Chaos, Solitons and Fractals 184 (2024) 115006

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

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

Forecasting the countries' gross domestic product growth: The case of Technological Fitness

Orazio Angelini [a,*], Andrea Gabrielli [b,d], Andrea Tacchella [b], Andrea Zaccaria [c,b], Luciano Pietronero [b,c], T. Di Matteo [a,b,e]

[a] Department of Mathematics, King's College London, The Strand, London, WC2R 2LS, United Kingdom
[b] Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi, Via Panisperna 89 A, 00184 Rome, Italy
[c] Istituto dei Sistemi Complessi, Consiglio Nazionale delle Ricerche, UOS Sapienza, P.le Aldo Moro, 5, 00185 Rome, Italy
[d] Dipartimento di Ingegneria Civile, Informatica e delle Tecnologie Aeronautiche, Università degli Studi "Roma Tre", Via Vito Volterra 62, 00146 Rome, Italy
[e] Complexity Science Hub Vienna, Josefstädter Straße 39, Vienna, 1080, Austria

ARTICLE INFO

Keywords:
Dynamical systems
Macroeconomic forecasting
Economic complexity
Fitness
Patents

ABSTRACT

The cornerstone of Economic Complexity (EC) studies is the assumption that most of the fundamental information about countries' capabilities can be extracted from the products they export. This extreme dimensionality reduction is evident in the typical models used in EC for Gross Domestic Product Per Capita (GDPpc) forecasting, in which only two dimensions – Economic Fitness (EF) and GDPpc – are considered. In this work, we consider adding a third dimension, Technological Fitness (TF), which is computed from the measured patenting activity of countries. We find this improves the GDPpc forecast by disambiguating the different growth patterns of countries. The effect is clearer for those advanced-development countries that already export most of the products present in customs' ontologies, saturating along the EF dimension. Importantly, we show that a higher dimensional model is not necessarily better for all countries and at all times. We illustrate a finding that exemplifies this: while for China adding TF information improves the GDPpc predictions, this is not true for India, a country that according to traditional metrics is very similar. We suggest that future work targeted at introducing new information in EC should exercise care in tailoring the observable quantities employed to each country being examined.

1. Introduction

The Economic Complexity (EC) framework describes economic systems in terms of the diversity of their outputs. In simple terms, the Fitness [1] of an economic actor is the sum of how many different activities it can competitively perform, weighted by how complex such activities are. Usually, in the Economic Complexity framework, the activities are exported products. The definition of Complexity of the activity is self-consistent with that of the fitness of actors: more complex activities are those that are only performed by actors with high Fitness.

It has been shown [1–5] that Fitness is an extremely effective indicator of how good the capability structure of an economic actor is. Despite being quite concise (i.e. it consists of a single number), it provides an accurate estimation of the potential for economic growth, especially when used in combination with more traditional indicators of economic performance such as GDP per capita [6]. When used at the country level to describe industrial capabilities, Fitness delivers very accurate forecasts, that improve on the accuracy of mainstream approaches by up to 25% [7].

The idea, introduced in [6–8], is that the economic dynamics of countries can be modelled as a dynamical system. While in principle the dimensionality of such a dynamical system can be very high, and its trajectory is potentially dependent on several economic indicators (such as e.g. interest rates, employment, median age, inequality, etc.) that can be seen as mutually interacting, it has been observed that projecting the dynamics on the two-dimensional plane given by $\log(GDPpc)$ and $\log(Fitness)$ the trajectories appear smooth and well behaved for a large number of countries [2]. The interpretation for this observation is that the effective dimension of the dynamical system of countries' economies is much lower than one would expect and most of the dynamics, especially for developing countries, is constrained to a low dimensional slice of the space, well described by the GDPpc and Fitness dimensions [7].

More mainstream treatments of the prediction problem in the economic literature normally are in stark contrast to EC's approach by involving a much larger number of factors. In a recent paper, Muller et al. [9] introduce a Bayesian factor model of GDPpc growth with a

* Corresponding author.
E-mail address: orazio.angelini@kcl.ac.uk (O. Angelini).

https://doi.org/10.1016/j.chaos.2024.115006
Received 14 April 2024; Accepted 12 May 2024
Available online 25 May 2024

very long time horizon (up to 100 years), motivated by the need for long-run planning in issues like policy evaluation and climate change. The model leverages more than 800 parameters and builds a hierarchical structure of factors by dividing the countries into correlated nested clusters [10]. The resulting uncertainties are relatively wide (see [11] for an estimation methodology). As noted in [7], combining the two methodologies improves the overall forecasts, pointing out their complementarity.

Economic Complexity's extremely reductionist picture, which replaces the aforementioned standard analysis based on hundreds of economic indicators with a two-dimensional one, has its drawbacks, as some second-order effects are of course lost. However, it presents also advantages when one is aiming to construct a forecasting tool that learns from the past to predict the future. In short, learning from a high-dimensional past is much harder than learning from a low-dimensional one: all forms of statistical learning (including machine learning) rely, in one way or another, on the basic principle of finding analogues, i.e. past examples that resemble the starting point of our prediction and of which we know the future evolution, that can be employed to build our expectation. Finding analogues is much easier in low-dimensional spaces as the curse of dimensionality makes high-dimensional spaces 'empty' of analogues to look for [12]. The case of forecasting countries' growth stands out since the number of analogues is notably small, consisting of just about 200 countries that exist in the world.

Considering only two dimensions as has been traditionally done in the Economic Complexity literature, however, is the most restrictive choice that could be made. It is possible that, by releasing this constraint, one could improve to some extent the predictive accuracy of this approach. In this work, we explore the effect of increasing the dimensionality to 3, by including the Technological Fitness [13–16] dimension, alongside the trade-based Fitness that has been used in the literature so far.

Technological Fitness is an indicator that can be computed for each country by applying the Fitness-Complexity algorithm [1] to patent data. More specifically, the input is a binary matrix whose element is 1 if a country is introducing a significant amount of new *patents* in a given patent category in a given year, and 0 otherwise. This can be compared to the "traditional" Fitness, which in the following we will either call just Fitness or *Economic Fitness*, computed on the country-products matrix, whose element is 1 if a country is exporting a significant amount of a given product in a given year.

While Fitness is very accurate at predicting the growth of developing countries [7], it provides less information regarding countries with high Economic Fitness. These developed countries tend to have a relatively slow dynamics in terms of new industrial products, and there is little distinction between them in terms of Economic Fitness as well as in GDPpc (*Gross Domestic Product per capita*). For this reason, they are very close on the Fitness-GDPpc (FG) plane, and their motion shows great uniformity. Because of this, GDPpc predictions for developed countries made by applying the SPSb algorithm to the FG plane are very similar, and less accurate than the predictions for developing countries [17].

