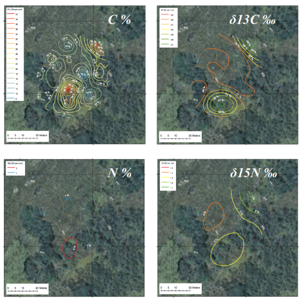
Sampling strategies in a forest environment for the elaboration of Isoscapes.†

Marco Ciolfi,[a] Francesca Chiocchini,[a] Giuseppe Russo[a] Luciano Spacchino,[a] Michele Mattioni,[a] Mauro Lauteri.[a]



This document introduces a sampling technique especially developed for the field collection of soil and plant samples in a complex forest environment. The technique consists in the definition of a spiders' web georeferenced lattice to be used as a reference frame on the field collection. The web lattice points are equally distributed in polar coordinates, thus providing a finer coverage of the central area, maintaining a good rate up to the boundary. Such a spatial distribution is specially suited to many geostatistical algorithms including the widespread Kriging or Inverse Distance Weighting, which are commonly used in the production of Isoscapes (isotopic abundances maps).
A dedicated application, *Webnodes*, has been developed for the production of web sampling lattices. It is also included in the document.
Keywords: Sampling strategies, Isoscapes, Software

Introduction

The so-called **Isoscapes** (from ***Iso**topic land**scape***) are thematic maps interpolated from the spatial distribution of the relative abundances of stable isotopes.[1] In forest ecology, in particular, the hydrogen (δD or δ^2H), carbon (δ 13C), nitrogen (δ^{15}N) and oxygen (δ^{18}O) isotopes have a widespread use. Isoscapes can be the starting point for the analysis of the isotopic fractioning which is met in many areas of research.[2–4]

The distribution maps are usually generated by geostatistics modules within an GIS (Geographic Information System) package, as Qgis, GRASS, ESRI's ArcGIS, and R, which is not itself (better, not only) an GIS, but a complete statistical ecosystem.[5–7] It is also possible to build more complex models, including the time variability of the measured quantities, combining time series and time interpolation at the same level.[8–14]

1 Sampling needs

The best known spatial interpolation techniques include IDW (Inverse Distance Weighting, exact interpolation) and Kriging (stochastic interpolation),[15] anyway, some mandatory requirements for any good spatial model,[16] include:

1. Measurements have to be georeferenced (spatial coordinates);
2. Measurements should have a time reference (timestamp)
3. The spatial distribution of the collected samples must allow the application of the geostatistical technique of choice;
4. The geolocalisation can be performed on the field (GPS) or on the desk, at a later time, from orthophotogrammetric surveys or topographic maps.

The forest environment, on the other hand, poses some constraints due to the very character of the forests themselves, as:

1. The GPS coverage, due to the tree crown and countless poter obstacles, is often poor or below the precision needs;
2. The complex morphology of the terrain does not allow, in most cases, the pre-definition on the map (or GIS) of a complete set of sampling sites which are really accessible during the data collection.
3. generally, more than an hair-thin, absolute spatial precision, it is the reciprocal position of the elements which is mandatory. The relative position of the samples can be easily measured regardless the GPS, using elementary field instruments such as a tape metre, protractor and compass;
4. The sampled areas, for the sake of the maximum spatial precision, should be circular in shape, or at least roughly polygonal, but the nature of the terrain can constraint the sampling to non-optimal areas;
5. Data from older researches, traditionally collected along linear transects, is not always fit for spatial interpolation (from a transect one can obtain profiles but not full two-dimensional maps).

It is clear from the above points that the ideal sampling geometry is a circle since it is the convex shape with the maximum *internal area / perimeter* ratio. Using such a circular distribution as a guide, we developed a technique that consists in the

[a] *CNR - Istituto di Ricerca sugli Ecosistemi Terrestri (ex CNR-IBAF), Viale G. Marconi, 2 – 05010 Porano (TR), Italia*
Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale
† The development of this technique and the accompanying software Webnodes is part of the CNR-IBAF project grant IBAF003/2013/POR (2013-2017) and the CNR-IBAF *Geostatistical Computing Facility* special grant (2015-2016).

