

Machine Learning Predictors for Sustainable Urban Planning

Sarojini Devi Nagappan¹, Salwani Mohd Daud²

Advanced Informatics School
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia

Abstract—While essential for economic reasons, rapid urbanization has had many negative impacts on the environment and the social wellbeing of humanity. Heavy traffic, unexpected geohazards are some of the effects of uncontrollable development. This situation points its finger to urban planning and design; there are numerous automation tools to help urban planners assess and forecast, yet unplanned development still occurs, impeding sustainability. Automation tools use machine learning classification models to analyze spatial data and various trend views before planning a new urban development. Although there are many sophisticated tools and massive datasets, big cities with colossal migration still witness traffic jams, pollution, and environmental degradation affecting urban dwellers' quality. This study will analyze the current predictors in urban planning machine learning models and identify the suitable predictors to support sustainable urban planning. A correct set of predictors could improve the efficiency of the urban development classification models and help urban planners to enhance the quality of life in big cities.

Keywords—Urban planning; sustainable development; urban development classification model; machine learning; urban development predictors

I. INTRODUCTION

Growing cities, due to socioeconomic activities, witness the challenges of rapid urbanization much faster. Rapid urbanization, once seen as a positive change, has brought many challenges to sustainable development. Some of the research cited challenges include environmental degradation, social equality, climate issues [1]. The United Nations has projected an increase of 60 per cent increase in urbanization in developing countries, and the situation could lead to unmanaged urban development if not handled wisely [2]. The author in [2] has pointed out urbanization is shaped by spatial and urban planning. However, the rapid growth in the city has brought in much unmannered urban development. The continuous development of facilities to feed the needs of the city population and fuel economic growth has led to challenges from various dimensions, including pollution, health issues, infrastructure congestion and massive traffic.

While much care is taken to ensure the ecosystem during city planning, the increase in geohazards such as flash floods and unbearable climate change is alarming [3-5]. Unmanageable city development has also led to overcrowding cities with uncontrollable traffic conditions and disoriented land development.

The increase of people and vehicles on the road increases the traffic index daily in cities, contributing to accidents and health issues such as anxiety and pollution. An improved sustainable urban planning methodology can control uncontrollable development and minimize its negative impact.

Although numerous urban planning support systems and various large amounts of data, unplanned development still happens in many major cities. The author in [6] have pointed out that urban planners can better support city planning by analyzing a city's density and compactness. The algorithmic analysis will help urban planners better study the city development patterns to help the decision-making process for the further development of a busy town. Some quantifiable metrics mentioned by [6] are used by urban planners for decision making include density, compactness, clustering and connectivity. However, the evidence of these metrics on their efficiency and relationship is still inconclusive.

Discussion on land development often leads to the topic of land use land cover (LULC). Modelling the land use change (LUC) process is essential in urban planning as it enables the urban planners and policymakers to make an informed decision in planning a city's infrastructure [7]. Many urban planning automation tools use machine learning models to classify the current development in a given area in urban planning simulation [8]. Models developed with these machine learning algorithms talks about the correlation between spatial indicators and land use classes but does not expand to the study of estimating LUC for forecasting or prediction [7].

The author in [7] stressed that a good machine learning framework with a suitable prediction model is necessary for urban planning. Many scholars study the classification of the attributes on the land cover using ML techniques to classify the characteristics on lands such as forest, water, buildings, and the intensity of land use change based on human activities for a given area. Recent research by [9] has indicated remote sensing data can provide information on natural attributes of ground components, but it does not show information of socio-economic features caused by human activities, which is contributing to challenges in urban land use classification.

The classification machine learning algorithms require the right set of features to classify urban development accurately. An excellent urban planning model allows the urban planners and policymakers to decide to plan a city's infrastructure [7].

The main objective of this paper is to explore machine learning urban planning classification models and propose new urban predictor categories to improve sustainable urban development. Section 2 of this paper will discuss the concepts in three domains, namely sustainability, urban planning and machine learning. Section 3 outlines the methodology used for this review. Section 4 presents the finding of the review coupled with the discussion of the findings.

II. PROBLEM BACKGROUND

Environment degradation, poor health and spread of disease, increasing urban crime are some of the challenges faced with rapid urbanization. Urbanization does equally provide a good transformation for the economic growth and citizens wellbeing; however, with the large scale migrating to the busy cities, the positive impact is now slowly fading. The author in [3] has highlighted adverse effects to environments are due to urban activities. An increase in urban population also speeds up the expansion for physical built up which exacerbate the warming climate of city dwellers [4]. The author in [10] also pointed out the fact that rapid urbanization causes problems to climate change, social inequalities and land scarcity. Sustainable development goals spearheaded by the United Nations [2] outlines the required services or necessary resources to support the wellbeing of a certain percentage of the population. Sustainable development aims to achieve these goals via various activities, and one of them is sustainable urban planning.

Sustainable urban planning is related to the physical and spatial planning to optimize the distribution of land allocation to support human activities [11]. In an urban context, this implies creating both resource-efficient systems and good, engaging urban design for attractive cities with good quality of life [12].

