Computers, Environment and Urban Systems 74 (2019) 244–256

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Review

Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis

George Grekousis

Sun Yat-Sen University, School of Geography and Urban Planning, Department of Urban and Regional Planning, No 135, Xingang Xi Road, Guangzhou, Haizhu 510275, PR China

ARTICLE INFO

Keywords:
Artificial Neural Networks
Deep Learning
Urban Geography
Meta-analysis
Trends
Guidelines on Reporting results

ABSTRACT

Artificial neural networks (ANNs) and their latest advancement in deep learning are blooming in computer science. Geography has integrated these artificial intelligence techniques, but not with the same enthusiasm. The main reason for hesitation is that ANNs are still confronted as complex and black boxes. However, ANNs might be more solid methods than conventional approaches when dealing with complex geographical problems. This study considers the great potential of ANNs for research in urban geography. First, using the PRISMA protocol, it provides a statistical review of 140 papers on studies that employed ANNs in urban geography between 1997 and 2016. Second, it performs a quantitative meta-analysis using non-parametric bootstrapping. 45 (of the 140) papers were assessed regarding ANNs' overall accuracy (OA) achieved when used for urban growth prediction or urban land-use classification. Third, a new guideline for reporting ANNs is proposed. Statistical review indicated that ANNs performed better in 75.7% of case studies compared to conventional methods. Meta-analysis found that on bootstrapped averages, the median OA achieved when using, ANNs was higher than the median OA achieved by other techniques by 2.3% ($p < .001$). ANNs also performed better when used for classification compared to prediction. Analysis also identified inadequate presentation of ANNs and related results when used in urban studies. For this reason, a new guideline for reporting ANNs is suggested in this work to ensure consistency and easier dissemination of individual lessons learned. These findings aim to motivate further studies on ANNs and deep learning in urban geography.

1. Introduction

More than twenty years have passed since Openshaw's influential book, entitled Artificial Intelligence in Geography, was published (Openshaw & Openshaw, 1997). Openshaw was one of the first geographers to strongly support the use of Artificial Intelligence (AI) and Artificial Neural Networks (ANNs) in geography, and his legacy is paramount in the field. ANNs are now used to respond to many geographical problems, although their use lag behind some other disciplines mainly because ANNs are (1) considered a “black box” technology, (2) believed to be extremely complicated, and (3) rarely taught to geography students.

On the contrary, AI is flourishing in other fields. For example, ANNs and, specifically, deep neural networks have recently proven their abilities to handle numerous complex problems, such as speech recognition (Hinton et al., 2012), visual object recognition (Krizhevsky, Sutskever, & Hinton, 2012), image analysis (Lecun, Bengio, & Hinton, 2015), language translation (Bengio, 2012), and self-driving vehicles (Goodfellow, Bengio, & Courville, 2016). Deep learning is a blooming field and many financial reports have highlighted the influences of AI and deep learning in global and national economies. The global deep learning market is expected to grow at a compound annual growth rate of 65.3% between 2016 and 2022 to achieve a worth of USD 1772.9 million by 2022 (Markets and Markets, 2016). Some companies, such as Google (US), IBM (US), Microsoft (US), and Baidu (China), are investing huge amounts of money in the research and development of deep learning ANNs to solve complex problems. For example, major North American companies spent USD 80 million on average during 2015 on AI research and technologies (TATA Consultancy Service, 2017). AI is projected to dramatically influence economic growth by boosting productivity as much as 40% by 2035 (Purdy & Daughtery, 2016). Openshaw and Openshaw (1997) had successfully predicted the AI and ANNs boom since 1997. Referring to ANNs, he stated, “A second computer revolution is under way with wide application in many areas of Geography.” In the context of the urban geography, the applications of ANNs seem almost endless, particularly now that urban geographers face two major challenges namely big data and data complexity.

Regarding big data, urban geographers should extract valuable

E-mail addresses: graikousis@mail.sysu.edu.cn, geograik@gmail.com.

https://doi.org/10.1016/j.compenvurbsys.2018.10.008
Received 3 July 2018; Received in revised form 17 September 2018; Accepted 29 October 2018
Available online 14 November 2018
0198-9715/

information from and uncover hidden patterns in currently available vast databases. As locational intelligence becomes increasingly relevant to most aspects of geographical analysis, exploitation and analysis of big geospatial data will play an increasingly important role (GMC, 2017). Combined with non-spatial big data derived from censuses, social media databases, environmental databases, health databases, and private companies' data repositories, these efforts will create a vast pool of data for deeper urban geographical analysis. This type of analysis can be accomplished mainly through new methods based on AI and machine learning, which, in many cases, surpasses conventional methods in efficiency through their superior capacity for handling big data (GFM, 2017).

A weakness of having an abundance of data is the increasing data complexity. The complexity of the data gets even higher when dealing with interdisciplinary problems (in urban geography context) such as accelerating urbanization, land resources scarcity, global warming, climate change, poverty, inequality. To briefly present an overview of the 21st century megatrends magnitude and their associated data complexity that urban geographers and planners are facing, related numbers and facts are listed below. An estimated 54.5% of the world's population lived in cities in 2016 (UN, 2016). This percentage is expected to rise to 60% by 2030 with a trend for more and more people concentrating in large cities (UN, 2016). Rapid urbanization is also associated to environmental degradation, extensive land cover change, and global warming (Bloom, Canning, & Fink, 2008; Grekousis, Mountrakis, & Kavouras, 2016). In 2014, more than one-half (56%) of the world's cities with populations over 300,000 were at high risk of experiencing at least one of the following disasters: cyclone, flood, drought, earthquake, landslide, and volcanic eruption (UN, 2016). Therefore, natural disasters increase the risk of high mortality and economic loss in cities. These problems threaten the sustainability of natural and urban environments, are highly complex, and call for advanced methods to be used (Grekousis & Mountrakis, 2015).

Altogether, the above mega-trends and challenges (big data, data complexity) sketch a bewildering backdrop to urban evolution. A more demanding analysis in the context of contemporary urban geography and interdisciplinary studies is necessary. As such, urban geographers have a quest for contemporary methods and concepts that will help them to better analyze complex problems. The analysis of 21st-century emerging urban mega-trends and the exploitation of very complex data using automation through ANNs can benefit the discipline of urban geography.

Therefore, this study emphasizes the need for adopting in urban geography more widely used methods and techniques to successfully handle big data and data complexity. The scope and purposes of this work are as follows:

(1) to perform a statistical review of literature published between 1997 and 2016 using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol. 140 papers are extracted from the ISI Web of Science database that make use of ANNs in urban studies grouped in three thematic groups namely urbanization, environment (urban), and socioeconomic. This statistical analysis aims to identify past and present trends and common practices of adopting ANNs in urban geography for geographers' use as a reference in their future research endeavors.
(2) to perform a quantitative meta-analysis using nonparametric bootstrapping. The meta-analysis has two parts. First, it identifies if ANNs perform with the same accuracy whether used for urban prediction or urban land use classification. Second, it compares ANNs performance to other techniques. Results aim to motivate further studies on ANNs and deep learning in urban geography.
(3) to propose a guideline of how ANNs' settings and results should be reported when integrated into urban geography and specifically for urban growth modeling to better disseminate individual lessons learned to the scientific community.

It should be emphasized that this review is not a complete review on ANNs and their applications in general, nor a review reporting all ANNs related papers including / analyzing geospatial data. It presents only a small fraction of ANNs usage in urban geography related problems. It does not focus on the technicalities behind ANNs but on their high performance and importance as tools for geographical analysis. As such it does not indent to unveil all the architectural details of ANNs or the current research gaps in artificial intelligence, but it aims to emphasize on their appropriateness when applied in a geographical context. It also intends to highlight existing problems on the use of ANNs and identify research gaps and future trends in their application when used to solve urban related problems. For a more thorough review on machine learning and deep learning advances in geosciences one can refer to Lary, Alavi, Gandomi, & Walker, 2016 and Zhu et al., 2017.

The structure of the remainder is as follows: Section 2 briefly presents the theoretical background of ANNS and deep learning and also reports the most widely used architectures of ANNs found specifically in this review related to urban analysis. This presentation aims to attract a wider audience that is not accustomed to the ANN jargon and make these methods more approachable for potential use in future geographical analysis. Section 3 focuses on the approach, methods and data used for conducting the systematic review and meta-analysis. Section 4, presents the results and offers a concrete analysis of the findings. It also suggests a schema for reporting ANNs architecture, hyperparameters and results. Finally, section 4 discusses the wider meaning of the overall analysis.

2. ANNs and deep learning

An ANN is a computational system inspired by the way that human brain works, aiming to solve problems through extensive learning from data. ANNs are powerful tools for modeling complex behaviors and patterns (Pijanowski, Shellito, Pithadia, & Alexandridis, 2002). Briefly, an ANN receives inputs and produces outputs as a network of nodes arranged in layers with connections and weights. The way that nodes are layered and connected is typically referred to as ANN architecture. In general, there is an input layer, an output layer, and one or more layers between them, called hidden layers. Each layer has nodes that are called neurons, and neurons are linked through connections. Each connection has a weight. Through an iterative process, ANN adjusts the weights in order to minimize some sort of error and/or to meet specific termination/ stopping criteria. In essence, an ANN is a system that through a special architecture and relatively simple mathematics, adjusts weights in order to identify patterns and valuable information of the input data.

