

## Structural equation models - PLS in engineering sciences: a brief guide for researchers through a case applied to the industry

Villalva A. Juan Enrique

ORCID: <https://orcid.org/0000-0002-6273-1154>

Juanev99@gmail.com

UNEXPO Puerto Ordaz - CVG Ferrominera Orinoco  
Edo. Bolívar, Venezuela

Received (08/04/21), Accepted(18/05/21)

**Abstract:** Modeling using structural equations, is a second generation statistical data analysis technique, it has been positioned as the methodological options most used by researchers in various fields of science. The best known method is the covariance-based approach, but it presents some limitations for its application in certain cases. Another alternative method is based on the variance structure, through the analysis of partial least squares, which is an appropriate option when the research involves the use of latent variables (for example, composite indicators) prepared by the researcher, and where it is necessary to explain and predict complex models. This article presents a brief summary of the structural equation modeling technique, with an example on the relationship of constructs, sustainability and competitiveness in iron mining, and is intended to be a brief guide for future researchers in the engineering sciences.

**Keywords:** Competitiveness, Structural equations, Iron mining, Sustainability.

## Modelos de ecuaciones estructurales - PLS en ciencias de la ingeniería: una breve guía para investigadores a través de un caso aplicado a la industria

**Resumen:** El modelado mediante ecuaciones estructurales, es una técnica de análisis de datos estadísticos de segunda generación, se ha posicionado como la opción metodológica más utilizada por investigadores en diversos campos de la ciencia. El método más conocido es el basado en la covarianza, pero presenta algunas limitaciones para su aplicación en determinados casos. Otro método alternativo se basa en la estructura de varianzas, mediante el análisis de mínimos cuadrados parciales, que es una opción adecuada cuando la investigación implica el uso de variables latentes (por ejemplo, indicadores compuestos) elaboradas por el investigador, y donde es necesario explicar y predecir modelos complejos. Este artículo presenta un breve resumen de la técnica de modelado de ecuaciones estructurales, con un ejemplo sobre la relación de constructos, sostenibilidad y competitividad en la minería del hierro, y pretende ser una breve guía para futuros investigadores en las ciencias de la ingeniería.

**Palabras Clave:** Competitividad, Ecuaciones estructurales, Minería de hierro, Sostenibilidad.



## I. INTRODUCTION

Latent variables or constructs are present in everyday life more than we realize, although we use them daily, examples of which are: happiness, intelligence, poverty, etc. Also at the level of engineering disciplines, constructs are present, examples of which are sustainability, environmental performance, competitiveness, corporate social responsibility, quality of service, capacity for innovation, among others, which need to be measured and evaluated for the problem diagnosis and decision making.

The structural equation model (Structural Equation Modeling, SEM) is a multivariate method that allows simultaneously evaluating the dependency relationships between observable and unobservable variables (constructs). With this technique, research models are carried out through the transformation of theoretical concepts into unobservable variables and the transformation of empirical concepts into indicators, both are related through the hypothesis expressed graphically by path diagrams. The SEM method can be applied using two alternatives: SEM based on the structure of covariance (BC) or SEM based on the structure of variance, through partial least squares analysis (PLS).

The origin of the BC SEM dates back to 1973, when Karl Jöreskog introduced a maximum likelihood algorithm for estimating models of covariance structures [1]. The Swedish professor Herman Wold, criticized the dependence of the distribution assumptions, which affects the validity of the empirical results, and proposed an alternative approach, Partial Least Squares (PLS), and in 1977 he developed the algorithm NIPALS (Nonlinear Iterative Partial Least Squares) [2]. The BC SEM is based on the assumption of normality of the variables and uses a maximum likelihood estimate, requires a generally large sample and is focused on the “reproduction” of the structure of relationships between variables.

This article is based on the PLS SEM, which has advantages for its application. This modeling method is more flexible by not requiring rigorous parametric assumptions. PLS SEM does not assume normality and is estimated by recursive least squares, it is applicable with small samples and is focused on prediction. The mathematical and statistical procedures underlying the PLS SEM are rigorous and robust [2].

PLS SEM is the iterative combination of principal component analysis, path analysis, and Ordinary Least Squares (OLS) regression. The principal component analysis links the observable variables with the constructs, the path analysis allows the construction of the structure of the system of variables and the OLS regression allows the estimation of the parameters. It is important to highlight that PLS SEM can be used for both explanatory (confirmatory) and predictive (exploratory) research [1], [3].

