Cellular Automata-based Simulation of Urban Growth Using Satellite Imagery and Geospatial Data

Colin Shen, Clear Lake High School, Houston, TX

All graphics credited to the researcher, TXDOT, and Landsat 5, 7, 8

INTRODUCTION

Houston is the largest metropolitan city without official zoning regulations, resulting in irregular growth when compared to other US metropolitan cities. As the city has expanded, forested areas have been replaced with residential areas, biodiversity has decreased drastically, and the climate crisis has grown.

Public policy plays a large role in the growth of a large city like Houston and its resulting environmental impacts. Since the results of public policy may take years to show, it is difficult to propose such policies and measure their sustainability. However, by forecasting the growth of a city due to external factors and public policy, this feat is achievable.

In order to forecast the growth of Houston, this project simulates land use changes based on a novel cellular automata model and a satellite imagery analysis algorithm. This cellular automata algorithm is known as SLEUTH, which considers six factors when determining urban growth— slope, land use, excluded areas, urban extent, transportation, and hillshade.

Areas of major forest loss can be predicted using a non-regulated scenario. Then, a second simulation can test the effectiveness of combatting forest loss in these areas by factoring in government regulation. This way, effective public policy and sustainability can be achieved.

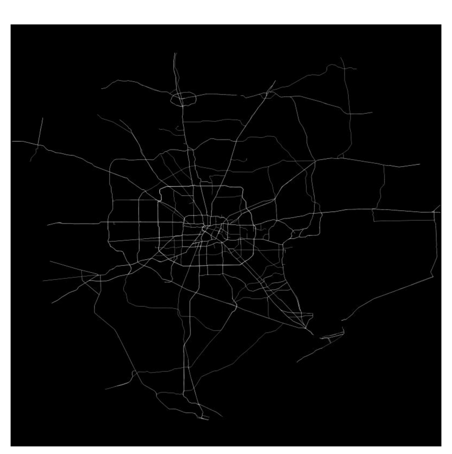
AIM

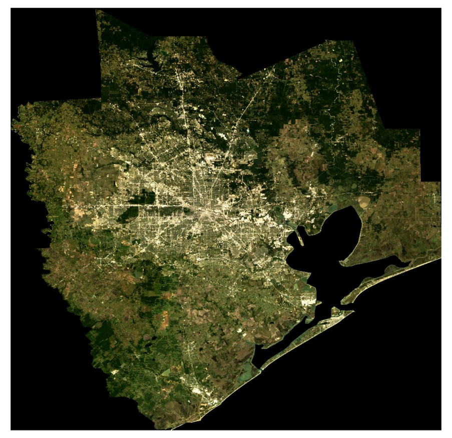
This project aims to:

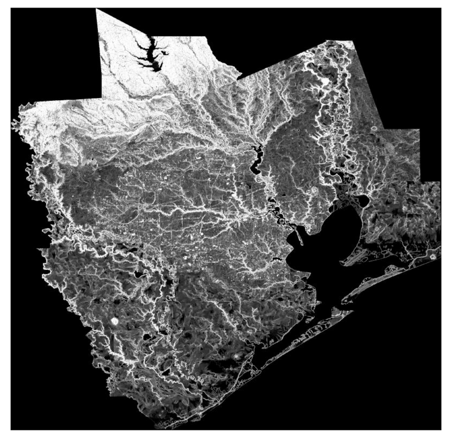
1) reliably simulate Houston's future land cover using aerial data
2) determine areas of unsustainable growth and areas susceptible to biodiversity loss
3) propose conservation policies and test their effectiveness

MATERIALS

1) Python 3.9.6
2) GeoJSON data files
   a) Houston city outline
   b) Houston current and historical road networks
3) Google Earth Engine (GEE)
   a) Access to Landsat 5, 7, 8 satellites
   b) Access to National Elevation Datasets
4) GIS/Data tools: Excel, QGIS, Notepad++
5) Image processing tools: OpenCV2, PillowPIL, scikit-image







