Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Analyzing Temporal Changes in an Urbanized Area Using Densely Staked Image Classification and Multinomial Logistic Regression (MLR) Technique

Abhinav Wadhwa
Vellore Institute of Technology
Pavan Kumar Kummamuru (pavankumar1978@gmail.com)
Vellore Institute of Technology https://orcid.org/0000-0002-1839-5819

Research

Keywords: Supervised and Unsupervised learning, Multinomial regression, Euclidean distance, Transition factors

Posted Date: June 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-32434/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Monitoring transformation of non-built-up area to urban spread via densely-stacked Land-Use-Land-Cover (LULC) classification offers a catalogue of spatio-temporal statistics to evaluate discrepancies instigated by transition factors. Impacts of major transition apparatuses in an area persuading the haphazard urbanization pattern are evaluated for Vellore acts a major contribution to Smart city project. Implications of causative factors: i) Population density; ii) proximity from rail-road-network; and iii) commercial areas are scrutinized with respect to urbanization upsurge. Multi-variate correlation is established using trend analysis and Multinomial Regression (MLR) technique for individual and homogeneous amalgamation of the aforementioned factors. Resulting equations obtained is formally used to detect closeness of urban extent from several landscapes. Research outcomes exhibited that the built-up straggling occurs from 30 to 232 m along the landscapes with a maximum of 336 m. Illustration of this study can also be assessed for various social and economic causative factors against urbanization for other smart cities.

Introduction

In India, for last two decades, increasing pace of urban trends has made it necessary to identify the causative factors leading to drastic changes in natural landscapes. Thus, in order to find an appropriate solution to minimize the damages to flora and fauna inhabiting the affected natural landscapes, trend analysis on the transition factors need to be evaluated. The swelling population density and outspreading urban sprawls, majorly in metropolitan municipalities, leads to intensifying demand of natural resources like water, energy, land surface thereby further results to deforestation, desertification and rigorous loss in agricultural lands. These land parcel variations mutually contribute in changing global environment and near-surface temperatures (Osgouei et al., 2019). Urbanization and changing life patterns during last three decades necessitated urban planners to refurbish an effective methodology and estimate the spatial extent of urbanization. Irregular shapes/sizes of urban features triggers hinderance to evaluate precisely the urban extent and its causative factors. In India, the evolving urbanized area resemble the fallow farmland because of their equivalent reflectance values (Long et al., 2009; Webster, 2001). Consequently, densely time stacked image analysis solves the aforementioned delinquent by categorizing urban and fallow farmland features appropriately. With saturated development in urban core, pre-urbanized segments located far away (ranges can be estimated using threshold proximity analysis) from city premises experiences the urban sprawl. VHR (1-4m) imagery is used to keenly introspect the urban changes occurring far away from peripheries of the densely populated areas (Ban et al., 2010; Del Frate et al., 2007).

Accuracy in urban area identification and modeling is of substantial interest to the municipal authorities for applications on urban planning such as resource allocation, management and distribution, facility provision and promotional policies (Jat et al., 2008). Non-parametric techniques like machine learning classification, decision tree algorithms and knowledge-based classifiers are used extensively to classify Landsat Imagery (Osgouei et al., 2019). Analysis and prediction modelling of impervious area using classification techniques consume high computational power and time. An alternative technique to demarcate the urban areas is by point sampling in addition to a supervised, unsupervised or knowledge based systematic learning technique (Bradley, 1997; Lu and Weng, 2005; Mundia and Aniya, 2005; Reddy et al., 2019; Stuckens et al., 2000; Vogelmann et al., 1998). The basic radiometric, geometric and gap filling (particularly for Landsat 7 images) corrections are not required for densely stacked images. The major factors influencing image classification techniques are: algorithms to be used to classify images, and what dataset is used (multi-spectral or multi-temporal or multi-fusor). Typically, the accuracy to classify satellite images for 3 major classes (vegetation, water and urban) is higher (over 85%) but while identifying more number of features, it becomes difficult and time consuming (Herold et al., 2003; Jia et al., 2014; Li et al., 2014; Lu et al., 2005). Out of various machine learning algorithms, support vector machines (SVM) and maximum likelihood classification (MLC) have been found as the most effective techniques (Osgouei et al., 2019). Various image enhancement procedures like principal component analysis (PCA), independent component analysis (ICA) or multiple marginal fisher analysis are used to eradicate the dependency of bands and correlation between them. The enhanced image can further be used for cataloguing in order to simplify and improve the classification processes for multi-spectral images. Landsat archive and open source datasets streamlines the evaluation of urban areas through remote sensing environment. With seasonal variations, the fallow farmland changes from barren to shallow vegetation and vice-versa whereas the built-up area remains the identical for all the seasons (Schneider, 2012). The availability of dense time stacked images in this archive facilitates the periodic (seasonal or yearly) analysis of images to clearly differentiate variations in minute features (similar features in case of fallow farmland and new built-up area).

Even though the urbanized areas are characterized, the question to isolate the reasons for increasing urban sprawl still remains unanswered. The evaluation of numerous factors contributing to increase in urban sprawl plays a vital role for discrete planning authorities towards computing development strategies intended for laid-back resource allocation and guaranteed forthcoming supplies of natural resources (Jat et al., 2008). Cihlar 2000; Weng 2001; Wang et al. 2003; Sudhira et al. 2004; Alsharif & Pradhan 2014; Abdullahi et al. 2015; Alsharif et al. 2015; Al-sharif & Pradhan 2016; Amini Parsa et al. 2016, addressed the primary causative factors responsible for increasing spatial extent of urbanization. The factors are broadly categorized as: geomorphological (elevation, slope), demographical (zoning status, Euclidian distance to national, state and local highways, railway stations, hefty communities, industrial, commercial or residential complexes), economic (employment rate, richness index), social and cultural (spiritual, tourism, large gathering areas, historic structures) (Mustafa et al., 2018). In India, marketing strategies and religious parameters show complex association with built-up expansion by reason of asymmetrical commercialization and conviction. Traditional methods to evaluate these factors includes manual mapping that necessitates employment, time and huge investment.

