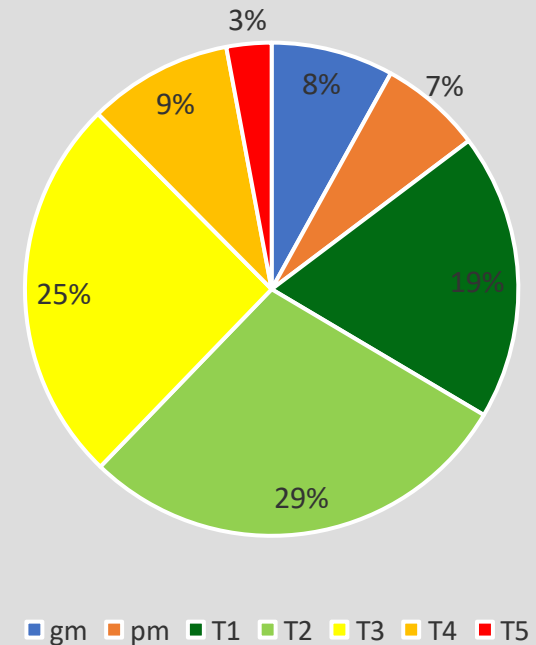


The valorisation of urban residential space. Questions and approaches for the quantitative geographer.

Giovanni FUSCO, Alessandro VENERANDI

The housing market on the French Riviera

- Essential to understand the capacity of our cities to allow **upward social mobility** (access to home ownership), **social mix** and **generational mix**.
- A particularly **tense housing market** in the last decades (second only to the one in Paris)
- Not a market, but **several markets**, according to housing typology (studios, 2-room apts, 3-room apts, houses, etc.)
- **114 K transactions on the French Riviera** from 2010 to 2017 (out of 150 K, French notaries BD)



From the value of dwellings to the value of spaces

Our research topic: Micro-geography of residential valorization
Carpenter & Peponis (2010): the interest of fine-grained local approaches.

Census areas: too large, do not correspond to morphological entities



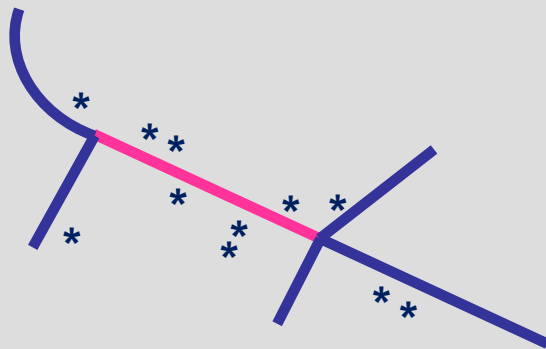
- **The street segment:**

- essential morphological component of the physical city (Conzen 1960, Caniggia & Maffei 2001)
- the basis of the functioning of the city as a social system (Jacobs 1960, Hillier 1996, Porta et al. 2006).
- A local ecosystem of the residential function

Local statistical description of residential valorization

Geolocation of transactions on the urban street network.

Normalization of transactions values by housing type and year to understand « **value of the street** »: market statistics on each street segment.



ventile:	1	2	3	4	5	6	7	8	9
T1 2013			*			**					
T1 2016				*							
T2 2012				*				*			
T2 2016			*								
T3 2015	*										
T4 2012			*		*						
pm 2014					*						