The numbers of hidden layers, connections, and neurons of an ANN depend on the data's complexity. The more complex the data, the more likely is that the ANN will need additional hidden layers and neurons. The process by which the weights of the connections change is performed through the options of a learning method. There are four main categories of learning methods: a) supervised (labeled data), b) unsupervised (unlabeled data), c) semi-supervised (labeled data only for a small portion of the training dataset), d) reinforcement (via a system of sensors that interact with a dynamic environment to obtain and analyze data on the fly). Selecting the appropriate learning option depends on a given problem, how it is conceptualized (e.g., predictive model, clustering, classification), and the available data.

Learning also could be classified into two main categories: shallow learning and deep learning (Deng & Yu, 2014). The distinction between them lies in the complexity of the network's architecture and the depth of the analysis. A shallow learning method (and its corresponding shallow ANN) is associated with one or two hidden layers, whereas deep learning is associated with ANNs that have more than two processing layers and more elegant architecture than shallow learning. Deep learning models can identify patterns in data at multiple levels of abstraction and can be used for pattern analysis, classification, and supervised or unsupervised feature extraction (Lecun et al., 2015). However, deep learning is not simply about the number of layers it uses. It is about the entire architecture, processing functions, and regularization techniques that literally and dramatically change the ANN scenery. Deep learning algorithms are typically applied to problems related to, for example, image classification, face analysis, audio analysis, and speech (Goodfellow et al., 2016).

Shallow ANNS prevailed before 2006 because it was believed that ANNs with two hidden layers could handle most problems. However, it became clear that complex problems could not be addressed with shallow ANNs. Shallow ANNs mainly use the back-propagation (BP) learning algorithm, which is very capable when layers are few, but, as more hidden layers stack up, it suffers from the vanishing (also named exploding) gradient problem (i.e., too small or too large gradients that cause algorithm to get stuck at a local minima). The BP algorithm can perform more efficiently in deeper ANNs that have more sophisticated architectures, better regularization techniques, and new activation functions.

Deep learning also emerged from the need to analyze big data not always successfully handled by shallow ANNs. Deep ANNs' special architectures comprise many hidden layers, which allow them to discover structures in big data and extract valuable knowledge (Lecun et al., 2015). A deep ANN has significantly more parameters to adjust (i.e., weights or biases), and it needs more data to get trained than a typical shallow ANNs need (i.e., 10 × the number of parameters). Building bigger models using more data improves the performance of deep learning algorithms. These functions are like two sides of a coin: Deep learning and deep ANNs can handle big data and big data are essential for successful deep ANNs. A deep learning revolution has been in progress since 2006, and many experimental results support deep ANNs' superiority over shallow ANNs (Bengio, 2012; Glorot & Bengio, 2010; Hinton, 2010; Hinton & Salakhutdinov, 2006; Lecun et al., 2015).

Out of the numerous ANNs existing, the three most commonly used ANNs in this review are the Multilayer Perceptron (MLP), the Self-Organizing Maps (SOMs) and the Convolutional NNs (CNNs), and are briefly explained below (Table 1, Fig. 1).

An MLP has a feed forward architecture, and it is composed of perceptrons arranged in layers. The single perceptron was the first ANN, built by Rosenblatt (Rosenblatt, 1958), that received many inputs and produced one output, which is either zero or one (Fig. 1). In feed-forward ANNs, information moves in one direction, which is forward from the input nodes to the output nodes. The output value is based on the input and a set of weights. MLP is the most common ANN, of which MLPs with two or fewer hidden layers is a shallow ANN, and MLPs with more than two hidden layers are deep MLPs (Fig. 1).

Self-Organizing Maps (SOMs) are popular ANNs that use the unsupervised learning method to project high-dimensional data onto (usually) two-dimensional space (referred to as a "map" or "grid") by mapping similarities of the input data based on topological relationships (Fig. 1) (Kohonen, 1982, 2001). Grids are lower dimensional arrays of neurons arranged as rectangles or hexagons (Gopal, 2017). Each neuron is associated with a weight vector of the same dimension as the input data vector and its location on the map. Because the training is unsupervised, a target vector is not needed (there is no external supervision). SOMs' topological structures are closer to the geographical concepts than some other methods, such as k-means or hierarchical classification.

Convolutional Neural Networks (CNNs) were introduced by Lecun et al. (1989) as specialized for analyzing data (mostly pattern recognition and image analysis) through multiple arrays named "tensors." The architecture of a typical CNN is a structure of stacks of different operational blocks that use a version of the BP algorithm (Fig. 1). The first stacks are convolutional layers, non-linearity layers, and pooling layers. The other stacks are fully connected layers just like typical MLPs. Convolution blocks extract features from an input image. They have the ability to preserve the spatial arrangements of pixels by

Table 1
Common ANN's found on this review, learning algorithm that they use, main characteristic and their architecture and main applications used in an urban context. Scientists who introduced them as well as highly cited papers that describe these ANNs are also presented.

ANN	Learning Algorithm	Main characteristic of architecture	Main applications	Introduced	Highly cited paper
MLP	Supervised	Feed forward	Classification, Prediction	Resenblat 1958 (inspired)	Rumelhart, Hinton, & Williams, 1985
SOM	Unsupervised	Neurons in the output layer have topological relations	Clustering, Dimensionality reduction	Kohonen, 1982	Kohonen, 2001
CNN	Supervised	Deep NN architecture using many layers of filtering and pooling	Image classification, pattern recognition	LeCun et al. (1989)	LeCun, Yoshua, Hinton (2015).

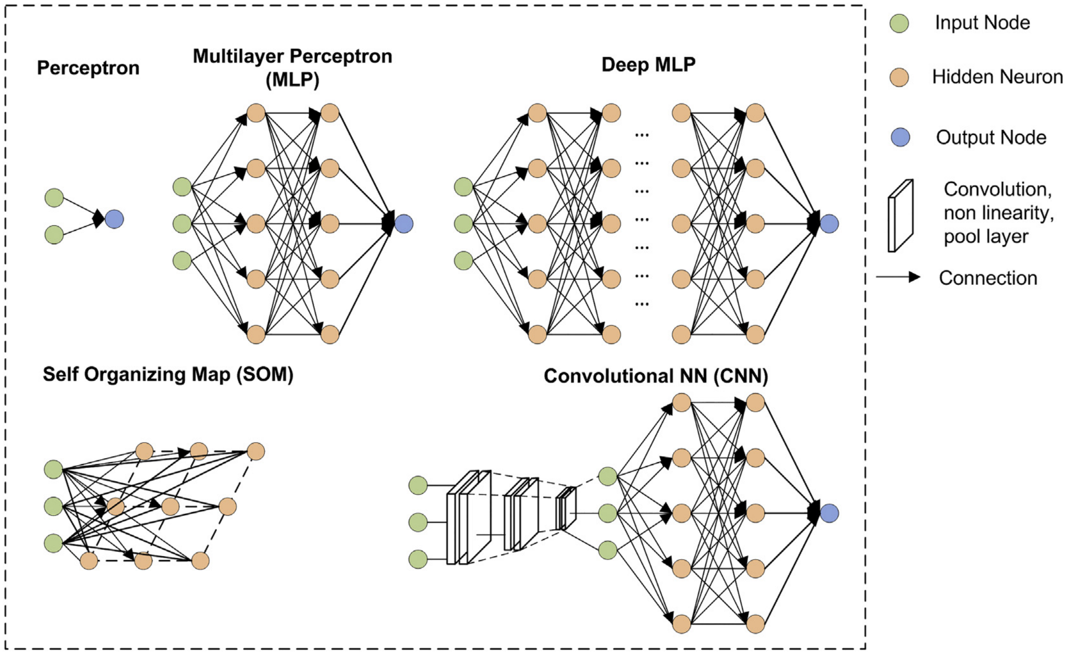

Fig. 1. Most common ANNs in this review and their architectures.

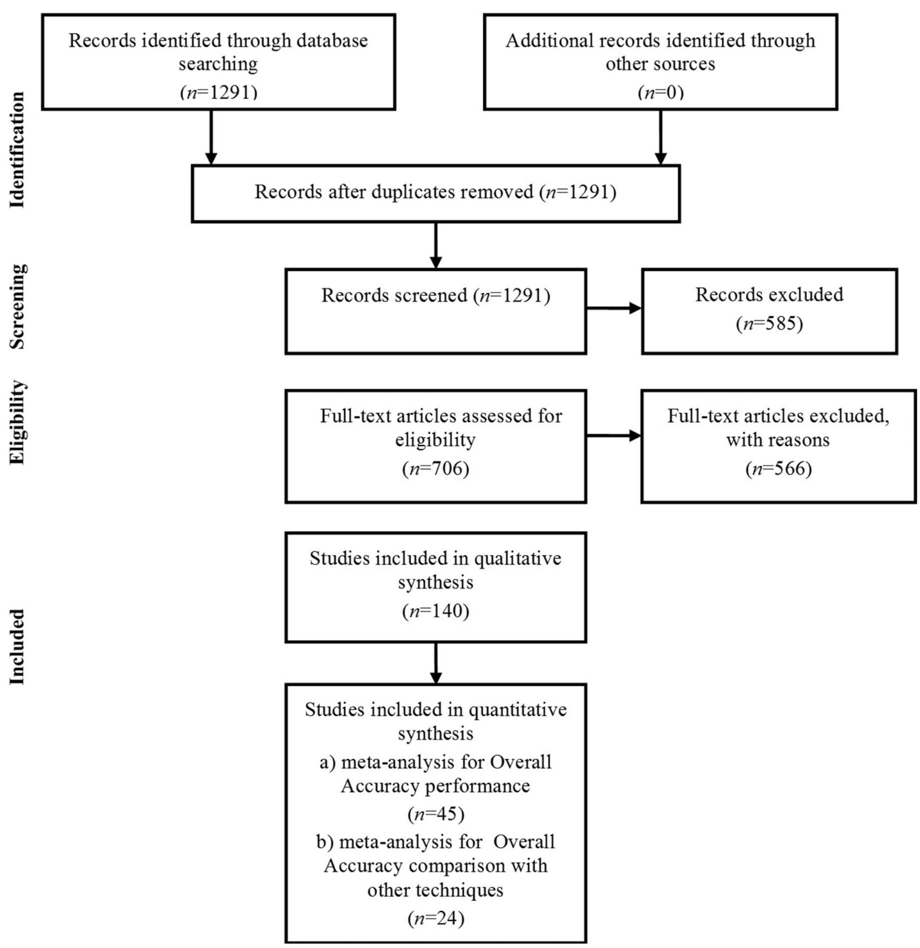

Fig. 2. PRISMA protocol used for sampling.

