

## Article

# Spatial Modelling and Prediction with the Spatio-Temporal Matrix: A Study on Predicting Future Settlement Growth

Zhiyuan Wang <sup>1,\*</sup> , Felix Bachofer <sup>1</sup> , Jonas Koehler <sup>1</sup> , Juliane Huth <sup>1</sup> , Thorsten Hoerer <sup>1</sup> ,  
Mattia Marconcini <sup>1</sup> , Thomas Esch <sup>1</sup>  and Claudia Kuenzer <sup>1,2</sup>

<sup>1</sup> German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Oberpfaffenhofen, D-82234 Wessling, Germany; felix.bachofer@dlr.de (F.B.); jonas.koehler@dlr.de (J.K.); juliane.huth@dlr.de (J.H.); thorsten.hoerer@dlr.de (T.H.); mattia.marconcini@dlr.de (M.M.); thomas.esch@dlr.de (T.E.); claudia.kuenzer@dlr.de (C.K.)

<sup>2</sup> Department of Remote Sensing, Institute of Geography and Geology, University of Wuerzburg, Am Hubland, D-97074 Wuerzburg, Germany

\* Correspondence: zhiyuan.wang@dlr.de; Tel.: +49-8153-28-4134

**Abstract:** In the past decades, various Earth observation-based time series products have emerged, which have enabled studies and analysis of global change processes. Besides their contribution to understanding past processes, time series datasets hold enormous potential for predictive modeling and thereby meet the demands of decision makers on future scenarios. In order to further exploit these data, a novel pixel-based approach has been introduced, which is the spatio-temporal matrix (STM). The approach integrates the historical characteristics of a specific land cover at a high temporal frequency in order to interpret the spatial and temporal information for the neighborhood of a given target pixel. The provided information can be exploited with common predictive models and algorithms. In this study, this approach was utilized and evaluated for the prediction of future urban/built-settlement growth. Random forest and multi-layer perceptron were employed for the prediction. The tests have been carried out with training strategies based on a one-year and a ten-year time span for the urban agglomerations of Surat (India), Ho-Chi-Minh City (Vietnam), and Abidjan (Ivory Coast). The slope, land use, exclusion, urban, transportation, hillshade (SLEUTH) model was selected as a baseline indicator for the performance evaluation. The statistical results from the receiver operating characteristic curve (ROC) demonstrate a good ability of the STM to facilitate the prediction of future settlement growth and its transferability to different cities, with area under the curve (AUC) values greater than 0.85. Compared with SLEUTH, the STM-based model achieved higher AUC in all of the test cases, while being independent of the additional datasets for the restricted and the preferential development areas.

**Keywords:** spatio-temporal analysis; time series; EO data; settlement growth; machine learning; urban modelling; future prediction



**Citation:** Wang, Z.; Bachofer, F.; Koehler, J.; Huth, J.; Hoerer, T.; Marconcini, M.; Esch, T.; Kuenzer, C. Spatial Modelling and Prediction with the Spatio-Temporal Matrix: A Study on Predicting Future Settlement Growth. *Land* **2022**, *11*, 1174. <https://doi.org/10.3390/land11081174>

Academic Editor: Deodato Tapete

Received: 15 June 2022

Accepted: 25 July 2022

Published: 28 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



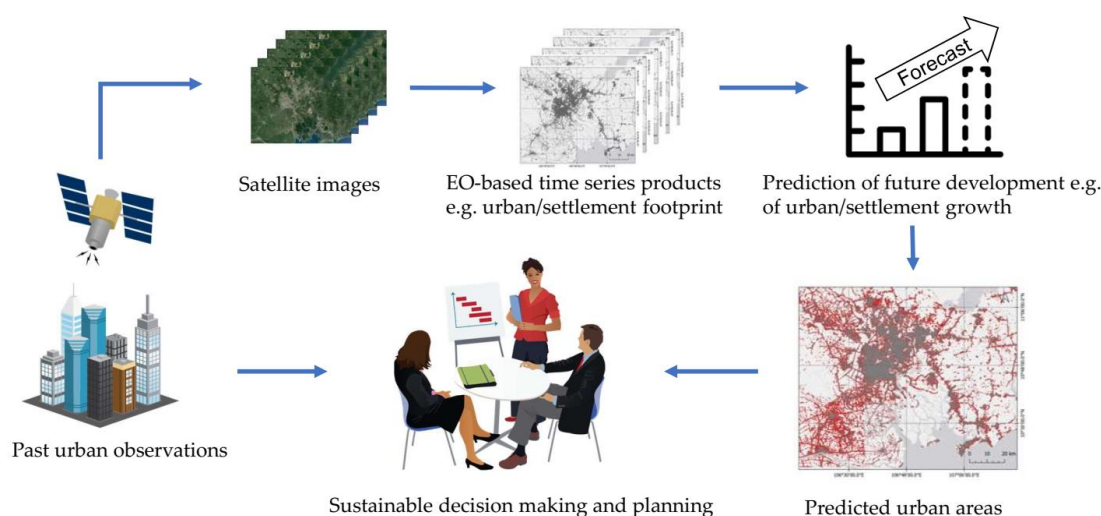
**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Remote sensing technology has been providing Earth observation (EO) data for decades by recording information on the surface of the Earth. The EO data from satellites has the advantage of providing globally-covered long-term measurements as follows: several satellite missions, such as Landsat [1], advanced very high resolution radiometer (AVHRR) [2], and moderate resolution imaging spectroradiometer (MODIS) [3], which have been operated for years and are able to provide historical long-term observations for creating long-term land use and land cover (LULC) time series products. In recent years, various thematic LULC time series products have been generated for water [4], snow cover [5], vegetation index [6], urban/settlement extents [7,8], and general LULC [9]. With the help of the EO-based time series products, historical records of land surface dynamics can be derived and used in order to forecast the future trends and dynamics. According to

Koehler and Kuenzer [10], a variety of EO-based forecast applications have been carried out for LULC, crop yield, vegetation cover, flood, and urbanization. Although many approaches have been developed for forecast applications in different fields, most of them require a variety of complex data in addition to EO-based time series data products, e.g., the precipitation data for vegetation cover prediction [11] and additional local geometric data, such as road networks [12] and population data [13], for future urbanization prediction. The external data sources acquisition increases the complexity of applying such approaches for the forecast applications.

Since the predictive modelling in the settlement growth context is well established in the scientific community, the methodological approach of this study has been illustrated and evaluated with urban/built-settlement test cases. Figure 1 illustrates the general workflow of the prediction of urban/settlement growth that is based on EO-based time series products.



**Figure 1.** Prediction of future urban/settlement growth based on EO-based time series products. Source of the symbols is the Integration and Application Network ([ian.umces.edu/media-library](http://ian.umces.edu/media-library)).

The global population growth, migration, and socio-economic development have led to rapidly accelerated urbanization processes and have, therefore, brought challenges; many middle- and small-sized cities face the challenge of lacking public infrastructure and institutions [14]. Since urban areas become hubs of transportation and industries, the increased urbanization level causes a rise in carbon emissions [15] and energy demand [16]. Rapid settlement growth also causes land-use conflicts; as one of the most affected land-use types, cultivated lands have been rapidly transformed into settlement areas [17]. In addition, with cities turning into centers for social and economic developments, unplanned settlement growth can take place in areas that are prone to natural hazards, such as flash floods [18], sea-level rise [19], and landslides [20]. The maintenance of urban areas brings also challenges for sustainable development. Artificially sealed surfaces and a decrease in urban green coverage lead to a decrease in and a deterioration of ecosystem services, e.g., heat mitigation, carbon storage, stormwater retention, and oxygen production [21]. The information products that are derived from EO data have been widely used for urban modelling for decades and have proven their potential for predicting urban growth [22]. In the last ten years, high-resolution layers outlining the global settlement extent have been generated from satellite imagery, thus proving the value of remote-sensing-based products for large-scale urban modelling applications. These include the global human settlement layer (GHSL) [23], the global annual impervious area (GAIA) [24], and the global impervious surface area (GISA) [25] products. In particular, the World Settlement Footprint (WSF) evolution dataset was released [26], which outlines the global settlement extent at 30 m resolution on a yearly basis, from 1985 to 2015.

So far, various methods have been employed for future urban/built-settlement growth prediction. The CLUE-S model [27–29] and the future land use simulation (FLUS) model [30] are popular for LULC modelling. These models aim for multi-class land cover change prediction. For prediction focusing on urban development, Gao and O’Neil [13] developed a data-driven approach, which was the spatially-explicit, long-term, empirical city development (SELECT) model. The cellular automata (CA) is another popular method for spatially explicit urban growth modelling [31,32]. A prominent variant of rule-based approaches in CA-based urban modelling is the slope, land use, exclusion, urban, transportation, hillshade (SLEUTH) model [33], whose acronym is derived from its main input variables as follows: slope, land use, exclusion, urban, transportation, and hillshade. This method has been widely used for future urban growth modelling [12,34–38].

For future urban/settlement prediction, two factors are essential, namely the spatial factor and the temporal factor. In most approaches, the spatial factor can be described as the spatial structure of the neighbors of the target [39]. Unlike the spatial factor, the temporal factor is usually achieved by using historical urban layers as inputs for the calibration process of a model [31]. For the EO-based data, the representation of the temporal factor remains to be a challenge. So far, only a few studies have explored the temporal factors for predicting urban development. Wang et al. [40] modeled the temporal evolution of each pixel. Schneider et al. [41] defined four spatial-lagged variables based on the inverse-distance weighted leave-one-out cross-validated (IDW-LOOCV) approach. Wang et al. [42] adapted a smoothing process in order to integrate the different transition rules for different historical periods.

Recently, studies have concentrated on the continuous temporal trend of historical urban growth. Inspired by the following assumption: “Recently developed regions have a higher likelihood to be developed in the near future” from Liu et al. [30], Li et al. [43] developed a Logistic-Trend-CA, which integrates the temporal weights that are calculated from the historical pathways of urban sprawls, based on a window that defines the neighborhood of the target pixel. This study is the first of its kind, as it implements continuous urban observation time series into a CA model and has achieved a good accuracy. Later, this method was employed by Johnson et al. [44] in order to predict the future urban growth for flood exposure estimation. Li et al. [45] further applied this model for the hindcast and the future projection of urban dynamics on a global scale. However, the Logistic-Trend-CA still requires various external spatial proxies in order to provide extra spatial constraints.

Similar to the other forecast applications, the first challenge of future urban/built-settlement growth modelling is that the multitude of drivers and factors have resulted in the abundant use of data-driven probability-based models. This brings about a dependence on various explanatory variables and necessitates the accurate and timely input of data. While nowadays reliable maps of the past yearly settlement extent can be generated by means of EO data, additional information, such as the infrastructure networks, socio-economic statistics, etc., may be hard to come by, especially when historical data is needed [31]. The second aspect of potential model improvement pertains to the integration of spatial factors and temporal factors with high temporal resolution. In most cases, the transition probabilities of CA-based models are computed based on the differences between the discrete observations. Koehler and Kuenzer [10] found that long satellite time-series data are seldom used in urban/settlement growth modelling and that the temporal information that is stored in urban time-series products, such as the WSF-evolution [8], is yet unexploited.

Consequently, this paper proposes the spatio-temporal matrix (STM), which is a novel pixel-based approach that tackles the aforementioned challenges in thematic future development modelling. Generated purely from EO-based time-series products, the STM makes full use of the spatial and temporal information that is stored in time-series data equally by considering not only the current state of the neighborhood, but also the historical state of these pixels. Hence, the probability of the future development of the pixel thus depends on the degree of the development and the spatial structure in its neighborhood during the past development. This study utilizes the STM-based approach for future

urban/settlement growth prediction, which is used to evaluate the approach and to discuss its advantages and limitations in the frame of the current state of the research. Besides utilizing the STM-based approach for urban growth prediction, it can presumably also be used for prediction in other thematic domains.