In engineering sciences, relatively few researchers have also begun to successfully exploit the potential of PLS SEM to obtain relevant results in their analyzes. In this article, after proposing the theoretical definitions and the procedure, to illustrate the application of the PLS SEM, a case of an multidisciplinary nature of mining, industrial and environmental engineering is developed, on the relationship of the constructs Sustainability and Competitiveness in the iron mining industry, demonstrating the high applicability of this novel technique for the development of models in the field of engineering.

## II. THEORETICAL ASPECTS

The general model of structural equations consists of a measurement model, also called an external model, and a structural model or internal model. The measurement model specifies the relationships between the observable variables and the latent variables that underlie them. On the contrary, the structural model specifies the relationships between the latent variables, which in turn consist of exogenous variables or constructs ( $\eta$ ) and endogenous variables or constructs ( $\xi$ ). In Figure 1, a schematic of the general model of a PLS SEM is presented. In the context of PLS SEM, you can work with two types of measurement models: (1) the reflective model; and (2) the training model.

The measurement model is governed by two equations; one that measures the relationships between endogenous latent variables and their observable variables.

$$y = \Lambda_y \eta + \varepsilon \quad (1)$$

Where:

$y$  is the vector of  $p$  observable variables ( $px1$ )

$\lambda y$  is the matrix of coefficients that show the relationships between the latent and observed variables ( $pxm$ ).

Also called the charge matrix ( $\lambda$ ).

$\epsilon$  is the error vector ( $px1$ ).

The second equation of the measurement model is the one that governs the relationships between the exogenous latent variables and their observable variables:

$$x = \Lambda_x \xi + \delta \quad (2)$$

Where:

$x$  is the vector of  $p$  observable variables ( $qx1$ )

$\lambda x$  is the matrix of coefficients that show the relationships between the latent and observed variables ( $qxm$ ).

Also called a matrix of weights ( $\pi$ ).

$\delta$  is the error vector ( $qx1$ ).

The structural model is defined by the equation:

$$\eta = \beta \eta + \tau \xi + \zeta \quad (3)$$

Where:

$\eta$  represents the vector of endogenous latent random variables of dimension  $m \times 1$ .

$\xi$  represents the vector of exogenous latent random variables of dimension  $n \times 1$ .

$\beta$  represents the matrix of coefficients that govern the relationships between the endogenous variables  $m \times m$ .

$\tau$  represents the matrix of coefficients that govern the exogenous relationships and each of the endogenous ones, or in other words, the effects of  $\xi$  on  $\eta$ . Its dimension is  $m \times n$ .

$\zeta$  represents the vector of disturbances or errors.

