

Computers, Environment and Urban Systems
Simulating large-scale urban land-use patterns and dynamics using the U-Net deep learning architecture
 --Manuscript Draft--

Manuscript Number:	CEUS-D-21-00877
Article Type:	Research Paper
Keywords:	Deep Learning; Urban expansion simulation; Cellular automata; Spatial pattern
Corresponding Author:	jinzhu wang Deakin University - Melbourne Burwood Campus AUSTRALIA
First Author:	jinzhu wang
Order of Authors:	jinzhu wang Michalis Hadjikakou Richard J. Hewitt Brett A. Bryan
Abstract:	Cellular automata (CA) models are widely used to simulate real-world urban dynamics, but many studies have oversimplified the spatial complexity of the driving factors in the process of objectifying urban transition functions and abstracting urban development rules. Advances in deep learning technology enable complex urban spatial patterns and dynamics to be captured. In this study, we selected the U-Net deep learning algorithm to assimilate historical urban development, describe the pattern-extraction process, validate the model against a reference map, and apply the model to predict urban layouts for 2030 in the North China Plain. The results showed that: 1) U-Net can gradually abstract high-level spatial features and refine those patterns into precise urban development shapes, 2) the Figure of Merit (FoM) of the simulation map was close to those of previous studies that went through a lengthy calibration process, and 3) the landscape patterns of the simulation and reference maps were well aligned. U-Net was able to learn complex urban development patterns such as gravity effects and linear development, which can complement CA models to integrate spatial features to project future urban development.
Suggested Reviewers:	Haijun Wang landgiswhj@163.com He integrated historical urban development with CA models to simulate urbanization, which enables him to judge my paper from the perspective of innovative urban simulation. Keith C. Clarke kcclarke@ucsb.edu He calibrated the SLEUTH model to project land use to 2100. So he is capable of criticising my research regarding the improvement of urban simulation. Markus Reichstein mreichstein@bgc-jena.mpg.de He published a paper on geospatial big data and machine learning on Nature, he can evaluate my study of using deep learning to simulate urban land use.
Opposed Reviewers:	



13th Dec 2021

To the Editor-in-Chief, *Computers, Environment and Urban Systems*

Dear Professor T. H. Grubesic,

We are delighted to submit our manuscript entitled "***Simulation of large-scale urban land-use using U-Net***" for consideration for publication as a **Research Article** in *Computers, Environment and Urban Systems*. The primary objective is to simulate urban dynamics with deep learning (U-Net model) technology. We described and assessed the pattern-identification process of the U-Net algorithm and discussed the advantages and limitations of applying advanced deep learning in the simulation of urban development.

Previous Cellular Automata (CA) models are built on transition suitability, neighborhood status, constraint variables, and stochastic factors (Roodposhti, Hewitt, & Bryan, 2020; Wang, Guo, Zhang, & Zeng, 2021). Such models are flexible given the large array of parameter settings. However, 1) the difficulty in calibrating the many parameter choices (Feng & Tong, 2020) and 2) the oversimplifying of the spatial complexity to the driving factors (Gao et al., 2020; Newland, Zecchin, Maier, Newman, & van Delden, 2018) challenges their ability to mimic complex urban dynamics. More recent studies applied machine learning algorithms, Convolutional Neuron Network (CNN), or special simulation techniques (e.g., geographical zoning, temporal profile incorporation) to retrieve transition rules automatically or integrate complex spatial features to simulate urban development. But those studies did not take full advantage of the pattern-recognition ability (e.g., reconstruct a human face from linear, circular, or triangular (Krizhevsky, Sutskever, & Hinton, 2017)) of Deep Learning technology.

We applied a complete deep learning model (U-Net) to automatically simulate the large-scale urban dynamics. U-Net is a unique deep learning structure that first abstracts low-level features (lines, circles, or triangles, etc.) into high-level features (human faces, fruits, or animals), and then reconstruct the high-level features to precise shapes/patterns using the learned low-level features (Ronneberger, Fischer, & Brox, 2015). U-Net has been deployed in multiple geographical applications (Ji, Wei, & Lu, 2019; Nezla, Mithun Haridas, & Supriya, 2021; Singh et al., 2021), but have seldom been used in urban dynamic simulation.

The simulation urban maps show that U-Net learned the gravitational effects and linear development of urban dynamics. The landscape metrics reveal that the shapes of the urban patch in the simulation map are similar to the reference map. The accuracy metrics indicate that the simulation map has a close performance to previous CA models that went through a length calibration process. U-Net succeeded in identifying urban dynamic patterns automatically, with a limitation of difficult to explain the transition rules.

This research has not been published, accepted for publication, or under consideration for publication in another journal or book. All authors have agreed to the submitted version of the manuscript, and all persons entitled to authorship have been so named.

Jinzu WANG

Jinzu WANG, on behalf of all co-authors

Centre for Integrative Ecology, School of Life and Environmental Sciences &
Deakin-SWU Joint Research Centre on Big Data, Faculty of Science, Engineering and Built Environment,
Deakin University, VIC 3125, Australia

Email: wangjinz@deakin.edu.au

This study is instructed by **Brett A. Bryan**,

Centre for Integrative Ecology, School of Life and Environmental Sciences

Deakin University, VIC 3125, Australia

Email: b.bryan@deakin.edu.au

References

- Feng, Y., & Tong, X. (2020). A new cellular automata framework of urban growth modeling by incorporating statistical and heuristic methods. *International Journal of Geographical Information Science*, 34(1), 74–97. <https://doi.org/10.1080/13658816.2019.1648813>
- Gao, C., Feng, Y., Tong, X., Lei, Z., Chen, S., & Zhai, S. (2020). Modeling urban growth using spatially heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR. *Computers, Environment and Urban Systems*, 81, 101459. <https://doi.org/10.1016/j.compenvurbsys.2020.101459>

-
- Ji, S., Wei, S., & Lu, M. (2019). A scale robust convolutional neural network for automatic building extraction from aerial and satellite imagery. *International Journal of Remote Sensing*, 40(9), 3308–3322. <https://doi.org/10.1080/01431161.2018.1528024>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Newland, C. P., Zecchin, A. C., Maier, H. R., Newman, J. P., & van Delden, H. (2018). Empirically derived method and software for semi-automatic calibration of Cellular Automata land-use models. *Environmental Modelling & Software*, 108, 208–239. <https://doi.org/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, 28–33. <https://doi.org/10.1109/ICACCS51430.2021.9441804>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation, 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- Roodposhti, M. S., Hewitt, R. J., & Bryan, B. A. (2020). Towards automatic calibration of neighbourhood influence in cellular automata land-use models. *Computers, Environment and Urban Systems*, 79, 101416. <https://doi.org/10.1016/j.compenvurbsys.2019.101416>
- Singh, M., Kumar, B., Rao, S., Gill, S. S., Chattopadhyay, R., Nanjundiah, R. S., & Niyogi, D. (2021). Deep learning for improved global precipitation in numerical weather prediction systems.
- Wang, H., Guo, J., Zhang, B., & Zeng, H. (2021). Simulating urban land growth by incorporating historical information into a cellular automata model. *Landscape and Urban Planning*, 214, 104168. <https://doi.org/10.1016/j.landurbplan.2021.104168>

Highlights

- U-Net learns both large and fine scale spatial features of urban dynamic
- Transition rules are created automatically in the U-Net
- The Landscape Shape Index of simulated and actual urban maps are close
- U-Net achieves similar accuracies with CA models with lengthen calibration

1 **Simulating large-scale urban land-use patterns and dynamics using**
2 **the U-Net deep learning architecture**

3

4 Jinzhu Wang^{a,b*}, Michalis Hadjikakou^a, Richard J.Hewitt^{c,d,e}, Brett A. Bryan^{a,b}

5 ^a *Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin*
6 *University, VIC 3125, Melbourne, Australia.*

7 ^b *Deakin-SWU Joint Research Centre on Big Data, Faculty of Science, Engineering and Built*
8 *Environment, Deakin University, VIC 3125, Australia.*

9 ^c *Transport, Infrastructure, and Territory Research Group (tGIS), Geography Department,*
10 *Faculty of Geography and History, Universidad Complutense de Madrid (UCM), C/*
11 *ProfesorAranguren, s/n, Ciudad Universitaria, 28040, Madrid, Spain.*

12 ^d *Observatorio para una Cultura del Territorio (OCT), Calle del Duque de Fernán Núñez 2,*
13 *1, 28012, Madrid, Spain.*

14 ^e *Informational and Computational Sciences Group, The James Hutton Institute,*
15 *Craigiebuckler, Aberdeen AB15 8QH, Scotland UK.*

16 *corresponding author at wangjinz@deakin.edu.au; Centre for Integrative Ecology, School
17 of Life and Environmental Sciences, Deakin University, VIC 3125, Melbourne, Australia;

18 **Abstract**

19 Cellular automata (CA) models are widely used to simulate real-world urban dynamics, but many
20 studies have oversimplified the spatial complexity of the driving factors in the process of objectifying
21 urban transition functions and abstracting urban development rules. Advances in deep learning
22 technology enable complex urban spatial patterns and dynamics to be captured. In this study, we
23 selected the U-Net deep learning algorithm to assimilate historical urban development, describe the
24 pattern-extraction process, validate the model against a reference map, and apply the model to
25 predict urban layouts for 2030 in the North China Plain. The results showed that: 1) U-Net can
26 gradually abstract high-level spatial features and refine those patterns into precise urban
27 development shapes, 2) the Figure of Merit (FoM) of the simulation map was close to those of
28 previous studies that went through a lengthy calibration process, and 3) the landscape patterns of
29 the simulation and reference maps were well aligned. U-Net was able to learn complex urban
30 development patterns such as gravity effects and linear development, which can complement CA
31 models to integrate spatial features to project future urban development.

32 **Keywords:** Deep Learning; Urban expansion simulation; Cellular automata; Spatial pattern

1 2 Simulating large-scale urban land-use patterns and dynamics using 1 3 the U-Net deep learning architecture 1 4

1 4 1. Introduction

1 5 Urbanization is a complex process that is influenced by a range of social, cultural, economic,
1 6 geographic, environmental, and political factors (P. Fan et al., 2018; Kipfer, 2018; Shaw, van Vliet, &
1 7 Verburg, 2020; Yeh & Chen, 2020). Understanding future urban patterns and their dynamics is
1 8 essential for ensuring sustainable development due to its profound environmental and socio-
1 9 economic impacts (Newland, Zecchin, Maier, Newman, & van Delden, 2018; Zheng, Shen, Wang,
1 10 Hong, & Li, 2017). Modeling urban patterns and dynamics has been undertaken in many parts of the
1 11 world as a basis for mitigating air pollution (H. Fan, Zhao, & Yang, 2020), habitat destruction (Planillo
1 12 et al., 2021), and loss of arable land (Qiu, Li, Tang, Chen, & Berry, 2020). Some urban modeling
1 13 studies have focused on participatory modeling and scenario analysis for engaging stakeholders and
1 14 experts in the modeling process to balance competing interests and facilitate co-learning processes
1 15 related to future urban development (Clarke & Johnson, 2020; Mansour, Al-Belushi, & Al-Awadhi,
1 16 2020; Peng et al., 2020). However, most urban modeling studies have focused primarily on the
1 17 accurate prediction of the future spatial layout of cities based on historical dynamics and changes in
1 18 key driving forces under various scenarios (Gantumur et al., 2020; Shafizadeh-Moghadam, Asghari,
1 19 Tayyebi, & Taleai, 2017). Despite significant recent advances in urban modeling, accurately capturing
1 20 the complex spatial dynamics, patterns, and stochasticities of cities remains a significant challenge
1 21 (Liu, Batty, Wang, & Corcoran, 2021).