## 2. Materials and Methods

### 2.1. Data

#### 2.1.1. Global Urban/Settlement Time Series Data: The WSF-Evolution

The World Settlement Footprint (WSF) suite is an unprecedented collection of open-and-free global datasets that aim to advance the understanding of urbanization at the planetary scale. In this framework, the first layer to be released was the WSF2015, a 10 m resolution binary mask outlining the 2015 global settlement extent, which was derived by jointly exploiting multitemporal optical Landsat-8 and radar Sentinel-1 imagery [8]. This layer provides highly accurate and reliable settlement information, which was quantitatively assessed by an extensive validation based on 900,000 ground-truth samples labelled by crowd-sourcing photointerpretation of very high resolution (VHR) satellite imagery. Since a proper understanding of the past growth is essential for characterizing ongoing trends, a novel iterative approach has been implemented, starting backward from 2015, that effectively outlines on a yearly basis the settlement extent based on Landsat data alone (given the lack of systematically available archived high-resolution radar imagery). Out of all Landsat scenes available for the given study region, the minimum, maximum, mean, and standard deviation over time per pixel of different spectral indices are computed for each year in the past. Among others, these indices include the normalized difference built-up index (NDBI), normalized difference vegetation index (NDVI), and modified normalized difference water index (MNDWI). Then, using the WSF2015 as reference, settlement and non-settlement training samples for the year  $t$  are extracted, starting from 2015, by first applying morphological filtering to the settlement mask generated for the year  $t + 1$ , then adaptively thresholding the corresponding temporal mean of NDBI, NDVI, and MNDWI. A random forest (RF) classification is eventually applied over the sole pixels marked as the settlement at time  $t + 1$ . After an extensive test phase, the approach has been ultimately employed for generating the WSF-evolution, i.e., a dataset outlining the global annual settlement extent at 30 m spatial resolution from 1985 to 2015 [8] with binary properties of “settlement” and “non-settlement”. In particular, the WSF-evolution has proven to be the most accurate product of its type, as assessed by means of an extensive campaign, similar to that carried out for the WSF2015, where overall ~1.2 M samples have been labelled for the years 1990, 1995, 2000, 2005, 2010 and 2015.

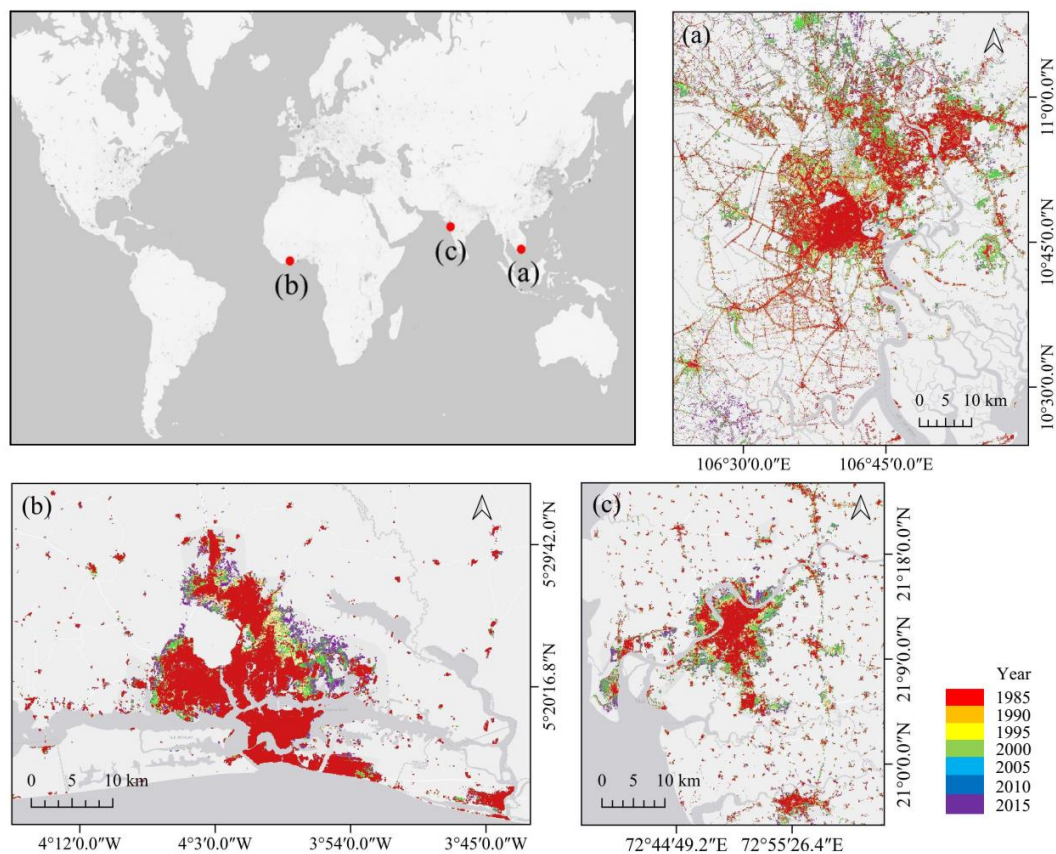
Having a long-term coverage (from 1985 to 2015) with annual data layers, the WSF-evolution was selected as the underlying EO-time series product for the prediction of future settlement growth in this study.

#### 2.1.2. Study Sites

The following three dynamic urban agglomerations with differing growth patterns have been selected as test sites: Ho-Chi-Minh City (HCMC) in Vietnam, Abidjan in the Ivory Coast, and Surat in India (Figure 2).

HCMC is the largest city of Vietnam and has been expanding rapidly since 1985. The population was 2,820,000 in 1985 and reached 7,348,000 in 2015 [46]. The expansion of the city shows an irregular growth pattern, with densification and expansion of its fringes, as well as sprawl along the road network (Figure 2a).

Abidjan is the capital of the Ivory Coast and is the largest city in the country. Being one of the six largest cities in Africa, Abidjan experienced an enormous population growth, leading to a rapid urban expansion. From 1985 to 2015, the population has increased from 1,716,000 to 4,533,000 [47]. Abidjan has been characterized by an uneven growing pattern for the last several decades (Figure 2b). Because of its location on the coast and the wetlands in the south, the city has expanded mainly in the northern direction.



**Figure 2.** WSF-evolution: overview maps; (a) HCMC in Vietnam; (b) Abidjan in the Ivory Coast; and (c) Surat in India.

Surat is a coastal city located in the Indian state of Gujarat. As the economic and commercial center of Gujarat, the city is under rapid development and has become one of the fastest-growing cities in India. Surat's population increased from 1,094,480 in 1985 to 5,401,214 in 2015 [46]. The city has shown a regular growth pattern in the last decades, where growth happened mostly along its fringes (Figure 2c).

The study sites are located in the Global South and have experienced a rapid development in recent decades, which seems to carry on in the near and mid-term future [46].

Moreover, the selected regions have experienced different types of growing behavior in the past, which makes them attractive for this study. HCMC faces urban sprawl in a scattered distributed pattern, due to its widespread and interconnected traffic network. The growth of Abidjan was directed by physical barriers, terrain, and restricted areas. Surat has experienced a uniform extension in the past decades, with most of the growth along its edges and fringes.

## 2.2. Spatio-Temporal Matrix (STM)

The proposed approach targets specific land surface developments represented by discrete values that can be interpreted as a continuous spatio-temporal process. Spatially, the target land surface develops itself along a spatial trend, e.g., expansion and shrinkage; temporally, the change process is continuous, with a high temporal dependency. For example, this kind of land surface state change takes place in urbanization, deforestation, and desertification.

The STM contains both spatial and temporal information of the target pixel and the pixels of a defined neighborhood. Being a pixel-based feature, the STM can serve as an input for machine learning (ML) algorithms, providing solid spatial and temporal constraints for the training and prediction phases of the modelling.

### 2.2.1. Assumptions

The STM condenses the spatio-temporal information of the target's neighborhood into one single matrix, based on the assumption that the probability of the change on the target pixel is affected by the spatial structure of the target's neighborhood [48], as well as the historical development of the pixels. Being a pixel-based approach, the STM is based on the following assumptions: (i) The status of a pixel will not change back to the former state once it has been changed. (ii) The target pixel is more likely to change if its neighboring pixels have changed in the recent past observation period.

In an urban context, these assumptions can be interpreted as follows: (i) A pixel will remain an urban pixel once it is urbanized. (ii) If a pixel is adjacent to or enclosed by newly developed urban pixels, it is more likely that this pixel is concurrently under development and will be urbanized in the near future. This assumption introduces an attraction effect by newly developed regions and describes the continuity of the urban development process. (iii) If the neighboring pixels of a non-urbanized pixel have been urbanized for a long time, the target pixel is assumed to have a low probability of change in the future. This premise makes use of the intrinsic information on persistence and introduces a resistance effect by considering legal, physical, or other restrictions preventing expansion and will continue to do so in the future (e.g., natural preserved areas, water bodies, areas with steep slopes, transport infrastructure, military areas, etc.).

Assumption (i) holds true for the test arrangement of this study. However, in some special cases, the change can be reversible. For example, after big catastrophes, e.g., severe natural hazards and financial crises, urban areas may decline. The currently existing globally-covered settlement dataset based on EO data, e.g., WSF-evolution [26] and GHSL [7], indicate only settlement growth. For this reason, a decline function is not considered in the following STM set-up but can be introduced if necessary.

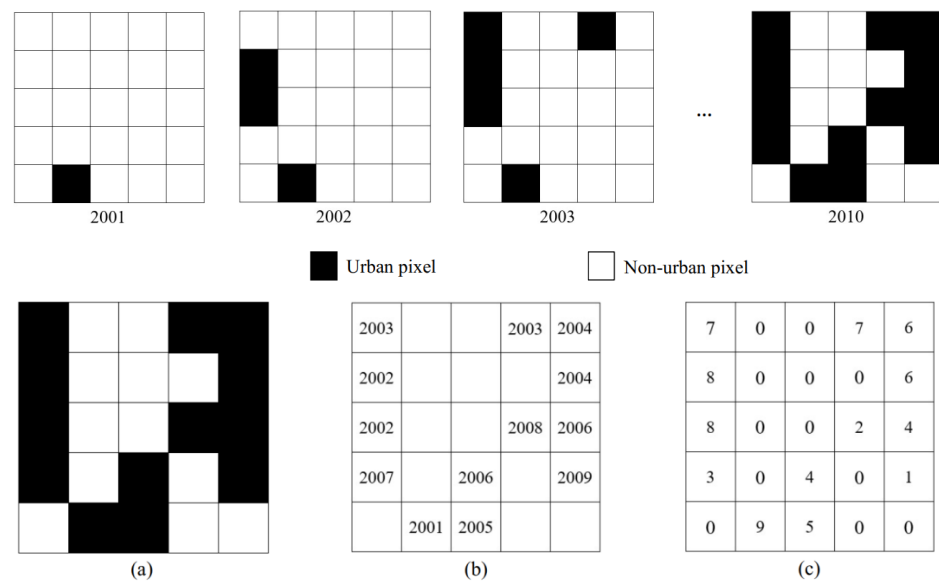
### 2.2.2. Matrix Design

The temporal information of the historical development is extracted from the EO-based time series product for a defined time span of the past. The length of the time span is defined by the user. A start year  $t_{start}$  and an end year  $t_{end}$  are set to define the time span of the historical settlement layers that are used for STM generation, with an annual temporal resolution given by the WSF product. The year for which a given pixel is marked as a settlement pixel is denoted as  $t_u$ . This value is set to 0 if the pixel is denoted as "non-settlement" throughout the entire considered time span. A parameter  $P$  named "Continuation Period" is defined, which specifies the duration of a given pixel being categorized as a settlement, as follows:

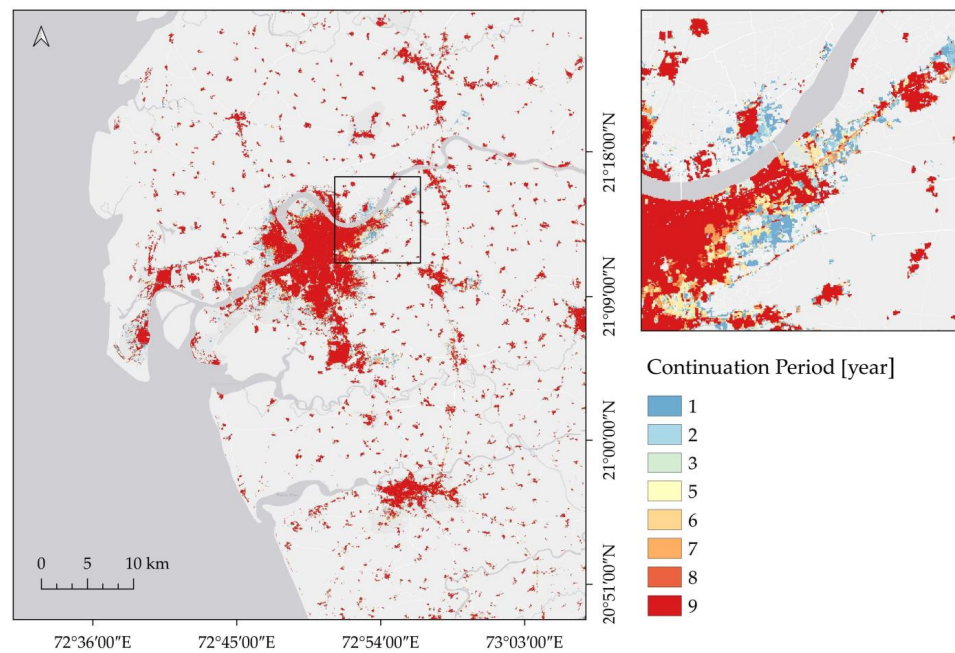
$$P = \begin{cases} t_{end} - t_u, & t_u \geq t_{start} \\ t_{end} - t_{start}, & t_u < t_{start} \\ 0, & t_u = 0 \end{cases} \quad (1)$$

Figure 3 illustrates an example of the calculation of  $P$ . The start year  $t_{start}$  is defined as 2001 and the end year  $t_{end}$  as 2010. The information of the  $t_u$  of each pixel is derived from the WSF-evolution.

$P$  is the key to integrating the observations with high temporal resolution. With this parameter, the temporal information of the settlement extents is incorporated. Instead of indicating the class type of the given pixel, the temporal constraint provided by  $P$  can be used to predict the urban/built-settlement growth scenarios described by assumptions (ii) and (iii). Figure 4 illustrates the continuation period layer of Surat.



**Figure 3.** Generation of continuation period layer, with the start year of 2001 and the end year of 2010. Pixels that have been urbanized are blackened, un-urbanized pixels stay white: (a) subset of a raster binary image of the urban extent in 2010; (b) the year of urbanization  $t_u$  is derived from each pixel; (c) the difference between the end year and the year of urbanization depicts the continuation period.



**Figure 4.** Continuation period layer for Surat, India, generated from the annual WSF-evolution layers of Surat (2001 to 2010).

A squared window is used to fill a matrix that includes the  $P$  of the target pixel and its neighbors. For pixels at the edge of an image, a padding value of zero will be added. The formed matrix of each pixel is named as the spatio-temporal matrix (STM) (Equation (2)).

The window size for forming the matrix  $s_w$  can be defined by the user based on the resolution of the input images and the observed land cover changes. The window size is critical for this approach; if the window size is too small, it will not consider the effects from more distant neighbors; if it is too big, the STM might consider too much information from non-relevant pixels. The optimal window size in the following section was chosen by empirically testing different window sizes in different implementation strategies.

After being generated, the STM of each pixel is transformed into a single STM vector (Equations (2) and (3)). This transformation is performed for the following three reasons:

1. The continuation period  $P$  of each neighboring pixel will be preserved;
2. The spatial position of each neighboring pixel will be recorded, as the sequence of  $P$  from all elements in the STM will remain in the formed vector;
3. The STM vector of each pixel can be used as the input for further training and prediction processes.

$$STM = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \tag{2}$$

$$STM \text{ vector} = \begin{bmatrix} P_{11} \\ P_{12} \\ \vdots \\ P_{1n} \\ \vdots \\ P_{n1} \\ \vdots \\ P_{nn} \end{bmatrix} \tag{3}$$

where  $n$  is the window size and  $P$  represents the corresponding continuation period of each pixel.

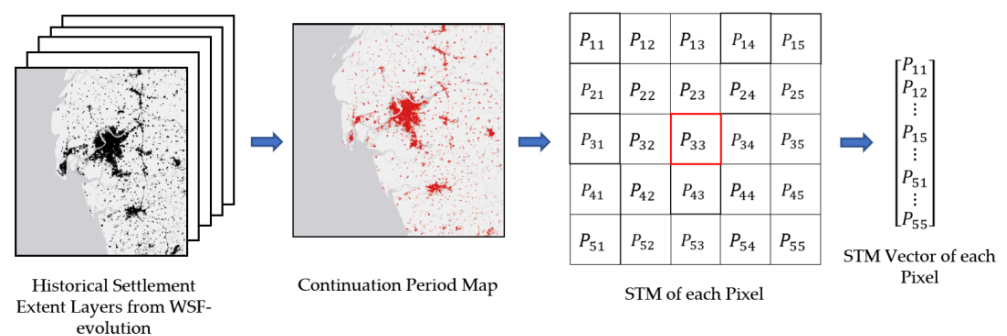
The generated STM vectors of each pixel can be directly used as inputs for the machine learning algorithm. During training, a feature space will be established by the STM vector, and the transition rules between features and the probability of the future settlement development of the target pixel will be studied. For the training and prediction, the class of the future target pixel is defined as  $y_{i,j}$  as follows:

$$y_{i,j} = \begin{cases} 0, & \text{No change (non urbanized/urbanized)} \\ 1, & \text{Newly urbanized} \end{cases} \tag{4}$$

where  $i, j$  are the row and column indices of the target pixel in the raster map.

To conclude,  $i \times j$  STM vectors are generated. Compared to the currently established time-series layers stacking that would be used as inputs for predictive models [32], this feature can reduce the data volume and redundancy by integrating annual settlement extent layers into one layer.

The workflow of STM generation is shown in Figure 5.



**Figure 5.** Workflow of generation of STM, illustrated based on a window size  $s_w$  of 5.

### 2.3. Machine-Learning Algorithms

The prediction of future urban/built-settlement growth can be described as a binary classification task by defining two classes of pixels at a fixed date in the future, repre-



senting “newly urbanized” and “non-urbanized”. In this paper, RF and a multi-layer perceptron (MLP) were selected for testing the performance of the STM as inputs for future growth prediction.

The RF is built by assembling numbers of decision trees [49], generated by samples drawn from the training set using a bootstrap strategy. The probability of the predicted class of the whole forest is calculated as the mean predicted class probabilities of every tree in the forest. The MLP is a type of feedforward artificial neural network (ANN). It consists of at least three layers as follows: an input layer, a hidden layer, and an output layer. In the pixel-based classification context, the perceptron input layer contains features of each pixel. The hidden layer consists of neurons that utilize an activation function. Common activation functions are the step function, logistic sigmoid function, or rectified linear unit (ReLU). In this study, the ReLU was selected as the activation function for the implementations because of its fast and efficient training speed. The MLP is popular with other future urban growth simulation and prediction approaches [16,50,51].

Both RF and MLP have been selected for the prediction of future settlement growth of Surat for evaluating the STM as the training input. Additionally, the MLP was operated on HCM city and Abidjan to test the transferability of STM to other cities. The trained RF had the number of trees set to 100. Other parameters of the scikit-learn module were set to default [52]. The MLP in this study was designed with one input layer, one output layer, and one hidden layer, with the number of neurons set to 100.

#### 2.4. Implementation of STM-Based Urban/Settlement Growth Prediction

The STM contains the local spatial and temporal characteristics that will be learned by the ML algorithms during the training phase.

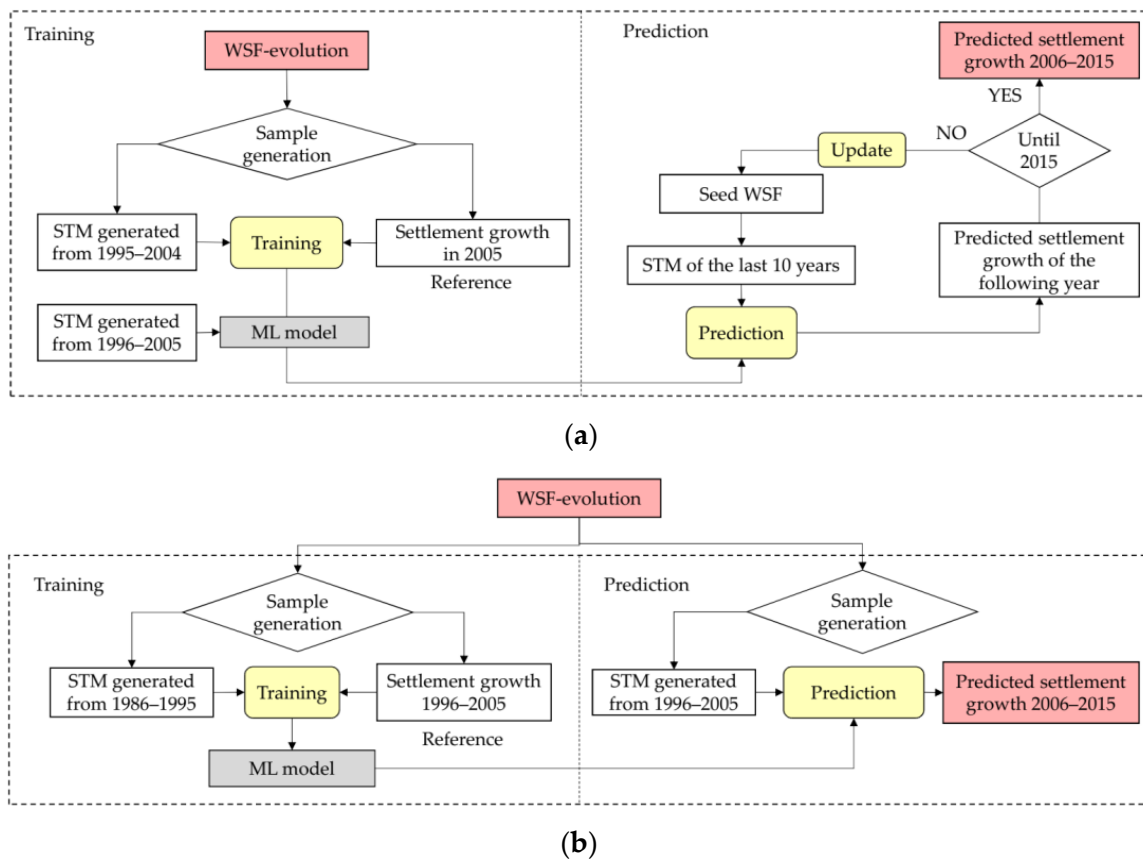
The objective of the following tests is to assess the performance of STM with future urban/built-settlement growth prediction. According to the aforementioned assumptions, the spatial and temporal constraints determined from the growth behavior of the past years that are stored in the STM can provide implicit attraction and resistance constraints. Consequently, the use of constraints provided by additional features, such as terrain information, local population density, distance to transportation infrastructures, and distance to economic centers, etc., for future urban/built-settlement growth prediction was deliberately not considered. For the evaluation, the global variables that could be used to control the total growth for a case city were also omitted.