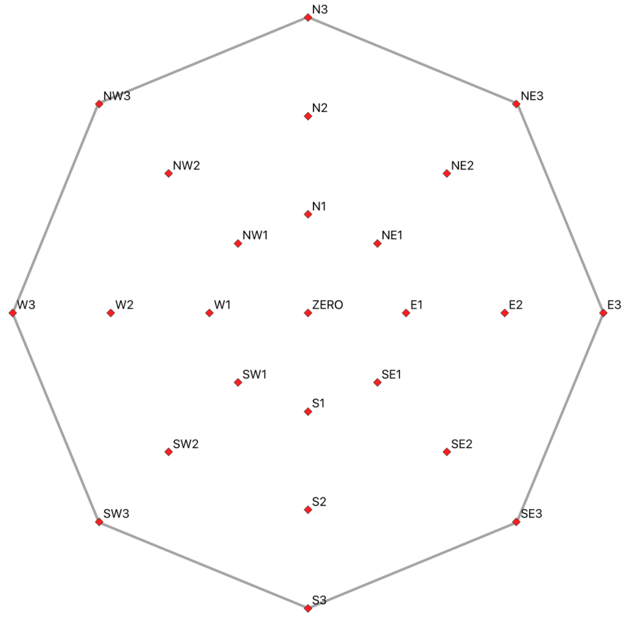


Fig. 1

predisposition of approximately circular areas,[i] easily measurable on the field, flexible enough to be changed on the filed to match the actual terrain.

2 Georeferenced sampling technique

The proposed sampling pattern consists in defining a circumference centered about the point (or the area) of interest. Several points need to be placed within the circle boundary.[ii] Setting the points along a family of concentric octagons one obtains a regular starry pattern as in the figure; the points are located along the principal (N, E, S, W) and intermediate (NE, SE, SW, NW) directions, plus the center point. Practically, one has to generate 8n+1 points within the circle, n per direction. As for attains the areas, the underestimate relative to the whole circle is very low. [iii]

The useful surface for the geostatistical elaboration of the Isoscapes is the Convex Envelope (CE) of the pois set, i.e. the whole octagon.

Employing this technique, one has not to measure the coordinates of all the points while on the field; only the coordinates of the center (ZERO in figure 1) has to be measured, typically by a GPS device, the points on the arms can be easily found with a simple compass and a tape metre. It could be advisable to shift the sampling points only if the terrain morphology requires so, i.e. just in case of unpredicted obstacles. This sampling pattern is particularly suited for soil, grasses and other uniformly distributed specimens.

It is very easy to prepare some field sampling sheets, as the one depicted in figure 2 below, for the recording of the space-related informations.[iv]

i The circle is compatible with the forse inventory standard plots. If one needs a map which includes all the plot, it is advisable to extend the sampling for a few metres beyond the limits of the plot, or about 5%, normally one or two metres.
ii Consider at least thirty points for a good Kriging procedure.
iii An octagon occupies a bit more than the 90% of the area of the surrounding circle.
iv A sample sheet is included as extra material.

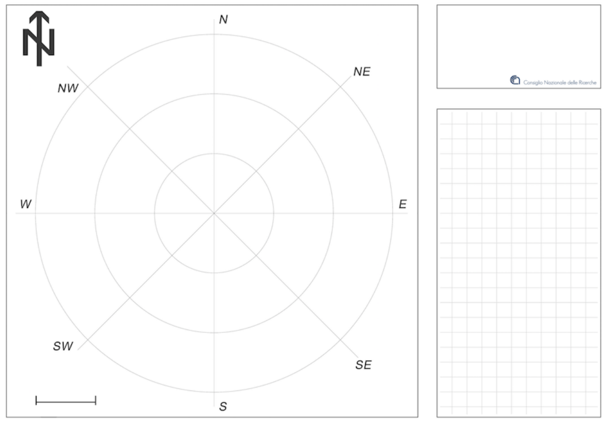


Fig. 2

Given d, the unit distance along the direction axes, as d = R / n, where R is the radius of the circumference and n is the number of points per arm, the 8n+1 points are easily obtained.[v] Arm by arm:

- Center: ZERO = [XZERO, YZERO]
- North: Nk = [XZERO, YZERO + k d]
- North-East: NEk = [XZERO + γ k d, YZERO + γ k d]
- East: Ek = [XZERO + k d, YZERO]
- South-East: NEk = [XZERO + γ k d, YZERO - γ k d]
- South: Sk = [XZERO, YZERO - k d]
- South-West: NEk = [XZERO - γ k d, YZERO - γ k d]
- West: Wk = [XZERO - k d, YZERO]
- North-West: NWk = [XZERO - γ k d, YZERO + γ k d]

where $\gamma = 1/\sqrt{2}$ is the common value of the sine and cosine of $\pi/4$ ($45°$), circa 0.71.

3 The Webnodes application

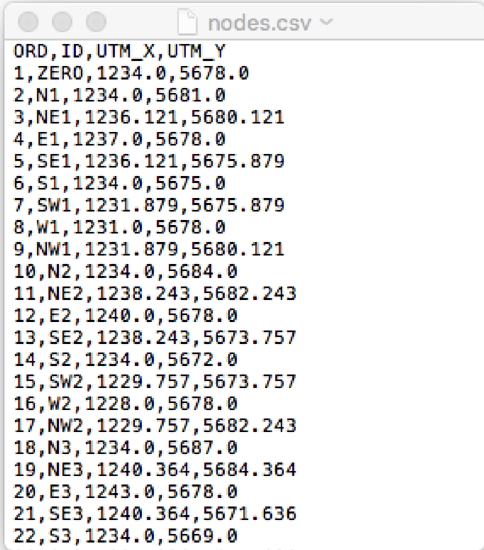
A simple stand-alone application, **Webnodes**, has been developed to speed up the simple but tedious calculations described in the previous section.

The **Webnodes** application has been developed in Java, so it runs on every operating system running a Java Virtual Machine.[vi] The program consists in a single class, which opens a three-tabbed window: SQUARE, SPIDER WEB and NODES. The SQUARE tab allows the definition of a regular squared lattice; the SPIDER WEB tab allows the definition of the octagonal "spider's web" described above, with or without the intermediate arms; the NODES tab shows the points generated by the application, allowing the clipboard copy and the ASCII .csv export. Such points can be imported in any spreadsheet or GIS, specifying the coordinates reference system.

The application output is a plain ASCII .csv text; the decimal separator is the dot (.) while the field separator is the comma (,). The fields are:

- ORD: Progressive number, starting from one.
- ID: Alphanumeric id, following the schema ZERO (central point), or DIRECTION n for the arms.
- UTM_X: horizontal coordinate.

v For example, when n = 4 one has 33 sampling points, enough for a good Kriging.
vi The full code is given in the appendix and as extra material as well. The application executable is also included as extra material.



```
ORD,ID,UTM_X,UTM_Y
1,ZERO,1234.0,5678.0
2,N1,1234.0,5681.0
3,NE1,1236.121,5680.121
4,E1,1237.0,5678.0
5,SE1,1236.121,5675.879
6,S1,1234.0,5675.0
7,SW1,1231.879,5675.879
8,W1,1231.0,5678.0
9,NW1,1231.879,5680.121
10,N2,1234.0,5684.0
11,NE2,1238.243,5682.243
12,E2,1240.0,5678.0
13,SE2,1238.243,5673.757
14,S2,1234.0,5672.0
15,SW2,1229.757,5673.757
16,W2,1228.0,5678.0
17,NW2,1229.757,5682.243
18,N3,1234.0,5687.0
19,NE3,1240.364,5684.364
20,E3,1243.0,5678.0
21,SE3,1240.364,5671.636
22,S3,1234.0,5669.0
```


Fig. 3

- UTM_Y: vertical coordinate.

Where UTM is referred to the *Universal Transverse Mercator* projection, but many other projections can be used, like the Lambert conic, used in many European projects, and many other projection systems of local interest. It is also possible, in principle, to use angular coordinates, but intros case the distance between points on the same arm is not constant and the diagonal arms do not point exactly towards the right directions.

