The Determinants of Landscape and Cultural Heritage Among Italian Regions in the Period 2004-2019

Angelo Leogrande (angelo.economics@gmail.com)
LUM University Giuseppe Degennaro
Alberto Costantiello (costantiello@lum.it)
Lum University Giuseppe Degennaro
Laureti Lucio (Laureti@lum.it)
Lum University Giuseppe Degennaro
Domenico Leogrande
Università degli Studi di Bari Aldo Moro

Research Article

Keywords: Environmental Economics, Valuation of Environmental Effects, Pollution Control Adoption and Costs, Sustainability, Government Policy

Posted Date: November 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1112258/v1

Angelo Leogrande[1], Alberto Costantiello[2], Lucio Laureti[3], Domenico Leogrande[4]

The Determinants of Landscape and Cultural Heritage among Italian Regions in the Period 2004-2019

We estimate the Landscape and Cultural Heritage among Italian regions in the period 2004-2019 using data from ISTAT-BES. We use Panel Data with Fixed Effects, Panel Data with Random Effects, Pooled OLS, WLS, Dynamic Panel. We found that the Landscape and Cultural Heritage is negatively associated with "*Dissatisfaction with the landscape of the place of life*", "*Illegal building*", "*Density and relevance of the museum heritage*", "*Internal material consumption*", "*Erosion of the rural space due to abandonment*", "*Availability of urban green*", and positively associated with "*Pressure from mining activities*", "*Erosion of the rural space by urban dispersion*", "*Concern about the deterioration of the landscape", "Diffusion of agritourism farms*", "*Current expenditure of the Municipalities for culture"*. Secondly, we have realized a cluster analysis with the k-Means algorithm optimized with the Silhouette Coefficient and we found two clusters in the sense of "*Concern about the deterioration of the landscape*". Finally, we use eight different machine learning algorithms to predict the level of "*Concern about the deterioration of the landscape*" and we found that the Tree Ensemble Regression is the best predictor.

JEL CODE: Q50; Q51; Q52; Q56; Q58
Keywords: Environmental Economics; Valuation of Environmental Effects; Pollution Control Adoption and Costs; Sustainability; Government Policy.

1. Introduction

In this article we have analyzed the determinant of Landscape and Cultural Heritage among Italian regions in the period 2004-2019. We use data from ISTAT-BES. The role of Landscape and Cultural Heritage has acquired a growing interest among population and policy makers as a result either of the actions against climate change either of a re-discovery of ethno-linguistics and monumental traditions also at a local level.

(Della Spina, 2017) afford the question of the relevance of a multi-methodological and multi-stakeholder approach in the context of Historical Urban Landscape with an active participation of communities and local experts in preserving either landscape either historical heritage. (Gravagnuolo & Girard, 2017) propose a series of metrics to evaluate the level of Historical Urban Landscape. (Bulian, 2021) considers the role of multi-locality and the interdependence of cultural heritage and landscape in Japan. (Rouhi, 2017) affords the question of the anthropological meaning of cultural heritage in the context of universal values. (Vallerani & Visentin, 2018) show the relevance of waterway as a cultural and socio-economic tool for civilization and landscape valuation. (Cicinelli, Salerno, & Caneva, 2018) apply a multidisciplinary approach to the preservation of Medieval Benedictine Monastery of San Vincenzo al Volturno considering natural, archeological, and agricultural elements.

(Ravankhah, Schmidt, & Will, 2021) develop a methodology to evaluate the environmental risk for cultural heritage size such as in the case of earthquakes. The authors propose the "*Cultural Heritage*

[1] Assistant Professor of Economics, Lum University Giuseppe Degennaro, Casamassima, Bari, Puglia, Italy, Eu.
[2] Professor of Economics, Lum University Giuseppe Degennaro, Casamassima, Bari, Puglia, Italy, Eu.
[3] Professor of Economics, Lum University Giuseppe Degennaro, Casamassima, Bari, Puglia, Italy, Eu.
[4] Professor of Statistics, University of Bari Aldo Moro, Bari, Puglia, Italy, Eu.

Risk Index" with an application for the World Heritage site of Bam in Iran. (Antonson, Buckland, & Nyqvist, 2021) analyze the question of the relationship among cultural heritage, climate change and active governmental policies in Sweden. (Assandri, Bogliani, & Pedrini, 2018) show the relevance of agriculture in creating the conditions to preserve landscape and biodiversity with an application to wine production in Italian region Trentino-Alto Adige.

(Cai, Fang, Zhang, & Chen, 2012) consider the application of a connection between digital technologies and the cultural heritage protection in the case of Mount Lushan in China. To preserve Mount Lushan in China from the negative externalities of massive tourism the authors have promoted a model of virtual tourism based on 3D laser scanning, oblique aerial photography and 360 degrees panorama technology. (Foster, 2020) suggest new methodologies based on circular economy to cultural heritage buildings in the context of environmental sustainability. (Shirvani Dastgerdi, Sargolini, Broussard Allred, Chatrchyan, & De Luca, 2020) consider the role of climate change in worsening the condition of central Italy regions in terms of rainfall patterns suggesting a deeper coordination between the European Landscape Convention and local and territorial planning to pursue the objective of conservation.

(Guzman, Fatorić, & Ishizawa, 2020) propose some methodology to mitigate the risk of climate change for World Heritage-WH sites at a global level suggesting a multidisciplinary and multistakeholder approach to promote resilience. (Li, Krishnamurthy, Roders, & Van Wesemael, 2020) consider the role of community participation in heritage management considering a people-centered approach with an application in China.

The article continues as follows: the second paragraph presents the econometric model, the third paragraph present the cluster analysis, the fourth paragraph contains the machine learning and prediction algorithms, the fifth paragraph concludes.

2. The Econometric Model

We estimate the following econometric model:

$$\begin{aligned}
\boldsymbol{LandscapeAndCulturalHeritage_{it}} \\
&= \boldsymbol{a_1 + b_1(CurrentExpenditureOfTheMunicipalitiesForCulture)_{it}} \\
&+ \boldsymbol{b_2(DensityAndRelevanceOfTheMuseumHeritage)_{it}} \\
&+ \boldsymbol{b_3(IllegalBuilding)_{it}} \\
&+ \boldsymbol{b_4(ErosionOfTheRuralSpaceByUrbanDispersion)_{it}} \\
&+ \boldsymbol{b_5(ErosionOfTheRuralSpaceDueToAbandonment)_{it}} \\
&+ \boldsymbol{b_6(PressureFromMiningActivities)_{it}} \\
&+ \boldsymbol{b_7(DiffusionOfAgritourismFarms)_{it}} \\
&+ \boldsymbol{b_8(DissatisfactionWithTheLandscapeOfThe\ placeOfLife)_{it}} \\
&+ \boldsymbol{b_9(ConcernAboutTheDeteriorationOfTheLandscape)_{it}} \\
&+ \boldsymbol{b_{10}(InternalMaterialConsumption)_{it}} \\
&+ \boldsymbol{b_{11}(AvailabilityOfUrbanGreen)_{it}}
\end{aligned}$$

Where $i = 20$ and $t = 2004 - 2019$. We perform different regression models i.e.: Panel Data With Random Effects, Panel Data With Fixed Effects, Dynamic Panel Data, WLS, and Pooled OLS. We use data from ISTAT-BES.

We found that the "*Landscape And Cultural Heritage"* is positively associate with:

- *Current Expenditure of the Municipalities for Culture:* is defined as "*payments in accountability for the protection and enhancement of cultural assets and activities, in euros per capita*". There is a positive relationship between the Landscape and Cultural Heritage and the level of the "*Current Expenditure of the Municipalities for Culture*" i.e. the greater the local expenditures in cultural event the greater the level of Landscape and Cultural Heritage. This means that there is an effective role that policy makers, even at the local level, can play in promoting a culture more oriented toward landscape preservation.
- *Diffusion of Agritourism Farms:* is defined as the number of farms per 100 km2. There is a positive relationship between "*Diffusion of Agritourism Farms*" and the level of "*Landscape and Cultural Heritage". The* presence of agritourism can improve the culture of respect for environmental public goods and public artistic goods among the population creating the conditions even for collective actions in defence of green and cultural commons.
- *Concern About the Deterioration of the Landscape:* is defined as the percentage of people aged 14 and over who indicate the ruin of the landscape caused by excessive building construction among the five problems environmental issues more worrying than the total number of people aged 14 and over. There is a positive relationship between "*Concern About the Deterioration of the Landscape*" and "*Landscape and Cultural Heritage*" suggesting that if the population is emotionally engaged in the deterioration of the landscape than the condition of environmental and artistic goods can improve.
- *Erosion of the Rural Space by Urban Dispersion:* is defined as "*Percentage incidence of the agricultural regions concerned from the phenomenon on the total of the regional surface*". There is a positive relationship between Landscape and Cultural Heritage and the Erosion of the Rural Space by Urban Dispersion.
- *Pressure from Mining Activities:* is defined as the Volume of resources non-energy minerals extracted (cubic meters) per km2. There is a positive relationship between "*Pressure from Mining Activities*" and "*Landscape and Cultural Heritage*".

Results of the Econometrics Models for the Estimation of the Composite Index Landscape And Cultural Heritage											
Label	Variables	Dynamic Panel		Fixed Effects		Random Effects		POOLED OLS		WLS	
A101	*Composite index Landscape and cultural heritage*	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
	Costant	0,0447119	*	99,277	***	98,973	***	98,973	***	98,8187	***
A90	*Current expenditure of the Municipalities for culture*	0,429403	***	0,438627	***	0,441137	***	0,441137	***	0,437829	***
A91	*Density and relevance of the museum heritage*	-0,141787	*	-0,22987	***	-0,226927	***	-0,226927	***	-0,253816	***
A92	*Illegal building*	-0,159717	***	-0,139001	***	-0,138063	***	-0,138063	***	-0,131504	***
A93	*Erosion of the rural space by urban dispersion*	0,0357363	***	0,0318745	***	0,0311435	***	0,0311435	***	0,0316906	***
A94	*Erosion of the rural space due to abandonment*	-0,0299848	***	-0,0394458	***	-0,0383875	***	-0,0383875	***	-0,0392682	***
A95	*Pressure from mining activities*	0,0128409	***	0,0121328	***	0,0123628	***	0,0123628	***	0,012349	***
A97	*Diffusion of agritourism farms*	0,261535	***	0,264261	***	0,263085	***	0,263085	***	0,275149	***
A99	*Dissatisfaction with the landscape of the place of life*	-0,624804	***	-0,647361	***	-0,642133	***	-0,642133	***	-0,65687	***
A100	*Concern about the deterioration of the landscape*	0,105488	***	0,112826	***	0,117746	***	0,117746	***	0,137727	***
A102	*Internal material consumption*	-0,0603181	***	-0,0570269	***	-0,0581511	***	-0,0581511	***	-0,0561594	***
A108	*Availability of urban green*	-0,00716397	***	-0,00998954	***	-0,0099123	***	-0,00991231	***	-0,0104049	***
A101(-1)		-0,0122381	*								

Figure 1. Synthesis of the main results of different econometric models.

