

# Post Sim Run

This takes the simulation output and does some maths to get it to look exactly like the GPS output. This is intended to allow both sim and gps datasets to run through the same code so that 1:1 comparisons can be made.

## Config and Functions

### Read Sim Output

Read several CSV files that have been created using the "make benchmark" command. These files are created using a couple of different config files. The goal here is to compare the different parameters in the benchmark config files.

For this experiment one month of data was generated using different "movement methods" for each run. The different movement methods are:

- Random Movement: No behaviour states, no landscape interaction, just basic step and turn movement
- HMM Movement: Multiple behaviour states that are set the step and turn parameters. Transitions between states is a result of the landscape data.
- DLD movement: Complex movement influenced by behaviour states, home ranges, landscape info etc

The csv's are read into a dataframe in chunks and then saved to a parquet file. This should allow larger simulations to be processed on a smaller machine (like my laptop) without running into memory errors... up to a point.

Convert the parquet file into a geopandas dataframe and eventually a geopackage. This will allow it to be easily added to a QGIS project

### Read GPS Data

Once all the preprocessing has been done, let's read the smaller/compressed parquet file into memory, do some analysis, and start plotting it.

# Calculate HomeRange Timeseries

How much area is covered every week/month/year? Is it similar between the sim and GPS data?

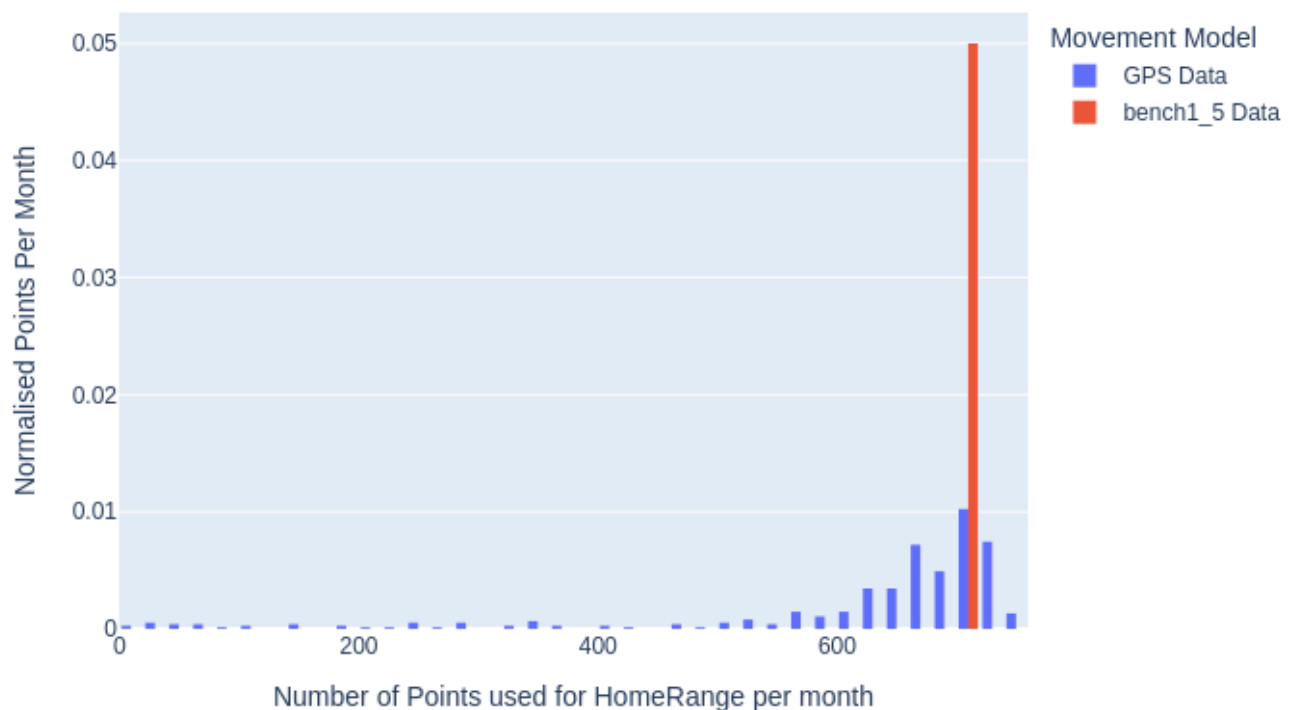
Let's take both the GPS data and SIM data and group it by Agent ID, Month and, in the case of the simulation data, the Filename of the data. This is in order to keep the different benchmark results separate.