This is where Technological Fitness can help. A developing country's growth in competitiveness can be measured by looking at its progression in the space of exportable products (which are already classified by customs). For developed countries, competitiveness consists in their ability to innovate by introducing new products or improving their efficiency in manufacturing already-exported products. These efforts may be measured by investigating the complexity-weighted diversification in different technological sectors, that is Technological Fitness. We use the number of yearly registered patents in different categories as a proxy for technological capabilities [18–20]. In [21], Ye et al. use the patent network to forecast GDPpc with the method of analogues, but differently from this work they compute centrality measures in the network. We instead derive an index of Technological fitness in a manner analogous to what one does for the amount of exported goods in different categories.

We expect the Technological Fitness indicator to be able to resolve developed countries (i.e. those with high Economic Fitness) better than just Economic Fitness. We also expect to see no improvement in prediction performance for developing (i.e. medium-low Economic Fitness) countries. This will be quantified by comparing the performances of models using different indicators in forecasting GDPpc growth. We model GDPpc growth using Bootstrapped Selective Predictability Scheme (SPSb) models which proved effective in [22].

This work's contribution consists in exploring the results of forecasting forecast countries' GDPpc growth with Technological Fitness and determines under which conditions, and for which countries, this indicator is useful. We also investigate the consequences of relaxing the low-dimensionality constraint typical of the Economic Complexity framework. We examine for the first time the possibility of using three-dimensional SPSb models, and the trade-offs of adding and removing observable quantities from such models. We find strong indication that improving prediction needs additional information, and at the same time new information does not improve predictions uniformly for all countries. This has implications for the direction of future work in the field.

2. Materials and methods

2.1. Fitness computation

Economic Fitness and Complexity measures are calculated from the M_{cp}, a binary matrix, with $M_{cp} = 1$ if country c exports a significant amount of product p, and 0 elsewhere. To meaningfully measure the significance of p exports for country c, first we define the *export matrix* EXM_{cp} as the value in dollars of product p exported by country c on a given year. Following the standard treatment in the literature, we then use the *Revealed Comparative Advantage*, or Balassa index [23] RCA_{cp}, defined as:

$$RCA_{cp} = \frac{\frac{EXM_{cp}}{\sum_j EXM_{cj}}}{\frac{\sum_i EXM_{ip}}{\sum_{kl} EXM_{kl}}}. \tag{1}$$

Notably, this is the ratio between the exports of p done by country c and total exports of c, divided by the same ratio computed over all world. Traditionally, the thresholding of this matrix at a given level (normally 1) returns the M_{cp}:

$$M_{cp} = \begin{cases} 1 & \text{if} \quad RCA_{cp} \geq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

As mentioned later in this section, instead, we instead obtain M_{cp} by applying a Hidden Markov Model to regularise its entries across time.

The algorithm for computing the Fitness and Complexity measures is a map iterated to convergence on the M_{cp}. The equations of the map are:

$$F_c^{(0)} = 1 \quad \forall c, \qquad C_p^{(0)} = 1 \forall p. \tag{3}$$

$$\tilde{F}_c^{(n)} = \sum_p M_{cp} C_p^{(n-1)}, \qquad \tilde{C}_p^{(n)} = \frac{1}{\sum_c M_{cp} \frac{1}{F_c^{(n-1)}}} \tag{4}$$

$$F_c^{(n)} = \frac{\tilde{F}_c^{(n)}}{\langle \tilde{F}_c^{(n)} \rangle_c}, \qquad Cp^{(n)} = \frac{\tilde{C}_p^{(n)}}{\langle \tilde{C}_p^{(n)} \rangle_p}. \tag{5}$$

Fitness of country c is plainly the sum of Complexities of the products it exports. Complexity of a product p is instead constrained by the map to always be less than the lower Fitness found among the exporters of p. Additionally, the more exporters of p there are, the less its Complexity.

The Fitness/Complexity iterative map can prove unstable on noisy datasets. For Technological Fitness, we use a stable version of the map [24], where Eq. (4) is changed this way:

$$F_c^{(n)} = \phi_c + \sum_p M_{cp} C_p^{(n-1)}, \qquad C_p^{(n)} = \pi_p + \frac{1}{\sum_c M_{cp} \frac{1}{F_c^{(n-1)}}} \tag{6}$$

and Eq. (5) is skipped. In our case, we set $\phi_c = 10^{-6} \quad \forall c$ and $\pi_p = 1 \quad \forall p$.

2.2. Datasets

In this work we used three datasets:

GDPpc is taken from the World Development Indicators [25] dataset. Specifically, we use the `NY.GDP.PCAP.PP.KD` indicator, which consists in GDPpc, based on purchasing power parity, re-scaled in constant 2017 international dollars.

We compute Fitness using the COMTRADE dataset (1996–2018) [26]. Following [7] we reconcile bilateral declarations of trade flows from COMTRADE and smooth the M_{cp} time series with Hidden Markov Models. In order to avoid convergence issues in 2017, we removed monopolies from the M_{cp} matrices. We then compute Fitness with the well-known Fitness-Complexity algorithm from [2].

We compute Technological Fitness using the Patstat dataset, which starts in 1900 and ends in 2017. Patstat classifies patents with a 8-digit number using a hierarchical system. The first 3 digits define the lowest level of the hierarchy, which is then refined into sub-classes with one further digit. Each of these sub-classes has further subdivisions, each assigned 4 more digits. We use the highest granularity level of classification (8 digits) to define technologies. The country-technology matrices are defined as number of patents that contain an 8-digit classification code for each country. The allocation of patents to countries is done on the basis of the address of the inventor. We use fractional counting whenever inventors from more than one country are present. We binarise the patent matrix by setting its entries to 1 if they are ≥1, and 0 otherwise. We then compute Technological Fitness with the aforementioned "stable" Fitness and Complexity algorithm proposed by Servedio et al. [24], which improves convergence. Technological fitness is defined for a country only if the country has at least one patent in that given year. Some less developed countries repeatedly go from zero to non-zero patents across the years, which makes them alternate between finite and undefined values of Technological Fitness. This is a manifestation of noise in data, which can potentially be fixed in several different ways. We smooth it out by interpolating the value of Technological fitness in the years where it is undefined, in order to obtain a complete time-series.

2.3. Data cleaning

We remove from the data any analogues with Economic Fitness values less than −6. This is a threshold under which most countries for which the Fitness computation fail to converge are found, and the indicators starts misbehaving (see section on Fitness computation). We compute the base-10 logarithm for Economic Fitness, GDPpc, and Technological Fitness.

Economic Fitness fails to converge in 307 (8.3%), and GDPpc is missing for 18 (0.05%) analogues. To ensure that the comparison of any two models is always done on the same set of analogues, we remove from the data any analogue for which at least one of the observables (Economic Fitness, GDPpc, and Technological Fitness) is not present. If e.g. for a given country-year couple we have Technological Fitness and GDPpc measurements but we are missing the value of Economic Fitness, an analogue is available for the Technological Fitness-GDPpc model but not for the Fitness-GDPpc model: in this case we would discard that analogue so that the comparison between models is fair.