Remote sensing environment along with regression operation provides competent practice which is not time consuming and offers more accuracy for long-term outcomes (Haack and Rafter, 2006; Sudhira et al., 2004; Yang and Liu, 2005). Regression analysis of the causative factors and increasing urban spread using remote sensing is looked-for dynamically in this study. In this study, an effort is undertaken to inspect spatial and temporal applications of GIS and Remote Sensing to classify built-up area in Vellore city and adjoining areas. Population density and Euclidean distance from varied features were considered as the causative factors for urban sprawl. For this purpose, cloud free time series Landsat images from 2002 to 2020 were obtained from USGS database for Landsat and Sentinel images. Remote sensing and GIS techniques were used to obtain data for land parcel occupied by impervious extent. Initially, the unsupervised classification algorithms (either ISO or K-means) were implemented for each Landsat imagery after enhancing them through PCA or ICA to

endure handler precision just before selecting distinctive pixel distribution. Training data in the order of 10 times the number of bands (10n) were carefully chosen by means of image to image comparison technique (comparing classified image with historical archive of Google Earth images), and ground truth data obtained from toposheets and survey maps. Further, MLC or SVM approaches intended for supervised classification are applied to classify the subsequent enhanced images and urban area are validated through Shannon entropy or patchiness change matrix taking place on landscape levels. The statistical analysis such as MLR and Trend analysis for identification of relationship between urban spread and causative factors was performed.

2 Methodology

2.1 Study area

Vellore city (boundary obtained from Google maps) has an area of 98.3 km^2 is located in North Eastern part of Tamilnadu. The city experiences tropical savanna (semi-arid) climate with high temperatures ranging from 29 to 40 ^oC with a water shortage for almost entire summer season from March to July. Vellore lies on the banks of Palar river which is an underground river opening at Bethamangala town. The major source of water for the city is groundwater (current ground water level ranges from 0.3 m to 8 m bgl) and overhead tanks provided by municipal Palar and Karungamputhur water works. The geographical location, as shown in Fig. 1 of the city lies between 12°53′30″ to 12°57′30″ N as latitude and 79°3′30″ to 79°10′30″ E as longitude, 220 m above mean sea level. Vellore city has been identified as one of the 27 upcoming Smart cites by Government of India. Ease in proximity from the two major metropolitan cities of India namely, Chennai and Bangalore, makes Vellore susceptible for urbanization. VIT campus and CMC are the major urbanized or commercial regions in the city. Vellore is believed to be a highly spiritual place with more than seven temples, three churches, and three mosques which attracts tourism and increase urban sprawl.

The other land use features comprise of vegetative areas, open grounds and water bodies (permanent and perineal). Intended for the Smart city projects, Vellore is prone to increasing development that triggers the cumulative urban sprawl and property demand for residential and commercial purposes. In order to fulfil this land requirement vegetative lands are converted to built-ups in that way encroaching the pervious land usage.

2.2 Image processing and causative factor identification:

To model the spatial and temporal changes in the aforementioned features with respect to increasing urban extent densely time stacked cloud free Landsat images (Landsat MSS, TM and ETMT) from 2002 to 2020 are used extensively to develop LULC maps. The preliminary image corrections like radiometric and geometric corrections and enhancement procedures like Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are performed. Various learning techniques like supervised and unsupervised techniques are applied on the enhanced images to cluster similar spectral pixels with high degree of objectivity (Yand & Lo 2002 (a)). To ease the selection of training samples and cluster the similar spectral signatures, firstly unsupervised classification techniques (k-means and iso data) is performed. Image to image comparison is performed in which the classified image is contrasted to Google Earth achieve of historical images. A multiband image is obtained and each unknown feature is assigned with the mentioned 4 classes. Various indexes like NDWI or NDVI, NDBI, BI, DBI are performed to cross validate features obtained using unsupervised classification.

Table 1
List of Data used for the study

Map Type	Spatial Resolution	Year of map developed
SOI Toposheets	1:25000	1997 and 2003
Google Earth Images	60 m (for 1995–2001), 5 m (for 2002–2020)	1995–2020
Landsat images	TM (60 m), ETM+ (28.5 m), OLI/TIRS (28.5 m)	1995–2011 (TM), 1999–2019 (ETM+, SLC failed), 2013–2020 (OLI/TIRS C1 Level 1)
Digital Elevation Models (DEMs)	30 m	2000 (SRTM 1-Arc Second)
VIT Campus Maps	1cm:10 m (Drawings and AutoCAD Files)	1995, 2005, 2019
Municipal Boundary Map	1:25000	2000
Census data	Decadal	1991, 2001 and 2011
OSM Road Network		2006
Sewer Network	1:25000	2000
Soil Maps	1:500000	1992, 1996
Master Plan for Study Area	1:6000	2011

Signature files or training dataset are correspondingly generated using Google earth and ground reference data by signature editor option in ERDAS imagine tool. Training samples are selected in the order of *10n* (where, n is the number of classes acknowledged) along with ground truth data obtained by means of survey maps, toposheets obtained from SOI and city development plans as shown in Table 1. SVM classification technique is then performed for each set of images subjected to training samples obtained in the preceding phase. SVM tool typically incorporates variance and covariance of diverse signatures while assigning the feature category. The classified images are then subjected to accuracy assessment which involves selection of ground samples for each feature to compare with classified pixels, mathematically represented as Eq. 1. Accuracy assessment is done for each classified image considering 263 sample

points. Figure 2 shows the stepwise methodology to extract built-up pixels from densely time stacked series of classified images and the effect of different causative factors on urbanization change.

