Jones et al 2014)
  - parasitic pressure in sheep (Ikurior et al 2020)
  - early indicators for diseases or illness (e.g. joint diseases, mastitis, see e.g., review by Chapa et al 2020)

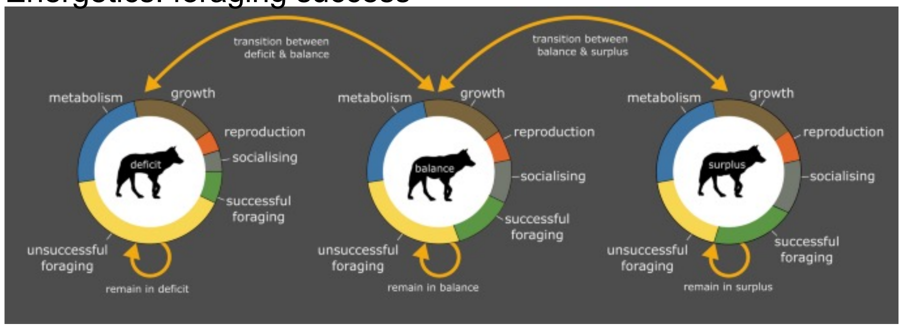
- Acceleration-based proxy for metabolic rate/energy expenditure (commonly no validation)
  - high/low activity (e.g., English et al 2024; Trondrud et al 2023)
  - e.g. overall dynamix body acceleration (OBDA)

- Behavioural specific energetics: e.g., foraging success which in turn can be linked to fitness, survival and reproduction

- Behaviour-specific habitat selection

Chapa et al 2020

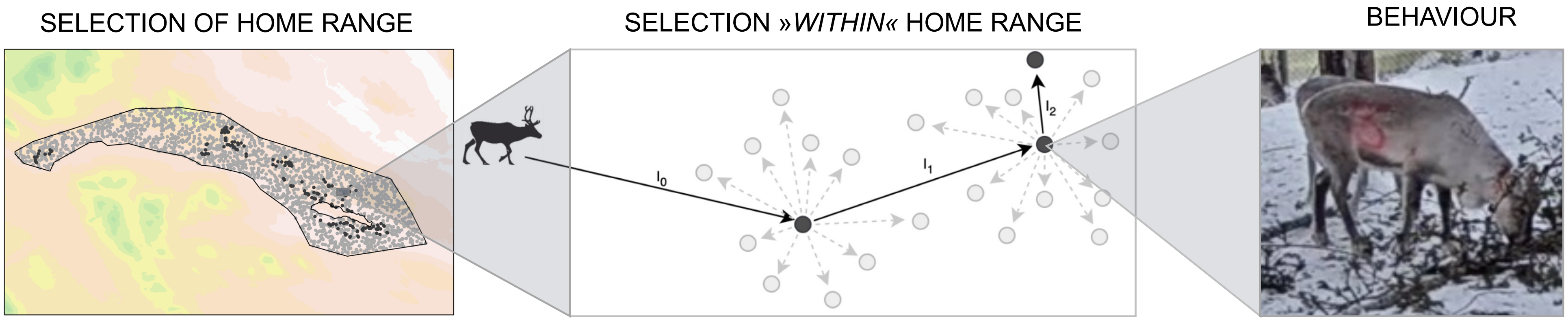
Energetics: foraging success



English et al 2024

Habitat selection at different scales

- understanding the behavioural mechanisms that drive selection processes

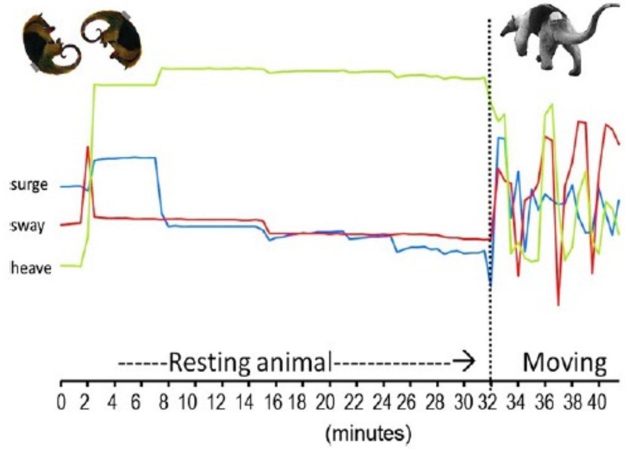


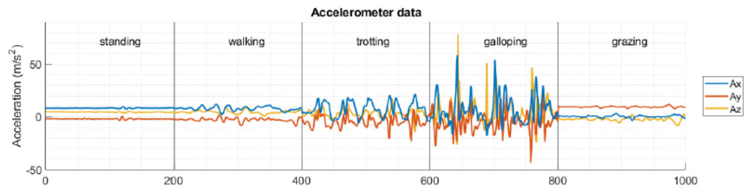
within the seasonal (winter) range

- HSF assumes constant proportion of behaviours – but proportion of behaviour may vary temporally, seasonally and change environmental conditions – and some behaviours may contribute to fitness more than others
- HSF may underestimate selection for habitats that are linked to specific behaviours (see e.g., Roever et al 2013)
  - risk of weak/incorrect inference

Behavioural classification

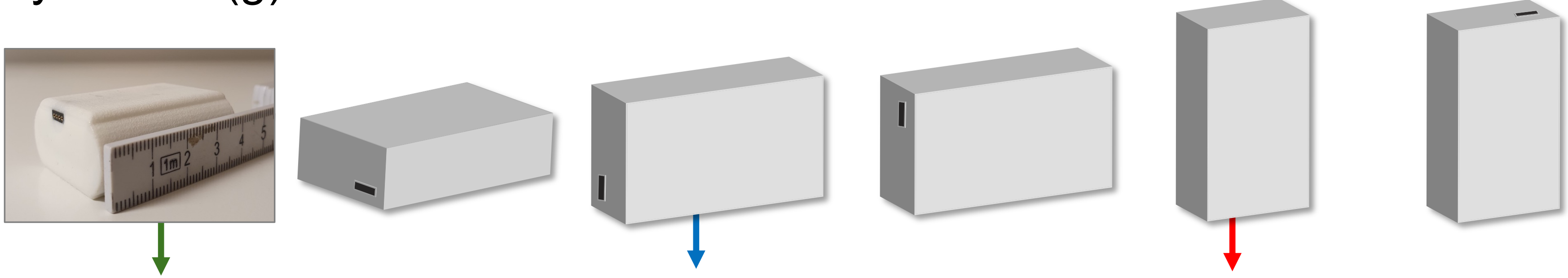
- Sampling frequency: 8 – 100 Hz most common (but 0.5 – 10.000 Hz also used)
- Sampling: continously or in bursts (e.g., GPS interval)
- Behavioural resolution
  - resting vs movement
  - fine-scale: grazing, browsing, resting, walking, running etc.
- Parameters used from acceleration (X, Y, X)
  – signal processing
- Segmentation
- Validation, accuracy:
  Examples of performace metrics: accuracy, precision, sensitivity, F1





Accelerometers (under static circumstances)

