land

Article

Integrating ANNs and Cellular Automata–Markov Chain to Simulate Urban Expansion with Annual Land Use Data

Tingting Xu [1,*], Dingjie Zhou [2] and Yuhua Li [3]

[1] School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2] Yunnan Institute of Surveying and Mapping Engineering, Kunming 650033, China; djzhou@whu.edu.cn
[3] The Institute of Chongqing Planning and Nature Resource Surveying and Monitoring, Chongqing 401121, China; liyuhua202206@163.com
* Correspondence: xutt@cqupt.edu.cn; Tel.: +86-13206167597

Abstract: Accurately simulating and predicting the urban expansion process, especially in expeditious urbanization areas, is an important aspect of managing limited land resources and adjusting flawed land use policies. This research was conducted on the basis of a high-temporal-resolution land use dataset to precisely model urban expansion in a rapidly developing zone by integrating the Artificial Neural Network (ANN), cellular automata (CA), and Markov Chain (MC). An urban suitability index (USI) map was created using ANN and fed to CA–MC to identify possible changed-to-urban cells. Two ANNs, multiple-layer perceptron (MLP) and long short-term memory network (LSTM), were implemented as simulation models for comparison. Due to its ability to capture more temporal information, LSTM outperformed MLP in modeling urban expansion dynamics over a short temporal interval. The simulated results were validated by (fuzzy) kappa simulation and the results revealed that the combination of ANN and CA–MC can precisely model the urban development locations due to its strength in revealing the nonlinear relationship between the expansion process and its drivers. The same model was applied to southern Auckland, and the compared results show that the most simulated variance is caused by the land use policies implemented by different types of governments.

Keywords: urban expansion; artificial neural network; multiple-layer perceptron; long short-term memory network; cellular automata

Citation: Xu, T.; Zhou, D.; Li, Y. Integrating ANNs and Cellular Automata–Markov Chain to Simulate Urban Expansion with Annual Land Use Data. *Land* **2022**, *11*, 1074. https://doi.org/10.3390/land11071074

Academic Editors: Sheng Zheng, Yuzhe Wu and Ramesh P. Singh

Received: 8 June 2022
Accepted: 12 July 2022
Published: 14 July 2022

1. Introduction

Since the economic reform in the late 1970s, China has been experiencing rapid economic development and urbanization, driven partially by a dramatic increase in the urban population [1–5]. However, the fast but unsustainable urban development in the past few decades has caused many pressing issues that have become obvious in recent years. The massive and chaotic urban expansion has led to the inefficient and irrational use of limited land resources, polluted the environment, damaged or destroyed the sensitive ecosystem, and degraded agricultural and forest land [6–9]. Naturally, much research has been conducted to monitor and model such changes [10–14].

Current urban expansion modeling is limited by a coarse temporal resolution owing to a lack of suitable data. Many empirical studies selected a temporal increment of five-year intervals or longer for the urban expansion model, due to the normal life cycle to form a new urban area from an existing one [8,9,12,15,16]. However, this temporal resolution does not allow urban expansion in rapidly developing cities to be detected in a timely manner, especially in some industrial districts [13,17]. Thus, it is impossible to assess the unanticipated effects of land use policies abruptly promulgated by central and local governments [3,18,19] or to remedy the problems arising from them. Therefore, a shorter temporal interval is highly desirable in modeling and analyzing urban land use changes to reveal more frequently updated details both spatially and non-spatially, and hence to better explain the dynamic process within a rapidly growing urban area [20]. Another impactful issue, besides land use policy, is the transportation factor, including both already-built transportation facilities and transport planning, which could also exert significant effects on the urban expansion process. Early in 2004, Wegener and Fuerst [21] explored the interactions between transport planning and land use changes by implementing several models. In 2013, Wegener provided a more detailed explanation of the integrated models of urban land use and transport and discussed how to improve the models to simulate the most appropriate level of spatial resolution and substantive disaggregation [22]. Based on this research, Dingli et al. [23] conducted a systematic review of the socio-technical factors that impact urban transportation performance, hence modeling the urban land use status. All these previous studies demonstrated the critical role of the transportation factor. However, transport planning is always performed for a long time gap, which may reduce its impact when the simulation time is relatively short. To offer compensation for this, this research uses several proxy factors as transportation factors to include this important aspect.

Another issue in urban expansion modeling is the limited predictive ability of individual models. For the past two decades, static-data-analysis-based urban expansion research has been performed via advanced simulation modeling processes due to progress in geospatial technology. Such models as Cellular Automata (CA) [10,24–26], the Multi-Agent Model [27,28], the Land Transformation Model [29], and SLEUTH [30] have added the spatial scope into the urban expansion simulation processes. In 2015, Dr. Roger White, Dr. Guy Engelen, and Dr. Inga Ulgee, with their excellent book "Modeling Cities and Regions as Complex Systems: From Theory to Planning Applications", especially pointed out that to simulate the "complex, adaptive, self-organizing" urban system precisely, the traditional land use change models must be carefully incorporated with the spatial dynamics of human–environment interactions. They also provided a framework to combine the raw types of ANNs with CA to model population dynamics and economic activities [31,32]. In 2022, Kim et al. [33] and Yi et al. [34] also pointed out that CA models could combine with machine-learning-based algorithms to enhance simulation accuracy and provide more reliable results. From their work, in line with many other studies in the urban expansion modeling field, all simulation models need to capture the complex nonlinear relationships between various driving factors and urban expansion parcels, via the "bottom-up" model approach and different transition rules. Among these models, Cellular Automata with Markov Chain (CA–MC) is found to be one of the most popular CA-based models since it can easily predict the urban expansion demand based on the Markov Chain (MC) and implement the analytic hierarchy process (AHP) and logistic regression (LR) processes into the land use change iteration process to build the urban suitability index map. It can also allocate the amount of urban land spatially based on CA by taking fewer variables into consideration than other models. However, the self-adaptive ability of CA–MC to model the nonlinear relationships between drivers and the urban dynamic process is still not completely reliable [16,26,35].

Reliability can be improved by replacing AHP and LR with an Artificial Neural Network (ANN). ANNs can more efficiently consider the complex nonlinear relationships between the dependent and independent variables than AHP and LR, since they have advantages in simulating urban expansion dynamics [36,37]. Because of their self-learning ability, the unknown relationships among variables can be addressed with the automatic approximation of nonlinear functions by the ANN, which is better than linear-based regression models [38]. Moreover, due to its self-organizing and adaptive approach, the integration of an ANN can make the CA–MC more self-adaptive than the traditional models [39,40]. Therefore, ANNs have been integrated into many other spatial simulation models to simulate urban expansion, despite the difficulties in properly parameterizing and optimally configuring an ANN model [8,9,41,42]. Xu et al. [43,44] incorporated an ANN with CA–MC to simulate the urban expansion process in both the U.S. and Auckland, New Zealand, and to predict future possible urban building areas. Results from these studies have demonstrated that such innovative integration can improve the accuracy of simulations. To provide more detailed information on the ANN-based CA–MC and improve simulation accuracy, the major objective of this paper is to integrate two ANNs into CA–MC for the modeling of spatiotemporal urban expansion dynamics in a rapidly developing area and to assess the model's capability.