So now that the homeranges are written to geopackage files, they can be visualised in QGIS. Here's a quick example:

## Homerange sizes

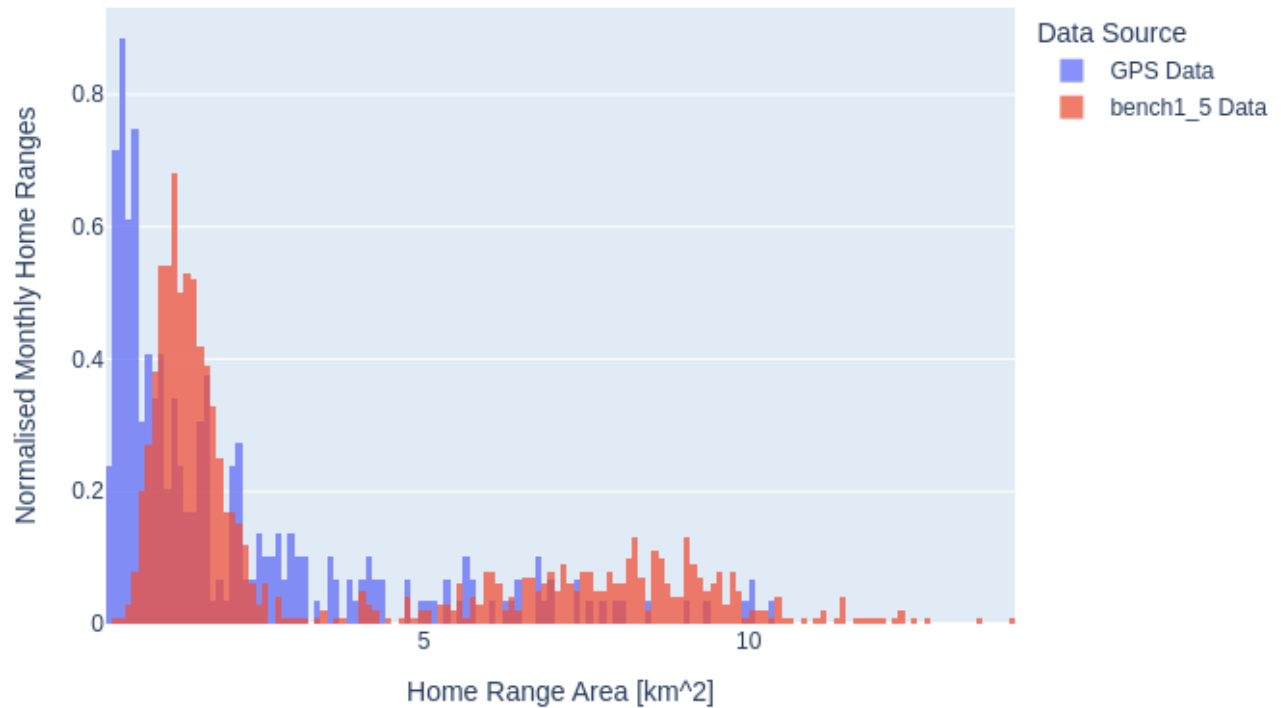
Are the monthly home range areas similar to the GPS data? Let's plot the monthly home range area (convex hull of GPS/Simulated points per agent, per month). This homerange area is not specified in the model parameters and so "should" be similar if the model is behaving as expected.

Monthly Number of Rows/Points in HomeRange



The above figure shows that the simulated data is, obviously, more regular than the GPS data. There are a few GPS devices that have far fewer samples in a given month than expected. These few GPS points would alter the expected monthly home range size. Let's drop all rows that have a points/month that is too low.

