

Spatial assignment of test sample

November 24, 2016

Contents

Input	1
Isotope values of test sample	1
Output	1
Model	1
Map of best fitted reference sample	2
Best fitted reference entries	3
Testing robustness of assignment: Wilcoxon signed rank test	3
P-values for the k nearest neighbours in Wilcoxon Test	3
Goodness of fit of test sample:	4

Input

Website Identifier: Blind 17_MW

Isotope values of test sample

Table 1: Isotope values of test sample

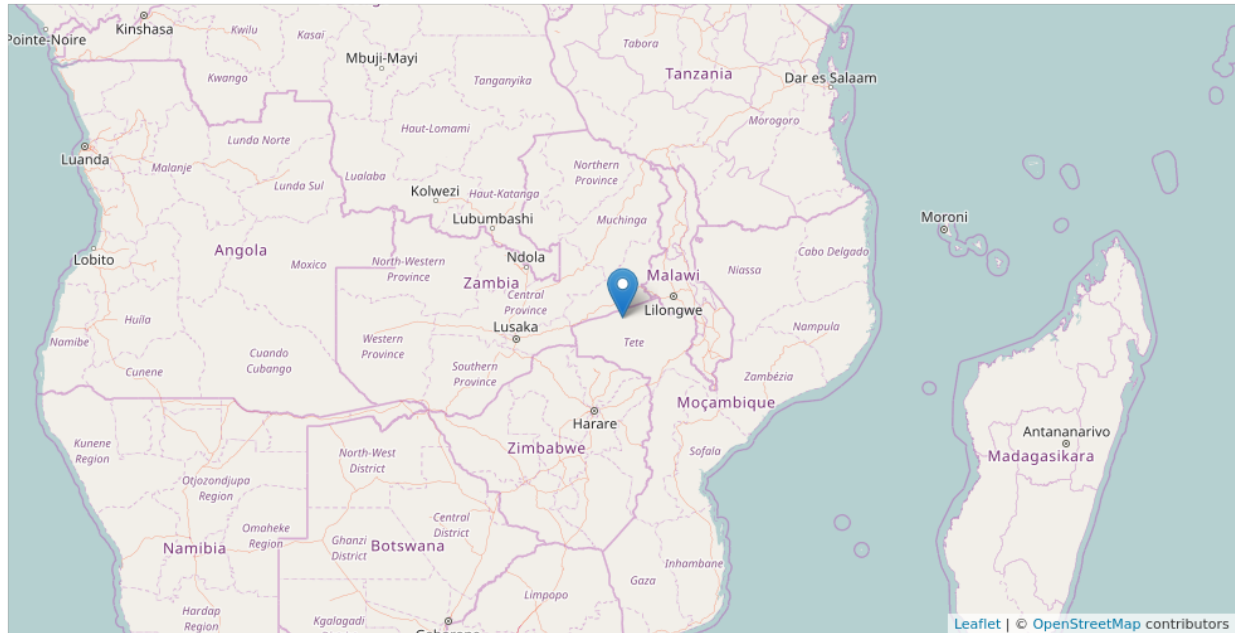
13C/12C	15N/14N	18O/16O	2H/1H	34S/32S
-20.8	7.7	15.7	-45.2	10.6

Output

Model

```
##
## Call:
## train.kknn(formula = fmla, data = ivory.train, kmax = 15, distance = 2, kernel = knl)
##
## Type of response variable: nominal
## Minimal misclassification: 0.3765867
## Best kernel: triangular
## Best k: 15
Classifier: country_code
```

