Developing Accurate Data Sources for Environmental Social Government Metrics: Environmental Metrics in Vancouver

Sommer Harris Liam Bao Mengdi Wei Alice Shen Cynthia Cheng

Ezra MacDonald Yvonne Coady
m.coady@northeastern.edu

Abstract

This paper was developed as a result of a collection of students who worked toward developing a proof of concept for combining several types of data sources to create a more accurate 'Environmental' metric for Environmental Social Governance (ESG) metrics in the city of Vancouver. We review the various phases of thought in this project, and where the next steps are for continuing this proof of concept.

1. Introduction

Environmental Social Governance metrics are a way of measuring an organization's performance in terms of these three areas. They are used in similar ways as business metrics, to assess performance and risk. There are multitudes of ESG metrics and some common ones include greenhouse gas emissions, air and water pollution, biodiversity, business circularity, deforestation, recycling, and waste management. Our collective is invested in developing ESG metrics for the city of Vancouver, and cultivating the infrastructure to build robust, data driven metrics that give Vancouver a ground truth from which to act. As this is a large scope, this initial survey focuses on 'E', the environmental component.

The need for refinement of ESG metrics has been illustrated in various studies and real life examples. Atkins, et. al. 2022 analyzed responses from corporations regarding the resilience of ESG metrics during COVID-19. The feedback they received stated that COVID-19 highlighted the interconnected nature of 'E', 'S', and 'G'. It also highlighted the link between climate change and human rights issues. They identified that organizations saw environmental management as a priority across the board, and a facet with great opportunity, and great risk if ignored. They also reflected a need to develop better standardization and accuracy in ESG metrics [3].

We have drawn together various sources of data which could be combined to develop a more robust, more accurate environmental snapshot of the city of Vancouver. We propose combining Light Detection and Randing (LiDar) data, city zoning data, Normalizd Difference Vegetation Index (NDVI), and Skywatch data to answer the question 'How green is the city of Vancouver?'. This work has dramatic implications, fostering a ground truth from which to build benchmarks for organizations and cities.

2. Light Detection and Ranging Data

Light Detection and Ranging (LiDAR) data is a way to see the city in terms of 3d data. LiDAR is a remote sensing method that uses an active optical sensor and a laser in the infrared, visible, or ultra-violet spectrum. The laser transmits a light pulse and a receiver measures the reflected light. The distance to the object is calculated using the time it takes the light pulses to be reflected back and the speed of light. The 3-D model is generated using a .las file. Refer to Figures 1 and 2 for examples of images generated from LiDAR.

Data from LiDAR gives us a unique snapshot that can include important environmental information such as the height of the trees, which is not obtainable through our Satellite data.

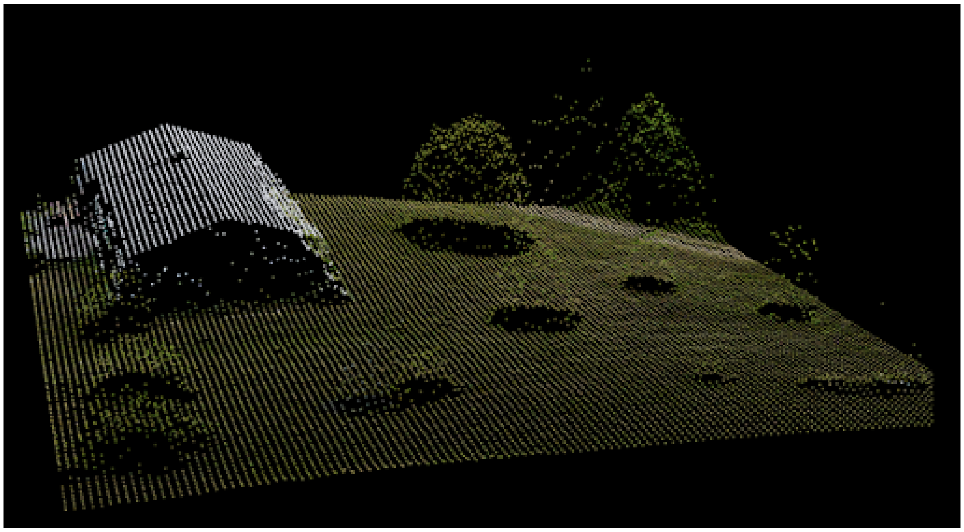

Figure 1. Example of LiDar data, build up of millions of data points. The model can be zoomed in or out or rotated.

