

Q4 Simulating large-scale urban land-use patterns and dynamics using the U-Net deep learning architecture

 The corrections made in this section will be reviewed and approved by a journal production editor.

Jin Zhu Wang^{a,b,*}, wangjinz@deakin.edu.au, Michalis Hadjikakou^a, Richard J. Hewitt^{c,d,e}, Brett A. Bryan^{a,b}

^aCentre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, VIC 3125, Melbourne, Australia

^bDeakin-SWU Joint Research Centre on Big Data, Faculty of Science, Engineering and Built Environment, Deakin University, VIC 3125, Australia

^cTransport, Infrastructure, and Territory Research Group (tGIS), Geography Department, Faculty of Geography and History, Universidad Complutense de Madrid (UCM), C/ Profesor Aranguren, s/n, Ciudad Universitaria, 28040 Madrid, Spain

^dObservatorio para una Cultura del Territorio (OCT), Calle del Duque de Fernán Núñez 2, 1, 28012 Madrid, Spain

^eInformational and Computational Sciences Group, The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, Scotland, UK

Q6 *Corresponding author at: Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, VIC 3125 Melbourne, Australia

Abstract

Accurate predictions of land-use change are important for supporting planning. Cellular automata (CA) models are widely used to simulate real-world urban land-use change but accurately modelling modeling complex spatial urban patterns and dynamics can be challenging due to the high degree of subjectivity involved in CA model parameterisation. Advances in deep learning enable complex spatial patterns such as urban development to be learned and simulated. In this study, we used the U-Net deep learning algorithm to capture historical urban development and simulate future patterns for the North China Plain, one of the most rapidly urbanizing regions on the planet. We validated the model against a reference map for 2018 then applied it in predicting patterns of urban expansion for 2030. The results showed that U-Net can accurately predict urban land-use and mimic real-world spatial patterns with very low requirement for model parameterization and forcing data. Visual inspection of the outputs revealed that U-Net was able to automatically learn complex urban development patterns and processes such as neighbourhood neighborhood influence, the gravity effects of large cities, and the tendency for linear development. Deep learning architectures such as U-Net provide a new parameter-free way to accurately capture and simulate spatial features in projections of future urban development and land-use change.

Keywords:

Deep learning, Urban expansion simulation, Cellular automata, Spatial pattern

Abbreviations

No keyword abbreviations are available

Data availability

07 1 Introduction

Urbanization is a complex process that is influenced by a range of social, cultural, economic, geographic, environmental, and political factors (Fan et al., 2018; Kipfer, 2018; Shaw, van Vliet, & Verburg, 2020; Yeh & Chen, 2020). Understanding future urban patterns and their dynamics is essential for ensuring sustainable development due to the profound environmental and socio-economic impacts of urban land transformation (Newland, Zecchin, Maier, Newman, & van Delden, 2018; Zheng, Shen, Wang, Hong, & Li, 2017). Modeling urban patterns and dynamics has been undertaken in many parts of the world as a basis for mitigating air pollution (Fan, Zhao, & Yang, 2020), habitat

08 destruction (Planillo et al., 2021), and loss of arable land (Qiu, Li, Tang, Chen, & Berry, 2020). Some urban modeling studies have focused on participatory modeling and scenario analysis for engaging stakeholders and experts in the modeling process to balance competing interests and facilitate co-learning processes related to future urban development (Clarke & Johnson, 2020; Mansour, Al-Belushi, & Al-Awadhi, 2020; Peng et al., 2020). However, most urban land-use modeling studies have focused primarily on the accurate prediction of the future spatial layout of cities based on historical dynamics and changes in key driving forces under various scenarios (Gantumur et al., 2020; Shafizadeh-Moghadam, Asghari, Tayyebi, & Taleai, 2017). Despite significant recent advances in urban land-use modeling, accurately capturing the complex spatial dynamics, patterns, and stochasticities of cities remains a significant challenge (Liu, Batty, Wang, & Corcoran, 2021).

Cellular automata (CA) have been widely used to model future urbanization patterns because of their ability to incorporate real-world urbanization processes (Li, Gong, Le, & Hu, 2017; Tong & Feng, 2020). CA models consider multiple factors such as land suitability, neighborhood status, constraint variables, and stochastic factors (Roodposhti,

09 Hewitt, & Bryan, 2020; Wang et al., 2021). Land suitability refers to the rescaled biophysical, geographic, and socio-economic driving factors that influence urbanization (Feng & Tong, 2020). Neighborhood status reflects the amount of urban or other land-use occurring in the immediate vicinity of each cell and is characterized by different structures, sizes, and weights (Roodposhti et al., 2020; Yu, Hagen-Zanker, Santitissadeekorn, & Hughes, 2021). Constraint variables and stochastic factors regulate and randomize future urban development, respectively (Zhai et al., 2020). Transition rules are a set of functions and parameters that control the scaling of land suitability factors, configuration of the neighborhood, constraints, and 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. Automatic rule extraction has included a suite of regression and machine learning (ML) based methods such as logistic regression (Mustafa et al., 2018), support vector machines (Kafy et al., 2021), tree-based methods (Shafizadeh-Moghadam et al., 2017), neural networks (Gantumur et al., 2020), heuristic methods (Cameiro & Oliveira, 2013), and dictionaries of trusted rules (Roodposhti, Aryal, & Bryan, 2019). Although the flexibility of process-based CA-based models with their large array of parameter settings makes them ideal for participatory 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 land-use 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 have subdivided their study area into separate regions, allowing independent transition rule sets to be constructed to align with specific conditions in each zone (Qian, Xing, Guan, Yang, & Wu, 2020; Xia & Zhang, 2021). Many studies have incorporated the shape and texture index to reflect the neighborhood spatial configurations of urban land-use (Ruiz Hernandez &

10 Shi, 2018; Zhai et al., 2020). Wang, Feng, et al. (2021) incorporated historical urban development as a temporal process to simulate future urban land-use, while Peng et al. (2020) integrated evolutionary and swarm algorithms to mimic urban land-use. Despite having successfully constructed transition rules, these studies identified significant mismatches when compared to real-world urbanization. The source of these errors has been identified as: 1) the introduction of subjectivities via prescribed parameters (Liu et al., 2021), and; 2) oversimplifying the spatial heterogeneity of the driving factors (Gao et al., 2020; Mustafa et al., 2018; Newland et al., 2018). Distances to geographical factors and the spatial configurations of these factors are key to urban development in the real world. The distances used in these studies however, are measured by preset decay functions, and subjectivities are inevitably raised, given the gap between empirical knowledge and real-world urban land-use. The spatial heterogeneity of the urbanization process is often oversimplified to a single shape or texture index that captures only part of the pattern information within the neighborhood (Akin & Erdoğan, 2020; Motlagh, Lotfi, Pourmanafi, Ahmadzadeh, & Soffianian, 2020), let alone the large-scale spatial features (e.g., the layout of the whole built city area) that typify real world urban development.

Deep learning algorithms are a type of machine learning technology that abstract naïve input variables into high-level features. Convolutional neural networks (CNNs) are a special kind of deep learning architecture designed to detect spatial patterns from both low- (e.g., lines, triangles, circles) and high-level features (e.g., human faces, animals) (Krizhevsky, Sutskever, & Hinton, 2017). CNNs are increasingly being used to extract patterns and insights from

geospatial data (Reichstein et al., 2019) which enables the spatial configurations of driving factors, rather than proxy variables such as decay distances, to be integrated directly into urban development simulations. For example, Zhai et al. (2020) used a CNN to retrieve the neighborhood spatial features and improve the Figure of Merit (FoM) of their simulation to 0.361 compared to a random-forest-based method (0.323). Qian et al. (2020) reported similar results, i.e., improving the FoM from 0.299 to 0.346 compared to a random-forest algorithm. Although the spatial features introduced by the CNN improved the simulation performance, these studies used CNNs as little more than an advanced decay function to rescale the driving factors and overlooked their ability to abstract high-level features. More advanced deep learning structures that take full advantage of the spatial pattern recognition ability of CNNs, i.e., integrating low-level spatial features into high-level patterns, have great potential for the spatial simulation of future urban development.

