



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

Multilevel-growth modelling for the study of sustainability transitions



Matteo Mura*, Mariolina Longo, Laura Toschi, Sara Zanni, Franco Visani, Silvia Bianconcini

University of Bologna, Italy

A B S T R A C T

Sustainability Transitions (ST) is a complex phenomenon, encompassing environmental, societal and economic aspects. Its study requires a proper investigation, with the identification of a robust indicator and the definition of a suitable method of analysis. To identify the most informative geographical boundaries for analysing ST pathways, we consider the Carbon Emission Intensity (CEI) and estimate a four-level growth model to study its pattern over time for all the EU regions. We apply this model to a novel longitudinal dataset that covers CEI data of European regions at four different geographical scales (state, areas, regions, and provinces) over a nine-year timespan. This approach aims at supporting the decision-makers in developing more effective sustainability transitions policies across Europe, especially focusing on regions and overcoming the well-known “one-size fits all” approach.

- *The unconditional growth model has been applied to a multi-level structure considering four levels, defined by three geographical scales and time.*
- *The ideal structure of the model would have required five levels, but the sample size of the dataset made the application computationally unfeasible;*
- *The application of the model allowed to identify patterns of stability and change over time of the variable amongst different geographical units.*

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

A R T I C L E I N F O

Method name: Four-level growth model for Sustainability Transition pathways

Keywords: Multilevel-growth model, Sustainability transitions, Carbon emission intensity, Time dependence, Transition Pathways

Article history: Received 3 March 2021; Accepted 14 April 2021; Available online 22 April 2021

DOI of original article: [10.1016/j.ecolecon.2021.106968](https://doi.org/10.1016/j.ecolecon.2021.106968)

* Corresponding author.

E-mail address: matteo.mura@unibo.it (M. Mura).

<https://doi.org/10.1016/j.mex.2021.101359>

2215-0161/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject Area:	Economics and Finance
More specific subject area:	Sustainability Transitions studies
Method name:	Four-level growth model for Sustainability Transition pathways
Name and reference of original method:	Bollen and Curran [3]. <i>Latent growth models: a structural equation perspective</i> . Wiley and Sons.
	Curran et al. [9]. <i>A multivariate growth curve model for three-level data</i> . In H. Cooper (Ed.) <i>APA Handbook of Research Methods in Psychology</i> , Volume 3 (pp.335–358). Washington, DC: American Psychological Association.
	Tasca et al. [23]. <i>Three-level multilevel growth models for nested change data: A guide for group treatment researchers</i> , <i>Psychotherapy Research</i> , 19: 4, 453 – 461.
Resource availability:	Data are available from the authors by request and in Mura et al. [19] “Industrial Carbon Emission Intensity: a comprehensive dataset of European Regions”, <i>Data in Brief</i> (submitted), whereas the R code used to perform the analysis is detailed in the paper (section “Model Code”).

Background

Sustainability Transitions (ST) are generated by technological, cultural, political and institutional forces that drive significant shifts of entire industries towards more sustainable development models [11,13,21]. ST are spatial processes [7] that depend on actions taken at different geographical scales: urban, regional, national, and global. This concept of geographical multiscale has been already investigated in the literature [18], but most of the studies have just focused on the country level, by analysing single countries or comparing ST on few countries [14,17,22]. Very limited research has investigated the role played by different geographical scales in affecting ST pathways.

This paper aims at presenting the method used by Mura et al. [19] to analyse how different geographical scales affect the development of ST. The paper is focused on the industrial sector, with a specific attention on the Carbon Emission Intensity (CEI) of European regions at four different geographical scales: NUTS 0¹ (country), NUTS 1 (macro-region), NUTS 2 (region), NUTS 3 (province), from 2008 to 2016.