learning features using small windows called filters or kernels that scan entire images. Each filter produces a unique activation map. The scope of a CNN is to learn the values of these filters during the training phase. The more filters used, the more objects are identified in an image.

3. Method used for systematic review and meta-analysis

3.1. Protocol used to sample the papers

The papers reviewed by this study were selected using the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* protocol (PRISMA) (Moher et al., 2009), which has been used in other geographically related works as well (Khatami, Mountrakis, & Stehman, 2016). Selected papers include studies in English language literature that applied ANN in urban geography between 1997 and 2016 extracted from the ISI Web of Science database. The goal was to evaluate the integration of ANNs in urban geography topics such as urbanization, socioeconomic, urban environment and sustainability. Fig. 2 illustrates the PRISMA selection process and supplementary Table S1 presents the PRISMA statement checklist.

3.2. Eligibility criteria

Due to the vast literature on ANNs, the following criteria to include a paper in the analysis should be met.

(1) Papers should make use of ANNs to solve problems related to urban geography and lead to conclusions valuable from the urban planning and policy perspective. Studies that only use a set of images to test the performance of new ANNs architectures and are focusing solely on the technical advances and not on wider aspects of urban geography are excluded. Additionally, many studies scrutinized referred to particular cities to address problems but not in a geographical context (e.g. using ANNs to predict earthquake casualties). In cases where studies did not consider any geographic concepts, such as location, distance, proximity, or neighborhood, were excluded.
(2) The initial keywords were "neural network," "deep learning," "urban" and "city." For complete list of keywords see supplementary Text S1-Table B. An example of the database search is provided in supplementary Text S1. Articles on studies that did not use ANNs were excluded.
(3) To keep workload reasonable, paper's main topic is further limited to three large thematic groups of urban geography:
    (a). Urbanization group: includes two types of studies: i) urban growth/ sprawl (e.g. predicting urban land growth in near future) and ii) urban land use/cover classification (e.g. impervious surfaces) and change (e.g. detection of urban change during a given period.
    (b). Environment group: includes studies related to urban environment and sustainability such as hazards (urban flooding), urban biomass, or air quality (PM10 concentration in urban environments).
    (c). Socioeconomic group: includes studies related to population growth (distribution, density, or demographic processes), economic activities of the residents, or societal spatial dynamics/behaviors.
(4) Papers whose studies referred to land cover or land-use changes not in urban or peri-urban areas (but focusing instead on large regions, counties, states, or continents) were excluded.
(5) Papers in remote sensing performing image analysis through ANNs (i.e. visual recognition, object detection, semantic segmentation, semantic labeling etc.), but not in the context of the three thematic groups described above (urban land cover change, urban environment, urban socioeconomic analysis) are not included. In addition, remote sensing pares focusing only to very small geographical areas (e.g. some building blocks) to test the proposed techniques are not included as they mainly focus on the methodological advancements of ANNs, which is not the scope of this review.
(6) If more than two methods are tested for prediction or classification then ANN accuracy is compared to the method that achieved the largest overall accuracy. This study used overall accuracy to evaluate the performance of a classifier or predictor.

3.3. Database search

The query used to search the database is presented in supplementary Text S1. The search was conducted on May 5, 2017, in ISI Web of Science, covering 1997 through 2016. Using the search criteria, 1291 papers were retrieved and refined as journal articles, reviews, or editorials (Fig. 2). The results were further refined using Web of Science categories to narrow the results to areas related to geography (e.g., geography, urban studies, and environmental studies), which yielded 706 papers (for details, see supplementary Text S1). These 706 papers were meticulously inspected manually, and 140 papers that met all of the criteria were retained (see Text S2 for full list of publications used). These 140 papers reported on studies that focused on applications of ANNs in urban geography in the three thematic groups. Of these, 140 papers were used for qualitative synthesis, and a sub-sample of 45 papers was used for the meta-analysis. Meta-analysis referred to the OA achieved by an ANN for a classification or prediction in urbanization thematic group studies. Comparisons of OA achieved by other methods are also made in 23 studies (subset of the 45 studies used for meta-analysis).

3.4. Meta-analysis

The scope of this meta-analysis is to summarize the research results across individual and independent studies related to the overall accuracy (OA) performance of ANNs when used in urban studies. Specifically, this meta-analysis focuses on studies dealing only with urban growth prediction or urban land use classification. The main reasons for focusing on these two topics are that: (1) modeling and monitoring urban evolution are critical to urban planners, decision-makers, environmental specialists, and other stakeholders aiming to achieve sustainable development; (2) these types of studies were dominant in the sample; and (3) they allowed for analysis of thematically-related studies in the same research group. Consequently, meta-analysis studies are pooled from the urbanization thematic group. OA is the standard metric for accuracy in this study. Although OA is not the best quantitative measure to assess the performance of a method, it is preferred in this review first, because it is most frequently reported in the specific sample in comparison to other metrics (i.e. other metrics were reported only to a few studies and as such a critical mass of studies could not be formed for further analysis) and second, because it is more widely used in general. No other effect size was calculated.

The meta-analysis is performed in two parts, as follows.

(1) The first level of analysis identifies differences in the OA values obtained when ANNs were used for urban prediction compared to those obtained when ANNs were used for urban land-use classification. This level of analysis does not compare ANN results to those of other algorithms.
(2) The second level of analysis compared the ANNs' performances to the performances of other algorithms and techniques.

Both levels of the meta-analysis are performed using nonparametric bootstrapping. Bootstrapping is widely used in many disciplines, including econometrics, biology, and engineering (Chernick, 2007). It has been also used in geography to estimate the global mean urban growth rate of 292 cities across the world (Seto, Fragkias, Güneralp, & Reilly, 2011). The main reason for using this method is that the validity of the

OA in the sample was limited to each study's scope and dataset. Comparisons could not directly be made across the studies because of the differences in images, ANN architectures, datasets, case studies, and methodological approaches. Therefore, we cannot directly answer some questions, such as “Are ANNs a relatively better technology to implement in urban geography?” and “What is the extent of improvement to urban modeling achieved by using ANNs and AI?” In addition, researcher bias regarding the choice of alternative technique used to compare to an ANN should be considered.

To overcome these problems, this study used nonparametric bootstrapping. Bootstrapping is a powerful statistical method for estimating quantities (e.g., variance, median, or mean) from a data sample by measuring those quantities when sampling from an approximating distribution (Brownlee, 2017). Bootstrap is actually a resampling procedure involving resampling from the original data set (Chernick, 2007) (for more details on the method see supplementary Text S1).

4. Results

4.1. Statistical review

In this section a statistical review of papers is conducted based on: (1) their regional, break down; (2) thematic breakdown; (3) cross-comparison between regional and thematic breakdown; (4) temporal breakdown; (5) types of ANN architecture and data used in the studies; (6) types of journals in which the articles were published; and (7) a comparison of the ANNs' performances compared to alternative methods.

4.1.1. Regional breakdown

In this work, there are more case studies than papers, as many papers include several case studies. Altogether, 161 case studies (on cities) are reported in 111 unique geographical locations in 34 countries on all continents except Antarctica (Fig. 3). There are more case studies than geographic locations as there may be multiple case studies for a single location. The most studied urban agglomerations are Beijing (9 case studies), Guangzhou (5 case studies), Hangzhou (5 case studies) and Rome (5 case studies) (See supplementary Text S1:Table B). The most studied countries are China (47 case studies (out of 161) 29.2%) and the USA (29 case studies 18.0%) were both account for nearly 50% of the works (See supplementary Text S1:Table C).

The regional breakdown of the case studies follows the United Nations' UN-M.49 classification scheme (UN, 2014) (Fig. 3). The first and second levels of classification were combined to meet this study's goals, and the regions used in the study were Africa, North America, Latin America and the Caribbean, Europe, Oceania, Asia, and China. China was treated as a separate region because of the large number of papers on studies conducted there (n = 47, 34%). China was followed by Europe (41 case studies, 25.5%). North America and Asia (except China) had similar numbers of case studies (32 and 31, respectively). Although most of the world's fastest growing cities are in Asia and Africa (UN, 2016), the least studied region in the sample was Africa. Only 2.5% of the articles were on studies of urban areas in Africa (Fig. 4A).