2.4. CAGR and MAE

Throughout the paper we refer to Compound Annualised Growth Rate (CAGR) predictions. CAGR is defined as the annualised growth of a certain quantity (here the country's GDPpc) over a certain period of time. Supposing that a country C starts at a certain level of GDPpc G_i and ends after t years at $G_{i+t} = G_i + \Delta_t(G_i)$, CAGR is defined as:

$$CAGR(C, i, t) = \sqrt[t]{\frac{G_{i+t}}{G_i}} - 1. \tag{7}$$

We always multiply CAGR by 100 to get a percentage annualised growth.

We make predictions for a number of countries, and aggregate them using Mean Average Error. Let us define a CAGR prediction for country C at time t as $\mathrm{pred}(C, t)$, and the actual measured growth as $\mathrm{gt}(C, t)$ we define the error as the difference of the two: $E(C, t) = \mathrm{gt}(C, t) - \mathrm{pred}(C, t)$, and we aggregate the predictions for countries $C_1, C_2, \ldots, C_n$ as:

$$MAE(C_{\{i\}}, t) = \sum_i \frac{\|E(C_i, t)\|}{n} = \sum_i \frac{\|\mathrm{gt}(C_i, t) - \mathrm{pred}(C_i, t)\|}{n}. \tag{8}$$

2.5. GDPpc forecast: the SPSb and vSPSb methods

Bootstrapped Selective Predictability Scheme (SPSb) is a prediction technique used to forecast GDPpc growth for a country by averaging the growth of the most similar countries observed in the past. In [7,8], similarity is translated to distance on the Fitness-GDPpc plane, but the method itself is defined on the euclidean space S defined by arbitrary metrics. Each country has a trajectory in the space S over time. Given the position of country $\hat{c}$ in S plane time $\hat{t}$, $\vec{x}_{\hat{c},\hat{t}}$, we want to forecast $\delta x_{\hat{c},\hat{t}}$, the future displacement of country $\tilde{c}$ from time $\hat{t}$ to $\hat{t} + \Delta t$. We consider the set of observed past observations $(\delta x_{c,t}, \vec{x}_{c,t})$ available in S, called *analogues*. Note that imposing backtesting only allows using the analogues for which $t < \hat{t}$. It is possible to estimate an empirical probability distribution for $\delta x_{\hat{c},\hat{t}}$ with this procedure:

1. Sample (with repetition) the N available analogues, weighted by a probability distribution p given by a Gaussian kernel centred on $x_{\hat{c},\hat{t}}$. More precisely, we define the probability of sampling the analogue displacement $\delta x_{c,t}$ with:

$$p(\delta x_{c,t} | x_{c,t}) = \mathcal{N}(\vec{x}_{\hat{c},\hat{t}} - \vec{x}_{c,t} | 0, \sigma), \tag{9}$$

where we defined $\mathcal{N}$ as the N-dimensional Gaussian distribution:

$$\mathcal{N}(\vec{z} | \vec{\mu}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(\vec{z} - \vec{\mu})^2}{2\sigma^2}\right). \tag{10}$$

Notably, the sampling probability depends only on the Euclidean distance between $\vec{x}_{\hat{c},\hat{t}}$ and the position of the analogue.

2. Sample $B = 1000$ times with the above procedure, each time averaging the sampled displacements. The global distribution of these averages is the empirical probability distribution for $\delta\vec{x}_{\hat{c},\hat{t}}$. The mean of the distribution is used as the prediction value and the standard deviation as the uncertainty on the forecast. This procedure is in effect equivalent to estimating the modes of the empirical distribution with a probability-weighted bootstrap.

In [27] it is shown that this algorithm converges to a Nadaraya–Watson kernel regression (NWKR) for large values of B. Therefore, all our predictions are done with the NWKR version of SPSb,[1] as in [27].

In this context, we define the *bandwidth* for the kernel regression in terms of the standard deviation of the Gaussian kernel as $b = 1/2\sigma^2$. In this work we use a Nadaraya–Watson kernel bandwidth of 0.3 in all the SPSb models along all directions.

The *Velocity-SPSb* method (vSPSb) is a modification proposed in [7] based on the observation that there is autocorrelation in GDPpc growth time series. vSPSb mixes the prediction produced by SPSb with the

[1] The code for the predictions can be found here: https://github.com/ganileni/ectools.

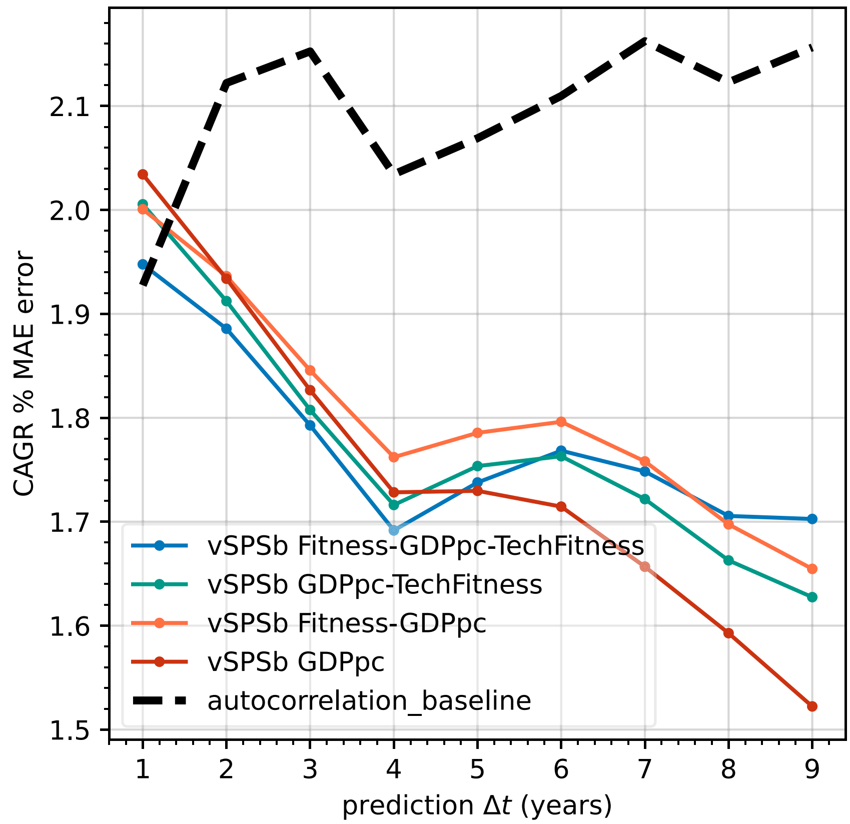


Fig. 1. Error for all backtested predictions at different Δt for all models. All vSPSb models vastly outperform the baseline at $\Delta t > 1$; the 3D model shows the lowest errors in the short-medium term; the GDPpc model is the best one at longest timescales.

