

Segmentation of Animal Displacement

Josh Cullen

14 April 2020

Background

While we have evaluated animal behavior (for snail kites and snow leopards) using step lengths and turning angles, these measurements typically result in relatively fine temporal scale behaviors. By comparison, net squared displacement (NSD) has been used to evaluate coarser scale behaviors, such as residency, dispersal, migration, partial migration, or nomadism (Bunnefeld et al., 2011). NSD is typically used instead of net displacement (ND) since the former increases linearly with time when animals move according to a random walk (Turchin, 1998; Borger et al., 2008), a property that is favorable to some analyses. However, we will only be analyzing ND since we do not expect the species we analyze to exhibit nomadism and this measure is more interpretable than NSD.

In this document, I will be showing how our time segmentation model can be slightly modified to analyze ND for snow leopards to identify coarse-scale behaviors. I will provide two different implementations for how this can be performed, with the possibility that some adjustments may need to be made. In general, this model will operate by completing two passes to segment the data: (1) the first pass identifies large changes in ND and generates new time segments accordingly, and (2) the second pass performs segmentation on the collated set of time segments per individual. This second pass is either performed on data that have been normalized to a new baseline, or uses the original dataset, but adjusts the value for the hyperparameter (α) to achieve a greater number of breakpoints.

The motivation behind testing two different methods by which to perform the segmentation is that territorial species like the snow leopard frequently return to the same location. And since the baseline is likely to remain the same when only generating a few large time segments, the values for normalized ND may not change much if at all. By adjusting α to achieve small and large numbers of breakpoints, the model can potentially identify a greater number of time segments than possible with the other method, but subjectively relies on tuning the hyperparameter.

Data Exploration

To understand the trends present in the data for each of the four snow leopards, below are time series plots of ND in addition to histograms of ND values per individual.

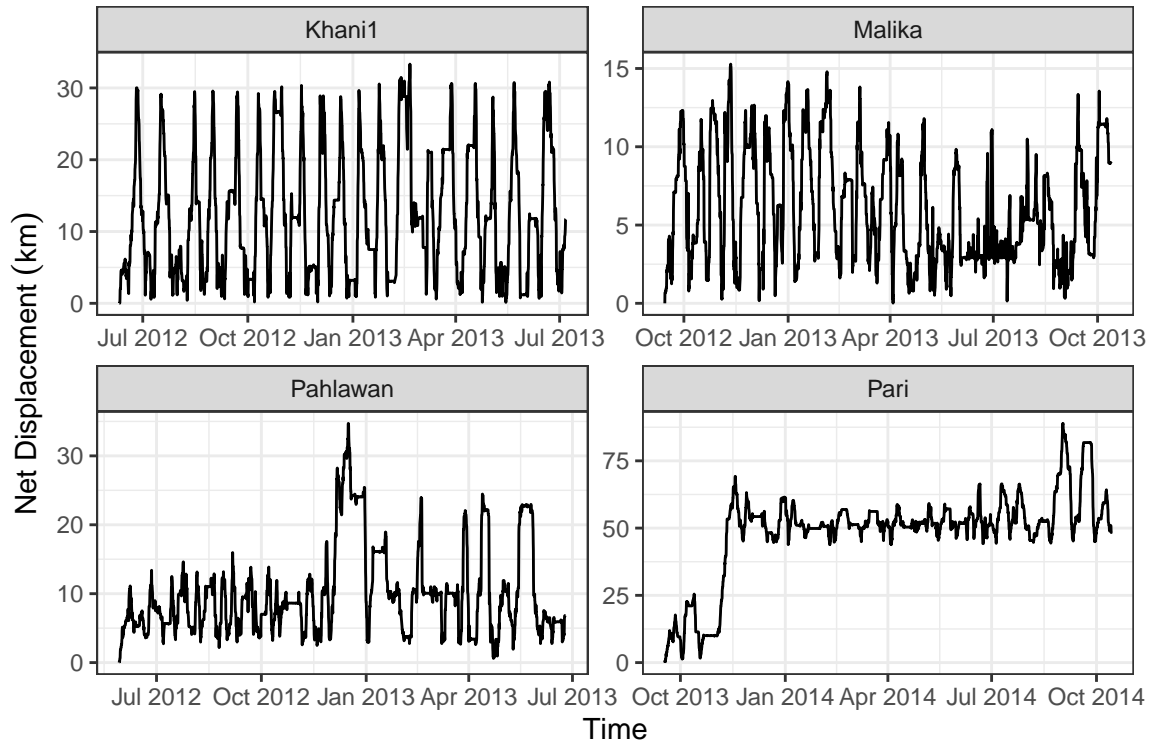


Figure 1: Time series of ND for each snow leopard. Cyclical pattern present in their movement, although note that some individuals have greater magnitudes of change as shown by the different scales of the y-axes.

