The Lao Journal of Agriculture and Forestry. 8, 2-13. 2005

Generation of soil property maps of the Luang Prabang province. Selection of optimal interpolation techniques.

Konngkeo PHACHOMPHON[1], Norbert SILVERA[1,2], Xaysana Xayarath[3], Vincent Chaplot[1,2*]

1. SSLCC_NAFRI, Vientiane, LAO-PDR
2. IRD-Ambassade de France, BP 06, Vientiane, LAO-PDR. Tel office: (+856) 520 50 77; Fax: (+856) 21 41 29 93; E-mail: chaplotird@laopdr.com
3. CIAT, Vientiane, LAO-PDR.

*** corresponding author**

ACKNOWLEDGMENTS
The authors are indebted to the SSLCC from NAFRI (Vientiane, LAOS) for providing data points. Special thanks go to IWMI, the international water management institute for funding the training workshop on "Selection of optimal interpolation techniques at different scales" (NAFRI, 4-14 November 2003) during which data analysis and interpretation were performed.

ABSTRACT
The knowledge of the spatial distribution of soil properties over large areas is crucial for a better inventory of natural resources. Soil property maps would allow for instance to improve land-use planning as well as environment modeling. In this study at a regional scale our objectives were (i) to generate soil property maps; (ii) to define the optimal interpolation technique for each soil property. The study was conducted in the Luang Prabang province (19,149 km^2) of Northern Laos. A generation (115 points) and a validation set (25 points) were randomly extracted from a total data set of 140 data points. From the generation set, soil property maps of pH, CEC, clay and carbon (C) content for the surface A horizon as well as of soil depth (SD) were generated using the ArcMap software. Interpolation techniques include inverse distance weighting (IDW) with 12 or 3 neighbors; spline; ordinary and universal kriging (OK; UK); global polynomial (GP); and radial basis (RB). The validation procedure was performed using Arc-View3.2. Statistical parameters of mean error (ME) and mean absolute error (MAE) allowed to access maps accuracy. IDW techniques were shown to give in all cases accurate spatial estimation of all study soil properties. In the case of lower sampling density only (45 data points instead of 115), GP and kriging were more accurate. Deep, organic and clayed soils are mostly observed in the North and the south. Greater pH and CEC values occurred in the south-west only. The described process allowed the authors to make an optimal valorization of existing soil observations and measurements. This process could now be applied at the Laos scale to support decision making in environment.
***Key words**: map; soil properties; interpolation; Laos*

1. INTRODUCTION

Because environmental scientists and planners need to be able to make informed decisions at the regional scale, there is a crucial need to obtain spatial information of natural resources. Among natural resources, soils and especially soil properties are key information. These informations allow for example the estimation of the land ability to some cropping systems, the soil water availability, the soil susceptibility to erosion as well as the optimal areas for water waste treatment. In addition to this, soil properties maps are an essential basis information for environmental modeling. Indeed, erosion or hydrologic models use soil property maps as input layers, for the ultimate estimation of the quantity and quality of water and sediments outputs from watersheds.

Quantitative information on the map accuracy is also a key issue. Indeed, it will allow to define possible limitations for further uses of maps and the confidence for this. In the case of erosion or hydrologic modeling, outputs depend to a great extent on how well model input spatial maps describe the relevant characteristics of the watershed. It is assumed that high quality maps would allow the accurate estimation of runoff and sediment yield models for instance. The estimation of erosion risks at the province level will request high precision maps of soil properties.

The improve of the map quality could be achieved by increasing the density of observation. This is the prevailing way. But, in many cases, increasing the sampling density is very costly especially in the case of large areas. An other possibility to improve the map accuracy is to use interpolation techniques that would minimize errors commonly produced during the map generation processes.

Many interpolation techniques are available and were used with or without success for soil properties (Odeh et al., 1994, 1995; Bourennane et al., 1996; Chaplot et al., 2000) or other attributes such as topography (e.g. Desmet, 1997). At the regional scale the authors' belief is that the optimal technique will mainly depend on the considered soil property.

In this study of the Luang Prabang province (19,149 km^2, northern Laos) we compared the quality of soil properties maps for soil depth, pH, CEC, clay and carbon content using several interpolation techniques including inverse distance weighting, spline, ordinary and universal kriging, global polynomial and radial basis. The validation procedure was performed by using an independent data set.

2. Materials and methods

The description of the study site

The study area is located in Northern Laos. It is a 19,149 km^2 area corresponding to the Luang Prabang (LPB) province. The perimeter of the province is 770 km. The UTM coordinates (zone 48) of this province ranges from 145,235 to 322,471 m for longitudes and from 2,102,456 to 2,338,915 m for latitudes. Altitudes in the province are presented in Figure 1. The mean value was 790 m, altitudes ranging between 206 and 2263 m. Higher altitudes are located in the eastern part where linear mountainous structures are oriented along a north-east/south-west axis. A large depression, similarly oriented, is observed in the western part where the Mekong river flows. These topographic structures are probably controlled by tectonism.

Mean slope gradients over the province are presented in figure 2. Slope gradients ranged from 0 to 108 % with a mean value of 18.5 %. Maximum values are located either in mountainous areas and at the border of main depressions.

The most common rock types in LPB are gabro, diorite, andesite and basic rock, schist, gneiss, sandstone (Department of Geology and Mines,1990-1991). A detailed soil survey conducted by The Soil Survey and Land Classification Center (SSLCC) in 1996 gathered that the most widespread soil groups are Acrisols, Alisols (FAO UNESCO/Systems) or Ultisols (US.Taxonomy). They are mainly found on the slopes ranking from 8% to 50%, which made up most of surveyed area.