We also find that the level of "*Landscape And Cultural Heritage"* is negatively associated with:

- *Availability of Urban Green:* is defined as "*Square meters of urban green space per inhabitant*". The greater the "*Availability of Urban Green*" the lower the level of "*Landscape and Cultural Heritage*".
- *Erosion of the Rural Space due to Abandonment* is defined as *"Incidence percentage of agricultural regions concerned from the phenomenon on the total of the regional surface"*. There is a negative relationship between the Erosion of the Rural Space due to Abandonment

and the level of Landscape and Cultural Heritage. The greater the level of the "*Erosion of the Rural Space due to Abandonment*" the lower the level of "*Landscape and Cultural Heritage*".

- *Internal Material Consumption:* is defined as the *"Quantity of materials processed into emissions, waste or new stocks of the system anthropic (in millions of tons)"*. There is a negative relationship between *"Internal Material Consumption"* and the level of *"Landscape and Cultural Heritage"*. The greater the level of Internal Material Consumption the lower the level of Landscape and Cultural Heritage.
- *Density and Relevance of the Museum Heritage:* is defined as the *"Number of permanent exhibition structures for 100 km2 (museums, archaeological areas and monuments open to audience), weighted by the number of visitors"*. There is a negative relationship between *"Density and Relevance of the Museum Heritage"* and the level of *"Landscape and Cultural Heritage"*. The greater the level of *"Density and Relevance of the Museum Heritage"* the lower the level of *"Landscape and Cultural Heritage"*. This result can appear counterfactual. But it can be better understood considering that many southern Italian regions that have a widespread diffusion of museum also have low levels of attention for the landscape and cultural heritage due to the lack of social and human capital.
- *Illegal Building:* is defined as *"Number of illegal buildings for 100 buildings authorized by the Municipalities"*. There is a negative relationship between Landscape and Cultural Heritage and Illegal Building. The greater the level of Illegal Building the lower the level of Landscape and Cultural Heritage.
- *Dissatisfaction with the Landscape of the Place of Life:* is defined as the "*Percentage of people aged 14 and over who declare that the landscape of the living place is affected by obvious degradation on the total of people aged 14 and over*." There is a positive relationship between "Dissatisfaction with the Landscape of the Place of Life" and the level of "Landscape and Cultural Heritage". The greater the level of "Dissatisfaction with the Landscape of the Place of Life" the lower the level of "Landscape and Cultural Heritage".

Current Expenditure of the Municipalities for Culture and Diffusion of Agritourism Farms have the main positive effects on Landscape and Cultural Heritage with a mean value among the different econometric models respectively equal to 0,4376 and 0,265. On the other side "*Dissatisfaction with the landscape of the place of life*" and *"Density and relevance of the museum heritage"* have the greater negative impact on Landscape and Cultural Heritage with a mean value respectively equal to -0.2158 and -0.6426.

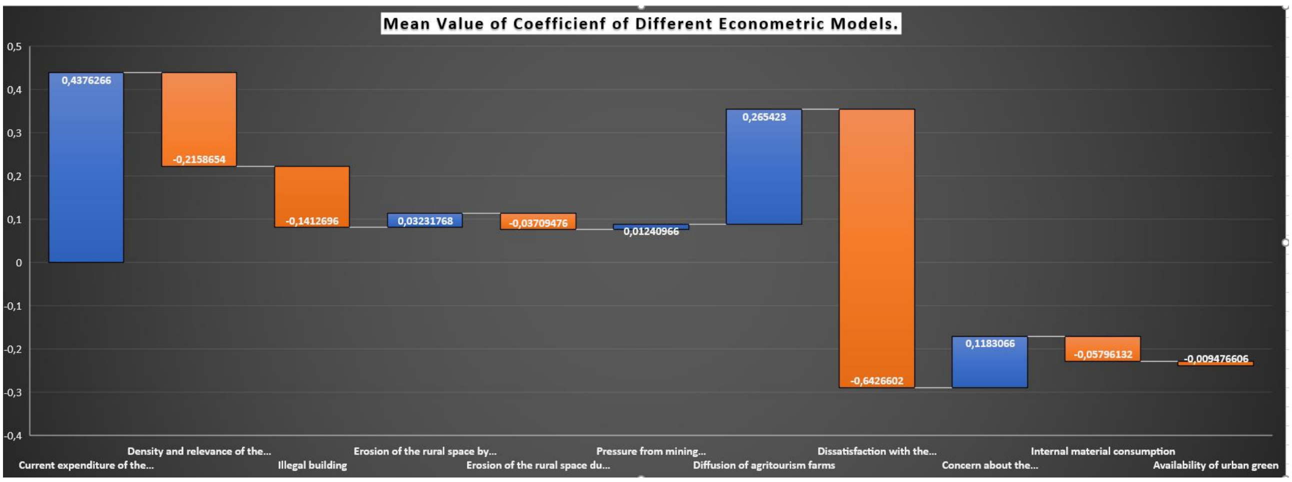


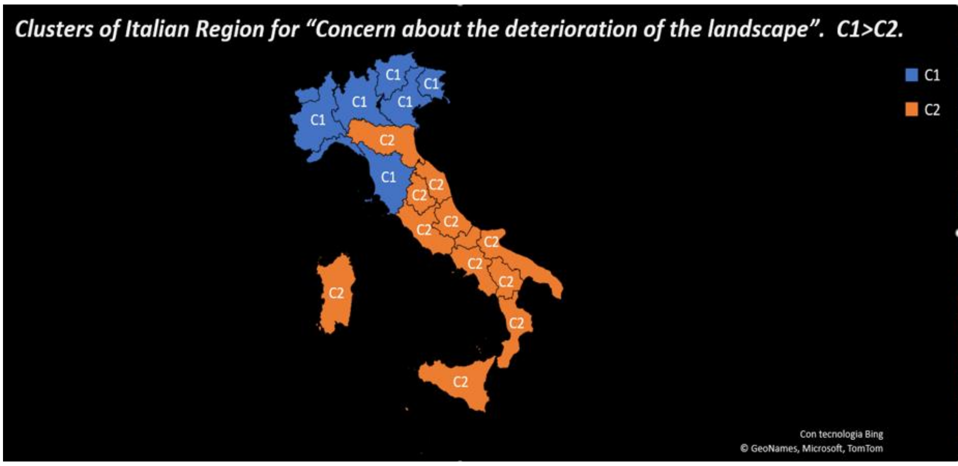
Figure 2. Mean value of the variables among different econometric models.

3. Clusterization

We have applied the k-Means algorithm optimized with the Silhouette Coefficient to investigate the presence of clusters in the sense of "*Concern about the deterioration of the landscape*". We found two clusters as follows:

- Cluster 1: *Piemonte, Valle d'Aosta, Liguria, Lombardia, Veneto, Friuli-Venezia Giulia, Toscana, Trentino-Alto Adige;*
- Cluster 2: *Emilia-Romagna, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna.*

Specifically, the median value of regions in the Cluster 1 is equal to 13,55, while the median value of regions in the Cluster 2 is equal to 10,6. As we can see there is a great divide between Southern and Northern Italy in the sense of "Concern about the deterioration of the landscape". Northern Italian regions that have generally greater Gdp Per Capita and greater human and social capital also show a greater concern about the deterioration of the landscape.



Concern about the deterioration of the landscape			
Value	Region	Cluster	Median value
13,1	Piemonte	C1	13,55
13,8	Valle D'Aosta	C1	
14,7	Liguria	C1	
14,8	Lombardia	C1	
13,3	Veneto	C1	
12,4	Friuli Venezia Giulia	C1	
11,7	Toscana	C1	
15,5	Trentino Alto Adige	C1	
12,6	Emilia-Romagna	C2	10,6
13,1	Umbria	C2	
10,6	Marche	C2	
13,4	Lazio	C2	
10,5	Abruzzo	C2	
9,3	Molise	C2	
10,4	Campania	C2	
9,7	Puglia	C2	
9,1	Basilicata	C2	
10,6	Calabria	C2	
11	Sicilia	C2	
12,7	Sardegna	C2	

4. Machine Learning and Predictions

We have estimated the level of "*Concern about the deterioration of the landscape*" using eight different machine learning algorithms to predict the future value of the observed values. We use 70% of the dataset as learning rate and the remaining 30% for the prediction. We have ranked the eight different algorithms based on their ability to minimize statistical errors such as "Mean Absolute Error", "Mean Squared Error", "Root Mean Squared Error", "Mean Signed Difference". The order of algorithms based on their ability to minimize errors is as follows:

- Tree Ensemble Regression with a payoff of 7;
- ANN-Artificial Neural Network Perceptron Multilayer with a payoff equal to 10;
- Liner Regression with a payoff equal to 13;
- PNN-Probabilistic Neural Network with a payoff equal to 14;
- Gradient Boosted Tree Regression with a payoff equal to 16;
- Random Forest Regression with a payoff equal to 25;
- Simple Regression Tree with a payoff equal to 29;

- Polynomial Regression with a payoff equal to 30.

Results of Algorithms in Terms of Minimization of Statistical Errors					
Algorithm	**Mean absolute error**	**Mean squared error**	**Root mean squared error**	**Mean signed difference**	**Sum**
Tree Ensemble Regression	1	1	1	4	7
ANN	3	2	2	3	10
Linear Regression	2	3	3	5	13
PNN	4	4	4	2	14
Gradient Boosted Tree Regression	5	5	5	1	16
Random Forest Regression	6	6	6	7	25
Simple Regression Tree	7	7	7	8	29
Polynomial Regression	8	8	8	6	30

Figure 3. Results of Algorithms in terms of minimization of statistical errors.

