DiVA
http://www.diva-portal.org

This is the published version of a paper published in .

Citation for the original published paper (version of record):

Devendran, A A., Gnanappazham, L. (2018)
Urban growth prediction using neural network coupled agents-based Cellular Automata model for Sriperumbudur Taluk, Tamil Nadu, India
*The Egyptian Journal of Remote Sensing and Space Science*, 21(3): 353-362
https://doi.org/10.1016/j.ejrs.2017.12.004

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-107371

The Egyptian Journal of Remote Sensing and Space Sciences 21 (2018) 353–362

HOSTED BY

Contents lists available at ScienceDirect

The Egyptian Journal of Remote Sensing and Space Sciences

journal homepage: www.sciencedirect.com

Research Paper

# Urban growth prediction using neural network coupled agents-based Cellular Automata model for Sriperumbudur Taluk, Tamil Nadu, India

Aarthi Aishwarya Devendran, Gnanappazham Lakshmanan *

*Department of Earth and Space Science, Indian Institute of Space Science and Technology, Trivandrum, India*

ARTICLE INFO

*Article history:*
Received 1 May 2017
Revised 31 October 2017
Accepted 28 December 2017
Available online 20 January 2018

*Keywords:*
Urbanization
CA model
Neural network coupled agents-based CA model
Sriperumbudur Taluk
Shannon's entropy
Urban modeling

ABSTRACT

Unplanned urbanization would pose serious threats to both environment and mankind. Hence, urban growth model (UGM) becomes mandatory to predict future growth of a city. In the current study, urban growth of Sriperumbudur Taluk, Tamil Nadu, India was predicted using three types of Cellular Automata (CA) model namely Traditional CA (TCA) model, Agents based Cellular Automata (ACA) Model and Neural Network coupled Agents- based Cellular Automata (NNACA) model. The urban maps of the study region for the years 2009, 2013 and 2016 along with the influencing agents of urbanization namely transportation, industries, elevation and also hotspot locations based on the Government policy were used in the modeling. Analytical Hierarchical Process (AHP) technique was adopted to estimate the weights of the agents for suitability map preparation in ACA model. On validating 2016 predicted outputs, NNACA model proved to be the better urban model (kappa coefficient - 0.72) when compared to TCA and ACA models (kappa coefficient - 0.6 each). Shannon's entropy measure revealed that the urbanization is concentrated in the north-east direction and it is predicted to have an urban sprawl area of 157 $km^2$ in 2020 using NNACA model while the observed urbanization is 113 $km^2$ of the area in 2016.

© 2018 National Authority for Remote Sensing and Space Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

Cities expand because of increase in population and migration of people from rural places to cities. Unplanned urbanization degrades the natural resources and affects the quality of life of the humans. Urban prediction models will help decision makers to plan the growth of their city in the future (Hill and Lindner, 2010). In recent days, the availability of high-resolution temporal satellite data and different types of GIS-based modeling tools help in the handling of spatial data and its analyses, making urban growth models (UGMs) more realistic.

CA models are widely used among all developed UGMs as they are found to perform well in predicting urban development more close to the reality than conventional mathematical models (Divigalpitiya et al., 2007; Mubea, 2014). CA consists of five basic elements namely Cell Space, Cell State, Cell Neighborhood, Transition Rule (TR) and Time. Cell space is the one in which the automaton is present. Cell state refers to the land use category occupied by the cell at a given time. Neighborhood includes the surrounding cells which could be square (3 x 3, 5 x 5), circular, hexagonal etc. Cells change their state based on the state of its neighborhood cells in discrete time steps through TR. Time is the temporal dimension in which a cellular automaton exists and thus the states of all cells are updated simultaneously at all iterations over time (Liu, 2009). A detailed review on CA-based UGMs is given by (Triantakonstantis and Mountrakis, 2012). SLEUTH, based on CA model, is one of the most popular simulation models to predict the urbanization and was implemented to predict the urbanization of Xiamen city of China in 2020 (Hua et al., 2014) while Chaudhuri and Clarke (2013) gave a detailed review of this model.

Apart from the historical land use data, various agents triggering urban growth is also considered in urban modeling. CA model along with various agents namely road, slope and exclusion layers were used to explore the urban growth of Tehran (Jokar Arsanjani et al., 2013), Nairobi city of Africa (Mubea et al., 2014) and Mumbai (Shafizadeh Moghadam and Helbich, 2013).

Recently, neural networks are integrated with GIS and CA models for predicting urban land uses. Unlike TCA and ACA models, NNACA model does not require users to provide TR, which are subjective (Yeh and Li, 2002). Also, NNACA models are capable of predicting the complex, non-linear, spatial-temporal process of urban

Peer review under responsibility of National Authority for Remote Sensing and Space Sciences.

* Corresponding author.
*E-mail addresses:* aarthiaishwarya.14@iist.ac.in (A.D. Aarthi), gnanam@iist.ac.in (L. Gnanappazham).

https://doi.org/10.1016/j.ejrs.2017.12.004
1110-9823/© 2018 National Authority for Remote Sensing and Space Sciences. Production and hosting by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

growth with less uncertainty (Basse et al., 2014). Artificial Neural Network (ANN) was integrated with CA model to predict the land use of Changping district of China in 2008 with different types of agents such as traffic data and urban and town areas (Yang et al., 2016).

In the present study, the performances of three types of CA models namely TCA, ACA, and NNACA are compared in predicting the urban growth in 2016 for Sriperumbudur Taluk, a southwest extension of Chennai, in the state of Tamil Nadu, India, using 2009 and 2013 urban covers as historical data. The most efficient model is used to predict the urbanization in 2020 for the study region using 2013 and 2016 urban land cover.

## 2. Study area

Chennai, the capital of the state of Tamil Nadu, is the India's fourth largest metropolitan cities in India after Mumbai, New Delhi and Kolkata. Owing to the development of many Information Technology centers and educational institutions, South Chennai in the recent decades is witnessing an exponential urban growth. In 1999, Hyundai commenced the automotive industries in Sriperumbudur which triggered the rate of urbanization added up with the development of new international airport (The Growth Story of Sriperumbudur Real Estate, 2013).

In the present study, the prediction model was implemented for Sriperumbudur Taluk of Kancheepuram district (Fig. 1). The study region encompasses 645 km$^2$ area and a population of 4,86,000 as per 2011 census. Chembarambakkam, Sriperumbudur, Manimangalam and Somangalam Lakes are some of the major water bodies located within this study region. Agriculture is the major occupation of the people. But recently, due to rapid industrialization, many agricultural lands are converted for residential plots and industries. Apart from Sriperumbudur Taluk Headquarters, Oragadam, Kundrathur, Mangadu, Manapakkam, Karasangal are other major locations in this taluk where continuous growth in settlements is observed in the past 10 years.

## 3. Data and methods

To model the urban growth, five major spatial parameters were used considering the nature of urban development in this area.

1. Urban maps of 2009, 2013 and 2016 were derived using satellite images of Landsat 5 TM of 6th May 2009 and Landsat 8 data of 17th May 2013 and 23rd April 2016 (https://earthexplorer.usgs.gov/).
2. Road network and road junctions (http://www.openstreetmap.org/) updated with Google Earth (Fig. 1).
3. SRTM (Shuttle Radar Topography Mission) DEM (Digital Elevation Model) of 1 arc-second (~30 m) resolution.
4. Categories of industries were obtained from Government of India MEPZ (Madras Export Processing Zone) Special Economic Zone (http://www.mepz.gov.in/Directory_SEZ.php?page=18&District=&Entity=&SEZ=&Unit=) (Fig. 1).
5. Google Earth integrated field knowledge was used for the validation of urban land cover mapping of the study period.

