A methodological approach for obtaining Sustainable development synthetic indicators for Venezuela

ANNA GABRIELA PÉREZ* MONTSERRAT HERNÁNDEZ LÓPEZ**
FRANCKLIN RIVAS ECHEVERRÍA^,^^,^^^
*Universidad de Los Andes.
Facultad de Ciencias Económicas y Sociales.
Departamento de Estadística.
Mérida – VENEZUELA
**Universidad de La Laguna.
Facultad de Economía, Empresa y Turismo.
Departamento de Economía Aplicada y Métodos Cuantitativos.
Tenerife – ESPAÑA
^Universidad de Los Andes.
Facultad de Ingeniería.
Laboratorio de Sistemas Inteligentes.
Mérida – VENEZUELA
^^ Escuela de Ingeniería
Pontificia Universidad Católica del Ecuador-sede Ibarra
Ibarra, Provincia Imbabura
ECUADOR
^^^ Programa Prometeo
Secretaría de Educación Superior, Ciencia, Tecnología e Innovación
Quito, Provincia Pichincha
ECUADOR
gabipm@ula.ve, mhdezl@ull.edu.es, rivas@ula.ve

Abstract: In this work two types of sustainable development synthetic indicators using a two-step methodology are proposed. There are first obtained partial indicators through the principal components method and then the synthetic indicator are found. The first proposed synthetic indicator is a simple aggregative indicator, and the others are based on the concept of distance, specifically the Ivanovic distance and DP2 distance.

Keywords: sustainable development, partial indicators, synthetic indicator, distance analysis, principal components analysis.

1 Introduction

The sustainable development concept is created after impact analysis on the environment given by human activities done searching better living conditions. Within the United Nations, in particular the World Commission on Environment and Development, has formulated the definition of sustainable development. In “Our Common Future” report (1987), also known as the Brundtland Report, the concept of sustainable development is outlined as "human progress path able to meet the needs and aspirations of the present generation without compromising the ability of future generations to meet their needs" [3, 13].

The measurement of this type of development is based on the use of indicators as tools for explaining, summarizing and reporting the progress or behavior of this phenomenon contextualized in a given territory and culture [14].

Sustainable development Indicators are tools for which there is no single methodology. One of the types of indicators that can be proposed are synthetic indicators, which are defined as a function of the variables that are part of the dynamics of a more complex process, which aims to simplify and provide a value that is easily interpretable.

In the case of sustainable development, the development of a synthetic indicator is a two stage process. In the first stage partial indicators of social, environmental, economic and institutional dimensions are obtained, and in the second stage, the aggregative development indicator, synthetizes the four partial indicators obtained in the previous step.

This article proposes an approach for obtaining sustainable development synthetic indicators for Venezuela using two methodologies: a principal components-based statistical methodology and another based on the distance definition.

2 Basic definitions before obtaining a sustainable development indicator

2.1 Synthetic Indicator

It is understood by synthetic indicator (I) a function with domain in the space that generates a υ-dimension vector that reflects or summarizes the information contained in it´s k partial indicators.

Both partial indicators and the final synthetic indicators are defined as variables that approach the concept to be measured. However, a very good approximation or fit goodness of measurements are required of such partial indicators in order that can minimize measurement error.

Generically, we can express the synthetic indicator I through the following expression:

$$I = \Psi\left(I_1, I_2, ..., I_k\right)$$

where each partial indicator I_i can be expressed as:

$$I_i = \left(I_1, I_2, ..., I_n\right) \quad i = 1, ..., k$$

The use of a synthetic indicator should take into account certain aspects of partial indicators:

- The aggregation of partial indicators and heterogeneity of measurement units.
- The need that information contained in the partial indicators is measurable.
- The weighting of the partial indicators.

According to [2, 4, 17, 7, 11, 15], among others, the synthetic indicator I, as a function of the partial indicators, must meet the following properties:

- Existence of the indicator and Determination the set of partial indicators.
- Monotony in variations of the partial indicators, if there is an improvement in any of the partial indicators and the rest remain constant, the synthetic indicator should be improved.
- Uniqueness for partial indicators, for a given situation the synthetic indicator assume a single value. The invariance to changes in sources and/or scale property must be met together.
- "Degree one" homogeneity of the function that defines the synthetic indicator such that if the partial indicators are multiplied by a given constant, the synthetic indicator is also multiplied by this constant.
- Transitivity: Let a, b and c three different situations of the measurable objective, and I(a), I(b) and I(c) values of the synthetic indicator, it should be verified that:
  If I(a)> I (b) and I (b)> I (c) then I (a)> I (c)

- Completeness in the use of the information provided by the partial indicators, avoiding duplication of information.