1 22 Cellular automata (CA) have been widely used to model future urbanization patterns because of
1 23 their ability to capture real-world urban patterns (Li, Gong, Le Yu, & Hu, 2017; Tong & Feng, 2020).
1 24 CA models are composed of the elements of transition suitability, neighborhood status, constraint
1 25 variables, and stochastic factors (Roodposhti, Hewitt, & Bryan, 2020; H. Wang, Guo, Zhang, & Zeng,
1 26 2021). Transition suitability refers to the rescaled biophysical, geographic, and socio-economic
1 27 driving factors that encourage or hinder the urbanization process (Feng & Tong, 2020).
1 28 Neighborhood status reflects the amount of urban or other land-use occurring in the immediate
1 29 vicinity of each cell, and is characterized according to different structures, sizes, and weights
1 30 (Roodposhti et al., 2020; Yu, Hagen-Zanker, Santitissadeekorn, & Hughes, 2021). Constraint variables
1 31 and stochastic factors regulate and randomize future urban development, respectively (Zhai et al.,
1 32 2020). Transition rules are a set of functions and parameters that control the scaling of the transition
1 33 suitability factors, the configuration of the neighborhood, the constraints, and the stochastic factors,
1 34 and then combine these elements into a spatial layer defining the probability of each cell becoming
1 35 urbanized in the future. While transition rules have typically been derived by trial and error or expert
1 36 knowledge, they are increasingly derived automatically to achieve the highest predictive accuracy.
1 37 Automatic rule extraction includes a suite of regression and machine learning (ML)-based methods
1 38 such as logistic regression (Mustafa et al., 2018), support vector machines (Kafy et al., 2021), tree-
1 39 based methods (Shafizadeh-Moghadam et al., 2017), neural networks (Gantumur et al., 2020),
1 40 heuristic methods (Carneiro & Oliveira, 2013), and dictionaries of trusted rules (Roodposhti, Aryal, &
1 41 Bryan, 2019). Although the flexibility of CA-based models with their large array of parameter settings
1 42 makes them ideal for participatory-based scenario modeling exercises, the difficulty in calibrating
1 43 the many parameter choices—still largely a manual process of trial and error—challenges their
1 44 ability to mimic complex urban dynamics and accurately capture future urban patterns.

45 More recent studies have adopted different techniques, such as geographical zoning, context
46 integration, and innovative algorithms, to increase the predictive accuracy of urban land-use
47 modeling. For example, some studies have subdivided their study area into separate regions,
48 allowing independent transition rule sets to be constructed to align with the specific conditions in
49 each zone (Qian, Xing, Guan, Yang, & Wu, 2020; Xia & Zhang, 2021). Many studies have incorporated
50 the shape and texture index to reflect the neighborhood spatial configurations of urban dynamics
51 (Ruiz Hernandez & Shi, 2018; Zhai et al., 2020). H. Wang et al. (2021) incorporated historical urban
52 development as a temporal process to simulate the future urban layout, while Peng et al. (2020)
53 integrated evolutionary and swarm algorithms to mimic urban dynamics. Despite having successfully
54 constructed transition rules, these studies identified significant mismatches when compared to real-
55 world urbanization: 1) introducing subjectivities via prescribed parameters (Liu et al., 2021), and 2)
56 oversimplifying the spatial heterogeneity of the driving factors (Gao et al., 2020; Mustafa et al.,
57 2018; Newland et al., 2018). Distances to and spatial configurations of geographical factors are key
58 to urban development in the real world. The distances used in these studies, however, are measured
59 by preset decay functions, and subjectivities are inevitably raised, given the gap between empirical
60 knowledge and real-world urban dynamics. The spatial heterogeneity of the urbanization process is
61 oversimplified to a single shape or texture index that capture only part of the pattern information
62 within the neighborhood (Akın & Erdoğan, 2020; Motlagh, Lotfi, Pourmanafi, Ahmadizadeh, &
63 Soffianian, 2020), let alone the large-scale spatial features (e.g., the layout of the whole built city
64 area) that drives urban development in the real world.

65 Deep learning algorithms are a type of machine learning technology that abstract naïve input
66 variables into high-level features. Convolutional neural networks (CNN) are a special kind of deep
67 learning architecture designed to detect spatial patterns from both low- (e.g., lines, triangles, circles)
68 and high-level features (e.g., human faces, visual patterns of different fruits or animals) (Krizhevsky,
69 Sutskever, & Hinton, 2017). CNN are increasingly being used to extract patterns and insights from
70 geospatial data (Reichstein et al., 2019), which enables the spatial configurations of driving factors,
71 rather than proxied variables such as decay distances, to be integrated directly into urban
72 development simulations. For example, Zhai et al. (2020) used CNN to retrieve the neighborhood
73 spatial features and improve the Figure of Merit (FoM) of their simulation from 0.323 to 0.361
74 compared to a random-forest-based method. Qian et al. (2020) reported similar results, i.e.,
75 improving the FoM from 0.299 to 0.346 compared to a random-forest algorithm. Although the
76 spatial features introduced by CNN improved the simulation performance, these studies took CNN as
77 an advanced decay function to rescale the driving factors and were therefore bound to detect only a
78 definite neighborhood spatial configuration, overlooking the ability of deep learning to abstract
79 high-level features. More advanced deep learning structures that take full advantage of CNN's
80 spatial pattern recognition ability, i.e., integrating low-level spatial features into high-level patterns,
81 have great potential for the simulation of future urban development.

82 U-Net, first introduced for biomedical image segmentation in 2015, is a unique type of CNN
83 architecture that not only abstracts spatial features to high-level patterns but also refines the high-
84 level patterns to precise shapes (Ronneberger, Fischer, & Brox, 2015). Its robust segmentation
85 performance enables U-Net to be used in multiple fields, such as improving global precipitation
86 estimation to improve weather prediction systems (Singh et al., 2021), detecting underwater objects
87 for oceanic ecosystem evaluation (Nezla, Mithun Haridas, & Supriya, 2021), and identifying buildings
88 from aerial and satellite imagery (Ji, Wei, & Lu, 2019). Unlike CA models that have a definite
89 neighborhood size, U-Net deploys a series of convolutional layers to extract spatial features and

90 then assimilates these features automatically to produce transition rules. This ability to learn spatial
1 91 patterns suggests that U-Net has the potential to identify and assimilate the spatial processes that
2 92 drive urban development and accurately capture the resulting patterns of cities.
3

4
5 93 In this study, we first assessed the ability of the U-Net deep learning architecture to accurately
6 94 project the spatial extent and pattern of urban development. We developed and applied U-Net to
7 95 project urban development in the North China Plain—China’s food bowl and one of the most rapidly
8 96 urbanizing areas on Earth. We next applied a U-Net model to learn the patterns of urban
9 97 development in the study area and used high-accuracy maps of urban land-use between 1993 and
10 98 2012 to train the model. We then projected the spatial distribution of urbanization for 2018 and
11 99 thoroughly tested the ability of the U-Net model to accurately capture the high-level spatial
12 100 features, shapes, and patterns of urban development. Finally, we used the trained model to project
13 101 urban development patterns in the study area for 2030 based on historical rates of urbanization. We
14 102 discuss both the advantages and limitations of U-Net for simulating urban development and, in
15 103 particular, its ability to learn typical urban development patterns such as neighborhood influences
16 104 and linear expansion along transport routes. We also discuss the implications of rapid urbanization
17 105 in the study area for food security and sustainability.
18
19

20 106

2. Methods

21

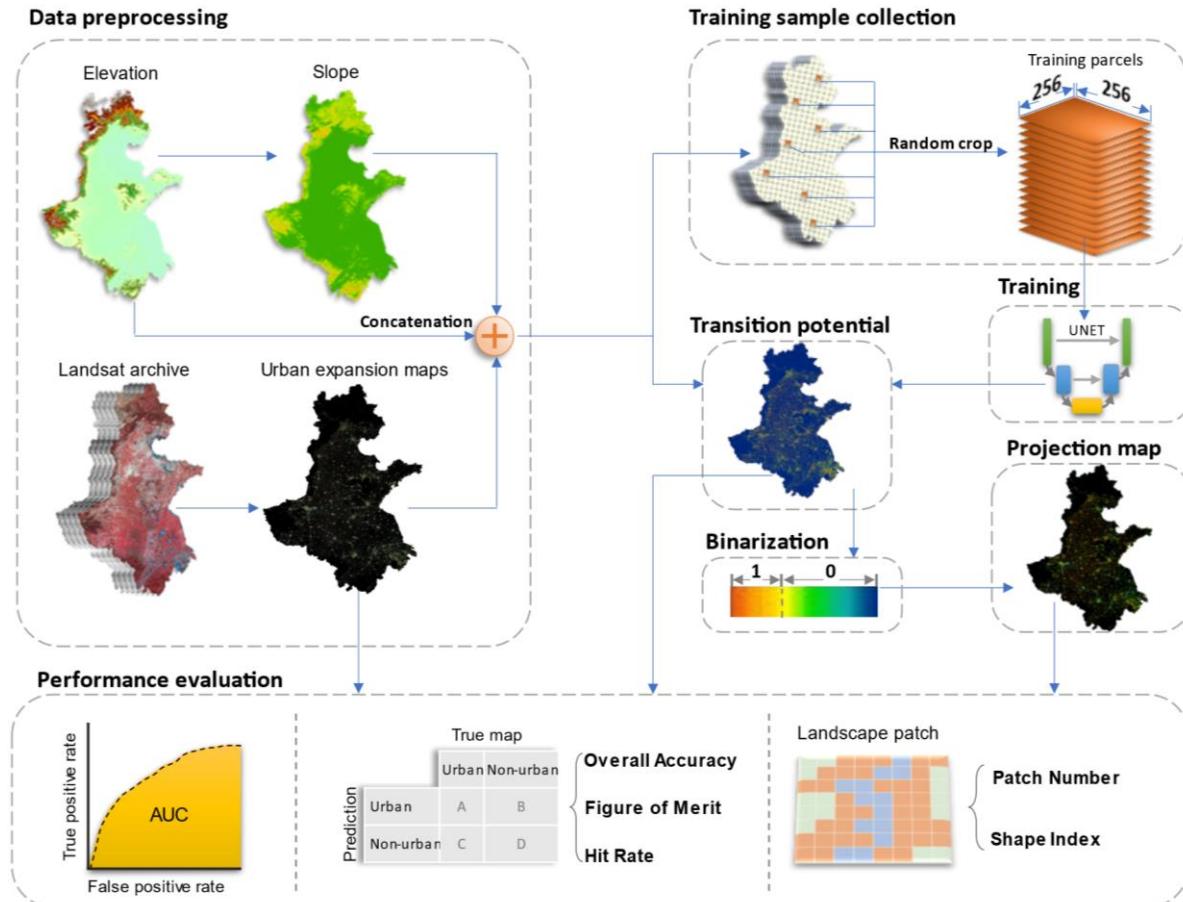
22 107

2.1. Method overview

23

24 108 We used Landsat data to map urban development in the study area for the years 1994, 2006, and
25 109 2018 (J. Wang, Hadjikakou, & Bryan, 2021) and combined them with elevation and slope information
26 110 to simulate urban developments (Qian et al., 2020; H. Wang et al., 2021; Xing, Qian, Guan, Yang, &
27 111 Wu, 2020). We created two U-Net models: one was trained with urban maps for the period 1994–
28 112 2006 and validated using the urban map for 2018; the other was trained on maps for the period
29 113 2006–2018 and was used to predict the urban layout in 2030. Both training phases (i.e., 1994–2006
30 114 and 2006–2018) are referred to as “*base*,” both prediction dates (2018 and 2030) are called “*target*,”
31 115 and the true 2018 urban map is called “*reference*” in this paper. Training samples were randomly
32 116 extracted and used to train U-Net. The trained model was then used to produce a transition
33 117 potential layer and create an urban land-use map. We then evaluated the accuracy of the transition
34 118 potential and urban land-use maps using a range of accuracy and pattern-based metrics. We
35 119 illustrate the use of the model in creating a future projection of urban land for 2030. The study
36 120 workflow is summarized and illustrated in Fig. 1 and described in more detail below.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

