Gridded multi-crop suitability mapping using public domain soil and related thematic data

Deependra Dhakal
Agriculture and Forestry University, Chitwan, Nepal
ddhakal@afu.edu.np

1 Introduction

Several research and planning processes in the past have acknowledged the logic of land evaluation (Rossiter 1996). In order to prevent exploitation of land or to use it to its potential, the rational basis of land-use planning is to determine what resources (or in agricultural terms intrinsic factors) the land can offer. FAO, 1993 details a comprehensive framework for suitability classification of land, based on several land "qualities". Broad qualities include availability of energy, availability of water, conditions for ripening, climatic hazards, sufficiency of oxygen in the root zone, sufficiency of nutrients, erosion hazard and toxicity (United Nations. Soil Resources, Service, and United Nations. Interdepartmental Working Group on Land Use Planning 1993). While the information on several qualities is justified for broad scale land-use planning, planning for a small unit (for example local administrative region), where various characteristics of climatic/erosion hazards and energy and water availability are similar should offer leniency as to their inclusion. However, depending upon the scale of interpretation, factors such as soil, topography, vegetation patterns, precipitation, temperature, soil parent material, land cover type, landform classes constitute important covariates for inferring about crop suitability to agricultural land.

Recent amendments of Nepal's Land Act, 1964 bases classification of land for scientific land-use "based on, inter alia, the nature and fertility of soil, geographical situation, environment and climate of the country." The wealth of geographical and remotely sensed data publicly available, nowadays, can serve as a starting point in making informed decisions about scientific land management (Huang et al. 2018). With the focus of all three tiers of government on scientific land mapping for various purposes, identification of crop pockets will assist in agriculture production planning and mechanization. Farmholds and local planners benefit by having a concrete visual map of what focus is to be laid where. Moreover, this data driven information support system can help identify most profitable ventures, while tapping geographical comparative advantage.

Land resource optimization and spatial allocation planning have been a subject of a number of studies in past. Much focus has been on either defining a, supposedly, advantage-index (Liu, Wang, and Chen 2019) and maximizing it with respect to variables such as land use, urbanization and industrialization, precipitation and current production levels, or using GIS based multi-criteria evaluation (Maleki et al. 2017). Current study is akin to the latter work, in that multiple-criteria are taken into account. However, as opposed to defining numeric weights of each criteria using hierarchical process to derive suitability values, rank based indicator is used to break ties when comparing any two crops.

2 Methodology

2.1 Data sources

Covariate layers were generated using satellite images (raster data). The soil profile data were collected from various government projects including National Land Use Project, Irrigation and Water Resources

Management Project, Central Agriculture Laboratory (previously Soil Management Directorate), and Nepal Agricultural Research Council (NARC). The maps were prepared using soil information from 23,273 soils samples, collected from 56 districts covering seven provinces. These soil properties are combined with a stack of 168 remote sensing-based soil covariates (SRTM DEM derivatives, climatic images, vegetation index etc.). Later the spatial predictions on 250x250m grids were generated using a machine learning method and the random forest (Hengl et al. 2017). These data are available as raster layers for use from NARC, Nepal[1]. Thematic soil data on **p**ercentage **N**itrogen **c**ontent (PNC) and absolute pH value (pH) and **el**evation (EL) data layers were acquired for processing. DSM uses advanced computational algorithms that use both soil sample data and environmental variables to generate maps. In addition to taking into account the spatial autocorrelation, the interpolated data also includes soil forming factors as covariates.

For purposes of overlaying, map drawing and masking, vector layer data (shapefiles) were obtained from Election Commission, Nepal ("Electoral and Geopolitical Map of Kailali" 2023) and the OpenStreetMap database (OpenStreetMap contributors 2023).

- **G**eopolitical **p**olygon **m**ap of **T**kapur municipality (CRS: Nepal_MUTM_Central_84_Everest_1830) (GPMT)
- **M**ulti**p**olygon spanning areas of **T**ikapur municipality having **c**ultivable **l**and (CRS: Nepal_MUTM_Central_84_Everest_1830) (MPTCL) constructed with querying of the original geo-polygon obtained from OSM database for various land forms and usage.
- **M**ulti**p**olygon **r**egular **t**iles (1645) forming a grid spanning **T**ikapur municipality (CRS: WGS 84, later transformed) (MPRTT)
- **L**ine and **m**ulti**l**ine features representing local, national and international **p**olitical **b**oundaries (CRS: Nepal_MUTM_Central_84_Everest_1830) (LMLPB)

Election Commission, Nepal maintains a repository of national, provincial and local geo-political maps. GPMT and LMLPB shapefiles were developed by cleaning and filtering required features from the electoral map. The shapefiles were processed using QGIS and open source packages available in R computing environment (R Core Team 2022; Dunnington 2022; Pebesma 2018; Bivand, Keitt, and Rowlingson 2022).