U-Net, first introduced for biomedical image segmentation in 2015, is a unique type of CNN architecture that not only abstracts spatial features to high-level patterns but also refines the high-level patterns to precise shapes (Ronneberger, Fischer, & Brox, 2015). The robust segmentation performance of U-Net enables it to be used in multiple fields such as improving weather prediction (Singh et al., 2021), detecting underwater objects for oceanic ecosystem evaluation (Nezla, Mithun Haridas, & Supriya, 2021), and identifying buildings from aerial and satellite imagery (Ji, Wei, & Lu, 2019). Unlike CA models that have a fixed neighborhood size, U-Net deploys a series of convolutional layers to extract spatial features and then assimilates these features automatically to produce transition rules. This ability to learn spatial patterns suggests that U-Net has the potential to identify and assimilate the spatial processes that drive urban development and accurately capture the resulting patterns of cities.

In this study, we used the U-Net deep learning architecture to project urban land-use in the North China Plain—China’s food bowl and one of the most rapidly urbanizing areas on Earth. We first applied a U-Net model to learn the patterns of urban land-use and land-use change between 1994 and 2006. We then projected the spatial distribution of urbanization for 2018 using this model and thoroughly tested and validated its ability to accurately capture the high-level spatial features, shapes, and patterns of urban development against a reference land-use map for 2018. Last, we trained another U-Net model on land-use maps for 2006 and 2018 to project urban land-use in 2030 based on the extrapolation of historical urbanization rates. We discuss the advantages and limitations of U-Net for simulating urban development and, in particular, its ability to learn urban land-use patterns such as neighborhood influences and linear expansion along transport routes. We also discuss the implications of future rapid urbanization on sustainable development in the study area.

2.2 Methods

2.1 Method overview

We used Landsat data to map urban land-use in the study area for the years 1994, 2006, and 2018 (Wang, Hadjikakou, & Bryan, 2021) and combined them with elevation and slope information to simulate urban development. We created two U-Net models (Table 1) where one was trained/validated by historical data to evaluate the model performance and the other one was trained on more recent data to predict future urban land-use. *U-Net-A* was trained for the purpose of validation by comparing its projected land-use map with the reference map using a range of accuracy and pattern-based metrics. *U-Net-B* was trained to predict future urban land-use for 2030. The training samples for both U-Net models were randomly extracted from the data for the training years. The trained models were then applied to produce a transition potential layer used to create an urban land-use map for the prediction year. The study workflow is summarized and illustrated in Fig. 1 and described in more detail below.

alt-text: Table 1

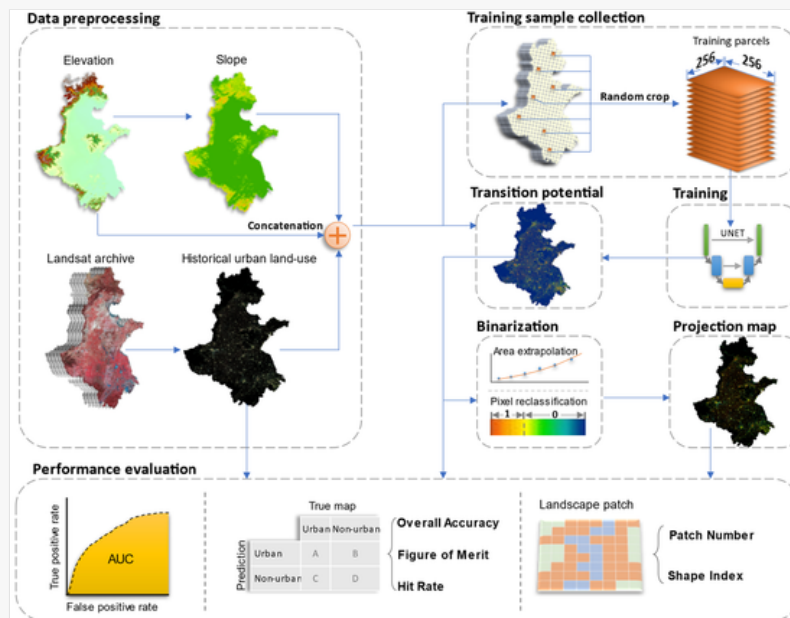
Table 1: Table 1

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

Brief description of the U-Net models trained in this study.

Model name	Training years	Prediction year
U-Net-A	1994, 2006	2018
U-Net-B	2006, 2018	2030


Images are optimised for fast web viewing. Click on the image to view the original version.



Research workflow.

2.2.2 Study area

The North China Plain (Fig. 2) includes 76 prefectures, spans an area of ~~more than~~ $\geq 780,000 \text{ km}^2$, and is home to ~~more than~~ ≥ 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 its built-up land coverage from approximately 5% in 1990 to approximately 15% in 2020 (Wang, Hadjikakou, & Bryan, 2021). 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 its national gross domestic product (GDP) and grain supply (National Bureau of Statistics of China, 2019a). Managing the tension between urbanization and agricultural land-use 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 socio-economic development.


 Images are optimised for fast web viewing. Click on the image to view the original version.



Study area of the North China Plain.

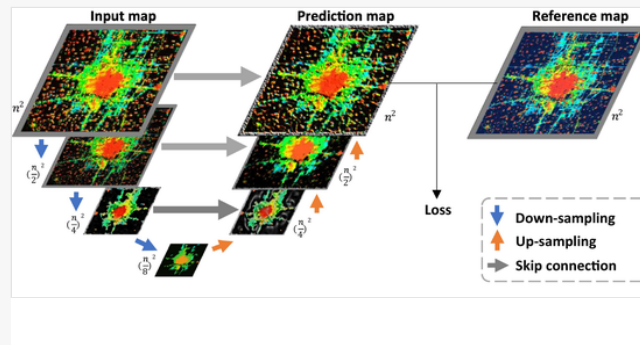
2.3.2.3 Structure of U-Net

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.

 Images are optimised for fast web viewing. Click on the image to view the original version.

alt-text: 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 $(\frac{n}{2})^2$) 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 reference maps, reflecting the performance of the U-Net.

The down-sampling layers include convolution, pooling, and rescaling processes (Table 2). Convolution is a cross-correlation operation that produces feature maps indicating the similarities between inputs and convolution filters (Kapinchev, Bradu, Barnes, & Podoleanu, 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 selected max-pooling to reduce dimensionality because it is adopted in most deep learning structures (Murray & Perronnin, 2014). The rescaling process rescales the pixel values of the feature maps to a specified range to optimize

the computational flow of the network. We used a rectified linear unit activation function (ReLU) for its simple, efficient, and robust performance (Agarap, 2018).

alt-text: Table 2

Table 2: Table 2

i The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

[Instruction: Table 2 appears too ahead in the PDF, and the following paragraph, "In eq (1), K is the the number of classes", is a "Table footnote" for Table 2. Separating Table 2 from its following paragraph will confuse readers.

Can you 1) move Table 2 to Section 2.3 or 2) make the following paragraph a "table footnote" to Table 2 in the PDF?

]Equations of the component layers of U-Net.