The recent paradigm shift on urbanization has also awakened the realization of the need for sustainable urban planning to support sustainability goals. Land development along with land transactions in the cities is a forced change by urbanization. Nevertheless, when global migration to significant cities happens rapidly, this change is viewed as a positive relationship between land development and land-based revenue growth [13]. Due to this, policymakers and urban planners are often bounded by the economic growth indicator to decide on further development. The development of residential, commercial and other amenities is continuously increased as the need of city dwellers increase. As a result, the infrastructure eco system becomes unstable, causing traffic congestion, climate change, urban crimes and many other negative effects.

These recent advancements have exposed the need for urban planners to support sustainable development is now increasingly evident. They hold the key responsibility to decide on the need and impact of further development in a city with rapid urbanization. Fortunately, with the rise of smart cities and big data, there is an explosion of different types of data from various sources for the urban planner to analyze and predict the need and impact of further development in a compact city. With the advancement in the planning support systems, big data and artificial intelligence (AI), urban planners are now

able to use data and urban models not only to evaluate the need but to predict the consequences of rapid development. Although there are many areas of focus in urban planning supporting sustainable planning, there is an increasing research interest in the study of land use land cover (LULC). Research by [14] mentions LULC is a major driver to globalization as well as a great support to sustainable development. This research has pointed out how the LULC change analysis employed in prediction models are helpful to monitor LULC to understand the historical events of change on landcover and forecast the impact of future development to mankind. LULC analysis largely depends on remote sensing data, and the accuracy depends on the machine learning model employed to process the data. While there were claims on the unavailability of spatial data previously [15], Google Earth has brought in many data sets on the cloud to make the LULC not only easier but also faster coupled with visualization, making it easier for urban planners to ingest the data [16]. Despite the availability of large datasets, there is little evidence on the classification, simulation or prediction of land use domain that involves human interaction in urban development [17]. This study aims to review existing classification models and predictors used in LULC classification to understand the missing predictors to support sustainable urban planning. Urban development models with the right set of predictors supporting sustainable urban development can help urban planners better decide on future developments and impacts in highly urbanized cities.

III. RELATED WORK

Related works in this study are derived from three different domains, including sustainability, urban planning and machine learning. Table I lists the concepts gathered in the literature review used directly or indirectly in this study.

A. Sustainable Development

Sustainability is initially defined as the mechanism of how natural systems produce and provide to the ecosystem in a balanced way. Sustainability originally defines as the mechanism of how natural systems produce and provide to the ecosystem in a balanced way. This concept has been evolving since its conceptualization in 1946 by Sir John Hicks [18]. According to [18] there is a high correlation between human activities and the stability of the natural ecosystem. Human activities influence on the natural ecosystem or environment can be from economic, social and political grounds. Hence to create a sustainable solution, it is crucial to have the three pillars: economic, social and environmental systems to work in harmonization [18].

Brundtland commission popularized the concept of sustainable development (SD) and defined it as the "development that meets the current needs without compromising the ability of future generation needs" [19]. Sustainable development indicators can be divided into economic, social and environmental indicators. There are three pillars of sustainability which is the foundation to drive sustainable development. Sustainable development indicators should be understandable, adaptable and measurable for future development. Indicators provide the means of performance measurement for urban sustainability. It is therefore important

to select the best indicators to achieve the optimum performance of urban sustainability.

This study concentrates on the indicators to achieve sustainable development goal 11(SDG 11), which has the target to develop safe, resilient, and sustainable cities. There are ten targets with 15 indicators to be addressed by SDG 11, and they are strongly related to the urban transformations to sustainability. This includes four focus areas of the SDG 11 indicators, namely, inclusivity, urban safety, resiliency and sustainability, with the following target indicators:

- Safe and Affordable Housing.
- Affordable and sustainable transport systems.
- Inclusive and sustainable urbanization.
- Protect the world's cultural and natural heritage.
- Reduce the adverse effects of natural disasters.
- Reduce the environmental impacts of cities.
- Provide access to safe and inclusive green and public spaces.
- Strong national and regional development planning.
- Implement policies for inclusion, resource efficiency and disaster risk reduction.

This study refers to indicators from the need for improved urban planning and management techniques to develop sustainable cities in SDG 11 and map it to the current features used in urban development model classifications use for LULC monitoring. Once the mapping is done, the missing indicator will be used in the classification to enhance the urban development classification to support sustainable urban planning.

B. Urban Planning Concepts

Urban planning is a cyclic process encompassing many phases and stakeholders' involvement at many intervals and entirely bounded by the legal framework. It is a domain associated with the process of open land development as well as redesigning the unutilized space effectively to promote sustainable urban development [20]. Urban planning is deemed an important tool to promote sustainable urban development and keep the authorities informed on infrastructure development, land investments and demand, or urban population growth [21].