Using the Tree Ensemble Regression we have predicted the following percentage variation of the variable "*Concern about the deterioration of the landscape"* :

- *Piemonte* +13,92%;
- *Lombardia* -2,13%;
- *Veneto* +27,59%;
- *Abruzzo* +87,67%;
- *Molise* 222,58%;
- *Puglia* +209,57%.

Finally the mean value of the prediction is equal to 93,20% in the observed regions.

5. Conclusions

We estimate the Landscape and Cultural Heritage Index among Italian regions in the period 2004-2019 using data from ISTAT-BES. We use Panel Data with Fixed Effects, Panel Data with Random Effects, Pooled OLS, WLS, Dynamic Panel. We present a brief literature review considering the role of communitarian and environmental issues on landscape and cultural heritage preservation. Climate change is one of the main threats for landscape and cultural heritage but also socio-political issues, such as communitarian engagement, have a relevant role in promoting tools for conservation. In the second paragraph we estimate the value of Landscape and Cultural Heritage and we find that it is negatively associated with "*Dissatisfaction with the landscape of the place of life*", "*Illegal building*", "*Density and relevance of the museum heritage*", "*Internal material consumption*", "*Erosion of the rural space due to abandonment*", "*Availability of urban green*", and positively associated with "*Pressure from mining activities*", "*Erosion of the rural space by urban dispersion*", "*Concern about the deterioration of the landscape", "Diffusion of agritourism farms*", "*Current expenditure of the Municipalities for culture".* In the third paragraph we propose the application of a cluster analysis with the k-Means algorithm optimized with the Silhouette Coefficient and we found two clusters in the sense of "*Concern about the deterioration of the landscape*". We found that the Italian regions are essentially divided in two main parts: the Northern Italy with high levels of concern and the Southern Italy with lower level of concern. This contraposition suggests that the economic divide in terms of Gdp per capita and social and human capital operates as a determinant for the better performance of Northern Italian regions in respect to Southern Italian regions.
Finally, we use eight different machine learning algorithms to predict the level of "*Concern about the deterioration of the landscape*" and we found that the Tree Ensemble Regression is the best predictor. Furthermore, the predicted values suggest that the level of "*Concern about the deterioration of the landscape*" is expected to growth significantly.

Our analysis suggests that to promote conservation of landscape and cultural heritage is essential either to invest in cultural expenditure at a local level either to promote economic activities that are more oriented to environmental sustainability such as in the case of agritourism farms. Furthermore, the lack of human and social capital significantly reduces the possibility of southern Italian regions to effectively protect their landscape and cultural heritage. If policy makers are oriented to promote landscape and cultural heritage conservation they should invest more in human and social capital, promoting a deeper consciousness of preservation of cultural and environmental common goods among the communities.

6. Bibliography

Antonson, H., Buckland, P., & Nyqvist, R. (2021). A society ill-equipped to deal with the effects of climate change on cultural heritage and landscape: a qualitative assessment of planning practices in transport infrastructure. *Climatic Change, 1*, 166.

Assandri, G., Bogliani, G., & Pedrini, P. B. (2018). Beautiful agricultural landscapes promote cultural ecosystem services and biodiversity conservation. *Agriculture, Ecosystems & Environment, 256*, 200-210.

Bulian, G. (2021). Landscape, Cultural Heritage and Environmental Sustainability in Japan. *Ca' Foscari Japanese Studies | Religion and Thought , 14*(4), 111-133.

Cai, Z., Fang, C., Zhang, Q., & Chen, F. (2012). Joint development of cultural heritage protection and tourism: the case of Mount Lushan cultural landscape heritage site. *Heritage Science, 9*(1), 1-16.

Cicinelli, E., Salerno, G., & Caneva, G. (2018). An assessment methodology to combine the preservation of biodiversity and cultural heritage: the San Vincenzo al Volturno historical site (Molise, Italy). . *Biodiversity and conservation, 27*(5), 1073-1093.

Della Spina, L. (2017). Integrated evaluation and multi-methodological approaches for the enhancement of the cultural landscape. *International Conference on Computational Science and Its Applications*(Springer, Cham.), 478-493.

Foster, G. (2020). Circular economy strategies for adaptive reuse of cultural heritage buildings to reduce environmental impacts. *Resources, Conservation and Recycling*(104507), 152.

Gravagnuolo, A., & Girard, L. F. (2017). Multicriteria tools for the implementation of historic urban landscape. *Quality Innovation Prosperity, 21*(1), 186-201.

Guzman, P., Fatorić, S., & Ishizawa, M. (2020). Monitoring climate change in world heritage properties: Evaluating landscape-based approach in the state of conservation system. *Climate, 3*(29), 8.

Li, J., Krishnamurthy, S., Roders, A. P., & Van Wesemael, P. (2020). Community participation in cultural heritage management: A systematic literature review comparing Chinese and international practices. *Cities, 96*.

Ravankhah, M., Schmidt, M., & Will, T. (2021). An indicator-based risk assessment framework for World Heritage sites in seismic zones: The case of "Bam and its Cultural Landscape" in Iran . *International Journal of Disaster Risk Reduction*.

Rouhi, J. (2017). DEFINITION OF CULTURAL HERITAGE PROPERTIES AND THEIR VALUESBY THE PAST. *Asian Journal of Science and Technology , 08*(12), 7109-7114.

Shirvani Dastgerdi, A., Sargolini, M., Broussard Allred, S., Chatrchyan, A., & De Luca, G. (2020). Climate change and sustaining heritage resources: A framework for boosting cultural and natural heritage conservation in Central Italy. *Climate, 8*(2), 26.

Vallerani, F., & Visentin, F. (2018). *Waterways and the cultural landscape* . London, New York: Routledge.

7. Appendix

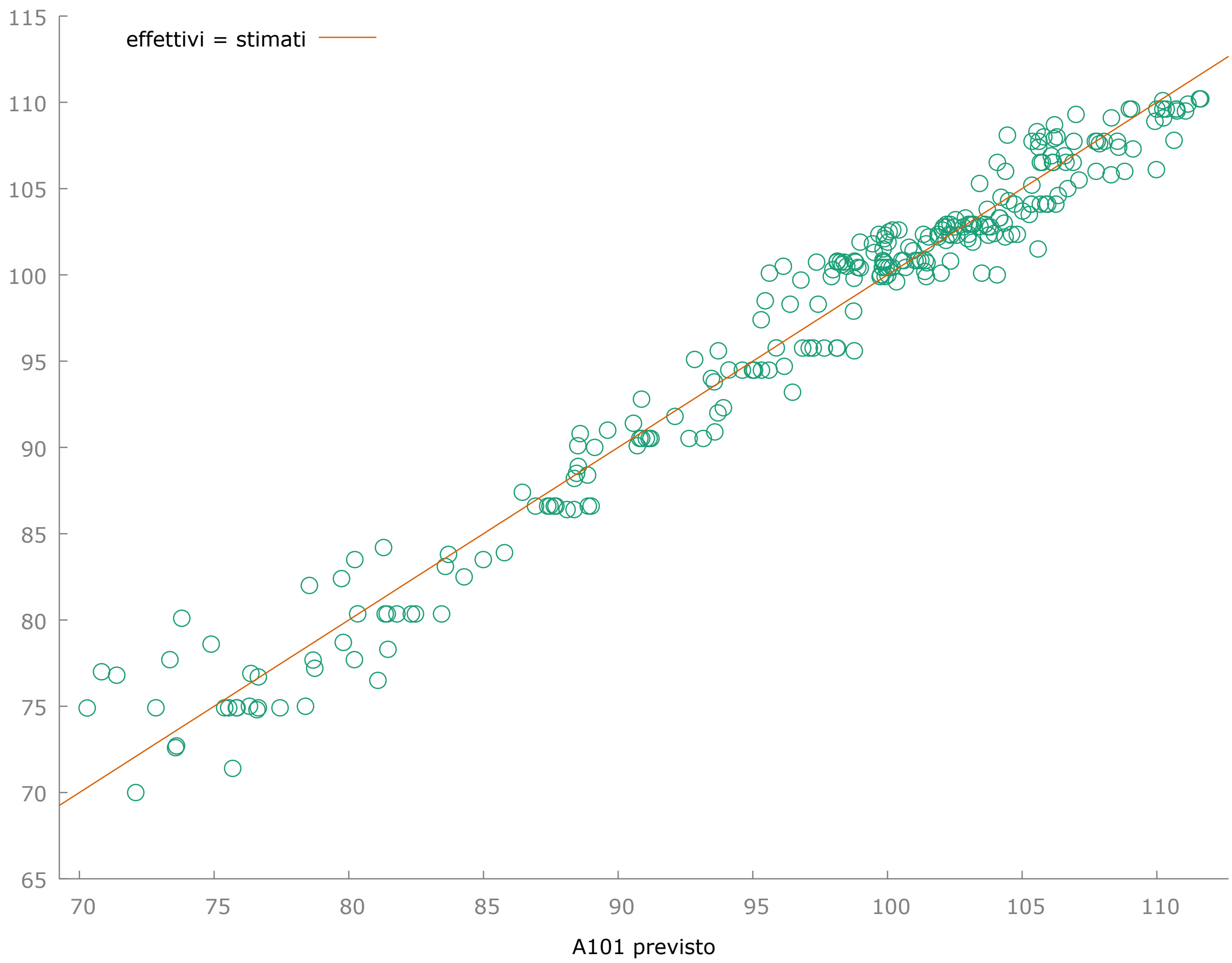
7.1 Econometric Models

Modello 19: Panel dinamico a un passo, usando 263 osservazioni
Incluse 20 unità cross section
Lunghezza serie storiche: minimo 12, massimo 14
Matrice H conforme ad Ox/DPD
Variabile dipendente: A101

	Coefficiente	*Errore Std.*	*z*	*p-value*	
A101(-1)	−0,0122381	0,00685145	−1,786	0,0741	*
const	0,0447119	0,0233032	1,919	0,0550	*
A90	0,429403	0,0163766	26,22	<0,0001	***
A91	−0,141787	0,0801861	−1,768	0,0770	*
A92	−0,159717	0,0208942	−7,644	<0,0001	***
A93	0,0357363	0,00635312	5,625	<0,0001	***
A94	−0,0299848	0,0105147	−2,852	0,0043	***
A95	0,0128409	0,00134943	9,516	<0,0001	***
A97	0,261535	0,0244100	10,71	<0,0001	***
A99	−0,624804	0,0396848	−15,74	<0,0001	***
A100	0,105488	0,0372153	2,835	0,0046	***
A102	−0,0603181	0,00846911	−7,122	<0,0001	***
A108	−0,00716397	0,00195655	−3,662	0,0003	***

