



Grupo Temático N° 20: Cuestiones conceptuales y metodológicas de los temas que abarcan los estudios del trabajo

Coordinadores: Cynthia Pok y Andrea Lorenzetti

La elaboración de índices sintéticos de bienestar social. Validación teórica y empírica del método de agregación/ponderación.

Autor/es: Eugenio Actis Di Pasquale

E – mails: edipasq@mdp.edu.ar

Pertenencia institucional: Grupo Estudios del Trabajo – Facultad de Ciencias Económicas y Sociales – Universidad Nacional de Mar del Plata

1. INTRODUCCION

Los índices sintéticos de bienestar social representan una medida que se obtiene mediante la agregación adecuada de un conjunto de indicadores sociales que pertenecen a las distintas dimensiones de ese concepto. Su elaboración no es sencilla, ya que los indicadores son de distinta naturaleza y provienen de diferentes fuentes de datos. En este sentido, en la literatura se recomienda seguir distintas etapas para su elaboración -selección de las dimensiones y variables, estandarización, ponderación y agregación-. Sin embargo, en cada una de las etapas hay una variedad de perspectivas, métodos y procedimientos, muchas veces antagónicos, que exigen la necesidad de elegir a través de algún criterio que sea adecuado para el estudio de un fenómeno en particular. Las principales dificultades que se presentan en cada etapa son:

1- *Selección de los indicadores.* Hay dos perspectivas epistemológicamente opuestas al momento de definir los indicadores. Una es la “técnica-metodológica”, que considera que se deben obtener datos de la observación empírica que se relacionan con un concepto. La otra es la “metodológica-conceptual”, que concibe al indicador como un equivalente empírico de una dimensión conceptual que pertenece a alguna teoría social. En esta línea, primero debemos saber a qué llamamos bienestar social, y luego podremos preguntarnos cómo lo medimos.



El concepto de bienestar social contiene múltiples acepciones y orientaciones, lo que induce a que su significado suela ser interpretado con cierta vaguedad e imprecisión. Para elaborar un constructo se tiene que hacer un trabajo analítico a nivel teórico, de tal forma que se obtenga un concepto no observable sin referentes empíricos inmediatos. Luego se especifican las dimensiones y con posterioridad los indicadores. Por ende, para determinar si la introducción de un indicador es relevante, se debe estudiar no sólo la confiabilidad -a través de criterios estadísticos y conocimiento de la calidad de los datos-, sino también la validez interna –tomando como referencia estudios previos que parten del mismo constructo o bien analizando el constructo elaborado-. Es decir, en este marco, en la selección de los indicadores también entran en juego un conjunto de criterios subjetivos explícitos que son los que permiten que el indicador cumpla con la función mediadora entre teoría y empiria.

Algunos autores elaboran un conjunto de hipótesis con respecto a la relación lógica entre los indicadores simples y el concepto de bienestar social para sortear estas dificultades, tales como la completitud, bondad de los indicadores simples y objetividad (Pena Trapero, 1977: 67). Sin embargo, consideramos que estas presunciones resultan en una pseudo-validación, dado que se estaría elaborando un modelo cercano a la visión empirista extrema, que considera a los datos como puros y completos, sin comprender la brecha que siempre se presenta entre teoría y empiria. La operacionalización en Ciencias Sociales tiene ciertas limitaciones que deben ser tenidas en cuenta en todo el proceso, lo que permitirá otorgar un grado de relatividad a la medición efectuada.

2- *Selección de la técnica de estandarización.* Los indicadores sociales se miden en distintas unidades de medida: años, personas, porcentajes, etc. La estandarización permite que un conjunto heterogéneo de indicadores pueda ser transformado a una unidad de medida común. Cada una de las distintas técnicas que se presentan en la literatura nos puede llevar a resultados diferentes. No obstante, no profundizaremos en este análisis dado que excede el objetivo del presente trabajo.

3- *Problemas en la ponderación simple y parcial.* No se puede dar la misma importancia a todos los indicadores y, por ende, es necesario ponderarlos. Encontrar un buen sistema de ponderación es difícil, por lo que se debe determinar la potencia relativa de un indicador con respecto a otro, a fin de determinar su peso en la medida específica que se intenta llevar a cabo.

4- *Selección del método de agregación.* En esta etapa se busca conjugar a los indicadores en un índice sintético a través de algún procedimiento estadístico, matemático o lógico. Resulta deseable que el valor que arroje la medida final represente niveles factibles de ser interpretados en



una escala ordinal -alto, medio, bajo o nulo-. Esto último dependerá de la técnica de estandarización utilizada y de que el procedimiento de agregación no haga perder esa propiedad. Entonces, estamos en presencia del problema de agrupar medidas de distintas dimensiones de un concepto, todo lo cual requiere de un trabajo de reflexión conceptual y, a la vez, técnico que construya la forma teóricamente más adecuada de resolver esta cuestión.

Estas cuestiones resultan los principales problemas a resolver al momento de la elaboración de un índice. No obstante, en esta ponencia nos enfocaremos solamente en estudiar los métodos de agregación/ponderación¹ para reconocer cuál es el más adecuado desde una perspectiva metodológica-conceptual, ¿hay alguna forma de comprobarlo? En la literatura hay un cierto consenso en que a través de la deducción de ciertas propiedades matemáticas se puede evaluar la bondad de estos procesos. Sin embargo, consideramos que algunas de éstas estarían aludiendo a una equivalencia de formas entre la estructura lógica del sistema numérico y la estructura de la realidad.

En virtud de ello, proponemos reformular desde una visión crítica algunas de estas propiedades, de tal forma que sean concluyentes para aquellas mediciones de conceptos o fenómenos sociales que son abordadas desde la perspectiva metodológica-conceptual. De esta manera, a partir de la reformulación pretendemos comparar teóricamente a cuatro métodos de agregación/ponderación de indicadores, destacando los que más propiedades satisfacen, y luego, validar empíricamente esas conclusiones, con el fin de establecer cuál es el método que permite generar un índice de bienestar social con validez interna y externa.

La ponencia está organizada en apartados. En el apartado 2, presentamos, discutimos y reformulamos las propiedades deseables de los índices sintéticos. Luego en el apartado 3, realizamos una exposición y comparación teórica de los métodos que se configuran como más relevantes y difundidos por su aplicabilidad práctica: la media ponderada, el análisis de componentes principales, la distancia P_2 y la agregación de conjuntos difusos. En el apartado 4 comparamos los resultados de las estimaciones con estos cuatro métodos con datos del año 2005. Finalmente, elaboramos las conclusiones.

¹ Si bien la ponderación y agregación son dos etapas separadas, algunos métodos tienen integradas ambas etapas en un único procedimiento. Por ende, decidimos utilizar el genérico agregación/ponderación que involucra a todos los métodos seleccionados para el análisis.

2. LAS PROPIEDADES EXIGIBLES AL ÍNDICE SINTÉTICO. UNA REFORMULACIÓN.

Las propiedades matemáticas que debe reunir un indicador sintético para dar cuenta de la bondad de los métodos utilizados en su elaboración que son las siguientes (Zarzosa Espina, 1996:67-81; Pena Trapero, 2009: 310-311):

1. *Existencia y determinación.* La función matemática que define el indicador sintético debe de ser tal que éste exista y no sea indeterminado para todo sistema de indicadores parciales.
2. *Monotonía.* El indicador sintético debe responder positivamente a una modificación positiva de los componentes y negativamente a una modificación negativa. Por ejemplo, si un indicador parcial indica una situación de mejora en el componente que mide y los demás indicadores permanecen constantes, entonces el indicador sintético debe reflejar esa mejora.
3. *Unicidad.* El indicador sintético debe ser tal que para una situación social determinada se obtenga un único valor, omitiendo toda apreciación personal en la selección de métodos de estandarización y ponderadores. Es decir, que esta propiedad se cumpliría en aquellos métodos integrados que se alimentan de los datos en bruto, y luego a través de una serie de procedimientos totalmente “rígidos” -que no permiten la posibilidad de incorporar apreciaciones personales-, se consigue un único resultado posible.
4. *Invariencia.* Esta propiedad está en parte derivada de la anterior, y significa que el indicador sintético debe ser invariante respecto a un cambio de origen y/o de escala en las unidades en que vengan expresados los valores –estados- de los componentes. Es decir, se puede suponer una transformación lineal $X_i' = a + bX_i$, donde a y b son dos parámetros –b positivo- y X_i es la variable que viene medida por el componente i. A partir de allí se analiza la posibilidad que el valor del índice se modifique.
5. *Homogeneidad.* La función matemática que defina el indicador sintético, en función de los indicadores simples, debe ser una función homogénea de grado uno. Así, si todos los indicadores parciales aumentan o disminuyen en la misma proporción, el indicador sintético aumentará o disminuirá en la misma proporción.
6. *Transitividad.* Si a, b y c son tres situaciones distintas del objetivo medido por el indicador sintético, e $I(a)$, $I(b)$, $I(c)$ son los valores que toma el indicador sintético para esas tres situaciones, entonces debe verificarse que: $[I(a) > I(b) \wedge I(b) > I(c)]$ implica que $[I(a) > I(c)]$.