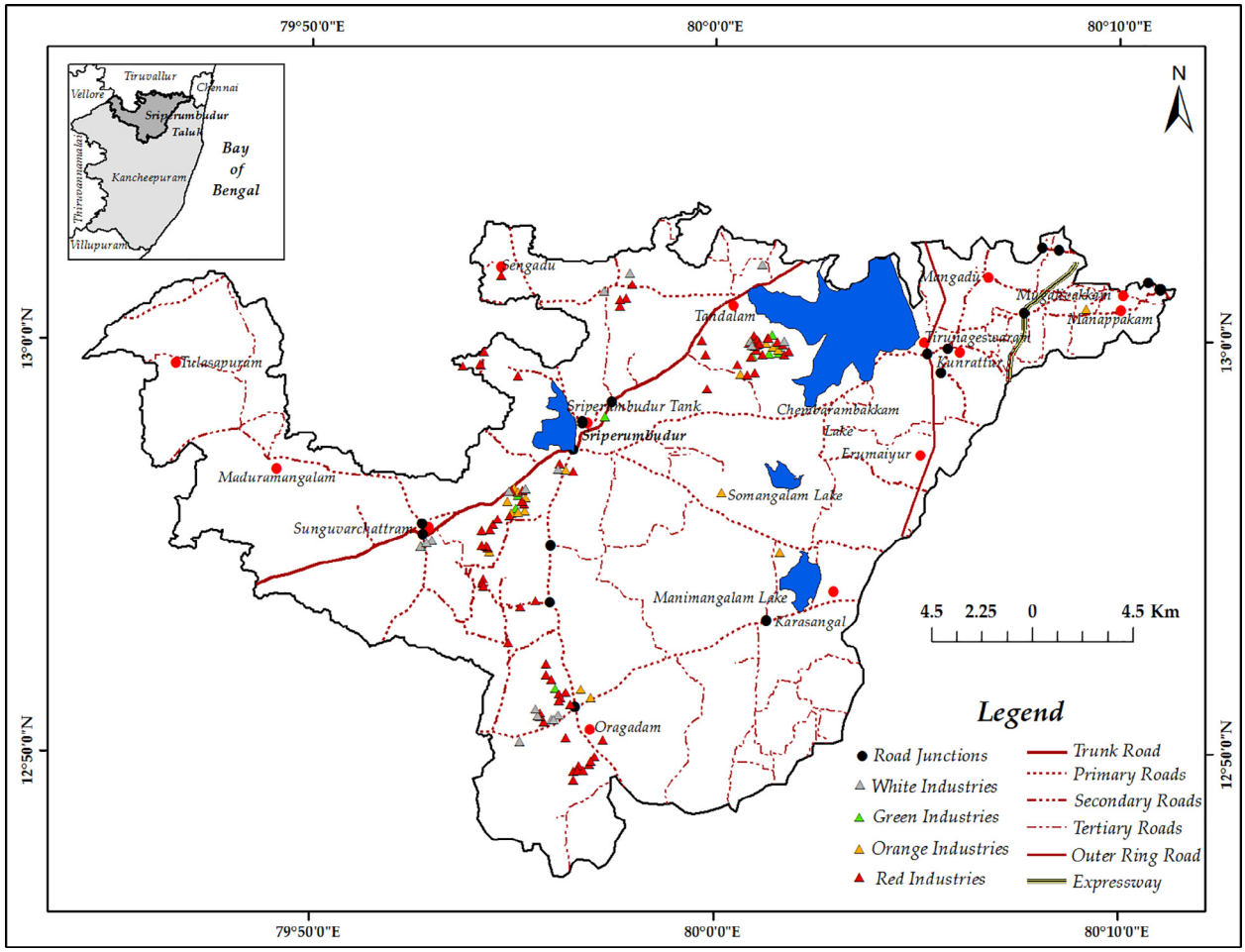


**Fig. 1.** Map of the Study Region.

### 3.1. Urban cover

Landsat data of the years 2009, 2013 and 2016 (as given in Section 3) were used in the land cover map preparation for the study area. Support Vector Machine (SVM) technique of Supervised Classification (Huang et al., 2002) was adopted to classify the land cover features such as Built-Up, Vegetation, Waterbody and Open land. Since the current study focuses on urban growth the land cover maps were converted to Built-up and Non-Built-up categories by combining last three land cover categories into a single category. Thus, the urban maps of the years 2009, 2013 and 2016 were obtained with two categories namely 'Built-Up' and 'Non-Built-Up'.

### 3.2. Identification of hotspots

New SIPCOT (State Industries Promotion Corporation of Tamil Nadu) Industrial Park was recently established in Vallam-Vadagal villages of Sriperumbudur Taluk over an area of 4.64 $km^2$ (Government of Tamil Nadu, 2014). In 2013, this region was identified as non-built-up. But in CA model, the state of a cell at any given time depends on its own state and the state of its neighbors at the previous step (Wolfram, 1984). Thus, built-up pixels of minimum size to initiate the urban growth are needed in a completely non-built-up cover. Hence, built-up areas of 3.65 $km^2$ at Vadagal and 0.99 $km^2$ at Vallam were used as built-up 'hotspots' in the urban map of 2013 to trigger the prediction of urban growth in 2016. Thus for the prediction of urbanization in 2016 through three types of CA models, two scenarios were considered. One scenario included these hotspot areas at Vallam, Vadagal villages in the 2013 urban map and the other without the hotspots.

### 3.3. Prediction modeling techniques

Once the urban maps of the study region were prepared, the prediction modeling was carried out using (i) TCA Model (ii) ACA Model (iii) NNACA model (Fig. 2) using 2009 and 2013 data to predict 2016 with the major water bodies as constraints.

#### 3.3.1. TCA model

In case of TCA model, the urban maps of the years 2009 and 2013 (with and without hotspots) were used for the urban prediction of 2016 of the study region. The urban maps were used to obtain the Transition Potential Matrix (TPM) between 2009 ($t_1$) and 2013 ($t_2$) and can be represented as