Urban planning tools employ land use land cover (LULC) models to study the change analysis in the landscape. This change analysis provides an insight into historical land use patterns as a base to forecast future developments [22]. Land cover defines the area covered by forest, wetlands, impervious surfaces, agriculture and other land and water types on earth surfaces and land use describe how the land has been used by socioeconomic activities to change the natural landscape into urban built land, cultivated or recreational areas [23]. LULC change analysis uses remote sensing geographical information system (GIS) data to simulate and forecast changes in many areas like deforestation, climate change and population

movement [23]. LULC monitoring plays an important role in urban planning [24]. The change analysis allows the urban planner to take an informed decision for future development for a designated area. LULC change analysis is also used in understanding the effects of urbanization, ecosystem degradation, the impact of pollution on the quality of life [23, 25, 26]. Urban development models use machine learning techniques like SVM, Random Forrest (RF) are widely used to study the change analysis LULC. These models use spatial data to classify LULC and are used by urban planners to evaluate the impact on domains like climate, pollution, traffic via GIS tools. To minimize the impact of rapid urban development, it is important to have the right set of features included in the machine learning models. The right set of features will classify LULC at a higher accuracy to give urban planners a better understanding of the study area before moving forward with a development decision.

Land use land cover monitoring plays a vital role in sustainable development planning [27] because it shows the changes to land cover historically. Revisiting the urban planning decision and management policies is crucial to achieving sustainable urban development. With an improved LULC classification, it would be easier to make informed decisions to drive sustainable development for a better future generation.

Many modelling techniques have been employed to measure the land use land cover change (LULCC). It provides the continuous assessment of the current situation as well as provides the capability to predict future land use change [28]. The author in [29] argue that the performance of the individual indicators or predictors plays a significant role in modelling the land use change (LUC) dynamics. Hence the ability to model LULCC and simulate LUC is firmly based on the predictors used in the urban simulation model.

C. Urban Planning Predictors

Urban development or urban expansion are terms used in urban planning to denote the development of a given land area. At present, the evaluation of land use change patterns has gained research popularity again. It was evident that environmental degradation was due to uncontrollable human activities on the land surface. Hence urban simulation using geospatial data is now popular again. Machine learning is currently being employed to achieve better classification and improve and improve the accuracy of existing machine learning models. The machine learning models' performance depends on the features used to measure the land use change simulation. In this study, predictors are referred to as the features used in the urban simulation model to simulate and predict a land use change. Simulating urban development patterns using these predictors provides an early assessment of urban planning before its implementation [30]. The role and selection of predictors are solely dependent on using the model in the defined area.

D. Machine Learning in Urban Planning

Many researchers have utilized machine learning in urban planning and design to develop and monitor city planning in recent years. Urban analytics and urban simulations employ machine learning is used for predictive analysis for a given

domain. Several techniques and algorithms have been employed in simulation to enable urban planners to improve decision-making for strategic planning. Recent research by [28] stressed the importance of high prediction accuracy to monitor land use change. The same study has also pointed out the need for a suitable model for different applications to ensure accuracy and fitness.

Urban planner uses GIS tools to collect, manage, analyze and visualize geospatial data [30]. Commonly GIS tools use a simulation model to evaluate, predict and analyze complex urban interactions and processors [31]. These urban development models can also identify the urban development trends and their impact with various views to help urban planners understand better the area they are working on. The efficiency of urban planning would depend on the predictors used in the urban classification model. A good predictor enables the urban development classification model to analyze the current geospatial data pattern and simulate future sustainability growth.

TABLE I. KEY CONCEPTS

| Concept | Definition |
|--|--|
| Sustainable Development | Sustainable development (SD) is the development that meets the current needs without compromising the ability of future generation needs.” |
| Urban Planning | Urban planning is a cyclic process encompassing many phases and stakeholders' involvement at many intervals and entirely bounded by the legal framework. |
| Urban Planning Predictors | Urban planning predictors are referred to as the features used in the urban simulation model to simulate and predict a land use change. |
| Urban development classification model | An urban development classification model would analyze the current pattern of geospatial data and simulate future growth with a good set of predictors |

IV. LITERATURE REVIEW METHODOLOGY

A literature review study was carried out to have a concise understanding of concepts and capture the relevant literature on sustainable development and machine learning predictors in urban planning. The purpose of the review is to understand what and how sustainable development goals play a role in urban planning and how the indicators can be used to define additional features or predictors for a better urban development classification model.

This study uses the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) to validate the relevancy of the literature’s collated before reporting the conclusion. PRISMA advocates four phases for literature review, starting from identification of the record source, screening of the records, setting the eligibility criteria and finally compiling the records to be included for the identified research [32]. Fig. 1 illustrates the PRISMA flowchart adopted for this study.

In the identification phase, the source of the records is determined to avoid the irrelevant search. For this study, literature will be extracted from five databases, including IEEE, Scopus, Springer, Science Direct and Web of Science.

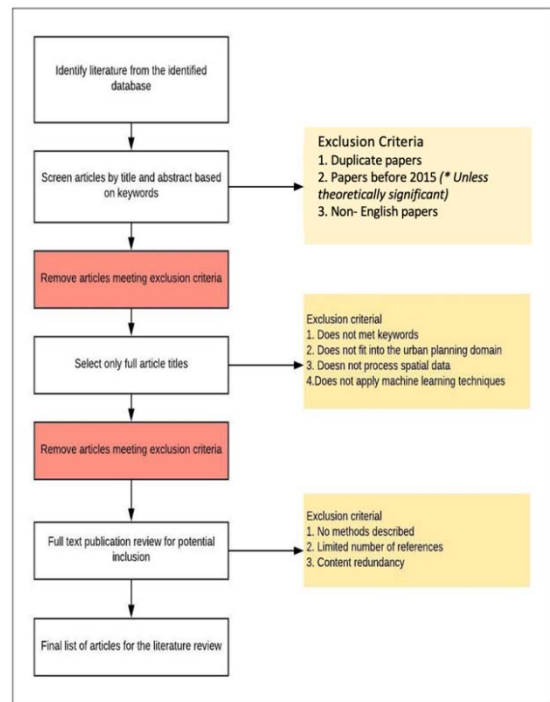


Fig. 1. PRISMA Flow Chart used for Literature Review.