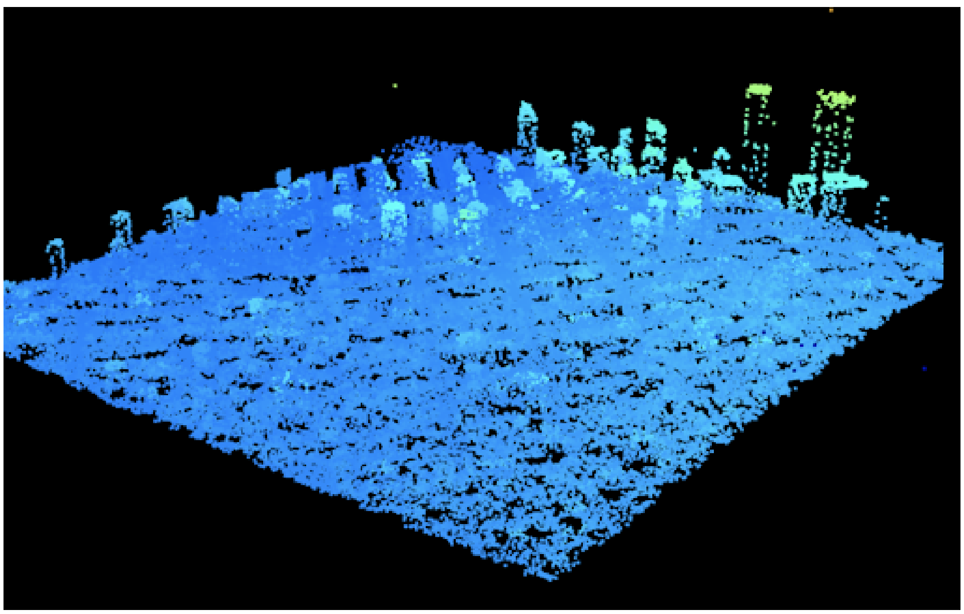

Figure 2. Example of LiDAR image for Vancouver based on 2018 data.

As of now, we currently only have 1 year of data, from 2018. However, if more years were collected we could show 3-D change over a period of time. With data from other cities, we can begin to develop benchmarks based on 3-D data to compare with other cities.

3. Map of City Zoning

City zoning data helps us to understand which green spaces are private and which are accessible to the public. In considering the interplay between environment, social issues, and governance, we know that green space impacts human health. However, the green space available to one class of person in a city may drastically differ from the green space available to someone of a different class. The questions arise: Does greenness matter to people who cannot access it? How do we measure the greenness of a city in an inclusive way? This data would also help us determine the level of access to green space on buildings- for example the top of the public library versus the top of a private apartment building.

Figure 3 explores percentage of green space that is parks and golf courses, using city zoning data overlaid on a map. The key displays a spectrum of 5 levels of access. We can note that the areas of Vancouver that contain forest (Pacific Spirit Park, Queen Elizabeth Park, Stanley Park) happen to be areas with the highest level of access to green space.