change in GDPpc observed for the country C in the previous time interval, $\Delta G_{c,t}$. Suppose we produce with SPSb a prediction pred(C,t) and a standard deviation for the prediction $\sigma_{sps}(C,t)$. If we define the standard deviation of past GDPpc changes $\Delta G_{c,t}$ for the country as $\sigma_{\Delta G}(C,t)$, then the vSPSb prediction is defined as:

$$\frac{\frac{\text{pred}(C,t)}{\sigma_{sps}(C,t)} + \frac{\Delta G_{c,t}}{\sigma_{\Delta G}(C,t)}}{\frac{1}{\sigma_{sps}(C,t)} + \frac{1}{\sigma_{\Delta G}(C,t)}}. \tag{11}$$

2.6. List of models considered

We define a forecast model as SPSb performed on a space defined by certain observables. Given our available observables (Economic Fitness, GDPpc, Technological Fitness), it is possible to construct single-, bi- and tri-dimensional models. The model normally considered in the literature is the bi-dimensional SPSb defined on the (Fitness, GDPpc) plane. This means analogues are couples of (Fitness, GDPpc) measurements for countries.

In this work we consider one single-dimensional model (GDPpc), two bi-dimensional ones (GDPpcFitness) and (GDPpc-Technological Fitness), and one tri-dimensional model (GDPpc-Fitness-Technological Fitness).

We additionally consider one baseline, which we call the *autocorrelation baseline*. This is a naive model that predicts future GDPpc growth by assuming it will be equal to the last observed growth rate over the same time interval. For example, supposing the United states grew at 1.05% over the period 2014–2015, the autocorrelation baseline would predict exactly 1.05% growth for 2015–2016.

2.7. GDPpc forecast: selection of analogues

As the number of analogues available to use for prediction decreases, the model error increases noticeably. Since we want to only use past analogues to forecast future growth, the number of available analogues is small for the first years of our time-series. The model performance is not accurately reflected when using too few analogues, so we only average over the last 5 periods of time available (ending in 2017, 2016, 2015, 2014, 2013). To clarify: suppose we select a model, say the GDP-only SPSb, then select a time interval, say $\Delta t = 2$. Then we consider the prediction error on the 2-year GDPpc growth for all countries over the last 5 time intervals available in the data (in this case 2015–2017, 2014–2016, 2013–2015, 2012–2014, 2011–2013).

3. Results

3.1. Global forecast errors

We first compare at average CAGR % MAE over all countries over different time intervals Δt (1 to 9 years). We obtain CAGR % MAE by comparing with the ground truth actual growth, and we plot it. The result can be seen in Fig. 1. This plot shows several interesting trends.

1. for $2 \leq \Delta t \leq 4$ the 3-dimensional SPSb model (including Fitness, GDPpc and Tech Fitness) has the lowest error of all. For $\Delta t = 1$ the autocorrelation baseline is comparable with most SPSb models.
2. For $\Delta t < 4$, the prediction error steadily decreases by $\sim 0.1\%$ per additional year added to the prediction time interval.
3. After a local minimum at $\Delta t = 4$, the average error increases and starts to decrease again at $\Delta t \geq 7$. The long-term decrease is caused by the fact that long-run predictions are generally easier than mid- and short-term ones, which are more affected by high-frequency noise.
4. For $\Delta t \geq 5$, the model with the lowest average error is the 1-dimensional GDPpc SPSb.

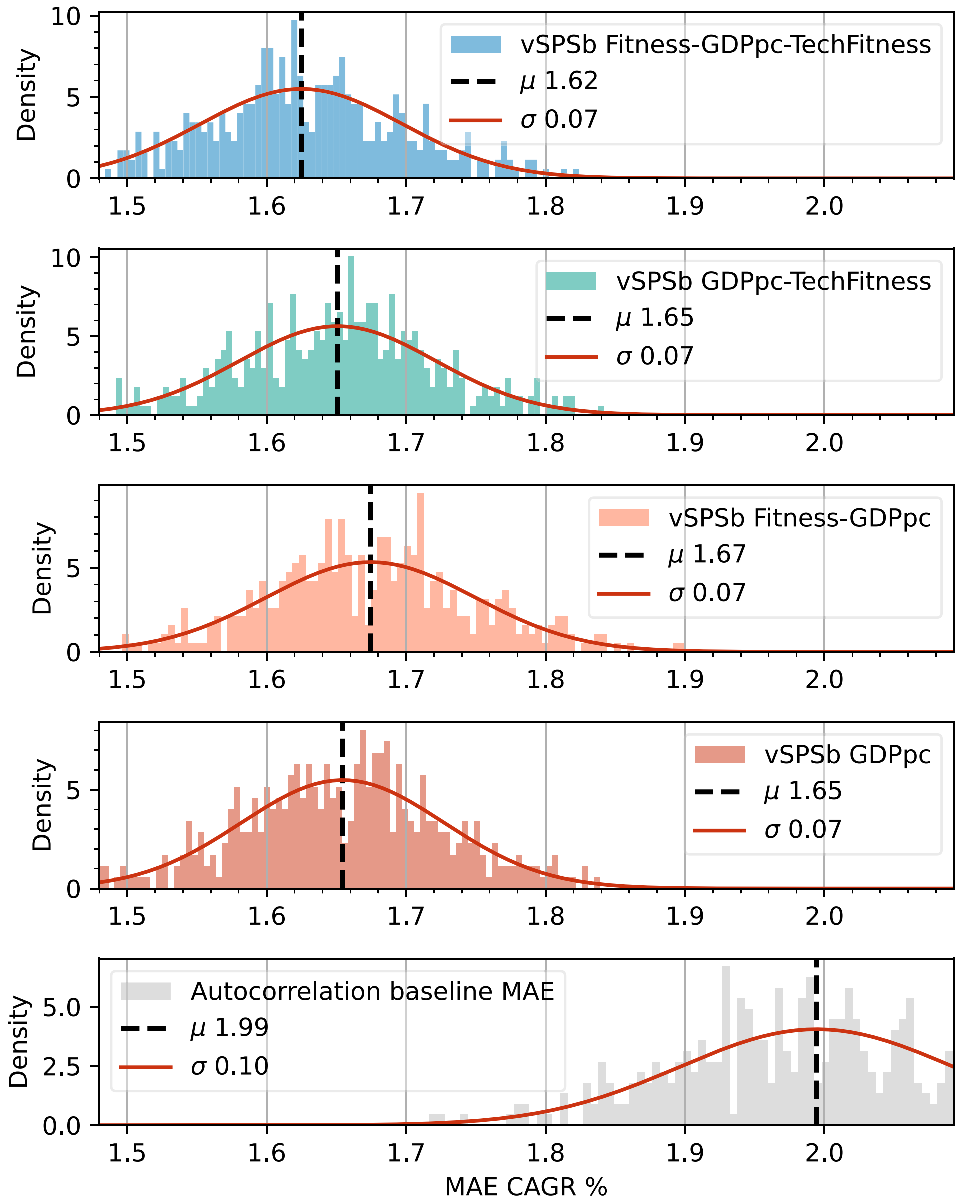


Fig. 2. Bootstrap of CAGR MAE % errors for different models. Errors are re-sampled N times, where $N = 406$ is the number of available predictions. The fluctuations between errors in the models are larger than the differences between the respective averages, except for baseline. This prompts us to examine difference at a more granular level.