The next process is screening the records. Records are screen based on the keywords found from the first level of record retrieval. The first screening exclusion criteria to be applied is as follows:

- Duplicate papers. As there are many sources of the database the records will be retrieved from, there are high chances of duplicate records.
- Papers which are before the year of 2015 will not be considered unless it defines a standard terminology or significantly prominent for the research. Also, for the purpose of this study.
- Non-English papers are to be removed.

In the first phase of screening, the articles will be screened by title and abstract based on the list of keywords. The list of keywords is listed in Table II.

The second round of screening only takes into account full text records. Records with only abstracts are not included in the screening. The third level of screening is based on the actual content. The content of the articles must be complete with the methods without content redundancy. Records with below ten references are omitted. The screening process includes setting the inclusion and exclusion criteria that have been listed earlier. The last stage of the flow is to compile the final list of records to be included in the final literature review.

The review was done to

- Gather the sustainable development goal indicators for urban planning.
- Analyze the concepts in urban planning and evaluate the human activities on land.

- Review the features used by different urban classification models.

TABLE II. KEYWORD USED IN THE LITERATURE SEARCH STRINGS

| Domain | Keywords |
|-------------------------|--|
| Sustainable Development | <ul style="list-style-type: none"> • sustainability • sustainable development • sustainable development goals • sustainable urban planning • sustainable urban planning indicators |
| Urban Planning | <ul style="list-style-type: none"> • Urban planning • Urban planning classification model • Urban planning indicators • Urban planning simulation • Urban planning assessment • Urban planning measurement metrics • Urban planning density |
| Machine Learning | <ul style="list-style-type: none"> • Machine learning models in urban planning • Machine learning features in urban planning • Machine learning algorithm performance • Machine learning accuracy • LULC classification |

V. FINDINGS AND DISCUSSION

The following six urban models have been analyzed from scholarly reviews to understand the predictors or indicators used to simulate urban expansion and tabled out in Table III. These models were actively used to simulate land use changes since 2011 cited in a review done by [30].

- Cellular Automata.
- Regression Models.
- Artificial Neural Networks.
- Agent Based Models (ABM).
- Tree Based Models.
- Support Vector Machines (SVM).

Some of the reviews have detailed out the criteria, like the indicators used in the SVM model. This gives an indication to know the popularity of the model employed in this field. Another interesting finding is the gaining interest in the Agent-Based Model (ABM) model to which is widely used in land-use change and transport line models. This model is gaining interest as it considers three main stakeholders in the urban planning process. However, this model has its limitation because its dependency on agents' predictors is highly variable.

TABLE III. URBAN DEVELOPMENT MODELS AND ITS PREDICTORS

| Urban Development Model | Predictor | Reference |
|------------------------------------|--|---|
| Cellular Automata model (CA) | proximation roads, expressways, railways and town centers, land use types and topography and population density | [33, 34] |
| Artificial neural networks (ANN) | elevation, slope, annual population growth rate, land use types and distance | [35, 36] |
| Linear Regression (LR) Predictors | easting and northing coordinates, land use types, slope, restricted areas, population density, distances to main active economic centres, a central business district, roads and urbanized area | [37, 38] |
| Agent Based Model (ABM) Predictors | <p>Dwellers -accessibility to education centers, favourite elevation, favourite slope, accessibility to health services, accessibility to metro stations, distance from disposal area, accessibility to fruit garden accessibility to sports centers, accessibility to the road network, accessibility to recreation areas, accessibility to business districts</p> <p>distance to railway</p> <p>Developers- residential density,employment density,commercial intensity/density,investment profit minimum unseen profit maximum profit</p> <p>Government river streams risk zone, roads network buffer, highways buffer, airports risk buffer, military facility risk zone, power facility risk zone, parks buffer on a suitable slope</p> | <p>[39]</p> <p>[40]</p> <p>[36]</p> <p>[41]</p> <p>[42]</p> |
| Decision Tree Model(DT) | development history, elevation details, land type, proximities to neighbourhood amenities and densities of residential, commercial | [43] |
| Support Vector Machines (SVM) | land-use types land use densities proximities to: roads, expressways, railways and town centres, as well as land-use types and topography neighbouring land characteristics built area unbuilt area | <p>[30]</p> <p>[44]</p> <p>[45]</p> <p>[46]</p> <p>[40]</p> |

The land-use change patterns, which are the main urban development metrics, cannot be just determined by only one factor. The relationship between various factors provides a better evaluation of urban development. The author in [36] explain the urban growth and development indicators differs based on the model and data used to evaluate or predict the urban pattern change.

