Accurate statistical tests for MWM swim-tracks

The Morris Water Maze (MWM) is one of the most frequently used tests in rodent research. One important application involves spatial testing, but the analysis of the resulting swim-tracks has never been perfected. The main question here is: “are some places in the maze visited more often than expected by chance?” As is often the case, the problem lies in the definition of the null-hypothesis: what is ‘chance’. One obvious solution would be to compare to the theoretical chance of a point being visited given the number of recorded points and the total number of points. However, this approach gives no information about the variation that may be expected and furthermore, testing each pixel separately invokes correcting for multiple comparisons.

An alternative method has been developed based on a similar problem in analyzing classification images (Chauvin et al. 2005), and we now apply it to the current problem. First, we have to decide what our random distribution will be; ideally, this would resemble the original data as close as possible, except for the spatial information (compare to swimming controls). To achieve this, we create for each swim-track a so called ‘mirror’-track. These were created using the same starting position and step sizes and combining these with random directions changes (taken from a distribution based on data from over 2000 tracks, with standard deviation depending on step size; method inspired by Bovet&Benhamou, 1988). These tracks consist of the same number of points, so path length is identical; we also preserved velocity, because we use the actual step sizes.

These tracks can be used to create a heat plot (2D histogram), in the same way as for the original data. In fact, after computing 100 of these ‘random’ heat plots, we obtain an average and standard deviation that is a valid estimate of the population average (mu) and standard deviation (sigma) that would exist if no spatial preference would exist. By using mu and sigma to convert the actual heat plot to Z scores, we obtain normalized heat maps that can be compared between groups. This is valid, because each group of tracks is now expressed relative to variation that exists in random data.

Optionally, we can apply the Cluster test as was used in Chauvin et al.; effectively detecting continuous areas that are visited more often than would be expected by chance. Using these thresholded images, we can compare heat plots from different groups by calculating amount of overlap between significant regions. It would also be informative to compare for example the amount of overlap with a fixed area around to platform as a measure for spatial learning.

Figure 1: Examples of tracks with their mirrored track below. A) Both random, indistinguishable B) Wall hugging vs random => clear difference C) Clear spatial preference vs random.

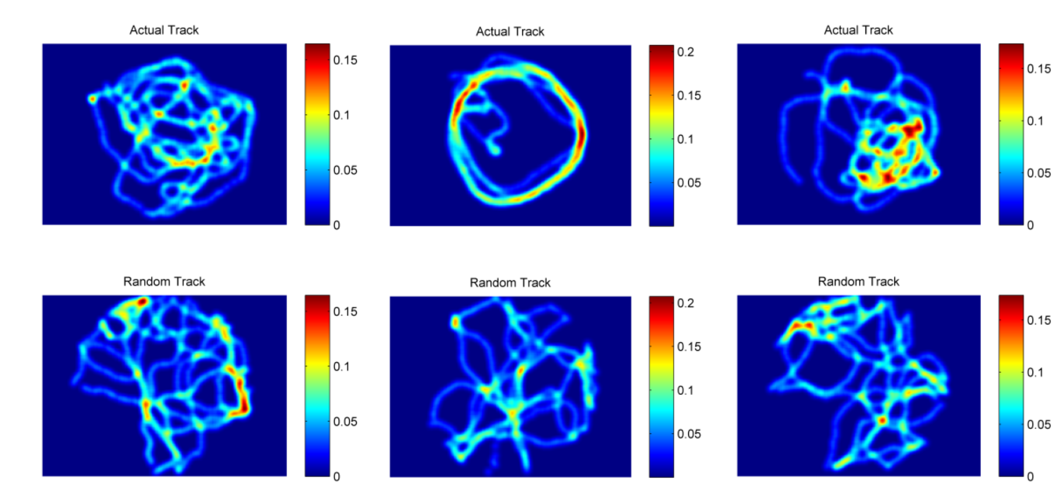

Figure 2: Resulting of the analysis for an experiment with two groups (HET versus WT) and two probe trials. A) Z scores maps each normalized to their respective random distributions and scaled to the maximum values of all 4 maps. B) Significance maps indicating area visited more (white) or less (black) then expected by chance (orange).