Principal coordinate analysis, non-metric multidimensional scaling and minimum spanning tree (the first question (1) is for exam, not the NMS question (2))

1. Produce a principal coordinate analysis of with overlain minimum spanning tree (MST) for the following data files (how many axes are needed to explain 75, 80 and 90 % of the variation?, how many axes should you retain according to the broken stick model?, is the result reasonable for Sealand distances?). Calculate the broken stick model values for the Europe flying times example.

a. Sealand distances (+ and - MST) (remember you should add a 0 km distance diagonal, and that you can make an mirror image by changing axis or by multiplying the scores by -1: MUL[-1] in transformation). Can PCO reform the MAP of Sealand by using the distances between cities? How relevant is the minimum spanning tree in this instance? Do you think it solves the traveling salesman problem? What will single linkage, average linkage and complete linkage clustering tell about Sealand?
b. Europe fly times (+ and - MST)
c. Sm.txt (compare SM with J and Y)

2. Produce a non-metric multidimensional scaling (NMS) on:

Horse shoe data (compare SM, J and Y)

Do you see a difference from the PCO data? What are the stress values?