

# El análisis multivariante aplicado al abastecimiento de Gas Licuado de Petróleo (GLP) en el Perú

BEATRIZ ADANIYA HIGA <sup>1</sup>

RECIBIDO: 30/01/2019 ACEPTADO: 25/04/2019

## RESUMEN

El artículo presenta un breve análisis de la seguridad de suministro de energía, el cual incide principalmente en las variables relacionadas con el abastecimiento de Gas Licuado de Petróleo (GLP) dentro del mercado peruano. Asimismo, a través del uso de data real, se explica e ilustra la aplicación de una de las técnicas multivariantes más utilizadas, el Análisis de Componentes Principales (ACP), cuyo resultado muestra la existencia de correlación entre las variables analizadas.

**Palabras-claves:** Análisis de Componentes Principales; técnica multivariante; abastecimiento de GLP; variables; seguridad energética.

## INTRODUCCIÓN

Este artículo desarrolla un primer análisis de las variables causales involucradas en el abastecimiento de **Gas Licuado de Petróleo (GLP)**, a fin de actuar, prever y contrarrestar cualquier afectación que pueda incidir en el mercado de GLP que impacte la economía de la población. El interés por desarrollar esta investigación es la escasa o nula existencia de estudio alguno en el país que realice un análisis multivariante de los factores para conseguir un abastecimiento permanente de este combustible.

En los primeros meses del 2004 era notorio que la producción de GLP se volvía insuficiente para abastecer la demanda nacional, lo cual hizo necesaria la importación; sin embargo, con el inicio de operaciones de la Planta de Fraccionamiento de Líquidos de Gas Natural de Pluspetrol, el Perú pasó a ser exportador de dicho producto. Actualmente, esta condición está siendo revertida por el acelerado aumento de la demanda de GLP y la lenta ejecución de proyectos de ampliación de capacidad de almacenamiento en plantas de abastecimiento. Ello justifica la elaboración de modelos para pronosticar futuros escenarios en el suministro de GLP; no obstante, factores legales, pendientes de atención por las autoridades del sector, impiden mejoras en la estrategia de abastecimiento de GLP, variables que deberían ser estudiadas dentro de un análisis más profundo.

El término «seguridad energética» es usado en diferentes contextos. Ello evidencia la diferencia de criterios a la hora de enfocarla, lo cual provoca diversas interpretaciones, generando desacuerdos conceptuales. La Agencia Internacional de la Energía (AIE) definió la «seguridad energética» como «un suministro adecuado de energía a un precio razonable» (1985, p. 9). Años después, en 2001, desarrolló nuevas definiciones, concibiéndola como «disponibilidad física de suministros para satisfacer la demanda a un precio dado» (p. 76). Posteriormente, esta misma agencia sostuvo que la «seguridad energética, definida en términos generales, significa suministros de energía adecuados, asequibles y fiables» (AIE, 2007, p. 160) con un enfoque de precios que pudieran ser «competitivos o no demasiado volátiles».

<sup>1</sup> Ingeniera petroquímica por la Universidad Nacional de Ingeniería. MBA y MSc. en Ingeniería de Petróleo y Gas Natural por la misma institución. Es profesora principal de la Escuela Profesional de Ingeniería Petroquímica de la Facultad de Ingeniería de Petróleo, Gas Natural y Petroquímica de la UNI. Asimismo, es jefa de producción y procesamiento de gas natural de la División de Supervisión de Gas Natural de Osinergmin.  
E-mail: badaniya@fip.uni.edu.pe

## METODOLOGÍA

La metodología para la recolección de datos incluyó la exploración del entorno mediante consultas a expertos, análisis de las transacciones comerciales del mercado peruano de GLP, informes emitidos por instituciones y agencias nacionales e internacionales de energía y la determinación de los constructos. Asimismo, la comprobación se realizó mediante un Análisis por Componentes Principales. Se tomó en consideración la evolución temporal del mercado peruano de GLP, pues en estudios posteriores es posible encontrar distintos modelos estructurales que muestren diversos estados para las variables consideradas, lo cual podría determinar la ampliación del estudio a un Análisis Factorial Confirmatorio y modelos que evalúen el efecto del tiempo en el desarrollo del mercado. Los resultados de estos estudios podrían incidir en la necesidad de cambios en la legislación para lograr un mayor ordenamiento del mercado de GLP y el cumplimiento de las obligaciones de parte de las empresas involucradas.

## LA SEGURIDAD ENERGÉTICA Y SU DIFERENCIA CON LA SEGURIDAD DE ABASTECIMIENTO

La denominación «seguridad energética» también es referida a la seguridad de abastecimiento (Kruyt *et al.*, 2009; Löschel *et al.*, 2009). Además, se considera que, en el primer caso, se trata del marco general que incluye a importadores y exportadores; mientras que, en el segundo caso, la seguridad de abastecimiento es considerada como un subtipo dentro de la seguridad energética y, como tal, es tratada por los analistas sopesando el corto y el largo plazo (AIE, 2007; Checchi *et al.*, 2009). En el corto plazo, señala Kruyt *et al.* (2009), la seguridad energética busca mitigar las potenciales interrupciones de suministro o contar con una estrategia de respuesta rápida ante una potencial escasez de energía. Esta estrategia tendría que considerar el reemplazo de los flujos de suministro, el uso de capacidad no utilizada de los productores (un símil de reserva fría, en generación eléctrica) y de las reservas de emergencia (inventarios mínimos, en el caso peruano) o las medidas coyunturales para regular la demanda. La finalidad es reducir el impacto de un desabastecimiento de energía en sectores críticos (AIE, 2001) y restringir la demanda para afrontar la falta física de energía (AIE, 2007), asunto de difícil aplicación en el país.

Si se analiza cuidadosamente los factores principales de la seguridad de abastecimiento energético, como lo es la disponibilidad física en volúmenes suficientes, sin interrupciones y con precios acce-

sibles, es importante destacar que la seguridad de abastecimiento en el largo plazo pretende garantizar la disponibilidad de energía suficiente para un desarrollo económico sostenible. Para ello, la estrategia busca asegurar una mayor eficiencia energética, disminuir la vulnerabilidad, contar con diversas fuentes mediante el uso de energías renovables, nuevos proveedores, así como el desarrollo de medios de transporte con nuevas tecnologías y el aumento de redes interconectadas de energía existentes.

La Commission of the European Communities (1990a, p.16 en Berrah, Feng, Priddle y Wang, 2007) explica que «la seguridad de suministro significa la capacidad de asegurar que las necesidades energéticas básicas futuras pueden ser cubiertas, tanto por medio de unos recursos domésticos adecuados producidos bajo condiciones aceptables económicamente o conservados como reservas estratégicas, así como recurriendo a fuentes externas accesibles y estables, complementadas con las reservas estratégicas, cuando sea apropiado». Años más tarde, se ha relacionado a la seguridad de abastecimiento de energía considerando que «la estrategia de la Unión Europea a largo plazo para la seguridad de abastecimiento energético debe dirigirse a asegurar, por el bienestar de sus ciudadanos y el correcto funcionamiento de la economía, la disponibilidad física ininterrumpida de productos energéticos en el mercado, a un precio asequible para todos los consumidores (particulares e industriales), a la vez que respete las preocupaciones medioambientales y vele por el desarrollo sostenible» (Commission of the European Communities, 2000a, p. 2).

En cuanto a la disponibilidad de energía, de forma permanente y en la cantidad requerida, es importante considerar las condiciones necesarias para permitir la satisfacción de la demanda que coadyuven al desarrollo de las actividades estratégicas, económicas y ordinarias de los países (acciones de defensa, producción, transporte, consumo doméstico, entre otros). Estas condiciones tienen su origen geológico, técnico, económico y geopolítico, ya que se refieren a la disponibilidad de los recursos, el buen funcionamiento de la infraestructura relacionada y la confiabilidad de las empresas y de cada grupo de interés comprometido con el sistema. Como es de esperar, las inversiones juegan un rol importante dentro de este sistema complejo.

