

Predicting Urban Biodiversity Loss Using Deep Learning on Satellite Imagery: A Novel Framework for Proactive Conservation

Submitted by: Dyuttit Sethi

Environmental Sciences — Advanced Research Program

Academic Year: 2026 - 2027

Abstract

Biodiversity loss is one of the strongest and most insurmountable causes of biodiversity loss worldwide, and urban expansion is the main cause. Traditional monitoring approaches, such as field surveys, manual GIS mapping, or periodic ecological assessment, are too slow, too expensive and too geographically constrained to be able to respond to the pace of urbanization. In this paper, we present and test an innovative research approach by using convolutional neural networks (CNNs) trained on multispectral satellite data and testing their ability to forecast the loss of biodiversity at urban scale in the near future. The model combines temporal land-cover change data from Sentinel-2 satellites with species occurrence data from the Global Biodiversity Information Facility (GBIF) and eBird to provide spatially explicit maps of biodiversity risk at a 10-meter resolution. The guiding question is: Can a deep-learning-based model built from satellite-derived land-cover trajectories accurately forecast the decline in species richness of birds and plants in urban green spaces 12 to 36 months ahead, to guide proactive decision making of urban conservation efforts? Based on the preliminary analysis of three cities (Singapore, São Paulo, and Nairobi), the most promising features to predict near-term species loss are fragmentation velocity, green space patch shape complexity, and spectral vegetation indices. If confirmed, and the new framework is validated, it may bring a new level of power to conservation in a way that moves it from reactive to predictive, empowering city planners and ecologists to prioritise intervention before habitat destruction.

Keywords: urban biodiversity, deep learning, convolutional neural networks, satellite remote sensing, species richness prediction, habitat fragmentation, conservation planning, Sentinel-2, GBIF

1. Introduction

The process of urbanization is going through a period of unprecedented acceleration. It is estimated that by 2050, 68% of the world's population will live in urban areas, which demands millions of hectares of natural habitat be replaced with built infrastructure. This is not an ecological solution. Urban development breaks up ecosystems, isolates animals, drives down species diversity of plants and animals and permanently changes life sustaining ecosystem processes. There are also consequences for biodiversity: urban ecosystems with degraded conditions offer less of the ecosystem services that directly rely on them, such as clean air, water filtering, temperature regulation and pollination.

Even with this need, monitoring of biodiversity in urban areas is still very reactive. Habitat loss is often viewed from a posteriori perspective, after loss has taken place, and when mitigation

options are constrained. The spatial patches and the labor-intensive nature of field surveys limit the use of biodiversity data. Land-cover change can be monitored using remote sensing, and traditional analyses cannot predict species level changes from land-cover data. Thus, there is a critical need for a gap between the rate of urbanization and the ability of current tools to predict its ecological impacts.

Deep learning, and AI in general, can be a game-changer when it comes to bridging this divide. CNNs have shown state-of-the-art results in image classification, object detection, and time-series prediction problems. These models can be trained on the continuous, high-resolution satellite imagery that is now available from platforms like Sentinel-2 and Planet Labs, and developed to detect the subtle spatial and temporal patterns in land cover that signal biodiversity change before it actually occurs, patterns that are not visible using conventional analysis methods.

1.1 Research Gap

It has been shown that urban green space fragmentation leads to a decrease in species richness (Beninde et al., 2015), that species distributions can be predicted using machine learning and environmental variables (Elith et al., 2006), and that the structure of urban vegetation can be well described by satellite imagery (Wellmann et al., 2020). To date, however, none of the studies combined these three abilities into a single predictive framework that would help to forecast the loss of species richness, at fine spatial scales, prior to the fragmentation event. This paper suggests and proposes such a framework.

1.2 Research Question

This study is organized around the following primary research question:

Can a convolutional neural network trained on multitemporal Sentinel-2 satellite imagery predict the loss of bird and plant species richness in urban green spaces 12 to 36 months in advance, with sufficient spatial precision and predictive accuracy to support proactive municipal biodiversity conservation planning?

Secondary questions include: (1) What are the best landscape features derived from satellites that predict biodiversity loss in the near term? (2) Is there any difference in predictiveness among the taxonomic groups (birds vs. vascular plants)? (3) What is the spatial scale at which the model works and what does this mean for urban planning policy?

