Degree to which natural gas (NG) has replaced other energy sources in the Principality of Asturias (Spain)

Grado en que el gas natural ha sustituido a las otras fuentes de energía en el Principado de Asturias (España)

Antonio José Gutiérrez Trashorras^{*}, Jorge Xiberta Bernat, Robert Arana Hurtado

Departamento de Energía, Universidad de Oviedo. C. Independencia 13, 33004 - Oviedo, España

(Recibido el 2 de mayo de 2009. Aceptado el 8 de enero de 2010)

Abstract

In order to determine the significant substitute products in each one of the consumption sectors in the Principality of Asturias, the correspondent bivariant and partial statistical correlations of NG against other energies in demand are calculated. The functions among the value-pair sequences of the analyzed variables are linear. The measurement of the dependence level is carried out by Pearson's linear correlation coefficient. To evaluate this coefficient and to confirm whether there exists or not a significant correlation between the variables (*p*-value<5%) a hypothesis test based on the *t*-Student distribution is done. The results obtained show a clear substitution of coal by NG.

----- *Keywords*: Natural gas, energy consumption, correlation, level of significance

Resumen

Con la finalidad de determinar los productos sustitutivos significativos en cada uno de los sectores del consumo del Principado de Asturias, se obtienen las correspondientes correlaciones estadísticas bivariadas y parciales del gas natural frente a los demás energéticos demandados. Las funciones entre las secuencias de pares de valores de las variables obtenidas son de tipo lineal. La medida del nivel de dependencia se lleva a efecto mediante el coeficiente de correlación lineal de Pearson. Para la valoración del coeficiente de correlación

^{*} Autor de correspondencia: teléfono: + 34 + 98 + 510 43 27, fax: + 34 + 98 + 510 43 22, correo electrónico: gutierrezantonio@uniovi.es (A. Gutiérrez).

y confirmar si existe o no una asociación estadísticamente significativa entre ambas variables (*p-valor* < del 5%), se realiza un test de hipótesis basado en la distribución *t* de *Student*. Los resultados obtenidos ponen de manifiesto una clara sustitución del carbón por el gas natural.

----- Palabras clave: Gas natural, consumo de energía, correlación, nivel de significación

Introduction

The increasing demand for energy at a global level shows the need to invest in the development of new energy sources that could support or substitute, in this case, these other declining energies. Coal, oil and NG represent the most consumed primary energies in the world but NG has undergone the highest degree of development in its history and the outlook for the future promises a similar growth passing from 23.3% in 2001 to 29% in 2010 according to some predictions [1]. In Spain NG holds the third place in energy consumption and is the one whose use is increasing most. NG has been the primary source of energy registering a greater relative increase in the regional energy balance, from 0.29% in 1989 to 8.2 % in 2005, since its introduction (1989) in Asturias [2].

This work analyses the evolution in the demand for this fuel against other sources of energy during the period 1989-2007. During these years along with the increase in consumption a heterogeneous variation of the demanded sources by the Industry and Diverse Uses has been shown too. This behaviour has been linked to the technological improvements implemented in these sectors together with the economical and social development in the Principality [3].

The main aims of this work are:

- 1. To identify the products that have been replaced by NG in the regional energy consumption.
- 2. To analyze the correlations between this fuel and the other sources in the energy balance of the Principality, in order to deal with future consumption forecasts in further investigations.

The concept of correlation has been applied to study the degree of joint variation that exists between two or more variables. In the particular case of the X and Y variables Pearson's linear correlation coefficient has been defined as the mathematical expression that is the quotient between the covariance and the product of the standard deviations of both variables [4].

$$r = \frac{S_{XY}}{S_X \cdot S_Y} \tag{1}$$

Methodology

If r = 1, the correlation is perfect and direct, and the functional relation between both variables is exact and positive, both varying in the same line. If r = -1, the correlation is perfect and inverse, both variables varying in the opposite line. If r= 0, the correlation is null and the variables are not associated, it being impossible to find any functional relation between them. There are two procedures that allow us to quantify the degree of linear relationship between two variables and to evaluate the degree of adjustment by a straight line, of the point cloud produced by the graphic representation of the data. These procedures are namely the partial correlations and the bivariant correlations, both based on the above mentioned concept of correlation [5].

