Big Data Analysis on Montane Geography and Urban Development in Nepal

Xinyi Li, MESc Candidate 2020
Advised by Prof. Karen Seto
12/15/2019

Abstract

The Hindu Kush Himalayan (HKH) region is home to about 210 million people and is undergoing rapid urban sprawl. This study analyzed the correlation between the urban settlement and montane geography with ArcGIS and regression models. This study discovered that 81% of urban settlements are on valley floors (gradient < 3°) than slopes. Slope gradient (p-value = 0.0036) and soil type are correlated with urban intensity, but slope aspect is not. In terms of soil type, Cambisols, especially Eutric and Ferralic Cambisols are more suitable for urban settlements than other soils. This quantitative study on the urbanization process in Nepal will help to inform area-specific land use and urban planning policies, assist hazard vulnerability assessment and management, and offer basic urbanization data necessary to other environmental studies.

1. Introduction

The United Nations projects that more than half of the world will live in urban areas by 2030[1]. To understand this ongoing trend of urbanization, researchers have raised land use and location theories. Most classic theories were based on an imaginary "featureless plain"[2] that without elevations and tested with coastal metropolises such as Tokyo[3] and New York[4], while fewer studies focus on urbanization in montane regions, which cover 26.5% of the world's land surface[5].

The Hindu Kush Himalayan (HKH) region is home to about 210 million people and is undergoing massive urbanization due to population growth, tourism, migration, transportation network construction, and urban-favored policies[6]. The development in the HKH countries is still in an early stage, with a high rate of urban expansion[7]. In Nepal alone, the urban population grew to 40.5% in 2015, which is more than 13 times than five decades ago[8]. The rapid and usually unplanned urbanization process in the HKH was accompanied with various form of urban settlements nationwide and significant land cover and land use change (LCLUC), which leads to further ecological and socioeconomic shifts. For instance, reduced vegetation[9] will affect the global carbon cycle and may destabilize local ecosystems[10], and increased agricultural land abandonment[11] may pose threats to food security[12] and reshape the local labor market[13]. Since the HKH has been recognized as a tectonically unstable[14], ecologically fragile[15], and economically underdeveloped[16, 17] montane system, many efforts have been devoted to examining the environmental effects of urbanization and how it is increasing the vulnerability of the HKH community[18, 19]. However, few publications acknowledged the diversity of urban land use[20] in the HKH, and few investigated the root-cause of these consequences: the urban development in montane regions.

This study aimed to understand the correlation between urban intensity and montane geography. This quantitative study on the urbanization process in the HKH will help to inform area-specific land use and urban planning policies, assist hazard vulnerability assessment and management, and offer basic urbanization data necessary to other environmental studies.

2. Objectives

a. Examine the spatial pattern of urban settlements in Nepal.
b. Test hypotheses about how montane geography correlates with the spatial pattern of urban settlement and its intensity in Nepal.

3. Research hypotheses

Key Hypothesis: Montane geography constrains the spatial pattern and the development of urban settlements.

H1: High mountains are the boundaries of the urban area.
Valley floors may be more suitable for urban sprawl than mountains due to their mild weather, fertile land, and proximity to rivers. This study anticipated more urban settlements on valley floors than slopes.

H2: Not all slopes are suitable for urban settlement.
The gradient (gentle or steep) and aspect (sunny or shady) of slopes can affect the land use decision of the montane community. This study anticipated a correlation between urban intensity and the slope gradient/aspect.

H3: Not all soil types are suitable for urban settlement.
Due to the difference in soil strength, moisture, or fertility, soil type can affect the land use decision of the montane community. This study anticipated a correlation between urban intensity and certain groups of soil.

4. Data Collection and Preparation

(1) Study Site

Nepal, whose capital is Kathmandu, covers a region of 147,181 km^2 and holds a population of 29.3 million. It lies between latitudes 26° and 31°N, and longitudes 80° and 89°E. Nepal contains three physiographic belts: Himal (mountain regions with snow), Pahad (mountain regions without snow), and Terai (lowland regions). Nepal ranges from 50 to 8000 meters in altitude and has five climate zones that correspond to the elevation: tropical, subtropical, temperate, subalpine, and alpine.