Somma quadr. residui	901,5848	E.S. della regressione	1,899036

Numero di strumenti = 112
Test per errori AR(1): z = -3,91613 [0,0001]
Test per errori AR(2): z = -1,65382 [0,0982]
Test di sovra-identificazione di Sargan: Chi-quadro(99) = 140,204 [0,0041]
Test (congiunto) di Wald: Chi-quadro(12) = 486432 [0,0000]



Modello 20: Effetti fissi, usando 303 osservazioni
Incluse 20 unità cross section
Lunghezza serie storiche: minimo 14, massimo 16
Variabile dipendente: A101

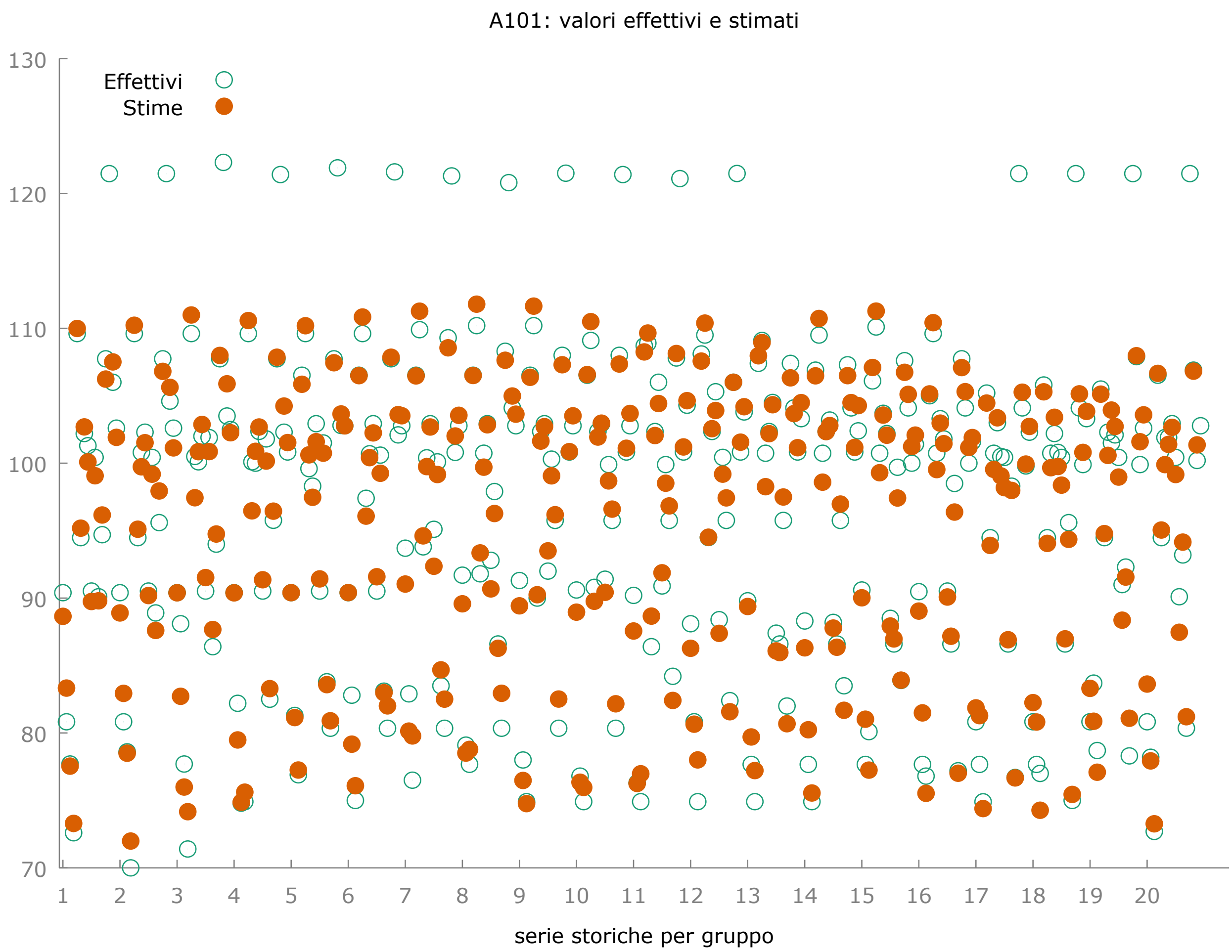
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	99,2770	1,18838	83,54	<0,0001	***
A90	0,438627	0,0181192	24,21	<0,0001	***
A91	−0,229870	0,0850603	−2,702	0,0073	***
A92	−0,139001	0,0116234	−11,96	<0,0001	***
A93	0,0318745	0,00983558	3,241	0,0013	***
A94	−0,0394458	0,00868176	−4,544	<0,0001	***
A95	0,0121328	0,00110433	10,99	<0,0001	***
A97	0,264261	0,0278792	9,479	<0,0001	***
A99	−0,647361	0,0288152	−22,47	<0,0001	***
A100	0,112826	0,0431383	2,615	0,0094	***
A102	−0,0570269	0,00795734	−7,167	<0,0001	***
A108	−0,00998954	0,000897042	−11,14	<0,0001	***

Media var. dipendente	95,21914	SQM var. dipendente	10,67528

Somma quadr. residui	576,9482	E.S. della regressione	1,456411
R-quadro LSDV	0,983236	R-quadro intra-gruppi	0,983134
LSDV F(30, 272)	531,7826	P-value(F)	4,6e-223
Log-verosimiglianza	−527,5074	Criterio di Akaike	1117,015
Criterio di Schwarz	1232,140	Hannan-Quinn	1163,073
rho	−0,093748	Durbin-Watson	2,026913

Test congiunto sui regressori -
Statistica test: F(11, 272) = 1441,4
con p-value = P(F(11, 272) > 1441,4) = 5,76456e-234

Test per la differenza delle intercette di gruppo -
Ipotesi nulla: i gruppi hanno un'intercetta comune
Statistica test: F(19, 272) = 0,438887
con p-value = P(F(19, 272) > 0,438887) = 0,981311



Modello 21: Effetti casuali (GLS), usando 303 osservazioni
Incluse 20 unità cross section
Lunghezza serie storiche: minimo 14, massimo 16
Variabile dipendente: A101

	Coefficiente	*Errore Std.*	*z*	*p-value*	
const	98,9730	1,14924	86,12	<0,0001	***
A90	0,441137	0,0176381	25,01	<0,0001	***
A91	−0,226927	0,0818594	−2,772	0,0056	***
A92	−0,138063	0,0112887	−12,23	<0,0001	***
A93	0,0311435	0,00952634	3,269	0,0011	***
A94	−0,0383875	0,00849684	−4,518	<0,0001	***
A95	0,0123628	0,00106783	11,58	<0,0001	***
A97	0,263085	0,0271491	9,690	<0,0001	***
A99	−0,642133	0,0278004	−23,10	<0,0001	***
A100	0,117746	0,0417925	2,817	0,0048	***
A102	−0,0581511	0,00773319	−7,520	<0,0001	***
A108	−0,00991231	0,000873439	−11,35	<0,0001	***

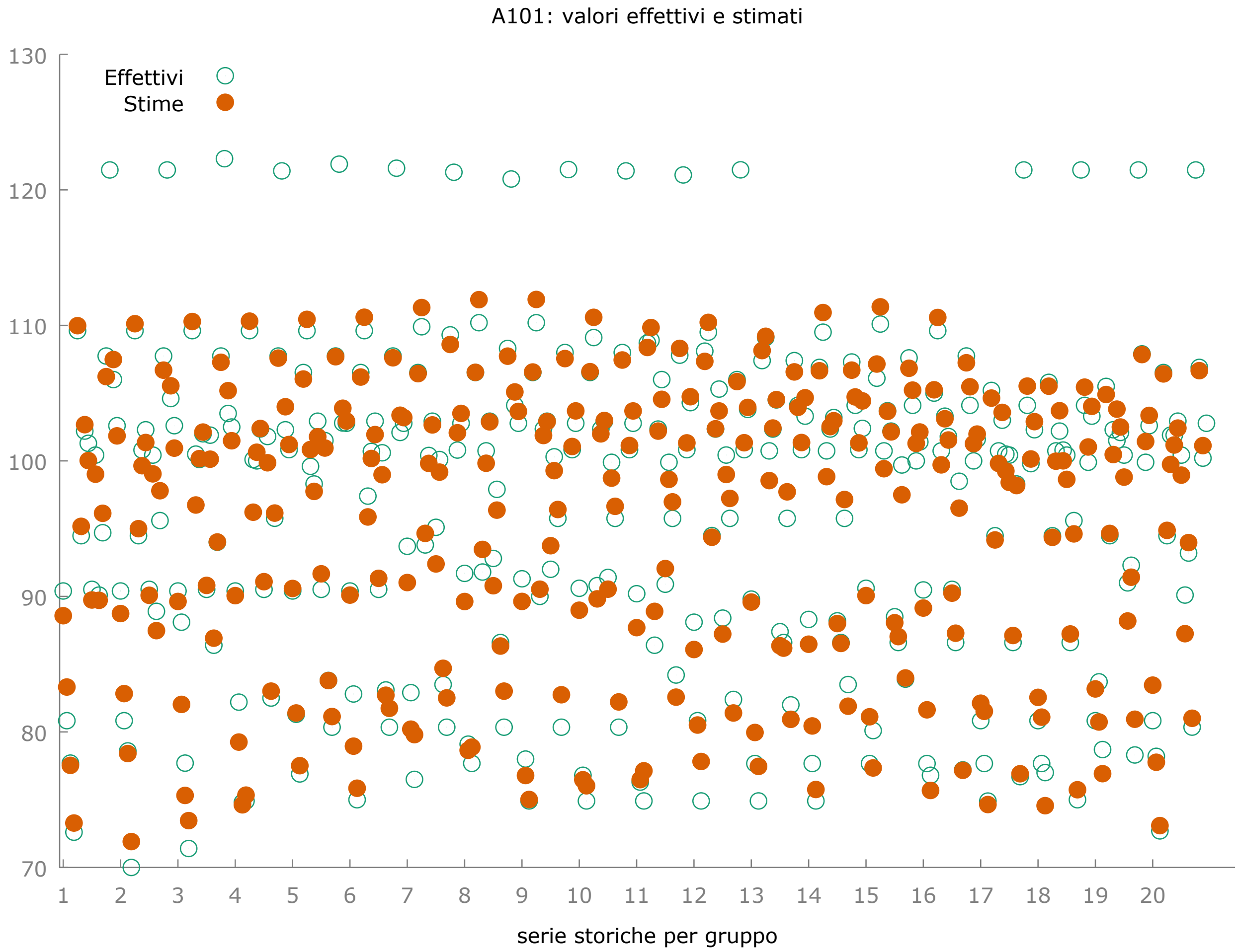
Media var. dipendente	95,21914	SQM var. dipendente	10,67528
Somma quadr. residui	594,6360	E.S. della regressione	1,427034
Log-verosimiglianza	−532,0822	Criterio di Akaike	1088,164
Criterio di Schwarz	1132,729	Hannan-Quinn	1105,993
rho	−0,093748	Durbin-Watson	2,026913