Bivariant correlations

This procedure uses Pearson's correlation coefficient, Spearman's rho or Kendall's τ_b with their significance levels. For quantitative, normally distributed variables, Pearson's correlation coefficient is used. If the data are not normally distributed or have ordered categories,

Kendall's τ_b or Spearman's rho are used, as they measure the association between rank orders.

Partial correlations

We calculate the partial correlation coefficient that measures the degree of association between two variables and allows us to control the effects of one or more additional variables (control variables). With partial correlation the effect of the set of controlling variables is removed.

Previous study

Linear relations were found in the analysis of the value pair frequencies of the considered energy variables. Because the available data [2, 3] was of a quantitative type and normally distributed, it was inferred that the analysis by Pearson's coefficient correlation was the most satisfactory form to deal with the problem. To evaluate this coefficient and establish if there exists or not significant association among the analyzed variables (*p*-value<5%) a *t*-Student based hypothesis test was carried out. The test of significance can be two-tailed or one-tailed probabilities. If the direction of association is known in advance, one-tailed is used. Otherwise, two-tailed is preferred.

Such a contrast of hypothesis uses the *t*-contrast statistic which is distributed following a student's *t*-probability function with *N*-2 degrees of freedom, being *N* the total number of observations:

$$t = r\sqrt{\frac{N-2}{1-r^2}} \tag{2}$$

Working plan

There are several computerized programmes for bivariant and partial correlations. The Statistical Package for Social Science (SPSS Inc) has been chosen, because it is one of the most complete and most widely used for this kind of surveys [6, 7].

The following working plan is:

• Identification of the variables to study

- Value assignation for these variables
- Creation of a data base by using the program EXCEL
- Analysis of the type of relation among variables. For this, we use the graphics of the related sequences.
- Selection of the program or programs of calculation.
- Information processing is carried out by the statistical analysis of correlations.
- Results evaluation that requires the quantitative analysis of the results and the posterior qualitative study by using the significant obtained correlations.

Calculation process

By the *bivariant correlation procedure*, the correlation coefficient is established and the significance level between NG and the other selected energy products.

With the *partial correlation procedure*, the partial correlation coefficients are obtained and the significance level considering the conditioning energy products (control variables) for each pair of energy products to compare.

For example, if the energy-producing pairs are NG against black-coal, anthracite, Hydro-energy and petroleum products (PP) respectively, the total number of matrixes will be 28 (7 forms of combination of the variable control multiplied by 4 pairs of comparisons).

Results and discussion

Table 1 shows the matrix of the bivariant correlations of NG against other energies whose consumptions are relevant in the Asturian Gross internal consumption (GIC): NG, black-coal, anthracite, hydro-energy and PP [8-14]. This table shows that Pearson's correlation coefficient for the variables NG and black-coal is 0.785, with a critical level p = 0.000 which allows us to affirm that the

coefficient is significantly different from zero, that is to say, there exists a very significant positive linear relationship with black-coal (*p*-value<1%) and a relatively high coefficient. Likewise it shows a very significant and positive correlation between NG and PP with a high correlation coefficient (r = 0.858). The correlation between NG and hydro-energy is significant (*p*-value< 5%) and coefficient r = -0.539, which indicates the possible replacement between both energy-producing, with certain limitations, the coefficient being relatively low. The production of hydraulic energy and, consequently, its consumption depends directly on the amount of rainfall gathered from the regional dams and a real relationship with NG consumption cannot be guessed.

	GIC	NG	Black-coal	Anthracite	Hydro-energy	PP
NG	Pearson correlation	1	0.785ª	0.259	-0.539 ^b	0.858ª
	Sig. (2-tailed)		0.000	0.315	0.026	0,000
Black-coal	Pearson correlation	0.785ª	1	0.276	-0.759ª	0.834ª
	Sig. (2-tailed)	0.000		0.284	0.000	0.000
Anthracite	Pearson correlation	0.259	0.276	1	-0.358	0.314
	Sig. (2-tailed)	0.315	0.284		0.159	0.219
Hydro-energy	Pearson correlation	-0.539 ^b	-0.759ª	-0.358	1	-0.568 ^b
	Sig. (2-tailed)	0.026	0.000	0.159		0.017
PP	Pearson correlation	0.858ª	0.834ª	0.314	-0.568 ^b	1
	Sig. (2-tailed)	0.000	0.000	0.219	0.017	

Table 1 Matrix of bivariant correlations in the Gross Internal Consumption (1989-2007)

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

^b Correlation is significant at the 0.05 level (2-tailed).

