

# Xây dựng chỉ số tổng hợp đánh giá mức độ phát triển kinh tế cấp tỉnh tại Việt Nam

Nguyễn Văn Thắng\*

## TÓM TẮT

Trên địa bàn cấp tỉnh tại Việt Nam, hiện nay tồn tại rất nhiều chỉ tiêu thống kê riêng lẻ phản ánh tình hình kinh tế - xã hội, gây khó khăn để nắm bắt, tóm tắt thông tin chung. Nhận thức về sự phát triển kinh tế của một địa phương được gắn liền với chỉ tiêu tăng trưởng Tổng sản phẩm trên địa bàn cấp tỉnh (GRDP) chưa thể hiện được tính đa chiều trong khái niệm phát triển kinh tế mang lại. Do vậy, xây dựng chỉ số tổng hợp nhằm đo lường mức độ phát triển kinh tế trên địa bàn cấp tỉnh là rất cần thiết góp phần bổ sung công cụ thống kê đánh giá sự phát triển kinh tế của cấp tỉnh tại Việt Nam. Chỉ số tổng hợp được hình thành từ 19 chỉ tiêu của 5 thành phần: Đặc điểm nhân khẩu, lao động – việc làm, nguồn lực, năng lực cạnh tranh và chất lượng cuộc sống. Nguồn dữ liệu được thu thập chủ yếu từ dữ liệu ngành Thống kê trong giai đoạn 2016-2020. Dữ liệu được chuẩn hóa theo Z-score với trọng số được ước lượng theo phương pháp thành phần chính (PCA) và chỉ số tổng hợp tính từ trung bình có trọng số của 5 thành phần chính với trọng số là phương sai giải thích của các thành phần. Kết quả chỉ số tổng hợp cho thấy, Thành phố Hồ Chí Minh, Hà Nội là trung tâm kinh tế cả nước khi có mức độ phát triển cao nhất, ngược lại các tỉnh vùng Trung du và miền núi phía Bắc có mức độ phát triển thấp nhất. Ngoài ra, nghiên cứu cho thấy sự chênh lệch khá rõ nét về mức độ phát triển kinh tế giữa 06 vùng tại Việt Nam. Bên cạnh đó, có bằng chứng thống kê về tương quan không gian giữa các tỉnh, thành Việt Nam, do vậy mô hình Durbin không gian được sử dụng để thay thế mô hình hồi quy OLS dữ liệu bảng. Phân tích hồi quy không gian cho thấy, tỷ trọng dân số, doanh nghiệp, HDI có tác động đến mức độ phát triển kinh tế của chính địa phương sở tại, ngoài ra tỷ trọng dân số còn tác động tích cực đến mức độ phát triển kinh tế của các địa phương lân cận.

**Từ khoá:** Chỉ số tổng hợp, phát triển kinh tế, hồi quy không gian

Cục Thống kê Thành phố Hồ Chí Minh, Việt Nam

## Liên hệ

**Nguyễn Văn Thắng**, Cục Thống kê Thành phố Hồ Chí Minh, Việt Nam

Email: thangnguyen.923102060056@st.ueh.edu.vn

## Lịch sử

- Ngày nhận: 08-5-2024
- Ngày sửa đổi: 12-7-2024
- Ngày chấp nhận: 21-8-2024
- Ngày đăng: 30-9-2024

## DOI:

<https://doi.org/10.32508/stdjelm.v8i3.1409>



## Bản quyền

© ĐHQG Tp.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



## GIỚI THIỆU

Trong bối cảnh toàn cầu hóa, sự thay đổi về địa chính trị, kinh tế, xã hội gây ra những khủng hoảng cho các quốc gia và khu vực. Tác động từ các yếu tố bên ngoài trở thành yếu tố chính cản trở mục tiêu phát triển kinh tế của một quốc gia. Chính vì vậy, việc theo dõi mục tiêu phát triển kinh tế cần phải thực hiện thường xuyên và kịp thời góp phần thúc đẩy phát triển kinh tế của quốc gia. Do đó, việc xây dựng chỉ số tổng hợp để đo lường sự phát triển kinh tế và sử dụng nó như một công cụ quản lý, hỗ trợ việc lập kế hoạch, làm cơ sở xây dựng chính sách phù hợp nhằm tăng cường cho sự phát triển kinh tế - xã hội.

Đo lường phát triển kinh tế cũng dần thay đổi theo thời gian, những nỗ lực ban đầu là thước đo tăng trưởng kinh tế, trong đó GDP được xem là tiêu chí chính cho phúc lợi và tiến bộ quốc gia<sup>1</sup>. Tuy nhiên, việc sử dụng GDP làm thước đo duy nhất để đánh giá, xếp hạng các quốc gia dẫn đến những tranh luận khi cho rằng GDP chưa đánh giá đến phúc lợi xã hội<sup>2,3</sup> và GDP được xem là di tích của một thời kỳ chi phối bởi

sản xuất<sup>1</sup>. HDI được xem là chỉ số thay thế GDP và trở nên phổ biến để đánh giá, so sánh giữa các quốc gia do tính đơn giản của khái niệm, cách tính toán, giải thích chỉ số<sup>4</sup>. Tuy nhiên, HDI cũng gặp phải chỉ trích về việc xác định trọng số<sup>5</sup> và HDI rất hữu ích cho các quốc gia có đặc điểm tương tự nhau nhưng đối với các quốc gia khác nhau thì không có bổ sung điều gì mới<sup>6</sup>. Trong vài thập kỷ qua đã có số lượng lớn nghiên cứu về đo lường chỉ số phát triển kinh tế - xã hội và đi kèm với nó là một số vấn đề tranh luận, tính vô tận của chủ đề này để lại tiềm năng lớn cho các nghiên cứu khoa học. Hiện tại không có cách tiếp cận thống nhất để đo lường mức độ phát triển kinh tế nhưng cách tiếp cận của các tổ chức quốc tế được xem phổ biến hơn và được công nhận rộng rãi<sup>7</sup>.

Tại Việt Nam, việc thực hiện nghị Quyết Đại hội đại biểu các cấp đều đi kèm là hệ thống các chỉ tiêu chủ yếu về phát triển kinh tế - xã hội, trong đó tốc độ tăng trưởng GDP/GRDP được xem là thước đo chính cho phát triển kinh tế nhiệm kỳ (5 năm). Bên cạnh đó, có nhiều hệ thống chỉ tiêu được ban hành như Bộ chỉ

**Trích dẫn bài báo này:** Thắng N.V. **Xây dựng chỉ số tổng hợp đánh giá mức độ phát triển kinh tế cấp tỉnh tại Việt Nam.** *Sci. Tech. Dev. J. - Eco. Law Manag.* 2024; 8(3):5463-5479.

tiêu thống kê phát triển bền vững (158 chỉ tiêu); Bộ chỉ tiêu tổng hợp báo cáo định kỳ và báo cáo thống kê kinh tế - xã hội phục vụ chỉ đạo, điều hành của Chính phủ (200 chỉ tiêu); hệ thống chỉ tiêu thống kê cấp tỉnh (154 chỉ tiêu). Các hệ thống chỉ tiêu gây khó khăn trong nỗ lực nắm bắt, tóm tắt thông tin chung. Trong những năm qua có nhiều nghiên cứu xây dựng chỉ số tổng hợp cấp quốc gia Việt Nam và ngày càng có nhiều chỉ số tổng hợp được công bố trên địa bàn cấp tỉnh. Các chỉ số được hình thành nhằm đáp ứng yêu cầu quản lý của các lĩnh vực khác nhau nên các chỉ tiêu sử dụng cũng khác nhau. Qua tổng quan các nghiên cứu trước, tác giả chưa phát hiện nghiên cứu về xây dựng chỉ số tổng hợp bao quát các khía cạnh phát triển kinh tế cấp tỉnh, cũng như cấp quốc gia Việt Nam. Vì vậy, việc hình thành chỉ số tổng hợp để nắm bắt tính đa chiều trong kinh tế một cách có hệ thống và có thể so sánh được giữa các địa phương tại Việt Nam là cần thiết.

Từ thực tiễn trên, bài viết thực hiện nhằm giải quyết 2 mục tiêu chính:

Thứ nhất, xây dựng chỉ số tổng hợp phản ánh mức độ phát triển kinh tế cấp tỉnh với các chỉ tiêu đo lường phù hợp với đặc điểm tại Việt Nam; đây là cách tiếp cận mới dựa trên xây dựng chỉ số tổng hợp phản ánh mức độ phát triển kinh tế đa chiều nhằm bổ sung, thay thế chỉ tiêu GRDP đang sử dụng trên địa bàn cấp tỉnh. Thứ hai, phân tích các yếu tố ảnh hưởng đến mức độ phát triển kinh tế cấp tỉnh nhằm nhận dạng những yếu tố chủ yếu ảnh hưởng đến mức độ phát triển kinh tế của cấp tỉnh trong bối cảnh có xét đến tương quan về mặt không gian địa lý.

## CƠ SỞ LÝ THUYẾT VÀ TỔNG QUAN CÁC NGHIÊN CỨU LIÊN QUAN

Tăng trưởng kinh tế và phát triển kinh tế là những thuật ngữ sử dụng phổ biến, chúng có khái niệm khác nhau nhưng có mối liên hệ chặt chẽ với nhau. Tăng trưởng kinh tế là sự gia tăng quy mô sản lượng của nền kinh tế trong một thời gian nhất định (thường là trong một năm). Chỉ tiêu thường được sử dụng để đo lường tốc độ tăng trưởng kinh tế là Tổng sản phẩm quốc dân (GNP) hay Tổng sản phẩm quốc nội (GDP). Phát triển kinh tế là sự tăng trưởng kinh tế đi kèm với sự hoàn thiện cơ cấu kinh tế theo hướng hiện đại, nâng cao chất lượng thể chế kinh tế, đồng thời nâng cao chất lượng cuộc sống và đảm bảo công bằng, tiến bộ xã hội. Có thể thấy nội dung của phát triển kinh tế có phạm vi rộng hơn so với tăng trưởng kinh tế.