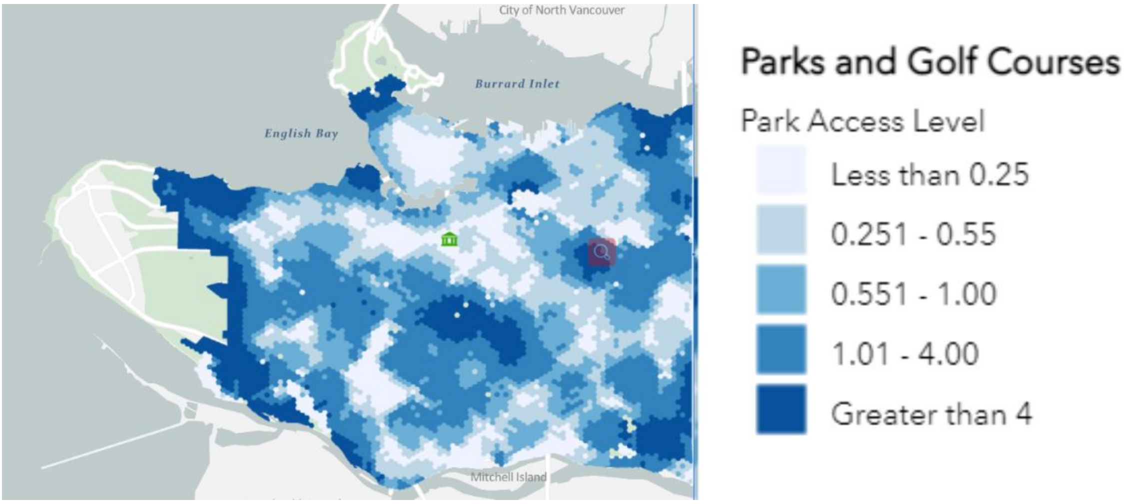

Figure 3. City zoning overlaid on a map to show range of private and public green space, generated from City of Vancouver VanMap.

4. Normalized Difference Vegetation Index

Figure 4. High resolution satellite photo (10m per pixel is the highest res available for free from Copernicus Open Access Hub)

The Normalized Difference Vegetation Index (NDVI) is a method to measure vegetation in a particular area. This method uses a high resolution satellite image, such as the one shown in Figure 4. We can retrieve different bands from the image (such as yellow or red), and then uses the image to calculate NDVI. The NDVI image shows the state of vegetation health based on how the plants reflect light at different bands in the light spectrum, as shown in Figure 5 [2]. The final result would be an image such as the one in Figure 6.

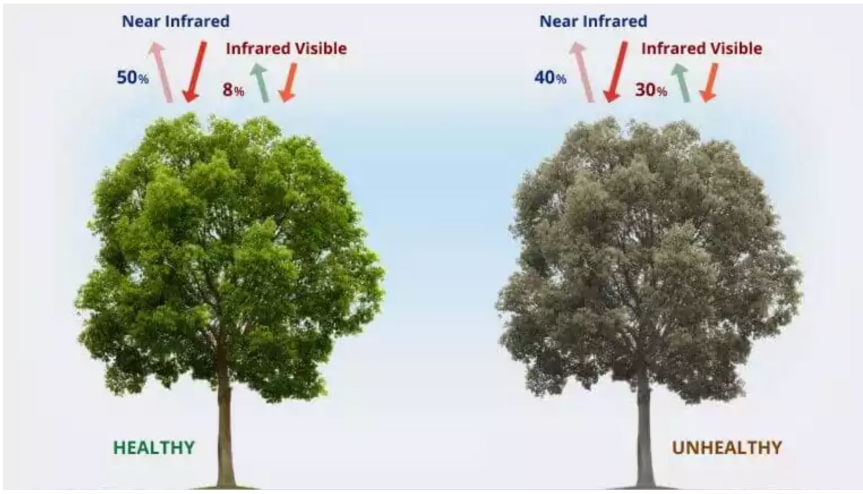

Figure 5. Image from Earth Daily Analytics demonstrating how light bands can show plant health.

Figure 6. Normalized Difference Vegetation Index (NDVI) for downtown Vancouver from 2022

We can use this data to compare vegetation in different areas. We can also combine it with the results of our other data analysis methods to build a more robust and accurate snapshot, and use a machine learning model to evaluate the environmental index of Vancouver.