The remainder of this paper is organized as follows: In Section 2, we describe the study area, explain the data used, and describe the methodologies that we have proposed to build the ANN–CA–MC model. After this, the results present the possible future urban expansion in both spatial and quantitative scopes. The model's results are further compared with a previous study to indicate the different expansion processes for developing and developed countries. Some limitations of this study are also discussed in this section. Finally, we conclude with the results and offer some future research directions. By doing this, we wish to reveal that a high temporal resolution of data could lead to a better result. Still, the model must have the ability to capture such characteristics. Furthermore, with the results, we seek to provide insightful information and a planning tool for the urban planners and decision makers that could be applied to compare the simulated results with land use plans to modify their work more effectively.

2. Data and Methods

2.1. Study Area

The Liangjiang New District of Chongqing, southwest China, was chosen as the study area because it has been experiencing rapid growth since 2009. It comprises three industrial zones (Longxing, Yufu, Shuitu), two duty-free ports (an airport and a harbor), plus several counties (Figure 1). Characterized by an area of 1172 km^2, it has been designated as the first hinterland development and a showcase zone for southwest China. Thus, this district plays a very critical role in regional economic development. It enjoys both local and national incentivized land use policies, as well as a high level of investment and a well-designed future development plan. All these political advantages have led this area toward unprecedented urban development, utilizing a significant amount of land resources. The projected population inside this area was 3.5 million within a 350 km^2 built-up urban area by 2020. However, the actual amount of the built-up areas has already surpassed the planned value, since the number had already reached 344 km^2 in 2014, indicating that the average annual growth rate for the urban land is at 7.5%. Thus, further development will have to rely on a more intensive use of the land. Therefore, it is critical to judiciously model, allocate, and plan new urban areas in the future.

2.2. Data

The National Annual Land Use Survey Database was used in this study, in which land use was mapped at different levels by the China Ministry of Land Resources. The database is in an annual temporal resolution and obtains data from a national–province–city multi-scaled land use transformation monitoring approach involving both remote sensing and field survey data. With this temporal resolution, the land use change can be periodically monitored so as to plan urban growth effectively. We simplified the data from the original 25 classes into 9 categories (farmland, orchard, forest, grassland, transport, water bodies, bare land, urban land, and other land uses).

In addition, for every grid cell in question during the modeling, its neighborhood effects in different years were also calculated because they would affect the cell status in future years. Other urban expansion drivers were all proxy factors, such as nearby existing urban areas, the distance to different types of roads, and utilities. They were also considered as independent variables for every cell during the simulation (Table 1).

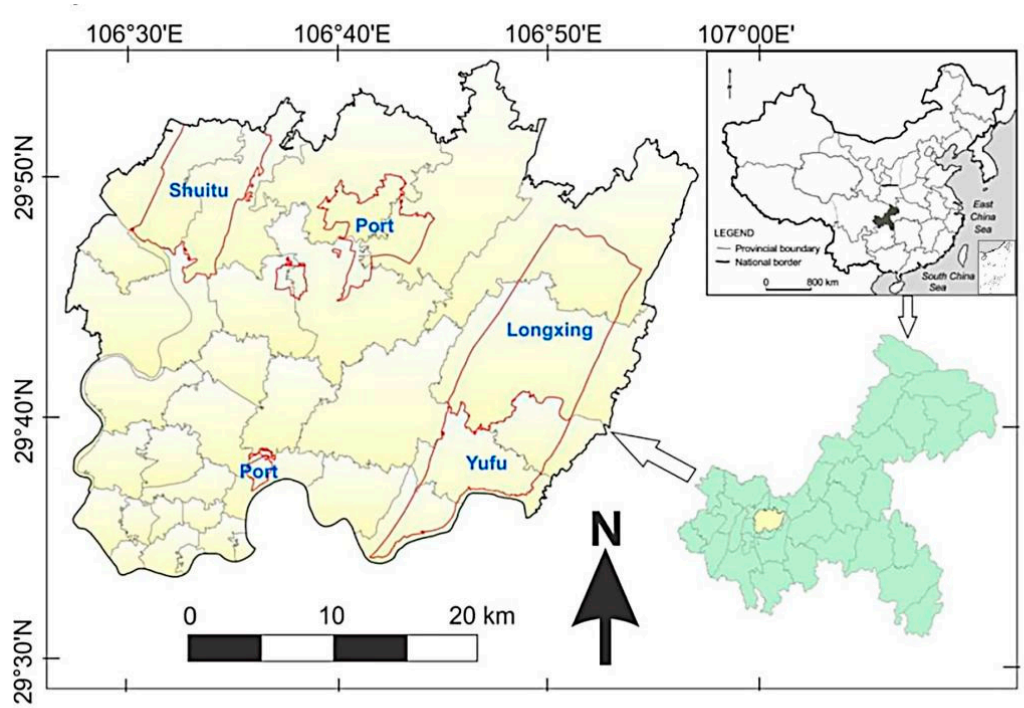


Figure 1. Study area: the Liangjiang New District, Chongqing, Southwest China. Red lines: boundaries of industrial zones and duty-free ports; grey lines: county boundaries.

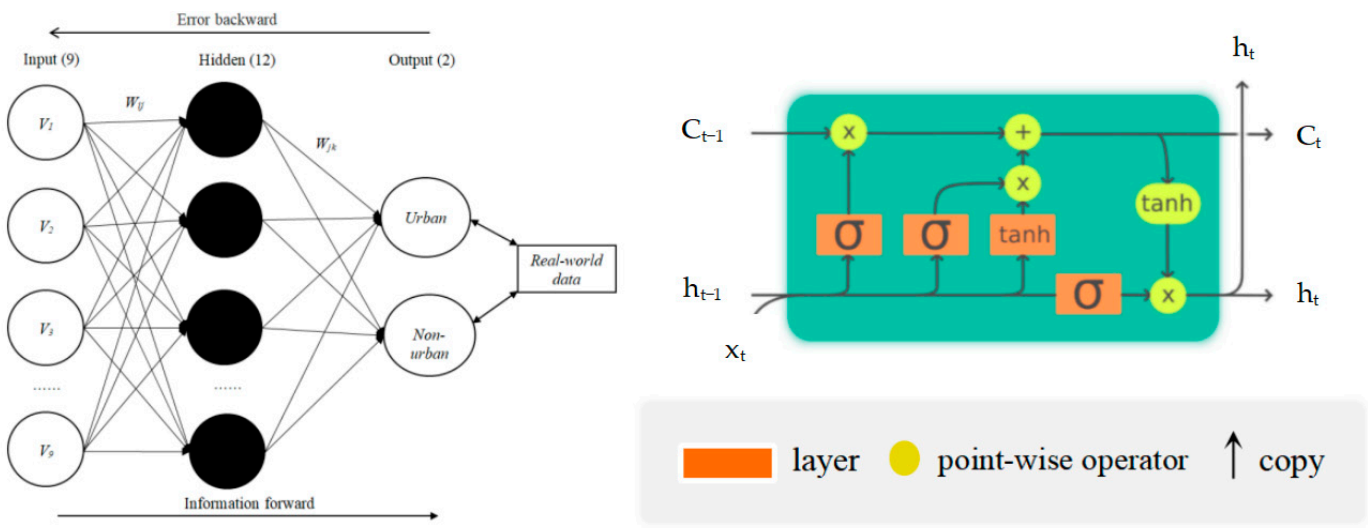
Table 1. Variables considered in the modeling.

