Liu, Nairui & Morawska, Lidia
(2020)
The relationship between social development and ambient particulate matter pollution: Can we predict the turning points?
Environmental Management, *66*(2), pp. 278-287.

https://doi.org/10.1007/s00267-020-01299-0

The relationship between social development and ambient particulate matter pollution: can we predict the turning points?

Nairui Liu[a] and Lidia Morawska[a*]

[a]International Laboratory for Air Quality and Health, Queensland University of Technology, Brisbane, QLD 4000, Australia

*Corresponding author
International Laboratory for Air Quality and Health, Queensland University of Technology
2 George St, Brisbane, QLD 4000, Australia
Email: l.morawska@qut.edu.au

Abstract:

Much research has been conducted to find evidence of the environmental Kuznets curve (EKC) in the relationship between air pollutant concentration and economic development. A major focus of EKC-related research has so far been to identify the turning point illustrated by EKC theory or to predict the moment when the turning point will occur. In our research, via analyzing the relationship between an aggregate social-development-representative variable (the Socio-demographic Index or SDI) and the population-weighted concentration of $PM_{2.5}$, we propose that the overall relationship takes the form of a tilted-S shape with two types of turning points rather than one. Additionally, our research shows that the EKC is highly country-specific, making it extremely difficult to predict the positions of both turning points based on the historical development or trajectories of other countries. Therefore, we conclude that EKC theory is not a useful tool to predict the location of the turning points. However, for short-term prediction of the relationship, we advocate the use of support vector regression, which can forecast the evolution, unless rapid changes are occurring. We suggest that policy makers should not anchor their hopes on predicting turning points from previous studies, but should put more effort into dealing with present particulate matter pollution.

Keywords:

Environmental Kuznets curve; social development; ambient particulate matter pollution.

Abbreviations:

EKC: environmental Kuznets curve; SDI: Socio-demographic Index;

PWC: population-weighted concentration; SVR: support vector regression;

GBD: Global Burden of Diseases, Injuries, and Risk Factors Study

1 Introduction

1.1 Introduction of research

Previous studies (Dong et al. 2018; Ji et al. 2018; Luo et al. 2018) have regarded particulate matter pollution as a type of environmental degradation, and therefore tested the environmental Kuznets curve theory with various panels of data. However, because these studies usually used income per capita and atmospheric concentrations as indices for economic development and environmental degradation respectively, they did not take into account that factors other than income can significantly influence life quality (such as the level of education received), and that a high concentration of pollutants will exert more harm to human health in regions with a denser population. Therefore, the aim of this research was to study the characteristics of the relationship between social development and ambient particulate matter pollution with indices more holistically representative of the actual situation (which will be described in section 1.3).

1.2 Introduction of the EKC

The environmental Kuznets curve (EKC) theory was first proposed in the middle of the last century (Dong et al. 2018), although heated discussion about it did not start until the publication of Grossman and Krueger's work (Grossman and Krueger 1991). The main logic behind EKC theory is that the economy develops at the cost of the environment in the early stages, but when a country becomes rich enough, it will put more effort into controlling environmental pollution, leading to a negative correlation between economic growth and environmental degradation in the later stages of development (Kaika and Zervas 2013). It is worth mentioning that EKC does not imply a causal relationship between economic development and environmental degradation, i.e. environmental degradation is not improved nor exacerbated as a result of economic development and vice versa. The EKC is widely used to depict the relationship between economic development and environmental degradation (Kaika and Zervas 2013). For example, one previous study (Wu et al. 2018) used 259 prefecture-level cities in China to investigate the relationship between $PM_{2.5}$ and urbanization. It concluded that economic urbanization and coal consumption were the two most important factors influencing $PM_{2.5}$ concentration (Wu et al. 2018). In particular, the correlation between economic urbanization and $PM_{2.5}$ concentration presented the shape of an inversed-U or inversed-N (Wu et al. 2018). Other researchers claim that the EKC effect is less important as it does not reduce pollution directly. A previous study (Stern and Zha 2016) investigated the particulate matter pollution data in 50 large Chinese cities from 2013 to 2014. Evidence was found for negative time effects and convergence effects in particulate matter concentration, while none of the obtained EKC curve's coefficients were statistically significant (Stern and Zha 2016). Another study (Stern 2017) also stressed the importance of time effects and convergence effects when considering air pollution emissions.

A schematic diagram of the EKC is shown in Figure 1. M in this figure denotes the turning point of the EKC. Before M, environmental pollution worsens as the economy grows (denoted as the L stage). After M, the environment improves as the economy grows (denoted as the R stage). Although EKC theory does not indicate that symmetry will exist in the curve, empirical studies usually use 2-degree polynomial regression to judge whether an EKC exists (Dinda 2004).

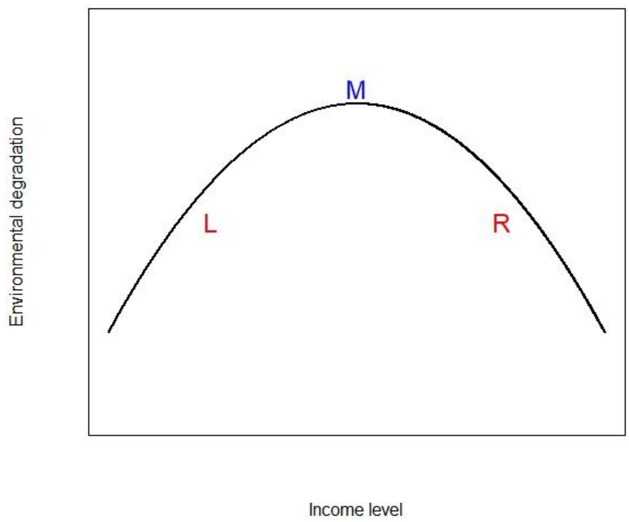


Figure 1 A schematic diagram of the environmental Kuznets curve (EKC), which contains two developing stages (L, R) and one turning point (M).

1.3 Introduction of the Socio-demographic Index and population-weighted concentration

The index used in this research to evaluate the social development of a region is the Socio-demographic Index (SDI). It is an aggregate indicator defined, calculated and deemed as most representative of a region's social development by the GBD (Global Burden of Diseases, Injuries, and Risk Factors Study group (Forouzanfar et al. 2016; Roth et al. 2018). The SDI is "the geometric mean of 0 to 1 indices of total fertility rate under the age of 25, mean education for those aged 15 and older, and lag distributed income per capita" (Roth et al. 2018). The value of the SDI ranges from 0 to 1, where "an index score of 0 represents the minimum level of each covariate input past which selected health outcomes can get no worse, while an index score of 1 represents the maximum level of each covariate input past which selected health outcomes cease to improve" (the GBD Study group uses "life expectancy at birth and under-5 mortality" as selected health outcomes) (Roth et al. 2018).

The index used in this research to evaluate the severity of particulate matter pollution in a country is the population-weighted concentration (PWC) of $PM_{2.5}$ (measured in micrograms per cubic meter). High levels of $PM_{2.5}$ are well known for their negative effects on human health, especially their strong links with cardiovascular and respiratory diseases (Pui et al. 2014). Society usually shows much concern to $PM_{2.5}$ pollution taking place in a densely populated region, such as metropolises. This makes sense because it does cost the society more if $PM_{2.5}$ pollution happens in a densely populated area instead of a sparsely populated area. Therefore, we used PWC rather than mean concentration to better describe the level at which people in a certain country are suffering from $PM_{2.5}$. Through using the approach of sampling more in a more densely populated area, PWC can reflect the influence of an unbalanced population distribution (Stanaway et al. 2018).

