Ecography

E7367
Basille, M., Fortin, D., Dussault, C., Ouellet, J.-P. and Courtois, R. 2012. Ecologically based definition of seasons clarifies predator–prey interactions. – Ecography 35: xxx–xxx.

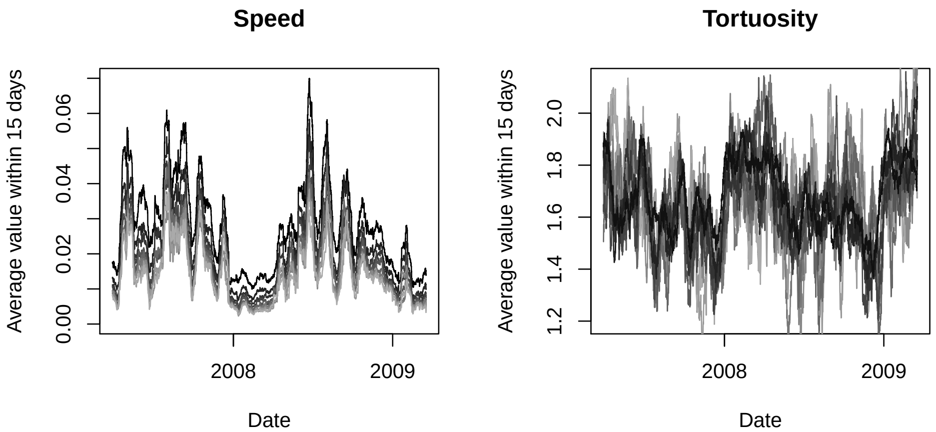
Supplementary material

Appendix 1–5

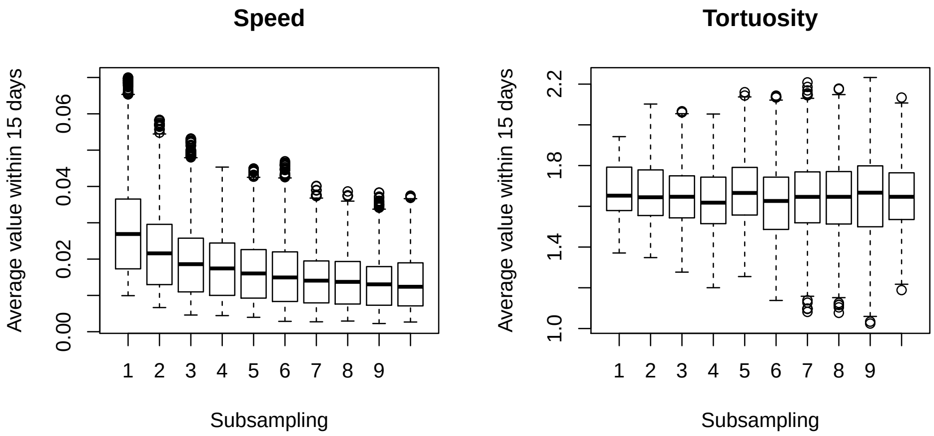
Supplementary material: Expanded results of seasonality

1 Impact of sampling frequency

We assessed variation in movement pattern due to the different monitoring schedules, by subsampling our finest-scale data (one location every 1 hour) to one location every 2 to 10 hours. We computed the average speed and tortuosity in a 15-day moving window as indicated in the manuscript. Temporal patterns were the same for every subsample. We illustrate the outcome based on one individual relocated every 1 hour. We first show the average speed (left) and tortuosity (right) in a 15-day moving window from the beginning to the end of the monitoring period. The subsamples are displayed with decreasing levels of grey, from black (one location every 1 hour) to light grey (one location every 10 hours):



The amplitude decreases for speed with higher subsampling (left), while the variance increases for angles (right):