Variable	Description
	Dynamic variables
Proxy: existing urban	Euclidean distance between existing urban cell and the target cell
Proxy: new road	Euclidean distance between the newly built roads (one year in advance) and the target cell
Land use	The proportion of all land use types within a Moore neighborhood (3 × 3 cells)
	Still variables
Digital Elevation Model (DEM)	Terrain factor, elevation value (height: meters)
Slope	Terrain factor, slope (in degrees)
Proxy: Motorway	Euclidean distance between the entrances or exits of motorway and the target cell
Proxy: Arterial Road	Euclidean distance between arterial roads and the target cell
Proxy: City major road	Euclidean distance between urban major roads and the target cell
Proxy: City medium road	Euclidean distance between urban medium roads and the target cell
Proxy: City minor road	Euclidean distance between urban minor roads and the target cell
Proxy: School	Euclidean distance between kindergarten, primary schools, and middle schools and the target cell
Proxy: Transports	Euclidean distance between bus stops, train stations, and ferry ports and the target cell
Proxy: Hospital	Euclidean distance between hospitals and the target cell
Proxy: Market	Euclidean distance between supermarkets and convenience stores and the target cell

2.3. Methods

2.3.1. Artificial Neural Networks

ANN is one of the most popular modeling techniques used today, with self-learning, self-organizing, and self-adapting abilities [40,41,45]. In this study, two ANNs were integrated with CA–MC: MLP and LSTM. The MLP is considered the most frequently used and is constructed as the information feed-forward and error Back-Propagation Three-Layer Perceptron (BP-TLP) network architecture (Figure 2a). Due to its simple structure, ease of training, reasonable memory consumption, and fast as well as relatively accurate prediction results, this neural network was adopted to simulate urban expansion [46]. On the other hand, the LSTM is explicitly designed to avoid the long-term dependency problem and has a strong focus on the temporal (sequential) scale; it can be combined with CA–MC to extract spatiotemporal information [47]. Instead of feed-forward, it is a recurrent neural network including a delayed input, also known as the feedback of the output, which creates a chain of repeating modules of the neural network (Figure 2b). For both ANNs, the most important aspect of their architecture is the number of hidden nodes in each hidden layer (also known as forgotten nodes and layers for LSTM), which significantly affects the ANN's performance [48]. Too few nodes will cause a significant prediction error, while too many will prolong the training process and lead to overfitting problems. Initially, the networks start training stepwise iteratively and will stop when the mean square error (MSE) meets the threshold value (0.01 in this study). The MSE is calculated from the difference between the network running output and the expected data value. However, insistence on meeting this supposed MSE threshold might lead to overfitting issues, which will decrease the model's accuracy. To avoid this problem, the two ANNs in this study applied an early stopping strategy by adjusting two training parameters: the number of training epochs and the maximum fail number. The time for a training epoch was set to 1000 and the maximum fail number was set to 50, which significantly reduced the possibility of overfitting in this study. In addition, the training batch size of LSTM also impacts the final result, so it is usually set as an integral time from the data.



(a) (b)

Figure 2. (**a**) The architecture of the BP-TLP ANN used in this study. V stands for the input variable, W_{ij} is the connecting weights between the input layer and hidden layer, and W_{jk} is the connecting weights between the hidden layer and output layer. (**b**) The architecture of LSTM. X is the vector (features) input; h is a hidden state vector, which is also known as the memory output vector of the LSTM unit; c is the cell state vector; σ and tanh are the activation functions. The green rectangle indicates the "forget" process for the LSTM.

2.3.2. ANNs' Integration with Markov Chain—Cellular Automata

The Markov Chain model determines the expected amount of non-urban land transitioned to urban use in future urban expansions. It calculates the land use change process through the transitions of different land use types at a temporal interval of one year. Two results are generated from MC: a transition area matrix, which shows the absolute amount of change from different non-urban land uses to urban land, and a transition possibility matrix, which reveals the likelihood of the transition.

CA uses the present cell status and the cell's neighborhoods to forecast the future cell status, as expressed in Equation (1):

$$S^{t+1} = f\ (S^{t}, N, TR) \tag{1}$$

in which S^{t+1} and S^{t} represent the cell state at time t and t + 1, respectively; N is the effect neighborhood for one cell; TR stands for the globe transition rule of the CA model; and f represents the cell state transition function. For regular MC-CA, R is the transition possibility matrix produced from MC, which shows only the numeric results from two land use datasets, without any information on the spatial distribution of the modeled changes. This problem is solved by using the *USI* map generated from *MLP* (Equation (2)) and *LSTM* (Equation (3)).

$$USI\_MLP = f_{MLP}\ (V_1, V_2, \ldots, V_i) \prod C \tag{2}$$

where f_{MLP} is the activation function of the ANN; V_i serves as the ith input variable, and C is a binary value of 0 or 1, indicating the constraint factors, such as water or ecological reserved zones.

$$USI\_LSTM_t = f_t \times USI\_LSTM_{t-1} + i_t \times S^{t} \tag{3}$$

where f_t is the forget function of *LSTM*; it is the scaling down function for input variables and cell state S^{t}, and it is usually sigmoid or tanh.

Hence, the final state (urbanized or not) of cell_{ij} in this ANN–CA–MC will be represented as Equation (4):

$$S^{t+1}{}_{ij} = f_{ca}\ (S^{t}{}_{ij}, A^{t+1}, USI_{ann}, N_{multiple}) \tag{4}$$

where A^{t+1} is the projected amount of the future urban expansion by MC, USI_{ann} is the *USI* generated from either the *MLP* or *LSTM* network, and $N_{multiple}$ represents different land use neighborhood effects at cell (i, j) (i-row, j-column). The CA transition rules are now the combination of *USI* and *N*.

2.4. Model Implementation

Figure 3 illustrates the detailed steps in the integration of these two machine-learning-based ANNs with CA–MC to model urban expansion. Annual land use data and spatial data such as DEM, slope, road network, and facility locations (Table 1) were collected to create a geodatabase in the GIS environment, from which the urban expansion areas and related driving factor values were extracted. Afterwards, both changed and unchanged cell samples were selected using the maximum dissimilarity distance algorithm (MDA) or random sampling. The selected samples were used to train MLP and LSTM and to create the USI map. In addition, the annual land use data were delivered to the MC to calculate the planning amount for future urban expansion. The ANN–CA–MC simulated the urban expansion at a certain time from the expected amount of expansion and the USI map output from either MLP or LSTM, taking into account the multiple neighborhood land use effect. The output model results were validated against the actual land use data using kappa simulation. Only when the modeled results passed the validation process was the properly configured and trained model then used to predict future urban expansion, which could provide useful information for the modification of land use policies and plans.

