Seasonal semi-supervised domain adaptation for linking population studies and Local Climate Zones

Basile Rousse[1,2], Sylvain Lobry[1], Géraldine Duthé[2], Laurent Wendling[1], Valérie Golaz[2,3]
[1]LIPADE, Université Paris Cité, Paris [2]INED, Aubervilliers [3]Aix-Marseille Univ., IRD, LPED, Marseille

Context

- Little documented impact of environmental change in demographic studies.
- Scarce demographic and environmental data in sub-Saharan countries.
- Large amount of Sentinel images with a high refresh rate that allows to monitor environmental change.

Linking environmental and demographic data

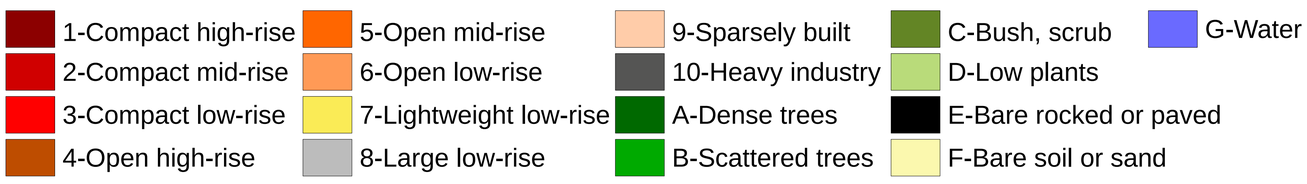
- Demographic and Health Survey Program [1] provide:
  - Individual and household demographic data (health, nutrition, wealth...) with some contextual environmental data (rainfall, NDVI...)

→ **A finer land classification scheme provides a richer analysis of the link between environment and population.**

Local Climate Zones

17 Local Climate Zone (LCZ) classes based on properties of surface structures [2]

So2Sat dataset available for training deep learning models [3]:
- 42 labeled cities mainly in Europe, Estern Asia and Americas.
- 400,673 pairs of Sentinel-1/2 images patches (320mx320m).



Semi Supervised Domain Adaptation (s-SSDA) using seasonal variations in sub-Saharan countries

Few sub-Saharan cities in **So2Sat → Ground truth seldom available**

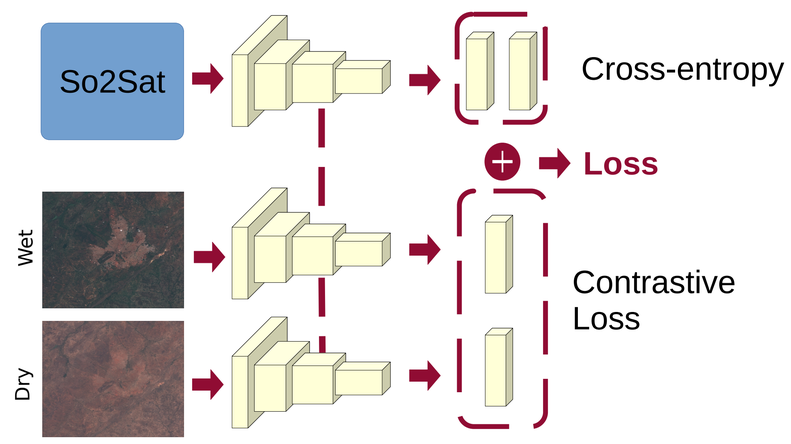
Spatial variations

Reduce domain gap between training regions and sub-Saharan Countries (urban and rural).

Seasonal variations

- High seasonal variations with wet and dry seasons. A given region can:
  - Have a different label for each season.
  - Be visually different while keeping the same label.
- How to learn both LCZ classification and seasonal variations?
  → Supervised learning and self-supervised **contrastive learning** (SimCLR [4], SeCo[5]).
  → Creation of a dataset : 2x250k patches over regions of interest of Burkina Faso for each season.

→ **Temporal regularization using a first order Markov process**



Linking Local Climate Zones and Malaria Indicator Survey data (MIS 2017-2018) in Burkina Faso

Burkina Faso MIS, 2017-2018 [6]

Objective: to provide demographic and malaria related health data.
- Households are sampled and malaria prevalence for 6-59 months old children is representative at regional level.
- 245 Enumeration Zones (EZs) - 6322 households.
- Malaria rapid tests (15 minutes) + lab confirmation
- Malaria rate for one EZ is the proportion of positive 6-59 months old children

Linking LCZs to Enumeration Zones

EZs' centroids geo-locations are provided in MIS with an offset (0-10 km):
We associate geo-locations to small areas (640x640m), intentionnaly focused on the centroid [7].