In the next step, built-up feature from respective years was extracted and landscape indices like Shannon's entropy, sparseness index and built-up map density were assessed by means of demographical and built-up area statistics from data sources mentioned in Table 1. Further, data for temporal causative factors is obtained from OSM layers to map roads and railways, Census reports for annual population data and Toposheets for urban core identification. Classified images are converted into polygons and Euclidean distance of urban features from the NHs, SHs, railway lines, kaccha roads and commercial landuse layers is calculated using near distance feature. The classified images for each year are shown in Fig. 3 for visual interpretation of changes in different features with time.

Individually, the causative factors are examined to identify their nature of regression towards increasing built-up pixel count. The attributes of individual causative factors and their effects on urbanization are further formulated using various regression algorithms like Linear, polynomial, logarithmic, powered and exponential. Nash Efficiency (NSE), Regression coefficient (R^2), and standard errors (SE) values are calculated for respective comparisons. Correlation amongst the transition or causative factors are assessed and the resultant MLR model is established to study urbanization changes due to combined effect of Euclidean distance from NHs, SHs, Railway lines, commercial areas and population density. A report in Deccan Chronicle (dated July 4, 2017) noted that by year 2030, Tamil Nadu will become the most urbanized state in the country with about 67% of its total population estimated to live in urban areas. Population growth and movement of people from villages and towns to cities initiates the expansion and development of urban regions.

3 Results And Discussions

3.1 Image Classification:

Georeferenced urban features for Vellore city were extracted using toposheets, google earth, manual survey, city master maps, and municipality water distribution maps. As per 2011 census, the rate of urbanization in Vellore was noted to be about 43.2% which is slightly below the average rate of urbanization for the state of Tamil Nadu. 78 landsat images from 2002 to 2020 (preferably for the days of cloud free and no-rainfall periods) with band description as shown in Table 1 are used to develop an urban growth model to describe the spatial and temporal variation of urban features and predict the impact on future locations, characteristics of commercial structures and consequences of increasing growth rate. Each image is tested for the signature verification for accurate pixel representation using following three measures: i) Histograms of standard deviation - to confirm unimodality, ii) Transformed divergence (TD) comparison, and iii) Contingency matrix - preferred during accuracy assessment. Necessary merging and redefinition operations are executed to maintain histogram's un-modality. As shown in Table 2, TD values more than 1900 represents good separation and further, the band composition for Landsat archive is done using bands with higher TD values.

Table 2
Average transformed divergence (TD) values for the images from 1989 to 2020.

Landsat Sensor	Year	Band composition	Spatial Resolution	Average TD Values	Indices Calculated
TM	1995–2000	1,3,4,5	57 m	1804	NDVI, NDWI
ETM+	1999–2019	1,2,3,4,5	28.5 m	1999	NDVI, NDWI, NDBI, BI, DBI
OLI and TIRS	2013–2020	2,3,4,5,6,7	28.5 m	1994	NDVI, NDWI, NDBI, BI, DBI

For urban features, the rendering mechanism performed fairly feeble because of similar reflectance values of urban settlement, exposed rocky structures and wet alluvial soil. The variation in reflectance values from series of densely stacked cloud free images collected for a cycle gap of 16 days assisted in separating urban settlement from the features with analogous characteristics. Image enhancement like PCA and ICA performed in ERDAS and ENVI improved the visual qualities and equalized band histograms for the satellite image. The enhanced images (as shown in Fig. 4a) attained are exposed to two unsupervised classification techniques: K-means and ISO-data, to cluster the pixels into user specified number (shown in Fig. 4b). The time staked reflectance values from classified images were obtained and re-grouped in the major 4 classes: Water, Built-up, Vegetation and Barren lands. The classified images are further compared with historical image dataset available in Google earth to obtain training datasets. The training clusters for 4 varied features are selected on the basis of following feature classes:

1. Training samples for Water - Water bodies surrounding Vellore Fort, BHEL industrial discharge points etc.
2. Training samples for Built-up - Institutional areas like VIT Campus, CMC college and hospital, Vellore Government hospital, Residential colony, Katpadi railway station and Central jail, Government offices.
3. Training samples for Vegetation - Mountainous range (Rocks with no mining activities are preferred), parks and lawns in VIT campus, Vellore fort and CMC college.
4. Training samples for Barren lands - Mining regions, Open areas like school and institutional grounds, officer's lane stadium, Airport runway (proposed) and Central Jail ground.

Selected training samples were verified using Google earth, SOI toposheets (Toposheet No.: D44N4, D44N8 and D44T5), archeological and historical backgrounds of buildings, survey sheets, expertise opinions and proposed maps of the features by government and private organizations (available in VIT

estate office and municipal authorities). Pixel variance and covariance of signatures obtained for supervised classifiers were revised to evaluate the image enhancement procedure. Almost 263 training classifiers were obtained for 4 features (Water, Built-up, Vegetation and Barren lands) later subjected to Supervised learning techniques (a sample image for the year 2020 is shown in Fig. 4c). The classified images were then exported to a knowledge-based system in which an ancillary information about various features from DEM, soil maps, municipal boundaries and water bodies were integrated. This technique proved most effective for resource allocation and accuracy assessment studies in remote sensing environment (2). Validation datasets were obtained using random stratified sampling technique from the google earth, ground truth data and toposheets.

3.2 Causative factors and Built-up area extraction:

The number of urban cells is calculated by integrating the supplementary features in one non-built-up class. The average percentage increase in built-up area was 51.45% from year 2002 to 2019, with an annual increasing pattern as shown in Table 3. The core urbanization for Vellore province increased due tourist attraction and commercialization surrounding NHs and SHs. Also, growing population trends were observed in the zones near to the institutional and market locations. According to expertise opinions and literature survey in various parts of India (Ajmer, Pune, Chennai, Delhi, Bangalore), population density, distance from roads, railways and commercial regions, were selected as causative factors contributing to urbanization increase. Population data for study area is acquired from the Census data and online portal (www.population.cit/India/Vellore) showing a growth rate of + 2.28% from 2001 to 2011. Decadal census statistics and the population growth were plotted for trend analysis in a regression model and a 3-degree polynomial proved to be best fit comparing to linear, exponential, logarithmic and powered distribution. Eq. 2 shows the variation of population for the study area vs. decadal variation with its correlation coefficient as 0.99.