An optimized training/prediction strategy can further support the implementation of the STM-based approach. The following two strategies were designed and compared for training and prediction: (1) A one-year strategy based on the full temporal resolution and one-year time steps for the prediction; (2) A ten-year strategy trained on the full temporal resolution and a ten-year time span for the prediction. In both strategies, the yearly binary settlement extent maps have been split into two subsets for training and prediction.

##### 2.4.1. One-Year Strategy

The basic idea of the one-year strategy is to use the existing settlement layers from the last ten years to generate the STM and to use it to predict settlement growth iteratively for the following years. The newly predicted settlement growth is then used to update the existing WSF-evolution. Afterwards, from the updated WSF-evolution, new STM will be generated, including the predicted result from the last time step. The iteration repeats until the updated WSF-evolution reaches the target year. In this study, the settlement growth from 2006 to 2015 was predicted.

The implementation of the one-year strategy is illustrated in Figure 6a, showing the use of the annual historical settlement layers of the WSF-evolution up to the year 2004. The settlement growth in the year 2005 was used as a binary reference layer. The STM was generated for each pixel based on the continuation period map of 1995–2004.



**Figure 6.** Workflow of two strategies: (a) workflow of the one-year strategy and (b) workflow of the ten-year strategy.

For the prediction phase, the iterative process continued annually until the prediction of the urban/built-settlement growth for 2015 was made. By stacking the predicted settlement growth layers of each year, the predicted settlement growth cumulation from 2006 to 2015 was obtained. The resulting layer was compared with the real urban/built-settlement growth from 2006 to 2015.

#### 2.4.2. Ten-Year Strategy

For the ten-year strategy, the time span of the reference data has been extended to ten years. The model based on this strategy predicts the future settlement growth with a single prediction and computes the future growth for ten years at once.

Figure 6b shows the workflow of the ten-year strategy, in which the STM for the training was generated based on historical urban layers from 1986 to 1995. The reference urban extent layer represents the urbanized pixel between 1996 and 2005. For the prediction phase, the STM generated with the annual settlement extent layers from 1996 to 2005 was used as the input, with the objective to predict the urban/built-settlement growth from 2006 to 2015.

#### 2.5. Baseline Approach: SLEUTH

With the widespread use, the maturity, and the long- and well-documented experiences of the SLEUTH model for urban growth modelling, it is predestinated to serve as the baseline for the performance evaluation of the STM approach. The model complies the following two coupled CA-models: one for urban growth modelling (UGM) and one for land-use change (LUC) prediction [33]. The SLEUTH model utilizes historical urban/built-settlement extent layers as references to predict the future, based on the assumption that the city will continue to grow at a similar pace to that between the historical training

layers. To adjust the changes of the growing speed, the SLEUTH model is coupled with a self-modification model. The urban/built-settlement growth prediction for 2006 to 2015 of the three case cities was conducted with the SLEUTH UGM, and the results were used as a baseline for comparing the results of the STM-based models.

### 2.6. Test Arrangement

The tests were arranged in three phases to test the STM's ability to serve as the performant input data structure for future settlement growth modelling. Firstly, to find out an optimal strategy for the training and prediction, a comparison test based on two strategies was applied on Surat and was conducted with the RF. Secondly, the ten-year strategy was applied with the MLP for Surat. The results from this test were compared with the results based on the ten-year strategy and the RF was used to test the STM's adaptability to the different ML algorithms. Lastly, to test the STM's transferability to different cities, the MLP was integrated with the ten-year strategy and was tested for HCMC and Abidjan. For the training process, several areas were selected from the reference map of each city to avoid the imbalance between the amount of developed and undeveloped pixels.

Additionally, the SLEUTH model was employed for settlement growth prediction for the selected cities from 2006 to 2015. The calibration of the model was implemented with the genetic algorithm (GA). Similar to the STM-based models, the urban extent layers were extracted from the WSF-evolution. The layer of 1985 was used as the seed layer for the calibration; the layers from 1990, 1995, 2000, and 2005 were used as the reference layers. In addition, the slope and hillshade layers were generated from Copernicus DEM [53], and the traffic layer was extracted from the global roads open access dataset (gROADS) [54]. Tests of SLEUTH were conducted on the business-as-usual (BaU) scenario.

Lastly, in order to further evaluate the performance of the STM-based model, future settlement growth predictions based on the WSF-evolution were conducted for Surat, HCMC, and Abidjan from 2016 to 2025 and were interpreted for conclusiveness.

All tests were implemented in a high spatial resolution of 30 m. The tests of the STM-based model were conducted with Python 3, using the ML algorithms provided by the scikit-learn package [52].

### 2.7. Evaluation

In the testing phases, the STM-based model predicted the settlement growth from 2006 to 2015. The actual settlement growth from 2006 to 2015, derived from the WSF-evolution, was used to evaluate the results. The receiving operative characteristic (ROC) and Cohen's kappa coefficient of the agreement were selected as the evaluation tools.

The ROC is an ideal evaluation tool for binary classification [55] and has been widely used for the evaluation of urban growth simulation [34,35,56]. The ROC curve is created by plotting and connecting the points of the true positive rate (TPR) against the false positive rate (FPR) concerning the outcomes from the different probabilities. The area under the curve (AUC) can be calculated from the ROC curve, ranging from 0 to 1. The optimal case would be a TPR of 1 and an FPR of 0, which leads to an AUC of 1. An AUC score equaling 0.5 is interpreted as a random classification. If the value is between 0.7 and 0.9, the classifier has a sufficient classification ability. A score higher than 0.9 indicates a high precision of the classified results. For the evaluation, the following nine different probability thresholds: 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% were selected for the decomposition of the results and the generation of the probability maps with a probability higher than the corresponding probability thresholds. These probability maps were laid over the real settlement growth from the corresponding years to create a ROC curve space.

Cohen's kappa coefficient is widely used for image classification tasks by revealing the agreement and shape similarity between the predicted settlement growth and the actual settlement growth [35,50,57]. The kappa coefficient ranges between  $-1$  and  $1$  and is interpreted as shown in Table 1.

**Table 1.** Kappa coefficients and the corresponding agreement level [8].

Kappa Coefficient	Agreement Level
<0	No agreement
0–0.2	Slight agreement
0.2–0.4	Fair agreement
0.4–0.6	Moderate agreement
0.6–0.8	Substantial agreement
0.8–1	Perfect agreement

Being under discussion [58], the kappa coefficient is controversial because it focuses on the agreement rather than the disagreement and establishes the comparison to a baseline of randomness. Therefore, in this study, the evaluation of the prediction quality was majorly based on AUC. The kappa coefficient was used for comparison and to select the best-performing probability threshold for illustrating the results. Additionally, a second method, the SLEUTH model, was introduced as a baseline.

The AUC values and kappa coefficients were calculated using the “sklearn.metrics” module from the scikit-learn package [52].

### 3. Results

#### 3.1. Surat

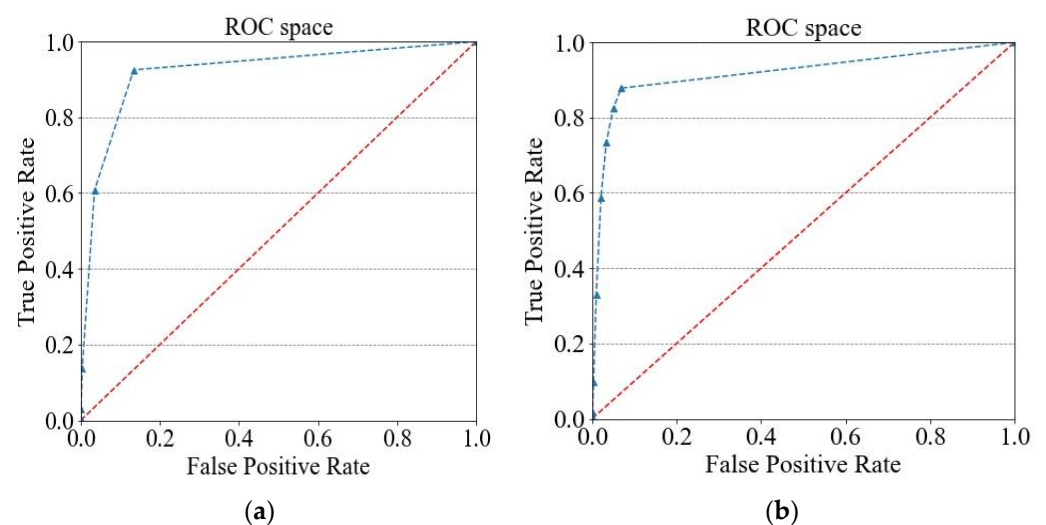
The AUC values from both of the strategies exceeded 0.9. Compared to the one-year strategy, the ten-year strategy resulted in a higher mean AUC. The statistical results of the ROC curves from two strategies are shown in Table 2.

**Table 2.** Statistical results of ROC curves from the two strategies applied to Surat.

	One-Year Strategy	Ten-Year Strategy
AUC	0.9216	0.9183
Mean AUC	0.5852	0.6827
Max AUC	0.8958	0.9045
TPR of Max AUC	0.9252	0.8781
FPR of Max AUC	0.1337	0.0691

Figure 7 illustrates the ROC curves of the two strategies.

In the second test phase, the MLP was then applied to Surat. Figure 8 shows the comparison between the results of MLP and RF.

**Figure 7.** ROC curves from the evaluation of the two different strategies based on the STM-based model on Surat from (a) the one-year strategy and (b) the ten-year strategy.

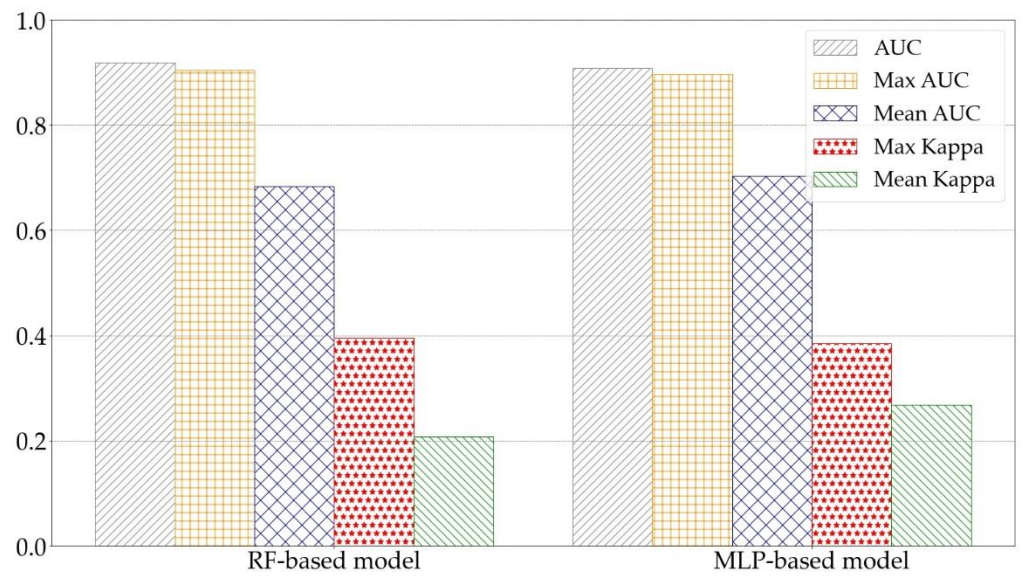


Figure 8. Comparison between the RF-based model and the MLP-based model.