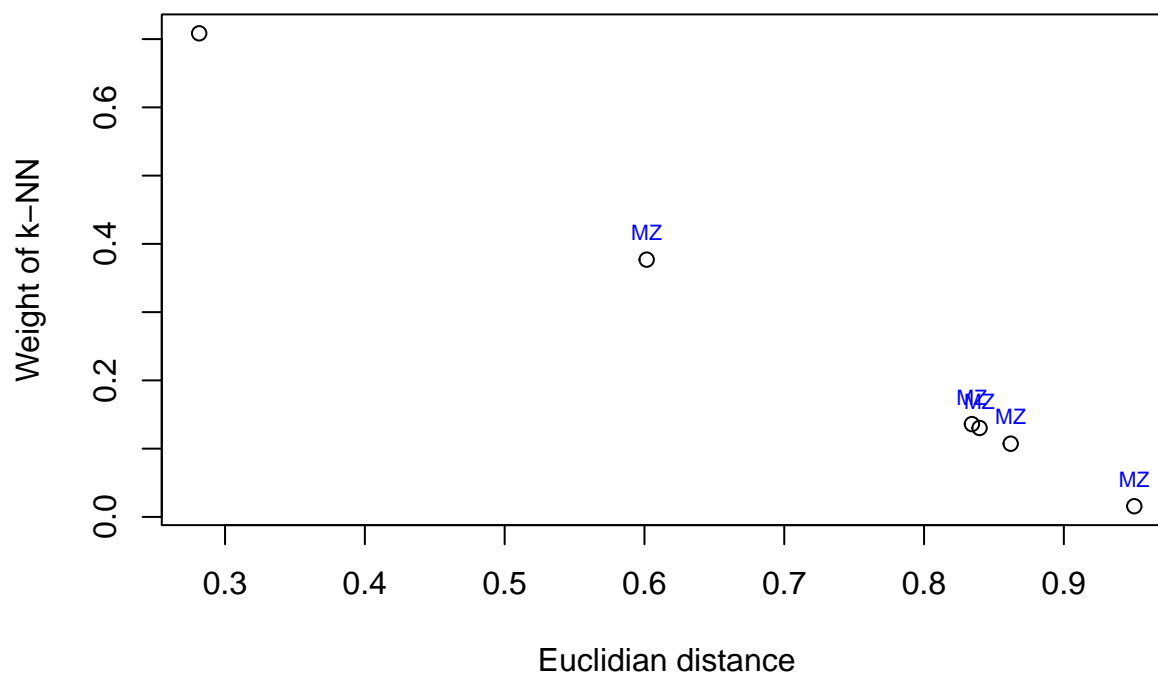
Map of best fitted reference sample



Best fitted reference sample:

- Site: Mozambique, Kambako
- Country: MZ
- Region: Southern Africa
- CITES: Appendix I
- Lat: -14.75
- Lon: 32

Assignment of test sample to nearest neighbours



Best fitted reference entries

Table 2: Details of best fitted reference entry (row 1) and other fitted entries within best classifier

lon	lat	location	13C/12C	15N/14N	18O/16O	2H/1H	34S/32S
32.00	-14.75	Mozambique, Kambako	-20.8	7.8	16.0	-45.0	9.8
32.10	-14.70	Mozambique, Kambako block - L8	-21.5	7.1	16.2	-42.3	10.8
36.86	-11.90	Mozambique, Chipuputa area (block L5)	-21.7	7.0	15.9	-39.0	10.3
38.22	-12.18	Mozambique, Lugenda south bank (block C)	-22.0	7.8	15.5	-39.5	10.1
31.15	-14.86	Mozambique, Bairro Gebeuza village	-19.7	7.3	16.4	-46.3	8.9
34.34	-18.66	Mozambique	-21.4	7.8	15.1	-50.3	8.2

Country of prediction: MZ

Testing robustness of assignment: Wilcoxon signed rank test

If $p\text{-value} > 0.05$, the test concludes that the isotope signature of the test sample is similar to the respective nearest neighbour reference sample.

P-values for the k nearest neighbours in Wilcoxon Test

“0.780, 0.451, 0.310, 0.270, 0.146, 0.057”

Goodness of fit of test sample:

- good fit: if $p > 0.05$ for at least two tested nearest neighbour reference samples;
- moderate fit: if $p > 0.05$ for at least one tested nearest neighbour reference samples;
- uncertain: if $p > 0.05$ for none of the tested nearest neighbour reference samples.

Assumption: At least two nearest reference samples are available.

Overall goodness of fit of test sample: “**good fit**”