# Abstract

Models are widely used to simulate real-world urban dynamics, but many studies have oversimplified the spatial complexity of driving factors. Advances in deep learning technology enable complex urban dynamic patterns to be more precisely captured. In this study, we selected the U-Net algorithm to assimilate historical urban development, described the pattern extraction process, validated the model with a reference map, and applied the model to predict the urban layout of 2030 in the North China Plain. The results showed that 1) U-Net can gradually abstract high-level spatial features and refined these patterns to precise urban development shapes, 2) the FoM (Figure of Merit) of the simulation map was close to previous studies that went through a lengthy calibration process, and 3) the landscape patterns between the simulation and the reference map were well aligned. U-Net was able to learn complex urban development patterns such as neighborhood effects and linear development, which supplement CA models to integrate spatial features to project future urban development.

# 1. Introduction

Urbanization is a complex process that is influenced by a range of social, cultural, economic, geographic, environmental, and political factors (Shaw et al. 2020; Yeh and Chen 2020; Kipfer 2018; Fan et al. 2018). Understanding future urban patterns and extent is essential to sustainable development because urbanization plays a key role in population change, economic growth, and climate change (Zheng et al. 2017; Newland et al. 2018). Modeling city dynamics to a future extent has been practiced in many regions of the world for mitigating air pollution (Fan et al. 2020), habitat destruction (Planillo et al. 2021), and loss of arable land (Qiu et al. 2020). Some urban modeling studies have focussed on participatory modeling and scenario analysis for engaging stakeholders and experts in the modeling process to balance competing interests and facilitate the co-learning processes on future urban development (Clarke and Johnson 2020; Mansour et al. 2020; Peng et al. 2020). However, most urban modeling studies have primarily focused on the accurate prediction of the future layout and patterns of urban areas based on historical urban dynamics and changes in key driving forces under various scenarios (Gantumur et al. 2020; Shafizadeh-Moghadam et al. 2017). Despite significant recent advances in urban modeling, accurately capturing the complex spatial dynamics, patterns, and stochasticities of cities remains a significant challenge (Liu et al. 2021).

Cellular Automata (CA) have been widely used to model future urbanization patterns because of their ability to capture real-world urban patterns (Tong and Feng 2020; Li et al. 2017). CA models are composed of transition suitability, neighborhood status, constraint variables, and stochastic factors (Roodposhti et al. 2020; Wang et al. 2021a). The transition suitability refers to the rescaled biophysical, geographic, and socio-economic driving factors that encourage or hinder the urbanization process (Feng and Tong 2020). Neighborhood status reflects the amount of urban or other land-use occurring in the immediate vicinity of each cell and is characterized according to different structures, sizes, and weights (Yu et al. 2021; Roodposhti et al. 2020). The constraint variables and stochastic factors forbid or randomize the future urban development from happening (Zhai et al. 2020) (e.g., no urban land can be constructed on water, and the grassland has a 1% chance to be converted to built-up land). The transition rules are a set of parameters that control the rescaling of the transition suitability factors, the configuration of the neighborhood, the constraints, and the stochastic factors, and then combine these elements into a spatial layer defining the probability of each cell becoming urbanized in the future. While transition rules have typically been derived by trial and error or expert knowledge, they are increasingly derived automatically to achieve the highest predictive accuracy. The automatic rule extraction includes a suite of regression and machine learning (ML) based methods such as logistic regression (Mustafa et al. 2018), support vector machine (Kafy et al. 2021), tree-based method (Shafizadeh-Moghadam et al. 2017), neural networks (Gantumur et al. 2020), heuristic methods (Carneiro and Oliveira 2013), and dictionary of trusted rules (Roodposhti et al. 2019). Although the flexibility of CA-based models with their large array of parameter settings makes them ideal for participatory-based scenario modeling exercises, the difficulty in calibrating the many parameter choices—still largely a manual process of trial and error—challenges their ability to mimic complex urban dynamics and accurately capture future urban patterns.

More recent studies have adopted different techniques, such as geographical zoning, context integration, and innovative algorithms, to increase the predictive accuracy of urban land-use modeling. For example, some studies subset their study area into separate regions, allowing independent transition rule sets to be constructed to align with the specific condition in each zone (Qian et al. 2020; Xia and Zhang 2021). Many studies incorporate the shape and texture index to reflect the neighborhood spatial configurations of urban dynamics (Zhai et al. 2020; Ruiz Hernandez and Shi 2018). Wang et al. (2021a) incorporate historical urban development as momentum to simulate the future urban layout; Peng et al. (2020) integrate evolutionary and swarm algorithms to mimic urban dynamics. Despite having successfully constructed transition rules, these studies presented mismatches compared to real-world urbanization: 1) introducing subjectivities via prescribed parameters, and 2) oversimplifying the spatial heterogeneity of driving factors. Distances to and spatial configurations of geographical factors are key to urban development in the real world. The distances used in these studies, however, are measured by pre-set decay functions, and subjectivities are inevitably raised given the gap between imperial knowledge and real-world urban dynamics. The spatial heterogeneity of the urbanization process is oversimplified to a single shape or texture index that hardly captures the complete pattern information within the neighborhood scope, let along the large-scale spatial features, e.g., the layout of the whole built city area, that drives the urban development in the real world. Therefore, the advances for existing urban simulation models should be 1) capturing the large-scale spatial heterogeneity of driving factors, and 2) assimilating geographical variables unattained to reduce subjectivity.