## EL ABASTECIMIENTO DE GLP DENTRO DEL CONTEXTO PERUANO

En el contexto peruano, el Reglamento para la Comercialización de Gas Licuado de Petróleo (GLP),

aprobado por Decreto Supremo 01-94-EM, regula el aseguramiento de su abastecimiento. En su artículo 8 se establece que toda planta de producción e importador de este combustible está obligada a mantener, como existencia media, un volumen equivalente a quince días del volumen promedio vendido en los últimos seis meses o del volumen promedio importado dentro del mismo periodo, según sea el caso. Sin embargo, un informe del estado situacional de comercialización muestra un análisis estadístico de las ventas a nivel nacional que permite observar la imposibilidad de supervisar el cumplimiento de esta obligación (División Planeamiento y Desarrollo de la Gerencia de Fiscalización de Hidrocarburos Líquidos, 2015). Además, este cargo deberá garantizar el abastecimiento, dado que al calcular el número de días de abastecimiento disponible se encontró una variabilidad de tres a doce días, sin considerar la posibilidad de presentarse oleajes anómalos que impidan a las embarcaciones descargar el producto en los tanques esféricos de las plantas de abastecimiento de dos grandes mayoristas que abastecen la capital y en donde se concentra la mayor demanda nacional. Posteriormente, esta norma sufrió una modificación sin que haya resultado en una solución permanente al problema de suministro. En este contexto, es apropiado para una investigación que desarrolle un análisis multivariante que identifique los factores de mayor impacto en el suministro de GLP. Por último, como resultado de este análisis se deberá plantear acciones y estrategias que coadyuven a lograr un suministro adecuado de este energético satisfaciendo la demanda nacional.

### APLICACIÓN DEL ANÁLISIS MULTIVARIANTE. ANÁLISIS DE COMPONENTES PRINCIPALES

El Análisis de Componentes Principales fue desarrollado, inicialmente, por Karl Pearson a principios del siglo XX y, posteriormente, estudiado por Harold Hotelling, hacia 1930 (Ganzo, 2004). Es particularmente útil para sintetizar la información, reduciendo el número de variables con la pérdida de la mínima cantidad posible de información. Los nuevos componentes principales resultantes del análisis, independientes entre sí, constituirán una combinación lineal de las variables originales y se construyen según su importancia en relación a la variabilidad total obtenida de la muestra.

Por un lado, se ha identificado trece variables relacionadas con el abastecimiento de GLP. Una vez identificadas las variables, se ha encontrado dificultad para visualizar las posibles relaciones existentes entre las variables, además de la posibilidad de

encontrar correlaciones que puedan presentarse entre ellas, lo que puede traer como consecuencia que algunas se relacionen o que midan lo mismo, pero bajo diferentes enfoques. Para la evaluación de las relaciones que puedan presentarse entre aquellas variables que miden la misma información, se requeriría cambiar el conjunto inicial de variables en un nuevo grupo que no las correlacione entre sí; es decir, en donde no se tenga repetición o información redundante. Ello se logra reduciendo el número de variables, estimando que solo pocos factores podrían explicar buena parte de la variabilidad identificada.

Por otro lado, la selección de las variables se realiza de manera que cada una de ellas recoja la mayor parte de la variabilidad inicial, tomando en cuenta que todas serán complementarias. Del total de variables obtenidas, se elegirán las que recojan una proporción suficiente de variabilidad, las cuales reciben el nombre de componentes principales que, una vez seleccionadas, son representadas como una matriz en la que cada elemento constituye los coeficientes factoriales de las variables; en otras palabras, las correlaciones entre las variables y los componentes principales. La matriz tendrá tantas columnas como componentes principales haya y, tantas filas como variables se tenga. Para cada caso, los componentes principales tienen puntuaciones factoriales que hacen posible su representación gráfica. Las puntuaciones son calculadas a partir de la siguiente fórmula:

$$X_{ij} = a_{i1} \cdot Z_{1j} + \dots + a_{ik} \cdot Z_{kj} = \sum_{s=1}^k a_{is} \cdot Z_{sj}$$

Donde,  $a$  representa a los coeficientes y  $Z$ , a los valores estandarizados de las variables.

En la etapa de interpretación, la dificultad radica en encontrar que los coeficientes factoriales sean próximos a la unidad; que cualquiera de las variables tenga coeficientes relacionados con una de ellas y que no existan variables con coeficientes similares. Luego, concierne hacer la comprobación mediante un análisis factorial confirmatorio o mediante los modelos de ecuaciones estructurales.

Para determinar los componentes principales, se hace uso del *software* estadístico SPSS 24, donde el número del nuevo conjunto de variables no directamente observables sea pequeño, con una pérdida mínima de información y la solución que se obtenga pueda ser interpretada. Las etapas de análisis incluyen la extracción de las variables y la

estimación de las puntuaciones factoriales. El criterio de rotación utilizado ha sido el Varimax, en rotación ortogonal; sin embargo, quedaría pendiente desarrollar una investigación en la que se pruebe distintas soluciones factoriales con diferentes criterios de rotación y, de acuerdo con los resultados obtenidos, seleccionar el criterio de rotación apropiado, considerando que el resultado debe ser una solución factorial simple y que aporte información.

Este artículo pretende ilustrar la aplicación del Análisis de Componentes Principales para la identificación de factores relevantes que pueden afectar el proceso normal de abastecimiento de GLP en el Perú. Por tanto, el análisis será adecuado si se presenta una alta correlación entre las variables, lo cual es posible cuando existen aspectos comunes. De esta forma, el primer paso será realizar el análisis de la matriz de correlaciones. De acuerdo con ello, los siguientes factores de cambio han sido identificados como aquellos que afectan directamente el abastecimiento de GLP a nivel nacional:

X1	Producción de petróleo.
X2	Producción de Líquidos de Gas Natural (LGN).
X3	Procesamiento de petróleo.
X4	Importación de petróleo.
X5	Reservas de petróleo.
X6	Reservas de LGN.
X7	Procesamiento de LGN.
X8	Capacidad de almacenamiento de GLP.
X9	Cierre de puertos.
X10	Producción de GLP.
X11	Demanda de GLP.
X12	Importación de GLP.
X13	Inventarios de GLP.

Una práctica usual es determinar primero el valor del coeficiente alfa de Cronbach a fin de estimar la fiabilidad de los constructos. Sin embargo, habiendo utilizado información real, no procedente de

encuestas o cuestionarios, el resultado ratificará la conclusión de que, en caso de uso de data real, el valor obtenido no descalifica la aplicación de una técnica multivariante (Lloret-Segura *et al.*, 2014). El análisis de fiabilidad ha sido realizado utilizando el *software* estadístico SPSS 24. Mediante un análisis semejante al de la búsqueda de consistencia entre las respuestas de un cuestionario, se utiliza este coeficiente para determinar la consistencia entre los factores (variables) involucrados.

En una primera corrida, se obtuvo un coeficiente alfa de Cronbach de 0,252; el valor mínimo aceptable para el coeficiente alfa de Cronbach es de 0,7. Un valor menor indica que la consistencia interna muestra una baja correlación entre los factores, esto debido a una débil relación entre ellos, mientras que un valor superior muestra una fuerte relación entre los factores. En corridas sucesivas, tras eliminar X13, X12, X6, X5, X3 y X1, se obtiene un coeficiente alfa de Cronbach de 0,829 (ver tabla 1). Por tanto, a fin de verificar si se ratifica la eliminación de estas variables, se realizó un análisis de componentes principales con siete factores consiguiendo un KMO de 0,652 (ver tabla 2). Por otro lado, las tablas 3 y 4 justifican las varianzas explicadas para los siete componentes.

La matriz de correlaciones muestra el 100% de los coeficientes de correlación son diferentes de cero por lo que es posible continuar con el análisis. El método de extracción es el de análisis de componentes principales.

La tabla 5 muestra que la rotación de los factores consigue que el primer factor explique el 65,217%; mientras que el segundo, el 28,347%. Con esta rotación se ha conseguido una ligera mejora en la distribución de la información explicada por cada uno de los factores. En conjunto, representan el 93,564%, lo que significa que esa diferencia (6,436%) equivale a una pérdida de información menor que en el caso de las sumas de extracción.