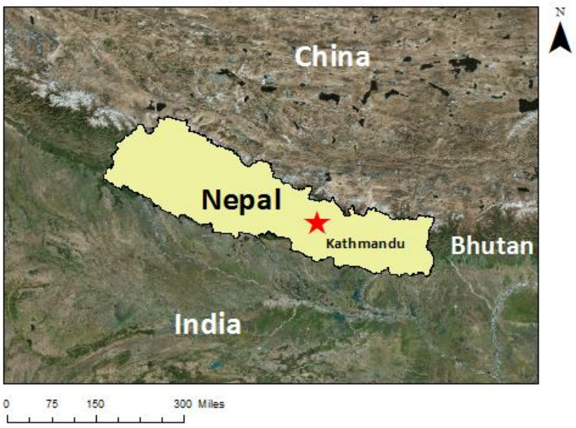

Fig 1. Study Site

(2) Dataset

a. GHSL[21]

The GHSL (Global Human Settlement Layer) dataset describes the presence of population and built-up infrastructure on the Earth's surface with data generated from continental satellite image archives, fine-scale satellite imagery, census data, and volunteered geographic information. This study used the GHSL Built-up product (250-meter resolution) as the dataset for urban intensity. The dataset, with a range of 0-100, reflects the diversity of urban intensity well. As shown in Fig 2, 0 (blue) means “no urban” and 100 (red) means “fully urban”.

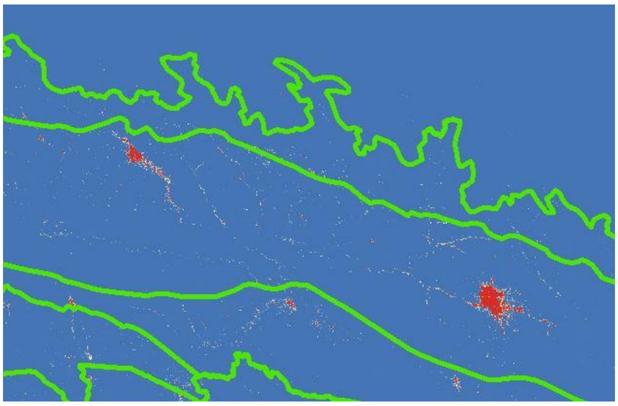

Fig 2. GHSL Dataset

b. DEM[22]

This study used a 30-arc-second (1-km) gridded, quality-controlled global Digital Elevation Model (DEM) data from the National Oceanic and Atmospheric Administration (NOAA) as the montane geography dataset. From this DEM dataset, this study further generated a slope gradient dataset (0-90°) and a slope aspect dataset (0-360°).

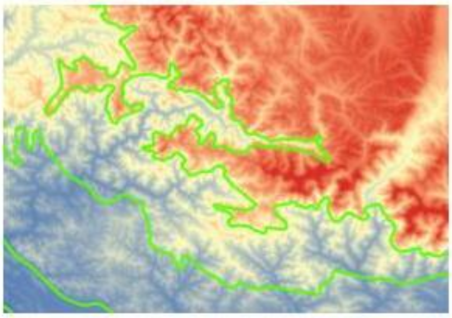

Fig 3. DEM Dataset

c. Soil type[23]

This study used the Soil and Terrain (SOTER) database as the Nepal soil type dataset. This database, with a scale of 1:1 million and 17 soil types (including glacier), provides generalized information on landform and soil properties. The SOTER program was initiated in 1986 by the Food and Agricultural Organization of the United Nations (FAO), the United Nations Environmental Program, and the International Soil Reference and Information Center (ISRIC).