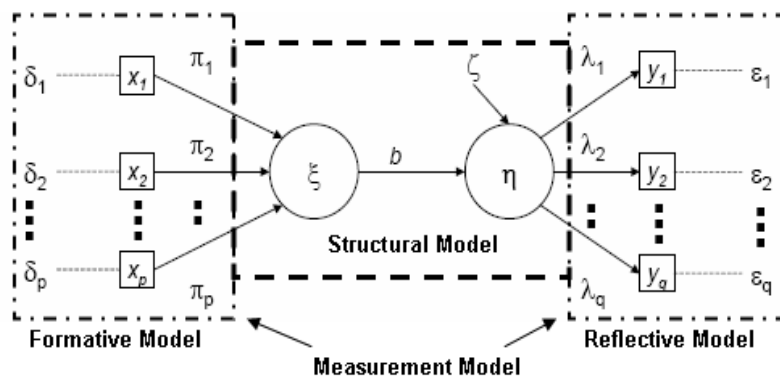


Fig. 1. General model schematic of a PLS SEM

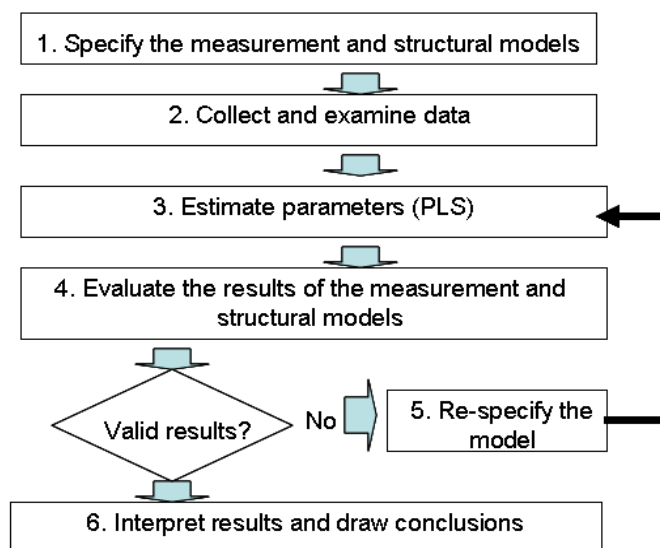
Source: Adapted from Cepeda and Roldán [4]

### III. PROCEDURE TO APPLY PLS SEM

The procedure to apply PLS SEM is illustrated in Figure 2, there are six steps:

1. Specify the measurement and structural models.
2. Collect and examine the data.

3. Estimate the PLS parameters.
4. Evaluate the results of the measurement and structural models.
5. Re-specify the model.
6. Interpret the results and draw conclusions.



**Fig. 2. cheme of the procedure to apply the PLS SEM.**

#### *A. Specify the measurement and structural models.*

The researcher applies the theoretical knowledge of the studied phenomena to the formulation of mathematical expressions related to the relationships between latent variables, and their relationships with their indicators or observable variables.

#### *B. Collect and examine data.*

The data collection and examination stage is very important in the SEM application and can avoid delay, especially when careful examination of the data manages to rid the data of outliers and identify missing data. The first step in dealing with outliers is to identify them, standard statistical software packages offer a multitude of statistical tools, which allow you to identify outliers. In general, each PLS SEM software offers ways to handle the missing data, the most common are: substituting the mean of the valid values of that indicator or eliminating the cases that include missing values.

#### *C. Estimate parameters (PLS).*

Once specified, the structural and measurement parameters of a PLS SEM model are estimated by the software (in our case the SmartPLS) iteratively using simple Ordinary Least Squares (OLS) and multiple regressions. In a much summarized way it would be the following sequence of iterations:

1. In the first iteration of PLS, you get an initial value for  $\eta$  (adding the values  $y_1 \dots, y_q$ ).
2. Estimation of the regression weights  $\pi_1 \dots, \pi_p$  (regression of the value of  $\eta$  with  $x_1 \dots, x_p$ ).
3. Estimates of  $\pi_1 \dots, \pi_p$  in linear combination with  $x_1 \dots, x_p$  resulting in an initial value for  $\xi$ .
4. Estimates of the charges  $\lambda_1 \dots, \lambda_q$  by a series of simple regressions of  $y_1 \dots, y_q$  on  $\xi$ .
5. The estimated charges  $\lambda_1 \dots, \lambda_q$ , in linear combination with  $y_1 \dots, y_q$ , obtain a new estimate of the value of  $\eta$ .

This procedure continues until the difference between consecutive iterations is extremely small, according to the criterion selected by the researcher [4].

#### *D. Evaluate the results of the measurement and structural models.*

To evaluate the results, it is necessary to verify and validate the goodness of fit of the models. There is no global fit coefficient available in PLS SEM and not all measures are appropriate to assess all types of fit [5]. The

validation of the SEM model through the PLS statistical tool requires a series of parameters that are estimated in two stages: the measurement model and the structural model [6], [7].

Validation of the Measurement Model. This is done with respect to the validity and reliability attributes of the model. This implies verifying: i) individual item reliability, ii) internal consistency, iii) convergent validity, and iv) discriminant validity [7], [8].

Item Reliability. The criterion for an item to be considered in the composition of the variables is that it must load at least 0.5 in the factor [9]. In this sense, it is considered that the individual reliability of the item is assessed by examining the loads ( $\lambda$ ) or simple correlations. Another more demanding criterion to accept an indicator is that it has a load equal to or greater than 0.707 ( $\lambda^2$ , 50% of the variance is explained) [4].