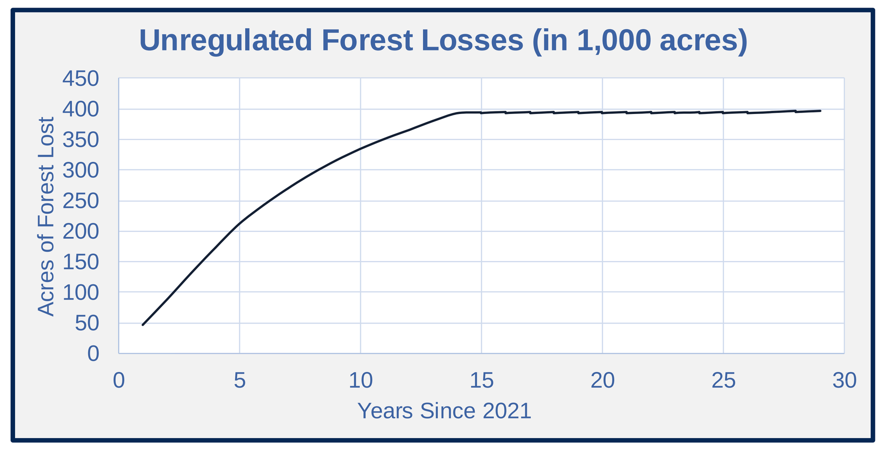
PROCEDURE FOR DATA COLLECTION

Parse Houston JSON boundary into GEE → Filter for <5% cloud cover, map from 1985 to 2020 → Download blue, green, NDVI, and SWIR 1 bands → Normalize images by matching histograms → Filter pixels individually based on luminance → Extract excluded layer from black pixels on NDVI band

PROCEDURE FOR CELLULAR AUTOMATA SIMULATION

Convert land cover, roads, excluded layer, slope, and hillshade into GIF images → Do brute Monte Carlo runs for calibration (approx. 9000+ iterations) → Gather Dispersion, Spread, Breed, Slope resistance, and Road coefficients. → Set coefficients and forecast model to 2050 → Map forest loss into excluded layer, to simulate maximum forest protection → Run model with same coefficients and updated exclusion layer

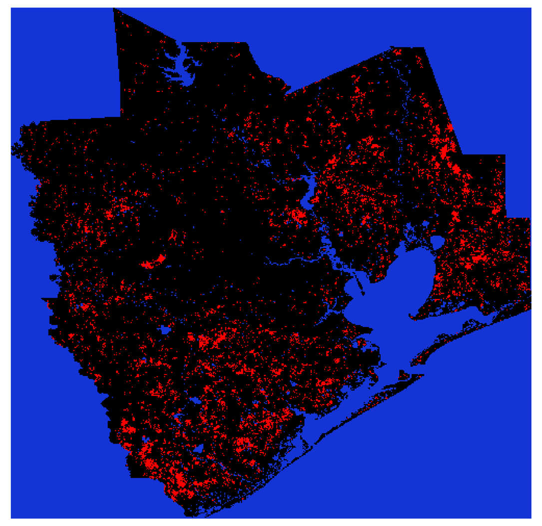
SIMULATION-PREDICTED DATA



The image and chart represent Houston's predicted forest loss from 2021-2050, without government conservation policies.

Losses are in red

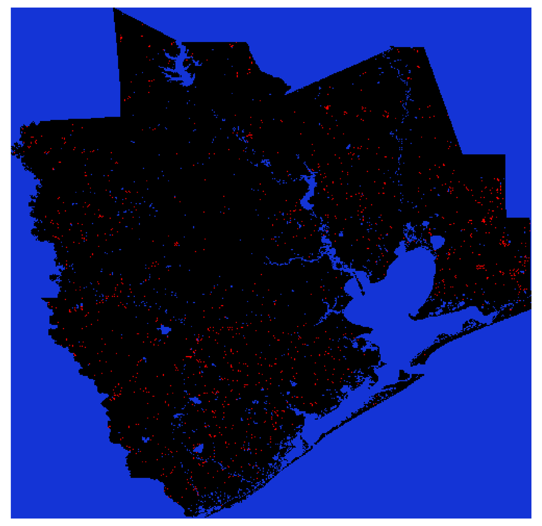
396,078 acres of forest are expected to be lost in 30 years.