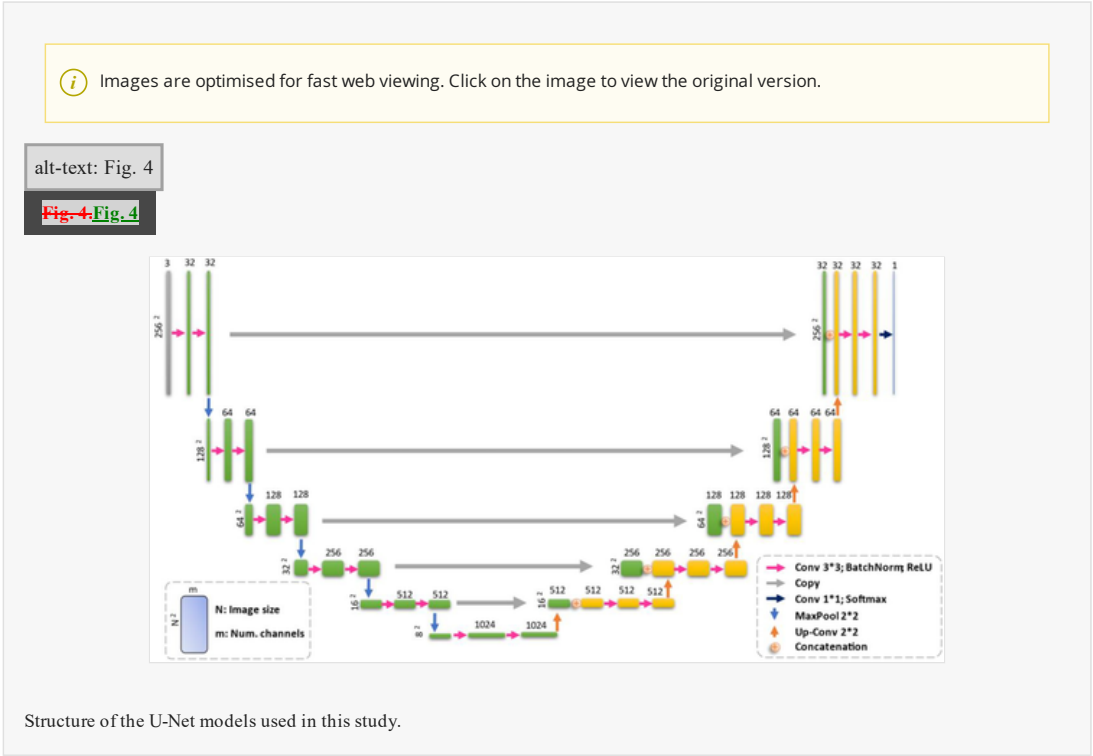
Layer	Equation	No.
Convolution	$C_{out\ j} = \text{bias}(C_{out}) + \sum_{k=0}^{C_{in}-1} \text{weight}(C_{out\ j}, k) \star \text{input}(k)$ $C_{out\ j} = \text{bias}(C_{out}) + \sum_{k=0}^{C_{in}-1} \text{weight}(C_{out\ j}, k) \star \text{input}(k)$ <p>[Instruction: The "$\star \text{input}(k)$" is nudged to another line. I can see there is enough space in the row. Is it possible to present this equation in one row?</p> <p>I tried to nudge the equation, but the summation mark became higher. Can you make the summation mark slimmer?]</p>	(1)
Max-pooling	$\text{out}(C_j, h, w) = \max_{m=0, \dots, kH-1} \max_{n=0, \dots, kW-1} \text{input}(C_j, \text{stride}[0] \times h + m, \text{stride}[1] \times w + n)$ $\text{out}(C_j, h, w) = \max_{m=0, \dots, kH-1} \max_{n=0, \dots, kW-1} \text{input}(C_j, \text{stride}[0] \times h + m, \text{stride}[1] \times w + n)$ <p>[Instruction: There is a blank space before "out ..." which should be removed.]</p>	(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	$\text{loss}(x, y) = -\log\left(\frac{\exp(x[y])}{\sum_{j=1}^K \exp(x[j])}\right)$	(5)

In eq^{uation} (1), the sizes of the input and output images are (C_{in}, H, W) and $(C_{out}, H_{out}, W_{out})$, C denotes the number of channels, H is the height of the input planes in pixels, W is the width in pixels, \star is the valid cross-correlation operator, and j is the j -th channel of the output feature map. In eq^{uation} (2), (kH, kW) denotes the kernel size of the pooling, h and w refer to the height and width of the output image, respectively. In eq^{uation} (3), x denotes the pixel values of the input feature map. In eq^{uation} (4), x_i is the i -th pixel value of the input feature map and K is the number of classes. In eq^{uation} (5), x and y refer to the predicted and reference pixel values, respectively, and K is the number of classes.

The up-sampling layers include transpose convolution and rescaling processes. The transpose convolution, an inverse convolutive operation, transforms the input image from a lower resolution to a higher one. 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 larger than the input size and the rescaling is the same ReLU operation in the down-sampling. Additional components used in the U-Net are batch normalization, softmax, and the cross-entropy algorithm (Table 2). Batch normalization is applied to standardize the weights that control the convolution and transpose convolution processes, which has been proven to be effective in improving deep learning performance (Ioffe & Szegedy, 2015). The softmax algorithm is applied to the last feature map of the U-Net to rescale the pixel values to the 0–1 range for a better comparison with the reference map (also composed of pixels of 0's and 1's). The cross-entropy algorithm is then 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 Fig. 4. The input image size was reduced by half after passing each down-sampling block until reaching a size of 8×8 pixels (i.e., 8^2 in Fig. 4), then the image was ultimately recovered to the original size through each up-sampling block. The channels of the feature maps underwent the inverse process: the number of feature maps was doubled in each down-sampling block and halved in each up-sampling block. As a result, the U-Net model gradually extracted more abstract spatial features over a larger field of view according to $k \times 2^{(d-1)}$, 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 had a field of view of 3, which became 96 in the block of depth 6, meaning that U-Net

was looking for spatial patterns of 96×96 pixels in the original image scale. Meanwhile, U-Net identified more sophisticated spatial patterns as the network block went deeper because the number of feature maps was doubled. The skip-connection links the down-sampling and up-sampling blocks, allowing U-Net to refine the overview patterns extracted by deeper network blocks with precise shapes and textures retrieved by shallower blocks. A total of ~~more~~ ~~than~~ ≥ 31 million parameters were included in the U-Net model, making it highly flexible for capturing the spatial patterns and stochasticity of the urban land-use process.



~~2.4~~~~2.4~~ Data preprocessing

We mapped urban land-use for the years 1994, 2006, and 2018 with a consistently high ($>94\%$) accuracy (for details see [Wang, Hadjikakou, and Bryan \(2021\)](#)). Terrain data was obtained from the Shuttle Radar Topography Mission, from which slope and elevation data were derived ([Table 3](#)). Elevation and slope information was used to assist U-Net in capturing 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](#); [Wang, Feng, et al., 2021](#); [Xing, Qian, Guan, Yang, & Wu, 2020](#)). While accessibility variables such as distance to roads and railways are frequently employed in urban land-use modeling ([Ronneberger et al., 2015](#); [Tripathy & Kumar, 2019](#); [Valencia, Levin, & Hansen, 2020](#)), U-Net is a pattern-sensitive model and in this case, these data layers provide little additional useful information (see Supplementary Material). Hence, to maintain model parsimony and reduce training times distance variables were not used in this study.

alt-text: Table 3

~~Table 3~~~~Table 3~~

The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

The data used to train U-Net models.

Data	Acquiring time	Source	Resolution
Urban land-use	1994, 2006, 2018	Wang, Hadjikakou, and Bryan (2021)	30 m
Elevation	2000*	Shuttle Radar Topography Mission	30 m
Slope	2000*	Derived from elevation	30 m

Table Footnotes

* Note, the elevation and slope data were used in 1994, 2006, and 2018 despite their being acquired in 2000.

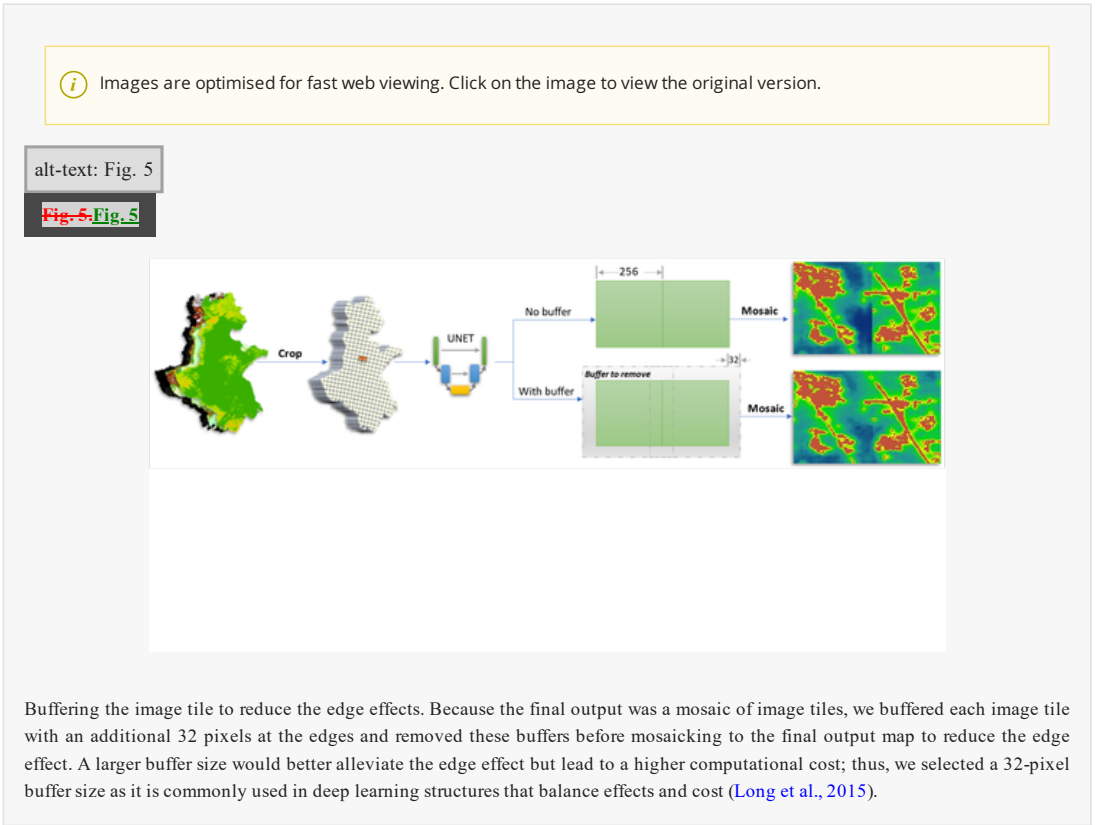
~~2.5~~~~2.5~~ Training the U-Net models