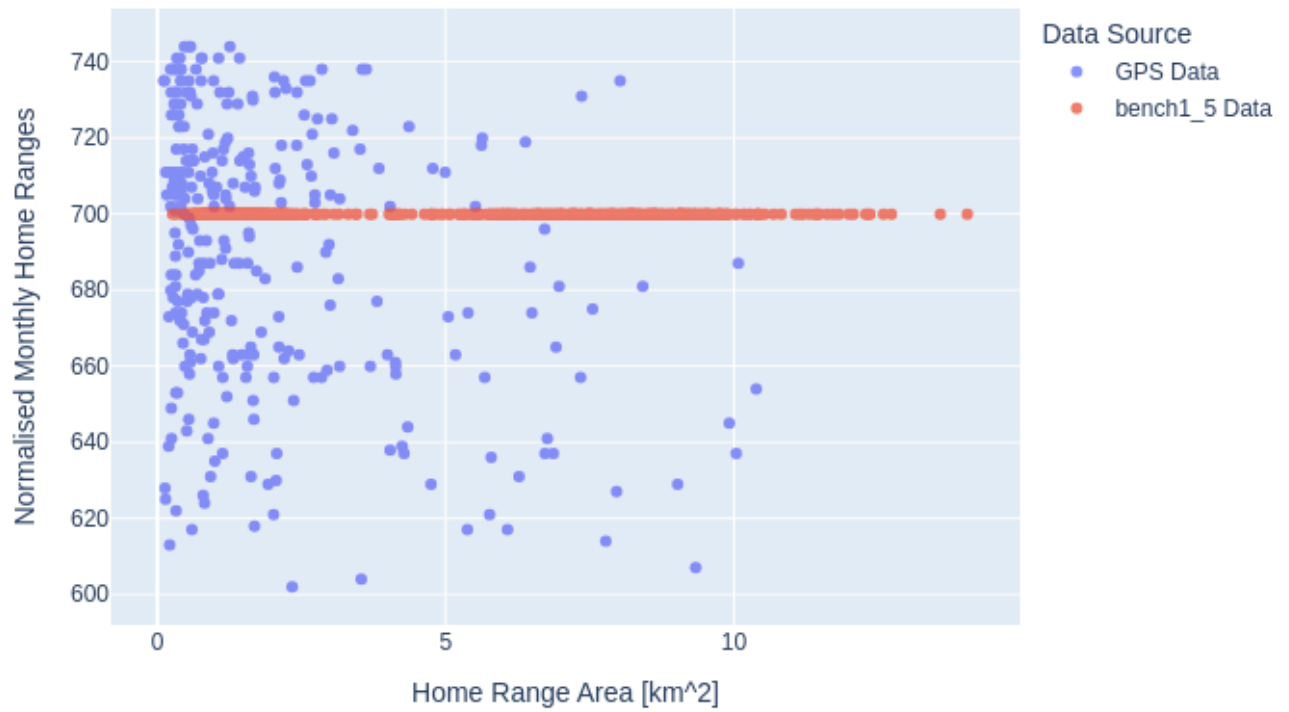
Monthly Home Range Area Histogram



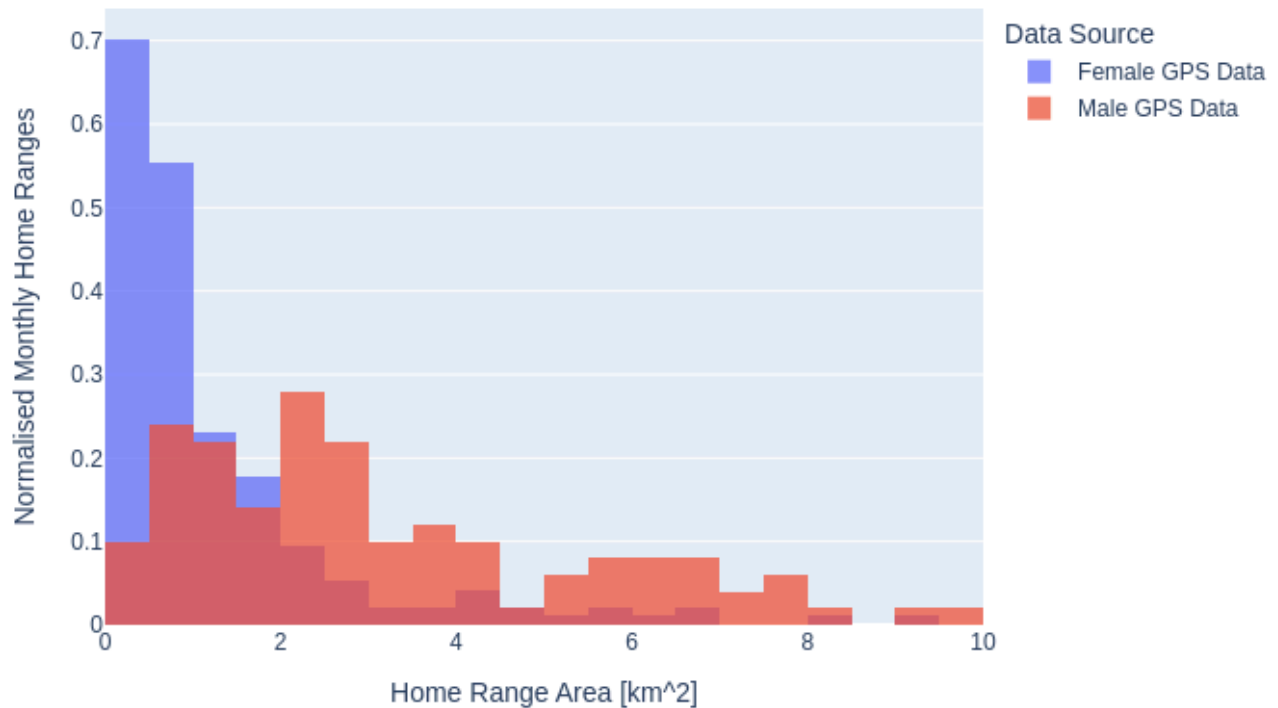
	Deer ID	time_group	0	samples_in_group	sex	area
1	5.0	2017-03	POLYGON ((1630270.227 1955816.227, 1629829.331...	729	f	1.394776
2	5.0	2017-04	POLYGON ((1630009.915 1955581.548, 1629650.187...	679	f	1.064420
3	5.0	2017-05	POLYGON ((1630377.678 1955945.985, 1629920.545...	687	f	0.722701
4	5.0	2017-06	POLYGON ((1630279.399 1956399.968, 1629956.121...	626	f	0.794714
5	5.0	2017-07	POLYGON ((1630054.77 1956285.764, 1629748.171 ...	674	f	0.859696
...	...	...	...	...	...	...
367	119.0	2018-06	POLYGON ((1633435.535 1964782.099, 1633419.265...	613	f	0.213922
370	124.0	2018-05	POLYGON ((1622365.745 1961452.582, 1622130.715...	681	f	0.315997
372	999.0	2018-03	POLYGON ((1628891.246 1963454.946, 1628816.475...	709	f	0.290556
373	999.0	2018-04	POLYGON ((1629022.295 1963442.164, 1628943.389...	717	f	0.330390
374	999.0	2018-05	POLYGON ((1629029.472 1963419.368, 1628983.163...	732	f	0.241843

294 rows × 6 columns

Home Range Area vs Number of Hours/Points in Area



## Monthly Home Range Area Histogram



The above figure shows that the male/female homeranges in the GPS data is significantly different. It would therefore make sense to model the male/female movement parameters differently.