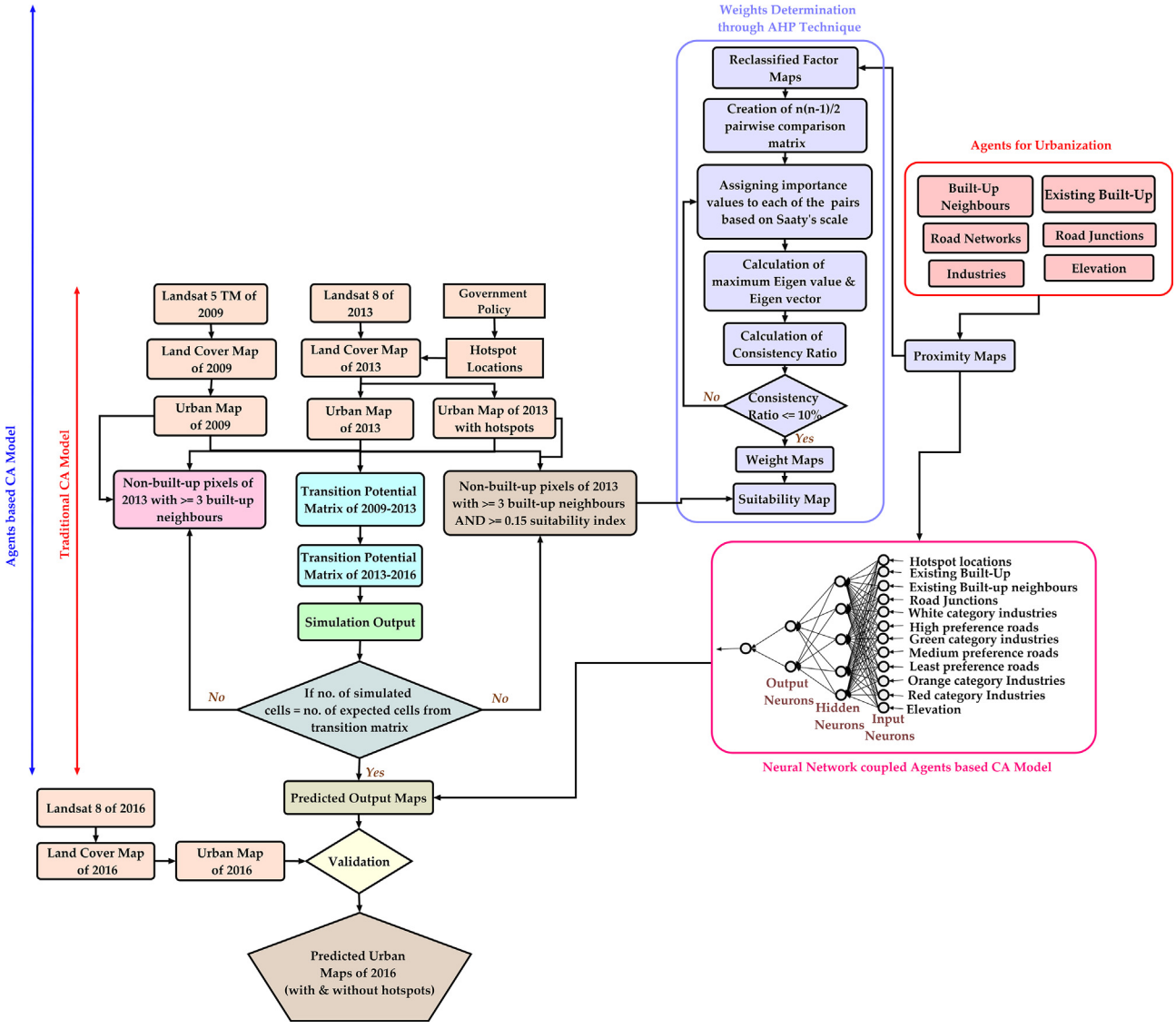


**Fig. 2.** Methodology adopted in the study for the prediction of urban areas using TCA, ACA and NNACA models.

$$X_n = \begin{bmatrix} p_{b-b} & p_{b-nb} \\ p_{nb-b} & p_{nb-nb} \end{bmatrix} \quad (1)$$

where, p: Probability of n time steps = $\frac{N_c}{N_T}$ ($N_c$: number of pixels in each category; $N_T$: Total number of pixels); *b*: State of Built-up; *nb*: State of Non-Built-Up. In the current study, to determine the TPM for one-time step, elements of equation 1 was divided by the number of time steps between the transition periods, i.e., n = 4. $p_1 = \frac{p}{n}$ with the following consideration (Chung, 1960).

i. Sum of row elements = 1
ii. Diagonal elements are changed to satisfy (i)
iii. $p_{b-nb} = 0$ (Assumption)
iv. Hence, $p_{b-b} = 1$

Thus the TPM for one time step ($X_1$) is represented as

$$X_1 = \begin{bmatrix} 1 & 0 \\ p_{1_{nb-b}} & 1 - p_{1_{nb-b}} \end{bmatrix} \quad (2)$$

Hence, TPM between 2013 and 2016 is given by

$$X_{n+1} = X_n (X_1)^m \quad (3)$$

where, $X_{n+1}$: Projected TPM between $t_2$ and $t_3$; m: number of time steps between $t_2$ and $t_3$ (here, m = 3)

**Table 1**
Overall Accuracy and kappa values of land cover maps prepared using SVM technique for the study region.

| | 2009 | 2013 | 2016 |
|---|---|---|---|
| OA | 99.19% | 94.91% | 92.80% |
| k | 0.98 | 0.91 | 0.87 |

A 3 × 3 neighborhood was used in the study for CA neighborhood configuration. The TR adopted in TCA model was *'to find all the non-built-up pixels with greater than or equal to 3 built-up neighbors'* (left of Fig. 2). The TCA model was run in ArcGIS environment till the expected numbers of pixels are reached in 2016 based on the TPM. The predicted urban maps based on TCA model with and without hotspots were validated against the observed urban map of 2016 and the accuracies were reported.

### 3.3.2. ACA model

Unlike TCA models, ACA models use agents of urbanization apart from time series urban maps for more realistic prediction outputs. In the current study, based on the inputs from experts from CMDA (Chennai Metropolitan Development Authority), urban planning departments and academicians, agents responsible for urban growth in Sriperumbudur Taluk were identified namely Road Networks, Road Junctions, Industries, Elevation, Built-Up Neighbours and Existing Built-Up (Upper right of Fig. 2). For each of these agents, proximity maps were derived and were reclassified into 5 classes (Table 3). Industries were categorized into four based on their Pollution Index Score (PIS) (>=60: Red, 41–59: Orange, 21–40: Green, and <=: 20 White Industries) (Ministry of Environment and Forests, 2016). Agents of 2013 and 2016 were used to predict urban growth of 2016 and 2020 respectively. Suitability map of urbanization was obtained using overlay analysis of all of these agents assigned with weights by adopting AHP (Analytical Hierarchical Process) Technique.

*3.3.2.1. Preparation of suitability map.* AHP was developed in the 1970s (Saaty, 1977) and has been implemented to solve various multi-criteria decision-making problems (Feizizadeh et al., 2012). In the current study, 12 agents responsible for urbanization were used (Table 3) and a Pairwise Comparison Matrix (PCM) was built