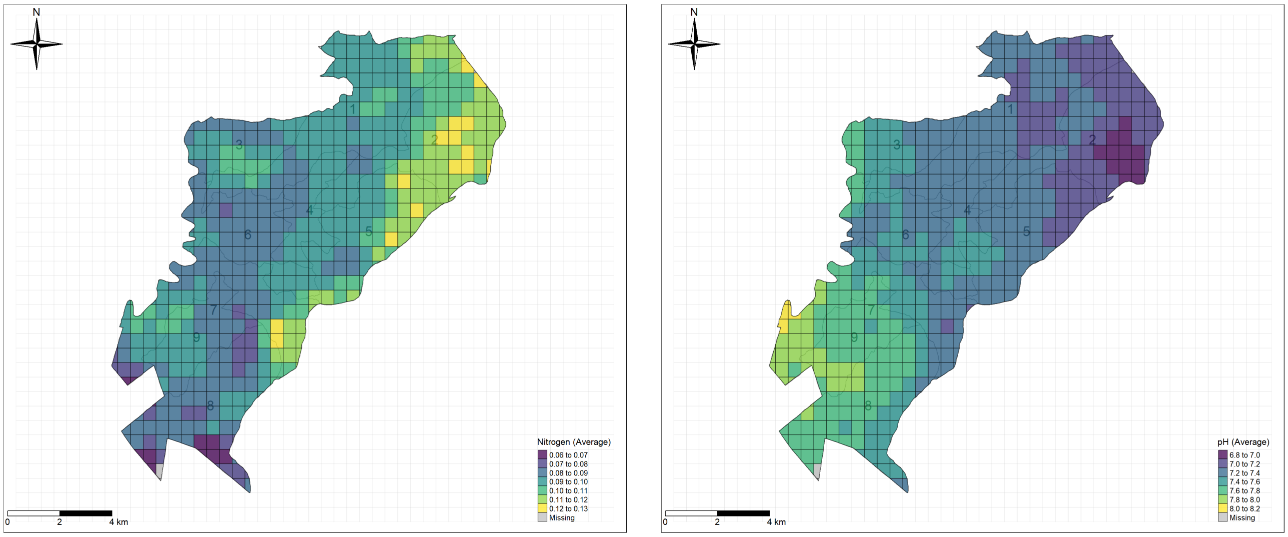
2.2 Spatial pre-processing

First step in processing of raster consisted of clipping raster to the geopolitical boundary of Tikapur municipality, excluding edge pixels. That followed obtaining zonal statistics for every raster layer in the input vector of a spatial square grid (MPTCL) of approximately 550m x 550m. The vector output from such processing containing zonal attribute with mean was used to classify the map with quantile based classifier (QGIS Development Team 2022) and display graduated information visually (Figure 1 (a) and (b), show the nitrogen and pH information for grids encompassing Tikapur).

2.3 Crop layers generation

Binary data, each layer corresponding to a unique crop and a thematic attribute combination, was constructed. Thus for 3 thematic attributes (PNC, PH and EL) and 19 crops (Banana, Blackgram, Broccoli, Cabbage, Cauliflower, Chickpea, Garlic, Groundnut, Lentil, Maize, Mustard, Onion, Pea, Pigeonpea, Potato, Rice, Sesame, Soybean and Wheat) a total of 57 2D binary layers were formed. Binary thresholding of each layer was accomplished by first defining the optimal ranges (defined as limiting values elsewhere in literatures) of thematic attribute as thresholding parameter. Any grid of a layer of particular crop, i.e. Cauliflower, for a theme, i.e. Nitrogen, would be labelled 1 (suitable) if the grid attribute (defined as zonal attribute, earlier) value more than the threshold value defined for the crop for the theme in any standard literature, and 0 otherwise. An extensive discussion of stucture of land suitability classification in general (development, land, and division 1976), and based on limitting factors (surface texture, pH, organic

[1] https://soil.narc.gov.np/data



(a) Nitrogen (b) pH

Figure 1. Gridded map of Tikapur at spatial resolution of 550m showing attribute information.

carbon and NPK regimen) as well as crop basis (Ramamurthy et al. 2020; Sys et al. 1993) has been documented. While the author acknowledges the vastness of resource available, for simiplification purpose, 'Krishi Diary', an annual publication of Government of Nepal was used for obtaining parameter range. Table 1, for example, is a tabulation of limiting values of pH for individual crops. All layers were overlayed with this binarization filter based on range thresholding and each grid was populated with binary values indicating suitability for the crop.