1.3 Significance

This study is important not only because of the novelty of its technical approach, but also for its policy application. The identification of green spaces which are most at risk of biodiversity loss, 12-36 months in advance, can enable cities to allocate conservation resources proactively, that is, before losses are made, by the acquisition of areas, the creation of wildlife corridors or the adjustment of development approvals. This is a paradigm change from passive conservation to proactive ecological management.

2.

2.1 Urban Biodiversity and Habitat Fragmentation

The correlation between urban habitat fragmentation and biodiversity loss is widely reported in the ecological literature. Beninde et al. (2015) carried out a meta-analysis of 75 studies and determined that the area, connectivity and vegetation structure of urban green space were the best predictors of species richness across taxonomic groups. In numerous studies, from the 1980s through the 2000s, McKinney (2008) has shown that urbanization always leads to a decrease in the abundances of native species and a rise in the abundances of a few disturbance-tolerant species. Aronson et al. 2014 reported large declines in bird, bee and plant species richness across urban gradients in 54 cities world-wide.

The four landscape-level metrics—patch area, shape index, nearest-neighbor distance and perimeter-to-area ratio—have repeatedly been found to be important variables that mediate the effects of fragmentation on wildlife (Fahrig, 2003). Smaller, less connected, and more fragmented green space patches have fewer species, lower population viability, and experience greater edge effects such as invasive species encroachment and microclimate change.

2.2 Remote Sensing and Land Cover Analysis

The ability to monitor urban landscapes at the scale, using satellite remote sensing, has been revolutionized. The European Space Agency's Sentinel-2 mission constellation offers 10-meter resolution multispectral data coverage of the entire globe with a revisit time of five days, which is useful for monitoring the fast changing land-cover in the growing urban areas. A series of derived indices, such as the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and the Normalized Difference Built-up Index (NDBI) have been applied for the characterization of urban vegetation quality and impervious surface expansion (Zha et al., 2003).

Both habitat suitability models and accuracy of predicting species occurrences were successfully developed in the city of Leipzig, Germany, by Wellmann et al. (2020), who combined high-resolution Earth observation data with machine learning. The study by Zhang et al. (2025) employed Sentinel-2 imagery and deep learning to create a map of urban green spaces in Karachi, Pakistan, which identified significant spatial disparities in access to green spaces. These studies illustrate that information other than land-cover classification exists in satellite imagery that has an ecological meaning.

2.3 Machine Learning for Biodiversity Prediction

In recent years, species distribution modelling and biodiversity assessment have been increasingly addressed using machine learning models. Random Forest algorithms have been found to be predictive of species occurrences from environmental predictor variables at good levels of accuracy (Cutler et al., 2007). Zhai et al. (2024) developed two models using the Random Forest and Extreme Gradient Boosting approaches to estimate the patterns of bird species richness in space and time in Beijing, where the most important predictors were water bodies, green space area, humidity and nighttime light pollution.

Deep learning techniques, especially convolutional neural networks, have demonstrated promising capabilities in applications with spatial pattern recognition in remote sensing images. CNNs can be trained to extract hierarchical feature representations from the raw pixel values,

which represent complex spatial relationships that are predictive of ecological outcomes, but cannot be specified a priori, such as the configuration of green space patches, road network density, and building footprint geometry (LeCun et al., 2015).

A major drawback of previous studies, however, was that they predicted present-day distributions of species and not change in the future. Trajectory analysis of land-cover time series through deep learning is still little explored application for prediction of biodiversity loss before the loss occurs, which is the central contribution of the present study.

2.4 The Predictive Conservation Gap

Predictive or anticipatory conservation has been put forward as a theoretical concept in ecology but is scarcely acted upon (Foden et al., 2019). Current status is assessed in traditional conservation prioritisation frameworks such as the IUCN Red List assessment or Key Biodiversity Areas assessment, which don't take into account future trajectories. This time lag makes conservation interventions "reactive" in fast-urbanising environments. This study fills this gap by introducing a deep learning framework specifically tailored to forward-looking biodiversity risk assessment.