Due to the natural conditions and especially due to topographic limitations, most of the province is mainly cultivated following slash and burn practices. Low land rice production is restricted to main depressions. According to the Department of Forestry, (MAF,1998) 53% of the hole population of 210,600 people is engaged in shifting cultivation. The province exhibit a wet-dry monsoon. The dry season occurs from November to March. This period is cold and mostly dry. The wet season characterize the months between April to October. It is hot and humid. Mean annual precipitation at the city of Luang Prabang over the last 30-years was 1403 mm.

The map generation using interpolation techniques

It is beyond the scope of this paper to give a review of all interpolation methods which can be used to construct soil maps. Interpolations chosen here are the most common ones found in the bibliography. They include inverse distance weighting (IDW) with 12 or 3 neighbors; spline; ordinary and universal kriging (OK; UK); global polynomial (GP); and radial basis (RB). All techniques were available in the ArcMap software.

Inverse Distance Weighted (IDW) assumes that each input point has a local influence that diminishes with distance. It weights the points closer to the processing cell greater than those farther away. A specified number of points, or optionally all points within a specified radius, can be used to determine the output value for each location. More distant locations have less influence. In this study we used 12 and 3 points. The case of three points correspond to a linear triangulation technique. The IDW was the regularized type with a power of 2. This power controls the significance of the surrounding points upon the interpolated value, a higher weight resulting in less influence from distant points.

Spline interpolator fits a minimum-curvature surface through the input points. Spline methods produce a function that passes as close as possible through the map generation points and still maintains a certain degree of smoothness. The Regularized method was considered here with a weight of 0.1, defining the weight of the third derivatives of the surface in the curvature minimization expression.

A second family of interpolation methods consists of geostatistical methods that are based on statistical models that include autocorrelation (statistical relationships among the measured points). They are exact interpolation techniques, i.e., the surface goes through each measured sample value. Like IDW interpolation, kriging forms weights from surrounding measured values to predict values at unmeasured locations. The closest measured values usually have also the most influence but Kriging weights come from a semivariogram that was developed by looking at the spatial structure of the data. This techniques makes the hypothesis that pairs of points that are close in distance should have a smaller difference than those farther away from one another. The extent that this assumption is true can be examined in the empirical semivariogram. The model that better fit the semivariogram is further used for data point interpolation. Semivariograms and fitted models for all soil properties were estimated through ArcMap spatial analyst extension. Ordinary kriging (OK) and universal kriging (UK) have been thoroughly described in the soil science literature (Burgess and Webster, 1981; Cressie, 1991). The predicted soil property using OK at an unknown point used only the observed values from the neighboring sample points and combined them linearly with weights derived from the inference of the experimental variogram. Ordinary kriging (OK) assumes a constant mean. The fitted model for all soil properties was spherical and 12 points were selected for radius settings. Universal kriging (UK) assumes the mean to be a

deterministic function. The fitted model for all soil properties was linear with linear drift and 12 points were identically selected for radius settings.
Global polynomial (gp) interpolation bases the prediction on the overriding trend, fitting a plane between the sample points. It is an inexact interpolators since the surface do not pass through the measured points. A plane is a special case of a family of mathematical formulas called polynomials. A first order gp interpolation was used for soil properties.

Radial basis (rb) interpolation is another an exact interpolation technique. Among the five different basis functions available in ArcMap (thin-plate spline, spline with tension, completely regularized spline, multiquadric function, and inverse multiquadric function) we used the completely regularized spline one. RB interpolation is conceptually similar to fitting a rubber membrane through the measured sample values while minimizing the total curvature of the surface. The selected basis function determines how the rubber membrane will fit between the values. Twelve neighbors were used for interpolation.

From a total data set of 140 points, a sub-set of 115 points was randomly selected for map generation, the remaining points aiming at validated the maps.

Soil property maps for soil depth, pH, CEC, clay and carbon content over the LPB province by using several interpolation techniques including inverse distance weighting, spline, ordinary and universal kriging, global polynomial and radial basis were generated using the ArcMap software (ESRI, 2002).

Quality assessment
Estimations for soil depth, pH, CEC, clay, carbon content using the 7 interpolation techniques at the 25 validation data points were compared to observations at the same data points. A quality assessment was performed using several statistical parameters: the mean error (*ME*), the mean absolute error (*MAE*) between estimated estimations and observations (Eq. 1 and 2).

$$ME = \frac{1}{n}\sum_{i=1}^{n} [(\mathrm{Var})^{*} - \mathrm{Var}] \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left|(Var)^{*} - Var_{20-m}\right| \quad (2)$$

MAE and ME decrease with increasing regression model accuracy. The MAE expresses the degree to which the interpolated value differs from the reference value whereas the ME evaluates the extent to which errors are around the zero value.

Finally, using the optimal interpolation techniques defined for each soil property, precise maps were established from the total data set of 140 points. These last maps are presented in this paper (see figures 3 to 7).

3. Results and discussion

3.1. The spatial distribution of soil property in the LPB

The soil depth
The spatial variations of the soil depth are presented in Figure 3. Soil depth varies from 0.45 to 1.25 m with an average of 1.07 m. The higher depths are located along a north-east axis. A large area with shallow soils characterizes the center of the region. Some spots with depths lower than 0.6 meter are encountered in the valley bottoms of the center and south. This is

surprising since valley bottoms are intended to accumulate sediments eroded from hillslopes. One explanation is that the presence of coarse elements such as big stones did not allow field scientists to reach the soil bottom with the auger hole.