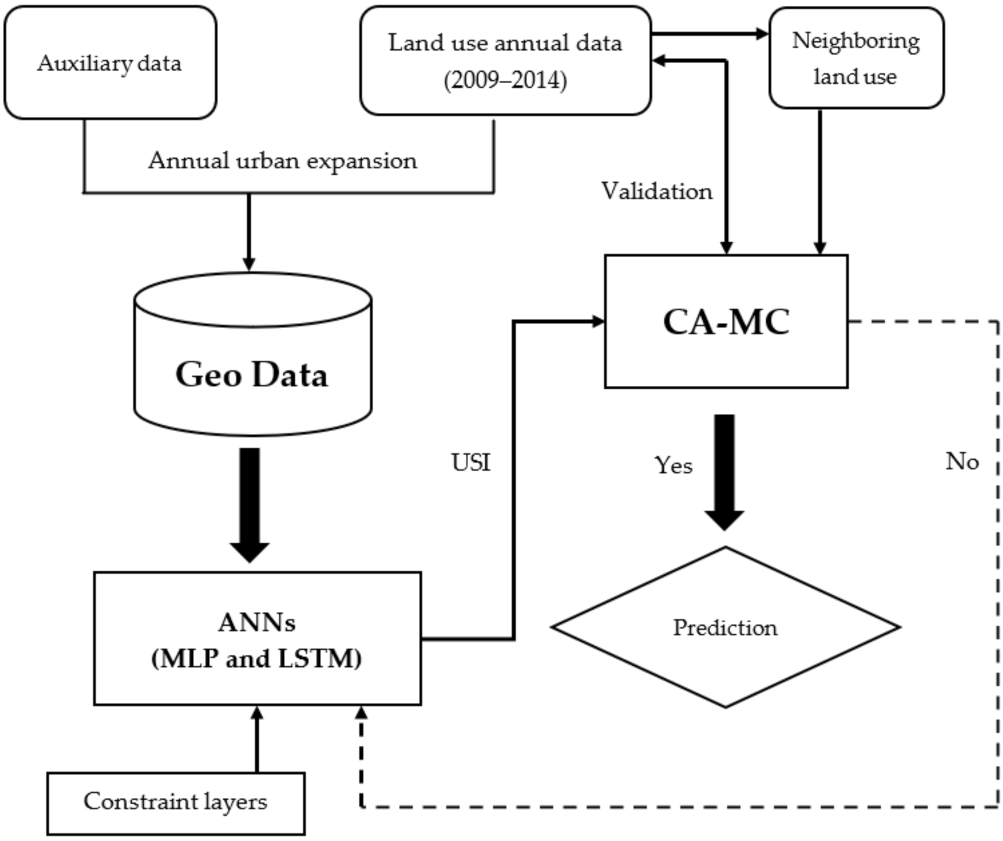


Figure 3. ANN–CA–MC framework for modeling urban expansion in the Liangjiang District of Chongqing, China.

3. Results and Discussion

3.1. Urban Suitability Index Map

Three USI maps (Figure 4) have been created based on two ANNs. Cell values in the USI maps ranged from 0 to 1, indicating the suitability of non-urban cells for the future urban development (new urban areas) process. This range was grouped into five classes using an equal interval to present a more obvious spatial distribution of these cells' future urbanization possibilities, namely very high (0.8–1, red), high (0.6–0.8), moderate (0.4–0.6), low (0.2–0.4), and very low (0.0–0.2, green) (Table 2, with both areas and percentage values). Based on the results from MLP with two different sampling methods, known as random selection and the most dissimilarity algorithm (MDA) (Figure 4A,B), we can see that the vast majority of non-urban areas in the Liangjiang New District are not a suitable choice for future urban development, as, generally, over 75% of the non-urban land cells have a suitability index value of less than 0.2. From Table 2, rows one and two, cells having a suitability > 0.6 are relatively small in quantity (approximately over 8%). They are located either close to or inside mature, built-up urban zones, indicating a possible compact development (e.g., infilling and edge development) in this study area called the "organic urban growth mode"; all of these areas were accurately modeled by the MLP-based ANN. The most suitable focal areas for future urban development are distributed in the government-delineated industrial development zones (red line in Figure 2), as well as next to the developed counties. The medium suitability value indicates possible future urban expansion. Both maps show that the Longxing and Shichuan counties (top right), having this value, are the next most promising areas to be urbanized. However, both the quantity (Table 2, row 3) and spatial pattern (Figure 4C) of the USI from LSTM are different to those of the MLP network. Over 95% of these non-urban cells are predicted to have a very low likelihood of urbanization in the future, while only 2.6% of them have a USI > 0.6, much lower than this value for the MLP. In addition, only a very small number of non-urban cells have a medium USI value (0.2–0.6), with less than 0.5% of this area predicted to be transitioned to urban uses based on the LSTM. The spatial patterns of high USI with the LSTM are still close to existing urban cells but are more dispersed, with a large amount of high USI cells only distributed in the three industrial zones, and very few outside their boundaries (Figure 4C). This zonal-based distribution shows that the LSTM has the ability to capture the land use planning information through the temporal learning process.