Table 3
Built-up pixel count, LULC Accuracy Assessment, Population data, Proximity from Different Features and Threshold Proximity.

Year	Built-up Pixels	Population (in thousands)	Average Proximity (m)	Proximity from commercial landuse (m)	Threshold proximity (m)
2002	30490	394.54	103.54	154.285	390.68
2003	31495	403.33	142.61	175.008	392.39
2004	31581	412.41	142.81	192.101	393.22
2005	31984	421.77	152.65	199.589	409.11
2006	34307	431.44	156.43	205.172	409.20
2007	35570	441.44	161.11	209.685	414.28
2008	35758	451.76	166.05	210.313	422.29
2009	37118	462.44	187.56	215.438	426.23
2010	37211	473.48	187.78	217.200	428.99
2011	38108	484.91	191.32	217.474	431.43
2012	38915	496.72	191.63	218.734	432.14
2013	40107	508.94	200.96	224.196	436.14
2014	43825	521.59	201.49	226.787	440.32
2015	43947	534.67	202.57	234.010	442.21
2016	44369	548.21	203.71	235.090	451.22
2017	46230	562.21	226.19	249.305	452.08
2018	50932	576.69	234.01	414.277	452.16
2019	60922	591.67	236.79	416.148	461.19
2020	62812	607.16	241.98	421.154	463.49

where, P is in thousand and x is the years in decade

The percentage increase in the population from 2002 to 2020 is estimated to be 53.8% which shows that by 2030, the increase in population can cross 67% (as given in deccan chronicle (dated July 4, 2017)). Logarithmic regression equation for the population data gives the lowest correlation with an R^2 value of 0.91 whereas for the other distributions (exponential, linear, quadratic polynomial and powered) values are as 0.9929, 0.9981, 0.9985 and 0.9761 respectively. Eq. 2 was further used to estimate the population density and growth rate for each year. Population (P) data (in thousands) and number of built-up pixels for each year is shown in Table 3. Built-up areas extracted from the classified images and the population density were then correlated and a regression model is formulated using excel as shown in Eq. 3.

where, is the number of built-up pixels in hundreds

The built-up area from 2002 to 2020 is increased by 80.82% approximately. In addition, pixel transformation analysis of built-up areas in classified images is done to study the distribution along the urban landuse like roads, railways and commercial areas. Urban sprawling, population density and economic indices alongside the commercial landuse was rapid as compared to the far-flung ranges further validated using census of India report, 2011. Layers of urban/non-urban variation, permanent commercial features proximity, proximity from roads and railways, and population density are prepared at local scale (as shown in Fig. 4d). On the other hand, factors associated with social, cultural and political influence on urbanization change cannot be evaluated quantitatively. The impact of population change is also analyzed against the commercial regions and found to positively correlated. Figure 3c shows the calculation of distance of various features from roads, railways and commercial regions. From the proximity analysis, the distance from national highways, state highways and railways are found to be the major factor for the population increase and urbanization change as compared to proximity through township roads and commercial regions.

where, is the Distance from roads and railways

is the Distance from tourist and commercial areas

Increasing built-ups in the study area were analyzed for the proximity from the roadways and railways by dividing the decision criteria as inert features responsible for their correlation. The inert features considered for this study are: permanent water bodies, existing landmarks or developed areas, national and state highways, railways, pilgrimage and tourist places. A slope of more than 15% is not considered for urban sprawl. The proximity from settlements and tourist places like Vellore Fort, Golden Temple etc. were considered as constraints responsible for conversion of barren or vegetative spaces to urban features and upsurge urbanization around the landscapes. The second causative factor is proximity from road and railway network which was responsible to transform the forest and barren area to new-fangled impervious surfaces parting the restricted spaces aside. The constraints on the proximity to road and rail network is: the area falling in between the equidistant lines of distance between 30 to 232 m are more prone to urbanization. The commercial regions like university campuses, hospitals and market areas showed a major contribution in the increasing sprawling where the development became motionless approximately 336 m away from these areas. In some places, land parcels adjacent to the national highways are provided with trivial shrubberies aiding to sustainability and smart city requirements. Altogether, the above factors considered during the study helps to develop a MLR (shown in Fig. 4e) for analyzing the change of urban parcels due to a combined effect of increasing population, proximity through roads, rails and commercial area. Table 4 shows the combined MLR models for four combinations for causative factors vs. the built-up pixel variation with SE and NSE calculations. The first correlation for population, proximity of roads, railways and commercial regions with built-up pixel gives the maximum NSE value and can be selected for the future studies on identification of increased urban extent with time.

Table 4
Multivariate regression equations w.r.t dependent variables as Population (P), Proximity from roads and railways (D_r) and from commercial area

Causative Factors	Dependent Variable	Multivariate Linear Regression Equation
P, D_r, and D_B	Built-up Area (in km^2) - PB	
P, D_r, D_B, $P*D_r$, $P*D_B$, D_r*D_B		
P, D_r, D_B, $P*D_r$, $P*D_B$, D_r*D_B and $P*D_r*D_B$		
P, D_r, D_B, and $P*D_r*D_B$		