Si se dejara de eliminar los factores X3 y X12, considerando nueve componentes, la determinación del coeficiente KMO, al tomarse en cuenta los factores X2, X3, X4, X7, X8, X9, X10, X11 y X12, después de eliminaciones sucesivas de los factores, genera un KMO igual a 0,702, valor que supera la restricción

**Tabla 1.** Estadísticas de fiabilidad.

Alfa de Cronbach	Alfa de Cronbach basado en elementos estandarizados	N.º de elementos
0,829	0,951	7

Fuente: resultados de corrida en SPSS 24.

**Tabla 2.** Prueba de KMO y Bartlett.

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		0,652
Prueba de esfericidad de Bartlett	Aproximación chi-cuadrado	189,948
	Grados de libertad	21
	Significación	0,000

Fuente: resultados de corrida en SPSS 24.

**Tabla 3.** Matriz de correlaciones para siete factores<sup>a</sup>.

		X2	X4	X7	X8	X9	X10	X11
Correlación	X2	1,000	0,404	0,941	0,765	0,835	0,999	0,977
	X4	0,404	1,000	0,508	0,747	0,234	0,419	0,349
	X7	0,941	0,508	1,000	0,886	0,731	0,944	0,928
	X8	0,765	0,747	0,886	1,000	0,614	0,772	0,713
	X9	0,835	0,234	0,731	0,614	1,000	0,828	0,834
	X10	0,999	0,419	0,944	0,772	0,828	1,000	0,979
	X11	0,977	0,349	0,928	0,713	0,834	0,979	1,000
Significación (unilateral)	X2		0,067	0,000	0,000	0,000	0,000	0,000
	X4	0,067		0,027	0,001	0,201	0,060	0,101
	X7	0,000	0,027		0,000	0,001	0,000	0,000
	X8	0,000	0,001	0,000		0,007	0,000	0,001
	X9	0,000	0,201	0,001	0,007		0,000	0,000
	X10	0,000	0,060	0,000	0,000	0,000		0,000
	X11	0,000	0,101	0,000	0,001	0,000	0,000	

Fuente: resultados de corrida en SPSS 24.

Determinante = 2,43E-008.

**Tabla 4.** Matriz de correlaciones reproducidas.

		X2	X4	X7	X8	X9	X10	X11
Correlación reproducida	X2	0,978 <sup>a</sup>	0,399	0,942	0,787	0,873	0,978	0,969
	X4	0,399	0,949 <sup>a</sup>	0,552	0,798	0,189	0,413	0,338
	X7	0,942	0,552	0,942 <sup>a</sup>	0,860	0,804	0,944	0,920
	X8	0,787	0,798	0,860	0,923 <sup>a</sup>	0,601	0,796	0,745
	X9	0,873	0,189	0,804	0,601	0,815 <sup>a</sup>	0,870	0,877
	X10	0,978	0,413	0,944	0,796	0,870	0,978 <sup>a</sup>	0,968
	X11	0,969	0,338	0,920	0,745	0,877	0,968	0,965 <sup>a</sup>
Residuo <sup>b</sup>	X2		0,005	-0,001	-0,022	-0,038	0,021	0,007
	X4	0,005		-0,044	-0,052	0,045	0,006	0,011
	X7	-0,001	-0,044		0,027	-0,074	0,000	0,007
	X8	-0,022	-0,052	0,027		0,013	-0,024	-0,032
	X9	-0,038	0,045	-0,074	0,013		-0,042	-0,043
	X10	0,021	0,006	0,000	-0,024	-0,042		0,011
	X11	0,007	0,011	0,007	-0,032	-0,043	0,011	

Fuente: resultados de corrida en SPSS 24.

a. Comunalidades reproducidas.

b. Los residuos se han calculado entre las correlaciones observadas y las reproducidas. Se ha encontrado dos (9,0%) residuos no redundantes con valores absolutos mayores que 0,05.



para el KMO. Por lo tanto, ello lleva finalmente a trabajar con nueve factores.

La tabla 6, como en el caso anterior, muestra que el 100% de los coeficientes de correlación son diferentes de cero, lo que implica que es posible continuar con el análisis. El método de extracción ha sido el de análisis de componentes principales.

La tabla 7 muestra la comprobación de la bondad del modelo obtenido, el cual se realiza a través

de un análisis de los residuos en el que se observa las diferencias entre la matriz de correlaciones inicial y la reproducida por el modelo. De esta forma, en la medida que sean próximos a cero en valor absoluto, será indicativo de un buen ajuste, tal y como se observa en la tabla. La existencia del 11% de residuos no redundantes con valor absoluto mayor que 0,05 implicaría un bajo nivel de bondad del modelo, lo que podría ser considerado como relativamente bueno.

**Tabla 5.** Varianza total explicada.

Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	5,527	78,957	78,957	5,527	78,957	78,957	4,565	65,217	65,217
2	1,022	14,606	93,564	1,022	14,606	93,564	1,984	28,347	93,564
3	0,256	3,655	97,218						
4	0,155	2,211	99,429						
5	0,030	0,432	99,861						
6	0,009	0,134	99,995						
7	0,00	0,005	100,000						

Fuente: resultados de corrida en SPSS 24.

El método de extracción es de análisis de componentes principales.

**Tabla 6.** Matriz de correlaciones para nueve factores.

		X2	X3	X4	X7	X8	X9	X10	X11	X12
Correlación	X2	1,000	-0,392	0,404	0,941	0,765	0,835	0,999	0,977	-0,877
	X3	-0,392	1,000	0,484	-0,316	0,033	-0,413	-0,380	-0,487	0,056
	X4	0,404	0,484	1,000	0,508	0,747	0,234	0,419	0,349	-0,549
	X7	0,941	-0,316	0,508	1,000	0,886	0,731	0,944	0,928	-0,825
	X8	0,765	0,033	0,747	0,886	1,000	0,614	0,772	0,713	-0,799
	X9	0,835	-0,413	0,234	0,731	0,614	1,000	0,828	0,834	-0,682
	X10	0,999	-0,380	0,419	0,944	0,772	0,828	1,000	0,979	-0,881
	X11	0,977	-0,487	0,349	0,928	0,713	0,834	0,979	1,000	-0,792
	X12	-0,877	0,056	-0,549	-0,825	-0,799	-0,682	-0,881	-0,792	1,000
Significación (unilateral)	X2		0,074	0,067	0,000	0,000	0,000	0,000	0,000	0,000
	X3	0,074		0,034	0,126	0,453	0,063	0,081	0,033	0,422
	X4	0,067	0,034		0,027	0,001	0,201	0,060	0,101	0,017
	X7	0,000	0,126	0,027		0,000	0,001	0,000	0,000	0,000
	X8	0,000	0,453	0,001	0,000		0,007	0,000	0,001	0,000
	X9	0,000	0,063	0,201	0,001	0,007		0,000	0,000	0,003
	X10	0,000	0,081	0,060	0,000	0,000	0,000		0,000	0,000
	X11	0,000	0,033	0,101	0,000	0,001	0,000	0,000		0,000
	X12	0,000	0,422	0,017	0,000	0,000	0,003	0,000	0,000	

Fuente: resultados de corrida en SPSS 24.

La tabla 8 muestra que la rotación de los factores consigue que el primero explique el 70,23%; mientras que el segundo, el 20,46%. Con la rotación se ha conseguido una ligera mejora en la distribución de la información explicada por cada uno de los factores. En conjunto, representan el 90,69%, siendo la diferencia (9,31%) una pérdida de información

menor que en el caso de las sumas de extracción, pero mayor al caso de siete componentes.

