Unwilling Urban Growth and Demographic Trends Through Supervised Classification Techniques

Muzammil Aktar, Umer Farooq, Iqra Ramzan, Khizar Ahmed, Syeda Minha Tirmizi

Abstract

Karachi is the capital of Sindh and supports the country with 65% of its revenue, it is one of the world's megacities with a population of more than 20 million and an inferior management style that has caused its sustainability to be at risk. Two sustainable indicators, urban sprawl and demography are used to analyze the present conditions in the city.

In this study, we used a change detection technique to access the urban sprawl through different Landsat images from the years 2000 to 2024, With the help of Google Earth different land features of Karachi divisions are digitized and classified based on their land uses and housing units. Karachi has extended its built-up area by 15% in the last 24 years. Its urban expansion has expanded at the rate of about 2.2% per year to accommodate its increasing population, which is growing with a more than 4% annual population growth rate.

Even though there has been a serious problem with the law and order situation, the population of Karachi has expanded extensively for not only natural reasons but also because of the heavy immigration. This study provides a thorough understanding for the planner to determine a better future growth and development plan for a leading sustainable Karachi.

Introduction

Karachi has a population of 20.3 million according to the 2023 census, it is the largest city within Pakistan and it is the economic core of Pakistan. It is located on the coastline of Sindh province in southern Pakistan with the coordinates 24.8607° N, and 67.0011° E, and is most often cited as a strong international trade centre. Karachi is said to be a hub of activity, blending many different traditions, languages, and cultures. The growth of urban centres in an area has both benefits and drawbacks. The Indian Ocean made it easier for trade to grow, making it not just a shipping port but also a trading port.

Karachi seems to have undergone rapid growth in both numbers and size, primarily due to those from other regions moving into the area, rural-urban migration was said to be a factor to consider as well. The economy is fast attracting migrants, causing Karachi's population to grow wildly, further aiding the expansion of the urban area. However, it is important to note that a negative side of migration is that rural areas tend to be depopulated leading to generations of informal settlements. In most cases these areas are known for poor sanitary conditions, resulting from infrastructure degradation, shortage of green labs, and environmental disruption.

To tackle these urbanisation challenges, supervised classification, a remote sensing technique, offers critical insights into land use and land cover changes. This approach classifies urban areas, vegetation, and water bodies such as lakes and rivers according to a set of classes guided by satellite images. This approach facilitates the assessment of the spatial extent of urban growth, vegetation encroachment, and the change in the earth's surface.

Supervised classification has been applied to Karachi to:

- Determine the geographical extent of urbanization and its trends over time.
- Provide indicators of problems such as shrinkage of vegetation cover and proliferation of squatter settlements.
- Encourage management of cities and resources through data-based decisions.

This methodology was presented at the NASA Space Apps Challenge where a group of participants attempted to integrate the perspective of people with remote sensing of Karachi's urban growth. The project highlighted the role of spatial information in exploring urban issues by incorporating remote sensing and GIS within the same framework and showed the potential for addressing urbanisation pressures in large cities like Karachi through a creative model.

Urbanisation in Karachi provides massive opportunities for economic development but also marks a necessity to balance this progress with preservation. Supervised classification emerges as a priceless approach toward understanding such trends; it yields actionable insights, guiding policymakers toward appropriate urban management and planning strategies.

Study Area

Karachi is the capital of the province of Sindh and the largest city of Pakistan. Karachi, the world's first megacity to have a population of more than 100 million, got this status in 2000. In 2000, it ranked 17th in the world; in 2005, it reached 13th position and was expected to be in 12th position by 2015 (Biswas, 2006; Afsar et al., 2013). It was the first federal capital of Pakistan from 1947 to 1958. Administratively, Karachi is governed by the City District Government, covering 3,527 square kilometres with a population of 9,856,318 (Afsar et al., 2013). It is located 129 km west of the Indus mouth, between 24°45'–25°38'N and 66°40'–67°34'E. It is bounded by the Jamshoro and Thatta districts, the Arabian Sea, and the Lasbela district of Balochistan. Karachi is situated on the north coast of the Arabian Sea and has a marine climate. It was a potential port and situated in a perfect location where it could link Europe, Africa, and Asia together due to its protected harbour. The metropolis is made up of the city, suburbs, farming land, and government wasteland.

Almost half of the population is Muhajir, referring to those who migrated (or whose families migrated) from India in 1947, while a quarter are Pashtun. Ethnic groups like Sindhis and Punjabis are represented in smaller numbers. Further, Karachi has a large number of undocumented immigrants, mainly from Afghanistan, Bangladesh and India, along with a steady influx of new residents from different parts of the country. Karachi is organised into seven districts, these are: Korangi, East, West, South, Malir, Central and Karachi District West. (city-spotlight).

The attractive physical surroundings in Karachi have a great impact not only on the city's appearance but also on the standard of living, a fact which promotes environmental and urban growth by providing a natural coastal belt and mountains that surround it along with a favourable topography and moderate climate. The environment consists of low flat hills, parallel ridges, vast open areas, dry riverbeds, and a variety of different water canals, which have their origin from the Tethys Sea and were uplifted by the Himalayas. Significant locations include Mango Pir Hills, Clifton, Manora, Ghizri Hills, and Drigh Road Dome, where tertiary rocks, indicating possible petroleum reserves, are found. The elevations are getting higher as we go south where cliffs go up to 400 feet over the sea while Al-Kamari is getting lower in elevation. Being soft rocks, these deposits are easily weathered, and the action is noticeable from the masses of debris which are commonly found. The surface springs are appearing in denuded anticlines, meanwhile, and the soils are eroding because of these reductions. However, the tourism industry will have interesting stories to tell visitors about the region's rich past and the evolution of different parts of the earth through the times. This has resulted in beaches such as Clifton Beach, the shores of which have been reshaped due to the rising sea levels. The urban policy is to direct expansion towards the stable plains, although many live on the peak terrain adjacent to the ridge where they are in danger of being buried beneath a landslide but at the same time are close to urban amenities. (Pithawala,1949).

