

# Spatial assignment of test sample

November 24, 2016

## Contents

<b>Input</b>	<b>1</b>
Isotope values of test sample . . . . .	1
<b>Output</b>	<b>1</b>
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: . . . . .	3

## Input

Website Identifier: 099-L

### Isotope values of test sample

Table 1: Isotope values of test sample

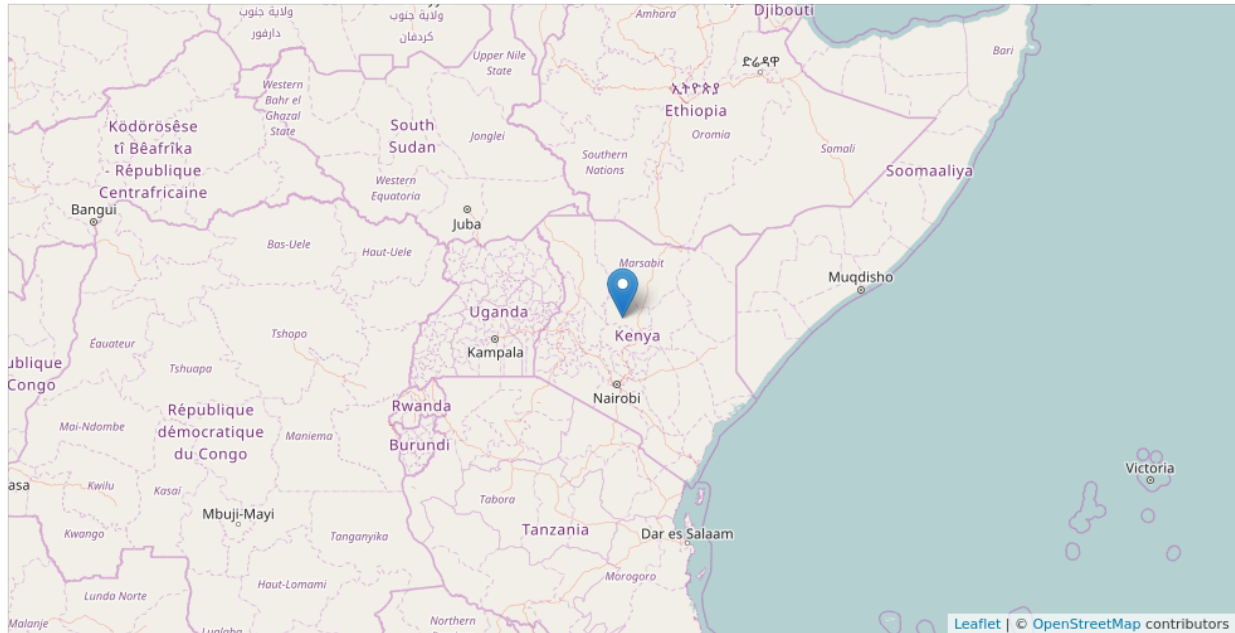
13C/12C	15N/14N	18O/16O	2H/1H	34S/32S
-19	13	21.2	-23.4	6.2

## 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
```

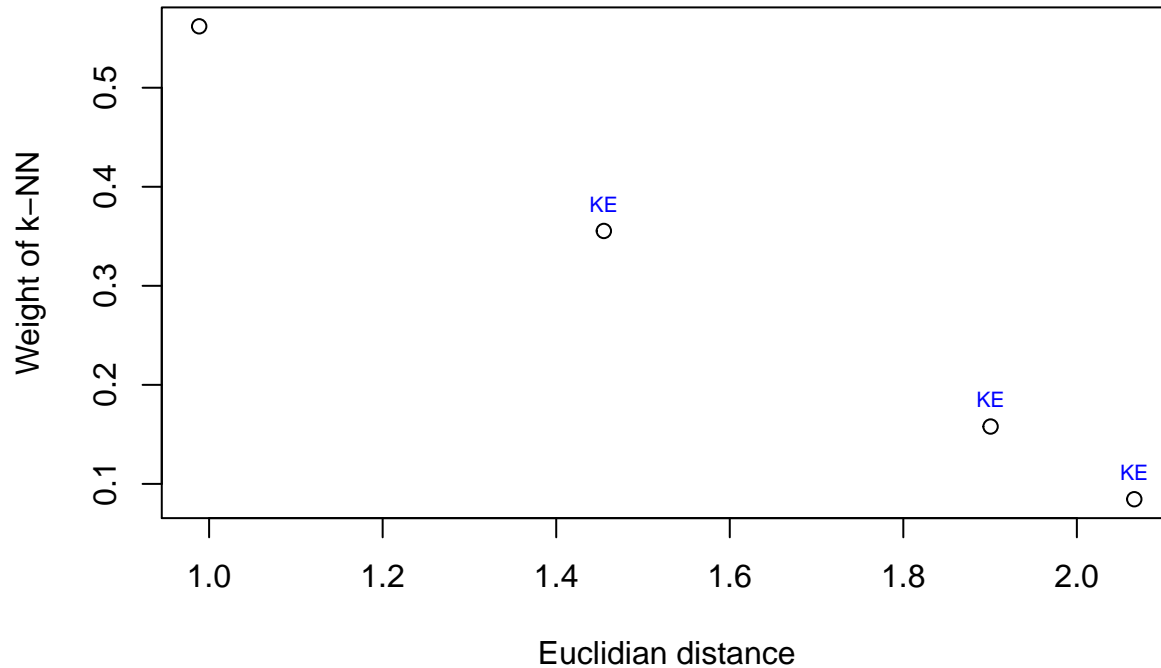
## Map of best fitted reference sample



### Best fitted reference sample:

- Site: Kenya
- Country: KE
- Region: East Africa
- CITES: Appendix I
- Lat: 1
- Lon: 37

## 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
37.00	1.00	Kenya	-18.1	12.3	20.2	-22.1	8.3
37.00	0.00	Kenya	-18.2	12.0	19.7	-25.5	9.8
39.50	0.90	Kenya, Amala River or Lorian Swamp]	-19.6	14.7	18.3	-31.0	7.8
37.35	0.71	Kenya	-17.5	12.1	17.7	-29.2	4.8

Country of prediction: **KE**

### Testing robustness of assignment: Wilcoxon signed rank test

If  $p$ -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

“0.0932157, 0.0091146, 0.0003413, 0.0000096”

### 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: “**moderate fit**”