The SQUARE point lattice is far from optimal from a geostatistical point of view, but it its very useful to prepare the sampling along North-South pr West-East transects.

The user interface consists of three tabs: two input tabs (**SPIDER WEB** and **SQUARE**) and an output one (NODES). Users should provide only the coordinates of the centre, labelled *ZERO-X* and *ZERO-Y*, the number of points for each arm *NUM* and the lattice distance between points *SIZE*.

The *GENERATE* button starts the evaluation of the lattice points, switching automatically the interface to the **NODES** tab; from the latter it is possible to export the results on a *.csv* formatted file or into the system clipboard.

3.1 The SPIDER WEB tab

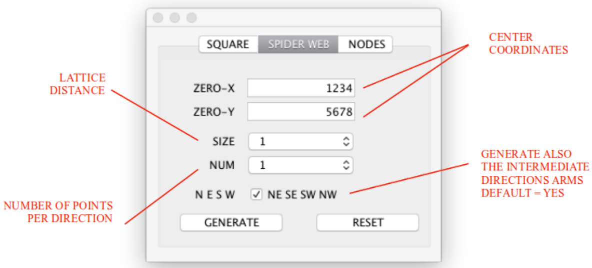


Fig. 4

3.2 The SQUARE tab

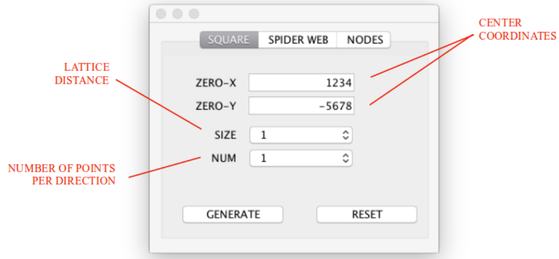


Fig. 5

3.3 The NODES tab

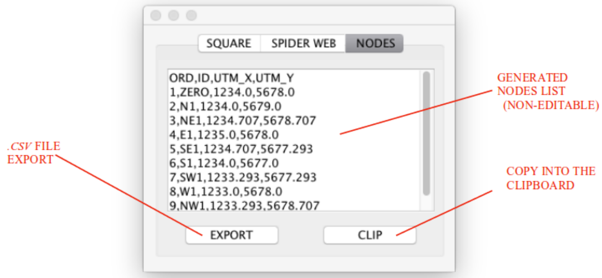


Fig. 6

3.4 Lattice points import

The points evaluated by **Webnodes** can be read almost automatically importing them into a spreadsheet. According to the local settings, some software packages (Microsoft Excel in particular) can misinterpret the dot and comma separators.

The main goal of **Webnodes** is the generation of points to be plugged into an GIS workflow, where a .csv import is generally straightforward. As an example, in the following figure it is shown the detailed import procedure in Qgis (versione 3.0).

The import into a different GIS environment follows the same logic, with unimportant local variations. Note that while importing the lattice points it is mandatory to specify the correct coordinates projection system (step 5 of figure 8). **XWebnodes** evaluates the numerical values only and it is completely projection-blind.

The result of the import procedure is a new layer, generally it is labelled as a *temporary or scratch* layer: it is advisable to save it in a suitable format as a shapefile or as a geodatabase object, according to the user's taste, following the chosen GIS procedures.

4 Use cases

The following examples ave been chosen to show the versatility of the above described sampling technique. All these use cases come from actual CNR-IBAF projects involving the stable isotopes sampling and statistical elaboration.

ORD	ID	UTM_X	UTM_Y
1	ZERO	1234	5678
2	N1	1234	5681
3	NE1	1236	5680
4	E1	1237	5678
5	SE1	1236	5676
6	S1	1234	5675
7	SW1	1232	5676
8	W1	1231	5678
9	NW1	1232	5680
10	N2	1234	5684
11	NE2	1238	5682
12	E2	1240	5678
...	...	...	...
...	...	...	...

