Spectral Separability of Local Climate Zones in Landsat 8 Imagery and its Application in WUDAPT Framework

Parth Bansal

▶ **To cite this version:**

Parth Bansal. Spectral Separability of Local Climate Zones in Landsat 8 Imagery and its Application in WUDAPT Framework. 2019. hal-02265553

Spectral Separability of Local Climate Zones in Landsat 8 Imagery and its Application in WUDAPT Framework

Parth Bansal

Graduate School of Environmental Studies, Seoul National University

Abstract: Local Climate Zones (LCZs) classification of urban area under World Urban Database and Portal Tool (WUDAPT) framework has amassed about 100 regions' LCZ catalog by an international collaboration of researchers. With Landsat 8 (L8) being the current standard of imagery for most of the LCZ classification process, the question of ability to distinguish LCZs from the spectral range of L8 imagery need to be answered. Using sample of training areas, Jeffries-Matusita distance are computed as the measure of separability to investigate if any pairs of LCZ class consistently show similar spectral signature and it's utility to predict the quality of final LCZ classification.

1. Introduction

Based on the LCZ scheme developed by Stewart & Oke (2012), WUDAPT has been building the database of LCZ classification of cities to address the need of information on urban form and function of cities (Mills et al., 2015). Currently in 'level 0' state, global collaboration has produced LCZ maps for the number of cities using user developed regions of interests (ROIs) (training areas (TAs)) which are used to train a classifier, with Random Forest being the choice of supervised classifier under WUDAPT. Detailed workflow and current progress have been provided by Bechtel et al. (2019).

1.1 Quality control in WUDAPT

To measure the quality of LCZ classification result, two-step procedure is followed. Firstly, ROIs are bootstrapped into two equal parts with LCZ class based stratification, and 25 different random forests are trained using half of ROIs. The remaining half is used for generating confusion matrix to determine if they were correctly classified. If the average accuracy exceeds 50%, the classification is labeled as 'passed'. Secondly, manual review is done in which an experienced reviewer visually checks the classification in comparison with high resolution imagery. The classification is then assigned three alternative labels; accepted (no improvement required) / minor (minor refinement required) / major (significant disagreement between classification and ground truth).

To attain high quality results, Bechtel et al.(2019) recommends iterating between the creations of training areas and running the classifier. Hence, by measuring the spectral separability of ROI pairs, researchers can identify ROI which may be causing the classification result to not achieve quality requirement.

1.2 Spectral Separability

Essentially, to differentiate between two classes, any classifier uses variance in spectral response across the bands. Thus a pair of LCZ (for example, Compact High Rise and Compact Mid-rise) while being structurally different, can have similar spectral response, due to similar material. Thus the integration of surface model has been proposed as a solution to this issue in future 'levels' of WUDAPT (Benjamin et al., 2015). Hence there is a need to investigate if the current classification under WUDAPT has this issue across the study area and if yes, which pairs of LCZs lacks separability and its effects on final classification quality.

Several measures of spectral separability have been developed to statistically measure the distance between pairs of signature with Transformed Divergence and Jeffries-Matusita (JM) Distance being more popular and available in commercial and open source remote sensing packages. For this study, JM Distance is used which is defined as:

$$J_{ij} = 2(1 - e^{-B}) \quad (1)$$

$$B = \frac{1}{8}(u_i - u_j)^T \left(\frac{C_i + C_j}{2}\right)^{-1} (u_i - u_j) + \frac{1}{2}\ln\left(\frac{|(C_i + C_j)/2)|}{\sqrt{|C_i C_j|}}\right) \text{ (Jensen, 1996)} \quad (2)$$

Where:

i and j = two classes being compared

u_i and u_j = spectral signature vector of i and j

C_i and C_j = Covariance Matrix of i and j

The negative exponent in (1) means a decrease in weight as spectral separation increase and JM distance is asymptotic to 2.0. In other words, the value of 2.0 indicates that the signature can be separated with 100% accuracy with 1.8 often used as the threshold for acceptable level of separation (Yang et al., 2017).