4.1.2. Thematic breakdown

The thematic breakdown reveals that almost two thirds (65%) of the papers (91 papers) deal with the urban thematic group and 20% (28 papers) concerned urban environmental issues (Fig. 4B). A smaller proportion of papers 15% (21 papers) were on socioeconomic issues (See Text S1:Table D). Of the 28 urban environmental studies, eight are about urban hazards, seven concerned urban air quality, and 13 covered other urban environmental topics. Of the 91 papers on urbanization, 34 of them were about urban growth/sprawl modeling/prediction and 57 of them concerned urban land use/cover classification and change detection.

4.1.3. Cross-comparison between regional and thematic breakdown

The thematic group that a paper belongs is further analyzed relatively to the region where the case study was conducted. Table 2 reports the percentages of papers in the socioeconomic, environmental and urbanization thematic groups in the seven regions.

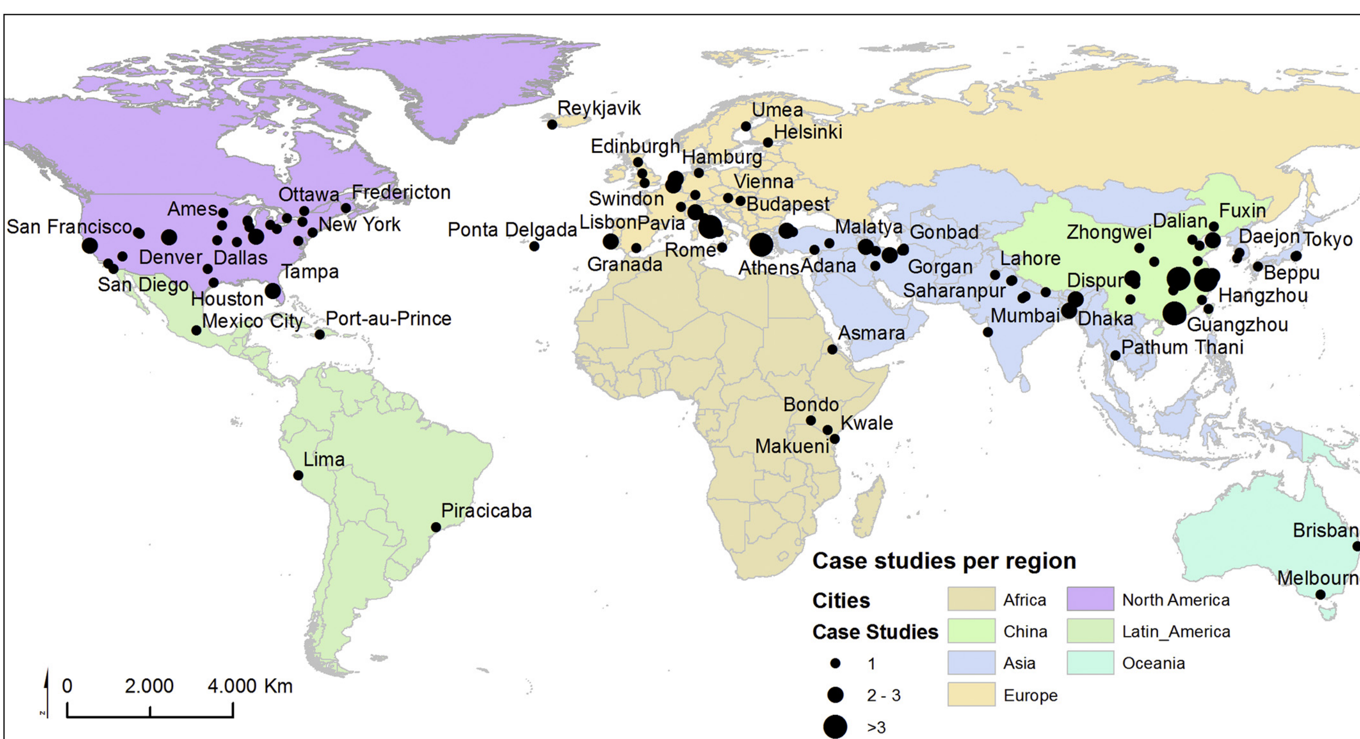

Fig. 3. Geographical distribution of 161 case studies over seven regions as defined by the UN-M.49 classification scheme.

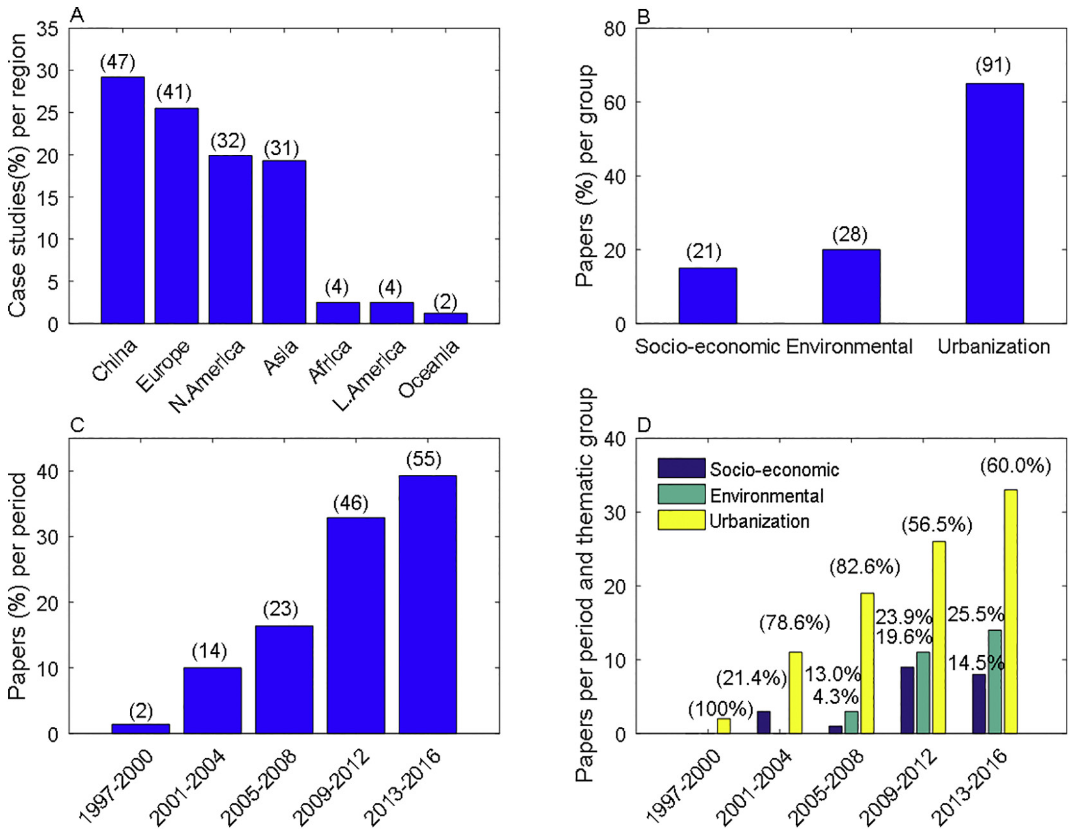

Fig. 4. (A) Percentage of case studies by region, with numbers of case studies in parentheses; the total number of case studies are 161. (B) Percentage of papers by thematic group, with numbers of case studies in parentheses; the total number of papers are 140. (C) Percentages of papers over five four-year periods, with numbers of papers in parentheses. (D) Numbers of papers by thematic group and by study period, with percentages in parentheses representing the percent of each thematic group during a given period; e.g., 19.6% of the articles in 2009–2012 concerned socioeconomics, 23.9% were on the urban environment, and 56.5% were on urbanization.

Table 2
Thematic group breakdown by region. Each line adds up to 100.0%. Numbers in parentheses are the absolute numbers by thematic group and by region. Each column (numbers in parentheses) adds up to the total number of papers per thematic group also reported in Fig. 3B.

Region	Socioeconomic	Environmental	Urbanization	Total
China	11.1% (4)	22.2% (8)	66.7% (24)	100.0%
Europe	25.8% (8)	22.6% (7)	51.6% (16)	100.0%
N_America	9.1% (3)	15.2% (5)	75.8% (25)	100.0%
Asia	9.4% (3)	15.6% (5)	75.0% (24)	100.0%
Africa	50.0% (1)	0.0% (0)	50.0% (1)	100.0%
L_America	25.0% (1)	50.0% (2)	25.0%(1)	100.0%
Oceania	50.0% (1)	50.0% (1)	0.0% (0)	100.0%
Total Papers	(21)	(28)	(91)	(140)

The results for Africa, Latin America and Oceania are not further analyzed because the percentages and numbers of papers are too small to provide valuable information. China has a large percentage (66.7%) on studies focusing on urbanization issues. On the other hand, Europe is the least studied region regarding urbanization (51.6%) but the first in both socioeconomic (25.8%) and environmental topics (22.6%).