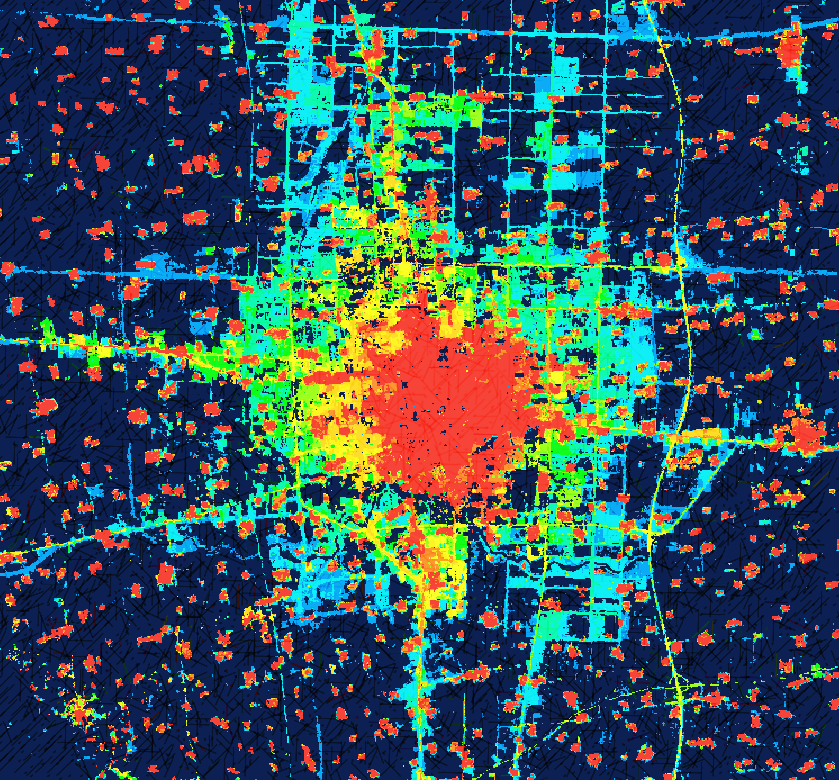
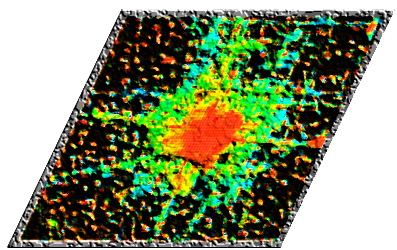
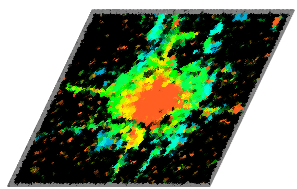
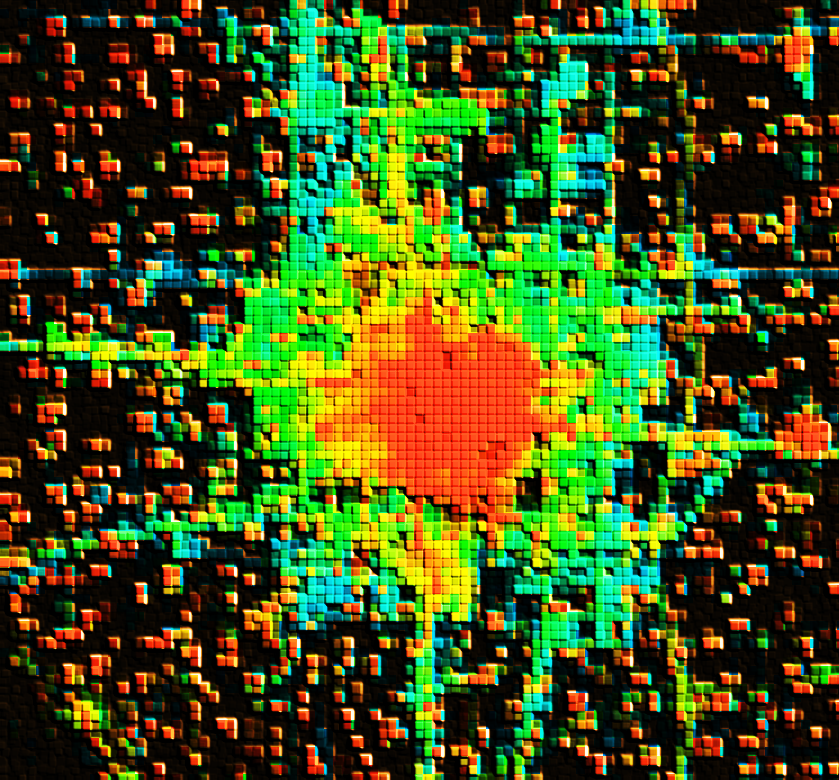
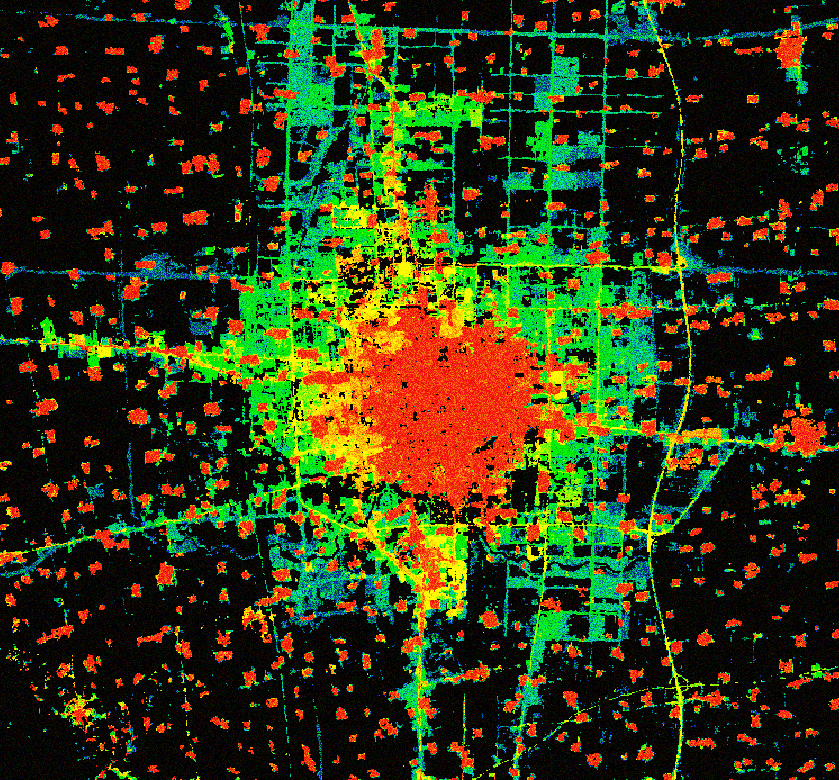
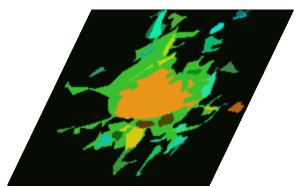
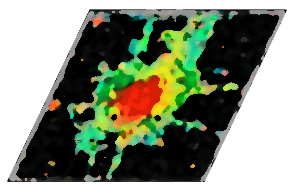
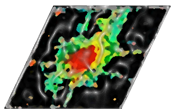
The Deep Learning (DL) algorithm is a kind of ML technology that abstracts naïve input variables to high-level features. Convolutional neural networks (CNN) are a special structure of DL that is designed to detect spatial patterns from both low- and high-level features (Krizhevsky et al. 2017). DL technologies are increasingly used to extract patterns and insights from geospatial data (Reichstein et al. 2019), which enables the spatial configurations, rather than proxied variables like decay distances, of driving factors to be integrated directly into urban development simulation. For example, Zhai et al. (2020) used CNN to retrieve the neighborhood spatial features and improved the Figure of Merit (FoM) of the simulation from 0.323 to 0.361 compared to the random-forest-based method. Qian et al. (2020) reported similar results that integrating spatial features abstracted by CNN improved the FoM from 0.299 to 0.346 comparing the random-forest algorithm. Although the spatial features introduced by CNN improved the simulation performance, these studies took CNN as an advanced decay function to rescale driving factors and were bound to detect only a definite neighborhood spatial configuration, whereas the high-level features abstracting capability of DL was overlooked. Therefore, more advanced deep learning structures that fully take advantage of CNN’s spatial pattern recognition ability, i.e., integrating low-level spatial features to high-level patterns, should be introduced to urban development simulation.

U-Net, first introduced for biomedical image segmentation in 2015, is a unique type of CNN-based deep learning structure that not only abstracts spatial features to high-level patterns but also refines the high-level patterns to precise shapes (Ronneberger et al. 2015). The robust segmentation performance enables U-Net to be used in multiple fields such as improving the global precipitation estimation to improve the weather prediction system (Singh et al. 2021), detecting underwater objects for oceanic ecosystem evaluation (Nezla et al. 2021), and identifying buildings from aerial and satellite imagery (Ji et al. 2019). Unlike CA models that have a definite neighborhood size, U-Net deploys a series of convolutional layers to extract spatial features and then assimilate these features automatically to produce transition rules. The intuition of using U-Net is that shape and spatial pattern matter when mapping real-world urbanization. This ability to learn spatial patterns suggests that U-Net has the potential for identifying and assimilating the spatial processes that drive urban development and accurately capture the resulting patterns of cities.

In this study, we assess the ability of the U-Net deep learning architecture to accurately project the spatial extent and pattern of urban development. We developed and applied the U-Net to project urban development in the North China Plain—China’s food bowl and one of the most rapidly urbanizing areas on the planet. We applied a U-Net model to learn patterns of urban development in the study area and used highly accurate maps of urban land-use between 1993 and 2012 to train the model. We then projected the spatial distribution of urbanization for 2019 and thoroughly tested the ability of the U-Net model to accurately capture the high-level spatial features, shapes, and patterns of urban development. Lastly, we used the trained model to project urban development patterns in the study area for 2030 based on historical rates of urbanization. We discuss both the advantages and limitations of the U-Net for simulating urban development and in particular its ability to learn typical urban development patterns such as neighborhood influences and linear expansion along transport routes. We also discuss the implications of rapid urbanization in the study area for food security and sustainability in the study area.

# 2. The U-Net structure

The U-Net structure includes down-sampling layers which extract the general context from the input data, up-sampling layers which refine these contexts to precise shapes, and skip-connections that balance the generalization of down-sampling and the refinement of up-sampling (Ronneberger et al. 2015). A conceptual U-Net structure (Figure 1) demonstrates its pattern recognition capability.



**Input map**

**Prediction map**

**Reference map**

**Loss**

**Down sampling**

**Skip connection**

**Up sampling**

*Figure 1. The conceptual structure of a 4-layer U-Net. In this study, the input map was resized to half of its input size (e.g., n2 to ) in each down-sampling layer, and then expanded to the original size in the up-sampling process. The loss denotes the difference between the prediction and the reference map, reflecting the performance of the U-Net.*

The down-sampling layers include convolution, pooling, and rescaling processes (Table 1). The convolution is a cross-correlation operation that produces feature maps indicating the similarity between the inputs and the convolution filters (Kapinchev et al. 2015). By applying multiple filters, the different spatial patterns can be retrieved independently. For example, Zeiler and Fergus (2013) reported that horizontal, vertical, and circular patterns were identified by different filters applied to the input image. The pooling process reduces the size of inputs. We select max-pooling to reduce the dimensionality because it is adopted in most DL structures (Murray and Perronnin 2014). The rescaling process rescales the pixel values of the feature maps to a specified range to optimize the computation flow of the network. We use rectified linear unit function (ReLU) for its simple, efficient, and robust performance (Agarap 2018).