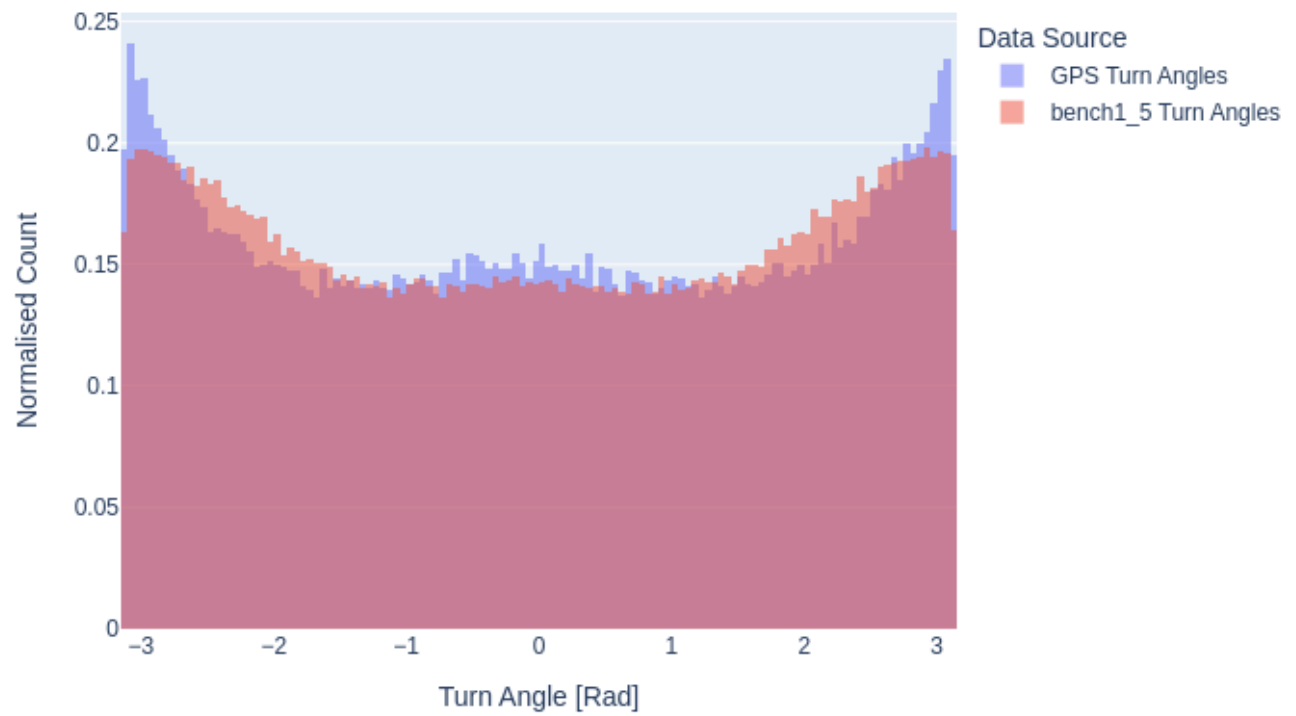
## Those same step and turn calcs again

Below is the method used to calculate step and turn values from a timeseries of points. It must be known that the values calculated are often for the previous points, so when looking at it with agents mixed up it'll just look plain wrong

## Look at Movement Params

Similar to the initial data analysis, let's take a look at the step and turn distributions and then some other statistical measures like home range size.

Normalised Histogram of Turn Angles



	timestamp	ID	Is Male	x	y	x_centroid	y_centroid
1000	2000-01-01 02:00:00	680fbfc2-053d-4c88-b522-13a7df0c2a85	False	1.601859e+06	1.965749e+06	1.601901e+06	1.965769e+06
1001	2000-01-01 02:00:00	4e5f6d7a-6462-4318-a078-abe564c52ba8	False	1.600513e+06	1.968072e+06	1.600553e+06	1.968094e+06
1002	2000-01-01 02:00:00	92a4fd6a-f2ec-4bbe-b49b-8a953d12c6ca	True	1.602121e+06	1.967023e+06	1.602072e+06	1.967010e+06
1003	2000-01-01 02:00:00	38fa7e71-3e90-4a7b-92ec-270729ba4446	False	1.602087e+06	1.965864e+06	1.602072e+06	1.965859e+06
1004	2000-01-01 02:00:00	d85986c7-22c9-45d0-aa6f-d2394ca6ce5d	True	1.601539e+06	1.965962e+06	1.601542e+06	1.966076e+06
...	...	...	...	...	...	...	...
698995	2000-01-30 03:00:00	997c6c6a-24b6-401e-b193-c60b5892b53b	False	1.608299e+06	1.964896e+06	1.607526e+06	1.964519e+06
698996	2000-01-30 03:00:00	cf21dd11-df2c-4386-b4ad-db2286db4ecf	True	1.608246e+06	1.964379e+06	1.607307e+06	1.964376e+06
698997	2000-01-30 03:00:00	7eed6b3c-9367-4c6e-b457-fa88e8a90057	True	1.606807e+06	1.963385e+06	1.607405e+06	1.962373e+06
698998	2000-01-30 03:00:00	4d2a8e83-5bb2-4f8b-a6d8-148977961ba0	True	1.609085e+06	1.964552e+06	1.608605e+06	1.963879e+06
698999	2000-01-30 03:00:00	c3882ce9-e7e2-4a2f-9217-87614cc8ed71	False	1.606717e+06	1.964936e+06	1.606561e+06	1.964735e+06