Chỉ số tổng hợp là tổng hợp các chỉ tiêu riêng lẻ dựa trên một mô hình cơ bản về khái niệm đa chiều được đo lường. Chỉ số tổng hợp trở nên phổ biến ở mọi

quốc gia, lĩnh vực vì chúng dễ dàng giải thích xu hướng của hiện tượng hơn là giải thích trên nhiều chỉ số riêng biệt<sup>8</sup> và giúp đơn giản hóa các cấu trúc phức tạp<sup>9</sup>. Quy trình xây dựng chỉ số tổng hợp được<sup>10</sup> để xuất tập trung vào các vấn đề chính: Xác định hiện tượng cần đo, lựa chọn chỉ tiêu, chuẩn hóa dữ liệu, trọng số và tổng hợp chỉ số.

### Tổng quan các nghiên cứu liên quan

HDI do UNDP đề xuất, được xem là chỉ số đo lường toàn diện và được sử dụng nhiều nhất nhằm đánh giá trình độ phát triển kinh tế - xã hội của các quốc gia<sup>7</sup>. HDI đo lường phát triển con người trên 3 phương diện: Sức khỏe, giáo dục và thu nhập của quốc gia/vùng. HDI được tổng hợp từ trung bình nhân giản đơn của 3 chỉ số thành phần.

Các nghiên cứu độc lập về xây dựng chỉ số tổng hợp cấp quốc gia có thể đề cập như phân tích kinh tế - xã hội của các quốc gia Châu Âu thông qua hiệu chỉnh HDI bằng cách sử dụng 20 chỉ số của 5 trụ cột<sup>11</sup>. Chỉ số tổng hợp là trung bình cộng giản đơn của các yếu tố; Nghiên cứu về xây dựng chỉ số phát triển kinh tế - xã hội của 65 quốc gia trong sáng kiến vành đai và con đường bằng việc sử dụng 17 chỉ số của 3 chỉ số thành phần, chỉ số chung phát triển kinh tế - xã hội được tính từ trung bình cộng của 3 chỉ số thành phần (3); xây dựng chỉ số tổng hợp phát triển kinh tế - xã hội của 59 quốc gia bằng cách tích hợp 12 chỉ số có sẵn vào 3 chỉ số phụ và chỉ số chung được hình thành từ trung bình số học của căn bậc hai của các chỉ số phụ<sup>12</sup>.

Đối với chỉ số tổng hợp cấp khu vực có các nghiên cứu như nghiên cứu về bất bình đẳng trong tăng trưởng kinh tế của các quận của Romania bằng cách dựa vào 3 chỉ số: GDP/người, năng suất lao động và tuổi thọ; chỉ số tổng hợp từ trung bình học giản đơn của 3 chỉ số thành phần<sup>13</sup>; hay đo lường các mục tiêu phát triển bền vững ở cấp độ khu đô thị của Romania bằng sử dụng 36 chỉ báo để hình thành một chỉ số tổng hợp chung và chỉ số tổng hợp được tính từ trung bình cộng giản đơn với trọng số được gán bằng nhau<sup>14</sup>; hay đánh giá sự phát triển chênh lệch các vùng cấp huyện tại Ấn Độ bằng dữ liệu thứ cấp với 21 chỉ số của 3 thành phần và chỉ số chung là trung bình cộng giản đơn<sup>15</sup>.

Tại Việt Nam, nghiên cứu về xây dựng chỉ số tổng hợp cấp quốc gia có thể đề cập đến như: Đánh giá phát triển bền vững của Việt Nam bằng cách sử dụng 30 chỉ tiêu của 4 thành phần: Kinh tế, xã hội, môi trường và chỉ tiêu tổng hợp. Các chỉ tiêu được sử dụng xây dựng chỉ số tổng hợp từ bình quân nhân giản đơn và bình quân nhân gia quyền với trọng số được tính bằng

phương pháp bán ma trận<sup>16</sup>; xây dựng chỉ tiêu đo lường chất lượng tăng trưởng kinh tế Việt Nam bằng dữ liệu thứ cấp với 12 chỉ tiêu phản ánh chất lượng tăng trưởng kinh tế Việt Nam của 3 nhóm: Kinh tế, xã hội và môi trường. Chỉ số tổng hợp là bình quân nhân với trọng số bằng nhau và trọng số chuyên gia<sup>17</sup>. Đối với cấp tỉnh của Việt Nam ngày càng có nhiều chỉ số công bố như Chỉ số năng lực cạnh tranh cấp tỉnh (PCI) do Văn phòng Thương mại và công nghiệp Việt Nam (VCCI) thực hiện bằng cách sử dụng 141 chỉ tiêu của 10 chỉ số thành phần. Chỉ số thành phần được tổng hợp từ trung bình cộng gia quyền với quyền số 40% từ dữ liệu thứ cấp và 60% từ dữ liệu khảo sát. Tổng cục Thống kê là cơ quan công bố chỉ số phát triển con người (HDI) cấp tỉnh tại Việt Nam. HDI được tổng hợp từ trung bình nhân giản đơn của 3 thành phần. Gần đây, Bộ khoa học và Công nghệ đã công bố chỉ số đổi mới sáng tạo địa phương (PII) bao gồm 52 chỉ số chia làm 7 trụ cột, dữ liệu khai thác từ dữ liệu hành chính và chỉ số tổng hợp dựa trên trung bình nhân có trọng số bằng nhau.

Trên cơ sở tổng quan các nghiên cứu trước, các hạn chế có thể được khái quát như sau: (1) hầu hết các chỉ số phát triển kinh tế được áp dụng ở cấp quốc gia; (2) chỉ số phát triển kinh tế cấp tỉnh chưa được xây dựng do chưa có chỉ tiêu phù hợp hoặc thiếu dữ liệu; (3) chưa có sự thống nhất về hệ thống chỉ tiêu đo lường phát triển kinh tế cấp địa phương do tính phức tạp của thước đo đa chiều; (4) thiếu phân tích sâu tác động đến sự chênh lệch phát triển kinh tế cấp địa phương hoặc cấp vùng. Nghiên cứu này sẽ thực hiện nhằm giải quyết các hạn chế nêu trên góp phần trong xây dựng chỉ số tổng hợp đo lường phát triển kinh tế cấp tỉnh tại Việt Nam.

## PHƯƠNG PHÁP NGHIÊN CỨU

Bài viết sử dụng dữ liệu thứ cấp được thu thập chủ yếu từ Niên giám Tổng Cục Thống kê, Cục Thống kê của 63 tỉnh, thành Việt Nam trong giai đoạn 2016-2020. Ngoài ra, nghiên cứu sử dụng dữ liệu hành chính từ Cục Sở Hữu Trí Tuệ để bổ sung biến đại diện.

### Phương pháp xây dựng chỉ số tổng hợp đo lường phát triển kinh tế cấp tỉnh Việt Nam

Nghiên cứu áp dụng quy trình xây dựng chỉ số tổng hợp phản ánh mức độ phát triển kinh tế cấp tỉnh tại Việt Nam trong giai đoạn 2016-2020, trong đó:

Lựa chọn chỉ số: Trên cơ sở lý thuyết và tổng quan các nghiên cứu và tính sẵn có dữ liệu để hình thành các chỉ tiêu đo lường phát triển kinh tế cấp tỉnh tại Việt Nam.

Chuẩn hóa Z-score: Được tính bằng cách lấy giá trị chỉ báo trừ cho giá trị trung bình và chia cho độ lệch

chuẩn. Sau khi thực hiện chuẩn hóa, dữ liệu có thang đo chung với trung bình bằng 0 và phương sai bằng 1. Công thức tính:

$$N_{ij} = \frac{X_{ij} - \mu}{\sigma}$$

Trong đó:

$N_{ij}$ : giá trị chuẩn hóa chỉ tiêu i, đơn vị j;

$X_{ij}$ : giá trị của chỉ tiêu i, đơn vị j;

$\mu$ : giá trị trung bình của chỉ tiêu i, đơn vị j;

$\sigma$ : độ lệch chuẩn chỉ tiêu i, đơn vị j.

Trọng số thu được bằng phương pháp PCA: Phân tích thành phần chính (Principal Component Analysis – PCA) là phương pháp dùng để giảm số chiều của một bộ dữ liệu mà trong đó số lượng lớn các biến có tương quan nhau trong khi vẫn giữ được lượng thông tin nhiều nhất có thể có từ bộ dữ liệu gốc. Việc giảm chiều được thực hiện bởi việc biến đổi tập hợp các biến gốc sang một tập hợp các biến mới gọi là thành phần chính. Các thành phần này không có tương quan với nhau và có thứ tự sao cho một vài thành phần đầu tiên giữ lại được nhiều nhất thông tin của bộ dữ liệu. Có thể diễn tả dưới dạng ma trận:

$$X_{n \times p} = Y_{n \times p} \cdot P$$

Trong đó:  $X$  là ma trận dữ liệu gốc, ma trận chứa các điểm thành phần và  $P$  là ma trận hệ số tải.

Các bước phân tích thành phần chính:

Bước 1: Xây dựng ma trận dữ liệu ban đầu và chuẩn hóa dữ liệu.

Bước 2: Tính ma trận tương quan và hệ số KMO.

Bước 3: Xác định số thành phần chính và tỷ lệ phương sai giải thích.

Bước 4: Tính điểm số tổng hợp căn cứ điểm thành phần chính và trọng số là tỷ lệ phương sai giải thích của các thành phần chính được giữ lại.