Finalmente, la matriz de componentes rotados, mostrada en la tabla 9, indica que la rotación ha convergido en tres iteraciones, lo cual nos lleva a un modelo con dos componentes principales, don-

**Tabla 7.** Matriz de correlaciones reproducidas.

		X2	X3	X4	X7	X8	X9	X10	X11	X12
Correlación reproducida	X2	0,980 <sup>a</sup>	-0,412	0,413	0,948	0,799	0,855	0,980	0,968	-0,857
	X3	-0,412	0,903 <sup>a</sup>	0,553	-0,285	0,083	-0,482	-0,399	-0,489	0,100
	X4	0,413	0,553	0,898 <sup>a</sup>	0,512	0,754	0,238	0,427	0,326	-0,621
	X7	0,948	-0,285	0,512	0,934 <sup>a</sup>	0,838	0,808	0,950	0,923	-0,869
	X8	0,799	0,083	0,754	0,838	0,892 <sup>a</sup>	0,627	0,807	0,742	-0,849
	X9	0,855	-0,482	0,238	0,808	0,627	0,766 <sup>a</sup>	0,853	0,858	-0,704
	X10	0,980	-0,399	0,427	0,950	0,807	0,853	0,980 <sup>a</sup>	0,967	-0,862
	X11	0,968	-0,489	0,326	0,923	0,742	0,858	0,967	0,966 <sup>a</sup>	-0,817
	X12	-0,857	0,100	-0,621	-0,869	-0,849	-0,704	-0,862	-0,817	,843 <sup>a</sup>
Residuo <sup>b</sup>	X2		0,020	-0,009	-0,007	-0,034	-0,020	0,019	0,009	-0,019
	X3	0,020		-0,069	-0,030	-0,049	0,069	0,019	0,002	-0,044
	X4	-0,009	-0,069		-0,004	-0,007	-0,004	-0,008	0,023	0,072
	X7	-0,007	-0,030	-0,004		0,048	-0,077	-0,006	0,004	0,044
	X8	-0,034	-0,049	-0,007	0,048		-0,013	-0,035	-0,030	0,050
	X9	-0,020	0,069	-0,004	-0,077	-0,013		-0,025	-0,024	0,022
	X10	0,019	0,019	-0,008	-0,006	-0,035	-0,025		0,012	-0,019
	X11	0,009	0,002	0,023	0,004	-0,030	-0,024	0,012		0,025
	X12	-0,019	-0,044	0,072	0,044	0,050	0,022	-0,019	0,025	

Fuente: resultados de corrida en SPSS 24.

a. Comunalidades reproducidas.

b. Los residuos se han calculado entre las correlaciones observadas y las reproducidas. Se ha observado cuatro (11,0%) residuos no redundantes con valor absoluto mayor que 0,05.

**Tabla 8.** Varianza total explicada para nueve factores.

Componente	Autovalores iniciales			Sumas de extracción de cargas al cuadrado			Sumas de rotación de cargas al cuadrado		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	6,402	71,138	71,138	6,402	71,138	71,138	6,321	70,231	70,231
2	1,760	19,554	90,692	1,760	19,554	90,692	1,841	20,461	90,692
3	0,324	3,598	94,290						
4	0,255	2,835	97,125						
5	0,150	1,662	98,787						
6	0,089	0,985	99,772						
7	0,012	0,135	99,907						
8	0,008	0,090	99,997						
9	0,000	0,003	100,000						

Fuente: resultado de corrida en SPSS 24.

de el primer componente es la variable con máxima varianza con solo dos coeficientes negativos.

$$Y1 = 0,990X2 - 0,421X3 + 0,413X4 + 0,956X7 + 0,804X8 + 0,864X9 + 0,990X10 + 0,978X11 - 0,864X12$$

$$Y2 = 0,006X2 + 0,852X3 + 0,853X4 + 0,138X7 + 0,495X8 - 0,138X9 + 0,022X10 - 0,091X11 - 0,310X12$$

Se puede observar que X2 (producción de LGN), X10 (producción de GLP), X8 (capacidad de almacenamiento de GLP) y X7 (procesamiento de LGN) son variables altamente correlacionadas, puesto que el procesamiento de líquidos de gas natural dependerá de la producción de gas natural y, a su vez, de la producción de GLP. Asimismo, esta producción estará determinada por su capacidad de almacenamiento.

La tabla 9 muestra la matriz de componente rotado considerando que el método de extracción fue el de análisis de componentes principales. El método de rotación utilizado ha sido el Varimax con normalización Kaiser, en la que la rotación ha convergido en tres iteraciones.

**Tabla 9.** Matriz de componente rotado<sup>a</sup>.

	Componente	
	1	2
X2	0,990	0,006
X10	0,990	0,022
X11	0,978	-0,091
X7	0,956	0,138
X9	0,864	-0,138
X12	-0,864	-0,310
X8	0,804	0,495
X4	0,413	0,853
X3	-0,421	0,852

Fuente: resultados de corrida en SPSS 24.

a. La rotación ha convergido en tres iteraciones.

## CONCLUSIONES

1. El Análisis de Componentes Principales (ACP) ha analizado la estructura de las interrelaciones entre once variables relacionadas con el abastecimiento de GLP, en el que se ha utilizado datos reales para cada una, encontrándose un ajuste satisfactorio a los supuestos de un modelo del ACP clásico. Si se analiza la matriz de correlaciones para verificar que exista una correlación entre las variables, se encuentra que el 100% de

los coeficientes de correlación son diferentes de cero, lo que implica que es posible continuar con el análisis y obtener un mejor ajuste.

- La medida de adecuación muestral de Kaiser-Meyer-Olkin (KMO) indica que la muestra de datos utilizada es adecuada.
- La prueba de esfericidad de Bartlett indica un valor de significancia de 0,000. Por ello, no se trata de una matriz identidad, y se concluye con que existe correlación entre las variables.
- La tabla de varianza total explicada muestra que existe solo 9,31% de pérdida de información cuando se consideran nueve factores; y, 6,44%, cuando se consideran siete factores.
- Se ha comprobado la bondad del modelo, dado que las diferencias entre las matrices de correlación inicial y la reproducida por el modelo son próximos a cero en valor absoluto, lo cual es indicativo de un buen ajuste.
- Con las variables relacionadas se podría realizar un Análisis Factorial Confirmatorio (AFC) y posterior modelamiento de ecuaciones estructurales (Exploratory Structural Equation Modeling, ESEM), lo que permitiría comprobar que el modelo estructural que se proponga se ajusta adecuadamente a los datos.
- Con estos resultados se comprueba que las variables producción de GLP (X10) y cierre de puertos (X9) tienen un aporte significativo en el modelo. Asimismo, el abastecimiento de GLP a nivel nacional depende de la producción (X2) y procesamiento de LGN (X7), por lo cual, cualquier problema relacionado a estas actividades puede afectarlo.
- Las variables producción de petróleo (X1), reservas de petróleo (X5), reservas de LGN (X6) e inventarios de GLP (X13) han mostrado tener un menor nivel de influencia en el abastecimiento de GLP a nivel nacional.
- El resto de las variables —producción de LGN (X2), procesamiento de petróleo (X3), importación de petróleo (X4), procesamiento de LGN (X7), capacidad de almacenamiento de GLP (X8), cierre de puertos (X9), producción de GLP (X10), demanda de GLP (X11) e importación de GLP (X12)— han mostrado tener un alto grado de correlación con el abastecimiento de GLP a nivel nacional; por tanto, cualquier perturbación que afecte a una de estas variables causará un impacto en este abastecimiento.



10. El modelo de dos componentes refleja el comportamiento real de las variables, de acuerdo a la data histórica utilizada.
11. Los componentes Y1 e Y2 relacionan el abastecimiento de GLP de acuerdo a su origen, por lo cual Y1 se designa  $GLP_{Gas\ Natural}$  e Y2 corresponde a  $GLP_{Petróleo}$ ; por tanto:

$$GLP_{Gas\ Natural} = 0,990 \text{ Producción de LGN} - 0,421 \text{ Procesamiento de petróleo} + 0,413 \text{ Importación de petróleo} + 0,956 \text{ Procesamiento de LGN} + 0,804 \text{ Capacidad de almacenamiento de GLP} + 0,864 \text{ Cierre de puertos} + 0,990 \text{ Producción de GLP} + 0,978 \text{ Demanda de GLP} - 0,864 \text{ Importación de GLP.}$$

$$GLP_{Petróleo} = 0,006 \text{ Producción de LGN} + 0,852 \text{ Procesamiento de petróleo} + 0,853 \text{ Importación de petróleo} + 0,138 \text{ Procesamiento de LGN} + 0,495 \text{ Capacidad de almacenamiento de GLP} - 0,138 \text{ Cierre de puertos} + 0,022 \text{ Producción de GLP} - 0,091 \text{ Demanda de GLP} - 0,310 \text{ Importación de GLP.}$$