2.2. Indicators based on distance

The distance between two cases or two individuals refers to the similarity or dissimilarity between both of them. For calculating the distance there are diverse procedures that consider the nature of the variables that have been measured on the considered cases.

It has been named as dissimilarity or distance between two individuals I and j (d_{ij}) the measures of difference degree between the two individuals in relation to a certain number of quantitative and/or qualitative features. The d_{ij} value is always non-negative and the higher is the value, the greater the difference between i and j individuals [16, 10, 5, 8].

The synthetic indicators based on distance, focus on the calculation of the difference between the value of one indicator and other taken as a reference, and have the goodness of solving the heterogeneity problem on the units of measurement [11, 6, 10, 18]. The comparison by difference is performed in absolute or squares terms, between each indicator in different territorial units or in relation to a considered baseline.

Mathematically, the approach can be expressed as follows: Let X be a nxp data matrix where n represents the territorial units and p the indicators. The x_{ij} value represents the state of the i-th indicator (i = 1, 2, ..., p) in the territorial unit j (j = 1, 2, ..., n).

It is denoted by $X^*=[X_1^*, X_2^*, ..., X_p^*]$ the reference indicators vector. The p-metric distance generally is defined by:

$$D_p=\left\{\sum_i |X_{ji}-X_i|^p\right\}^{\frac{1}{p}}$$

The families of indicators based on the latter case are called synthetic distance indicators, verifying the conditions of the distance in a metric space (no negativity, commutativity and triangular condition).

Within this family of indicators are included indicators based on the Frechet distance [9, 6] which is defined as follows:

$$DF=\sum_{j=1}^{p}\left|\frac{X_j-X^*}{\sigma_j}\right|$$

where σ_i represents the standard deviation of the values taken by the i-th indicator.

This distance measurement has the disadvantage of not weighing the importance of each indicator, or considering the possible independence between them, which is essential for a synthetic indicator being calculated as the sum of the partial indicators. Also, functional dependency condition must be fulfilled, in the sense that if an indicator is a function of others, it would collect information concerning all of them, so the latter could be eliminated. Finally, the synthetic indicator should be invariant with respect to the baseline and should also test the additivity hypothesis.

Although the synthetic indicator based on Frechet distance cannot be considered entirely acceptable, it can serve as a basis for defining further distances. In fact, based on the Frechet distance, Ivanovic (1963) has formulated Ivanovic distance (ID) as follows:

$$ID(r,k)=\sum_{i=1}^{p}\left|\frac{X_j-X^*}{\sigma_j}\right|\prod(1-r_{ji})$$

In the Ivanovic distance a term that reflects the product of the partial correlation coefficients complement between the i indicator and j indicator is introduced. That factor is introduced for taking care of duplicate information, because r_{ji} measures the association degree between the new introduced indicator and the earlier one, after removing the effects of the other included indicators.

For this indicator it should be found a criterion for establishing the entry order of the partial indicators and must be taken into account for this the amount of new information that contribute to the overall index. Ivanovic proposes an iterative method for achieving a hierarchization. It stars accepting that each individual indicator importance can be expressed by the degree of dependence between it and the target to be measured.

The application problem of this principle is that the value of ID is unknown and hierarchy would be previous to the calculation. Therefore it must be started the hierarchization process using the Frechet distance (FD), and the procedure is as follows:

- It is calculated the FD
- Partial indicators are ordered according to the absolute value of the simple correlation coefficient between each indicator and FD.
- ID is calculated by introducing the partial indicators according to the order established in the previous step
- The procedure is iterated

According to Peña-Trapero [11], synthetic indicators based on ID meet all the properties of a good synthetic indicator except the completeness property, since the weighting factor does not eliminate the double information. However, the DP2 distance methodology, attempts to correct the dependence of the partial indicators through an iterative process that starts from the assumption of linear dependence.

In the P2 distance (DP2) method, weighting factor w_j is given by the complement of the determination coefficients (R^2) of the regressions between all the partial indicators, i.e $w_j=1-R^2$, which corresponds to the unexplained variation of the model. The researcher must fit the best models and then obtain the R^2 values that are going to be used for calculating the weighting factor.

So, the general form of sustainable development indicator based on DP2 is:

$$DP2=\sum_{j=1}^{4}\left|\frac{X_j-X^*}{\sigma_j}\right|\left(1-R_j^2\right)$$

$$DP2=\sum_{j=1}^{4}\frac{d_j}{\sigma_j}w_j$$

2.3. Principal component-based indicators

The principal component analysis (PCA) is one of the most widespread multivariate statistical analysis techniques that allows structuring a set of multivariate data obtained from a population whose probability distribution does not need be known.