Table 1. Optimum pH ranges of crops as defined in Krishi Diary, 2079.

Crop	pH	Crop	pH
Rice	5-6.5	Onion	6-7.5
Sorghum	5.5-7	Lettuce	6-7
Maize	5.5-7	Cabbage	6-7
Wheat	5.5-7.5	Cauliflower	6-7
Oat	6-7	Guard	6-7
Barley	6.5-8	Beans	6-7.5
Groundnut	5-6.0	Lady's finger	6-7.5
Mustard	6-7	Cucumber	6-8
Rye	6-7	Spinach	6.5-7
Sunflower	6-7	Mint	6-8
Sugarcane	6-7	Ginger	6-7.5
Sugarbeet	6-7	Watermelon	5-6.0
Cotton	5-6.0	Pineapple	5.5-6.5
Tea	5-5.5	Mango	5.5-6.5
Coffee	5-6.0	Apple	5.5-6.5
Pepper	6-6.5	Pummelo	5.5-6.5
Lucerne	6-7	Guava	5.5-7
Clover	6-7	Lemon	6-7
Potato	5-5.5	Fig	6-7
Sweet potato	5-6.0	Banana	6-7
Carrot	5.5-6.5	Papaya	6-7.5
Bitter gourd	5.5-6.5	Grape fruit	6-7.5
Brinjal	5.5-6.5	Lime	6-7.5

Pea	5.5-6.5	Pomegranete	6-7.5
Radish	5.5-6.5	Grapes	6-8
Chillies	5.5-6.5	Pear	6-8
Turnip	5.5-6.5	Walnut	6-8
Tomato	5.5-6.0	Date palm	6-8.5
Cowpea	5.5-7	Bougainvillea	5.5-7
		Rose	6.6-7
		Chrysanthemum	6-7.5

Figure 2 is a single layer of information using one decision criterion – Nitrogen. Decision variables were given equal weights to form composite sum, which essentially are the suitability values for individual crops.

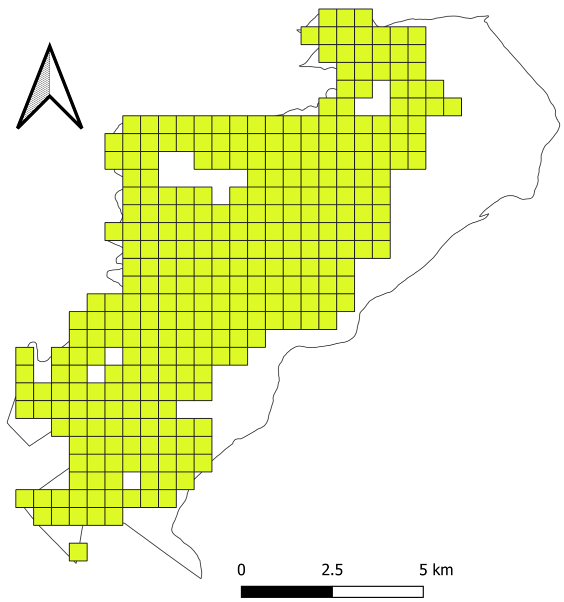


Figure 2. Binary map showing the result of thresholding on gridded raster representing nitrogen profile for suitability determination of Rapeseed/mustard. Grid polygons filled with color are the areas suitable for cultivation of the crop.

This high resolution map was subject to 3 x 3 rolling grid aggregation for obtaining frequency weights. Finally, a lower resolution gridded map of multi-crop suitability ranked by cumulative frequency was obtained. This information was plotted, using R, to display crop suitability information (Figure 3).

2.4 Grid averaging with non-overlapping windows and dimension reduction

The mean aggregation operation on a 2-dimensional grid (a matrix) using a fixed-size two-dimensional window can be described as computing the mean of all elements within each window as it moves across the matrix. Given a matrix A of size $m \times n$ and a window size $w \times h$, the grid averaging operation can be expressed as follows:

1. Let A be the matrix:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

2. Define the window size as $w \times h$.

3. Padding with zeros to ensure the matrix dimensions are multiples of the window dimensions.

- If m is not a multiple of W, pad A with $w - (m \mod w)$ rows of zeros.
- If n is not a multiple of h, pad A with $h - (n \mod h)$ columns of zeros.