### Các yếu tố tác động đến mức độ phát triển kinh tế cấp tỉnh Việt Nam bằng hồi quy không gian.

#### Ma trận trọng số

Trong phân tích không gian, ma trận trọng số đóng vai trò quan trọng nhằm xác định tương quan không gian về vị trí địa lý giữa các địa phương/quốc gia. Ma trận  $W$  là một ma trận vuông để mô tả khoảng cách địa lý giữa các địa phương:

$$W = (w_{ij}) = \begin{pmatrix} 0 & \dots & w_{1j} \\ \dots & 0 & \dots \\ w_{i1} & \dots & 0 \end{pmatrix}$$

Trong nghiên cứu thực nghiệm, ma trận trọng số thường được lập dựa trên khoảng cách địa lý hoặc dựa

trên sự tiếp giáp ranh giới thực tế giữa các địa phương. Các ma trận trọng số thông dụng như ma trận trọng số liên kế, ma trận lân cận theo khoảng cách, ma trận tỷ lệ đường biên chung. Trong nghiên cứu này sử dụng ma trận trọng số khoảng cách lũy thừa được đo theo khoảng cách đường bộ (gọi là ma trận đường bộ), trong đó khoảng cách giữa các địa phương được xác định với khoảng cách đường bộ được thu thập từ w ww.distance.to và giả định rằng các địa phương càng gần nhau thì càng sự tương quan không gian càng chặt chẽ.

$w = (w_{ij})$  với  $w_{ij} = d_{ij}^{-\alpha}$ , Trong đó:  $\alpha=1$  và  $d_{ij}$  là khoảng cách đường bộ giữa 2 địa phương i và j.

### Kiểm định tương quan không gian

Các địa phương gần nhau thường có mối quan hệ với nhau thông qua trao đổi hàng hóa, lao động, hiệu ứng lan tỏa chính sách, công nghệ.... Mối quan hệ này gọi là tương quan không gian. Để phát hiện sự tồn tại hiện tượng tương quan không gian, công cụ thường được sử dụng là hệ số Moran's I. Đối với tương quan không gian của biến X thì hệ số Moran's I được tính từ công thức:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n [w_{ij} (X_i - \bar{X}) (X_j - \bar{X})]}{(\sum_{i=1}^n \sum_{j=1}^n w_{ij}) \sum_{i=1}^n (X_i - \bar{X})^2}$$

Trong đó:

$X_i$ : Là biến nghiên cứu của đối tượng thứ i;

$X_j$ : Là biến nghiên cứu của đối tượng thứ j;

$\bar{X}$ : Là giá trị trung bình của tất cả các đối tượng quan sát;

$w_{ij}$ : Trọng số không gian giữa hai đối tượng và đã được chuẩn hóa;

$n$ : Là số quan sát (số đối tượng khảo sát).

Hệ số Moran's I có giá trị từ -1 đến 1 ( $-1 \leq I \leq 1$ ). Nếu  $I < 0$ , cho biết có sự tương quan âm về mặt không gian giữa các đại lượng nghiên cứu và ngược lại  $I > 0$ , cho biết có sự tương quan dương về mặt không gian.

### Mô hình hồi quy không gian dữ liệu bảng

Hồi quy không gian với dữ liệu bảng sẽ góp phần hạn chế được vấn đề đa cộng tuyến so với dữ liệu chéo, tuy nhiên vẫn còn tồn tại hiện tượng phương sai thay đổi hay tự tương quan. Do vậy, khi ước lượng với dữ liệu bảng thì hai vấn đề trên được khắc phục bằng các mô hình tác động cố định (FEM) hay mô hình tác động ngẫu nhiên (REM). Mô hình kinh tế lượng không gian với dữ liệu bảng ngày càng phát triển mạnh cả về phương pháp luận và thực tiễn. Hiện có nhiều mô hình kinh tế lượng không gian được áp dụng, trong đó có 03 dạng mô hình phổ biến:

Mô hình tự hồi quy không gian (SAR) với dữ liệu bảng:

$$y_{it} = \rho W y_i + X_{it} \beta + \alpha_i + \varepsilon_{it}$$

Mô hình sai số không gian (SEM) với dữ liệu bảng:

$$y_{it} = X_{it} \beta + \alpha_i + u_{it}$$

$$u_{it} = \lambda W u_t + \varepsilon_{it}$$

Mô hình Durbin không gian (SDM) với dữ liệu bảng:

$$y_{it} = \rho W y_i + X_{it} \beta + W X_t \delta + \alpha_i + \varepsilon_{it}$$

Trong đó:

$y_{it}$ : Là vector của biến phụ thuộc ở kỳ t của i đối tượng;

$X_{it}$ : Là ma trận các giá trị cụ thể của biến độc lập;

$\alpha_i$ : Là vector thể hiện sự không đồng nhất của đối tượng gây ra bởi các yếu tố là đặc điểm riêng không đổi theo thời gian;

$u_{it}$ : Là sai số tự tương quan không gian;

$\varepsilon_{it}$ : Là sai số thỏa điều kiện;

W: Là ma trận trọng số thể hiện mối liên hệ không gian giữa các đối tượng;

Wy: Biến trễ không gian của biến phụ thuộc;

Wx: Biến trễ không gian của biến độc lập.

## KẾT QUẢ NGHIÊN CỨU

### Xây dựng chỉ số tổng hợp đo lường phát triển kinh tế cấp tỉnh Việt Nam

#### Chỉ tiêu đo lường

Trên cơ sở lý thuyết, tổng quan các nghiên cứu liên quan về xây dựng chỉ số tổng hợp đo lường phát triển kinh tế đã hình thành nên các thành phần của chỉ số và các chỉ tiêu đo lường được sử dụng tại các quốc gia. Nghiên cứu kết hợp với hệ thống các chỉ tiêu thống kê cấp tỉnh từ Bộ chỉ tiêu phát triển bền vững theo Thông tư 03/2019/TT-BKHDĐT ngày 22 tháng 01 năm 2019 Bộ Kế hoạch và Đầu tư; Quyết định số 293/QĐ-TTg ngày 24 tháng 02 năm 2020 của Thủ tướng Chính phủ Ban hành Bộ chỉ tiêu tổng hợp báo cáo định kỳ và báo cáo thống kê kinh tế - xã hội phục vụ sự chỉ đạo, điều hành của Chính phủ, Thủ tướng Chính phủ; Quyết định 05/2023/QĐ-TTg ngày 24 tháng 02 năm 2023 Ban hành hệ thống chỉ tiêu thống kê cấp tỉnh, cấp huyện, cấp xã. Trên cơ sở đó, nghiên cứu hình thành 21 chỉ tiêu phục vụ cho xây dựng chỉ số. Bước tiếp theo tác giả tham vấn 03 ý kiến chuyên gia, kết quả là có 05/21 chỉ tiêu để nghị không sử dụng vì trùng lặp thông tin và ý nghĩa chưa phù hợp (tỷ suất di cư thuần; số con binh quân của hộ; tỷ trọng lao động ngành nông nghiệp; tỷ lệ dân số trên 15 tuổi không hoạt động kinh

tế, giá tiêu dùng CPI) và 03 chỉ tiêu đề nghị bổ sung (tỷ lệ đô thị hóa; chỉ số giá không gian; số lượng đơn bảo hộ sáng chế, giải pháp hữu ích, kiểu dáng công nghiệp). Cuối cùng có 19 chỉ tiêu (Bảng 1) phản ánh 05 thành phần được sử dụng để xây dựng chỉ số đo lường mức độ phát triển kinh tế cấp tỉnh tại Việt Nam.

Trong đó, yếu tố nhân khẩu học được xem là nét đặc trưng, nguồn lực tác động đến phát triển kinh tế cấp tỉnh; yếu tố lao động, việc làm phản ánh sức khỏe của nền kinh tế, khi yếu tố lao động, việc làm tăng lên đồng nghĩa năng lực sản xuất được tăng cường; yếu tố nguồn lực phản ánh tiềm lực kinh tế của địa phương; yếu tố năng lực cạnh tranh phản ánh sức cạnh tranh thị trường trong nước và quốc tế; yếu tố chất lượng cuộc sống, phản ánh mục đích phát triển kinh tế là đem đến phúc lợi cho người dân.

### Chuẩn hóa dữ liệu

Có nhiều phương pháp chuẩn hóa, trong đó chuẩn hóa Min-Max là phương pháp được sử dụng nhiều nhất (30,5%) và chuẩn hóa Z-score (12,6%) (18). Chuẩn hóa Min-Max được thiết lập các biến đều có phạm vi giống hệt nhau, tuy nhiên, phương pháp này dựa trên giá trị lớn nhất, giá trị nhỏ nhất có thể là giá trị ngoại lai (outlier) nên các phạm vi này ảnh hưởng lớn đến kết quả cuối cùng. Đối với chuẩn hóa Z-Score có trung bình bằng 0, phương sai bằng 1, điều này không làm biến dạng giá trị trung bình và điều chỉnh được thang đo khác nhau với phương sai khác nhau. Đối với dữ liệu của 63 tỉnh, thành tại Việt Nam nhận thấy có sự chênh lệch lớn giữa các tỉnh, do vậy chuẩn hóa theo tiêu chuẩn Z-score là phù hợp với đặc thù riêng của cấp tỉnh tại Việt Nam.

### Xây dựng trọng số tổng hợp

Việc lựa chọn trọng số có tác động đáng kể đến chỉ số tổng hợp. Theo số tay<sup>8</sup> trọng số gồm: Trọng số chủ quan, phản ánh ý kiến cá nhân/nhóm người; trọng số khách quan thường được phân tích bằng các công cụ phân tích thống kê. Trong các nghiên cứu thực nghiệm thì phương pháp thống kê (trọng số khách quan) được sử dụng nhiều nhất (30,5%), trong đó phương pháp phân tích thành phần chính (PCA) chiếm 21,1%<sup>18</sup>. Vì vậy, nghiên cứu xây dựng trọng số dựa trên phương pháp sử dụng nhiều nhất là PCA.