Puesto que los indicadores sintéticos que se utilizan habitualmente son números, esta propiedad no planteará problemas.

7. *Exhaustividad*. El indicador sintético debe ser tal que aproveche al máximo la información suministrada por los indicadores parciales, y cuando sea necesario, eliminar aquella que se encuentra duplicada.

Las condiciones 1, 2, 5 y 6 son sencillas de entender y comprobar, y por lo general, la cumplen todos los métodos. Sin embargo, para el resto de las propiedades realizaremos algunas salvedades. Con respecto a la invarianza, si bien en la bibliografía se presenta como una propiedad de los índices sintéticos, consideramos que debe verificarse en el procedimiento de estandarización utilizado, ya que es en los datos originales –indicadores simples o parciales- donde deben comprobarse los posibles cambios frente a las modificaciones de origen o escala.

En cuanto a la unicidad, consideramos que su cumplimiento es posible solamente cuando se emplean métodos “duros” que prescindan de toda posibilidad de apreciaciones personales, ya sea en la determinación de los ponderadores como también en la selección de procedimientos de agregación. Un método que esté preestablecido y no permita integrar la discrecionalidad del investigador ni tampoco la participación de informantes calificados y expertos, atenta no sólo contra la posibilidad de ponderar la calidad relativa de los datos sino también de incorporar la comprensión de la realidad en la elaboración del índice. Los métodos supuestamente objetivos no son neutrales, dado que en su elaboración se encuentra la “carga” ideológica de cómo se relacionan los datos, lo que implica una forma de percibir el funcionamiento de la propia sociedad. En definitiva, la existencia de unicidad puede resultar en el no cumplimiento de la validez interna y externa del índice, lo que atenta no sólo contra la comprensión de la realidad, sino también del avance del conocimiento científico.

Nuestra propuesta es cambiar la unicidad por la *pluralidad*, en el sentido que ante una determinada situación social se puedan obtener diferentes resultados, dependiendo de las distintas apreciaciones que se pueden realizar en cada una de las etapas de elaboración. Así por ejemplo, a través del cálculo de la media ponderada, uno de los métodos de agregación que analizaremos, se pueden conseguir distintos resultados de acuerdo a los ponderadores utilizados.

Por último, la propiedad de exhaustividad, hace referencia a la maximización en el uso de un tipo de información que es puramente estadística y que surge de relaciones lineales entre variables. Es lógico que un método “no puede saber” si un indicador posee la misma información del tipo



conceptual que otro, o si posee sesgos propios de su recolección o cálculo, sino solamente si las relaciones estadísticas entre indicadores presentan variabilidades ya explicadas en algunos de éstos. Por ende, destacamos el carácter relativo de lo que implica la información estadística, y en su remplazo proponemos estudiar la *información operativo-conceptual* del índice, dado que resulta más significativa desde un punto de vista metodológico.

En virtud de ello, reformulamos la propiedad como *exhaustividad conceptual*. Para verificarla no se necesita un método matemático-estadístico, sino el pleno conocimiento por parte del investigador del concepto teórico y de los indicadores disponibles, de tal forma que se obtenga una medida válida. Y tampoco se comprueba una sola vez, sino en dos momentos del proceso de elaboración del índice sintético, siendo el primero una condición necesaria para que se dé el segundo: 1- cuando se seleccionan y elaboran los indicadores simples para cada dimensión, de tal manera que no dupliquen la información operativo-conceptual que se encuentra en las otras dimensiones, y 2- con el método de ponderación y/o agregación empleado, de forma que el procedimiento no distorsione el constructo teórico, sino más bien que lo aproxime para obtener una medida válida. Esto último no significa que un método cumpla o no con la exhaustividad, sino que permita que no se pierda la información conceptual de los indicadores simples o parciales. Si lo primero está relacionado con la búsqueda de una operacionalización que sea válida para cada indicador, entonces con lo segundo, se pretende que esa validez no se debilite por el o los métodos de estandarización, ponderación y agregación.

En definitiva, las propiedades originales pueden ser exigibles desde el punto de vista matemático, pero no lo son desde una concepción metodológica-conceptual como la explicitada al comienzo de esta ponencia. En particular, las propiedades 3 y 7 presuponen que las reglas de la estadística se corresponden con las de la realidad, lo que implica la existencia de un isomorfismo extremo con la estructura de cálculo del índice sintético. Nuestra reformulación da cuenta del carácter relativo que implica el proceso de operacionalización en las Ciencias Sociales.

3. ANALISIS DE LOS METODOS DE AGREGACION/PONDERACION

El siguiente análisis lo centramos en aquellas técnicas más relevantes y difundidos por su aplicabilidad práctica: la media ponderada, el análisis de componentes principales –propuesto por Pearson en 1901-, la distancia P_2 (Pena Trapero, 1977) y la agregación de conjuntos difusos (Zadeh, 1965). Los cuatro métodos poseen características diferenciales. Algunos son más rígidos en su cálculo, dado que no permiten incorporar apreciaciones personales y suponen ciertas relaciones

entre los indicadores, principalmente lineales. Otros poseen un mayor grado de discrecionalidad que permite al investigador elegir los pesos o bien la forma de agregarlos. No obstante, ningún método es neutral a los juicios personales, dado que la sola elección de uno de los métodos implica preferencias individuales. A continuación realizamos una exposición sintética de cada uno y luego verificaremos el cumplimiento de las propiedades reformuladas en el apartado anterior.

3.1. Media ponderada –MP–.

Este es el único método de los que presentamos que realiza solamente la agregación, con lo cual primero hay que estandarizar las variables y luego decidir y aplicar los factores de ponderación. Se puede efectuar de dos formas, como la media aritmética ponderada –MP–, o bien, como la media geométrica ponderada –GP–. La primera es la más utilizada y sencilla en su elaboración (PNUD, 1990-2009; Mirabella de Sant, 2000; Distaso, 2007), y la segunda menos difundida (PNUD, 2010-2014):

$$MP_j = \sum_1^n w_{ij} I_{ij} = w_{1j} I_{1j} + \dots + w_{nj} I_{nj} \quad GP_j = \prod_{i=1}^p (I_{ij})^{w_i} = \left[(I_{1j})^{w_1} \dots (I_{pj})^{w_p} \right]$$

Con:

$$\sum_1^n w_i = 1 \quad \text{y} \quad 0 < w_{ij} < 1; \forall j, 1 \leq j \leq N_{país}$$

Donde:- el subíndice j corresponde al país o ámbito geográfico

- el subíndice i corresponde al indicador perteneciente a la dimensión i-ésima
- w_{ij} es el ponderador del indicador i-ésimo
- I_{ij} es el indicador simple i del país j

Es común, en aquellos casos que el ponderador w_{ij} sea el mismo para todos los indicadores, confundirlo con un método neutral o sin ponderación. Esta igualdad de pesos implica el reconocimiento de un estatuto de igualdad para todos los indicadores, que podría ser el resultado de un conocimiento insuficiente de las relaciones de causalidad, o la ignorancia sobre el modelo correcto para aplicar, o incluso de la falta de consenso sobre soluciones alternativas (Saltelli et al, 2007: 13). No obstante, en muchas ocasiones puede ser conveniente usar ponderadores idénticos cuando las variables estén altamente correlacionadas entre sí, sin que ello implique la presencia de redundancia explicativa puesto que esos elementos correlacionados explican aspectos diferentes de fenómeno en cuestión (Schuschny y Soto, 2009).