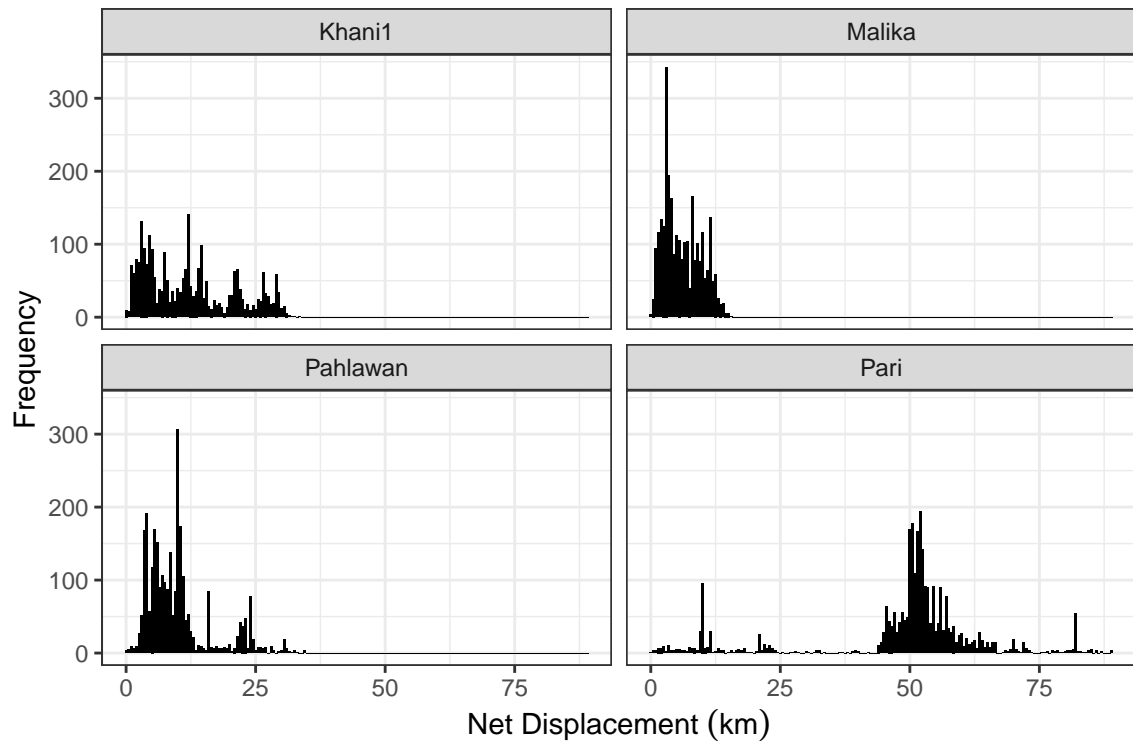


Figure 2: Histograms of ND per individual. Pari has the greatest number of high ND values since she dispersed from her initial tagging location and the others did not.

To segment time series of ND, we must first discretize the variable into bins (and ideally a large number to maintain fidelity to the pattern of change). For this implementation, I use a binwidth of 0.5 km on which to discretize ND for all individuals. This resulted in 109 distinct bins. Next, we will implement the Gibbs sampler to identify