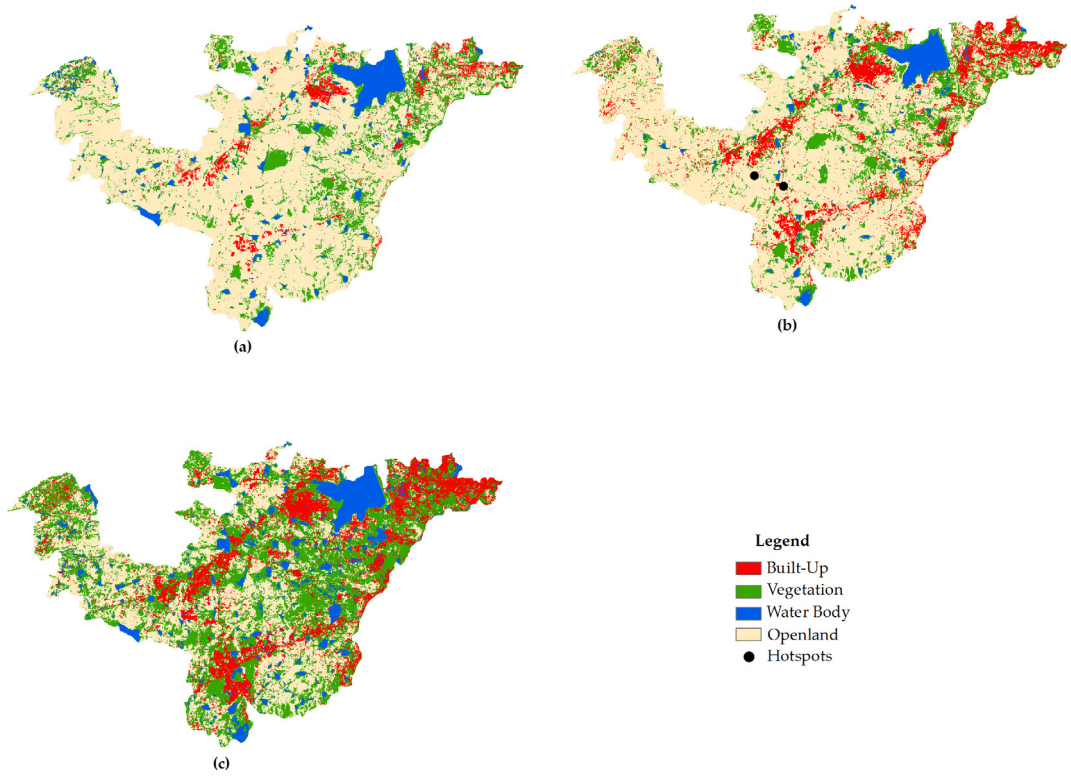


**Fig. 3.** Land Cover Maps of Sriperumbudur Taluk during the study periods. (a) In May 2009; (b) In May 2013; (c) In April 2016. (Increased vegetation observed in 2016 was mainly due to the occurrence of flood in December 2015).

among $\frac{n(n-1)}{2}$ number of pairs of agents, i.e., the PCM contained comparison between 66 number of agent pairs, where n is number of agents of urbanization.

$$\mathrm{PCM}[V] = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1j} \\ V_{21} & V_{22} & & V_{2j} \\ \vdots & & \ddots & \vdots \\ V_{j1} & V_{j2} & \cdots & V_{jj} \end{bmatrix} \quad (4)$$

where,V: Importance value of each agent pair as given by Saaty (2008); $V_{ij}$ = 1 where i = j.

The relative weights of each of the agents i.e., the priority vectors are calculated through the Eigen Vectors ($\lambda$: $\lambda_1$, $\lambda_2 \ldots \lambda_n$) of PCM. The permissible weights of the agents are assessed using Consistency Ratio (CR).

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \quad (5)$$

where, CI: Consistency Index = $\frac{\lambda_{max}-n}{n-1}$; RI: Random Index provided by Saaty (2008); $\lambda_{max}$: Maximum Eigenvalue of PCM; n: Number of agents used for comparison

$$\text{Based on CR, Priority Vector of PCM} = \lambda \begin{cases} \text{Accepted}, CR \leqslant 10\% \\ \text{Modified based on[V]}, CR > 10\% \end{cases} \quad (6)$$

The weights obtained using AHP technique for each of these 12 agents and for their corresponding 5 classes are reported in Table 3. From these weight maps (Fig. 4), Suitability Maps for urbanization with and without hotspot locations were obtained (Fig. 5).

Thus in ACA model, both urban maps of 2009 and 2013 and the suitability maps of the agents obtained based on AHP technique were used for the urban prediction of 2016. The TR applied here was '*to find all the non-built-up pixels with >=3 built-up neighbors AND >= 0.15 suitability index*'. The model was run in ArcGIS environment until the expected number of pixels of built-up and non-built-up is achieved in 2016. The predicted urban maps of 2016 with and without hotspots were validated against the observed urban map of 2016 and the accuracies were reported.

### 3.3.3. NNACA model

NNACA model makes use of the Multi-Layer Perceptron (MLP) neural network of LCM (Land Change Modeler) (https://clarklabs.

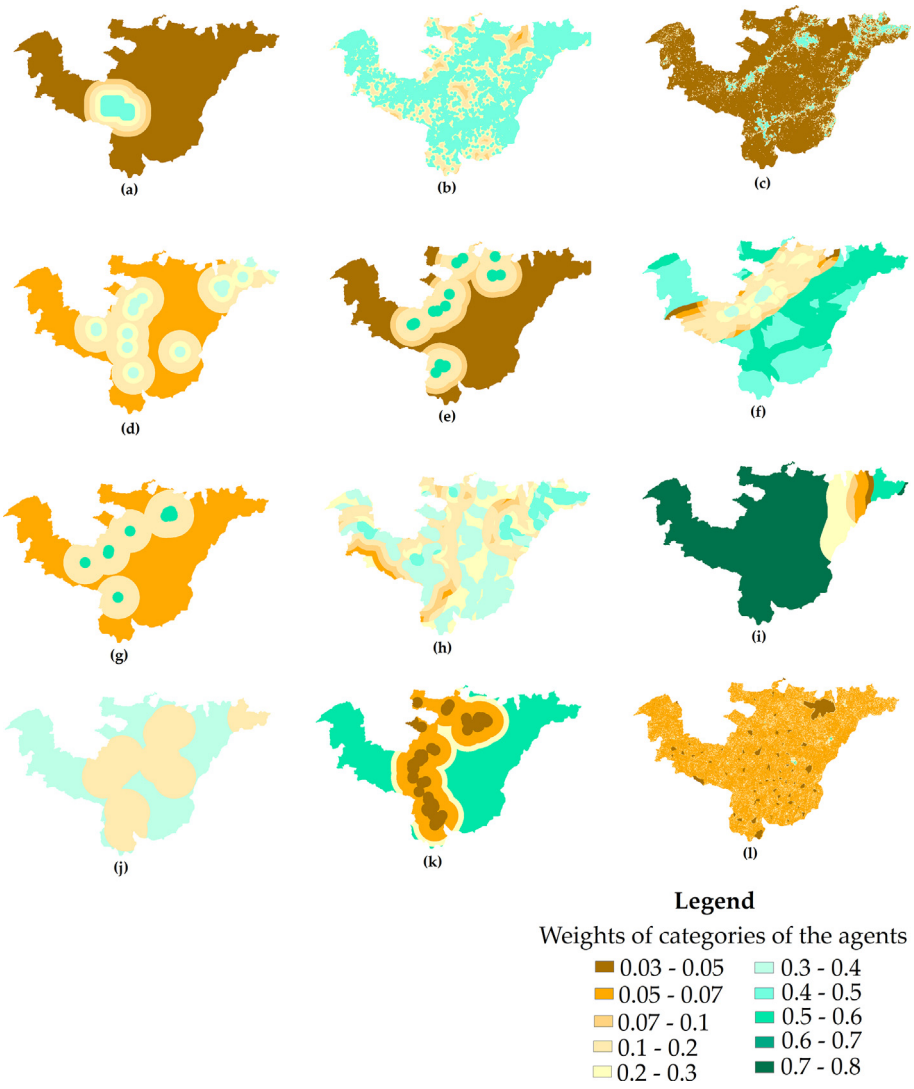


**Fig. 4.** Agents of urbanization – Weights assigned using AHP technique (Weights given in the legend correspond to different categories of the agents summed up to 1 while the weights specified within the brackets against every agent is the weight of that particular agent which sums up to 1 to prepare the suitability map). (a) Hotspot Locations (0.243); (b) Existing Built-Up (0.186); (c) Existing Built-Up Neighbours (0.142); (d) Road Junctions (0.11); (e) White Industries (0.085); (f) High Preference Roads (0.065); (g) Green Industries (0.05); (h) Medium Preference Roads (0.038); (i) Least Preference Roads (0.029); (j) Orange Industries (0.023); (k) Red Industries (0.017); (l) Slope (0.013).