Control samples were assembled using Google Earth Engine (Gorelick et al., 2017). The *NeighborhoodToArray* module was used to crop 256×256 pixel tiled samples following common data science practices (Ronneberger et al., 2015). We collected 20,000 tile samples for training and 5,000 for validation for both *U-Net-A* and *U-Net-B*. Each sampled tiled image included three-layer input data (e.g., *U-Net-A* included urban land-use map for 1994, elevation, and slope) and a single-layer future urban land-use map (e.g., *U-Net-A* included urban land-use map for 2006).

Both U-Net models were trained for 200 epochs (an epoch refers to the U-Net model completely updating its weights using all 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 (Fig. 4). During the validation process, the mean squared error (MSE) was used to compute the difference between the output and target images (i.e., the loss) of the model. Finally, the model with the lowest MSE was determined to be the best model for projecting the target urban layout.

2.6.2.6 Producing the simulation map

The input data (e.g., for *U-Net-A* a three-layer raster of urban land-use for 2006 and the elevation/slope data) was split into tiles, supplied to the trained U-Net model to produce separate outputs, and then the outputs were mosaicked 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 at the desired projection date. To reduce the tile edge effects, image tiles were cropped to a size that was 32-pixel larger than the training image tiles (Fig. 5). Although the U-Net model was trained on image tiles of 256×256 pixels, it could process the buffered 320×320 (original size of 256 plus buffers at edges of size 32×2) image tiles because of the dimensional insensitivity of CNN structures (Long, Shelhamer, & Darrell, 2015).



A classified urban land-use map was created by binarizing the transition potential map using a threshold. First, we ranked the pixel values of the transition potential map from highest to lowest, accumulated the pixel count in this order, and classified them to a value of 1 (indicating urban land-use pixels) until the accumulated count value matched the urban pixel number in the target year. We then allocated the remaining pixels a value of 0 (indicating non-urban land-use).

Q15 The historical urban areas were derived using the data from Wang, Hadjikakou, and Bryan (2021). Predicted urban areas are listed by province in Table 4 and an exponential extrapolation (Fig. 6) was carried out to binarize the transition potential map for 2030. To allocate the total projected urban land-use area to each province we assumed that the share of urban area within each province remained the same as in 2018, and then binarized each province independently to reduce the bias caused by different regional development levels.

alt-text: Table 4

~~Table 4~~ Table 4

--

i The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

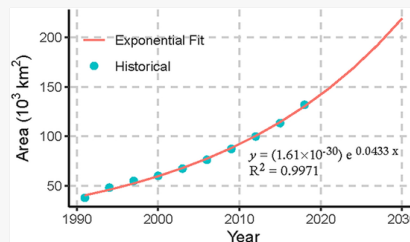
Predicted extents of urban areas in each province in 2030.

Province	2018 (km ²)	2030 (km ²)	Increase (%)
Anhui	21,070.961	29,463.614	39.83
Beijing	2630.09	3338.759	26.94
Hebei	20,454.29	26,440.38	29.27
Henan	28,282.24	38,850.77	37.37
Jiangsu	24,210.08	33,731.722	39.33
Shandong	32,005.14	42,758.619	33.60
Tianjin	2925.686	3754.655	22.08
Total	131,578.48	178,338.49	35.54

i Images are optimised for fast web viewing. Click on the image to view the original version.

alt-text: Fig. 6

Fig. 6, Fig. 6



Exponential regression on historical urban areas in the North China Plain. The historical urban areas are computed using the data from Wang, Hadjikakou, and Bryan (2021).

Q1

2.7.2.7 Validation and accuracy assessment

We selected a series of accuracy metrics and spatial pattern metrics to quantitatively assess the ability of U-Net to accurately project the spatial distributions and patterns 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 its diagnostic ability to discriminate urban and non-urban land-use pixels under various thresholds (Fawcett, 2006). The predicted urban land-use map was evaluated via map overlay and landscape-level spatial pattern metrics (McGarigal & Marks, 1995). The map-overlay metrics selected in this study were overall accuracy (OA), hit rate, and FoM because they not only quantified how the prediction agrees with the reference but also assessed how the reference was erroneously misrepresented by the prediction. The spatial pattern metrics selected in this study were patch number (PN) and landscape shape index (LSI) because they covered the general landscape features and were capable of assessing the typical shapes and spatial patterns of urban development. The selected metrics are described more fully in Table 5. Validation was performed for each prefecture independently, yielding a total of 76 records for each metric.

alt-text: Table 5

Table 5, Table 5

i The table layout displayed in this section is not how it will appear in the final version. The representation below is solely purposed for providing corrections to the table. To preview the actual presentation of the table, please view the Proof.

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 between urban and nonurban land-use pixels (Tong & Feng, 2020).
OA	$(A + D) / (A + B + C + D)$	<i>Overall accuracy</i> is the ratio of correctly identified urban and nonurban land-use pixels to the total number of predictions.
[Instruction: Is it possible to expand the width of the "Name" column so that the "Hit rate" occupied a single row in the PDF?]Hit rate		<i>Hit rate</i> is the ratio of correctly identified urban land-use pixels (i.e., hits) to the number of urban land-use 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 land-use pixels with reference urban land-use 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 square patch is simple (low value), whereas a linear patch is complex (high value).

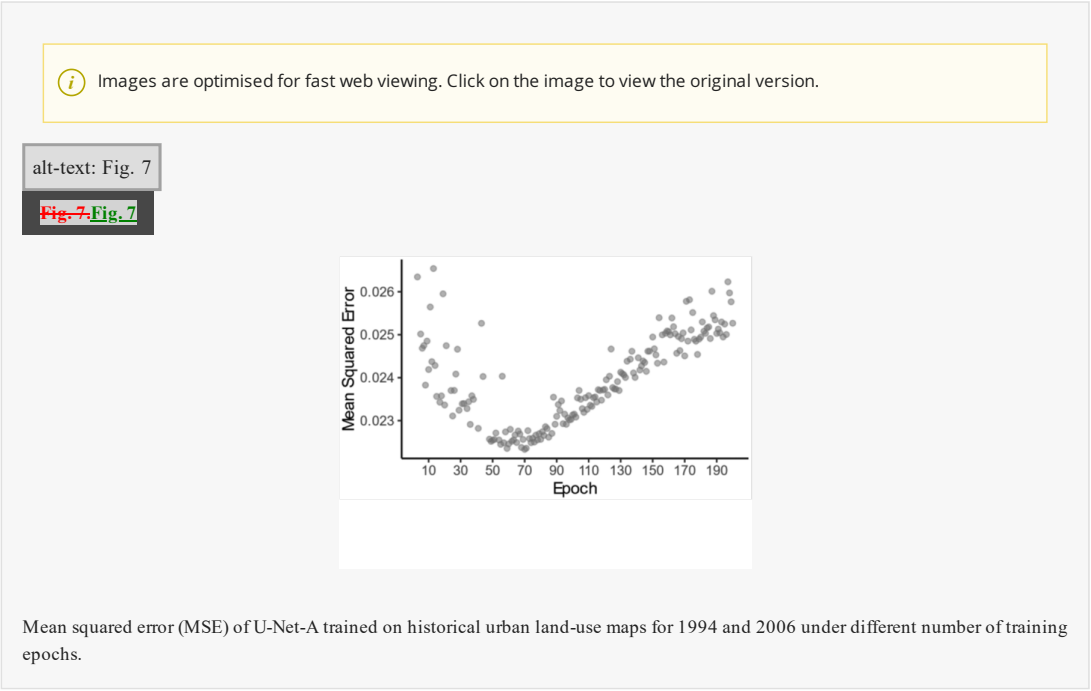
Note: *TPR* is the true positive rate, and *FPR* is the false positive rate of the ROC from the transition potential map. A represents the correctly predicted urban land-use pixels (hit), B represents the incorrectly predicted urban land-use pixels (false alarm), C represents the incorrectly predicted nonurban land-use pixels (miss), and D represents the correctly predicted nonurban land-use pixels (correct rejection). n 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 area (ha) of the i th urban patch.