2.0 Methodology

For the first step, metadata for each region in WUDAPT were collected and compiled. Since the objective was to determine LCZ pair separability, we filtered out only those regions which have been accepted after passing bootstrapping and manual review. Secondly, although utilization of Sentinel, both Sentinel-1 (synthetic aperture radar (SAR)) and Sentinel-2 (optical) imagery has been explored (Bechtel et al., 2016) (Kaloustian et al., 2017), all but one region in WUDAPT currently uses Landsat 8 as imagery source. Hence the one region developed using Sentinel was excluded from the set. In the third step, 5 sample regions were selected for each review status (Accepted/Minor/Major) since the objective was to study the relation between JM distance and quality of final classification. Since both 'accepted' and 'major' category has only 5 entries, the sampling was only required for 'minor' status and was done using rand and index function in Excel.

For selected regions in step 3, KML/KMZ training files were converted into ROI files (with minor edits to the LCZ field for some samples to streamline the terminologies). Landsat 8 images for the dates specified by the researchers in metadata were downloaded. Individual bands used (B1-7, 11,12) from multi date imagery (when specified in metadata) were combined using 'Layer Stacking' tool and associated ROI were used to calculate JM distance using ENVI 5.3, for the LCZ pairs specified for a given region. ENVI is a commercial image processing tool by Harris Geospatial, however equivalent tools are also available for open source QGIS in Semi-Automatic Classification (SCP) plugin.

3.0 Results

With the objective of studying spectral separability of LCZ pairs under the best classification scenario, the results of JM distance for 5 regions with accepted status are shown fig 1. On the other hand fig 2 shows the distribution of JM distance with respect to all review status and classification score. Key observations:

1) No LCZ pair has JM distance bellow 1.8 across all 5 regions with 'accepted' review status. This is significant because it shows that with a careful and iterative selection of training, it can be possible for all LCZ pairs in a study area to have spectral separability.

2) That being said, some LCZ pairs seem to have more instances of spectral similarity. Especially LCZ pairs; compact high rise – compact low rise, open high rise – open midrise / open low rise, open midrise – open low rise, large low rise – heavy industry seems to have low spectral separability in more than half of five regions. This indicates that, while the horizontal dimension of urban form is separatable, vertical differences may not produce separable spectral signatures. However, natural and urban classes and natural LCZ classes among themselves have good separability across majority of cases.