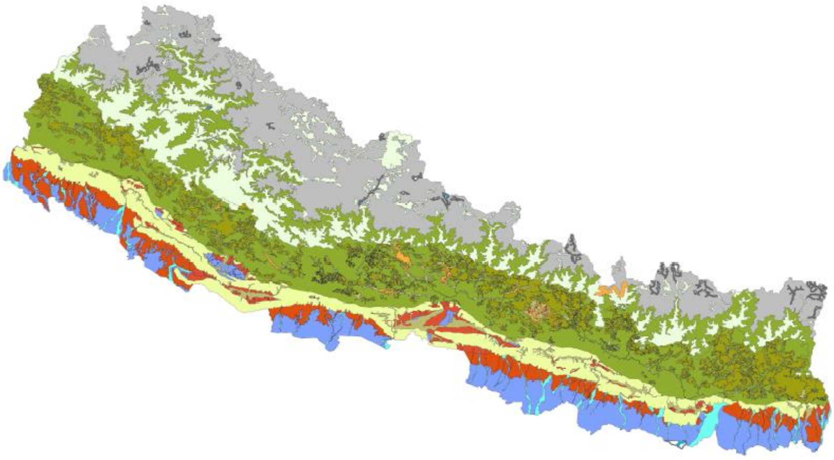

Fig 4. SOTER Dataset

General Type	Acronym	Full Name
Cambisols	CMg	Gleyic Cambisols
	CMo	Ferralic Cambisols
	CMx	Chromic Cambisols
	CMe	Eutric Cambisols
	CMu	Humic Cambisols
Fluvisols	FLc	Calcaric Fluvisols
Gleysols	GLe	Eutric Gleysols
Leptosols	LPi	Gelic Leptosols
Luvisols	LVx	Chromic Luvisols
Phaeozems	PHc	Calcaric Phaeozems
	PHh	Haplic Phaeozems
Regosols	RGd	Dystric Regosols
	RGe	Eutric Regosols

Table 1. Valid Soil Type List

(2) Data Preparation

a. DEM

- Slope Gradient

  This study used the “Slope” tool[24] in ArcGIS 10.6 to generate a slope dataset from DEM. For each cell, the “Slope” tool calculates the maximum rate of change in value from that cell to its neighbors using the formula below:

  $$slope\_degrees = \text{ATAN} \left(\sqrt{([dz/dx]^2 + [dz/dy]^2)} \right) * 180/pi$$

  where

  (dz/dx) is the horizontal directions from the center cell

  (dz/dy) is the vertical directions from the center cell

- Aspect

  The aspect dataset preparation required two steps. First, this study used the “Aspect” tool[25] in ArcGIS 10.6 to generate an original aspect dataset from DEM. For each cell, the tool identifies the downslope direction of the maximum rate of change in value from each cell to its neighbors. For example, to calculate the aspect of cell “e” in Fig 5, the formulas will be:

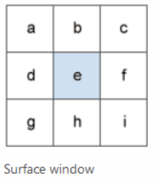

Fig 5. Aspect Cell Calculation

$aspect = 180/pi * atan2([dz/dy], -[dz/dx])$

where

rate of change in x direction: $[dz/dx] = ((c + 2f + i) - (a + 2d + g)) / 8$

rate of change in y direction: $[dz/dy] = ((g + 2h + i) - (a + 2b + c)) / 8$

The output of this first step would be a dataset ranged from 0° to 360°.
Next, this study rescaled the original dataset into a new one with a range of -1 to 1, where -1 means North/Shady (0° or 360°) and +1 means South/Sunny (180°). This conversion aimed to directly describe the extent of received sunshine for every cell, in order to test *Hypothesis 2*.

b. Generate Point Data
This study generated 4528 points with 6 values: longitude, latitude, urban intensity, slope, aspect, and soil. The process took 4 steps:

- Random Selection from GHSL
  With the “Create random points” tool in ArcGIS 10.6, this study randomly picked thousands of points in Nepal. As this study focused on the urban settlement, points with urban intensity values below 5 were omitted.
- Overlay with DEM and soil dataset
  With the “Sample” tool in ArcGIS 10.6, points from the last step were added with gradient, aspect, soil information based on their location.
- Data Cleaning
  This study omitted NA points due to their negative gradient or aspect value, or NA for soil type. After data cleaning, this study obtained 4534 points.
- Output
  A subset of the cleaned point dataset is shown below:

X	Y	Urban Intensity	Aspect	Slope	Soil Type
-176668	3209638	10.5488	38.44634	3.989838	CMe
-173744	3222048	5.3664	125.1651	1.970186	CMo
-173654	3222291	5.5792	44.21841	2.971969	CMe

Fig 6. Point Dataset

5. Methodology and Results

This study used Excel 2016 and R 3.2.5 for regression and boxplot analysis, and used Tableau for data visualization.

(1) *H1: High mountains are the boundaries of urban area*

To test the first hypothesis, this study analyzed the slope gradient of all selected urban settlements. As anticipated, more urban settlements are on valley floors than slopes. In fact, 80.9% of urban settlements are on slopes with gradients fewer than 3°, which are valley floors as other literature[26, 27] defined. In fact, only 1% of urban settlements are on slopes steeper than 15°. This proves that our hypothesis, that High mountains are the boundaries of the urban area, is correct.