Only 14,300 units with at least 10 transactions

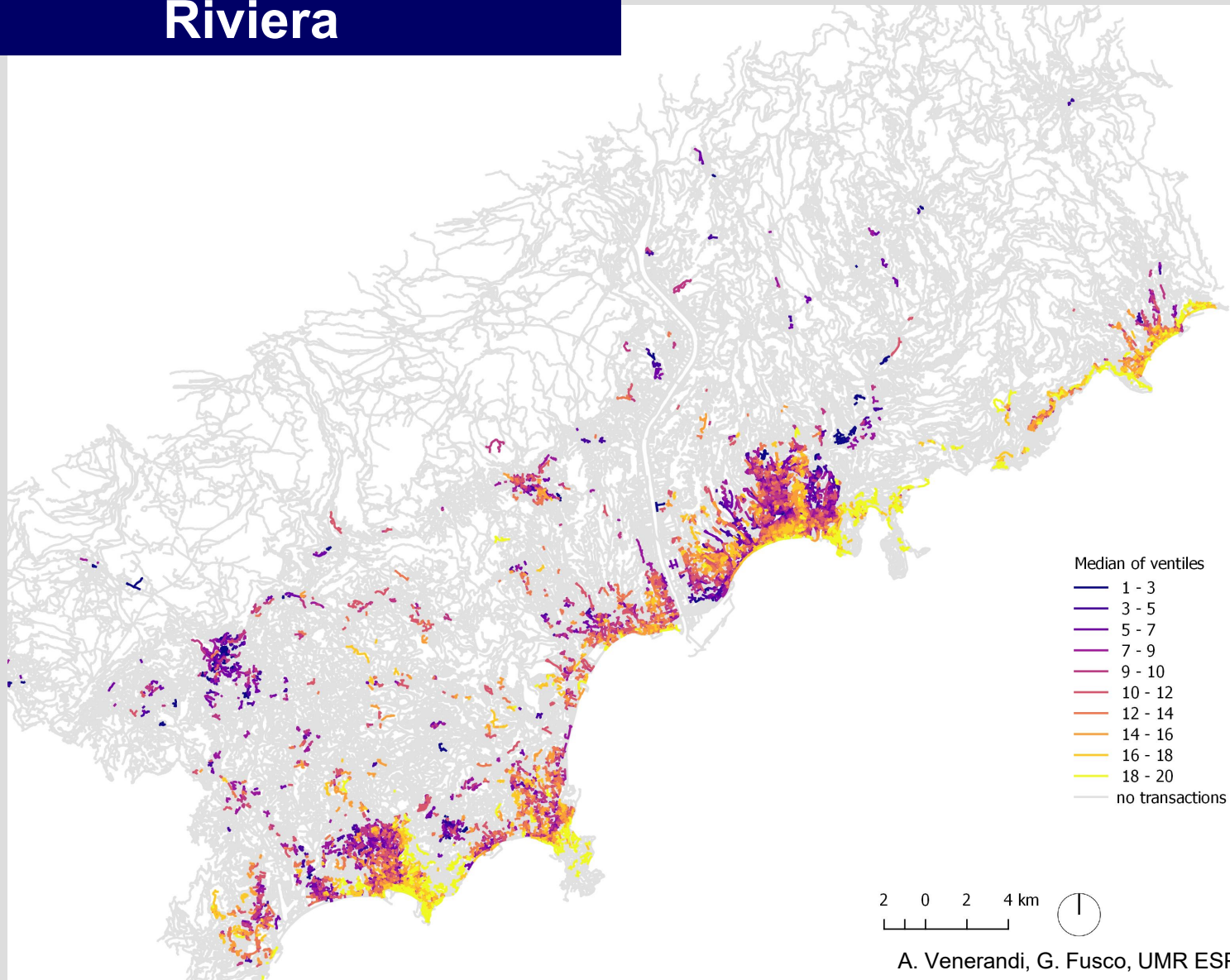
D5 (median) = 4

D9 = 6

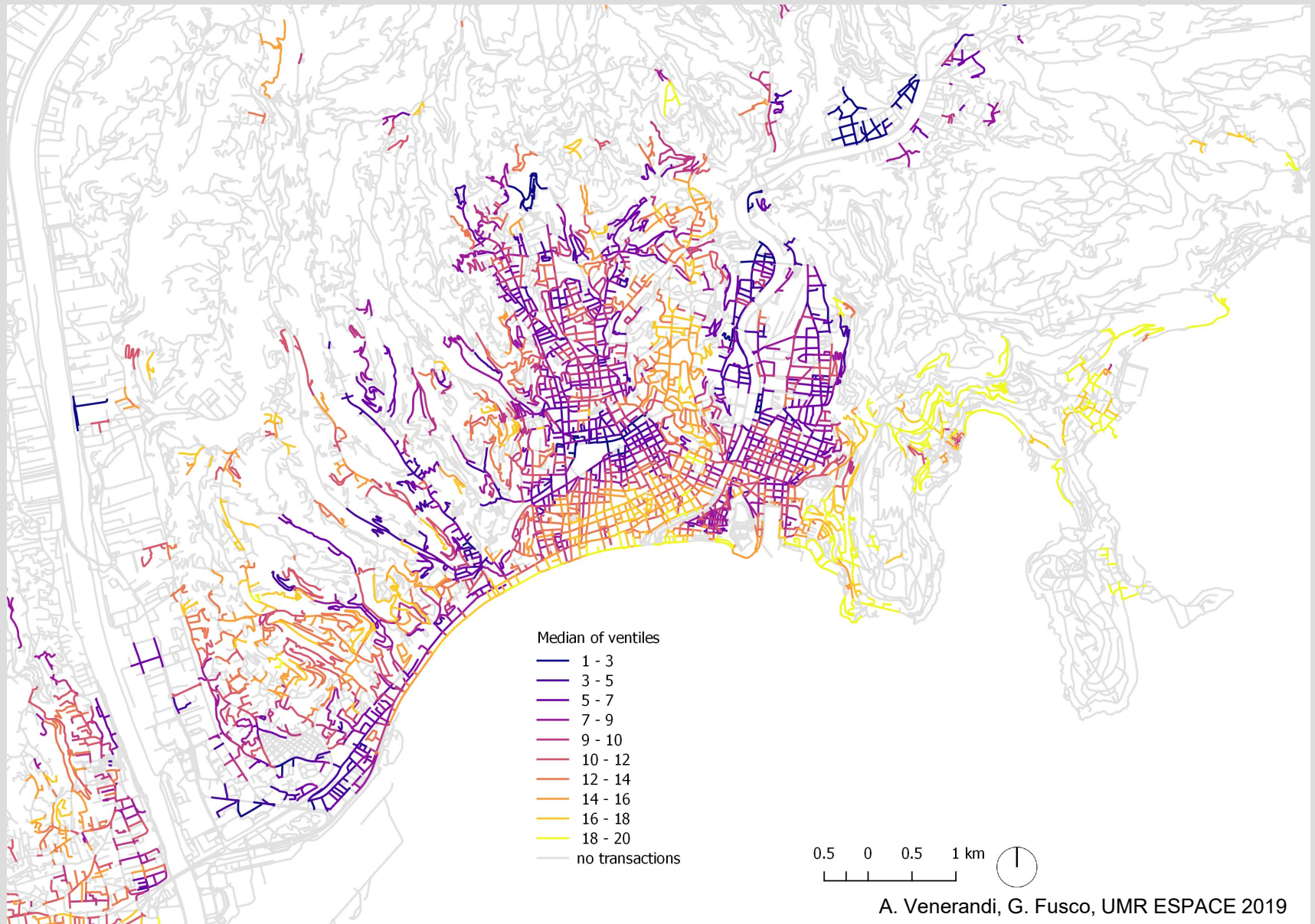