3) Fig 2 indicates to a correlation between the percentage of LCZ pairs with JM distance greater than 1.8 and the final quality of classification obtained.
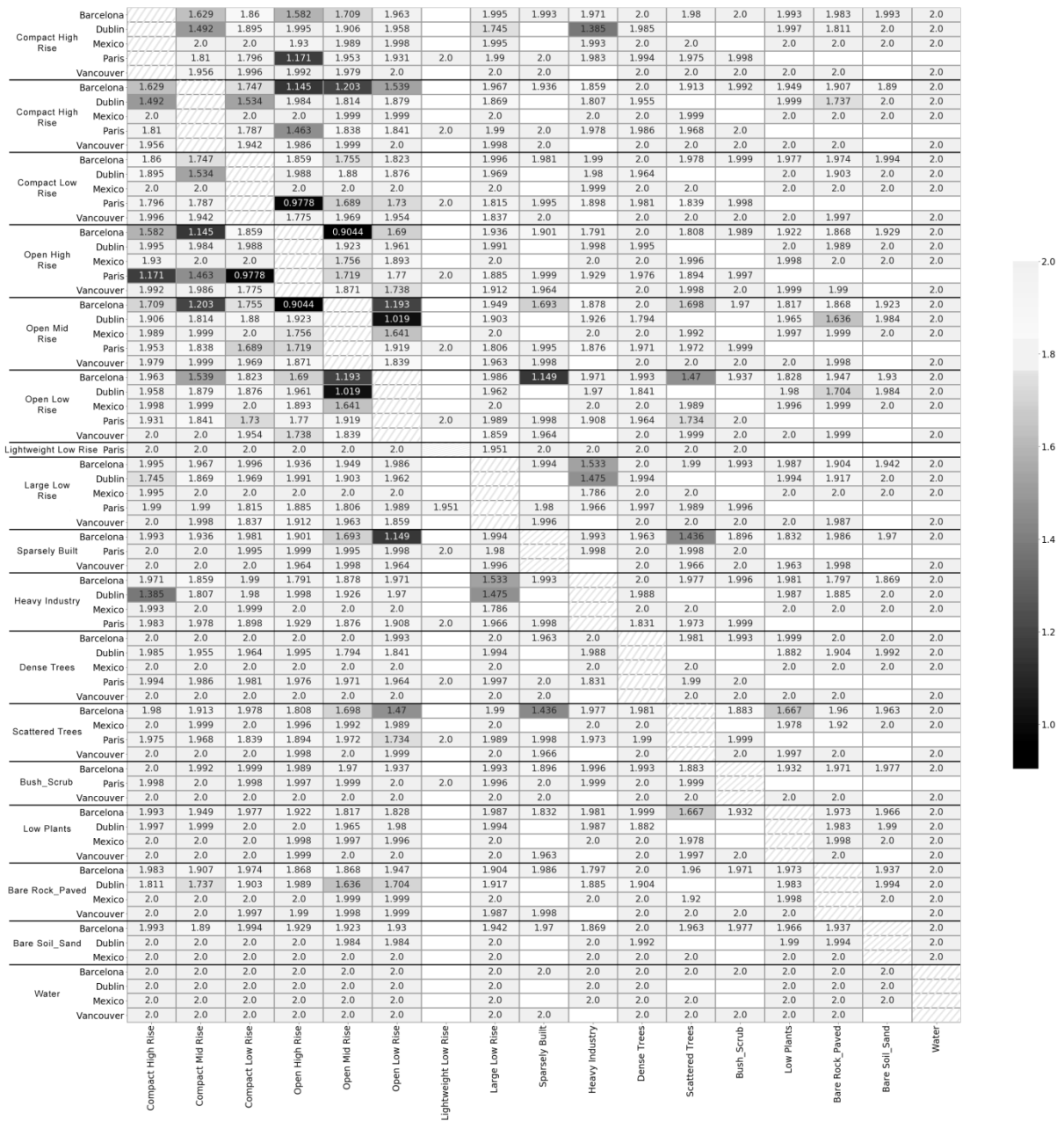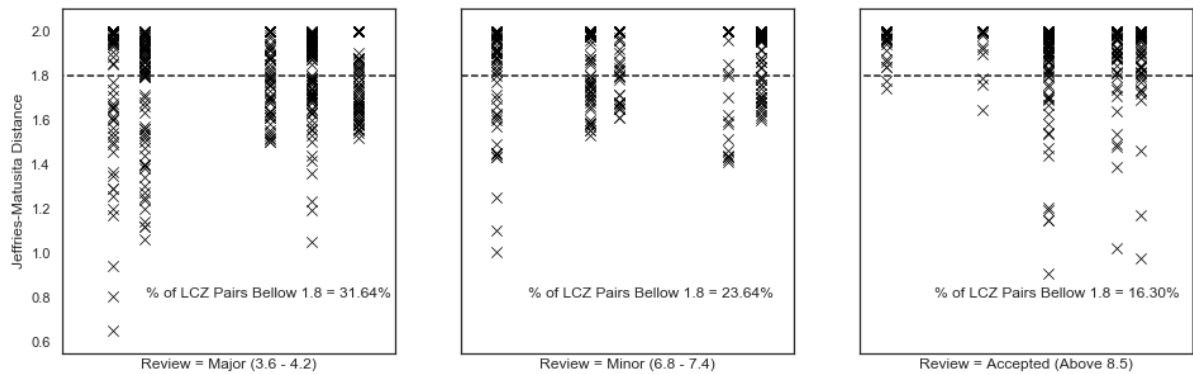
Figure 1 table — JM Distance for each LCZ pair by region. Row groups are the reference LCZ (with rows per region: Barcelona, Dublin, Mexico, Paris, Vancouver); columns are the compared LCZ. The diagonal (self-pair) cells are hatched/blank.