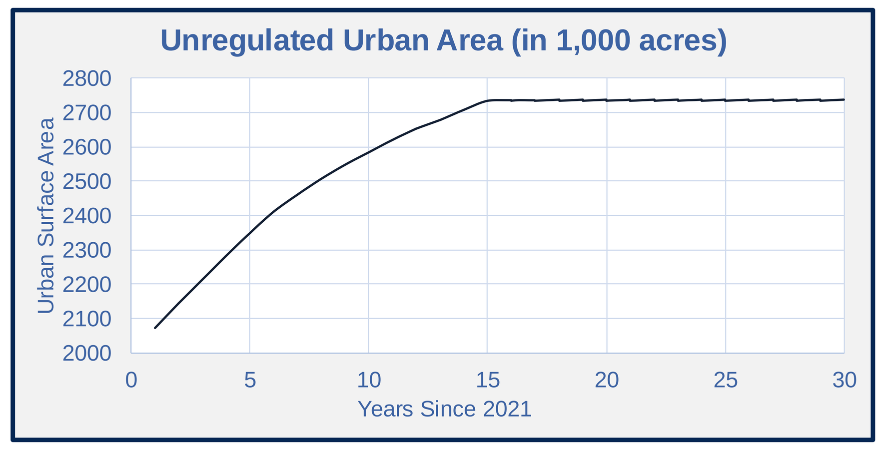


The image and chart represent Houston's predicted forest loss from 2021-2050, with strong government conservation policies.

Losses are in red

58,004 acres of forest are expected to be lost in 30 years.

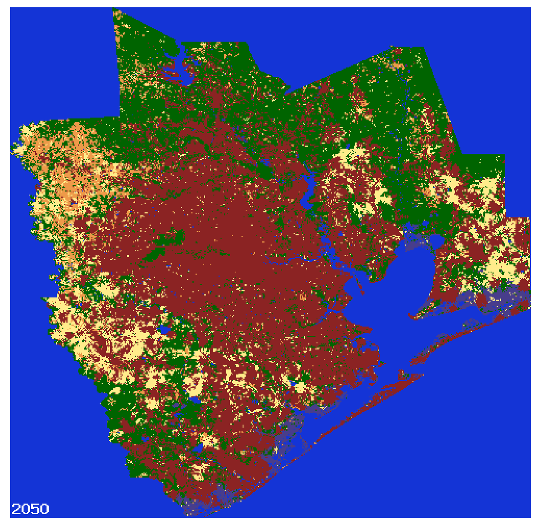


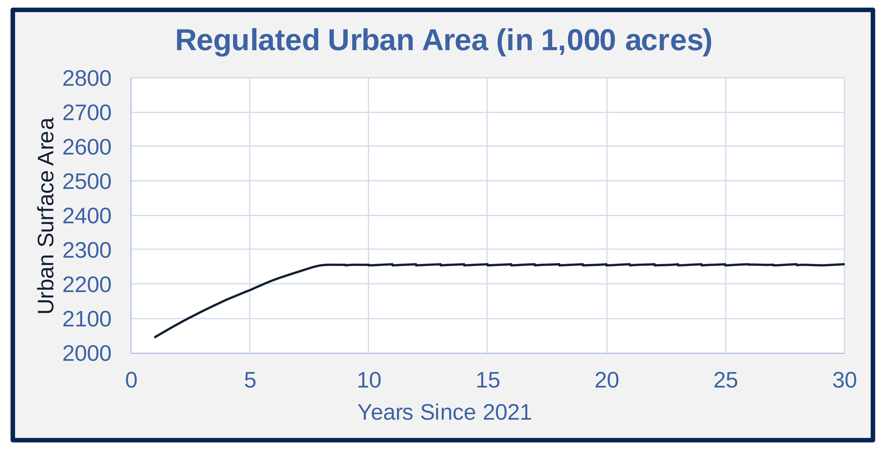


The chart represents Houston's urban growth from 2021-2050. The image represents Houston's projected land cover in 2050 without regulation.

Urban areas are in red

2,738,000 acres of urban land in 2050

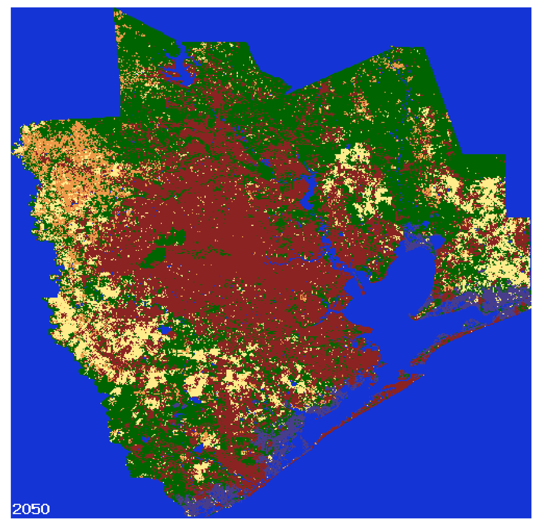




The chart represents Houston's urban growth from 2021-2050. The image represents Houston's projected land cover in 2050 with regulation.