A. Urban Development Models Predictor Evaluation

According to [30], predictors in urban development modelling can be further categorized as site specific proximity and neighbourhood characteristics. Site specific characteristics describe built and unbuilt areas as well as commercial and residential buildings. Proximity predictors refer to the distance of access to a specific facility, like the distance to roads, highways, urban built areas, green space and water bodies. The neighborhood characteristics refer to predictors which will affect LULC in future including the number of cells corresponding to wetlands, forest, barren lands, developed areas. Table IV below shows a summary mapping between the existing urban land indicator categories by [47] and predictor categories by [30]. If the predictors match the indicators, it is indicated as “complies”.

Table V shows that the urban development models are commonly used by tools or applications to model urban growth. All the models have used site-specific and proximity predictors. Remote sensing data gives a good source for site-specific and proximity indicators, which is one reason why it's a commonly used predictor category in all models. The neighborhood category defines its attribute in the number of cells usually found on maps. This measurement is very suitable for fractal development models as it works best with maps. In a fractal model, structures are produced by iterations of the same

principle in a given pattern of points, lines or surfaces. The model's accuracy depends on the model. The last two categories are population, and crowd mobility is the least used predictor category in the urban development models analyzed in this study. Mobility of the city dwellers and the population indicators are not used in the urban expansion models. This information is not easily made available in a single source data source like remote sensing data. Research by [9] has pointed out that although including socioeconomic data would improve the accuracy of classification accuracy, there is very little evidence to show its usage. However, due to rapid urbanization, the increase of traffic and environmental hazards warrants the inclusion of these predictors in the urban development model. research has used these data to improve accuracy. Agent based models are now considering factoring these predictors as it may improve the transit-oriented model.

TABLE IV. URBAN DEVELOPMENT MODEL PREDICTOR CATEGORIES COMPLIANCE TO LAND USE

| Land Use Predictor Categories | Urban Development Model Predictor Categories | | |
|-------------------------------|--|-------------|---------------|
| | Site Specific | Proximities | Neighbourhood |
| Neighbourhood | - | - | - |
| Employment | - | - | - |
| Infrastructure | Complies | Complies | Complies |
| Access to Basic Facilities | - | Complies | Complies |
| Legislation | - | - | - |
| Environment | Complies | - | - |

TABLE V. PREDICTORS CATEGORIES USED IN SELECTED URBAN DEVELOPMENT MODELS

| Urban Development Models | Urban Development Model Predictor Derived from Research Article (1986 – 2018) | | | | | |
|-----------------------------|---|---------------|-------------|---------------|--------------------|----------------|
| | Application | Site Specific | Proximities | Neighbourhood | Population Density | Crowd Mobility |
| Cellular Automata | Large Scale Urban Simulation | √ | √ | | | |
| Artificial Neural Network | Modelling of Urban Growth | √ | √ | | √ | |
| Fractals | Urban Growth Experience | √ | √ | √ | | |
| Linear/ Logistic Regression | Prediction of urban growth pattern | √ | √ | | √ | |
| Agent Based Model | Modelling and simulation of land use and land cover dynamics | √ | √ | | √ | |
| Decision Tree | Prediction of Urban Growth | √ | √ | | | |
| Support Vector Machine | Prediction of Urban Expansion | √ | √ | √ | | |

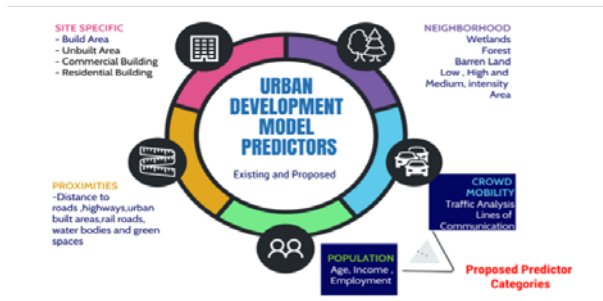


Fig. 2. Urban Development Model Predictors: Existing and Proposed.

The above finding concludes that urban development models have used various predictors to simulate urban expansion. However, the current extensive list of predictors for these research resources does not include population and crowd mobility as key predictor categories in the urban development models. Fig. 2 shows the existing and new categories along with their predictors for the urban development model.

Rapidly urbanizing countries often show rapid land usage; however, population factors are not explicitly included in the urban models [48]. Urbanization warrants more urban built up in dense cities and converts prime land to support the needs of the growing population [17]. This scenario takes place because the spatial analysis of the built-up area is limited to building-like structures. Urban development models do not capture the external factors such as traffic, population maturity in LULC classification. Traffic congestion is a cycle effect, a higher population growth rate in an area warrants more development, and more development increases traffic. However, urban development models do not capture this effect, limiting the LULC analysis for the urban planners. Sustainable urban design framework includes indicators for mobility and transport indicators are for a better environment and economy. The author in [49] highlighted how mobility and transport indicators are key for health and the environment in a sustainable city, supporting the need to have these predictors to assess urban planning. The traffic layer or properties and urban growth included in the urban development model can improve urban-built in LULC monitoring. Instead of classifying the urban built-up based on infrastructure properties, the inclusion of vehicular information for a period of time gives the urban planners insight into how dense the area is both in traffic and built-up area.

A sustainable city development should not forget the social aspects to improve quality of life. Hence, this study proposes that population and crowd mobility predictors should be included to improve the urban development classification model.