Fig. 7

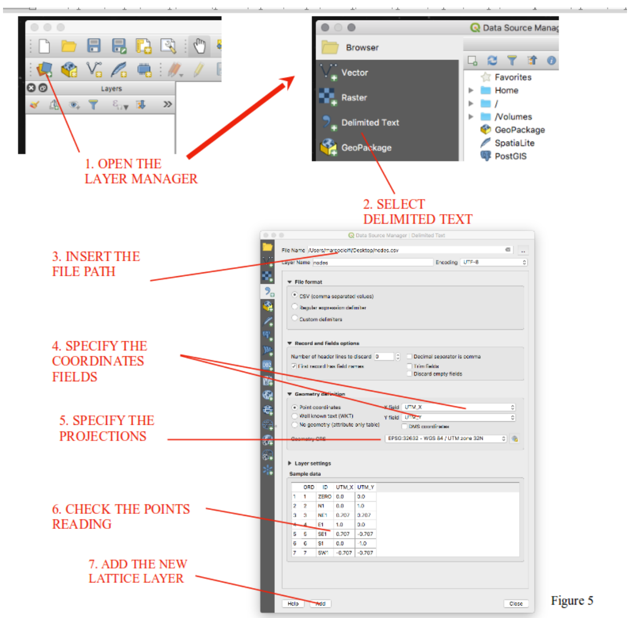


Fig. 8

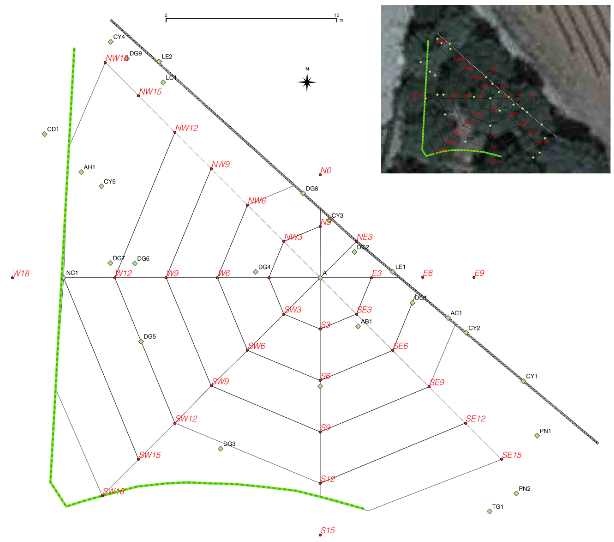


Fig. 9

4.1 Sampling lattice for a single *Ailanthus altissima* tree

The sampling has been conducted on an Ailanthus altissima, (main tree and its resprouts, soil, other plants nearby) placed in a flat but hardly reachable area, surrounded by a hedgerow (West and South bounds, in green) and a fence (Northeast boundary, in grey) which are obviously limiting collection of the samples.

his is an example of an extremely detailed scale, needing a sub-decimetre precision, so that the GPS localisation is not enough. The location of the ZERO point has been evaluated by surveying techniques, than the nodes on the arms have been generated by Webnodes (step 2m, nodes in red). Prior to the sampling the nodes falling off the boundaries have been deleted.

4.2 $\delta^{15}N$ multiple Isoscapes of the forest soil

This sampling has been conducted within a complex forest ecology study. Among the other materials, soil has been sampled on nine different spots, centred about pivotal trees, which coordinates have been taken by GPS (area T1 to T9 of figure 8). The areas aere about 12 to 25 metres wide, the pivotal trees are tens to hundreds metres apart.

The spiders web lattice nodes have been actually signalled by planting poles in the ground. The step size employed is five metres. The nodes have been recalculated with **Webnodes** to build up the samples geodatabase (other than the soil, fungi and tree leaves have been collected in each area).

This demonstrates that the web lattice can also be calculated after that the sampling has taken place, limiting the within-area localisation error to a few centimetres, while the area-to-area error is about two metres, according to the GPS reading.

Using multiple webs on smaller areas the researchers can focus on the actual interesting spots, thus saving on time and sampling budget, without affecting the statistical quality of the results.

