

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 12_MW

Isotope values of test sample

Table 1: Isotope values of test sample

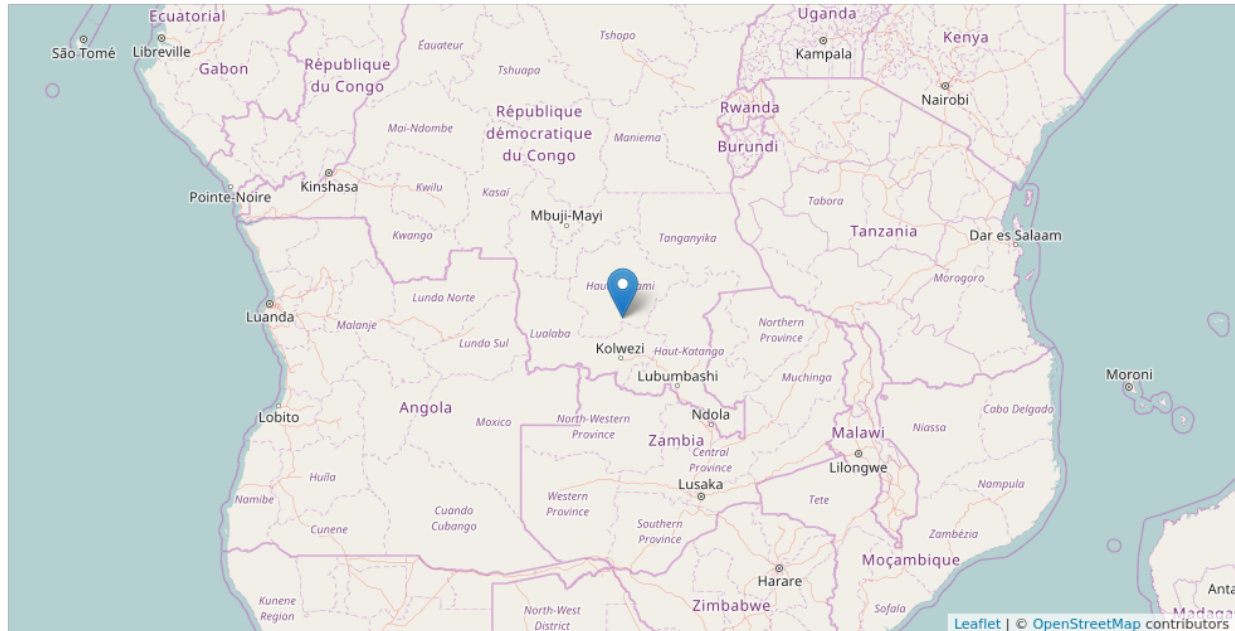
13C/12C	15N/14N	18O/16O	2H/1H	34S/32S
-18.9	7.5	16.5	-39.2	8

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
```

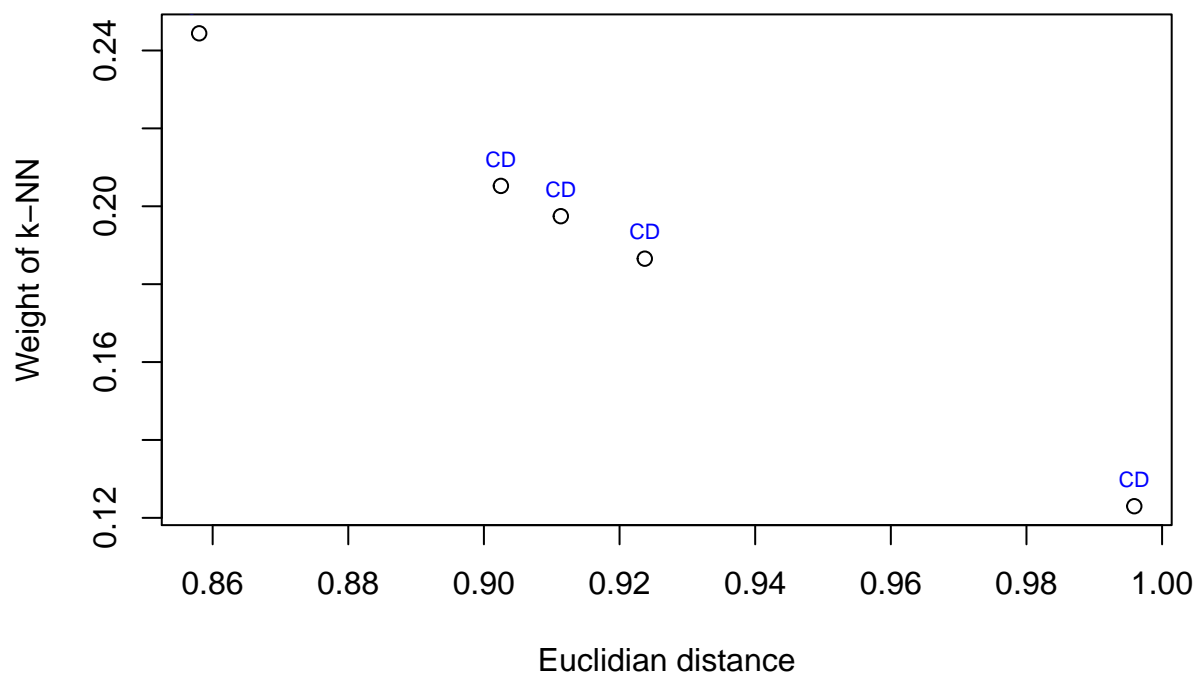
Map of best fitted reference sample



Best fitted reference sample:

- Site: Dem. Rep. Congo, Bukama
- Country: CD
- Region: Central Africa
- CITES: Appendix I
- Lat: -9.35
- Lon: 25.52

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
25.52	-9.35	Dem. Rep. Congo, Bukama	-19.2	7.3	17.4	-46.7	7.6
14.35	-5.17	Dem. Rep. Congo, Lukungu Vallei	-20.0	7.9	17.5	-43.5	7.3
19.52	4.14	Dem. Rep. Congo, Bosobolo	-20.0	7.9	17.5	-38.8	6.4
29.12	3.71	Dem. Rep. Congo, Gangala	-20.1	7.5	17.0	-33.0	7.1
29.50	0.45	Dem. Rep. Congo, Beni	-18.2	8.6	17.4	-34.3	6.7

Country of prediction: CD

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

“0.0043372, 0.0003413, 0.0002757, 0.0000406, 0.0000012”

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: “**uncertain**”