121

122
123 Fig. 1. Research workflow.
3233
124 2.2. Study area

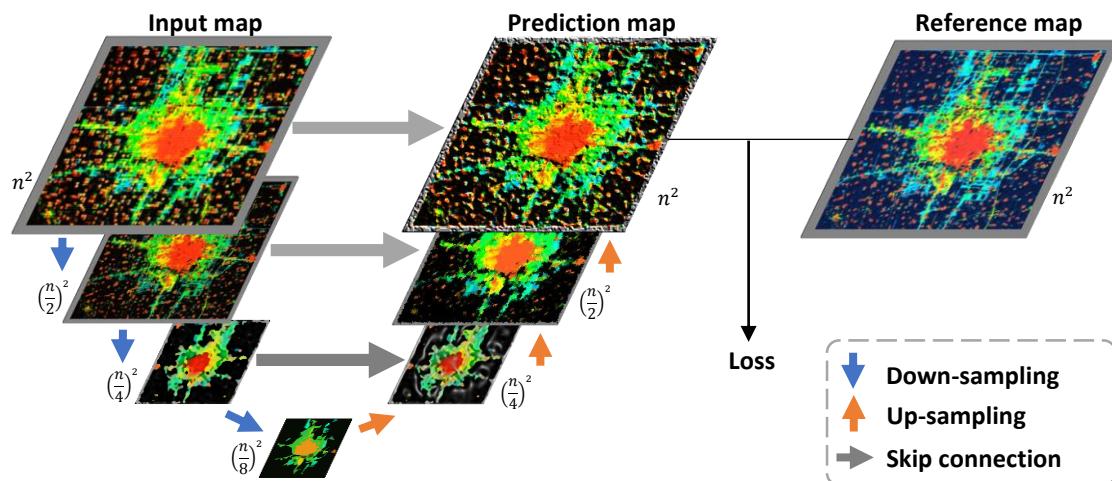
35
125 The North China Plain (Fig. 2) includes 76 prefectures, spans an area of more than 780,000 km², and
36 is home to more than 450 million people (National Bureau of Statistics of China, 2019b). This area is
37 one of the most rapidly urbanizing regions in China and the world, tripling the built-up land coverage
38 from approximately 5% in 1990 to approximately 15% in 2020 (J. Wang et al., 2021). This region is
39 crucial to China's economic development and holds a strategic role in safeguarding China's food
40 security, generating over one-third of its national gross domestic product (GDP) and grain supply
41 (National Bureau of Statistics of China, 2019a). Managing the tension between urbanization and
42 agricultural land-use in the study area requires accurate, spatially explicit projections of future urban
43 development to address the interconnected challenges of food security, environmental protection,
44 urbanization, and socio-economic development.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64



Fig. 2. Study area of the North China Plain.

2.3. Structure of U-Net

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, which balance the generalization of down-sampling and the refinement of up-sampling (Ronneberger et al., 2015). A conceptual U-Net structure (Fig. 3) demonstrates its pattern recognition capability.



144 Fig. 3. Conceptual structure of a four-layer U-Net. In this study, the input map was resized to half of its input size (e.g., n^2 to
 145 $\left(\frac{n}{2}\right)^2$) in each down-sampling layer, and then expanded to the original size in the up-sampling process. The loss denotes the
 146 difference between the prediction and reference maps, reflecting the performance of the U-Net.

147 The down-sampling layers include convolution, pooling, and rescaling processes (Table 1).
 148 Convolution is a cross-correlation operation that produces feature maps indicating the similarities
 149 between inputs and convolution filters (Kapinchev, Bradu, Barnes, & Podoleanu, 2015). By applying
 150 multiple filters, the different spatial patterns can be retrieved independently. For example, Zeiler
 151 and Fergus (2013) reported that horizontal, vertical, and circular patterns were identified by
 152 different filters applied to the input image. The pooling process reduces the size of inputs. We
 153 selected max-pooling to reduce dimensionality because it is adopted in most deep learning
 154 structures (Murray & Perronnin, 2014). The rescaling process rescales the pixel values of the feature
 155 maps to a specified range to optimize the computational flow of the network. We used a rectified
 156 linear unit activation function (ReLU) for its simple, efficient, and robust performance (Agarap,
 157 2018).

19 158 Table 1. Equations of the component layers of U-Net.

Layer	Equation	No.
Convolution	$C_{\text{out}j} = \text{bias}(C_{\text{out}}) + \sum_{k=0}^{C_{\text{in}}-1} \text{weight}(C_{\text{out}j}, k) * \text{input}(k)$	(1)
Max-pooling	$\text{out}(C_j, h, w) = \max_{m=0, \dots, kH-1} \max_{n=0, \dots, kW-1} \text{input}(C_j, \text{stride}[0] \times h + m, \text{stride}[1] \times w + n)$	(2)
ReLU	$\text{ReLU}(x) = \max(0, x)$	(3)
Softmax	$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_{j=1}^K \exp(x_j)}$	(4)
Cross-entropy	$\text{loss}(x, y) = -\log \left(\frac{\exp(y)}{\sum_{j=1}^K \exp(x[j])} \right)$	(5)

42 159 In equation (1), the sizes of the input and output images are (C_{in}, H, W) and $(C_{\text{out}}, H_{\text{out}}, W_{\text{out}})$, C denotes the number of
 43 160 channels, H is the height of the input planes in pixels, W is the width in pixels, $*$ is the valid cross-correlation operator, and j
 44 161 is the j -th channel of the output feature map. In equation (2), (kH, kW) denotes the kernel size of the pooling, h and w refer
 45 162 to the height and width of the output image, respectively. In equation (3), x denotes the pixel values of the input feature
 46 163 map. In equation (4), x_i is the i -th pixel value of the input feature map and K is the number of classes. In equation (5), x and
 47 164 y refer to the predicted and reference pixel values, respectively, and K is the number of classes.

48 165 The up-sampling layers include transpose convolution and rescaling processes. The transpose
 49 166 convolution, an inverse convolutive operation, transforms the input image from a lower resolution
 50 167 to a higher one. This process is assisted by skip-connections that bring additional spatial information
 51 168 from the down-sampling layers. The transpose convolution is similar to the convolution except that
 52 169 the output size is larger than the input size and the rescaling is the same ReLU operation in the
 53 170 down-sampling. Additional components used in the U-Net are batch normalization, softmax, and the
 54 171 cross-entropy algorithm (Table 1). Batch normalization is applied to standardize the weights that
 55 172 control the convolution and transpose convolution processes, which has been proven to be effective
 56 173 in improving deep learning performance (Sergey Ioffe & Christian Szegedy, 2015). The softmax

174 algorithm is applied to the last feature map of the U-Net to rescale the pixel values to the 0–1 range
 175 for a better comparison with the reference map (also composed of pixels of 0s and 1s). The cross-
 176 entropy algorithm is then applied to calculate the difference between the prediction and the
 177 reference map and reflects the performance of the U-Net.

178 The complete U-Net structure used in this study is shown in Fig. 4. The input image size was reduced
 179 by half after passing each down-sampling block until reaching a size of 8×8 pixels (i.e., 8^2 in Fig. 4),
 180 then the image was ultimately recovered to the original size through each up-sampling block. The
 181 channels of the feature maps underwent the inverse process: the number of feature maps was
 182 doubled in each down-sampling block and halved in each up-sampling block. As a result, the U-Net
 183 model gradually extracted more abstract spatial features over a larger field of view according to $k \times$
 184 $2^{(d-1)}$, where k is the kernel size (3 in this study) and d is the depth of the layer block. For example,
 185 the block with depth 1 had a field of view of 3, which became 96 in the block of depth 6, meaning
 186 that U-Net was looking for spatial patterns of 96×96 pixels in the original image scale. Meanwhile,
 187 U-Net identified more sophisticated spatial patterns as the network block went deeper because the
 188 number of feature maps was doubled. The skip-connection links the down-sampling and up-
 189 sampling blocks, allowing U-Net to refine the overview patterns extracted by deeper network blocks
 190 with precise shapes and textures retrieved by shallower blocks. A total of more than 31 million
 191 parameters were included in the U-Net model, making it highly flexible for capturing the spatial
 192 patterns and stochasticity of the urban dynamic process.

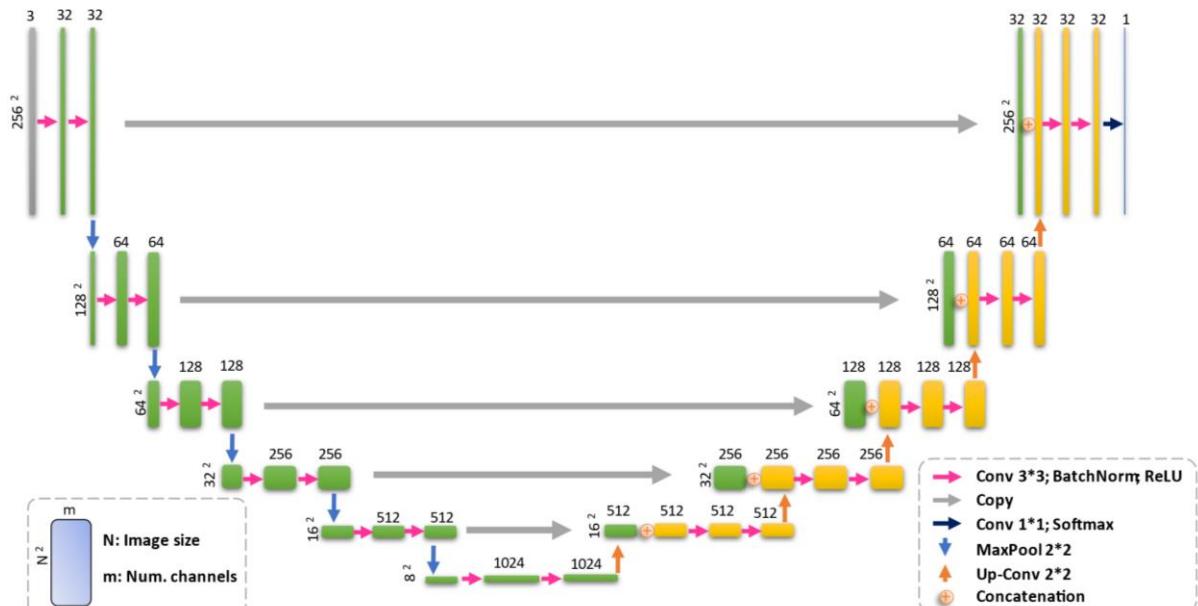


Fig. 4. Structure of the U-Net used in this study.