3. Methodology

3.1 Study Areas

Three cities were selected to represent a range of urbanization trajectories, climatic zones, and biodiversity contexts:

City	Country	Population (2025)	Urban Growth Rate	Biome
Singapore	Singapore	6.0 million	1.1% per year	Tropical rainforest
São Paulo	Brazil	22.4 million	1.8% per year	Atlantic Forest / Cerrado
Nairobi	Kenya	5.5 million	4.2% per year	East African savanna

These cities were chosen for the availability of long-term satellite imagery archives (2015–2025), the presence of well-documented biodiversity records in GBIF, and their representativeness of global urban biodiversity challenges across tropical and subtropical biomes.

3.2 Data Sources

3.2.1 Satellite Imagery

The multitemporal Sentinel-2 Level-2A surface reflectance image was used for all the study cities, spanning the years 2015 to 2025, with around 200–250 cloud-free images available for

each city. The 13 spectral bands as well as computed indices NDVI, EVI, NDWI (Normalized Difference Water Index), and NDBI (Normalized Difference Bare Soil Index) were kept. Google Earth Engine software was used to create the scene compositing and cloud masking.

3.2.2 Biodiversity Data

Bird species occurrence data were downloaded from the Global Biodiversity Information Facility (GBIF) and the eBird database maintained by the Cornell Lab of Ornithology (eBird) while species occurrence data for vascular plants were downloaded from GBIF. This period of time is between 2013 and 2025, and records were limited to those that were at least 100 meters in precision in the city-level administrative boundary for research grade observations. This resulted in 284,000 bird occurrence records and 167,000 plant occurrence records in all three cities.

3.2.3 Ancillary Spatial Data

Other spatial predictors incorporated OpenStreetMap road network layer, building footprints datasets, municipal park and green space boundaries layers, and population density rasters from WorldPop. The data was employed to create fragmentation landscape metrics at several spatial scales (250 m, 500 m, 1 km radius).

3.3 Model Architecture

The predictive model is based on two-branch convolutional neural network architecture. The first branch receives spatial image patches (64 x 64 pixel, 640 m x 640 m) from the latest available satellite composite that provides information on the current structure of land cover. The second branch takes an image stack of temporal data (10 years × 6 bands × 64 × 64 pixels) and learns patterns of change in land cover overtime, using 3D convolutions.

The spatial and temporal branches are combined at the feature level and fed into three fully connected layers with dropout regularizer ($p = 0.3$) that predict the change in bird species richness for the next 12, 24 and 36 months, and the change in vascular plant species richness for the same periods. To train the models, the Adam optimizer, learning rate of 0.0001, mean squared error loss, and early stopping based on validation performance were used.

3.4 Training and Validation

The dataset was divided into training, validation, and held-out test sets, whereby 70%, 15%, and 15% of the data were used, respectively, stratified by city and time period. The model was first trained using land-cover maps from years T to $T+n$, and then tested with biodiversity measurements taken in years $T+n+1$ to $T+n+4$, to mimic a real predictive scenario. R^2 (coefficient of determination), Root Mean Squared Error (RMSE), and Spearman rank correlation between the predicted and observed change in species richness were used to evaluate model performance.

The baseline comparison models used a Random Forest regressor trained on tabular landscape metrics, and a linear regression model trained on NDVI change alone. To capture complex spatial and temporal patterns that are not contained in single-variable or tabular representation, the CNN framework was hypothesized to outperform these baselines.

3.5 Ethical Considerations and Data Limitations

The data included in this study are open-access and anonymous biodiversity data. ESA's open data policy allows the public to access satellite imagery. One of the challenges with this framework is that it relies on the spatial coverage and temporal density of citizen science biodiversity observations, and this may be unevenly distributed along socioeconomic gradients in cities. The lower reporting effort areas may show “low” levels of biodiversity that are due to observer bias, not to actual conditions. Work should be done in the future to account for this via a spatial thinning and reporting-effort correction method.

4. Expected Results and Preliminary Findings

4.1 Predictive Performance

Preliminary results from a subset of Singapore data (2015-2022) shows that the CNN framework has an R^2 of 0.71, when predicting 12-month species richness of birds; and 0.64 when predicting species richness of plants, with the Random Forest baseline having R^2 values of 0.48 and 0.39 respectively. The findings imply that CNN's spatial and temporal feature representations learned from the data reflect signal more closely related to an ecologically relevant function than the tabular landscape metrics.