The pH of A horizons

The spatial variations of the pH of A horizons are presented in Figure 4. The pH values range from 4 to 8 m with a mean of 5.6. Values gradually decrease from 8 in a limited area of the south-west to 4 in the north-east where deep soils are encountered. Beyond expectations, limits between pH values are globally perpendicular to the axis of main topographic structures. However, such a gradient could be associated to a gradient of agricultural; intensification along a similar axis, calcium inputs being preferentially added in the south.

The clay content of A horizons

The spatial variations of the clay content of A horizon are presented in Figure 5. Clay content varies from 5.5 to 51.2 % with an average of 25.2 %. Minimal values (less the 20 %) are mostly situated in the center east of the province where shallow soils were also observed. Greater contents occur in the south-west and the north-east of the region. Similarly, a high correlation exist between deep and clayed soils.

The carbon content of A horizons

The spatial variations of the carbon content of A horizons are presented in Figure 6. Carbon content varies from 0.015 to 3.99 % with an average of 1.83 %. The general distribution of carbon content is similar to this of the soil depth distribution, i.e. high concentrations along the north-east and a large area with low C content in the center of the region.

The CEC of A horizons

The spatial variations of the CEC of A horizons are presented in Figure 7. CEC varies from 4.45 to 42.8 with an average of 15.2. The higher concentration is situated in the south-west of the region where high pH values were also observed.

3.2. The map accuracy as function of the interpolation technique

The accuracy of interpolation techniques for the estimation of the spatial distribution of soil properties is presented in Table 1.

For pH, CEC, clay and carbon (C) content for the surface A horizon as well as of soil depth (SD), mean errors (ME) and mean absolute errors (MAE) were the lower for IDW techniques. For instance, MAE for SD was of only 1.15 cm for IDW12. In the case of the carbon content MAE was 0.03%, i.e., 2% only of C average at the province scale.

Better results for IDW in the case of soil properties was not expected. ME for

Indeed, many authors already demonstrated the interest of geostatistics but on smaller areas or at lower sampling density. This is confirmed by additional results (not presented here) at lower sampling density (45 data points) showing that IDW was not the most accurate technique. In this case, global polynomial and kriging gave better results for all study soil properties.

CONCLUSION

In this study of a regional scale, our main objective was to generate soil property maps from 140 existing data points by using optimal interpolation techniques. The study was conducted in the Luang Prabang province (19,149 km^2) of northern Laos where soil property maps of pH, CEC, clay and carbon (C) content for the surface A horizon as well as of soil depth (SD) were established using interpolation techniques including inverse distance weighting (IDW); spline; ordinary and universal kriging (OK; UK); global polynomial (GP); and radial basis (RB). The validation of these maps revealed that at an observation density of 140 data points, IDW was the most accurate technique. When decreasing the sampling density until for instance 45 data points, global polynomial and kriging became the best interpolation techniques.

In this study we used a process making the best possible use and at low cost of existing soil observations and laboratory measurements. Such maps would allow to support decision making in environment at the province scale.

Soil property maps could now be generated at the country scale since a point data base is already available and GIS tools allow to consider large areas.

As an other perspective, the improvement of the interpolation process could also be cited. For example, this interpolation process could benefit from a better knowledge of the relationship between the soil distribution in landscapes and environmental factors such as topography or climate. These relationships could directly be included during the map generation through techniques lik e as co-kriging or kriging with an external drift (Odeh et al., 1994, 1995; Bourennane et al., 1996; Chaplot et al., 2000).

REFERENCES

Bourennane, H., King, D., Chery, P., Bruand, A., 1996. Improving the kriging of a soil variable using slope gradient as external drift. European Journal of Soil Science, 47, 473-483.

Chaplot V., Walter C., Curmi P. 2000. Improving soil hydromorphy prediction according to DEM resolution and available pedological data. Geoderma. 97, 405-422.

Desmet, P. J. J. 1997. Effects of interpolation errors on the analysis of DEMs. Earth surface processes and landforms. 22.563-580.

Division of Agriculture and Forestry Office, 1998. Agricultural Statistic of Luangprabang district and province.

Department of Geology and Mines ,1990-91. Geological map of Lao PDR, at scale 1:000.000. Atlas des Ressources Physique Economique et Social du Bassin Iferieur du Mekongand 1.500.000, scale 1:500.000.

Odeh, I. O. A., McBratney, A. B., Chittleborough, D. J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma, 63, 197-214.

Odeh, I. O. A., McBratney, A. B., Chittleborough, D. J., 1995. Further results on prediction of soil properties from terrain attributes-heterotropy cokriging and regression-kriging. Geoderma, 67, 215-226.

Table 1. Mean error (ME) and mean absolute error (MAE) between estimations and observations for soil pH, CEC, clay and carbon (C) content of the surface A horizon as well as soil depth (SD) as function of the used interpolation technique: inverse distance weighting (IDW) with 12 or 3 neighbors; spline; ordinary and universal kriging (OK; UK); global polynomial (GP); and radial basis (RB). ME and MAE were computed from the validation set of 25 data points.

ME		IDW12	IDW3	spline	OK	UK	GP	RB	mean
	SD	0.12	4.25	-0.15	0.65	0.22	1.41	-0.36	107
	pH	-0.03	-0.03	-0.01	0.06	0.10	0.13	0.03	5.48
	Clay	-0.28	-0.31	-0.34	1.80	1.54	2.27	0.51	23.64
	C	0.02	0.02	-0.01	0.18	0.23	0.16	0.07	1.64
	CEC	-0.11	-0.13	0.24	-0.12	0.60	0.11	-0.02	8.96
MAE									
	SD	1.15	23.02	43.61	7.40	22.97	27.64	10.33	107
	pH	0.04	0.03	0.12	0.43	0.49	0.55	0.22	5.48
	Clay	0.64	0.56	1.88	6.33	6.47	7.55	3.04	23.64
	C	0.03	0.03	0.15	0.56	0.54	0.59	0.25	1.64
	CEC	0.22	0.17	0.62	2.22	4.09	4.64	0.73	8.96