*Internal Consistency (Construct Reliability).* The reliability of a construct makes it possible to check the internal consistency of all the indicators when measuring the concept, that is, it is evaluated how rigorously the observable variables are measuring the same latent variable (Roldán, 2004). Construct reliability can be verified using composite reliability and Cronbach's alpha. Composite reliability is a preferred alternative to Cronbach's alpha as a test of convergent validity in a reflective model; Cronbach's alpha may overestimate or underestimate the reliability of the scale. In a model suitable for exploratory purposes, the composite reliabilities should be equal to or greater than .6 [10], [11]; equal to or greater than .70 for a suitable model for confirmation purposes [12]; and equal to or greater than .80 is considered good for confirmatory research [13]. The compound reliability measure (ICC) is given by the following mathematical expression:

$$ICC = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum_i var(\varepsilon_i)} \quad (4)$$

donde  $\lambda_i$  = carga estandarizada del indicador i,  $\lambda_i$  = error de medida del indicador i, y  $var(\varepsilon_i) = 1 - \lambda_i^2$  [14].

Convergent Validation. It determines if the different items destined to measure a concept or construct really measure the same thing, then the adjustment of these items will be significant and they will be highly correlated [6]. The assessment of convergent validity is carried out by means of the measure developed by Fornell and Larcker (1981) called the mean extracted variance (Average Variance Extracted: AVE) [4]. AVE measures the amount of variance that a construct obtains from its indicators in relation to the amount of variance due to the measurement error, its formula being the following:

$$AVE = \frac{(\sum \lambda_i^2)}{(\sum \lambda_i^2) + \sum_i var(\varepsilon_i)} \quad (5)$$

Where,  $\lambda_i$  = standardized load of indicator i,  $\lambda_i$  = measurement error of indicator i, and  $var(\varepsilon_i) = 1 - \lambda_i^2$  [14].

This statistic can be interpreted as a measure of construct reliability and as a measure of the evaluation of discriminant validity [15]. The mean extracted variance is recommended to be greater than 0.50, which establishes that more than 50% of the variance of the construct is due to its indicators [14].

*Discriminant validity.* It indicates to what extent a given construct is different from others in a research model [6]. Therefore, establishing discriminant validity implies that a construct is unique and captures phenomena not represented by other constructs in the model [1]. For there to be discriminant validity in a construct, there must be weak correlations between it and other latent variables that measure different phenomena [7].

Traditionally, researchers have relied on two measures of discriminant validity: the Fornell-Larcker discriminant validity criterion and crossloads. According to the Fornell-Larcker criterion, for any latent variable, the square root of the AVE must be greater than its correlation with any other latent variable. In a good model, the indicators load well in their expected factors and the cross loads with other factors that they should not measure should be low, as a general rule it is understood that the expected loads should be greater than .7 (some use .6)

and cross loads must be below .3 (some use .4) [5].

*Validation of the Structural Model.* In this phase, it must be verified whether the amount of variance of the endogenous variable is explained by the constructs that predict it. The appropriate model fit criteria are summarized in the following aspects:

- R-square
- R-squared change and the f-squared effect of exogenous factors
- Structural path coefficients
- Predictive relevance (q-square)
- Multicollinearity

The R-square, also called the coefficient of determination, is the measure of the overall effect size for the structural model, it indicates the % of the variance in the variable that is explained by the model. The explained variance of the endogenous variables ( $R^2$ ) should be greater than or equal to 0.1 [16].

To assess the validity of the structural model, changes in  $R^2$  can also be explored to determine whether the influence of a particular latent variable on a dependent construct has a substantive impact [10]. The importance of the effect  $f^2$  can be calculated with the following expression:

$$f^2 = \frac{R^2_{included} - R^2_{excluded}}{1 - R^2_{included}} \quad (6)$$

where  $R^2_{included}$  and  $R^2_{excluded}$  represent the  $R^2$  provided by the dependent latent variable when the predictor variable is used or omitted in the structural equation respectively [10]. The  $f^2$  levels of 0.02, 0.15 and 0.35 are respectively a small, medium or large effect.

The structural path coefficients (loads) vary from 0 to 1, for standardized data. These loads must be significant. The significance level is determined from the Student t value derived from the resampling or bootstrapping process, which is a non-parametric technique (there are no initial parameters; it is tested if the paths between variables are feasible) [17].

The predictive relevance check is performed using a procedure called "blindfolding", to determine the  $Q^2$  coefficients. This procedure omits part of the data when estimating a dependent latent variable from other independent latent variables, and then attempts to estimate those data using the previously estimated parameters. The process is repeated until each omitted data has been estimated.

