

ỨNG DỤNG PHƯƠNG PHÁP PHÂN TÍCH THÀNH PHẦN CHÍNH TRONG ĐÁNH GIÁ TỔNG HỢP TRÌNH ĐỘ PHÁT TRIỂN KINH TẾ CỦA CÁC TỈNH THÀNH VIỆT NAM

Hồ Thị Linh Trang

Tóm tắt

Dựa trên ý tưởng ứng dụng của phân tích thành phần chính (PCA) vào bài toán đánh giá năng lực tổng hợp, và nhận thấy việc nhìn nhận trình độ phát triển kinh tế của địa phương theo một chỉ tiêu đơn có những hạn chế nhất định, bài viết này sử dụng phương pháp PCA dựa trên bộ dữ liệu gồm 11 chỉ tiêu kinh tế của 63 tỉnh, thành phố Việt Nam tiến hành đánh giá, xếp hạng và so sánh trình độ phát triển kinh tế tổng hợp của các địa phương. Kết quả cho thấy có sự khác biệt rõ ràng trong trình độ phát triển của các tỉnh, thành. Đa phần các tỉnh thuộc vùng Đồng bằng sông Hồng và Đông Nam Bộ có xếp hạng cao hơn. Và các tỉnh lạc hậu chủ yếu thuộc khu vực Trung du và miền núi phía Bắc.

Tóm tắt: Phân tích thành phần chính, trình độ phát triển kinh tế, các tỉnh Việt Nam.

APPLYING THE PRINCIPAL COMPONENT ANALYSIS METHOD TO THE INTEGRATED ASSESSMENT OF THE ECONOMIC DEVELOPMENT LEVEL OF THE PROVINCES IN VIETNAM

Abstract

Based on the application concept of Principal Component Analysis (PCA) to general capacity assessment, and in recognition of some limitations in evaluating local economic development levels according to a single indicator, this article uses PCA method which is based on the data of 11 economic indicators of 63 provinces in Vietnam to conduct reviews, ratings and compare the development level of the local economy. The results show that there is a clear difference in economic development level of the provinces. Most provinces in the Red River Delta and Southeast regions have higher rankings. And the backward provinces are mainly in the Northern midland and mountainous area.

Keywords: Principal Component Analysis, economic development levels, provinces of Vietnam.

JEL classification: C; C38

1. Giới thiệu

Nền kinh tế sau cải cách của Việt Nam đã đạt được những thành tựu đáng ghi nhận. Tuy nhiên có sự khác biệt lớn trong trình độ phát triển giữa các vùng, các tỉnh. Do đó việc nhận thức và phân tích sự chênh lệch này có ý nghĩa quan trọng trong việc thực hiện các chính sách phát triển vùng, liên kết vùng và chính sách phát triển bền vững của cả nước.

Theo Nghị định 92/2006/NĐ - CP của Chính phủ về lập, phê duyệt và quản lý quy hoạch tổng thể phát triển kinh tế - xã hội, nước ta được phân thành 6 vùng kinh tế - xã hội gồm vùng Trung du và Miền núi phía Bắc, Đồng bằng sông Hồng, Bắc Trung Bộ và Duyên hải miền Trung, Tây Nguyên, Đông Nam Bộ và vùng Đồng bằng sông Cửu Long.

Vùng trung du phía Bắc và miền núi bao gồm 14 tỉnh, có nguồn tài nguyên khoáng sản phong phú, cũng là nơi có nhiều đồng bào dân tộc thiểu số sinh sống. Nhưng đây cũng là khu vực có tỷ lệ hộ nghèo cao nhất, thu nhập bình quân đầu người thấp nhất cả nước. Vùng Đồng bằng sông Hồng gồm 9 tỉnh và 2 thành phố trực thuộc Trung ương, là vùng có diện tích nhỏ nhất và dân cư đông đúc nhất. Đồng bằng sông Hồng có lịch sử phát triển lâu đời, là vùng trung tâm chính trị, kinh tế, văn hóa và công nghệ của cả nước. Vùng Bắc Trung Bộ và duyên hải miền Trung gồm 13 tỉnh và 1 thành phố trực thuộc Trung ương. Vùng này chủ