4.1. Deep Learning Transformer Architecture Applications to Remote Sensing

We can also manipulate NDVI data with deep learning transformer. Refer to Figure 7, for another display of NDVI data. The transformer architecture was introduced in the famous 2017 paper Attention is All You Need, as a successor to recurrent and convolutional neural network layers. This model was originally applied to machine translation problems and was less computationally expensive and more accurate than state-of-the-art machine translation models [10]. Researchers were then able to successfully use this architecture to achieve state of the art results in imaging classification by performing some additional transformations on flattened strings of image inputs [5]. This layer architecture has since been used in a variety of model architectures ranging in their complexity - that have shown further promise for applications to imaging tasks [4], [6]. Remote sensing has a clear use case for advanced model architectures that perform classification, transformation, and segmentation tasks; and this model architecture can be used to contribute to the advancement of precision in this field [1].

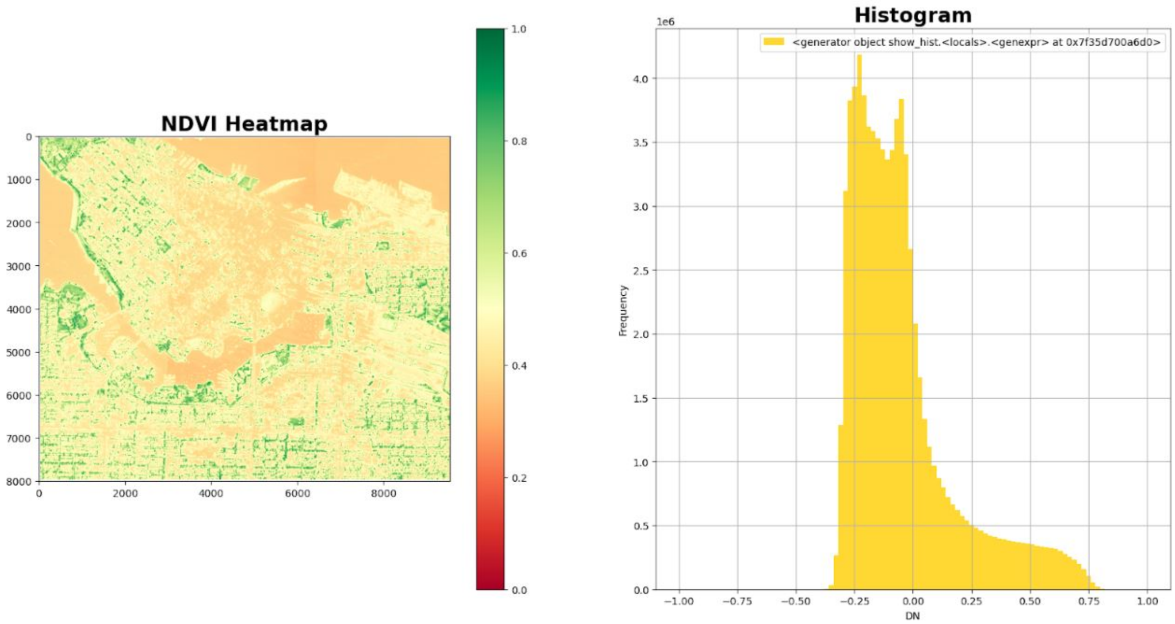

Figure 7. A heatmap for NDVI calculations, and Histogram displaying DN and Frequency Data

5. Skywatch data

We can use Skywatch data to visualize the building age and size for buildings in Vancouver. This data is relevant to environmental metrics because the age of a building tells us what year's codes that building adheres to. Often newer buildings will have greener building codes. We generated a map in Figure 8, using data from city of Vancouver's open data portal. The datasets we used were "Property Parcel Polygons," and "Property Tax Report." Property Parcel Polygons gives us a polygon on the map, the parcel identification number (pID), and location in latitude and longitude. Property Tax Report gives us the pID, which we can use to map to the Property Parcel Polygons dataset, the year the building was built, the improvement year, and some property value and tax information. To visualize this data, we retrieved the datasets as json and combined them in Tableau. The resulting image displays the year of the building as color (bluer for older buildings and redder for newer buildings). We can also use the information in Property Parcel Polygons to estimate the size of the building, which can be used with building policies to measure greenness of buildings by percentage.