It is important to show that a high correlation coefficient does not imply causality. Two variables can be linearly related (even highly related) without one causing the other. This happens to the same extent in the relations between NG and black-coal consumption and PP. In both cases causality does not exist, because the use of NG is due to environmental reasons, efficiency, cleanliness, and comfort use whereas black-coal consumption is mainly caused by the demand of the thermal stations for power generation. On the other hand, oil product consumption is mostly produced by the demand of the Transport sector, owing to the increase in the standard level and the improvement in the infrastructure in communications within the region and with the rest of the country.

The correlation between NG and black-coal decreases slightly when anthracite and hydroenergy are taken as determinant factors (control variables) with *p-values*<1% as is shown in the results of the partial correlations (Table 2). The correlation coefficient changes from 0.785 to 0.685 when the hydro-energy effect is controlled. This shows that eliminating the control effects of these variables on the analysed relationship: 1) there is a linear significant relationship between the NG black-coal variables and 2) such a relation does not result essentially altered (decreases slightly) after controlling the anthracite and the hydro-energy variable effects. This small variation is due to the existent relationship between black-coal and hydro-energy (r = -0.759) which are

inversely consumed for power generation (Table 1). In this case the results obtained by using PP as a control variable are not included because they are not significant.

Table 2 Matrix of the significant partial correlations in the GIC: NG vs Black-coal (1989-2007)

			Control Variables	
GIC		Anthracite Hydro-energy	Anthracite	Hydro-energy
NG & Black-coal	Pearson correlation Sig. (2-tailed)	0.687ª 0.005	0.768ª 0.001	0.685ª 0.003

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

^b Correlation is significant at the 0.05 level (2-tailed).

The results obtained after eliminating the other energy effects (partial correlations) in the relationship between NG and PP indicate a significant relationship between them but a smaller correlation coefficient is shown, changing from 0.858 to 0.595 in the most unfavourable case (black-coal). All the partial correlations are positive and significant (Table 3).

Table 3 Matrix of the significant partial correlations in the GIC: NG vs PP (1989-2007)

		Control Variables						
GIC		Anthacite Hydro-energy Black-coal	Anthacite Hydro-energy	Hydro-energy Black-coal	Anthacite Black-coal	Black-coal	Anthacite	Hydro-energy
NG&PP	Correlation Sig.	0.578 ^b 0.030	0.795ª 0.000	0.585⁵ 0.022	0.593⁵ 0.020	0.595⁵ 0.015	0.847ª 0.000	0.796ª 0.000

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

^b Correlation is significant at the 0.05 level (2-tailed).

The partial correlations corresponding to the relationship between NG and hydro-energy have resulted non-significant, for all the possibilities of combination among the control variables. This indicates that a linear relationship does not exist between both variables.

In order to determine the NG correlations in the final consumption the main types of energy with relevant final consumption: NG, coal, PP and electricity have been considered [2,3,10]. Significant correlation coefficients of all these energies have been obtained by using the bivariant correlation procedure (Table 4).

Final c	onsumption	NG	Coal	PP	Electricity
NG	Pearson correlation	1	-0.714ª	0.817ª	0.973ª
	Sig. (2-tailed)		0.001	0.000	0.000
Coal	Pearson correlation	-0.714ª	1	-0.618ª	-0.584 ^b
	Sig. (2-tailed)	0.001		0.008	0.014
PP	Pearson correlation	0.817ª	-0.618ª	1	0.743ª
	Sig. (2-tailed)	0.000	0.008		0.001
Electricity	Pearson correlation	0.973ª	-0.584 ^b	0.743ª	1
	Sig. (2-tailed)	0.000	0.014	0.001	

Table 4 Matrix of the bivariant correlations in the final consumption (1989-2007)

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

^b Correlation is significant at the 0.05 level (2-tailed).