yếu tập trung phát triển tổng hợp kinh tế biển bao gồm thủy sản, du lịch biển và dịch vụ vận tải biển. Tây Nguyên bao gồm 5 tỉnh, có mật độ dân số thấp nhất, tại đây chủ yếu phát triển các loại cây công nghiệp như cà phê, hồ tiêu, cao su, đồng thời cũng là vùng có tài nguyên rừng phong phú nhất cả nước, tuy nhiên phát triển kinh tế còn nhiều khó khăn. Đông Nam Bộ gồm 5 tỉnh và 1 thành phố trực thuộc Trung ương, có thu nhập bình quân cao nhất và tỷ lệ hộ nghèo thấp nhất, đồng thời cũng là vùng phát triển nhất cả nước. Đồng bằng sông Cửu Long bao gồm 12 tỉnh và 1 thành phố trực thuộc Trung ương và là khu vực sản xuất lương thực lớn nhất của đất nước, là vùng xếp thứ 3 về thu nhập bình quân đầu người. Trong việc đánh giá trình độ phát triển kinh tế của một khu vực hay địa phương thường nhắc tới các chỉ tiêu đơn về tổng lượng hoặc chỉ số bình quân như: tổng sản phẩm quốc nội (GDP), tổng sản lượng quốc gia GNP, GDP bình quân đầu người, thu nhập bình quân, chi tiêu bình quân ... Tuy nhiên việc chỉ sử dụng một chỉ tiêu có những hạn chế nhất định khi nó chỉ phản ánh một phương diện nhất định của nền kinh tế. Mặt khác nếu lựa chọn chỉ tiêu về tổng lượng sẽ bỏ qua các yếu tố về dân số, mật độ dẫn đến có thể đánh giá quá cao trình độ phát triển của một khu vực. Nếu lựa chọn chỉ tiêu bình quân nhiều khi sẽ dẫn đến kết luận không sát với thực

tế nếu khu vực đó có dân cư quá ít trên một diện tích lớn (Tong Huili, 2019). Do đó cần có phương pháp để đánh giá tổng hợp nhiều chỉ tiêu cùng lúc đối với kinh tế địa phương và khu vực. Phân tích thành phần chính (Principal Component Analysis – PCA) là một trong những phương pháp phân tích dữ liệu nhiều biến đơn giản nhất. Phương pháp PCA xây dựng một không gian mới ít chiều hơn nhưng lại có khả năng đại diện cho bộ dữ liệu gốc, đảm bảo độ biến thiên của dữ liệu trên mỗi chiều mới (Thủy, 2015). PCA được ứng dụng trong nhiều lĩnh vực, một trong số đó là giải bài toán xếp hạng năng lực tổng hợp các quan sát thông qua tính toán tổng điểm của các quan sát đó trên các trục chính. Liu Yuan Chao (2003), sử dụng phương pháp phân tích thành phần chính, dựa trên dữ liệu về lao động, GDP, GDP bình quân đầu người, sản lượng của 3 thành phần kinh tế, vốn đầu tư toàn xã hội, tổng doanh thu hàng hóa bán lẻ, thu ngân sách tiến hành đánh giá toàn diện sự phát triển kinh tế của 40 quận và huyện thuộc Trung Khánh, Trung Quốc. Tương tự, có rất nhiều các nghiên cứu khác về ứng dụng của PCA trong phân tích kinh tế tổng hợp như của Gao Wei (2005) đối với trường hợp của tỉnh Giang Tô, nghiên cứu của Zhao Xu (2010) đối với tỉnh Sơn Đông... Gần đây có nghiên cứu của Wang Ning (2019) lựa chọn 11 chỉ tiêu phát triển kinh tế thông qua PCA đánh giá toàn diện trình độ phát triển kinh tế của sáu tỉnh miền Trung Trung Quốc. Hay các nghiên cứu khác như Lin Yu Jin (2020) đánh giá về mức độ phát triển kinh tế của 9 thành phố thuộc tỉnh Phúc Kiến, Chen Hong Shou (2021) so sánh mức độ phát triển của 34 thành phố thuộc tỉnh Quảng Đông và Giang Tô... Có thể thấy Trung Quốc là quốc gia có mức độ chênh lệch giữa các vùng, các tỉnh thậm chí là cấp huyện, xã khá nghiêm trọng. Do đó các nghiên cứu về đánh giá tình trạng kinh tế nói chung và ứng dụng PCA trong đánh giá chênh lệch kinh tế tổng hợp nói riêng đều khá phong phú, và được ứng dụng rộng rãi. Tuy nhiên các nghiên cứu về lĩnh vực này ở Việt Nam còn hạn chế. Trong bài viết này tác giả tiến hành thiết lập hệ thống gồm 11 tiêu chí kinh tế của 63 tỉnh thành phố. Thông qua phân tích PCA sẽ cho ra một chỉ tiêu tổng hợp mới đại diện cho hệ thống nhiều chỉ tiêu đơn trong bộ dữ liệu gốc, để đánh giá xếp hạng và phân loại trình độ phát triển kinh tế của các tỉnh thành Việt Nam.