## REFERENCIAS BIBLIOGRÁFICAS

- [1] Berrah, N.; Feng, F.; Priddle, R. & Wang, L. (2007). *Sustainable energy in China: The closing window of opportunity*, p. 273. Washington, DC, International Bank for Reconstruction and Development, The World Bank. ISBN-13: 978-0-812-6754-4.
- [2] Checchi, A.; Behrens, A. y Egenhofer, C. (2009). Long-term energy security risks for Europe: a sector-specific approach. *Working Document*, 52(309). Recuperado de <http://aei.pitt.edu/10759/1/1785.pdf>.
- [3] Commission of the European Communities (2000a). Green paper: towards a European strategy for the security of energy supply. *Office for Official Publications of the European Communities*, 90. Recuperado de <https://op.europa.eu/en/publication-detail/-/publication/0ef8d03f-7c54-41b6-ab89-6b93e61fd37c/language-en>.
- [4] Cuadras, C. (2018). *Nuevos métodos de análisis multivariante*. Barcelona, España: CMC Editions.
- [5] Ganzo, A. (2004). Análisis de componentes principales. Introducción al análisis multivariante. Recuperado de [http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat\\_50140155\\_AnalisisMultivariante.pdf](http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat_50140155_AnalisisMultivariante.pdf).
- [6] IEA (1985). *Energy technology policy*. París, Francia: OECD Publishing.
- [7] IEA (2001). *Towards a sustainable energy future*. París, Francia: OECD. Recuperado de [https://www.oecd-ilibrary.org/energy/towards-a-sustainable-energy-future\\_9789264193581-en](https://www.oecd-ilibrary.org/energy/towards-a-sustainable-energy-future_9789264193581-en).
- [8] IEA (2007). *Energy security and climate policy-assessing interactions*. París, Francia: OECD. Recuperado de [https://www.oecd-ilibrary.org/energy/energy-security-and-climate-policy-assessing-interactions\\_9789264109940-en](https://www.oecd-ilibrary.org/energy/energy-security-and-climate-policy-assessing-interactions_9789264109940-en).
- [9] *Informe situacional de la comercialización del Gas Licuado de Petróleo (GLP) en el Perú* (2015). Lima, Perú: Gerencia de Fiscalización de Hidrocarburos Líquidos de Osinergmin.
- [10] Kruyt, B.; Vuuren, D. P. van; Vries, H. J. M. de y Groenenberg, H. (2009). Indicators for energy security. *Energy Policy*, 37(6), 2166-2181. Recuperado de <https://www.sciencedirect.com/science/article/pii/S0301421509000883?via%3Dihub>.
- [11] Lloret, S.; Ferreres, A.; Hernández, A. y Tomás, I. (2014). Exploratory item factor analysis: a practical guide revised and up-dated. *Annals of Psychology*, 30(3), 1151-1169. Recuperado de <https://revistas.um.es/analesps/article/view/analesps.30.3.199361/217271>.
- [12] Löschel, A.; Johnston, J.; Delucchi, M.; Demayo, T.; Gautier, D.; Greene, D.; Ogdén, J.; Rayner, S. y Worrell, E. (2009). Stocks, flows, and prospects of energy. En T. E. Graedel y E. van der Voet (Eds.), *Linkages of sustainability*. Oxford, Reino Unido: MIT Press Scholarship.

# Multivariate analysis applied to the supply of liquefied petroleum gas (LPG) in Peru

BEATRIZ ADANIYA HIGA <sup>1</sup>

RECIBIDO: 30/01/2019 ACEPTADO: 25/04/2019

## RESUMEN

This article provides a brief analysis on energy supply security, focusing mainly on the variables related to the supply of liquefied petroleum gas (LPG) within the Peruvian market. Furthermore, via the use of real data, principal component analysis (PCA), one of the most common multivariate techniques, is explained and illustrated, whose result indicates that a correlation exists between the analyzed variables.

**Palabras-claves:** Principal component analysis; multivariate statistical technique; LPG supply; variables; energy security.

## INTRODUCTION

This article features an initial analysis of the causal variables involved in the supply of liquefied petroleum gas (LPG), in order to act upon, forecast, and counteract any issue that may affect the LPG market and have an impact on the national economy. The interest to conduct this research arose from the fact that no or limited multivariate analysis studies of the factors to secure a permanent supply of said fuel had been conducted in the country.

In the early months of 2004, it became evident that national LPG production was insufficient to satisfy national demand, necessitating importation. However, after the Pluspetrol Natural Gas Liquids (NGLs) Fractionation Plant began operating, Peru became an exporter of said product. This condition is currently being reversed by the rapid increase in LPG demand and the slow execution of projects to expand storage capacity within supply plants. This justifies the development of models to forecast future scenarios regarding LPG supply; however, legal factors, awaiting the attention of sector authorities, prevent improvements in the LPG supply strategy, variables that should be subject to a more in-depth analysis.

The term “energy security” is used in various contexts. This evidences the difference in approach that causes multiple interpretations and conceptual disagreements. The International Energy Agency (IEA) defined “energy security” as “an adequate supply of energy at a reasonable price” (1985, p. 9). Years later, in 2001, it redefined it as the “physical availability of supplies to satisfy demand at a given price” (p. 76). Later, the same agency stated that “energy security, defined in general terms, means adequate and reliable supplies of energy at affordable prices” (IEA, 2007, p. 160), with a focus on prices that would be “competitive and not too volatile”.

<sup>1</sup> Petrochemical Engineer from the Universidad Nacional de Ingeniería (UNI). MBA and MSc in Petroleum and Natural Gas Engineering from the same university. Professor at the Department of Petrochemical Engineering of the School of Petroleum, Natural Gas and Petrochemistry of the UNI. She is also manager of natural gas production and processing for the División de Supervisión de Gas Natural at Osinergmin.  
 E-mail: badaniya@fip.uni.edu.pe

## METHODOLOGY

The methodology for data collection included exploring the field through expert opinions, analysis of commercial transactions in the Peruvian LPG market, reports issued by national and international energy agencies and institutions, and the formulation of the constructs. Principal component analysis was used for verification. The evolution of the Peruvian LPG market over time was taken into account, as it is possible to find several structural models in previous research that show various states for the variables considered in this analysis, which could determine an expansion of the study to a confirmatory factor analysis (CFA) and models that evaluate the effect of time on fuel market development. The findings of these studies might stress the need for changes in legislation in order to achieve greater order in the LPG market and to ensure companies comply with regulations.

## ENERGY SECURITY AND ITS DIFFERENCE FROM SUPPLY SECURITY

The term “energy security” is also used to refer to security of supply (Kruyt *et al.*, 2009; Löschel *et al.*, 2009). The first is seen as the general framework that includes importers and exporters; whereas the second is considered a subtype of energy security, and as such it is examined by analysts in the short and long term (IEA, 2007; Checchi *et al.*, 2009). As stated by Kruyt *et al.* (2009), in the short term, energy security aims to mitigate potential supply interruptions or to have a rapid response strategy to confront potential energy scarcity. This strategy would have to consider the replacement of supply flow, the use of spare capacity of producers (a simile for cold standby in power generation) and emergency reserves (minimum inventories in the case of Peru) or immediately relevant measures to regulate demand. The purpose is to reduce the impact a shortage of energy supply would have on key sectors (IEA, 2001) and to restrain demand to address the physical lack of energy (IEA, 2007), a difficult task in the country.

If the main factors associated with security of energy supply—such as physical availability in sufficient volumes, without interruptions and at affordable prices—are carefully analyzed, it is possible to note that long-term security of supply intends to ensure sufficient energy availability for sustainable economic development. To that end, the strategy aims to ensure greater energy efficiency, reduce vulnerability, rely on diverse energy sources via the use of renewable energy and new suppliers, as well as

the development of new transportation technologies and the expansion of existing interconnected energy networks.