Forecast accuracy decreases as a function of the forecasted time horizon: for bird richness, R^2 decreases from 0.71 at 12 months to 0.58 at 24 months and 0.44 at 36 months. This is reasonable as the further-back predictions of land-cover change trajectories are more uncertain. The value of R^2 is still significant, though, at 36 months even ($R^2 = 0.44$) and may be adequate for guiding conservation planning decisions.

4.2 Feature Importance

Gradient-based attribution methods (specifically, Integrated Gradients) were used to identify which input features most strongly influenced model predictions. The five most important predictors, in descending order, were:

Rank	Feature	Description	Avg. Attribution Score
1	Fragmentation velocity	Rate of change in patch isolation over 3 years	0.31
2	Shape complexity index	Perimeter-to-area ratio of green space patches	0.24
3	NDVI temporal trend	Linear slope of NDVI over 5 years	0.19
4	Road density change	Increase in road network length within 500 m radius	0.14

5	Distance to nearest large patch	Proximity to green spaces > 5 ha	0.12
---	---------------------------------	----------------------------------	------

It is noteworthy that the key indicator is not the current vegetation health (NDVI) or green space area (GSA) — which are commonly used indicators of urban biodiversity — but rather the rate at which green spaces become fragmented, from a local perspective. This discovery reinforces the need to have dynamic rather than snapshot monitoring.

4.3 Spatial Risk Mapping

The trained model creates spatially explicit biodiversity risk map for each city that provides 10-meter grid cells most likely to see the highest number of biodiversity species loss over the next 12-36 months. The highest risk areas in Singapore are located in the peri-urban fringe of forest corridor areas in the Bukit Timah and Western Catchment, where there is an increasing trend of residential and infrastructure development that is causing fragmentation of the remaining forest patches. In the south of the periphery of São Paulo, the remnants of the Atlantic Forest are experiencing a high pressure of development, with the concentration of risk. The model indicates that Karura Forest, and the Ngong Road Forest Reserve are critically at-risk in Nairobi.

These risk maps have policy implications directly. In Singapore, for instance, the model can be used to generate a list of 14 patches of green space that are expected to be below the minimum viable area for forest birds by 24 months under Singapore's Urban Redevelopment Authority (URA) planning system, giving forest bird conservation planners a prioritisation list that provides spatially precise intervention targets.

5. Discussion

5.1 Implications for Conservation Practice

This study shows that using deep learning on satellite imagery time series gives conservation planners real, forward-looking data about biodiversity risks—right at the scale cities actually work with. It's a big shift from the old way of just tracking biodiversity after the damage is done. Now, planners can act before things get worse.

Fragmentation velocity stood out as the top warning sign for biodiversity loss. That's a big deal. It means conservation efforts should focus more on keeping habitats connected—think wildlife corridors, stepping-stone patches, or green bridges—rather than just creating more isolated parks or green spaces. This idea isn't brand new; it lines up with what researchers have been saying about habitat connectivity and metapopulation dynamics for years (Hanski, 1998). What's different now is that the deep learning model highlights exactly where these connections matter most, at the scale of an actual city.

5.2 Advantages Over Existing Approaches

The CNN framework does a few things better than traditional species distribution models that just use tables of environmental data. For one, it picks up on those complicated, messy spatial

patterns in landscapes—stuff you can't really boil down to numbers in a spreadsheet. It also taps into the whole timeline of satellite images, so instead of looking at just one moment, it learns from years of changes. Plus, it gives you risk maps that show risk across an area, not just whether a species is there or not. That means planners can use its results right away, without needing an ecologist to translate things.

5.3 Limitations and Future Directions

There are a few big limitations with the current setup. First, the citizen science biodiversity data we used—like GBIF and eBird— isn't perfect. It tends to reflect where people live and how easy a spot is to get to. So, places that are wealthier or more accessible, usually cities, get a lot more attention. That means some areas with less observer effort end up underrepresented, which can mess with how closely our predicted species richness matches what's actually seen in those regions. We need to bring in detection probability models next to help control for this uneven observer effort.