LCZ	Region	Compact High Rise	Compact Mid Rise	Compact Low Rise	Open High Rise	Open Mid Rise	Open Low Rise	Lightweight Low Rise	Large Low Rise	Sparsely Built	Heavy Industry	Dense Trees	Scattered Trees	Bush_Scrub	Low Plants	Bare Rock_Paved	Bare Soil_Sand	Water
Compact High Rise	Barcelona		1.629	1.86	1.582	1.709	1.963		1.995	1.993	1.971	2.0	1.98	2.0	1.993	1.983	1.993	2.0
	Dublin		1.492	1.895	1.995	1.906	1.958		1.745		1.385	1.985			1.997	1.811	2.0	2.0
	Mexico		2.0	2.0	1.93	1.989	1.998		1.995		1.993	2.0	2.0		2.0	2.0	2.0	2.0
	Paris		1.81	1.796	1.171	1.953	1.931	2.0	1.99	2.0	1.983	1.994	1.975	1.998				
	Vancouver		1.956	1.996	1.992	1.979	2.0		2.0	2.0		2.0	2.0	2.0	2.0	2.0		2.0
Compact Mid Rise	Barcelona	1.629		1.747	1.145	1.203	1.539		1.967	1.936	1.859	2.0	1.913	1.992	1.949	1.907	1.89	2.0
	Dublin	1.492		1.534	1.984	1.814	1.879		1.869		1.807	1.955			1.999	1.737	2.0	2.0
	Mexico	2.0		2.0	2.0	1.999	1.999		2.0		2.0	2.0	1.999		2.0	2.0	2.0	2.0
	Paris	1.81		1.787	1.463	1.838	1.841	2.0	1.99	2.0	1.978	1.986	1.968	2.0				
	Vancouver	1.956		1.942	1.986	1.999	2.0		1.998	2.0		2.0	2.0	2.0	2.0	2.0		2.0
Compact Low Rise	Barcelona	1.86	1.747		1.859	1.755	1.823		1.996	1.981	1.99	2.0	1.978	1.999	1.977	1.974	1.994	2.0
	Dublin	1.895	1.534		1.988	1.88	1.876		1.969		1.98	1.964			2.0	1.903	2.0	2.0
	Mexico	2.0	2.0		2.0	2.0	2.0		2.0		1.999	2.0	2.0		2.0	2.0	2.0	2.0
	Paris	1.796	1.787		0.9778	1.689	1.73	2.0	1.815	1.995	1.898	1.981	1.839	1.998				
	Vancouver	1.996	1.942		1.775	1.969	1.954		1.837	2.0		2.0	2.0	2.0	2.0	1.997		2.0
Open High Rise	Barcelona	1.582	1.145	1.859		0.9044	1.69		1.936	1.901	1.791	2.0	1.808	1.989	1.922	1.868	1.929	2.0
	Dublin	1.995	1.984	1.988		1.923	1.961		1.991		1.998	1.995			2.0	1.989	2.0	2.0
	Mexico	1.93	2.0	2.0		1.756	1.893		2.0		2.0	2.0	1.996		1.998	2.0	2.0	2.0
	Paris	1.171	1.463	0.9778		1.719	1.77	2.0	1.885	1.999	1.929	1.976	1.894	1.997				
	Vancouver	1.992	1.986	1.775		1.871	1.738		1.912	1.964		2.0	1.998	2.0	1.999	1.99		2.0
Open Mid Rise	Barcelona	1.709	1.203	1.755	0.9044		1.193		1.949	1.693	1.878	2.0	1.698	1.97	1.817	1.868	1.923	2.0
	Dublin	1.906	1.814	1.88	1.923		1.019		1.903		1.926	1.794			1.965	1.636	1.984	2.0
	Mexico	1.989	1.999	2.0	1.756		1.641		2.0		2.0	2.0	1.992		1.997	1.999	2.0	2.0
	Paris	1.953	1.838	1.689	1.719		1.919	2.0	1.806	1.995	1.876	1.971	1.972	1.999				
	Vancouver	1.979	1.999	1.969	1.871		1.839		1.963	1.998		2.0	2.0	2.0	2.0	1.998		2.0
Open Low Rise	Barcelona	1.963	1.539	1.823	1.69	1.193			1.986	1.149	1.971	1.993	1.47	1.937	1.828	1.947	1.93	2.0
	Dublin	1.958	1.879	1.876	1.961	1.019			1.962		1.97	1.841			1.98	1.704	1.984	2.0
	Mexico	1.998	1.999	2.0	1.893	1.641			2.0		2.0	2.0	1.989		1.996	1.999	2.0	2.0
	Paris	1.931	1.841	1.73	1.77	1.919		2.0	1.989	1.998	1.908	1.964	1.734	2.0				
	Vancouver	2.0	2.0	1.954	1.738	1.839			1.859	1.964		2.0	1.999	2.0	2.0	1.999		2.0
Lightweight Low Rise	Paris	2.0	2.0	2.0	2.0	2.0	2.0		1.951	2.0	2.0	2.0	2.0	2.0				
Large Low Rise	Barcelona	1.995	1.967	1.996	1.936	1.949	1.986			1.994	1.533	2.0	1.99	1.993	1.987	1.904	1.942	2.0
	Dublin	1.745	1.869	1.969	1.991	1.903	1.962				1.475	1.994			1.994	1.917	2.0	2.0