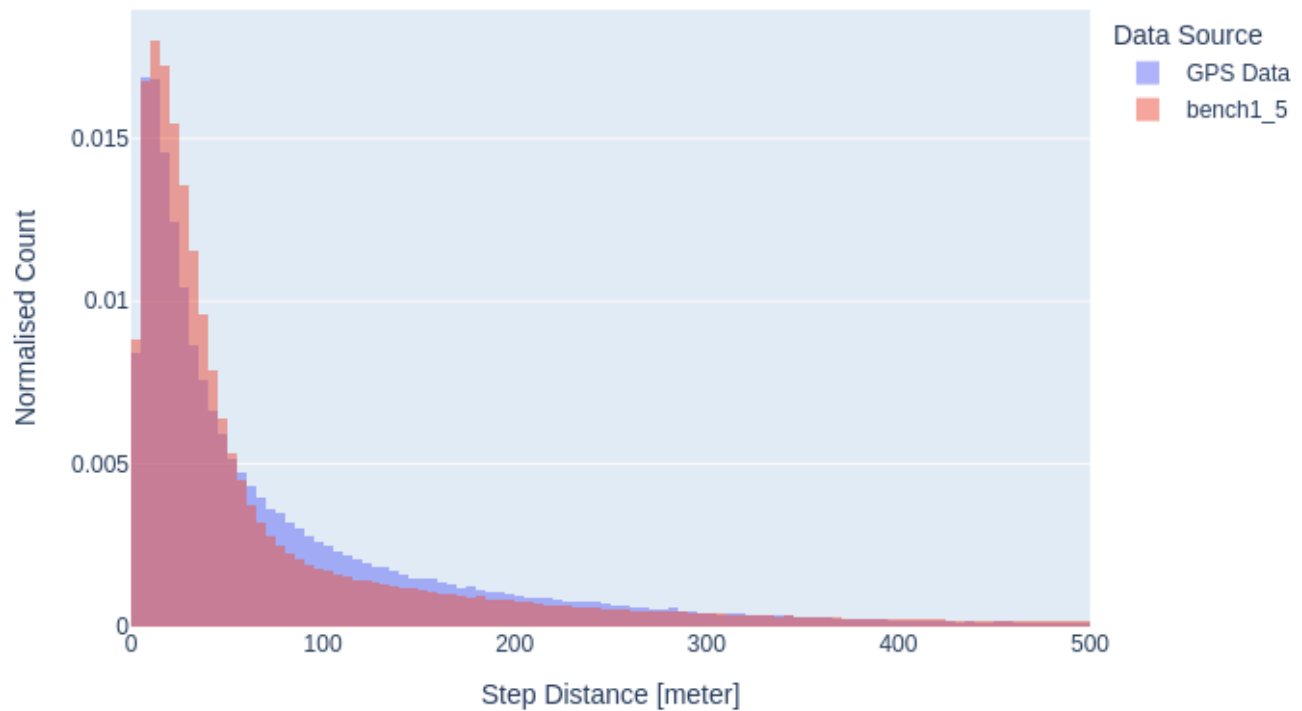
698000 rows × 19 columns

Note:



There is a spike at 0 degrees turn angle, this happens when the agent slides up against the edge of the grid boundary, the x/y of the grid value get's set to 0 (or max) and resulting turning angle becomes different from model's chosen turning angle (and is also often 0)

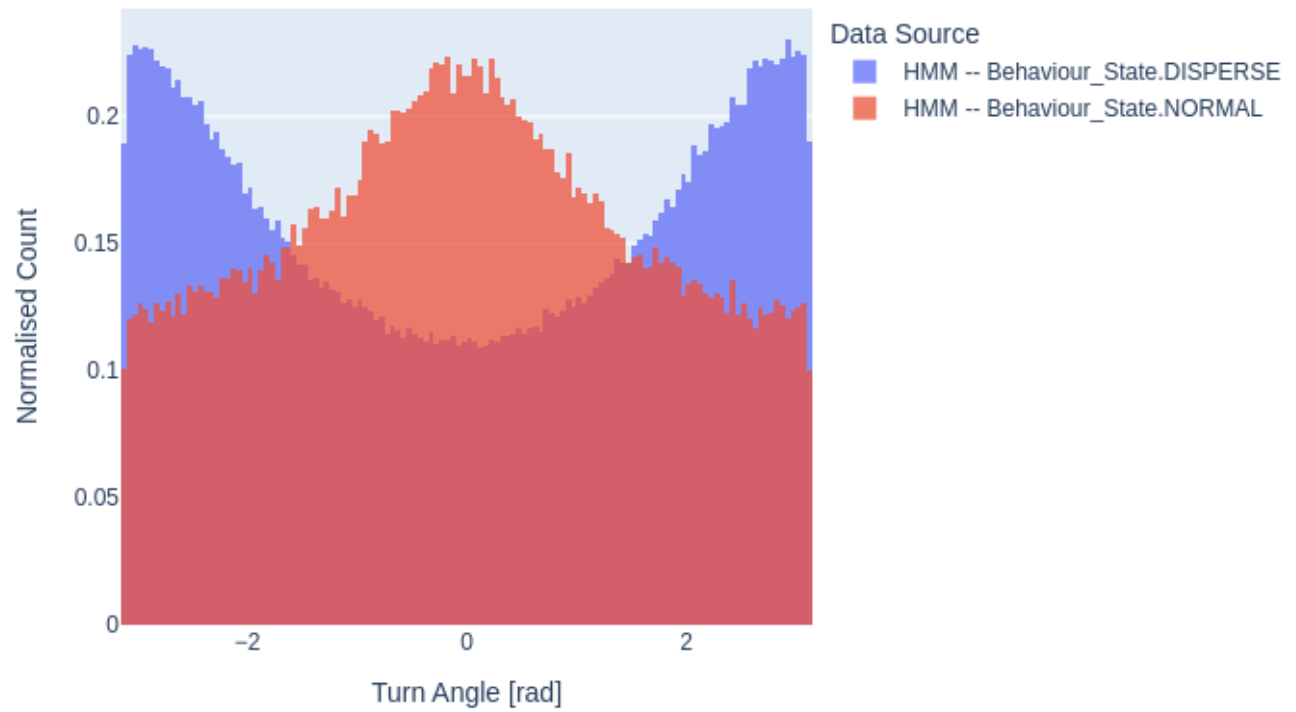
### Normalised Histogram of Step Distances