2. Phương pháp phân tích thành phần chính

2.1. Ý tưởng của phân tích thành phần chính

Phân tích thành phần chính (Principal Component Analysis, PCA) là một phương pháp thống kê đa biến để kiểm tra mối tương quan giữa nhiều biến. Dựa trên ý tưởng giảm chiều dữ liệu

từ bộ dữ liệu có nhiều chiều nhưng bảo toàn lượng thông tin nhiều nhất có thể. Tức là PCA tiến hành chuyển đổi tổ hợp nhiều biến ban đầu thành một vài (số ít) biến tổng hợp trên nguyên tắc thông tin dữ liệu mất đi ít nhất.

Phương pháp này không chỉ có thể loại bỏ mối tương quan giữa các chỉ số (biến) đánh giá, giảm khối lượng công việc mà còn có thể thực hiện đơn giản nhanh chóng, chính xác thông qua các phần mềm toán học.

Thành phần chính F_1 có giá trị phương sai càng lớn thì vai trò của nó trong việc đánh giá tổng hợp càng lớn, lượng thông tin mà nó biểu đạt càng nhiều. Nếu thành phần chính thứ nhất (F_1) chưa đủ để biểu đạt tốt lượng thông tin của dữ liệu m biến ban đầu thì tiếp tục lựa chọn thành phần chính thứ hai (F_2), cứ như vậy cho đến m thành phần chính (Lín HảiMing, 2013)

2.2. Các bước phân tích thành phần chính

Bước 1: Số liệu ban đầu được sắp xếp dưới dạng một ma trận $X = (x_{ij})_{n \times m}$. Trong đó n là số quan sát, m là số biến. Vì các biến ban đầu có đơn vị khác nhau nên tiến hành tiêu chuẩn hóa dữ liệu thành $Z = (z_{ij})_{n \times m}$.

Bước 2: Từ ma trận số liệu đã được tiêu chuẩn hóa, tính hệ số ma trận tương quan $R = (r_{ij})_{n \times m}$. Tính m giá trị riêng λ_p ($p=1,2,3,...m$), các vector riêng $L_p = L_{p1}, L_{p2}, ..., L_{pm}$. Rút ra công thức biểu diễn thành phần chính:

$$F_p = L_{p1}Z_1 + L_{p2}Z_2 + \dots + L_{pm}Z_m$$

Trong đó $p = 1, 2, 3 \dots m$

Bước 3: Xác định tỷ lệ đóng góp của các thành phần chính: Lựa chọn các giá trị riêng $\lambda_p > 1$. Tỷ lệ đóng góp tích lũy của các thành phần chính là tỷ lệ phần trăm lượng thông tin mà các thành phần chính đó bao hàm. Tỷ lệ đóng góp tích lũy càng cao, càng gần với tình hình thực tế. Phạm vi chấp nhận được của tỷ lệ đóng góp tích lũy của các thành phần chính là 70% đến 90% (Wang ErPeng, 2021)

Bước 4: Tiến hành tính tổng trọng số (Weighted Sum Method) đối với k thành phần chính đã chọn thu được tổng điểm đánh giá của các thành phần chính (principal component score)

$$F = \sum_{p=1}^k \left(\frac{\lambda_p}{\sum_{p=1}^m \lambda_p} \right) F_p$$

3. Ứng dụng PCA trong đánh giá tổng hợp trình độ phát triển của 63 tỉnh thành

3.1. Thiết lập hệ thống chỉ tiêu về trình độ phát triển kinh tế địa phương

Việc lựa chọn chỉ tiêu dựa trên các nguyên tắc sau: Thứ nhất, phải phản ánh các khía cạnh chính của nền kinh tế một cách tổng hợp và toàn

diện, vừa có tính đại diện vừa có thể so sánh với các địa phương khác. *Thứ hai*, đảm bảo tính khả thi, nghĩa là số liệu có thể thu thập được, đã được tính toán rõ ràng. *Thứ ba*, đảm bảo tính khoa học, xem xét tình hình thực tế các quy luật khách quan kết hợp với cơ sở lý luận, kết quả phân tích các chỉ tiêu cần phù hợp với tình hình thực tế.

Dựa trên các nguyên tắc lựa chọn số liệu trên kết hợp với tham khảo các kết quả của những nghiên cứu trước, tác giả lựa chọn các chỉ tiêu sau:

Tổng sản phẩm trên địa bàn hay GRDP địa phương (X_1), tổng thu ngân sách địa phương (X_2), tổng đầu tư toàn xã hội địa phương (X_3), tổng doanh thu bán lẻ hàng hóa (X_4), đầu tư trực tiếp nước ngoài (X_5), tổng kim ngạch xuất nhập khẩu (X_6), GRDP bình quân đầu người (X_7), cơ cấu công nghiệp trong GRDP địa phương (X_8), tỷ lệ dân thành thị (X_9), thu nhập bình quân đầu người nhóm 1 của địa phương (X_{10}), thu nhập bình quân đầu người nhóm 5 của địa phương (X_{11}). Trong đó nhóm 1 tương ứng với nhóm nghèo nhất và nhóm 5 tương ứng với nhóm giàu nhất. Các chỉ

tiêu được lựa chọn nhằm phản ánh các phương diện của quy mô kinh tế, mức độ giàu có của địa phương, độ mở thương mại, cơ cấu kinh tế, trình độ đô thị hóa và mức sống dân cư.