	Mexico	1.995	2.0	2.0	2.0	2.0	2.0				1.786	2.0	2.0		2.0	2.0	2.0	2.0
	Paris	1.99	1.99	1.815	1.885	1.806	1.989	1.951		1.98	1.966	1.997	1.989	1.996				
	Vancouver	2.0	1.998	1.837	1.912	1.963	1.859			1.996		2.0	2.0	2.0	1.987			2.0
Sparsely Built	Barcelona	1.993	1.936	1.981	1.901	1.693	1.149		1.994		1.993	1.963	1.436	1.896	1.832	1.986	1.97	2.0
	Paris	2.0	2.0	1.995	1.999	1.995	1.998	2.0	1.98		1.998	2.0	1.998	2.0				
	Vancouver	2.0	2.0	2.0	1.964	1.998	1.964		1.996			2.0	1.966	2.0	1.963	1.998		2.0
Heavy Industry	Barcelona	1.971	1.859	1.99	1.791	1.878	1.971		1.533	1.993		2.0	1.977	1.996	1.981	1.797	1.869	2.0
	Dublin	1.385	1.807	1.98	1.998	1.926	1.97		1.475			1.988			1.987	1.885	2.0	2.0
	Mexico	1.993	2.0	1.999	2.0	2.0	2.0		1.786			2.0	2.0		2.0	2.0	2.0	2.0
	Paris	1.983	1.978	1.898	1.929	1.876	1.908	2.0	1.966	1.998		1.831	1.973	1.999				
Dense Trees	Barcelona	2.0	2.0	2.0	2.0	2.0	1.993		2.0	1.963	2.0		1.981	1.993	1.999	2.0	2.0	2.0
	Dublin	1.985	1.955	1.964	1.995	1.794	1.841		1.994		1.988				1.882	1.904	1.992	2.0
	Mexico	2.0	2.0	2.0	2.0	2.0	2.0		2.0		2.0		2.0		2.0	2.0		2.0
	Paris	1.994	1.986	1.981	1.976	1.971	1.964	2.0	1.997	2.0	1.831		1.99	2.0				
	Vancouver	2.0	2.0	2.0	2.0	2.0	2.0		2.0	2.0			2.0	2.0	2.0	2.0		2.0
Scattered Trees	Barcelona	1.98	1.913	1.978	1.808	1.698	1.47		1.99	1.436	1.977	1.981		1.883	1.667	1.96	1.963	2.0
	Mexico	2.0	1.999	2.0	1.996	1.992	1.989		2.0		2.0	2.0			1.978	1.92	2.0	2.0
	Paris	1.975	1.968	1.839	1.894	1.972	1.734	2.0	1.989	1.998	1.973	1.99		1.999				
	Vancouver	2.0	2.0	2.0	1.998	2.0	1.999		2.0	1.966		2.0		2.0	1.997	2.0		2.0
Bush_Scrub	Barcelona	2.0	1.992	1.999	1.989	1.97	1.937		1.993	1.896	1.996	1.993	1.883		1.932	1.971	1.977	2.0
	Paris	1.998	2.0	1.998	1.997	1.999	2.0	2.0	1.996	2.0	1.999	2.0	1.999					
	Vancouver	2.0	2.0	2.0	2.0	2.0	2.0		2.0	2.0		2.0	2.0		2.0	2.0		2.0
Low Plants	Barcelona	1.993	1.949	1.977	1.922	1.817	1.828		1.987	1.832	1.981	1.999	1.667	1.932		1.973	1.966	2.0
	Dublin	1.997	1.999	2.0	2.0	1.965	1.98		1.994		1.987	1.882				1.983	1.99	2.0
	Mexico	2.0	2.0	2.0	1.998	1.997	1.996		2.0		2.0	2.0	1.978			1.998	2.0	2.0
	Vancouver	2.0	2.0	2.0	1.999	2.0	2.0		2.0	1.963		2.0	1.997	2.0		2.0		2.0
Bare Rock_Paved	Barcelona	1.983	1.907	1.974	1.868	1.868	1.947		1.904	1.986	1.797	2.0	1.96	1.971	1.973		1.937	2.0
	Dublin	1.811	1.737	1.903	1.989	1.636	1.704		1.917		1.885	1.904			1.983		1.994	2.0
	Mexico	2.0	2.0	2.0	2.0	1.999	1.999		2.0		2.0	2.0	1.92		1.998		2.0	2.0
	Vancouver	2.0	2.0	1.997	1.99	1.998	1.999		1.987	1.998		2.0	2.0	2.0	2.0			2.0
Bare Soil_Sand	Barcelona	1.993	1.89	1.994	1.929	1.923	1.93		1.942	1.97	1.869	2.0	1.963	1.977	1.966	1.937		2.0
	Dublin	2.0	2.0	2.0	2.0	1.984	1.984		2.0		2.0	1.992			1.99	1.994		2.0
	Mexico	2.0	2.0	2.0	2.0	2.0	2.0		2.0		2.0	2.0	2.0		2.0	2.0		2.0
Water	Barcelona	2.0	2.0	2.0	2.0	2.0	2.0		2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	
	Dublin	2.0	2.0	2.0	2.0	2.0	2.0		2.0		2.0	2.0			2.0	2.0	2.0	
	Mexico	2.0	2.0	2.0	2.0	2.0	2.0		2.0		2.0	2.0	2.0		2.0	2.0	2.0	
	Vancouver	2.0	2.0	2.0	2.0	2.0	2.0		2.0	2.0		2.0	2.0	2.0	2.0	2.0		