Finally, we selected 11 cities of varying sizes from across the study area for visual inspection to qualitatively evaluate the ability of U-Net to simulate realistic spatial urban land-use patterns and development characteristics.

3.3 Results


3.3.1 Model training

The MSE of the *U-Net-A* model trained on the urban land-use maps of 1994 and 2006 is shown in Fig. 7. The lowest MSE was 0.022 at the 70th epoch, which was determined as the best training epoch to simulate urban land-use for 2018.



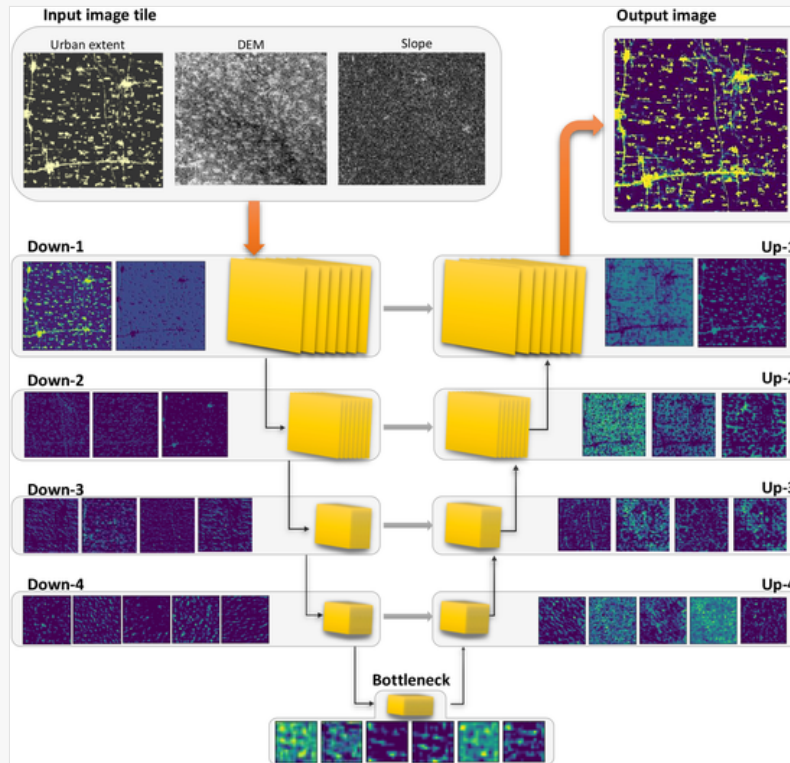
U-Net-A successfully learned and captured different aspects of the complex spatial patterns of urban land-use in the study area (Fig. 8). For example, in the illustrative sample presented in Fig. 8, the Down-1 layer broadly distinguished between urban and non-urban land-use pixels. The Down-2 layer recognized simple patterns such as horizontal and vertical roads and learned to distinguish larger towns from small villages. The Down-3 layer learned to associate adjacent urban clusters and identify potential urban development corridors connecting discrete towns/villages. The Down-4 layer allocated higher urban development probability to the pixels near existing towns/villages, making 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 land-use pixels around larger towns while maintaining the refined patterns.

 Images are optimised for fast web viewing. Click on the image to view the original version.

alt-text: Fig. 8


Fig. 8



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

3.2.3.2 Validation maps


The transition potential map (Fig. 9) illustrates the gravity effect where pixels near larger towns and cities were associated with higher urban transition potential than those near smaller villages. For example, a large area surrounding Suixian was identified as having high transition potential while much smaller areas at the edges of nearby villages exhibited high transition potential. Similar patterns were found near Yangyuan and Duling. Linear development patterns were also captured well in the transition potential map. For example, development along roads in Yangkou Zhen and Huiji Zhen were correctly identified despite only appearing as discrete linear segments in the initial urban land-use map for 2006.

 Images are optimised for fast web viewing. Click on the image to view the original version.



Urban transition potential for 2018 based on the 2006 urban land-use map.

The classified urban land-use map (Fig. 10) 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 planned new developments. Many of the missed urban areas occurred some distance from the original urban areas, such as the newly developed areas of Binhai Zhen, Yangkou Zhen, and Xuancheng City that were separated 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 Duling Xiang, Gaogou Zhen, and Huiji Zhen, despite the absence of explicit accessibility variables in the input dataset.


 Images are optimised for fast web viewing. Click on the image to view the original version.



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

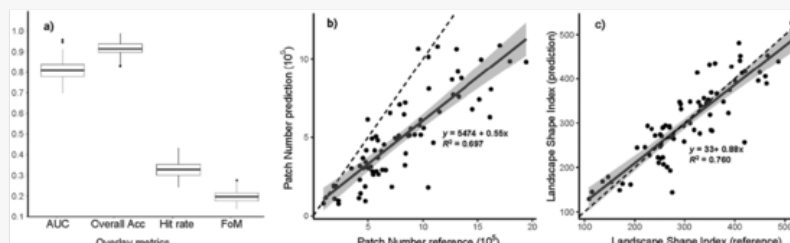
3.3.3.3 Validation metrics

A comparison of accuracy metrics between the output maps and the reference map for 2018 is shown in Fig. 11. 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.

 Images are optimised for fast web viewing. Click on the image to view the original version.

alt-text: Fig. 11

Fig. 11




Assessment metrics for the urban prediction in 2018: a) accuracy metrics; b) patch number; and c) landscape shape metrics. Each dot refers to a record computed from a prefecture in the study area. The dashed line refers to $y = x$ and the ribbon is the confidence interval (95%) of the mean for the fitted linear regression.

The landscape-scale spatial pattern metrics of the classified urban land-use map compared to the reference map are also shown in Fig. 11. The patch number 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 actual urban development depicted in the reference map but produced urban shapes very close to the real-world urban patterns.

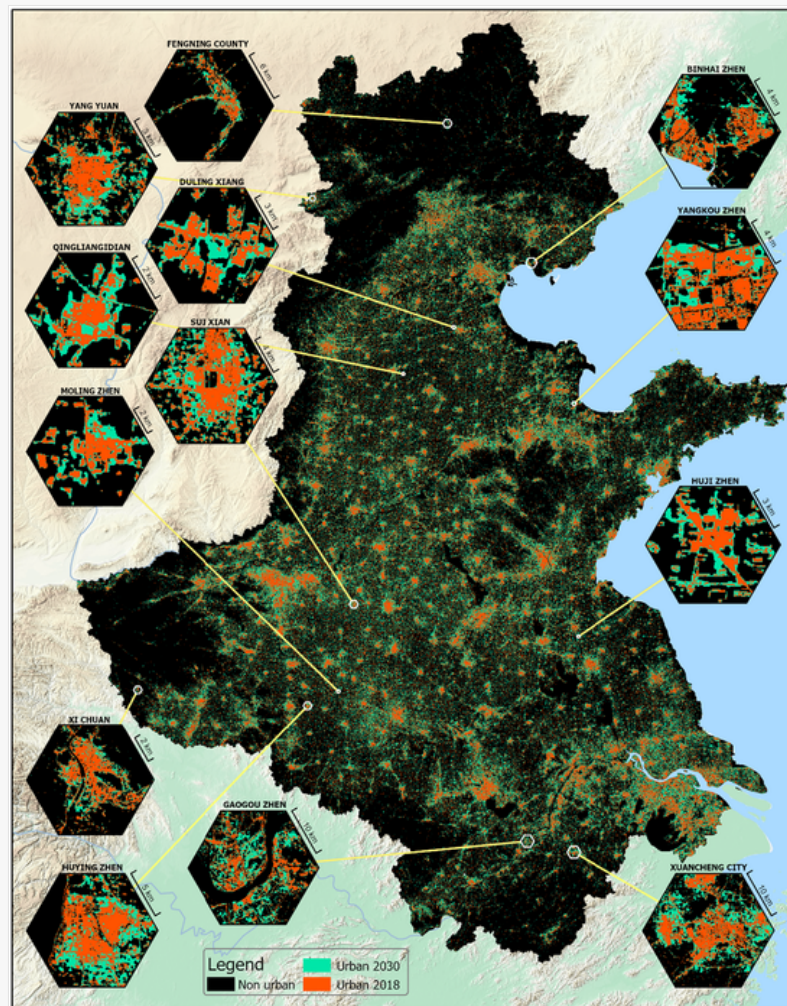
3.4.3.4 Predicted urban area for 2030

U-Net-B was developed with the same process as *U-Net-A* except for taking urban land-use in 2006 and 2018 for training and predicting urban land-use for 2030 (Fig. 12). Newly predicted urban areas tended to be concentrated around large cities. In Suixian, for example, the predicted urban expansion was much larger around large cities/towns than that around smaller villages. The predicted urban area followed specified patterns rather than sprawling in all directions. For example, the predicted urban land areas in Duling Xiang and Yangkou Zhen filled in gaps in the spatial distribution of existing urban areas and maintained the general shape of the original urban layout. Urban 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.