## HMM Step and Turns vs Behavioural State

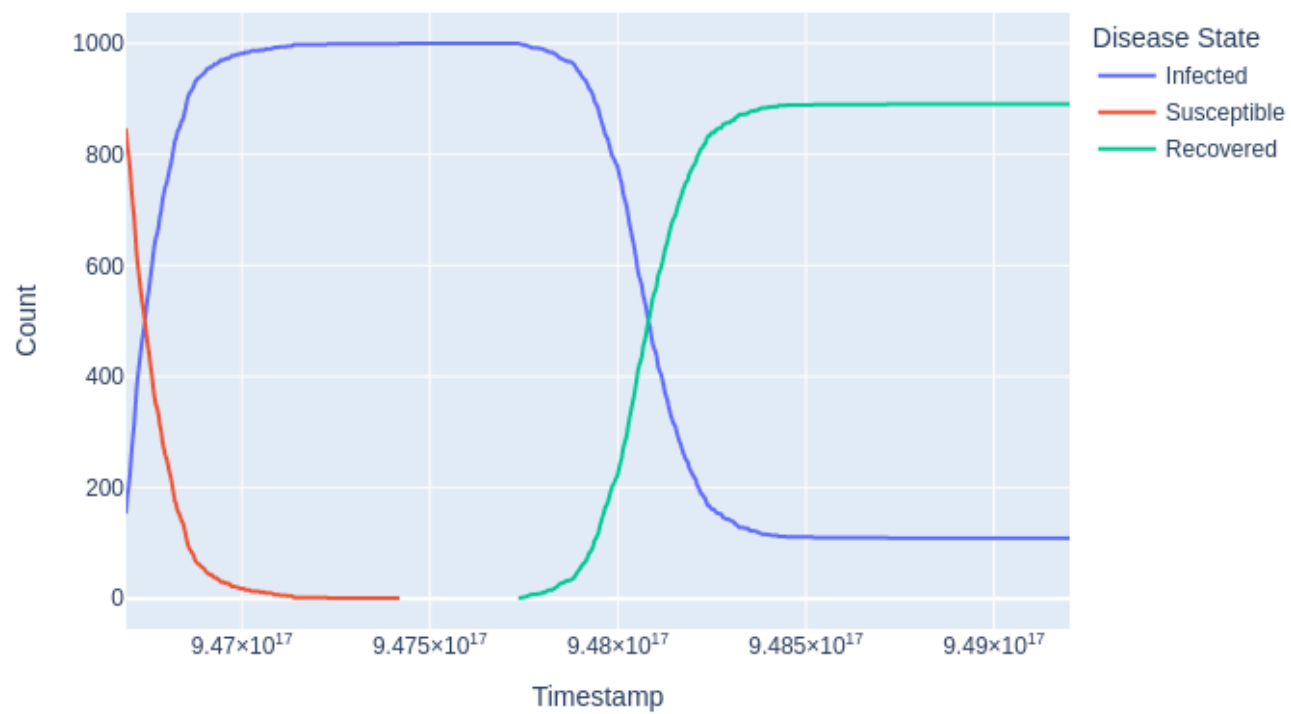
Let's compare the step and turn vs behavioural state for a specific run