Số liệu về xuất nhập khẩu được lấy từ báo cáo xuất nhập khẩu 2018 bản chính thức của Tổng cục hải quan, số liệu về thu nhập bình quân của nhóm 1 và 5 lấy từ trang web của Tổng cục thống kê tại địa chỉ <https://www.gso.gov.vn/>. Các số liệu khác được lấy từ cuốn *Tư liệu kinh tế – xã hội 63 tỉnh, thành phố trực thuộc trung ương* của Tổng cục thống kê Việt Nam (2020), riêng số liệu GRDP bình quân đầu người được tác giả tính toán bằng cách chia GRDP cho tổng dân số.

3.2. Phân tích mức độ phù hợp của bộ dữ liệu

Trước khi thực hiện phân tích nhân tố, sử dụng phần mềm SPSS để nhập dữ liệu liên quan và thực hiện kiểm định KMO và Bartlett. Kết quả kiểm định giá trị KMO là 0,808 và kiểm định Bartlett có Sig = 0,000 cho thấy giữa các chỉ tiêu có tương quan với nhau. Dữ liệu gốc đáp ứng các yêu cầu kiểm định và phù hợp để tiến hành phân tích thành phần chính.

Bảng 1: Kết quả kiểm định KMO và Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0,808
Bartlett's Test of Sphericity	Approx. Chi-Square	1,15E+03
	df	55
	Sig.	0

Nguồn: Tính toán của tác giả từ SPSS

Bảng 2: Giá trị riêng và tích lũy phương sai của các thành phần chính

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,315	66,496	66,496	7,315	66,496	66,496	5,751	52,28	52,28
2	1,836	16,695	83,192	1,836	16,695	83,192	3,4	30,911	83,192
3	0,813	7,393	90,585						
4	0,422	3,836	94,421						
5	0,235	2,14	96,561						
6	0,157	1,429	97,99						
7	0,14	1,274	99,264						
8	0,038	0,343	99,607						
9	0,022	0,195	99,803						
10	0,013	0,122	99,924						
11	0,008	0,076	100						

Ghi chú: Fi là thành phần chính thứ i, Variability (%) là phần trăm của phương sai và Cumulative (%): phần trăm tích lũy của phương sai.

3.3. Lựa chọn thành phần chính

Sau khi tiêu chuẩn hóa dữ liệu gốc và sử dụng phần mềm SPSS tiến hành phân tích thành phần chính thu được kết quả như bảng 2.

Qua Bảng 2 có thể thấy các giá trị riêng của hai thành phần thứ nhất và thứ hai lớn hơn 1 và

Nguồn: Tính toán của tác giả từ SPSS chiếm 83,192% phần trăm tích lũy của phương sai, tức là 2 thành phần đầu tiên có thể đại diện tới 83,192% thông tin dữ liệu gốc. Cụ thể thành phần chính đầu tiên chứa đựng 66,496%, thành phần chính thứ hai chứa đựng 16,695 % thông tin dữ liệu gốc.

Bảng 3: Ma trận thành phần xoay

	Component (Các thành phần)	
	1	2
Zscore(X_4)	0,968	0,126
Zscore(X_3)	0,963	0,196
Zscore(X_2)	0,96	0,174
Zscore(X_1)	0,952	0,263
Zscore(X_5)	0,878	0,321
Zscore(X_6)	0,702	0,527
Zscore(X_9)	0,532	0,485
Zscore(X_8)	-0,156	0,826
Zscore(X_7)	0,337	0,82
Zscore(X_{10})	0,406	0,805
Zscore(X_{11})	0,456	0,791

Nguồn: Tính toán của tác giả từ SPSS

Bảng ma trận thành phần xoay thể hiện mối tương quan giữa các biến và các trục chính (thành phần chính). Bảng 3 cho thấy các biến tổng doanh thu bán lẻ hàng hóa (X_4), tổng đầu tư toàn xã hội (X_3), thu ngân sách nhà nước (X_2), GDP (X_1), đầu tư trực tiếp nước ngoài (X_5), tổng kim ngạch xuất nhập khẩu (X_6) có đóng góp lớn cho

thành phần chính thứ nhất (F1). Tỷ lệ dân thành thị (X_9) mặc dù thiên về F1 nhưng hệ số tương quan không quá mạnh, với hệ số 0,538. Các biến cơ cấu công nghiệp (X_8), GDP bình quân đầu người (X_7), thu nhập bình quân của nhóm 1 (X_{10}) và thu nhập bình quân của nhóm 5 (X_{11}) có tương quan mạnh mẽ với thành phần chính thứ 2.