Table 1. The equation of the component layers of U-Net.

|  |  |  |
| --- | --- | --- |
| **Layer** | **Equation** | **No.** |
| Convolution |  | (1) |
| Max-pooling |  | (2) |
| ReLU |  | (3) |
| Softmax |  | (4) |
| Cross-entropy |  | (5) |

For equation (1), the (*C*in, *H*, *W*) and (*C*out, *H*out, *W*out) refers to the size of input and output image, *C* denotes the number of channels, *H* is the height of input planes in pixels, and *W* is the width in pixels, ⋆ is the valid cross-correlation operator, *j* is the *j*-th channel of the output feature map. For equation (2), (*kH*, *kW*) denotes the kernel size of the pooling, h and w refer to the height and width of the output image. For equation (3), *x*denotes the pixel values of the input feature map. For equation (4), the *xi* is the *i*-th pixel value of the input feature map, *K* is the number of classes. For equation (5), *x* and *y* refer to the predicted and the reference pixel values, *K* is the number of classes.

The up-sampling layers include transpose-convolution and rescaling processes. The transpose-convolution, an inversed convolutive operation, transforms the input image from a lower resolution to a higher resolution. This process is assisted by skip-connections that bring additional spatial information from the down-sampling layers. The transpose-convolution is similar to the convolution except that the output size is bigger than the input size, and the rescaling is the same ReLU operation as the down-sampling. Additional components used in the U-Net are batch normalization, soft-max, and cross-entropy algorithm (Table 1). The batch normalization is applied to standardize the weights that control the convolution and transpose-convolution process, which have been proven to be effective in improving the DL performances (Sergey Ioffe and Christian Szegedy 2015). The softmax algorithm is applied to the last feature map of the U-Net to squash the pixel values to the 0 – 1 range for a better comparison to the reference map that is composed of pixels of 0s and 1s. The cross-entropy algorithm is applied to calculate the difference between the prediction and the reference map and reflects the performance of the U-Net.

The complete U-Net structure used in this study is shown in Figure 2. The input image size is reduced by half after passing each down-sampling block until reaching the size of 8 × 8 (i.e., 82 in Figure 2), then the image is recovered to the original size through each up-sampling block. The channels of the feature maps are undergoing an inversed process: the number of feature maps is doubled in each down-sampling block and halved in each up-sampling block. As a result, the U-Net gradually extracted more abstract spatial features on a bigger field-of-view according to *k × 2(d-1)*, where *k* is the kernel size (3 in this study) and *d* is the depth of the layer block. For example, the block with depth 1 has a field-of-view of 3, while it becomes 96 in the block of depth 6, meaning the U-Net is looking for spatial patterns of 96 × 96 in the original image scale. On the other hand, the U-Net is identifying more sophisticated spatial patterns as the network block goes deeper because the number of feature maps is doubled. The skip-connection links the down-sampling and up-sampling blocks, allowing the U-Net to refine the overview patterns extracted by deeper network blocks with precise shapes and textures retrieved by shallower blocks. A total of >31 million parameters were included in the U-Net model, making it highly flexible to capture the spatial patterns and stochasticity of the urban dynamic process.

Figure 2. The structure of the U-Net used in this study.

# 3. Methods

We used Landsat data to map urban development in the study area for the years 1994, 2006, and 2018 (Wang et al. 2021b), and combined them with elevation and slope information to simulate urban development. We created two U-Net models in this study: one was trained with urban maps of 1994-2006 and validated using the urban map of 2018; the other was trained on 2006-2018 and was used to predict the urban layout of 2030. Both training phases (i.e., 1994-2006 and 2006-2018) are referred to as “base”, both prediction data (2018 and 2030) are called “target,” and the true 2018 urban map is called as “reference” in this paper. Training samples were randomly extracted and used to train the U-Net. The trained model was used to produce a transition potential layer and create an urban land-use map. We then evaluated the accuracy of the transition potential map and urban land-use map using a range of accuracy and pattern-based metrics. We illustrate the use of the model in creating a future projection or urban land for 2030. The research workflow is summarized in Figure 3 and described in more detail below.

Figure 3. The research workflow.

## 3.1 Study area

The North China Plain (Figure 4) includes 76 prefectures, spans over 780,000 km2 area, and is home to >450 million people (National Bureau of Statistics of China 2019b). This area is one of the most rapidly urbanizing regions in China and the world, tripling the built-up land coverage from ~5% in 1990 to ~15% in 2020 (Wang et al. 2021b). This region is crucial to China's economic development and holds a strategic role in safeguarding China's food security, generating over one-third of the national gross domestic product and grain supply (National Bureau of Statistics of China 2019a). Managing the tension between urbanization and agricultural land-uses in the study area requires accurate, spatially-explicit projections of future urban development to address the interconnected challenges of food security, environmental protection, urbanization, and economic development.

Figure 4. The study area.

## 3.2 Data preprocessing

We mapped urban land-use (Figure 5) for the years 1994, 2006, and 2018 with a consistently high (>94%) accuracy. Full details of this work can be found in Wang et al. (2021b). The terrain data was obtained from the Shuttle Radar Topography Mission from which slope and elevation data were derived. Elevation and slope information was used to assist the U-Net in simulating the topographic control of urban development (i.e., urban expansion is more likely on flatter vs hilly and mountainous landscapes), a strategy which has proven to be effective in previous studies (Xing et al. 2020; Wang et al. 2021a; Qian et al. 2020). Accessibility variables such as distance to roads and railways were often employed in previous urban dynamic modeling (Tripathy and Kumar 2019; Valencia et al. 2020; Ronneberger et al. 2015). However, U-Net is a pattern-sensitive model. Distance factors hardly provide any useful pattern information. Hence, it was not used in this study.

Figure 5. The urban dynamic map of the North China Plain.

## 3.3 Training the U-Net

Control samples were assembled using Google Earth Engine (Gorelick et al. 2017). The *neighborhoodToArray* module was used to crop 256 × 256 pixel tiled samples from the base image. We set the tile size to 256 following common data science practices (Ronneberger et al. 2015), and collected 20,000 samples for training and 5,000 for validation. The samples included 3-layer input data (e.g., urban map of 1994, elevation, and slope) and a single-layer later year image (e.g., urban map for 2006).

The U-Net was trained for 200 epochs (an epoch refers to the U-Net complete updating its weights using all the 20,000 training samples). We saved the model produced at each epoch and tested its performance on the 5,000 validation samples. During the training process, the tiled input images were resized to 8 × 8 pixels after five down-sampling operations and then converted to a single-layer output image that was the same size as the original input image tile with another five up-sampling operations (Figure 2). During the validation process, the mean squared error (MSE) was used to compute the difference between the output image and the target image (i.e., the loss) of the model. Finally, the model that has the least MSE was determined as the best model to project the target urban layout.

## 3.4 Producing the simulation map

The input image was split into tiles, supplied to the trained U-Net to produce separate outputs, and then the outputs were mosaiced into a single transition potential map. The pixel values of the transition potential map ranged from 0 to 1, indicating the probability of being an urban pixel in the projection date. To reduce the "edge effect," the image tiles were cropped at the size of 256 × 256 pixels and then supplied with a 32-pixel width buffer to be removed after being processed by the U-Net (Figure 6). Although the U-Net was trained on image tiles of 256 × 256 pixels, it can process the buffered 320 × 320 (original size of 256 plus buffers at edges of size 32 × 2) image tiles because of the dimension insensibility of CNN structures (Long et al. 2015).

Figure 6. Buffering image tile to reduce the "edge effect." Because the final output was a mosaic of image tiles, we buffered each image tile with additional 32 pixels at the edges and removed these buffers before mosaicking to the final output map to reduce the “edge effect.” A bigger buffer size will better alleviate the “edge effect” but lead to a higher computation cost, thus we selected a 32 buffer size for its common usage in deep learning structures that balances effect and cost (Long et al. 2015).