VI. PROPOSED FUTURE WORK

Developing geospatial classification models using Google Earth Engine has been the recent research trends [16, 50-52] as it allows to process large sets earth imagery data without the need to have high end software or hardware requirements. The availability historical data sets and the ability to preprocess the data in Google Earth Engine allows fast simulation of LULC change analysis therefore developing enhanced application with improved feature is faster using the web interface.

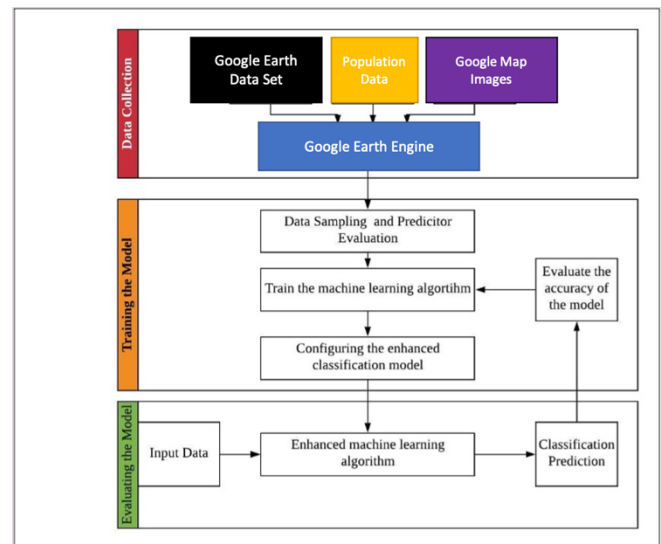


Fig. 3. Proposed Enhanced Urban Development Classification Model Development Framework.

An enhanced urban development classification model will be developed using the development methodology in Fig. 3. The proposed development framework will derive its dataset for Google Earth and use the Google Earth Engine as a platform to develop and train the classified model. The enhanced urban development classification model will evaluate the impact of the newly identified indicator in this study on the LULC classification. Specifically, the classification model will use the existing features from the land use land cover properties, traffic or mobility data, and urban demographic data to increase the accuracy of LULC classification, particularly the urban built classification. Similar research employing the traffic layer classified as spatial static temporal dynamic data for intelligent transportation [53] will be closely adopted in the urban classification model development framework.

VII. CONCLUSION

Effects of rapid urbanization are simply inevitable, and however, with appropriate measures, it can be controlled. The recent paradigm shift on urbanization has also awakened the need for sustainable urban planning to support sustainability goals. This study has pointed out the shortfall in urban development models, which could be the factor for uncontrolled urban planning with the exponential population growth. The shortfall of the urban development models include:

- Urban development models do not capture traffic intensity and population attributes in land use land cover monitoring.
- Urban development models use predictors available only directly from remote sensing data and do not combine data from different sources, limiting the analysis.
- The urban development model classifies LULC, but its accuracy highly depends on the data available in the area of interest. Thus, the urban planning forecast relies

very much on what type of classification algorithm is used in the model.

- Urban development classifies the urban built-up land as residential, commercial and services, industrial, communication and utilities, mixed urban or built-up land but does not include vehicular properties as part of the classification predictor.

This study proposes two new predictor categories: population and crowd mobility, also known as socioecological indicators, to enhance the urban development classification model further. Traffic analysis will provide the crowd mobility predictors for urban development models to simulate the optimum crowd movement, land change, and population growth. Many urban development models concentrate on classifying the attributes on the surface land to affect future urban expansion; however, based on the discussions earlier on sustainable development, it is pointless only to analyze the land cover changes. LUC is beneficial from economics as it promotes industrial development; however, to achieve urban sustainability, the models should consider the social, economic indicator as it affects the human wellbeing and the stability of the environment. Moving forward, data for this predictor will be gathered to assess the accuracy of the urban development classification model.

ACKNOWLEDGMENT

The authors would like to express greatest appreciation to Universiti Teknologi Malaysia (UTM) and Ministry of Higher Education (MOHE), Malaysia for the financial support (grant number: Q.K130000.2456.08G28). Last but not least, we also appreciate Advanced Informatics at Razak Faculty of Technology and Informatics (FTIR), UTM for their support in conducting this research.