The authors specify a multilevel-growth model [23] as the most appropriate for their study, being the dataset characterised by a nested geographical structure and temporal dependant observations. Multilevel-growth modelling is regarded as one of most powerful and informative approaches for analysing unbalanced repeated measures [4,9]. Indeed, multilevel-growth models allow researchers to test the developmental trajectories of constructs across time [8], while simultaneously accounting for the within- and between-individual differences that stem from changes of longitudinal variables [5]. Thus, compared to conventional repeated measures analyses, multilevel-growth modelling can provide more adequate information about temporal changes in a specific variable. In this specific case, the multilevel growth model technically involves specifying a ST trajectory for each NUTS: that is, estimating NUTS-specific intercepts and slopes, under the main assumption that NUTS ST trajectories share a common functional form [3]. A peculiar characteristic of these models is that they allow to capture differences in both the mean and variance structures that are related to patterns of stability and change over time amongst the several NUTS.

Method details

In order to identify the effects of multiple geographical scales on ST pathways, the authors estimate a four-level growth model for (CEI) over time [9]. CEI is defined as the ratio between Carbon Dioxide equivalent emissions² and Gross Domestic Product. It has been indicated as a robust indicator of Sustainability Transitions at the level of countries, regions, and industrial values chains [1,6,10,24]. The study of ST is performed by estimating an overall average ST trajectory common to the whole

¹ From French: Nomenclature des Unités Territoriales Statistiques

² Carbon dioxide equivalent emission (CO₂e) is the measure of greenhouse gases emissions (GHGs) reported in terms of tonnes CO₂, considering characterisations factors based on the global warming potential of each gas in respect to CO₂.

EU and then analysing at which geographical scale the ST pathways deviate from that overall mean trajectory. Once it is evaluated that the unit trajectories at one specific level differ from the overall mean pattern, the model can be extended to identify systematic factors that drive the ST patterns at each level.

The structure of the dataset built over five levels (i.e., NUTS 0-1-2-3 and Time) would require the implementation of a five-level unconditional growth model. However, given the large number of observations in the dataset, the estimation of a five-level model is computationally unfeasible. Hence, the authors have fitted three different four-level models, based on the following nested structures: NUTS 1-2-3 and Time (Model 1); NUTS 0-1-2 and Time (Model 2); and NUTS 0-2-3 and Time (Model 3).

Models 1–3 are specified as follows

$$y_{tjki} = \pi_{0jki} + \pi_{1jki} \text{time}_{tjki} + e_{tjki} \quad t = 1, \dots, T_{jki}; j = 1, \dots, N_k; k = 1, \dots, N_i; i = 1, \dots, N$$

where t represents the time, and j , k , and i denote the different NUTS levels taken into consideration. In particular, y_{tjki} is the CEI measured at time t for:

- NUTS 3 (j) nested in NUTS 2 (k) nested in NUTS 1 (i) for Model 1;
- NUTS 2 (j) nested in NUTS 1 (k) nested in NUTS 0 (i) for Model 2;
- NUTS 3 (j) nested in NUTS 2 (k) nested in NUTS 0 (i) for Model 3.

π_{0jki} and π_{1jki} are the intercept and slope specific for each NUTS 3 (j) (or NUTS 2 (j) for Model 2), nested in NUTS 2 (k) (or NUTS 1 (k) in Model 2), nested in NUTS 0 (i) (or NUTS 1 in Model 1).

time_{tjki} refers to the time occasion t observed at NUTS 3 (j) (or NUTS 2 (j) in Model 2) nested in the NUTS 2 (k) (or NUTS 1 (k) in Model 2) nested in NUTS 0 (i) (or NUTS 1 (i) for Model 1). For our specific analysis (i.e., for all j , k and i), t refers to each year in the span 2008–2016.

Finally, e_{tjki} is the residual specific to time t , NUTS 3 (j) (or NUTS 2 (j) in Model 2), NUTS 2 (k) (or NUTS 1 (k) in Model 2), and c NUTS 0 (i) (or NUTS 1 (i) for Model 1). It is assumed to be normally distributed with a zero mean and homoscedastic variance.