### Phân tích tương quan và kiểm định Bartlett

Để kiểm chứng khả năng thích ứng của dữ liệu với phân tích PCA, nghiên cứu tiến hành đánh giá hệ số tương quan của các biến. Ma trận 19x19 chỉ tiêu cung cấp giá trị tương quan trung bình các biến đạt 0,43 (lớn hơn 0,3) cho thấy tập dữ liệu có tương quan với

nhau. Mặt khác, hệ số KMO = 0,76 (lớn hơn 0,5) và kiểm định Bartlett có giá trị  $P < 0,05$ , đủ cơ sở thống kê để kết luận có sự tương quan trong tập dữ liệu. Các đánh giá trên đều chỉ ra rằng dữ liệu trong tập nghiên cứu có tương quan với nhau và điều này là phù hợp cho phân tích PCA.

### Xác định thành phần chính

Để quyết định số lượng thành phần chính được giữ lại trong phân tích PCA, nghiên cứu sử dụng quy tắc Kaiser. Ý tưởng cơ bản của quy tắc Kaiser là giữ lại các thành phần chính có giá trị riêng (eigenvalue) lớn hơn 1 (nếu giá trị riêng lớn hơn 1 điều đó ngụ ý rằng thành phần đó giải thích được lượng thông tin lớn hơn so với một biến gốc riêng lẻ).

**Bảng 2: Giá trị riêng, phương sai giải thích của các thành phần chính**

ST	Thành phần	Độ lệch chuẩn	Eigenval	Phương sai giải thích	Phương sai giải thích tích lũy
1	Comp.1	2,926	8,562	0,435	0,435
2	Comp.2	1,521	2,315	0,118	0,553
3	Comp.3	1,318	1,737	0,088	0,641
4	Comp.4	1,247	1,555	0,079	0,720
5	Comp.5	1,037	1,074	0,055	0,774

Nguồn: Tác giả tính toán từ dữ liệu thu thập

Căn cứ kết quả Bảng 2, có 5 thành phần chính có giá trị riêng lớn hơn 1, đồng thời tỷ lệ phương sai giải thích của 5 thành phần đầu tiên chiếm 77,4% sự biến thiên của tập dữ liệu nghiên cứu. Trong đó, thành phần chính thứ 1 giải thích 43,5% thông tin dữ liệu, thành phần chính thứ 2 giải thích 11,8%, thành phần chính thứ 3 giải thích 8,8%, thành phần chính thứ 4 giải thích 7,9% và thành phần chính thứ 5 giải thích 5,5%.

### Hệ số tải và chỉ số tổng hợp

Để thể hiện đóng góp của các biến lên từng thành phần chính được giữ lại, nghiên cứu tiến hành ước lượng hệ số tải của các biến và giá trị của các thành phần chính được hình thành từ tổ hợp tuyến tính của các biến với hệ số tải được ước lượng.

Căn cứ vào điểm của các thành phần chính, chỉ số tổng hợp đo lường mức độ phát triển kinh tế hình thành từ trung bình có trọng số của 5 thành phần chính, trong đó trọng số là tỷ trọng của phương sai giải thích của từng thành phần chính so với tổng số phương sai của 5 thành phần chính được giữ

**Bảng 1: Chỉ tiêu đo lường phát triển kinh tế cấp tỉnh tại Việt Nam**

Đặc điểm	Ký hiệu	Nội dung	Đơn vị tính	Nguồn	Năm
Nhân khẩu học	NK1	Tỷ trọng dân số của tỉnh so với cả nước	%	TCTK	2016-2020
	NK2	Mật độ dân số	Người/km <sup>2</sup>	TCTK	2016-2020
	NK3	Tỷ suất nhập cư	‰	TCTK	2016-2020
	NK4	Tỷ lệ đô thị hóa	%	TCTK	2016-2020
Lao động và việc làm	VL1	Tỷ trọng lao động từ 15 tuổi trở lên của tỉnh so với cả nước	%	TCTK	2016-2020
	VL2	Tỷ trọng lao động khu vực công nghiệp xây dựng trong tỉnh	%	CTK	2016-2020
	VL3	Tỷ trọng lao động khu vực dịch vụ trong tỉnh	%	CTK	2016-2020
	VL4	Tỷ lệ lao động từ 15 tuổi trở lên đã qua đào tạo	%	TCTK	2016-2020
Nguồn lực	TC1	GRDP bình quân đầu người	Triệu đồng/người	TCTK	2016-2020
	TC2	Thu ngân sách bình quân đầu người	Triệu đồng/người	CTK	2016-2020
	TC3	Chỉ số giá không gian	Triệu đồng/người	CTK	2016-2020
	TC4	Năng suất lao động của tỉnh	Triệu đồng/người	CTK	2016-2020
Năng lực cạnh tranh	NL1	Số doanh nghiệp có kết quả sản xuất kinh doanh trên 10.000 dân	DN	TCTK	2016-2020
	NL2	Tổng mức bán lẻ hàng hóa và doanh thu dịch vụ tiêu dùng bình quân đầu người.	Triệu đồng/người	TCTK	2016-2020
	NL3	Giá trị xuất khẩu bình quân đầu người	USD/người	CTK	2016-2020
	NL4	Số lượng đơn bảo hộ sáng chế, giải pháp hữu ích, kiểu dáng công nghiệp và nhãn hiệu trên 10000 dân	Đơn	CỤC SHTT	2016-2020
Chất lượng cuộc sống	CLCS1	Tỷ lệ dân số từ 15 tuổi trở lên biết chữ	%	TCTK	2016-2020
	CLCS2	Tỷ trọng bác sĩ của tỉnh so với cả nước	%	TCTK	2016-2020
	CLCS3	Tỷ lệ trẻ em dưới 1 tuổi được tiêm đầy đủ các loại vaccine (%)	%	TCTK	2016-2020
	CLCS4	Tỷ lệ dân số đô thị được cung cấp nước sạch qua hệ thống tập trung (%)	%	TCTK	2016-2020

Nguồn: Tác giả tổng hợp

**Bảng 3: Chỉ số tổng hợp phát triển kinh tế cấp tỉnh tại Việt Nam**

STT	Tỉnh	Điểm số tổng hợp					Xếp hạng		2017	2018	2019	2020
		2016	2017	2018	2019	2020	2016	2017				
A		1	2	3	4	5	6	7	8	9	10	
1	Hà Nội	6,54	6,08	6,14	6,10	5,55	2	2	2	2	2	
2	Hà Giang	-1,81	-1,49	-1,64	-1,98	-1,84	62	62	62	61	61	
3	Cao Bằng	-1,40	-1,16	-1,34	-1,61	-1,50	57	54	54	59	59	
4	Bắc Kạn	-1,56	-1,53	-1,44	-1,48	-1,50	60	63	63	57	58	
5	Tuyên Quang	-1,21	-1,39	-1,19	-1,21	-1,22	54	61	61	53	54	
6	Lào Cai	-0,97	-0,69	-0,95	-1,01	-1,01	50	40	40	50	52	
7	Điện Biên	-1,67	-1,18	-1,59	-1,98	-1,86	61	55	55	62	62	
8	Lai Châu	-2,12	-1,32	-1,66	-2,07	-2,02	63	58	58	63	63	
9	Sơn La	-1,52	-1,33	-1,43	-1,71	-1,57	59	59	59	60	60	
10	Yên Bái	-1,28	-1,23	-1,14	-1,31	-1,22	56	57	57	54	53	
11	Hoà Bình	-0,88	-0,89	-0,93	-0,91	-0,87	46	49	49	48	48	
12	Thái Nguyên	0,32	0,32	0,27	0,67	0,41	15	13	13	12	16	
13	Lạng Sơn	-0,98	-0,85	-0,84	-1,02	-0,96	52	46	46	51	50	
14	Quảng Ninh	1,07	1,43	1,28	1,30	1,28	9	9	9	9	9	
15	Bắc Giang	-0,21	-0,05	0,08	0,30	0,43	24	20	20	18	15	
16	Phú Thọ	-0,35	-0,71	-0,53	-0,33	-0,34	32	41	41	31	32	
17	Vĩnh Phúc	0,73	0,55	0,53	0,80	0,77	11	11	11	11	11	
18	Bắc Ninh	1,73	1,68	2,08	2,45	2,60	8	7	7	5	4	
19	Hải Dương	0,48	0,20	0,65	0,63	0,65	13	15	15	13	12	
20	Hải Phòng	1,87	1,88	1,81	1,98	2,02	6	6	6	7	6	
21	Hưng Yên	0,25	0,07	0,19	0,42	0,51	17	18	18	15	13	
22	Thái Bình	0,08	-0,49	-0,08	-0,05	-0,09	19	34	34	21	21	
23	Hà Nam	-0,31	-0,53	-0,14	0,00	0,10	30	36	36	20	19	
24	Nam Định	-0,23	-0,59	-0,27	-0,08	-0,12	25	37	37	22	23	
25	Ninh Bình	-0,08	-0,07	0,11	0,14	0,06	20	21	21	19	20	