As stated above, China and Europe are the most studied regions with 29.2% and 25.5% of case studies. Still, the focus in these two regions is quite different. China is continuing its rapid and accelerating urbanization, which has focused research attention on urban sprawl/urban change (Du, Xia, & Feng, 2015; Gong, Chen, Liu, & Wang, 2014; Yang, Zhao, Chen, & Zheng, 2016) and environmental issues (Liu et al., 2015). As a consequence, socioeconomic analysis is only struggled to a low 11.1% (Yao, 2012) of the studies. On the other hand, in Europe with a relatively stable population and not high migration fluxes, urbanization is not a hot topic during the study period, and only about one-half of the studies were about this topic. Possibly due to 2008 Global Financial Crisis, research interest has shifted mainly to socioeconomic aspects and urban geographers directed their attention towards complicated problems such as geodemographic segmentation (Gorricha & Lobo, 2012; Grekousis, Manetos, & Photis, 2013), housing markets (Kauko, 2009), and urban healthcare (Grekousis & Photis, 2014; Photis & Grekousis, 2012).

Asia (omitting China) is another region that focuses on urbanization issues. Asia is also experiencing huge urbanization rates and research overwhelmingly concerned urban growth and urban land use/cover detection and change (75.0%). Hence, there is an evident lack of crucial studies on the urban environment (15.6%) and socioeconomics (9.4%), which are important because urbanization tends to create environmental problems and to significantly influence urban residents' socioeconomic status and potential. North America has similar percentages with Asia and the largest parentage of all regions (75.8%) in urbanization studies of the seven regions. North America is not experiencing urban growth at the same rate as Asia because the United States and Canada have initiated smart growth policies for controlling urban growth. In fact, most of the studies on the urbanization thematic group, improve urban land use/ impervious surface classification methods by using high resolution images and object based classification (Pacifici, Chini, & Emery, 2009; Pu, Landry, & Yu, 2011; Salehi, Sahebi, & Maghsoudi, 2014), combine various methods (Tayyebi & Pijanowski, 2014), or use deep learning artificial neural networks (Zhang, Gong, Su, Liu, & Li, 2016) and they deal less with typical urban growth analysis.

Table 3
Percentages of papers studied by model type, ANN architecture and data categories. Numbers in parentheses are the absolute numbers of papers. Percentages are not referred for data type as some studies combine more than one data types.

	Model Type		ANN Architecture					Data				
	Classification	Prediction	MLP 1 layer	MLP 2 layers	KSOM	CNN	Other	Landsat	Other RS	GIS	SE + GIS	RS + SE + GIS
Papers	60% (84)	40% (56)	65.0% (91)	7.1% (10)	11.4% (16)	1.5%(2)	15.0% (21)	48	46	23	28	5

4.1.4. Temporal breakdown

A temporal breakdown to track how ANNs are integrated into urban geography is also carried out by dividing the 20-year study period (1997–2016) into five four-year groups (Fig. 4C). Growing interest was evident during 2009–2016 years because 72.2% ($n = 101$) of the studies using ANNs were conducted during the two most recent periods (2009–2012 and 2013–2016). In particular, the number of studies using ANNs doubled between 2005 and 2008 ($n = 23$) and 2009–2012 ($n = 46$). When the temporal breakdown was cross-checked to the thematic breakdown, a consistent interest across the years in using ANNs in urban geography was revealed, mostly regarding urbanization (Fig. 4D). However, after 2008, socioeconomic and environmental studies had growing proportions of studies using ANNs. For example, before 2008, < 20% of the studies were on socioeconomic or environmental issues, whereas, after 2008, the proportion doubled to about 40%).

4.1.5. ANN architecture and data

From a technical perspective, 60.0% of the papers ($n = 84$) are of studies on classification problems (land use land cover, socioeconomic segmentation) and the rest 40.0% are dealing with prediction (urban growth, pollution etc.) (Table 3). The most commonly used ANN architecture is the one hidden layer Multilayer Perceptron (MLP) (65.0%, 91 studies). MLP with two hidden layers is used in 10 studies (7.1%) and Kohonen Shelf Organizing Map (KSOM) in 16 (11.4%) studies. Deep Convolutional NN (CNN), Adaptive Resonance Theory MAP (ARTMAP), Fuzzy ARTMAP, Radial Basis Function NN (RBNN), Generalized Regression NN (GRNN) are other types of architecture used. Only two studies (1.5%) make use of deep learning through CNN for image classification related to urban land cover/ land change (Längkvist, Kiselev, Alirezaie, & Loutfi, 2016; Zhang et al., 2016). Thus, the focus currently is on more conventional ANNs' architectures as for example MLP.

The use of ANNs in urban geography tends to be accompanied by the use of remote sensing (RS) data, such as Landsat, Worldview, Quickbird, Aster, and other types of satellite imaging. Landsat images are the most widely used (48 papers) which is more than all of the other types of satellite imaging combined ($n = 46$). GIS data are exclusively used in 23 papers. GIS data typically refer to spatial data such as boundaries, or data extracted from GIS process that might include distances from given points/areas, slope, altitude, area, land use land cover, and topographic maps. Socioeconomic data and GIS data are combined in 28 studies, and five studies, all published after 2010, combined remote sensing, GIS, and socioeconomic data in studies of interdisciplinary research (Bhatti, Tripathi, Nitivattananon, Rana, & Mozumder, 2015; Gong et al., 2014; Mozumder, Tripathi, & Losiri, 2016; Pijanowski, Tayyebi, Delavar, & Yazdanpanah, 2010; Wang, Shao, & Kennedy, 2014).

4.1.6. Types of journals

To gain a detailed understanding of ANNs' applications in urban geography, it is useful to know the scientific journals that are publishing these types of papers. This analysis is presented in supplementary Text S1 (Tables E,F) and the main finding is that the majority of the studies referring to ANNs and urban geography are published on Remote Sensing related journals and not on urban studies journals.

4.1.7. Comparisons with other methods

Of the 140 papers, 44 reported comparisons of ANNs to other methodologies (logistic regression, random forests, linear/stepwise regression, maximum likelihood, logistic decision trees, support vector machines (SVM), Fuzzy SVM, and K-nearest neighbor). However, the comparisons were made using different criteria (e.g., OA or kappa). ANNs were assessed as better methods in 64.4% ($n = 29$) of the studies. When compared to relatively conventional methods only (logistic regression, random forests, linear/stepwise regression, maximum likelihood, logistic decision trees), the percentage was even higher 75.7% (28 of 37).

This analysis is based on whether an ANN is better than an alternative method regardless of the metric used. A more thorough analysis based on bootstrapping method is performed in the meta-analysis section next, which considers only those studies that reported OA values on the ANNs and on the other methods used.

4.2. Meta-analysis results

4.2.1. Comparison of papers on prediction to those on classification

Out of 91 papers in the urbanization group, a set of 45 papers are selected that report OA. The other 46 papers in this thematic group did not report OAs or they reported other accuracy measures, such as Kappa statistics. These 45 papers were further sorted into a group of studies that performed urban land-use classifications ($n = 31$) and a group that predicted urban growth ($n = 14$). The bootstrap method is used to measure uncertainty over the mean OA reported in each group and the groups' results are compared to each other. For each group, 1000 resampled datasets are drawn by replacement, keeping the size of each new dataset equal to the size of the original dataset. Then, the mean OA value of each subsample is calculated. Finally, the OA mean of the original dataset is estimated, and the quartiles of the bootstrapped distribution of means are reported.

The bootstrapped OA mean value of the classification group is 87.10%. The 95% confidence interval of the mean is (87.01%, 87.17%). This result provides strong quantitative evidence that the OA result when ANNs are used for classification of urban land use is high. The Kernel probability density estimate of the bootstrapped means is also calculated (Fig. 5A).

The bootstrapped mean of the prediction group is 85.44%. The 95% confidence interval of the mean is (85.32%, 85.58%), which is strong quantitative evidence that the OA when ANNs are used to predict urban growth is high. The Kernel probability density estimate of the bootstrapped means is also calculated (Fig. 5A).

To test for a statistically significant difference between the classification group and the prediction group, the Mann-Whitney median U test is applied (Mann & Whitney, 1947). The null hypothesis is H_0: There is no difference in the median OA value between the ANNs used to classify urban land and the ANNs used to predict urban growth. The calculated p-value is 7.2e-86, which is less than the 0.001 significance level. There is a compelling evidence to reject H_0 and conclude that these two distributions are statistically different (Fig. 5B). The OA of the prediction studies is, on the bootstrapped averages, 1.7% less than that of the classification studies, suggesting that ANNs are more accurate when they are used in urban land-use classifications than when they are used for urban growth predictions.