Varianza 'between' = 0
Varianza 'within' = 2,12113
theta medio = 0

Test congiunto sui regressori -
Statistica test asintotica: Chi-quadro(11) = 16551,5
con p-value = 0

Test Breusch-Pagan -
Ipotesi nulla: varianza dell'errore specifico all'unità = 0
Statistica test asintotica: Chi-quadro(1) = 3,34751
con p-value = 0,0673066

Test di Hausman -
Ipotesi nulla: le stime GLS sono consistenti
Statistica test asintotica: Chi-quadro(11) = 7,02169
con p-value = 0,797324

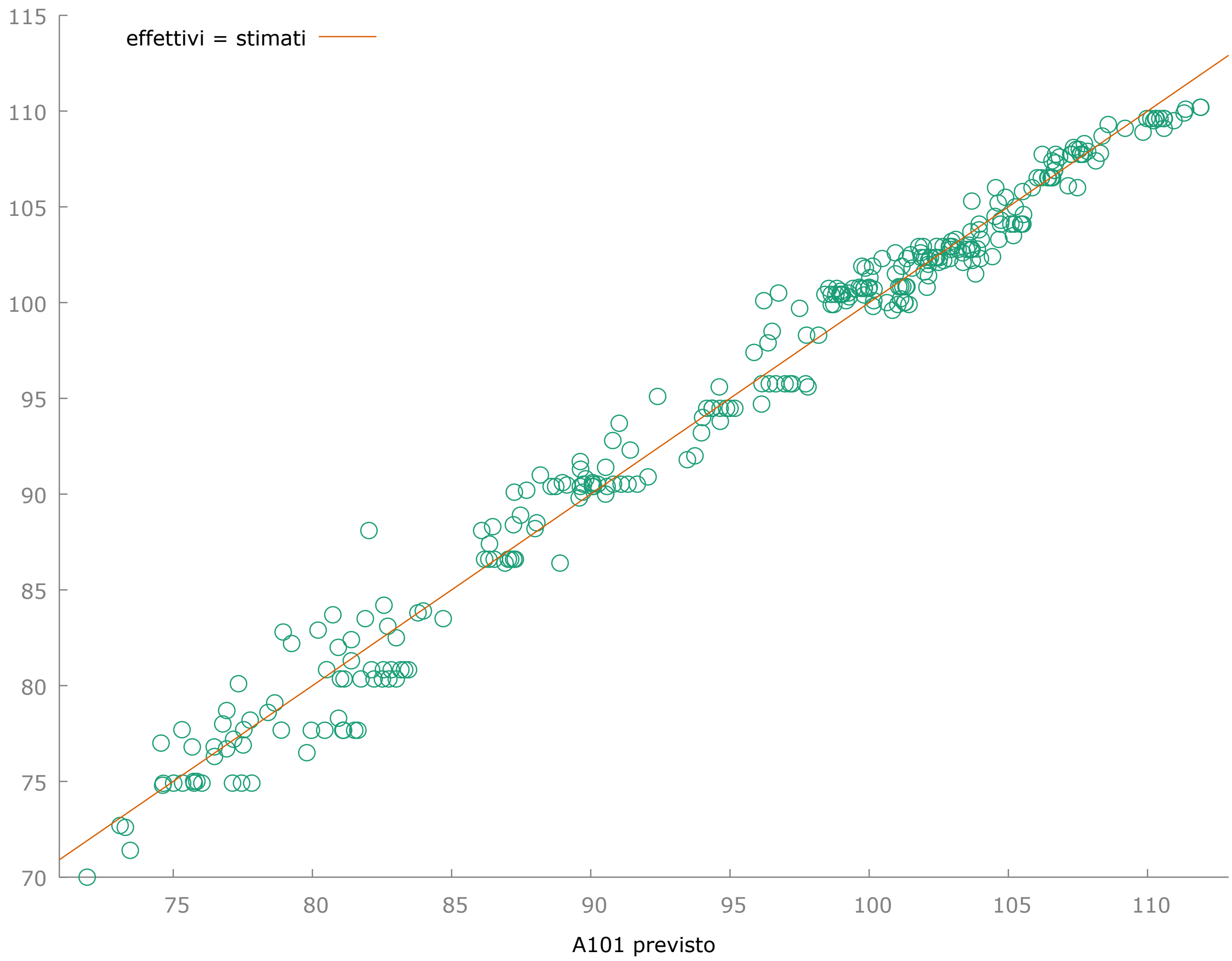


Modello 22: Pooled OLS, usando 303 osservazioni
Incluse 20 unità cross section
Lunghezza serie storiche: minimo 14, massimo 16
Variabile dipendente: A101

	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	98,9730	1,14924	86,12	<0,0001	***
A90	0,441137	0,0176381	25,01	<0,0001	***
A91	−0,226927	0,0818594	−2,772	0,0059	***
A92	−0,138063	0,0112887	−12,23	<0,0001	***
A93	0,0311435	0,00952634	3,269	0,0012	***
A94	−0,0383875	0,00849684	−4,518	<0,0001	***
A95	0,0123628	0,00106783	11,58	<0,0001	***
A97	0,263085	0,0271491	9,690	<0,0001	***
A99	−0,642133	0,0278004	−23,10	<0,0001	***
A100	0,117746	0,0417925	2,817	0,0052	***
A102	−0,0581511	0,00773319	−7,520	<0,0001	***
A108	−0,00991231	0,000873439	−11,35	<0,0001	***

Media var. dipendente	95,21914	SQM var. dipendente	10,67528

Somma quadr. residui	594,6360	E.S. della regressione	1,429483
R-quadro	0,982722	R-quadro corretto	0,982069
F(11, 291)	1504,684	P-value(F)	3,7e-249
Log-verosimiglianza	−532,0822	Criterio di Akaike	1088,164
Criterio di Schwarz	1132,729	Hannan-Quinn	1105,993
rho	−0,059945	Durbin-Watson	1,968098



Modello 23: WLS, usando 303 osservazioni
Incluse 20 unità cross section
Variabile dipendente: A101
Pesi basati sulle varianze degli errori per unità

	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	98,8187	1,04717	94,37	<0,0001	***
A90	0,437829	0,0159993	27,37	<0,0001	***
A91	−0,253816	0,0735367	−3,452	0,0006	***
A92	−0,131504	0,0103080	−12,76	<0,0001	***
A93	0,0316906	0,00858405	3,692	0,0003	***
A94	−0,0392682	0,00773810	−5,075	<0,0001	***

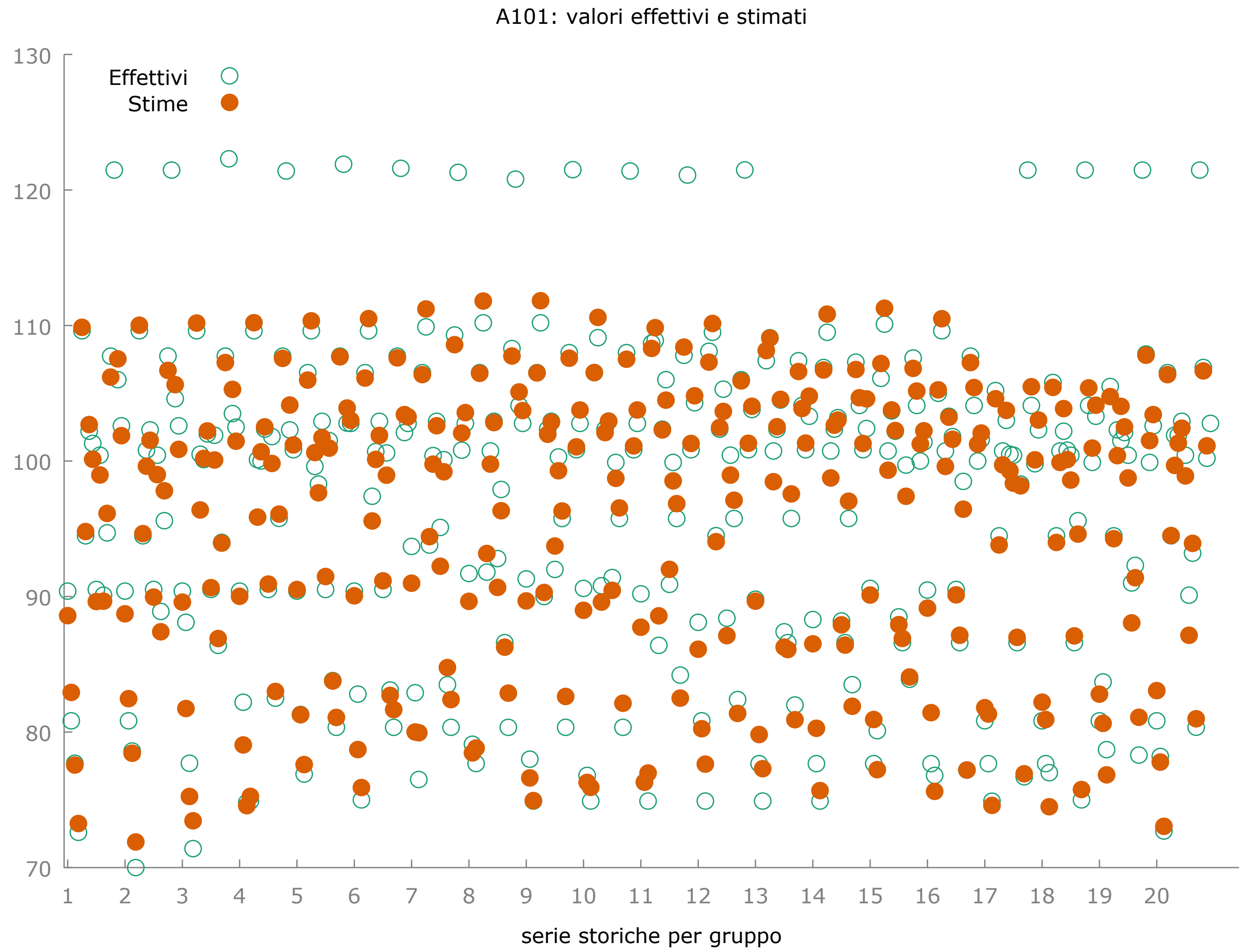
A95	0,0123490	0,000997927	12,37	<0,0001	***
A97	0,275149	0,0243874	11,28	<0,0001	***
A99	−0,656870	0,0251274	−26,14	<0,0001	***
A100	0,137727	0,0394509	3,491	0,0006	***
A102	−0,0561594	0,00719014	−7,811	<0,0001	***
A108	−0,0104049	0,000824957	−12,61	<0,0001	***