Two Kriging calculation strategies have been cross-checked: one the one hand, an overall Kriging, bounded by the blue line of figure 11 (the convex envelope of all the points), has been

Fig. 10

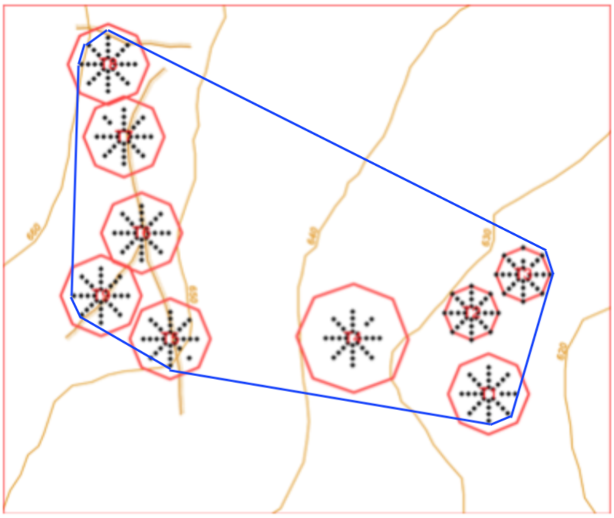


Fig. 11

computed.[vii] On the other hand a separate Kriging has been computed for each of the octagonal areas (the dots in the red areas show the position of the poles).

Clipping sub-maps from the overall Kriging and confronting them with the single maps obtained area by area has shown no significant differences of the values, thus showing the stability and reliability of the sampling technique. Figure 12 shows the result of the area by area Kriging.

4.3 Carbon and Nitrogen soil characterisation

The last case is is about an high mountain field sampling for the isotopic characterisation of the soil and the plants (*Polylepis reticulata* trees and *Calamagrostys intermedia* "pajonal" meadows). The site, in the Equatorial Andes, is challenging by a climatic, goemorphological and logistic point of view, so the collection had to be as fast and precise as possible. Also in this case, only a few pivotal coordinates have been measured by

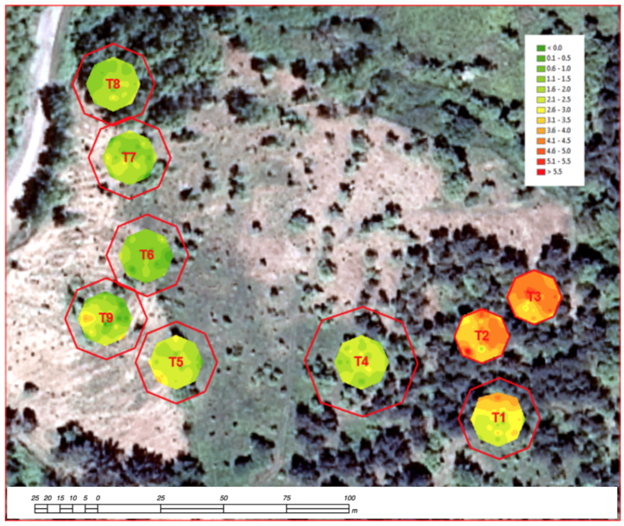


Fig. 12

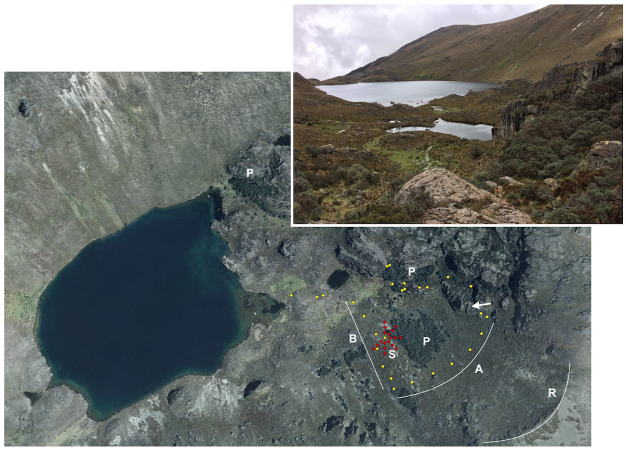


Fig. 13

GPS, operating with an above-average precision. The lack of any previous knowledge of the area did not allow the preparation of the sampling lattices before the campaign.