The classified urban land-use map was created by binarizing the transition potential map. We ranked the pixel values of the transition potential map from the highest to the lowest, accumulated the pixel count in this order, and classified them to the value of 1 (indicating an urban pixel) until the accumulating count value met the urban pixel number in the target year, then allocated the rest pixels with the value of 0 indicating non-urban pixels. The urban pixel number of 2018 was computed from the reference map to binarize the transition potential map of 2018, and an exponential extrapolation was carried out to (Figure 7) determine the urban pixel count of 2030 that was used to binarize the transition potential map of 2030. The binarization was carried out independently at each prefecture to reduce the bias caused by different regional development levels.

Figure 7. The exponential regression on historical urban areas.

## 3.5 Validation and accuracy assessment

We selected a series of accuracy metrics and spatial pattern metrics to assess the ability of U-Net to accurately project the spatial distribution and pattern of urban areas in the study area. The transition potential map was assessed using the area under the curve (AUC) of the receiver operating characteristic curve (ROC), which illustrates the diagnostic ability to discriminate urban and non-urban pixels under varied thresholds (Fawcett 2006). The classified urban land-use map was evaluated via map-overlay and landscape-level spatial pattern metrics (McGarigal and Marks 1995). The map-overlay metrics selected in this study were overall accuracy (OA), the hit rate, and the figure of merit (FoM). The spatial pattern metrics selected in this study were patch number (PN) and landscape shape index (LSI). The selected metrics are described in Table 2. Validation was performed for each prefecture independently, yielding a total of 76 records for each metric.

Table 2. Validation metrics to evaluate the classified urban land-use map.

|  |  |  |
| --- | --- | --- |
| **Name** | **Equation** | **Explanation** |
| AUC |  | AUC measures a model's aggregated performance under different thresholds to discriminate urban and non-urban pixels (Tong and Feng 2020). |
| OA | (A + D) / (A + B + C +D) | Overall accuracy is the ratio of correctly identified urban and non-urban pixels to the total number of predictions. |
| Hit rate | A / (A + C) | The hit rate is the ratio of correctly identified urban pixels (i.e., hits) to the number of urban pixels in the reference map. |
| FoM | A / (A + B + C) | FoM is the ratio of the intersection to the union when overlay predicted urban pixels with reference urban pixels (Pontius et al. 2008). |
| PN | *n* | The patch number is the number of patches of the urban landscape. |
| LSI |  | Reflect the complexity of urban landscape patches. For example, a squared patch is deemed simple (low value), whereas a linear patch is complex (high value). |

*Note, TPR is the true positive rate, FPR is the false positive rate of the ROC from the transition potential map. A refers to correctly predicted urban pixels (hit), B refers to the incorrectly predicted urban pixels (false alarm), C is the incorrectly predicted non-urban pixels (miss), and D is the correctly predicted non-urban pixels (correct rejection). n is the total number of landscape patches. is the perimeter (m) of the ith path of the urban patches, is the area (hectare) of the ith patch of the urban patches.*

We first validated the U-Net using historical urban maps, then trained the other U-Net model to project future urban land-use. First, urban land-use maps for 1994 and 2006 were used to train the U-Net model that was used to simulate the spatial distribution of urban land-use for 2018, which was assessed, compared, and validated against the reference urban land-use map for 2018. Lastly, the other U-Net was trained on the urban land-use maps for 2006 and 2018 to predict the urban land-use for 2030.

# 4. Results

## 4.1 Model training

The MSE of the U-Net trained on the urban images of 1994 and 2006 is shown in Figure 8. The lowest MSE was 0.022 at the 70th epoch, which was determined as the best training epoch to produce the U-Net model that simulated the urban layout of 2018.

Figure 8. The mean squared error (MSE) of applying the training U-Net on the validation samples at different epochs. We saved the intermedia model of each epoch and the dot is the MSE of such model on the validation samples. The line and ribbon are the fitting and confidence interval of the dots to better reflect the trend.

The structure of the U-Net successfully learned and captured different aspects of the complex spatial patterns of urban areas in the study area (Figure 9). For example, in the illustrative sample presented in Figure 9, the Down-1 layer broadly distinguished between urban and non-urban pixels. The Down-2 layer had recognized simple patterns such as the horizontal and vertical roads and learned to highlight larger towns among small villages. The Down-3 layer had learned to associate adjacent urban clusters, and potential urban development corridors connecting discreet towns/villages were identified. The Down-4 layer allocated higher urban development probability to the pixels near existing towns/villages, making the urban development corridors more concentrated and intensive. The Bottleneck layers captured the general pattern of urban development in the target year. The up-sampling layers integrated both high-level and low-level features. The Up-4 layer refined the overview patterns in the Bottleneck layer by combining the Bottleneck layer and the Down-4 layer. The Up-3 layer further refined the spatial features in the Up-4 layer by assimilating urban development corridors identified in the Down-3 layer. Lastly, the Up-1 layer associated the urban/non-urban activation maps in the Down-1 layer with the Up-2 layer producing the final output image tile, which allocated more urban pixels around bigger towns and maintained the refined patterns.

**Down-1**

**Down-2**

**Down-3**

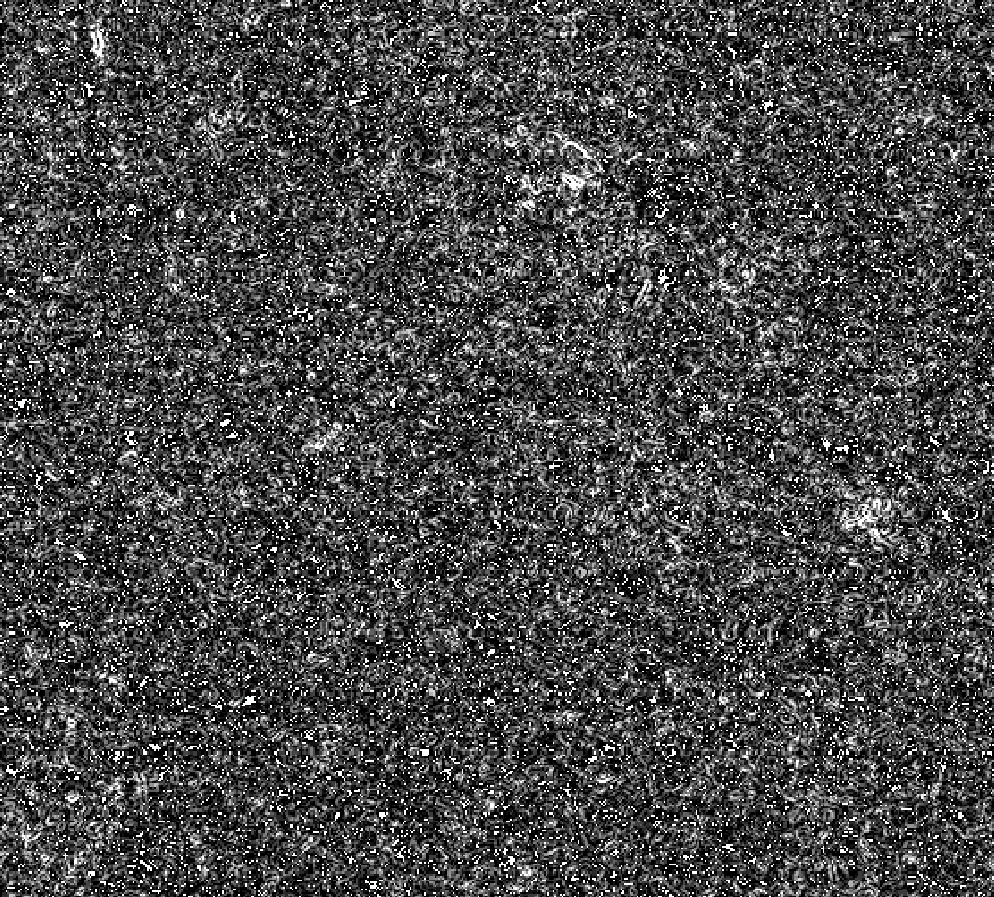
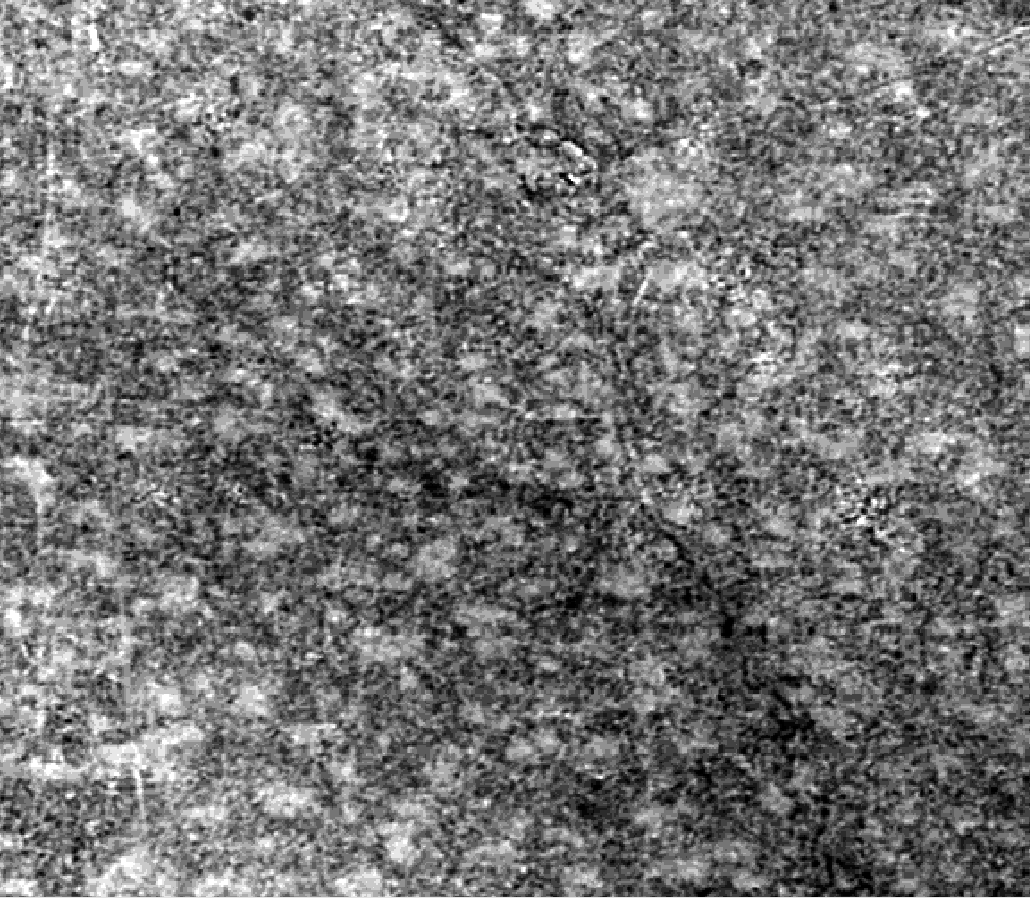
**Down-4**