Statistiche basate sui dati ponderati:

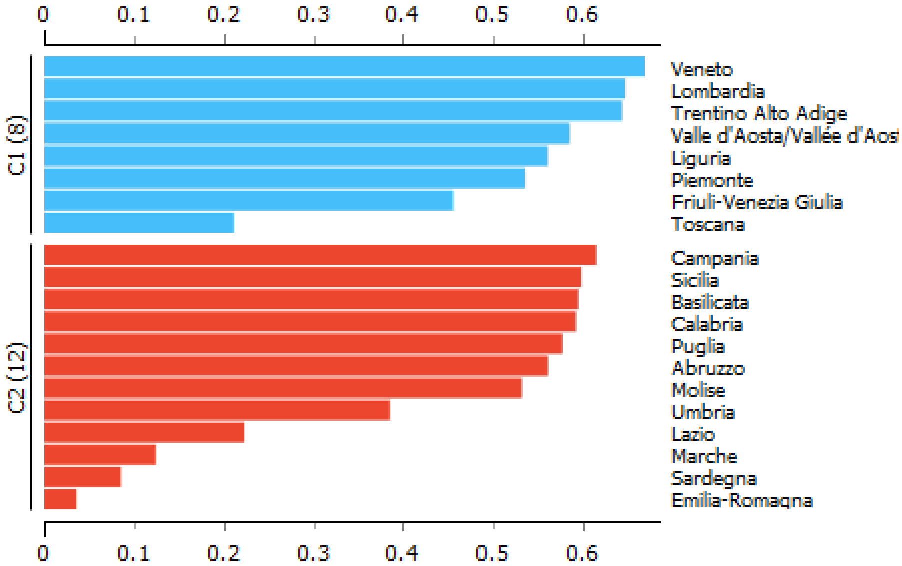
Somma quadr. residui	300,2794	E.S. della regressione	1,015819
R-quadro	0,985821	R-quadro corretto	0,985285
F(11, 291)	1839,346	P-value(F)	1,2e-261
Log-verosimiglianza	−428,5719	Criterio di Akaike	881,1438
Criterio di Schwarz	925,7086	Hannan-Quinn	898,9728

Statistiche basate sui dati originali:

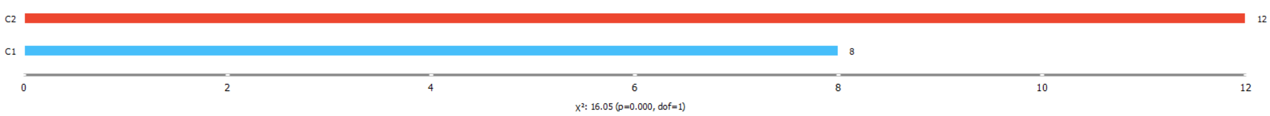
Media var. dipendente	95,21914	SQM var. dipendente	10,67528
Somma quadr. residui	599,4535	E.S. della regressione	1,435262