Gravity values (g) for different orientations



Z	1	-1	0	0	0	0
Y	0	0	1	-1	0	0
X	0	0	0	0	1	-1

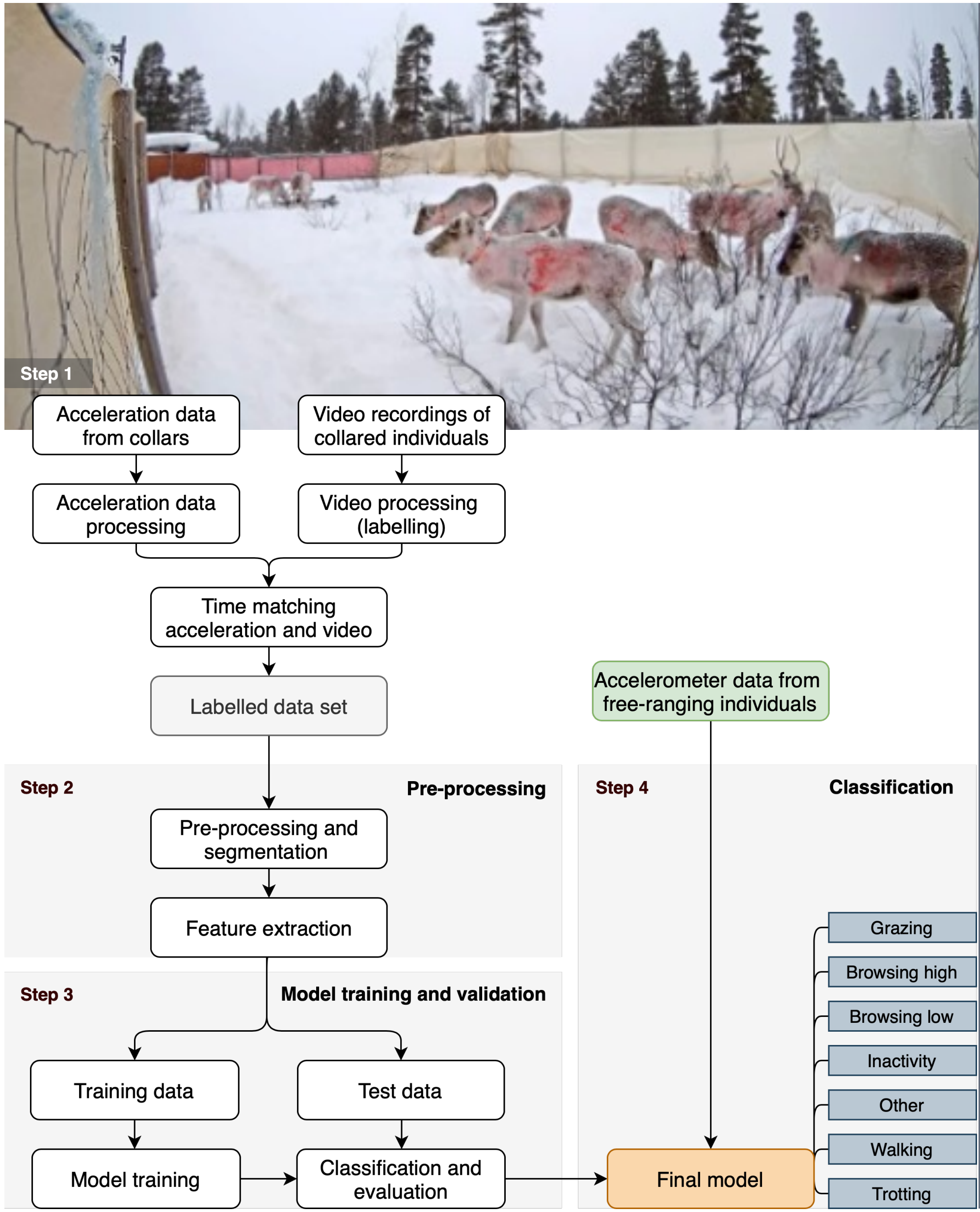
Change in velocity = 1 g = 9.82 m/s^2

Static and dynamic acceleration

- Raw output may be used as it is (static and dynamic) or converted to actual acceleration (dynamic acceleration)
- **Static acceleration** = gravititional force (measuerd in 1 g = 9.82 m/s^2)
- **Dynamic acceleration** = animal movement

Automated classification of reindeer behaviour

Identification of reindeer fine-scale foraging behaviour using tri-axial accelerometer data



Main steps

1. **Data aqusition**
   - Creating labelled data set

2. **Signal processing**
   - Segmentation
   - Feature extraction
   - Feature selection

3. **Model training**
   - validation
   - model performance on new individuals

4. **Classification of behaviour of free-ranging individuals**

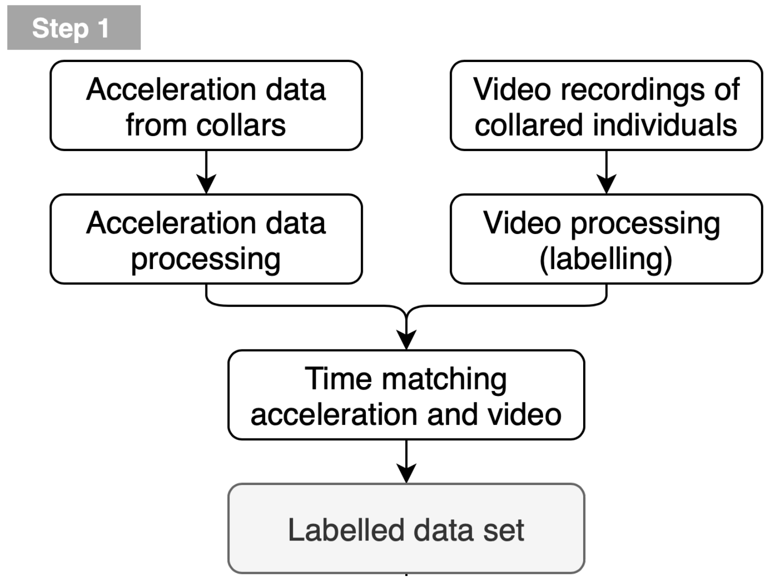
Step 1 Creating labelled data set

Ethogram

Video recordings of individuals fitted with accelerometers

Behavioural annotation (labelling)

Time matching of acceleration and behaviour



Ethogram

Behavior	General description	Sub-group
Browsing	Moving lips towards a tree branch or bush	a) Standing on all four legs, stretching the neck upwards, head level above shoulder height b) Standing on the hind legs, stretching the neck upwards (minimum 45º) c) Standing on all four legs, moving the head forward or downwards, but without mouth touching the ground
Grazing	Lower the head to the ground; foraging from the ground. Mouth close to the ground.	From ground while standing still; lägg till ang. rörelse med ett ben när huvudet är på samma ställe – om den har huvudet nedåt på samma ställe, men flyttar ett/två ben så är det fortfarande grazing From a hole digged hole: the head positioned below the level of lowest point of front leg While walking slowly and foraging from the ground, Mouth close to the ground. Changing head position and walks at least one step.
Digging	Standing and repetitively scratching on ground with one front leg at least two times in a row	
Lying	Belly or side on the ground with folded or extended legs and head in different positions	Resting: folded legs with head raised from the ground facing forward or with the neck bent on the side Sleeping: Head close to (leaning on) ground (on ground or against body) in the same position Ruminating: lying with legs folded and belly on the ground, head raised from the ground facing forward or with the neck bent on the side while chewing Grooming: lying with legs folded and belly on the ground, head moving against legs/body
Standing	Standing on all four legs without moving forward, without chewing	Not chewing
Movement	Moving forward by alternately moving the legs from one point to another	a) Walking: lifting all four legs in a symmetric movement and moving forward, with mouth up from the ground (not grazing) b) Trotting: simultaneous movement of hoof paired two by two diagonally c) Running: three-beat gait faster than the average trot
Agonistic behavior	Pushing away an individual or being pushed away by another individual	
Scratching head against tree	Repeated head movement against branches on trees without having contact with the lips on branch	
Missing data	Animal out of sight	Or when head is not visible for grazing behavior
Other		E.g., standing chewing

Browsing (A)