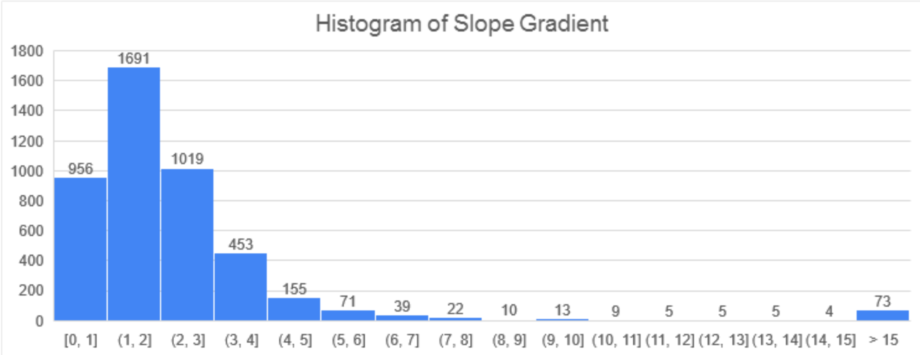

Fig 7. Slope Gradient Distribution Histogram

(2) *H2: Not all slopes are suitable for urban settlement*

a. Gradient

Since the point dataset is not evenly distributed in respect of gradients, this study divided slope into 3 bins: 0°-10°, 11°-30°, 31°-90°. To test this hypothesis, this study randomly chose 20 samples in each bin and ran a regression model between urban intensity and slope gradient with 60 points.

The result is shown below. As p-value < 0.05, there is evidence to believe that urban intensity and slope gradient are negatively correlated. The steeper the slope, the less unlikely for mature urban settlements to be built or developed. This is aligned with our hypothesis.

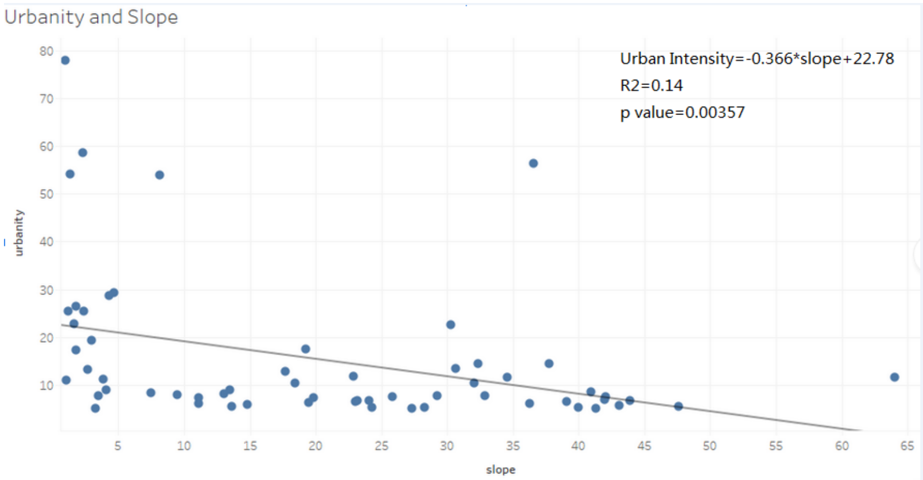

Fig 8. Regression between urban intensity and slope gradient

b. Aspect

To determine whether it is reasonable to do a similar "bin division" as in gradient analysis, this study started with a preliminary data visualization to see the distribution of aspect from thousands of point data.

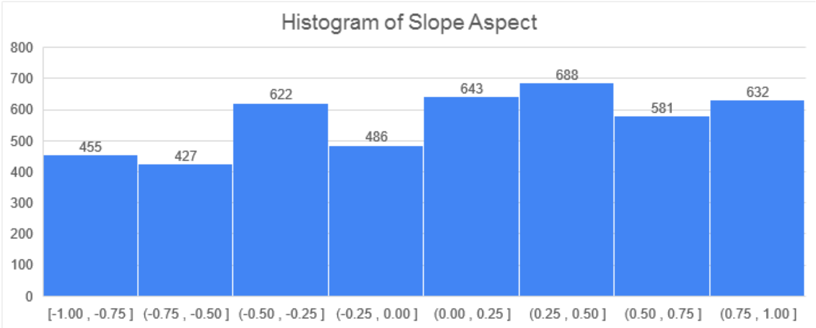

Fig 9. Aspect Distribution

As Fig 9 suggested, the aspect distribution seemed to be not unevenly distributed. Therefore, this study randomly picked 250 points to test the correlation between urban intensity and aspect.