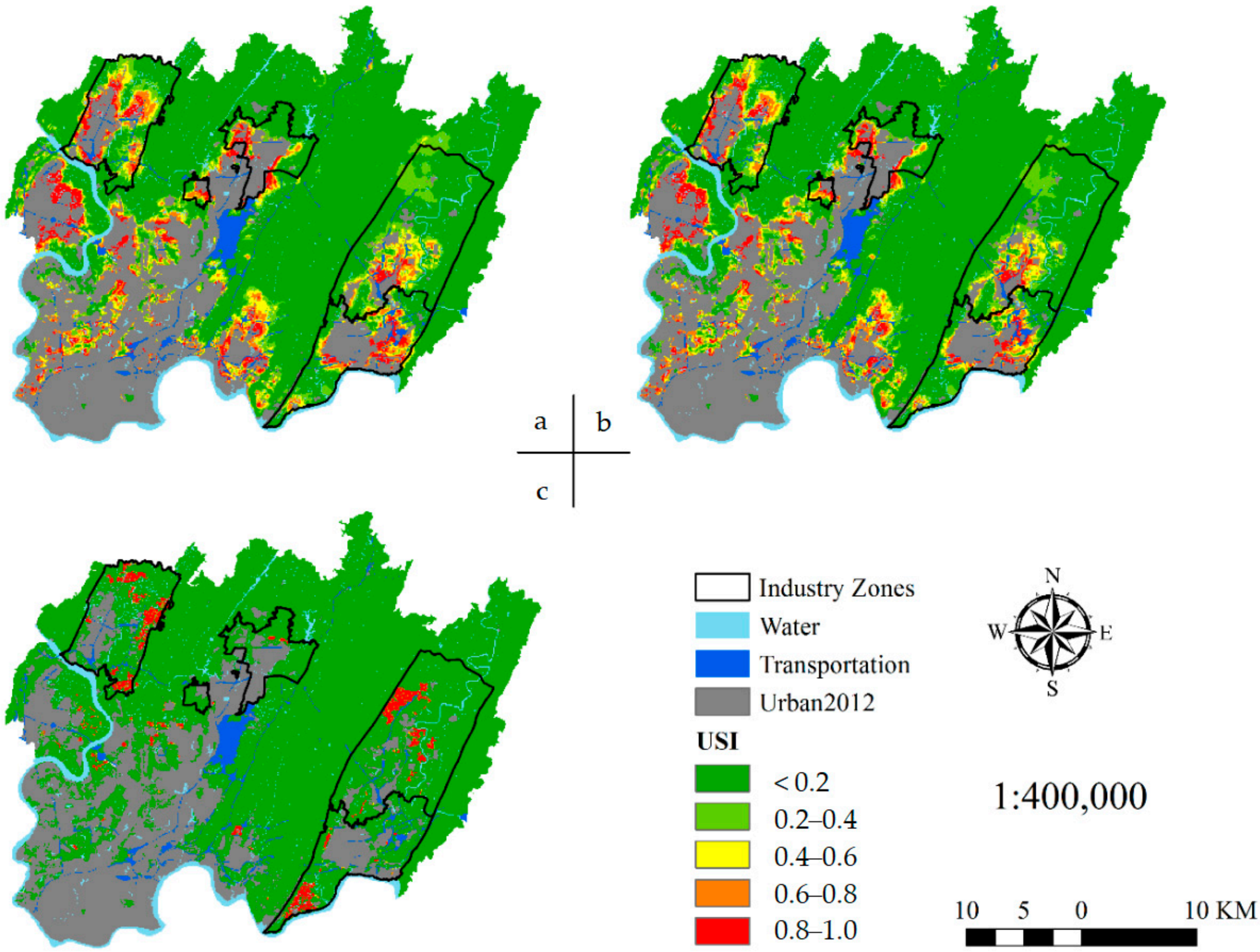


Figure 4. USI map of the study area from ANNs. (**a**) MLP with random sampling; (**b**) MLP with MDA sampling; and (**c**) LSTM.

Table 2. Proportions of areas for five groups of USI values predicted with three methods (unit: km^2/%).

Range	<0.2	0.2–0.4	0.4–0.6	0.6–0.8	≥0.8
MLP–Random	596.38/75.3%	79.38/10%	47.94/6.0%	38.11/4.8%	31.27/3.9%
MLP–MDA	612.23/77.2%	73.30/9.2%	44.98/5.7%	33.92/4.3%	28.65/3.6%
LSTM	769.00/96.9%	2.43/0.3%	0.87/0.1%	1.54/0.2%	19.27/2.4%

3.2. Model Validation

To validate the model's accuracy, the simulated urban expansion result for 2014 (Figure 5) was obtained by feeding the annual land cover changes from 2009 to 2013 into the integrated ANN–CA–MC model. After this, the model results were compared to the ground truth urban land data for 2014, and the validated results are shown in Table 3. MC analysis resulted in a projected future urban expansion value of 46.3 km^2 in 2014, compared to the actually expanded area of 48.78 km^2, leading to a maximum simulation accuracy of 95%. As we can see from the spatial urban expansion patterns in Figure 5, MLR-Random predicted fewer expansion areas than the other three models in the upper northern region of the study area. In addition, the infilling and edging expansions were more accurately simulated than the isolation expansion due to the model's capability. All the models showed a high degree of local agreement (kappa > 80%) between the model-simulated and the ground truth urban areas, which indicates that both MLP (regardless of the sampling method, either random or MDA) and LSTM could precisely simulate the urban area in 2014. Judged against (fuzzy) kappa simulation, which takes future possible land transitions into account, as well as modifying the agreement between two urban land use maps for the size of class changes and comparing the cell states' variations instead of the static status [49], LR has the lowest values, indicating that it has the weakest ability to model the nonlinear relationships among all the algorithms. LSTM has the highest validation values among the models, and thus the most accurately modeled results, because of its capability to learn within a short time interval, which is important in modeling fast-growing urban areas.

Figure 5. Simulation results of year 2014, Liangjiang New District.

Table 3. Validation results of (fuzzy) kappa simulation of four types of models.

Method	Kappa	Kappa Simulation	Fuzzy Kappa Simulation
LR	0.818	0.27	0.51
MLP–Random	0.829	0.36	0.64
MLP–MDA	0.827	0.32	0.53
LSTM	0.857	0.48	0.70

The significant differences between MLP and LSTM are due to differences in their treatment of temporal aspects. As Figure 6 illustrates, if the annual land use data (2009–2013 in this study) were used to simulate the most recent land usage (2014 in this study), the MLP would not consider the time differences of each factor but only use the data from the previous year as the entire training dataset to train the network, since it cannot successfully reveal the temporal changes. However, on the other hand, the LSTM will consider every annual change and changes between different time gaps separately and add this time-scaled effect into the final result.

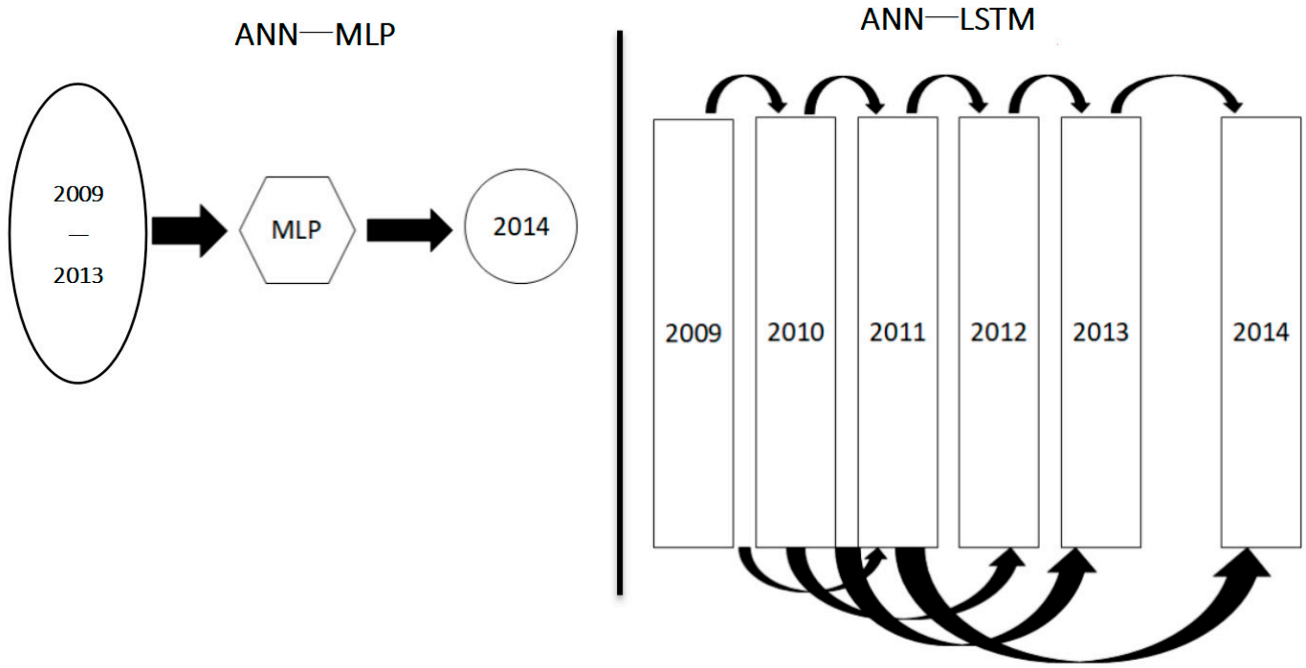


Figure 6. Differences between MLP and LSTM in dealing with annual land use data.