D1 = 3

D9-D1 (dispersion) = 3

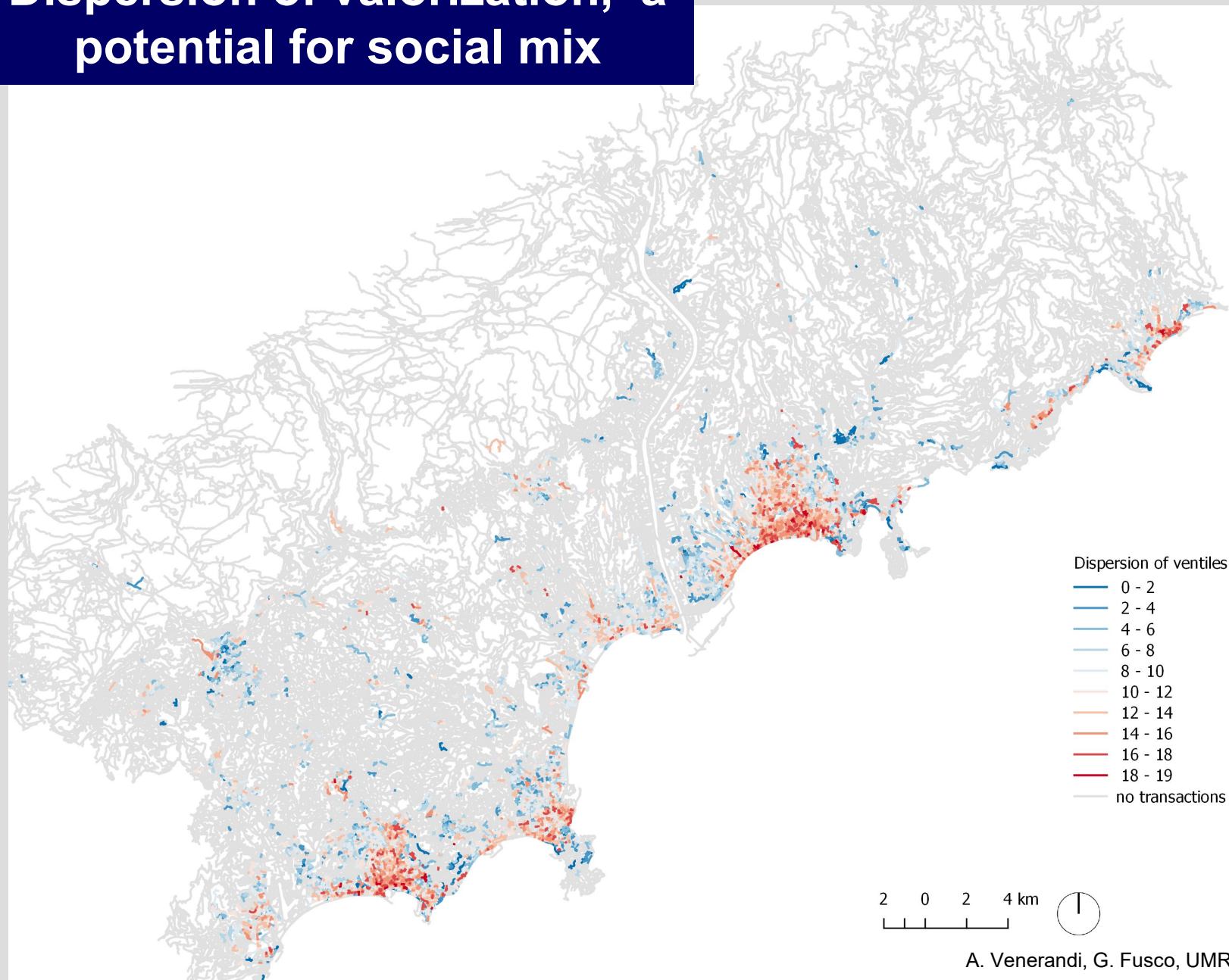
Median residential valorization on the French Riviera