The result is shown below. As p-value > 0.05, there is no evidence to believe that urban intensity and slope aspect are correlated. This is not aligned with the hypothesis.

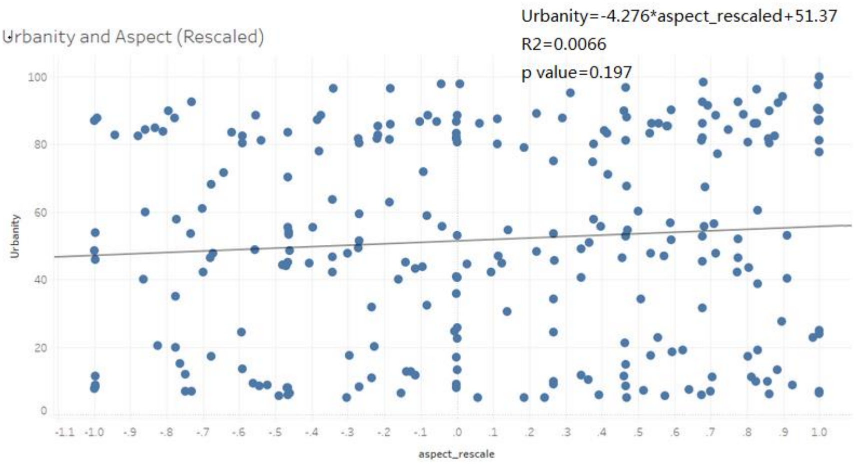

Fig 10. Regression between Urban Intensity and Aspect (rescaled)

In summary, this study proved that urban intensity and slope gradient are negatively correlated, and urban intensity and slope aspect are not correlated. Hence, this partially proves that our hypothesis, that not all slopes are suitable for the urban settlement, is correct.

(3) *H3: Not all soil types are suitable for urban settlement*

To test this hypothesis, this study started with the 7 general types of soil: Cambisols, Fluvisols, Gleysols, Leptosols, Luvisols, Phaeozem, and Regosols. The first analysis is interested to see what types of soil the urban settlements are on, regardless of the urban intensity. From 4500+ data points, the general soil type distribution is highly disproportionate, shown as follows:

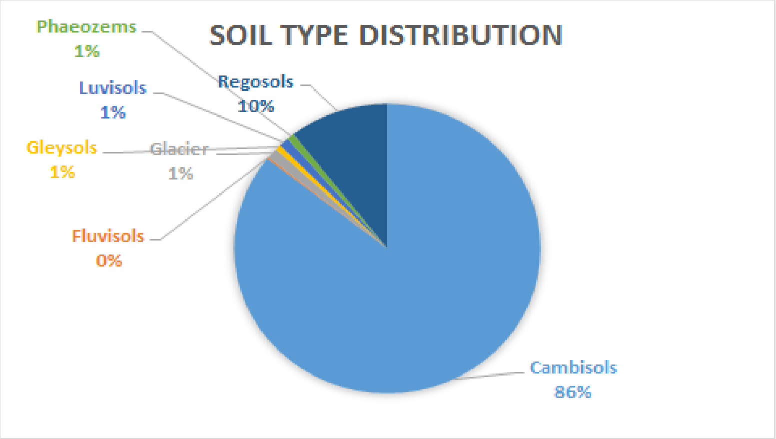

Fig 11. General Soil Type Distribution

From Fig 11, it is clear that Cambisols is the majority (86%) of urban settlement soil. Then, this study took a further analysis to breakdown urban settlements on Cambisols, as shown in Fig 12.

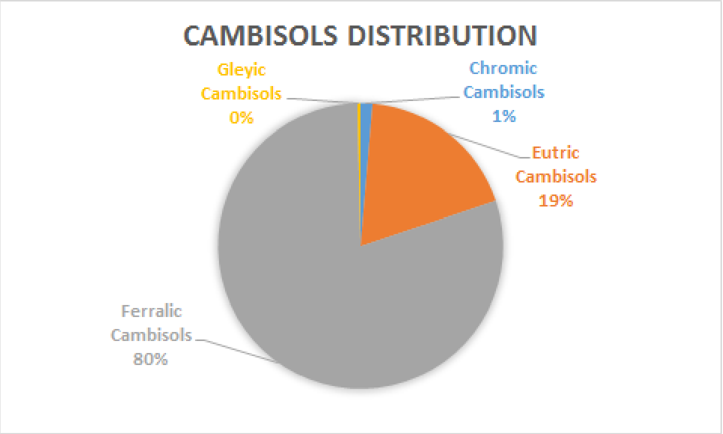

Fig 12. Cambisols Urban Soil Breakdown

As Fig 12 suggested, within urban Cambisols soil, the majority (80%) is Ferralic Cambisols, making 69% of nationwide urban soil.