Figure 8. Map generated in Tableau that gives us building age. Bluer buildings are older and redder buildings are newer.

6. Conclusion

As stated in the discussion for each datatype, these sources reveal unique information that can be combined in meaningful ways to build a robust and accurate approach for measuring greenness. The implications of building this multifaceted data picture are the capacity to set objective benchmarks, track progress toward sustainability, and measure the greenness of one city or organization against another. These data approaches also support meaningful intersections for overlaps between 'E' m 'S', and 'G' such as mental health, and private versus public access to green space. Combining multiple data sources also gives us the capacity to properly contextualize machine learning models. For example, if machine learning is used to fill in data gaps from one source, the other sources can be used to verify the accuracy of generated data. This data picture can also support 21st century innovations in the path toward sustainable economies. Doughnut economics is just one model that reflects the drive toward healthy, sustainable economies in the 21st century. It frames economic goals in terms of staying within our 'ecological ceiling' and meeting the baseline of our 'social foundation'. Figure 9 is Kate Raworth's visualization of the doughnut economy[8].

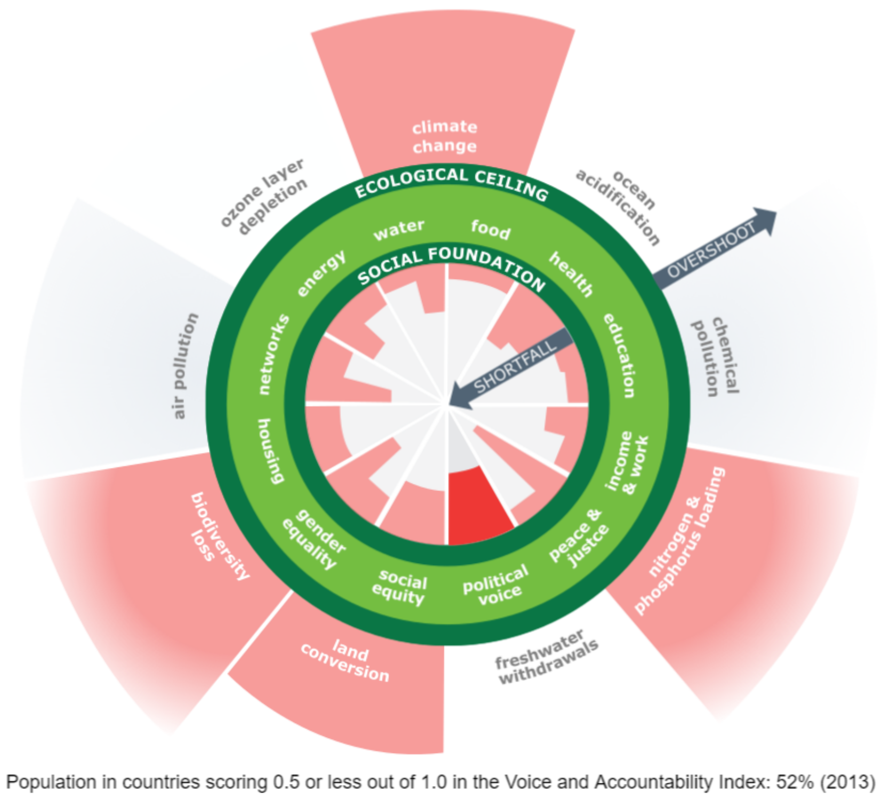

Figure 9. Doughnut Economy as conceived by Kate Raworth

Turner and Wills, 2022 express the difficulty in adapting the doughnut model to local contexts while aligning with larger scale applications. They identify the following directions for applications of doughnut economics: develop adaptiveness that allows governments to look past institutional path dependencies, build momentum for local applications that frames challenges in ways that are non partisan and generate solutions, research on how local communities can be mobilized in their relationship to place through local institutions, trans-disciplinary approaches where scientists, decision-makers, and citizens collaborate [9].