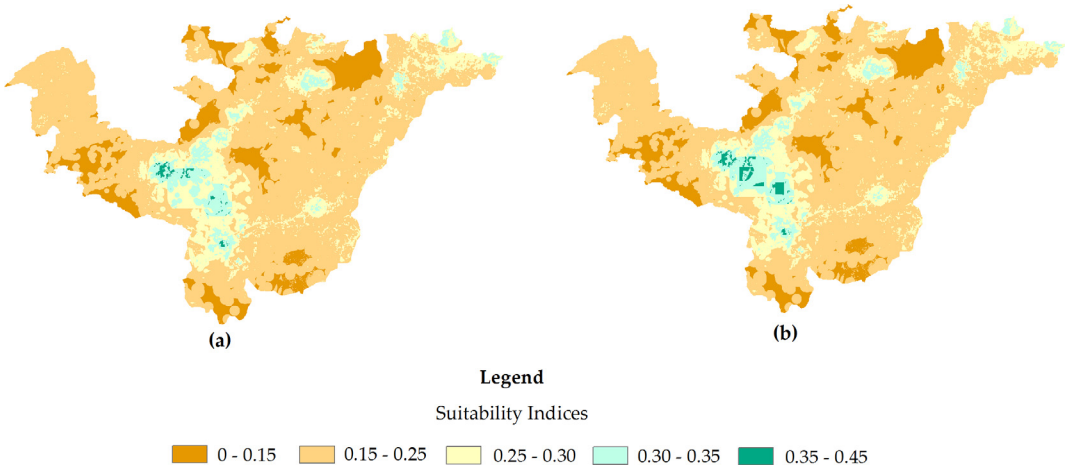


**Fig. 5.** Suitability Maps of urbanization of the study region. (a) Without Hotspot Locations; (b) With Hotspot Locations.

org/terrset/land-change-modeler/), TerrSet for the prediction of urban growth in 2016 with the urban maps of 2009 and 2013 and the proximity maps of the agents (Table 3) as inputs (Lower Right box of Fig. 2). A typical MLP consists of an input layer, one or more hidden layers and an output layer. In the current study, all the 12 agents were fed into the MLP as input neurons. 2n/3 numbers of hidden neurons are appropriate for modeling complex scenarios (Yeh and Li, 2002), where n = 12 is the number of agents of urbanization. Hence, 8 hidden neurons were used and the output layer had 2 neurons namely built-up and non-built-up. The model was run for 10,000 iterations with dynamic learning rate. An accuracy rate of 80% is considered to be acceptable and it can

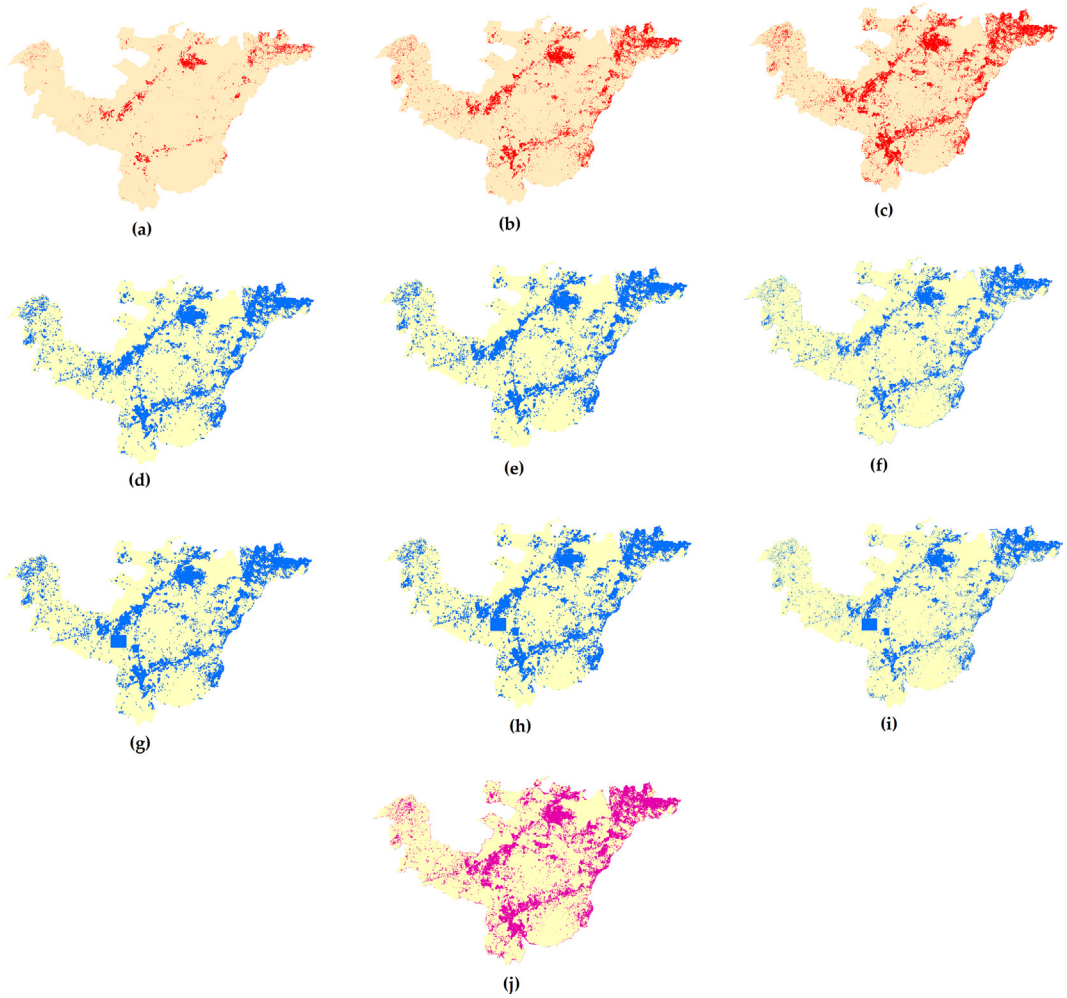


**Fig. 6.** Predicted Urban in 2016 and 2020 using TCA, ACA and NNACA Models. (a)–(c) Observed Urban Cover in 2009, 2013 and 2016 respectively; (d)–(f) Predicted Urban Cover of 2016 using TCA, ACA and NNACA models without hotspots; (g)–(i) Predicted Urban Cover of 2016 using TCA, ACA and NNACA models with hotspots; (j) Predicted Urban Cover of 2020 using NNACA model.

be understood that the MLP has learned sufficiently well from the input dataset (Eastman et al., 2012). In the present study, an accuracy rate of 90% is reported.

### 3.4. Validation

Validation is an important process which enables the users to understand the accuracy of the prediction model. The most widely adopted accuracy assessment technique is 'Error Matrix' through which Overall Accuracy (OA) and Kappa coefficient (k) are derived (Ahmed et al., 2013).