The following level-2 equations allow us to evaluate differences in the ST trajectories at the j level (NUTS3 and NUTS2 for Models 1 and 3 and Model 2, respectively):

$$\pi_{0jki} = \beta_{0ki} + r_{0jki} \tag{1}$$

$$\pi_{1jki} = \beta_{1ki} + r_{1jki} \tag{2}$$

β_{0ki} and β_{1ki} are the intercept and slope specific to the NUTS 2 (or NUTS 1) linear trajectory. They represent the means of the intercepts and slopes of the growth trajectories for all of the NUTS 3 (or NUTS 2) nested within NUTS 2 (or NUTS 1) k grouped into NUTS 1 (NUTS 0) i . As such, the residuals r_{0jki} and r_{1jki} represent the deviation of each NUTS 3 (or NUTS 2) intercept and slope around their group-specific mean values. More formally, this is given as $[r_{0jki}, r_{1jki}]$ is assumed to follow a bivariate normal distribution with null mean vector and full covariance matrix T_π .

Differences in the ST temporal patterns at the k level are studied by analysing the estimates of the parameters involved in the level-3 equations, specified as

$$\beta_{0ki} = \alpha_{0i} + s_{0ki} \tag{3}$$

$$\beta_{1ki} = \alpha_{1i} + s_{1ki} \tag{4}$$

As for the previous levels, α_{0i} and α_{1i} are the intercept and slope specific to the NUTS 1 (or NUTS 0) linear trajectory. They represent the means of the intercepts and slopes of the growth trajectories for all of the NUTS 2 (or NUTS 1) k nested within NUTS 1 (or NUTS 0) i . As such, the residuals s_{0ki} and s_{1ki} represent the deviation of each NUTS 2 (or NUTS 1) intercept and slope around their group-specific mean values. More formally, this is given as $[s_{0ki}, s_{1ki}]$ is assumed to follow a bivariate normal distribution with null mean vector and full covariance matrix T_β .

Finally, given the four-level structure of the data, the group-specific intercepts and slopes (e.g. α_{0i} and α_{1i}) themselves vary randomly across groups. The level-4 equations are thus

$$\alpha_{0i} = \gamma_0 + u_{0i} \quad (5)$$

$$\alpha_{1i} = \gamma_1 + u_{1i} \quad (6)$$

where γ_0 and γ_1 represent the grand mean intercept and slope pooling over all the European Community. For the specific available data, γ_0 represents the starting point of CEI at time 2008 for all the European countries, whereas γ_1 is their average growth rate over time. The residual terms u_{0i} and u_{1i} capture the deviation of each value from the grand means that is specific to NUTS 1 (or NUTS 0), while $[u_{0i}, u_{1i}]$ follows a bivariate normal distribution with null mean vector and full covariance matrix T_α .

In sum, γ_0 and γ_1 represent the “fixed effects” of the model, which estimate the expected trajectories for the overall sample. This common trajectory is identified by estimating the expected intercept γ_0 of the CEI in 2008 and the expected slope γ_1 for the whole EU. On the other hand, the variance and covariance of the “random effects” (T_π , T_β , and T_α) estimates, for each geographical scale included in the model, if the individual trajectories at that specific level significantly differ from the overall one. Thus, looking at the corresponding p-values, we can see if there is significant variability at each geographical scale included in the model with respect to the overall European ST pattern. In other words, the “random effects” allow to test the statistical significance for both intercept and slope at the different scales considered.