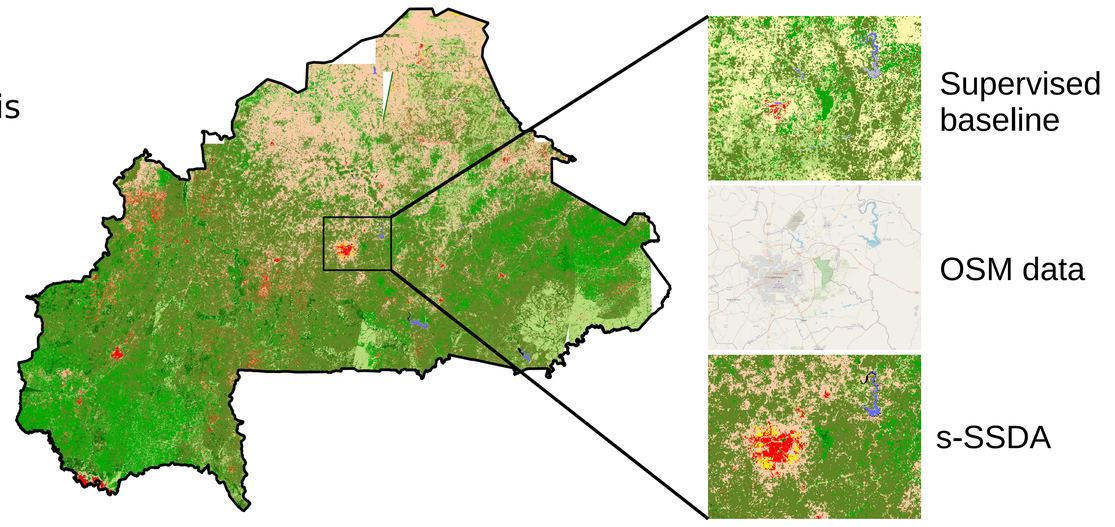
Characterizing the environment

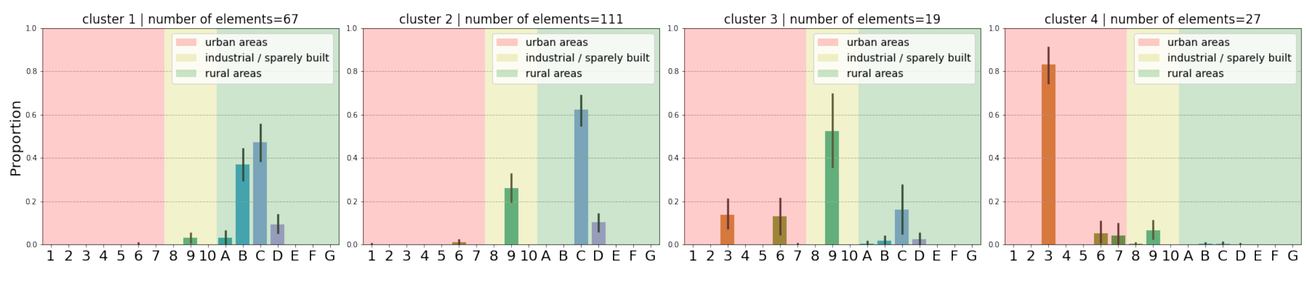
We group EZs in clusters according to their LCZ distributions.
Clusters represent different types of environment:
- Cluster 1: Bush/scrub and scattered trees (very rural).
- Cluster 2: Bush/scrub, sparsely built and low plants (rural).
- Cluster 3: Sparsely built, compact low-rise, open low-rise, bush/scrub (between rural and urban).

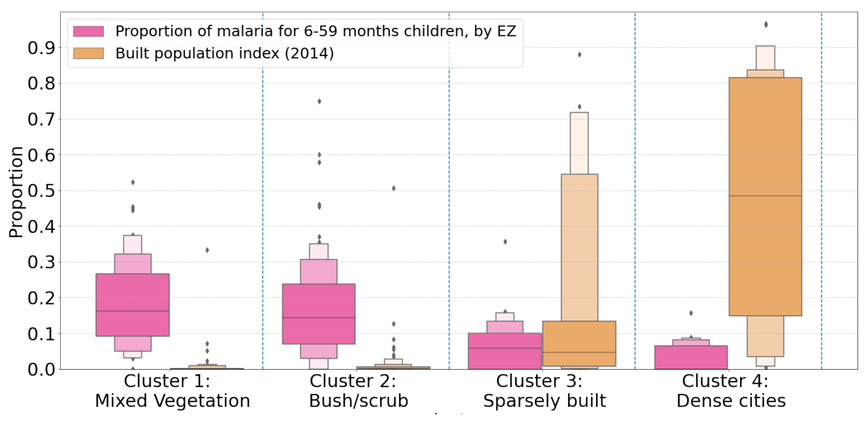
Cluster 4: Compact low-rise and urban classes (very urban)

Linking malaria to Local Climate Zones

→ Differences in malaria rates in both urban and rural clusters.







Conclusion and perspectives

→ Training strategy for mapping sub-Saharan Countries.
→ Differenciation of several environmental structures according to their malaria rates.
→ Next step: Applying s-SSDA to Madagascar and linking LCZs to mortality data in Antananativo.

This work was performed using HPC resources from GENCI-IDRIS (Grant 2021-AD011013527).

References

[1] https://dhsprogram.com/
[2] Stewart and Oke, Local climate zones for urban temperature studies. Bulletin of the American Meteorological Society, 93(12), 1879-1900, 2012
[3] Zhu et al., "So2Sat LCZ42: A Benchmark Data Set for the Classification of Global Local ClimateZones", IEEE GRSS Magazine, vol. 8, pp. 76–89, 2020.
[4] Chen et al. , "A Simple Framework for Contrastive Learning of Visual Representations", Proceedings of the 37th ICML, vol. 119, pp. 1597–1607, 2020.
[5] Mañas et al. ,"Seasonal Contrast: Unsupervised Pre-Training From Uncurated Remote Sensing Data", Proceedings of ICCV, pp. 9414–9423, 2021.
[6] INSD et al., "Enquête sur les indicateurs du paludisme au Burkina Faso", 2018,https://dhsprogram.com/pubs/pdf/MIS32/MIS32.pdf
[7] Grace et al., "Integrating Environmental Context into DHS Analysis While Protecting Participant Confidentiality: A New Remote Sensing Method", Population and Development Review, pp. 197-218, 2019.