The Commission of the European Communities explains that “security of supply means the ability to ensure that future essential energy needs can be met, both by means of adequate domestic resources worked under economically acceptable conditions or maintained as strategic reserves, and by calling upon accessible and stable external sources supplemented where appropriate, by strategic stocks” (European Commission, 1990a, p.16, as cited in Berrah, Feng, Priddle & Wang, 2007). Years later, it was related to security of energy supply considering that “the European Union’s long-term strategy for energy supply security must be geared to ensuring, for the well-being of its citizens and the proper functioning of the economy, the uninterrupted physical availability of energy products on the market, at a price which is affordable for all consumers (private and industrial), while respecting environmental concerns and looking towards sustainable development” (Commission of the European Communities, 2001a, p. 2).

As for the permanent and sufficient availability of energy, it is important to consider conditions needed to satisfy demand while contributing to the development of the strategic economic and ordinary activities of each country (defense, production, transportation, and domestic consumption actions, among others). These conditions originate from geological, technical, economic and geopolitical aspects, as they refer to the availability of resources, proper operation of infrastructure and reliability of each company and each interest group committed to the system. As expected, investments play a major role within this complex system.

## LPG SUPPLY IN THE PERUVIAN CONTEXT

In Peru, the Reglamento para la Comercialización de Gas Licuado de Petróleo [Regulation for Commercialization of Liquefied Petroleum Gas] passed by Supreme Decree No. 01-94-EM, regulates LPG supply security. Section 8 of the same decree states that every LPG production plant and importer is compelled to maintain, as average supply, a volume equivalent to fifteen days of the average volume sold in the last six months or the average volume imported during the same period, as appropriate. Unfortunately, a report on the situation of commercialization shows a statistical analysis of sales nationwide that reveals that it is impossible to monitor compliance with that obligation (División Planeam-

iento y Desarrollo de la Gerencia de Fiscalización de Hidrocarburos Líquidos, 2015). The fulfillment of said obligation would guarantee fuel supply. According to the report, the available supply ranged from three to twelve days, without considering the possibility of abnormal waves that prevent ships from unloading the product in the spherical tanks of the supply plants of two major wholesalers that supply the capital, where the national demand is concentrated. This regulation was later amended, nonetheless, it did not signify a permanent solution to the supply problem. In this case, a research study involving a multivariate analysis that identifies the factors with greater impact on the supply of LPG is necessary. Finally, actions and strategies that contribute to secure an adequate LPG supply to meet domestic demand will be proposed as a result of this analysis.

### MULTIVARIATE ANALYSIS APPLICATION. PRINCIPAL COMPONENT ANALYSIS

The Principal Component Analysis was initially developed by Karl Pearson at the beginning of the 20<sup>th</sup> century, and later studied by Harold Hotelling around 1930 (Ganzo, 2004). It is particularly useful to synthesize information, reducing the number of variables while ensuring the smallest loss of information possible. The new principal components resulting from the analysis, independent from each other, will constitute a linear combination of the original variables and are created according to their importance in relation to the total variance obtained from the sample.

Firstly, thirteen variables related to LPG supply were identified. After identification, possible existing relationships between variables, in addition to the potential correlations that may arise between them, were difficult to visualize. This might result in the correlation of some variables or in some having the same weight, but under different approaches. In order to evaluate the correlation that may arise between variables that measure the same data, it would be necessary to replace the initial group of variables with one where the variables are not correlated, thus, a group with no repetitions or redundant information. This is achieved by reducing the number of variables, estimating that only few factors may explain a significant part of the identified variability.

Secondly, the variables are selected in such a manner that each of them includes the largest amount possible of the initial variability, considering that all of them will be complementary. Of the total

variables obtained, those that include a sufficient proportion of variability will be chosen. These are the principal components. These components will be represented as a matrix, where each element will constitute the variables' factorial coefficients; in other words, the correlations between the variables and the principal components. The matrix will have as many columns as principal components chosen and as many rows as there are variables. In each case, the principal components have factor points that allow them to be represented graphically. Factors points are calculated using the following formula:

$$X_{ij} = a_{i1} \cdot Z_{1j} + \dots + a_{ik} \cdot Z_{kj} = \sum_{s=1}^k a_{is} \cdot Z_{sk}$$

Where  $a$  represents the coefficients and  $Z$ , the standardized values of the variables.

For the interpretation stage, factor coefficients must be close to one, a variable must have high coefficients with one factor, and there must not be variables with similar coefficients. Thereafter, verification by means of a confirmatory factor analysis or structural equations models is necessary.

Statistical software SPSS 24 is used to determine the principal components, where a small new set of not directly observable variables is obtained with minimum information loss, and the resulting solution can be interpreted. The analysis stages include the extraction of variables and estimation of factor points. The Varimax criterion for orthogonal rotation was used for factor rotation. However, a research study that analyzes various factor solutions through several rotation criteria in order to select the appropriate rotation criterion, bearing in mind that the result must be a simple factor solution that provides information, is yet to be conducted.

This article attempts to illustrate the application of the Principal Component Analysis to identify relevant factors that may affect the usual LPG supply process in Peru. Therefore, the analysis will be deemed adequate if a high correlation between variables is found, which is possible when there are common aspects. Thus, the first step is to analyze the correlations matrix. The following factors have been identified as those that directly affect LPG supply at national level:

X1 Petroleum production.



- X2 Natural Gas Liquids (NGL) production.
- X3 Petroleum processing.
- X4 Petroleum importation.
- X5 Petroleum reserves.
- X6 NGL reserves.
- X7 NGL processing.
- X8 LPG storage capacity.
- X9 Ports closure.
- X10 LPG production.
- X11 LPG demand.
- X12 LPG importation.
- X13 LPG inventories.

A common practice is to first determine the Cronbach's alpha coefficient value in order to estimate the reliability of constructs. However, having used real data, and not data collected via surveys and questionnaires, the outcome will confirm the conclusion that the value obtained does not invalidate the application of a multivariate technique when real data is used (Lloret-Segura *et al.*, 2014). The reliability analysis was performed using statistical software SPSS 24. Through an analysis similar to the search for consistency among the responses to a questionnaire, the resulting coefficient is used to determine consistency among the involved factors (variables).

On the first run, Cronbach's alpha coefficient was 0.252. The minimum acceptable value for Cronbach's alpha coefficient is 0.7. Lower values indicate that internal consistency reflect low correlation between factors, due to a weak relationship between them; whereas, higher values reflect strong relationship between factors. After removing variables X13, X12, X6, X5, X3 and X1, a Cronbach's alpha coefficient of 0.829 was obtained on subsequent runs (see Table 1). Consequently, a principal component analysis with seven factors was conducted in order to confirm the elimination of said variables. It showed a KMO value of 0.652(see Table 2). Ad-

ditionally, Tables 3 and 4 justify the explained variances for seven factors (variables).

The correlation matrix shows that 100% of the correlation coefficients are different from zero, therefore it is possible to continue the analysis. The extraction method used was the principal component analysis.

It can be observed in Table 5 that the first factor explains 65.217% of the variance, whereas the second explains 28.347%. A slight improvement in the distribution of the information explained by each factor was obtained applying factor rotation. Together, they represent 93.564% of the variance, which signifies that the difference (6.436%) corresponds to a smaller information loss compared to the extraction sums case.

If factors X3 and X12 were not eliminated, examining nine factors, the determination of the KMO coefficient, considering factors X2, X3, X4, X7, X8, X9, X10, X11 and X12 upon successive elimination of factors, would be equal to 0.702, a value that exceeds the minimum acceptable KMO values. Consequently, nine factors are considered in the analysis.

Table 6, as before, shows that 100% of the correlation coefficients are different from zero, therefore it is possible to continue the analysis. The extraction method used was the principal component analysis.

Table 7 shows the verification of the goodness of fit of the obtained model, which is performed through an analysis of the residuals, where the differences between the matrix of initial correlations and the matrix reproduced by the model are observed. Thus, absolute value close to zero indicates a good fit, as observed in the table. The existence of 11% of nonredundant residuals with absolute value greater than 0.05 indicates a low goodness of fit, which may be considered relatively good.

Table 8 shows that the first factor explains 70.23% of the variance, whereas the second explains 20.46%. A slight improvement in the distribution of the information explained by each factor was obtained applying factor rotation. Together, they represent 90.69% of the variance, which signifies that the difference

**Table 1. Reliability statistics.**

Cronbach's alpha	Cronbach's alpha based on standardized items	No. of items
0.829	0.951	7

Source: Results of analysis ran on SPSS 24.