The first description of this technique was performed by Karl Pearson in 1901, in a published work on orthogonal settings where he proposed a multi spatial points system adjusting into a line or a plane. Pearson's original work was based on those linear combinations of the original variables for which the unexplained variance was minimal. These linear combinations or principal components, generate a function plane of the original variables. This work was later taken again by Hotteling (1933), who was the first to formulate the principal component analysis as has been spread until the present day ([16, 10, 5, 8]).

The main PCA objective is to reduce the dimensionality of the data set, which consists of a large number of related variables and retain the maximum variation in that set. This is accomplished by a transformation into a new set of variables called uncorrelated principal components and independent from each other, which are arranged such that the first components retain the most of the variation present in all the original variables.

When there is a high positive correlation between all variables, the first principal component has all of its coordinates with the same sign and can be interpreted as a weighted average of all variables or as a size global factor. The remaining principal components are interpreted in terms of positive and negative coefficients presenting and involving the opposition of a group of variables against another or others [16].

The problem that arises is how to define the number of components to be retained. Among the suggested criteria are the following:

- Proportion of variance explained by components: Kaiser in 1958 suggests retaining the principal components that explain at least 80% of the total variation.
- Cattell Procedure (1966) or Sedimentation graphic.
- Kaiser Average Criterion (1972): Once the eigenvalues are calculated, the average is calculated and retained only those values who are above average.

3 Obtaining sustainable development partial and synthetic indicators for Venezuela

For obtaining the synthetic indicators there are going to be employed two methodologies: principal components and distance.

With the method of principal components, sustainable development partial indicators for each dimension are obtained, and three synthetic indicators are obtained: one based on principal components and two of them based on the concept of distance, specifically Ivanovic distance and DP2 distance.

The information used in this study has been provided by the National Statistical Institute of Venezuela.

3.1. Partial indicators

For obtaining sustainable development partial indicators on social, environmental, economic and institutional dimensions, the principal components method is used and the correlation matrix R is decomposed, variables are standardized or typed in order to avoid the effects of measurement units and magnitudes of values. For each dimension the first component is used as synthetic indicator, although the first two components are retained for creating the graphics.

For using principal components as indicators, it´s usually employed the first principal component identifying the component's cases score as a factor or synthetic indicator [1].

In Table 1, the variables considered in each dimension for obtaining partial indicators are shown.

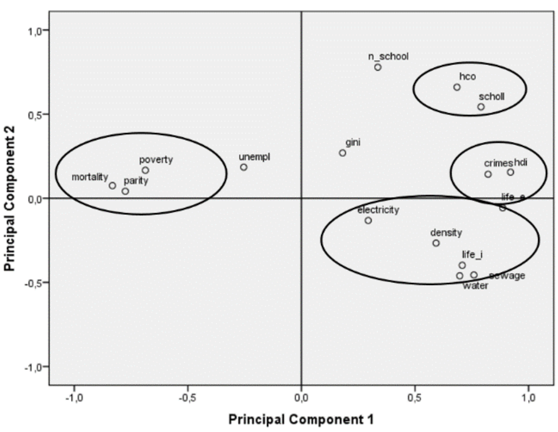
Using SPSS statistical software version 20, analysis for each dimension is performed. Regarding measures of sampling adequacy, the statistic coefficient KMO indicates that in all cases it is appropriate the principal component analysis, since the value in each dimension exceeds 0.5. The sphericity test is even stronger, $H_{0:}$ $|R|=1$ hypothesis is contrasted, which means that the correlations between the variables included in each dimension are not significant, but for the four dimensions this hypothesis is rejected, it is concluded that the R matrices are different from the identity matrix and correlations between variables are significant.

Social Dimension Partial Indicator: To build this partial indicator it was used information concerning 16 variables associated with poverty, mortality, sanitary facilities, drinking water, health provision, education level, housing conditions, crime and population. In the first component those states with better living and development conditions (urban) oppose against those associated with poverty, higher infant mortality and where the average number of sons on women younger than 20 years are higher than in urban areas.

Looking at Graphic 1, and analyzing each quadrant, it can be distinguished in the first quadrant those states with the highest human development index, however, these states are directly related to a high number of crimes reported. There also can be distinguished in this quadrant houses with critical overcrowding and scholar attendance or not. The Gini index appears very close to the origin, which can be interpreted as a very similar behavior concerning this variable in all states and correlation with other variables is low.

In the second quadrant those states with the highest rate of unemployment, high poverty, high mortality and more children in women younger than 20 years (parity) are presented. In the fourth quadrant those states with higher population density, life expectancy, living conditions, are directly related to the potable water provision and electricity in homes, and also have proper black water disposal (sewage disposal).

