

URBAN CO-CREATION DATA LAB

#2 Waste management Provincial Council of Badajoz



Co-financed by the Connecting Europe Facility of the European Union

Objective

Test the methodology applied for waste management in **Lisbon** with waste data from the **Provincial Council of Badajoz**



Business understanding

- The waste collection is made for 95 villages of the province and a specific waste collection circuit can be made in more than one village.



Service



Development of a service to **predict mixed waste** production for a set of waste collection circuits for **every week of the following month**



Data understanding



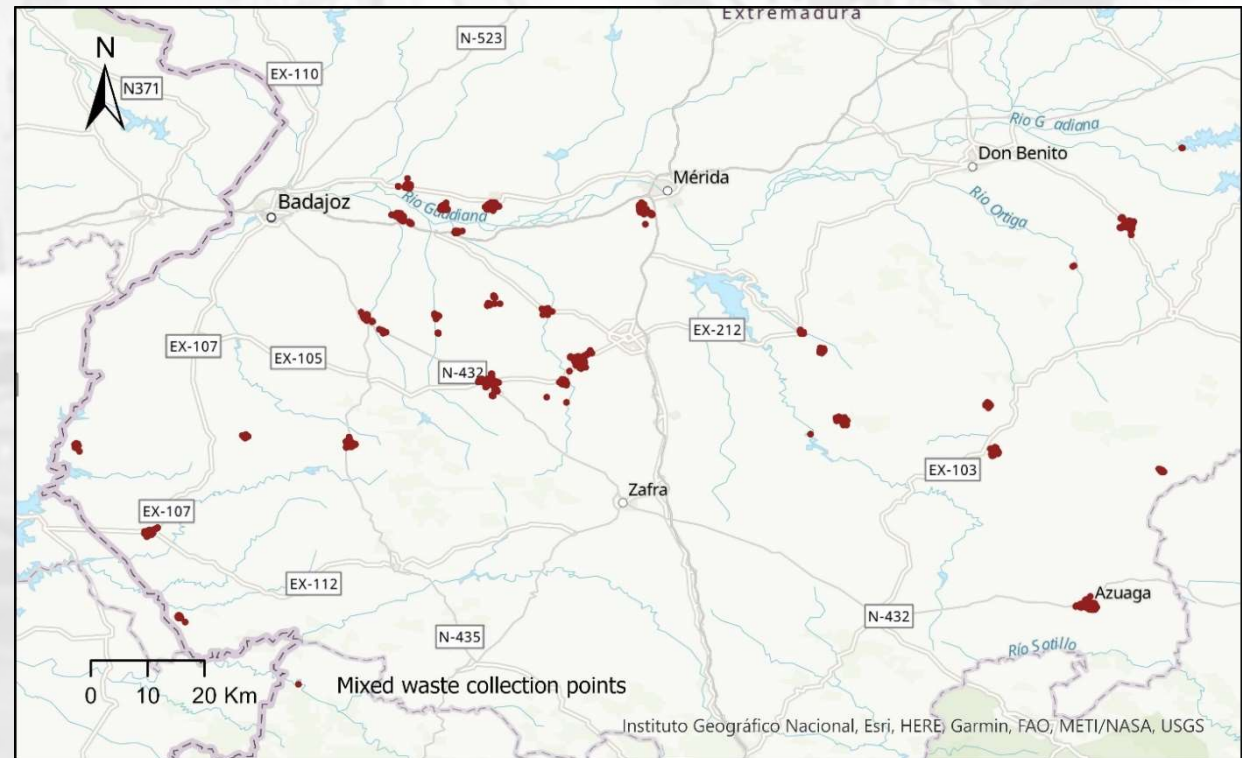
Mixed waste loads (07/2018 – 06/2021) for 16 circuits (dataset 1) and 4 circuits (dataset 2)



Holidays



Employment (proxy for income)