**Table 2. KMO and Bartlett's test.**

<b>Kaiser-Meyer-Olkin measure of sampling adequacy</b>		<b>0.652</b>
Bartlett's test of sphericity	Approx. Chi-square	189.948
	Degrees of freedom	21
	Significance	0.000

Source: Results of analysis ran on SPSS 24.

**Table 3. Correlation matrix for seven factors<sup>a</sup>.**

		X2	X4	X7	X8	X9	X10	X11
Correlation	X2	1.000	0.404	0.941	0.765	0.835	0.999	0.977
	X4	0.404	1.000	0.508	0.747	0.234	0.419	0.349
	X7	0.941	0.508	1.000	0.886	0.731	0.944	0.928
	X8	0.765	0.747	0.886	1.000	0.614	0.772	0.713
	X9	0.835	0.234	0.731	0.614	1.000	0.828	0.834
	X10	0.999	0.419	0.944	0.772	0.828	1.000	0.979
	X11	0.977	0.349	0.928	0.713	0.834	0.979	1.000
Significance (one-tailed)	X2		0.067	0.000	0.000	0.000	0.000	0.000
	X4	0.067		0.027	0.001	0.201	0.060	0.101
	X7	0.000	0.027		0.000	0.001	0.000	0.000
	X8	0.000	0.001	0.000		0.007	0.000	0.001
	X9	0.000	0.201	0.001	0.007		0.000	0.000
	X10	0.000	0.060	0.000	0.000	0.000		0.000
	X11	0.000	0.101	0.000	0.001	0.000	0.000	

Source: Results of analysis ran on SPSS 24.

a. Determinant = 2.43E-008.

**Table 4. Reproduced correlations matrix.**

		X2	X4	X7	X8	X9	X10	X11
Reproduced correlation	X2	0.978 <sup>a</sup>	0.399	0.942	0.787	0.873	0.978	0.969
	X4	0.399	0.949 <sup>a</sup>	0.552	0.798	0.189	0.413	0.338
	X7	0.942	0.552	0.942 <sup>a</sup>	0.860	0.804	0.944	0.920
	X8	0.787	0.798	0.860	0.923 <sup>a</sup>	0.601	0.796	0.745
	X9	0.873	0.189	0.804	0.601	0.815 <sup>a</sup>	0.870	0.877
	X10	0.978	0.413	0.944	0.796	0.870	0.978 <sup>a</sup>	0.968
	X11	0.969	0.338	0.920	0.745	0.877	0.968	0.965 <sup>a</sup>
Residual <sup>b</sup>	X2		0.005	-0.001	-0.022	-0.038	0.021	0.007
	X4	0.005		-0.044	-0.052	0.045	0.006	0.011
	X7	-0.001	-0.044		0.027	-0.074	0.000	0.007
	X8	-0.022	-0.052	0.027		0.013	-0.024	-0.032
	X9	-0.038	0.045	-0.074	0.013		-0.042	-0.043
	X10	0.021	0.006	0.000	-0.024	-0.042		0.011
	X11	0.007	0.011	0.007	-0.032	-0.043	0.011	

Source: Results of analysis ran on SPSS 24.

a. Reproduced communalities.

b. Residuals are computed between observed and reproduced correlations. There are two (9.0%) nonredundant residuals with absolute values greater than 0.05.

(9.31%) corresponds to a smaller information loss compared to the extraction sums case, but greater than that considering seven components.

Lastly, the rotated component matrix exhibited in Table 9 indicates that rotation converged in three iterations, which results in a model with two principal components, where the first component is the variable with the highest variance and only two negative coefficients.

$$Y1 = 0.990X2 - 0.421X3 + 0.413X4 + 0.956X7 + 0.804X8 + 0.864X9 + 0.990X10 + 0.978X11 - 0.864X12$$

$$Y2 = 0.006X2 + 0.852X3 + 0.853X4 + 0.138X7 + 0.495X8 - 0.138X9 + 0.022X10 - 0.091X11 - 0.310X12$$

It is observed that X2 (NGL production), X10 (LPG production), X8 (LPG storage capacity) and X7 (NGL processing) are highly correlated variables,

**Table 5.** Total variance explained.

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	% cumulative	Total	% of variance	% cumulative	Total	% of variance	% cumulative
1	5.527	78.957	78.957	5.527	78.957	78.957	4.565	65.217	65.217
2	1.022	14.606	93.564	1.022	14.606	93.564	1.984	28.347	93.564
3	0.256	3.655	97.218						
4	0.155	2.211	99.429						
5	0.030	0.432	99.861						
6	0.009	0.134	99.995						
7	0.00	0.005	100.000						

Source: Results of analysis ran on SPSS 24.

Extraction method: Principal component analysis.

**Table 6.** Correlation matrix for nine factors.

		X2	X3	X4	X7	X8	X9	X10	X11	X12
Correlation	X2	1.000	-0.392	0.404	0.941	0.765	0.835	0.999	0.977	-0.877
	X3	-0.392	1.000	0.484	-0.316	0.033	-0.413	-0.380	-0.487	0.056
	X4	0.404	0.484	1.000	0.508	0.747	0.234	0.419	0.349	-0.549
	X7	0.941	-0.316	0.508	1.000	0.886	0.731	0.944	0.928	-0.825
	X8	0.765	0.033	0.747	0.886	1.000	0.614	0.772	0.713	-0.799
	X9	0.835	-0.413	0.234	0.731	0.614	1.000	0.828	0.834	-0.682
	X10	0.999	-0.380	0.419	0.944	0.772	0.828	1.000	0.979	-0.881
	X11	0.977	-0.487	0.349	0.928	0.713	0.834	0.979	1.000	-0.792
	X12	-0.877	0.056	-0.549	-0.825	-0.799	-0.682	-0.881	-0.792	1.000
Significance (one-tailed)	X2		0.074	0.067	0.000	0.000	0.000	0.000	0.000	0.000
	X3	0.074		0.034	0.126	0.453	0.063	0.081	0.033	0.422
	X4	0.067	0.034		0.027	0.001	0.201	0.060	0.101	0.017
	X7	0.000	0.126	0.027		0.000	0.001	0.000	0.000	0.000
	X8	0.000	0.453	0.001	0.000		0.007	0.000	0.001	0.000
	X9	0.000	0.063	0.201	0.001	0.007		0.000	0.000	0.003
	X10	0.000	0.081	0.060	0.000	0.000	0.000		0.000	0.000
	X11	0.000	0.033	0.101	0.000	0.001	0.000	0.000		0.000
	X12	0.000	0.422	0.017	0.000	0.000	0.003	0.000	0.000	

Source: Results of analysis ran on SPSS 24.

because natural gas liquids processing depends both on natural gas production and LPG production. Furthermore, the latter is determined by LPG storage capacity.

Table 9 shows the rotated component matrix using the principal component analysis as extraction method. The rotation method used was Varimax with Kaiser normalization. Rotation converged in three iterations.