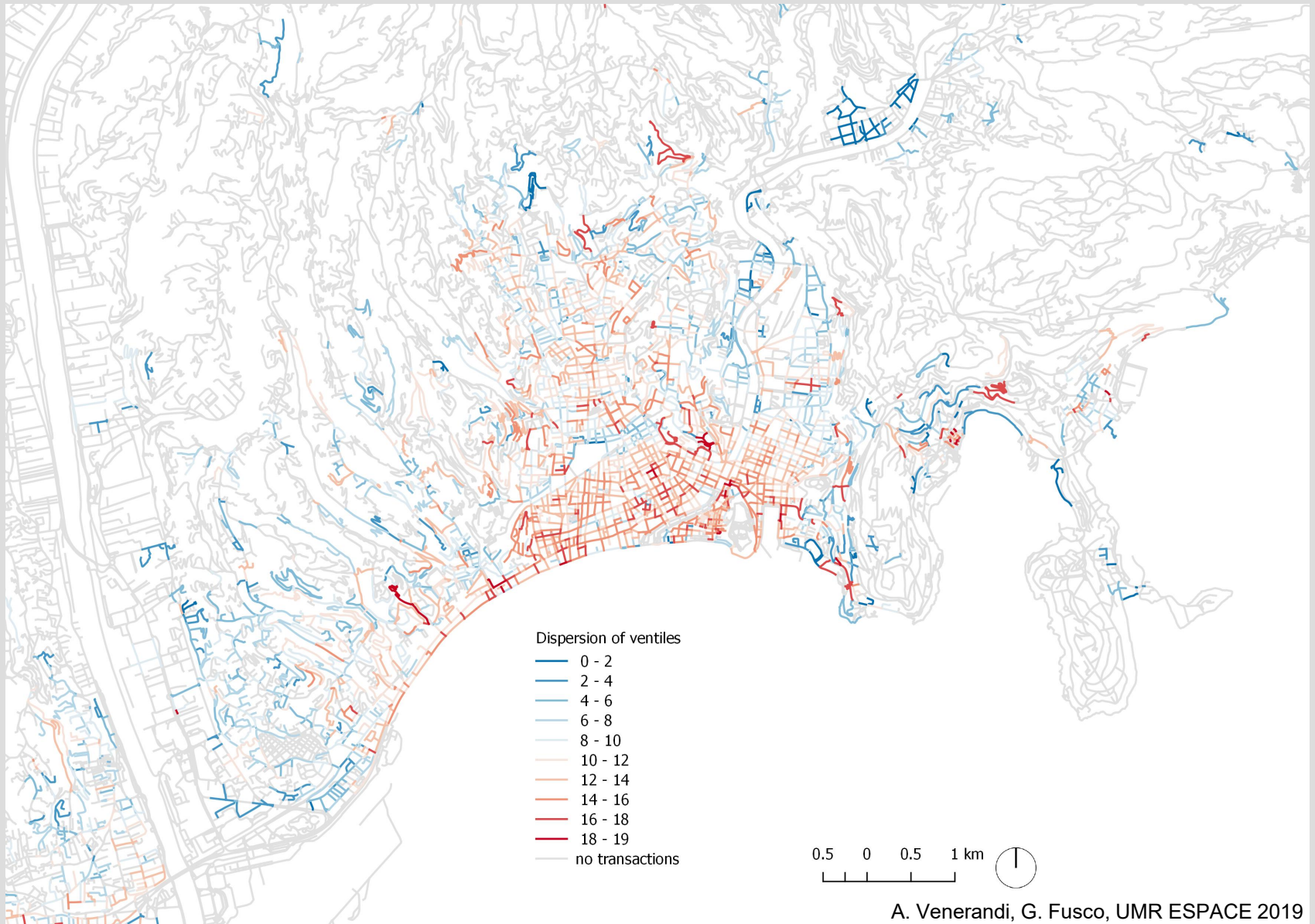
Median residential valorization in Nice



Dispersion of valorization, a potential for social mix



Social Mix potential in Nice





Bayesian Clustering

Score optimization (log-likelihood + penalization for number of clusters... in order to avoid overfitting).

Random walk in solution space under constraints :

- Min % of street segments in each cluster (1%)
- Max number of clusters (20)
- Minimum average cluster purity (90%)
- ...
- Pseudo-random initialization with given seed