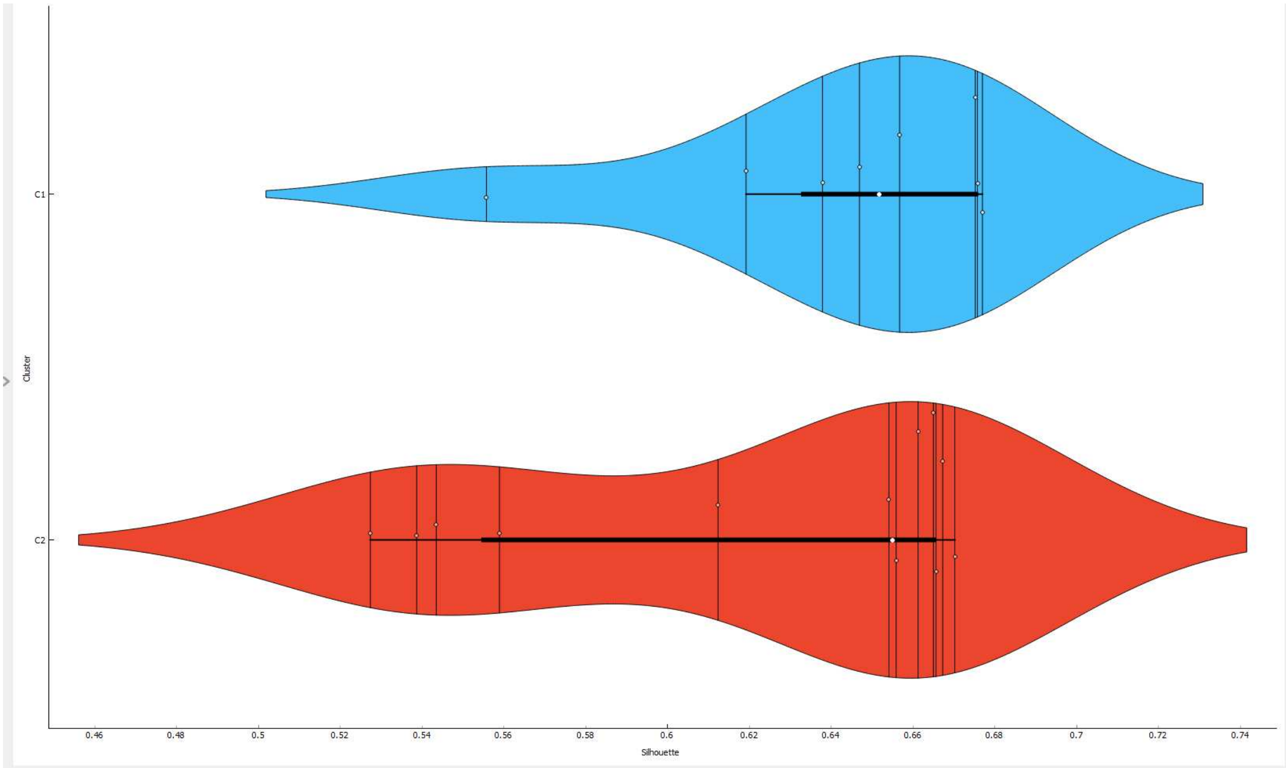


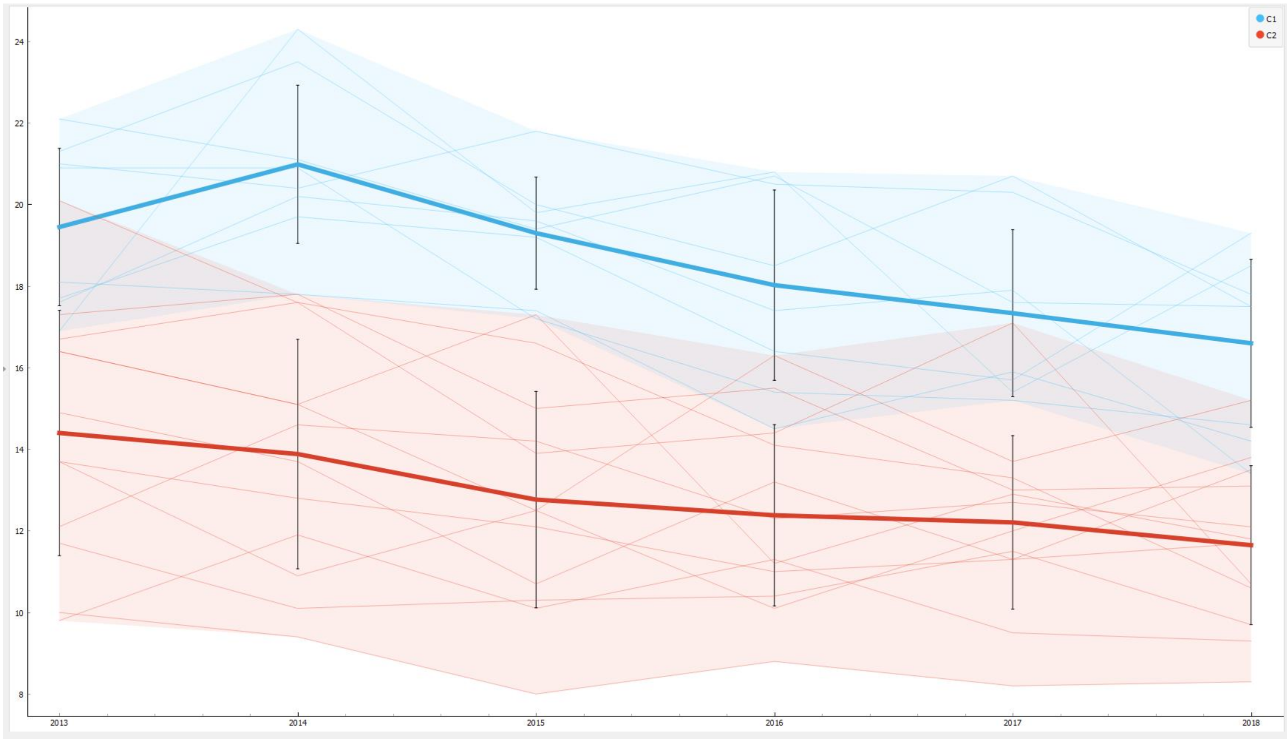
7.2 Clusterization

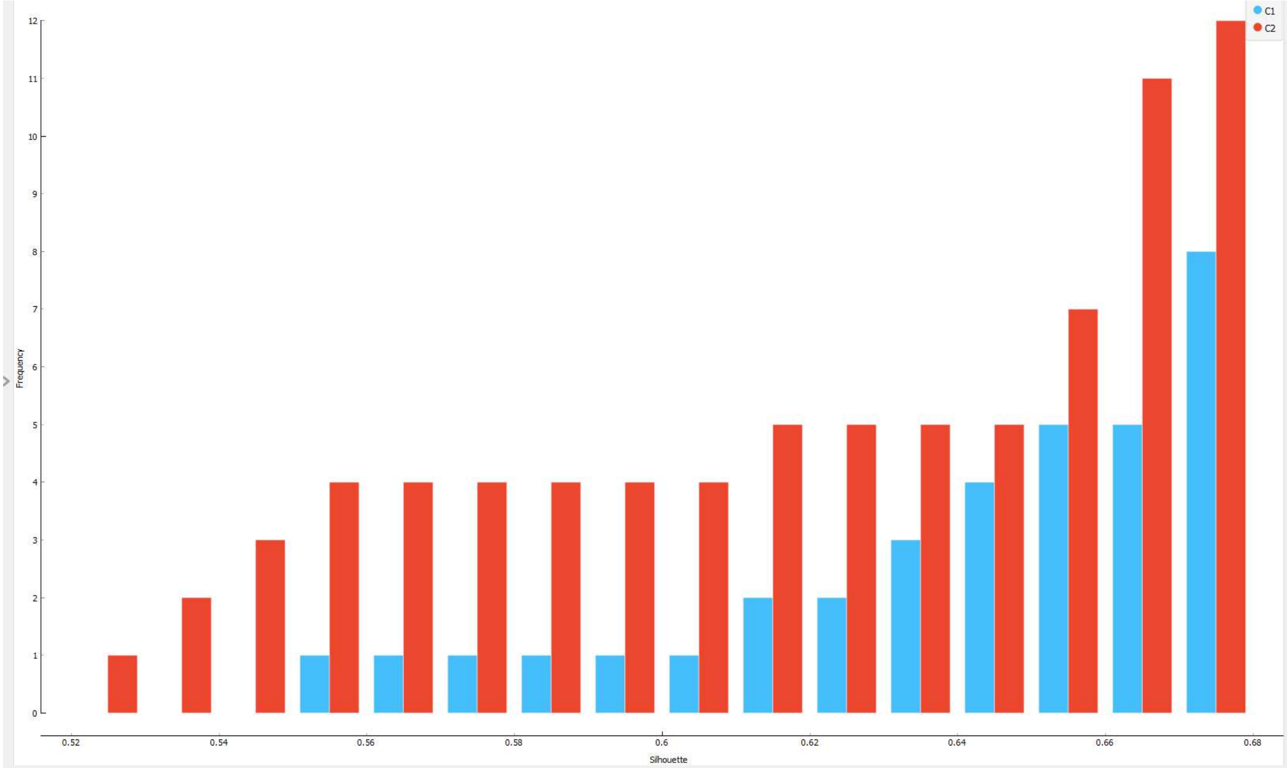


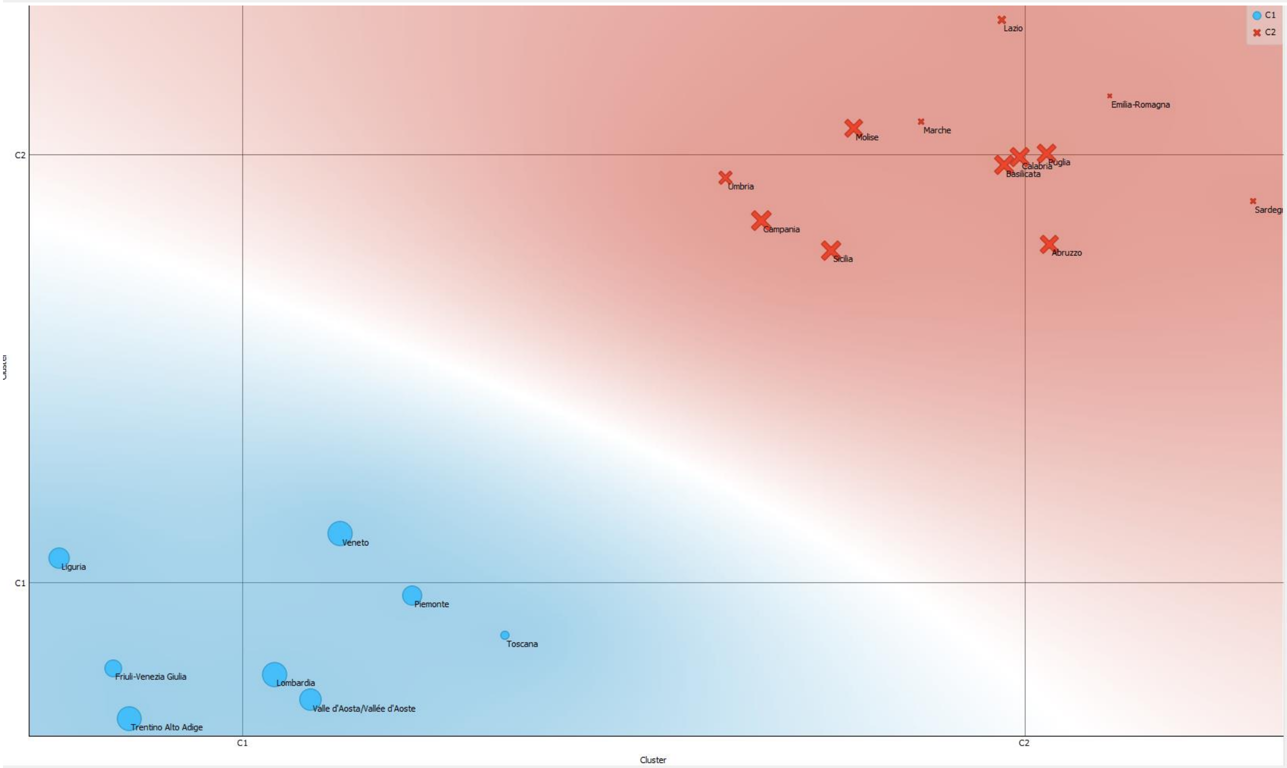
	2019	Regions	Cluster
1	13.1	Piemonte	C1
2	13.8	Valle d'Aosta/V...	C1
3	14.7	Liguria	C1
4	14.8	Lombardia	C1
5	13.3	Veneto	C1
6	12.4	Friuli-Venezia G...	C1
8	11.7	Toscana	C1
20	15.5	Trentino Alto A...	C1
7	12.6	Emilia-Romagna	C2
9	13.1	Umbria	C2
10	10.6	Marche	C2
11	13.4	Lazio	C2
12	10.5	Abruzzo	C2
13	9.3	Molise	C2
14	10.4	Campania	C2
15	9.7	Puglia	C2
16	9.1	Basilicata	C2
17	10.6	Calabria	C2
18	11.0	Sicilia	C2
19	12.7	Sardegna	C2

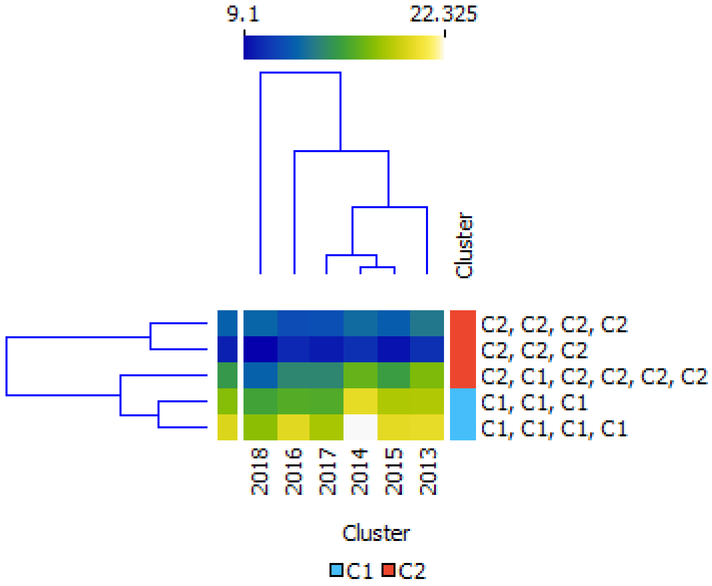



C1
C2
Cluster
0.46
0.48
0.5
0.52
0.54
0.56
0.58
0.6
0.62
0.64
0.66
0.68
0.7
0.72
0.74
Silhouette