**Table 7. Reproduced correlations matrix.**

		X2	X3	X4	X7	X8	X9	X10	X11	X12
Reproduced correlations	X2	0.980 <sup>a</sup>	-0.412	0.413	0.948	0.799	0.855	0.980	0.968	-0.857
	X3	-0.412	0.903 <sup>a</sup>	0.553	-0.285	0.083	-0.482	-0.399	-0.489	0.100
	X4	0.413	0.553	0.898 <sup>a</sup>	0.512	0.754	0.238	0.427	0.326	-0.621
	X7	0.948	-0.285	0.512	0.934 <sup>a</sup>	0.838	0.808	0.950	0.923	-0.869
	X8	0.799	0.083	0.754	0.838	0.892 <sup>a</sup>	0.627	0.807	0.742	-0.849
	X9	0.855	-0.482	0.238	0.808	0.627	0.766 <sup>a</sup>	0.853	0.858	-0.704
	X10	0.980	-0.399	0.427	0.950	0.807	0.853	0.980 <sup>a</sup>	0.967	-0.862
	X11	0.968	-0.489	0.326	0.923	0.742	0.858	0.967	0.966 <sup>a</sup>	-0.817
	X12	-0.857	0.100	-0.621	-0.869	-0.849	-0.704	-0.862	-0.817	0.843 <sup>a</sup>
Residual <sup>b</sup>	X2		0.020	-0.009	-0.007	-0.034	-0.020	0.019	0.009	-0.019
	X3	0.020		-0.069	-0.030	-0.049	0.069	0.019	0.002	-0.044
	X4	-0.009	-0.069		-0.004	-0.007	-0.004	-0.008	0.023	0.072
	X7	-0.007	-0.030	-0.004		0.048	-0.077	-0.006	0.004	0.044
	X8	-0.034	-0.049	-0.007	0.048		-0.013	-0.035	-0.030	0.050
	X9	-0.020	0.069	-0.004	-0.077	-0.013		-0.025	-0.024	0.022
	X10	0.019	0.019	-0.008	-0.006	-0.035	-0.025		0.012	-0.019
	X11	0.009	0.002	0.023	0.004	-0.030	-0.024	0.012		0.025
	X12	-0.019	-0.044	0.072	0.044	0.050	0.022	-0.019	0.025	

Source: Results of analysis ran on SPSS 24.

a. Reproduced communalities.

b. Residuals are computed between observed and reproduced correlations. There are four (11.0%) nonredundant residuals with absolute values greater than 0.05.

**Table 8. Total variance explained for nine factors.**

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	% cumulative	Total	% of variance	% cumulative	Total	% of variance	% cumulative
1	6.402	71.138	71.138	6.402	71.138	71.138	6.321	70.231	70.231
2	1.760	19.554	90.692	1.760	19.554	90.692	1.841	20.461	90.692
3	0.324	3.598	94.290						
4	0.255	2.835	97.125						
5	0.150	1.662	98.787						
6	0.089	0.985	99.772						
7	0.012	0.135	99.907						
8	0.008	0.090	99.997						
9	0.000	0.003	100.000						

Source: Results of analysis ran on SPSS 24.

## CONCLUSIONS

1. The principal component analysis (PCA) analyzed the structure of the interrelationships between eleven variables related to LPG supply. Real data was provided for each variable, and a satisfactory fit to the assumptions of the classic PCA model was found. When the correlation matrix is analyzed to verify that correlation exists between variables, it is found that 100% of the correlation coefficients are different from zero; therefore, it is possible to continue the analysis and obtain a better fit.
2. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy indicates that the data sample used is adequate.
3. Bartlett's test of sphericity indicates a significance value of 0.000, therefore, it is not an identity matrix. It is concluded that there is correlation between the variables.
4. The total explained variance table shows that there is only 9.31% of information loss when nine factors are considered, and 6.44% when seven factors are considered.
5. The goodness of fit of the model has been demonstrated, as the differences between the matrix of initial correlation and the matrix reproduced by the model are close to zero in absolute value, which is indicative of a good fit.
6. It may be possible to conduct a confirmatory factor analysis (CFA) and a subsequent exploratory structural equation modeling (ESEM) using the correlated variables, which would make it possible to verify that the proposed structural model fits the data appropriately.
7. These results demonstrate that variables LPG production (X10) and ports closure (X9) significantly contribute to the model. Additionally, nationwide LPG supply depends on NGL production (X2) and NGL processing (X7); thus, any issue related to these activities may affect it.
8. Variables petroleum production (X1), petroleum reserves (X5), NGL reserves (X6) and LPG inventories (X13) showed less influence on the LPG supply nationwide.
9. The remaining variables—NGL production (X2), petroleum processing (X3), petroleum importation (X4), NGL processing (X7), LPG storage capacity (X8), ports closure (X9), LPG production (X10), LPG demand (X11) and LPG impor-

tation (X12)—have a high degree of correlation with nationwide LPG supply; consequently, any disturbance that affects one of these variables will have an impact on supply.

10. The two component model reflects the real behavior of the variables, based on the historical data used.
11. Components Y1 and Y2 represent LPG supply according to origin, whereby Y1 corresponds to  $LPG_{\text{Natural Gas}}$  and Y2 corresponds to  $LPG_{\text{Petroleum}}$ ; therefore:

$$LPG_{\text{Natural Gas}} = 0.990 \text{ NGL production} - 0.421 \text{ petroleum processing} + 0.413 \text{ petroleum importation} + 0.956 \text{ NGL processing} + 0.804 \text{ LPG storage capacity} + 0.864 \text{ Ports closure} + 0.990 \text{ LPG production} + 0.978 \text{ LPG demand} - 0.864 \text{ Importation of LPG.}$$

$$LPG_{\text{Petroleum}} = 0.006 \text{ NGL production} + 0.852 \text{ petroleum processing} + 0.853 \text{ petroleum importation} + 0.138 \text{ NGL processing} + 0.495 \text{ LPG storage capacity} - 0.138 \text{ ports closure} + 0.022 \text{ LPG production} - 0.091 \text{ LPG demand} - 0.310 \text{ LPG importation.}$$

## REFERENCES

- [1] Berrah, N.; Feng, F.; Priddle, R. & Wang, L. (2007). *Sustainable energy in China: The closing window of opportunity*, p. 273. Washington, DC, International Bank for Reconstruction and Development, The World Bank. ISBN-13: 978-0-812-6754-4.
- [2] Checchi, A.; Behrens, A. & Egenhofer, C. (2009). Long-term energy security risks for Europe: a sector-specific approach. *Working Document*, 52(309). Retrieved from <http://aei.pitt.edu/10759/1/1785.pdf>.
- [3] Commission of the European Communities (2001a). Green paper: towards a European strategy for the security of energy supply. *Office for Official Publications of the European Communities*, 90. Retrieved from <https://op.europa.eu/en/publication-detail/-/publication/0ef8d03f-7c54-41b6-ab89-6b93e61fd37c/language-en>.
- [4] Cuadras, C. (2018). *Nuevos métodos de análisis multivariante*. Barcelona, Spain: CMC Editions.
- [5] Ganzo, A. (2004). Análisis de componentes principales. Introducción al análisis multivariante. Retrieved from <http://eio.usc.es/>

- eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat\_50140155\_AnalisisMultivariante.pdf.
- [6] IEA (1985). *Energy technology policy*. Paris, France: OECD Publishing.
- [7] IEA (2001). *Towards a sustainable energy future*. Paris, France: OECD. Retrieved from [https://www.oecd-ilibrary.org/energy/towards-a-sustainable-energy-future\\_9789264193581-en](https://www.oecd-ilibrary.org/energy/towards-a-sustainable-energy-future_9789264193581-en).
- [8] IEA (2007). *Energy security and climate policy-assessing interactions*. Paris, France: OECD. Retrieved from [https://www.oecd-ilibrary.org/energy/energy-security-and-climate-policy-assessing-interactions\\_9789264109940-en](https://www.oecd-ilibrary.org/energy/energy-security-and-climate-policy-assessing-interactions_9789264109940-en).
- [9] *Informe situacional de la comercialización del Gas Licuado de Petróleo (GLP) en el Perú* (2015). Lima, Peru: Gerencia de Fiscalización de Hidrocarburos Líquidos de Osinergmin.
- [10] Kruyt, B.; Vuuren, D. P. van; Vries, H. J. M. de & Groenenberg, H. (2009). Indicators for energy security. *Energy Policy*, 37(6), 2166-2181. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0301421509000883?via%3Dihub>.
- [11] Lloret, S.; Ferreres, A.; Hernández, A. & Tomás, I. (2014). Exploratory item factor analysis: a practical guide revised and up-dated. *Annals of Psychology*, 30(3), 1151-1169. Retrieved from <https://revistas.um.es/analesps/article/view/analesps.30.3.199361/217271>.
- [12] Löschel, A.; Johnston, J.; Delucchi, M.; Demayo, T.; Gautier, D.; Greene, D.; Ogden, J.; Rayner, S. & Worrell, E. (2009). Stocks, flows, and prospects of energy. In T. E. Graedel & E. van der Voet (Eds.), *Linkages of sustainability*. Oxford, Reino Unido: MIT Press Scholarship.