The average error keeps steadily decreasing for as long as we have data; we have directly observed the trend up to $\Delta t = 27$.

Apart from these general trends for the average errors, it is difficult to discern large differences in the aggregated predictions from the models. In Fig. 2 we show the CAGR MAE distributions relative to the five prediction models, obtained by bootstrapping the predictions. The typical standard deviations are larger than the difference between the means of any couple of vSPSb models. The autocorrelation benchmark shows, instead, a distribution that is widely different. In order to more deeply understand the differences among the different vSPSb models, we will have to look at disaggregated, more granular, country-level data, which is done in the next section.

The granular results we will present in the following section refer to the $2 \le \Delta t \le 5$ regime, for which both kinds of Fitness are relevant. In this regime, granular results are qualitatively similar, and change smoothly as one changes the value of Δ by a small amount. We will show results at $\Delta t = 4$, the interval for which the differences among models are more pronounced. At longer timescales ($\Delta t \ge 6$), we observe that GDPpc is overwhelmingly more relevant for prediction compared to Fitness. The results we will discuss in the following section do not apply to the $\Delta t \ge 6$ regime, and we will not discuss it in this work.

3.2. Country-level comparison

It is a known feature of the SPS approach that the dynamical properties of the system are heterogeneous across the various areas of the GDPpc-Fitness plane [6]. So it is natural to investigate the heterogeneity of the prediction performances by visualising the prediction errors in this plane, as already done for example in [6,7]. In particular, within each pane of Fig. 3 we plot the difference in CAGR MAE % between the 3-dimensional SPSb model (Fitness, GDPpc, Tech Fitness) and one other SPSb model. Let us for instance consider the classic 2-dimensional SPSb model (Fitness, GDPpc), which whe show in the bottom left pane of Fig. 3. For each country c and year t we compute the difference in error between the two models:

$$D(c,t) = Err[SPSb_{(Fitness,GDPpc\,TechFitness)}(c,t)] - Err[SPSb_{(Fitness,GDPpc)}(c,t)] \tag{12}$$

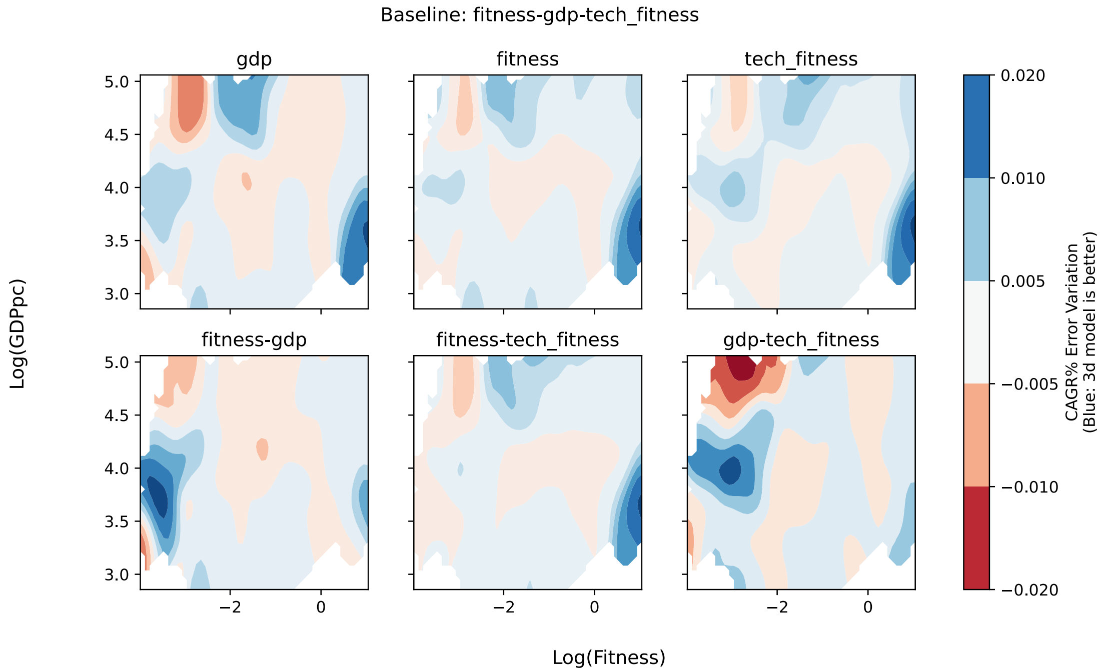


Fig. 3. Error difference between the 3-dimensional SPSb model (Fitness, GDPpc, Tech Fitness) and the other models. A negative value means the 3-dimensional model predicts with lower error (blue areas). Please refer to Section 3.2 for a detailed analysis of the results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $Err[SPSb_{(\mathcal{M})}(c,t)]$ is the difference between the prediction of the model $\mathcal{M}$ and the actual GDPpc growth of the country c at time t. As explained before, we use predictions with $\Delta t = 4$, and close values of Δ provide comparable results. Defined this way, D is negative when the 2-dimensional model makes an error larger than the 3-dimensional model (blue areas in the plot). We then compute a nonparametric regression of $D(c,t)$ on the familiar $\log(GDPpc) - \log(Fitness)$ plane, with a Gaussian kernel of bandwidth 0.2.

In Fig. 3, we also compute the difference in error between the 3D model and all the other possible 1- and 2-dimensional models we can obtain with our 3 observables. Let us start with the comparison between the 3D model and the 2D GDPpc-Fitness model. Some trends are evident:

- The 3d model does better for advanced high-GDPpc and high-fitness economies.
- There is an intermediate area of $\log(Fitness) \approx 0$ and $\log(GDPpc) \lessapprox 4.5$ for which the 2D model predicts with lower error.
- The previous area connects with another band of countries with $\log(Fitness) \approx -2$ and $3.5 \lessapprox \log(GDPpc) \lessapprox 4.5$
- There is one band of countries with middling fitness $\log(Fitness) \lessapprox -1$ and mid-to-low GDPpc $\log(GDPpc) \lessapprox 4$ for which the 3D model has lower error on average.
- $\log(GDPpc) \gtrapprox 4.5$ and $\log(Fitness) \lessapprox -1$ countries, strongly dependent on raw materials export, are all predicted better by the 2D model, with the exception of Bahrain and Oman.
- All other $\log(Fitness) \lessapprox -2$ countries are better predicted by the 3D model.