## Turn Angles vs Behavioural State

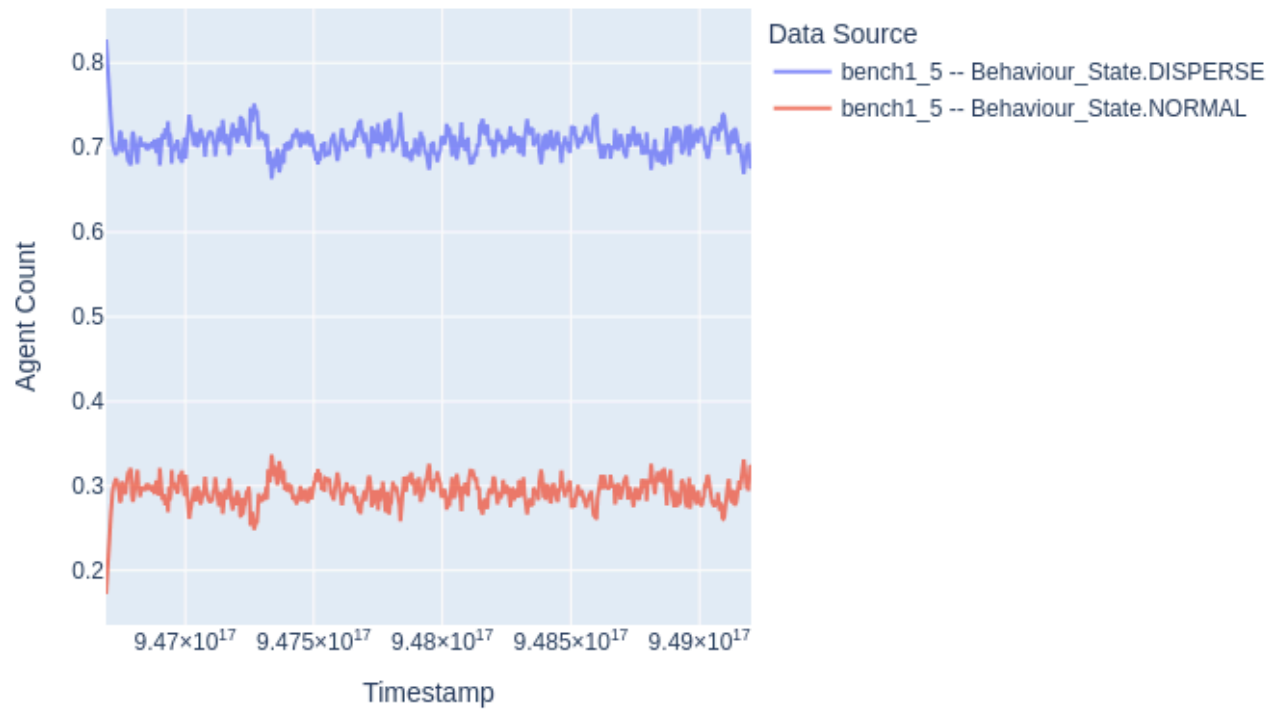


## Timeseries Plots

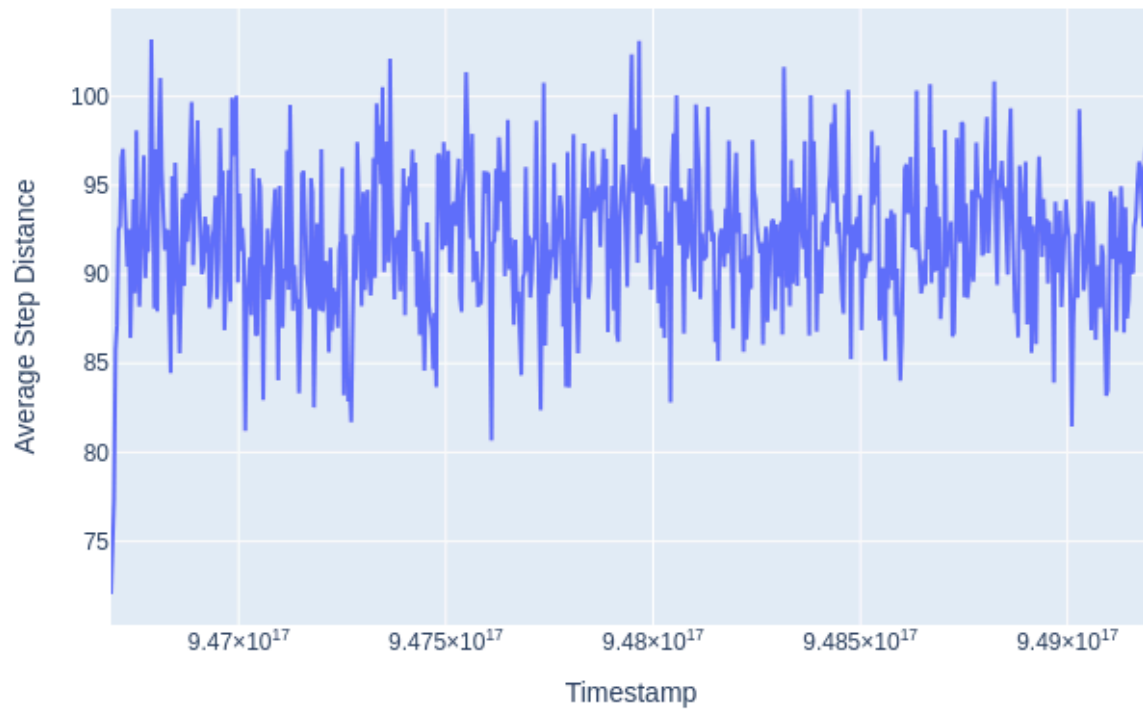
Count of Categories Over Time



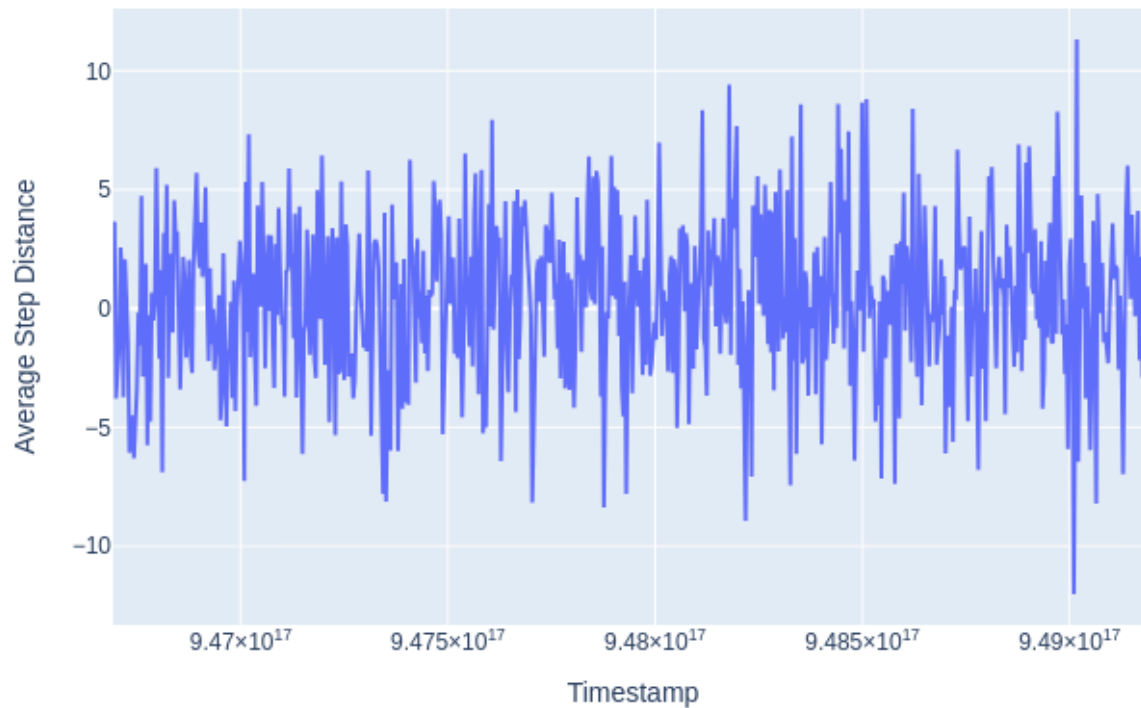
## Timeseries of Behaviour State Counts



Average Step Distance Over Time



## Average Turn Angle Over Time



```
[NbConvertApp] Converting notebook 04-Post ABM Run.ipynb to webpdf
[NbConvertApp] WARNING | Alternative text is missing on 4 image(s).
[NbConvertApp] Building PDF
[NbConvertApp] PDF successfully created
[NbConvertApp] Writing 149144 bytes to 04-Post ABM Run.pdf
```

	timestamp	ID	Is Male	x	y	x_centroid	y_centroid
0	2000-01-01T01:00:00	680bfc2-053d-4c88-b522-13a7df0c2a85	False	1.601887e+06	1.965746e+06	1.601901e+06	1.965769e+06
1	2000-01-01T01:00:00	4e5f6d7a-6462-4318-a078-abe564c52ba8	False	1.600540e+06	1.968083e+06	1.600553e+06	1.968094e+06
2	2000-01-01T01:00:00	92a4fd6a-f2ec-4bbe-b49b-8a953d12c6ca	True	1.602054e+06	1.967028e+06	1.602072e+06	1.967010e+06