### 3.5. Assessing entropy of urban growth

To quantify the direction and distribution of urban sprawl, normalized Shannon's entropy was calculated (Mosammam et al., 2017). For this purpose, urban maps of the years 2009, 2013, 2016 and projected urban map of 2020 were used. The study area was divided into eight directional zones with Sriperumbudur Taluk headquarters as the center and based on the urbanization in each of these zones normalized entropy ($H_n$) was calculated and entropy is given by

$$H_n = \frac{1}{n}\sum_{i=1}^{n} p(x_i)\log\left\{\frac{1}{p(x_i)}\right\} \quad (7)$$

where, n: Number of zones (n = 8); p(x): Probability of built-up in each zone; $H_n$ values range from 0 to log (n). Values closer to 0 indicate a compact distribution of urbanization and values closer to log (n) indicates that the urban distribution is dispersed.

## 4. Results and discussions

### 4.1. Land cover map

Accuracies of the classified land cover maps are reported in Table 1. The land cover maps of 2009, 2013 and 2016 (Fig. 3) revealed a significant increase in built-up from 2009 to 2016. Out of the total study area of 645 $km^2$, built-up areas of 2009, 2013 and 2016 occupied 28.99 $km^2$, 74.48 $km^2$, and 113.16 $km^2$ respectively. Hotspot locations were included into the urban map of 2013 as rectangular and square shapes of the proposed industrial area for the prediction of urban cover in 2016.

### 4.2. Agents for urbanization

Based on AHP technique, weights were assigned to the categories of agents and to the agents themselves (Fig. 4).

A maximum weight of 0.24 was assigned to hotspots as urban growth was expected in these locations in 2016 and least weight of 0.01 was assigned to the slope as the study region experiences a gentle slope. The summation of weights of all the agents equals to one. Suitability map of urbanization (Fig. 5) had suitability indices between 0.07 and 0.42. The northeastern part of the study region which lies in close proximity to Chennai district, including Kundrattur, Moulivakkam, Mangadu had the maximum urban suitability index, indicating higher urban growth in these regions.

### 4.3. Simulated urban growth

#### 4.3.1. TCA and ACA models

Urban growth prediction for 2016 for without and with hotspot regions (Fig. 6) using TCA and ACA models are found to be overpredicted especially with hotspots when compared with the observed built-up area of 113.16 $km^2$ in 2016. The predicted built-up regions are accumulated around the existing built-up area since the prediction is based on TPM (Table 2) and the neighborhood configuration. Weights assigned to the categories of the agents are given in Table 3. Validation of the models against observed urban of 2016 using TCA and ACA are given in Fig. 7. It is to be noted here, even after the inclusion of agents into the model, the prediction outputs did not improve its efficiency and it remained the same.

**Table 2**
Observed and simulated Transition Probability Matrix between 2009 and 2013, 2013 and 2016 and its corresponding areas.

| | Observed transition probability between 2009 & 2013 | | Simulated transition probability between 2013 & 2016 | |
|---|---|---|---|---|
| | Built-Up | Non Built-Up | Built-Up | Non Built-Up |
| Built-Up | 1 (28.98 $km^2$) | 0 | 1 (74.78 $km^2$) | 0 |
| Non Built-Up | 0.0738 (45.50 $km^2$) | 0.9262 (570.93 $km^2$) | 0.1242 (70.91 $km^2$) | 0.8758 (499.72 $km^2$) |

**Table 3**
Weights generated through AHP technique for urbanization agents in the study region.

| Factors of Urbanization | Weights of Criteria | Alternatives and its weights | | | | |
|---|---|---|---|---|---|---|
| Hotspots | 0.2430 | 0–1 km | 1–2 km | 2–3 km | 3–4 km | >4 km |
| | | 0.4630 | 0.2560 | 0.1490 | 0.0860 | 0.0460 |
| Existing Built-Up | 0.1860 | 0–0.25 km | 0.25–0.5 km | 0.5–1 km | 1–2 km | 2–3 km |
| | | 0.4481 | 0.2607 | 0.1517 | 0.0882 | 0.0513 |
| No. of Existing Built-Up Neighbours | 0.1420 | 0 | 1–2 | 3–4 | 5–6 | 7–8 |
| | | 0.0392 | 0.0824 | 0.1443 | 0.2633 | 0.4708 |
| Road Junctions | 0.1100 | 0–1 km | 1–2 km | 2–3 km | 3–4 km | >4 km |
| | | 0.3500 | 0.2195 | 0.1830 | 0.1818 | 0.0657 |
| White Category Industries | 0.0850 | 0.5615 | 0.1802 | 0.1454 | 0.0748 | 0.0381 |
| High Preference Roads (Trunk, Primary & Residential Roads) | 0.0650 | 0.5713 | 0.1441 | 0.1097 | 0.0645 | 0.1104 |
| Green Category Industries | 0.0500 | 0.5683 | 0.1420 | 0.1291 | 0.1818 | 0.0657 |
| Medium Preference Roads (Secondary & Tertiary Roads) | 0.0380 | 0.4675 | 0.2054 | 0.0898 | 0.0578 | 0.1795 |
| Least Preference Roads (Outer ring road & Expressway) | 0.0290 | 0.0804 | 0.0603 | 0.0653 | 0.0503 | 0.7437 |
| Orange Category Industries | 0.0230 | 0.1371 | 0.1560 | 0.1817 | 0.1841 | 0.3411 |
| Red Category Industries | 0.017 | 0.0466 | 0.0628 | 0.0667 | 0.2761 | 0.5478 |
| Slope | 0.0130 | 1°–3° | 3°–10° | 10°–20° | 20°–32° | 0° |
| | | 0.5160 | 0.2939 | 0.1033 | 0.0554 | 0.0314 |

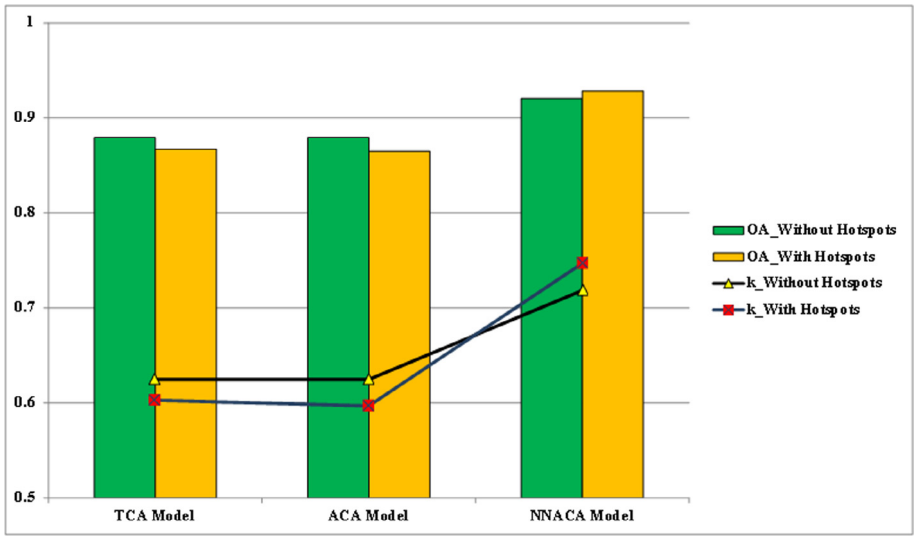


**Fig. 7.** Prediction accuracies of TCA, ACA and NNACA Models.

*4.3.2. NNACA model*

Urban growth for 2016 without and with hotspots (Fig. 6) predicted 106.12 km$^2$ and 107.13 km$^2$ of built-up areas with OA of 92% and 93% and a k value of 0.72 and 0.75 respectively. This implies that the efficiency of NNACA model is higher than those of TCA and ACA models.