 Images are optimised for fast web viewing. Click on the image to view the original version.

alt-text: Fig. 12

Fig. 12. Fig. 12



Predicted urban land areas in 2030.

4.4 Discussion

4.4.1 Mimicking real-world urban patterns and dynamics

U-Net was able to model urban land-use change at high accuracy as well as capture and assimilate high-level spatial patterns of urban development in the North China Plain. First, the model captured neighborhood effects. The transition potential map revealed that lands near existing urban areas were more likely to transition to urban land-use in the future and that urbanization ~~amongst~~among larger expanses of non-urban land would be unlikely (Fig. 9). Second, U-Net was able to assimilate small-scale neighborhood effects into large-scale gravity effects. Larger areas of land near larger cities were predicted to become urbanized in the future, while much smaller areas were predicted to become urbanized around smaller, more remote villages (Fig. 10). Third, U-Net, rather than just simply buffering spatial features, captured the tendency of urban expansion to follow linear patterns. For example, the elongated development trend of Fengning County controlled by valley terrain was identified in the prediction (Fig. 12), and areas along major transport routes

were allocated high transition potentials (Fig. 9). However, the U-Net model was unable to predict new planned developments that sprouted up some distance from existing urban areas, such as Binhai Zhen, Yangkou Zhen, Huying Zhen, and Xuancheng City (Fig. 10).

4.2.4.2 Reducing subjectivity by automatically constructing transition rules

The large-scale spatial patterns identified by U-Net, in particular the gravity effect of large cities, would not have been captured by existing CA models. A large neighborhood is required for a CA model to incorporate information at a large scale which tends to decrease simulation performance. Wang, Feng, et al. (2021) found that the best neighborhood size to simulate urban development in Beijing was 25×25 pixels, while Roodposhti et al. (2020) observed that a 9×9 pixel neighborhood outperformed other settings. However, the unique design of the U-Net model is capable of incorporating more extensive information than previous CA model structures, enabling the neighborhood influence to be learned from the data and urban development to be simulated with more refined spatial configurations.

U-Net did not require extensive amounts of forcing data or parameterization to simulate urban land-use changes, except for the learning rate and the size of the training images. The transition rules were automatically generated in the form of learned weights, and training was an automatic process that required few subjective decisions to be made. In contrast, CA model structures have often required 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 urban growth in Shanghai, whereas an array of parameters including inertia weights, decay magnitude, spatial heterogeneity, and variable scaling, needed to be specified before running the model. Many studies determine these parameters according to expert knowledge (Chen et al., 2020; Mustafa, Cools, Saadi, & Teller, 2017; Tripathy & Kumar, 2019), potentially leading to subjective biases. Other studies went through a systematic parameter selection process to find the best parameters (Roodposhti et al., 2020; Yu et al., 2021), but this process is time-consuming and impractical given the many unknown parameters.

4.3.4.3 Accurate prediction and robustness in capturing spatial patterns

There are challenges to comparing model performance across studies. Pontius et al. (2008) reported two factors that profoundly influence urban land-use simulation accuracy: 1) the area of urban expansion, and; 2) the spatial resolution. A positive relationship exists between the FoM and observed land-use change. Prediction errors vanished when the simulation maps were resampled into coarser resolution (Pontius et al., 2008). Given these sensitivities, to enable a fair comparison of the accuracy of our U-Net model outputs with CA model outputs, we selected two CA-based studies with similar historical urbanization areas and the same 30 m spatial resolution as our study. Wang, Guo, Zhang, and Zeng (2021) used a particle swarm optimization algorithm, iterated a range of parameter settings such as a self-recognition component, a social component, inertia weights, and the number of particles, to finally arrive at the best model with an FoM of 0.193 for Zhuji, China. Wang, Feng, et al. (2021) developed 17 sub-models incorporating four periods of historical urban land-use and tested eight different neighborhood sizes (5×5 to 41×41) for each sub-model to ultimately identify the best simulation which achieved an FoM of 0.219 for Beijing, China. The FoM (computed from 76 prefectures) in our study ranges from 0.177 to 0.215 (interquartile range) which covers the 0.193 reported by Wang, Guo, et al. (2021) but is slightly lower than the 0.219 found by Wang, Feng, et al. (2021). This higher FoM reported by Wang, Feng, et al. (2021) may be explained by the very large amount of urbanization in their study area (i.e., Beijing) (Pontius et al., 2008). This comparison demonstrates that U-Net achieved similar predictive accuracies to comparable CA-based urban land-use models.

4.4.4.4 Predicting urban land-use change for formulating policy

An additional area of 46,760 km² was predicted to become urbanized in the North China Plain between 2018 and 2030. The three provinces with the highest rates of urban area increase (Anhui, Jiangsu, and Henan) were all in the southern part of the region (Fig. 2). Our projections for 2030 also captured urban development in the three megacity groups (Beijing-Tianjin-Hebei, the Yangtze River Delta, and the Central Plains) that account for one-third of China's GDP (National Bureau of Statistics of China, 2019b). China's strategic development planning process can benefit from the predictions arising out of this study in many ways. For example, urban area predictions can be used to plan for infrastructure to support socio-economic development. Simulations can assist policy formulation that is tailored to the expected rate, location, and patterns of urbanization.

In addition, the competition between urban development and food production in the study area is a critical issue for China's food security (Jin et al., 2019). Pressure on arable land could be alleviated by making use of our highly granular predictions of future urban development because they enable the development of plans/policies to reduce food production losses that would otherwise arise from urban expansion (Jin et al., 2019). Ecosystem protection, biodiversity, and environmental conservation can also greatly benefit from an accurate prediction of future urbanization. For example, urban land is an important proxy for domestic water consumption (Hoekstra, Buurman, & van Ginkel, 2018) and pollution (Zeller, Towa, Degrez, & Achten, 2019), which can be quantified based on urban land-use maps.

Finally, future urban layouts provide a baseline for local governments to address land-use conflicts and settle competing interests among different stakeholders in the urbanization process.

~~4.5~~4.5 Limitations and prospects

Spatial patterns of urban development are highly complex. Very few tools currently exist to quantitatively compare these patterns. The shape index employed in this study revealed a high level of agreement between the U-Net projection of urban areas and the reference land-use map. However, our assessment of the ability of the U-Net to mimic specific patterns was largely based on visual inspection. No quantitative metrics currently exist to objectively assess whether projected urban patterns and shapes look plausible. Expert human interpretation such as the visual inspection method used in this study can effectively assess the realism of projection urban patterns but it is a qualitative and imprecise process. Better tools and metrics are required to quantitatively assess these complex spatial shapes and patterns to complement qualitative assessment based on human interpretation.

U-Net automatically learned the spatial patterns from historical urban land-use change, while CA methods model urban development based on the influence of the driving factors and known urbanization processes. As a result, U-Net has a limited capacity to support participatory modeling that requires a series of ‘what-if’ trials undertaken with stakeholders involving the tweaking of the model parameters and driving factors and inspection of the results. Conversely, the U-Net model focuses on predictive accuracy rather than supporting participatory modeling. The transition rules of U-Net are hidden in its large number of learned weights and are not human-readable, whereas CA models are usually explainable and intuitive, i.e., the key driving factors can be identified. Hence, the utility of U-Net as a participatory planning tool, where co-learning of the driving factors of urbanization is a key objective, is limited at present. However, emerging developments in extracting human-recognizable knowledge from deep learning architectures (Dosilovic, Breie, & Hupic, 2022; Dosilovic, Brcic, & Hupic, 2018; Gunning et al., 2019) could be applied to U-Net models for a better understanding of the mechanisms behind urban land-use dynamics.