Methodology

Data Collection

The data in this study was collected through USGS. Landsat 4, Landsat TM/ETM and Landsat OLI are the satellites whose data was used. Bands 4, 3, and 2 were used and a blend was formed which was used further in the supervised classification.

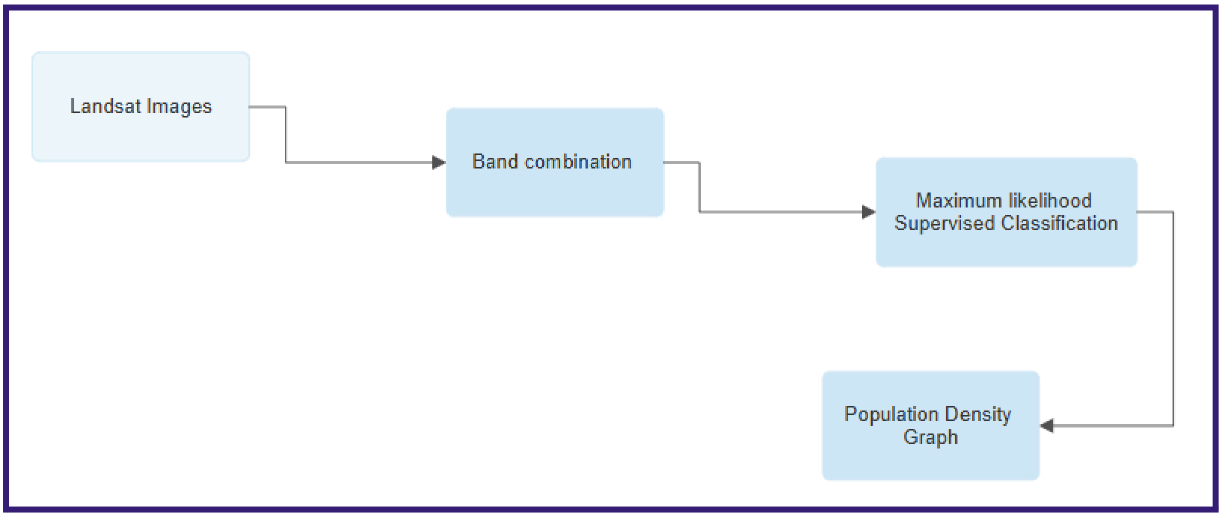


Image Classification Techniques

The maximum likelihood (ML) procedure is a supervised statistical approach to pattern recognition. The probability of a pixel belonging to each of a predefined set of classes is calculated, and the pixel is then assigned to the class for which the probability is the highest. ML is based on the Bayesian probability formula

$P(x,w)=P(w|x) P(x) =P(x|w)P(w)$. (1)

In equation (2), where 'x' and 'w' are generally called 'events'. P(x,w) is the probability of coexistence (or intersection) of events 'x' and 'w.' ' P * (x) deg and ' P(w) are the prior probabilities of events 'x' and 'w', and ' P(w |x)^ * is the conditional probability of event 'x' given event 'w'. P(w |x)^ - is interpreted in the same manner. If event x ^ 1* is the (th pattern vector and 'w' is information class 'then, according to the above equation, the probability that xi belongs to class wj is given by

$P (w^{i} |x^{j})= P (x^{i}|w^{j}) \backslash P(w^{i})$,.........................(2)

Since, in general, P(x) is set to be uniformly distributed (i.e., the probability of occurrence is the same for all pixel features), the equation (2) can be rewritten as

$P (w^{j} |x^{i}) \infty P(x^{i} |w^{j}) P(w^{j})$ ....(3)

One can thus allocate pixel 'I' to the class '*k*', which has the largest value of the term P(w ^ k |x^ i) in equation (3). The classification criterion can be expressed as

w* = argv⋁^wj max[P(x^i|w^j) P(wj)................................(4)

where 'arg' denotes 'argument. The criterion shown in the above equation (4) is called the Maximum A Posteriori (MAP) solution, which maximizes the product of conditional probability and prior probability. However, on some occasions, the prior probability ' P * (w)' is also set to be uniformly distributed (because of a lack of prior knowledge or because one does not know the true distribution). In this case, above equation (4) reduces to

P(w ^ j |x^ i) ∞ P(x^ i |w^ j)....................................(5)

If we allocate pixel 'I' to that class 'k' that maximizes the expression of the above eq., the result is called the maximum likelihood solution. Normally, the conditional probability P(x ^ i |w ^ j) is assumed to follow a Gaussian (or normal) distribution assumption.

The ML classifier assumes that the information class prior probability P(w) is uniformly distributed. However, if one can model P(w) properly, the classification accuracy could be increased. For example, one may model prior probability as different weights associated with each class. A higher weight for a given class implies that there is a higher probability of a pixel receiving the label associated with that class. In recent years, there has been a trend toward modelling the prior probability using a smoothness assumption based on the concept of context. This method is called the smoothness prior. By using the smoothness prior together with the class conditional probability P(x)w), one can attempt to perform a MAP classification (Paul et al., 2009).

Urban Growth Analysis

Land use/ land cover patterns for 2000, 2008 and 2016 were mapped by the use of Landsat 4 & 5 data. For 2020 & 2024 landsat 9 were used. Six land use/landcover types are identified and used in this study which include, 1. Cloud Cover 2. Water Body 3. Vegetation 4. Residential Area 5. Industrial Area 6. Barren Land. A supervised classification with the maximum likelihood algorithm was conducted to classify the Landsat images using bands 2 (Green), 3 (Red), and 4 (near infra-red). The acquired data will show the changes that have been taking place over the years and the changes that have occurred through urbanisation.