Continued on next page

Table 3 continued

26	Thanh Hoá	0,71	0,21	0,30	0,31	0,31	12	14	14	17	17
27	Nghệ An	0,10	-0,19	-0,22	-0,14	-0,10	18	23	23	24	22
28	Hà Tĩnh	-0,24	-0,16	0,08	-0,12	-0,22	27	22	22	23	27
29	Quảng Bình	-0,83	-0,65	-0,77	-0,85	-0,76	45	39	39	45	44
30	Quảng Trị	-0,56	-0,48	-0,51	-0,66	-0,61	39	33	33	39	39
31	Thừa Thiên Huế	-0,10	-0,04	-0,18	-0,21	-0,17	22	19	19	25	25
32	Đà Nẵng	2,47	2,99	2,49	2,67	2,40	4	3	3	4	5
33	Quảng Nam	-0,26	-0,20	-0,21	-0,24	-0,30	28	24	24	26	29
34	Quảng Ngãi	-0,38	-0,60	-0,58	-0,46	-0,32	34	38	38	35	31
35	Bình Định	-0,24	-0,35	-0,23	-0,24	-0,14	26	25	25	27	24
36	Phú Yên	-0,91	-0,99	-0,97	-0,92	-0,84	47	50	50	49	46
37	Khánh Hoà	0,42	0,33	0,29	0,42	0,20	14	12	12	16	18
38	Ninh Thuận	-0,97	-0,85	-0,80	-0,85	-0,89	49	47	47	46	49
39	Bình Thuận	-0,49	-0,39	-0,57	-0,44	-0,44	35	29	29	34	35
40	Kon Tum	-1,24	-1,06	-1,17	-1,50	-1,31	55	52	52	58	55
41	Gia Lai	-1,15	-1,08	-1,28	-1,36	-1,33	53	53	53	55	56
42	Đắk Lắk	-0,54	-0,50	-0,64	-0,81	-0,77	37	35	35	43	45
43	Đắk Nông	-1,48	-1,38	-1,34	-1,39	-1,40	58	60	60	56	57
44	Lâm Đồng	-0,51	-0,36	-0,52	-0,57	-0,58	36	26	26	37	37
45	Bình Phước	-0,70	-0,41	-0,66	-0,65	-0,60	41	30	30	38	38
46	Tây Ninh	-0,37	-0,37	-0,39	-0,34	-0,24	33	27	27	32	28
47	Bình Dương	2,50	2,74	3,33	3,24	3,22	3	4	4	3	3
48	Đồng Nai	1,74	1,56	1,23	1,60	1,55	7	8	8	8	8
49	Bà Rịa - Vũng Tàu	2,33	2,71	1,97	2,20	1,83	5	5	5	6	7
50	TP Hồ Chí Minh	9,45	8,75	8,25	8,44	8,39	1	1	1	1	1
51	Long An	0,26	0,16	0,34	0,46	0,44	16	16	16	14	14
52	Tiền Giang	-0,10	-0,47	-0,32	-0,27	-0,32	21	32	32	29	30
53	Bến Tre	-0,63	-0,77	-0,73	-0,71	-0,72	40	43	43	41	41
54	Trà Vinh	-0,97	-1,00	-1,01	-1,02	-0,99	48	51	51	52	51

Continued on next page



Table 3 continued

55	Vĩnh Long	-0,55	-0,72	-0,57	-0,50	-0,53	38	42	42	36	36
56	Đồng Tháp	-0,29	-0,38	-0,31	-0,38	-0,35	29	28	28	33	33
57	An Giang	-0,13	0,11	-0,05	-0,25	-0,21	23	17	17	28	26
58	Kiên Giang	-0,35	-0,42	-0,21	-0,33	-0,35	31	31	31	30	34
59	Cần Thơ	0,83	1,30	1,12	1,07	0,94	10	10	10	10	10
60	Hậu Giang	-0,97	-1,21	-0,94	-0,90	-0,85	51	56	56	47	47
61	Sóc Trăng	-0,82	-0,89	-0,73	-0,80	-0,70	44	48	48	42	40
62	Bạc Liêu	-0,80	-0,79	-0,78	-0,84	-0,73	43	44	44	44	42
63	Cà Mau	-0,71	-0,84	-0,68	-0,69	-0,75	42	45	45	40	43

Nguồn: Tác giả tính toán từ dữ liệu thu thập.

lại (Trong số các thành phần chính được tính toán:  
Thành phần chính 1: 0,56; thành phần chính số 2:  
0,15; thành phần chính số 3: 0,11, thành phần chính  
số 4: 0,10; thành phần chính số 5: 0,07; tổng của 5  
thành phần bằng 1). Cụ thể tại Bảng 3.

Kết quả chỉ số tổng hợp đo lường mức độ phát triển kinh tế của các tỉnh năm 2020 được thể hiện trong Bảng 3 (cột 5) cho thấy: Trong 63 tỉnh, thành tại Việt Nam thì Thành phố Hồ Chí Minh, Thành phố Hà Nội có mức độ phát triển kinh tế cao nhất, tiếp đến tỉnh Bình Dương được xem là trung tâm công nghiệp của cả nước; ngược lại các tỉnh Hà Giang, Sơn La, Lai Châu có mức độ phát triển kinh tế thấp nhất và đây đều là những tỉnh thuộc vùng Trung du và miền núi phía Bắc còn nhiều khó khăn trong phát triển kinh tế. Để có thông tin đánh giá sự thay đổi về phát triển kinh tế trong nhiệm kỳ thực hiện nghị quyết Đảng bộ của cấp tỉnh tại Việt Nam (nhiệm kỳ 2016-2020). Nghiên cứu thực hiện tính toán bổ sung năm 2016- 2019, trên cơ sở phương pháp luận, số lượng chỉ tiêu, quy trình thực hiện như năm 2020 và thu được kết quả được thể hiện tại Bảng 3 (cột 1-4). Kết quả trong 5 năm qua thì Thành phố Hồ Chí Minh, Thành phố Hà Nội vẫn khẳng định được vai trò là đầu tàu kinh tế cả nước, riêng Bắc Ninh là tỉnh có mức phát triển nhanh từ vị trí thứ 8 (năm 2016) lên vị trí thứ 4 (năm 2020). Ngược lại, các tỉnh Hà Giang, Điện Biên, Lai Châu là các tỉnh thuộc vùng Trung du và miền núi phía Bắc có mức độ phát triển thấp nhất cả nước.

Giữa các vùng kinh tế - xã hội của Việt Nam có mức độ phát triển kinh tế khá chênh lệch. Trong đó, vùng Đông Nam Bộ vẫn được xem là trung tâm kinh tế của cả nước có mức độ phát triển kinh tế ở mức cao nhất (Hình 1); Vùng Đông Bằng Sông Hồng với trung tâm là Thủ đô Hà Nội, đây được xác định là vùng có ý nghĩa chiến lược quan trọng về chính trị, văn hóa và có mức độ phát triển kinh tế chỉ sau vùng Đông Nam Bộ; tiếp theo là vùng Trung bộ và Duyên hải Miền Trung, đây là vùng chiến lược trong phát triển kinh tế biển, có nhiều điều kiện thuận lợi phát triển du lịch; Vùng Đông Bằng Sông Cửu Long là trung tâm nông nghiệp cả nước nhưng ngày càng có tốc độ tăng trưởng chậm lại so với đà tăng trưởng các vùng khác và khi chuyển dịch cơ cấu kinh tế chậm, mô hình nông nghiệp không còn nhiều không gian tăng trưởng dẫn đến mức độ phát triển kinh tế hiện khá thấp; Vùng Tây Nguyên với tiềm năng lớn về chuyên canh cây công nghiệp, ngày càng phát triển lĩnh vực năng lượng, khai khoáng, du lịch, tuy nhiên hạ tầng kết nối giao thông còn hạn chế đã ảnh hưởng đến mức độ phát triển kinh tế của vùng; Vùng Trung du và miền núi phía Bắc, là vùng có tiềm năng phát triển thủy điện, du lịch với thiên nhiên phong phú, đa dạng khí hậu, đa dạng sắc tộc. Tuy nhiên, vùng được xem là vùng trũng trong phát triển kinh tế khi năng suất lao động thấp và hạ tầng giao thông còn nhiều khó khăn.

## Các yếu tố tác động đến mức độ phát triển kinh tế cấp tỉnh Việt Nam

### Chọn biến đại diện cho nhóm yếu tố

Hàm sản xuất Cobb-Douglass được sử dụng rộng rãi, phổ biến trong việc phân tích tăng trưởng và năng suất của nền kinh tế. Trong kinh tế học, hàm sản xuất được thể hiện như sau:  $Y = A L^\alpha K^\beta$

Trong đó: Y là sản lượng; L là yếu tố lao động; K là yếu tố vốn; A là năng suất nhân tố tổng hợp;  $\alpha, \beta$  lần lượt là hệ số co giãn đầu ra của lao động và vốn.

Trên cơ sở nghiên cứu thực nghiệm về tác động đến tăng trưởng kinh tế, dựa trên tính sẵn có của dữ liệu và tham vấn ý kiến chuyên gia, nghiên cứu hình thành các nhóm yếu tố tác động đến mức độ phát triển kinh tế cấp tỉnh tại Việt Nam (Bảng 4).

Để tránh hiện tượng đa cộng tuyến, nghiên cứu chọn biến đại diện cho mỗi nhóm yếu tố thông qua giá trị Shapley. Hồi quy giá trị Shapley là kỹ thuật xác định tầm quan trọng tương đối của các biến dựa vào dự đoán trong hồi quy tuyến tính<sup>19</sup>. Đầu tiên, hồi quy tuyến tính với tất cả các kết hợp của biến độc lập và thu được giá trị  $R^2$ . Nếu có 3 biến độc lập (A, B, C) thì các mô hình hồi quy tuyến tính với biến độc lập là (A), (B), (C), (A, B), (A, C), (B, C) và (A, B, C). Sau đó, giá trị Shapley của mỗi biến độc lập được tính bằng trung bình có trọng số của chênh lệch  $R^2$  khi biến này được thêm vào từng tập con. Trong đó, gán trọng số bằng nhau cho mô hình có một biến độc lập, 2 biến độc lập và 3 biến độc lập. Mô phỏng giá trị Shapley của biến A được tính:

$$\begin{aligned} Shapley_A &= \frac{1}{3} (R_A^2 - R_0^2) \\ &+ \frac{1}{3} \left( \frac{(R_{AB}^2 - R_B^2) + (R_{AC}^2 - R_C^2)}{2} \right) \\ &+ \frac{1}{3} (R_{ABC}^2 - R_{BC}^2) \end{aligned}$$

Với sự hỗ trợ phần mềm các giá trị Shapley được tính toán nhanh chóng. Căn cứ trên tỷ lệ phần trăm dự đoán làm cơ sở kết luận chọn biến tỷ trọng dân số (Pop\_ratio) đại diện cho yếu tố lao động, số lượng doanh nghiệp (lnCompany) đại diện cho nhóm yếu tố vốn và chỉ số phát triển con người cấp tỉnh (HDI) đại diện cho nhóm yếu tố khác.