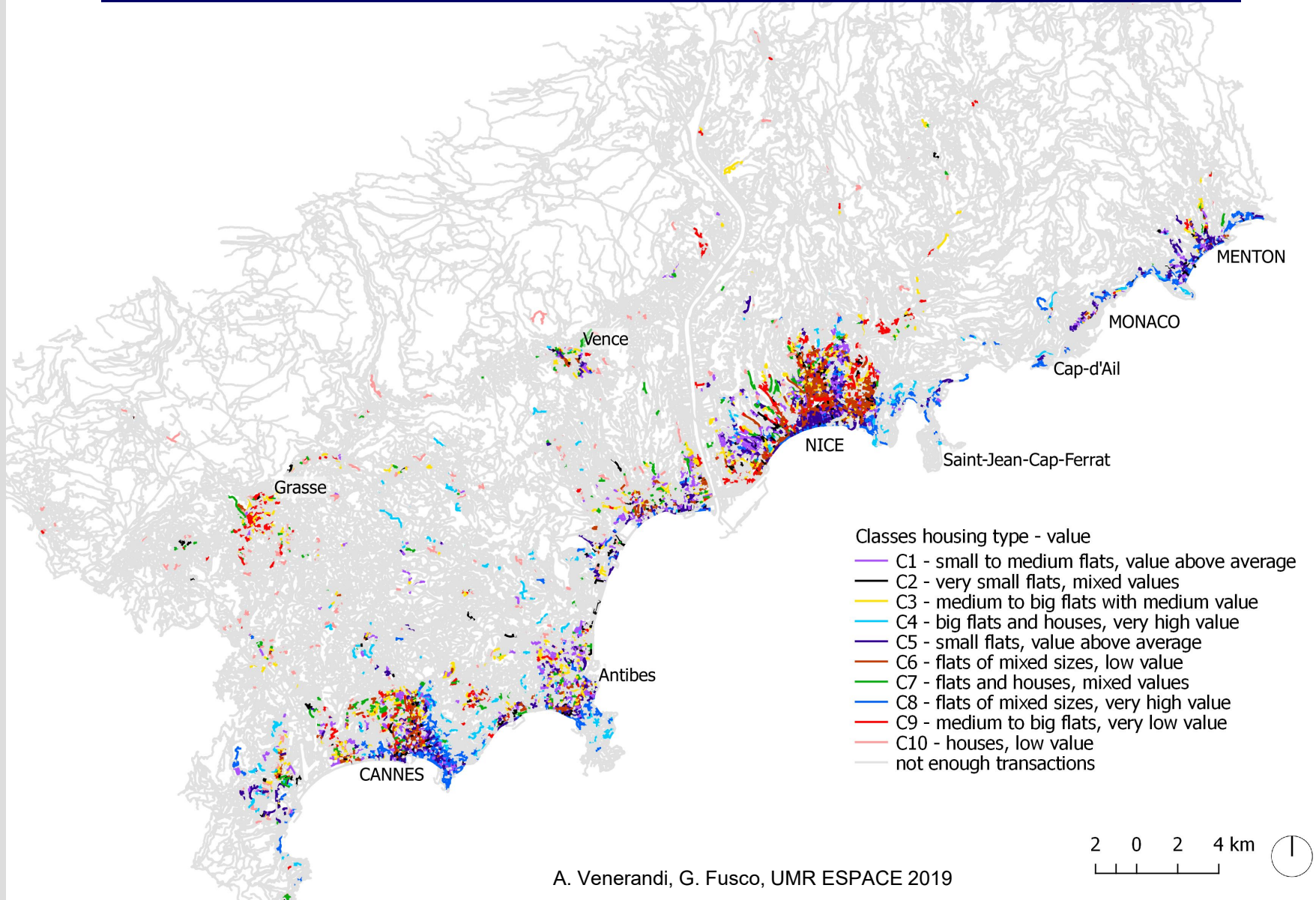
Different search parametrizations → different clustering results