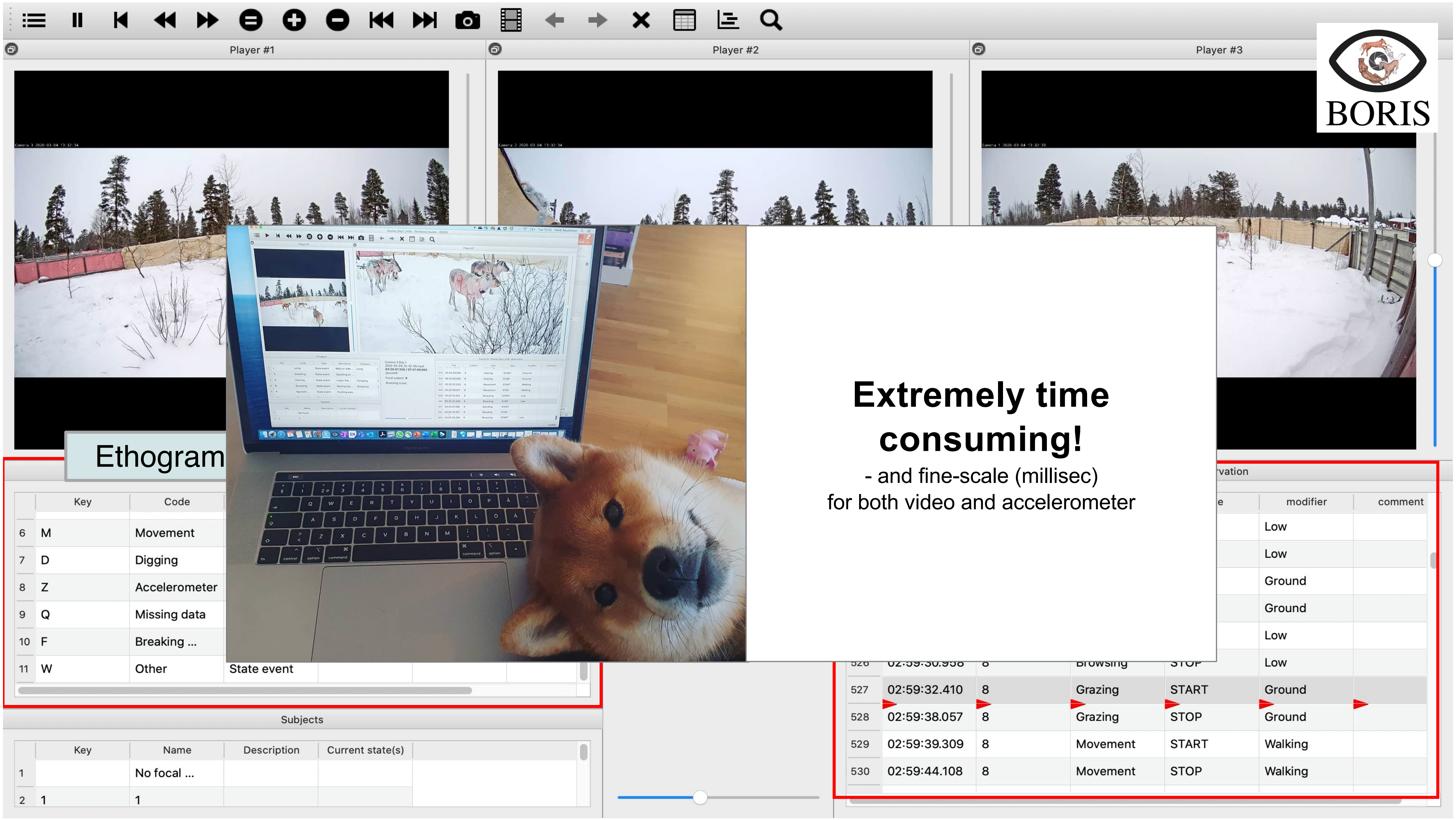
Individual 2, Study area 2, Day 1
Behaviour: Browsing “High”

General description:

Moving lips towards a tree branch or bush

Browsing – high (A):

Standing on all four legs, stretching the neck upwards, head level above shoulder height



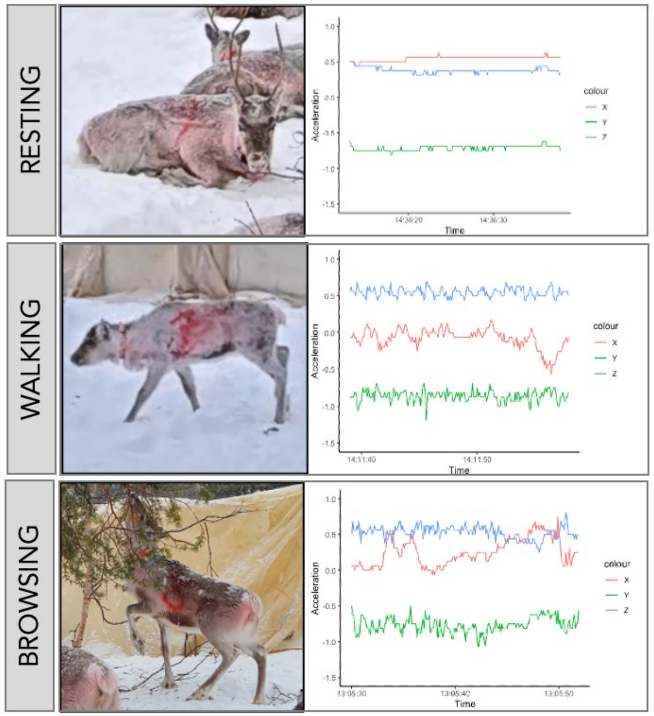
Player #1
Player #2
Player #3
BORIS
Ethogram
Key
Code
6 M Movement
7 D Digging
8 Z Accelerometer
9 Q Missing data
10 F Breaking ...
11 W Other State event
Subjects
Key
Name
Description
Current state(s)
1 No focal ...
2 1 1
modifier
comment
Low
Low
Ground
Ground
Low
Low
527 02:59:32.410 8 Grazing START Ground
528 02:59:38.057 8 Grazing STOP Ground
529 02:59:39.309 8 Movement START Walking
530 02:59:44.108 8 Movement STOP Walking
Extremely time consuming!
- and fine-scale (millisec)
for both video and accelerometer