Let the new dimensions of A after padding be $m' \times n'$, where:

$$m' = \begin{cases} m + (w - (m \mod w)), & \text{if } m \mod w \neq 0 \\ m, & \text{otherwise} \end{cases}$$

$$n' = \begin{cases} n + (h - (n \mod h)), & \text{if } n \mod h \neq 0 \\ n, \text{ otherwise} \end{cases}$$

4. Aggregation

Define the resulting matrix B of size $s \times t$ such that:

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,t} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ b_{s,1} & b_{s,2} & \cdots & b_{s,t} \end{bmatrix}$$

where $s = \frac{m'}{w}$ and $t = \frac{n'}{h}$.

- Each element $b_{i,j}$ is the mean of the elements within the $w \times h$ window starting at $(w(i-1)+1, h(j-1)+1)$ in the padded matrix A':

$$b_{i,j} = \frac{1}{wh} \sum_{p=0}^{w-1} \sum_{q=0}^{h-1} a_{w(i-1)+p+1, h(j-1)+q+1}$$

The process is repeated for all $s \times t$ elements in B.

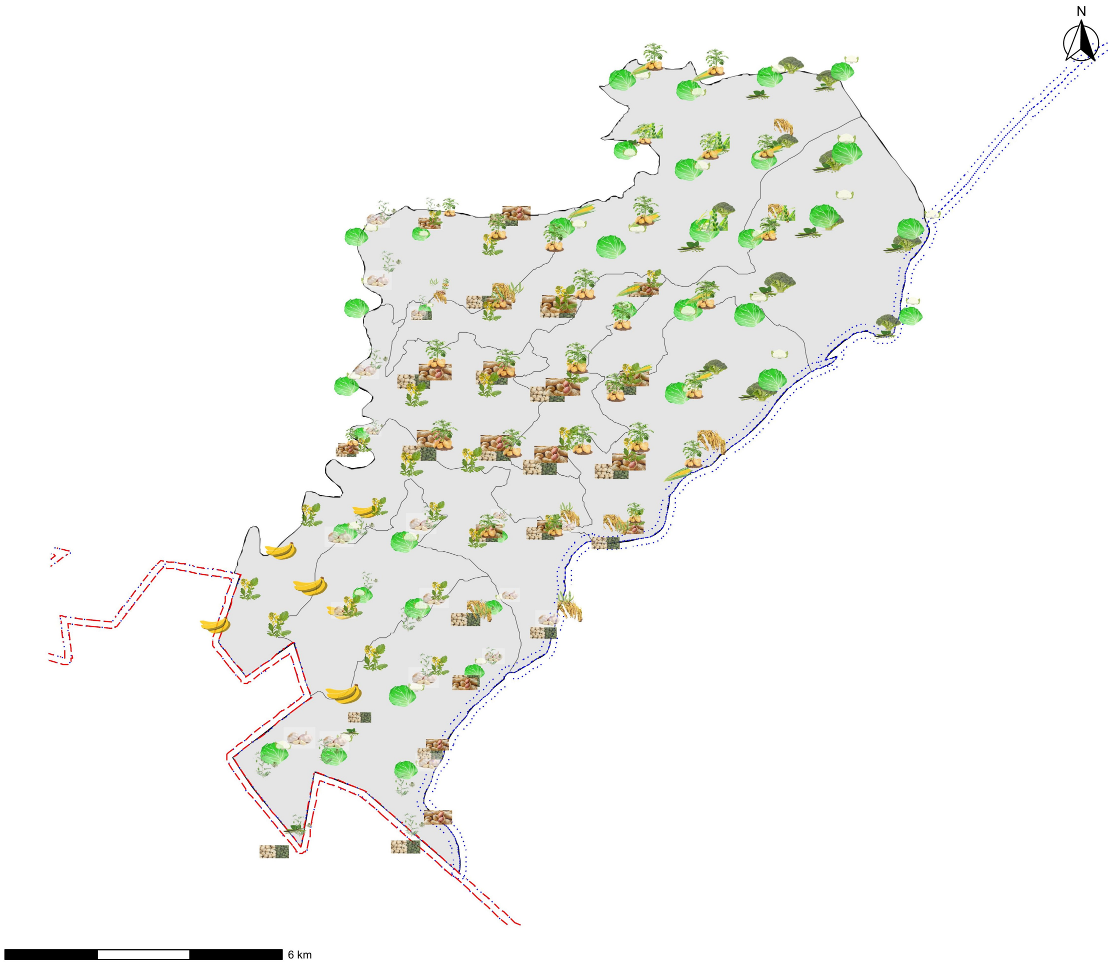


Figure 3. Crop suitability map showing visual depiction of crops with the image scaling proportionate to their suitability ranking.