Graphic 1. Social Dimension Components

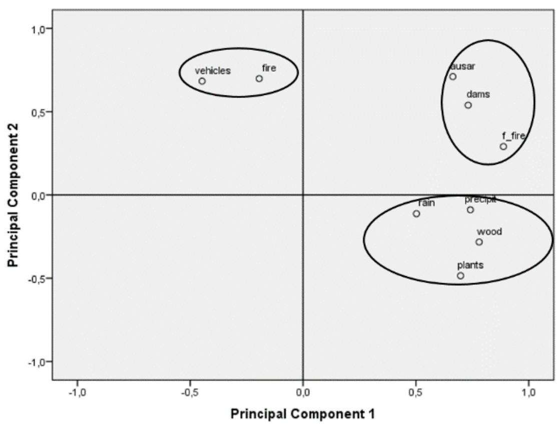


Source: Own preparation

Environmental Dimension Partial Indicator: For building this partial indicator, information was available from 9 variables associated to the natural physical environment (air quality, plantations, amount of water, protected areas). In the first component those states with better conditions of natural physical environment oppose to those who are associated with impaired air quality.

In the first quadrant (Graphic 2) those states with the largest number of areas under special administration regimen (abraes) and drinking water reservoirs are projected, but also are most affected by forest fires (f_fire). In the second quadrant are located those states with high number of vehicles and with high number of not forest fires. Finally, in the fourth quadrant those states associated with the largest area of forest plantations (plants) and timber plantations (wood) are represented; abundant rainfall are presented in these states and are therefore more noticeable rainy days.

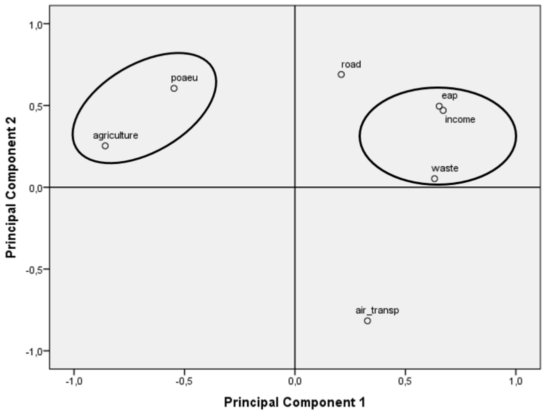
Graphic 2. Environmental Dimension Components



Source: Own preparation

Economic Dimension Partial Indicator: For building this partial indicator it was available information from 7 variables associated to economic, trade, management and solid waste generation and transportation. The first component opposes those states with higher economically active population, higher income and which has an important road network, to the states where the population is occupied mainly in agriculture and in which there is a higher percentage of open active economic units. The second component is associated primarily with the air transport (Air_trans) variable located in the fourth quadrant.

Graphic 3. Economic Dimension Components

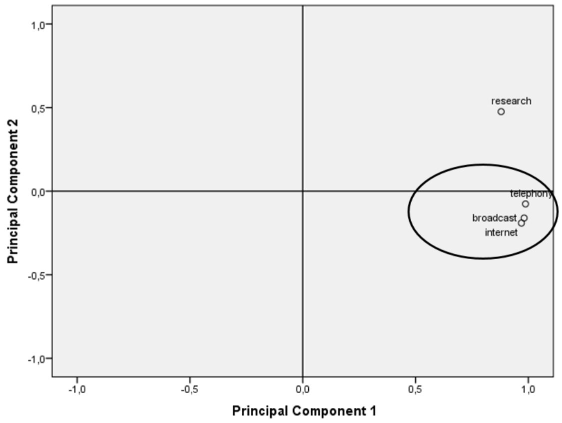


Source: Own preparation

Institutional Dimension Partial Indicator: For this dimension is only available information concerning 4 variables associated with issues of information access, communications infrastructure and science and technology.

The first component is a size component, i.e., the relationship between the four variables of this dimension is linear direct (positive ratio), therefore, the first component corresponds to a weighted sum of these variables and in the graphic all variables are located on the right side between the first and fourth quadrant. The second component is a shape component and opposes those states with the highest number of researchers and innovators registered in the National Observatory of science and technology (ONCTI) to states with more houses related to information access variables (telephone, internet homes and broadcasting service subscription), as shown in Graphic 4.