Method application

Data about CO₂e were collected from the EU Emission Trading System (EU ETS) register. It is the biggest dataset about CO₂e industrial emissions in Europe, accounting for 45% of EU greenhouse gases (GHGs) emissions. Data are reported at the plant level, so they were then aggregated at the NUTS 3-2-1-0 levels, based on the plant’s location. Data about GDP were collected from the Eurostat database at the NUTS 3 level and then aggregated at higher levels (NUTS 0,1 and 2). The CEI data were then obtained as the ratio between CO₂e data and GDP data at different geographical scales. Therefore, the final dataset consisted of longitudinal data on CEI from 2008 to 2016, covering four different scales (i.e., NUTS 0-1-2-3): 28 NUTS 0, 103 NUTS 1, 279 NUTS 2 and 1248 NUTS 3. The total number of observations over the eight years covered by the analysis was 14,433.

Based on the empirical evidence, the growth curve presents a common functional form for all our units (i.e. NUTS) which is linear. To further corroborate the appropriateness of the linear growth model, the authors considered the empirical growth curves of all the observed NUTS2 within each country (NUTS0), and in Fig. 1 those related to Germany, Italy, Spain, and United Kingdom are reported. For simplicity, only these countries are reported since they are the most important ones, in terms of total emissions and time series of data available, and because similar conclusions can be also drawn for all the other countries. It is evident that at this level of analysis, the average CEI patterns appear linear.

By applying the model, it was possible to estimate the expected slope and intercept for each EU country trajectory [12,15]. While the slopes explain the convergence towards low-carbon economy, the intercepts represent the steady-state of a country when the remaining variables are equal to 0. For each country, these expected slopes and intercepts are illustrated in Figs. 2 and 3, respectively.

Both Figs. 2 and 3 show relevant differences between different countries. Looking at the slopes we can see countries with high negative values (Bulgaria, Romania and Slovenia, for instance, with values from -15.4 to -3.3) and others with slightly positive figures (as for Germany, Spain and Portugal, with values from 5.3 to 7.3). The first group shows a higher decrease of CEI compared to the overall European trend, while the second group shows a lower decrease.

As shown by Mura et al. [19] and reported by Fig. 4, it is possible to trace groups of countries with similar behaviour. Looking at the intercepts (Fig. 4), for example, we have a group of countries with very high values (Poland, Czech Republic, Greece, Estonia and Bulgaria, with values ranging from 170

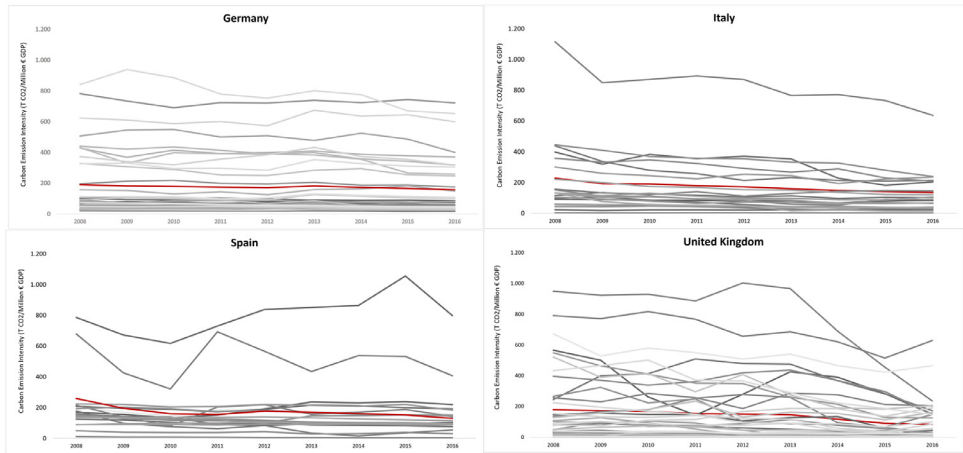


Fig. 1. Empirical growth curves of CEI (T CO₂ / Million € GDP) at NUTS 2 level for Germany, Italy, Spain and France. Each line, for each country, represents the trajectory of CEI for each region (NUTS 2). On average the patterns appear linear, thus supporting the development of a linear model.