2.4. Data preprocessing

We mapped urban land-use (Fig. 5) for the years 1994, 2006, and 2018 with a consistently high (>94%) accuracy. Full details of this work can be found in J. Wang et al. (2021). The terrain data were obtained from the Shuttle Radar Topography Mission, from which slope and elevation data were derived. Elevation and slope information was used to assist 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 that has proven to be effective in previous studies (Qian et al., 2020; H. Wang et al., 2021; Xing et al., 2020). Accessibility variables, such as distance to roads and

203 railways, are frequently employed in urban dynamic modeling (Ronneberger et al., 2015; Tripathy &
1
2 Kumar, 2019; Valencia, Levin, & Hansen, 2020). However, U-Net is a pattern-sensitive model, and so,
3 in this case, distance factors provide little additional useful information. Hence, it was not used in
4 this study as it would have compromised model parsimony and led to a longer training time.
5



52
53 Fig. 5. Urban dynamic map of the North China Plain.
54

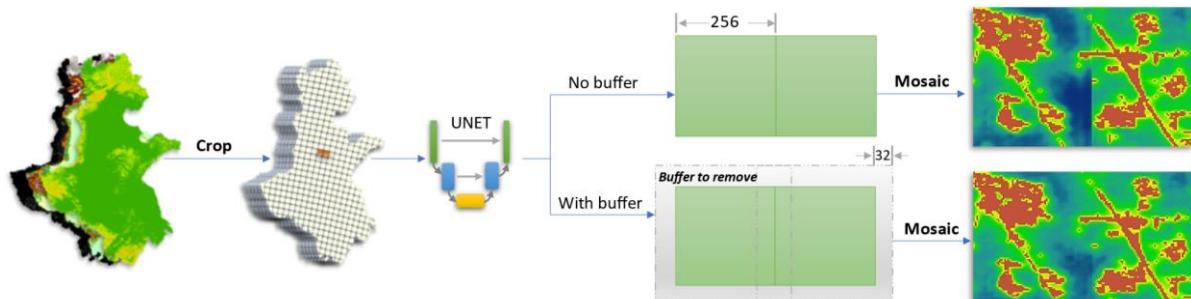
55 2.5. Training the U-Net model

56 Control samples were assembled using Google Earth Engine (Gorelick et al., 2017). The
57 *NeighborhoodToArray* module was used to crop 256 × 256 pixel tiled samples from the base image.
58 We set the tile size to 256 following common data science practices (Ronneberger et al., 2015) and
59
60
61

213 collected 20,000 samples for training and 5,000 for validation. The samples included three-layer
1 input data (e.g., urban map of 1994, elevation, and slope) and a single-layer later year image (e.g.,
2 urban map for 2006).
3
4
5 U-Net was trained for 200 epochs (an epoch refers to the U-Net model completely updating its
6 weights using all 20,000 training samples). We saved the model produced at each epoch and tested
7 its performance on the 5,000 validation samples. During the training process, the tiled input images
8 were resized to 8×8 pixels after five down-sampling operations and then converted to a single-layer
9 output image that was the same size as the original input image tile with another five up-sampling
10 operations (Fig. 4). During the validation process, the mean squared error (MSE) was used to
11 compute the difference between the output and target images (i.e., the loss) of the model. Finally,
12 the model with the lowest MSE was determined to be the best model for projecting the target urban
13 layout.
14
15
16
17

18 2.6. Producing the simulation map

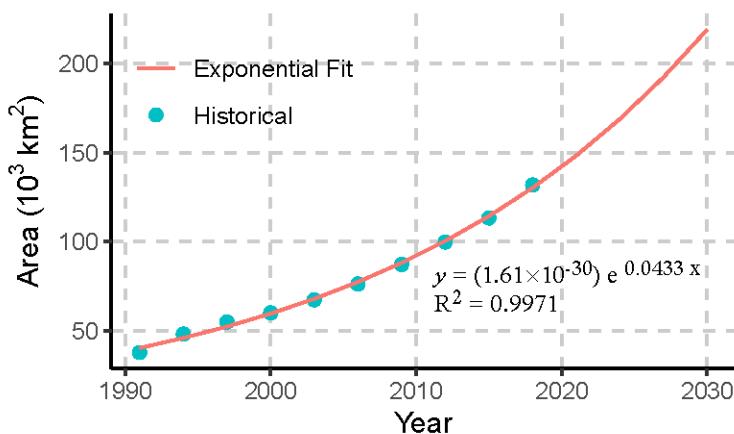
19
20 The input image was split into tiles, supplied to the trained U-Net model to produce separate
21 outputs, and then the outputs were mosaicked into a single transition potential map. The pixel
22 values of the transition potential map ranged from 0 to 1, indicating the probability of being an
23 urban pixel at the projection date. To reduce the tile edge effects, image tiles were cropped to a size
24 of 256×256 pixels and then supplied with a 32-pixel width buffer to be removed following
25 processing by U-Net (Fig. 6). Although the U-Net model was trained on image tiles of 256×256
26 pixels, it could process the buffered 320×320 (original size of 256 plus buffers at edges of size $32 \times$
27 2) image tiles because of the dimensional insensitivity of CNN structures (Long, Shelhamer, & Darrell,
28 2015).
29
30
31
32



33
34
35 Fig. 6. Buffering the image tile to reduce the edge effects. Because the final output was a mosaic of image tiles, we
36 buffered each image tile with an additional 32 pixels at the edges and removed these buffers before mosaicking to the final
37 output map to reduce the edge effect. A larger buffer size would better alleviate the edge effect but lead to a higher
38 computational cost; thus, we selected a 32-pixel buffer size as it is commonly used in deep learning structures that balance
39 effects and cost (Long et al., 2015).
40
41

42
43 A classified urban land-use map was created by binarizing the transition potential map. We ranked
44 the pixel values of the transition potential map from highest to lowest, accumulated the pixel count
45 in this order, and classified them to a value of 1 (indicating urban pixels) until the accumulated count
46 value matched the urban pixel number in the target year, then allocated the remaining pixels to a
47 value of 0 (indicating nonurban pixels). The urban pixel number for 2018 was computed from the
48 reference map to binarize the transition potential map for 2018, and an exponential extrapolation
49 (Fig. 7) was carried out to determine the urban pixel count for 2030 that was used to binarize the
50 transition potential map for 2030. The historical urban areas were derived using the data from J.
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

249 Wang et al. (2021), and the binarization was carried out independently at each prefecture to reduce
1 250 the bias caused by different regional development levels.
2
3



18 251 Fig. 7. Exponential regression on historical urban areas in the North China Plain. The historical urban area are computed
19 252 using the data from J. Wang et al. (2021).
20 253
21

22 254 2.7. Validation and accuracy assessment

23

24 255 We selected a series of accuracy metrics and spatial pattern metrics to assess the ability of U-Net to
25 256 accurately project the spatial distributions and patterns of urban areas in the study area. The
26 257 transition potential map was assessed using the area under the curve (AUC) of the receiver
27 258 operating characteristic curve (ROC), which illustrates its diagnostic ability to discriminate urban and
28 259 nonurban pixels under various thresholds (Fawcett, 2006). The predicted urban land-use map was
29 260 evaluated via map overlay and landscape-level spatial pattern metrics (McGarigal & Marks, 1995).
30 261 The map-overlay metrics selected in this study were overall accuracy (OA), hit rate, and FoM
31 262 because they quantified the urban simulation performance in both the prediction and the actual
32 263 perspectives. The spatial pattern metrics selected in this study were patch number (PN) and
33 264 landscape shape index (LSI) because they covered the general landscape features and were capable
34 265 of reflecting the shape patterns that delineate U-Net from CA models. All of the selected metrics are
35 266 described more fully in Table 2. Validation was performed for each prefecture independently,
36 267 yielding a total of 76 records for each metric.

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

268 Table 2. Validation metrics for evaluating the classified urban land-use maps.

Name	Equation	Explanation
AUC	$\int_{x=0}^1 TPR(FPR^{-1}(x))dx$	AUC measures a model's aggregated performance under different thresholds to discriminate urban and nonurban pixels (Tong & Feng, 2020).
OA	$(A + D) / (A + B + C + D)$	<i>Overall accuracy</i> is the ratio of correctly identified urban and nonurban pixels to the total number of predictions.
Hit rate	$A / (A + C)$	<i>Hit rate</i> is the ratio of correctly identified urban pixels (i.e., hits) to the number of urban pixels in the reference map.
FoM	$A / (A + B + C)$	<i>FoM</i> is the ratio of the intersection to the union upon overlaying predicted urban pixels with reference urban pixels (Pontius et al., 2008).
PN	n	<i>Patch number</i> is the number of patches in the urban landscape.
LSI	$\frac{\sum p_i}{4\sqrt{\sum a_i}}$	<i>Landscape shape index</i> reflects the complexity of urban landscape patches. For example, a squared patch is deemed simple (low value), whereas a linear patch is complex (high value).

269 Note: *TPR* is the true positive rate, *FPR* is the false positive rate of the ROC from the transition potential map. A represents
 270 the correctly predicted urban pixels (hit), B represents the incorrectly predicted urban pixels (false alarm), C represents the
 271 incorrectly predicted nonurban pixels (miss), and D represents the correctly predicted nonurban pixels (correct rejection). *n*
 272 is the total number of landscape patches, p_i is the perimeter length (m) of the *i*th path of the urban patches, and a_i is the
 273 area (ha) of the *i*th urban patch.

274 We first validated the U-Net model using the assessment metrics and then produced a future urban
 275 map for 2030. Urban land-use maps for 1994 and 2006 were used to train the U-Net model that
 276 simulated urban land-use for 2018, which was assessed, compared, and validated against the
 277 reference urban land-use map for 2018. Next, the other U-Net model was trained on the urban land-
 278 use maps for 2006 and 2018 to predict the future urban map for 2030.

42 279 3. Results

45 280 3.1. Model training

47 281 The MSE of the U-Net model trained on the urban images of 1994 and 2006 is shown in Fig. 8. The
 48 282 lowest MSE was 0.022 at the 70th epoch, which was determined as the best training epoch to
 50 283 produce the U-Net model that simulated the urban layout of 2018.

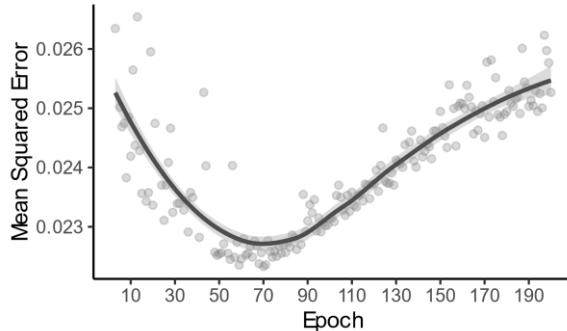


Fig. 8. Mean squared errors (MSEs) were obtained by applying the U-Net to historical urban land-use maps for the period 1994–2006 at different epochs. We saved the intermediate model of each epoch, and the dots represent the MSEs of such model on the validated samples. The line and ribbon represent the fitting and confidence interval (95%) of the dots, respectively, to better reflect the trend.

The U-Net successfully learned and captured different aspects of the complex spatial patterns of urban areas in the study area (Fig. 9). For example, in the illustrative sample presented in Fig. 9, the Down-1 layer broadly distinguished between urban and nonurban pixels. The Down-2 layer recognized simple patterns such as horizontal and vertical roads and learned to highlight larger towns among small villages. The Down-3 layer learned to associate adjacent urban clusters, and potential urban development corridors connecting discrete 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-sampled layers integrated both high-level and low-level features. The Up-4 layer refined the overview patterns in the Bottleneck layer by combining the latter with 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/nonurban feature maps in the Down-1 layer with the Up-2 layer, producing the final output image tile, which allocated more urban pixels around larger towns while maintaining the refined patterns.