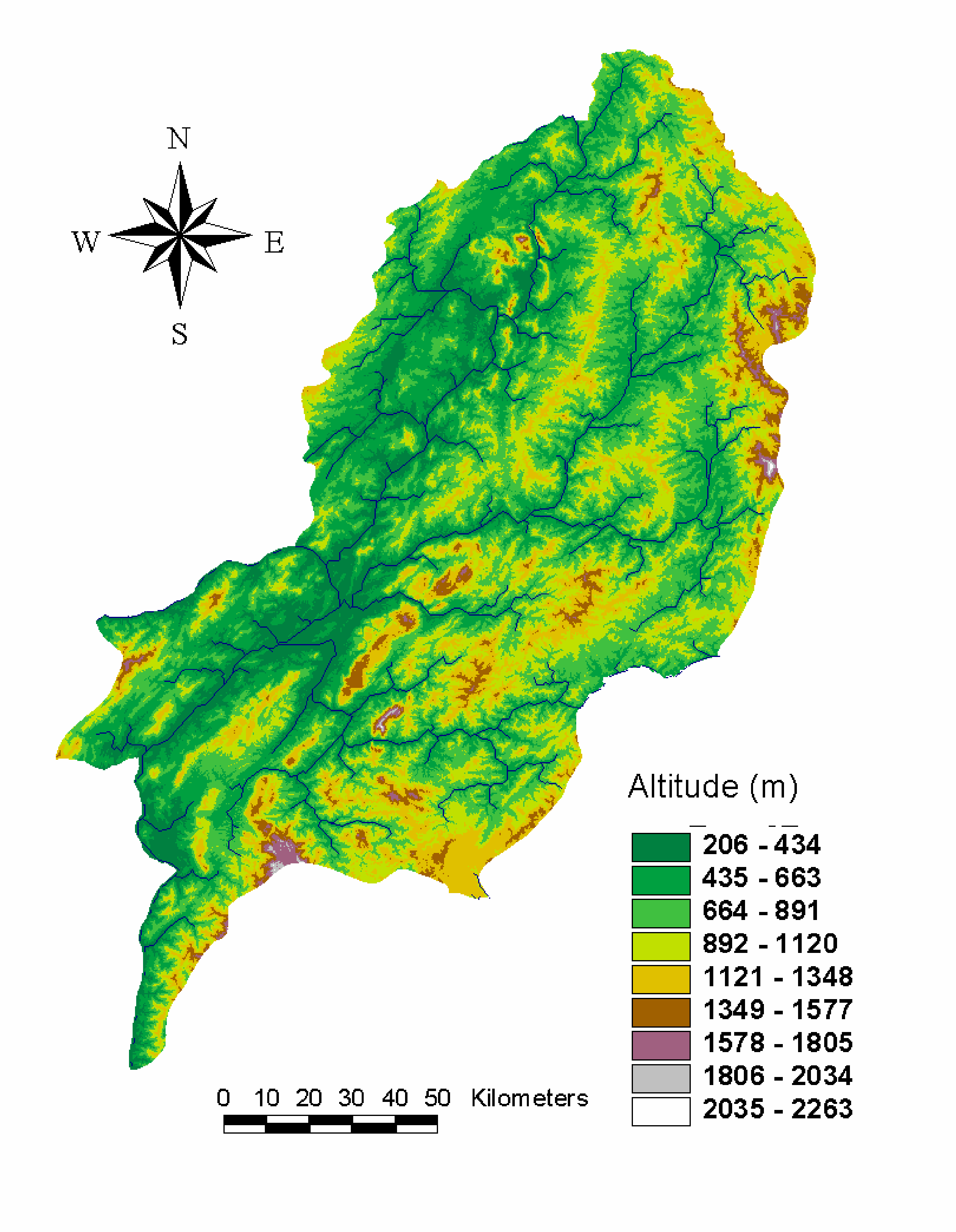

Figure 1. Digital elevation model of the Luang Prabang province with a 50-m mesh and location of streams.