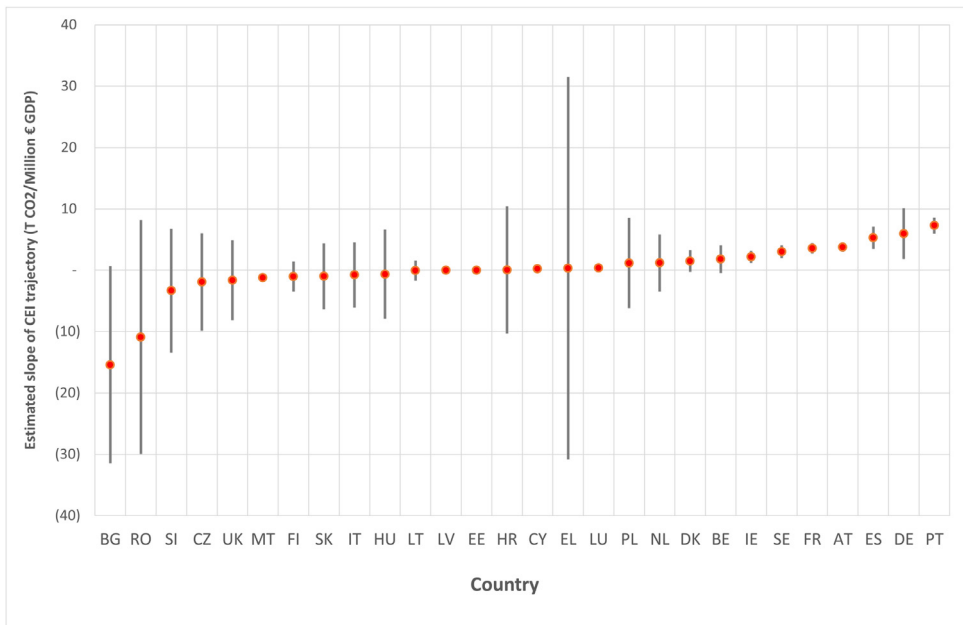


Fig. 2. Expected slope for each country ($\hat{\alpha}_{1i}$). The red dot represents the mean value of the estimated slope, while the black bar accounts for the variability of the estimated slope within the country (length equal to twice the standard deviation of the values at the NUTS 2 level). The countries are ordered on the basis of the average estimated slope, from the lowest one to the highest one. The countries with negative values show a decreasing trend for CEI, whereas the countries with positive values show a growing trend.