**Up-4**

**Up-3**

**Up-2**

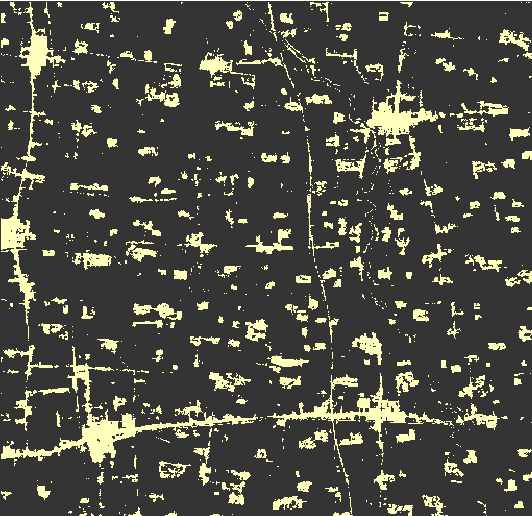
**Up-1**



Urban extent

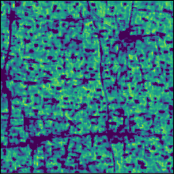
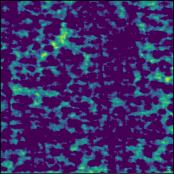
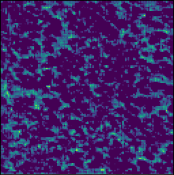
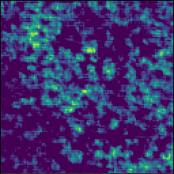
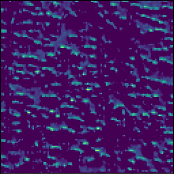
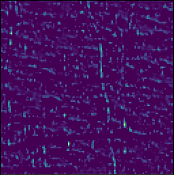
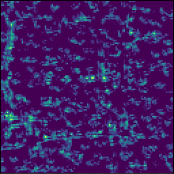
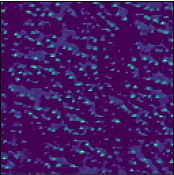
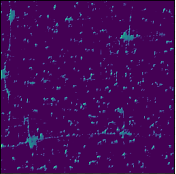
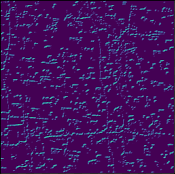
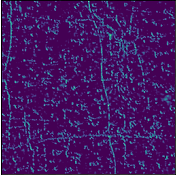
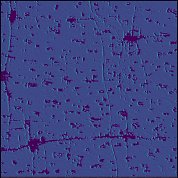
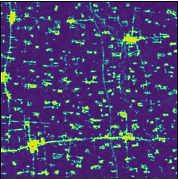
DEM

Slope



**Input image tile**

**Output image tile**



**Bottleneck**

Figure 9. The visualization of an image tile process by different U-Net layers. Note we only selected a few activation maps for the visualization given a limited page space. A randomly selected image tile was used to visualize the U-Net's pattern recognition capability. The last activation map of each layer was used to visualize its pattern recognition capability.

## 4.2 Validation maps

The transition potential map is shown in Figure 10. The pixels near larger towns and cities were allocated with higher transition potential value than small villages. For example, a large area surrounding Sui`Xian was identified as having high transition potential, while such phenomena were only appeared at the thin edges of smaller nearby villages. Similar patterns were shown in Yang`Yuan and Du`Ling. Linear development patterns were also well captured in the transition potential map. For example, the roads in Yang’Kou Zhen and Hui`Ji Zhen were identified despite only discrete linear segments shown in the urban map of 2006.

Figure 10. The transition potential to urban in 2018 given the urban map of 2006.

The classified urban land-use map (Figure 11) created by binarizing the transition potential map correctly identified new urban areas (i.e., hits) particularly those adjacent to existing cities demonstrating the ability to capture the neighborhood influence on urban development. It also incorrectly identified many of these areas as new urban areas (i.e., false alarms) and failed to identify many new urban areas particularly in areas designated as new developments (i.e., misses). Much of the missed urban projections occurred some distance from the original urban areas, such as the newly developed land in Bin’Hai Zhen, Yang'Kou Zhen, Hu’Ying Zhen, and Xuan’Cheng City that were remote from the old town center. Importantly, the model was exceptional at identifying urban expansion along with linear features such as the new development along major roads in Du’Ling Xiang, Gao’Gou Zhen, and Hui`Ji Zhen.

Figure 11. The classified urban land-use map includes hits, misses, and false alarms.

## 4.3 Validation metrics

The overlay metrics of the validation maps compared to the reference map are shown in Figure 12. The AUC the transition potential ranges from 0.70 to 0.97 with a median value of 0.81. The median values of the OA, the hit rate, and the FoM of the classified urban land-use map were 0.91, 0.33, and 0.20, respectively.

Figure 12. The accuracy metrics for the classified urban land use map for 2018.

The landscape-scale spatial pattern metrics of the classified urban land-use map compared to the reference map are shown in Figure 13. The Patch Number of the prediction map were systematically less than the reference map, while the Landscape Shape Index of both maps were close. As a result, the U-Net tends to predict fewer urban land patches than the real urban development but produced very close urban shapes to the real-world urban patterns.

Figure 13. The landscape metrics for the hard classified map of 2018. The dashed line refers to y = x, the ribbon is the confidence interval of the linear fitting to the scatter points.