Bảng 4: Ma trận trọng số của các thành phần

	Component (Các thành phần)	
	1	2
Zscore(X_1)	0,192	-0,063
Zscore(X_2)	0,211	-0,102
Zscore(X_3)	0,208	-0,093
Zscore(X_4)	0,222	-0,124
Zscore(X_5)	0,163	-0,024
Zscore(X_6)	0,081	0,096
Zscore(X_7)	-0,066	0,289
Zscore(X_8)	-0,192	0,383
Zscore(X_9)	0,045	0,11
Zscore(X_{10})	-0,045	0,27
Zscore(X_{11})	-0,03	0,255

Nguồn: Tính toán của tác giả từ SPSS

Từ bảng 4 có thể biểu diễn các thành phần chính như sau:

$$F1 = 0,192X_1 + 0,211X_2 + 0,208X_3 + 0,222X_4 + 0,163X_5 + 0,081X_6 - 0,066X_7 - 0,192X_8 + 0,045X_9 - 0,045X_{10} - 0,03X_{11}$$

Tương tự với F2

$$F2 = -0,063X_1 - 0,102X_2 - 0,093X_3 - 0,124X_4 - 0,024X_5 + 0,096X_6 + 0,289X_7 + 0,383X_8 + 0,11X_9 + 0,27X_{10} + 0,255X_{11}$$

Dựa vào công thức biểu diễn các thành phần chính như trên phần mềm đã tự động tính được

điểm số của các quan sát (ở đây chính là các tỉnh, thành phố) trên các trục chính (Các biến lúc này ở dạng đã được tiêu chuẩn hóa). Từ kết quả điểm số các quan sát trên trục chính tiếp tục tính tổng điểm của các quan sát theo công thức:

$$F = \frac{\lambda_1}{\lambda_1 + \lambda_2} F1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} F2$$

$$= \frac{7.315}{9.151} F1 + \frac{1.836}{9.151} F2$$

4. Kết quả phân tích

Tác giả tiến hành tính toán tổng điểm F của các tỉnh, thành theo công thức trên, thực hiện xếp hạng và thu được bảng số liệu sau.

Bảng 5: Kết quả điểm thành phần, tổng điểm và xếp hạng của các tỉnh, thành phố

Tỉnh, thành	F1	F2	Tổng điểm	Xếp hạng
TP. Hồ Chí Minh	6,229	-0,239	4,931	1
Hà Nội	4,269	-0,008	3,411	2
Bình Dương	0,340	3,091	0,892	3
Hải Phòng	0,765	1,011	0,815	4
Đồng Nai	0,359	1,925	0,673	5
Bà Rịa Vũng Tàu	-0,380	3,572	0,413	6
Bắc Ninh	-0,314	2,990	0,349	7
Đà Nẵng	-0,030	1,424	0,261	8
Quảng Ninh	-0,059	1,461	0,246	9
Thanh Hóa	0,242	-0,279	0,138	10
Cần Thơ	-0,106	0,823	0,081	11
Thái Nguyên	-0,244	1,236	0,053	12
Kiên Giang	0,170	-0,478	0,040	13
Khánh Hòa	-0,032	-0,089	-0,043	14
Nghệ An	0,181	-0,945	-0,045	15
Long An	-0,287	0,858	-0,057	16
Vĩnh Phúc	-0,225	0,583	-0,063	17
Thừa Thiên Huế	-0,046	-0,203	-0,077	18
Đắk Lắk	0,139	-0,994	-0,088	19
Sóc Trăng	-0,022	-0,361	-0,090	20
An Giang	0,025	-0,551	-0,091	21
Hải Dương	-0,257	0,572	-0,091	22
Tây Ninh	-0,297	0,725	-0,092	23
Lâm Đồng	-0,055	-0,249	-0,094	24
Tiền Giang	-0,133	-0,077	-0,122	25
Bình Định	-0,103	-0,322	-0,147	26
Quảng Nam	-0,142	-0,235	-0,161	27
Đồng Tháp	-0,108	-0,416	-0,170	28
Bình Thuận	-0,225	-0,017	-0,183	29
Ninh Bình	-0,313	0,244	-0,201	30
Bạc Liêu	-0,045	-0,861	-0,209	31
Vĩnh Long	-0,125	-0,594	-0,219	32
Bắc Giang	-0,371	0,371	-0,222	33
Bến Tre	-0,147	-0,537	-0,226	34
Bình Phước	-0,355	0,255	-0,233	35
Nam Định	-0,264	-0,135	-0,238	36
Thái Bình	-0,250	-0,202	-0,240	37
Gia Lai	-0,162	-0,553	-0,240	38
Hà Tĩnh	-0,280	-0,127	-0,250	39
Ninh Thuận	-0,135	-0,712	-0,251	40
Quảng Ngãi	-0,361	0,169	-0,254	41
Hậu Giang	-0,247	-0,313	-0,260	42
Hưng Yên	-0,517	0,746	-0,263	43
Phú Thọ	-0,268	-0,329	-0,280	44
Phú Yên	-0,251	-0,414	-0,284	45
Cà Mau	-0,245	-0,476	-0,291	46
Trà Vinh	-0,274	-0,427	-0,305	47
Lạng Sơn	-0,121	-1,064	-0,310	48
Quảng Trị	-0,242	-0,655	-0,325	49
Đắk Nông	-0,222	-0,742	-0,327	50
Quảng Bình	-0,242	-0,707	-0,335	51
Hà Nam	-0,612	0,654	-0,358	52
Lào Cai	-0,390	-0,238	-0,360	53
Yên Bái	-0,262	-0,752	-0,360	54
Kon Tum	-0,295	-0,633	-0,363	55
Sơn La	-0,257	-0,890	-0,384	56
Bắc Kạn	-0,191	-1,176	-0,389	57