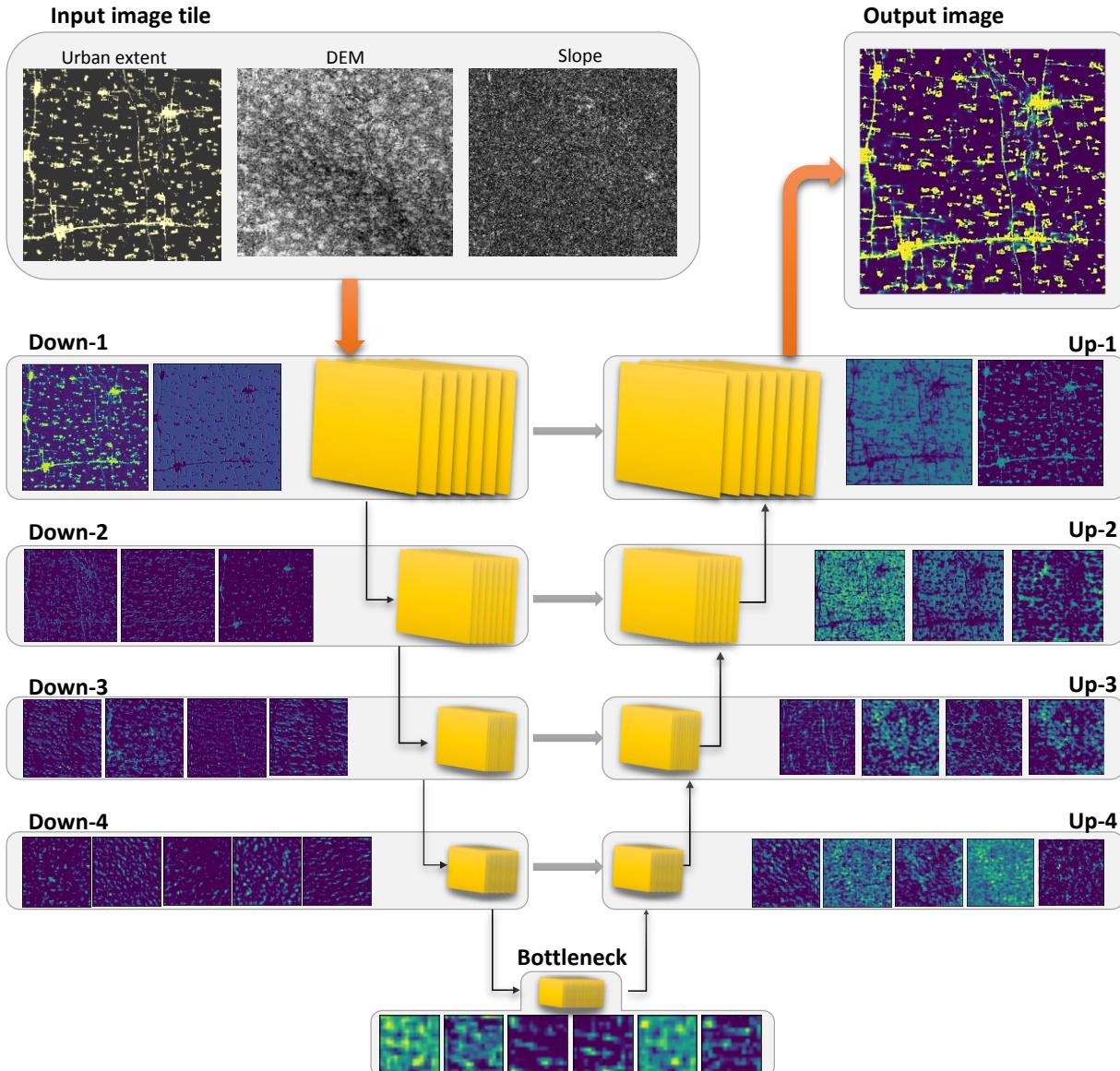


Fig. 9. Visualization of an image tile process by different U-Net layers. Note that we selected only a few activation maps for the visualization, given limited page space. The last activation map of each layer was used to visualize its pattern recognition capability.

3.2. Validation maps

The transition potential map is shown in Fig. 10. Pixels near larger towns and cities were associated with higher transition potential values than those of small villages. For example, a large area surrounding Sui'Xian was identified as having high transition potential, while such phenomena appeared only at the edges of smaller nearby villages. Similar patterns were found in Yang'Yuan and Du'Ling. Linear development patterns were also captured well in the transition potential map. For example, the roads in Yang'Kou Zhen and Hui'Ji Zhen were correctly identified despite only appearing as discrete linear segments in the initial urban map for 2006.

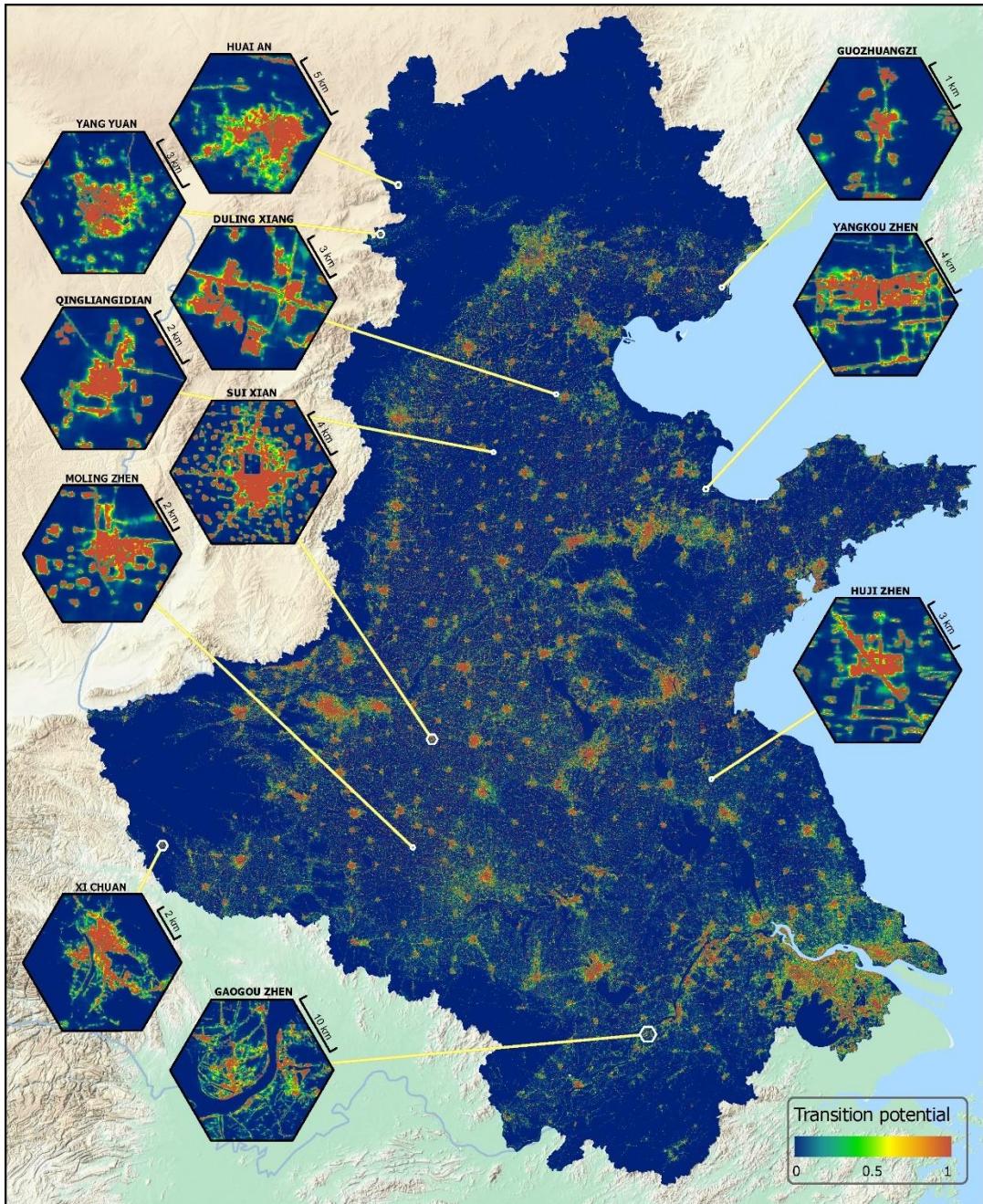


Fig. 10. Transition potential to urban in 2018 based on the 2006 urban map.

The classified urban land-use map (Fig. 11) created by binarizing the transition potential map correctly identified new urban areas (i.e., hits), particularly those adjacent to existing cities, demonstrating its ability to capture the effect of neighborhood on urban development. It also incorrectly identified some areas as new urban areas (i.e., false alarms) and failed to identify some new urban areas (i.e., misses), particularly in areas designated as new developments. Much of the missed urban projections occurred some distance from the original urban areas, such as the newly developed areas of land in Bin’Hai Zhen, Yang’Kou Zhen, and Xuan’Cheng City that were remote from the old town centers. Importantly, the model was exceptionally good at identifying urban expansion along linear features, such as the new development along major roads in Du’Ling Xiang, Gao’Gou Zhen, and Hui’Ji Zhen, despite the absence of explicit accessibility variables in the input dataset.

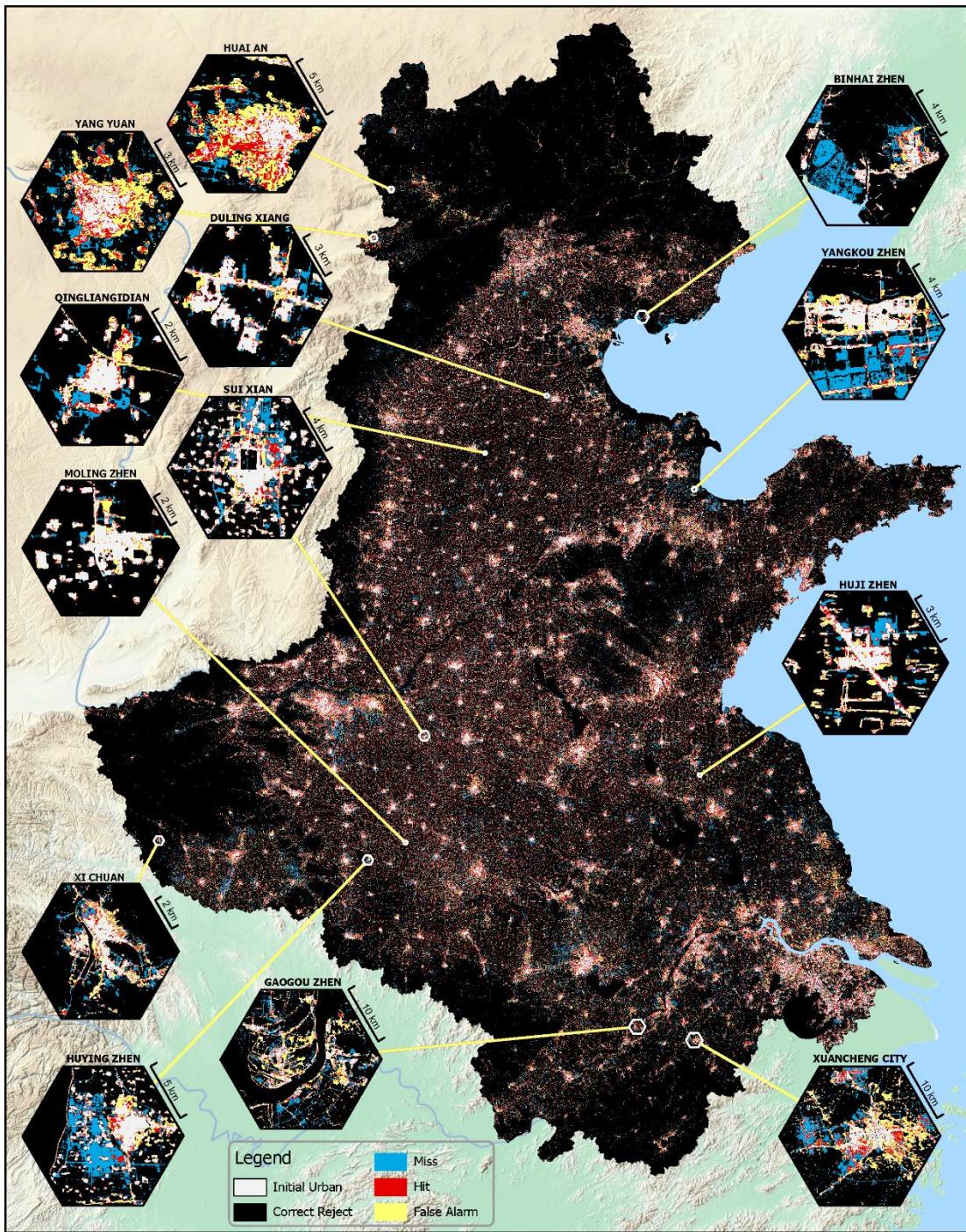


Fig. 11. Classified urban land-use map including hits, misses, and false alarms.