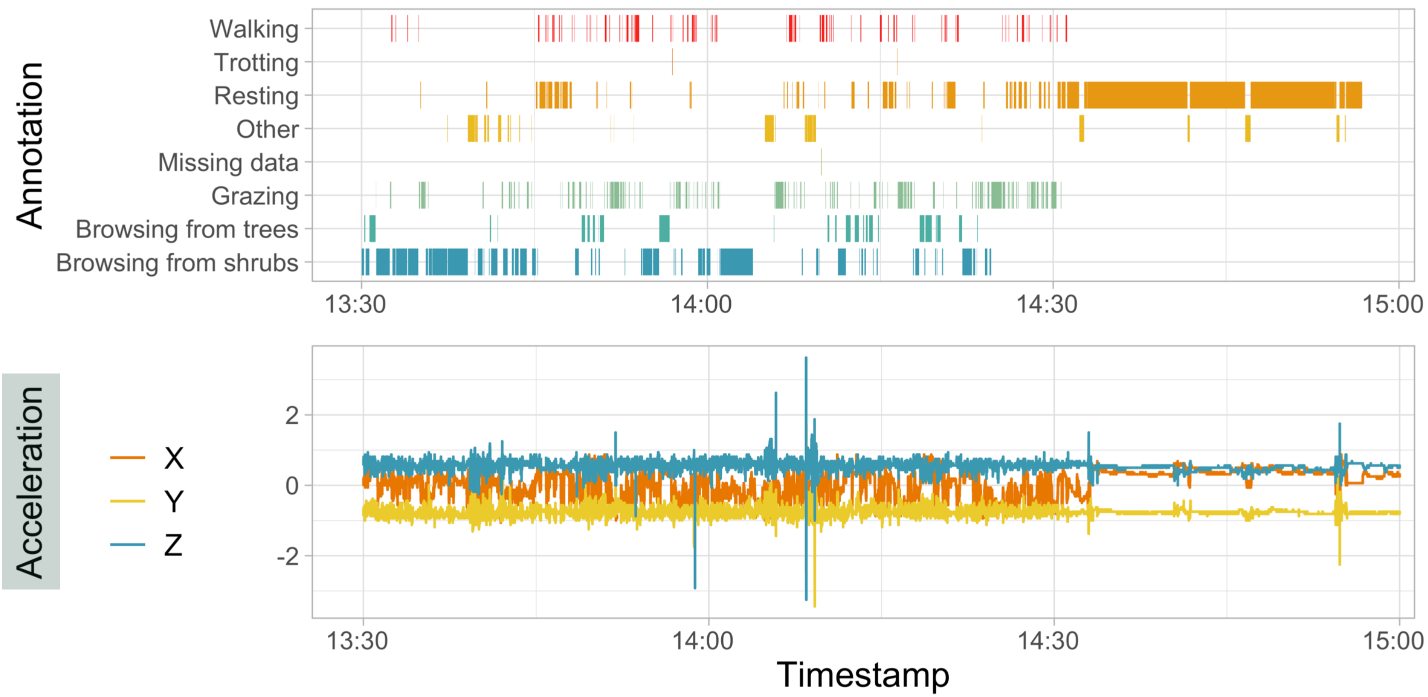
Step 1 Labelled dataset

Video recordings

Acceleration

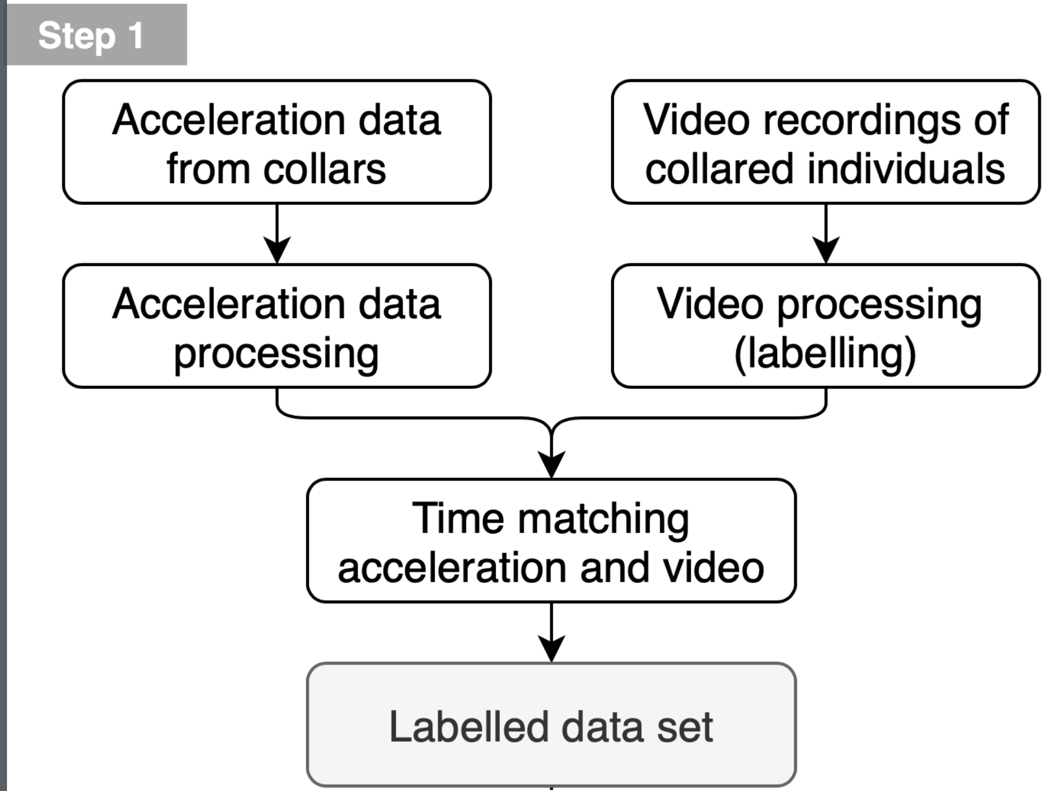
Continuous recordings of **7 behaviours** and corresponding acceleration (19 individuals)





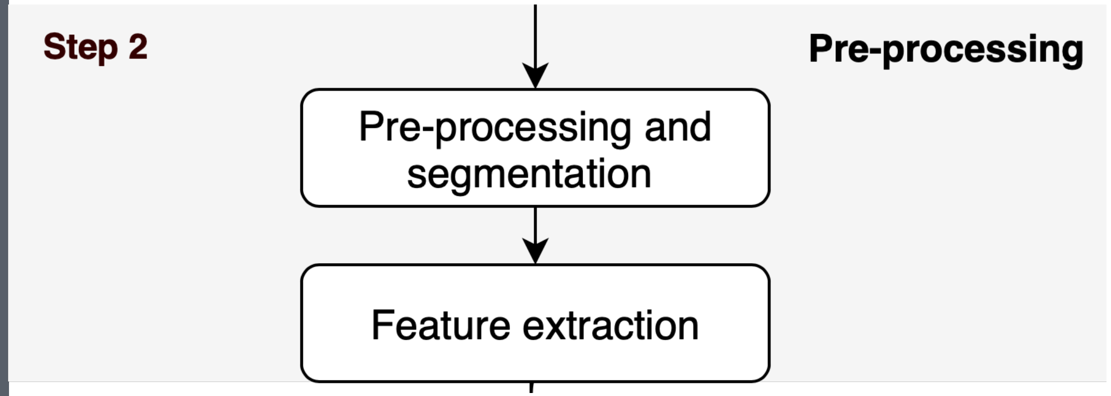
Step 1 Final labelled dataset

	TagID	Timestamp	X	Y	Z	Modifiers
1	Acc 1	2020-02-27 11:18:06	0.125	-0.688	0.188	Trotting
2	Acc 1	2020-02-27 11:18:06	0.625	-1.000	0.750	Trotting
3	Acc 1	2020-02-27 11:18:06	0.250	-0.375	0.375	Trotting
4	Acc 1	2020-02-27 11:18:06	0.188	-1.313	1.500	Trotting
5	Acc 1	2020-02-27 11:18:06	1.000	-1.688	0.750	Trotting
6	Acc 1	2020-02-27 11:18:06	0.188	-0.813	0.438	Trotting
7	Acc 1	2020-02-27 11:18:06	0.250	-0.875	0.750	Trotting
8	Acc 1	2020-02-27 11:18:07	0.313	-0.375	0.500	Standing
9	Acc 1	2020-02-27 11:18:07	0.250	-0.438	0.500	Standing
10	Acc 1	2020-02-27 11:18:07	0.313	-0.625	0.563	Standing
11	Acc 1	2020-02-27 11:18:07	0.313	-0.688	0.625	Standing
12	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.563	Standing
13	Acc 1	2020-02-27 11:18:08	0.375	-0.625	0.563	Standing
14	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.563	Standing
15	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.625	Standing
16	Acc 1	2020-02-27 11:18:08	0.375	-0.750	0.688	Standing
17	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.625	Standing
18	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.625	Standing
19	Acc 1	2020-02-27 11:18:08	0.375	-0.688	0.625	Standing
20	Acc 1	2020-02-27 11:18:08	0.375	-0.750	0.563	Standing



Step 2 Pre-processing

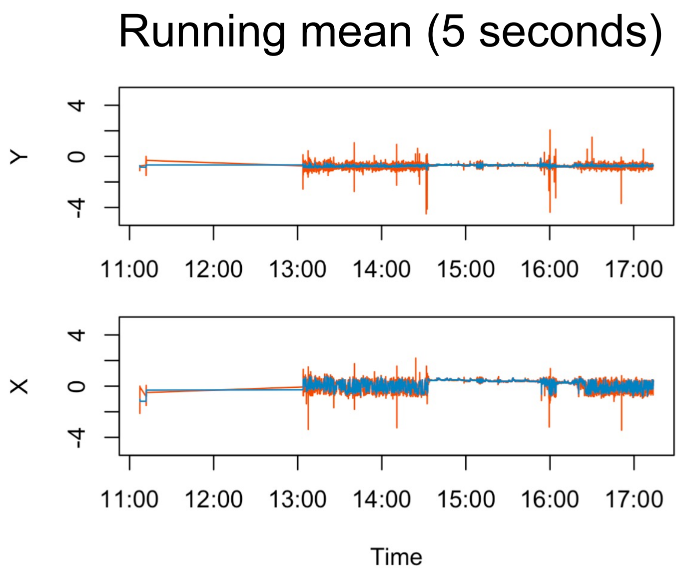
- Separate the dynamic accelration from static acceleration
  - i.e., separate acceleration caused by movement from acceleration caused by gravity
- Find metrics to use as discrete representation of the acceleration data