Mô hình hóa 3 yếu tố tác động đến mức độ phát triển kinh tế cấp tỉnh (CI) có dạng:

$$\begin{aligned} CI &= \beta_0 + \beta_1 Pop\_ratio \\ &+ \beta_2 \ln Company + \beta_3 HDI + u \end{aligned}$$

Trong đó: CI là chỉ số đo lường mức độ phát triển kinh tế cấp tỉnh của Việt Nam được tính toán cho kết quả tại Bảng 3.

**Bảng 4: Các yếu tố ảnh hưởng đến phát triển kinh tế cấp địa phương**

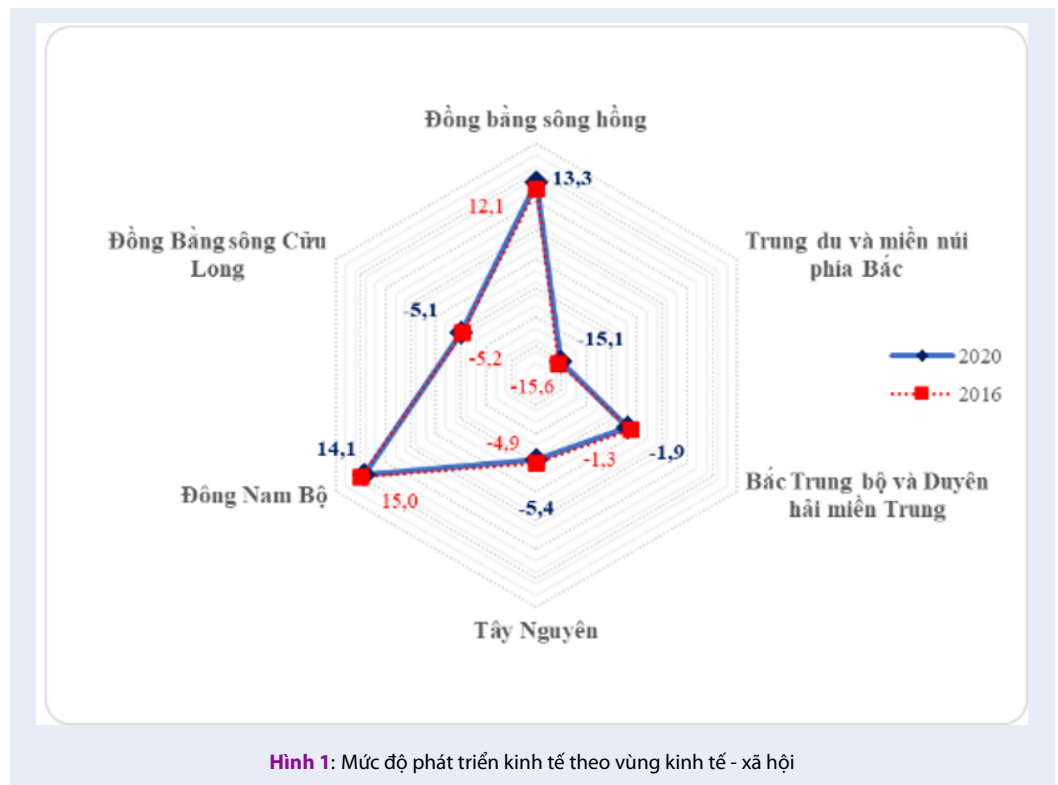
Yếu tố	Tên biến	Nội dung biến	Giá trị Shapley	Phần trăm (%)
Lao động	lnPopulation	Tổng dân số	0,1279	13,88
	lnUrban	Dân số thành thị	0,1654	17,95
	lnRural	Dân số nông thôn	0,0498	5,40
	Pop_ratio	Tỷ trọng dân số	0,1990	21,59
	Urban_rate	Tỷ lệ đô thị hóa	0,0978	10,61
	Birth_rate	Tỷ suất sinh	0,0378	4,10
	lnlabor	Quy mô lao động	0,0798	8,65
	labor_level	Tỷ lệ lao động được đào tạo	0,1642	17,82
	Total			100,00
Vốn	lnBudget	Thu ngân sách	0,1346	16,02
	lnExpen	Chi ngân sách	0,0877	10,44
	lnCapital	Vốn đầu tư thực hiện	0,1266	15,08
	lnCompany	Số lượng doanh nghiệp	0,2215	26,37
	lnRetail	Tổng mức bán lẻ và doanh thu dịch vụ	0,1072	12,76
	lnTravel	Doanh thu du lịch	0,0833	9,92
	lnIm_Ex	Kim ngạch xuất nhập khẩu	0,0790	9,41
	Total			100,00
Yếu tố khác	PCI	Chỉ số năng lực cạnh tranh cấp tỉnh	0,0530	8,95
	HDI	Chỉ số phát triển con người cấp tỉnh	0,2336	39,47
	Scoli	Chỉ số sinh hoạt không gian	0,1260	21,30
	LnProduction	Năng suất lao động	0,1791	30,28
	Total			100,00

Nguồn: Tác giả tính toán từ dữ liệu thu thập.

**Bảng 5: Ma trận hệ số tương quan các biến**

	CI	Pop_ratio	LnCompany	HDI
CI	1			
Pop_ratio	0,8372***	1		
LnCompany	0,9064***	0,8256***	1	
HDI	0,6871***	0,4499***	0,6902***	1

Nguồn: Tác giả tính toán từ dữ liệu thu thập



Hình 1: Mức độ phát triển kinh tế theo vùng kinh tế - xã hội

### Ma trận tương quan các biến

Hệ số tương quan giữa các biến cho kết quả tại Bảng 5. Kết quả sig của tương quan Pearson (Bảng 5) giữa các biến tỷ trọng dân số, số lượng doanh nghiệp, HDI với biến mức độ phát triển kinh tế đều nhỏ hơn 1%. Cho thấy, các biến độc lập có mối tương quan đến biến phụ thuộc và mức độ tương quan này là khá mạnh ( $>0,6$ ). Ngoài ra, tương quan giữa các biến độc lập có giá trị sig đều nhỏ hơn 1% cho thấy giữa các biến tương quan với nhau, do vậy vấn đề đa cộng tuyến hoàn hảo trong mô hình cần được kiểm tra.

### Ước lượng mô hình hồi quy chưa xét đến tương quan không gian

Nhằm kiểm tra tính ổn định của ước lượng, nghiên cứu tiến hành ước lượng hồi quy OLS với dữ liệu bảng bằng các mô hình POLS, mô hình tác động cố định (FEM) và mô hình tác động ngẫu nhiên (REM) được thể hiện trong Bảng 6. Nghiên cứu thực hiện tuân tự các kiểm định từ đó đủ cơ sở kết luận rằng mô hình FEM là phù hợp hơn so với mô hình POLS và REM. Mặc dù mô hình FEM không xảy ra hiện tượng đa cộng tuyến hoàn hảo nhưng xảy ra hiện tượng tự tương quan và phương sai thay đổi dẫn đến ước lượng không hiệu quả, do vậy sử dụng mô hình GLS nhằm khắc phục các hiện tượng trên. Kết quả ước lượng

GLS cho thấy 03 yếu tố đều tác động đến mức độ phát triển kinh tế của cấp tỉnh tại Việt Nam.

Nhược điểm của ước lượng này là chưa xét đến sự liên hệ không gian giữa các tỉnh thành của Việt Nam, nếu tồn tại về tương quan không gian giữa các tỉnh thành thì ước lượng này sẽ không hiệu quả và không vững. Do vậy, cần tiến hành kiểm định sự tồn tại tương quan không gian.

### Kiểm định tương quan không gian

Các nghiên cứu trước đây của tác giả đã kiểm chứng rằng đối với dữ liệu cấp tỉnh, thành tại Việt Nam thì ma trận trọng số đường bộ được xem là phù hợp hơn so với ma trận trọng số liên kế. Kết quả kiểm định Moran's I bằng ma trận trọng số đường bộ (Bảng 7) cho thấy, mức độ phát triển kinh tế (CI), số lượng doanh nghiệp (lnCompany) thu được giá trị Moran's I đều có ý nghĩa thống kê ở mức 1%, riêng HDI có tương quan không gian ở một số năm nhất định và đối với tỷ trọng dân số không có ý nghĩa thống kê ở mức 5%. Trên cơ sở đó khẳng định về sự tồn tại tương quan không gian giữa các địa phương tại Việt Nam.