Data preparation

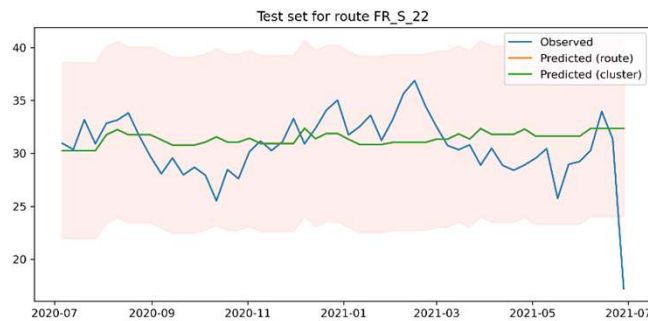
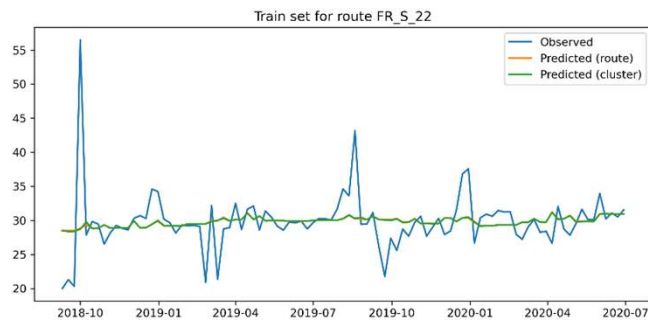
- Mixed waste data was **aggregated weekly** by waste collection circuit
- Created a flag to **identify weeks with holidays**
- Added data from **employment**

Feature	Description
date	Date identifying the first day of the week
waste_t	Quantity of mixed waste in tons collected in a specific week
holiday	Flag identifying holidays
employment	Variation coefficient of the annual employment rate

Modelling



Modelling – dataset 1



Circuit	RMSE (t)	MAE (t)	MAPE
FR_C_03	12,08	9,78	0,29
FR_C_04	11,25	8,22	0,21
FR_N_05	9,34	6,43	0,19
FR_N_10	9,30	5,86	0,13
FR_N_13	17,20	12,63	0,24
FR_N_14	13,35	11,45	0,32
FR_N_15	12,06	8,18	0,20
FR_N_19	6,54	4,13	0,15
FR_S_03	14,20	10,56	0,18
FR_S_04	15,12	11,72	0,23
FR_S_09	11,07	6,20	0,10
FR_S_20	3,95	2,61	0,16
FR_S_21	8,47	5,79	0,14
FR_S_22	3,28	2,33	0,07
All (mean)	10,51	7,56	0,19

- MAPE values are reasonable
- The model is not able to capture the trends in waste production
- No significant seasonal effects were found

Modelling – dataset 2



Circuit	RMSE (t)	MAE (t)	MAPE
FR_C_01	8,37	4,95	0,15
FR_C_03	3,65	2,68	0,09
FR_C_04	5,82	4,42	0,12
FR_C_05	6,71	4,99	0,10
All (mean)	6,14	4,26	0,12

- The adjustment of predicted values is **more consistent** (FR_C_05)
- The observed values **fit inside the confidence interval** (FR_C_05)

Evaluation

Use case	RMSE (t)	MAE (t)	MAPE (t)
Badajoz (dataset 1)	10,51	7,56	0,19
Badajoz (dataset 2)	6,14	4,26	0,12
Lisbon	4,6	5,3	0,16

- MAPE is **higher for dataset 1 than dataset 2**
- RMSE and MAE are **lower for dataset 2**
- The evaluation metrics are **similar** for Lisbon and the Provincial Council of Badajoz (**dataset 2 – just 4 circuits**)

Conclusion

- No seasonal effects were identified
- The fitted model don't fit too well to the observed values
- The data from the Provincial Council of Badajoz is substantially different than the one for Lisbon
- Need to understand factors that impact waste production in the Province of Badajoz to identify the explanation of trends in data

Thank you

<http://urbandatalab.pt/>



Co-financed by the Connecting Europe
Facility of the European Union