By using the bivariant correlation procedure a negative correlation coefficient (-0.714) and a very significant *p*-value (0.001) are obtained between NG and coal. The results show a minimal influence of PP and electricity joint-action, according to the correlation coefficient (-0.769) and *p*-value (0.001). A similar result is reached when electricity is the determining factor (correlation coefficient of -0.781 and

p-value of 0.000). In all these significant cases, the correlation coefficients between NG and coal are negative, contrary to what happens with NG and the other energies (Table 5). It is to be noted that 70% of the final gas consumption is used in the industrial sector. In this field it competes against coal replacing it to a great extent because of its efficiency and for environmental reasons.

Table 5 Matrix of the partial correlations in the final consumption (1989-2007)

			Control Variables	
Final co	nsumption	PP Electricity	Electricity	PP
	Pearson correlation	-0.769ª	-0.781ª	-0.462
NG & Coal	Sig. (2-tailed)	0.001	0.000	0.072
		Coal Electricity	Coal	Electricity
	Pearson correlation	0.586 ^b	0.682ª	0.609 ^b
NG & PP	Sig. (2-tailed)	0.022	0.004	0.012
		Coal PP	PP	Coal
	Pearson correlation	0.973ª	0.948ª	0.978ª
	Sig. (2-tailed)	0.000	0.000	0.000

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

^b Correlation is significant at the 0.05 level (2-tailed).

Between NG and PP there is a high positive correlation coefficient (r = 0.817) and a significant p-value (0.000) by means of the bivariant correlation procedure. In this case, the degree of correlation varies with the ot influence exerted by other energies (Table 5), decreasing to 0.586 when the coal and electricity joint-effect is eliminated, with a significant *p-value* of 0.022. This indicates a certain linear relationship between NG and PP but with a relatively low correlation coefficient. As explained above, this relationship does not imply causality.

The bivariant correlation (Table 4) between NG and electricity is very high and significant

(r = 0.973 and p-value = 0.000). There is a high, significant correlation between NG and electricity independent of the other energies also under consideration.

The main types of energy with relevant final consumption in the industrial sector are: NG, coal, PP and electricity. The *p-value* correspondent to the relationship between NG and PP is 0,181, that means that it is greater than 0.05, indicating the lack of significant correlation between both energy producers (Table 6). Similar results are obtained by calculating the partial correlations of this case with all the control variables showing that no significant correlation exists.

Industria	l consumption	NG	Coal	PP	Electricity
NG	Pearson correlation	1	-0.687ª	-0.340	0.929ª
	Sig. (2-tailed)		0.002	0.181	0.000
Coal	Pearson correlation	-0.687ª	1	0.243	-0.418
	Sig. (2-tailed)	0.002		0.347	0.095
PP	Pearson correlation	-0.340	0.243	1	-0.401
	Sig. (2-tailed)	0.181	0.347		0.111
Electricity	Pearson correlation	0.929ª	-0.418	-0.401	1
	Sig. (2-tailed)	0.000	0.095	0.111	

Table 6 Matrix of the bivariant correlations in the industrial sector consumption (1989-2007

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

For NG and coal very significant *p*-values and negative correlations are obtained both by the bivariant procedure and when the partial correlations were used (Tables 6 and 7). In the latter, the high correlations and significant *p*-values indicate that the link between NG and coal increases when the electricity effect is eliminated. This replacing effect is because the higher industrial gas demand source is from Iron and Steel Metallurgy and Casting. New factories, using NG as a fuel, have been created in this subsector and others using coal have been reformed to consume NG because of its higher efficiency and lesser pollution, that better meets the Kyoto Protocol requirements. [2, 3, 15, 16].

Between NG and electricity a high correlation coefficient (0.929) is recorded and is very significant (*p*-value = 0.000) (Table 6). In the process of partial correlations, with the other energies as determining factors the *p*-values are maintained and high correlation coefficients are obtained as well (Table 7). The analysis of the obtained results shows the lack of significant

correlations between coal and PP, between coal and electricity as well as between PP and electricity (Table 6). In the Diverse Use sector the main types of energy with a relevant final consumption are: NG, coal, PP, electricity, LPG and gas-oils.