Compared with the indices used in previous studies (usually income per capita and pollution concentration), the SDI has the advantage of describing social development more comprehensively, and PWC has the advantage of considering the influence of population density. Unlike many previous studies related to the EKC, we do not include multiple variables, such as energy consumption or the structure of the economy. This is because our research is aimed at directly depicting the relationship between social development and ambient particulate matter pollution.

This paper is organized as follows. Section 2 describes the data and methods used in this research. Section 3 presents and discusses the results. Section 4 summarizes the conclusions and limitations of this research. Section 5 provides its policy implications.

2 Methods

2.1 Data sources

The data used in this research came from the database of the Global Burden of Diseases, Injuries, and Risk Factors (GBD) Study group. The GBD Study group is sponsored by the Bill

& Melinda Gates Foundation and its main aim is to “improve the health of the world’s populations by providing the best information on population health” (http://www.healthdata.org/gbd). The research of the GBD group represents the highest quality in its field and has led to many publications in high-impact factor journals (Forouzanfar et al. 2016; Gakidou et al. 2017; Roth et al. 2018; Stanaway et al. 2018).

The SDI data for different countries and regions used in this research were extracted from the appendix of Roth et al.’s paper (Roth et al. 2018), while the PWC data were downloaded from the Global Health Data Exchange’s website (http://ghdx.healthdata.org/record/global-burden-disease-study-2017-gbd-2017-covariates-1980-2017). In this research we used the SDI and PWC information of 195 countries in the period from 1990 to 2017. To summarize the typical characteristics of the SDI–PWC relationship, this research disregarded countries whose SDI did not keep increasing from 1990 to 2017, or in other words, countries that had experienced social development regression. This is because such countries usually went through tremendous political turmoil or abrupt policy change within this time period. One of the examples is Syria whose SDI decreased after the outbreak of war. Not hard to imagine, such factors would certainly complicate the analysis. Therefore, we deem them beyond our current research scope. A total of 155 countries were left for further analysis after this filtering procedure.

2.2 Empirical models

Polynomial regression was used to analyze the characteristics of the relationship between SDI and PWC. To align with previous studies on the EKC, polynomial regression in this research was restricted to 2-degrees. Support vector regression (SVR) was chosen to predict the PWC for a given country for its flexibility.

The most distinguishable feature of SVR is that it applies no penalty for residuals less than a predetermined tolerance (ε), so its solution relies only on a subset of training data (the so-called support vectors) (Jain et al. 2014). This method non-linearly maps the input space (x) into a high-dimensional feature space ($\varphi(x)$) (Yeganeh et al. 2012). Thus, the method’s prediction f(x) can be expressed with equation (1):

$$f(x) = \omega * \varphi(x) + b \qquad (1)$$

where ω is the weight vector and b is the bias (Dong et al. 2005; Yeganeh et al. 2012). The purpose of this method is to solve for ω and b through minimizing the term (2):

$$\frac{1}{2}||\omega||^2 + C\frac{1}{l}\sum_{i=1}^{l} L_\varepsilon(y_i, f(x_i)) \quad (2)$$

where $||\omega||^2$ is the regularized term, C is the regularization term determined by the user, l is the number of observations (y_i) outside the ε-tolerance range for the corresponding f(x), and $L_\varepsilon(y_i, f(x_i))$ is defined via equation (3):

$$L_\varepsilon\big(y_i, f(x_i)\big) = |y_i - f(x_i)| - \varepsilon \quad . \quad (3)$$

(Dong et al. 2005) Note that ε is selected by the user and $L_\varepsilon\big(y_i, f(x_i)\big)$ is set to 0 (no penalty) if $|y_i - f(x_i)|$ is less than ε (Dong et al. 2005). This method usually involves the use of a kernel function ($K(x_i,x_j) = \varphi(x_i) \cdot \varphi(x_j)$) based on the user's choice, such as the Gaussian radial basis function (Dong et al. 2005). This is because such functions can limit the computation within the input space rather than spread to the feature space (Dong et al. 2005).

SVR is a flexible method. It can be trained for either prediction or estimation purposes, depending on the type of training it receives (Dong et al. 2005). Moreover, it applies structural risk minimization (SRM) principles, which means this method balances between "the quality of the approximation of the given data and the complexity of the approximating function", as one can see from term (2) (Vapnik 2013). This feature helps SVR reduce the risk of overfitting to training data (Dong et al. 2005).

In this research, SVR was built and run in *RStudio* with the *e1071* package (Leisch 2018; R Core Team 2017). 80% of the total data was used for training while the remaining 20% was used for validation. To predict the PWC for a given country in the next year, the model takes the SDI and PWC in the current year and the last year as input. The Gaussian radial basis function was chosen to be the kernel function for SVR in this research, and the corresponding (C, ε) was decided through a stepwise search as described in Dong et al.'s work (Dong et al. 2005). In brief, this involves searching for the optimal value of one parameter while the other is fixed, and repeating this process until two consecutive searches result in a negligible difference in MSE (mean squared error). To the authors' knowledge, this is the first time that SVR has been applied in EKC-related research.

The criteria used in this research to judge the goodness of fit in regressions were the coefficient of determination (R^2), adjusted R^2, the results of the Kolmogorov-Smirnov test, and root mean squared error (RMSE).

3. Results and discussion

3.1 The relationship between the SDI and PWC

3.1.1 Overall relationship

Twenty-six countries that have maintained an annual SDI growth rate of more than 1% from 1990 to 2017 constituted the group of fast development countries for our initial analysis. The upper three plots in Figure 2 demonstrate the SDI–PWC curves of three countries whose shape is typical (ignoring local variations): a "V" shape, a monotonic increasing curve, and an "inverted-V" shape. The lower three plots in Figure 2 present the curves of three developed countries in the western world with similarly shaped monotonic decreasing curves. (Figure S1 shows the curves of all 26 countries.) Notice that we purposely did not keep the scales of PWC in every country the same. This is for the reason that the baseline value of PWC in each country shall be different because it is dependent on topographical and meteorological conditions. Since it is extremely hard to determine the baseline value for each country, we adopted the approach of identifying features of the relationship based on "shapes".