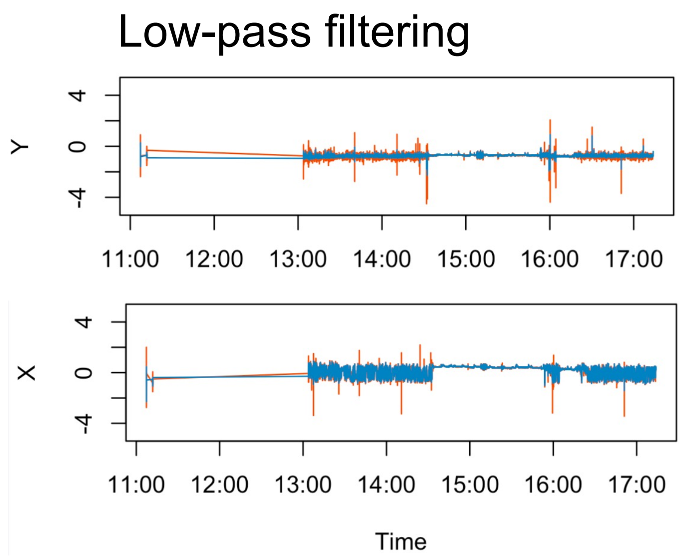


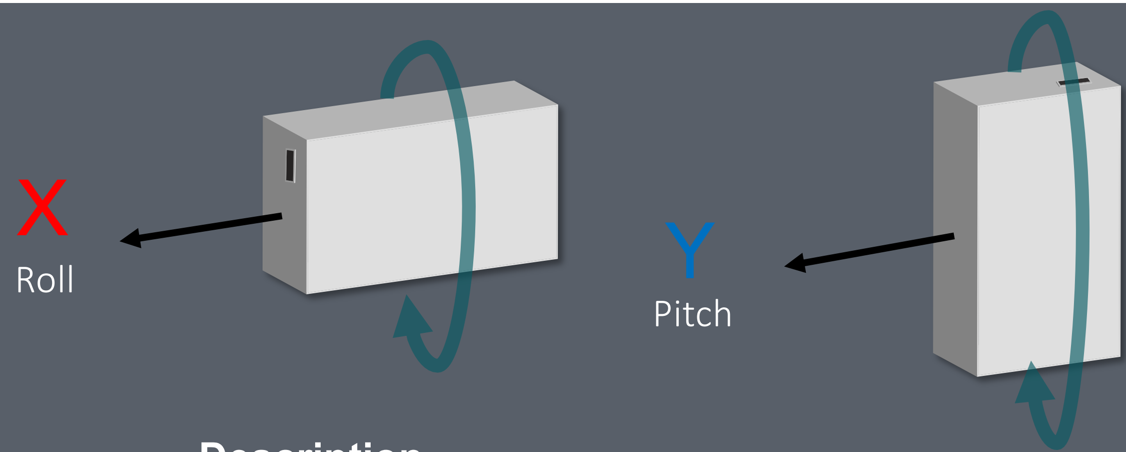
Signal processing

Filtering

- Noise removal/spikes
- Static acceleration
- Dynamic acceleration
- E.g., running mean or high/low pass-filtering

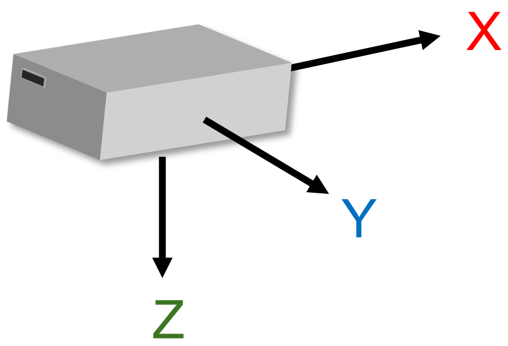






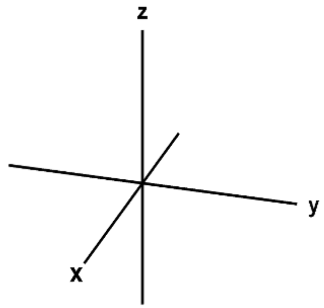
Signal processing

	Term	Equation	Description
Static acceleration	sX, sY, sZ	$\mathrm{sX_i} = \frac{1}{51}\sum_{i-25}^{i+25} \mathrm{X_i}$	Gravitational component of acceleration (9.81 m/s^2 = 1g)
Dynamic acceleration	dX, dY, dZ	$\mathrm{dX_i} = \lvert \mathrm{X_i} - \mathrm{sX_i} \rvert$	Dynamic acceleration measures acceleration caused by animal movements where the gravitational component is removed
Roll	*φ*	$atan2(sY, sZ)$	*Rotation around the X-axis (roll)*
Pitch	*θ*	$-atan(\frac{sX}{\sqrt{sY^2 + sZ^2}})$	*Rotation around the Y-axis (pitch)*
ℓ^2-norm of raw accelerometer axes	Norm	$\sqrt{\mathrm{X}^2 + \mathrm{Y}^2 + \mathrm{Z}^2}$	Orientation-independent measure of acceleration magnitude [



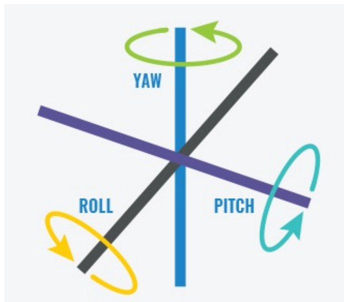
Drawbacks

Accelerometers
– linear (gravity and movement)
– collar orinetation and rotation can be accounted for to some extent
– gyroscopes or magnetometers can provide angles for rotation correction



Magnetometers
– measures earths' magnetic field, *compass*
– provide **absolute rotation** around X, Y and Z
– helpful for dynamic behaviours (grooming, grazing, walking)

Gyroscope
– angular velocity i.e. **rotation speed** (°/s)
– gives rotation specific movements (e.g. head shaking)
– battery consumption



Signal transformations

Time-domain transformations:

- Summary statistics (example: mean, IQR, median, min, max)

Frequency-domain transformations: frequency patterns

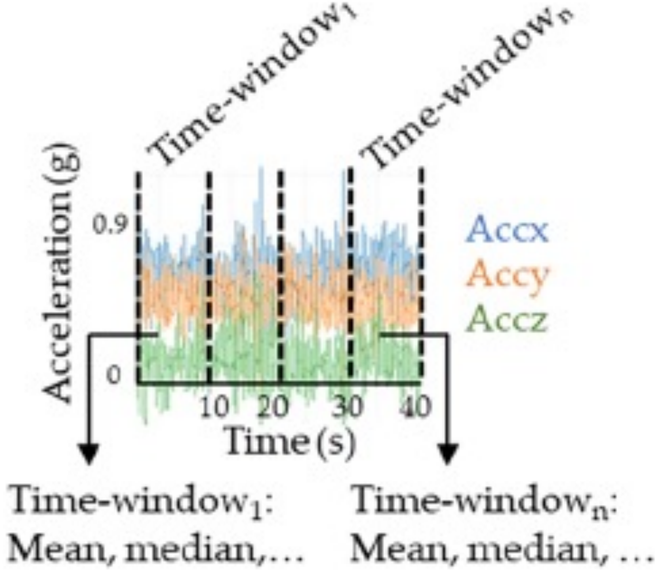
- Continuous frequency domain features such as
  - Power spectrum densities (PSD),
  - usually require Fourier transform (FFT)

Time-frequency / wavelet methods:

- Continuous Wavelet Transform (CWT) or Discrete Wavelet Transform (DWT) captures transient behaviours.
- Short-Time Fourier Transform (STFT) for non-stationary signals

Segmentation, considerations

- All transformations require windowing
- Small windows may not fully capture slow movements
- Large windows may not capture short behaviours
- Window size: < 20 seconds common – but not suitable for short behaviours
- Adaptive windowing an option
- Unique segements or non-overlapping windows:
  - overlapping windows common when annotated data is limited
  - risk of information leakage (if data is placed in both training and test sets) and will give over-optimistic results



Feature extraction: Time-domain transformations

Find metrics to use as discrete representation of the acceleration data

Mean	Mean value for each axis in each window
Minimum	Minimum value for each axis in each window
Maximum	Maximum value for each axis in each window
Median	Median for each axis in each window
Interquartile range	Third quantile (Q3) subtracted by the first quantile (Q1) for each axis in each window
Standard deviation	Standard deviation for each axis in each window

Signal processing

```
 [1] "TagID"     "Timestamp" "X"         "Y"         "Z"         ”Behaviour" "stX"       "stY"       "stZ"       "dyX"
[11] "dyY"       "dyZ"       "roll"      "pitch"     "rotX"      "rotY"      "rotZ"      "rot_dyX"   "rot_dyY"   "rot_dyZ"
[20] "rot_stX"   "rot_stY"   "rot_stZ"   "Norm"
```

…and 51 features!