Several new, advanced deep learning structures built on computer vision technology are emerging which could be used to retrieve a wide spectrum of features (e.g., distance to roads, and/or the temporal characteristics in historical urban development) and improve simulation performance. There is significant potential to explore the ability of other deep learning architectures to accurately model the spatial distribution and pattern of urban development.

~~5~~5 Conclusion

For our case study region of the North China Plain, the U-Net deep learning architecture was able to accurately simulate urban development as well as reveal key urbanization processes. The U-Net model successfully captured neighborhood effects whereby new urban areas are more likely to arise near existing urban areas. U-Net also captured gravity effects where new urban development is more likely to occur near large cities rather than small villages, and it also captured the tendency for linear expansion of urban land-use along transportation routes but it was not able to predict new planned urban development distant from existing urban areas. Nonetheless, the spatial patterns of simulated urban land-use maps matched the reference map closely, indicating that U-Net was capable of simulating urban land-use change at a highly granular level. In addition, the model had a very low requirement for parameterization thereby limiting the subjective decisions and biases in training the model. It also required very little foregoing limited input data thus making it particularly suitable for data-scarce regions. The more accurate identification of urban patterns and dynamics is a useful addition to urban land-use simulation studies to incorporate the various complex spatial driving factors. The resulting projected 2030 urban land-use map provides key information for planning China's strategic socioeconomic development, and can benefit policy formulation and decision-making concerning food security, biodiversity conservation, and environmental protection.

CRediT authorship contribution statement

Jinzhu Wang: Conceptualization, Data curation, Methodology, Visualization, Writing – original draft. **Michalis Hadjikakou:** Investigation, Validation, Writing – review & editing. **Richard J. Hewitt:** Investigation, Writing – review & editing. **Brett A. Bryan:** Supervision, Writing – original draft, Writing – review & editing, Methodology, Visualization.

Declaration of Competing Interest

None.


Acknowledgment

Q23 A Deakin University Postgraduate Research scholarship funded this research.

~~Appendix A~~Appendix A Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.compenvurbsys.2022.101855>.

References

 The corrections made in this section will be reviewed and approved by a journal production editor. The newly added/removed references and its citations will be reordered and rearranged by the production team.

Agarap, A.F. (2018). ~~Deep Learning using Rectified Linear Units (ReLU)~~Deep learning using rectified linear units (ReLU).

Akın, A., & Erdoğan, M.A. (2020). Analysing temporal and spatial urban sprawl change of Bursa city using landscape metrics and remote sensing. *Modeling Earth Systems and Environment*, 6, 1331–1343.

Cameiro, M.G., & Oliveira, G.M.B. (2013). Synchronous cellular automata-based scheduler initialized by heuristic and modeled by a pseudo-linear neighborhood. *Natural Computing*, 12, 339–351.

Chen, S., Feng, Y., Tong, X., Liu, S., Xie, H., Gao, C., & Lei, Z. (2020). Modeling ESV losses caused by urban expansion using cellular automata and geographically weighted regression. ~~The Science of the total environment~~The Science of the Total Environment, 712, 136509.

Clarke, K.C., & Johnson, J.M. (2020). Calibrating SLEUTH with big data: Projecting California's land use to 2100. *Computers, Environment and Urban Systems*, 83, 101525.

Q26

Dosilovic, F.K., Brcic, M., & Hlupic, N. (20222018). Explainable artificial intelligence: A survey. (pp. 210–215). 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)[Instruction: The font size of this citation looks extra larger in the PDF.]

Fan, H., Zhao, C., & Yang, Y. (2020). A comprehensive analysis of the spatio-temporal variation of urban air pollution in China during 2014–2018. *Atmospheric Environment*, 220, 117066.

Fan, P., Chen, J., Ouyang, Z., Groisman, P., Loboda, T., Gutman, G., ... Qi, J. (2018). ~~Urbanization and sustainability under transitional economies: a synthesis for Asian Russia~~Urbanization and sustainability under transitional economies: A synthesis for Asian Russia. *Environmental Research Letters*, 13, 95007.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874.

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, 74–97.

Gantumur, B., Wu, F., Vandansambuu, B., Tsegmid, B., Dalaibaatar, E., & Zhao, Y. (2020). Spatiotemporal dynamics of urban expansion and its simulation using CA-ANN model in Ulaanbaatar, Mongolia. *Geocarto International*, 1–16.

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.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). ~~Google Earth Engine: Planetary-scale geospatial analysis for everyone~~Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202. <https://www.sciencedirect.com/science/article/pii/S0034425717302900>.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI-Explainable artificial intelligence. *Science robotics*, 4.

Hoekstra, A.Y., Buurman, J., & van Ginkel, K.C.H. (2018). Urban water security: A review. *Environmental Research Letters*, 13, 53002.

Ioffe, S., & Szegedy, C. (2015). ~~Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift~~[Batch normalization: Accelerating deep network training by reducing internal covariate shift](http://proceedings.mlr.press/v37/ioffe15.html). International conference on machine learning. <http://proceedings.mlr.press/v37/ioffe15.html>.

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, 3308–3322.

Jin, G., Chen, K., Wang, P., Guo, B., Dong, Y., & Yang, J. (2019). ~~Trade-offs in land-use competition and sustainable land development in the North China Plain~~[Trade-offs in land-use competition and sustainable land development in the North China plain](#). *Technological Forecasting and Social Change*, 141, 36–46.

Kafy, ~~A.-A.-A.~~, Naim, M.N.H., Subramanyam, G., Faisal, ~~A.-A.-A.~~, Ahmed, N.U., Rakib, A.A., ... Sattar, G.S. (2021). ~~Cellular Automata approach in dynamic modelling of land cover changes using RapidEye images in Dhaka, Bangladesh~~[Cellular automata approach in dynamic modelling of land cover changes using RapidEye images in Dhaka, Bangladesh](#). *Environmental Challenges*, 4, 100084.

Kapinchev, K., Bradu, A., Barnes, F., & Podoleanu, A. (2015). GPU implementation of cross-correlation for image generation in real time. (pp. 1–6).

Kipfer, S. (2018). Pushing the limits of urban research: Urbanization, pipelines and counter-colonial politics. *Environment and Planning D: Society and Space*, 36, 474–493.

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60, 84–90.

Li, X., Gong, P., Le, ~~YuY.~~, & Hu, T. (2017). A segment derived patch-based logistic cellular automata for urban growth modeling with heuristic rules. *Computers, Environment and Urban Systems*, 65, 140–149.

Liu, Y., Batty, M., Wang, S., & Corcoran, J. (2021). Modelling urban change with cellular automata: Contemporary issues and future research directions. ~~Progress in human geography~~[Progress in Human Geography](#), 45, 3–24.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. (pp. 3431–3440).

Mansour, S., Al-Belushi, M., & Al-Awadhi, T. (2020). Monitoring land use and land cover changes in the mountainous cities of Oman using GIS and CA-Markov modelling techniques. ~~Land-use policy~~[Land Use Policy](#), 91, 104414.

McGarigal, K., & Marks, B.J. (1995). ~~FRAGSTATS: spatial pattern analysis program for quantifying landscape structure~~[FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure](#). ~~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.

Motlagh, Z.K., Lotfi, A., Pourmanafi, S., Ahmadizadeh, S., & Soffianian, A. (2020). ~~Spatial modeling of land-use change in a rapidly urbanizing landscape in central Iran: integration of remote sensing, CA-Markov, and landscape metrics~~[Spatial modeling of land-use change in a rapidly urbanizing landscape in Central Iran: Integration of remote sensing, CA-Markov, and landscape metrics](#). ~~Environmental monitoring and assessment~~[Environmental Monitoring and Assessment](#), 192, 695.

Murray, N., & Perronnin, F. (2014). ~~Generalized Max Pooling~~[Generalized max pooling](#).

Mustafa, A., Cools, M., Saadi, I., & Teller, J. (2017). Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM). ~~Land-use policy~~[Land Use Policy](#), 69, 529–540.

Mustafa, A., Heppenstall, A., Omrani, H., Saadi, I., Cools, M., & Teller, J. (2018). Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Computers, Environment and Urban Systems*, 67, 147–156.

National Bureau of Statistics of China . Announcement of the 2019 grain output . http://www.gov.cn/xinwen/2019-12/07/content_5459250.htm, 2019 .

National Bureau of Statistics of China (2019b). ~~China Statistical Yearbook~~[China statistical yearbook](#). Beijing, China: China Statistics Press.