4. Conclusion

One of the potential challenges for environmental protection developers, designers and planning agencies is to optimally utilize and allocate the natural resources in developing parcels. Decision-making in infrastructures initiatives, also depends on the potential harm to the environment due to the type of land transformation. Thus, identification of trend in urbanization and its causative factors becomes a prime importance for various authorities. So, the study performed deals firstly with the identification of urban areas using satellite images using supervised and unsupervised learning techniques. Urban pixel count extracted using densely stacked cloud free Landsat images were calculated. The annual comparison of the classified images showed almost 51% increase in urban pixel count from 2002 to 2020, while, official reports affirm an urban growth of 43.2% till 2017. The major reasons for increase in urban area were recognized as increasing population, resident requirements and commercialization. In this study the major conversion from vegetative to built-up expanse was

apparent in peripheries of the road-rail networks and commercial/income generating areas. Pixelwise analysis of classified images and comparing the classified images with historical archive of google imagery showed that the vegetative and barren feature classes were converted to settlements. Proximity matrix and increasing population were evaluated as per different causative factors apart from social and political tools (because the quantitative investigation of these parameters is unattainable). The growth rate in population was obtained for Vellore city and the average growth rate of 2.87 was observed with a polynomial regression variation (R^2 = 0.99, SE = 2.57). Furthermore, various regression techniques like linear, exponential, logarithmic, polynomial and powered, were performed for each causative factor out of which polynomial exhibited the minimum standard error (SE = 2.072, NSE = 0.87, R^2 = 0.96 for PB vs. P, SE = 2.61, NSE = 0.54, R^2 = 0.94 for PB vs. D_r and SE = 22.94, NSE = 0.81, R^2 = 0.88 for PB vs. D_B). Individual analysis for each causative factor showed that the distance from major roads and railways plays the most important role for urbanization increase which can be directly correlated to population change also. To analyze the combined effect of each causative factor on urbanization change and to develop a set of transition rules for prediction models, MLR for P, D_r and D_B was adopted. This multivariate relationship developed for Vellore city is useful for the local municipal authorities to easily quantify areas necessitated for resource allocation, daily needs scheduling, land acquisition, fund allocation for zonal or ward wise distribution, regional planning, designing for stormwater and sewage drainage system and so on. The uncertainty analysis of causative factors like tourism, religious places, public government partnerships, physical barriers, political domiciles, proper drainage system etc. may aid to improve the urban growth modelling as long as data availability is not an intricate factor.

Declarations

Availability of data and materials

Not Applicable.

Acknowledgements

The authors are thankful to USGS Earth Explorer for providing all kinds of support on open source satellite during compilation of the present work.

Funding

We would also like to acknowledge that there was no funding agency involved in the study and all the data used are open source and freely available on the websites mentioned in the study.

Contributions

Mr. Abhinav Wadhwa and Dr. Pavan Kumar K have been working on Remote Sensing and GIS for the project and contributes as conceiving of the presented idea, designing the theory of the project, and verification of the analytical methods. The designed methodology and the conclusions have been discussed thoroughly by the authors and the outcomes can be useful for many research purposes. Both the authors helped shaping the manuscript for research and analysis, along with discussed the results, provided critical feedback and contributed to the final manuscript. All authors read and approved the final manuscript.

Ethics declarations

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest (such as personal or professional relationships) in the subject matter or materials discussed in this manuscript. Finally, we thank the re-viewers and the editor for their profound and constructive comments.

Competing interests

The authors declare that they have no competing interests.

References

1. Abdullahi, S., Pradhan, B., Jebur, M.N., 2015. GIS-based sustainable city compactness assessment using integration of MCDM, Bayes theorem and RADAR technology. Geocarto Int. 30, 365–387. https://doi.org/10.1080/10106049.2014.911967
2. Al-sharif, A.A.A., Pradhan, B., 2016. Spatio-temporal Prediction of Urban Expansion Using Bivariate Statistical Models: Assessment of the Efficacy of Evidential Belief Functions and Frequency Ratio Models. Appl. Spat. Anal. Policy 9, 213–231. https://doi.org/10.1007/s12061-015-9147-1
3. Alsharif, A.A.A., Pradhan, B., 2014. Urban Sprawl Analysis of Tripoli Metropolitan City (Libya) Using Remote Sensing Data and Multivariate Logistic Regression Model. J. Indian Soc. Remote Sens. 42, 149–163. https://doi.org/10.1007/s12524-013-0299-7
4. Alsharif, A.A.A., Pradhan, B., Mansor, S., Shafri, H.Z.M., 2015. Urban expansion assessment by using remotely sensed data and the relative Shannon entropy model in GIS: A case study of Tripoli, Libya. Theor. Empir. Res. Urban Manag. 10, 55–71.