Uno de los trabajos pioneros que utiliza esta metodología estuvo dirigido al cálculo de indicadores sintéticos de infraestructuras (Biehl, 1986). En su elaboración parte de indicadores parciales por categorías -transporte, sanidad, educación, etc.- como media aritmética de las variables que aportan información de cada una. Luego construye el indicador sintético de infraestructuras calculando la media geométrica de todos ellos. La elección de la media aritmética y de la media geométrica se justifica por el mayor o menor grado de sustituibilidad, respectivamente.

En definitiva, por ser valores promedio, cumplen con las propiedades de existencia y determinación, monotonía, homogeneidad y transitividad. Asimismo cumple con la propiedad reformulada de pluralidad, dado que la posibilidad de optar por un criterio de ponderación genera que la medida no sea una sola para una situación dada -unicidad-, lo que refuerza la validez interna y permite la inclusión de apreciaciones de expertos. Esto permite a su vez, que se pueda cumplir con la propiedad de exhaustividad conceptual. Cabe destacar que el método no contempla ningún cálculo para eliminar información estadística redundante -propiedad original de exhaustividad-, que como ya explicitamos, ello no reduce la bondad del índice obtenido. Por el contrario, el método es válido para medir niveles de bienestar, dado que agrega todas las dimensiones que componen al concepto elegido, lo que también facilita su interpretación; cuestión que juzgamos muy importante.

3.2. Análisis de componentes principales –ACP–.

El objetivo del ACP es explicar la mayor parte de la variabilidad total observada de un conjunto de variables a través del menor número de componentes posibles (Uriel Jimenez, 1985). Esto se logra transformando las variables originales en un conjunto de variables ortogonales de varianza máxima -componentes principales-, que están relacionadas linealmente con las primeras. El requisito es que los indicadores estén correlacionados entre sí. De lo contrario será poco lo que se pueda hacer, ya que el número de componentes no se verá sustancialmente reducido (OECD, 2008).

Asimismo, se recomienda partir de un número de casos suficientemente grande. Ello supone la adopción de alguna de las siguientes reglas empíricas de relación entre número de casos y cantidad de variables respectivamente: 3 a 1, 5 a 1, 10 a 1 o, la más ambiciosa, 100 a 1 (Saisana y Tarántola, 2002: 12-15; OECD, 2008: 63-72).

El método es versátil dado que se lo puede utilizar en forma conjunta o independiente para la selección, estimación de pesos y/o agrupación de índices simples, particularmente en los casos en los que no existe un criterio de consenso entre los especialistas sobre la importancia relativa de cada componente (Domínguez Serrano et al, 2011). Las diversas aplicaciones que se encuentran en la

bibliografía demuestran la mencionada versatilidad dado que permite obtener los valores de un indicador sintético a través de diferentes modalidades². Los dos grupos de procedimientos principales son: 1) formulación de una escala aditiva y, 2) utilización de los valores de las componentes principales seleccionadas.

1) En este caso el objetivo es seleccionar los indicadores representativos del sistema original, tomando como criterio las correlaciones con la componente principal. Por ende, no se utilizan los valores de las componentes ni tampoco se deduce algún tipo de ponderación. Esto permite seleccionar los indicadores del sistema inicial más representativos para la medición, descartándose el resto. Luego, se procede a la definición de una variable representativa de cada componente principal a partir de una combinación lineal de sus indicadores suplentes –media aritmética simple o simplemente suma-. Por último, el indicador sintético se obtiene como el promedio simple de las variables representativas (Mcintyre et al., 2002; Liou et al., 2004).

2) El segundo grupo de procedimientos incluye dos formas diferentes de proceder: obteniendo una medida única de forma directa, o bien, ponderando los componentes principales. La primera puede ser a partir de los valores de la componente principal uno (INEGI, 1990; Lai, 2000; Vyas y Kumaranayake, 2006), o bien de la componente que mejor permita realizar dicha evaluación.

En cambio, la segunda forma tiene una doble función: limitar la cantidad de componentes principales y estimar sus ponderadores. La cantidad de componentes que integran la medida final, se delimitan a través de distintos criterios: el de raíz latente (Kaiser, 1958), contraste de caída (Catell, 1965) y porcentaje de varianza explicada (Schuschny y Soto, 2009) que determinan el umbral a partir del cual la varianza explicada adicional no aporta mucho más.

Luego, los pesos se calculan a través de la cuantía de los autovalores asociados a cada componente - λ_h -, o lo que es lo mismo, la varianza de los valores que adopta cada componente para cada territorio (Castro Bonaño, 2002). Una vez determinadas las componentes principales más significativas y sus pesos, se agregan a través de la media aritmética ponderada por la cuantía de los autovalores (Peters y Butler, 1970, Wubneh, 1987; Castro Bonaño, 2002), es decir:

$$ACP_i = \frac{\sum_{h=1}^m Z_{hi} \cdot \sqrt{\lambda_h}}{\sum_{h=1}^m \sqrt{\lambda_h}} ; i= 1, 2, \dots, 29.$$

² Para profundizar el estudio de estas dos modalidades se recomienda la lectura en Domínguez Serrano et al (2011).

Donde:

- Z_{hi} es la puntuación del componente h-ésimo para el caso i-ésimo, y
- λ_h el autovalor para dicho componente.

Es decir, esta segunda forma de cálculo realiza los siguientes algoritmos. En un primer momento, se consideran los componentes que son funciones lineales de los indicadores que aportan mayor variabilidad al sistema, con lo cual se minimiza la contribución de los indicadores menos correlacionados con el resto. Esto puede provocar que se pierda cierta cantidad de información valiosa relacionada con el objetivo de medir el nivel de bienestar social, dado que las correlaciones -que son la base de su cálculo- pueden no necesariamente representar la “influencia real” de las componentes sobre el fenómeno que quiere medirse, sino un efecto estadístico espurio (Schuschny y Soto, 2009). Por ende, la propiedad de exhaustividad conceptual reformulada no se cumple.

En segundo lugar, los ponderadores utilizados son atenuados a través del cálculo de su raíz cuadrada. De esta manera, cada componente seleccionada es introducida en el índice sintético con un peso geoméricamente menor al porcentaje de varianza explicada por el autovalor. En otras palabras, las puntuaciones de los componentes que explican una mayor variabilidad son incluidas en la medida final con un peso menor, y viceversa. Esto resulta en parte contradictorio, dado que con esta fórmula se intenta reducir el impacto de aquella/s variable/s ortogonal/es que el sistema generó como principal/es.

Si bien se le atribuye una cierta neutralidad a este proceso debido a que las ponderaciones no las determina el analista sino el propio método, hay que tener presente que desde el momento en que el analista adopta por aplicar un criterio o procedimiento de los ya enumerados, tal neutralidad resulta aparente.

La versatilidad del ACP atenta contra las propiedades de monotonía y homogeneidad, pero lo favorece en cuanto a la pluralidad, dado que los resultados dependen de los procedimientos elegidos (Pena Trapero, 2009: 316). Asimismo, la transitividad puede verse afectada en los casos que se modifique la matriz inicial de indicadores, dado que se puede afectar a las ordenaciones finales en las puntuaciones de los factores.