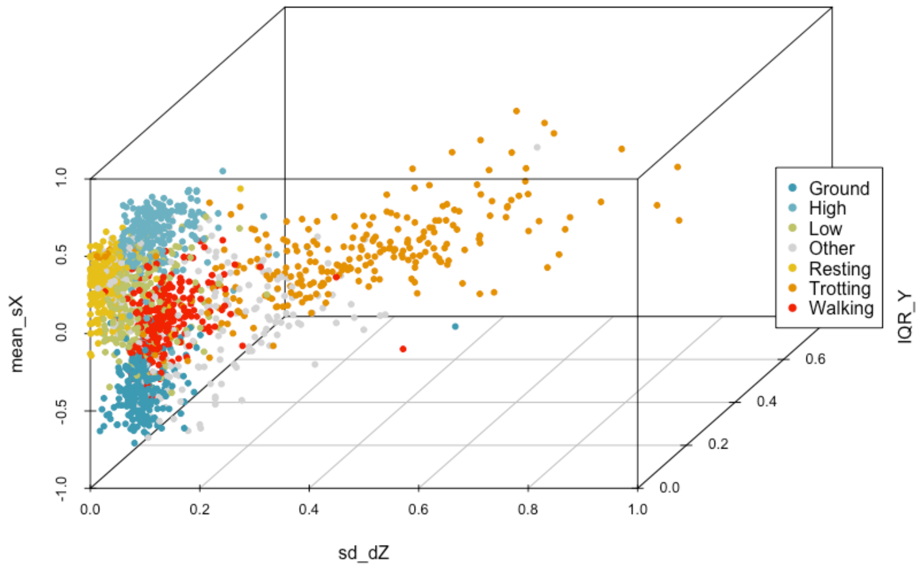
```
 [1] "samp"          "mrot_X"        "drot_X"        "mrot_Y"        "drot_Y"        "mrot_Z"        "drot_Z"
 [8] "mrot_stX"      "drot_stX"      "mrot_stY"      "drot_stY"      "mrot_stZ"      "drot_stZ"      "mrot_dyX"
[15] "drot_dyX"      "mrot_dyY"      "drot_dyY"      "mrot_dyZ"      "drot_dyZ"      "meanrot_stx" "meanrot_sty"
[22] "meanrot_stz" "meanrot_dyx" "meanrot_dyy" "meanrot_dyz" "minrot_stx"  "minrot_sty"  "minrot_stz"
[29] "maxrot_stx"  "maxrot_sty"  "maxrot_stz"  "minrot_dyx"  "minrot_dyy"  "minrot_dyz"  "maxrot_dyx"
[36] "maxrot_dyy"  "maxdrot_yz"  "sdrot_stx"   "sdrot_sty"   "sdrot_stz"   "sdrot_dyx"   "sdrot_dyy"
[43] "sdrot_dyz"   "sdnorm"      "meannorm"    "maxnorm"     "mroll"       "droll"       "mpitch"
[50] "dpitch"      "ID"          "Timestamp"   ”Behaviour"
```

Step 2 Feature selection

Removing highly correlated variables

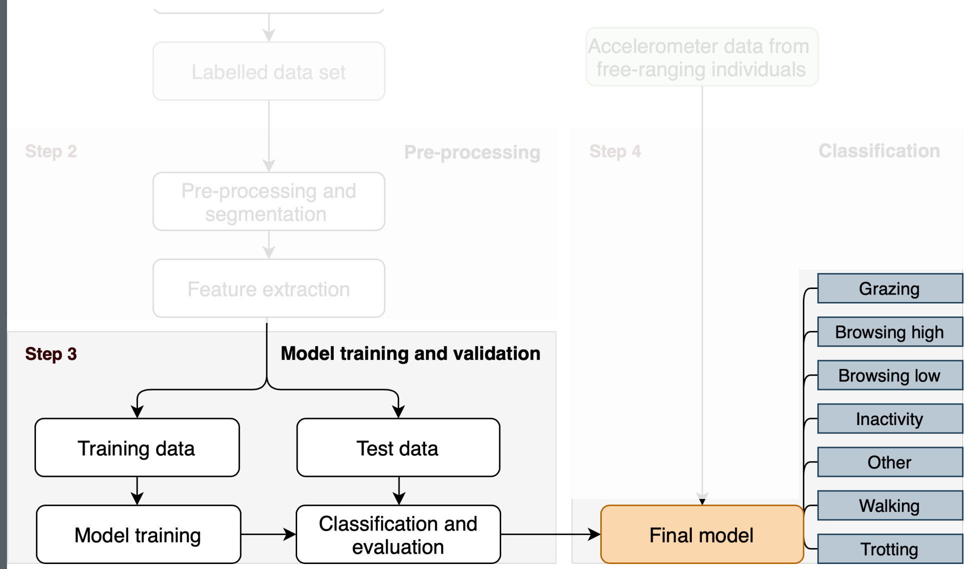
Select the most influential variables for classification (to avoid overfitting the model)

Can also be implemented using e.g., forward feature selection



Step 3 Model training and validation

Overview and considerations



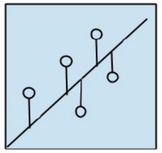
Unsupervised learning

- No response, only "covariates" – only inputs are provided, and the model finds patterns or structures
- Identify **patterns, clusters, or structure** in data
- No ground truth
- E.g., Principal Component Analysis, K-means clustering, hidden-Markov models
- Example covariates: acceleration, or more commonly, GPS (turn angles, step lengths)

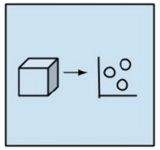
Supervised learning

- For each covariate value x_i (feature) there is also a response value y_i (behaviour) – requires labelled data (known behaviours)
- E.g, linear regression, logistic regression, neural networks, random forests… and hidden-Markov models!

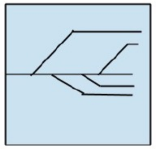
Short about supervised learning…



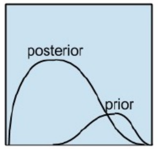
Regression
linear and non-linear
models, geostatistics



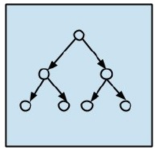
Dimension reduction
PCA, PLS



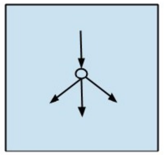
Regularisation Shrinkage
Lasso



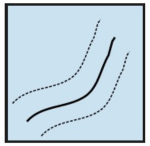
Bayes methods



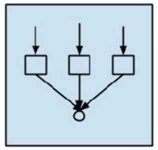
Decision trees
CART



Neuronal networks

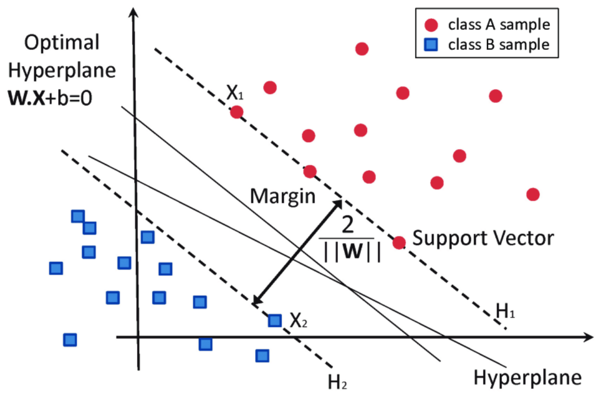


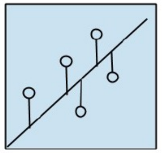
Support vector machines
kernel methods



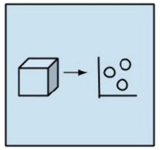
Ensembles
bootstrap, boosting,
model averaging
random forests