Graphic 4. Institutional Dimension Components



Source: Own preparation

3.2 Sustainable Development Synthetic indicator (SDSI) for Venezuela

The SDSI is defined as an additive or aggregative indicator of the scores obtained in each of the dimensions that define the sustainable development partial indicators, as has been presented by [9]. It was not considered any weight for partial indicators, since obtaining them through principal components method, the variables that make these indicators were already weighted. The general formula of the SDSI can be expressed by:

$$SDSI = I_{Social} + I_{Environmental} + I_{Economic} + I_{Institutional}$$

Table 2 shows the scores of the partial indicators and SDSI for each state. It is noteworthy that during 2005 over 50% of states have negative scores on the partial indicators, being the institutional dimension the worst performer.

Likewise, during the year 2005 was observed only one state with all positive partial indicators (Carabobo). Cojedes and Sucre states recorded all negative partial indicators, resulting in states that require immediate attention in social, economic, environmental and institutional aspects.

Moreover, the states of the Venezuelan flat places or "llanos" (Apure, Barinas, Guárico, Portuguesa and Yaracuy) together with the states of Amazonas, Delta Amacuro, Monagas and Trujillo have only one positive score corresponding to the environmental dimension and the rest of the dimensions have negative scores. This group of states has special characteristics: they have the highest percentage of population employed in agriculture because agriculture and animal husbandry are their main economic activity, and higher rates of poverty, infant mortality and average parity occur in women under 20 years. In Amazonas and Delta Amacuro state is settled most of the indigenous population of Venezuela, and feature large protected areas due to biodiversity living there.

It can also be observed another group of states (Anzoategui, Aragua, Distrito Capital, Miranda and Zulia) having negative scores in the environmental dimension partial indicator and positive scores on the other three dimensions partial indicators. These states are characterized by the highest population densities in the country, and for being the most important from a political point of view by the main public organizations concentration (Capital District) and economically by the intense industrial activity, especially the oil industry (Anzoategui and Zulia).

The Nueva Esparta, Tachira and Vargas states have a very similar behavior in the partial indicators as they had positive scores on the economic and social dimension and negative scores on environmental and institutional dimensions. For these states policies should be geared to the conservation of the natural physical environment and the strengthening of science and technology.

Bolivar State presented positive scores in three dimensions (had the highest score in the environmental dimension), and negative in the institutional dimension. Merida also introduced three partial indicators positive and negative on the economic dimension. Falcon and Lara states have three negative and one positive partial indicator: in the case of Falcon the positive partial indicator corresponds to the economic dimension and in Lara to the institutional dimension.

Regarding SDSI scores, only 9 of the 24 states (37.5%) had positive scores, the higher was recorded by Miranda state followed by the Capital District. Anzoategui, Aragua, Bolivar, Carabobo, Zulia, Mérida and Táchira had also positive SDSI, however for the last two states this value is slightly higher than 0 as shown in Table 2.

For states with negative SDSI scores should be designed special policies to promote sustainable development in all dimensions, policies that can bring economic activity to impact the environment, and the reduction of poverty and the problems associated with it. Although this study was conducted using data from 2005, the current situation does not differ much from that moment.

While it is interesting to analyze the states with SDSI scores less than 0, it is worth examining the optimal value of this indicator, so the optimal vector of scores is given by

$$X^{*}=[I_{Social}\quad I_{Environmental}\quad I_{Economic}\quad I_{Institutional}]$$
$$X^{*}=[2{,}7621\quad 1{,}795\quad 0{,}5445\quad 2{,}708]$$

The optimal values vector is different to maximum values vector. The optimal values are those that maximize the positive effects of a variable and minimize the negative effects of another depending on the context (the optimum is obtained in each dimension).

3.3 Sustainable Development Indicators for Venezuela based on Distance

These indicators were constructed according to the methodological specifications presented by Zarzosa [17] and Peña-Trapero [11] which are basically used to measure social welfare and in this work are adapted to measure sustainable development.

Indicators based on distance use the similarity concept. For calculating these distances or similarity measures is essential to determine the reference point, which is given in this application by the partial indicators optimal values, that were previously obtained by the principal components method.

The partial indicators optimal values vector was given by (1), i.e.:

$X^{*}=[I_{Social}\ \ I_{Environmental}\ \ I_{Economic}\ \ I_{Institutional}]$
$X^{*}=[2{,}7621\quad 1{,}795\quad 0{,}5445\quad 2{,}708]$

From here, there have been identified two sustainable development indicators based on the distance concept: Ivanovic distance (ID) and P2 distance (DP2). In both methodologies is desired to obtain an indicator of the form:

$$I=\sum_{j=1}^{4} d\left|X_j-X_j^{*}\right| w_j \qquad (2)$$

The $d\left|X_j-X_j^{*}\right|$ term of equation (2), represents the distance of the sustainable development partial indicator of dimension j from the optimal indicator of that dimension, and the w_j term represents the weight associated with each dimension, according to the used method (Ivanovic or DP2).