Por último, consideramos que los resultados obtenidos generan otros inconvenientes. Por un lado, los componentes obtenidas de la transformación lineal no son directamente interpretables empíricamente, dado que no constituyen un índice per se. Éstas quedan definidas a partir de la combinación lineal de variables y por ello es discutible que su valor pueda ser asimilado a un

determinado nivel de bienestar social. Es decir, la clasificación que puede efectuarse de las unidades de análisis es ordinal a pesar de obtener valores numéricos en los resultados (OECD, 2008).

3.3. Métodos basados en el concepto de distancia: la DP₂.

Este tipo de medidas se basan en la comparación por diferencia, en términos absolutos o cuadráticos, entre cada indicador perteneciente a unidades territoriales distintas y una base de referencia considerada para cada componente. Su particularidad es que a diferencia de los métodos anteriores, los algoritmos de cálculo no dejan ningún margen de discrecionalidad al investigador para participar en el procedimiento de estimación, y a su vez, su cálculo integra todas las etapas de estandarización, ponderación y agregación en una sola fórmula. Estos aspectos redundan en una serie de limitaciones a las cuales arribamos en este análisis.

Uno de los indicadores de distancia más relevante es la F de Frechet (Ivanovic, 1973), que se define como la sumatoria de las distancias relativas de cada indicador simple:

$$F = \sum_i^n \frac{d_i}{\sigma_i} ; \quad i = 1, \dots, n$$

Donde: - $d_i = |x_{ji} - x_i^*|$

- x_{ji} el estado en que se encuentra el componente i -ésimo en la unidad territorial j .

- x_i^* es el estado o norma de referencia del indicador i

- σ_i es la desviación típica de los valores tomados por el componente i -ésimo.

En este caso, la σ_i permite no sólo expresar el indicador parcial del componente i -ésimo en unidades abstractas, sino que a su vez, cumple la función de ponderador de d_i , de forma que la contribución de d_i al indicador sintético es inversamente proporcional a σ_i . Este tipo de estandarización-ponderación integrada posee ciertas limitaciones dado que al expresar la distancia en términos relativos a la variabilidad del resto de los espacios territoriales, puede traer sesgos en la evaluación del nivel de bienestar (Actis Di Pasquale, 2013).

Basado en la Distancia-F se encuentra el Índice Sintético DP₂ de Pena Trapero, que logra eliminar totalmente la doble información estadística a través de un factor corrector, lo cual lo convierte en una medida muy difundida por aquellos autores que ven en esta propiedad una ventaja para dar cuenta de la bondad de los resultados obtenidos (Pena Trapero, 1977). En el apartado 2 ya explicitamos nuestros argumentos acerca de las limitaciones que ello puede provocar.

Se define como:

$$DP_2 = \sum_i^n \frac{d_i}{\sigma_i} (1 - R^2_{i,i-1,i-2,\dots,1})$$

Donde $R^2_{i,i-1,\dots,1}$ es un coeficiente de determinación denominado coeficiente de correlación lineal múltiple al cuadrado en la regresión de x_i sobre $x_{i-1}, x_{i-2}, \dots, x_1$; y expresa la parte de varianza o variación de x_i explicada linealmente por las variables sobre $x_{i-1}, x_{i-2}, \dots, x_1$. Este coeficiente es un número abstracto independiente de las unidades de medida en que vengan expresados los distintos componentes. $R^2_1 = 0$, porque la primera componente aporta toda la información al no existir un componente previo. Por eso su ponderación es la unidad.

El procedimiento consta de realizar un método iterativo donde primero se jerarquizan los indicadores parciales en base a la correlación simple entre éstos y la Distancia F –como solución inicial- y luego se calculan la distancia P_2 , repitiendo el proceso hasta lograr la convergencia.

Este índice sintético verifica las propiedades matemáticas originales (Pena Trapero, 1977). Sin embargo, de las propiedades reformuladas, no cumple con la exhaustividad conceptual y la pluralidad. Respecto a la primera, al aplicar el factor corrector que permite eliminar la información estadística ya contenida en los indicadores simples precedentes, no se estaría cumpliendo con esta propiedad. Hay que tener presente que el tipo de información que se logra eliminar es la que surge de relaciones lineales entre variables, las cuales son puramente estadísticas. Los resultados de diversas aplicaciones son claros ejemplos de nuestra argumentación, que ubican al indicador “porcentaje de hogares sin teléfono” como la primera (Sánchez Domínguez y Rodríguez Ferrero, 2003) o la segunda variable más importante (Zarzosa Espina, 1996: 180) para explicar el bienestar social, incluso por encima de los componentes de empleo y salud.

Asimismo, producto de este proceso de depuración, las distancias obtenidas no son directamente interpretables empíricamente, dado que no constituyen un indicador o índice del nivel de bienestar, sino una distancia relativa refinada estadísticamente.

En cuanto a la propiedad reformulada de pluralidad, no puede ser satisfecha por esta medida dado que para un conjunto determinado de datos se obtendrá un único resultado, omitiendo toda apreciación personal en la selección de métodos de estandarización y ponderadores. Aquellos autores que adhieren a las propiedades matemáticas originales, aluden que el índice es “neutral”. Un cálculo preestablecido que no admita aportaciones respecto a las diferentes calidades de los datos, la comprensión del medio social a investigar y los aportes de expertos e informantes calificados, atenta contra la validez interna de la medida. En este sentido, el índice DP_2 presenta un fuerte isomorfismo, dado que en su construcción se alude implícitamente a que lo válido es lo que determina la estructura matemática, y que esta representa la estructura de la realidad social.



3.4. Agregación de conjuntos difusos –ACD-

La Teoría de los Conjuntos Difusos –TCD o FST por su sigla en inglés Fuzzy Sets Theory- se ocupa de modelar situaciones en las cuales los instrumentos tradicionales no conducen a resultados satisfactorios debido a aspectos como la imprecisión o complejidad en la definición de los conceptos implicados³ (Zadeh, 1965), tal como el constructo de bienestar social.

A través de la idea clásica de conjunto, la TCD traduce este tipo de problemas a uno en términos de conceptos difusos –incertidumbre-, asociada al evento en sí mismo y no a la ocurrencia de un evento, como sería el caso de la probabilidad. En este sentido, la diferencia sustancial respecto a los tres métodos anteriores es que la TCD busca medir el grado de verdad de que una expresión sea cierta –“en la región x hay un alto nivel de bienestar social”-, lo cual nos permite crear un ranking de distintos territorios según el grado de verdad de dicha expresión.

La lógica difusa intenta reducir la brecha existente entre teoría y empiria, a través de modelos lingüísticos en forma de predicados que incorporan el conocimiento de especialistas, y luego transforma ese conocimiento en algoritmos de cálculo. Es decir, que traslada sentencias elaboradas en “lenguaje natural” a un “lenguaje matemático formal” (Meschino, Andrade y Ballarin, 2008: 2).

Si bien en los análisis del bienestar y desarrollo ya se han efectuado algunas aplicaciones de esta metodología (Cornelissen et al, 2001; Castro Bonaño, 2002; Chiappero, 2006), aún no se encuentra muy difundida en las Ciencias Sociales. En virtud de ello, realizamos a continuación una breve introducción a los conceptos clave. Al mismo tiempo, se destaca el posible cumplimiento de las propiedades exigibles a los índices sintéticos.

Algunos elementos preliminares de la TCD⁴

En primer lugar se elabora lo que se denomina *predicado* o proposición, es decir, lo que se afirma o niega de un objeto. Un predicado clásico, es aquél que al aplicarlo a los elementos de un universo, lo divide en dos subconjuntos: el de los elementos que verifican dicho predicado, y el de los que no lo verifican. Los predicados difusos, son aquellos que al aplicarlos a los elementos de un universo, no lo dividen perfectamente en dos subconjuntos., sino que permite cuantificar niveles de

³ Zadeh propuso la lógica difusa como disciplina en la década de 1970. En esta combina conceptos de la lógica y de los conjuntos de Lukasiewicz mediante la definición de grados de pertenencia. Las principales aplicaciones de esta teoría son cuando: la información, el concepto a medir o las reglas de decisión son imprecisas, o bien se desconocen los mecanismos internos del sistema.