The second analysis is interested to see the correlation between urban intensity and soil type. The analysis was visualized in boxplots. As Fig 13 suggested, Cambisols and Regosols tend to have more high-intensity urban settlements. Making one step further to breakdown Cambisols, as shown in Fig 14, Eutric and Ferralic Cambisols tend to have more high-intensity urban settlements than other Cambisols.

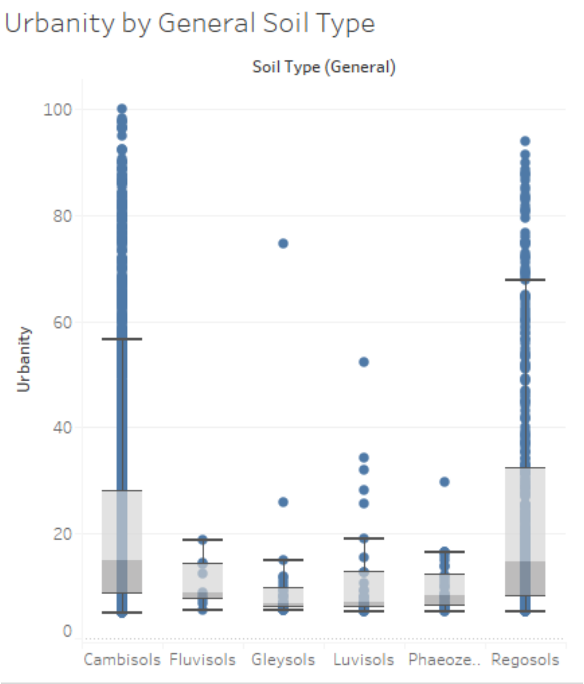

Fig 13. Urban Intensity by General Soil Type

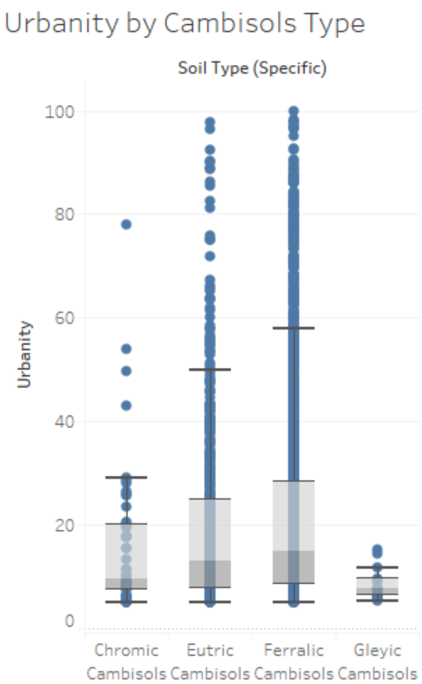

Fig 14. Urban Intensity by Cambisols Type

In summary, this study discovered that:

- The majority of urban soil is Cambisols (86%)
- Within Cambisols, the majority of urban soil is Ferralic Cambisols (80%)
- Cambisols and Regosols tend to have more high-intensity urban settlements than other general soil types
- Eutric and Ferralic Cambisols tend to have more high-intensity urban settlements than other Cambisols

6. Discussion

(1) Contributions

This study analyzed the correlation between the urban settlement and montane geography. The results showed that even in a montane country like Nepal, 90% of urban settlements are on valley floors (gradient < 3°) than slopes. Not all slopes are suitable for mature urban settlements with high intensity, as slope gradient and soil type clearly play a role. In terms of soil type, Cambisols, especially Eutric and Ferralic Cambisols are good matches for urban settlements. Surprisingly, slope aspect does not seem to correlate with the urban settlement. One possible explanation is that, although Nepal is an agriculture-dominated country, the majority of the farmlands are on the southern plain Terai, rather than in valleys[28]. Therefore, the extent of sunshine and rainfall in valleys, which should be affected by aspects, are not the priorities for the majority Nepal farmers who live on the plain.