In the Ivanovic distance method weights are calculated such that they can eliminate the dependence between different partial indicators, and therefore the information contained in the previous indicators when it's introduced a new one.

The effect of removing dependency or information duplicity is achieved by introducing as weighting factor the complement to the unit of the partial correlation coefficient, however, in practice there is hardly found a perfect independence. This distance can be calculated in stages as follows:

- Frechet distance is calculated. In the procedure presented in this article, Frechet distance has been simplified, since the partial indicators have standard deviation equal to unity ($\sigma_j=1$), then $F=\sum_{j=1}^{4}\left|X_j-X^{*}\right|$
- It must be set the order of entry of the partial indicators, which is determined by examining the correlation coefficients between the partial indicators of each dimension and Frechet distance. In to obtain these coefficients and perform their respective hierarchization, the following order of entry was determined: Environmental, institutional, economic and social dimension.
- The second term will be that partial indicator whose correlation with Frechet distance is the smallest (having already selected the first term) and its weighting factor will be the correlation coefficient complement with the partial indicator selected in the previous step. The second term corresponds to the institutional dimension indicator.
- The third term will be the partial indicator whose correlation is lower (in this step only two partial indicators remains) and its weighting factor corresponding to the partial correlation coefficient between this third term and the previous one, controlling the effect of the partial indicator obtained as first term. The third term corresponds to the economic dimensión.
- The last term is the partial indicator with the highest Frechet distance correlation, and its weight is the

partial correlation coefficient with the indicator of the third term, controlling with it the effects of the first two indicators. In this application the last term corresponds to the social dimension indicator.

Indicators based on Ivanovic distance have a disadvantage that is associated with the entry order of the partial indicators; because the values that can be taken by the correlation coefficients can make that the distance do not present uniqueness. To correct this problem it can be hierarchized the partial indicators according to the degree of information provided. It can also be used the multiple correlation coefficient between the indicator that provides more information and the other indicators [6].

In the synthetic indicators based on the DP2 distance, partial indicators have standard deviation equal to unity ($\sigma_j=1$), an then $DP2=\sum_{j=1}^{4} d_j w_j$

Table 3 shows the w_j weights assigned to each partial indicator. It was found that environmental dimension has the biggest influential in the distance dimension calculation. However, the weights assigned to the Ivanovic distance are slightly higher than those assigned to the DP2 distance.

Observing the distances obtained by both methods it can be found that the coefficients are similar. The greater distances were recorded by Cojedes and Sucre states applying both methods, as can be seen in Table 4. The Ivanovic distance also stand the greatest distance presented by Falcon, Vargas, Nueva Esparta and Apure states, while the DP2 distance stand out with greater distances: Apure, Barinas, Trujillo and Portuguesa states. States with smaller distance corresponds to Miranda, Capital District, Bolivar and Carabobo and coincide in both methods.

4 Conclusions

In this work the production of sustainable development synthetic indicators for Venezuela was presented, using two statistical methods such as principal component analysis and distance analysis. At the moment of finishing this study it has not been found any other sustainable development synthetic indicator for the Venezuelan reality.

For these synthetic indicators, two-stage methodology was used. The first stage was to obtain partial indicators for each sustainable development dimension, based on principal components generated from variables associated with each dimension, and the subsequent found of Sustainable Development Indicator in two ways: first as a sum of partial indicators and second based on two types of distances: Ivanovic distance (ID) and the P2 distance (DP2).

In the first way, the SDSI has been found as a simple aggregative index of each dimension partial indicators obtained by principal component and has the properties that should have any good indicator.

For the second approach, the SDSI is obtained as the distance between the optimal partial indicators vector. These distance indicators are weighted, in ID by the partial correlation coefficient complement, which happens to be a factor controlling the effect that a dimension can have on one or others and their relation with the final indicator. In the DP2 by the complement of the determination coefficient (1 - R2), this term represents the unexplained variation in the adjusted regression, and so would eliminate the dependency between a partial indicator and the rest, and eliminates the information duplication.

Examining the results it can be found that Venezuelan states with the highest SDSI score are those with less distance to the optimal indicators vector: Aragua, Bolivar, Carabobo, Capital District, Miranda and Zulia. States with lower SDSI score present greater distances and are Apure, Barinas, Cojedes and Sucre states.

Examining ID and DP2 scores, it can be found that ID scores are slightly lower than those obtained in the DP2, and are evident subtle variations in the states scores order. With both procedures (ID and DP2) has been obtained a sustainable development synthetic indicator for Venezuela based on the distance concept. By the reported experience in literature and in this application it's appropriate to apply the DP2 method.