5. Amini Parsa, V., Yavari, A., Nejadi, A., 2016. Spatio-temporal analysis of land use/land cover pattern changes in Arasbaran Biosphere Reserve: Iran. Model. Earth Syst. Environ. 2, 1–13. https://doi.org/10.1007/s40808-016-0227-2
6. Ban, Y., Hu, H., Rangel, I.M., 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land-cover mapping: Object-based and knowledge-based approach. Int. J. Remote Sens. 31, 1391–1410. https://doi.org/10.1080/01431160903475415
7. Bradley, G.A., 1997. Analyzing the Data. Chinese Econ. 30, 47–56. https://doi.org/10.2753/ces1097-1475300547
8. Cihlar, J., 2000. Land cover mapping of large areas from satellites: Status and research priorities. Int. J. Remote Sens. 21, 1093–1114. https://doi.org/10.1080/014311600210092
9. Del Frate, F., Pacifici, F., Schiavon, G., Solimini, C., 2007. Use of neural networks for automatic classification from high-resolution images. IEEE Trans. Geosci. Remote Sens. 45, 800–809. https://doi.org/10.1109/TGRS.2007.892009
10. Haack, B.N., Rafter, A., 2006. Urban growth analysis and modeling in the Kathmandu Valley, Nepal. Habitat Int. 30, 1056–1065. https://doi.org/10.1016/j.habitatint.2005.12.001
11. Herold, M., Gardner, M.E., Roberts, D.A., 2003. Spectral Resolution Requirements for Mapping Urban Areas. IEEE Trans. Geosci. Remote Sens. 41, 1907–1919.
12. Jat, M.K., Garg, P.K., Khare, D., 2008. Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. Int. J. Appl. Earth Obs. Geoinf. 10, 26–43. https://doi.org/10.1016/j.jag.2007.04.002
13. Jia, K., Wei, X., Gu, X., Yao, Y., Xie, X., Li, B., 2014. Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. Geocarto Int. 29, 941–951. https://doi.org/10.1080/10106049.2014.894586
14. Li, C., Wang, J., Wang, L., Hu, L., Gong, P., 2014. Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. Remote Sens. 6, 964–983. https://doi.org/10.3390/rs6020964
15. Long, H., Liu, Y., Wu, X., Dong, G., 2009. Spatio-temporal dynamic patterns of farmland and rural settlements in Su-Xi-Chang region: Implications for building a new countryside in coastal China. Land use policy 26, 322–333. https://doi.org/10.1016/j.landusepol.2008.04.001
16. Lu, D., Mausel, P., Batistella, M., Moran, E., 2005. Land-cover binary change detection methods for use in the moist tropical region of the Amazon: A comparative study. Int. J. Remote Sens. 26, 101–114. https://doi.org/10.1080/01431160410001720748
17. Lu, D., Weng, Q., 2005. Urban classification using full spectral information of Landsat ETM+ imagery in Marion County, Indiana. Photogramm. Eng. Remote Sensing 71, 1275–1284. https://doi.org/10.14358/PERS.71.11.1275
18. Mundia, C.N., Aniya, M., 2005. Analysis of land use/cover changes and urban expansion of Nairobi city using remote sensing and GIS. Int. J. Remote Sens. 26, 2831–2849. https://doi.org/10.1080/01431160500117865
19. Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., Teller, J., 2018. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. Comput. Environ. Urban Syst. 67, 147–156. https://doi.org/10.1016/j.compenvurbsys.2017.09.009
20. Osgouei, P.E., Kaya, S., Sertel, E., Alganci, U., 2019. Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery. Remote Sens. 11, 1–24. https://doi.org/10.3390/rs11030345
21. Reddy, K.R., Devaraj, S., Biradar, S., Yarrakula, K., Srinivas Kumar, K., 2019. Spatial distribution of land use/ land cover analysis in Hanamkonda taluk, Telangana - A case study. Indian J. Geo-Marine Sci. 48, 1761–1768.
22. Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sens. Environ. 124, 689–704. https://doi.org/10.1016/j.rse.2012.06.006
23. Stuckens, J., Coppin, P.R., Bauer, M.E., 2000. Integrating contextual information with per-pixel classification for improved land cover classification. Remote Sens. Environ. 71, 282–296. https://doi.org/10.1016/S0034-4257(99)00083-8
24. Sudhira, H.S., Ramachandra, T. V., Jagadish, K.S., 2004. Urban sprawl: Metrics, dynamics and modelling using GIS. Int. J. Appl. Earth Obs. Geoinf. 5, 29–39. https://doi.org/10.1016/j.jag.2003.08.002
25. Vogelmann, J.E., Sohl, T., Howard, S.M., 1998. Regional characterization of land cover using multiple sources of data. Photogramm. Eng. Remote Sensing 64, 45–57.
26. Wang, W. wu, Zhu, L. zhong, Wang, R. chao, Shi, Y. jun, 2003. Analysis on the spatial distribution variation characteristic of urban heat environmental quality and its mechanism - A case study of Hangzhou City. Chinese Geogr. Sci. 13, 39–47. https://doi.org/10.1007/s11769-003-0083-7
27. Webster, D.R., 2001. On the Edge: Shaping the Future of Peri-urban East Asia. Stanford Univ. Pacific Res. Cent. 1–53.
28. Weng, Q., 2001. Modeling urban growth effects on surface runoff with the integration of remote sensing and GIS. Environ. Manage. 28, 737–748. https://doi.org/10.1007/s002670010258
29. Yang, X., Liu, Z., 2005. Use of satellite-derived landscape imperviousness index to characterize urban spatial growth. Comput. Environ. Urban Syst. 29, 524–540. https://doi.org/10.1016/j.compenvurbsys.2005.01.005