### Mô hình hồi quy không gian

Trong nghiên cứu thực nghiệm có 03 mô hình hồi quy thường được sử dụng: Mô hình sai số không gian

**Bảng 6: Ước lượng mô hình chưa xét tương quan không gian**

Biến phụ thuộc: CI	POLS	FEM	REM	GLS	VIF
Biến độc lập	1	2	3	4	5
Pop_ratio	0,414***	-0,698**	0,687***	0,601***	5,24
LnCompany	0,777***	0,409**	0,448***	0,473***	3,44
HDI	6,717***	-3,806**	-0,605	3,915***	2,09
Hệ số chặn	-11,61***	0,334	-4,366***	-7,531***	
Số quan sát	315	315	315	315	
Kiểm định so sánh với POLS		63,12***	454***		
Kiểm định Hausman			587,29***		
Kiểm định tự tương quan	16,831***				
Kiểm định phương sai thay đổi	13870,35***	16135,45***	482,51***		
Nguồn: Tác giả tính toán từ dữ liệu thu thập					
*, **, *** tương ứng mức ý nghĩa 10%, 5%, 1%.					

**Bảng 7: Hệ số Moran'I với ma trận trọng số đường bộ**

Năm	CI		Pop_ratio		LnCompany		HDI	
	Hệ số Moran' I	P-value	Hệ số Moran' I	P-value	Hệ số Moran' I	P-value	Hệ số Moran' I	P-value
2016	0,197	0,003	0,035	0,244	0,187	0,008	0,136	0,041
2017	0,218	0,001	0,038	0,234	0,183	0,009	0,113	0,069
2018	0,207	0,002	0,040	0,223	0,189	0,008	0,109	0,076
2019	0,225	0,001	0,043	0,212	0,190	0,008	0,123	0,055
2020	0,236	0,001	0,046	0,199	0,203	0,005	0,132	0,045

Nguồn: Tác giả tính toán từ dữ liệu thu thập.

(SEM), mô hình tự hồi quy không gian (SAR) và mô hình Durbin không gian (SDM). Để lựa chọn mô hình không gian phù hợp nên bắt đầu từ mô hình Durbin không gian<sup>20</sup> và từ SDM có thể chuyển thành mô hình SAR hay SEM. Trong đó, mô hình Durbin không gian (SDM) với các biến nghiên cứu có dạng:

$$CI = \beta_0 + \lambda CI + \beta_1 Pop\_ratio + \beta_2 LnCompany + \beta_3 HDI + \delta_1 Pop\_ratio + \delta_2 LnCompany + \delta_3 HDI + u$$

Kết quả ước lượng mô hình SEM, SAR và SDM với ma trận đường bộ tại Bảng 8.

Nếu  $\theta=0$  thì mô hình SDM chuyển thành SAR và kết quả kiểm định tại Bảng 8 cung cấp giá trị thống kê Chi2(3)=13,23 với P-value=0,0042, nghĩa là mô hình SDM phù hợp hơn mô hình SAR.

Nếu  $\theta+\rho\beta=0$  thì mô hình SDM chuyển thành SEM và kết quả kiểm định cung cấp giá trị thống kê Chi2(3)=9,08 với P-value=0,0283, nghĩa là mô hình

SDM phù hợp hơn mô hình SEM. Từ các kiểm định trên cho thấy mô hình SDM là phù hợp với dữ liệu nghiên cứu. Mặt khác kiểm định Hausman cung cấp giá trị Chi2(3)=10,16 với P-value=0,1796 cho thấy mô hình SDM-REM là phù hợp hơn để phân tích dữ liệu. Trên cơ sở mô hình SDM-REM được chọn, nghiên cứu tiến hành phân tích tác động của các yếu tố đến mức độ phát triển kinh tế các tỉnh tại Việt Nam (Bảng 9).

## THẢO LUẬN KẾT QUẢ

Quá trình xây dựng chỉ số tổng hợp mức độ phát triển kinh tế cấp tỉnh tại Việt Nam với 19 chỉ tiêu được sử dụng đã đáp ứng tương quan giữa các biến và việc áp dụng phân tích thành phần chính PCA là phù hợp với dữ liệu nghiên cứu. Trong đó, có 5 thành phần chính được giữ lại đã giải thích 77,4% sự biến thiên của tập dữ liệu gốc. Điểm của các thành phần chính được tổng hợp dựa trên đóng góp của tất cả các biến và làm

**Bảng 8: Kết quả ước lượng với ma trận đường bộ**

Biến phụ thuộc: CI	SEM		SAR		SDM	
	FEM	REM	FEM	REM	FEM	REM
Biến độc lập	1	2	3	4	5	6
Pop_ratio	-0,705***	0,525***	-0,706***	0,689***	-0,570**	0,414***
LnCompany	0,401***	0,597***	0,415***	0,417***	0,646***	0,779***
HDI	-3,859**	4,432***	-3,802**	-1,231	3,263	6,777***
Hệ số chặn		-8,753***		-3,696***		-3,664***
Spatial						
Lambda	-0,119	1,420***				
rho			-0,194	-0,163	-0,126	-0,19
Variance						
sigma2_e	0,0236***	0,0337***	0,0235***	0,0325***	0,0226***	0,0297***
ln_phi		2,435***				
lgt_theta				-2,195***		-1,868***
Spatial Lag						
Pop_ratio					-1,024	1,697**
LnCompany					-0,337	-0,846**
HDI					-6,722	-5,431
Số quan sát	315	315	315	315	315	315
Hausman		73,31***		26,57***		10,16
Test SDM_FEM và SAR_FEM					13,23***	
Test SDM_FEM và SEM_FEM					9,08***	
Nguồn: Tác giả tính toán từ dữ liệu thu thập						
*, **, *** tương ứng mức ý nghĩa 10%, 5%, 1%.						

cơ sở để hình thành chỉ số tổng hợp với trọng số là phương sai giải thích của từng thành phần chính. Mức độ phát triển kinh tế cấp tỉnh tại Việt Nam trong nhiệm kỳ 2016-2020 cho thấy không có sự thay đổi đáng kể giữa các tỉnh có mức độ phát triển cao nhất hoặc thấp nhất cả nước, trong đó Thành phố Hồ Chí Minh, Thành phố Hà Nội, tỉnh Bình Dương vẫn là 3 đơn vị có mức độ phát triển cao nhất, ngược lại các tỉnh Hà Giang, Điện Biên, Lai Châu là các tỉnh thuộc vùng Trung du và miền núi phía Bắc nơi có điều kiện phát triển kinh tế còn nhiều khó khăn. Tuy nhiên, trong thời gian nghiên cứu tỉnh Bắc Ninh, Bắc Giang là hai địa phương “thay da đổi thịt” trong phát triển kinh tế nhờ thu hút đầu tư của các tập đoàn công nghiệp, điện tử lớn trên thế giới. Vì thế, Bắc Ninh

có mức độ phát triển kinh tế đứng thứ 8/63 tỉnh vào năm 2016 thì đến năm 2020 đã vươn lên vị trí thứ 4/63 tỉnh. Tương tự, Bắc Giang là tỉnh có mức độ phát triển kinh tế khiêm tốn khi năm 2016 đứng vị trí 24/63 thì đến năm 2020 đã cải thiện đáng kể khi đứng vị trí thứ 14/63 tỉnh.

Ngoài ra, giữa các vùng kinh tế - xã hội của Việt Nam có mức độ phát triển kinh tế khá chênh lệch và có thể phân thành 3 mức độ: Nhóm 1 gồm vùng Đông Nam Bộ và Đồng Bằng Sông Hồng là những vùng có mức độ phát triển cao của cả nước; nhóm 2 gồm vùng Trung bộ và Duyên hải Miền Trung, Vùng Đồng Bằng Sông Cửu Long, Vùng Tây Nguyên là những vùng có mức độ phát triển ở mức trung bình; riêng nhóm 3 là Vùng Trung du và miền núi phía Bắc có mức độ phát

**Bảng 9: Tác động trực tiếp, gián tiếp và tổng tác động mô hình SDM-REM**

	Hệ số (Coef)	Sai số chuẩn (Std.Err)
Biến giải thích		
Pop_ratio	0,414***	0,076
LnCompany	0,779***	0,110
HDI	6,777***	1,538
Biến trễ không gian		
W.Pop_ratio	1,696**	0,663
W.LnCompany	-0,846**	0,412
W.HDI	-5,430	4,874
Tổng tác động		
Pop_ratio	1,896***	0,544
LnCompany	-0,059	0,374
HDI	1,199	4,442
Tác động trực tiếp		
Pop_ratio	0,409***	0,077
LnCompany	0,777***	0,111
HDI	6,985***	1,485
Tác động gián tiếp		
Pop_ratio	1,486***	0,536
LnCompany	-0,836**	0,383
HDI	5,786	4,625
Nguồn: Tác giả tính toán từ dữ liệu thu thập		
*, **, *** tương ứng mức ý nghĩa 10%, 5%, 1%.		

triển kinh tế thấp so với vùng còn lại.

Kết quả phân tích hồi quy không gian cho thấy tác động trực tiếp các yếu tố nội tại đến mức độ phát triển kinh tế của chính địa phương đó. Trong đó, tỷ trọng dân số, số lượng doanh nghiệp và chỉ số phát triển con người có tác động tích cực đến mức độ phát triển kinh tế của chính địa phương sở tại. Tác động trực tiếp có chênh lệch so với hệ số ước lượng từ chính mô hình là do hiệu ứng phản hồi (feedback) của chính các biến nghiên cứu với biến trễ không gian của chính nó.

Bên cạnh đó, tác động gián tiếp cho thấy hiệu ứng lan tỏa không gian của các yếu tố đến các địa phương có khoảng cách gần nhau. Trong đó, tỷ trọng dân số của tỉnh có tác động tích cực đến mức độ phát triển của các tỉnh gần nhau. Tuy nhiên, số lượng doanh nghiệp, HDI chưa đủ bằng chứng cho thấy có tác động không gian đến mức độ phát triển kinh tế của các tỉnh gần nhau do hệ số tác động tổng hợp không có ý nghĩa thống kê.