## 4.4 Predicted urban area to 2030

The predicted urban map of 2030 is in Figure 14. The newly predicted urban area tended to be concentrated around large cities. In Sui`Xian for example, the predicted urban expansion was much larger around big cities/towns than those near smaller villages. The predicted urban area followed specified patterns rather than sprawling in every direction. For example, the predicted urban lands in Du`Ling Xiang and Yang'Kou Zhen infilled the gaps in the spatial distribution of existing urban areas and maintained the general shape of the original urban layout. The prediction of Feng`Ning County followed the city's elongated development trend, and urban lands emerged along transport routes and at major road intersections illustrated by Hu'Ji Zhen and Gao`Gou Zhen.

Figure 14. The predicted urban lands in 2030.

The predicted area of each region is shown in Table 3. The average urban area increase rate is 35.54% compared to 2018. Beijing, Tianjin, and Hebei had the lowest area increment of <30%, Anhui, Jiangsu, and Henan had the highest predicted area increase of >37%. Shandong was predicted with medium increases of 33.60%.

Table 3. The prediction urban area of each province in 2030.

|  |  |  |  |
| --- | --- | --- | --- |
| **Region** | **2018 (km2)** | **2030 (km2)** | **Increment (%)** |
| Anhui | 21070.96 | 29463.61 | 39.83 |
| Beijing | 2630.09 | 3338.75 | 26.94 |
| Hebei | 20454.29 | 26440.38 | 29.27 |
| Henan | 28282.24 | 38850.77 | 37.37 |
| Jiangsu | 24210.08 | 33731.72 | 39.33 |
| Shandong | 32005.14 | 42758.61 | 33.60 |
| Tianjin | 2925.68 | 3754.65 | 22.08 |
| **Sum** | **131578.48** | **178338.49** | **35.54** |

# 5. Discussion

## 5.1 Mimicking the real-world urban dynamic patterns

The U-Net had captured and assimilated high-level spatial patterns to mimic the urban development. First, the model captured the neighborhood effects in urban development. The transition potential map revealed that lands near existing urban areas are prone to be transited to urban in the future, and the abundant non-urban vicinity will refrain urbanization from happening (Figure 10). Second, the U-Net was able to assimilate the learned small-scale neighborhood effects into large-scale gravity effects. Many near-town-center lands were predicted to be urbanized in the future, while the small remote villages only have a few such predictions (Figure 11). Third, the U-Net, rather than just simply buffering spatial features, had captured the urban expansion momentum as well as the precise linear patterns. For example, the elongating development trend of Feng`Ning County controlled by a valley terrain was identified in the prediction (Figure 14), and areas along with transport routes were allocated with high transition potentials (Figure 10). The U-Net model was not able to predict new planned developments that sprouted remote from the old town centers such as Bin’Hai Zhen, Yang'Kou Zhen, Hu’Ying Zhen, and Xuan’Cheng City (Figure 11).

## 5.2 Reducing subjectivity by automatically constructing transition rules

The large-scale spatial patterns identified by the U-Net were hard to be captured by previous CA models. A large neighborhood is required for a CA model to incorporate information at a large scale. However, the simulation performance decreased as the neighborhood reached a specific size. Wang et al. (2021a) found that the best neighborhood size to simulate urban development in Beijing is 25 × 25. Roodposhti et al. (2020) also observed that the 9 × 9 neighborhood outperformed other window sizes settings. Therefore, the unique design of the U-Net is capable of ingesting more extensive information than previous CA models, enabling the urban development to be simulated with more refined spatial configurations.

The U-Net did not require preset parameters to simulate urban dynamics except for the learning rate and the size of training images. The transition rules were automatically generated in the form of learned weights, and the training is an unattained process that reduces subjective biases. In contrast, previous CA models often require a complex calibration process to determine suitable parameters. For example, Feng and Tong (2020) developed a framework that integrated three algorithms and different neighborhood settings to simulate the urban growth in Shanghai, whereas an array of parameters like inertia weights, decay magnitude, spatial heterogeneity, and variable scaling were required to be specified before running the model. Many studies determine these parameters according to expert knowledge (Mustafa et al. 2017; Chen et al. 2020; Tripathy and Kumar 2019), potentially leading to subjective biases. Other studies went through a1 systematic parameter selection process to find the best parameters (Roodposhti et al. 2020; Yu et al. 2021), but this process is time-consuming and could be infeasible given many undetermined parameters.

## 5.3 Accurate prediction and robust in capturing spatial patterns

The FoM in this study, ranging from 0.12 to 0.27 with the median value of 0.19, was close to existing studies: the FoM of the best urban simulation map on Zhuji was 0.19 (Wang et al. 2021c), 0.21 in Beijing (Wang et al. 2021a). Some studies reported a higher FoM because the urban development was simulated with a coarser resolution (Pontius et al. 2008; Peng et al. 2020; Valencia et al. 2020), or the simulation was performed in a relatively smaller area (Pramanik et al. 2021). The landscape metrics revealed that the simulation map produced by the U-Net aligned well with the reference map. The landscape shape index between the simulation and the reference was very close as the linear regression slope is close to 1 (Figure 13). Such correlation was also embodied in the classified urban maps, where the linear structures like roads and the general development trend of cities were captured in the simulation map (Figure 11, Figure 7). It is hard to compare the landscape metrics across studies because most of them only focus on one region, making the metric susceptible to the specific setting of each research. However, general outward expansion patterns, rather than precise patterns identified by the U-Net, were easy to observe in many previous urban simulations (Gao et al. 2020; Shafizadeh-Moghadam et al. 2017; Pérez-Molina et al. 2017). Therefore, the U-Net had learned the inherent patterns of urban expansion compared to previous models.

## 5.4 Predicting urban land-use change for policy formulation

A total of 4.67 × 104 km2 of lands were predicted to be urbanized in the North China Plain from 2018 to 2030 (Table 3), and the three provinces with the highest urban area increase rate (Anhui, Jiangsu, and Henan) were all in the southern part (Figure 4). The prediction map infers an imbalanced urbanization future in the next decade, which could bring challenges for the social-economic development and ecological conservations in the study area. Our precise 2030 urban layout prediction provides spatial-explicit references to reflect the key economic performance of China because three mega-city-group (Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Central Plains ) that accounted for one-third of the gross domestic product (GDP) in China are included in the study region (National Bureau of Statistics of China 2019b). The strategical development plan of China benefits from the prediction of this study in identifying potential social-economic hazards originating from unbalanced urban expansion, and our exact simulation offers an accurate map to assist the policy formulation that adapted to the explicit urbanization conditions. The competition between urban development and food production in the study area has an essential impact on China’s food security (Jin et al. 2019), which can be alleviated with our exact future urban predictions because this accurate prediction provides in advance alerts to develop plans/policies that reduce the food production loss from farmland occupation. The ecosystem and environment conservation also benefit from our precise urban predictions. For example, urban land is an important proxy for water consumption (Hoekstra et al. 2018) and waste production (Zeller et al. 2019), and such consumption and production in the study area in the future can be determined given the accurate urban map. Lastly, the future urban layout provides a baseline for the local government to address the land-use conflicts and settle the competing interests among different stakeholders in the urbanization process.

## 5.5 Limitations and prospects

The U-Net learned the spatial patterns from historical urban dynamics, while the previous CA methods modeled urban development from the influence of driving variables. As a result, U-Net has a limited capacity to support participatory modeling that requires a series of trials on the driving factors. The U-Net model focus on predictive accuracy rather than explaining the urbanization process. The transition rules of the U-Net are hidden in the vast weights. However, the CA models are usually explainable, i.e., the key driving factors can be identified. Thus, in the field of deliberately controlling a city’s development, such as city planning or scenario projection, U-Net could not be well applied. The following study will explore explainable U-Net models.

# 5. Conclusion

In this study, we applied the U-Net to simulate urban development in 76 cities of the North China Plain. The AUC, overlay, and landscape metrics were used to assess the simulation. The urban development of 1994-2006 was used to train a validation model that projected the urban map to 2018; this map was then evaluated by a reference map. The other model, trained on urban maps of 2006-2018, was used to project urban development to 2030. The U-Net model successfully captured the neighborhood effects, the gravity effects, and the linear expansion along transportation routes in urban dynamics. The overlay metrics showed that the simulated map from U-Net achieved similar performance to the previous models that went through a length calibration process. The landscape metrics showed that the predictions were well aligned with the reference map. The U-Net model identifies precise urban development patterns but has limited capacity to explain the urbanization process, which is a useful addition DL tool to the existing urban land-use change model collections.

Publication bibliography

Agarap, Abien Fred (2018): Deep Learning using Rectified Linear Units (ReLU).