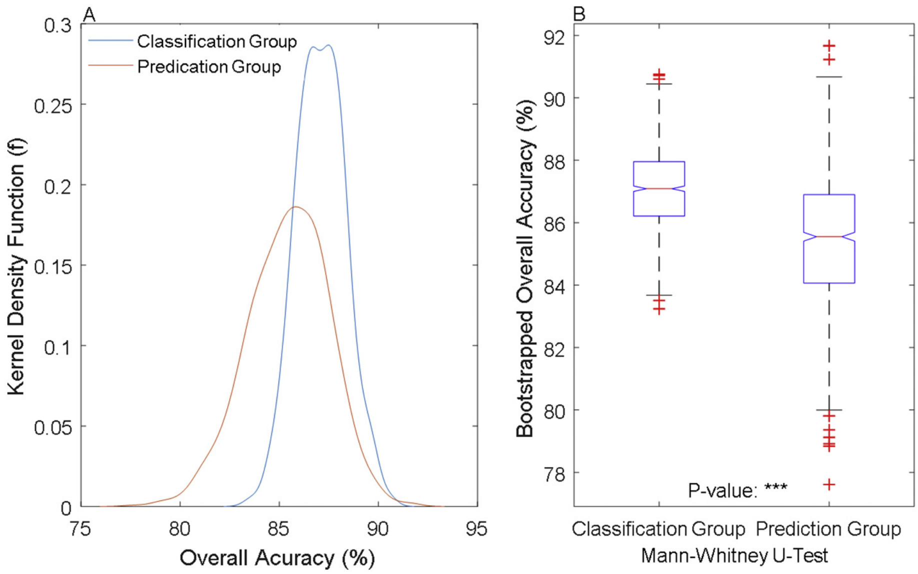

Fig. 5. (A) Kernel probability density estimate of bootstrapped means and (B) boxplots of the classification and prediction groups, including median, 25th, and 75th percentiles (marked at upper and lower edges of boxes). Red crosses indicate outlier values. *** = $p \leq .001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

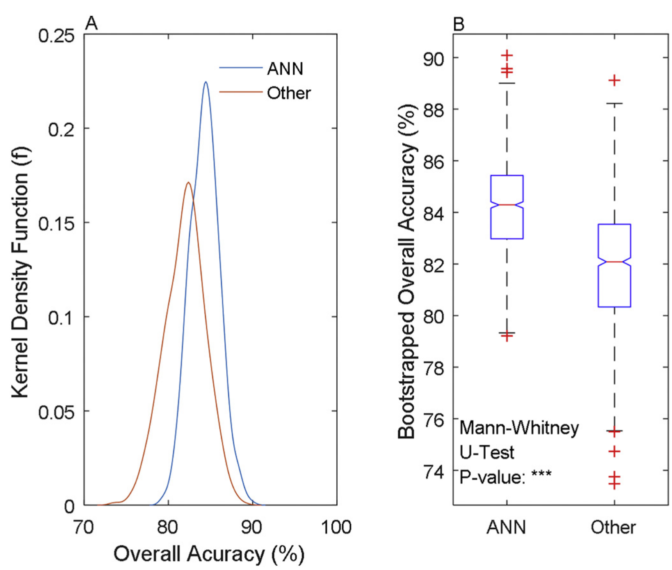

Fig. 6. (A) Kernel probability density estimate of bootstrapped means and (B) boxplots of the OA achieved by ANNs and the OA achieved by other techniques. *** = $p \leq .001$.

4.2.2. Comparison of ANNs to other methods

Of the 45 papers above, only 24 reported comparative results of the OA to other methods. In the following analysis, no distinction is made between the classification and prediction studies. The bootstrapped mean value of the OA derived from the ANNs is 84.28%. The 95% confidence interval of the mean is (84.17%, 84.38%), which is strong quantitative evidence that the OA achieved when using ANNs for classification or prediction is high. The Kernel probability density estimate of bootstrapped means is also calculated (Fig. 6A). The bootstrapped mean value of the OA of the same 24 studies using other techniques is 81.95%. The 95% confidence interval of the mean is (81.81%, 82.11%). The Kernel probability density estimate of the bootstrapped means also calculated (Fig. 6A).

The Mann-Whitney median *U* test is used to test for a statistically significant difference between the two distributions. The *p*-value is 1.7e-104, which is smaller than the 0.001 significance level, which is compelling evidence to reject the null hypothesis and conclude that these two distributions are significantly different (Fig. 6B).

On bootstrapped averages, the median OA achieved when using ANNs is higher (statistically significant different) than OA achieved by other techniques by 2.3% ($p < .001$), indicating that ANNs performed significantly better in these studies.

4.3. Guidelines for reporting ANNs

A major finding of the above review and meta-analysis is that most of the papers, do not provide sufficient information on the processes used to build the architecture of an ANN and then on their evaluation. Close and critical inspection of ANNs' performance is crucial to avoid over-fitting or under-fitting. Training error, testing error, validation error, or hyper-parameters settings, are scarcely mentioned if not mentioned at all. Specifically, hyperparameters' optimization is a key research topic regarding ANNs (Snoek, Larochelle, & Adams, 2012). To build a successful ANN, many model hyper-parameters must be set outside the learning algorithm. This tuning, demands experience, rules-of-thumb, and/or a brute-force search, and successful results are not guaranteed. Therefore, reporting hyper-parameters is very important to guide future researchers with successful or unsuccessful settings, particularly for geographical problems where complexity is intrinsically increasing. In addition, most of the papers did not refer to other ANN architectures tested before the final architecture definition. As a result, individual lessons learned through the setting up of ANN architecture are not diminished in the scientific community leaving a substantial gap.

For this reason, this study suggests some guidelines for reporting ANNs information and related graphs when used especially in urban growth modeling (Table 4). The suggested guidelines take the form of a table that can be included in publications in the main body of a manuscript or as a supplementary file (Supplementary Table S2 offers an empty version of the table that can be reproduced in other studies).

These guidelines aim: (1) to offer a standard way of presenting results making thus easier for other researchers to delve deeper into the details of the adopted architecture and the performance of the ANN used; (2) to assist in better identifying relationships of selected hyperparameters values to ANNs' behaviors by systematically reporting them in published papers; (3) to be included in benchmark datasets for the evaluation of results for urban growth modeling (more on this point in the discussion section next).

5. Discussion and conclusions

The purpose of this review was to summarize the extent and nature

Table 4
Guidelines on reporting ANNs. The first column refers to the metrics that should be included in a published report related: (1) to ANN and data; (2) to hyperparameters' selection; and (3) to ANN's evaluation. The second column reports the values of the metrics of the first column. Here it is used as an indicative example. Finally, the third column offers additional information related to the graphical representation of the metrics.

Suggested listing	Example	Additional Info
Data and ANN		
Name of ANN used and learning algorithm	Multilayer Perceptron, Supervised	In case of user defined ANN, system architecture should be presented in a graphical way (See for example Fig. 2)
Original dataset	Number of instances: 1000 spatial units. Each unit stands for a line.	Related link to download data
Training dataset	Number of instances used and percentage of total dataset: 700 objects, 70% of the dataset	
Testing dataset	Number of instances used and percentage of total dataset: 150 objects, 15% of the dataset	
Validation dataset	Number of instances used and percentage of total dataset: 150 objects, 15% of the dataset	
Cost function	Mean Square Error (MSE)	
Activation function	ReLU	
Hyperparameters		
Number of hidden layers	2	
Learning rate	0.1	
Momentum	0.9	
Mini-batch size	1	
Number of epochs	1000	
Neurons per hidden layer	250,250	
Regularization coefficient	$\lambda = 0.5$	
Dropout probability	0.7	
Evaluation		
Training error	0.005	A graph depicting training error versus validation error should be used to locate potential overfitting
Validation error	0.025	
Testing error	0.055	
Correlation coefficient	0.80	
Stopping criteria	1000 epochs	In case of early stopping method, a relevant graph should be used to locate the point where validation and training error began diverging

Table 5
Key findings of review and meta-analysis along with suggestions, gaps and future trends.

	Key findings	Suggestions/Gaps/Future trends
Performance perspective	ANNs are more accurate than other methods used for urban growth prediction and urban land-use classification in 75.7% of the cases.	This high 75.7% is a strong evidence that ANNs performed better than other methods
	On bootstrapped averages, the mean OA of the ANNs is higher than that of the other techniques by 2.3%.	This is an indication that ANNs can handle geographical data better than other techniques and can be further integrated into geographical and urban studies.
Thematic perspective	Currently, the majority of studies deal with urban growth or urban land use analysis while there is less interest in environmental studies and even less in socioeconomic studies.	As urban problems are intrinsically interdisciplinary, more research should be guided towards the use of ANNs and deep learning in solving environmental and socioeconomic problems of urban life.
Technical perspective	MLP is the most commonly used ANN in urban geography studies	Many newer ANN architectures have emerged, and it is essential to experiment with them. Although deep-learning ANNs are widely used in other scientific disciplines, urban geography lacks such use, which is a clear gap to be filled in the coming years.
Data perspective	Most of the ANNs reported in this review used satellite data	Large caches of socioeconomic data and big data derived from mobile phones or social media are readily available and should be integrated into urban research as ANNs are capable of handling complex data.
Publishing perspective	Most of the papers are published in remote sensing journals	There is a need for publishing more works on journals related to urban analysis.
Benchmark perspective	Benchmarking is missing especially when referring to urban growth studies. Creating benchmark datasets and benchmark frameworks is challenging in urban geography context	Labeled data at various spatial and temporal scales refereeing to urban growth for various geographical areas around the world could be established in a similar way to the benchmarks related to remote sensing
Report perspective	There is a need for better reporting ANN architecture, hyper-parameters, and results in urban studies	Guidelines for reporting results are suggested in this paper. This will also allow for creating better benchmark frameworks.

of the use of ANNs and deep learning in urban geography topics such as urbanization (e.g. urban growth, urban land change), urban environment (e.g. hazards, pollution) and socioeconomic aspects, allowing common threads to emerge. Key research gaps and future trends that identified in this study, are presented below and are also briefly summarized in Table 5. To keep this section reasonable in size, detailed analytical results, such as regional or temporal differences, are discussed above in the results section.