Urban areas are in red

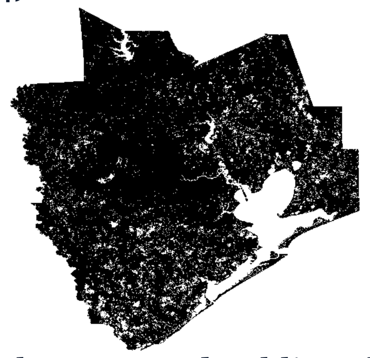
2,258,000 acres of urban land in 2050



DISCUSSION

Note the large disparities between the forest losses of the non-regulated forecasting and regulated forecasting. Non-regulated growth-

1) has much more pronounced forest losses
2) has much larger urban sprawl
3) has additional wetland habitat losses
4) faces forest loss for a longer time period (15 years)



Proposed public policy plan (Protected forest areas in white)

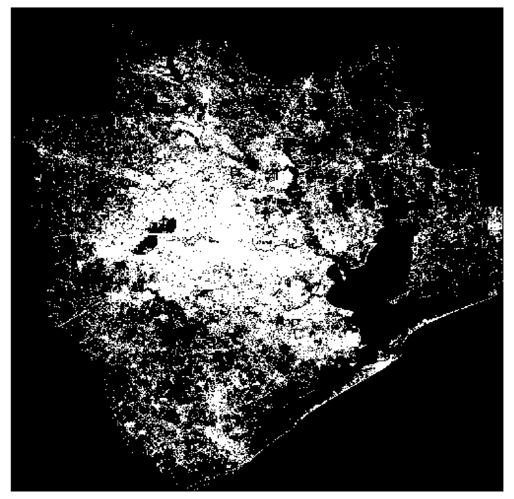
The proposed public policy plan and protected areas are derived from the non-regulated scenario. This policy has been demonstrated feasible through a second regulated simulation, involving more intense exclusion areas. Use of the new policy decreased forest losses from 396,078 acres to 58,004 acres— a 85.36% decrease. The process used can be generalized to other cities and land classes, such as costal habitats or farmland.

VALIDATION

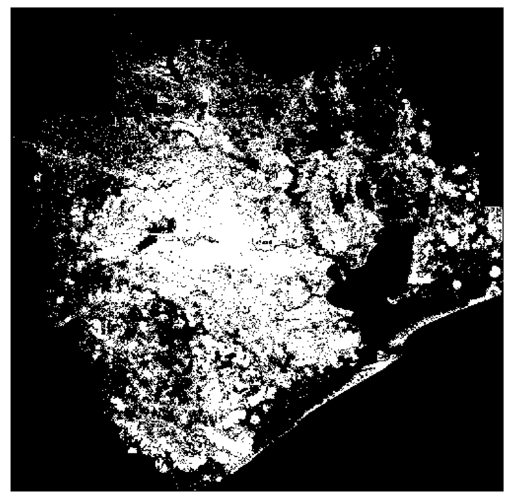
Final Coefficients: 1, 6, 91, 25, 17 (Diffusion, Breed, Spread, Slope, Road)

To validate the derived growth coefficients and SLEUTH model, the model was projected to 2020 and compared with 2020 ground truth data.

Total Accuracy: 90.78%



2020 Actual Growth



2020 Predicted Growth

CONCLUSION

Overall, these simulations were highly accurate and effective at predicting land cover changes, identifying areas of potential environmental concern, and testing the implementation of new sustainable policies.

For the satellite imagery processing algorithm developed, the land use classification was feasible for up to 95-98% of pixels. By downloading each band individually and filtering pixels by luminance, land cover features could be extracted easily. In addition, the growth coefficients were found to have a 90.78% accuracy rate. This accurate simulation model was later used to successfully test the impact of public policy.

This process may be generalized beyond Houston's case of forest loss, allowing it to be applicable to a wide variety of cities and habitats. In the future, the model can be expanded to add additional urban metrics, such as population density, land costs, upward construction, and socioeconomic data. Other possible areas for exploration include estimating the carbon footprint as cities expand, and the addition of a travel model.