			Control Variables	;
Industria	l consumption	PP Electriciy	PP	Electricity
NG&Coal	Pearson correlation Sig. (2-tailed)	-0.903ª 0.000	-0.663ª 0.005	-0.887ª 0.000
		Coal PP	PP	Coal
NG&Electricity	Pearson correlation Sig. (2-tailed)	0.974ª 0.000	0.920ª 0.000	0.972ª 0.000

Table 7 Matrix of the partial correlations in the industrial sector consumption (1989-2007)

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

^b Correlation is significant at the 0.05 level (2-tailed).

There is a high significant correlation between NG and coal, because the inverse correlation coefficient is reasonably high (-0.847) with a

very significant high-value (0.000) confirmed by results after eliminating the gas-oil and LPG influence (Tables 8 and 9).

Table 8 Matrix of bivariant correlations in the diverse use sector consumption (1989-2007)

Diverse	use consumption	NG	Coal	Gas-oil	Electricity	LPG
NG	Pearson correlation	1	-0.847ª	-0.656ª	0.978ª	-0.656ª
	Sig. (2-tailed)		0.000	0.004	0.000	0.004
Coal	Pearson correlation	-0.847ª	1	0.224	-0.885ª	0.255
	Sig. (2-tailed)	0.000		0.388	0.000	0.324
Gas-oil	Pearson correlation	-0.656ª	0.224	1	-0.564 ^b	0.826ª
	Sig. (2-tailed)	0.004	0.388		0.018	0.000
Electricity	Pearson correlation	0.978ª	-0.885ª	-0.564 ^b	1	-0.593 ^b
	Sig. (2-tailed)	0.000	0.000	0.018		0.012
LPG	Pearson correlation	-0.656ª	0.255	0.826ª	-0.593 ^b	1
	Sig. (2-tailed)	0.004	0.324	0.000	0.012	

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

^b Correlation is significant at the 0.05 level (2-tailed).

		Control Variables						
Dive consi	rse use umption	Gas-oil Electricity LPG	Electricity LPG	LPG Gas-oil	Gas-oil Electricity	Electricity	Gas-oil	LPG
NG&	Correlation	-0.475	-0.208	-0.960ª	-0.432	0.201	-0.952ª	-0.931ª
Coal	Sig.	0.086	0.457	0.000	0.108	0.456	0.000	0.000
		LPG Coal Electricity	Coal Electricity	Electricity LPG	LPG Coal	Coal	Electricity	LPG
NG&	Correlation	-0.602 ^b	-0.685ª	-0.460	-0.675ª	-0.900ª	-0.612 ^b	-0.268
Gas-oil	Sig.	0.023	0.005	0.085	0.006	0,000	0.012	0.315
		LPG Coal Gas-oil	Coal Gas-oil	Gas-oil LPG	LPG Coal	Coal	Gas-oil	LPG
NG&	Correlation	0.702ª	0.769ª	0.974ª	0.753ª	0.924ª	0.976ª	0.970ª
Electricity	Sig.	0.005	0.001	0.000	0.001	0.000	0.000	0.000
		Coal Gas-oil Electricity	Gas-oil Electricity	Electricity Coal	Coal Gas-oil	Gas-oil	Coal	Electricity
NG&	Correlation	-0.220	-0.004	-0.458	-0.483	-0.268	-0.856ª	-0.455
LPG	Sig.	0.450	0.989	0.086	0.068	0.315	0.000	0.077

Table 9 Matrix of partial correlations in the diverse use sector consumption (1989-2007)

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

^b Correlation is significant at the 0.05 level (2-tailed).

There is a positive. high and significant correlation between NG and electricity. with a correlation coefficient of 0.978 and *p-value* of 0.000. Such a coefficient decreases after eliminating the joint-influence of gas-oil. LPG and coal up down to the value of 0.702 and *p-value* of 0.005. As can be deduced from this and the last instance, there is a significant correlation between NG and electricity (Tables 8 and 9).

The improvement in household appliances and the increase in the number of connections to the gas pipeline and electricity grids give advantages to NG and electricity with respect to coal. The replacing effect of coal by NG increases after eliminating the gas-oil and LPG effect. because these energy-producers are still being used in rural areas where there is difficulty in accessing the gas supply lines.