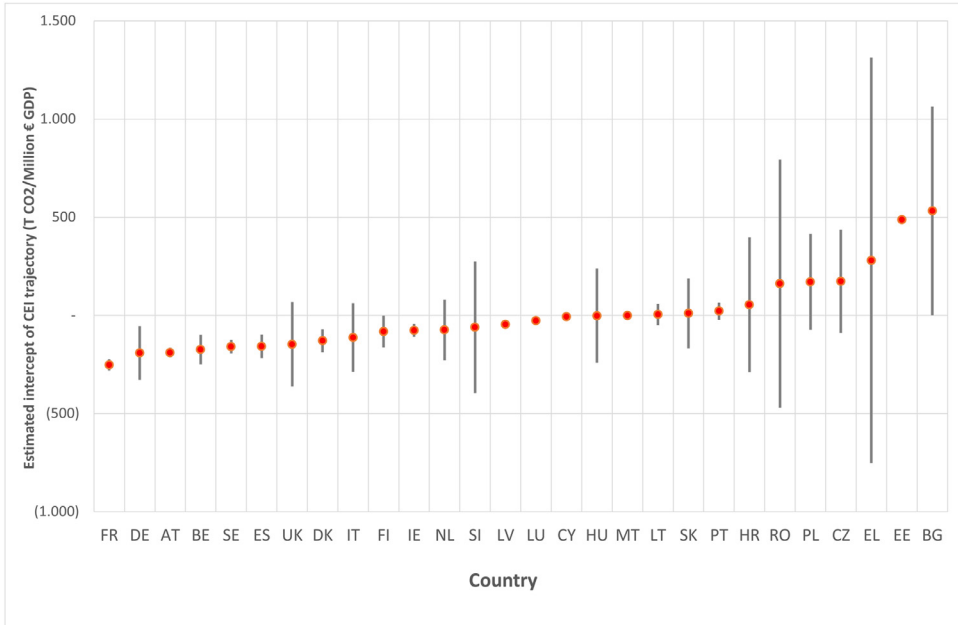


Fig. 3. Expected intercept for each country ($\hat{\alpha}_{0i}$). The red dots represent the mean value of the estimated intercept for each country, while the black bar represents the variability of the estimated intercept within the country (equal to twice the standard deviation of the values at the NUTS 2 level). The countries are ordered on the basis of the average estimated intercept, from the lowest one to the highest one. The countries with the lowest values show a lower estimated initial level of CEI.

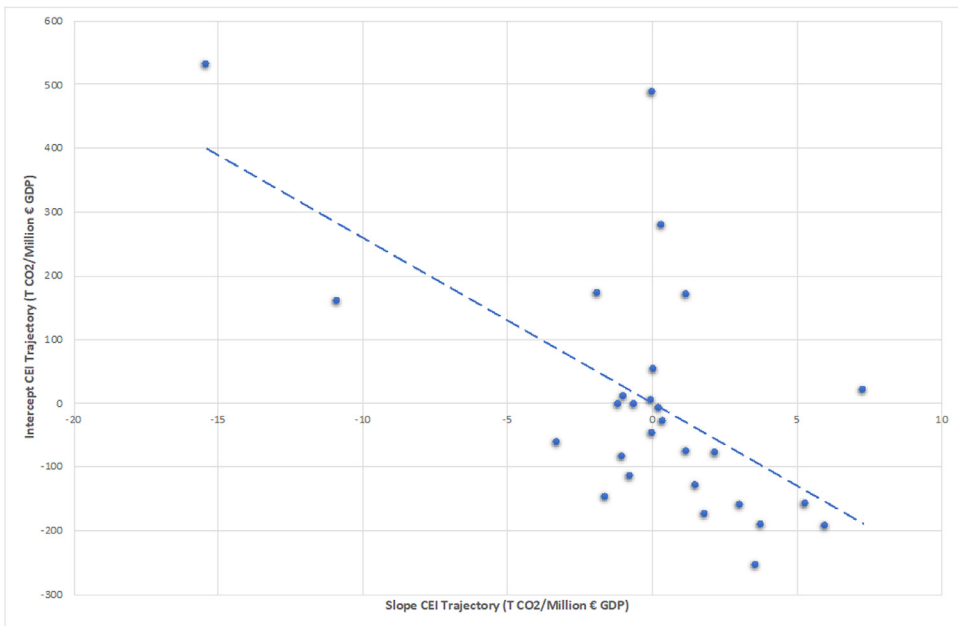


Fig. 4. (From [19]) Scatterplot of the estimated slope (Fig. 2) versus the estimated intercept (Fig. 3) of Carbon Emission Intensity for each country, derived from the multilevel-growth model. The figure clearly shows a negative correlation between the two values, so that the higher the estimated initial level of CEI (intercept) the lower the gradient of the line (slope of the trajectory).

to 530) and other countries showing negative values (for instance Austria, Germany and France with values ranging from -189 to -252).

Model code

All the analyses discussed by Mura et al. [19] have been performed using the *R* software (*R* [20]). In specific, the packages *lme4* [2] and *lmerTest* [16].

The data set has been provided in the person-period form, or univariate format, in which each (NUTS3) unit has multiple records, one for each period in which it was observed.