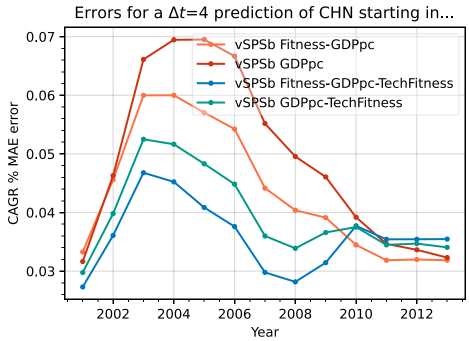
Comparing the 3-dimensional SPSb model (Fitness, GDPpc, Tech Fitness) with the 1-dimensional GDPpc-only model, we see a similar situation (see Fig. 3, top left).

In summary, The 3-dimensional model (and in general adding Technological Fitness to the observables of a model) seems to enable better predictions on high-GDPpchigh-Fitness countries like we expected. Additionally, it seems to always improve predictions on China and India (which are in the bottom-right part of the GDPpc-Fitness plane). There seem to be improvements in the bottom-left parts of the plane: namely low-GDPpc low-Fitness countries.

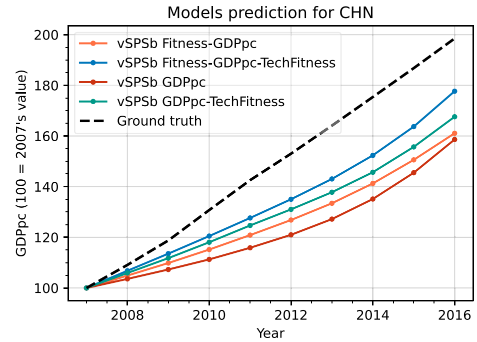
In Fig. 3 we also show the 2-dimensional GDPpc-Technological Fitness model, which shows similarity to the Fitness-GDPpc model. We explain this as the result of the correlation present between Fitness and Technological Fitness so that when each measure is taken individually it encodes similar information to the other. Finally in the figure we have the 1-dimensional Fitness and Technological Fitness models, the 2-dimensional Fitness-Technological Fitness model. We report these results for completeness, but they are not very informative, since any kind of Fitness is only effective at prediction when paired with another metric representing the income of countries.

Let us now consider some examples of predictions, relative to specific countries. A prominent case is China, whose growth period fuelled a long-standing debate. Fig. 4(a) shows the forecast errors at fixed Δt: all models show similar trends (with higher errors around 2004 and progressively lower errors towards 2012), and the 3-dimensional model overperforms the other models most of the time. From Fig. 4(b) it emerges that all models predict a lower growth than the one that actually occurred. However, the 3-D model predicts the highest growth.

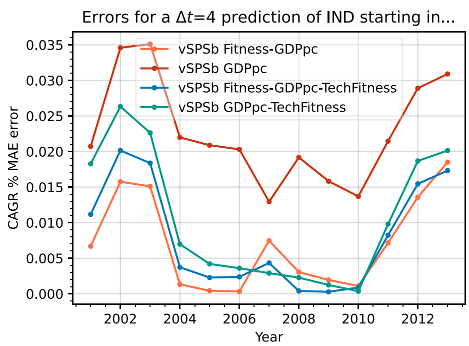
For India, instead, in Figs. 4(c) and 4(d) we observe that all models perform similarly with the exception of the 1-dimensional GDPpc-only one, which predicts once again a lower growth with respect to the growth actually occurred. Given the closeness of India and China in the Fitness-GDPpc plane, the 2-dimensional Fitness-GDPpc model tends to produce similar forecasts for the two countries. However, note that the 2-dimensional GDPpc-Tech Fitness model alone delivers worse forecasts than the standard 2-dimensional Fitness-GDPpc model. The fact that the 3-dimensional model outperforms all others in predicting China's growth lets us conclude that the addition of Technological Fitness is essential to disambiguate China from India's forecasts. In practice, this is due to the fact that in the 3D plane, the high Technological Fitness of China, and the comparably lower one of India, have the effect of pulling the former closer to previous analogues of high-growth China, and the latter away from them. This assigns better analogues to the prediction of each. Adding Technological fitness better reflects the different situation of the two countries: China has almost saturated the



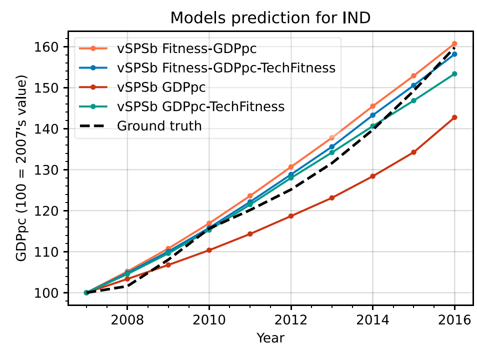
(a) Forecast errors for a $\Delta t = 4$ prediction on China's GDPpc growth as a function of the start year. The 3-dimensional model outperforms all others by a significant amount.



(b) Comparison: full series of predictions for China's GDPpc growth of each model. The prediction starts in 2007 and we show the forecast error at various Δt values.



(c) Forecast errors for a $\Delta t = 4$ prediction on India's GDPpc growth as a function of the start year. Differently from China, the 3-dimensional model here doesn't have significantly different performance than the others.



(d) Comparison: full series of predictions for India's GDPpc growth of each model. The prediction starts in 2007 and we show the forecast errors at various Δt values.

Fig. 4. Forecast errors as a function of time for China (top) and India (bottom).

space of industrial production, and is closer than India to the point where frontier technological innovation matters more than expanding the industrial portfolio (incidentally this would foretell a slow-down of Chinese growth in the medium term). While the two countries are near each other in the GDPpc-Fitness plane, they should be further away in the space defining the "true" dynamical system generating the evolution of countries, and the 3-D model is better than 2-D ones because it does increase their distance.

3.3. Model bias

In this section we discuss a systematic deviation of the models' predictions with respect to the realised GDPpc growth. This can be visually represented by plotting the forecast errors for all country-year couples as a function of Technological Fitness. This is shown in Fig. 5, where each blue dot represents a forecast error of the 3D model, while the green ones refer to the 2D Fitness-GDPpc model. The dashed lines represent the respective averages, while the continuous lines represent a nonparametric regression.

First of all, we find that both models tend to overestimate growth. However, while for low (<1.5) and high (>2.5) $\log(TechFitness)$ values the 2D model overestimates growth more than the 3D one, this tendency is inverted for medium $\log(TechFitness)$ values.