Results

Segmentation using data normalization

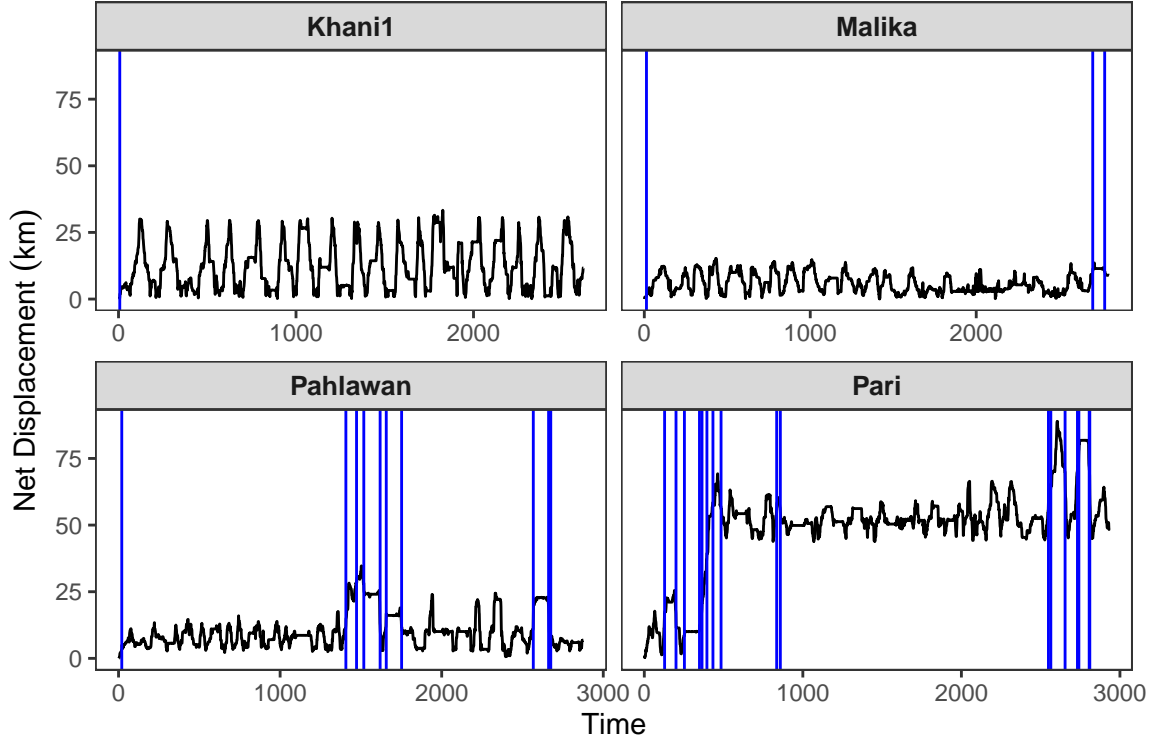


Figure 3: Time series of ND with breakpoints from the first pass of the segmentation model.

The number of breakpoints varies widely among individuals, which denotes a level of inter-individual variability in movement over each 13-month period. These data were analyzed with $\alpha = 1$, which will also be used on the second pass of the segmentation model.

To get a better sense of how the data are changed once each time segment is normalized to its new baseline, the plot below shows a direct comparison:

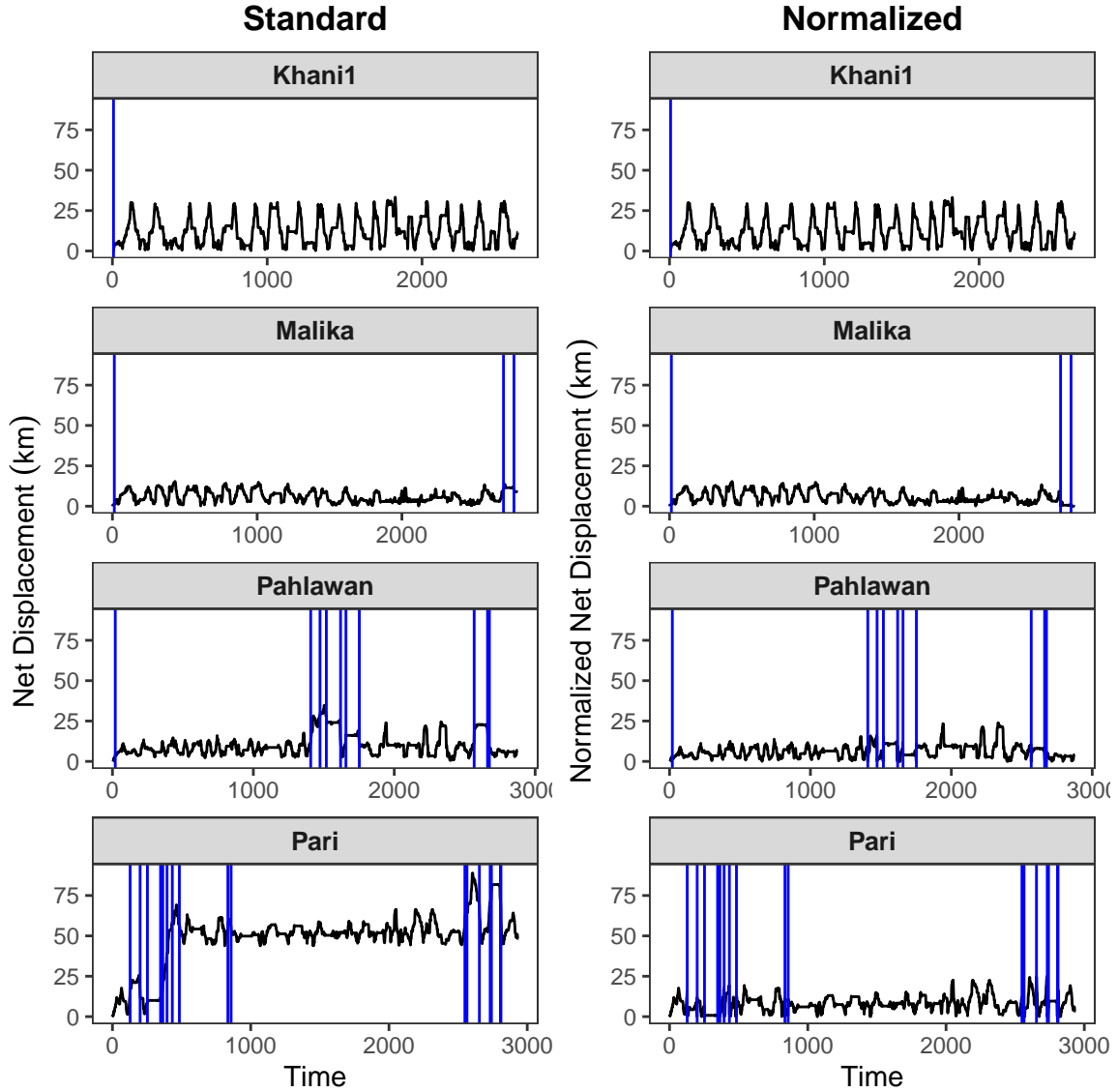


Figure 4: Time series of regular and normalized ND with breakpoints from the first pass of the segmentation model. Although the magnitude of the pattern appears to change, this is just a result of the scale of the y-axis, which decreases from the original to the normalized data due to Pari having a dispersal phase.

This set of comparisons demonstrates how normalization may not be particularly useful in the case of the snow leopard dataset (except for Pari), although it may perform better for the snail kite data. Since we would like to use a single model to analyze any species' ND pattern, it appears that this method may not be the best approach or would require modifications to improve its performance. Now, let's evaluate the results of the second pass of the model after normalization of ND:

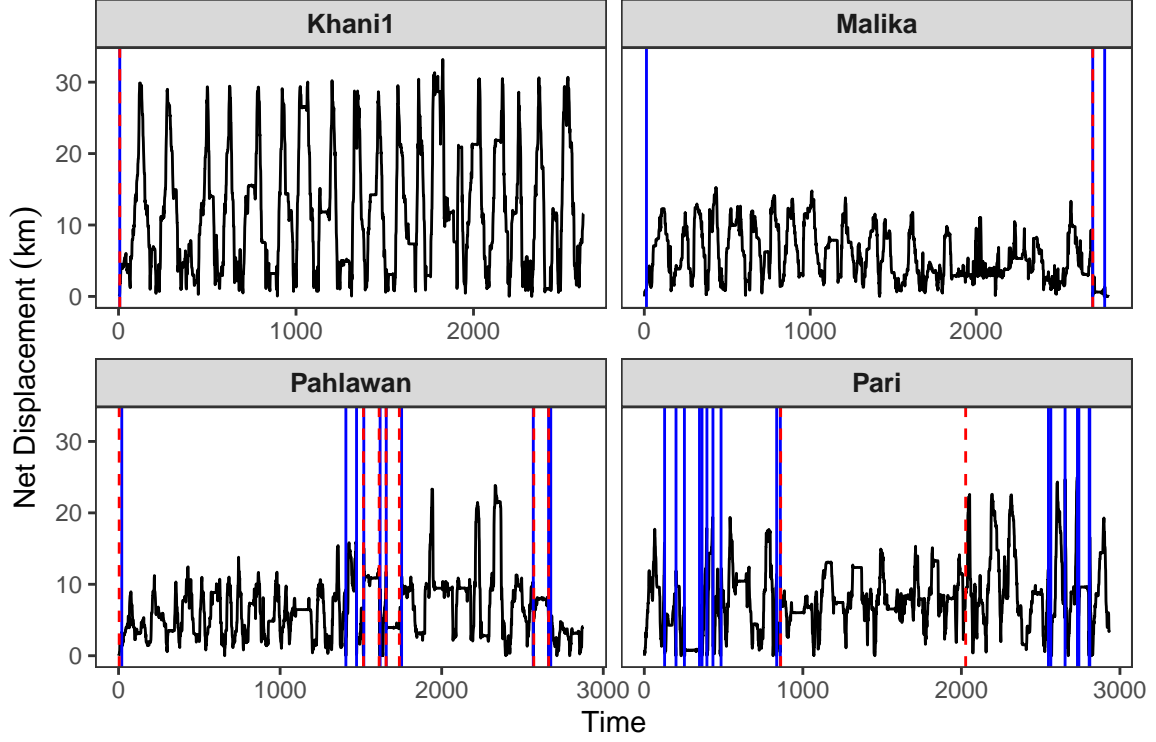


Figure 5: Time series of ND with breakpoints from the first (blue) and second pass (red) of the segmentation model.

We can see that the breakpoints proposed by the second pass of the model are essentially identical to that of the first pass, with the exception of a single new breakpoints for Pari. For comparison, I will now implement a very similar procedure, but will only adjust α to generate fewer or more breakpoints on the original ND data without any normalization for the second pass.

Segmentation by changing value of hyperparameter

For this example of the 2-pass model α will first be set to 2 to acquire a small number of breakpoints and then set to 0.5 to acquire a larger number of breakpoints.

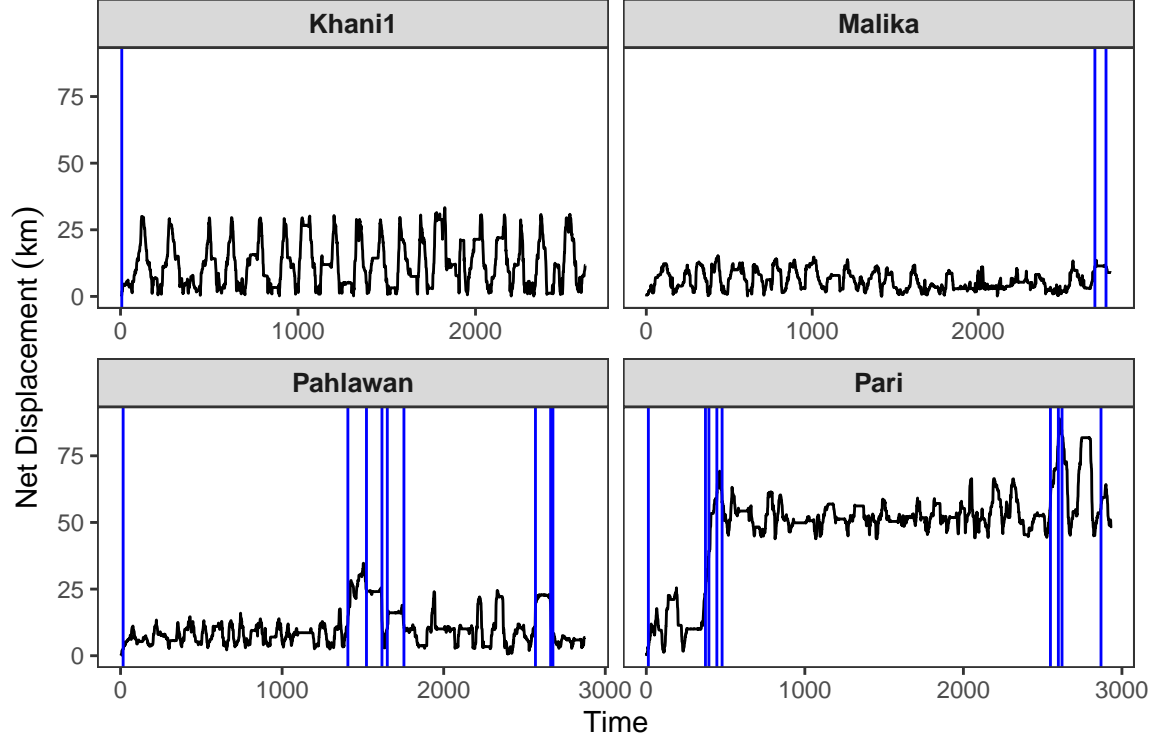


Figure 6: Time series of ND with breakpoints from the first pass of the model where alpha is set to 2.