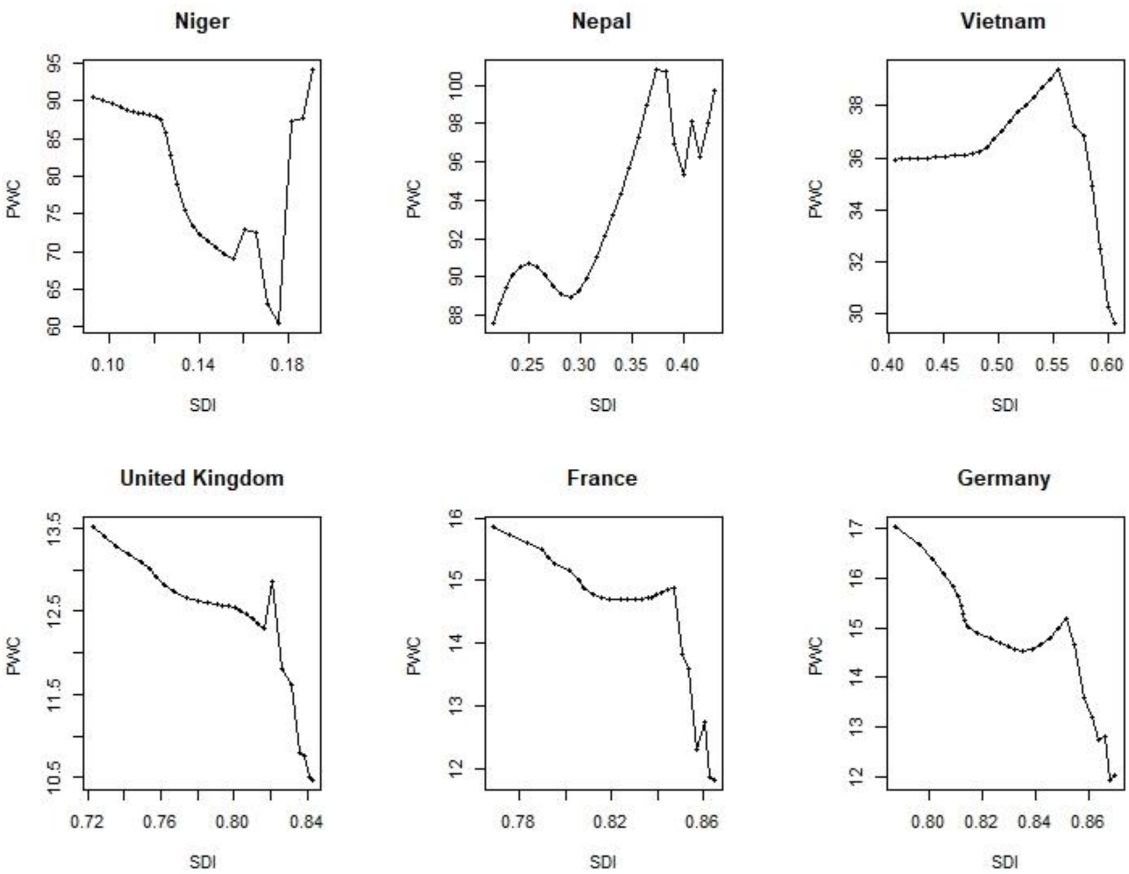


Figure 2 (Upper) SDI–PWC curves of three typical nations that have maintained an SDI growth rate of more than 1% from 1990 to 2017. (Lower) SDI–PWC curves of three western developed countries. The PWC values are expressed in micrograms per cubic meter. (SDI: Socio-demographic Index; PWC: population-weighted concentration)

If the relationship between social development and ambient particulate matter pollution is similar among different countries, or in other words, if developing countries adopt a similar evolution path as today's developed countries, then it is feasible to generalize a universal pattern (such as the shape of the EKC) through analyzing development modes in various countries. The data used in this research have a time span of 28 years. Because different countries have experienced different developmental stages within this time period, we are unable to investigate the overall relationship between SDI and PWC for even one country. In fact, it takes over 100 years for a country to elevate its SDI from 0.2 to 0.8 at an annual growth rate of 1%, a rate that is difficult to maintain in the later stages of development. However, through "connecting" the SDI–PWC relationships in developing and developed countries within the same period while ignoring local variations (please refer to the different curve shapes in Figure 2), we propose an overall relationship as shown in Figure 3. The lines in Figure 3 are indicators of schematic monotonicity, and there is no symmetry between adjacent lines.

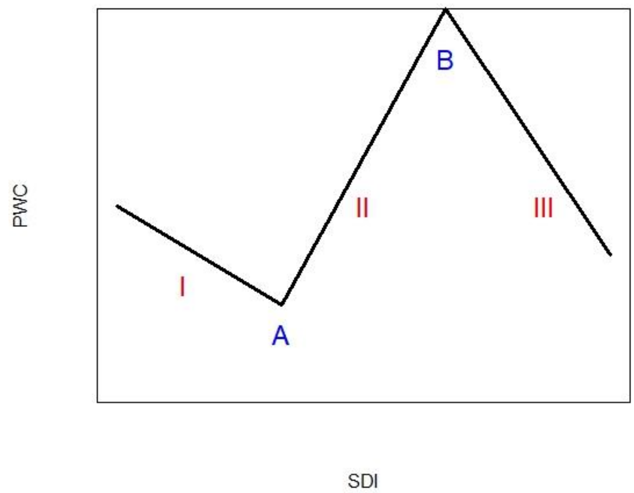


Figure 3 A schematic diagram of the proposed SDI–PWC overall relationship, which contains three developing stages (I, II, III) and two turning points (A, B). (SDI: Socio-demographic Index; PWC: population-weighted concentration)

There are two main differences between the proposed relationship and the EKC. Firstly, the proposed relationship divides the L stage in the EKC into two stages (stage I and stage II), where stage I indicates that at the early stage of social development, PWC is negatively correlated with SDI. Secondly, there are two turning points rather than one in the proposed relationship. Turning point A denotes the point where the monotonically decreasing relationship turns into a monotonically increasing relationship, while B is the same as the turning point M in the EKC. Besides, it is worth mentioning that the beginning of our proposed overall relationship denotes a time when industrialization has gained preliminary scale in a certain country. This is because our relationship starts from a relative high level of PWC and one should not expect such levels for countries whose economy relies dominantly on agriculture.

The reasons for the transition from stage II to III are similar to those that explain the shape of the EKC. The most prevalent reasons are:

1. In contrast with the transition from an agricultural to an industrial economy, the transition from an industrial to a services and technology economy will reduce pollutant emissions.

2. People in richer countries are more sensitive to their environmental quality. As a result, indigenous industries that generate severe pollution may have to relocate to poorer countries where environmental regulations are less strict for higher profit. For example, a previous study (Solarin et al. 2017) investigated the relationship between foreign direct investments and CO_2 emissions in Ghana and found evidence that there was a positive impact from the former on the latter. It indicated that foreign capital helped to relocate high-pollution industries to poor countries like Ghana where the environmental restrictions were weaker.

3. Governments and societies in richer countries are willing to pay more for environmental protection and pollution reduction.

4. Increasing returns to scale occur in pollution abatement. (Dinda 2004; Kaika and Zervas 2013; Managi 2006).

The formation of stage I has been observed by previous studies related to the EKC, with the following explanatory insights:

1. When the per capita income level is below a certain threshold, industries in that country can adopt only primitive "dirty" technologies. However, as the economy grows, high-polluting technologies will be gradually replaced by clean technologies.

2. Less developed countries are usually less polluted, which means the costs to mitigate pollution in these areas are much lower than those of countries whose industries are booming.

3. Poorer countries usually own fewer pollution sources (such as power plants or factories). Thus, there is less pressure on the government to regulate the sources. (Lapinskiene et al. 2014; Liu 2008; Managi 2006).

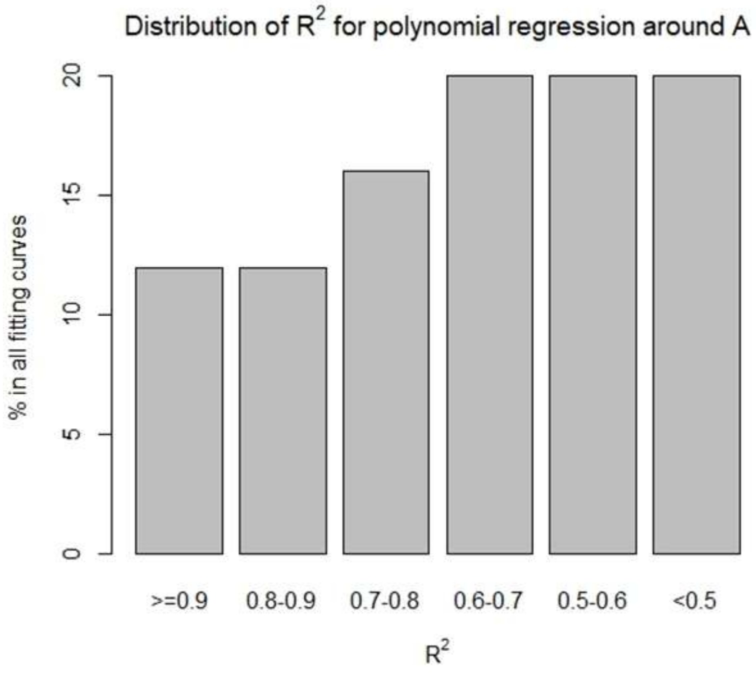
Previous GBD studies (Forouzanfar et al. 2016; Gakidou et al. 2017) failed to establish a common overall relationship between social development and ambient $PM_{2.5}$ pollution because they adopted the approach of plotting the data of all countries in one coordinate system and then fitting a cloud of data. Instead, we considered the relationship as being highly country-specific and focused more on the shape and trend shown by each country's data rather than their value.