C1
C2
24
22
20
18
16
14
12
10
8
2013
2014
2015
2016
2017
2018


C1
C2
Frequency
Silhouette
0.52
0.54
0.56
0.58
0.6
0.62
0.64
0.66
0.68

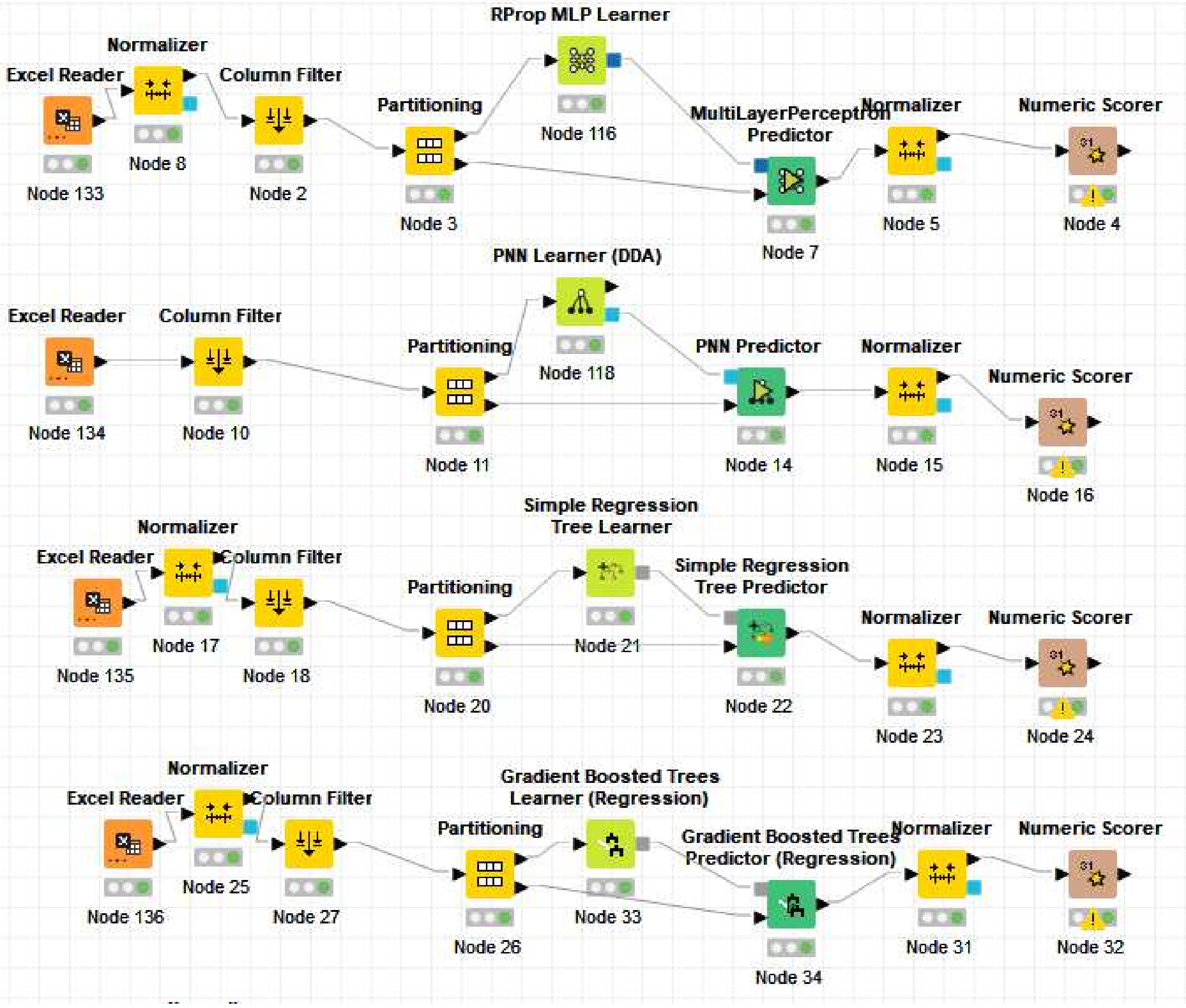

C1
C2
Lazio
Emilia-Romagna
Marche
Molise
Umbria
Calabria
Basilicata
Puglia
Campania
Sicilia
Abruzzo
Veneto
Liguria
Piemonte
Toscana
Friuli-Venezia Giulia
Lombardia
Valle d'Aosta/Vallée d'Aoste
Trentino Alto Adige
Cluster

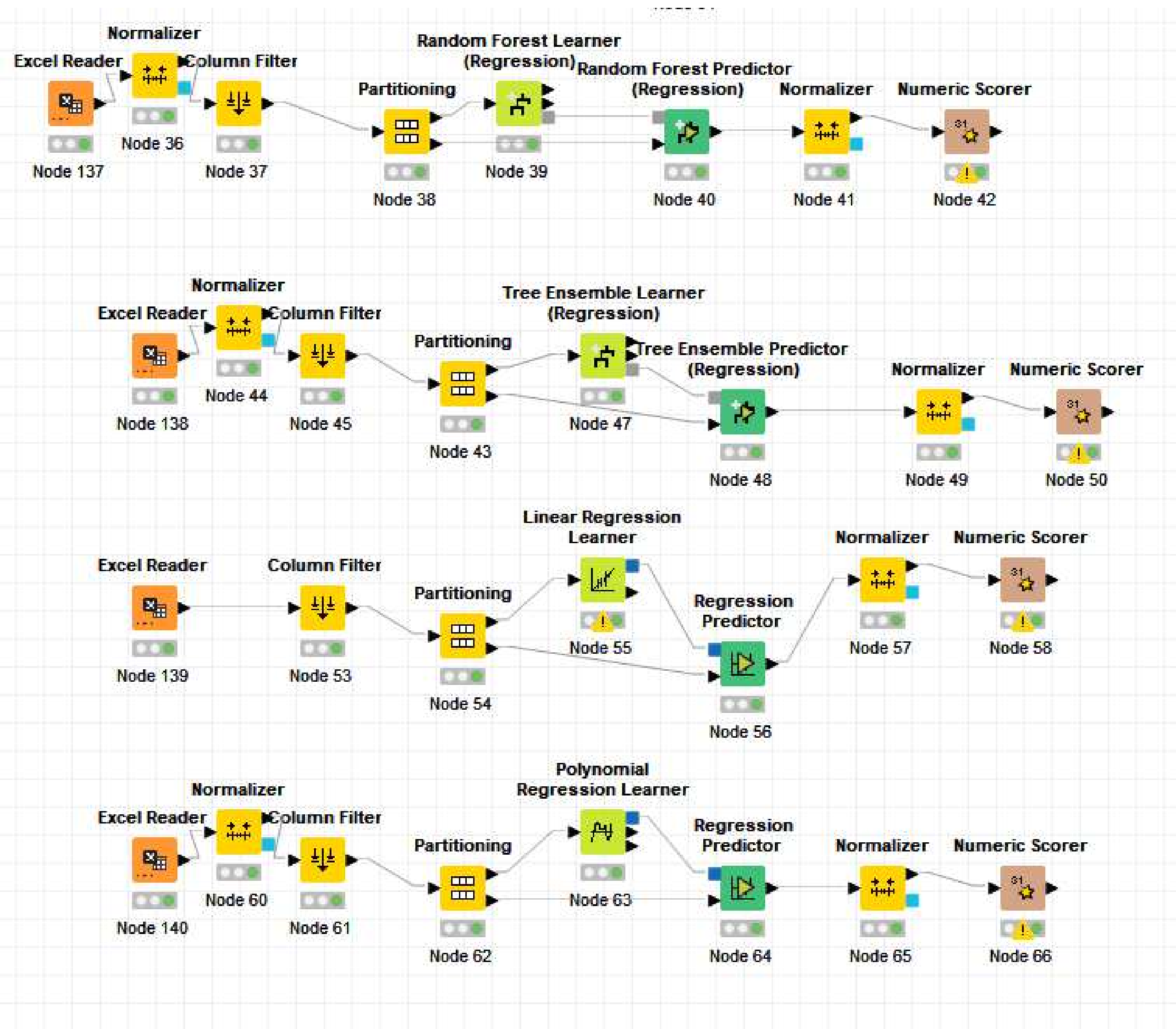


7.3 Machine Learning and Predictions

Tree Ensemble Learner Prediction			
	2019	Prediction	Var %
Piemonte	0,625	0,712	13,92
Lombardia	0,891	0,872	-2,13
Veneto	0,656	0,837	27,59
Abruzzo	0,219	0,411	87,67
Molise	0,031	0,1	222,58
Puglia	0,094	0,291	209,57
Mean			93,20

Results of Algorithms in Terms of Statistical Errors				
	ANN	PNN	Simple Regression Tree	Gradient Boosted Tree Regression
Mean absolute error	0,163427393	0,202020202	0,258397109	0,207428896
Mean squared error	0,042202598	0,053165969	0,132856902	0,080853627
Root mean squared error	0,205432708	0,23057747	0,364495407	0,284347721
Mean signed difference	0,104947845	-0,094043887	0,231186224	0,029165527
	Random Forest Regression	Tree Ensemble Regression	Linear Regression	Polynomial Regression
Mean absolute error	0,257273171	0,114827563	0,154190232	0,485000451
Mean squared error	0,088266818	0,021129288	0,043904432	0,315632674
Root mean squared error	0,297097321	0,145359169	0,209533845	0,561811956
Mean signed difference	0,193613004	0,114827563	0,135876482	0,151667118


RProp MLP Learner
Normalizer
Excel Reader
Column Filter
Partitioning
Node 116
MultiLayerPerceptron Predictor
Normalizer
Numeric Scorer
Node 8
Node 133
Node 2
Node 3
Node 5
Node 4
Node 7
PNN Learner (DDA)
Excel Reader
Column Filter
Partitioning
Node 118
PNN Predictor
Normalizer
Numeric Scorer
Node 134
Node 10
Node 11
Node 14
Node 15
Node 16
Simple Regression Tree Learner
Normalizer
Excel Reader
Column Filter
Simple Regression Tree Predictor
Partitioning
Normalizer
Numeric Scorer
Node 17
Node 21
Node 135
Node 18
Node 20
Node 22
Node 23
Node 24
Normalizer
Gradient Boosted Trees Learner (Regression)
Excel Reader
Column Filter
Partitioning
Gradient Boosted Trees Predictor (Regression)
Normalizer
Numeric Scorer
Node 25
Node 136
Node 27
Node 33
Node 26
Node 31
Node 32
Node 34


Normalizer
Excel Reader
Column Filter
Random Forest Learner (Regression)
Random Forest Predictor (Regression)
Partitioning
Normalizer
Numeric Scorer
Node 137
Node 36
Node 37
Node 38
Node 39
Node 40
Node 41
Node 42
Normalizer
Excel Reader
Column Filter
Tree Ensemble Learner (Regression)
Tree Ensemble Predictor (Regression)
Partitioning
Normalizer
Numeric Scorer
Node 138
Node 44
Node 45
Node 43
Node 47
Node 48
Node 49
Node 50
Linear Regression Learner
Excel Reader
Column Filter
Partitioning
Regression Predictor
Normalizer
Numeric Scorer
Node 139
Node 53
Node 54
Node 55
Node 56
Node 57
Node 58
Polynomial Regression Learner
Normalizer
Excel Reader
Column Filter
Partitioning
Regression Predictor
Normalizer
Numeric Scorer
Node 140
Node 60
Node 61
Node 62
Node 63
Node 64
Node 65
Node 66