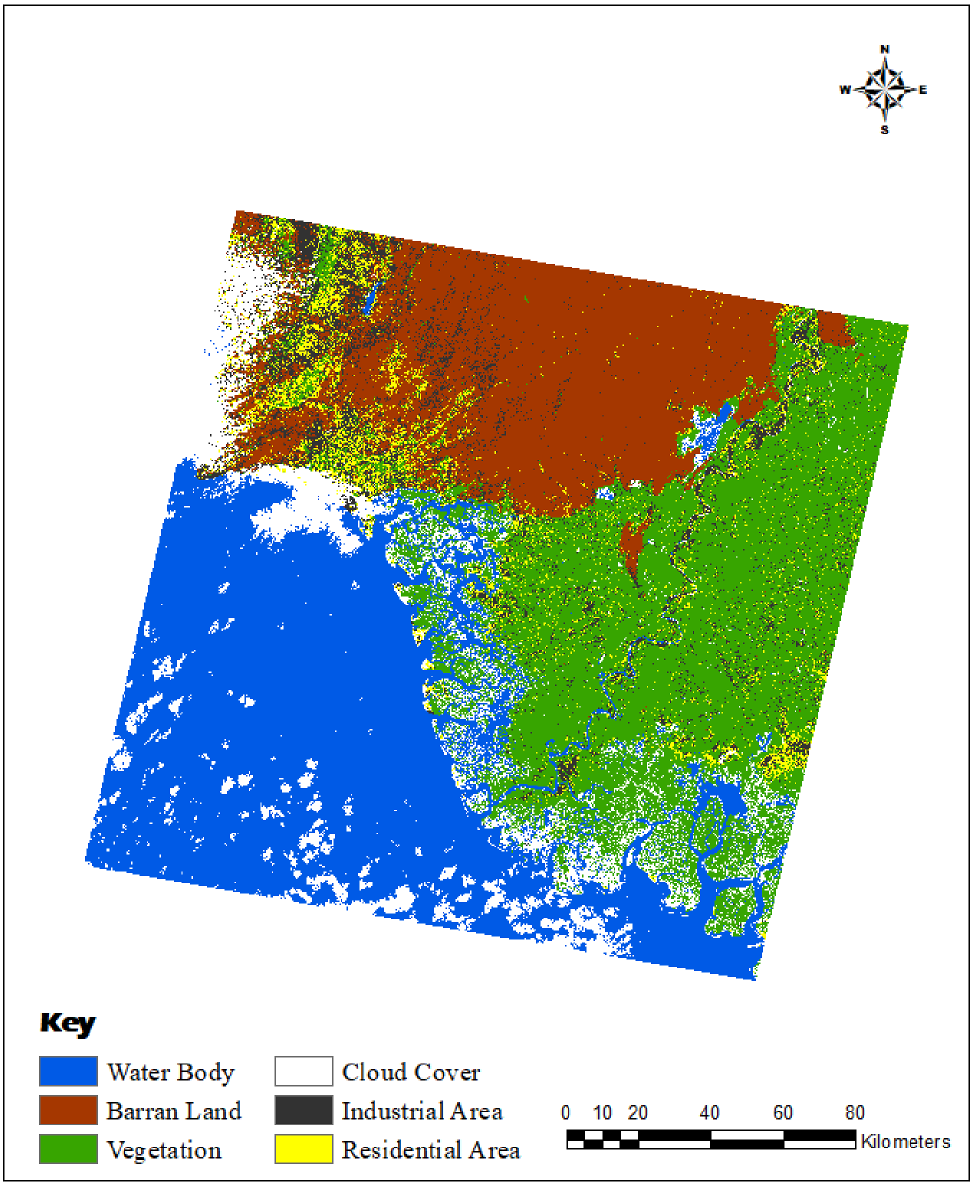
A population density map was also made using the data taken from the different censuses, this data was taken by the Pakistan Census Department. A dot density map was formed using these data, which shows the changes in the population that have taken place over the years.