A few areas in three distinct high altitude catchments have been sampled, at an elevation of 3800 to 4300 m. All the areas have been chosen at the *Polylepis* woods - pajonal meadows interface.

As a typical example, figure 13 shows the sampling web (S) in red. Note that the sampling is incomplete but the resulting dataset was enough for some good statistics. The soil has also been sampled along maximum slope (B) and same-elevation (A) transects, shown by yellow dots.[viii] The lattice size was 10 metres.

After the sampling, te web lattices have been reconstructed with **Webnodes**, centered about the sampled centres of the areas.

The result of the Kriging is shown in figure 14. Both the elemental composition (C% and N%) and the isotopic abundance ($\delta^{13}C$ and $\delta^{15}N$) of the soil have been interpolated.[ix]

vii Not all the points within the convex envelope have reliable values of the Kriginig: the North-East area far from the sampling spots, although within the C.E., cannot be trusted geostatistically.

viii The arrow in the orthophoto shows the inset picture point of view. P = Polylepis reticulata woods, R is part of the catchment upper rim.

ix These images show the equal-value curves derived from the Kriging, which output is normally a raster format GIS layer.

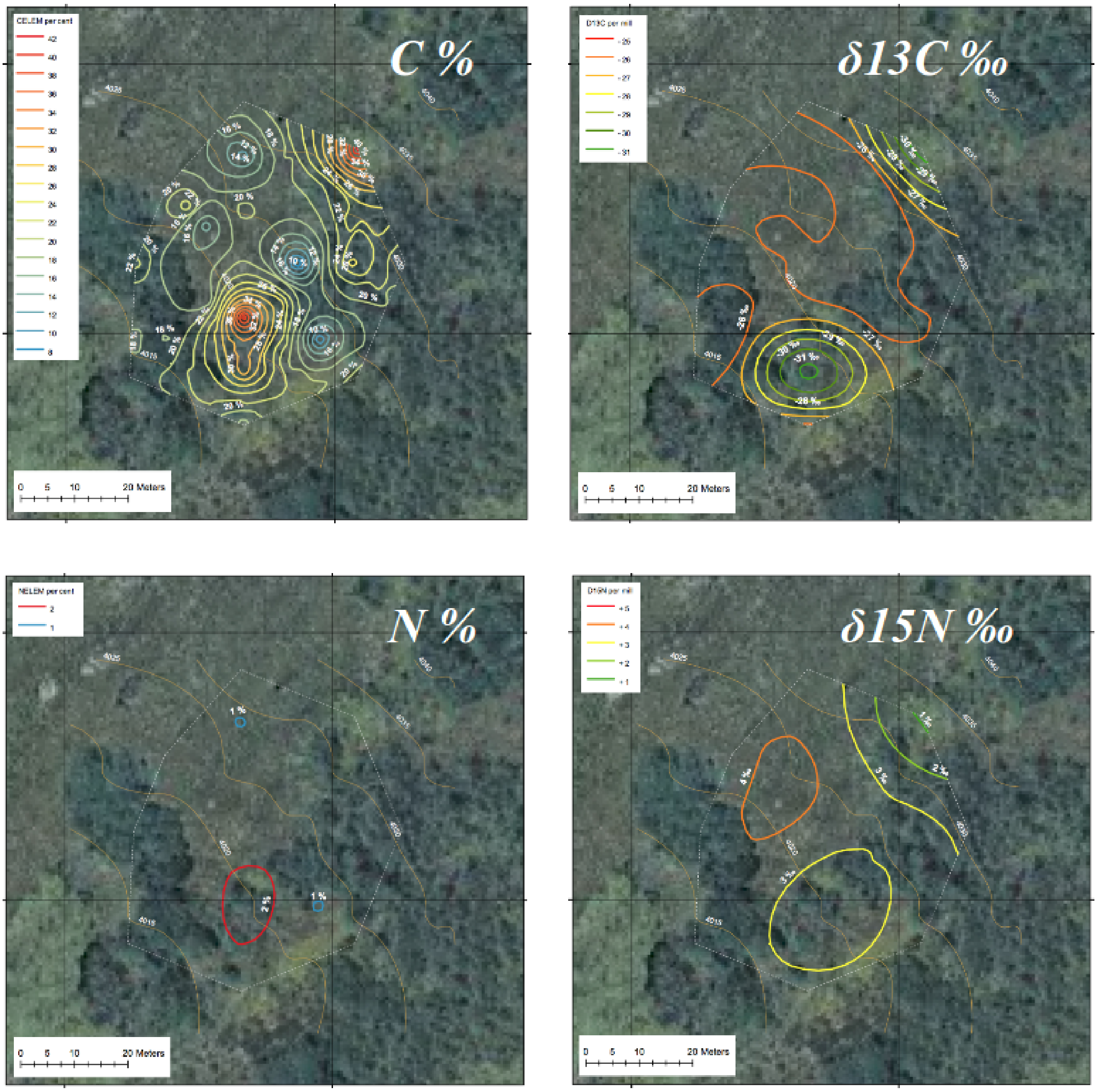


Fig. 14

Note that, due to the partial or complete lack of the western arms' samples, the interpolated region is slightly smaller than the complete octagon that could have been obtained given a complete coverage. Nonetheless, the results are satisfactory, thus proving again the reliability of the spiders' web sampling technique.

References

1 J. B. West, G. J. Bowen, T. E. Dawson, K. P. Tu, Understanding movement, pattern, and process on Earth through isotope mapping, Springer, 2010. doi:10.1007/978-90-481-3354-3.

2 E. Brugnoli, G. D. Farquhar, Photosynthetic Fractionation of Carbon Isotopes, Springer Netherlands, Dordrecht, 2000, Ch. 17, pp. 399–434. doi:10.1007/0-306-48137-5_17.