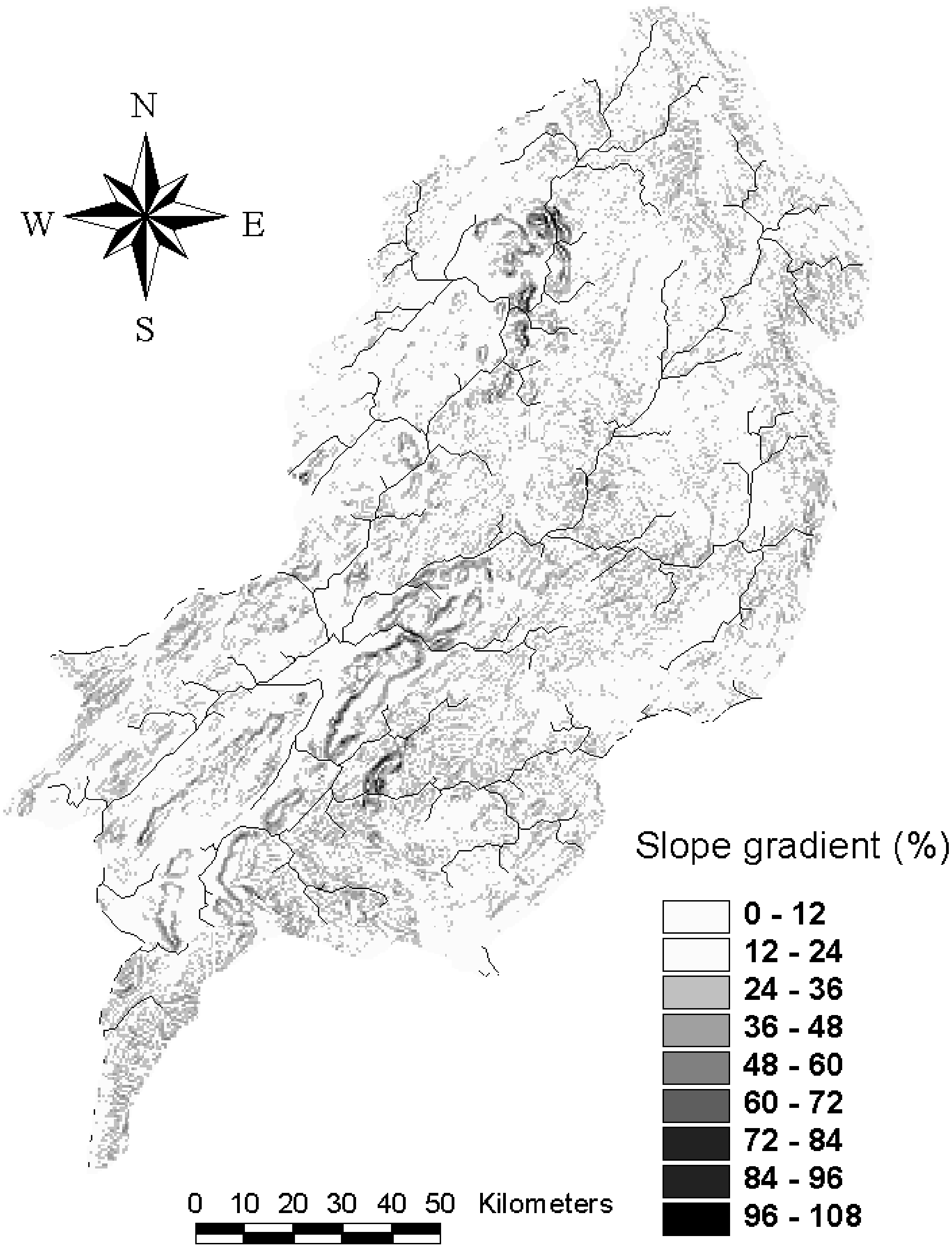

Figure 2. Mean slope map of the Luang Prabang province derived from a 500-m DEM. Location of streams.

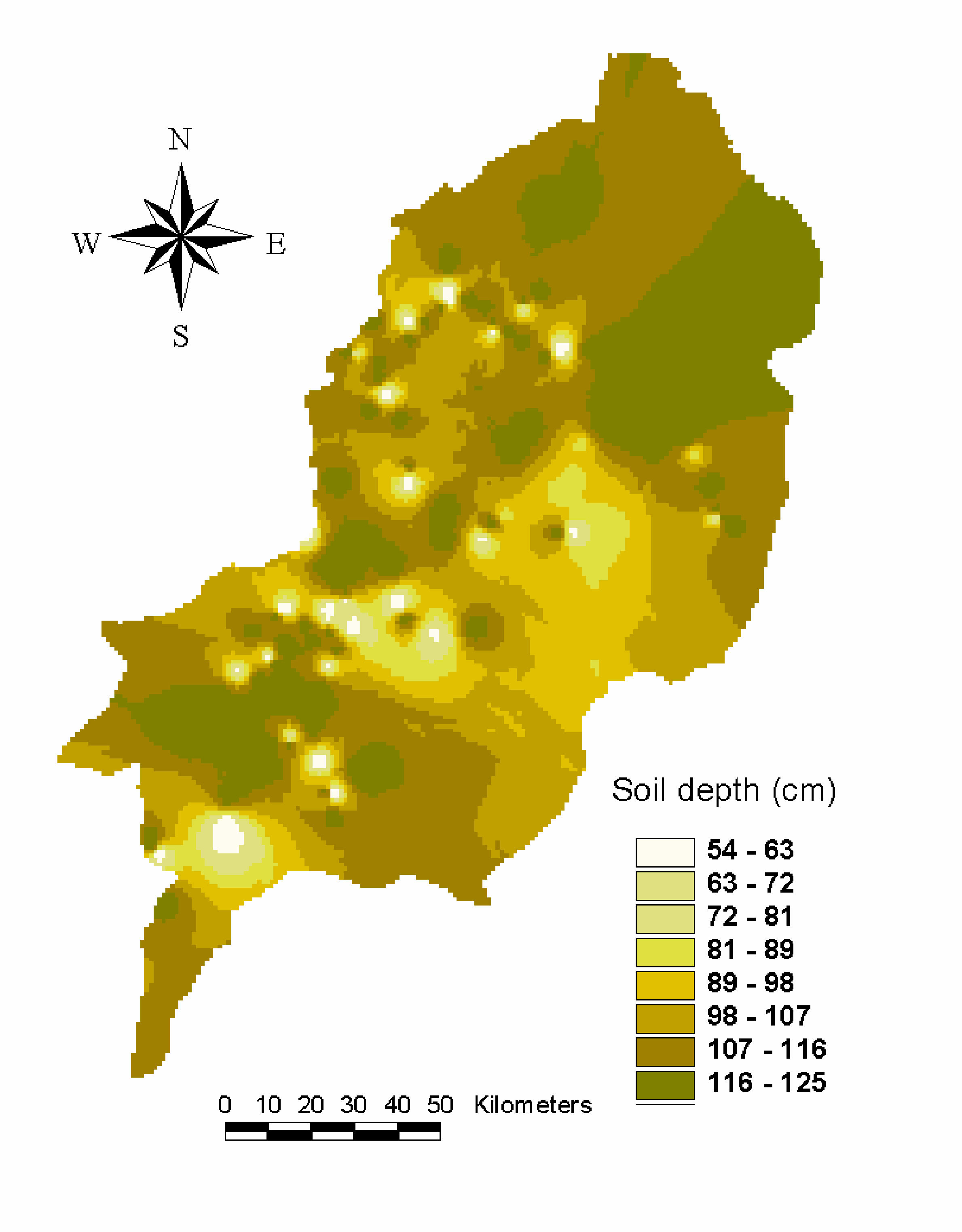

Figure 3. Soil depth map of the Luang Prabang province interpolated using 140 available data points by using inverse distance weighting (IDW) with 12 neighbors. Grid with a 1000-m mesh. Location of streams.