3.1.2 Turning point A

The dataset contains 40 countries and regions whose highest SDI is less than 0.5, which means they were situated at the early stage of social development from 1990 to 2017. Thus, it is unlikely that they were experiencing the R stage in the EKC. Twenty-five of these countries demonstrated the existence of stage I and turning point A, which is judged by three criteria: firstly, the lowest PWC value in the whole time span is located neither in 1990 nor 2017; secondly, the maximum difference between the PWC at turning point A (denoted as PWC_A) and the PWC before A is larger than PWC_A*10%; and thirdly, the time span between 1990 and turning point A is at least 10 years. Because more than half of the investigated countries showed indications of stage I and turning point A, and given that countries that failed the criteria may have had a higher PWC before 1990, it is sensible to regard stage I and turning point A as common characteristics rather than anomalies.

Figure 4 (upper) shows the distribution of the coefficient of determination (R^2) of 2-degree polynomial regression of the data around turning point A in these 25 countries. The regression was performed in the SDI range corresponding to the two highest PWC values before and after turning point A. Since degrees of freedom in polynomial regression are related to the number of data points, the values of degrees of freedom in this regression analysis varied from 3 to 25, depending on the length of the investigated SDI range in each country. Thus, we calculated the value of adjusted R^2 to eliminate the influence of different degrees of freedom. The results for adjusted R^2 are shown in Figure 4 (lower).

According to our results, 2-degree polynomial regression seems to be an acceptable choice for capturing the main features around turning point A. This is because only 20% of all R^2 and adjusted R^2 values are below 0.5. However, it is worth mentioning that there are several countries whose R^2 and adjusted R^2 are much lower than 0.5 with the lowest value being 0.31 and 0.23 respectively. The Kolmogorov-Smirnov test can also be used for judging the goodness of fit (Gaddis and Gaddis 1990). We ran this test on each country's fitted data, and the results showed the null hypothesis (that the actual data and the fitted data are from the same distribution) should be rejected for 28% of countries under a confidence level of 0.9. Therefore, it is safer to conclude that the relationship around turning point A is highly country-specific, and it is difficult to find a form of function that can accurately depict the features of all countries although 2-degree polynomial regression can serve as an initial attempt.



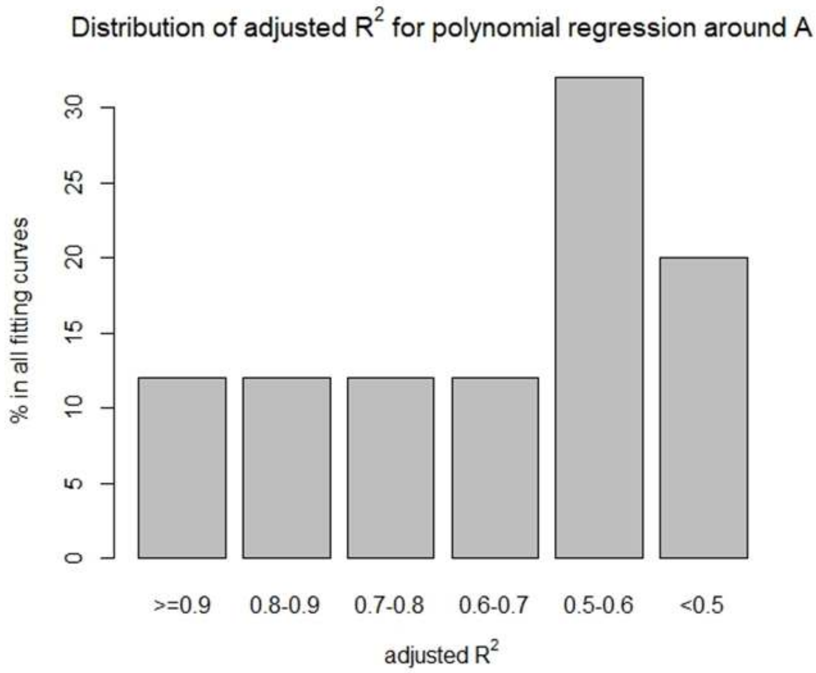


Figure 4 (Upper) The distribution of the R^2 values of 2-degree polynomial regression performed in the SDI range corresponding to the two highest PWC values before and after turning point A. (Lower) The same as upper plot but for adjusted R^2 values. (SDI: Socio-demographic Index; PWC: population-weighted concentration)

Figure 5 shows the distribution of the corresponding SDI of turning point A (denoted as SDI_A) in these 25 countries. Maximum and minimum SDI_A values were 0.48 and 0.18 respectively. The mean SDI_A value is 0.37 while the standard deviation is 0.085. Figure 5 demonstrates that turning point A is not likely to be reached when the SDI is less than 0.2. However, it is not possible to identify the most probable SDI zone for turning point A to appear. This is because the p-value of the chi-squared test conducted in the SDI range from 0.2 to 0.5 is 0.22, which means that the null hypothesis that SDI_A is uniformly distributed within this range cannot be rejected.