4. Conclusions

Within the Economic Complexity framework, GDPpc forecasting is based on the construction and analysis of a two-dimensional GDPpc-Fitness plane. In this severe dimensionality reduction, one assumes that the dynamics of heterogeneous and complex systems like countries' economies can be effectively described by only two variables. While this hypothesis is verified by the forecasting performance of this scheme, it is natural to ask whether adding more dimensions might be useful to better describe the complex dynamics of economic development. In this work, we attempted just that, using Technological Fitness as a third dimension for the Economic Complexity's approach to forecasting countries' GDPpc growth. We found that, as we hypothesised, it does improve forecasts for developed economies which have reached the frontier of industrial growth and must rely on technological innovation to keep growing. Additionally, there are signals that it does improve predictions for countries that have very low technological fitness — the ones for which new patent activity represents an important economic signal. For countries in the middle of the spectrum, which typically rely on the expansion of industrial capabilities for their development, it seems to make predictions slightly worse. On one hand, these results confirm the correctness of the dimensionality reduction described above, since the resulting projections are in general quite

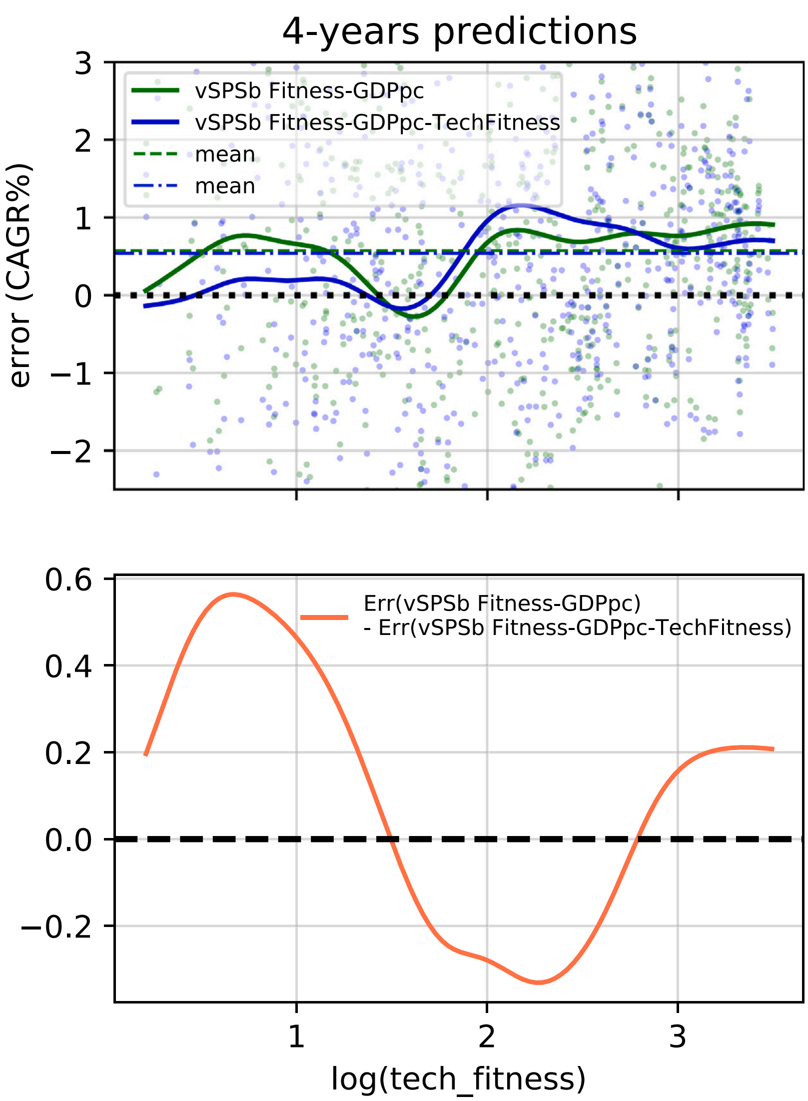


Fig. 5. 3D vs. 2D (Fitness, GDPpc) models forecast errors as a function of Tech Fitness ($\Delta t = 4$). The vertical axis cuts off large error values to make the patterns in the regression more visible. We also show the mean for both models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stable; on the other hand, adding the patent data improves the understanding of the growth mechanism of both developed and developing countries.

The point is clearly shown in Fig. 3. Interestingly, we described in that figure the intermediate area of $\log(\text{Fitness}) \approx 0$ and $\log(\text{GDPpc}) \lessapprox 4.5$ for which the 2D model predicts with lower error than the 3D model. This area is reminiscent of the threshold described in [3]. The finding could be interpreted in the spirit of that work: in the development of history a country, once it reaches a certain fitness threshold that propels it out of the poverty trap, the relative technological development of a country becomes less relevant, only adding noise to the predictions, until the country reaches a higher level of fitness where technology capabilities matter again. The area included $\log(GDPpc) \gtrapprox 4.5$ and $\log(Fitness) \lessapprox -1$, where the 2D model again outperforms 3D, also has a clear interpretation. The GDPpc of these countries depends on raw materials export, therefore knowing their technological level is irrelevant for the purpose of forecasting. An example of a country for which Fitness is relevant, and Tech fitness improves the prediction is Bahrain. This is in contrast with Saudi Arabia, for which the addition of Technological fitness actually worsens predictions. And indeed Bahrain's economy is arguably more dynamical and less dependent on raw materials exports than Saudi Arabia [26,28]. As a rule of thumb, we could propose technological fitness improves results overall, but mostly if in a 3D model, and is more useful for countries that export few products and work more with services. It is instead irrelevant or deleterious for countries that are on a clear path to development due to the expansion of industrial capabilities.

Another interesting result is the fact that China's forecasts show great improvement from the addition of Technological Fitness to the model. As shown in Fig. 4(b) the 3D model outperforms all others during the period in which China was growing at outlier-level rates. This stays true until the 2008's global financial crisis, after which the performance of all the models examined becomes comparable, probably due to a change in the growth dynamics of the country (it is possible to observe the error increasing after 2008 for almost all countries). This is in stark contrast for what happens with India. In Fig. 4(c) we can see that the single-dimensional GDPpc-only model performs significantly worse, whereas the other ones have comparable performance. This could be due to the fact that, while Technological Fitness does not add information relevant to the specific country, it is nonetheless strongly correlated with Fitness. The common information they share seems to be relevant to the forecasts for India, which is why all the models with at least one Fitness observable perform much better than GDPpc only.

As a consequence of highlighting these differences, we note that one of the main original results of this work is that while a higher-dimensional model can improve predictions, it is not necessarily better than a lower-dimensional model all around. We should not expect adding a new observable to always improve predictions for every country in every context. Instead, specific combinations of observables might prove better for specific classes of countries and at specific time intervals. This is not due to the reason, often mentioned in the literature, that increasing the number of observables adds to the total noise (which would lead, in the limit of a large number of dimensions, to completely unreliable predictions — which is not observed in practice). The reason is, instead, that some kinds of information are relevant to a country's growth only conditionally to specific conditions of that country. This is an important point to establish for future research programs into potential new observables and metrics that might conceivably aid with forecasting. Assuming this insight is correct, the most promising approach for improving the Economic Complexity framework lies in tailoring the information used to make a prediction to the particular area of the phase space of the countries' evolution dynamical system — or in other words to the kind of country one is examining.