Interaction plots as illustrated in Fig. 1 are derived by using the package *ggplot2* [25], and using the function *xyplot*, as detailed in the following box

```
xyplot(dat$CO.GDP ~ dat$time | NUTS0, data=dat,
prepanel = function(x, y) prepanel.loess(x, y, family="gaussian"),
xlab = "time", ylab = "CO.GDP",
panel = function(x, y) {
panel.xyplot(x, y)
panel.loess(x,y, family="gaussian",col="red",lwd=2)},
ylim=c(0, 1100),col="black", as.table=T)
```

The models have been fitted using the function *lmer* of the package *lme4*, whereas the significance of the random components of the models have been assessed using the *step* and *ranova* functions of the *lmerTest*, as shown below.

```
mo.123.nolog<-lmer(CO.GDP ~ time + (time | NUTS1/NUTS2/NUTS3),
data=dat,REML=F,control=lmerControl(optimizer = "optimx", calc.derivs = FALSE,optCtrl =
list(method = "nlnmb", starttests = FALSE, kkt = FALSE)))
summary(mo.123.nolog)
step(mo.123.nolog)
ranova(mo.123.nolog)
```

Conclusion

This article provides a detailed description of how to apply unconditional multilevel growth models to evaluate Sustainability Transition pathways, by accounting for both the nested structure induced by the different geographical scales and the time dependence present in the considered longitudinal dataset. Results on the application of the method to a comprehensive dataset of CEI developed over a nine years time span (2008–2016) and including all the European Regions have been briefly reported.

The results show a decreasing trajectory of CEI in the European Union between 2008 and 2016, thus demonstrating the effectiveness of European (Directive 2003/87/EC) and national policies in enabling the transitions towards re-industrialisation and sustainability. However, much is still needed to reach the targets set by the EU Green Deal (-50% CO_{2e} by 2030). The results also highlight very different starting points (intercepts) and trajectories (slopes) for different countries within the EU and for specific subregions (NUTS 1,2 and 3) within each country. It means that the transition towards sustainability cannot be managed through a top-down approach at an aggregated level (European Union and countries), but should also act bottom-up, by considering the contingent features of each geographical area. In more details, focusing on regional and subregional entities allows to fully

exploit the specific competences available and to effectively approach different institutional contexts. Future research may consider to include additional sources of data, e.g. satellite data from Copernicus programme or air pollution data from E-PRTR. As alternative, the model could be extended to account for potential latent drivers that could affect the ST dynamics (e.g. wealth, etc.) through the inclusion of a second-order factor model.

Declaration of Competing Interest

The authors confirm that there are no conflicts of interest.

Acknowledgments

We would like to acknowledge EIT Climate KIC, a body of the European Union, SPS Flagship ReIndustrialise project for providing financial support to this study.