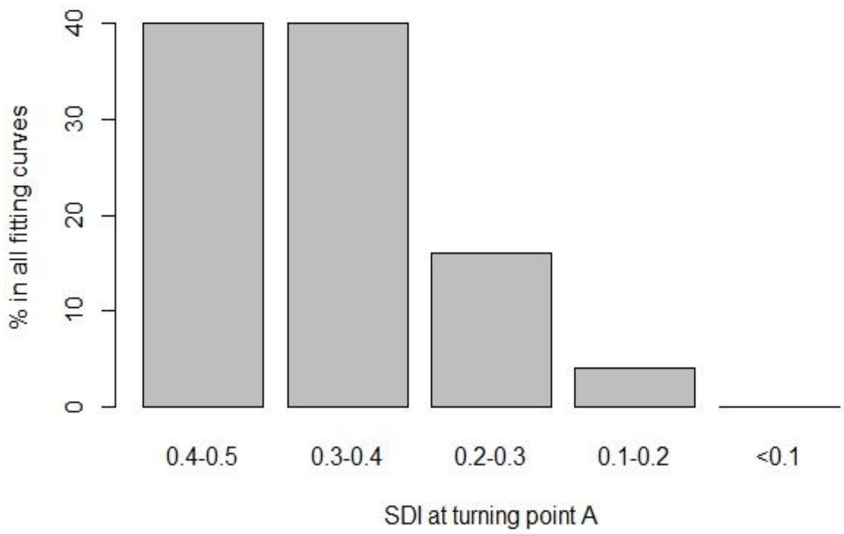


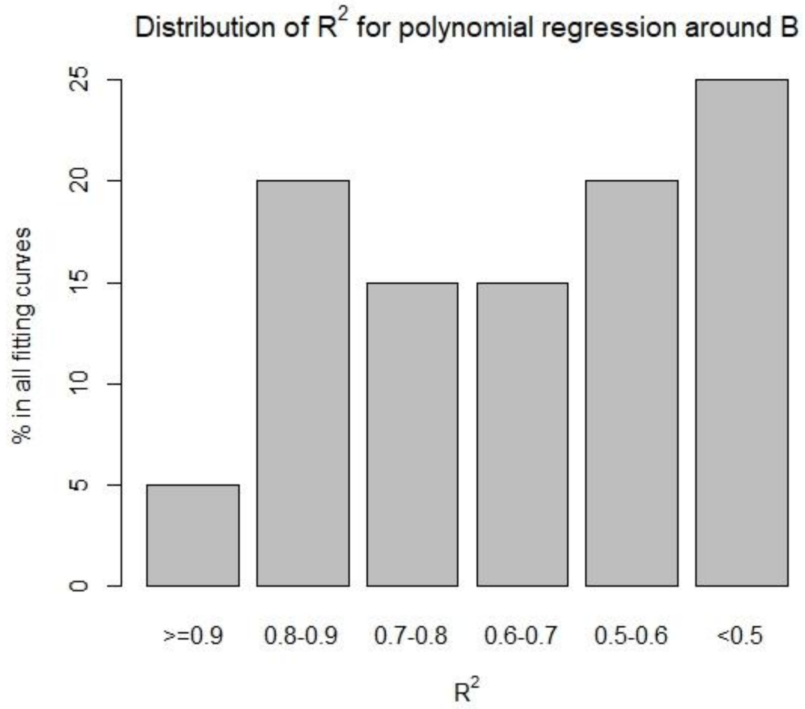
Figure 5 Distribution of the SDI at turning point A. (SDI: Socio-demographic Index)

3.1.3 Turning point B

The three criteria used to identify turning point B are: firstly, considering the fact that only when a country is developed enough can it experience stage III, the maximum value of PWC in 1990–2017 (denoted as PWC_{MAX}) will correspond to an SDI value higher than 0.5; secondly, PWC_{MAX} is located at neither 1990 nor 2017; thirdly, PWC_{MAX} is at least 10% higher than the two minimum values before and after PWC_{MAX}. Based on the first two criteria, 66 countries were selected out. By adding the third criterion, 20 countries in the dataset were identified as experiencing turning point B between 1990 and 2017. Although the ratio of countries experiencing B does not seem very high at first glance, this can be attributed to the threshold value of SDI that we decided. As a matter of fact, if the SDI threshold were raised up to 0.75, which means turning point B shall occur only in a well-developed country, 12 countries would meet the first two criteria, and 50% of them would meet the third criterion. Threshold elevation is reasonable because at least one of the key parameters related to $PM_{2.5}$ pollution, i.e. the number of vehicles usually keeps increasing as the economy grows, and it would take a lot of resources to compensate for this influence. Furthermore, it should be noticed that more countries will potentially fulfil the third criterion if the investigated time span is lengthened. Therefore, it is reasonable to admit that turning point B is a common feature in the SDI-PWC overall relationship.

Figure 6 (upper) shows the distribution of the coefficient of determination (R^2) of 2-degree polynomial regression for these 20 countries. The regression was performed in the SDI range corresponding to the two lowest PWC values before and after turning point B. Dependent on the length of the investigated SDI range, the degrees of freedom in this regression analysis varied from 4 to 25. Same as the process for turning point A, we calculated adjusted R^2 to eliminate the influence of different degrees of freedom. The results for adjusted R^2 are shown in Figure 6 (lower).

Similar to the results for turning point A, it is shown that 2-degree polynomial regression is generally acceptable for capturing the main features around turning point B, because 25% of all R^2 values and 40% of all adjusted R^2 values are less than 0.5. However, this choice is far from being universally correct because there are countries whose R^2 and adjusted R^2 are quite low with the lowest to be 0.31 and 0.26 respectively. Additionally, the Kolmogorov-Smirnov test results show that the null hypothesis (the actual data and the fitted data are from the same distribution) should be rejected for 15% of countries under a confidence level of 0.9. Thus, even though we propose that the overall curve would take the shape of a tilted-S and 2-degree polynomial regression can serve as an initial attempt for depicting features around the turning points, we do not regard 3-degree polynomial regression as a good choice to depict the major parts of the three stages. This is because a good fit to three-degree polynomial regression indicates a high degree of regularity and symmetry existing in the data, which is rarely observed during our research.



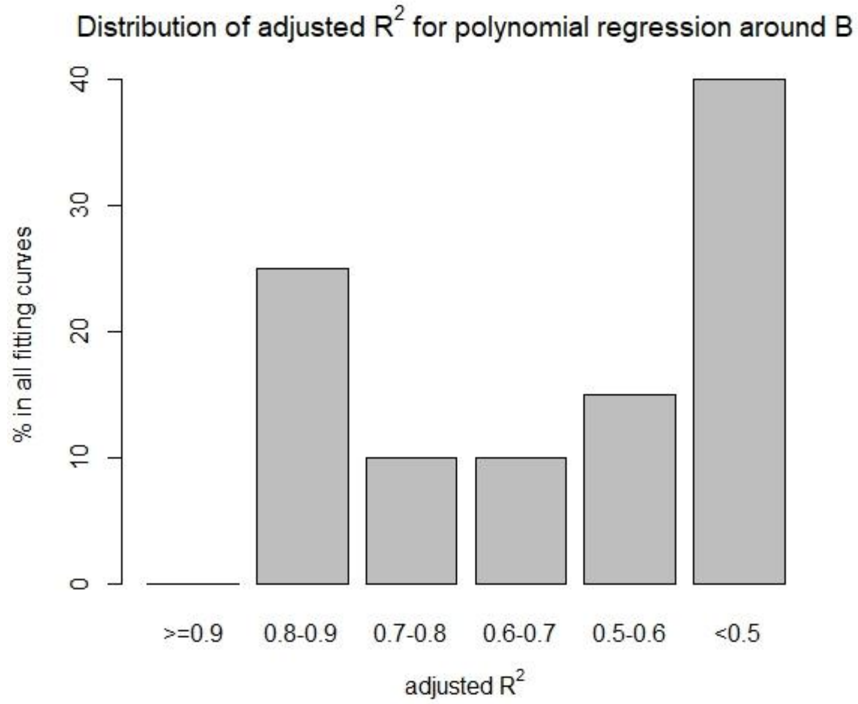


Figure 6 (Upper) The distribution of the R^2 values of 2-degree polynomial regression in the SDI range corresponding to the two lowest PWC values before and after turning point B. (Lower) The same as upper plot but for adjusted R^2 values. (SDI: Socio-demographic Index; PWC: population-weighted concentration)

Figure 7 shows the distribution of the corresponding SDI of turning point B (denoted as SDI_B) in these 20 countries. Maximum and minimum SDI_B values are 0.86 and 0.50 respectively. The mean value of SDI_B is 0.66 while the standard deviation is 0.12. Figure 7 demonstrates that turning point B is likely to be reached before the SDI exceeds 0.9. However, it is not possible to further narrow down the most probable SDI zone for turning point B to occur. This is because the p-value of the chi-squared test conducted in the SDI range from 0.5 to 0.9 is 0.22, which means that the null hypothesis that SDI_B is uniformly distributed within this range cannot be rejected.