3	2000-01-01T01:00:00	38fa7e71-3e90-4a7b-92ec-270729ba4446	False	1.602077e+06	1.965856e+06	1.602072e+06	1.965859e+06
4	2000-01-01T01:00:00	d85986c7-22c9-45d0-aa6f-d2394ca6ce5d	True	1.601552e+06	1.966058e+06	1.601542e+06	1.966076e+06
...	...	...	...	...	...	...	...
699995	2000-01-30T04:00:00	cf21dd11-df2c-4386-b4ad-db2286db4ecf	True	1.608287e+06	1.964410e+06	1.607307e+06	1.964376e+06
699996	2000-01-30T04:00:00	7eed6b3c-9367-4c6e-b457-fa88e8a90057	True	1.606915e+06	1.963262e+06	1.607405e+06	1.962373e+06
699997	2000-01-30T04:00:00	4d2a8e83-5bb2-4f8b-a6d8-148977961ba0	True	1.609057e+06	1.964608e+06	1.608605e+06	1.963879e+06
699998	2000-01-30T04:00:00	c3882ce9-e7e2-4a2f-9217-87614cc8ed71	False	1.606711e+06	1.965009e+06	1.606561e+06	1.964735e+06
699999	2000-01-30T04:00:00	faee93bd-1611-47fc-bf5e-6d8d01561ce8	False	1.606700e+06	1.964542e+06	1.606657e+06	1.964031e+06

700000 rows × 16 columns