⁴ La exposición de los conceptos preliminares está basada en Tanaka (1997) y Trillas y Cubillo (1999). Los ejemplos son de elaboración propia.



verdad entre 0 y 1. De esta manera, intenta aproximar los conceptos teóricos a la realidad por medio del valor de verdad de ciertas *proposiciones al conjunto difuso*.

Lo anterior se puede efectuar a través de una doble interpretación: a) el grado de verdad en que la expresión “x pertenece a A” sea cierta; o bien, b) el grado de pertenencia del elemento x al conjunto A. En este trabajo efectuaremos interpretaciones del primer tipo, es decir, la cuantificación del grado de verdad de predicados lógicos.

La *función de pertenencia* de un conjunto nos indica el grado en que cada elemento de un universo dado pertenece a dicho conjunto. Es decir, relaciona la variable con un valor numérico entre 0 y 1 que indica el grado en que la variable satisface el predicado. Hay una infinidad de formas posibles de funciones de pertenencia, entre las que se destacan como más comunes la triangular, trapezoidal, función-S y Gaussiana (Tanaka, 1997; Trillas y Cubillo, 1999), a partir de las que se pueden efectuar derivaciones modificando ciertos parámetros.

En muchas ocasiones puede no haber consenso respecto al grado de verdad en que el conjunto difuso se asocia a los valores de X con un cierto atributo. En estos casos, algunos autores proponen acotar el valor entre dos funciones de pertenencia, para dar lugar a un intervalo de pertenencia. A esto se lo denomina Lógica Difusa tipo 2 de intervalos (Zadeh, 1975; Mendel, 2007), y será la que aplicaremos en la medición del valor de verdad de los predicados que aludan al bienestar social.

Por lo tanto, al ser una construcción que es totalmente subjetiva y dependiente del problema concreto a estudiar, las funciones de pertenencia cumplen la propiedad de pluralidad. Por otra parte, no cumplen con la invariancia de origen, pero sí con la correspondiente a escala.

Operaciones entre conjuntos difusos

De la misma manera que se realizan operaciones con los conjuntos clásicos, se definen tres operaciones básicas para los conjuntos difusos: intersección \cap , unión \cup y complemento \neg (Zadeh, 1965). Para modelar la intersección y la unión hay una gran variedad de operadores que se denominan T-Normas y S-Normas respectivamente. Todos ellos generalizan las operaciones tradicionales para poder ser efectuadas utilizando valores de verdad en el intervalo $[0, 1]$. Sin embargo, presentan algunas limitaciones en cuanto a la validez como operador en aquellos estudios en que se obtienen formulaciones complejas y que derivan en predicados compuestos. Esto se debe

a que los valores de verdad obtenidos en los predicados compuestos deben poseer la sensibilidad⁵ a los cambios de los valores de verdad de los predicados básicos que lo integran.

Para darle solución a este problema, se ha desarrollado la Lógica Difusa Compensatoria, que es un modelo lógico multivalente que renuncia a varios axiomas clásicos para lograr un sistema idempotente y sensible, al permitir la compensación de los predicados (Meschino, Andrade y Ballarin, 2008; Bouchet et al, 2011). De esta manera, un aumento o disminución del valor de verdad de la conjunción provocado por el cambio del valor de verdad de una de sus componentes, puede ser “compensado” con la correspondiente disminución o aumento de la otra. Esto significa que la propiedad de homogeneidad no siempre se cumple en la ACD, ya que dependerá de los operadores y de los supuestos que lo sustentan, y por ende, iguales variaciones de los componentes, no se traducen en variaciones equiporcionales en la medida sintética.

En particular, nos basamos en la conjunción de Lógica Difusa Compensatoria basada en la Media Aritmética (Bouchet et al, 2011)⁶, que se define como:

$$C: [0,1]^n \rightarrow [0,1], \text{ está definida por:}$$

$$C(u_1, \dots, u_n) = \left[\min(u_1, \dots, u_n) \cdot \frac{1}{n} \sum_{i=1}^n u_i \right]^{\frac{1}{2}}$$

Donde:

- C es la conjunción que se aplica a los valores de pertenencia;
- n la cantidad de conjuntos difusos;
- u_n el valor de pertenencia correspondiente al conjunto “n”;
- $\min(u_1, \dots, u_n)$, representa el mínimo valor de pertenencia;

Al utilizar funciones de pertenencia tipo 2 de intervalos (Zadeh, 1975; Mendel, 2007), los valores de pertenencia están definidos por intervalos de pertenencia. En estos casos, los operadores se aplican individualmente entre los extremos de los intervalos. Por ejemplo, para el caso de una conjunción entre dos predicados simples:

$$P(X) = P_1(X) \wedge P_2(X) = [a_1, b_1] \wedge [a_2, b_2] = [C(a_1, a_2), C(b_1, b_2)]$$

Donde:

⁵ Cuando se habla de sensibilidad en el contexto de la lógica difusa compensatoria, se hace referencia a que el resultado de la intersección no varíe tan bruscamente ante la variación de algunas de las componentes.

⁶ De acuerdo a estos autores, los operadores de Lógica Difusa Compensatoria son considerados como tales, siempre y cuando verifiquen ciertas propiedades (Bouchet et al, 2011: 232). Hasta el momento se han desarrollado muy pocos operadores, siendo el mencionado un avance por parte de este grupo investigadores.

- $P(X)$ es un predicado difuso del ámbito geográfico X que surge de la conjunción
- $P_1(X)$ es el predicado difuso del conjunto 1 del ámbito geográfico X
- $P_2(X)$ es el predicado difuso del conjunto 2 del ámbito geográfico X
- a_1 y a_2 son los valores mínimos de los intervalos de pertenencia correspondientes al conjunto 1 y 2;
- b_1 y b_2 son los valores máximos de los intervalos de pertenencia correspondientes al conjunto 1 y 2;
- C es la conjunción definida anteriormente.

Por lo tanto, el operador se aplica individualmente a los extremos de los intervalos, lo que dará como resultado una conjunción de los mínimos y otra conjunción de los máximos, es decir, un intervalo con dos grados de verdad (Zadeh, 1975; Mendel, 2007). $P(X)$ indicará, a través de un intervalo de pertenencia, el grado de verdad del predicado. Para poder establecer un orden entre diferentes ámbitos geográficos cuando se evalúan predicados con este tipo de lógica, debe definirse una manera de establecer el ámbito geográfico que tiene mayor valor de verdad en el predicado que se evalúa. Por esto, para determinar la medida final, utilizamos la siguiente expresión:

Sea $E = [a, b]$, un intervalo de pertenencia que define el grado de verdad de un predicado, una medida de intervalos de pertenencia $\lambda(E) = g(a, b)$, con $g: [0, 1]^2 \rightarrow [0, 1]$, se define como⁷:

$$g(a, b) = \begin{cases} \alpha(1 - |b - a|) + (1 - \alpha) \max\{a, b\} & \text{si } (a, b) \neq (0, 0), \alpha \in [0, 1] \\ 0 & \text{otro caso} \end{cases}$$

Donde:

- a y b son los valores mínimos y máximos respectivamente obtenidos para el predicado;
- α es un parámetro definido entre 0 y 1. En nuestra propuesta tiene el valor de 0,5. Esto se debe a que:
 Si $\alpha = 1$, $\lambda(E)$ mide la incertidumbre del intervalo; si $\alpha = 0$, $\lambda(E)$ permite evaluar el grado de verdad del intervalo; en cambio si α entre 0 y 1 permiten tomar en consideración tanto la incertidumbre del intervalo como sus grados de verdad.

En definitiva, y a diferencia de la DP_2 , este método presenta un alto grado de discrecionalidad que permite la inclusión de apreciaciones en todo el proceso, con lo cual la ACD se posiciona en el otro extremo. Al ser una teoría basada en la teoría de los conjuntos clásica y que aplica operadores lógicos, no se efectúa ningún tipo de corrección que elimine la información correlacionada. Como se ha venido argumentando, esto puede significar una mayor validez de la medida obtenida.