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~~[Empirically derived method and software for semi-automatic calibration of cellular automata land-use models](#). Environmental Modelling & Software, 108, 208–239.

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~~[Semantic segmentation of underwater images using UNet architecture based deep convolutional encoder decoder model](#). (pp. 28–33).

Peng, K., Jiang, W., Deng, Y., Liu, Y., Wu, Z., & Chen, Z. (2020). ~~Simulating wetland changes under different scenarios based on integrating the random forest and CLUE-S models: A case study of Wuhan Urban Agglomeration~~[Simulating wetland changes under different scenarios based on integrating the random forest and CLUE-S models: A case study of Wuhan urban agglomeration](#). Ecological Indicators, 117, 106671.

Planillo, A., Kramer-Schadt, S., Buchholz, S., Gras, P., von der Lippe, M., & Radchuk, V. (2021). Arthropod abundance modulates bird community responses to urbanization. Diversity and Distributions, 27, 34–49.

Pontius, R.G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., ... Verburg, P.H. (2008). Comparing the input, output, and validation maps for several models of land change. The Annals of Regional Science, 42, 11–37.

Qian, Y., Xing, W., Guan, X., Yang, T., & Wu, H. (2020). Coupling cellular automata with area partitioning and spatiotemporal convolution for dynamic land use change simulation. ~~The Science of the total environment~~[The Science of the Total Environment](#), 722, 137738.

Qiu, B., Li, H., Tang, Z., Chen, C., & Berry, J. (2020). How cropland losses shaped by unbalanced urbanization process? ~~Land use policy~~[Land Use Policy](#), 96, 104715.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019). ~~Deep learning and process understanding for data-driven Earth system science~~[Deep learning and process understanding for data-driven earth system science](#). Nature, 566, 195–204.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net. Convolutional Networks for Biomedical Image Segmentation, 9351, 234–241.

Roodposhti, M.S., Aryal, J., & Bryan, B.A. (2019). A novel algorithm for calculating transition potential in cellular automata models of land-use/cover change. Environmental Modelling & Software, 112, 70–81.

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.

Ruiz Hernandez, I.E., & Shi, W. (2018). ~~A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis~~[A random forests classification method for urban land-use mapping integrating spatial metrics and texture analysis](#). International Journal of Remote Sensing, 39, 1175–1198.

Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017). Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. Computers, Environment and Urban Systems, 64, 297–308.

Shaw, B.J., van Vliet, J., & Verburg, P.H. (2020). The peri-urbanization of Europe: A systematic review of a multifaceted process. Landscape and Urban Planning, 196, 103733.

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.

Tong, X., & Feng, Y. (2020). A review of assessment methods for cellular automata models of land-use change and urban growth. *International Journal of Geographical Information Science*, 34, 866–898.

Tripathy, P., & Kumar, A. (2019). ~~Monitoring and modelling spatio-temporal urban growth of Delhi using Cellular Automata and geoinformatics~~ Monitoring and modelling spatio-temporal urban growth of Delhi using cellular automata and geoinformatics. *Cities*, 90, 52–63.

Valencia, V.H., Levin, G., & Hansen, H.S. (2020). ~~Modelling the spatial extent of urban growth using a cellular automata-based model: a case study for Quito, Ecuador~~ Modelling the spatial extent of urban growth using a cellular automata-based model: A case study for Quito, Ecuador. *Geografisk Tidsskrift-Danish Journal of Geography*, 120, 156–173.

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.

Wang, J., Hadjikakou, M., & Bryan, B.A. (2021). ~~Consistent, accurate, high resolution, long time-series mapping of built-up land in the North China Plain~~ Consistent, accurate, high resolution, long time-series mapping of built-up land in the North China plain. *GIScience & Remote Sensing*, 58, 982–998.

Wang, R., Feng, Y., Wei, Y., Tong, X., Zhai, S., Zhou, Y., & Wu, P. (2021). A comparison of proximity and accessibility drivers in simulating dynamic urban growth. *Transactions in GIS*, 25, 923–947.

Xia, C., & Zhang, B. (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~~ 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. *GIScience & Remote Sensing*, 1–24.

Xing, W., Qian, Y., Guan, X., Yang, T., & Wu, H. (2020). A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. *Computers & Geosciences*, 137, 104430.

Yeh, A.G.-O., & Chen, Z. (2020). From cities to super mega city regions in China in a new wave of urbanisation and economic transition: Issues and challenges. *Urban Studies*, 57, 636–654.

Yu, J., Hagen-Zanker, A., Santitissadeekorn, N., & Hughes, S. (2021). Calibration of cellular automata urban growth models from urban genesis onwards - a novel application of Markov chain Monte Carlo approximate Bayesian computation. *Computers, Environment and Urban Systems*, 90, 101689.

M.D. Zeiler , R. Fergus . ~~Visualizing and Understanding Convolutional Networks~~ Visualizing and understanding convolutional networks . <https://arxiv.org/pdf/1311.2901> , 2013 .

Zeller, V., Towa, E., Degrez, M., & Achten, W.M.J. (2019). Urban waste flows and their potential for a circular economy model at city-region level. *Waste management (New York, N.Y.)*, 83, 83–94.

Zhai, Y., Yao, Y., Guan, Q., Liang, X., Li, X., Pan, Y., ... Zhou, J. (2020). Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata. *International Journal of Geographical Information Science*, 34, 1475–1499.

Zheng, W., Shen, G.Q., Wang, H., Hong, J., & Li, Z. (2017). Decision support for sustainable urban renewal: A multi-scale model. ~~Land use policy~~ Land Use Policy, 69, 361–371.

- U-Net deep learning architecture was able to simulate complex urban development patterns.
 - Urban transition rules were learned automatically in the U-Net with minimal data requirements.
 - The complex spatial patterns of simulated urban land-use closely matched actual patterns.
 - U-Net achieved similar accuracies to CA-based models with a comparable urbanization context.
 - U-Net learned the neighborhood, gravity, and linear effects of urban growth.
-

~~Appendix A~~ Appendix A Supplementary data



[Multimedia Component 1](#)

Supplementary material

alt-text: Image 1

Queries and Answers

Q1

Query: The **citation** ‘Wang et al. (2021b)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q2

Query: Please check the **layout of all Tables** and correct if necessary.

Answer: Reviewed

Q3

Query: The **citation** ‘Wang et al. (2021b)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q4

Query: Your article is being processed as a regular item to be included in a **regular issue**. Please confirm if this is correct or if your article should be published in a special issue using the responses below.

Answer: Yes

Q5

Query: Please review the **given names and surnames** to make sure that we have identified them correctly and that they are presented in the desired order. Carefully verify the spelling of all authors’ names as well. If changes are needed, please provide the edits in the author section.

Answer: Yes

Q6

Query: Please check the **address for the corresponding author** that has been added here, and correct if necessary.

Answer: Reviewed

Q7

Query: Please check whether the **hierarchy of the section headings** are as expected. If changes are needed, please use the comment option to indicate the required changes

Answer: Reviewed

Q8

Query: The **citation** ‘Planillo et al. , 2021’ has been changed to match the author name in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q9

Query: The **citation** ‘Wang et al., 2021a’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q10

Query: The **citation** ‘Wang et al. (2021a)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q11

Query: The **citation** ‘Wang et al., 2021b’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q12

Query: The **citation** ‘Wang et al., 2021b’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q13

Query: The **citation** ‘Wang et al. (2021b)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q14

Query: The **citation** ‘Wang et al., 2021a’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q15

Query: The citation ‘Wang et al. (2021b)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: The change is correct.

Q16

Query: The **citation** ‘Wang et al. (2021a)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q17

Query: The **citation** ‘Wang et al. (2021c)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q18

Query: The **citation** ‘Wang et al. (2021a)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q19

Query: The **citation** ‘Wang et al. (2021c)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q20

Query: The **citation** ‘Wang et al. (2021a)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q21

Query: The **citation** ‘Wang et al. (2021a)’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q22

Query: The **citation** ‘Dosilovic et al.’ has been changed to match the date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Answer: Reviewed

Q23

Query: Have we correctly interpreted the following funding source(s) and country names you cited in your article: "Deakin University Postgraduate Research".

Answer: Yes

Q24

Query: Please provide a **caption** for the Supplementary data. You can use the comment option to indicate the caption text.

Answer: Done

Q25

Query: Please provide a **caption** for the Supplementary data. You can use the comment option to indicate the caption text.

Answer: Done

Q26

Query: Please check the publication year for this following references.

Answer: Done