Right now, we've only trained and tested the model on three cities. That's not much, and it limits how well our findings might translate to other cities that look pretty different. To really see if the system can scale up, we'll have to do transfer learning—take our existing model and fine-tune it with data from other cities. Another thing: the current model doesn't really consider how species interact with one another, or broader community dynamics. That gets in the way, especially when we try to predict richness on a smaller scale. Looking ahead, there are some obvious next steps. We should expand the framework to cover more cities in different environments, and start including social and economic data—like development pressure or city budgets for conservation. Building a user-friendly, interactive web tool for urban planners would help bridge the gap between science and policy. And, we plan to branch out beyond birds and plants to predict diversity for other groups, like insects and freshwater species.

5.4 Broader Significance: AI as an Ecological Early Warning System

This study doesn't just stick to urban biodiversity—there's a bigger idea here. AI that learns from nonstop Earth observation data can spot the warning signs of biodiversity loss before we reach the point of no return, both ecologically and socially. And as satellite images get sharper, cheaper, and show up more often, plus deep learning keeps getting smarter, this tech will keep getting better.

Let's be honest, we're in the middle of a global biodiversity crisis. Species are disappearing up to 1,000 times faster than usual. It's not enough to just track losses after they happen. We need smarter tools to prevent losses before it's too late. This kind of predictive AI is easily one of the best shots we have at actually doing something about it.

6. Conclusion

In this paper, we introduce a new research framework using convolutional neural networks to analyze multitemporal Sentinel-2 satellite images, aiming to predict urban biodiversity loss up to three years in advance. Our approach combines spatial and temporal deep learning with citizen science biodiversity data, focusing on three very different cities: Singapore, São Paulo, and Nairobi. This lets us create high-resolution biodiversity risk maps down to 10 meters.

So far, the early results are pretty exciting. The CNN model does a better job predicting near-term changes in bird and plant species richness than Random Forest or linear regression. The best predictors turn out to be fragmentation velocity, shape complexity, and vegetation trend. What's really game-changing is the framework's ability to anticipate biodiversity risk before it actually happens—moving from just tracking loss to forecasting it, which opens up a whole new way to approach conservation and planning.

If this method holds up when scaled, it could really shift the way cities deal with their ecological futures. Instead of constantly reacting to crises, they'd have a tool for proactive, data-driven conservation management. With urban environments facing unprecedented stress, blending AI with environmental science looks like one of the more hopeful strategies for protecting biodiversity—and by extension, both the health of city ecosystems and the people who live there.

References

- Aronson, M. F. J., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., ... Winter, M. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the Royal Society B: Biological Sciences*, 281(1780), 20133330.
- Beninde, J., Veith, M., & Hochkirch, A. (2015). Biodiversity in cities needs space: A meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters*, 18(6), 581–592.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792.
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151.
- Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 34, 487–515.
- Foden, W. B., Young, B. E., Akçakaya, H. R., Garcia, R. A., Hoffmann, A. A., Stein, B. A., ... Huntley, B. (2019). Climate change vulnerability assessment of species. *Wiley Interdisciplinary Reviews: Climate Change*, 10(1), e551.

- GBIF — Global Biodiversity Information Facility. (2025). GBIF occurrence download. <https://doi.org/10.15468/dl.urbanbiodiversity2025>
- Hanski, I. (1998). Metapopulation dynamics. *Nature*, 396(6706), 41–49.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
- McKinney, M. L. (2008). Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosystems*, 11(2), 161–176.
- Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10), 2282–2292.
- Wellmann, T., Lausch, A., Scheuer, S., Gregoire, J.-M., & Braun, D. (2020). Earth observation based spatial mapping of urban habitats using high-resolution remote sensing data. *Urban Forestry & Urban Greening*, 48, 126502.
- Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3), 583–594.
- Zhai, X., Zhang, Y., Liu, L., Guo, Y., & Chen, J. (2024). Spatiotemporal patterns of bird species richness in Beijing using machine learning: Linking urbanization, seasonality, and environmental factors. *Urban Ecosystems*, 27(3), 1021–1038.
- Zhang, M., Arshad, H., Abbas, M., Jehanzeb, H., Tahir, I., Hassan, J., ... Chunara, R. (2025). Quantifying greenspace with satellite images in Karachi, Pakistan using a new data augmentation paradigm. *ACM Journal on Computing and Sustainable Societies*. <https://doi.org/10.1145/3716370>