3.5. Grado de cumplimiento de las propiedades deseables reformuladas

Los cuatro procedimientos de ponderación y agregación presentados parten de presupuestos bien diferenciados. Ello nos permite obtener medidas que van desde aquella más simple en su cálculo e interpretación -MP-, hasta la que supone un cambio de paradigma -ACD-. Los otros dos métodos -ACP y DP₂- están basados en criterios estadísticos, en particular, respecto a la correlación entre los indicadores, lo cual presenta ciertas restricciones al momento de sintetizar la información.

La existencia y determinación, monotonía y transitividad deben darse de manera perentoria, debido a que son condiciones necesarias para obtener medidas ordinales y poder comparar el bienestar de los territorios. El ACP no cumple con estas dos últimas, lo que lo coloca en una desventaja adicional respecto al resto de los procedimientos. En cambio, la propiedad de homogeneidad no siempre debe ser exigible, ya que dependerá del efecto potencial que tengan las dimensiones que componen al constructo que se está midiendo. Esta propiedad la cumplen la MP y DP₂. Respecto a la pluralidad de posibles estimaciones en la medición del bienestar social en oposición al resultado único -como lo hace sólo la DP₂-, permite manejar cierta discrecionalidad en el cálculo, lo cual es sumamente necesario para acortar la brecha entre teoría y empiria.

En cuanto a la exhaustividad conceptual, solo puede ser satisfecha por la MP y el ACD, ya que permiten incorporar apreciaciones en el cálculo de los ponderadores o modificadores, respectivamente. Un método “neutral” puede llevarnos a resultados erróneos en cuanto a cuáles son las variables de mayor jerarquía en la medición. En definitiva, la MP cumple con las seis propiedades, la ACD con cinco, la DP₂ con cuatro y el ACP con tan solo dos -Cuadro 1-.

En el próximo apartado agregaremos una propuesta de indicadores de bienestar social (Actis Di Pasquale, 2013) a través de los cuatro procedimientos. De esta manera logramos un doble propósito: primero, testear las posibles diferencias de resultados de acuerdo a las interpretaciones teóricas vertidas en este capítulo; y segundo, encontrar la medida con mayor validez externa.

⁷ Esta medida fue desarrollada por los integrantes del Grupo de Procesamiento Digital de Imágenes (Facultad de Ingeniería, UNMDP).

Cuadro 1 - Comparación de métodos de agregación/ponderación

PROPIEDAD	METODO DE AGREGACION/PONDERACION			
	MP	ACP	DP ₂	ACD
Existencia y determinación	SI	SI	SI	SI
Monotonía	SI	NO	SI	SI
Pluralidad	SI	SI.	NO -unicidad-	SI.
Homogeneidad	SI	NO -ordinal-	SI	NO
Transitividad	SI	NO	SI	SI
Exhaustividad conceptual	SI	NO, depura correlaciones.	NO	SI

Fuente: elaboración propia

4. VALIDACIÓN EMPÍRICA DE LOS MÉTODOS DE AGREGACIÓN/PONDERACIÓN

La propuesta de indicadores fue realizada en el marco de una perspectiva metodológica-conceptual. En este sentido, se comienza por definir al bienestar social como “*una especie de orden social, en el cual todas las personas alcanzan logros –específicos para su edad- que favorecen tanto el desarrollo individual como social*” (Actis Di Pasquale, 2013). Esos logros son los siguientes: 1- Gozar de buena salud; 2- Gozar de seguridad ciudadana; 3- Alcanzar un nivel educativo apropiado; 4- Poseer un trabajo decente; 5-Habitar un hogar digno parte de la definición. El total de indicadores -índices simples- son trece y como algunas dimensiones contienen más de un índice simple, se conformaron seis índices parciales -Cuadro 2-. Las estimaciones se realizan para el año 2005 a nivel de los 29 aglomerados/departamentos en todos los indicadores.

Cuadro 2 - Índices simples y parciales propuestos

LOGRO DE BIENESTAR SOCIAL	INDICES PARCIALES	INDICES SIMPLES	FUENTE DE DATOS
I- Gozar de buena salud	1) Índice de salud	1- Índice de Años Potenciales de Vida Perdidos	DEIS - Ministerio de Salud
II- Gozar de seguridad ciudadana	2) Índice de seguridad ciudadana	2- Índice de Homicidios Dolosos	DNPC - Ministerio de Justicia y Derechos Humanos
		3- Índice de Homicidios Culposos en Accidentes de Tránsito	
III- Alcanzar un nivel educativo adecuado	3) Índice de nivel educativo total	4- Índice de Educación 6 a 17 años	Encuesta Permanente de Hogares - INDEC
		5- Índice de Educación de 18 a 24 años	
		6- Índice de Educación mayores 25 años	
IV- Poseer un trabajo decente	4) Índice de Trabajo Decente	7- Índice de Trabajo Decente	
V- Habitar un hogar adecuado	5) Índice de condiciones habitacionales	8- Índice de techo-piso	
		9- Índice de acceso al agua	
		10- Índice sobre disposición de excretas y saneamiento	
		11- Índice de personas por dormitorio	
	12 Índice sobre seguridad jurídica en la tenencia		
6) Índice de ingreso relativo	13- Índice de ingreso relativo		

Fuente: elaboración propia

Es decir, que para realizar la síntesis final partimos de una matriz de datos que contiene el valor medio de estos seis índices parciales para cada uno de los 29 aglomerados. A través de cada uno de los cuatro métodos de agregación -la media ponderada, el análisis de componentes principales, la distancia P_2 de Pena Trapero y la agregación de conjuntos difusos- generamos igual número de rankings de aglomerados.

4.1. Algunos aspectos específicos de las estimaciones

a) *Media ponderada –MP-*: para determinar el valor de los ponderadores, consideramos las limitaciones de las fuentes de datos, y a su vez, la naturaleza de los indicadores. En particular, a los índices de salud, seguridad y trabajo decente se les asignó un peso menor. A los dos primeros debido a que dan cuenta solamente de la peor situación en la escala de bienestar, las muertes. Al último, porque incluye solamente a los asalariados. Por lo tanto, decidimos asignarle a salud y seguridad un peso del 25%, a trabajo decente un 70% y al resto, los introducimos con el 100%.

Los resultados de esta síntesis de información estadística nos permiten evaluar no sólo el ranking sino también el valor numérico alcanzado. En este trabajo presentamos sólo el ranking.

b) *Análisis de componentes principales –ACP-*. La estructura de correlaciones dio alta entre los índices de trabajo decente, condiciones habitacionales e ingresos -entre 0,75 y 0,91-.

Se consiguió reducir a dos componentes, considerando en forma conjunta los criterios de raíz latente, contraste de caída y porcentaje de varianza explicada -los dos primeros componentes ordenados llegan a explicar el 72,59% de la varianza total observada-.

Para el cálculo del índice, utilizamos la ponderación de las componentes principales por la cuantía de los autovalores, dado que estaríamos explicando un mayor porcentaje de la variabilidad total que al hacerlo con un solo componente -ver fórmula en apartado 3.2-. Como ya se explicó, esto permite obtener solamente un ranking de aglomerados.

c) *Distancia P_2 -DP2-*. El cálculo del índice sintético DP_2 requiere de un proceso iterativo que consta de varias etapas. Con el fin de sistematizarlo y poder realizar diferentes pruebas, elaboramos un código en el programa informático Matlab que permite introducir una matriz de datos de n variables y m unidades geográficas⁸. La convergencia la conseguimos en la iteración 15. Mediante

⁸ El código en Matlab fue elaborado por el Dr. Juan Ignacio Pastore -Grupo de Procesamiento Digital de Imágenes, Facultad de Ingeniería, UNMdP-, a quien agradezco por su valiosa colaboración.

la aplicación de este procedimiento se obtiene una medida que sólo sirve para rankear aglomerados y medir su distancia relativa, ya que no se puede conocer el nivel de bienestar de cada uno.

d) *Agregación de conjuntos difusos –ACD–*: Los predicados lingüísticos los elaboramos para cada uno de los trece índices iniciales, lo cual define trece conjuntos difusos tipo 2 de intervalos con su respectivo predicado difuso básico. El predicado “*El aglomerado urbano X tiene un nivel alto de bienestar social*” es lógicamente equivalente -tiene el mismo valor de verdad- que el predicado compuesto “*El aglomerado urbano X tiene un nivel alto de salud y una baja tasa de homicidios totales y un alto nivel educativo y trabajo decente y condiciones habitacionales adecuadas y un nivel alto de ingresos*”. Luego se generaron las funciones de pertenencia y seguidamente los operadores de lógica compensatoria definidos en el apartado 3.4⁹. Este procedimiento nos permite obtener un ranking de aglomerados de acuerdo al grado de verdad de las proposiciones.