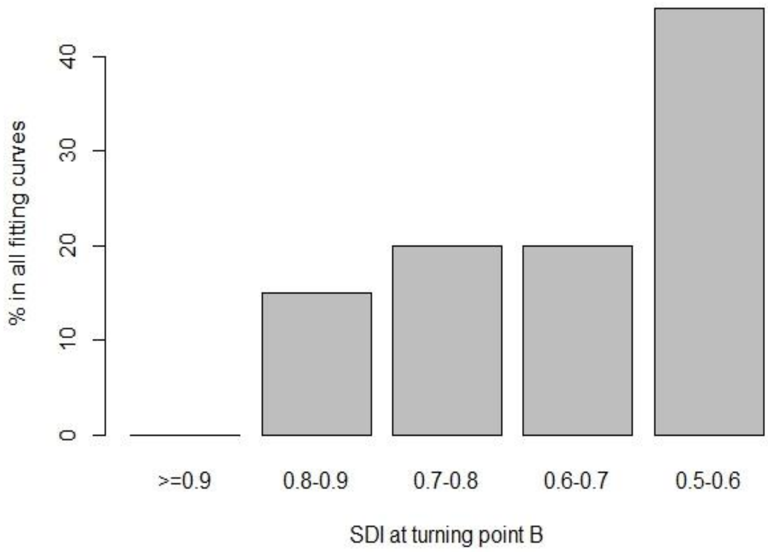


Figure 7 Distribution of the SDI at turning point B. (SDI: Socio-demographic Index)

3.2 Prediction of PWC

As the discussion in section 3.1 shows, the location of both turning points A and B are difficult to predict. In other words, these two important indices determining the shape of the SDI–PWC curve are highly country-specific. Moreover, considering that the average annual SDI growth rate for a given country is also difficult to estimate because it is related to politics, it is even more difficult to predict the advent of turning points in terms of years.

One aspect that is especially important for developing countries and can be investigated further is the short-term (i.e., next year) PWC prediction. The results of applying SVR to PWC short-term prediction are quite promising. The model's MSE and RMSE for the validation datasheet are 4.84 and 2.20 respectively (the mean and standard deviation value of the validation datasheet are 33.57 and 15.79 respectively). Although our research is aimed at predicting PWC in the next year, it is not difficult to use a similar method for prediction in the next two or three years. Certainly, the accuracy will decrease as the prediction goes further.

Figure 8 shows two typical countries' retrospective predictions of short-term PWC via SVR. Figure 8 demonstrates that SVR is generally able to predict the yearly temporal evolution of PWC. However, when the trend of PWC evolution is changing, especially when rapid changes occur, SVR cannot always recognize the symptoms. Notice that the horizontal axis in figure 8 is in the unit of Years rather than SDI values. This is because it is more helpful for policy makers to be aware of the expectation of PWC in the coming years. In comparison, PWC expectation of a future SDI is not that straightforward because one will also need to estimate SDI growth to know when the expectation will become true.

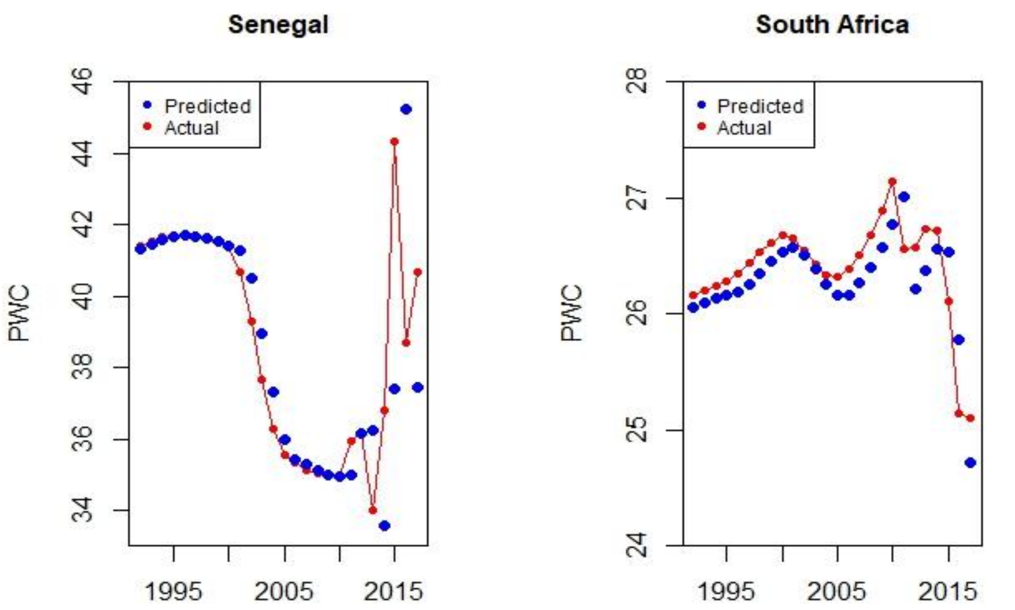


Figure 8 Two typical countries' retrospective predictions of short-term PWC via SVR. The red lines in the plots are actual PWC values, while the blue dots are predictions via SVR with the previous two years' SDI and PWC values as historical information. (SVR: support vector regression; SDI: Socio-demographic Index; PWC: population-weighted concentration)

4 Conclusions and limitations

The focus of this research was on the relationship between social development and ambient particulate matter pollution using indices more representative of the actual situation. The main conclusions of this research are as follows:

1. Compared with the traditional EKC shape, a new general relationship in the shape of a tilted-S is proposed. The proposed relationship contains two turning points (A and B) and three stages (I, II, and III).

2. Two-degree polynomial regression can serve as an initial choice to capture the main features around the turning points in the relationship, but this is far from universal. Three-degree polynomial regression may not be effective in depicting the overall relationship.

3. Because the location of turning points is highly country-specific, it is difficult to narrow down the SDI zone of the turning points. Thus, it is not possible to predict PWC over the long term.

4. SVR is a powerful tool for short-term PWC prediction.

This research has some limitations: firstly, our analysis process neglects the errors and uncertainties that can exist in our data sources (although the GBD's data product represents the highest quality level among its peers); secondly, the identification criteria for turning points do not completely rule out subjective effects; and thirdly, because the time span in our research was not long enough to show the whole SDI–PWC curve, it is reasonable to suspect that some of the identified turning points are not turning points, but just local minimums or maximums.

5 Policy implications

Based on the results of our research, we advise policy makers to treat predictions related to turning points in previous studies with caution, and to invest more resources in addressing present particulate matter pollution rather than attempting to achieve a more accurate forecast about turning points. Since economic development will not automatically solve the problem of particulate matter pollution, we also advise policy makers to act more aggressively when managing this issue. Possible measures include investing in environmental techniques, enacting stricter emission standards for vehicles and factories, promoting the use of sustainable energy, and enhancing the supervision of pollution sources. Because no health impact threshold is observed for $PM_{2.5}$ concentration, policy makers should continue to strive to mitigate $PM_{2.5}$

pollution to improve the quality of public health. Finally, since education is a key factor in elevating citizens' environmental awareness, which can help to cultivate a society prioritizing environmental protection, policy makers should also reinforce the investments in education as a long-term strategy against particulate matter pollution.