Result And Discussion

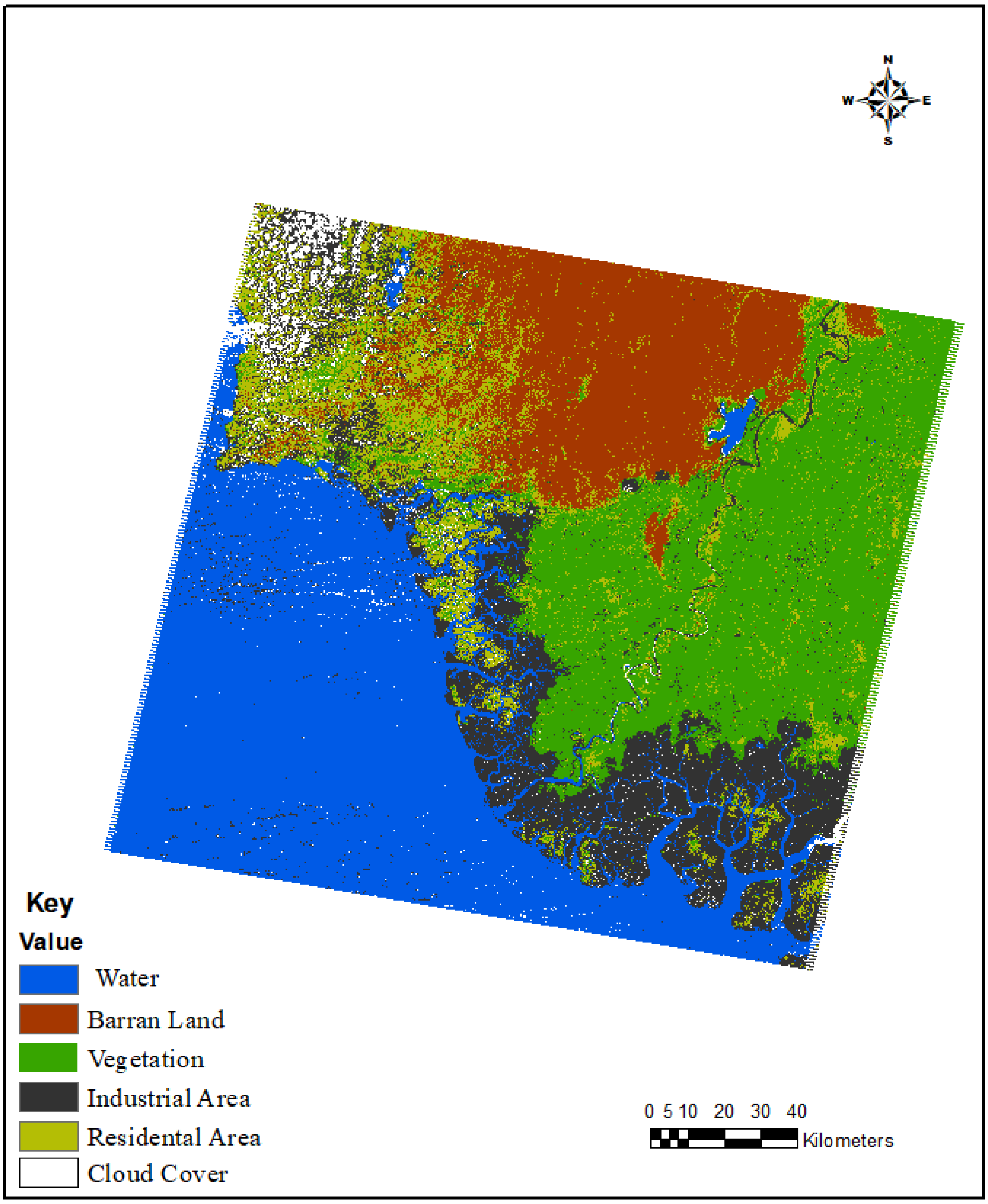
The Karachi city has an area of 3,780 km^2. More than of this area is vacant land. The land-cover features such as water and vegetation change over the year due to the climate effect at that time. But the urban area has extended, there has been a valuable change of 25 km in the last 24 years, which also means that there has been a decrease of 100 km of open land that has been converted into built-up land. Through this study, we can see that there has been a 15% increase in the built-up land and a 3% decrease in vegetation coverage. We can also see that after 2020 there has been a slight difference in the number of Industrial Areas. This growth in urbanisation and decrease in vegetation coverage is a meaningful observation for the stakeholders, environmentalists, engineers, town planners and scientists to plan and suggest future development plans, and strategies.

According to the census of 1998, there was a population of 9.85 million persons. The average annual growth rate is 3.52%. The total urban population of the city jumped from 1,068,459 in 1951 to 14,500,000 in 2007 which means that in the last fifty years, Karachi’s population has continuously increased. Karachi's population has reached 20 million, and there has been an increase of 25% over the years. This study also reveals the use of geo-informatics in urban and regional planning. Especially, in developing countries like Pakistan where tremendous population pressures evolve new cities at very high peace.