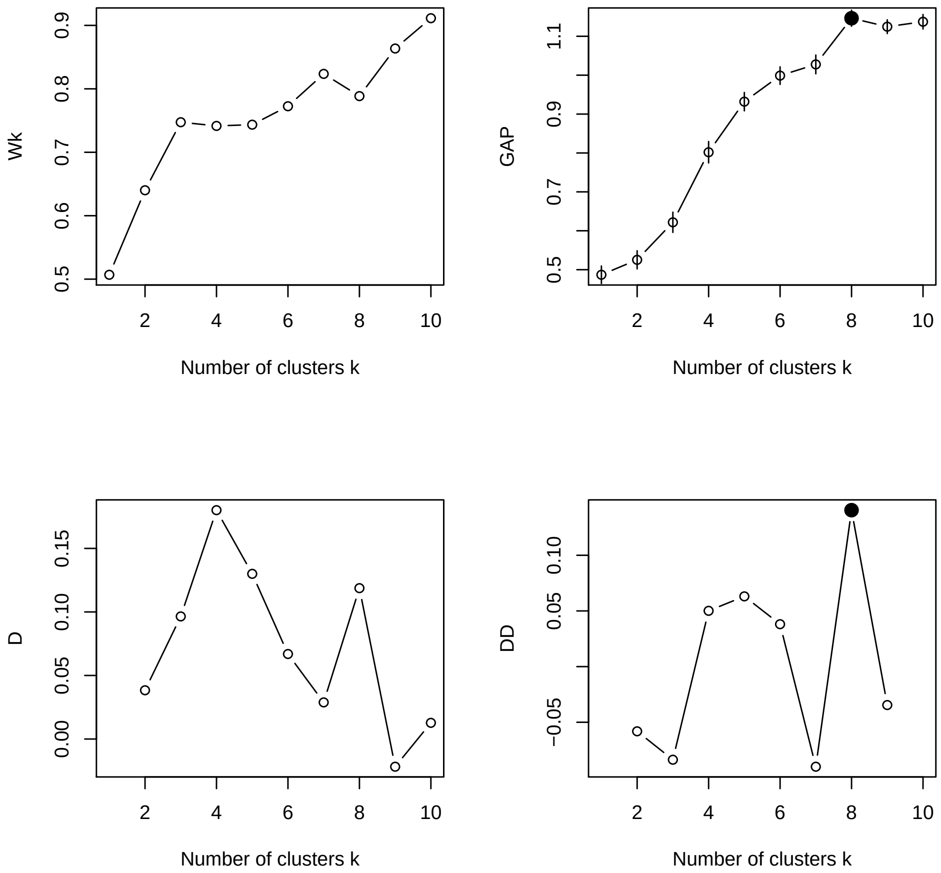


Nevertheless, this was not an issue in our analyses, since the range standardisation levelled out temporal patterns, so that subsamples at 1 hour or 8 hours provided the same results. In the end, this shows that the approach is robust to our level of variation in sampling frequency.

2 A detailed analysis of seasonality: the example of caribou

We start with a 365×9 data frame, with one row per Julian day and one column per space-use variable. The K-means is first run iteratively on this data frame, from 1 to 10 clusters. The "multilayer" approach is then used to choose the number of clusters (see the section *Statistical analyses* of the paper). We first compute W_k as $W_k = \sum_{m=1}^{k} \frac{1}{2n_m(n_m-1)} D_m$ where $D_m = \sum_{i,i' \in m} d_{i,i'}$ is the sum of pairwise distances. The weighted gap statistic is then computed as $\text{gap}_k = E^*\{\log(W_{kb})\} - \log(W_k) = (1/B)\sum_{b=1}^{B} \log(W^*_{kb}) - \log(W_k)$. A structure in the data set is revealed if $\text{gap}_1 < \text{gap}_2 - sd_2\sqrt{(1+1/B)}$, where $sd_2 = [(1/B)\{\sum_{b=1}^{B} \log(W^*_{2b}) - (1/B)\sum_{b=1}^{B} \log(W^*_{2b})\}^2]^{1/2}$. Finally, differences of weighted gap, and differences of difference weighted gap are defined as $D\text{gap}_k = \text{gap}_k - \text{gap}_{k-1}$, and $DD\text{gap}_k = D\text{gap}_k - D\text{gap}_{k+1}$.

To guide the decision, the results can be presented as follows:

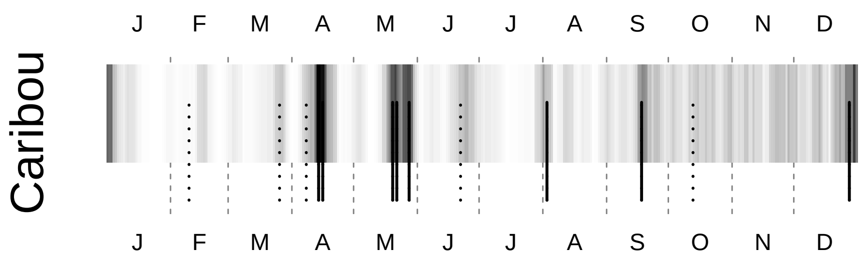


For each plot, the X-axis gives the number of clusters used in the K-means, and the corresponding W_k, weighted gap statistic, $D\text{gap}_k$, and $DD\text{gap}_k$. For the weighted gap statistic, the black dot indicates the estimated number of clusters (in this case, 8 clusters), such as $\text{gap}_k > \text{gap}_{k+1} - sd_{k+1}\sqrt{(1 + 1/B)}$. As the gap statistic tends to overestimate the real number of clusters, the "multilayer" approach provides another criterion of choice: $DD\text{gap}_k$ should be maximised when k is equal to the true number of clusters. The black dot for the $DD\text{gap}_k$ plot thus indicates the maximum value (in this case, also at 8 clusters).

We then proceed with the K-means with 8 clusters. Each Julian day is associated with a given cluster, and consecutive days from the same cluster define a season. The clustering gives the following breakpoints between seasons:

February 10, March 26, April 08, April 14, April 16, May 20, May 22,
May 28, June 22, August 03, September 18, October 13, December 28

We then computed the bootstrap-weight distribution to evaluate the robustness of the delineation, and checked whether the previous breakpoints fell in the top 20 % part of the weight distribution. In the following figures, the weight for each day is displayed with a level of grey, the darker the higher. Seasons that were retained, or dropped, are symbolised by a solid line, or a dashed line, respectively:

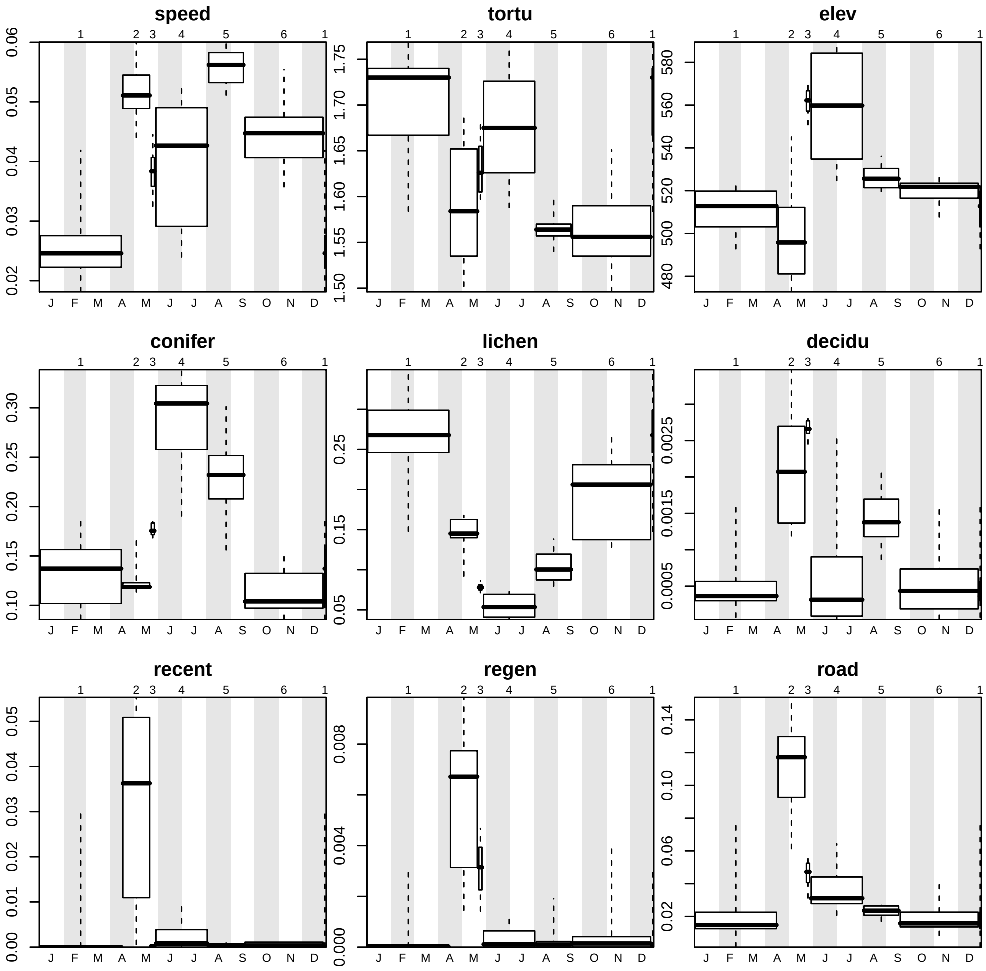


Among every season, only 8 started at days corresponding to the 20 % uppermost bootstrap-weight distribution and were thus retained, while seasons starting at March 26, April 08, April 14, June 22, and October 13 were dropped. Finally, we then remove every season shorter than 5 consecutive days. In this case, two 2-day periods were removed in April (April 14–April 15) and in May (May 20–May 21). We ended up with six biological seasons delimited by the following breakpoints:

April 16, May 22, May 28, August 03, September 18, December 28

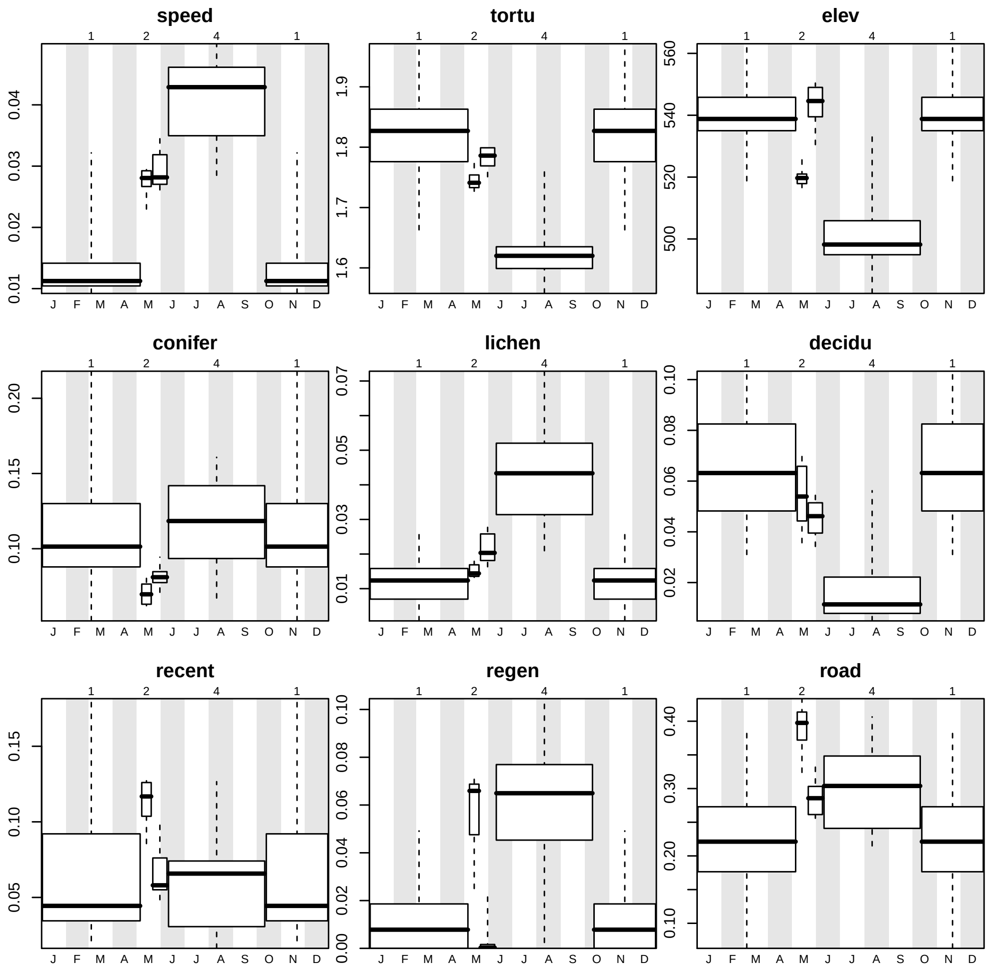
3 Detailed characteristics of sites used by caribou

The K-means associated with the multilayer approach results in 6 seasons during the year for caribou, which characteristics can be seen on the following plot (note that the winter season overlaps the end and the beginning of the year, so that the last and the first boxplot should be read as only one season):



4 Detailed characteristics of sites used by moose

The K-means associated with the multilayer approach results in 4 seasons during the year for moose, which characteristics can be seen on the following plot (note that the winter season overlaps the end and the beginning of the year, so that the last and the first boxplot should be read as only one season):



5 Detailed characteristics of sites used by wolves

The K-means associated with the multilayer approach results in 5 seasons during the year for wolves, which characteristics can be seen on the following plot (note that the winter season overlaps the end and the beginning of the year, so that the last and the first boxplot should be read as only one season):