Acknowledgements

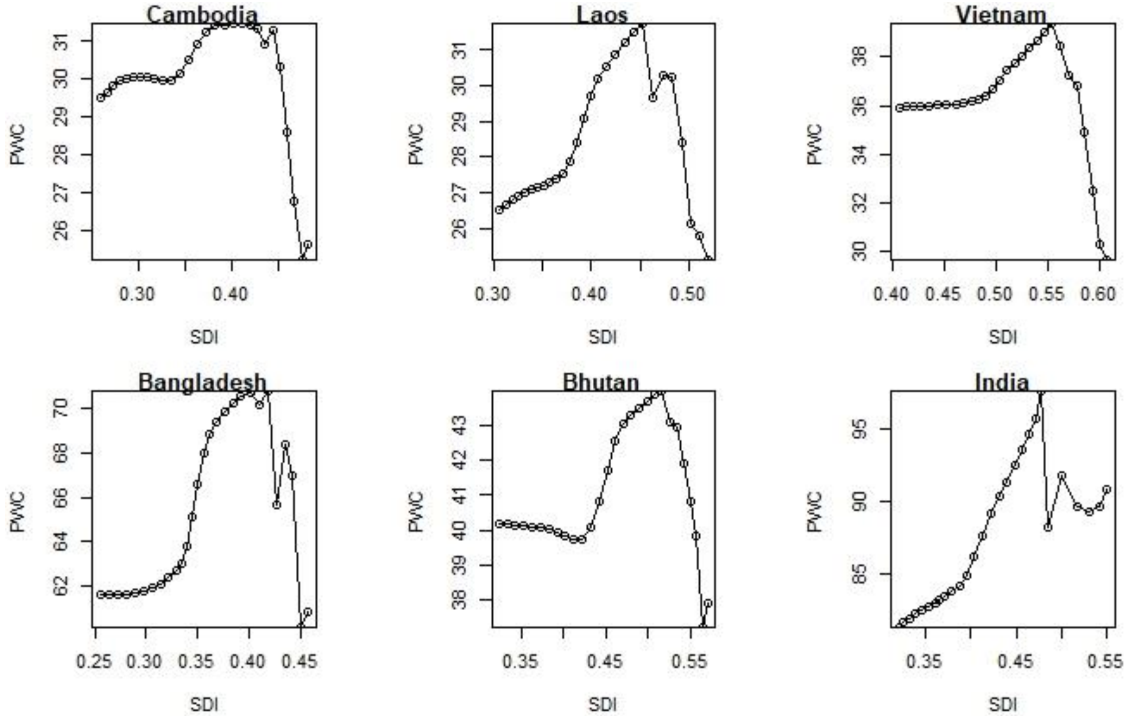
Nairui Liu is supported by the Chinese Scholarship Council for his PhD study. The authors also gratefully acknowledge colleagues in the International Laboratory for Air Quality and Health, Dr. Helen Thompson from Queensland University of Technology, and Professor Lana Friesen from The University of Queensland for their advice. We would like to thank the two anonymous reviewers as well for their suggestions have greatly helped to improve this paper.

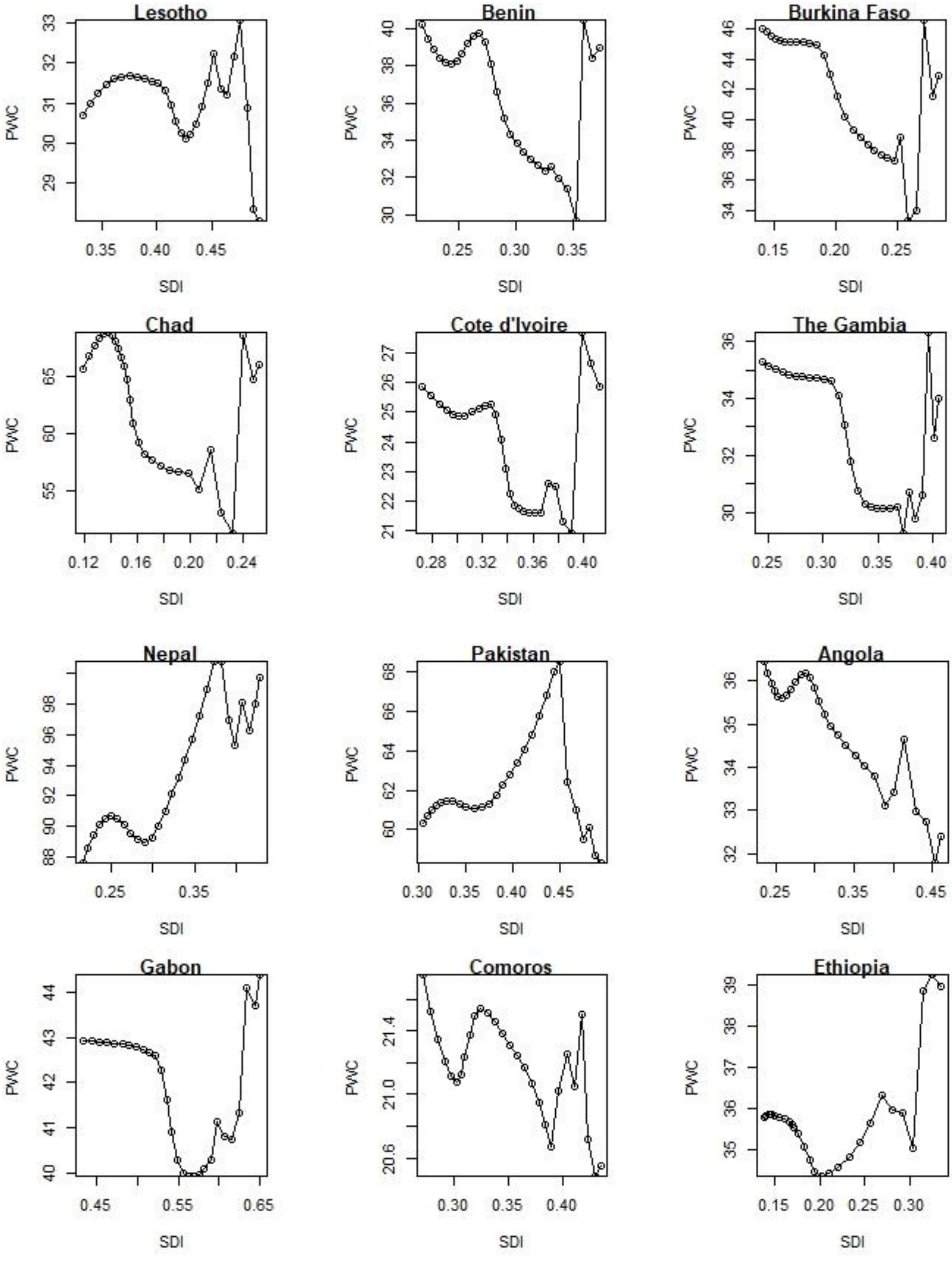
Funding sources

This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary figure

Figure S1 shows the SDI–PWC curve of all the countries (26 in total) that have maintained an annual SDI growth of more than 1% from 1990 to 2017. Ignoring local variations, their shapes conform to one of three categories: "V" shape, monotonic increasing shape, or "inverted-V" shape.


Cambodia
Laos
Vietnam
Bangladesh
Bhutan
India
PWC
SDI


Lesotho
Benin
Burkina Faso
Chad
Cote d'Ivoire
The Gambia
Nepal
Pakistan
Angola
Gabon
Comoros
Ethiopia
PWC
SDI