Bảng 5: Kết quả điểm thành phần, tổng điểm và xếp hạng của các tỉnh, thành phố (tiếp)

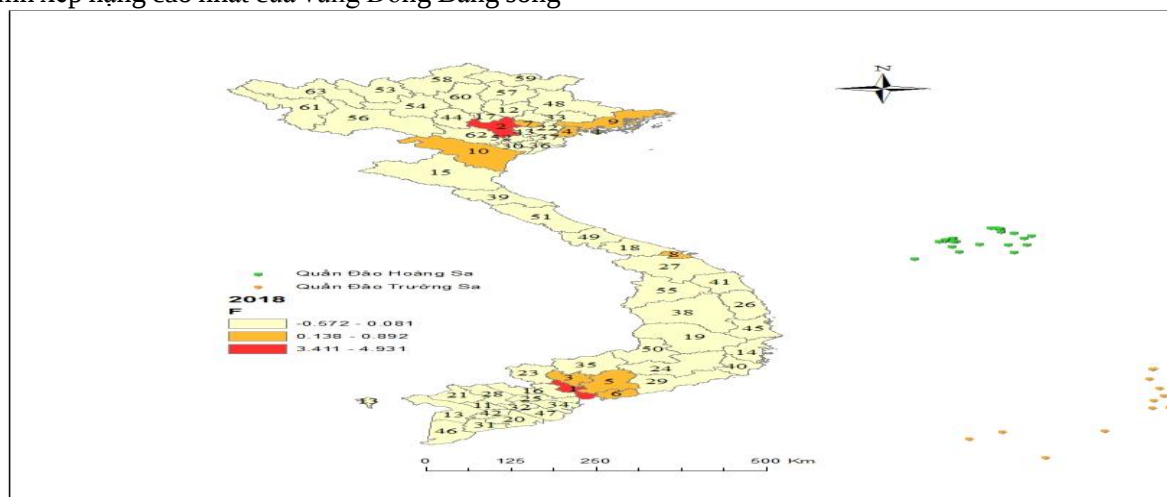
Tỉnh, thành	F1	F2	Tổng điểm	Xếp hạng
Hà Giang	-0,241	-1,014	-0,396	58
Cao Bằng	-0,261	-0,954	-0,400	59
Tuyên Quang	-0,358	-0,754	-0,437	60
Điện Biên	-0,239	-1,299	-0,452	61
Hòa Bình	-0,507	-0,247	-0,455	62
Lai Châu	-0,605	-0,441	-0,572	63