REFERENCES

- [1] K. Anarfi, R. A. Hill, and C. Shiel, "Highlighting the sustainability implications of urbanisation: A comparative analysis of two urban areas in ghana," *Land*, vol. 9, no. 9, 2020.
- [2] D. o. E. a. S. A. United Nations, Population Division "World urbanization prospects: The 2018 revision ", New York, 2019 2019. [Online]. Available: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf>.
- [3] A. Mersal, "Sustainable urban futures: Environmental planning for sustainable urban development," *Procedia Environmental Sciences*, vol. 34, pp. 49-61, 2016.
- [4] Q. Cao, D. Yu, M. Georgescu, J. Wu, and W. Wang, "Impacts of future urban expansion on summer climate and heat-related human health in eastern china," *Environ Int*, vol. 112, pp. 134-146, Mar 2018.
- [5] S. H. Mahmoud and T. Y. Gan, "Long-term impact of rapid urbanization on urban climate and human thermal comfort in hot-arid environment," *Building and Environment*, vol. 142, pp. 83-100, 2018.
- [6] J. Ogle, D. Delparte, and H. Sanger, "Quantifying the sustainability of urban growth and form through time: An algorithmic analysis of a city's development," *Applied Geography*, vol. 88, pp. 1-14, 2017.
- [7] M. Samardžić-Petrović, M. Kovačević, B. Bajat, and S. Dragičević, "Machine learning techniques for modelling short term land-use change," *ISPRS International Journal of Geo-Information*, vol. 6, no. 12, 2017.
- [8] M. Batty, "Big data and the city," *Built Environment*, vol. 42, no. 3, pp. 321-337, 2016.
- [9] W. Mao, D. Lu, L. Hou, X. Liu, and W. Yue, "Comparison of machine-learning methods for urban land-use mapping in hangzhou city, china," *Remote Sensing*, vol. 12, no. 17, 2020.
- [10] Y. Tan, H. Xu, and X. Zhang, "Sustainable urbanization in china: A comprehensive literature review," *Cities*, vol. 55, pp. 82-93, 2016.
- [11] D. Geneletti, D. La Rosa, M. Spyra, and C. Cortinovis, "A review of approaches and challenges for sustainable planning in urban peripheries," *Landscape and Urban Planning*, vol. 165, pp. 231-243, 2017.
- [12] C. Haaland and C. K. van den Bosch, "Challenges and strategies for urban green-space planning in cities undergoing densification: A review," *Urban Forestry & Urban Greening*, vol. 14, no. 4, pp. 760-771, 2015.
- [13] J. Chen, F. Guo, H. Wang, Z. Wang, and Y. Wu, "Urban land revenue and sustainable urbanization in china: Issues and challenges," *Sustainability*, vol. 10, no. 7, 2018.
- [14] C. Liping, S. Yujun, and S. Saeed, "Monitoring and predicting land use and land cover changes using remote sensing and gis techniques-a case study of a hilly area, jiangle, china," *PLoS One*, vol. 13, no. 7, p. e0200493, 2018.
- [15] D. Li, S. Wang, and D. Li, *Spatialdatamining theory and application*. SpringerNature, 2015.
- [16] R. Goldblatt, K. Deininger, and G. Hanson, "Utilizing publicly available satellite data for urban research: Mapping built-up land cover and land use in ho chi minh city, vietnam," *Development Engineering*, vol. 3, pp. 83-99, 2018/01/01/ 2018.
- [17] J. Gao and B. C. O'Neill, "Data-driven spatial modeling of global long-term urban land development: The select model," *Environmental Modelling & Software*, vol. 119, pp. 458-471, 2019.
- [18] N. R. Khalili, *Theory and concept of sustainability and sustainable development*. Palgrave Macmillan, New York, 2011.
- [19] B. Purvis, Y. Mao, and D. Robinson, "Three pillars of sustainability: In search of conceptual origins," *Sustainability Science*, vol. 14, no. 3, pp. 681-695, 2018.
- [20] R. Wennersten, "Development of new sustainable urban areas: Horizontal or vertical planning systems for resource efficient cities," in *An overview of urban and regional planning*, 2018, ch. Chapter 6.
- [21] O. M. Dede, "The analysis of turkish urban planning process regarding sustainable urban development," in *Sustainable urbanization*, 2016, ch. Chapter 12.
- [22] N. Hashem and P. Balakrishnan, "Change analysis of land use/land cover and modelling urban growth in greater doha, qatar," *Annals of GIS*, vol. 21, no. 3, pp. 233-247, 2014.
- [23] M. Mohamed, J. Anders, and C. Schneider, "Monitoring of changes in land use/land cover in syria from 2010 to 2018 using multitemporal landsat imagery and gis," *Land*, vol. 9, no. 7, 2020.
- [24] C. B. Karakuş, "The impact of land use/land cover (lulc) changes on land surface temperature in sivas city center and its surroundings and assessment of urban heat island," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 55, no. 4, pp. 669-684, 2019/11/01 2019.
- [25] H. Bagan and Y. Yamagata, "Land-cover change analysis in 50 global cities by using a combination of landsat data and analysis of grid cells," *Environmental Research Letters*, Article vol. 9, no. 6, 2014, Art no. 064015.
- [26] A. Navara and R. Vedamuthu, "Ecosystem services-based approach to sustainable development in a peri-urban area of chennai, india," *Environment, Development and Sustainability*, 2021/06/09 2021.
- [27] F. Ahmad, L. Goparaju, and A. Qayum, "Lulc analysis of urban spaces using markov chain predictive model at ranchi in india," *Spatial Information Research*, vol. 25, no. 3, pp. 351-359, 2017.
- [28] M. Ahmadlou, M. R. Delavar, A. Basiri, and M. Karimi, "A comparative study of machine learning techniques to simulate land use changes," *Journal of the Indian Society of Remote Sensing*, vol. 47, no. 1, pp. 53-62, 2018.
- [29] M. C. Cheruto, M. K. Kauti, P. D. Kisangau, and P. Kariuki, "Assessment of land use and land cover change using gis and remote sensing techniques: A case study of makueni county, kenya " *Assessment of Land Use and Land Cover Change Using GIS and Remote Sensing Techniques: A Case Study of Makueni County, Kenya*, 2016.