From the thematic perspective, this overview reveals that:

- During the period 1997–2016 the majority of research efforts (65% of the papers) have been focused on the use of ANNs to address urban growth predictions and urban land use classifications. In contrast, there is less interest in environmental studies (20%) and even less in socioeconomic studies (15%).
- However, interest in those areas is increasing. Before 2008, < 20% of the studies concerned socioeconomic and environmental problems, whereas, after 2008, this proportion doubled to about 40% (Fig. 4D).

- This finding highlights the increasing interest in interdisciplinary studies that combine data, methods, and concepts across environmental and social disciplines to address various urban phenomena.
- It is suggested in this paper that as urban problems are intrinsically interdisciplinary, more research should be guided towards the use of ANNs and deep learning in solving environmental and socio-economic problems of urban life.

From the technical perspective:

- The MLP is the most common ANN in urban geography, and its wide application to urban studies indicates a strong confidence in its performance for modeling complex nonlinear problems.
- However, many newer ANN architectures have emerged, and it is essential to experiment with them. For example, deep-learning ANNs are widely used in some scientific disciplines, but urban geography lacks such use, which is a clear gap to be filled in the coming years.

From the data perspective:

- Most of the ANNs reported here used satellite data.
- However, large caches of socioeconomic data are readily available and offer invaluable information for urban studies. For example, there is a lack of using ANNs to predict urban growth and urban change, integrating socioeconomic big data (Grekousis et al., 2013). Such data, derived from mobile phones or social media, should be integrated into urban research as ANNs are capable of handling complex data.

From the publishing perspective:

- Most of the papers are published in remote sensing journals (supplementary Text S1: Tables E and F). There is a lack of these studies in geography-related scientific journals, although urban studies offer a foundation for many uses of ANNs.
- This paper emphasizes the need for publishing more works on journals related to geography and urban analysis. Related journals should consider dedicating special issues to AI and machine learning in geography. These actions would foster the circulation of ideas creating an open environment for future submissions and advancements in the field.

From the reporting and benchmarking perspective:

- This study reveals a need for better reporting on ANN architecture, hyper-parameters, and results in urban studies as in most cases, an ANN is used as a methodological tool without thorough analysis or comment (see more on the discussion below). This study highlights the importance of a standardized way to report ANNs' performance and provides detailed guidelines so that knowledge gained related to ANNs' neural network architecture, hyper-parameters settings and results during the research process is better disseminated to the scientific community. Successful ANNs may be then reproduced/tested to other settings.
- Moreover, benchmarking is something that is missing especially when referring to urban growth studies.

In more detail, benchmark data-sets are used to evaluate and compare the performance of various machine learning techniques to solve problems in several domains as for example image classification (Rottensteiner et al., 2014). In a wider context, benchmark analysis tests a model's performance against predefined references (i.e. benchmarks) to evaluate the model's strong and weak points (Luo et al., 2012).

Benchmarking can be regarded as a substantial reason of the success of deep learning especially when it comes to image analysis. For example, the ISPRS benchmark on urban object detection provides a data repository of training images, along with evaluating and ranking results of various approaches through common statistical strategies.[1] This and other benchmarks offer a way to analyze common datasets, evaluate results and discuss on new machine learning approaches proposed, fostering thus discussion and further debate.

Similar well organized and defined benchmarks for urban growth could be quite helpful. In fact, the need for benchmarking land models has been already highlighted and frameworks for evaluating land model designed to predict future states of ecosystems and climate have been suggested (Luo et al., 2012). Urban land growth prediction is far more complicated than for example image analysis. Such modeling attempts to offer valuable understanding on how humans' activities change the urban landscape by identifying the driving forces of urban land cover change summarized into four main groups: Population, Economy, Transportation and Governance (policies, regulations) (Seto & Ramankutty, 2016). As such, the complexity increases and existing urban growth models vary significantly, both conceptually (perspective of analysis i.e. drivers they select to analyze) and methodologically (data used, methods applied, geographical and temporal scale). Consequently, creating benchmark datasets and benchmark frameworks is challenging in this context. Although these type of benchmark datasets might not get as vast as those in the remote sensing field, still they will offer valuable information in the evaluation of results in measuring complex phenomena as urban growth, and this work strongly emphasizes on this need.

In this respect, labeled data at various spatial and temporal scales refereeing to urban growth for various geographical areas around the world could be established in a similar way to the benchmarks related to remote sensing (i.e. ISPRS). For example, historical satellite images, temporal census data, and administrative boundaries, along with developed ANNs and associated results could be available in a harmonized way to encourage participation of scholars mainly utilizing more 'traditional' methods. Such benchmarks will allow for better results evaluation dissemination and progress in the field.

A first step towards this direction is the suggested guidelines for reporting results as described in the previous section. These guidelines could be supplementary to existing statistical procedures and tests used in current benchmarks, having a focus mostly on the ANNs structure and performance. These guidelines can be also followed when publishing results in reports or scientific journals.

According to these guidelines, there should be at least some basic metrics used to report assessments of an ANN's performance beyond the OA of a classified image or the prediction accuracy of urban growth. This would improve the comprehensibility of various ANNs' topologies and provide guidance on ways to improve urban and geographical research with ANNs. The guidelines given in this study aim to significantly assist researchers by proposing a standard way of presenting the setting up of architecture, hyper-parameter optimization and reporting results of the adopted ANN.

From the performance perspective:

- Meta-analysis reveals that the ANNs' OA achieved in the studies on urban land-use classification is 1.7% higher on bootstrapped averages (statistically significant $p < .001$) than the OA achieved in the studies predicting urban growth. However, prediction is considered a relatively more complex problem than classification, and the bootstrapped mean value of 85.44% is quite promising for getting larger with future research efforts.
- The review also reveals that ANNs achieved higher OAs compared to

[1] http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html. Other benchmarks datasets can be found at http://www2.isprs.org/commissions/comm2/wg6/bench.html.

other conventional techniques. Specifically, ANNs are more accurate than other methods used for urban growth prediction and urban land-use classification in 75.7% of the cases. Although a researcher's selection of a comparison method might relate to researcher bias, this high 75.7% is a strong evidence that ANNs performed better than other methods.
- Meta-analysis provides a similar argument. On bootstrapped averages, the mean OA of the ANNs is higher (statistically significant, $p < .001$) than that of the other techniques by 2.3%, indicating that the ANNs performed significantly better.
  This is an indication that ANNs can handle geographical data better in some cases than other techniques and can be further integrated into geographical and urban studies.

Although the present review and meta-analysis provide evidence of good performance of ANNs, still they cannot be considered as a remedy to every problem. For example, when problems have determinate solutions, then conventional approaches may yield similar or even better results than ANNs. In addition, when the dataset used is not adequate to support the learning process (i.e. not many training examples), then ANNs are incapable of producing reliable results. These two points should be taken into consideration when a researcher decides to utilize an artificial intelligence approach, as many times ANNs are inappropriately used.

From the urban geography point of view, ANNs could be useful in solving large complex problems when big data are available. A city is a massive, interconnected and complex system, which involves almost an infinite number of interactions among various agents and elements (Lai & Han, 2016). Therefore, complex and interdisciplinary urban problems such as simulating urban growth and dynamics, analyzing urban pollution-climate change, studying urban poverty, urban unemployment, urban health, analyzing urban resilience and urban sustainability can be addressed through comprehensive artificial intelligence modeling (Grekousis & Gialis 2018). The reason is that ANNs have the ability to learn, adapt and improve on the basis of their experience gained through extensive data analysis (Shai & Shai, 2014). When the goal is to identify nonlinear relationships and patterns in data, as in many complex urban problems, then ANNs can be used as they intrinsically handle nonlinear highly complex data better than conventional methods (Pijanowski et al., 2014).

Still, it depends on the scope of the analysis whether ANNs are suitable within the urban geography context. When the focus lies on the explanatory power of the model and on the identification of casual relationships, then ANNs might not be as appropriate as other techniques, as their explanatory capabilities are weak. For example, when the objective is to explain the causes of urban growth (or in general a cause and effect relationship), ANNs are not particularly descriptive as they do not provide equations that link input to output variables, so causal direction is not clear. In this case, some conventional methods, as for example logistic regression or spatial econometrics, offer a better explanation in the identification of the driving forces behind urban land use dynamics. On the other hand, when the interest is on the predictive capability of the model (either for classification or prediction) and in combination with the existence of big data, then ANNs and deep learning offer strong evidence for reliable modeling.

There is a large debate on this topic (selecting between reliable modeling and explanatory capabilities). Many scientists support the idea that it is better to have strong models, although it might not be perfectly clear how the job is done, than to have a clear understanding of an inadequate model. If the research problem is to model future urban growth, it is probably a reasonable approach to obtain accurate predictions even if we partially sacrifice understanding of the underlying driving forces. However, if we want to analyze the drivers of various urban phenomena, conventional statistical approaches, exploratory statistics, and spatial statistics can be combined with ANNs to build explanatory models. Although a conceptual framework is preferred from many scientists to deal with complex urban problems in contrast to models that simulate and explain urban complexity, still these models can provide quantifiable evidence for existing relationships and future trends (Lai et al. 2016, Batty, 2007).

Concluding, this systematic review and meta-analysis, provide numerical proofs of high performance of ANNs in an urban geography context, especially when used for urban growth modeling and urban land use classification/ change. Overall findings along with the suggested guidelines for reporting results, will hopefully assist other practitioners and provide research directions regarding the implementation of ANNs and deep learning by a systematic way in urban geography.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compenvurbsys.2018.10.008.