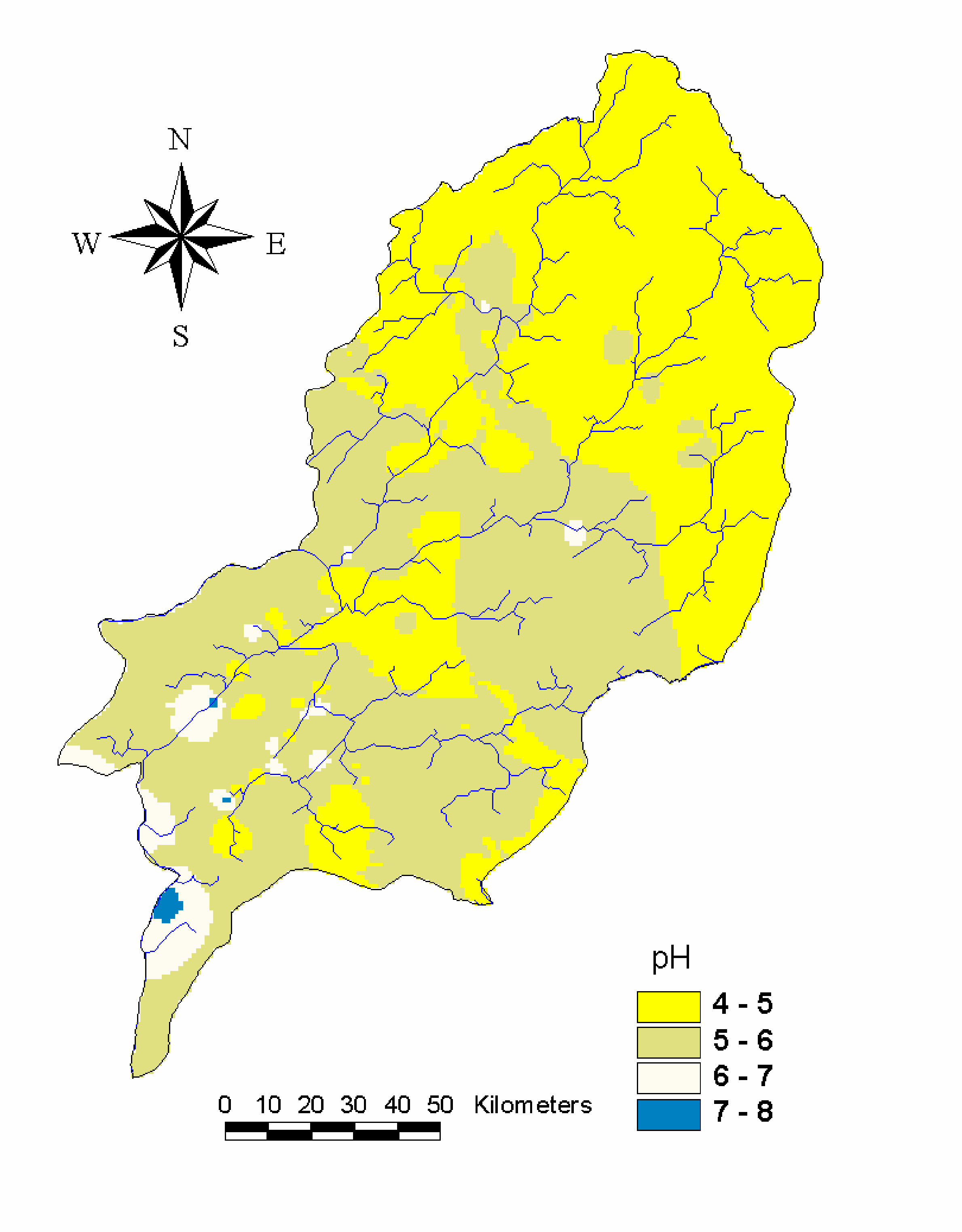

Figure 4. pH map of the Luang Prabang province interpolated using 140 available data points for surface A horizons by using inverse distance weighting (IDW) with 12 neighbors. Grid with a 1000-m mesh. Location of streams.