Short about supervised learning…

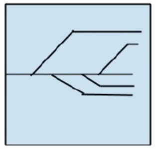




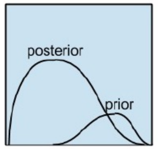
Regression
linear and non-linear models, geostatistics



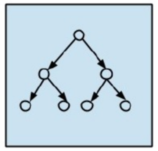
Dimension reduction
PCA, PLS



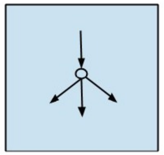
Regularisation Shrinkage
Lasso



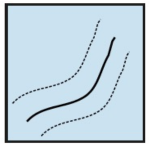
Bayes methods



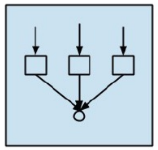
Decision trees
CART



Neuronal networks



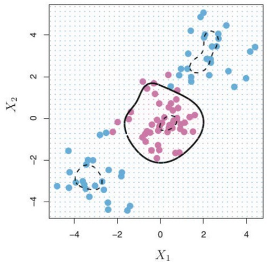
Support vector machines
kernel methods



Ensembles
bootstrap, boosting, model averaging
random forests

Complex, non-linear learning

- acceleration/behaviour rarely linear



Example litterature: Kuhn, M. & Johnson, K. (2013). *Applied Predictive Modeling*. 1st edition. New York: Springer.

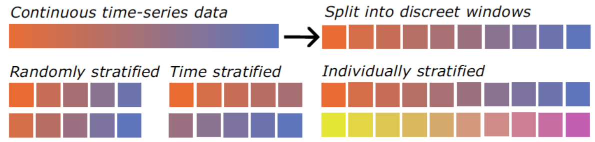
Model workflow, supervised learning

1) Data partitioning -- split data and resample:
   split data into **training, validation** and **test** set
2) Specify formula: Choose response and predictor variables
3) Pre-processing: transformations, normalization (depending on model)
4) Model training and tuning: optimize/tune hyperaparameters of the model
5) Find best model: Check model performance using performance metrics
6) Try model on test data

Some important considerations

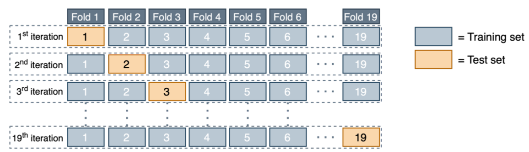
- **Attachment**: fixed deployment ensures constant measures – and deployed similar between individuals!
- **Placement** – compare harness (no foraging) and collar (no digging)
- **Configuration** settings appropriate (sampling interval)
- **Windowing** and feature extraction for best representation of given behaviour
  - but optimal window will also depend on model (HMM better with shorter)
- **Number of behaviours** in a model
- **Model performance and validation**
  - e.g., only 9 % of wildlife studies reported model performance in a review by Brown et al. (2013)
  - Cases of studies where test- and training data share information = will give over-optimistic results!
    E.g. when overlapping windows are used, random split
  - Majority rely on random splitting – not suitable for time series data or when the aim is to predict on *new* individuals
  - 79 % (of 94 papers) do not validate models in a way that could reveal overfitting (Wilson et al 2025)

X-fold CV be good for predicting behaviours on same IDs, but we want to predict behaviours on new individuals

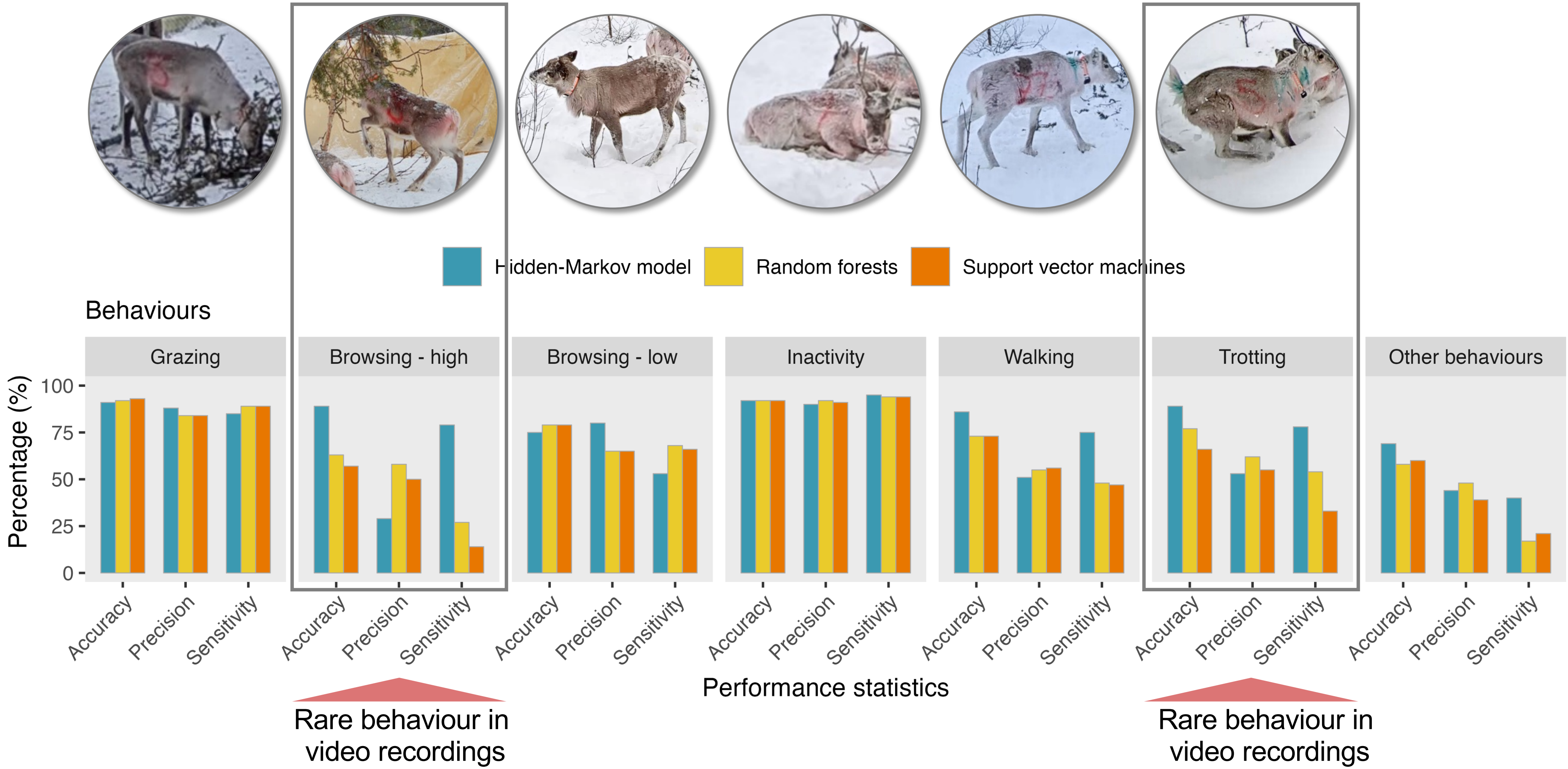


e.g. leave-one-out-CV

Leave-one-out-CV



- To make predictions on **new individuals**
- To not train and validate the model using the same individual; no information leakage
- Performance metrics drops, but gives more generalisable and trustful results
- Drawback: you need more data!