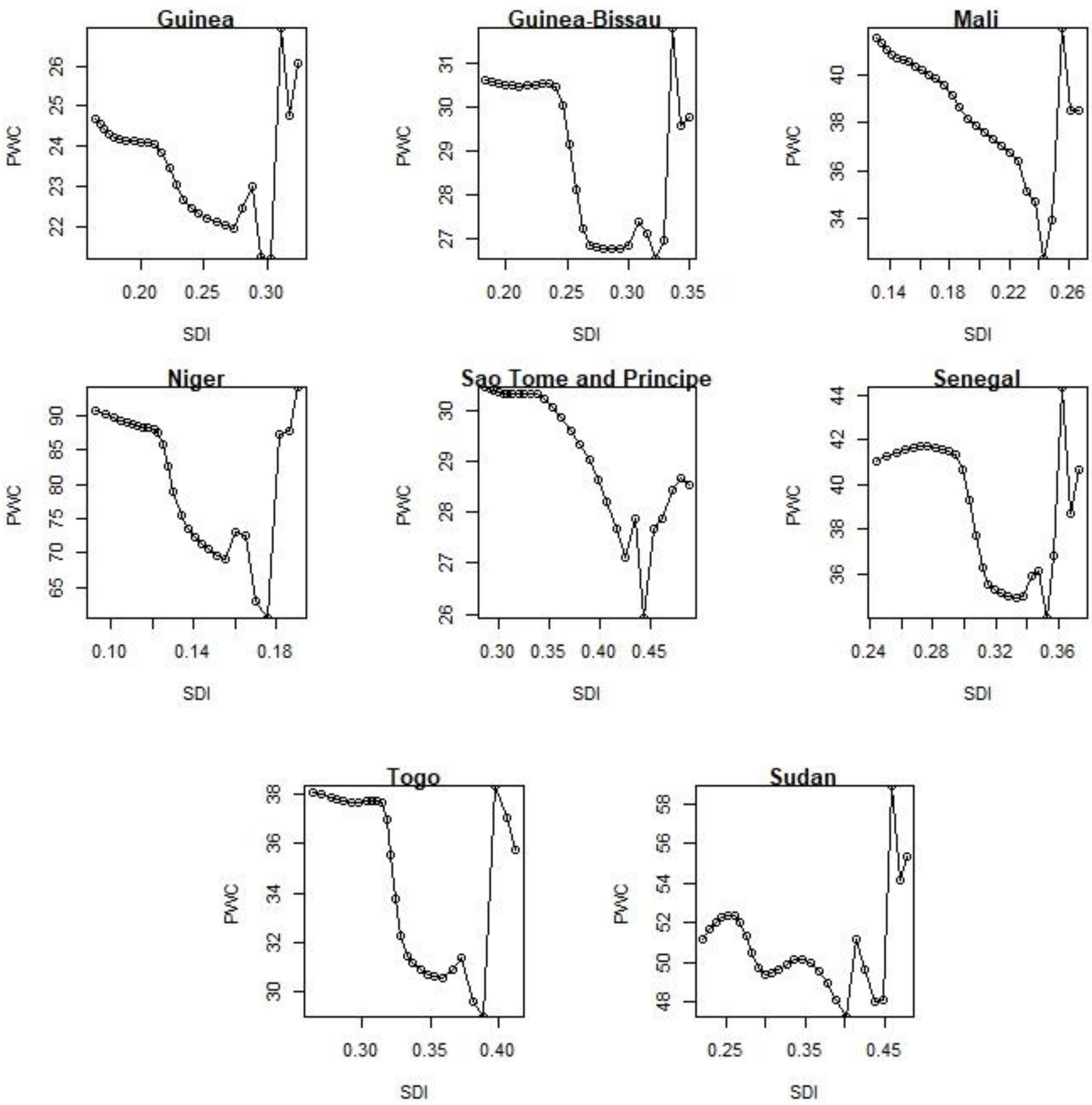


Figure S1. The SDI–PWC curve of all the countries (26 in total) that have maintained an annual SDI growth of more than 1% from 1990 to 2017. The PWC values are expressed in micrograms per cubic meter. (SDI: Socio-demographic Index; PWC: population-weighted concentration)

References

Dinda S (2004) Environmental Kuznets Curve hypothesis: A survey Ecological Economics 49:431-455 doi:10.1016/j.ecolecon.2004.02.011

Dong B, Cao C, Lee SE (2005) Applying support vector machines to predict building energy consumption in tropical region Energy and Buildings 37:545-553 doi:10.1016/j.enbuild.2004.09.009

Dong K, Sun R, Dong C, Li H, Zeng X, Ni G (2018) Environmental Kuznets curve for PM2.5 emissions in Beijing, China: What role can natural gas consumption play? Ecological Indicators 93:591-601 doi:10.1016/j.ecolind.2018.05.045

Forouzanfar MH et al. (2016) Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015 The Lancet 388:1659-1724 doi:10.1016/S0140-6736(16)31679-8
Gaddis GM, Gaddis ML (1990) Introduction to biostatistics: Part 5, statistical inference techniques for hypothesis testing with nonparametric data Annals of Emergency Medicine 19:1054-1059 doi:https://doi.org/10.1016/S0196-0644(05)82571-5
Gakidou E et al. (2017) Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016 The Lancet 390:1345-1422 doi:10.1016/S0140-6736(17)32366-8
Grossman GM, Krueger AB (1991) Environmental impacts of a North American free trade agreement. National Bureau of Economic Research,
Jain RK, Smith KM, Culligan PJ, Taylor JE (2014) Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy Applied Energy 123:168-178 doi:https://doi.org/10.1016/j.apenergy.2014.02.057
Ji X, Yao Y, Long X (2018) What causes PM2.5 pollution? Cross-economy empirical analysis from socioeconomic perspective Energy Policy 119:458-472 doi:10.1016/j.enpol.2018.04.040
Kaika D, Zervas E (2013) The Environmental Kuznets Curve (EKC) theory-Part A: Concept, causes and the CO2 emissions case Energy Policy 62:1392-1402 doi:10.1016/j.enpol.2013.07.131
Lapinskiene G, Tvaronavičiene M, Vaitkus P (2014) Greenhouse gases emissions and economic growth - evidence substantiating the presence of environmental Kuznets curve in the EU Technological and Economic Development of Economy 20:65-78 doi:10.3846/20294913.2014.881434
Leisch DMaEDaKHaAWaF (2018) e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien
Liu L (2008) Sustainability efforts in China: Reflections on the environmental Kuznets curve through a locational evaluation of "Eco-Communities" Ann Assoc Am Geogr 98:604-629 doi:10.1080/00045600802013452
Luo K, Li G, Fang C, Sun S (2018) PM2.5 mitigation in China: Socioeconomic determinants of concentrations and differential control policies Journal of Environmental Management 213:47-55 doi:10.1016/j.jenvman.2018.02.044
Managi S (2006) Are there increasing returns to pollution abatement? Empirical analytics of the Environmental Kuznets Curve in pesticides Ecological Economics 58:617-636 doi:10.1016/j.ecolecon.2005.08.011
Pui DYH, Chen S-C, Zuo Z (2014) PM2.5 in China: Measurements, sources, visibility and health effects, and mitigation Particuology 13:1-26 doi:https://doi.org/10.1016/j.partic.2013.11.001
R Core Team (2017) R: A Language and Environment for Statistical Computing
Roth GA et al. (2018) Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017 The Lancet 392:1736-1788 doi:https://doi.org/10.1016/S0140-6736(18)32203-7
Solarin SA, Al-Mulali U, Musah I, Ozturk I (2017) Investigating the pollution haven hypothesis in Ghana: An empirical investigation Energy 124:706-719 doi:https://doi.org/10.1016/j.energy.2017.02.089

Stanaway JD et al. (2018) Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017 The Lancet 392:1923-1994 doi:10.1016/S0140-6736(18)32225-6

Stern DI (2017) The environmental Kuznets curve after 25 years Journal of Bioeconomics 19:7-28 doi:10.1007/s10818-017-9243-1

Stern DI, Zha D (2016) Economic growth and particulate pollution concentrations in China Environmental Economics and Policy Studies 18:327-338 doi:10.1007/s10018-016-0148-3

Vapnik V (2013) The nature of statistical learning theory. Springer science & business media,

Wu J, Zheng H, Zhe F, Xie W, Song J (2018) Study on the relationship between urbanization and fine particulate matter (PM2.5) concentration and its implication in China Journal of Cleaner Production 182:872-882 doi:https://doi.org/10.1016/j.jclepro.2018.02.060

Yeganeh B, Motlagh MSP, Rashidi Y, Kamalan H (2012) Prediction of CO concentrations based on a hybrid Partial Least Square and Support Vector Machine model Atmos Environ 55:357-365 doi:https://doi.org/10.1016/j.atmosenv.2012.02.092