Our proposed data snapshot of Vancouver supports the aforementioned challenges in the following ways. An objective dataset and data driven recommendations gives governments a path to follow as an alternative to institutional path dependencies. Data can also support nonpartisan solutions through giving the various stakeholders a shared ground on which to make decisions. The snapshot can help build momentum for local applications, because each snapshot would be highly unique to the location in question. It can help mobilize local communities by identifying the locations that already have the highest level of access and inclusion, and learning from those successes. Lastly, our proposed snapshot demonstrates an avenue for interdisciplinary collaboration; we have identified avenues where the unique technical skills of computer scientists and the creativity of out of the box thinkers can be combined, and where those gifts can be greater than the sum of their parts, in service to our city at large.

One of the biggest challenges moving forward is the question of how we measure greenness in a diversity of bio-regions. Our suggested measurements such as vegetation cover, and plant health only work for Vancouver because we are a temperate rain forest. If we were in a desert, the approach to measuring greenness might look dramatically different. More work can be done on refining what measurement elements would be the same from city to city and what elements would be different.

A second challenge pertains to the question of data sovereignty and indigenous knowledge. Just because we have the capacity to develop a data snapshot of greenness on a particular piece of land, that does not imply that it is the most ethical or effective way to pursue sustainability. There is opportunity to develop decolonized approaches to ESG that respect data sovereignty and look to indigenous leadership. Recent analysis of ESG practices reflect that there is not enough consideration of indigenous truth and reconciliation in many corporate ESG approaches [7]. The development of a data snapshot could be an opportunity to introduce new and better practices.

Next steps moving forward are to prototype quantifiable metrics for 'E' and develop a testing suite for them, generating preliminary reports on potential data sources for 'S' and 'G', and conducting a literature review toward best practices for generating combined ESG metrics.

7. Acknowledgements

We would like to thank Northeastern University, Will Parkinson at Earth Data Analytics, Dr Derek Jacoby from the Digital Supercluster/National Research Council of Canada/University of Victoria and Nina Abizadeh from Vancouver City Studio for their support.

References

[1] A. A. Aleissaee, A. Kumar, R. M. Anwer, S. Khan, H. Cholakkal, G.-S. Xia, et al. Transformers in remote sensing: A survey. *arXiv preprint arXiv:2209.01206*, 2022.

[2] E. D. Analytics. Ndvi mapping in agriculture, index formula, and uses, Sep 2022.

[3] J. Atkins, F. Doni, A. Gasperini, S. Artuso, I. La Torre, and L. Sorrentino. Exploring the effectiveness of sustainability measurement: which esg metrics will survive covid-19? *Journal of Business Ethics*, pages 1–18, 2022.

[4] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar. Masked-attention mask transformer for universal image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1290–1299, 2022.

[5] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[6] K. He, C. Gan, Z. Li, I. Rekik, Z. Yin, W. Ji, Y. Gao, Q. Wang, J. Zhang, and D. Shen. Transformers in medical image analysis: A review. *arXiv preprint arXiv:2202.12165*, 2022.

[7] D. Lawrence. Indigenous action plans lack prominence in corporate esg efforts, May 2022.

[8] K. Raworth. Doughnut: Kate raworth, Sep 2020.

[9] R. A. Turner and J. Wills. Downscaling doughnut economics for sustainability governance. *Current Opinion in Environmental Sustainability*, 56:101180, 2022.

[10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.