Sensitivity analysis to parametrization

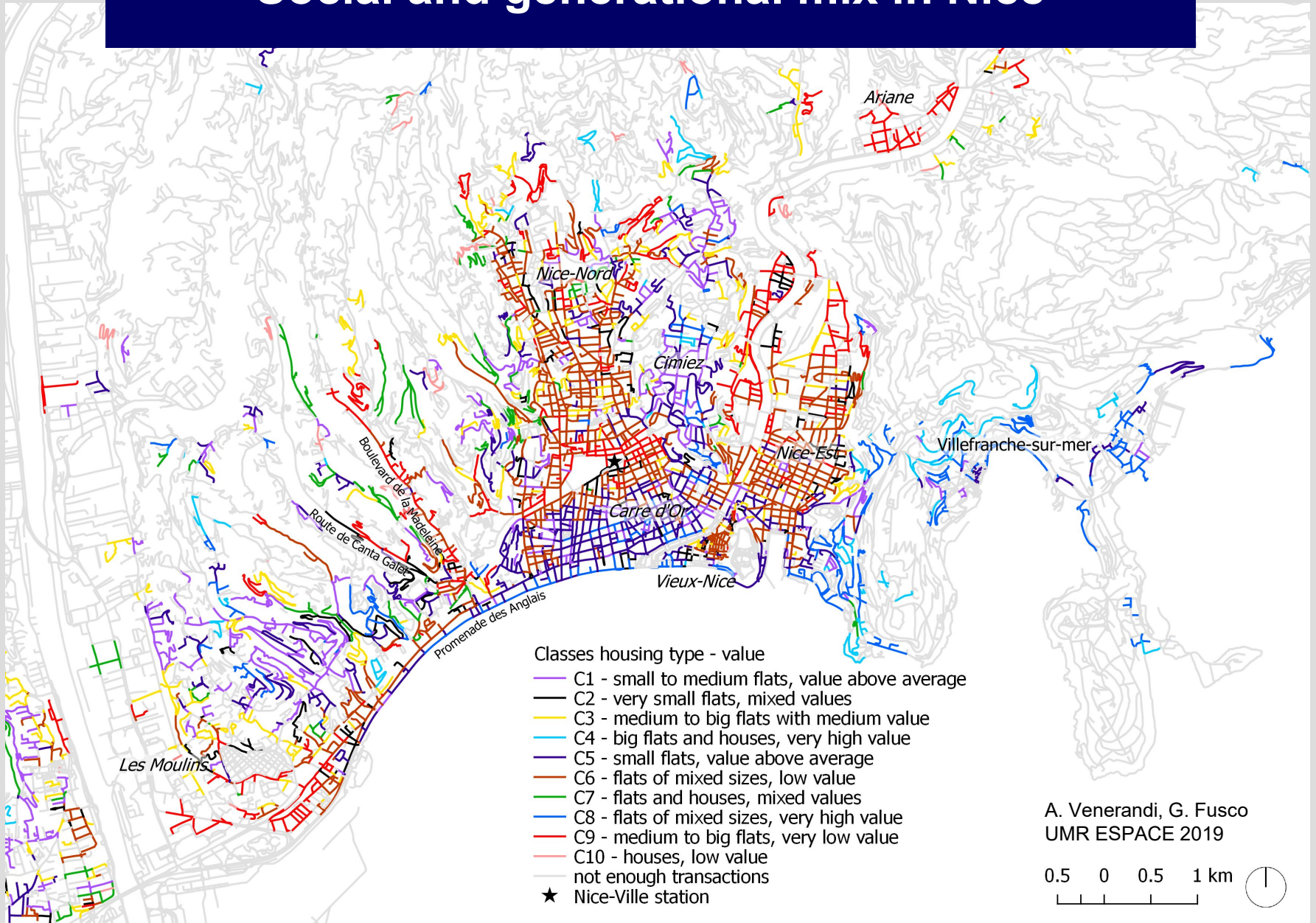
Final result: 10 clusters

(contingency table fit score of 29,4%)

Combining valorization and housing types: social and generational mix



Social and generational mix in Nice





What explanatory model for the value of the street?

Urban economics: the hedonic regressive model.

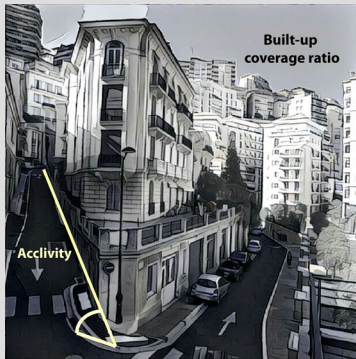
- Based on theory of price formation of a dwelling
- Difference between intrinsic and extrinsic characteristics
- Linearity (overcome through Box-Cox transforms), monotonicity, additivity (substitutability), independence
- Hofman et al. (2017): model specification based on unbiased estimation of significant parameters respecting the theoretical model – no attention to predictive power.

A different stance (urban micro-geography):

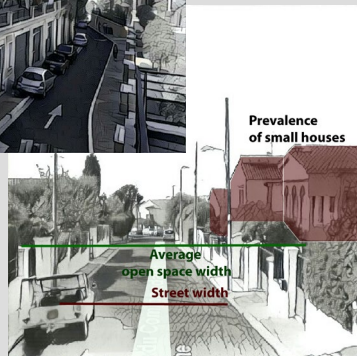
- Use transaction prices to obtain a statistical description of valorisation of spatial units.
- Explain valorisation statistics using only extrinsic characteristics (intrinsic ones are part of the phenomenon to be explained).

Hypotheses behind median valorisation

Configurational properties of the street
(through-movement, accessibility, tranquility)

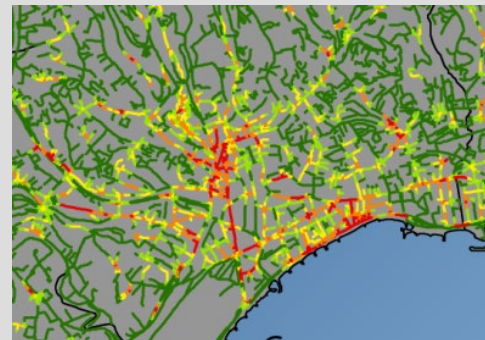


Form of the
urban fabric



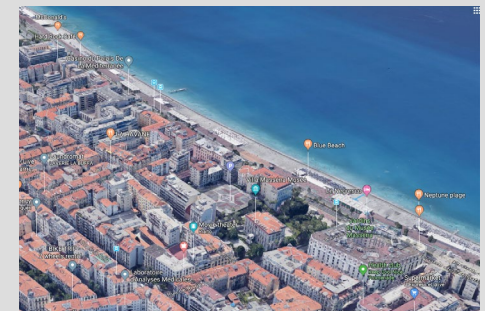
Topography
and orientation

Proximity to infrastructures



Proximity and
accessibility to
the Sea

Position within the spatial
structure of functions (jobs,
commerce, service, transit)





A Machine Learning Approach

Valorization of street segments:

- Absence of well-established theory (no formal model to adjust, only hypotheses)
- Multi-dimensionality (with interdependencies)
- Non-linearity, non-substitutability