En similar sentido a los ponderadores utilizados en la media ponderada, se aplicaron modificadores o *hedges* en los conjuntos difusos de salud, seguridad y trabajo decente.

4.2. Resultados preliminares

Los aglomerados que evidencian una mayor variabilidad en la posición alcanzada en los índices sintéticos -Cuadro 3- también presentan ubicaciones muy disímiles en el ranking de cada índice parcial. Por lo tanto, los resultados se van a ver influenciados por las características propias de cada procedimiento, principalmente, la estandarización y ponderación implicada, que le otorgan una importancia o jerarquía diferente a cada índice parcial.

Así por ejemplo, analicemos el caso del aglomerado Río Cuarto, que presenta posiciones muy bajas en salud -28º-, seguridad -24º-, y educación -28º-, y niveles superiores en trabajo decente -12º-, condiciones habitacionales -4º- e ingresos -5º-. Ahora observemos los diferentes resultados en cada método de agregación:

- a) En el cálculo de la MP resulta el 9º aglomerado, dado que los pesos de los dos primeros índices parciales son del 25%, y no generan una gran influencia en el resultado final.
- b) En el ACP, el proceso de normalización implicado genera que aquellos aglomerados con índices parciales por debajo de la media, y que tengan menor desvío típico –salud y

⁹ Agradezco a los integrantes del Grupo de Procesamiento Digital de Imágenes (Facultad de Ingeniería, UNMdP) por la colaboración en el cálculo de los algoritmos correspondientes. Esta estimación resulta de una aplicación de sus aportes a la lógica difusa tipo 2, que la han modelado a través de la elaboración de un *Type 2 Fuzzy Logic Toolbox* para el software Matlab.

educación- obtengan valores negativos en la estandarización interna propia del método. Por ello consigue una peor ubicación en el ranking: 16°.

- c) En la DP₂, dado que la estandarización es a través de los valores absolutos de los desvíos con respecto al valor de referencia, entonces en los aglomerados con índices con las características mencionadas, obtendrán mayores valores. Seguidamente, el factor corrector (1-R²) elimina gran parte de la información estadística contenida por el índice de salud de la DP₂ inicial, y que luego lo posiciona en la primera jerarquía. Por esto último, le otorga un peso mayor en las siguientes DP₂ y la final. Entonces, Río Cuarto queda 6°.
- d) Por último, en el ADP, la posición obtenida -11°- es más cercana a la MP, ya que se utilizan modificadores que disminuyen la importancia del predicado -salud y seguridad- en la relación global de los conjuntos difusos. -Cuadro 3-.

En este sentido, la MP y la ACD aparecen como las menos sensibles a las variaciones que son consecuencia de ponderaciones y estandarizaciones, y esto se debe a que ambos métodos respetan los valores ya estandarizados, y se le asignan pesos o modificadores que son discrecionales por parte del investigador. En cambio, ACP y DP₂ tienen procedimientos de estandarización integrados a su cálculo.

Cuadro 3 –Rankings obtenido con cada método.

AGLOMERADOS	MP	ACP	DP2	ACD	σ^2	Media
Ushuaia - Río Grande	1	2	1	2	0,25	1,50
CABA	2	1	3	1	0,69	1,75
Río Gallegos	3	4	2	9	7,25	4,50
Comodoro Rivadavia - Rada Tilly	4	8	4	3	3,69	4,75
Gran La Plata	5	6	5	4	0,50	5,00
Santa Rosa - Toay	6	3	8	7	3,50	6,00
Bahía Blanca - Cerri	7	5	9	6	2,19	6,75
Mar del Plata - Batán	8	7	7	5	1,19	6,75
Río Cuarto	9	16	6	11	13,25	10,50
Neuquén - Plottier	10	9	12	12	1,69	10,75
Gran Paraná	11	11	14	13	1,69	12,25
Gran Rosario	12	14	10	8	5,00	11,00
Gran Córdoba	13	13	11	10	1,69	11,75
San Luis - El Chorrillo	14	10	15	14	3,69	13,25
Gran Santa Fe	15	18	16	15	1,50	16,00
Gran Mendoza	16	15	13	20	6,50	16,00
Gran Catamarca	17	12	20	19	9,50	17,00
La Rioja	18	17	17	16	0,50	17,00
Partidos del GBA	19	19	18	17	0,69	18,25
Gran San Juan	20	21	19	18	1,25	19,50
Salta	21	20	23	23	1,69	21,75
Corrientes	22	22	28	27	7,69	24,75

Santiago del Estero - La Banda	23	24	22	22	0,69	22,75
Concordia	24	27	21	21	6,19	23,25
Jujuy - Palpalá	25	23	25	25	0,75	24,50
Gran Tucumán - Tafí Viejo	26	26	24	24	1,00	25,00
Posadas	27	28	26	26	0,69	26,75
Formosa	28	29	27	28	0,50	28,00
Gran Resistencia	29	25	29	29	3,00	28,00

Fuente: elaboración propia

Para determinar la validez externa se debe cotejar el instrumento de medición con algún criterio externo. Consideramos que si el criterio externo es un indicador que ya se encuentra validado, no tendría mucho sentido elaborar una nueva medida. En virtud de ello, realizaremos la validación externa a partir de un análisis de correlación entre las cuatro aproximaciones y su promedio. En este sentido, la más correlacionada evidenciaría una validez externa mayor.

Los coeficientes empleados son dos medidas no paramétricas: el coeficiente de correlación *rho* de Spearman y el de concordancia *tau-b* de Kendall. La primera es la versión no paramétrica del coeficiente de correlación de Pearson, que se basa en los rangos de los datos por lo que resulta apropiada para datos ordinales. Y la segunda, es una medida no paramétrica de asociación para variables ordinales o de rangos que tiene en consideración los empates. En ambos casos los valores posibles se encuentran entre -1 y +1.

Los resultados demuestran que todas las asociaciones son significativas al 1%, positivas y elevadas. Las correlaciones de Spearman más altas están dadas entre los rankings obtenidos con MP y cada uno de los otros métodos: con ACP 0,957; con DP₂ 0,970 y con ACD 0,960. Esto da cuenta de cierta superioridad de la MP con respecto al resto de las ordenaciones obtenidas. De hecho, al correlacionar cada índice con el promedio, la MP posee el rho de Spearman más cercano a 1, es decir, 0,995 y 0,996, respectivamente. Luego le siguen, también con correlaciones altas, la DP₂, la ACD y por último el ACP –Cuadro 4-.

Cuadro 4 – Coeficiente de correlación rho de Spearman.

	MP	ACP	DP ₂	ACD	Media de los rankings
MP	1				
ACP	0,957**	1			
DP ₂	0,970**	0,896**	1		
ACD	0,960**	0,908**	0,961**	1	
Media de los rankings	0,995**	0,951**	0,984**	0,970**	1,000

Fuente: elaboración propia

** . Correlation is significant at the 0.01 level (2-tailed).

El coeficiente de concordancia de Kendall confirma las mismas tendencias ya descriptas. La diferencia es que los valores de los coeficientes son más bajos, es rigor, menor a 0,90 en la mayor parte de las relaciones. Esto se debe a que la medida de la similitud de las ordenaciones se refina al tener en cuenta los empates. No obstante, pese a esta depuración, la asociación entre la MP, el promedio de los índices y su mediana es muy elevada, de 0,970 y 0,971 respectivamente -Cuadro 5-

Cuadro 5 – Coeficiente de concordancia tau-b de Kendall.