3.3. Model Result Comparison between the Liangjiang New Distract and Southern Auckland

We applied the same MLP-structured ANN–CA–MC model to southern Auckland, New Zealand, with a similar geodata set collected in 2018, to simulate urban expansion in this area. In this project, besides the ANN model, AHP-CA and logistic-CA were also applied to compare the simulation results and the ANN-based model that presented the best results as an intelligent urban model. More details for the Auckland-based research can be found in our earlier paper published in 2019 [42]; this former research provides us with the opportunity to perform a comparison. As a result, when comparing the results between Auckland (west) and Liangjiang New District (east), all the simulation model's validation results were higher in Auckland than in Liangjiang New District. The areas of simulated urban expansion in 2016 were validated by land cover change data from 1996 to 2006. Compared to Lianjiang New District, the simulation in Auckland showed a longer temporal discrepancy between the year of simulation and the year of the latest validation data (10 years for Auckland and 1 year for Liangjiang New District). However, all the model validation values were higher in Auckland than in Liangjiang New District (Table 4). With southern Auckland, the ANN–CA–MC could simulate urban expansion with a very high kappa (0.94) and acceptable kappa simulation (>0.50), demonstrating that both the urban pattern and urban expansion could be modeled at a high certainty level [42]. In Liangjiang New District, China, the model was less accurate. The validation results, though still considered acceptable, were less satisfactory compared to those for Auckland. The kappa was only slightly higher than 0.8, and the kappa simulation of MLP was below the medium level (0.4–0.6), while it was 0.48 for LSTM. These values suggest that we could still locate the major urban areas (kappa) within the study area, but the model is less accurate in simulating the changes in urban areas (kappa simulation).

Table 4. Model validation results between southern Auckland and Lianjiang New District.

Area	Model with Best Performance	Kappa	Kappa Simulation
Southern Auckland	ANN	0.941	0.547
Liangjiang New District	LSTM	0.857	0.48

The most significant reason for this discrepancy is the role of land use policies and planning in the simulation, which were not taken into account in the model, an issue that has also been identified by other researchers [50–52]. As is widely known, the governments in China act as the only legal landowners, and governments at different levels tend to have different perspectives on the prioritization of land utilization purposes [1,3,53]. The aforementioned Liangjiang New District plays a very important role in the national development strategy of China [54], where urbanization needs to be fast-paced and most of the new urban areas are pre-designated and planned by governments at both the local (Administration Committee of Liangjiang New District) and national levels (The State Council). Many policies and rules are used to restrict the location and distribution of future urban areas, such as the "Grain for Green", "Farmland Protection", and "Low Hilly Slope Zone" policies at the state level and "Destocking" and "Expropriation-Transfer Separation" at the local level. It is often found that the designated aims or outcomes of such multi-level policies conflict with their local or national counterparts [44]. Under this circumstance, the location of urban areas is more definitive, and their spatial distribution is more dispersed; these factors create difficulties in simulating and predicting their future patterns. Another possible reason for the less accurate modeled results is the quality of data supplied to the model. As a result, the urban expansion mode under government directions and guidance is much more complex than that in developed countries such as New Zealand. On the contrary, for cities such as Auckland, many have been experiencing slow development since the 1990s. However, the southern part of Auckland maintained significant growth, with an average annual urban development of almost 800.0 km^2. The simulated directional expansion of urban Auckland is in agreement with the government's desire and plan to build a compact city. Initially, the development in the south was mostly the continuation of the area's previous development, due to its low slope value (8.24 percent rise) compared to the other three directions. In 1996, the Metropolitan Urban Limit (MUL) was proposed and adopted to build a more compact city and to protect the coast and surrounding natural environment [55]. Therefore, the development in the first decade has moved from being primarily peripheral to infilling and intensification within the MUL. The land developers and new residents must obey the rules implemented by the government to match the land use, so the urban expansion over the last four decades in Auckland has been closely correlated to the government's 352 development strategies and policies [56]. The new urban areas are more regulated and clustered in distribution than those of fast-developing countries such as China.

Other factors affect the simulation accuracy, such as image spatial resolution: the urban information of Auckland was extracted from aerial photos with a very high spatial resolution (0.5 m), but it decreased to 30 m for Liangjiang New District. In addition, the precision of auxiliary data, such as DEM, had been degraded deliberately before they were released for public access from overseas. However, compared to policy factors, the impacts of these factors are relatively limited, especially in modeling urban development between western and eastern countries [57–60]. Furthermore, as many transportation components have been fed into the models, the relationship between transportation-related factors and future urban land expansion should also be analyzed. For example, does the urban area expand more quickly as more transportation facilities are created, or does the urban expansion follow the rules of transport plans? However, this paper classifies the urban area as one type rather than adopting transportation infrastructure as a single class. It does not calculate the different weights for all the impact factors, which reveals the ANN-based method cannot calculate the relative impacts as regression models [23]. Therefore, this is a future study direction in that we will seek to clarify the impact weights for these significant factors in our future studies with ANN-based algorithms.

4. Conclusions

In this paper, we presented two ANN-based CA–MC models to simulate the urban expansion inside a very rapidly developing area—Liangjiang New District, Chongqing. The obtained results confirm the possibility of using machine learning methods, such as MLP and LSTM, to improve the capability of CA–MC simulation models for the urban expansion field. The ANN–CA–MC outperformed LR–CA–MC, obtaining higher kappa and kappa simulation values. In the simulation, the most important layer, the USI map, can be efficiently created by MLP or LSTM, and provides critical information on potential future urban areas. LSTM is more powerful at capturing temporal information, and hence is more stable as well as predictable than the other two MLP-based CA–MC models with different sampling methods in simulating the fast annual urban expansion process for the study area. The results could provide spatial and qualitative information on urban expansion. Researchers could compare it with existing plans to check the expansion trends, implement the correct modifications and plan out new directions. For Liangjiang New District, we can see from the results that a higher temporal resolution could lead to more accurate results, and only a model that captures the time characteristics could yield a reliable result. Therefore, the LSTM-based CA–MC can be a powerful tool to add value to the urban expansion modeling field. In addition, different urbanization policies may highly affect the model's accuracy since a developing country may yield a more dynamic and rapidly changing result, so any policy should be considered carefully, especially for China and India. In general, the designed urban expansion simulation approaches can be applied to yield valuable and insightful information for local government decision makers and urban planners.

However, the simulation results obtained for Liangjiang New District were not as successful as those in other urban areas, even with the same model and similar datasets. Unlike the urban class patterns, the changes could not be well simulated, possibly because of the important role of land use policies from different levels of governments in urban development, which were not considered in the modeling. The effects of the government-directed land use directives were randomly distributed land use patterns with an ambiguous expansion mode or trend that was very difficult to model and predict. Therefore, an integrated model is still needed to overcome the limitations in simulating human behavior/policies at different levels. Therefore, in future studies, multiple-scenario simulations and predictions should be considered to address the uncertainty in development problems relating to land use policies more appropriately [8,16,61]. Furthermore, socioeconomic factors that also affect the urban expansion process need to be included because human behaviors and their interactions with natural and social changes also play critical roles in this dynamic process. In the future, Gross Domestic Product (GDP), personal income, population density, human movements, religion, employment rates, and so on should all be considered to produce more reliable and reasonable modeling outcomes. Lastly, focal/local transition rules should be considered, including the the global transition rule for CA models, in which different areas may be subjected to different urban expansion policies and dynamics.