3.3. Validation metrics

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

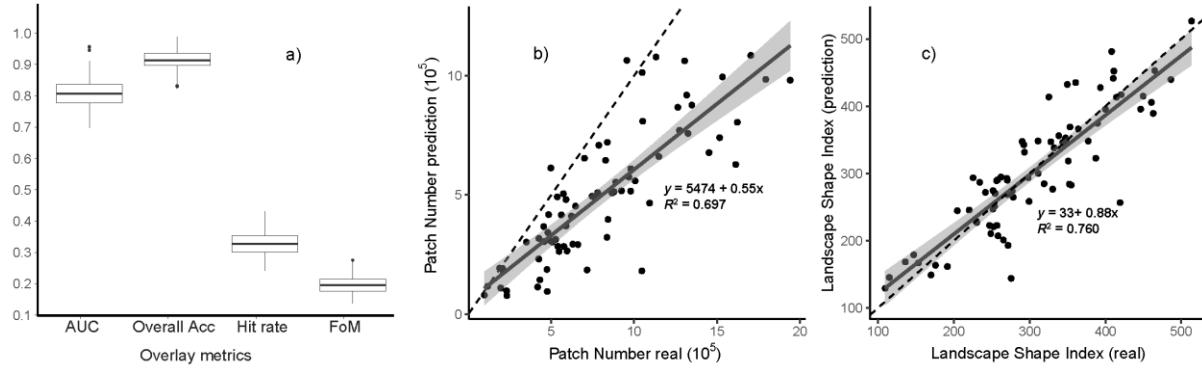


Fig. 12. Assessment metrics for the urban prediction in 2018: a) accuracy metrics; b) patch numbers; and c) landscape shape metrics. The dashed line refers to $y = x$ and the ribbon is the confidence interval (95%) of the linear fit to the scatter points.

The landscape-scale spatial pattern metrics of the classified urban land-use map compared to the reference map are shown in Fig. 12. The PN of the prediction map was systematically lower than that of the reference map, while the LSIs of both maps were very similar. This suggests that U-Net tended to predict fewer, more connected urban land patches than the real urban development but produced urban shapes very close to the real-world urban patterns.

3.4. Predicted urban area for 2030

The predicted urban map for 2030 is shown in Fig. 13. The newly predicted urban area tended to be concentrated around large cities. In SuiXian, for example, the predicted urban expansion was much larger around large cities/towns than that near smaller villages. The predicted urban area followed specified patterns rather than sprawling in every direction. For example, the predicted urban land areas in DuLing Xiang and YangKou Zhen infilled the gaps in the spatial distribution of existing urban areas and maintained the general shape of the original urban layout. The prediction for FengNing County followed the city's elongated development trend, and urban lands emerged along transport routes and at major road intersections, as illustrated by the cases of HuJi Zhen and GaoGou Zhen.

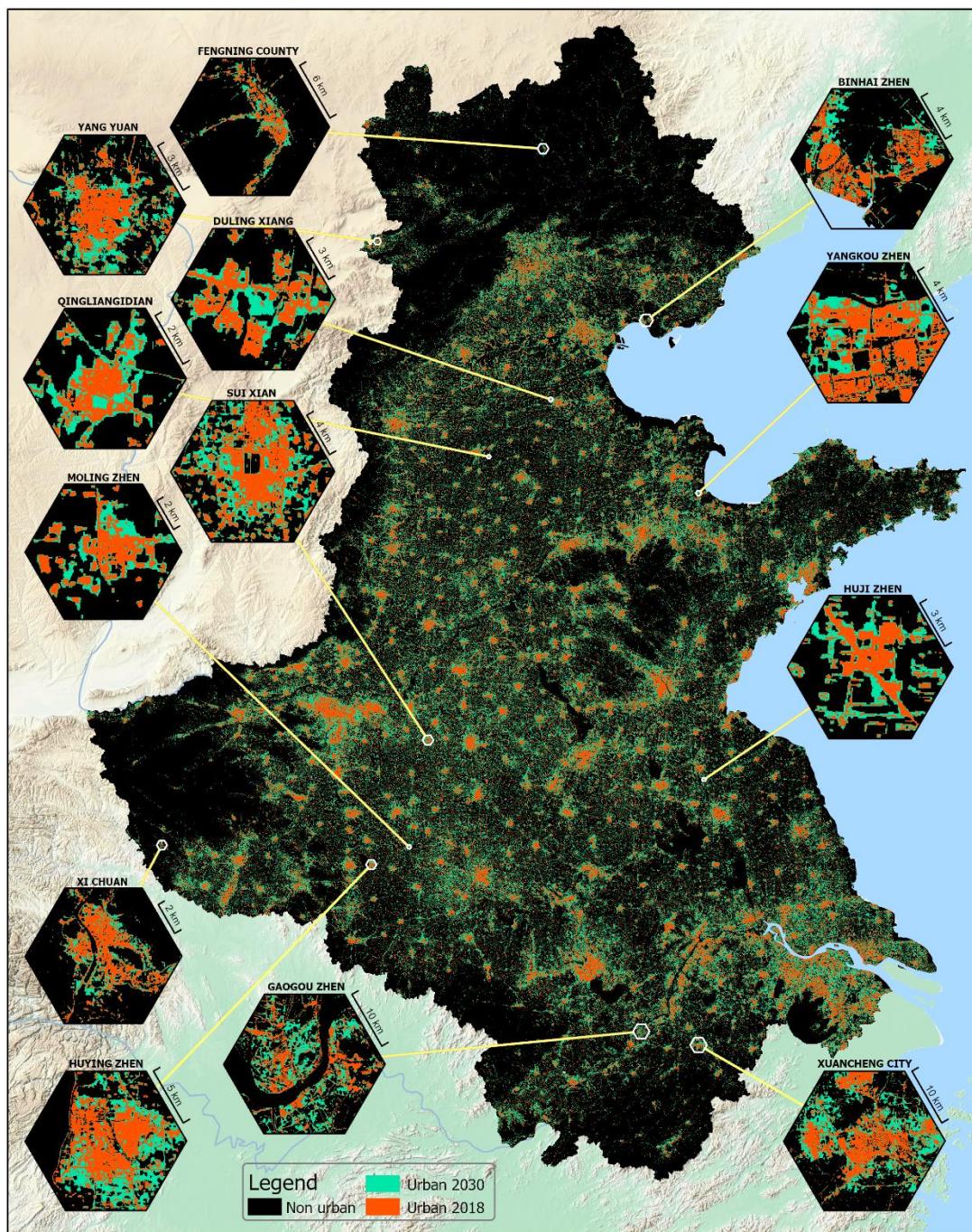


Fig. 13. Predicted urban land areas in 2030.

Predicted urban areas by province are listed in Table 3. The average rate of increase of urban areas was 35.54% compared to 2018. Beijing, Tianjin, and Hebei had the lowest area increase of <30%, while Anhui, Jiangsu, and Henan had the highest predicted area increases of >37%. Shandong was predicted to have a medium-level increase of 33.60%.

360 Table 3. Predicted extents of urban areas in each province in 2030.

Province	2018 (km²)	2030 (km²)	Increase (%)
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
Total	131578.48	178338.49	35.54

361 362 4. Discussion

363 4.1. Mimicking real-world urban patterns and dynamics

364 U-Net captured and assimilated high-level spatial patterns of urban development in the North China
 365 Plain. First, the model captured neighborhood effects. The transition potential map revealed that
 366 lands near existing urban areas were more likely to transition to urban land use in the future and
 367 that urbanization amongst larger expanses of nonurban land would be unlikely (Fig. 10). Second, U-
 368 Net was able to assimilate the learned small-scale neighborhood effects into large-scale gravity
 369 effects. Larger areas of land near larger cities were predicted to become urbanized in the future,
 370 while much smaller areas were predicted surrounding smaller, more remote villages (Fig. 11). Third,
 371 U-Net, rather than just simply buffering spatial features, captured the momentum of urban
 372 expansion as well as precise linear patterns. For example, the elongated development trend of
 373 FengNing County controlled by valley terrain was identified in the prediction (Fig. 13), and areas along
 374 major transport routes were allocated high transition potentials (Fig. 10). The U-Net model was
 375 unable to predict new planned developments that sprouted remote from old town centers, such as
 376 BinHai Zhen, YangKou Zhen, HuYing Zhen, and XuanCheng City (Fig. 11).

377 4.2. Reducing subjectivity by automatically constructing transition rules

378 The large-scale spatial patterns identified by U-Net would not have been captured by existing CA
 379 models. A large neighborhood is required for a CA model to incorporate information at a large scale.
 380 However, the simulation performance decreased as the neighborhood reached a specific size. H.
 381 Wang et al. (2021) found that the best neighborhood size to simulate urban development in Beijing
 382 was 25 × 25 pixels, while Roodposhti et al. (2020) observed that a 9 × 9 pixel neighborhood
 383 outperformed other settings. However, the unique design of the U-Net model is capable of
 384 incorporating more extensive information than previous CA model structures, enabling the
 385 neighbourhood influence to be learned from the data and urban development to be simulated with
 386 more refined spatial configurations.