Conclusions

In Asturias, the historical consumption of NG has undergone the highest relative continuous increase with regard to the other energies, with an annual average of 24% during the period 1989-2007. since its incorporation into the regional market. Its behaviour shows very significant negative correlation with coal, that means that the latter is being significantly replaced by NG.

Significant correlation coefficients (*p-values* <5%) between NG and the currently used energies in the GIC (coal. PP and electricity) were obtained by using the bivariant correlations procedure.

These relationships do not imply causality. In all the cases, the correlation coefficients between NG and coal are negative, contrary to what is happening between NG and other energies. It is evident that NG is replacing coal in the final regional consumption.

In the industrial sector, analysing the correlations between NG and coal very significant *p*-values and negative correlations were obtained, both with the bivariant procedure and with the partial correlations making it clear the replacement of coal by NG was due to the better industrial benefits brought by NG.

In the diverse use sector, there is a high significant correlation between NG and coal. with an inverse-type correlation coefficient of -0.8 and almost null *p-values*. Such correlation decreases slightly under the influence of gas-oil and LPG as determining factors, owing to their use in the most remote rural areas, so, coal is being replaced by NG in this sector too. An increase in the replacement of coal by NG in GIC is expected in the near future when a regasifier and various combined cycles are to be implemented.

References

- CNE. (National Commission of the Energy). *Basic* information on the energy sectors 2002. Madrid. 2003. pp. 273.
- J. Xiberta, M.P. Gancedo, H. Llaneza, S. Gonzales. *Energetic balances in Asturias-Spain- for the years* 1985-2007. Spain: Department of Energy. University of Oviedo. 1986-2008. (C/ Independencia. 13-33004 Oviedo. Spain.). pp. 324.
- M. P. Gancedo. Evolution of the energetic plan of Asturias-Spain 1995-2005. Spain: Department of Energy. University of Oviedo. 2006. (C/Independencia. 13-33004 Oviedo. Spain.). pp. 433.

- G. Canavos. Applied probability and statistical methods. Ed. Mc Gaw-Hill. México. 1988. pp. 53-97.
- L. Meyers. Applied multivariate research: design and interpretation. Ed. SAGE cop. London. 2006. pp. 12-143.
- 6. Q. Trends. *Statistical package for social sciences*, SPSS Inc. Chicago. 1994. pp. 15-69.
- C. Pérez López. Statistical techniques tools with SPSS 12. Ed. Prentice Hall. Madrid: 2004. pp. 33-157.
- IEA (International Energy Agency). *Statistical data.* 2003. Paris. 2003. (9 rue de la Fédération-75015. Paris. France). pp. 32.
- IEA (International Energy Agency). Statistical data. 2004. Paris. 2004. (9 rue de la Fédération-75015. Paris. France). pp. 36.
- J. Xibertat, H. Llaneza, S. Gonzales, E. Fernández. *Energetic plan of Asturias-Spain 1995-2005*. Spain: Department of Energy. University of Oviedo. 1996. (C/ Independencia. 13-33004 Oviedo. Spain.). pp. 242.
- 11. Fernández-Galiano L. *Spain 2006-Yearbook.* Ed. Arquitectura Viva. Madrid. 2006. pp. 34-41.
- J. Suarez García. Energetic balance of Asturias. natural gas incidence. Sedigas. Barcelona. 1988. pp. 7-59.
- CNE (National Commission of the Energy). Framework information of electric energy and NG and its coverage. 2001. Madrid. February-2002. (C/ Alcalá. 47-28014 Madrid. Spain.). pp. 395.
- CNE (National Commission of the Energy). Framework information of electric energy and NG and its coverage. 2002. Madrid. February-2003. (C/ Alcalá. 47-28014 Madrid. Spain.). pp. 321.
- SEDIGAS (Spanish Industrial Technical Association) Gas yearbook 2002. Barcelona. 2003. (C/ Balmes. 357-08006 Barcelona. Spain). pp. 104.
- SEDIGAS (Spanish Industrial Technical Association) Gas yearbook 2003. Barcelona. 2004. (C/ Balmes. 357-08006 Barcelona. Spain). pp. 115.