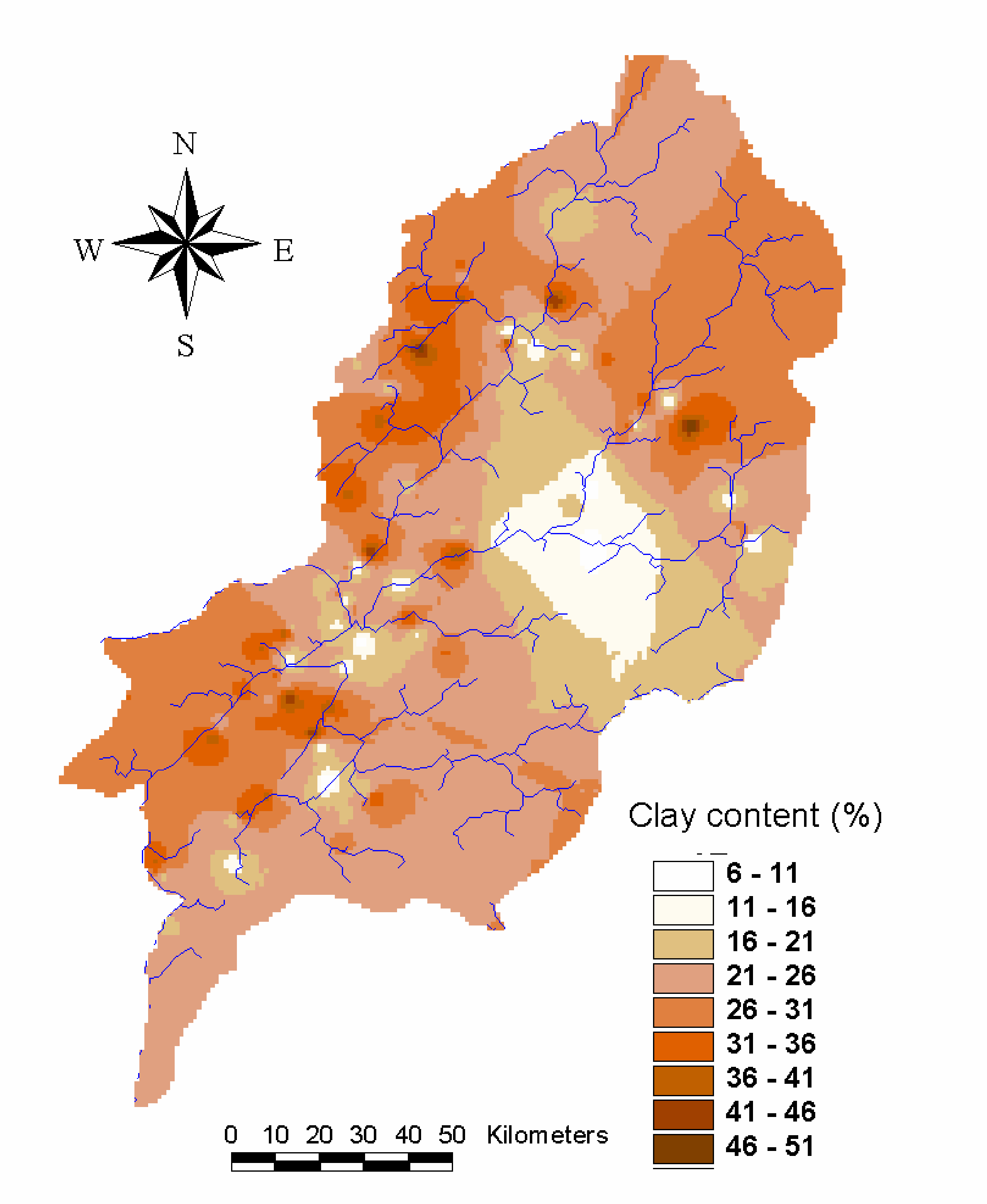

Figure 5. Clay content map of the Luang Prabang province interpolated using 140 available data points for surface A horizons by using inverse distance weighting (IDW) with 12 neighbors. Grid with a 1000-m mesh. Location of streams.