387 U-Net did not require preset parameters to simulate urban dynamics, except for the learning rate
 388 and the size of the training images. The transition rules were automatically generated in the form of
 389 learned weights, and the training was an automatic process that reduced subjective biases. In
 390 contrast, CA model structures have often required a complex calibration process to determine
 391 suitable parameters. For example, Feng and Tong (2020) developed a framework that integrated

392 three algorithms and different neighborhood settings to simulate urban growth in Shanghai,
1 whereas an array of parameters, including inertia weights, decay magnitude, spatial heterogeneity,
2 and variable scaling, needed to be specified before running the model. Many studies determine
3 these parameters according to expert knowledge (Chen et al., 2020; Mustafa, Cools, Saadi, & Teller,
4 2017; Tripathy & Kumar, 2019), potentially leading to subjective biases. Other studies went through
5 a systematic parameter selection process to find the best parameters (Roodposhti et al., 2020; Yu et
6 397 al., 2021), but this process is time-consuming and impractical given the many undetermined
7 398 parameters.
8 399

11 400 4.3. Accurate prediction and robustness in capturing spatial patterns

12 401 The FoM in this study, ranging from 0.12 to 0.27, with a median value of 0.19, was similar to the
13 values reported in previous studies: the FoM of the best urban simulation map for Zhiji was 0.19 (R.
14 Wang et al., 2021) and 0.21 for Beijing (H. Wang et al., 2021). Several studies reported a higher FoM
15 because urban development was simulated with a coarser resolution (Peng et al., 2020; Pontius et
16 al., 2008; Valencia et al., 2020) or the simulation was performed on a smaller area (Pramanik,
17 Butsch, & Punia, 2021). Landscape metrics revealed that the simulation map produced by U-Net
18 aligned well with the reference map. The LSI of the simulated and the reference maps were very
19 similar and the linear regression slope was close to 1, indicating the suitability of the approach for
20 capturing the subtleties in urban transition patterns (Fig. 12). Such correlation was also embodied in
21 the classified urban maps, where linear structures such as roads and the general development trend
22 of cities were captured in the simulation map (Fig. 11). Although it is difficult to compare landscape
23 metrics across different studies because the metrics are susceptible to the specific settings of each
24 study, many studies have simulated urban land-use patterns which tend to simply buffer existing
25 urban areas and transport routes, rather than mimic the specific and complex shapes and patterns
26 of urban development (Gao et al., 2020; Pérez-Molina, Sliuzas, Flacke, & Jetten, 2017; Shafizadeh-
27 Moghadam et al., 2017).

36 417 4.4. Predicting urban land-use changes for formulating policy

37 418 A total area of $4.67 \times 10^4 \text{ km}^2$ is predicted to be urbanized in the North China Plain between 2018
38 and 2030 (Table 3). The three provinces with the highest rates of urban area increase (Anhui,
39 Jiangsu, and Henan) were all in the southern part of the region (Fig. 2). Our projections for 2030 also
40 capture urban development in the three megacity groups (Beijing-Tianjin-Hebei, the Yangtze River
41 Delta, and the Central Plains) that account for one-third of China's GDP (National Bureau of Statistics
42 of China, 2019b). The strategic development plan for China can benefit from the predictions arising
43 out of this study in identifying potential socioeconomic hazards originating from projected urban
44 expansion, and our simulation can assist policy formulation that is tailored to the expected rate,
45 location, and patterns of urbanization. The competition between urban development and food
46 production in the study area is a critical issue for China's food security (Jin et al., 2019). Pressure on
47 arable land could be alleviated by making use of our more granular predictions of future urban
48 development because they enable the development of plans/policies to reduce food production
49 losses that would otherwise arise from urban expansion (Jin et al., 2019). Ecosystem protection,
50 biodiversity, and environmental conservation can also greatly benefit from our urbanization
51 projections. For example, urban land is an important proxy for domestic water consumption
52 (Hoekstra, Buurman, & van Ginkel, 2018) and pollution (Zeller, Towa, Degrez, & Achteren, 2019), which
53 can be quantified based on urban maps. Finally, future urban layouts provide a baseline for local
54 433
55 434
56 435
57 436
58 437
59 438
60 439
61 440
62 441
63 442
64 443
65 444

435 governments to address land-use conflicts and settle competing interests among different
1 436 stakeholders in the urbanization process.

2

3

4 **437 4.5. Limitations and prospects**

5 438 U-Net learned the spatial patterns from historical urban dynamics, while previous CA methods
6 439 modeled urban development from the influence of the driving variables. As a result, U-Net has a
7 440 limited capacity to support participatory modeling that requires a series of 'what-if' trials
8 441 undertaken with stakeholders involving the tweaking of the model parameters and driving factors
9 442 and inspection of the results. Conversely, the U-Net model focuses on predictive accuracy rather
10 443 than explaining the urbanization process or supporting participatory modelling. The transition rules
11 444 of U-Net are hidden in its large number of learned weights, whereas CA models are usually
12 445 explainable and intuitive, i.e., the key driving factors can be identified. Hence, the utility of U-Net as
13 446 a participatory planning tool, where co-learning of the driving factors of urbanization is a key
14 447 objective, may be limited.

15
16
17
18
19
20 448 Another uncertainty is that U-Net might not be the most efficient/accurate deep learning model to
21 449 capture urban dynamic patterns. Many new advanced deep learning structures built on computer
22 450 vision technology could be used to retrieve a wide spectrum of features (e.g., distance to roads,
23 451 and/or the temporal characteristics in historical urban development) and improve the simulation
24 452 performances. There is significant potential to explore the ability of other deep learning
25 453 architectures to accurately model the spatial distribution and pattern of urban development.

26

27

28

29

30 **454 5. Conclusion**

31

32 455 In this study, we applied U-Net to simulate urban development in the North China Plain. The U-Net
33 456 model successfully captured neighborhood effects, gravity effects, and linear expansion along
34 457 transportation routes in the urban dynamics. The shapes of simulated urban land-use maps were
35 458 well aligned with the reference map, meaning that U-Net was capable of simulating urban dynamics
36 459 at a granular level. The better identification of urban patterns and dynamics is a useful addition to
37 460 urban land-use simulation studies to incorporate the various complex spatial driving factors. The
38 461 resulting projected 2030 urban land-use map provides key information for planning China's strategic
39 462 socioeconomic development and can benefit policy formulation and decision-making concerning
40 463 food security, biodiversity conservation, and environmental protection.

41

42

43

44

45

46 **464 References**

- 47
- 48
- 49 465 Agarap, A. F. (2018, March 22). *Deep Learning using Rectified Linear Units (ReLU)*.
- 50 466 Akin, A., & Erdoğan, M. A. (2020). Analysing temporal and spatial urban sprawl change of Bursa city
51 467 using landscape metrics and remote sensing. *Modeling Earth Systems and Environment*, 6(3),
52 468 1331–1343. <https://doi.org/10.1007/s40808-020-00766-1>
- 53
- 54 469 Carneiro, M. G., & Oliveira, G. M. B. (2013). Synchronous cellular automata-based scheduler
55 470 initialized by heuristic and modeled by a pseudo-linear neighborhood. *Natural Computing*, 12(3),
56 471 339–351. <https://doi.org/10.1007/s11047-013-9375-8>
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65

- 472 Chen, S., Feng, Y., Tong, X., Liu, S., Xie, H., Gao, C., & Lei, Z. (2020). Modeling ESV losses caused by
1 urban expansion using cellular automata and geographically weighted regression. *The Science of*
2 *the Total Environment*, 712, 136509. <https://doi.org/10.1016/j.scitotenv.2020.136509>
- 3
- 4 Clarke, K. C., & Johnson, J. M. (2020). Calibrating SLEUTH with big data: Projecting California's land
5 use to 2100. *Computers, Environment and Urban Systems*, 83, 101525.
6 <https://doi.org/10.1016/j.compenvurbsys.2020.101525>
- 7
- 8 Fan, H., Zhao, C., & Yang, Y. (2020). A comprehensive analysis of the spatio-temporal variation of
9 urban air pollution in China during 2014–2018. *Atmospheric Environment*, 220, 117066.
10 <https://doi.org/10.1016/j.atmosenv.2019.117066>
- 11
- 12 Fan, P., Chen, J., Ouyang, Z., Groisman, P., Loboda, T., Gutman, G., . . . Qi, J. (2018). Urbanization and
13 sustainability under transitional economies: a synthesis for Asian Russia. *Environmental Research*
14 *Letters*, 13(9), 95007. <https://doi.org/10.1088/1748-9326/aadb8>
- 15
- 16 Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874.
17 <https://doi.org/10.1016/j.patrec.2005.10.010>
- 18
- 19 Feng, Y., & Tong, X. (2020). A new cellular automata framework of urban growth modeling by
20 incorporating statistical and heuristic methods. *International Journal of Geographical Information*
21 *Science*, 34(1), 74–97. <https://doi.org/10.1080/13658816.2019.1648813>
- 22
- 23 Gantumur, B., Wu, F., Vandansambuu, B., Tsegmid, B., Dalaibaatar, E., & Zhao, Y. (2020).
24 Spatiotemporal dynamics of urban expansion and its simulation using CA-ANN model in
25 Ulaanbaatar, Mongolia. *Geocarto International*, 1–16.
26 <https://doi.org/10.1080/10106049.2020.1723714>
- 27
- 28 Gao, C., Feng, Y., Tong, X., Lei, Z., Chen, S., & Zhai, S. (2020). Modeling urban growth using spatially
29 heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR.
30 *Computers, Environment and Urban Systems*, 81, 101459.
31 <https://doi.org/10.1016/j.compenvurbsys.2020.101459>
- 32
- 33 Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth
34 Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202,
35 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- 36
- 37 Hoekstra, A. Y., Buurman, J., & van Ginkel, K. C. H. (2018). Urban water security: A review.
38 *Environmental Research Letters*, 13(5), 53002. <https://doi.org/10.1088/1748-9326/aaba52>
- 39
- 40 Ji, S., Wei, S., & Lu, M. (2019). A scale robust convolutional neural network for automatic building
41 extraction from aerial and satellite imagery. *International Journal of Remote Sensing*, 40(9),
42 3308–3322. <https://doi.org/10.1080/01431161.2018.1528024>
- 43
- 44 Jin, G., Chen, K., Wang, P., Guo, B., Dong, Y., & Yang, J. (2019). Trade-offs in land-use competition
45 and sustainable land development in the North China Plain. *Technological Forecasting and Social*
46 *Change*, 141, 36–46. <https://doi.org/10.1016/j.techfore.2019.01.004>
- 47
- 48 Kafy, A. - A., Naim, M. N. H., Subramanyam, G., Faisal, A.-A., Ahmed, N. U., Rakib, A. A., . . .
49 Sattar, G. S. (2021). Cellular Automata approach in dynamic modelling of land cover changes
50 using RapidEye images in Dhaka, Bangladesh. *Environmental Challenges*, 4, 100084.
51 <https://doi.org/10.1016/j.envc.2021.100084>
- 52
- 53 Kapinchev, K., Bradu, A., Barnes, F., & Podoleanu, A. (2015). GPU implementation of cross-
54 correlation for image generation in real time, 1–6. <https://doi.org/10.1109/ICSPCS.2015.7391783>
- 55
- 56 Kipfer, S. (2018). Pushing the limits of urban research: Urbanization, pipelines and counter-colonial
57 politics. *Environment and Planning D: Society and Space*, 36(3), 474–493.
58 <https://doi.org/10.1177/0263775818758328>
- 59
- 60
- 61
- 62
- 63
- 64
- 65