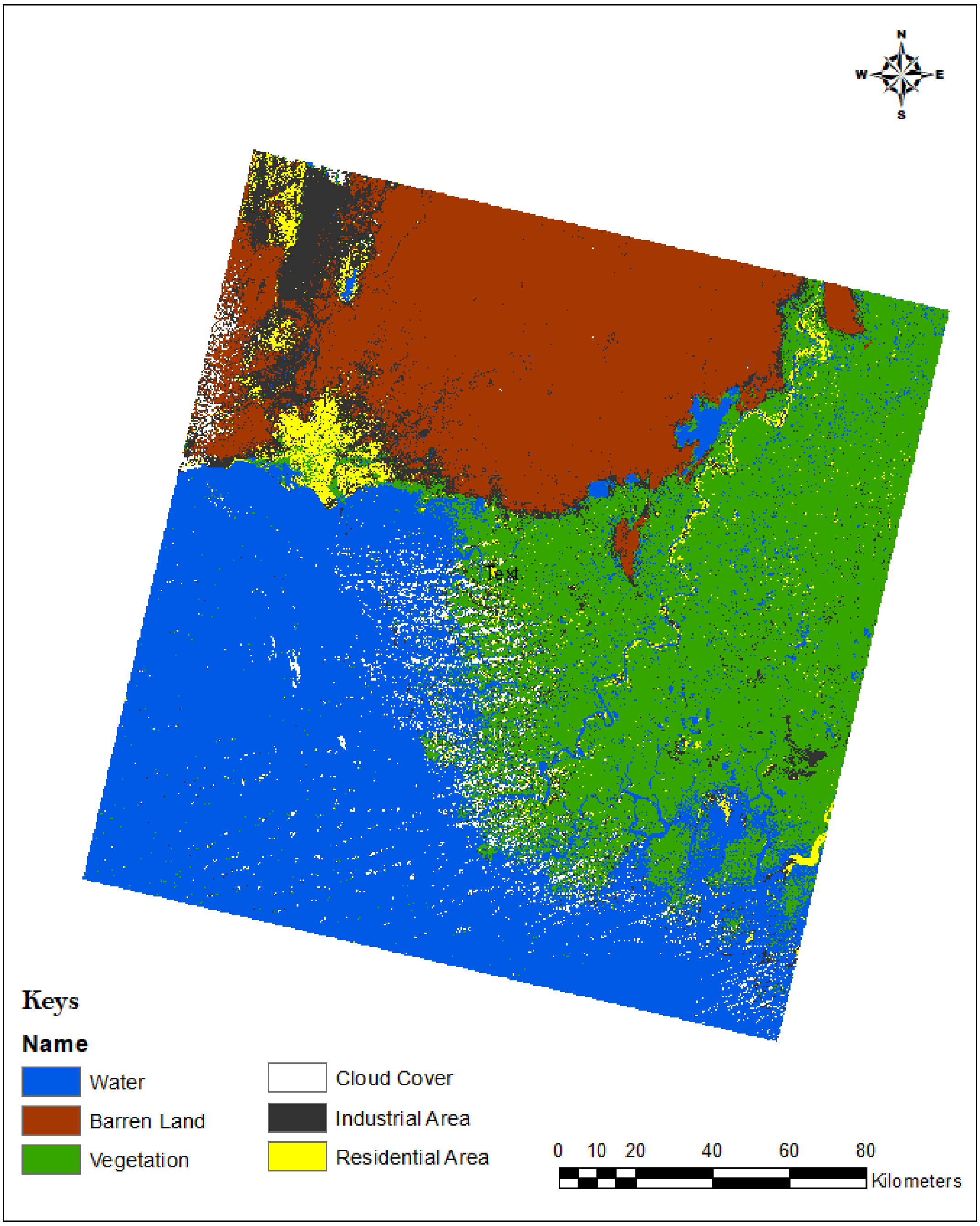
Urbanization In Karachi In 2000



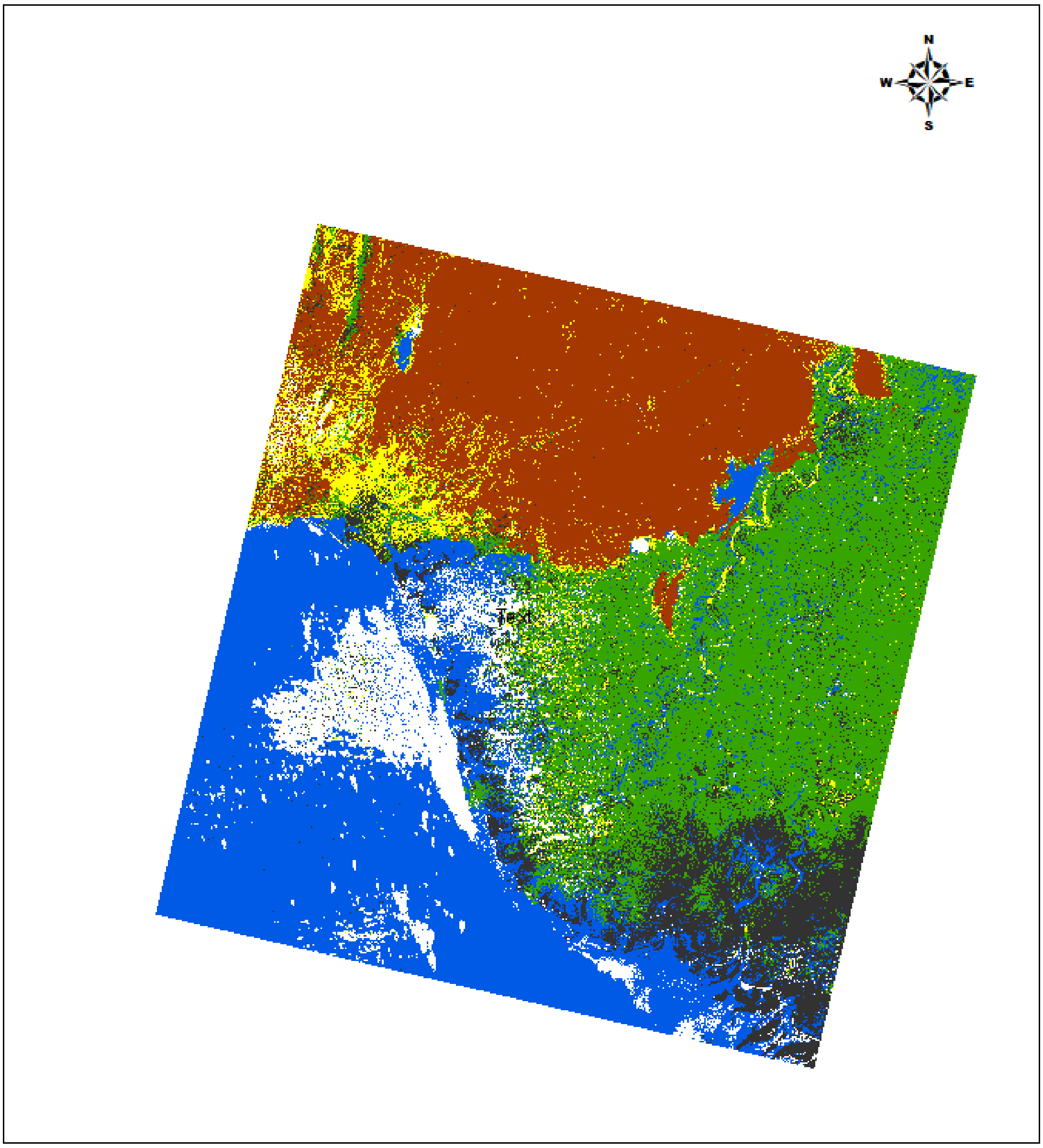
Urbanization In Karachi In 2008



Urbanization In Karachi In 2016



Urbanization In Karachi In 2020

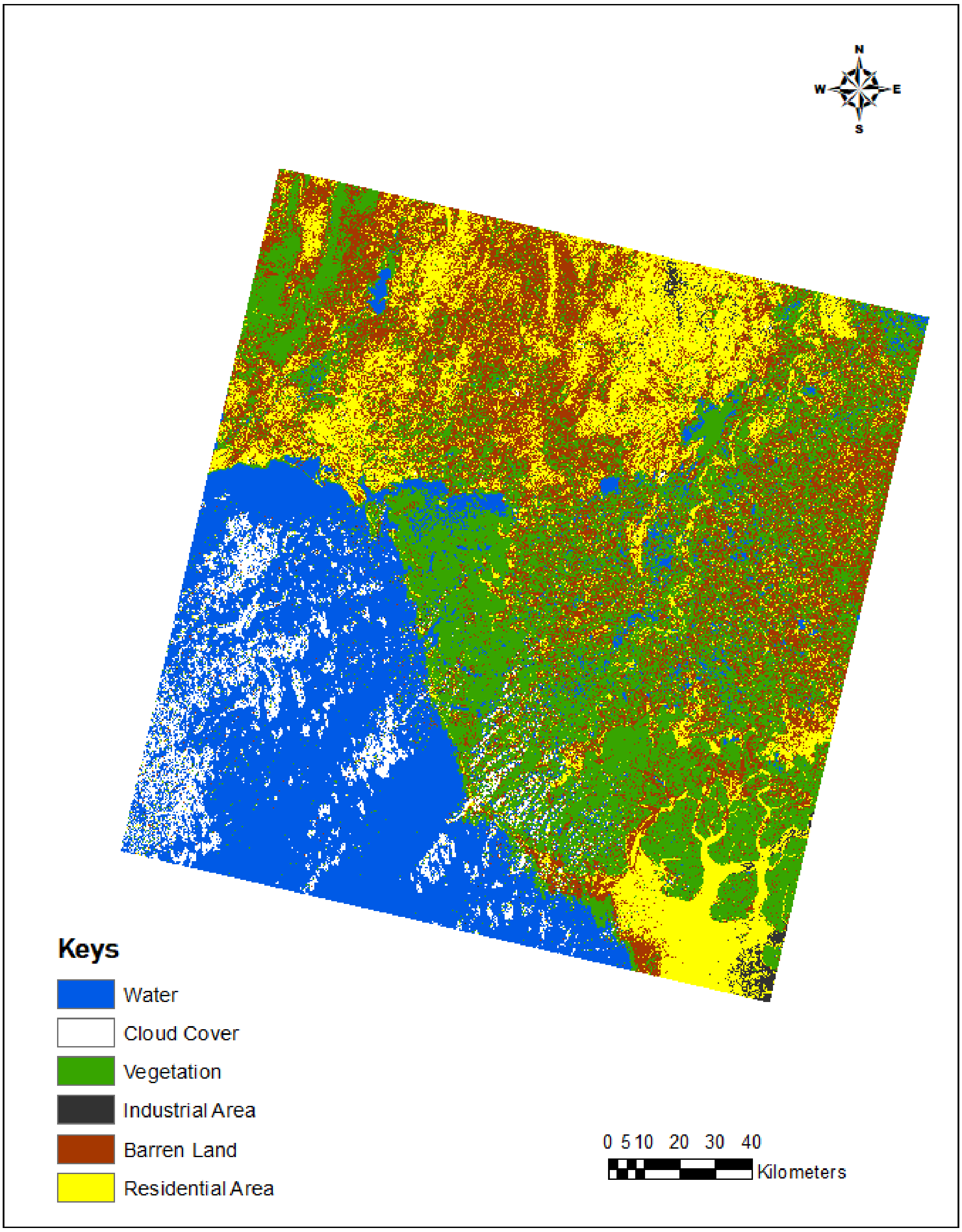


Key

- Water
- Barran Land
- Cloud Cover
- Vegetation
- Industrial Area
- Residential Area

0 510 20 30 40
Kilometers

Urbanization In Karachi In 2024



Year 2000

Feature Name	Area	Percentage
Water	101605923	27.84714712