To further study is recommended to analyze the temporal evolution of the indicators proposed in this research.

5 References

[1] Azqueta, D. Escobar, L. 2004. Calidad de Vida Urbana. Ekonomiaz 57:216 – 239. Retrieved from http://dialnet.unirioja.es/servlet/articulo?codigo=1373292

[2] Escobar, L. 2008. Indicadores Ambientales Sintéticos: una aproximación conceptual desde la estadística multivariante. Gestión y Ambiente 11:1:121–140.

[3] Gabaldón, A. 2006. Desarrollo Sustentable: La salida de América Latina. Caracas – Venezuela: Editorial Grijalbo.

[4] García, S. 1996. Indicadores para el desarrollo sostenible de la pesca. FAO. Taller de trabajo sobre indicadores de la sostenibilidad pesquera. Brisbane – Australia. Retrieved from http://www.fao.org/docrep/004/w4745s/w4745s13.htm

[5] Hair, J. Anderson, R. Thatam, R. Black, William. 1999. Análisis Multivariante. Quinta edición. Prentice Hall. Madrid – España.

[6] Instituto de Análisis Económico y Empresarial de Andalucía. 1997. Indicador Sintético de Bienestar Municipal de Andalucía. Retrieved from http://www.economiaandaluza.es/publicaciones/indicador-sintetico-bienestar-municipal-andalucia

[7] Jaramillo, L. 2008. Indicadores Ambientales Sintéticos: una aproximación conceptual desde la estadística multivariante. Gestión y Ambiente 11:1:121–140. Retrieved from http://www.redalyc.org/articulo.oa?id=169414452008

[8] Johnon, R. Wichern, D. 1998. Applied Multivariate Statistical Analysis. Prentice Hall, INC. New Jersey – United States of America.

[9] Pedregal, B. Torres, J. Zoido, F. 2006. Propuesta Metodológica para la Medición del Desarrollo y las Desigualdades Territoriales. Aplicación al Territorio Andaluz. Revista Electrónica de Geografía y Ciencias Sociales. 10, 220. Retrieved from http://www.ub.es/geocrit/sn/sn-220.htm

[10] Peña, D. 2002. Análisis de Datos Multivariantes. Mc Graw Hill Interamericana de España. Madrid – España.

[11] Peña – Trapero, B. 2009. La Medición del Bienestar Social: una revisión crítica. Estudios de Economía Aplicada 27:2:299 – 324. Retrieved from http://www.redalyc.org/articulo.oa?id=30117056001

[12] Peña – Trapero, B. 1977. Problemas de la Medición del bienestar y Conceptos Afines. Una Aplicación al Caso Español.

Instituto Nacional de Estadística. Madrid – España.

[13] Pierri, N. 2005. Desacuerdos sobre el Desarrollo Sustentable – Historia del Concepto de Desarrollo Sustentable. Universidad Autónoma de Zacatecas. México.

[14] Salas, A. Bartón, S. 2008. Propuesta de Indicadores de Desarrollo Sustentable para las Comunas de Ancud, Puqueldón y Quemchi. Pontificia Universidad Católica de Chile. Instituto de Estudios Urbanos y Terriotriales. Santiago de Chile. Chile.

[15] Schuschny, A. Soto, H. 2009. Guía Metodológica: Diseño de Indicadores Compuestos de Desarrollo Sostenible. División de Desarrollo Sostenible y Asentamientos Humanos (DDSAH) de la Comisión Económica para América Latina y el Caribe (CEPAL). Santiago de Chile – Chile.

[16] Uriel, E. Aldás, J. 2005. Análisis Multivariante Aplicado. Internacional Thomson Editores Spain. Madrid – España.

[17] Zarzosa, M. 1996. Aproximación a la Medición del Bienestar Social. Idoneidad del Indicador Sintético “Distancia P_2”. Aplicación al Caso Español. Cuadernos de Economía 24:139-163. Retrieved from https://repositorio.uam.es/bitstream/handle/10486/5105/32640_7.pdf?sequence=1

[18] Zarzosa, P. 2009. Estimación de la Pobreza en las Comunidades Autónomas Españolas Mediante la Distancia DP2 de Peña. Estudios de Economía Aplicada 27:2:397-415. Retrieved from http://www.redalyc.org/articulo.oa?id=30117056005.