- 517 Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional
518 neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- 519 Li, X. [Xuecao], Gong, P., Le Yu, & Hu, T. (2017). A segment derived patch-based logistic cellular
520 automata for urban growth modeling with heuristic rules. *Computers, Environment and Urban*
521 *Systems*, 65, 140–149. <https://doi.org/10.1016/j.compenvurbsys.2017.06.001>
- 522 Liu, Y. [Yan], Batty, M., Wang, S., & Corcoran, J. (2021). Modelling urban change with cellular
523 automata: Contemporary issues and future research directions. *Progress in Human Geography*,
524 45(1), 3–24. <https://doi.org/10.1177/0309132519895305>
- 525 Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation,
526 3431–3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- 527 Mansour, S., Al-Belushi, M., & Al-Awadhi, T. (2020). Monitoring land use and land cover changes in
528 the mountainous cities of Oman using GIS and CA-Markov modelling techniques. *Land Use Policy*,
529 91, 104414. <https://doi.org/10.1016/j.landusepol.2019.104414>
- 530 McGarigal, K., & Marks, B. J. (1995). *Fragstats: Spatial pattern analysis program for quantifying*
531 *landscape structure: Spatial Pattern Analysis Program for Quantifying Landscape Structure*.
532 Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.
533 <https://doi.org/10.2737/PNW-GTR-351>
- 534 Motlagh, Z. K., Lotfi, A., Pourmanafi, S., Ahmadizadeh, S., & Soffianian, A. (2020). Spatial modeling of
535 land-use change in a rapidly urbanizing landscape in central Iran: Integration of remote sensing,
536 CA-Markov, and landscape metrics. *Environmental Monitoring and Assessment*, 192(11), 695.
537 <https://doi.org/10.1007/s10661-020-08647-x>
- 538 Murray, N., & Perronnin, F. (2014). *Generalized Max Pooling*. <https://doi.org/10.1109/cvpr.2014.317>
- 539 Mustafa, A., Cools, M., Saadi, I., & Teller, J. (2017). Coupling agent-based, cellular automata and
540 logistic regression into a hybrid urban expansion model (HUEM). *Land Use Policy*, 69, 529–540.
541 <https://doi.org/10.1016/j.landusepol.2017.10.009>
- 542 Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., & Teller, J. (2018). Modelling built-up
543 expansion and densification with multinomial logistic regression, cellular automata and genetic
544 algorithm. *Computers, Environment and Urban Systems*, 67, 147–156.
545 <https://doi.org/10.1016/j.compenvurbsys.2017.09.009>
- 546 National Bureau of Statistics of China (2019a). Announcement of the 2019 grain output. Retrieved
547 from http://www.gov.cn/xinwen/2019-12/07/content_5459250.htm
- 548 National Bureau of Statistics of China (2019b). *China Statistical Yearbook*. Beijing, China: China
549 Statistics Press.
- 550 Newland, C. P., Zecchin, A. C., Maier, H. R., Newman, J. P., & van Delden, H. (2018). Empirically
551 derived method and software for semi-automatic calibration of Cellular Automata land-use
552 models. *Environmental Modelling & Software*, 108, 208–239.
553 <https://doi.org/10.1016/j.envsoft.2018.07.013>
- 554 Nezla, N. A., Mithun Haridas, T. P., & Supriya, M. H. (2021). Semantic Segmentation of Underwater
555 Images using UNet architecture based Deep Convolutional Encoder Decoder Model, 28–33.
556 <https://doi.org/10.1109/ICACCS51430.2021.9441804>
- 557 Peng, K., Jiang, W., Deng, Y., Liu, Y. [Yinghui], Wu, Z., & Chen, Z. [Zheng] (2020). Simulating wetland
558 changes under different scenarios based on integrating the random forest and CLUE-S models: A
559 case study of Wuhan Urban Agglomeration. *Ecological Indicators*, 117, 106671.
560 <https://doi.org/10.1016/j.ecolind.2020.106671>

- 561 Pérez-Molina, E., Sliuzas, R., Flacke, J., & Jetten, V. (2017). Developing a cellular automata model of
562 urban growth to inform spatial policy for flood mitigation: A case study in Kampala, Uganda.
563 *Computers, Environment and Urban Systems*, 65, 53–65.
564 <https://doi.org/10.1016/j.compenvurbsys.2017.04.013>
- 565 Planillo, A., Kramer-Schadt, S., Buchholz, S., Gras, P., Lippe, M. von der, & Radchuk, V. (2021).
566 Arthropod abundance modulates bird community responses to urbanization. *Diversity and*
567 *Distributions*, 27(1), 34–49. <https://doi.org/10.1111/ddi.13169>
- 568 Pontius, R. G., Boersma, W., Castella, J.-C., Clarke, K., Nijs, T. de, Dietzel, C., . . . Verburg, P. H. (2008).
569 Comparing the input, output, and validation maps for several models of land change. *The Annals*
570 *of Regional Science*, 42(1), 11–37. <https://doi.org/10.1007/s00168-007-0138-2>
- 571 Pramanik, S., Butsch, C., & Punia, M. (2021). Post-liberal urban dynamics in India – The case of
572 Gurugram, the ‘Millennium City’. *Remote Sensing Applications: Society and Environment*, 22,
573 100504. <https://doi.org/10.1016/j.rsase.2021.100504>
- 574 Qian, Y., Xing, W., Guan, X., Yang, T., & Wu, H. (2020). Coupling cellular automata with area
575 partitioning and spatiotemporal convolution for dynamic land use change simulation. *The Science*
576 *of the Total Environment*, 722, 137738. <https://doi.org/10.1016/j.scitotenv.2020.137738>
- 577 Qiu, B., Li, H., Tang, Z., Chen, C., & Berry, J. (2020). How cropland losses shaped by unbalanced
578 urbanization process? *Land Use Policy*, 96, 104715.
579 <https://doi.org/10.1016/j.landusepol.2020.104715>
- 580 Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019).
581 Deep learning and process understanding for data-driven Earth system science. *Nature*,
582 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- 583 Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image
584 Segmentation, 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
- 585 Roodposhti, M. S., Aryal, J., & Bryan, B. A. (2019). A novel algorithm for calculating transition
586 potential in cellular automata models of land-use/cover change. *Environmental Modelling &*
587 *Software*, 112, 70–81. <https://doi.org/10.1016/j.envsoft.2018.10.006>
- 588 Roodposhti, M. S., Hewitt, R. J., & Bryan, B. A. (2020). Towards automatic calibration of
589 neighbourhood influence in cellular automata land-use models. *Computers, Environment and*
590 *Urban Systems*, 79, 101416. <https://doi.org/10.1016/j.compenvurbsys.2019.101416>
- 591 Ruiz Hernandez, I. E., & Shi, W. (2018). A Random Forests classification method for urban land-use
592 mapping integrating spatial metrics and texture analysis. *International Journal of Remote Sensing*,
593 39(4), 1175–1198. <https://doi.org/10.1080/01431161.2017.1395968>
- 594 Sergey Ioffe, & Christian Szegedy (2015). Batch Normalization: Accelerating Deep Network Training
595 by Reducing Internal Covariate Shift. *International Conference on Machine Learning*, 448–456.
596 Retrieved from <http://proceedings.mlr.press/v37/ioffe15.html>
- 597 Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017). Coupling machine learning,
598 tree-based and statistical models with cellular automata to simulate urban growth. *Computers,*
599 *Environment and Urban Systems*, 64, 297–308.
600 <https://doi.org/10.1016/j.compenvurbsys.2017.04.002>
- 601 Shaw, B. J., van Vliet, J., & Verburg, P. H. (2020). The peri-urbanization of Europe: A systematic
602 review of a multifaceted process. *Landscape and Urban Planning*, 196, 103733.
603 <https://doi.org/10.1016/j.landurbplan.2019.103733>

- 604 Singh, M., Kumar, B., Rao, S., Gill, S. S., Chattopadhyay, R., Nanjundiah, R. S., & Niyogi, D. (2021).
605 Deep learning for improved global precipitation in numerical weather prediction systems.
- 606 Tong, X., & Feng, Y. (2020). A review of assessment methods for cellular automata models of land-
607 use change and urban growth. *International Journal of Geographical Information Science*, 34(5),
608 866–898. <https://doi.org/10.1080/13658816.2019.1684499>
- 609 Tripathy, P., & Kumar, A. (2019). Monitoring and modelling spatio-temporal urban growth of Delhi
610 using Cellular Automata and geoinformatics. *Cities*, 90, 52–63.
611 <https://doi.org/10.1016/j.cities.2019.01.021>
- 612 Valencia, V. H., Levin, G., & Hansen, H. S. (2020). Modelling the spatial extent of urban growth using
613 a cellular automata-based model: a case study for Quito, Ecuador. *Geografisk Tidsskrift-Danish
614 Journal of Geography*, 120(2), 156–173. <https://doi.org/10.1080/00167223.2020.1823867>
- 615 Wang, H. [Haijun], Guo, J., Zhang, B., & Zeng, H. (2021). Simulating urban land growth by
616 incorporating historical information into a cellular automata model. *Landscape and Urban
617 Planning*, 214, 104168. <https://doi.org/10.1016/j.landurbplan.2021.104168>
- 618 Wang, J., Hadjikakou, M., & Bryan, B. A. (2021). Consistent, accurate, high resolution, long time-
619 series mapping of built-up land in the North China Plain. *GIScience & Remote Sensing*, 1–17.
620 <https://doi.org/10.1080/15481603.2021.1948275>
- 621 Wang, R., Feng, Y., Wei, Y., Tong, X., Zhai, S., Zhou, Y., & Wu, P. (2021). A comparison of proximity
622 and accessibility drivers in simulating dynamic urban growth. *Transactions in GIS*, 25(2), 923–947.
623 <https://doi.org/10.1111/tgis.12707>
- 624 Xia, C., & Zhang, B. (2021). Exploring the effects of partitioned transition rules upon urban growth
625 simulation in a megacity region: a comparative study of cellular automata-based models in the
626 Greater Wuhan Area. *GIScience & Remote Sensing*, 1–24.
627 <https://doi.org/10.1080/15481603.2021.1933714>
- 628 Xing, W., Qian, Y., Guan, X., Yang, T., & Wu, H. (2020). A novel cellular automata model integrated
629 with deep learning for dynamic spatio-temporal land use change simulation. *Computers &
630 Geosciences*, 137, 104430. <https://doi.org/10.1016/j.cageo.2020.104430>
- 631 Yeh, A. G.-O., & Chen, Z. [Zifeng] (2020). From cities to super mega city regions in China in a new
632 wave of urbanisation and economic transition: Issues and challenges. *Urban Studies*, 57(3), 636–
633 654. <https://doi.org/10.1177/0042098019879566>
- 634 Yu, J., Hagen-Zanker, A., Santitissadeekorn, N., & Hughes, S. (2021). Calibration of cellular automata
635 urban growth models from urban genesis onwards - a novel application of Markov chain Monte
636 Carlo approximate Bayesian computation. *Computers, Environment and Urban Systems*, 90,
637 101689. <https://doi.org/10.1016/j.compenvurbsys.2021.101689>
- 638 Zeiler, M. D., & Fergus, R. (2013, November 13). *Visualizing and Understanding Convolutional
639 Networks*. Retrieved from <https://arxiv.org/pdf/1311.2901>
- 640 Zeller, V., Towa, E., Degrez, M., & Achtern, W. M. J. (2019). Urban waste flows and their potential for
641 a circular economy model at city-region level. *Waste Management (New York, N.Y.)*, 83, 83–94.
642 <https://doi.org/10.1016/j.wasman.2018.10.034>
- 643 Zhai, Y., Yao, Y., Guan, Q., Liang, X., Li, X. [Xia], Pan, Y., . . . Zhou, J. (2020). Simulating urban land use
644 change by integrating a convolutional neural network with vector-based cellular automata.
645 *International Journal of Geographical Information Science*, 34(7), 1475–1499.
646 <https://doi.org/10.1080/13658816.2020.1711915>

647 Zheng, W., Shen, G. Q., Wang, H. [Hao], Hong, J., & Li, Z. (2017). Decision support for sustainable
1 648 urban renewal: A multi-scale model. *Land Use Policy*, 69, 361–371.
2 649 <https://doi.org/10.1016/j.landusepol.2017.09.019>
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65