Author Contributions: Conceptualization, T.X.; methodology, T.X.; software, T.X. and D.Z.; validation, T.X., D.Z. and Y.L.; formal analysis, T.X.; investigation, D.Z. and Y.L.; resources, T.X. and Y.L.; data curation T.X.; writing—original draft preparation, T.X.; writing—review and editing, T.X., D.Z. and Y.L.; visualization, T.X.; supervision, T.X.; project administration, T.X.; funding acquisition, T.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the "Overseas students' innovation and entrepreneurship plan", Chongqing, grant number cx2021065.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, Tingting Xu, upon reasonable request.

Acknowledgments: The authors would like to acknowledge the Chongqing Land Resource Information Center and Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences, for providing significant and valuable data for the research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Demurger, S.; Sachs, J.D.; Woo, W.T.; Shuming, B.A.O.; Chang, G. The relative contributions of location and preferential policies in China's regional development: Being in the right place and having the right incentives. *China Econ. Rev.* **2002**, *13*, 444–465. [CrossRef]
2. Deng, X.Z.; Huang, J.K.; Rozelle, S.; Uchida, E. Growth, population and industrialization, and urban land expansion of China. *J. Urban Econ.* **2008**, *63*, 96–115. [CrossRef]
3. Zhang, Q.; Wallace, J.; Dong, X.Z.; Seto, K.C. Central versus local states: Which matters more in affecting China's urban growth? *Land Use Policy* **2014**, *38*, 487–496. [CrossRef]
4. Engelfriet, L.; Koomen, E. The impact of urban form on commuting in large Chinese cities. *Transportation* **2018**, *45*, 1269–1295. [CrossRef]
5. He, J.Q.; Wang, S.J.; Liu, Y.; Ma, H.T.; Liu, Q. Examining the relationship between urbanization and the eco-environment using a coupling analysis: Case study of Shanghai, China. *Ecol. Indic.* **2017**, *77*, 185–193. [CrossRef]
6. Jat, M.K.; Garg, P.K.; Khare, D. Modelling of urban growth using spatial analysis techniques: A case study of Ajmer city (India). *Int. J. Remote Sens.* **2008**, *29*, 543–567. [CrossRef]
7. Huang, B.; Zhang, L.; Wu, B. Spatiotemporal analysis of rural–urban land conversion. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 379–398. [CrossRef]
8. Shafizadeh Moghadam, H.; Helbich, H. Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. *Appl. Geogr.* **2013**, *40*, 140–149. [CrossRef]
9. Shafizadeh-Moghadam, H.; Helbich, H. Spatiotemporal variability of urban growth factors: A global and local perspective on the megacity of Mumbai. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35 Pt B*, 187–198. [CrossRef]
10. Sudhira, H.S.; Ramachadra, T.V.; Jagadish, K.S. Urban sprawl: Metrics, dynamics and modelling using GIS. *Int. J. Appl. Earth Obs. Geoinf.* **2004**, *5*, 29–39. [CrossRef]
11. Liu, J.; Liu, M.; Tian, H.; Zhuang, D.; Zhang, Z.; Zhang, W.; Deng, X. Spatial and temporal patterns of China's cropland during 1990–2000: An analysis based on Landsat TM data. *Remote Sens. Environ.* **2005**, *98*, 442–456. [CrossRef]
12. Deng, J.S.; Wang, K.; Hong, Y.; Qi, J.G. Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. *Landsc. Urban Plan.* **2009**, *92*, 187–198. [CrossRef]
13. Seto, K.C. Exploring the dynamics of migration to mega-delta cities in Asia and Africa: Contemporary drivers and future scenarios. *Glob. Environ. Chang.* **2011**, *21* (Suppl. S1), S94–S107. [CrossRef]
14. Jiang, L.; Deng, X.Z.; Seto, K.C. Multi-level modeling of urban expansion and cultivated land conversion for urban hotspot counties in China. *Landsc. Urban Plan.* **2012**, *108*, 131–139. [CrossRef]
15. Ma, Y.; Xu, R. Remote sensing monitoring and driving force analysis of urban expansion in Guangzhou City, China. *Habitat Int.* **2010**, *34*, 228–235. [CrossRef]
16. Arsanjani, J.; Helbich, M.; Kainz, W.; Boloorani, A.D. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 265–275. [CrossRef]
17. Xu, C.; Liu, M.; Hong, C.; Chi, T.; An, S.; Yang, X. Temporal variation of characteristic scales in urban landscapes: An insight into the evolving internal structures of China's two largest cities. *Landsc. Ecol.* **2012**, *27*, 1063–1074. [CrossRef]
18. Liu, Y.; He, Q.; Tan, R.; Zhou, K.; Liu, G.; Tang, S. Urban growth modeling based on a game between farmers and governments: Case study of urban fringe in Wuhan, Hubei Province in China. *J. Urban Plan. Dev.* **2015**, *142*, 04015018. [CrossRef]
19. Zhan, S.H. Hukou Reform and Land Politics in China: Rise of a Tripartite Alliance. *China J.* **2017**, *78*, 25–49. [CrossRef]
20. Yue, W.Z.; Wang, R.L.; Fan, B.L. Spatial patterns analysis of urban expansion in Hangzhou city. *J. ZheJiang Univ. (Sci. Ed.)* **2013**, *40*, 596–605.
21. Michael, W.; Franz, F. Land-Use Transport Interaction: State of the Art. Available online: https://ssrn.com/abstract=1434678 (accessed on 1 January 2004).
22. Wegener, M. Land-Use Transport Interaction Models. In *Handbook of Regional Science*; Fischer, M., Nijkamp, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2014.
23. Dingil, A.E.; Rupi, F.; Esztergár-Kiss, D. An Integrative Review of Socio-Technical Factors Influencing Travel Decision-Making and Urban Transport Performance. *Sustainability* **2021**, *13*, 10158. [CrossRef]
24. Batty, M.; Xie, Y.; Sun, Z. Modeling urban dynamics through GIS-based cellular automata. *Comput. Environ. Urban Syst.* **1999**, *23*, 205–233. [CrossRef]
25. Li, X.; Yeh, A.G.O. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *Int. J. Geogr. Inf. Sci.* **2000**, *14*, 131–152. [CrossRef]