	MP	ACP	DP ₂	ACD	Media de los rankings
MP	1				
ACP	0,837**	1			
DP ₂	0,872**	0,719**	1		
ACD	0,842**	0,729**	0,892**	1	
Media de los rankings	0,970**	0,817**	0,906**	0,871**	1

Fuente: elaboración propia

** . Correlation is significant at the 0.01 level (2-tailed).

5. CONCLUSIONES

Al reformular las propiedades deseables de los indicadores desde una perspectiva metodológica-conceptual, y luego al comparar el grado de cumplimiento de las mismas en los cuatro métodos de agregación/ponderación seleccionados, la MP presenta ventajas en cuanto al cumplimiento de todas las propiedades reformuladas.

Por otra parte, la medición del bienestar a través de los cuatro métodos nos sirvió para confirmar empíricamente la validez de la MP, dado que permite obtener una medida que se aproxime al fenómeno bajo estudio con mayor grado de precisión. Asimismo, el nivel de correlación del resto de las medidas con respecto al promedio de los índices, está asociado de forma directa con respecto al grado de cumplimiento de las propiedades reformuladas. Recordemos que el ACD cumple con cinco, la DP2 con cuatro, y el ACP con dos, es decir, el método que satisface menos características exigibles.

Asimismo, la MP resulta ser el método más sencillo desde el punto de vista matemático lo que permite efectuar una interpretación del resultado obtenido, y no sólo un ranking de territorios. Esto último también lo cumple en cierta medida la ACD, pero en el resto de los métodos se pierde la posibilidad de efectuar un completo análisis del fenómeno social.



6. BIBLIOGRAFIA

- Actis Di Pasquale, E. (2013): *Bienestar Social, resignificación del concepto y de su operacionalización. Un aporte metodológico aplicado al caso argentino (1988-2005)*, Tesis doctoral. Universidad Nacional de Quilmes. Inédita.
- Biehl, D. (1986): "The contribution of infrastructure to regional development", *Final Report of the Infrastructure Study Group*, Commission of the European Communities, Luxemburgo.
- Bouchet, A., Pastore, J.I., Andrade, R.E., Brun, M. y Ballarin, V. (2011): "Arithmetic mean based compensatory fuzzy logic", *International Journal of Computational Intelligence and Applications*, 10, 2, pp. 231–243.
- Castro Bonaño, J.M. (2002): *Indicadores de Desarrollo Sostenible Urbano. Una Aplicación para Andalucía*, Tesis doctoral, Universidad de Málaga.
- Catell, R.B. (1965): Factor Analysis: an introduction to essentials. *Biometrics*, 21, pp. 190-215.
- Chiappero, E. (2006): "Capability approach and fuzzy set theory: description, aggregation and inference issues", en: Lemmi, A. and Betti, G. (eds.): *Fuzzy Set Approach to Multidimensional Poverty Measurement*. New York, Springer, pp.93–113.
- Cornelissen, A.M.G., van den Berg, J., Koops, W.J., Grossman, M. y Udo, H.M.J. (2001): "Assessment of the contribution of sustainability indicators to sustainable development: a novel approach using fuzzy set theory", *Agriculture, Ecosystems & Environment*, 86, pp. 173–185.
- Distaso, A. (2007): "Well-being and/or quality of life in EU countries through a multidimensional index of sustainability". *Ecological Economics*, 64, 1, pp. 163–180.
- Dominguez Serrano, M., Blancas Peral, F.J., Guerrero Casas, F.M. y González Lozano, M. (2011): "Una revisión crítica para la construcción de indicadores sintéticos", *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 11, Universidad Pablo de Olavide, Sevilla, pp. 41-70.
- INEGI (1990): *Cuaderno de Información para la Planeación*, INEGI, México.
- Kaiser, H.F. (1958): "The varimax criterion for analytic rotation in factor analysis", *Psychometrika*, 23, pp. 187-200.
- Lai, D. (2000): "Temporal Analysis of Human Development Indicators: Principal Component Approach", *Social Indicator Research*, 51, pp. 331–330.
- Liou, S., Lo, S. y Wang, S. (2004): "A Generalized Water Quality Index for Taiwan", *Environmental Monitoring and Assessment*, 96, pp. 35–52.
- McIntyre, D., Muirhead, D. y Gilson, L. (2002), "Geographic Patterns of Deprivation in South Africa: Informing Health Equity Analyses and Public Resource Allocation Strategies", *Health Policy and Planning*, 17, pp. 30–39.
- Mendel, J.M. (2007): "Type-2 fuzzy sets and systems: an overview", *IEEE Computational Intelligence Magazine*, 2, 1, pp. 20-29.



- Meschino, G.J. Andrade, R.E. y Ballarin, V.L. (2008): “A framework for tissue discrimination in Magnetic Resonance Brain images based on predicates analysis and Compensatory Fuzzy Logic”, *IC-Med*, TSI Press, 2, X, 1, pp: 1-16.
- Mirabella de Sant, M.C. (2002): “Diferencias de bienestar entre provincias de Argentina”. *XXXVII reunión de la AAEP*, Tucumán.
- OECD (2008): *Handbook on Constructing Composite Indicators. Methodology and user guide*, European Commission, OECD.
- Pena Trapero, B. (2009): “La medición del Bienestar Social: una revisión crítica”, *Estudios de Economía Aplicada*, 27-2, Universidad de Valladolid, pp. 299-324.
- Pena Trapero, B. (1977): *Problemas de medición del bienestar y conceptos afines. Una aplicación al caso español*. INE. Madrid.
- Peters, W.S. y Butler, J.Q. (1970): “The construction of regional economic indicators by principal components”, *Annals of Regional Science*, 4, pp. 1-14.
- PNUD (1990 a 2014): *Informe sobre Desarrollo Humano*. Programa de Naciones Unidas para el Desarrollo. Mundiprensa Libros. Madrid / Fondo de Cultura Económica. México.
- Saisana, M. y Tarantola, S. (2002): *State of the art Report on Current Methodologies and Practices for Composite Indicator Development*. Report EUR 20408 EN. Joint Research Centre. European Commission. Ispra.
- Saltelli, A., Jesinghaus, J. y Munda, G. (2007): "Well Being Stories", *Contribution to the Experts Workshop at the Beyond GDP Conference*, EU Parliament, Brussels.
- Sánchez Domínguez, M.A. y Rodríguez Ferrero, N. (2003): “El Bienestar Social en los Municipios Andaluces en 1999”, *Revista Asturiana de Economía*, 27, pp. 99-119.
- Schuschny, A. y Soto, H. (2009): *Guía metodológica. Diseño de indicadores compuestos de desarrollo sostenible*. Documento de proyecto. DDSAH/CEPAL. Santiago de Chile.
- Tanaka, K. (1997): *An introduction to fuzzy logic for practical Applications*, Springer, USA.
- Trillas, E. y Cubillo Villanueva, S. (1999): *Primeras Lecciones de Lógica Borrosa*, Facultad de Informática, Universidad Politécnica de Madrid, Madrid.
- Uriel Jiménez, E. (1995), *Análisis de Datos. Series temporales y análisis multivariante*, Editorial AC, Madrid.
- Vyas, S. y Kumaranayake, L. (2006): “Constructing Socio-Economic Status Indices: How to Use Principal Components Analysis”, *Health Policy and Planning*, 21, pp. 459–468.
- Zadeh, L.A. (1973): “Outline of a new approach to the analysis of complex systems and decision processes”, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3, 1, 28-44.
- Zadeh, L.A. (1965): “Fuzzy Sets”, *Information and control*, 8, 3, University of California, pp. 338-353.
- Zarzosa Espina, P. (1996): *Aproximación a la medición del Bienestar Social*. Universidad de Valladolid, Secretariado de Publicaciones, Valladolid.