3 B. Peterson, B. Fry, Stable isotopes in ecosystem studies, Annual Review of Ecology and Systematics 18 (1987) 293–320.

4 P. Renard, H. Demougeot-Renard, R. Froidevaux, Geostatistics for Environmental Applications, Proceedings of the Seventh European Conference on Geostatistics for Environmental Applications, Springer Nature, 2010. doi:10.1007/978-90-481-2322-3.

5 D. Renard, Roger s. bivand, edzer j. pebesma, virgilio gomez-rubio: Applied spatial data analysis with r, Mathematical Geosciences 43 (5) (2011) 607–609. doi:10.1007/s11004-011-9340-y.

6 D. Borcard, F. Gillet, P. Legendre, Numerical Ecology with R, Springer, 2011. doi:10.1007/978-1-4419-7976-6.

7 N. A. C. Cressie, Statistics for Spatial Data, John Wiley & sons, 1990. doi:10.1002/9781119115151.

8 N. Cressie, C. K. Wikle, Statistics for Spatio-Temporal Data, John Wiley & sons, 2011.

9 J. D. Cryer, K.-S. Chan, Time Series Analysis: With Applications in R, Springer-Verlag, 2008. doi:10.1007/978-0-387-75959-3.

10 M. Goodchild, R. Haining, S. W. et al., Integrating gis and spatial data analysis: problems and possibilities, International Journal of Geographical Information Systems 6 (5) (1992) 407–423. doi:10.1080/02693799208901923.

11 A. D. Jassby, T. M. Powell, Detecting changes in ecological time series, Ecology 71 (6) (1990) 2044–2052. doi:10.2307/1938618.

12 P. Legendre, L. Legendre, Numerical Ecology, Vol. 24, Elsevier, 2012.

13 W. Pedrycz, S.-M. Chen, Time Series Analysis, Modeling and Applications, Vol. 47, Springer-Verlag, 2013. doi:10.1007/978-3-642-33439-9.

14 P. Turchin, A. D. Taylor, Complex dynamics in ecological time series, Ecology 73 (1) (1992) 289–305. doi:10.2307/1938740.

15 M. R. Dale, M.-J. Fortin, Spatial analysis: a guide for ecologists, Cambridge University Press, 2011.

16 D. R. Visscher, Gps measurement error and resource selection functions in a fragmented landscape, Ecography 29 (3) (2006) 458–464. doi:10.1111/j.0906-7590.2006.04648.x.