Carneiro, Murillo G.; Oliveira, Gina M. B. (2013): Synchronous cellular automata-based scheduler initialized by heuristic and modeled by a pseudo-linear neighborhood. In *Nat Comput* 12 (3), pp. 339–351. DOI: 10.1007/s11047-013-9375-8.

Chen, Shurui; Feng, Yongjiu; Tong, Xiaohua; Liu, Song; Xie, Huan; Gao, Chen; Lei, Zhenkun (2020): Modeling ESV losses caused by urban expansion using cellular automata and geographically weighted regression. In *The Science of the total environment* 712, p. 136509. DOI: 10.1016/j.scitotenv.2020.136509.

Clarke, Keith C.; Johnson, J. Michael (2020): Calibrating SLEUTH with big data: Projecting California's land use to 2100. In *Computers, Environment and Urban Systems* 83, p. 101525. DOI: 10.1016/j.compenvurbsys.2020.101525.

Fan, Hao; Zhao, Chuanfeng; Yang, Yikun (2020): A comprehensive analysis of the spatio-temporal variation of urban air pollution in China during 2014–2018. In *Atmospheric Environment* 220, p. 117066. DOI: 10.1016/j.atmosenv.2019.117066.

Fan, Peilei; Chen, Jiquan; Ouyang, Zutao; Groisman, Pavel; Loboda, Tatiana; Gutman, Garik et al. (2018): Urbanization and sustainability under transitional economies: a synthesis for Asian Russia. In *Environ. Res. Lett.* 13 (9), p. 95007. DOI: 10.1088/1748-9326/aadbf8.

Fawcett, Tom (2006): An introduction to ROC analysis. In *Pattern Recognition Letters* 27 (8), pp. 861–874. DOI: 10.1016/j.patrec.2005.10.010.

Feng, Yongjiu; Tong, Xiaohua (2020): A new cellular automata framework of urban growth modeling by incorporating statistical and heuristic methods. In *International Journal of Geographical Information Science* 34 (1), pp. 74–97. DOI: 10.1080/13658816.2019.1648813.

Gantumur, Byambakhuu; Wu, Falin; Vandansambuu, Battsengel; Tsegmid, Bazarkhand; Dalaibaatar, Enkhjargal; Zhao, Yan (2020): Spatiotemporal dynamics of urban expansion and its simulation using CA-ANN model in Ulaanbaatar, Mongolia. In *Geocarto International*, pp. 1–16. DOI: 10.1080/10106049.2020.1723714.

Gao, Chen; Feng, Yongjiu; Tong, Xiaohua; Lei, Zhenkun; Chen, Shurui; Zhai, Shuting (2020): Modeling urban growth using spatially heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR. In *Computers, Environment and Urban Systems* 81, p. 101459. DOI: 10.1016/j.compenvurbsys.2020.101459.

Gorelick, Noel; Hancher, Matt; Dixon, Mike; Ilyushchenko, Simon; Thau, David; Moore, Rebecca (2017): Google Earth Engine: Planetary-scale geospatial analysis for everyone. In *Remote Sens Environ* 202, pp. 18–27. DOI: 10.1016/j.rse.2017.06.031.

Hoekstra, Arjen Y.; Buurman, Joost; van Ginkel, Kees C. H. (2018): Urban water security: A review. In *Environ. Res. Lett.* 13 (5), p. 53002. DOI: 10.1088/1748-9326/aaba52.

Ji, Shunping; Wei, Shiqing; Lu, Meng (2019): A scale robust convolutional neural network for automatic building extraction from aerial and satellite imagery. In *International Journal of Remote Sensing* 40 (9), pp. 3308–3322. DOI: 10.1080/01431161.2018.1528024.

Jin, Gui; Chen, Kun; Wang, Pei; Guo, Baishu; Dong, Yin; Yang, Jun (2019): Trade-offs in land-use competition and sustainable land development in the North China Plain. In *Technological Forecasting and Social Change* 141, pp. 36–46. DOI: 10.1016/j.techfore.2019.01.004.

Kafy, Abdulla -. Al; Naim, Md. Nazmul Huda; Subramanyam, Gangaraju; Faisal, Abdullah-Al; Ahmed, Nessar Uddin; Rakib, Abdullah Al et al. (2021): Cellular Automata approach in dynamic modelling of land cover changes using RapidEye images in Dhaka, Bangladesh. In *Environmental Challenges* 4, p. 100084. DOI: 10.1016/j.envc.2021.100084.

Kapinchev, Konstantin; Bradu, Adrian; Barnes, Frederick; Podoleanu, Adrian (2015): GPU implementation of cross-correlation for image generation in real time, pp. 1–6. DOI: 10.1109/ICSPCS.2015.7391783.

Kipfer, Stefan (2018): Pushing the limits of urban research: Urbanization, pipelines and counter-colonial politics. In *Environ Plan D* 36 (3), pp. 474–493. DOI: 10.1177/0263775818758328.

Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey E. (2017): ImageNet classification with deep convolutional neural networks. In *Commun. ACM* 60 (6), pp. 84–90. DOI: 10.1145/3065386.

Li, Xuecao; Gong, Peng; Le Yu; Hu, Tengyun (2017): A segment derived patch-based logistic cellular automata for urban growth modeling with heuristic rules. In *Computers, Environment and Urban Systems* 65, pp. 140–149. DOI: 10.1016/j.compenvurbsys.2017.06.001.

Liu, Yan; Batty, Michael; Wang, Siqin; Corcoran, Jonathan (2021): Modelling urban change with cellular automata: Contemporary issues and future research directions. In *Progress in human geography* 45 (1), pp. 3–24. DOI: 10.1177/0309132519895305.

Long, Jonathan; Shelhamer, Evan; Darrell, Trevor (2015): Fully convolutional networks for semantic segmentation, pp. 3431–3440. DOI: 10.1109/CVPR.2015.7298965.

Mansour, Shawky; Al-Belushi, Mohammed; Al-Awadhi, Talal (2020): Monitoring land use and land cover changes in the mountainous cities of Oman using GIS and CA-Markov modelling techniques. In *Land use policy* 91, p. 104414. DOI: 10.1016/j.landusepol.2019.104414.

McGarigal, Kevin; Marks, Barbara J. (1995): FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. Spatial Pattern Analysis Program for Quantifying Landscape Structure. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Murray, Naila; Perronnin, Florent (2014): Generalized Max Pooling.

Mustafa, Ahmed; Cools, Mario; Saadi, Ismaïl; Teller, Jacques (2017): Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM). In *Land use policy* 69, pp. 529–540. DOI: 10.1016/j.landusepol.2017.10.009.

Mustafa, Ahmed; Heppenstall, Alison; Omrani, Hichem; Saadi, Ismaïl; Cools, Mario; Teller, Jacques (2018): Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. In *Computers, Environment and Urban Systems* 67, pp. 147–156. DOI: 10.1016/j.compenvurbsys.2017.09.009.

National Bureau of Statistics of China (2019a): Announcement of the 2019 grain output. Beijing, China. Available online at http://www.gov.cn/xinwen/2019-12/07/content\_5459250.htm, checked on 12/20/2020.

National Bureau of Statistics of China (2019b): China Statistical Yearbook. Beijing, China: China Statistics Press, checked on 11/6/2020.

Newland, Charles P.; Zecchin, Aaron C.; Maier, Holger R.; Newman, Jeffrey P.; van Delden, Hedwig (2018): Empirically derived method and software for semi-automatic calibration of Cellular Automata land-use models. In *Environmental Modelling & Software* 108, pp. 208–239. DOI: 10.1016/j.envsoft.2018.07.013.

Nezla, N. A.; Mithun Haridas, T. P.; Supriya, M. H. (2021): Semantic Segmentation of Underwater Images using UNet architecture based Deep Convolutional Encoder Decoder Model, pp. 28–33. DOI: 10.1109/ICACCS51430.2021.9441804.

Peng, Kaifeng; Jiang, Weiguo; Deng, Yue; Liu, Yinghui; Wu, Zhifeng; Chen, Zheng (2020): Simulating wetland changes under different scenarios based on integrating the random forest and CLUE-S models: A case study of Wuhan Urban Agglomeration. In *Ecological Indicators* 117, p. 106671. DOI: 10.1016/j.ecolind.2020.106671.