Acknowledgements

This work was supported by Sun Yat Sen University Starting Research Grant for George Grekousis (37000-18821113).

References

Batty, M. (2007). *Cities and complexity: Understanding cities with cellular automata, agent-based models, and fractals.* The MIT press.
Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. *Neural networks: Tricks of the trade* (pp. 437–478). Berlin Heidelberg: Springer.
Bhatti, S. S., Tripathi, N. K., Nitivattananon, V., Rana, I. A., & Mozumder, C. (2015). A multi-scale modeling approach for simulating urbanization in a metropolitan region. *Habitat International, 50*, 354–365.
Bloom, D. E., Canning, D., & Fink, G. (2008). Urbanization and the wealth of nations. *Science, 319*(5864), 772–775.
Brownlee, J. (2017). *Master machine learning algorithms.* E-bookhttps://machinelearningmastery.com/master-machine-learning-algorithms/.
Chernick, M. R. (2007). *Bootstrap methods: A guide for practitioners and researchers.* 2007.
Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends® in Signal Processing, 7*(3–4), 197–387.
Du, P., Xia, J., & Feng, L. (2015). Monitoring urban impervious surface area change using China-Brazil Earth Resources Satellites and HJ-1 remote sensing images. *Journal of Applied Remote Sensing, 9*(1), 096094.
Geospatial Media and Communications (GMC) (2017). Global geospatial industry Outlook. http://geospatialmedia.net/global-geospatial-outlook-report-2017-download.html?utm_source=geospatialworld&utm_campaign=jan2017&utm_medium=banner.
Geospatial World Forum (GFM) (2017). Geospatial + Deep Learning: Shaping a smarter world. http://geospatialworldforum.org/forms/gwf-2017-full-report-download.html.
Glorot, X., & Bengio, Y. (2010, May). Understanding the difficulty of training deep feedforward neural networks. *Aistats, 9*, 249–256.
Gong, J., Chen, W., Liu, Y., & Wang, J. (2014). The intensity change of urban development land: Implications for the city master plan of Guangzhou, China. *Land Use Policy, 40*, 91–100.
Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Presshttp://www.deeplearningbook.org.
Gopal, S. (2017). *Artificial Neural Networks in Geospatial Analysis.* The International Encyclopedia of Geography.
Gorricha, J., & Lobo, V. (2012). Improvements on the visualization of clusters in geo-referenced data using Self-Organizing Maps. *Computers & Geosciences, 43*, 177–186.
Grekousis, G., & Gialis, S. (2018). More flexible yet less developed? *Spatio-temporal analysis of labor flexibilization and gross domestic product in crisis-hit European Union regions. Social Indicators Research.* https://doi.org/10.1007/s11205-018-1994-0.
Grekousis, G., Manetos, P., & Photis, Y. N. (2013). Modeling urban evolution using neural networks, fuzzy logic and GIS: The case of the Athens metropolitan area. *Cities, 30*, 193–203.
Grekousis, G., & Mountrakis, G. (2015). Sustainable development under population pressure: Lessons from developed land consumption in the conterminous US. *PLoS One, 10*(3), e0119675.
Grekousis, G., Mountrakis, G., & Kavouras, M. (2016). Linking MODIS-derived forest and cropland land cover 2011 estimations to socioeconomic and environmental indicators for the European Union's 28 countries. *GIScience and Remote Sensing, 53*(1), 122–146.
Grekousis, G., & Photis, Y. N. (2014). Analyzing high-risk emergency areas with GIS and neural networks: The case of Athens, Greece. *The Professional Geographer, 66*(1), 124–137.
Hinton, G. (2010). A practical guide to training restricted Boltzmann machines. *Momentum, 9*(1), 926.
Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82–97.
Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507.

Kauko, T. (2009). The housing market dynamics of two Budapest neighbourhoods. *Housing Studies, 24*(5), 587–610.
Khatami, R., Mountrakis, G., & Stehman, S. V. (2016). A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sensing of Environment, 177*, 89–100.
Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*(1), 59–69.
Kohonen, T. (2001). *Self-Organizing Maps.* Berlin: Springer.
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* (pp. 1097–1105). .
Lai, S. K., & Han, H. (2016). *Urban complexity and planning: Theories and computer simulations.* Routledge.
Längkvist, M., Kiselev, A., Alirezaie, M., & Loutfi, A. (2016). Classification and segmentation of satellite orthoimagery using convolutional neural networks. *Remote Sensing, 8*(4), 329.
Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers, 7*(1), 3–10.
Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.
Lecun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation, 1*(4), 541–551.
Liu, W., Li, X., Chen, Z., Zeng, G., León, T., Liang, J., ... Lai, M. (2015). Land use regression models coupled with meteorology to model spatial and temporal variability of NO 2 and PM 10 in Changsha, China. *Atmospheric Environment, 116*, 272–280.
Luo, Y. Q., Randerson, J., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., ... Friedlingstein, P. (2012). A framework of benchmarking land models. *Biogeosciences Discussions, 9*(2), 3857–3874.
Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics,* 50–60.
Markets and Markets (2016). Deep Learning Market by Application, End-User Industry, and Geography - Global forecasts to 2022. *Report*http://www.marketsandmarkets.com/Market-Reports/deep-learning-market-107369271.html?gclid=CMvU6Jfyr9MCFQoA0wodQswJjg.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097.
Mozumder, C., Tripathi, N. K., & Losiri, C. (2016). Comparing three transition potential models: A case study of built-up transitions in North-East India. *Computers, Environment and Urban Systems, 59*, 38–49.
Openshaw, S., & Openshaw, C. (1997). *Artificial intelligence in geography.* John Wiley & Sons, Inc.
Pacifici, F., Chini, M., & Emery, W. J. (2009). A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment, 113*(6), 1276–1292.
Photis, Y., & Grekousis, G. (2012). Locational planning for emergency management and response: an artificial intelligence approach. *International Journal of Sustainable Development and Planning, 7*(3), 372–384. https://doi.org/10.2495/SDP-V7-N3-372-384.
Pijanowski, B. C., Shellito, B., Pithadia, S., & Alexandridis, K. (2002). Forecasting and assessing the impact of urban sprawl in coastal watersheds along eastern Lake Michigan. *Lakes & Reservoirs: Research & Management, 7*(3), 271–285.
Pijanowski, B. C., Tayyebi, A., Delavar, M. R., & Yazdanpanah, M. J. (2010). Urban expansion simulation using geospatial information system and artificial neural networks. *International Journal of Environmental Research, 3*(4), 493–502.
Pijanowski, B. C., Tayyebi, A., Doucette, J., Pekin, B. K., Braun, D., & Plourde, J. (2014). A big data urban growth simulation at a national scale: Configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment. *Environmental Modelling & Software, 51*, 250–268.
Pu, R., Landry, S., & Yu, Q. (2011). Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing, 32*(12), 3285–3308.
Purdy, M., & Daughtery, P. (2016). Why artificial intelligence are the future of growth. *Report*https://www.accenture.com/t20170206T005353__w__/in-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.PDF#zoom=50.
Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*(6), 386.
Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J. D., Breitkopf, U., & Jung, J. (2014). Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing, 93*, 256–271. https://doi.org/10.1016/j.isprsjprs.2013.10.004.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). Learning internal representations by error propagation (no. ICS-8506). *California Univ San Diego La Jolla Inst for Cognitive Science*.
Salehi, M., Sahebi, M. R., & Maghsoudi, Y. (2014). Improving the accuracy of urban land cover classification using Radarsat-2 PolSAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*(4), 1394–1401.
Seto K, Navin Ramankutty. (2016). Hidden linkages between urbanization and food systems. Science 20 : 943–945.
Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A meta-analysis of global urban land expansion. *PLoS One, 6*(8), e23777.
Shai, S. S., & Shai, B. D. (2014). *Understanding Machine Learning: From Theory to Algorithms.* Cambridge University Press.
Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* (pp. 2951–2959). .
TATA Consultancy Service (2017). Getting smarter by the day: How AI is elevating the performance of global companies. http://sites.tcs.com/artificial-intelligence/.
Tayyebi, A., & Pijanowski, B. C. (2014). Modeling multiple land use changes using ANN, CART and MARS: Comparing tradeoffs in goodness of fit and explanatory power of data mining tools. *International Journal of Applied Earth Observation and Geoinformation, 28*, 102–116.
United Nations, Department of Economic and Social Affairs (2016). Population Division. *The World's Cities in 2016 – Data Booklet (ST/ESA/ SER.A/392)*.
United Nations, Department of Economic and Social Affairs, Population Division (2014). World Urbanization prospects: The 2014 Revision. *Highlights (ST/ESA/SER.A/352)*.
Wang, H., Shao, Y., & Kennedy, L. M. (2014). Temporal generalization of sub-pixel vegetation mapping with multiple machine learning and atmospheric correction algorithms. *International Journal of Remote Sensing, 35*(20), 7118–7135.
Yang, X., Zhao, Y., Chen, R., & Zheng, X. (2016). Simulating land use change by integrating landscape metrics into ANN-CA in a new way. *Frontiers of Earth Science, 10*(2), 245–252.
Yao, H. (2012). Simulating the total ecological footprint of Suzhou from 1990 to 2009 by BPANN. *Polish Journal of Environmental Studies,* (6), 21.
Zhang, P., Gong, M., Su, L., Liu, J., & Li, Z. (2016). Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing, 116*, 24–41.
Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine, 5*(4), 8–36.