This study provided further implications for urbanization studies. First, this study acknowledged the diversity of urban intensity instead of arbitrarily treating all urban settlements as a single land use type with a same degree of maturity. This recognizes the inconsistency and the dynamics of the urban development in Nepal and provides directions for further quantitative urban studies beyond. Second, this study took a specific focus on a rapidly developing montane region, which is not a popular yet meaningful study site. The study on the HKH will shed light on other montane region studies such as the Appalachians, the Alps, and the Andes regarding their land use pattern and urban planning process. Lastly, this study proved several empirical land use preferences for urban development, which will further assist hazard management and lead to a less vulnerable future for the montane community.

(2) Limitations

This study had two main limitations. First, due to the availability of data, the resolution could be coarse for an in-depth nation-level project. For example, since the DEM data is generated in a 1-km grid, the calculations of aspect and slope are also based on a kilometer level, which is relatively larger than a single Nepalese house.

Second, due to the limited scope of hypotheses, this study did not conduct a temporal pattern analysis that dived deep into the land use changes from satellite images. Besides knowing "where" the changes happened, getting information on "when" and "what kinds of changes" could provide extra insights for the urban planning and hazard management authorities.

(3) Next steps

By the end of April 2020, this study will accomplish next steps on: further analysis of soil type (e.g., moisture, fertility, strength), a discussion on the Nepalese urban development history based on literature review, publishing geospatial data and related codes, and thesis writing.

Acknowledgements

This study is funded by the Yale Hixon Center for Urban Studies and the Yale Institute for Biospheric Studies (YIBS).

Many thanks to my thesis committee members, Prof. Karen Seto and Prof. Timothy Gregoire, for their support and feedback. Special thanks to the members of the Seto Lab and the team of Urban Himalaya.

Signature:

Xinyi Li 数字签名者：Xinyi Li DN：cn=Xinyi Li, o, ou, email=lixinyi328@hotmail.com, c=CN 日期：2019.12.06 11:45:25 -05'00'

Hixon Student Research Fellow

Karen C. Seto Digitally signed by Karen C. Seto DN: cn=Karen C. Seto, o=Yale University, ou=School of Forestry and Environmental Studies, email=karen.seto@yale.edu, c=US Date: 2019.12.09 23:27:27 -05'00'