Barran Land	83050434	22.76164209
Vegetation	93258198	25.55928516
Cloud Cover	28954746	7.935630599
Industrial Area	39733425	10.8897444
Residential Area	18267408	5.006550632
Total Area	364870134	

Year 2008

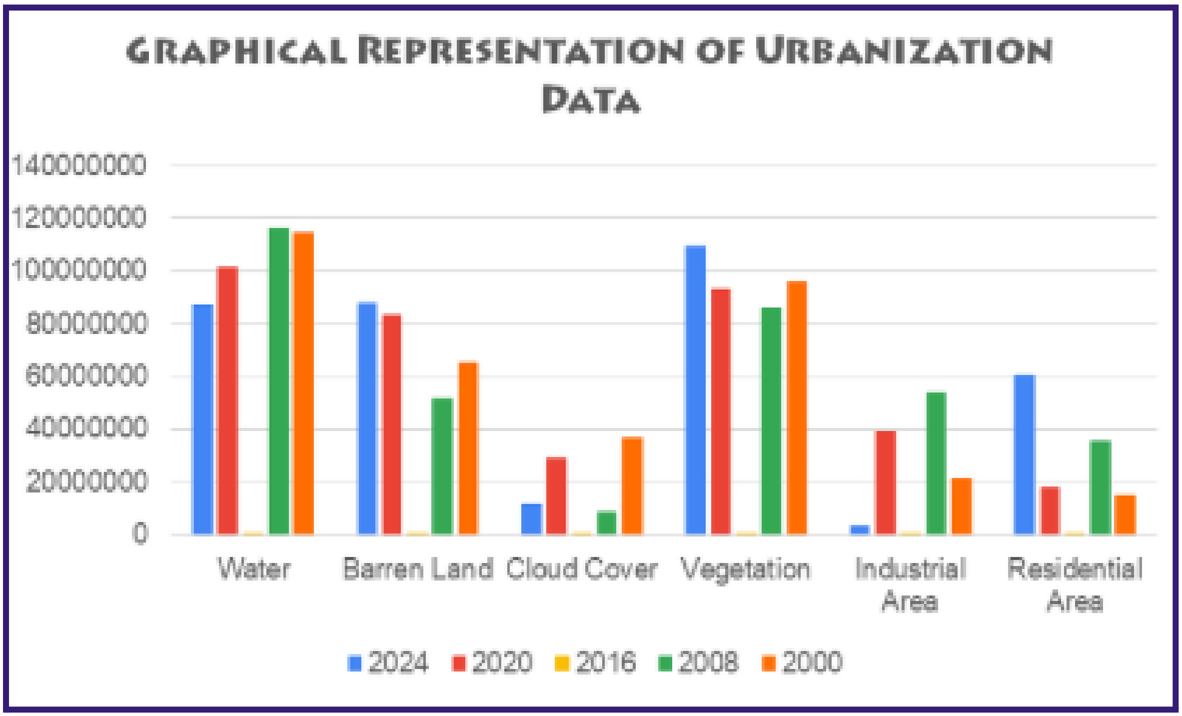
Feature Name	Area	Percentage
Water	116225559	31.85395245
Barran Land	51974964	14.24478442
Vegetation	86403429	23.68059782
Cloud Cover	9105255	2.495478295
Industrial Area	54421677	14.91535534
Residential Area	35409708	9.704742784
Total Area	353540592	