Nguồn: Tính toán của tác giả

Kết quả tính toán từ bảng 5 cho thấy chỉ có 13 tỉnh đạt điểm số lớn hơn 0, chiếm 20,63% tổng số tỉnh và 50 tỉnh có điểm âm, chiếm 79,37% tổng số tỉnh. Không ngạc nhiên khi Hồ Chí Minh là tỉnh thành có điểm cao nhất là đạt 4,931. Tỉnh xếp cuối cùng là Lai châu với số điểm -0,572 cho thấy khoảng cách tuyệt đối giữa các tỉnh là rất lớn, các tỉnh có trình độ phát triển kinh tế rất chênh lệch nhau. Ngay trong nội bộ các vùng cũng có sự chênh lệch rất nhiều. Trong 10 tỉnh cuối bảng xếp hạng thì có tới 9 tỉnh thuộc Trung du và miền núi phía Bắc. Tỉnh xếp hạng cao nhất của vùng này là Thái Nguyên ở vị trí thứ 12, thấp nhất là Lai Châu xếp thứ 63. Tỉnh xếp hạng cao nhất của vùng Đồng bằng Sông Hồng là Hà Nội ở vị trí thứ 2, thấp nhất là Hà Nam xếp thứ 52. Tỉnh xếp hạng cao nhất của vùng Bắc Trung Bộ và duyên hải Miền Trung là Đà Nẵng xếp thứ 8, tỉnh thấp nhất là Quảng Bình xếp thứ 51. Tỉnh xếp hạng cao nhất của Tây Nguyên là Đắk Lắk ở vị trí 19, thấp nhất là Kom Tum xếp thứ 55. Tỉnh xếp hạng cao nhất của vùng Đông Nam Bộ là TP. Hồ Chí Minh xếp thứ nhất và thấp nhất là Bình Phước ở vị trí 35. Tỉnh xếp hạng cao nhất của vùng Đồng Bằng sông

Cửu Long là Cần Thơ xếp 11, thấp nhất là Trà Vinh xếp thứ 47. So sánh với số liệu GRDP bình quân đầu người thấy rằng mặc dù Bà Rịa Vũng Tàu có GRDP bình quân đầu người lớn nhất cả nước do có lợi thế khai thác dầu khí nhưng ở xếp hạng tổng hợp tỉnh này đứng vị trí 6. Lào Cai cũng có GRDP/người ở mức khá nhưng các chỉ số khác như GRDP, đầu tư toàn xã hội, FDI, thu nhập nhóm 1 và 5 đều khá thấp dẫn đến bảng xếp hạng chung chỉ đứng thứ 53. Với trường hợp của Đắk Lắk, mặc dù có GRDP/người ở mức thấp, nhưng Zcore của các giá trị chính thuộc F1 âm ít hơn các giá trị của F2 dẫn đến các hệ số triệt tiêu nhau F1 có số điểm dương và tổng xếp hạng ở vị trí 19. Như vậy xét một cách tổng thể các phương diện kinh tế Đắk Lắk có những tiềm năng nhất định.

Để trực quan hóa kết quả năng lực kinh tế tổng hợp của các địa phương tác giả sử dụng phần mềm Arc gis bản đồ hóa kết quả thu được ở bảng 5. Sử dụng chức năng phân loại tự nhiên jenks natural breaks để phân các địa phương thành 3 nhóm: thấp, trung bình, cao tương ứng với ba nhóm trình độ phát triển kinh tế tổng hợp.



Hình 1: Trình độ phát triển kinh tế của các địa phương phân theo điểm và xếp hạng

Nguồn: Tính toán của tác giả từ Arc Gis

Như trong kết quả Bảng 5, nhờ lịch sử khai thác lâu đời, các nền tảng trong quá khứ, điều kiện tự nhiên thuận lợi, ưu ái về các chính sách kinh tế đã khiến cho Hà Nội và Thành phố Hồ Chí Minh

trở thành hai trung tâm dẫn đầu sự phát triển kinh tế của Việt Nam. Hình 1 cũng cho thấy chỉ có 2 địa phương này thuộc nhóm phát triển và các tỉnh có trình độ kinh tế nhóm 2 (mức trung bình) đa số

phân bố xung quanh hai trung tâm kinh tế này. Điều này phần nào minh chứng cho xu hướng hiệu ứng cực hóa lan tỏa, tác động lan tỏa chỉ ảnh hưởng đến các tỉnh có vị trí gần nơi phát triển. Có tới 54 tỉnh, thành phố thuộc nhóm 3 mức dưới mức trung bình cho thấy vấn đề phát triển vùng không đồng đều và thiếu đồng bộ vẫn còn nổi cộm, nó cũng là thách thức lớn nhất đối với việc thực hiện chiến lược phát triển bao trùm. Đặc biệt theo kết quả phân nhóm, miền Trung chỉ có duy nhất Đà Nẵng thuộc nhóm 2. Do đó trong tương lai cần thúc đẩy hơn nữa vai trò của hai đầu tàu kinh tế Hà Nội và Thành phố Hồ Chí Minh, đồng thời đầu tư cho “cực phát triển” tại miền Trung để trình độ phát triển giữa các khu vực trở nên đồng đều hơn.

5 Kết luận

Phương pháp phân tích thành phần chính khiến cho mối quan hệ tương tác phức tạp giữa các yếu tố được đơn giản hóa thành mối quan hệ tuyến tính giữa các thành phần chính, làm giảm mối quan hệ ảnh hưởng lẫn nhau giữa các yếu tố

khác nhau, tạo điều kiện thuận lợi đáng kể cho việc xử lý dữ liệu một cách có hệ thống.