3 Result and conclusion

The result is a map of local administrative unit, depicting grid based pockets of relative suitability of each crop that are generally suitable (based on limitting range) with various aesthetics (Figure 3, crop images are displayed with their relative suitability values mapped to size aesthetic and layer transparency). As for Tikapur, 58 such grid pockets have been mapped.

Current approach to is flexible with respect to the number of thematic variables that may be composited to obtain suitability ranking of individual crops, although its strictly required for them to be of same resolution; with respect to the number of crops for mapping as long as thresholding parameters are agreeably defined in literature. The method can also be scaled-up for any grid size for aggregation mapping as it relies on a simple rank based statistic. Information value of such maps can be extended by adding information on crop seasonality and agro-climatological variables.

Acknowledgement

The author is thankful to Rajesh Lamichhane, Riya Bhattarai, Janaki Pandey and Amrita Poudel, the third batch students of the agriculture faculty at College of Natural Resource Management, Tikapur, Kailali for

their assistance.

References

Bivand, Roger, Tim Keitt, and Barry Rowlingson. 2022. *Rgdal: Bindings for the 'Geospatial' Data Abstraction Library*. https://CRAN.R-project.org/package=rgdal.

development, Soil resources, conservation service land, and water development division. 1976. *A Framework for Land Evaluation. FAO Soils Bulletin*. Food; Agriculture Organization of the United Nations. https://www.fao.org/4/x5310e/x5310e00.htm.

Dunnington, Dewey. 2022. *Ggspatial: Spatial Data Framework for Ggplot2*. https://CRAN.R-project.org/package=ggspatial.

"Electoral and Geopolitical Map of Kailali." 2023. Civil Service. *Election Commission Nepal*. https://election.gov.np/np/.

Hengl, Tomislav, Jorge Mendes De Jesus, Gerard B. M. Heuvelink, Maria Ruiperez Gonzalez, Milan Kilibarda, Aleksandar Blagotić, Wei Shangguan, et al. 2017. "SoilGrids250m: Global Gridded Soil Information Based on Machine Learning." Edited by Ben Bond-Lamberty. *PLOS ONE* 12 (2): e0169748. https://doi.org/10.1371/journal.pone.0169748.

Huang, Yanbo, Zhong-xin Chen, YU Tao, Xiang-zhi Huang, and Xing-fa Gu. 2018. "Agricultural Remote Sensing Big Data: Management and Applications." *Journal of Integrative Agriculture* 17 (9): 1915–31.

Liu, Yating, Saige Wang, and Bin Chen. 2019. "Optimization of National Food Production Layout Based on Comparative Advantage Index." *Energy Procedia* 158: 3846–52.

Maleki, Fahimeh, Hossein Kazemi, Asieh Siahmarguee, and Behnam Kamkar. 2017. "Development of a Land Use Suitability Model for Saffron (Crocus Sativus l.) Cultivation by Multi-Criteria Evaluation and Spatial Analysis." *Ecological Engineering* 106: 140–53.

OpenStreetMap contributors. 2023. "Planet dump retrieved from https://planet.osm.org ." https://www.openstreetmap.org.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

QGIS Development Team. 2022. *QGIS Geographic Information System*. QGIS Association. https://www.qgis.org.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ramamurthy, V, O Challa, LGK Naidu, KS Anil Kumar, SK Singh, D Mamatha, K Ranjitha, and Bipin B Mishra. 2020. "Land Evaluation and Land Use Planning." *The Soils of India*. Springer.

Rossiter, David G. 1996. "A Theoretical Framework for Land Evaluation." *Geoderma* 72 (3-4): 165–90.

Sys, C. J., Eric Van Ranst, J. Debaveye, and F. Beernaert. 1993. "Crop Requirements." In *Land Evaluation - Part 3*. 7th Series. General Administration for Development Cooperation, Brussels, Belguim.

United Nations. Soil Resources, Agriculture Organization of the, Conservation Service, and Agriculture Organization of the United Nations. Interdepartmental Working Group on Land Use Planning. 1993. *Guidelines for Land-Use Planning*. Vol. 1. Food & Agriculture Org.