Year 2020

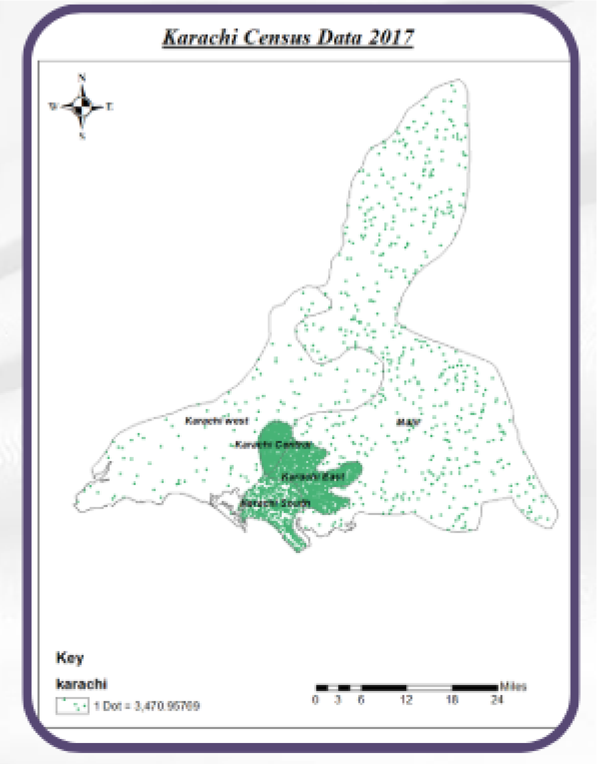
Feature Name	Area	Percentage
Water	114765480	32.83021909
Barran Land	65087505	18.61916187
Vegetation	96299307	27.54772033
Cloud Cover	37022607	10.59081789
Industrial Area	21373191	6.114090606
Residential Area	15024600	4.297990212
Total Area	349572690	

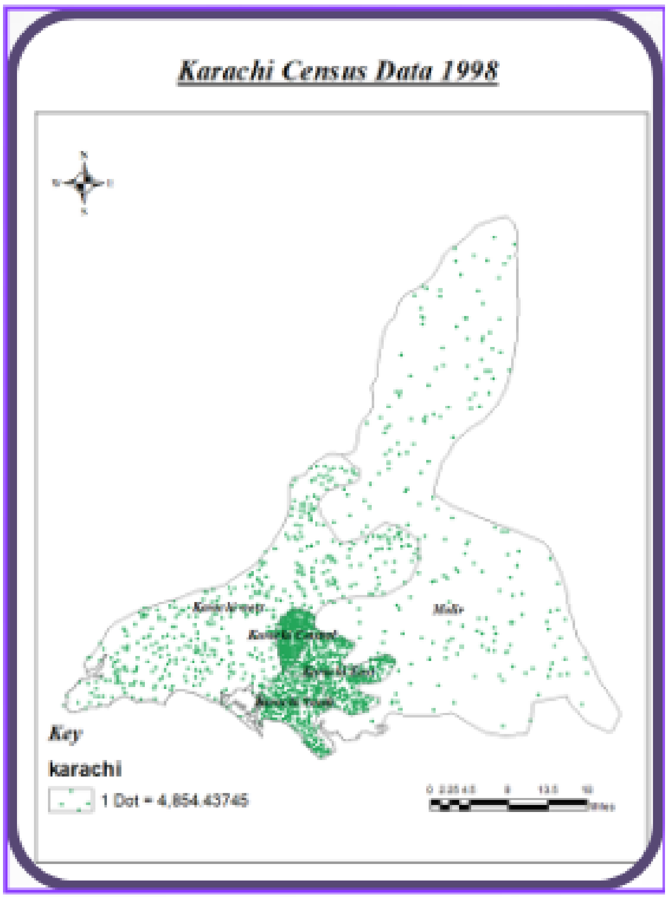
Year 2024

Feature Name	Area	Percentage
Water	87584121	24.26736106
Barran Land	88170165	24.42973914
Vegetation	109285164	30.28017525
Cloud Cover	11651778	3.228415156
Industrial Area	3726873	1.032622942
Residential Area	60495147	16.76168645
Total Area	360913248	




Karachi Census Data 2023
N
W
E
S
Karachi west
Karachi Central
Karachi East
Karachi South
Malir
Key
karachi
1 Dot = 2,325.77498
5 2.5 0 5 10 15 20
Miles


Karachi Census Data 2017
N
S
Karachi west
Karachi Central
Karachi East
Karachi South
Malir
Key
karachi
1 Dot = 3,470.95769
Miles
0 3 6 12 18 24


Karachi Census Data 1998
N
S
Malir
Key
karachi
1 Dot = 4,854.43745
Miles

Conclusion

Karachi, which is the largest city and also has the highest population, is an electric city that drives the economic, cultural, and social activity of Pakistan. Nevertheless, the fast growth of urban areas in the city has brought a lot of drawbacks such as environmental pollution, overloading the infrastructure, the emergence of informal settlements, and the conversion of green spaces into urban areas. To perform the study, the phenomenon utilizes supervised classification using ArcGIS which is a geospatial analysis tool, to identify land use and land cover modifications in Karachi. This technique, which employs satellite imagery and training datasets, offers a detailed account of the expansion of the city, vegetation loss, and demographic changes.