- [30] F. Karimi, S. Sultana, A. Shirzadi Babakan, and S. Suthaharan, "An enhanced support vector machine model for urban expansion prediction," *Computers, Environment and Urban Systems*, vol. 75, pp. 61-75, 2019.
- [31] D. Fecht, L. Beale, and D. Briggs, "A gis-based urban simulation model for environmental health analysis," *Environmental Modelling & Software*, vol. 58, pp. 1-11, 2014.
- [32] C. Pizzi et al., "Prevalence of cardiovascular medication on secondary prevention after myocardial infarction in china between 1995-2015: A systematic review and meta-analysis," *PLoS One*, vol. 12, no. 4, 2017.
- [33] S. Deep and A. Saklani, "Urban sprawl modeling using cellular automata," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 17, no. 2, pp. 179-187, 2014/12/01/ 2014.
- [34] E. d. N. Vaz, P. Nijkamp, M. Painho, and M. Caetano, "A multi-scenario forecast of urban change: A study on urban growth in the algarve," *Landscape and Urban Planning*, vol. 104, no. 2, pp. 201-211, 2012.
- [35] S. Mohammady, M. R. Delavar, and P. Pahlavani, "Urban growth modeling using an artificial neural network a case study of sanandaj city, iran," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-2/W3, pp. 203-208, 2014.
- [36] S. I. Musa, M. Hashim, and M. N. M. Reba, "A review of geospatial-based urban growth models and modelling initiatives," *Geocarto International*, vol. 32, no. 8, pp. 813-833, 2016.
- [37] A. A. A. Alsharif and B. Pradhan, "Urban sprawl analysis of tripoli metropolitan city (libya) using remote sensing data and multivariate logistic regression model," *Journal of the Indian Society of Remote Sensing*, vol. 42, no. 1, pp. 149-163, 2013.
- [38] P. Tripathy, A. Bandopadhyay, R. Raman, and S. K. Singh, "Urban growth modeling using logistic regression and geo-informatics: A case of jaipur, india," *Environment & We An International Journal of Science & Technology*, vol. 13, 2018.
- [39] H. Motieyan and M. S. Mesgari, "An agent-based modeling approach for sustainable urban planning from land use and public transit perspectives," *Cities*, vol. 81, pp. 91-100, 2018.
- [40] A. Shirzadi Babakan and M. Taleai, "Impacts of transport development on residence choice of renter households: An agent-based evaluation," *Habitat International*, vol. 49, pp. 275-285, 2015.
- [41] Q. Zhang, R. R. Vatsavai, A. Shashidharan, and D. Van Berkel, "Agent based urban growth modeling framework on apache spark," in *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, BigSpatial 2016*, 2016, pp. 50-59.
- [42] C. Wray and K. Cheruiyot, "Key challenges and potential urban modelling opportunities in south africa, with specific reference to the gauteng city-region," *South African Journal of Geomatics*, vol. 4, no. 1, 2015.
- [43] H. Jin and G. Mountrakis, "Integration of urban growth modelling products with image-based urban change analysis," *International Journal of Remote Sensing*, vol. 34, no. 15, 2013.
- [44] H. Kim, S. Sultana, and J. Weber, "A geographic assessment of the economic development impact of korean high-speed rail stations," *Transport Policy*, vol. 66, pp. 127-137, 2018.
- [45] M. Samardžić-Petrović, S. Dragičević, M. Kovačević, and B. Bajat, "Modeling urban land use changes using support vector machines," *Transactions in GIS, Article* vol. 20, no. 5, pp. 718-734, 2016.
- [46] Y. Deng and S. Srinivasan, "Urban land use change and regional access: A case study in beijing, china," *Habitat International*, vol. 51, pp. 103-113, 2016.
- [47] F. Samara, S. Tampekis, S. Sakellariou, and O. Christopoulou, "Sustainable indicators for land use planning evaluation: The case of a small greek island," presented at the International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE), Mykonos, Greece, 2013.
- [48] J. Gao and B. C. O'Neill, "Mapping global urban land for the 21st century with data-driven simulations and shared socioeconomic pathways," *Nat Commun*, vol. 11, no. 1, p. 2302, May 8 2020.
- [49] A. López Chao, A. Casares Gallego, V. Lopez-Chao, and A. Alvarelos, "Indicators framework for sustainable urban design," *Atmosphere*, vol. 11, no. 11, 2020.
- [50] T. N. Phan, V. Kuch, and L. W. Lehnert, "Land cover classification using google earth engine and random forest classifier—the role of image composition," *Remote Sensing*, vol. 12, no. 15, 2020.
- [51] P. Arévalo, E. L. Bullock, C. E. Woodcock, and P. Olofsson, "A suite of tools for continuous land change monitoring in google earth engine," *Frontiers in Climate*, vol. 2, 2020.
- [52] N. A. Wahap and H. Z. M. Shafri, "Utilization of google earth engine (gee) for land cover monitoring over klang valley, malaysia," *IOP Conference Series: Earth and Environmental Science*, vol. 540, p. 012003, 2020/08/05 2020.
- [53] H. Yuan and G. Li, "A survey of traffic prediction: From spatio-temporal data to intelligent transportation," *Data Science and Engineering*, vol. 6, no. 1, pp. 63-85, 2021