Faculty Advisor

Bibliography

[1] J. Bongaarts, "United nations department of economic and social affairs, population division world mortality report 2005," *Population and Development Review,* vol. 32, no. 3, pp. 594-596, 2006.
[2] R. Sinclair, "Von Thünen and urban sprawl," *Annals of the Association of American Geographers,* vol. 57, no. 1, pp. 72-87, 1967.
[3] P. Marcuse, and R. Van Kempen, *Globalizing cities: a new spatial order?*: John Wiley & Sons, 2011.
[4] R. P. Swierenga, "The settlement of the old northwest: ethnic pluralism in a featureless plain," *Journal of the Early Republic,* vol. 9, no. 1, pp. 73-105, 1989.
[5] D. Rodríguez-Rodríguez, B. Bomhard, S. H. Butchart, and M. N. Foster, "Progress towards international targets for protected area coverage in mountains: a multi-scale assessment," *Biological Conservation,* vol. 144, no. 12, pp. 2978-2983, 2011.
[6] B. Rimal, L. Zhang, N. Stork, S. Sloan, and S. Rijal, "Urban Expansion Occurred at the Expense of Agricultural Lands in the Tarai Region of Nepal from 1989 to 2016," *Sustainability (2071-1050),* vol. 10, no. 5, 2018.
[7] R. B. Thapa, and Y. Murayama, "Drivers of urban growth in the Kathmandu valley, Nepal: Examining the efficacy of the analytic hierarchy process," *Applied Geography,* vol. 30, no. 1, pp. 70-83, 2010.
[8] B. Rimal, L. Zhang, H. Keshtkar, N. Wang, and Y. Lin, "Monitoring and modeling of spatiotemporal urban expansion and land-use/land-cover change using integrated markov chain cellular automata model," *ISPRS International Journal of Geo-Information,* vol. 6, no. 9, pp. 288, 2017.
[9] N. B. Mishra, and G. Chaudhuri, "Spatio-temporal analysis of trends in seasonal vegetation productivity across Uttarakhand, Indian Himalayas, 2000–2014," *Applied Geography,* vol. 56, pp. 29-41, 2015.
[10] A. Schneider, C. Mertes, A. Tatem, B. Tan, D. Sulla-Menashe, S. Graves, N. Patel, J. Horton, A. Gaughan, and J. Rollo, "A new urban landscape in East–Southeast Asia, 2000–2010," *Environmental Research Letters,* vol. 10, no. 3, pp. 034002, 2015.
[11] A. Hussain, G. Rasul, B. Mahapatra, S. Wahid, and S. Tuladhar, "Climate change-induced hazards and local adaptations in agriculture: a study from Koshi River Basin, Nepal," *Natural Hazards,* vol. 91, no. 3, pp. 1365-1383, 2018.
[12] C. Munteanu, T. Kuemmerle, M. Boltiziar, V. Butsic, U. Gimmi, L. Halada, D. Kaim, G. Király, É. Konkoly-Gyuró, and J. Kozak, "Forest and agricultural land change in the Carpathian region—a meta-analysis of long-term patterns and drivers of change," *Land use policy,* vol. 38, pp. 685-697, 2014.
[13] T. Shi, X. Li, L. Xin, and X. Xu, "Analysis of farmland abandonment at parcel level: a case study in the mountainous area of China," *Sustainability,* vol. 8, no. 10, pp. 988, 2016.
[14] B. Rimal, H. Baral, N. E. Stork, K. Paudyal, and S. Rijal, "Growing city and rapid land use transition: assessing multiple hazards and risks in the Pokhara Valley, Nepal," *Land,* vol. 4, no. 4, pp. 957-978, 2015.
[15] M. K. Pandit, K. Manish, and L. P. Koh, "Dancing on the roof of the world: ecological transformation of the Himalayan landscape," *BioScience,* vol. 64, no. 11, pp. 980-992, 2014.
[16] K. C. Seto, B. Güneralp, and L. R. Hutyra, "Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools," *Proceedings of the National Academy of Sciences,* vol. 109, no. 40, pp. 16083-16088, 2012.
[17] P. Tiwari, "Land use changes in Himalaya and their impacts on environment, society and economy: A study of the Lake Region in Kumaon Himalaya, India," *Advances in Atmospheric Sciences,* vol. 25, no. 6, pp. 1029, 2008.
[18] N. Thakur, M. Rishi, D. A. Sharma, and T. Keesari, "Quality of water resources in Kullu Valley in Himachal Himalayas, India: perspective and prognosis," *Applied Water Science,* vol. 8, no. 1, pp. 20, 2018.
[19] E. Carabine, and A. Lemma, *The IPCC's Fifth Assessment Report: What's in it for South Asia*, IPCC, Budapest, Hungary, 2014.
[20] M. Reba, and K. Seto, "A systematic review and assessment of algorithms to detect, characterize, and monitor urban land use change," *Remote sensing of environment*, 2019. (in review)
[21] M. Pesaresi, D. Ehrlich, S. Ferri, A. Florczyk, S. Freire, M. Halkia, A. Julea, T. Kemper, P. Soille, and V. Syrris, "Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014," *Publications Office of the European Union*, pp. 1-62, 2016.
[22] T. GLOBE, "Team and others (Hastings, David A., Paula K. Dunbar, Gerald M. Elphingstone, Mark Bootz, Hiroshi Murakami, Hiroshi Maruyama, Hiroshi Masaharu, Peter Holland, John Payne, Nevin A. Bryant, Thomas L. Logan, J.-P. Muller, Gunter Schreier, and John S. MacDonald), eds., 1999," *The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version,* vol. 1, pp. 80305-3328.
[23] K. Dijkshoorn, and J. Huting, "Soil and terrain database for Nepal," *ISRIC–World Soil Information, Wageningen,* vol. 21, 2009.
[24] ESRI. "How Slope Works," http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-slope-works.htm.
[25] ESRI. "How Aspect Works," http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-aspect-works.htm.
[26] S. V. Gregory, G. A. Lamberti, and K. M. Moore, "Influence of valley floor landforms on stream ecosystems."
[27] W. Froehlich, E. Gil, I. Kasza, and L. Starkel, "Thresholds in the transformation of slopes and river channels in the Darjeeling Himalaya, India," *Mountain Research and Development*, pp. 301-312, 1990.
[28] D. Chaudhary, "A Critical Overview on Development-Welfare, Power or Domination," *Research Nepal Journal of Development Studies,* vol. 2, no. 1, pp. 43-55, 2019.