Figures

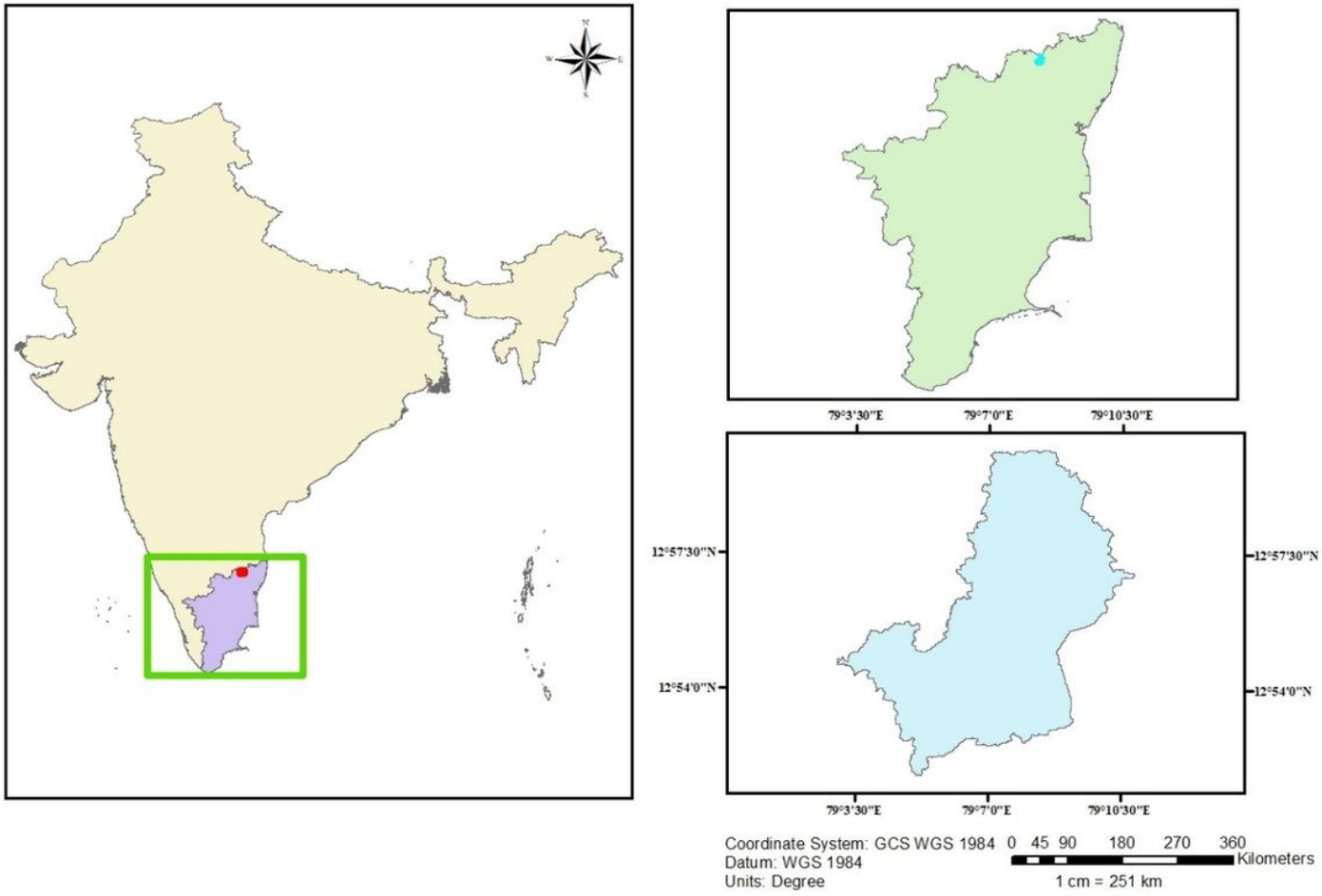

Figure 1

Vellore city and adjoining areas most likely influenced for urbanization

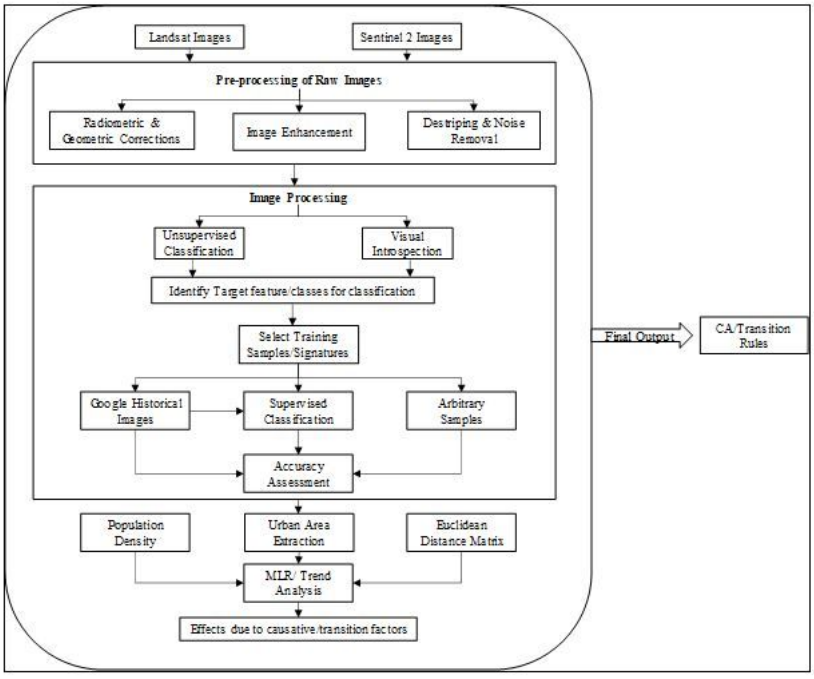

Figure 2

Detailed Methodology chart classified as Preprocessing of raw satellite images obtained, Image processing and identification of causative factors