Áp dụng phương pháp phân tích thành phần chính vẫn có những hạn chế nhất định, khi kết quả thu được sẽ phụ thuộc vào việc lựa chọn dữ liệu ban đầu. Do đó trong quá trình thu thập dữ liệu cần tuân thủ các yêu cầu về tính khách quan, tính đại diện, tính chính xác và phù hợp với mục tiêu nghiên cứu. Nếu không kết quả cho ra sẽ bị thiên lệch theo hướng chủ quan của người chọn dữ liệu.

Thông qua phân tích thành phần chính của dữ liệu thống kê của 63 tỉnh, thành phố năm 2018 đã loại bỏ được tính phiến diện khi chỉ sử dụng một chỉ tiêu kinh tế. Kết quả cung cấp một cách đánh giá toàn diện về trình độ phát triển kinh tế của các địa phương. Các tỉnh, thành có sự khác biệt đáng kể về trình độ phát triển tổng hợp. Tỉnh có trình độ phát triển cao vẫn tập trung chủ yếu ở vùng Đồng bằng sông Hồng và Đông Nam Bộ. Và các tỉnh lạc hậu hơn chủ yếu thuộc khu vực Trung du và miền núi phía Bắc.

TÀI LIỆU THAM KHẢO

- [1]. Chen HongShou. (2021). Phân tích thực nghiệm trình độ phát triển kinh tế của các khu vực của tỉnh Quảng Đông dựa trên phương pháp PCA. *Tạp chí Đại học Công nghệ Liêu Ninh (Ấn bản Khoa học Xã hội)* (01), 26-30.
- [2]. Gao wei, Zhou min. (2005). Phân tích thực nghiệm trình độ phát triển kinh tế của các khu vực tỉnh Giang Tô - Phương pháp đánh giá tổng hợp đa chỉ tiêu và ứng dụng. *Tạp chí Thống kê toán học và quản lý* (01), 5-9+14.
- [3]. Lin HaiMing & Du ZiFang. (2013). Các vấn đề cần lưu ý trong phân tích thành phần chính đánh giá tổng hợp. *Tạp chí Thống kê nghiên cứu*, 30(8), 25-31.
- [4]. Lin Yu Jin. (2020). Đánh giá trình độ phát triển kinh tế của các thành phố của tỉnh Phúc Kiến bằng phương pháp phân tích thành phần chính. *Tạp chí hướng dẫn giám sát chất lượng Quảng Tây* (08), 95-97.
- [5]. Liu Yuan Chao, Cao Yue Qun, Lu Hao. (2003). Phân tích thành phần chính trình độ phát triển kinh tế các quận huyện của thành phố Trùng Khánh. *Tạp chí học viện giao thông Trùng Khánh (Ấn bản Khoa học Xã hội)* (04), 31-34.
- [6]. Tong Huili. (2019). Nghiên cứu chênh lệch phát triển kinh tế giữa các huyện của tỉnh Hà Nam và các nguyên nhân ảnh hưởng. *Luận văn thạc sĩ, đại học sư phạm Chiết Giang*
- [7]. Thủy, N. M & Tuyên, N. T. M. (2015). Ứng dụng phương pháp phân tích thành phần chính, hồi quy logistic và giản đồ yếu thích trong đánh giá cảm quan sản phẩm sữa gạo. *Tạp chí Khoa học Trường Đại học Cần Thơ*, 37, 11-20.
- [8]. Wang ErPeng & Yao Lu. (2021). Phân nguyên nhân sự cố ở các mỏ than tỉnh Sơn Đông dựa trên phương pháp phân tích thành phần chính. *Tạp chí than đá Thiểm Tây* (01), 57-60.
- [9]. Wăng Ning, Wang YuFei & Chen Qian. (2019). Phân tích thực nghiệm trình độ phát triển của 6 tỉnh miền Trung dựa trên SPSS. *Tạp chí học viện đồ họa Bắc Kinh* (08), 85-87.
- [10]. Zhao Xu. (2010). Nghiên cứu chênh lệch về thời gian và không gian trong trình độ phát triển kinh tế tổng hợp của các huyện thuộc tỉnh Sơn Đông. *Luận văn thạc sĩ, đại học sư phạm Sơn Đông*.

Thông tin tác giả:

Hồ Thị Linh Trang

- Đơn vị công tác: Đại học Lanzhou, Cam Túc, Trung Quốc (Lanzhou University, Lanzhou City, Gansu Province, China)
- Địa chỉ email: thi2018@lzu.edu.cn

Ngày nhận bài: 03/2/2021

Ngày nhận bản sửa: 13/03/2021

Ngày duyệt đăng: 30/03/2021