CRediT authorship contribution statement

Orazio Angelini: Conceptualization, Data curation, Formal analysis, Supervision, Validation, Writing – review & editing. **Andrea Gabrielli:** Conceptualization, Writing – review & editing. **Andrea Tacchella:** Conceptualization, Data curation, Writing – review & editing. **Andrea Zaccaria:** Conceptualization, Data curation, Writing – review & editing. **Luciano Pietronero:** Conceptualization, Writing – review & editing. **T. Di Matteo:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

AZ acknowledges the PRIN project No. 20223W2JKJ "WECARE", CUP B53D23003880006, financed by the Italian Ministry of University and Research (MUR), Piano Nazionale Di Ripresa e Resilienza (PNRR), Missione 4 "Istruzione e Ricerca" - Componente C2 Investimento 1.1, funded by the European Union - NextGenerationEU.

AG acknowledges the PRIN 2022 PNRR project "C2T - From Crises to Theory: towards a science of resilience and recovery for economic and financial systems" funded by the Italian Ministry of University and Research (MUR) CUP: F53D23010380001

Italian PRIN 2022 PNRR P2022B5S5J: "TRIPLE T — Tackling a just Twin Transition: a complexity approach to the geography of capabilities, labour markets and inequalities".

TD acknowledges the 2023–2025 Project "Complexity in Natural, Social and Economic Sciences", Fund assigned to the Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi by the Italian Ministry of University and Research.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.chaos.2024.115006.

References

[1] Tacchella Andrea, Cristelli Matthieu, Caldarelli Guido, Gabrielli Andrea, Pietronero Luciano. A new metrics for countries' fitness and products' complexity. Sci Rep 2012;2.
[2] Cristelli Matthieu, Gabrielli Andrea, Tacchella Andrea, Caldarelli Guido, Pietronero Luciano. Measuring the intangibles: A metrics for the economic complexity of countries and products. PLoS One 2013;8(8).
[3] Pugliese Emanuele, Chiarotti Guido L, Zaccaria Andrea, Pietronero Luciano. Complex economies have a lateral escape from the poverty trap. PLoS One 2017;12(1):e0168540.
[4] Sbardella Angelica, Pugliese Emanuele, Zaccaria Andrea, Scaramozzino Pasquale. The role of complex analysis in modelling economic growth. Entropy 2018;20(11):883.
[5] Tacchella Andrea, Zaccaria Andrea, Miccheli Marco, Pietronero Luciano. Relatedness in the era of machine learning. Chaos Solitons Fractals 2023;176:114071.
[6] Cristelli Matthieu, Tacchella Andrea, Pietronero Luciano. The heterogeneous dynamics of economic complexity. PLoS One 2015;10(2):e0117174.
[7] Tacchella Andrea, Mazzilli Dario, Pietronero Luciano. A dynamical systems approach to gross domestic product forecasting. Nat Phys 2018;14(8):861.
[8] Cristelli Matthieu, Tacchella Andrea, Cader Masud, Roster Kirstin, Pietronero Luciano. On the predictability of growth. The World Bank; 2017.
[9] Müller Ulrich K, Stock James H, Watson Mark W. An econometric model of international growth dynamics for long-horizon forecasting. Rev Econ Stat 2022;104(5):857–76.
[10] Müller Ulrich K, Watson Mark W. Long-run covariability. Econometrica 2018;86(3):775–804.
[11] Müller Ulrich K, Watson Mark W. Measuring uncertainty about long-run predictions. Rev Econ Stud 2016;83(4):1711–40.
[12] Hosni Hykel, Vulpiani Angelo. Forecasting in light of big data. Philos Technol 2018;31(4):557–69.
[13] Pugliese Emanuele, Zaccaria Andrea, Pietronero Luciano. On the convergence of the fitness-complexity algorithm. Eur Phys J Spec Top 2016;225(10):1893–911.
[14] Sbardella Angelica, Perruchas François, Napolitano Lorenzo, Barbieri Nicolò, Consoli Davide. Green technology fitness. Entropy 2018;20(10):776.
[15] Barbieri Nicolò, Consoli Davide, Napolitano Lorenzo, Perruchas François, Pugliese Emanuele, Sbardella Angelica. Regional technological capabilities and green opportunities in Europe. J Technol Transfer 2023;48(2):749–78.
[16] Straccamore Matteo, Bruno Matteo, Monechi Bernardo, Loreto Vittorio. Urban economic fitness and complexity from patent data. Sci Rep 2023;13(1):3655.
[17] Castañeda Gonzalo, Pietronero Luciano, Romero-Padilla Juan, Zaccaria Andrea. The complex dynamic of growth: Fitness and the different patterns of economic activity in the medium and long terms. Struct Change Econ Dyn 2022;62:231–46.
[18] Balland Pierre-Alexandre, Rigby David. The geography of complex knowledge. Econ Geogr 2017;93(1):1–23.
[19] Pugliese Emanuele, Cimini Giulio, Patelli Aurelio, Zaccaria Andrea, Pietronero Luciano, Gabrielli Andrea. Unfolding the innovation system for the development of countries: coevolution of science, technology and production. Sci Rep 2019;9(1).
[20] de Cunzo Francesco, Petri Alberto, Zaccaria Andrea, Sbardella Angelica. The trickle down from environmental innovation to productive complexity. Sci Rep 2022;12(1):22141.
[21] Ye Yucheng, Xu Shuqi, Mariani Manuel Sebastian, Lü Linyuan. Forecasting countries' gross domestic product from patent data. Chaos Solitons Fractals 2022;160:112234.
[22] Tacchella Andrea, Mazzilli Dario, Pietronero Luciano. A dynamical systems approach to gross domestic product forecasting. Nat Phys 2018;1.
[23] Balassa Bela. Trade liberalisation and "revealed" comparative advantage. Manch Sch 1965;33(2):99–123.
[24] Servedio Vito, Buttà Paolo, Mazzilli Dario, Tacchella Andrea, Pietronero Luciano. A new and stable estimation method of country economic fitness and product complexity. Entropy 2018;20(10):783.
[25] World Bank International Economics Dept Development Data Group. World development indicators. World Bank; 2018.
[26] United Nations Statistics Division, UN COMTRADE. International Merchandise Trade Statistics. https://comtrade.un.org/.
[27] Angelini Orazio, Di Matteo Tiziana, Angelini Orazio, Di Matteo Tiziana. Complexity of products: The effect of data regularisation. Entropy 2018;20(11):814.
[28] Patelli Aurelio, Pietronero Luciano, Zaccaria Andrea. Integrated database for economic complexity. Sci Data 2022;9(1):628.