The model during its learning process learned from the past data and predicted the urban growth for 2016 thus reducing the overprediction as in the case of TCA and ACA models. TR identification which was considered difficult, highly subjective and a source of uncertainty (Yeh and Li, 2006) in the earlier two methods were handled by the neural network itself through its learning

**Table 4**
Validation outputs based on different types of CA model.

| | | Simulated Area (km$^2$) | | Validation Outputs (km$^2$) | | |
|---|---|---|---|---|---|---|
| | | Built-Up | Non-Built-Up | Hits | Misses | False Alarms |
| Without Hotspots | TCA Model | 145.39 | 500.02 | 567.54 | 22.82 | 55.05 |
| | ACA Model | 145.39 | 500.02 | 567.50 | 22.84 | 55.06 |
| | NNACA Model | 106.12 | 539.29 | 594.18 | 29.13 | 22.10 |
| With Hotspots | TCA Model | 158.48 | 486.93 | 559.62 | 20.23 | 65.56 |
| | ACA Model | 158.48 | 486.93 | 558.35 | 20.87 | 66.20 |
| | NNACA Model | 107.13 | 538.28 | **599.23** | 26.10 | **20.08** |

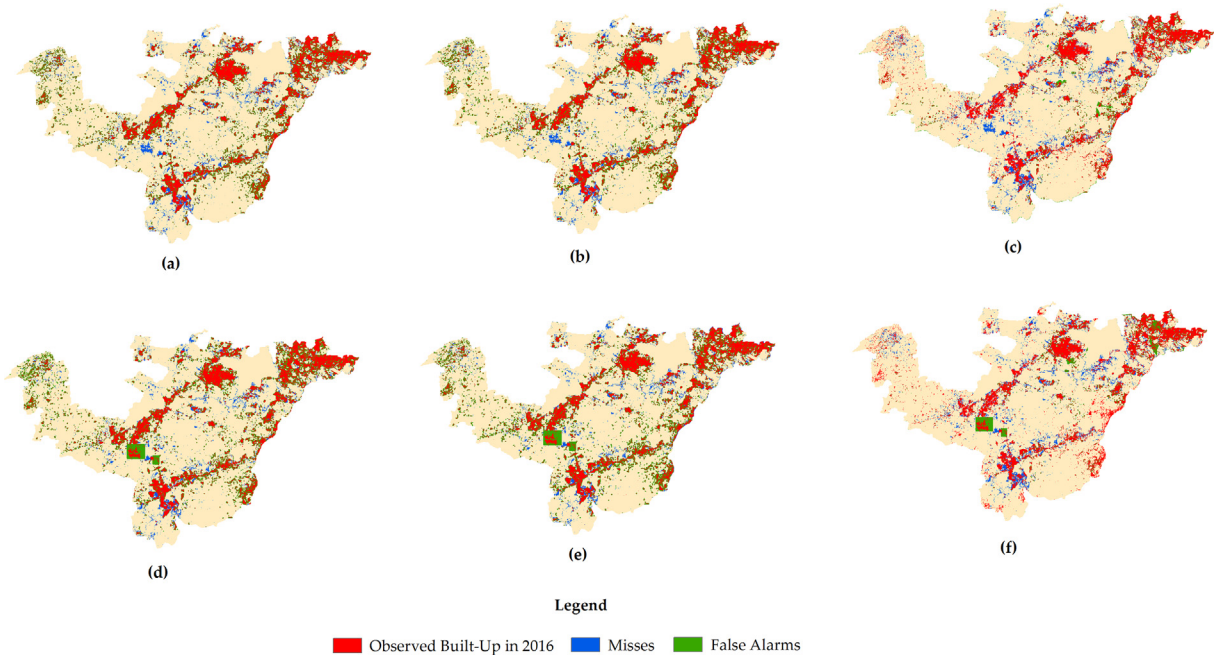


**Fig. 8.** Validation of predicted urban with observed urban in 2016 (a)–(c) Using TCA, ACA and NNACA models without hotspots; (d)–(f) Using TCA, ACA and NNACA models with hotspots.

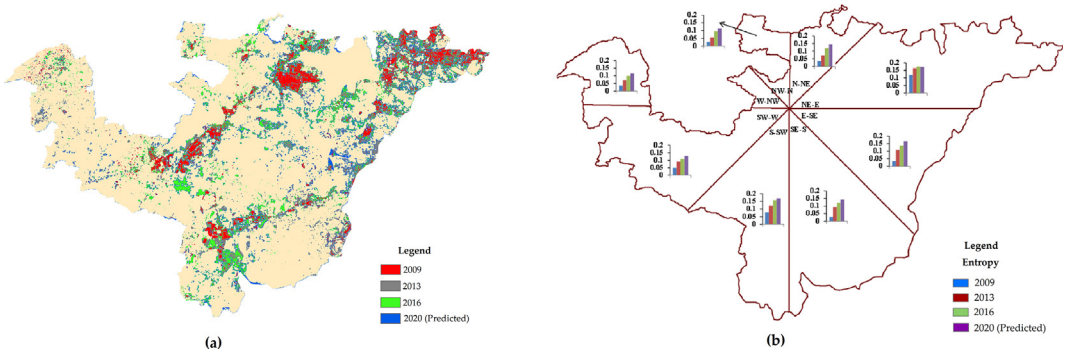


**Fig. 9.** Urban Sprawl and Entropy values of the study region (a) Urban Sprawl of observed 2009, 2013 and 2016 and predicted 2020 of the study region; (b) Entropy values of urbanization observed in 2009, 2013, 2016 and predicted in 2020.

algorithm. Thus, it could be inferred that along with including critical agents of urbanization in the CA model, when the neural network is also coupled, it would make a sophisticated tool in urban prediction modeling. With this expected accuracy levels (Fig. 7). NNACA predicts a higher built-up cover of 157.46 $km^2$ in the next four years by 2020.

Though all the three model outputs look almost similar the number of hits is higher using NNACA model (Table 4 and Fig. 8).

Hits quantify the correctly predicted pixels. Misses represent the underprediction and false alarms quantify the overprediction of the model. In the present study, false alarms are seen in the models, especially around Vallam and Vadagal locations, due to the shape assumed for the hotspots as only the extent of urban development in 2016 was proposed in the policy but not the actual shape. Hence, false alarms occurred higher in models with hotspot locations, whereas misses are found higher in models without hotspots. For an efficient model, apart from higher OA and k values, it is important to produce less number of false alarms. Table 4 shows that the overprediction is higher in the case of TCA and ACA models. Also, the maximum area of hits is found associated with NNACA model (Chellasamy et al., 2015).

### *4.4. Shannon's index for urban sprawl distribution type*

The urban sprawl kept increasing from 2009 to 2020 (Fig. 9(a)) and the entropy values of observed urbanization of 2009, 2013, 2016 and predicted 2020 were found to be 0.4107, 0.7812, 1.0172, and 1.1556 respectively. The maximum entropy value with eight zones is 2.0794 ($\log_e 8$). It can be seen that the entropy values in all eight zones kept increasing from 2009 to 2020 and was getting closer to the maximum entropy value.