26. Aburas, M.M.; Ho, Y.M.; Ramli, M.F.; Ash'aari, Z.H. The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 380–389. [CrossRef]
27. Arsanjani, J.J.; Helbich, M.; de Noronha Vaz, E. Spatiotemporal simulation of urban growth patterns using agent-based modeling: *Case Tehran*. *Cities* **2013**, *32*, 33–42. [CrossRef]
28. Zhang, H.; Jin, X.; Wang, L.; Zhou, Y.; Shu, B. Multi-agent based modeling of spatiotemporal dynamical urban growth in developing countries: Simulating future scenarios of Lianyungang city, China. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 63–78. [CrossRef]
29. Pijanowski, B.C.; Brown, D.G.; Shellito, B.A.; Manik, G.A. Using neural networks and GIS to forecast land use changes: A land transformation model. *Comput. Environ. Urban Syst.* **2002**, *26*, 553–575. [CrossRef]
30. Jat, M.K.; Choudhary, M.; Saxena, A. Urban growth assessment and prediction using RS, GIS and SLEUTH model for a heterogeneous urban fringe. *Egypt. J. Remote Sens. Space Sci.* **2017**, *10*, 1–19.
31. White, R.; Engelen, G.; Uljee, I. *Modeling Cities and Regions as Complex Systems*; MIT Press: Cambridge, MA, USA, 2015.
32. White, R. Cities and cellular automata. *Discret. Dyn. Nat. Soc.* **1998**, *2*, 111–125. [CrossRef]
33. Kim, Y.; Safikhani, A.; Tepe, E. Machine learning application to spatio-temporal modeling of urban growth. *Comput. Environ. Urban Syst.* **2022**, *94*, 101801. [CrossRef]
34. Yi, S.; Zhou, Y.; Li, Q. A New Perspective for Urban Development Boundary Delineation Based on the MCR Model and CA-Markov Model. *Land* **2022**, *11*, 401. [CrossRef]
35. Ghosh, P.; Mukhopadhyay, A.; Chanda, A.; Mondal, P.; Akhand, A.; Mukherjee, S.; Hazra, S. Application of Cellular automata and Markov-chain model in geospatial environmental modeling—A review. *Remote Sens. Appl. Soc. Environ.* **2017**, *5*, 64–77. [CrossRef]
36. Islam, K.; Rahman, M.F.; Jashimuddin, M. Modeling land use change using Cellular Automata and Artificial Neural Network: The case of Chunati Wildlife Sanctuary, Bangladesh. *Ecol. Indic.* **2018**, *88*, 439–453. [CrossRef]
37. Hagenauer, J.; Helbich, M. Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 963–982. [CrossRef]
38. Pijanowski, B.C.; Tayyebi, A.; Doucette, J.; Pekin, B.K.; Braun, D.; Plourde, J. A big data urban growth simulation at a national scale: Configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment. *Environ. Model. Softw.* **2014**, *51*, 250–268. [CrossRef]
39. Paliwal, M.; Kumar, U.A. Neural Networks and Statistical Techniques: A Review of Applications. *Expert Syst. Appl.* **2009**, *36*, 2–17. [CrossRef]
40. Berberoğlu, S.; Akın, A.; Clarke, K.C. Cellular automata modeling approaches to forecast urban growth for Adana, Turkey: A comparative approach. *Landsc. Urban Plan.* **2016**, *153*, 11–27. [CrossRef]
41. Li, X.; Yeh, A.G.O. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *Int. J. Geogr. Inf. Sci.* **2002**, *16*, 323–343. [CrossRef]
42. Xu, T.; Gao, J.; Coco, G. Simulation of urban expansion via integrating artificial neural network with Markov chain–cellular automata. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 1960–1983. [CrossRef]
43. Xu, T.; Gao, J.; Coco, G. Extraction of urban built-up areas from nighttime lights using artificial neural network. *Geocarto Int.* **2020**, *35*, 1049–1066. [CrossRef]
44. Xu, T.; Gao, J.; Li, Y. Machine learning-assisted evaluation of land use policies and plans in a rapidly urbanizing district in Chongqing, China. *Land Use Policy* **2019**, *87*, 30–40. [CrossRef]
45. Park, S.; Jeon, S.; Kim, S.; Choi, C. Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea. *Landsc. Urban Plan.* **2011**, *99*, 104–114. [CrossRef]
46. Rumelhart, D.; Hinton, G.; Williams, R. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*; Rumelhart, D.E., McClelland, J.L., Eds.; MIT Press: Cambridge, UK, 1986; Volume 1, pp. 318–362.
47. Lipton, Z.C.; Berkowitz, J.; Elkan, C. A critical review of recurrent neural networks for sequence learning. *arXiv* **2015**, arXiv:1506.00019.
48. Hagan, M.T.; Demuth, H.B.; Beale, M.H. *Neural Network Design*; PWS Publishing: Boston, MA, USA, 1996.
49. van Vliet, J.; Bregt, A.K.; Hagen-Zanker, A. Revisiting Kappa to account for change in the accuracy assessment of land-use change models. *Ecol. Model.* **2011**, *222*, 1367–1375. [CrossRef]
50. Deng, Y. *The Strategy Research for the Character of LiangJiang New District during the "Openning West"*; Xian Technology University: Xi'an, China, 2011.
51. Fu, F.C.; Zhou, B.T.; Liang, K.X. Research of the Ecosystem Services Value (ESV) of Land Use in Liangjiang New Area of Chongqing. *J. Southwest China Norm. Univ. (Nat. Sci. Ed.)* **2012**, *37*, 67–73.
52. Fu, J.P. Study on Land Use Change of Liangjiang New Area in Chongqing. *J. Green Sci. Technol.* **2017**, *2017*, 180–182.
53. Wang, J.; Chen, Y.Q.; Shao, X.M.; Zhang, Y.Y.; Cao, Y.G. Land-use changes and policy dimension driving forces in China: Present, trend and future. *Land Use Policy* **2012**, *29*, 737–749. [CrossRef]
54. Li, L. State rescaling and national new area development in China: The case of Chongqing Liangjiang. *Habitat Int.* **2015**, *50*, 80–89. [CrossRef]

55. Auckland City Council. The Auckland Plan. Available online: https://www.aucklandcouncil.govt.nz/plansprojects-policies-reports-bylaws/our-plans-strategies/Documents/auckland-plan-2012-fulldocument.pdf (accessed on 15 January 2012).
56. Auckland Regional Council. A Brief History of Auckland's Urban Form. Auckland Council: Scientific Report. 2010. Available online: http://knowledgeauckland.org.nz/assets/publications/Brief_history_of_Auckland_s_urban_form.pdf (accessed on 30 November 2010).
57. Chen, T.; Chiu, H.; Lin, Y. How do East and Southeast Asian Cities Differ from Western Cities? A Systematic Review of the Urban Form Characteristics. *Sustainability* **2020**, *12*, 2423. [CrossRef]
58. Cegielska, K.; Noszczyk, T.; Kukulska, A.; Szylar, M.; Hernik, J.; Dixon-Gough, R.; Jombach, S.; Valánszki, I.; Kovács, K.F. Land use and land cover changes in post-socialist countries: Some observations from Hungary and Poland. *Land Use Policy* **2018**, *78*, 1–18. [CrossRef]
59. Liu, Y.; Li, J.; Yang, Y. Strategic adjustment of land use policy under the economic transformation. *Land Use Policy* **2018**, *74*, 5–14. [CrossRef]
60. Zhu, J.; Simarmata, H.A. Formal land rights versus informal land rights: Governance for sustainable urbanization in the Jakarta metropolitan region, Indonesia. *Land Use Policy* **2015**, *43*, 63–73. [CrossRef]
61. Mustafa, A.; Cools, M.; Saadi, I.; Teller, J. Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM). *Land Use Policy* **2017**, *69*, 529–540. [CrossRef]