Từ đó tổng tác động trực tiếp và gián tiếp (tổng tác động) đã giải thích sự thay đổi một yếu tố của tỉnh sẽ tác động trực tiếp đến mức độ phát triển của chính tỉnh đó mà còn tác động đến mức độ phát triển của các tỉnh có khoảng cách gần nhau.

## KẾT LUẬN VÀ HÀM Ý NGHIÊN CỨU

Xây dựng chỉ số tổng hợp đo lường mức độ phát triển kinh tế cấp tỉnh được xem là bổ sung công cụ thống kê nhằm đánh giá những chuyển biến trong phát triển kinh tế cấp tỉnh trong thời gian 5 năm (nhiệm kỳ 2016-2020) không chỉ dựa vào chỉ tiêu duy nhất là tốc độ tăng trưởng GRDP. Kết quả chỉ số tổng hợp đã phản ánh được mức độ phát triển kinh tế của 63 tỉnh, thành tại Việt Nam và sự chênh lệch về mức độ phát triển kinh tế giữa 6 vùng kinh tế - xã hội của Việt Nam từ đó làm cơ sở xây dựng chính sách, giải pháp nhằm hỗ trợ các vùng khó khăn nhằm thu hẹp khoảng

cách phát triển giữa các vùng, các địa phương tại Việt Nam.

Nghiên cứu đã cung cấp đủ bằng chứng thống kê về tồn tại tương quan không gian giữa các tỉnh, thành của Việt Nam, do đó mô hình hồi quy không gian được sử dụng để ước lượng thay thế cho mô hình hồi quy dữ liệu bảng thông thường. Kết quả ước lượng cho thấy: Tỷ trọng dân số, doanh nghiệp, HDI có tác động tích cực đến mức độ phát triển kinh tế của chính địa phương sở tại, ngoài ra tỷ trọng dân số còn tác động tích cực đến mức độ phát triển kinh tế các địa phương lân cận. Qua đó góp phần hàm ý chính sách quan trọng trong việc xây dựng quy hoạch phát triển kinh tế - xã hội cấp tỉnh phải đặt trong bối cảnh chung của quy hoạch phát triển kinh tế - xã hội cấp vùng khi mà nguồn lực của từng địa phương có giới hạn.

Tuy nhiên, nghiên cứu hạn chế khi dựa vào các dữ liệu sẵn có nên chỉ dừng lại các biến đo lường cơ bản, vì vậy định hướng nghiên cứu tiếp theo sẽ mở rộng nhiều khía cạnh về xã hội, môi trường và năng lực quản trị địa phương.

## DANH MỤC CÁC TỪ VIẾT TẮT

FEM: Fixed Effect Model  
 GDP: Gross Domestic Product  
 GLS: Generalized Least Squares  
 GRDP: Gross Regional Domestic Product  
 OLS: Ordinary Least Squares  
 POLS: Pooled Ordinary Least Squared  
 REM: Random Effect Model  
 SAR: Spatial Autoregressive Model  
 SDM: Spatial Dubin Model  
 SEM: Spatial Error Model  
 VIF: Variance Inflation Factor

## XUNG ĐỘT LỢI ÍCH

Tác giả xin cam đoan rằng không có bất kỳ xung đột lợi ích nào trong công bố bài báo này.

## ĐÓNG GÓP CỦA TÁC GIẢ

Tác giả xin chịu toàn bộ trách nhiệm đối với toàn bộ nội dung bài báo.

## TÀI LIỆU THAM KHẢO

- Barnat N, MacFeely S, Cantu F, Peltola A, Khazhgerieva A, Pantelev A, et al. Compiling an inclusive growth index: Methodological challenges, considerations and conclusions. 2023;39(3):699-719; Available from: <https://doi.org/10.3233/SJI-230025>.
- Išlamović S, Jeremić V, Petrović N, Radojičić Z. Colouring the socio-economic development into green: I-distance framework for countries' welfare evaluation. 2015;49:617-29; Available from: <https://doi.org/10.1007/s11135-014-0012-0>.

- You Z, Shi H, Feng Z, Yang Y. Creation and validation of a socioeconomic development index: A case study on the countries in the Belt and Road Initiative. 2020;258:120634; Available from: <https://doi.org/10.1016/j.jclepro.2020.120634>.
- Mendola D, Scuderì R, Lacagnina V. Defining and measuring the development of a country over time: a proposal of a new index. 2013;47:2473-94; Available from: <https://doi.org/10.1007/s11135-012-9665-8>.
- Goletsis Y, Chletsos M. Measurement of development and regional disparities in Greek periphery: A multivariate approach. 2011;45(4):174-83; Available from: <https://doi.org/10.1016/j.seps.2011.06.002>.
- Silva R, Ferreira-Lopes A. A regional development index for Portugal. 2014;18:1055-85; Available from: <https://doi.org/10.1007/s11205-013-0455-z>.
- Podgorna I, Babenko V, Honcharenko N, Sáez-Fernández FJ, Fernández JAS, Yakubovskiy S. Modelling and analysis of socio-economic development of the European Union countries through DP2 method. 2020;17(44):454-66; Available from: <https://doi.org/10.37394/23207.2020.17.44>.
- OECD. Handbook on constructing composite indicators: methodology and user guide: Joint Research Centre-European Commission; 2008 2008 Aug 22;.
- Salvati L, Carlucci M. A composite index of sustainable development at the local scale: Italy as a case study. 2014;43:162-71; Available from: <https://doi.org/10.1016/j.ecolind.2014.02.021>.
- Mazziotta M, Pareto A. Methods for constructing composite indices: One for all or all for one. 2013;67(2):67-80;.
- Babiarz P, Grabiński T, Migala-Warchoł A, Szczygieł E. The application of customized human development index to the analysis of socio-economic development of the European Union member states. 2018;11(4):332; Available from: <https://doi.org/10.14254/2071-789X.2018/11-4/22>.
- Nuralina K, Baizholova R, Aleksandrova N, Konstantinov V, Biryukov A. Socio-economic development of countries based on the Composite country Development Index (CCDI). 2023;4(2):115-28; Available from: <https://doi.org/10.1016/j.regsus.2023.03.005>.
- Goschin Z. Territorial inequalities and economic growth in Romania. A multi-factor approach. 2015;22:690-8; Available from: [https://doi.org/10.1016/S2212-5671\(15\)00285-3](https://doi.org/10.1016/S2212-5671(15)00285-3).
- Nagy JA, Benedek J, Ivan K. Measuring sustainable development goals at a local level: A case of a metropolitan area in Romania. 2018;10(11):3962; Available from: <https://doi.org/10.3390/su10113962>.
- Karmakar S, Barman B, Roy R. Block-Wise Disparities in Socio-Economic Condition of Koch Bihar District, West Bengal. 2019;10; Available from: <https://doi.org/10.46587/JGR.2019.v06i01.019>.
- Thu NM. Nghiên cứu thống kê đánh giá phát triển bền vững ở Việt Nam 2013;.
- Huân ĐV. Xây dựng chỉ tiêu tổng hợp đo lường chất lượng tăng trưởng kinh tế Việt Nam 2018;.
- Moreira LL, de Brito MM, Kobiyama M. A systematic review and future prospects of flood vulnerability indices. 2021;21(5):1513-30; Available from: <https://doi.org/10.5194/nhess-21-1513-2021>.
- Tim Bock. What is Shapley Value Regression?[online] 2024; Available from: <https://www.displayr.com/shapley-value-regression/>.
- LeSage J, Pace RK. Introduction to spatial econometrics: Chapman and Hall/CRC; 2009; Available from: <https://doi.org/10.1201/9781420064254>.



# Composite index assessing the level of economic development at the provincial level in Vietnam

Nguyen Van Thang\*

## ABSTRACT

At the provincial and city levels in Vietnam, many individual statistical indicators reflect the socio-economic situation, making it difficult to capture and synthesize general information. The economic development of a province is often understood as the growth of the Gross Regional Domestic Product (GRDP). Still, it needs to demonstrate the multidimensionality of the concept of economic development. Therefore, building a comprehensive index to measure the level of economic development at the provincial level is very necessary, contributing to supplementing statistical tools to evaluate the economic development of the province in Vietnam. The composite index is formed from 19 indicators of 5 components: Demographic characteristics, labor-employment, resources, competitiveness, and quality of life. Data source is collected mainly from data from the Statistics for the period 2016-2020. Data are standardized by Z-score with weights estimated by the Principal Component Analysis (PCA) and composite index calculated from the weighted average of 5 main components with weight being the explained variance of the components. The results of the composite index show that Ho Chi Minh City and Hanoi are the economic centers of the country with the highest level of development. In contrast, the Northern midlands and mountain areas provinces have the lowest level of development. In addition, the study shows that a clear difference in the level of economic growth between the six regions in Vietnam. Besides, there is statistical evidence of spatial correlation between provinces in Vietnam, so the spatial Durbin model is used to replace the panel data OLS model. Spatial regression analysis shows that the proportion of population, businesses, and HDI have an impact on the level of economic development of the host locality, in addition, the proportion of the population also has a positive effect on the level of development economic development of neighboring localities.

**Key words:** Composite index, economic development, spatial regression

Ho Chi Minh City Statistics Office, Viet Nam

## Correspondence

**Nguyen Van Thang**, Ho Chi Minh City Statistics Office, Viet Nam

Email: thangnguyen.923102060056@st.uvh.edu.vn

## History

- Received: 08-5-2024
- Revised: 12-7-2024
- Accepted: 21-8-2024
- Published Online: 30-9-2024

DOI : <https://doi.org/10.32508/stdjelm.v8i3.1409>



## Copyright

© VNUHCM Press. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



Cite this article : Thang N V. **Composite index assessing the level of economic development at the provincial level in Vietnam.** *Sci. Tech. Dev. J. - Eco. Law Manag.* 2024, 8(3):5463-5479.