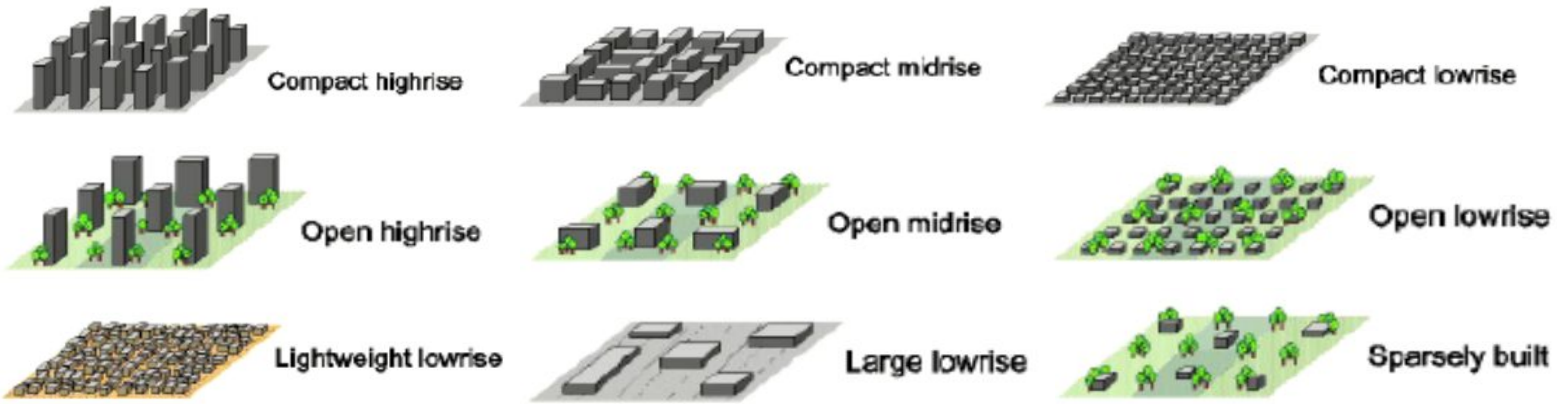
Leveraging machine learning to produce a cost effective national building height map of Ireland to create local climate zones

Eoghan Keany, Geoffrey Bessardon and Emily Gleeson

Motivation

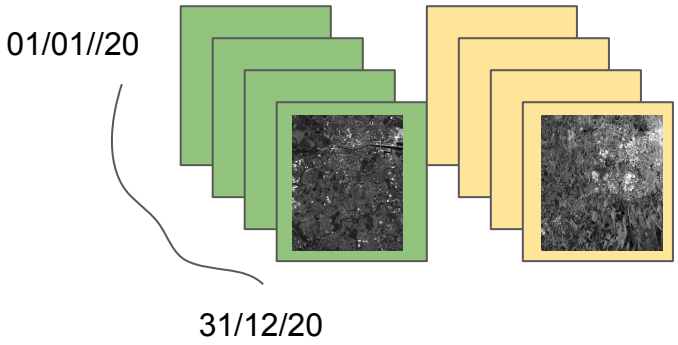
- Met eireann has recently produced a land cover map of Ireland using machine learning (Walsh et al. 2021) based on the Corine map.
- To make this map operational Met Eireann needs to be able to correctly classify urban areas. However in order to classify urban climate zones we need both the urban density and building heights.



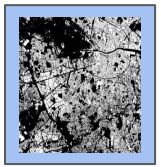
Data Source

- 3 distinct open source Datasets are needed to create this map.
- Time series data from both Sentinel 1 and 2 are used as our X values, the 2012 building heights map only covers dublin city centre and is our Y value for training the regression model.
- The 2015 European Settlement Map is used to mask our predictions to locations with actual buildings.