Interest of ML Approach

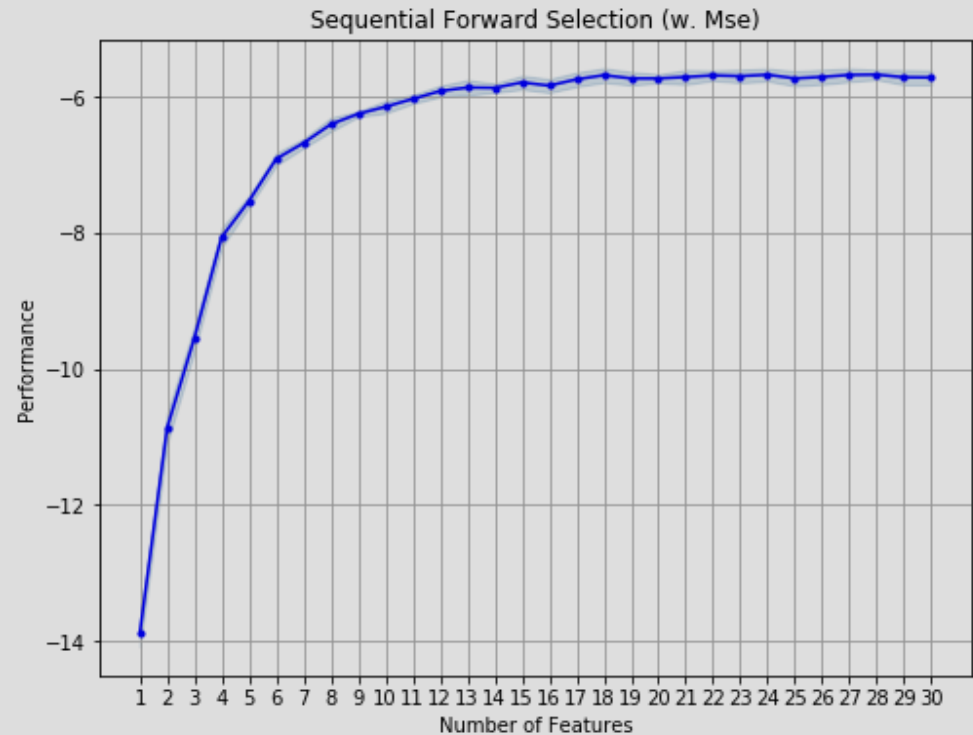
Ensemble ML method:

- Sequential Forward Selection
- Gradient-Boosting Decision Trees
- Network-based Moran I
- SHapley Additive exPlanations

Feature Selection

Sequential Forward Selection

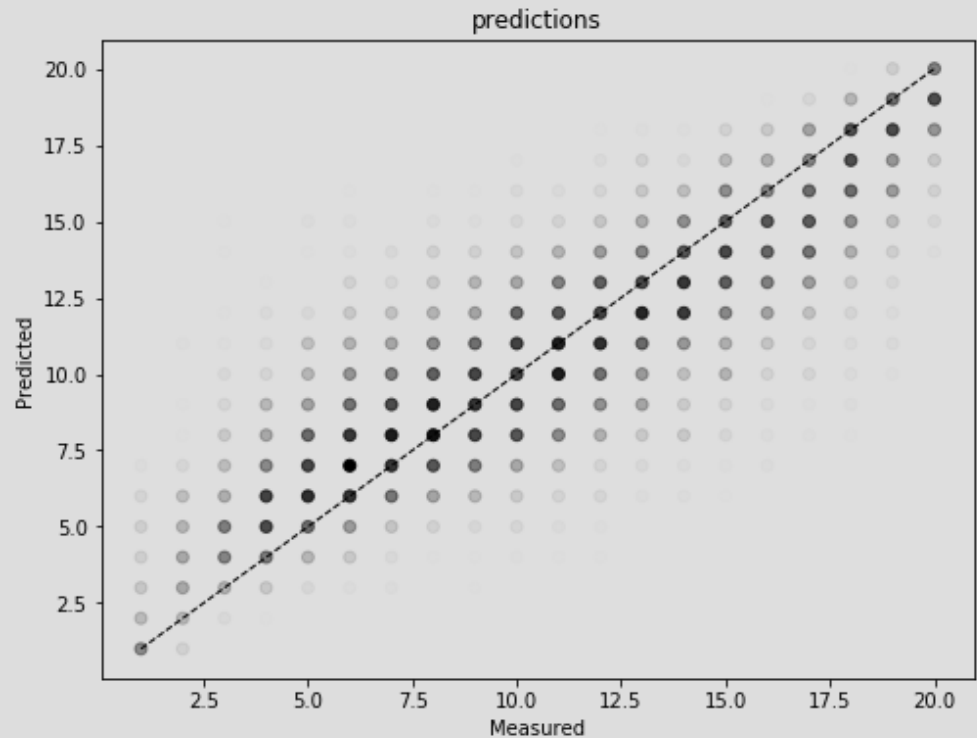
- From 116 metrics to 28 factors (optimum required between 10 and 30)
- 14,319 units (80% train, 20% test), 5-fold cross validation
- Average mean absolute error 2.38 (target variable between 1 and 20)



Gradient Boosting Model

A predictive model based on nested decision trees.

- 14,319 units (80% train, 20% test)
- Hyper-parameter learning, max 128 DT, 10-fold cross-validation
- Average mean absolute error :
 - 1.05 (train)
 - 2.43 (test)
- R^2 :
 - 0.95 (train)
 - 0.75 (test)



Network-based Moran I

Need to account for all possible spatial structure of target variable.