References

- [1] A. Acquaye, T. Ibn-Mohammed, A. Genovese, G.A. Africa, F.A. Yamoah, E. Opon, A quantitative model for environmentally sustainable supply chain performance measurement, *Eur. J. Oper. Res.* 269 (2018) 188–205.
- [2] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4, *J. Stat. Softw.* 67 (1) (2015) 1–48.
- [3] K.A. Bollen, P.J. Curran, *Latent Curve Models: A Structural Equation Perspective*, John Wiley & Sons, Hoboken, NJ, 2006.
- [4] B.M. Byrne, *Multivariate applications series, Structural Equation Modeling With Mplus: Basic concepts, applications, and Programming*, Routledge/Taylor & Francis Group, 2012.
- [5] B.M. Byrne, G. Crombie, Modeling and testing change: an introduction to the latent growth curve model, *Underst. Stat.* 2 (3) (2003) 177–203.
- [6] J. Cai, H. Yin, O. Varis, Impacts of industrial transition on water use intensity and energy-related carbon intensity in China: a spatio-temporal analysis during 2003–2012, *Appl. Energy* 183 (2016) 1112–1122.
- [7] L. Coenen, B. Truffer, Places and spaces of sustainability transitions: geographical contributions to an emerging research and policy field, *Eur. Plan. Stud.* 20 (2012) 367–374.
- [8] P.J. Curran, A.M. Hussong, The use of latent trajectory models in psychopathology research, *J. Abnorm. Psychol.* 112 (4) (2003) 526–544.
- [9] P.J. Curran, J.S. McGinley, D. Serrano, C. Burfeind, A multivariate growth curve model for three-level data, in: H. Cooper (Ed.), *APA Handbook of Research Methods in Psychology*, APA Handbook of Research Methods in Psychology, 3, Data Analysis and Research Publication, 2012 Editor-in-Chief.
- [10] F. Dong, B. Yu, T. Hadachin, Y. Dai, Y. Wang, S. Zhang, R. Long, Drivers of carbon emission intensity change in China, *Resour. Conserv. Recycl.* 129 (2018) 187–201, doi:10.1016/j.resconrec.2017.10.035.
- [11] A.R. Davies, Cleantech clusters: transformational assemblages for a just, green economy or just business as usual? *Glob. Environ. Chang.* 23 (5) (2013) 1285–1295.
- [12] A. Díaz Dapena, F. Rubiera-Morollón, D. Paredes, New approach to economic convergence in the EU: a multilevel analysis from the spatial effects perspective, *Int. Reg. Sci. Rev.* 42 (2018) 335–367.
- [13] T.J. Foxon, A coevolutionary framework for analysing a transition to a sustainable low carbon economy, *Ecol. Econ.* 70 (2011) 2258–2267.
- [14] F.W. Geels, F. Kern, G. Fuchs, N. Hinderer, G. Kungl, J. Mylan, M. Neukirch, S. Wassermann, The enactment of socio-technical transition pathways: a reformulated typology and a comparative multi-level analysis of the German and UK low-carbon electricity transitions (1990–2014), *Res. Policy* 45 (4) (2016) 896–913.
- [15] H. Goldstein, *Multilevel Statistical Models*, John Wiley, Chichester, UK, 2011.
- [16] A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen, lmerTest package: tests in linear mixed effects models, *J. Stat. Softw.* 82 (2017) 1–26.
- [17] A.C. Maassen, Heterogeneity of lock-in and the role of strategic technological interventions in urban infrastructural transformations, *Eur. Plan. Stud.* 20 (3) (2012) 441–460.
- [18] R. Martin, P. Sunley, Path dependence and regional economic evolution, *J. Econ. Geogr.* 6 (4) (2006) 395–437.
- [19] M. Mura, M. Longo, L. Toschi, S. Zanni, F. Visani, S. Bianconcini, The role of geographical scales in sustainability transitions: an empirical investigation of the European industrial context, *Ecol. Econ.* (2021) 106968, doi:10.1016/j.ecolecon.2021.106968.
- [20] R. Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [21] K. Skellern, R. Markey, L. Thornthwaite, Identifying attributes of sustainable transitions for traditional regional manufacturing industry sectors—a conceptual framework, *J. Clean. Prod.* 140 (2017) 1782–1793.
- [22] P. Späth, H. Rohrer, Local demonstrations for global transitions—dynamics across governance levels fostering socio-technical regime change towards sustainability, *Eur. Plan. Stud.* 20 (2012) 461–479.
- [23] G.A. Tasca, V. Illing, A.S. Joyce, J.S. Ogrodniczuk, Three-level multilevel growth models for nested change data: a guide for group treatment researchers, *Psychother. Res.* 19 (4) (2009) 453–461.
- [24] J. Wang, M. Hu, J.F.D. Rodrigues, An empirical spatiotemporal decomposition analysis of carbon intensity in China's industrial sector, *J. Clean. Prod.* 195 (2018) 133–144.
- [25] Hadley Wickham, *Ggplot2: Elegant Graphics for Data Analysis*, 2nd, Springer, 2009.