Sentinel 1 & 2
(01/01/20 -> 31/12/20)

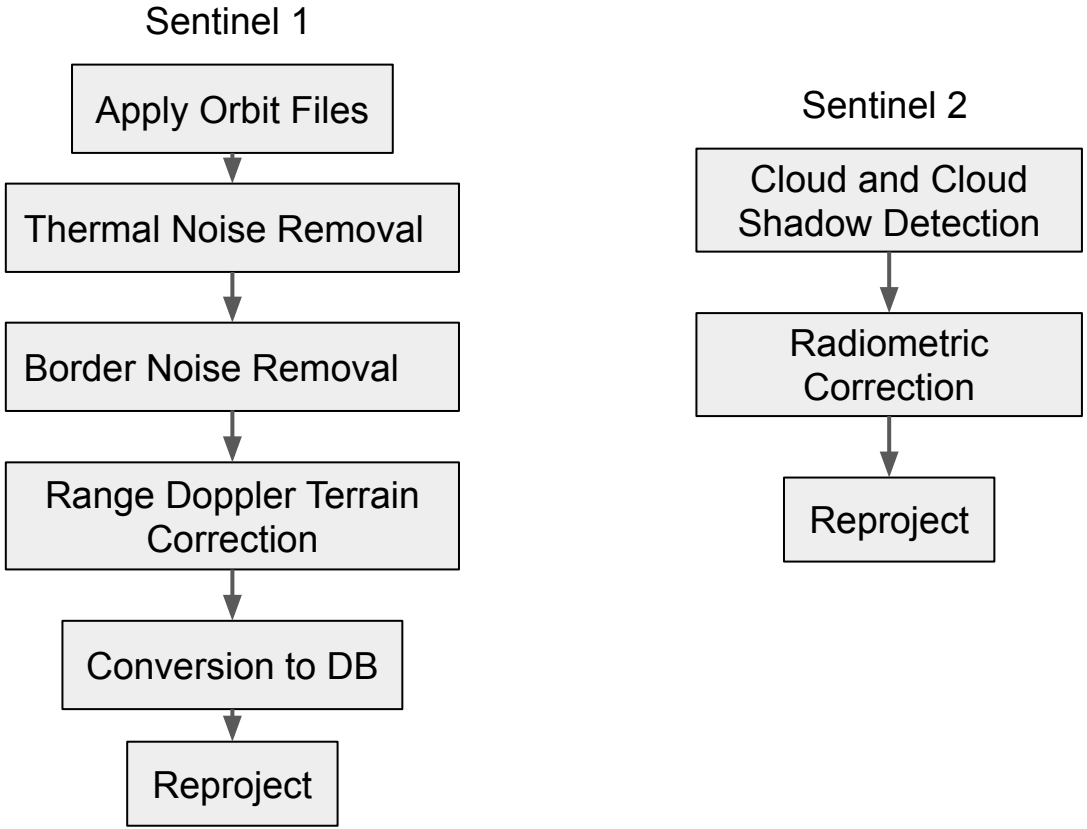


2015 European Settlement Map
(Copernicus)



2012 Building Heights
(Copernicus)

Pre-Processing Sentinel Scenes



Feature Engineering

1. Spectral

- Six new spectral indices (TCG, TCW, NDW, NDVI,NDB, TCB) were derived from the existing 10 Sentinel 2 Bands.
- In total we have 18 bands including both the vertical and horizontal polarizations from the Sentinel-1 constellation.

2. Temporal

- Thirteen statistical aggregate features were created. For each pixel in the image we calculate the mean, max, interquartile range … etc over for the entire time period.
- These features are not only robust and complete, but they capture seasonal patterns in reflectance, and shadows which are important features for determining building heights.

3. Spatial

- A structuring element with a 50m radius was used to calculate morphological operators. These operations capture the spatial context of the point of interest.