No spatial auto-correlation of model deviates

20 topological steps: $I = 0.08$ Significant (p-value = 0.0099)

10 topological steps: $I = 0.13$ Significant (p-value = 0.0099)

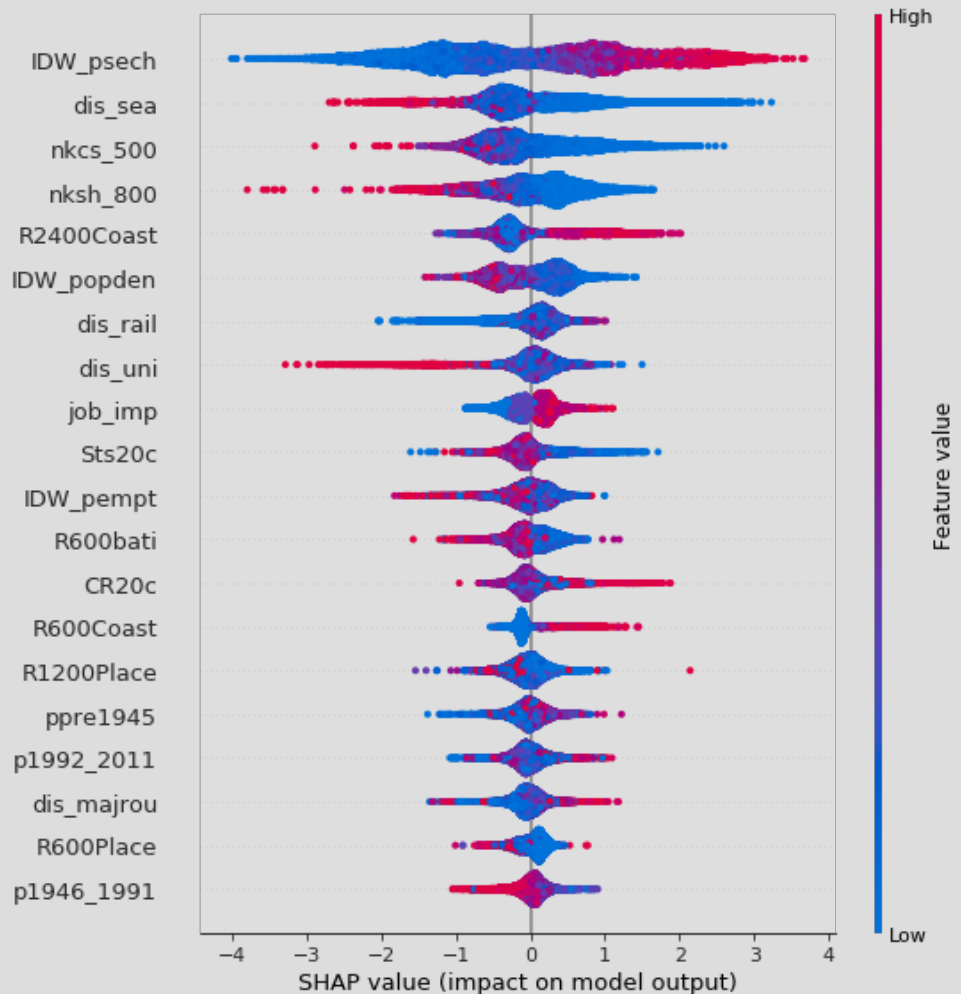
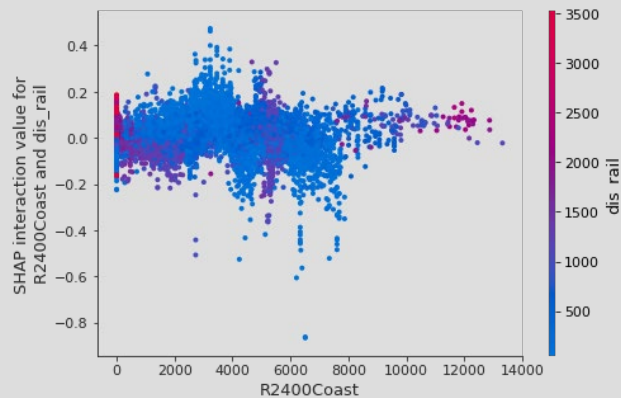
- Helpful to identify missing factors (ex. Distance to University Campus identified at this stage)
- Residual spatial structure at finer scale (not all data were available at the street-segment level)



Making the model understandable

SHapley Additive exPlanations (SHAP)

- Local analysis the I/O relationships in the model
- Possibility to analyse SHAP interaction values





Conclusions

- The value of streets, a question rooted in urban geography, no intrinsic/extrinsic dichotomy.
- The proposed ensemble method: effort to make ML approaches interpretable, without renouncing to predictive power.
- French Riviera: % second homes, distance to the sea, density of commerce/services and concentration of social housing, the most important factors (the first showing a positive relationship, the others an inverse one).
- Perspective of comparative analyses with other urban areas.



Limitations & Future Work

- Inaccuracies from the PERVAL DB
- Loss of information from streets with too few transactions
- Hyperparameters of the GB model set through trial and error by visual evaluation of learning curves... need of algorithmic solution.
- Residual spatial structure at local scale

Forthcoming:

- Integrate information from streets having too few transactions. Approaches of uncertain knowledge (Bayesian probabilities, possibility theory).



Thanks for your attention

giovanni.fusco@univ-cotedazur.fr
alessandro.venerandi@univ-cotedazur.fr

UNIVERSITÉ
CÔTE D'AZUR 