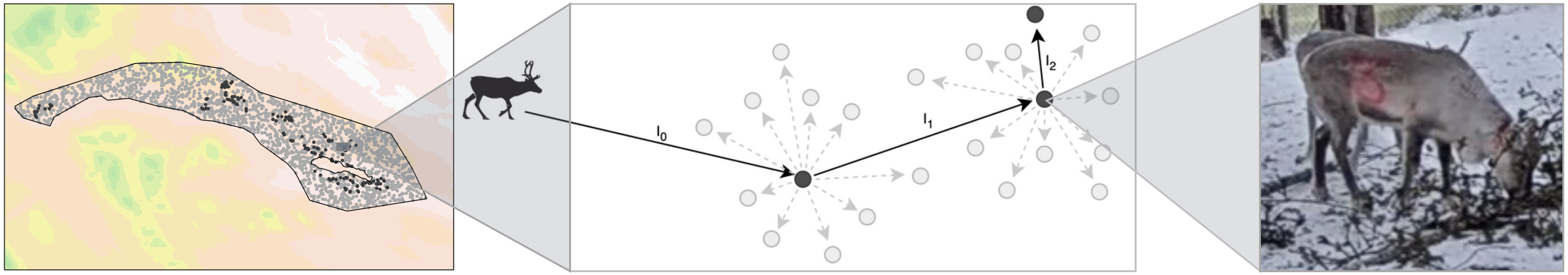
Hidden-Markov model
Random forests
Support vector machines
Behaviours
Grazing
Browsing - high
Browsing - low
Inactivity
Walking
Trotting
Other behaviours
Percentage (%)
100
75
50
25
0
Accuracy
Precision
Sensitivity
Performance statistics
Rare behaviour in video recordings
Rare behaviour in video recordings

Habitat selection at different scales

SELECTION OF HOME RANGE

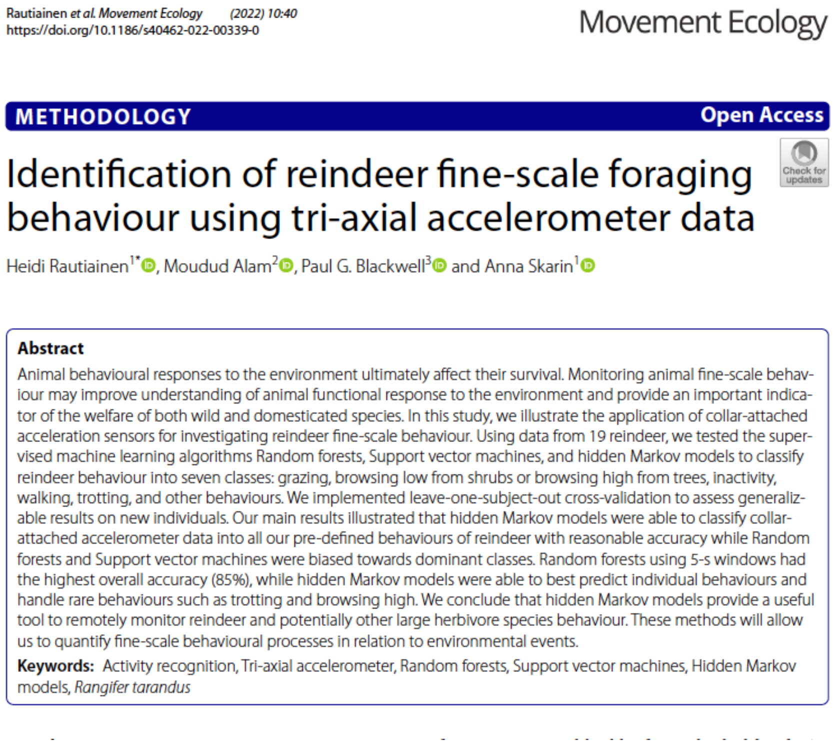
SELECTION »*WITHIN*« HOME RANGE

BEHAVIOUR



within the seasonal (winter) range

Example data


Rautiainen et al. Movement Ecology (2022) 10:40
https://doi.org/10.1186/s40462-022-00339-0

Movement Ecology

METHODOLOGY Open Access

Identification of reindeer fine-scale foraging behaviour using tri-axial accelerometer data

Heidi Rautiainen[1*], Moudud Alam[2], Paul G. Blackwell[3] and Anna Skarin[1]

Abstract

Animal behavioural responses to the environment ultimately affect their survival. Monitoring animal fine-scale behaviour may improve understanding of animal functional response to the environment and provide an important indicator of the welfare of both wild and domesticated species. In this study, we illustrate the application of collar-attached acceleration sensors for investigating reindeer fine-scale behaviour. Using data from 19 reindeer, we tested the supervised machine learning algorithms Random forests, Support vector machines, and hidden Markov models to classify reindeer behaviour into seven classes: grazing, browsing low from shrubs or browsing high from trees, inactivity, walking, trotting, and other behaviours. We implemented leave-one-subject-out cross-validation to assess generalizable results on new individuals. Our main results illustrated that hidden Markov models were able to classify collar-attached accelerometer data into all our pre-defined behaviours of reindeer with reasonable accuracy while Random forests and Support vector machines were biased towards dominant classes. Random forests using 5-s windows had the highest overall accuracy (85%), while hidden Markov models were able to best predict individual behaviours and handle rare behaviours such as trotting and browsing high. We conclude that hidden Markov models provide a useful tool to remotely monitor reindeer and potentially other large herbivore species behaviour. These methods will allow us to quantify fine-scale behavioural processes in relation to environmental events.

Keywords: Activity recognition, Tri-axial accelerometer, Random forests, Support vector machines, Hidden Markov models, *Rangifer tarandus*

Available from:

https://doi.org/10.1186/s40462-022-00339-0
https://doi.org/10.5061/dryad.8sf7m0cs7

Litterature

Reviews on best practices (focus on supervised learning from 3d accelerometers)

Brown, D.D., Kays, R., Wikelski, M., Wilson, R. & Klimley, A. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry,* 1(1), 1-16. https://doi.org/10.1186/2050-3385-1-20

Riaboff, L., Shalloo, L., Smeaton, A.F., Couvreur, S., Madouasse, A. & Keane, M.T. (2022). Predicting livestock behaviour using accelerometers: a systematic review of processing techniques for ruminant behaviour prediction from raw accelerometer data. *Computers and Electronics in Agriculture,* 192, 1–22. https://doi.org/10.1016/j.compag.2021.106610

Wilson, O., Schoeman, D., Bradley, A. & Clemente, C. (2025). Practical guidelines for validation of supervised machine learning models in accelerometer-based animal behaviour classification. *J Anim Ecol,* 94(7), 1322-1334. https://doi.org/10.1111/1365-2656.70054

Litterature cited

Chapa, J.M., Maschat, K., Iwersen, M., Baumgartner, J. & Drillich, M. (2020). Accelerometer systems as tools for health and welfare assessment in cattle and pigs - A review. Behav Processes, 181, 104262. https://doi.org/10.1016/j.beproc.2020.104262

English, H.M., Borger, L., Kane, A. & Ciuti, S. (2024). Advances in biologging can identify nuanced energetic costs and gains in predators. *Mov Ecol,* 12(1), 7. https://doi.org/10.1186/s40462-024-00448-y

Ikurior, S.J., Marquetoux, N., Leu, S.T., Corner-Thomas, R.A., Scott, I. & Pomroy, W.E. (2021). What are sheep doing? Tri-axial accelerometer sensor data identify the diel activity pattern of ewe lambs on pasture. Sensors, 21(20), 6816.

Jones, S., Dowling-Guyer, S., Patronek, G.J., Marder, A.R., Segurson D'Arpino, S. & McCobb, E. (2014). Use of accelerometers to measure stress levels in shelter dogs. *Journal of Applied Animal Welfare Science,* 17(1), 18-28.

Litterature cited

Trondrud, L.M., Pigeon, G., Król, E., Albon, S., Ropstad, E., Kumpula, J., Evans, A.L., Speakman, J.R. & Loe, L.E. (2023). A summer heat wave reduced activity, heart rate, and autumn body mass in a cold-adapted ungulate. Physiological and Biochemical Zoology, 96(4), 282–293. https://doi.org/10.1086/725363

Roever, C.L., Beyer, H.L., Chase, M.J., van Aarde, R.J. & Roura-Pascual, N. (2013). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions,* 20(3), 322-333. https://doi.org/10.1111/ddi.12164