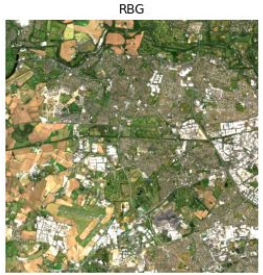
Modelling

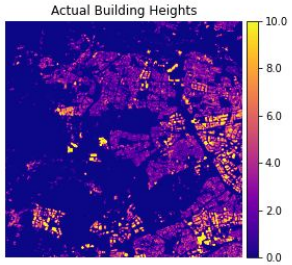
- After all of the pre processing and feature engineering, the dataset contained 1,638 features that can be used to train a regression model of over the capital city (where the true building heights are known)
- Before modelling the data was converted into a tabular format and the BorutaShap feature selection method was applied which returned 64 features, which will be used for further modelling.
- Three different types of machine learning models were applied to the data. These models can be categorized as tabular, cnn and encoder decoder. Each model was validated using 10 fold cross validation with RMSE, WRMSE metrics.

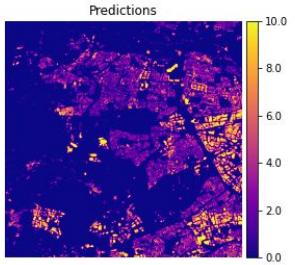
Model	RMSE	WRMSE	Earth Mover's Distance	TIME (Mins)
Ridge Regression	2.23	1.74	68.12	102
XGBoost Poisson	2.12	1.55	125.2	73
CNN	1.92	1.42	56.67	1,380

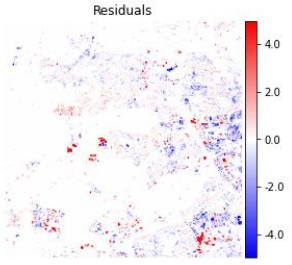
Resulting images

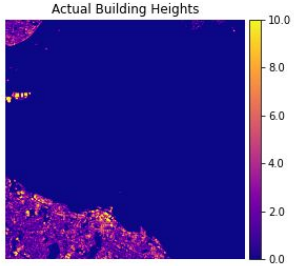
- Below are two out of sample 10km x 10km scenes with the actual, predicted and residuals.
- We can see the predictions histogram does an overall good job of fitting the underlying distribution despite over predicting the average building height.