Figure 1 JM Distance for each LCZ pair in regions with review status of accepted.

Figure 2 (three scatter panels of Jeffries-Matusita Distance with dashed line at 1.8):

- Review = Major (3.6 - 4.2): % of LCZ Pairs Bellow 1.8 = 31.64%
- Review = Minor (6.8 - 7.4): % of LCZ Pairs Bellow 1.8 = 23.64%
- Review = Accepted (Above 8.5): % of LCZ Pairs Bellow 1.8 = 16.30%

Figure 2 Distribution of LCZ Pair JM Distance with respect to classification score and review status.

4.0 Conclusion

The investigation into spectral separability allows the researcher to examine the possible LCZ pairs which might be causing classifier to produce unsatisfactory results. It shows that while spectral separability issues persist even in classification deemed acceptable, no pairs of LCZs are inseparable across all regions. However, the percentage of pairs bellow acceptable level of separability is a strong indicator of final classification result quality. Thus, it would be beneficial for researchers following WUDAPT methodology to add this step into their workflow while iterating between TAs modifications and running classifier.

Declarations of interest
None

References:

Bechtel, B. et al., 2019. Generating WUDAP Level 0 data - Current status of production and evaluation. *Urban Climate*, 27, pp.24-45.

Bechtel, B., See, L., Mills, G. & Foley, M., 2016. Classification of Local Climate Zones Using SAR and Multispectral Data in an Arid Environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(7), pp.3097-105.

Benjamin, B. et al., 2015. CENSUS of Cities: LCZ Classification of Cities (Level 0) – Workflow and Initial Results from Various Cities. In *9th International Conference on Urban Climate*., 2015.

Jensen, J.R., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. 2nd ed.

Kaloustian, N., Tamminga, M. & Bechtel, B., 2017. Local climate zones and annual surface thermal response in a Mediterranean city. In *2017 Joint Urban Remote Sensing Event*., 2017.

Mills, G. et al., 2015. An Introduction to the WUDAPT Project. In *Proceeding of the ICUC9*. France, Toulouse, 2015.

Stewart, I.D. & Oke, T.R., 2012. Local Climate Zones for Urban Temperature Studies. *Bulletin of the American Meteorological Society*, 93(12), pp.1879-900.

Yang, C., Wu, G., Ding, K. & et, a., 2017. Improving Land Use/Land Cover Classification by Integrating Pixel Unmixing and Decision Tree Methods. *Remote Sensing*, 9.