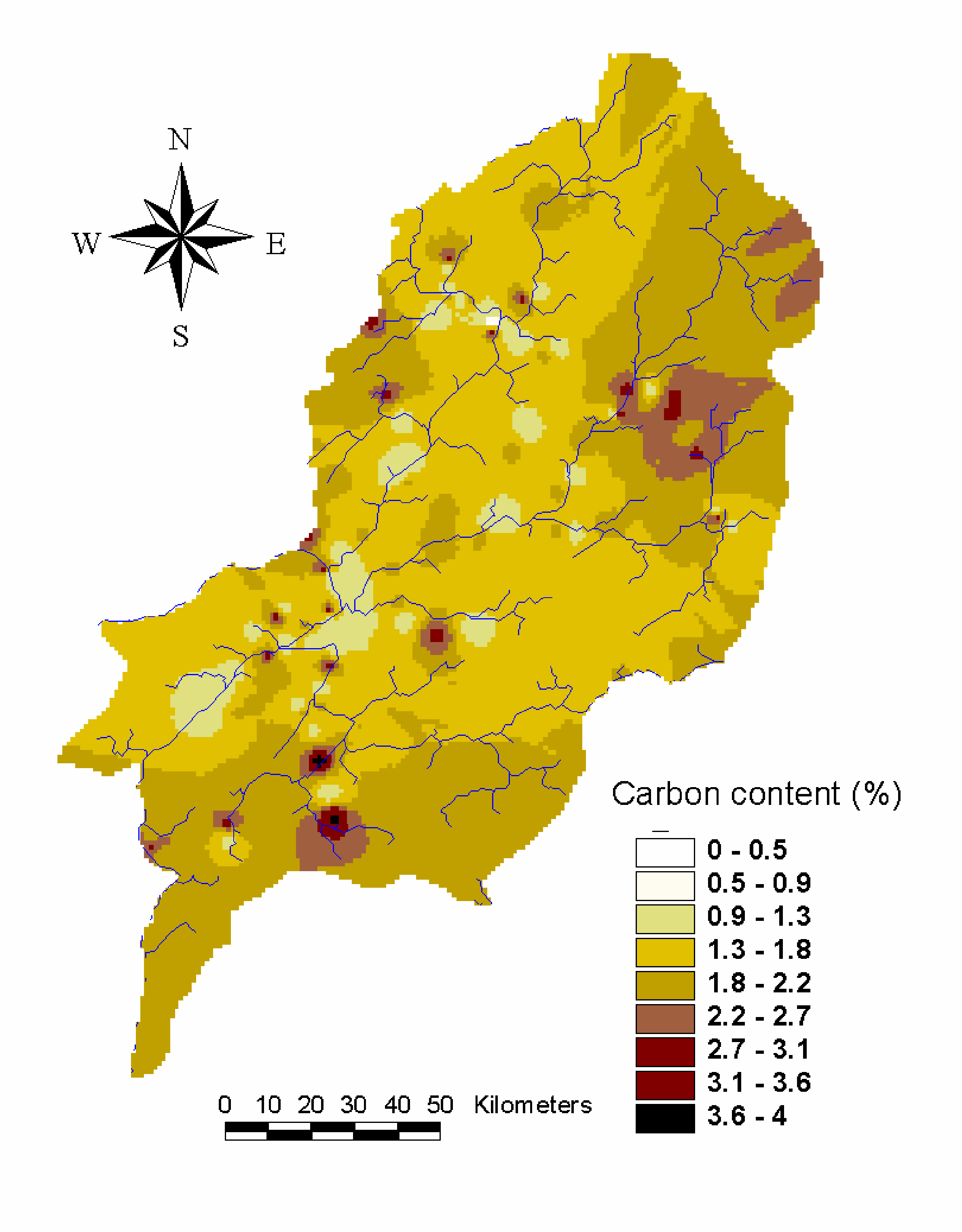

Figure 6. Carbon content map of the Luang Prabang province interpolated using 140 available data points for surface A horizons by using inverse distance weighting (IDW) with 12 neighbors. Grid with a 1000-m mesh. Location of streams.

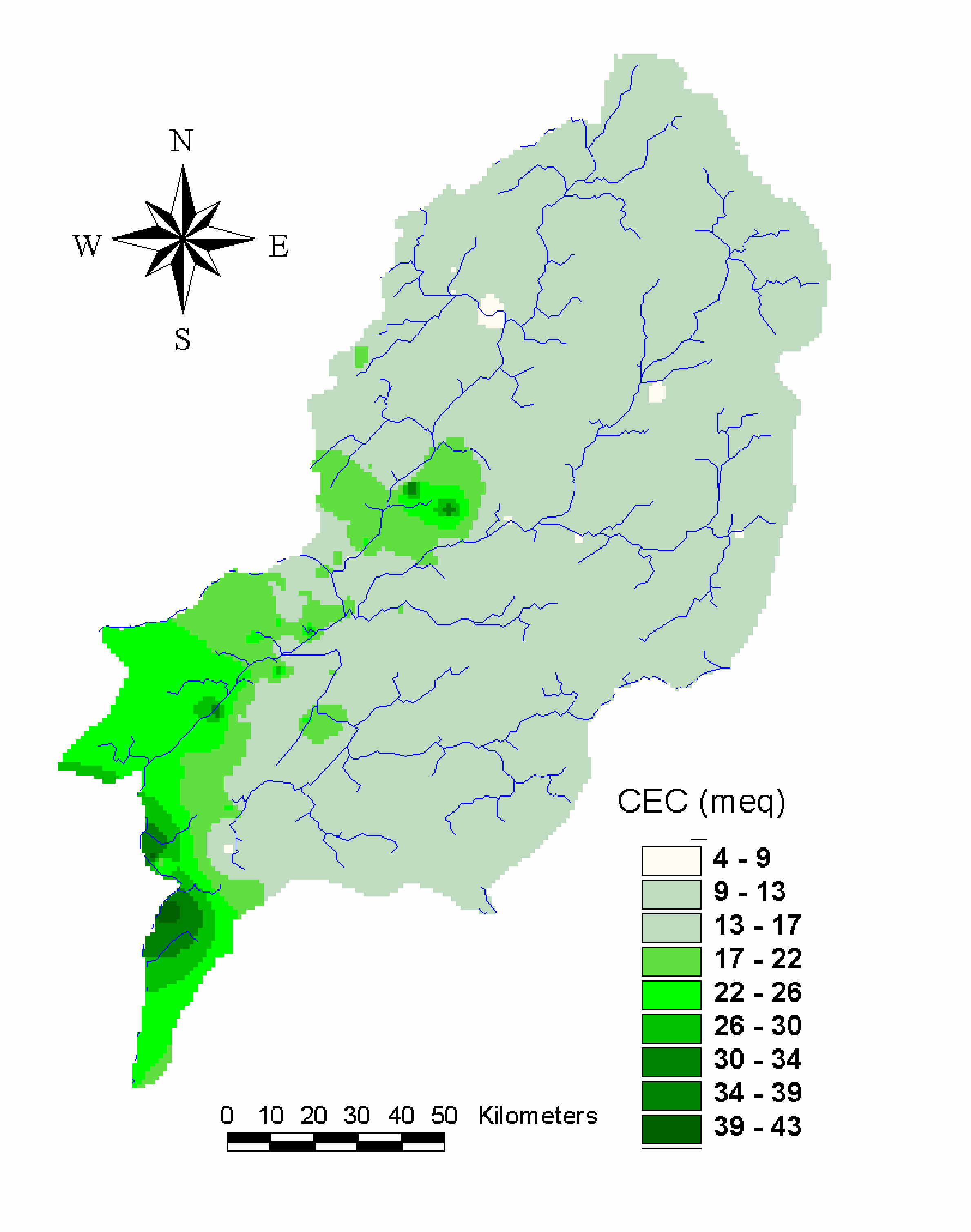

Figure 7. Total CEC map of the Luang Prabang province interpolated using 140 available data points for surface A horizons by using inverse distance weighting (IDW) with 3 neighbors. Grid with a 1000-m mesh. Location of streams.