This modified form of the segmentation model identifies slightly fewer breakpoints than when it was originally analyzed with $\alpha = 1$. These breakpoints are also at similar if not identical locations as this prior model. Next, we will evaluate the breakpoints for the second pass of the model.

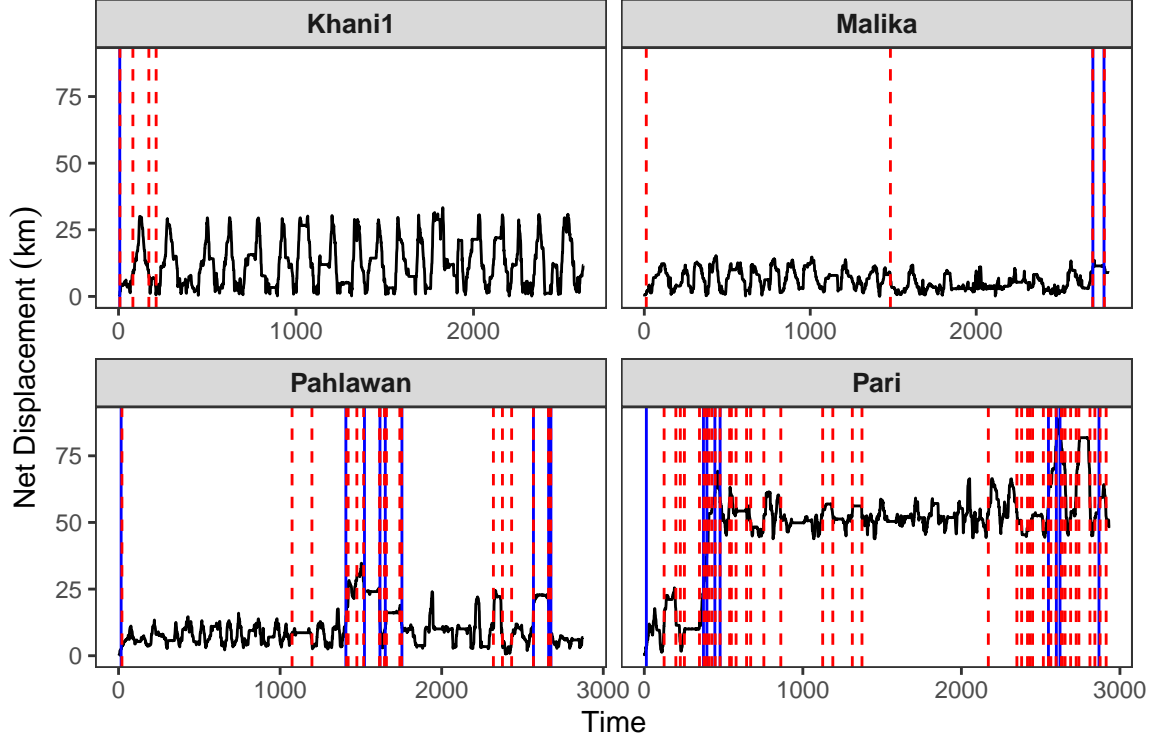


Figure 7: Time series of ND with breakpoints from the first (blue) and second pass (red) of the model when the value for alpha is changed for each pass.

We can see in this version of the second pass model that there are a greater number of unique breakpoints identified, although their ability to identify important changes at all locations along the time series remains to be seen. However, certain breakpoints such as breakpoint in the middle of the time series from the second pass model for Malika appears to denote the time when she had cubs with her. Other individuals, such as Pari, showed what appears to be a superfluous number and location of breakpoints by comparison. Alternatively, the model identified very few breakpoints for Khani, all of which did not appear to be informative. Therefore, further work likely needs to be done on this model to achieve results that match our intuition of how the model should operate and what positions would be considered appropriate breakpoints for the variable of interest.

Conclusions

The results from this process of segmenting the time series of ND are promising for the identification of coarse-scale behavioral changes, such as the timing of dispersal by Pari and the period of time where Malika is taking care of her cubs. However, there are some issues with each of the implementations of this segmentation method that need to be addressed before we can effectively make behavioral inference with these results. It also may be worth considering the merit of using a modification of this segmentation method that treats ND as a continuous rather than a discrete random variable to see if this improves our results. Once these issues have been resolved, this addition of segmentation of ND should qualify as a substantial improvement on standard behavior estimation models by allowing for inference to be made at multiple temporal scales on a variety of movement parameters.