Table 1. Variables considered in each dimension for obtaining partial indicators

Social	Environmental	Economic	Institutional
Percentage of houses with potable water (water supply) Percentage of houses with electricity (electricity) Percentage of houses with sewage disposal (sewage) School attendance (school) Children between 7 and 12 who do not attend school (not scholar) Houses in critical overcrowding (hco) Infant mortality rate (mortality) Life expectancy at birth (life_e) Average parity in women under 20 years (parity) Population density (density) Gini Coefficient (gini) Human Development Index (HDI) Living conditions index (life_i) Poverty (poverty) Reported crimes (crimes) Unemployment rate (unempl)	Areas under special administration regimen (ausar) Number of vehicles (vehicles) Number of reservoirs (dams) Total precipitation (precipitation) Timber plantations by State (wood) Appreciable number of raining days (rain) Number of reported Fires (fire) Area affected by forest fires (f_fire) Area of forest plantations (plants)	Percentage of the population employed in agriculture (agriculture) Road network (road) Production of solid waste (waste) Economically active population (EAP) Income per person (income) Percentage of open active economic units (poaeu) Passenger movement in Air transport: loading and unloading (Air_trans)	Houses that have fixed telephony (telephony) Internet users (internet) Houses with broadcasting service subscription (broadcast) Researchers and innovators registered in the ONCTI (research)

Source: Own preparation

Table 2. Sustainable Development Partial Indicators and Synthetic Indicator Scores

State	Social Indicator	Environmental Indicator	Economic Indicator	Institutional Indicator	SDSI
Amazonas	-1,9374	1,8381	-0,9671	-0,8462	-1,91
Anzoátegui	0,4037	-0,3653	0,6696	0,0609	0,77
Apure	-1,5387	0,3695	-0,5394	-0,6789	-2,39
Aragua	0,9192	-1,7661	1,0217	0,3385	2,10
Barinas	-0,8586	0,2514	-1,3077	-0,4605	-2,38
Bolívar	0,4186	2,5574	0,9753	-0,0083	3,94
Carabobo	1,2352	0,2385	1,4816	0,6621	3,62
Cojedes	-0,6828	-1,0047	-0,7809	-0,7150	-3,18
Delta Amacuro	-1,4038	1,0247	-1,2977	-0,8488	-2,53
Distrito Capital	2,5062	-1,1287	1,3978	2,2215	5,00
Falcón	-0,1149	-1,0082	0,4896	-0,3664	-1,00
Guárico	-0,4585	1,4058	-0,8526	-0,5419	-0,45
Lara	-0,0064	-0,6952	-0,0209	0,2186	-0,50
Mérida	0,0212	0,4516	-0,5418	0,1734	0,10
Miranda	1,5535	0,7248	1,1992	3,0368	5,06
Monagas	-0,0278	0,483	-0,6456	-0,521	-0,71
Nueva Esparta	0,6619	-1,1298	0,7077	-0,4811	-0,24
Portuguesa	-0,2461	0,2006	-1,3141	-0,567	-1,93
Sucre	-0,7155	-1,1602	-0,778	-0,426	-3,08
Táchira	0,261	-0,4881	0,3995	-0,0713	0,10
Trujillo	-0,5606	0,3191	-1,4752	-0,5264	-2,24
Vargas	0,3682	-0,8874	0,7287	-0,7239	-0,51
Yaracuy	-0,4124	0,532	-0,0846	-0,6427	-0,61
Zulia	0,6147	-0,9027	1,5347	1,713	2,96

Source: Own preparation

Table 3. Weights of the Partial Indicators

Dimension	Ivanovic distance weights	DP2 distance weights
Environmental	0,918	0,814
Institutional	0,657	0,443
Economic	0,360	0,333
Social	0,385	0,224

Source: Own preparation

Table 4. Sustainable Development Indicators based on Distance. Ivanovic Distance and DP2 Distance

State	SDSI Value based on Ivanovic Distance	SDSI Value based on DP2 Distance
Amazonas	4,73	5,47
Anzoátegui	4,68	4,46
Apure	5,58	6,00
Aragua	4,25	4,09
Barinas	5,56	6,07
Bolívar	3,54	3,54
Carabobo	3,70	3,71
Cojedes	6,62	6,76
Delta Amacuro	5,31	5,94
Distrito Capital	3,41	3,30
Falcón	5,72	5,45
Guárico	4,24	4,63
Lara	5,19	5,19
Mérida	4,34	4,60
Miranda	3,23	3,36
Monagas	4,83	4,98
Nueva Esparta	5,65	5,19
Portuguesa	5,44	5,79
Sucre	6,59	6,78
Táchira	4,94	4,72
Trujillo	5,49	5,97
Vargas	5,71	5,29
Yaracuy	4,81	4,87
Zulia	4,31	4,59

Source: Own preparation