Figure 9 illustrates the settlement growth prediction maps of the different strategies that are based on the probability maps with the highest kappa coefficient.

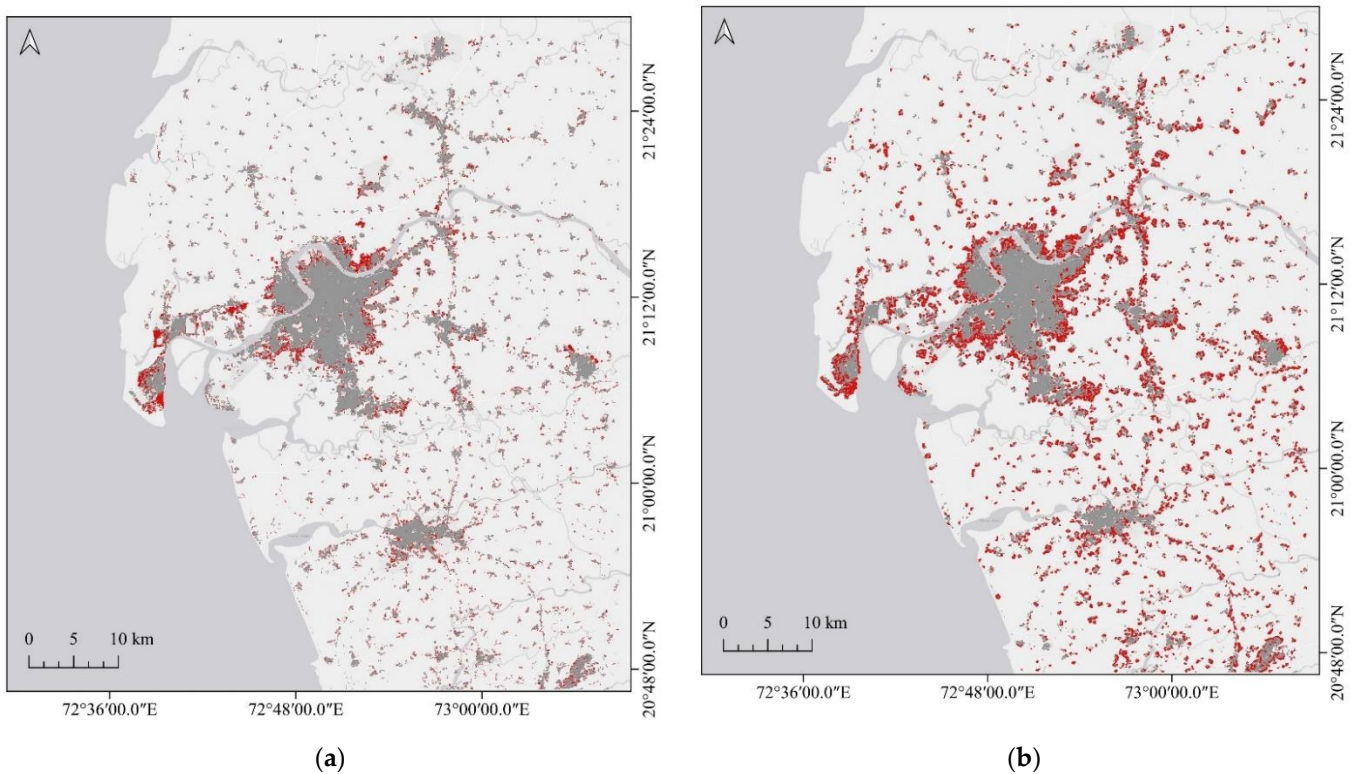
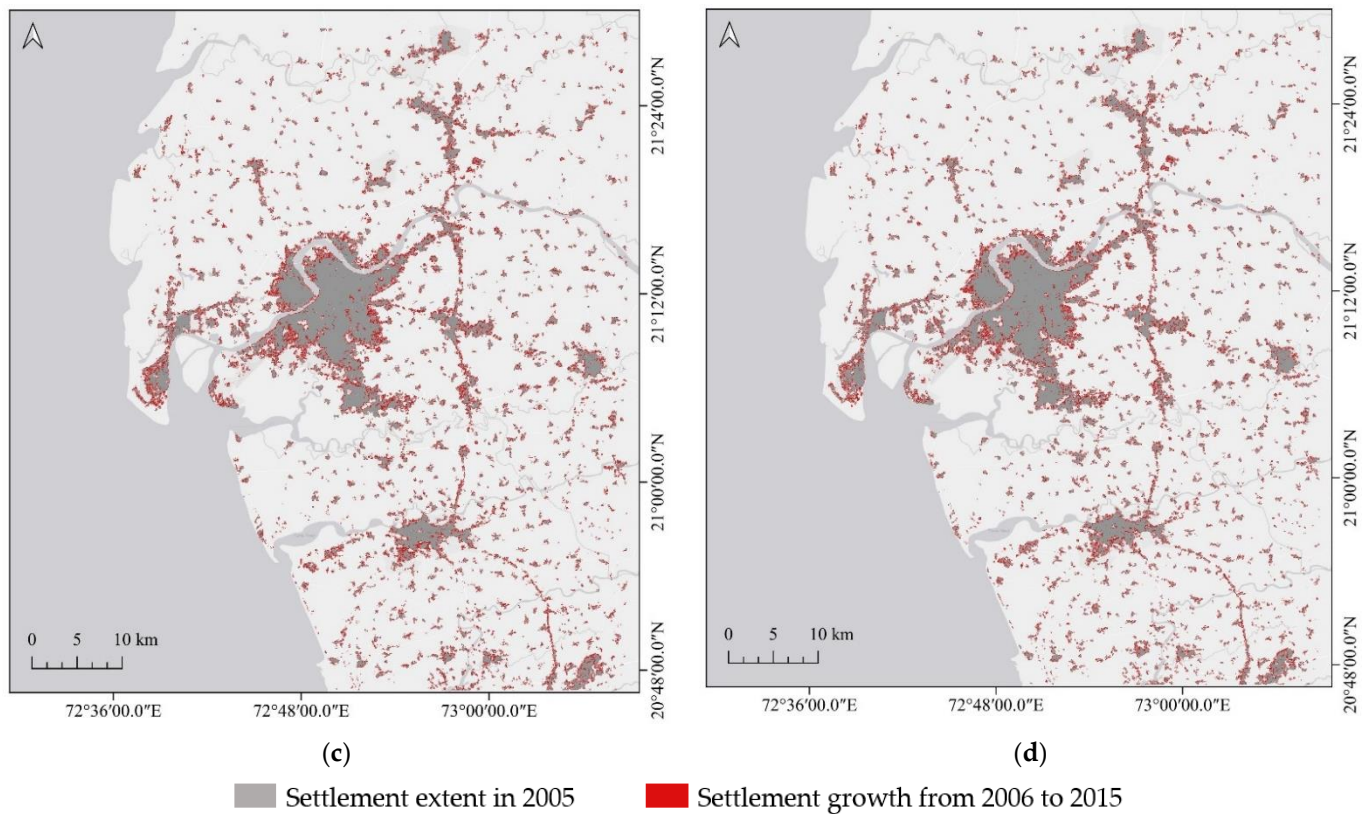


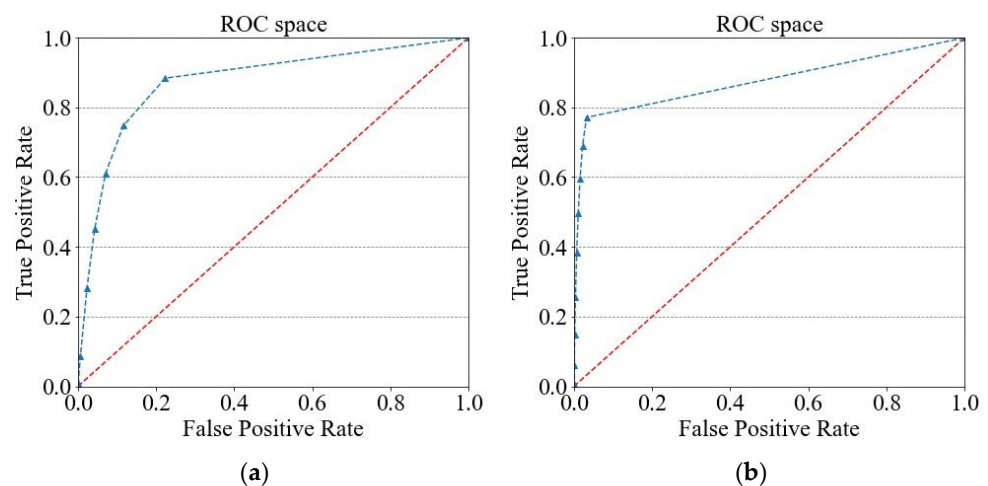
Figure 9. Cont.



**Figure 9.** Actual settlement growth map and predicted settlement growth maps (2006–2015) of Surat, based on the STM-based model: (a) actual settlement growth map of Surat, generated from the WSF-evolution; (b) predicted settlement growth map based on the one-year strategy and RF; (c) predicted settlement growth map based on the ten-year strategy and RF; (d) predicted settlement growth map based on the ten-year strategy and MLP.

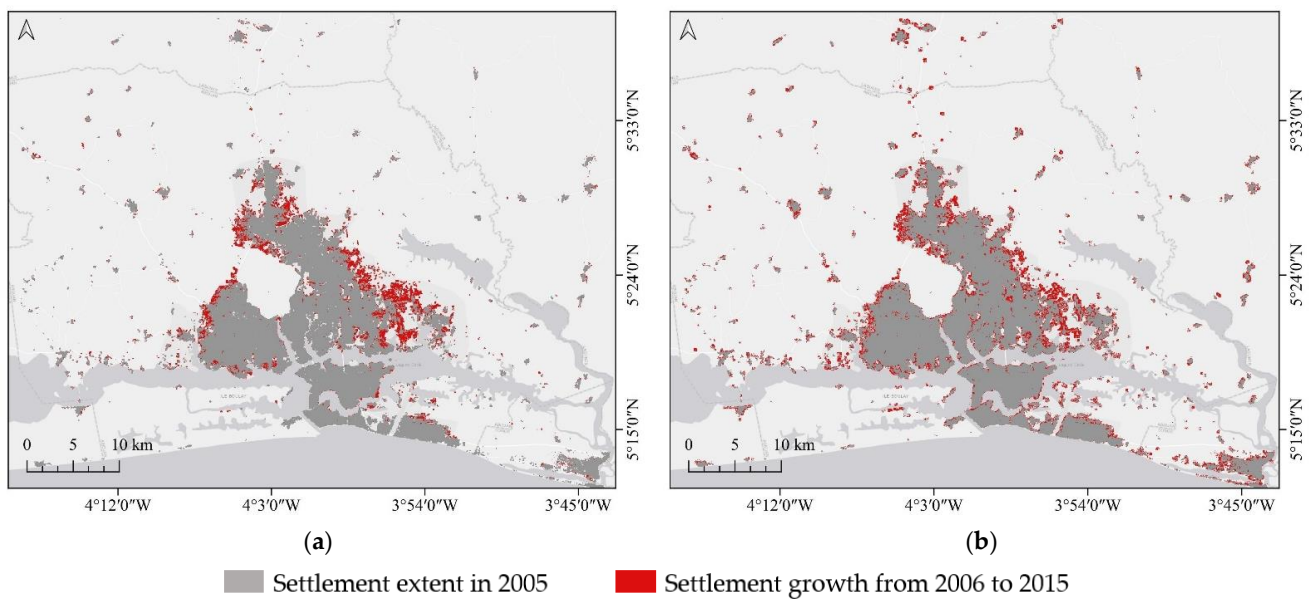
### 3.2. HCMC and Abidjan

In the third test phase, the STM-based model coupled with MLP was applied to HCMC and Abidjan with the ten-year strategy. The evaluations of both of the cities were carried out for the settlement growth from 2006 to 2015. The ROC curves are shown in Figure 10 and the statistical results of the ROC curve of these two cities are shown in Table 3.