This clearly indicates that the urban development is dispersive in nature. This dispersive nature of the urbanization could be the reason for TCA and ACA models not able to provide a realistic prediction outputs. During the years of 2009, 2013, 2016 and 2020, NE-E direction of the study region had the maximum entropy values followed by E-SE and then by N-NE (Fig. 9(b)). All these three directions lie in close proximity to Chennai district and also had high suitability index justifying the higher migration of people to these regions.

## 5. Conclusion

Thus, in the present study, NNACA model remains a better modeling tool in predicting the urbanization rather than TCA and ACA models as it could handle the uncertainties in the TR determination and also the prediction outputs are realistic. Though NNACA based urban models remain a popular technique, some features of CA including neighborhood configuration, TR determination, are always subjective and involves experts' knowledge and can be refined through sensitivity analysis which will be taken into account while modeling Chennai Metropolitan Area (CMA) in the bigger picture along with uncertainty component of the modeling in our future studies.

## Conflict of interest

The authors Aarthi Aishwarya Devendran and Gnanappazham Lakshmanan have no conflicts of interest to disclose for the submission of this manuscript.

## References

Ahmed, B., Ahmed, R., Zhu, X., 2013. Evaluation of model validation techniques in land cover dynamics. ISPRS Int. J. Geo-Information 2, 577–597. https://doi.org/10.3390/ijgi2030577.

Basse, R.M., Omrani, H., Charif, O., Gerber, P., Bódis, K., 2014. Land use changes modelling using advanced methods: cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. Appl. Geogr. 53, 160–171. https://doi.org/10.1016/j.apgeog.2014.06.016.

Chaudhuri, G., Clarke, K.C., 2013. The SLEUTH land use change model: a review. Int. J. Environ. Resour. Res. 1, 88–104.

Chellasamy, M., Chinnasamy, U., Ramaswamy, S.K., 2015. A neural-evidence pooling approach to predict urban sprawl using multi-temporal remote sensing data. Int. J. GEOMATICS Geosci. 5, 459–473.

Chung, K.L., 1960. Transition probabilities. In: Grammel, R., Hirzebruch, F., Hopf, E., Hopf, E., Maak, H., Magnus, W., Schmidt, F.K., Stein, K., van der Waerden, B.L. (Eds.), Markov Chains with Stationary Transition Probabilities. Springer, Berlin, Heidelberg, pp. 5–11. doi:10.1007/978-3-642-49686-8_2.

Divigalpitiya, P., Ohgai, A., Tani, T., Watanabe, K., Gohnai, Y., 2007. Modeling Land Conversion in the Colombo Metropolitan Area Using Cellular Automata. J. Asian Archit. Build. Eng. 6, 291–298. https://doi.org/10.3130/jaabe.6.291.

Eastman, J.R., 2012. IDRISI Selva Tutorial Clark Labs, Clark University, Worcester MA USA, pp. 270. Technical Manual. Retrieved from: <http://uhulag.mendelu.cz/files/pagesdata/eng/gis/idrisi_selva_tutorial.pdf> (accessed on 5.1.17).

Feizizadeh, B., Blaschke, T., Nazmfar, H., 2012. GIS-based ordered weighted averaging and Dempster-Shafer methods for landslide susceptibility mapping in the Urmia Lake Basin. Iran. Int. J. Digit. Earth 7, 688–708. https://doi.org/10.1080/17538947.2012.749950.

Government of Tamil Nadu, Ministry for Industries Policy Note 2014–2015, 2014, pp. 15 Technical Report. Retrieved from: <http://cms.tn.gov.in/sites/default/files/documents/industries_pn_e_2014> (accessed on 5.1.17).

Hill, A., Lindner, C., 2010. Modelling informal urban growth under rapid urbanisation - A CA-based land-use simulation model for the city of Dar es Salaam, Tanzania (Ph.D. Thesis) Technical University of Dortmund, Germany. p. 18. doi:http://hdl.handle.net/2003/27283.

Hua, L., Tang, L., Cui, S., Yin, K., 2014. Simulating urban growth using the SLEUTH model in a coastal peri-urban district in China. Sustain 6, 3899–3914. https://doi.org/10.3390/su6063899.

Huang, C., Davis, L.S., Townshend, J.R.G., 2002. An assessment of support vector machines for land cover classification. Int. J. Remote Sens. 23 (4), 725–749. https://doi.org/10.1080/01431160110040323.

Jokar Arsanjani, J., Helbich, M., de Noronha Vaz, E., 2013. Spatiotemporal simulation of urban growth patterns using agent-based modeling: the case of Tehran. Cities 32, 33–42. https://doi.org/10.1016/j.cities.2013.01.005.

Liu, Y., 2009. Modelling urban development with geographical information systems and cellular automata. CRC Press, Florida, USA.

Ministry of Environment and Forests, 2016. Environment Ministry releases new categorisation of industries. Ministry of Environment and Forests, Press Information Bureau, Government of India. Retrieved from: <http://pib.nic.in/newsite/PrintRelease.aspx?relid=137373 > (accessed 5.1.17).

Mosammam, H. M., Nia, J. T., Khani, H., Teymouri, A., Kazemi, M., 2017. Monitoring land use change and measuring urban sprawl based on its spatial forms. The case of Qom city. Egypt. J. Remote Sens. Space Sci. 20(1):103–116.

Mubea, K., Goetzke, R., Menz, G., 2014. Applying cellular automata for simulating and assessing urban growth scenario based in Nairobi, Kenya. Int. J. Adv. Comput. Sci. Appl. 5, 1–13. https://doi.org/10.14569/IJACSA.2014.050201.

Mubea, K.W., 2014. Scenarios of Urban Growth in Kenya Using Regionalised Cellular Automata based on Multi temporal Landsat Satellite Data (Ph.D. Thesis) University of Bonn, Germany. p. 19.

Saaty, T.L., 2008. Decision making with the analytic hierarchy process. Int. J. Serv. Sci. 1 (1), 83–98.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. J. Math. Psychol. 15, 234–281. https://doi.org/10.1016/0022-2496(77)90033-5.

Shafizadeh Moghadam, H., Helbich, M., 2013. Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. Appl. Geogr. 40, 140–149. https://doi.org/10.1016/j.apgeog.2013.01.009.

The Growth Story of Sriperumbudur Real Estate, 2013. Retrieved from: <http://www.etastar.com/The-Growth-Story-Sriperumbudur-Real-Estate> (accessed on 5.1.17).

Triantakonstantis, D., Mountrakis, G., 2012. Urban growth prediction: a review of computational models and human perceptions. J. Geogr. Inf. Syst. 4, 555–587. https://doi.org/10.4236/jgis.2012.46060.

Wolfram, S., 1984. Cellular automata as models of complexity. Nature 311, 419–424. https://doi.org/10.1038/311419a0.

Yang, X., Zhao, Y., Chen, R., Zheng, X., 2016. Simulating land use change by integrating landscape metrics into ANN-CA in a new way. Front. Earth Sci. 10, 245–252. https://doi.org/10.1007/s11707-015-0522-7.

Yeh, A.G.-O., Li, X., 2002. Urban Simulation Using Neural Networks and Cellular Automata for Land Use Planning. In: Advances in Spatial Data Handling. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 451–464. doi:10.1007/978-3-642-56094-1_33.

Yeh, A.G.-O., Li, X., 2006. Errors and uncertainties in urban cellular automata. Comput. Environ. Urban Syst. 30, 10–28. https://doi.org/10.1016/j.compenvurbsys.2004.05.007.