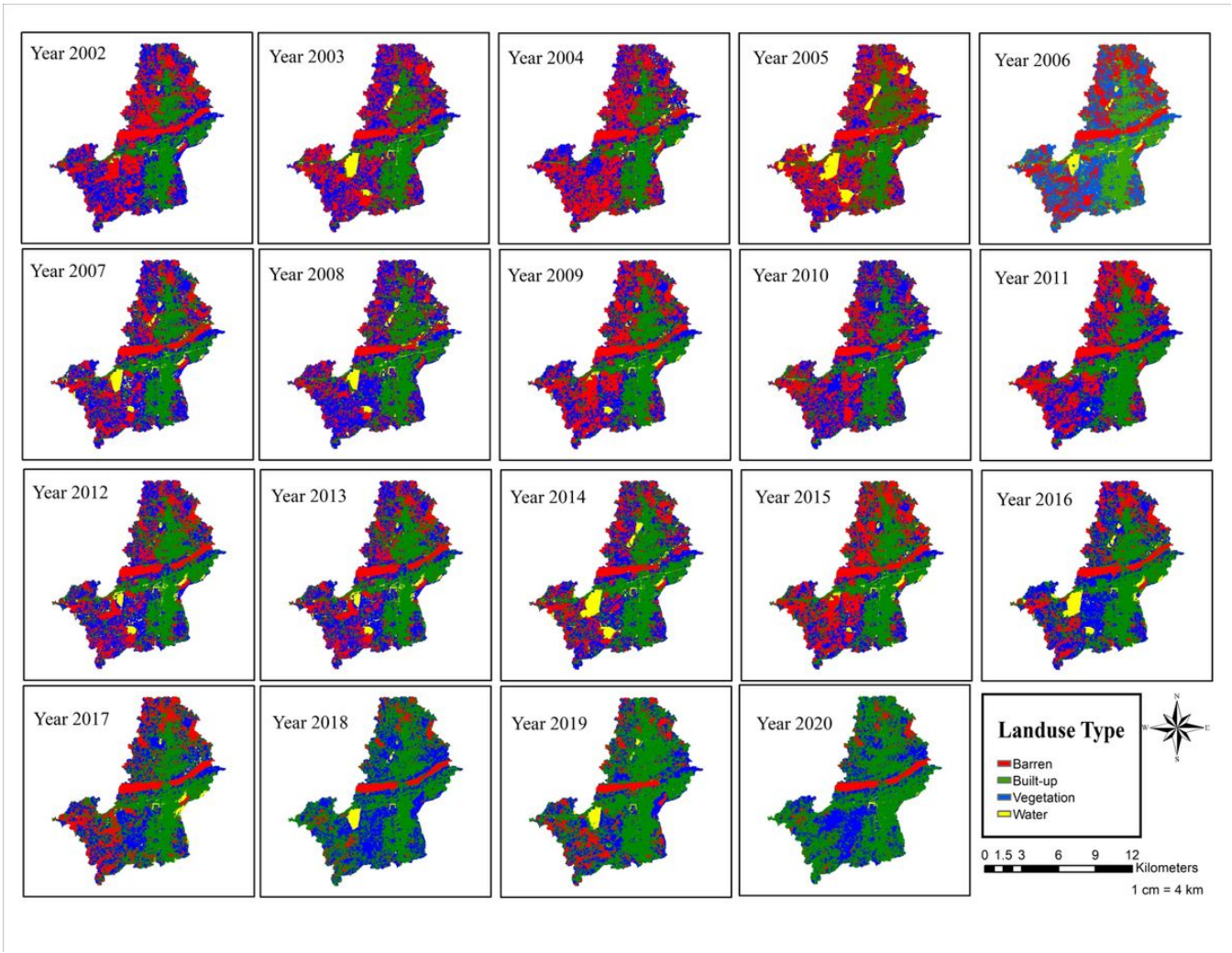

Figure 3

Year-wise classified landsat images for four major features

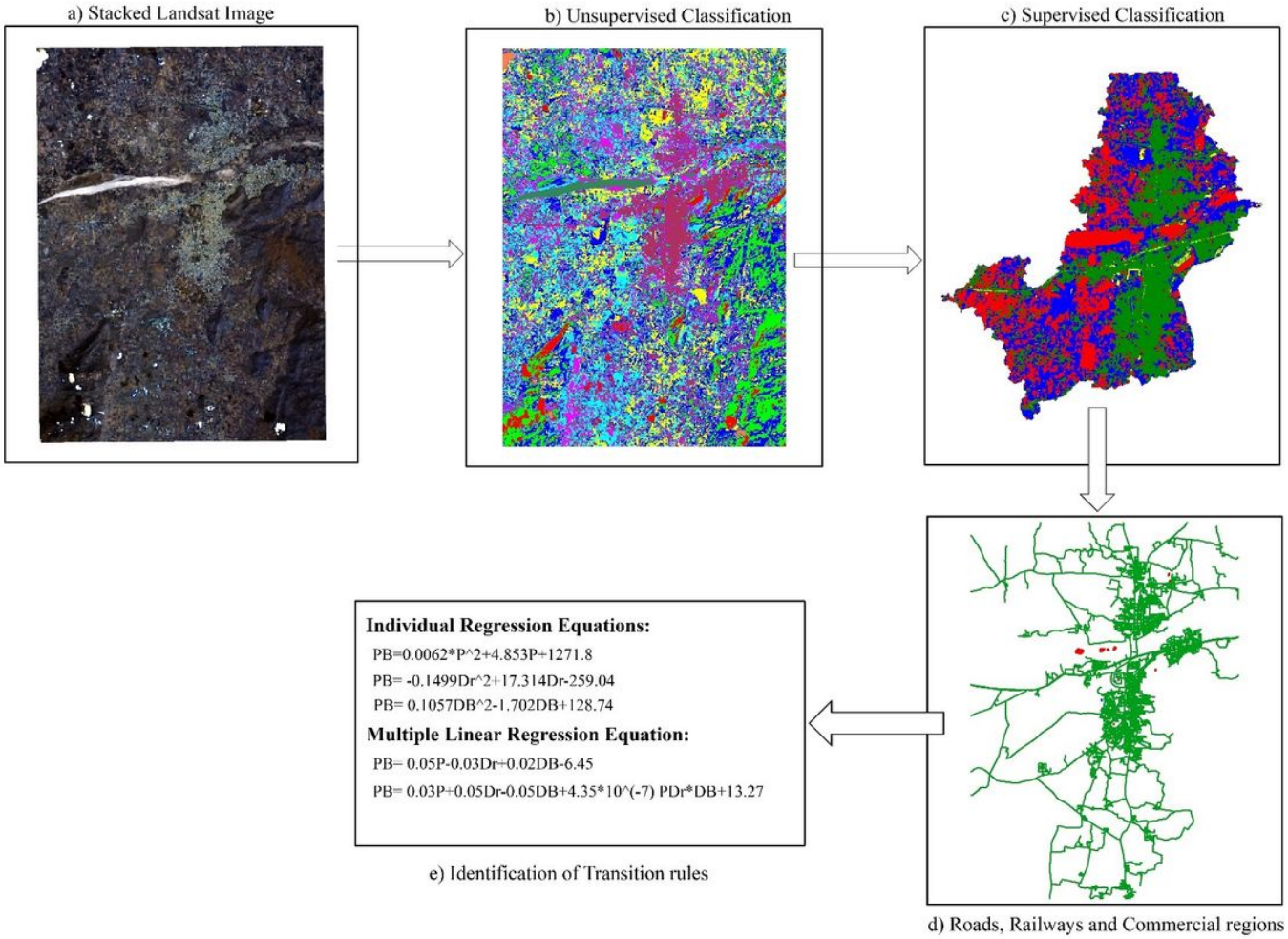

Figure 4

Stepwise procedure to evaluate the quantitative effect of transition factors