**Figure 10.** ROC curves generated from predicted settlement growth probability maps of HCMC and Abidjan. Results were generated from STM-based model, implemented with the MLP and the ten-year strategy. (a) ROC curve of HCMC, (b) ROC curve of Abidjan.





**Figure 12.** Actual settlement growth map and predicted settlement growth map of Abidjan based on the STM-based model from 2006 to 2015: (a) actual settlement growth map of Abidjan; (b) predicted settlement growth map from based on the ten-year strategy and MLP.

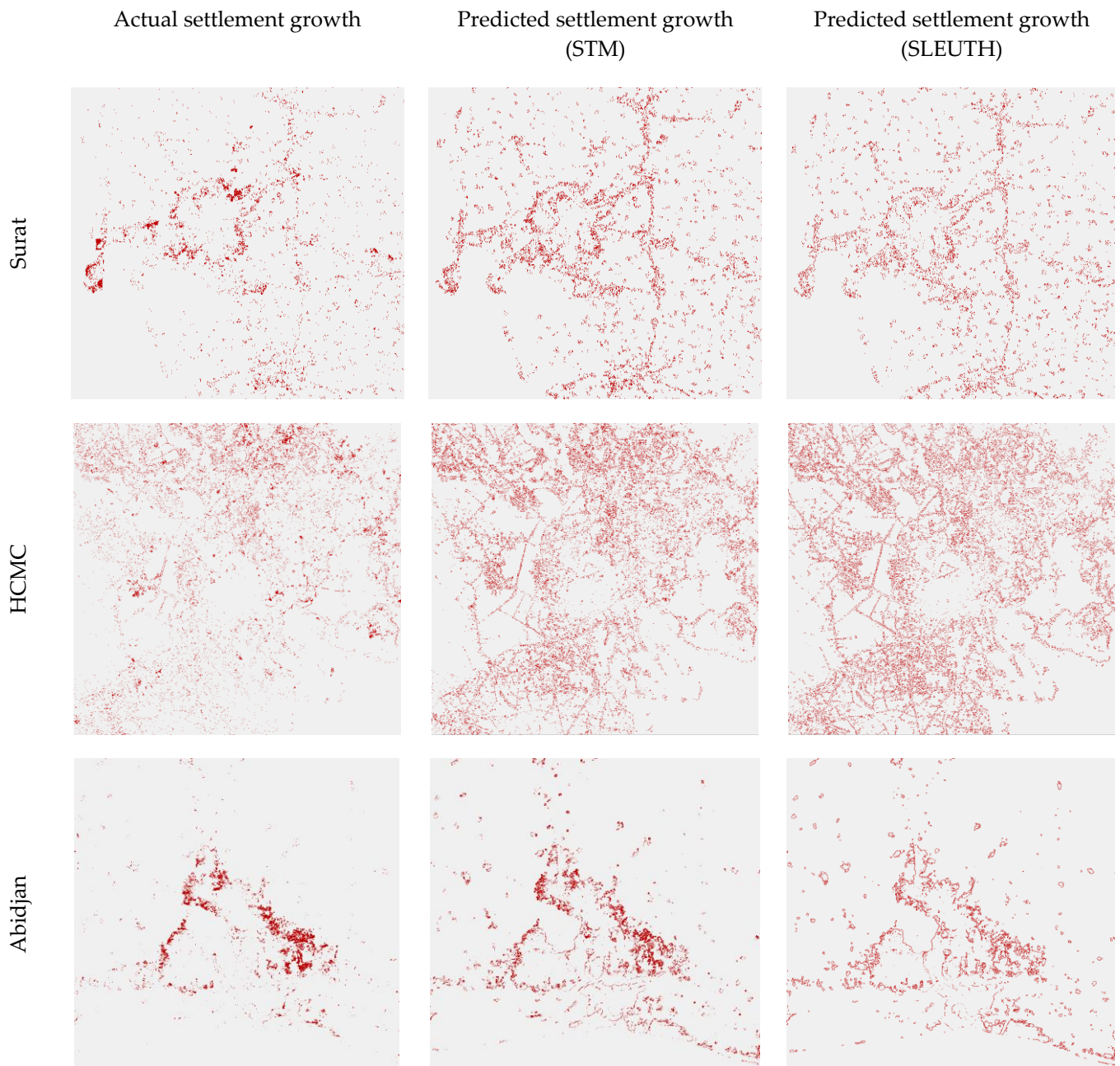
### 3.3. Comparison with SLEUTH

The results of the STM-based model integrated with the MLP for the three cities were compared with the results of the SLEUTH model. Based on the comparison, the differences in growth patterns of the two models were analyzed and discussed. In Figure 13, the predicted urban/built-settlement growth patterns from 2006 to 2015 of the STM-based model and the SLEUTH model are compared with the actual settlement growth. The AUC values and kappa coefficients were calculated and are displayed in Table 4, in order to further quantify the accuracy of the two prediction models.

**Table 4.** Comparison between results of the STM-based model and results of SLEUTH. Results from the STM-based model were generated with the ten-year strategy based on MLP.

City	Parameters	STM-Based Model	SLEUTH
Surat	AUC	0.9076	0.8851
	Kappa	0.3845	0.4391
	Probability threshold	0.40%	30%
HCMC	AUC	0.8753	0.8444
	Kappa	0.3182	0.4034
	Probability threshold	40%	30%
Abidjan	AUC	0.8745	0.8289
	Kappa	0.4524	0.4375
	Probability threshold	30%	10%



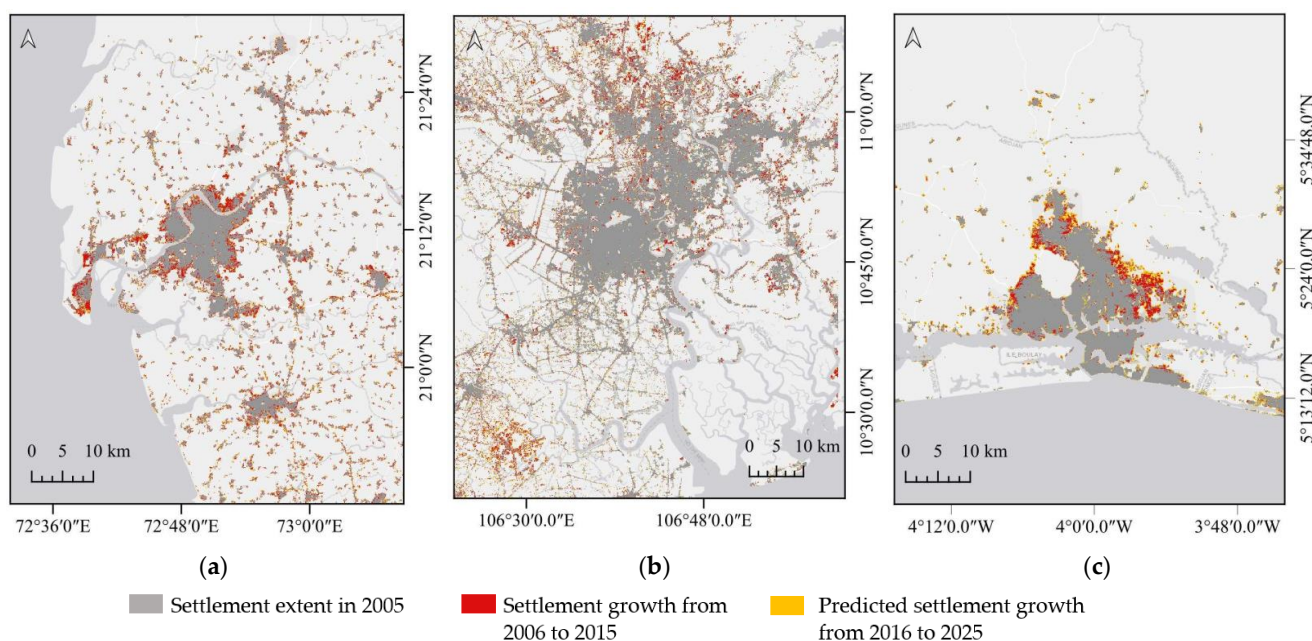


**Figure 13.** Comparison between the actual settlement growth patterns, the predicted settlement growth patterns from the STM-based model, and the predicted growth patterns from the SLEUTH model. Results are illustrated in binary images, with dark red pixels representing urbanized cells from 2006 to 2015.

#### 3.4. Settlement Growth Prediction to 2025

The ten-year strategy, in combination with MLP, was employed for the urban/built-settlement growth prediction to 2025.

The predicted results from 2016 to 2025 are shown in Figure 14. The statistical results are shown in Table 5.



**Figure 14.** Predicted settlement growth of the near future based on the STM-based model: (a) predicted growth of Surat; (b) predicted growth of HCMC; (c) predicted growth of Abidjan.

**Table 5.** Predicted settlement growth of the case cities from 2016 to 2025 (unit: number of pixels).

City	Settlement Pixel in 2015	Predicted Settlement Pixel	Predicted Growth	Predicted Growth Rate (%)
Surat	515,399	631,447	116,048	22.52%
HCMC	1,670,879	2,062,928	392,049	23.46%
Abidjan	406,340	456,712	50,372	12.40%

#### 4. Discussion

##### 4.1. Test Results

The STM-based approach demonstrated a good ability to predict the future urban/built-settlement growth. The AUC values of all of the tests exceeded 0.85, and the AUC value of Surat was above 0.9. The overall results indicated comprehensible shape similarities between the predicted settlement growth and the actual settlement growth (Figures 9, 11 and 12). The first test compared the two training strategies in the test city of Surat. The comparison between the two strategies demonstrated that the STM could be implemented variably. The results further showed that the ten-year strategy could achieve a better performance than the one-year strategy, because the one-year strategy presumably introduced an error-propagation effect. The second test investigated the use of STM with two different ML algorithms. Overall, both of the ML algorithms achieved promising results. The MLP showed a slightly better performance than the RF, when comparing the average kappa and the average AUC (Figure 8).

According to the annual settlement extent layers of the WSF-evolution, Surat showed a relatively regular growth pattern between 2006 and 2015, as most of the growth occurred along the city’s existing fringes. In comparison to the reference map, the STM-based model predicted a similar growth pattern. Notable growth along roads was also observed with the prediction. The visual interpretation of the future settlement growth prediction to 2025 showed reasonable results without any apparent anomalies. Compared with Surat, HCMC had a relatively heterogeneous settlement growth from 2006 to 2015. The reference map reveals the following two growing hotspots: the southeast part of Tan An, which is located in the south of HCMC, and the city agglomerations of Thuan An and Thu Dau

Mot, which are located in the north of HCMC. The STM-based model predicted a slightly over-estimated growth from 2006 to 2015. According to the STM-based model's forecast to 2025, the two hotspots in the north and south of the city will keep growing, especially Tan An, which seems reasonable when comparing it with the current development. Abidjan experienced irregular settlement growth from 2006 to 2015. According to the reference map, the city developed along edges on the eastern and western sides. Meanwhile, the southern parts of the city, where the city opens up towards the sea, hardly grew. The results from the STM-based model show a good performance on the settlement growth prediction from 2006 to 2015. The visual interpretation of the future settlement growth prediction of Abidjan to 2025 did not provide any inconclusive developments.