Pérez-Molina, Eduardo; Sliuzas, Richard; Flacke, Johannes; Jetten, Victor (2017): Developing a cellular automata model of urban growth to inform spatial policy for flood mitigation: A case study in Kampala, Uganda. In *Computers, Environment and Urban Systems* 65, pp. 53–65. DOI: 10.1016/j.compenvurbsys.2017.04.013.

Planillo, Aimara; Kramer‐Schadt, Stephanie; Buchholz, Sascha; Gras, Pierre; Lippe, Moritz von der; Radchuk, Viktoriia (2021): Arthropod abundance modulates bird community responses to urbanization. In *Divers Distrib* 27 (1), pp. 34–49. DOI: 10.1111/ddi.13169.

Pontius, Robert Gilmore; Boersma, Wideke; Castella, Jean-Christophe; Clarke, Keith; Nijs, Ton de; Dietzel, Charles et al. (2008): Comparing the input, output, and validation maps for several models of land change. In *Ann Reg Sci* 42 (1), pp. 11–37. DOI: 10.1007/s00168-007-0138-2.

Pramanik, Suvamoy; Butsch, Carsten; Punia, Milap (2021): Post-liberal urban dynamics in India – The case of Gurugram, the ‘Millennium City’. In *Remote Sensing Applications: Society and Environment* 22, p. 100504. DOI: 10.1016/j.rsase.2021.100504.

Qian, Yuehui; Xing, Weiran; Guan, Xuefeng; Yang, Tingting; Wu, Huayi (2020): Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation. In *The Science of the total environment* 722, p. 137738. DOI: 10.1016/j.scitotenv.2020.137738.

Qiu, Bingwen; Li, Haiwen; Tang, Zhenghong; Chen, Chongcheng; Berry, Joe (2020): How cropland losses shaped by unbalanced urbanization process? In *Land use policy* 96, p. 104715. DOI: 10.1016/j.landusepol.2020.104715.

Reichstein, Markus; Camps-Valls, Gustau; Stevens, Bjorn; Jung, Martin; Denzler, Joachim; Carvalhais, Nuno; Prabhat (2019): Deep learning and process understanding for data-driven Earth system science. In *Nature* 566 (7743), pp. 195–204. DOI: 10.1038/s41586-019-0912-1.

Ronneberger, Olaf; Fischer, Philipp; Brox, Thomas (2015): U-Net: Convolutional Networks for Biomedical Image Segmentation 9351, pp. 234–241. DOI: 10.1007/978-3-319-24574-4\_28.

Roodposhti, Majid Shadman; Aryal, Jagannath; Bryan, Brett A. (2019): A novel algorithm for calculating transition potential in cellular automata models of land-use/cover change. In *Environmental Modelling & Software* 112, pp. 70–81. DOI: 10.1016/j.envsoft.2018.10.006.

Roodposhti, Majid Shadman; Hewitt, Richard J.; Bryan, Brett A. (2020): Towards automatic calibration of neighbourhood influence in cellular automata land-use models. In *Computers, Environment and Urban Systems* 79, p. 101416. DOI: 10.1016/j.compenvurbsys.2019.101416.

Ruiz Hernandez, Ivan Elias; Shi, Wenzhong (2018): A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. In *International Journal of Remote Sensing* 39 (4), pp. 1175–1198. DOI: 10.1080/01431161.2017.1395968.

Sergey Ioffe; Christian Szegedy (2015): Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *International Conference on Machine Learning*, pp. 448–456. Available online at http://proceedings.mlr.press/v37/ioffe15.html.

Shafizadeh-Moghadam, Hossein; Asghari, Ali; Tayyebi, Amin; Taleai, Mohammad (2017): Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. In *Computers, Environment and Urban Systems* 64, pp. 297–308. DOI: 10.1016/j.compenvurbsys.2017.04.002.

Shaw, Brian J.; van Vliet, Jasper; Verburg, Peter H. (2020): The peri-urbanization of Europe: A systematic review of a multifaceted process. In *Landscape and Urban Planning* 196, p. 103733. DOI: 10.1016/j.landurbplan.2019.103733.

Singh, Manmeet; Kumar, Bipin; Rao, Suryachandra; Gill, Sukhpal Singh; Chattopadhyay, Rajib; Nanjundiah, Ravi S.; Niyogi, Dev (2021): Deep learning for improved global precipitation in numerical weather prediction systems.

Tong, Xiaohua; Feng, Yongjiu (2020): A review of assessment methods for cellular automata models of land-use change and urban growth. In *International Journal of Geographical Information Science* 34 (5), pp. 866–898. DOI: 10.1080/13658816.2019.1684499.

Tripathy, Pratyush; Kumar, Amit (2019): Monitoring and modelling spatio-temporal urban growth of Delhi using Cellular Automata and geoinformatics. In *Cities* 90, pp. 52–63. DOI: 10.1016/j.cities.2019.01.021.

Valencia, Victor H.; Levin, Gregor; Hansen, Henning Sten (2020): Modelling the spatial extent of urban growth using a cellular automata-based model: a case study for Quito, Ecuador. In *Geografisk Tidsskrift-Danish Journal of Geography* 120 (2), pp. 156–173. DOI: 10.1080/00167223.2020.1823867.

Wang, Haijun; Guo, Jiaqi; Zhang, Bin; Zeng, Haoran (2021a): Simulating urban land growth by incorporating historical information into a cellular automata model. In *Landscape and Urban Planning* 214, p. 104168. DOI: 10.1016/j.landurbplan.2021.104168.

Wang, Jinzhu; Hadjikakou, Michalis; Bryan, Brett A. (2021b): Consistent, accurate, high resolution, long time-series mapping of built-up land in the North China Plain. In *GIScience & Remote Sensing*, pp. 1–17. DOI: 10.1080/15481603.2021.1948275.

Wang, Rong; Feng, Yongjiu; Wei, Yongliang; Tong, Xiaohua; Zhai, Shuting; Zhou, Yilun; Wu, Peiqi (2021c): A comparison of proximity and accessibility drivers in simulating dynamic urban growth. In *Transactions in GIS* 25 (2), pp. 923–947. DOI: 10.1111/tgis.12707.

Xia, Chang; Zhang, Bin (2021): Exploring the effects of partitioned transition rules upon urban growth simulation in a megacity region: a comparative study of cellular automata-based models in the Greater Wuhan Area. In *GIScience & Remote Sensing*, pp. 1–24. DOI: 10.1080/15481603.2021.1933714.

Xing, Weiran; Qian, Yuehui; Guan, Xuefeng; Yang, Tingting; Wu, Huayi (2020): A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. In *Computers & Geosciences* 137, p. 104430. DOI: 10.1016/j.cageo.2020.104430.

Yeh, Anthony Gar-On; Chen, Zifeng (2020): From cities to super mega city regions in China in a new wave of urbanisation and economic transition: Issues and challenges. In *Urban Studies* 57 (3), pp. 636–654. DOI: 10.1177/0042098019879566.

Yu, Jingyan; Hagen-Zanker, Alex; Santitissadeekorn, Naratip; Hughes, Susan (2021): Calibration of cellular automata urban growth models from urban genesis onwards - a novel application of Markov chain Monte Carlo approximate Bayesian computation. In *Computers, Environment and Urban Systems* 90, p. 101689. DOI: 10.1016/j.compenvurbsys.2021.101689.

Zeiler, Matthew D.; Fergus, Rob (2013): Visualizing and Understanding Convolutional Networks. Available online at https://arxiv.org/pdf/1311.2901.

Zeller, Vanessa; Towa, Edgar; Degrez, Marc; Achten, Wouter M. J. (2019): Urban waste flows and their potential for a circular economy model at city-region level. In *Waste management (New York, N.Y.)* 83, pp. 83–94. DOI: 10.1016/j.wasman.2018.10.034.

Zhai, Yaqian; Yao, Yao; Guan, Qingfeng; Liang, Xun; Li, Xia; Pan, Yongting et al. (2020): Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata. In *International Journal of Geographical Information Science* 34 (7), pp. 1475–1499. DOI: 10.1080/13658816.2020.1711915.

Zheng, Wei; Shen, Geoffrey Qiping; Wang, Hao; Hong, Jingke; Li, Zhengdao (2017): Decision support for sustainable urban renewal: A multi-scale model. In *Land use policy* 69, pp. 361–371. DOI: 10.1016/j.landusepol.2017.09.019.