Karachi has been expanding its urban area since the analysis of it, showing an increase in its built-up area of about 53 square kilometres between 1989 and 2014, corresponding to a mean annual growth rate of 2 square kilometres. The urbanization process is in the southern part of the city and from the year 1989 to 2014 there has been more than 40% increase in the land area and around 60% increase in the population. The projected population of Karachi, in 2025, which is 18,266,000 will surpass Los Angeles' population of 16,899,000 by 1,367,000. In the case of Delhi, the population is projected to reach 31,834,000 which is a 63.62^ population increase when compared to that of 2021.

The city's population density is a big problem that the current infrastructure can't solve when every year around 500,000 people are moving into it. Under this pressure, the existing neighbourhoods get denser, and the informal upswell of the settlements occurs. The built-up area increased from 343 square kilometres in 1989 to 396 square kilometres in 2014. However, there have been discrepancies between the actual occupancy and the developed land with many housing societies built in the 1970s, 80s, and 90s being vacant or underutilized.
The development of villages into urban areas along with the sudden boom in the population is speeding the process by which urban problems are getting more acute such as overcrowding, the loss of the arable land, and pollution. The study underlines the necessity of constant control of urban extension because of eco-friendly planning. However, the applications of supervised classification techniques allow spatial data analysts to capture land cover changes, track

human population density over time, and observe the spatial extent of urban sprawl, thus, they can make better-informed decisions.

Recommendations

It is inevitable to implement urban planning strategies that are based on data for sustainable urban development with a proper balance between economic growth and nature conservation. Recommendations include the following:

- **Sustainable Development Planning:** Integrate GIS and remote sensing techniques to amass real-time data for time-sequenced analysis of urban growth thus identifying areas suitable for future development.
- **Policy Interventions:** The urban sprawl, the encroachment of agriculture on the metropolitan areas, and the utilization of new infrastructure are the problems that should be dealt with the help of policy interventions.
- **Infrastructure Development:** The goal is to have the focus shift to developing the industrial and residential zones in the less populated areas to reduce the pressure on the central city.
- **Resource Management:** Green spaces should be the highest concern and environmental resources should be maintained when planning for the city.
- **Urban Monitoring Systems:** To effectively monitor cities, we need to collect and analyse data using up-to-date methods such as GIS and remote sensing, which will help us track changes and appraise impacts.

Reference

Afsar, S., Ali, S. S., & Kazmi, S. J. H. (2013). Assessment of the quality of life in Karachi city through the integration of space and spatial technologies. *Journal of Basic & Applied Sciences*, *9*, 373-388.

Pithawalla, M. B. (Ed.). (1949). Bulletin of the Karachi Geographical Society.
Concrete Hell in Karachi, Pakistan

Mahboob, M. A., Atif, I., & Iqbal, J. (2015). Remote sensing and GIS applications for assessment of urban sprawl in Karachi, Pakistan. *Science, technology and development*, *34*(3), 179-188.

Rizvi, S. H., Fatima, H., Iqbal, M. J., & Alam, K. (2020). The effect of urbanization on the intensification of SUHIs: Analysis by LULC on Karachi. *Journal of Atmospheric and Solar-Terrestrial Physics*, *207*, 105374.

Shaikh, A. A., & GOTOH, K. (2007). A Satellite Remote Sensing Evaluation of Urban Land Cover Changes and Its Associated Impacts on Water Resources in Karachi, Pakistan. *Journal of the Japan society of photogrammetry and remote sensing*, *45*(6), 41-55.

Ahmed, M. W., Saadi, S., Ahmed, M., & Shaikh, A. A. (2024). Decoding Informal Settlements in Core Urban Areas of Karachi: Leveraging Machine Learning Algorithms for Classification and Analysis. *Remote Sensing in Earth Systems Sciences*, 1-14.

Raza, D., Karim, R. B., Nasir, A., Khan, S. U., Zubair, M. H., & Amir, R. (2019). Satellite-based surveillance of LULC with deliberation on urban land surface temperature and precipitation pattern changes of Karachi. *Pakistan. J. Geogr. Nat. Disast*, *9*(1037), 2167-0587.

Baqa, M. F., Lu, L., Chen, F., Nawaz-ul-Huda, S., Pan, L., Tariq, A., ... & Li, Q. (2022). Characterizing spatiotemporal variations in the urban thermal environment related to land cover changes in Karachi, Pakistan, from 2000 to 2020. *Remote Sensing*, *14*(9), 2164.

Ahmed, M. W., Ahmed, M., & Shaikh, A. A. (2024). Digitizing Karachi's Decades-Old Cadastral Maps: Leveraging Unsupervised Machine Learning and GEOBIA for Digitization. *Engineering, Technology & Applied Science Research*, *14*(5), 16404-16410.

Din, S. U., & Yamamoto, K. (2024). Urban spatial dynamics and geo-informatics prediction of Karachi from 1990–2050 using remote sensing and CA-ANN simulation. *Earth Systems and Environment*, *8*(3), 849-868.