Overall, the STM-based prediction results achieved higher AUC values than the results of the SLEUTH model (Table 5). The SLEUTH model achieved higher kappa coefficients than the STM-based model in the case of Surat and HCMC. In contrast, the performance of the STM-based model in Abidjan was better than that of the SLEUTH. In the case of Surat, the STM-based model and the SLEUTH model both generated similar growth by predicting further growth along the existing fringes, while the results from the SLEUTH model were slightly underestimated. In the case of Abidjan, the STM-based model managed to predict irregular growth patterns along the city, while the predicted patterns from the SLEUTH model appeared to be evenly distributed (Figure 13). The SLEUTH model controlled the growth of the entire study area by utilizing the same set of coefficients, which resulted in a similar growing speed along the entire region. The STM-based model overcame this issue by generating the urbanization probability of each pixel independently.

#### 4.2. Advantages and Limitations

The results of the first test phase confirmed the STM's adaptability to the different ML algorithms, with RF and MLP achieving similar accuracies. Because the STM contained intrinsic information on attracting and resisting the growth based on the spatial and temporal constraints that were derived from past developments, it allowed the predictive models to predict the growth without the need of including external spatial information layers. For example, the excluded area layers were used in the SLEUTH model in order to provide spatial resistant factors for settlement development, e.g., water masks, natural preserved regions, and planned infrastructures. The results demonstrated the STM's ability to detect such areas without having them as input data. In all of the tests, predicted future urban expansion into water bodies was barely found. In HCMC (Figure 11), minimal growth was predicted at the airport of Tan-Son-Nhat. In the case of Abidjan (Figure 12), the growth towards the National Park Banco has been avoided. Aside from the excluded areas, the road network is also an essential factor that can potentially affect the future settlement growth by attracting settlement growth. The STM-based results shared similar road-attracted growth patterns as the results from the SLEUTH model, with only the latter having road networks as input layers. At the same time, the STM condensed the historical settlement extent layers into one continuation period layer and further reduced the data redundancy. Without considering the external layers, the STM-based model utilized only the continuation period layer as the input. In the test case of Surat, the continuation period layer took up 33 MB. In comparison, the annual settlement extent layers from the same time period took up 490 MB.

Nonetheless, there are some shortcomings of STM-based approaches that could be observed. Based on the assumption that the information from the neighboring pixels is steering the probability of the urbanization of the target pixel, the STM-based models are not able to predict spontaneous urban/built-settlement growth. Therefore, without any further modification, it is challenging for STM-based models to introduce new sites that do not show adjacent building activities in administrative units with specific planning regulations. Another challenge is the definition of a meaningful window size for the STM. An adaptive solution has to be identified based on the resolution of the input images and the observed land cover changes. For the case cities of this study, different window sizes

have been tested empirically for the one-year and the ten-year strategies. For the examined cities, the resulting optimal window size for the one-year strategy is five and nine for the ten-year strategy. Another general limitation that all of the predictive models share is the total time span of data availability. For this study, the WSF-evolution product was available only from 1985 to 2015, which restricts the modelling and evaluation of the model outputs with real measurements to certain constellations. With the continuation of the EO missions, this challenge will gradually be reduced. The STM depends on continuous time series observations. This leads to a potential issue when data gaps exist in the time series. The WSF that was used in this study closed the gaps mainly resulting from a low coverage and cloud cover in the Landsat 5/7. When using other datasets, measures to overcome the issue of missing data have to be considered [59,60].

## 5. Conclusions

This paper introduced the spatio-temporal matrix (STM), which is a novel pixel-based approach for data prediction that is based on long-term EO-based time series data products.

The approach was applied and tested for urban/settlement growth prediction based on the WSF-evolution data, which was an EO-based time series data product that outlined the global settlement extents annually, from 1985 to 2015, at a high spatial resolution (30 m). The following three cities were selected as test sites: Ho-Chi-Minh City (HCMC) in Vietnam, Abidjan in the Ivory Coast, and Surat in India.

RF and MLP were employed in order to test STM's adaptability to different ML algorithms. The models were implemented with two training and predicting strategies that were based on different time spans, the one-year and ten-year time span strategies. The urban/settlement growth from 2006 to 2015 was predicted by the approach, using historical urban extent layers from the WSF-evolution and was compared with the real growth of that period for evaluation. The results were evaluated with AUC that was generated from the ROC and Cohen's kappa index. In addition, the SLEUTH model was selected as the baseline method for comparison. Furthermore, the STM-based model was utilized in order to predict the urban/settlement growth of the case cities from 2016 to 2025, based on STM that was generated from WSF-evolution from 2006 to 2015.

The AUC values of all of the tests exceeded 0.85, proving the good ability of the STM-based model to predict the future settlement growth. In the evaluation setting, the best results were achieved by the combination of MLP and ten-year strategy. The results of the predictive model that were based on this combination were further compared with the predicted results of the SLEUTH model. The AUC that was calculated from the STM-based approach of all three of the case cities surpassed that of the SLEUTH model, achieving 0.91 for Surat, 0.88 for HCMC, and 0.87 for Abidjan. In comparison, the SLEUTH model achieved an AUC of 0.89 for Surat, 0.84 for HCMC, and 0.83 for Abidjan (Table 5). The STM-based model was able to predict irregular growth patterns though its local pixel-based design, while the predictions of the SLEUTH model appeared to be evenly distributed (Figure 13).

According to the predicted results, the growth rate of HCMC from 2016 to 2025 will reach 23.46%. Surat can expect a growth rate of 22.52%. In the same time period, Abidjan will expand by 12.4%. The predicted urban/settlement growth map also indicates the growing hot spots of the three cities (Figure 14).

The STM brings several assets to the growth prediction based on the time-series products, as follows:

1. The utilization of the STM approach is highly flexible as it can be easily integrated with different ML algorithms. In this study, RF and MLP were selected, were integrated with STM, and were tested for Surat. Both of the ML models performed well, with similar accuracies;
2. STM makes full use of the spatio-temporal characteristics of the EO-based time series products by condensing the annual information into one continuation period map;

3. The matrix approach reduces the data volume and redundancy by integrating discrete EO-based time-series information into one feature vector;
4. The presented results show the self-sufficiency of the STM-based approaches by reducing the dependence on additional data layers. Without providing the corresponding layers, the settlement growth that was predicted by the STM-based models barely expanded in the areas that were restricted for growth or were beyond natural barriers.

Further investigations are needed in order to understand the full potential of the methodological approach. Firstly, the effect of different window sizes will be investigated. Secondly, the STM-based model can be enriched with additional social-economic features, e.g., population density maps, distance to existing cities, and distance to economic centers for the settlement growth prediction in order to improve the robustness of the approach. In order to overcome irregular planning rules and separated construction, an external weighting layer can be included as a controlling factor for the entire region, supporting the STM-based predictions. In addition, the ability of STM to be integrated with different growth scenarios will be evaluated.

The EO provides a wealth of time series data in the urban domain. The STM-based approaches have the potential to contribute to the predictive modelling of other land surface dynamics, as long as their general patterns and trends fit the general assumptions of the method. Allowing increasing and decreasing trends with adaptive continuation maps will enable applications in deforestation, shrub encroachment, and soil and landscape degradation.

**Author Contributions:** Conceptualization, Z.W., F.B. and C.K.; methodology, Z.W. and F.B.; software, Z.W.; validation, Z.W., F.B. and M.M.; formal analysis, Z.W. and J.K.; investigation, Z.W., F.B., J.H. and J.K.; data curation, Z.W., M.M., T.E. and T.H.; writing—original draft preparation, Z.W.; writing—review and editing, F.B., J.K., J.H., T.H., M.M., T.E. and C.K.; visualization, Z.W. and J.H.; supervision, F.B., J.H. and C.K.; funding acquisition, C.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The WSF-evolution dataset is publicly available and can be downloaded from <https://geoservice.dlr.de/web/maps> (accessed on 3 February 2021).

**Acknowledgments:** The authors would like to thank Stefanie Feuerstein for preparing the water mask layers for the SLEUTH model set-up and Michael Wang for the language checking and editing.

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

## References

1. *Landsat—Earth Observation Satellites*; Fact Sheet; U.S. Geological Survey: Reston, VA, USA, 2016; Volume 2015–3081, p. 4.
2. USGS EROS Archive—Advanced Very High Resolution Radiometer—AVHRR. Available online: <https://www.usgs.gov/centers/eros/science/usgs-eros-archive-advanced-very-high-resolution-radiometer-avhrr> (accessed on 1 November 2021).
3. National Aeronautics and Space Administration (NASA) MODIS Moderate Resolution Imaging Spectroradiometer. Available online: <https://modis.gsfc.nasa.gov/about/> (accessed on 15 March 2022).
4. Klein, I.; Gessner, U.; Dietz, A.J.; Kuenzer, C. Global WaterPack—A 250 m Resolution Dataset Revealing the Daily Dynamics of Global Inland Water Bodies. *Remote Sens. Environ.* **2017**, *198*, 345–362. [[CrossRef](#)]
5. Dietz, A.J.; Kuenzer, C.; Dech, S. Global SnowPack: A New Set of Snow Cover Parameters for Studying Status and Dynamics of the Planetary Snow Cover Extent. *Remote Sens. Lett.* **2015**, *6*, 844–853. [[CrossRef](#)]
6. Leon-Tavares, J.; Roujean, J.-L.; Smets, B.; Wolters, E.; Toté, C.; Swinnen, E. Correction of Directional Effects in VEGETATION NDVI Time-Series. *Remote Sens.* **2021**, *16*, 1130. [[CrossRef](#)]
7. European Commission; Joint Research Centre. *GHS-L Data Package 2019: Public Release GHS P2019*; Publications Office of the European Union: Luxembourg, 2019.
8. Marconcini, M.; Metz-Marconcini, A.; Üreyen, S.; Palacios-Lopez, D.; Hanke, W.; Bachofer, F.; Zeidler, J.; Esch, T.; Gorelick, N.; Kakarla, A.; et al. Outlining Where Humans Live, the World Settlement Footprint 2015. *Sci. Data* **2020**, *7*, 242. [[CrossRef](#)]
9. Damien, S.M.; Mark, F. *MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 2019*; NASA EOSDIS Land Processes DAAC: Sioux Falls, SD, USA, 2019.