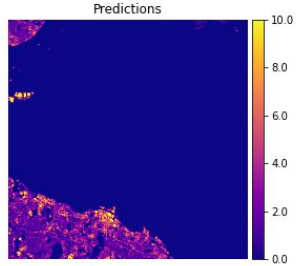


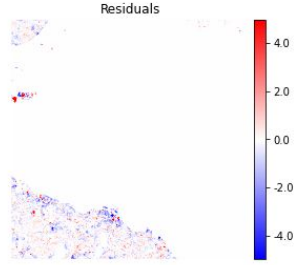


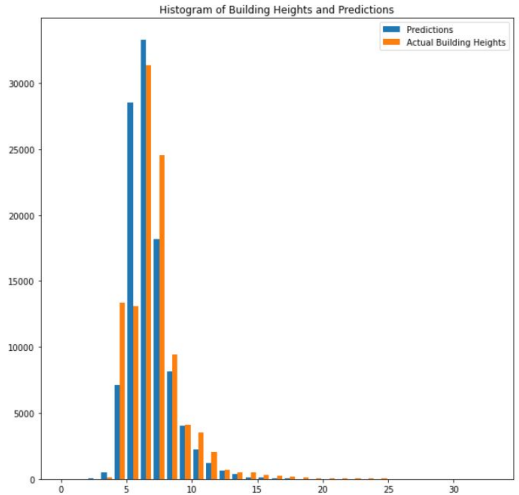








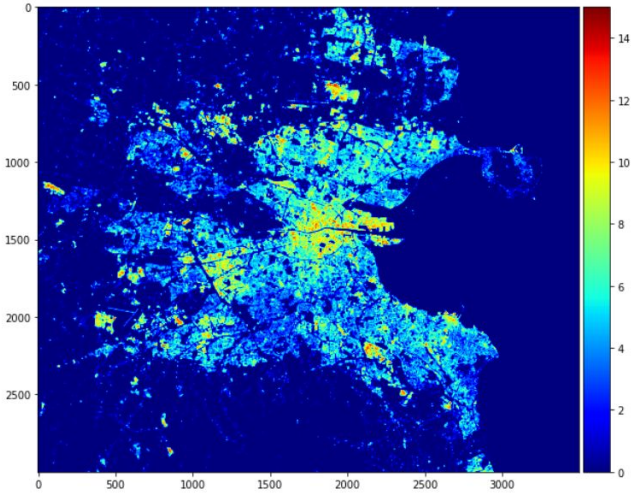




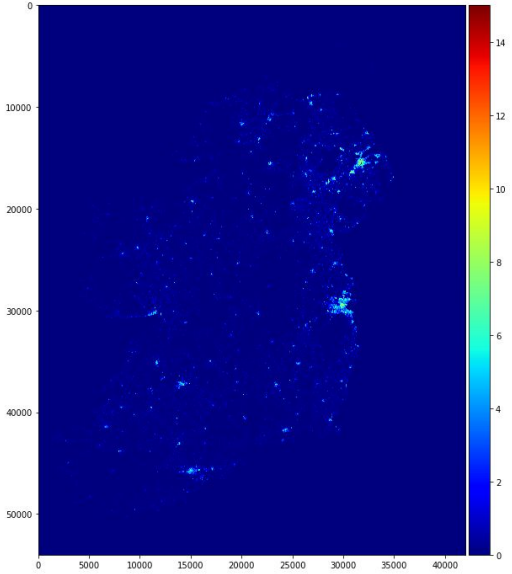
Final Map

- The resultant map is a 10m raster image that contains the building heights of Ireland.
- The color bar of 0-15 was chosen as this gave the best contrast as any building taller than that are sparse and stretch the colors.

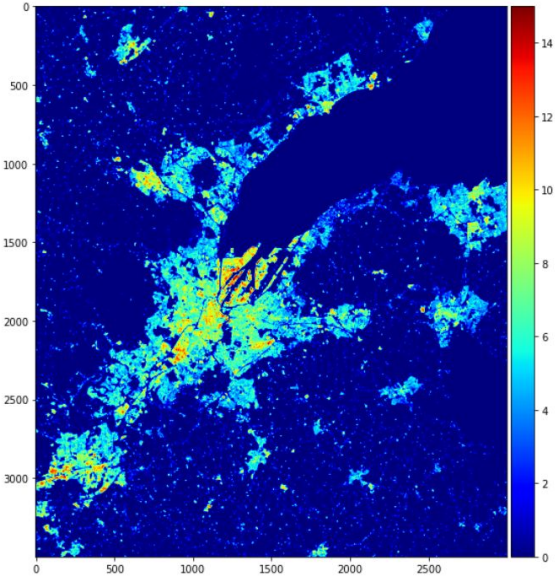
Dublin City



Entire Map



Belfast City



Next Steps

1. Produce an urban density map.
2. Use the density and building height map to classify local climate zones and update our existing land cover map.
3. Test the new land cover map with the correct climate zones in the Harmonie-Arome weather model.

References

1. Perera, Narein & Emmanuel, R. & Mahanama, P.. (2013). Projected urban development, changing 'Local Climate Zones' and relative warming effects in Colombo, Sri Lanka.
2. Frantz D, Schug F, Okujeni A, Navacchi C, Wagner W, van der Linden S, Hostert P. National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series. Remote Sensing of Environment. 2021 Jan;252:112128.
3. Walsh E, Bessardon G, Gleeson E, Ulmas P. Using machine learning to produce a very high resolution land-cover map for Ireland. Advances in Science and Research. 2021 May 11;18:65-87.
4. https://www.copernicus.eu/en
5. https://github.com/davidfrantz/force
6. https://sentinel.esa.int/web/sentinel/toolboxes