10. Koehler, J.; Kuenzer, C. Forecasting Spatio-Temporal Dynamics on the Land Surface Using Earth Observation Data—A Review. *Remote Sens.* **2020**, *12*, 3513. [[CrossRef](#)]
11. Demisse, G.B.; Tadesse, T.; Bayissa, Y.; Atnafu, S.; Argaw, M.; Nedaw, D. Vegetation Condition Prediction for Drought Monitoring in Pastoralist Areas: A Case Study in Ethiopia. *Int. J. Remote Sens.* **2018**, *39*, 4599–4615. [[CrossRef](#)]
12. Silva, E.A.; Clarke, K.C. Calibration of the SLEUTH Urban Growth Model for Lisbon and Porto, Portugal. *Comput. Environ. Urban Syst.* **2002**, *26*, 525–552. [[CrossRef](#)]
13. Gao, J.; O'Neill, B.C. Data-Driven Spatial Modeling of Global Long-Term Urban Land Development: The SELECT Model. *Environ. Model. Softw.* **2019**, *119*, 458–471. [[CrossRef](#)]
14. Cohen, B. Urbanization in Developing Countries: Current Trends, Future Projections, and Key Challenges for Sustainability. *Technol. Soc.* **2006**, *28*, 63–80. [[CrossRef](#)]
15. Zhang, Y.-J.; Yi, W.-C.; Li, B.-W. The Impact of Urbanization on Carbon Emission: Empirical Evidence in Beijing. *Energy Procedia* **2015**, *75*, 2963–2968. [[CrossRef](#)]
16. Lee, H.-Y.; Jang, K.M.; Kim, Y. Energy Consumption Prediction in Vietnam with an Artificial Neural Network-Based Urban Growth Model. *Energies* **2020**, *13*, 4282. [[CrossRef](#)]
17. Deng, X.; Huang, J.; Rozelle, S.; Zhang, J.; Li, Z. Impact of Urbanization on Cultivated Land Changes in China. *Land Use Policy* **2015**, *45*, 1–7. [[CrossRef](#)]
18. Bosseler, B.; Salomon, M.; Schlüter, M.; Rubinato, M. Living with Urban Flooding: A Continuous Learning Process for Local Municipalities and Lessons Learnt from the 2021 Events in Germany. *Water* **2021**, *13*, 2769. [[CrossRef](#)]
19. Kulp, S.A.; Strauss, B.H. New Elevation Data Triple Estimates of Global Vulnerability to Sea-Level Rise and Coastal Flooding. *Nat. Commun.* **2019**, *10*, 4844. [[CrossRef](#)]
20. Zhang, X.; Song, J.; Peng, J.; Wu, J. Landslides-Oriented Urban Disaster Resilience Assessment—A Case Study in Shenzhen, China. *Sci. Total Environ.* **2019**, *661*, 95–106. [[CrossRef](#)]
21. Endreny, T.; Santagata, R.; Perna, A.; Stefano, C.D.; Rallo, R.F.; Ulgiati, S. Implementing and Managing Urban Forests: A Much Needed Conservation Strategy to Increase Ecosystem Services and Urban Wellbeing. *Ecol. Model.* **2017**, *360*, 328–335. [[CrossRef](#)]
22. Donnay, J.-P.; Barnsley, M.J.; Longley, P.A. *Remote Sensing and Urban Analysis: GISDATA 9*; CRC Press: Boca Raton, FL, USA, 2014; ISBN 978-1-4822-6811-9.
23. Corbane, C.; Florczyk, A.; Pesaresi, M.; Politis, P.; Syrris, V. GHS Built-up Grid, Derived from Landsat, Multitemporal (1975–1990–2000–2014)—OBSOLETE RELEASE. European Commission, Joint Research Centre (JRC) [Dataset]. Available online: <http://data.europa.eu/89h/jrc-ghsl-10007> (accessed on 21 September 2020). [[CrossRef](#)]
24. Gong, P.; Li, X.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Liu, X.; Xu, B.; Yang, J.; Zhang, W.; et al. Annual Maps of Global Artificial Impervious Area (GAIA) between 1985 and 2018. *Remote Sens. Environ.* **2020**, *236*, 111510. [[CrossRef](#)]
25. Huang, X.; Li, J.; Yang, J.; Zhang, Z.; Li, D.; Liu, X. 30 m Global Impervious Surface Area Dynamics and Urban Expansion Pattern Observed by Landsat Satellites: From 1972 to 2019. *Sci. China Earth Sci.* **2021**, *64*, 1922–1933. [[CrossRef](#)]
26. Marconcini, M.; Metz-Marconcini, A.; Esch, T.; Gorelick, N. Understanding Current Trends in Global Urbanisation—The World Settlement Footprint Suite. *GI Forum* **2021**, *1*, 33–38. [[CrossRef](#)]
27. Chang, X.; Zhang, F.; Cong, K.; Liu, X. Scenario Simulation of Land Use and Land Cover Change in Mining Area. *Sci. Rep.* **2021**, *11*, 12910. [[CrossRef](#)] [[PubMed](#)]
28. Han, H.; Yang, C.; Song, J. Scenario Simulation and the Prediction of Land Use and Land Cover Change in Beijing, China. *Sustainability* **2015**, *7*, 4260–4279. [[CrossRef](#)]
29. Srichaichana, J.; Trisurat, Y.; Ongsomwang, S. Land Use and Land Cover Scenarios for Optimum Water Yield and Sediment Retention Ecosystem Services in Klong U-Tapao Watershed, Songkhla, Thailand. *Sustainability* **2019**, *11*, 2895. [[CrossRef](#)]
30. Liu, X.; Liang, X.; Li, X.; Xu, X.; Ou, J.; Chen, Y.; Li, S.; Wang, S.; Pei, F. A Future Land Use Simulation Model (FLUS) for Simulating Multiple Land Use Scenarios by Coupling Human and Natural Effects. *Landsc. Urban Plan.* **2017**, *168*, 94–116. [[CrossRef](#)]
31. Li, X.; Gong, P. Urban Growth Models: Progress and Perspective. *Sci. Bull.* **2016**, *61*, 1637–1650. [[CrossRef](#)]
32. Musa, S.I.; Hashim, M.; Reba, M.N.M. A Review of Geospatial-Based Urban Growth Models and Modelling Initiatives. *Geocarto Int.* **2017**, *32*, 813–833. [[CrossRef](#)]
33. Clarke, K.C.; Hoppen, S.; Gaydos, L. A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area. *Environ. Plan. B Plan. Des.* **1997**, *24*, 247–261. [[CrossRef](#)]
34. Wu, X.; Hu, Y.; He, H.S.; Bu, R.; Onsted, J.; Xi, F. Performance Evaluation of the SLEUTH Model in the Shenyang Metropolitan Area of Northeastern China. *Environ. Model. Assess.* **2009**, *14*, 221–230. [[CrossRef](#)]
35. Sakieh, Y.; Amiri, B.J.; Danekar, A.; Feghhi, J.; Dezhkam, S. Simulating Urban Expansion and Scenario Prediction Using a Cellular Automata Urban Growth Model, SLEUTH, through a Case Study of Karaj City, Iran. *J. Hous. Built Environ.* **2015**, *30*, 591–611. [[CrossRef](#)]
36. Chaudhuri, G.; Clarke, K.C. Modeling an Indian Megalopolis—A Case Study on Adapting SLEUTH Urban Growth Model. *Comput. Environ. Urban Syst.* **2019**, *77*, 101358. [[CrossRef](#)]
37. Zhou, Y.; Varquez, A.C.G.; Kanda, M. High-Resolution Global Urban Growth Projection Based on Multiple Applications of the SLEUTH Urban Growth Model. *Sci. Data* **2019**, *6*, 34. [[CrossRef](#)] [[PubMed](#)]
38. Clarke, K.C.; Johnson, J.M. Calibrating SLEUTH with Big Data: Projecting California's Land Use to 2100. *Comput. Environ. Urban Syst.* **2020**, *83*, 101525. [[CrossRef](#)]

39. Liu, H.; Zhou, Q. Developing Urban Growth Predictions from Spatial Indicators Based on Multi-Temporal Images. *Comput. Environ. Urban Syst.* **2005**, *29*, 580–594. [[CrossRef](#)]
40. Wang, C.; Lei, S.; Elmore, A.J.; Jia, D.; Mu, S. Integrating Temporal Evolution with Cellular Automata for Simulating Land Cover Change. *Remote Sens.* **2019**, *11*, 301. [[CrossRef](#)]
41. Schneider, R.; Vicedo-Cabrera, A.; Sera, F.; Masselot, P.; Stafoggia, M.; de Hoogh, K.; Kloog, I.; Reis, S.; Vieno, M.; Gasparrini, A. A Satellite-Based Spatio-Temporal Machine Learning Model to Reconstruct Daily PM<sub>2.5</sub> Concentrations across Great Britain. *Remote Sens.* **2020**, *12*, 3803. [[CrossRef](#)]
42. Wang, H.; Guo, J.; Zhang, B.; Zeng, H. Simulating Urban Land Growth by Incorporating Historical Information into a Cellular Automata Model. *Landscape Urban Plan.* **2021**, *214*, 104168. [[CrossRef](#)]
43. Li, X.; Zhou, Y.; Chen, W. An Improved Urban Cellular Automata Model by Using the Trend-Adjusted Neighborhood. *Ecol. Process.* **2020**, *9*, 28. [[CrossRef](#)]
44. Johnson, B.A.; Estoque, R.C.; Li, X.; Kumar, P.; Dasgupta, R.; Avtar, R.; Magcale-Macandog, D.B. High-Resolution Urban Change Modeling and Flood Exposure Estimation at a National Scale Using Open Geospatial Data: A Case Study of the Philippines. *Comput. Environ. Urban Syst.* **2021**, *90*, 101704. [[CrossRef](#)]
45. Li, X.; Zhou, Y.; Hejazi, M.; Wise, M.; Vernon, C.; Iyer, G.; Chen, W. Global Urban Growth between 1870 and 2100 from Integrated High Resolution Mapped Data and Urban Dynamic Modeling. *Commun. Earth Environ.* **2021**, *2*, 201. [[CrossRef](#)]
46. United Nations Department of Economic and Social Affairs. *World Urbanization Prospects 2018: Highlights*; UN: Geneva, Switzerland, 2019; ISBN 978-92-1-004313-7.
47. Abidjan, Cote D'Ivoire Metro Area Population 1950–2022. Available online: <https://www.macrotrends.net/cities/21602/abidjan/population> (accessed on 7 February 2022).
48. Tobler, W.R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* **1970**, *46*, 234. [[CrossRef](#)]
49. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
50. Aarathi, A.D.; Gnanappazham, L. Urban Growth Prediction Using Neural Network Coupled Agents-Based Cellular Automata Model for Sriperumbudur Taluk, Tamil Nadu, India. *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 353–362. [[CrossRef](#)]
51. Guan, Q.; Wang, L.; Clarke, K.C. An Artificial-Neural-Network-Based, Constrained CA Model for Simulating Urban Growth. *Cartogr. Geogr. Inf. Sci.* **2005**, *32*, 369–380. [[CrossRef](#)]
52. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
53. European Space Agency, Sinergise. Copernicus Global Digital Elevation Model. Distributed by Open Topography. 2021. Available online: <https://portal.opentopography.org/datasetMetadata?otCollectionID=OT.032021.4326.1> (accessed on 8 January 2021).
54. Center for International Earth Science Information Network—CIESIN—Columbia University, and Information Technology Outreach Services—ITOS—University of Georgia; Global Roads Open Access Data Set, Version 1 (gROADSv1); NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2013. [[CrossRef](#)]
55. Pontius, R.G.; Schneider, L.C. Land-Cover Change Model Validation by an ROC Method for the Ipswich Watershed, Massachusetts, USA. *Agric. Ecosyst. Environ.* **2001**, *85*, 239–248. [[CrossRef](#)]
56. Puertas, O.L.; Henríquez, C.; Meza, F.J. Assessing Spatial Dynamics of Urban Growth Using an Integrated Land Use Model. Application in Santiago Metropolitan Area, 2010–2045. *Land Use Policy* **2014**, *38*, 415–425. [[CrossRef](#)]
57. Martínez-Vega, J.; Díaz, A.; Nava, J.M.; Gallardo, M.; Echavarría, P. Assessing Land Use-Cover Changes and Modelling Change Scenarios in Two Mountain Spanish National Parks. *Environments* **2017**, *4*, 79. [[CrossRef](#)]
58. Pontius, R.G.; Millones, M. Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [[CrossRef](#)]
59. Mayr, S.; Kuenzer, C.; Gessner, U.; Klein, I.; Rutzinger, M. Validation of Earth Observation Time-Series: A Review for Large-Area and Temporally Dense Land Surface Products. *Remote Sens.* **2019**, *11*, 2616. [[CrossRef](#)]
60. Hirabayashi, S.; Kroll, C.N. Single Imputation Method of Missing Air Quality Data for I-Tree Eco Analyses in the Conterminous United States. 2017. Available online: [https://www.itreetools.org/documents/51/Single\\_imputation\\_method\\_of\\_missing\\_air\\_quality\\_data\\_for\\_i-Tree\\_Eco\\_analyses\\_in\\_the\\_conterminous\\_United\\_States.pdf](https://www.itreetools.org/documents/51/Single_imputation_method_of_missing_air_quality_data_for_i-Tree_Eco_analyses_in_the_conterminous_United_States.pdf) (accessed on 21 September 2020).