The  $Q^2$  (Stone-Geisser validated redundancy measures), predicts the punctual indicators in the endogenous reflexive measurement models and the constructs (the  $Q^2$  does not apply to the endogenous formative constructs). This criterion refers to the fact that the model must have the ability to predict reflective indicators of endogenous latent variables [18].

For its calculation of  $Q^2$ , an omission distance "D" is taken that is not a divisor of the sample size. "D" corresponds to the number of cases omitted in the sample that must be estimated. Generally in existing PLS software packages, the default distance is between 5-10.

A good model demonstrates predictive relevance when  $Q^2$  is greater than zero [10]. For values close to .02 it represents a "small" relevance size, .15 represents a "medium" relevance size and .35 represents a "high" relevance size [19]. The Stone-Geisser  $Q^2$  measurement is based on the following parameters:

$$Q^2 = 1 - \frac{SSE}{SSO} \quad (7)$$

Where:

SSE = sum of squares of prediction error

SSO = sum of squares of observations



*Multicollinearity* is a problem in the reflective or formative models, as well as the structural model, for the same reason that it is in the OLS regression models. To evaluate multicollinearity, it is performed through the variance inflation factor coefficients (VIF) and / or the tolerance is equal to 1.0 minus  $R^2$ . In a well-fitted model, the structural VIF coefficients should not exceed 4.0 (some use the more lenient criterion of 5.0), and tolerance  $<.20$  indicates possible multicollinearity [1]. This is equivalent to saying that  $R^2 > .80$  suggests a possible multicollinearity problem [5].

#### E. Re-specify the model.

On rare occasions, the proposed model is the one that best fits initially in the first run, so it is very common to re-specify it, which consists of adding or removing parameters from the model. These modifications must respond to theoretical justifications, and not to desirable empirical justifications.

#### F. Interpret results and draw conclusions.

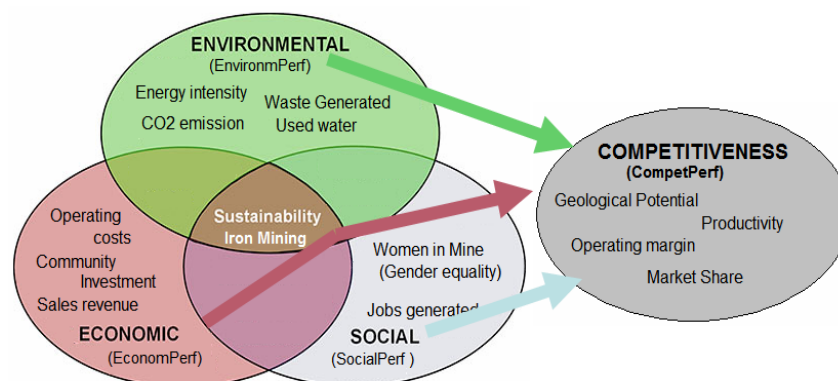
As a last step, the simple regression coefficients between the scores of the components of  $\xi$  and  $\eta$  are analyzed, where the results and statistical significance of the relationships between latent variables that make up the hypotheses are analyzed, in order to check whether they were accepted or not in the study.

### IV. Case Study: Evaluation Of The Effect Of Sustainability Dimensions On Competitiveness In Iron Mining

#### A. Measurement and structural models.

To study the effect of the dimensions of sustainability on competitiveness in iron mining, the dimensions specified in the model proposed by the ICM (‘‘International Council on Mining and Metals’’) and the GRI (‘‘Global Reporting Initiative’’). This model establishes three dimensions for sustainability in economic, environmental and social terms. The economic dimension refers to the impacts of the organization on the economic conditions of its stakeholders and on economic systems at the local, national and global levels. The environmental dimension refers to the impact of the organization on natural systems, including land, air, water, and ecosystems. The Environmental Category covers impacts related to energy, water, emissions and waste. The social dimension refers to the impacts that the organization has on the social systems in which it operates [20]. In Figure 3, the indicators and dimensions of the sustainability and competitiveness of iron mining are illustrated, the variables that were commonly reported by mining companies in their sustainability reports were selected as indicators.

To define the competitiveness construct of iron mining, the variables used by prominent authors in the area of business competitiveness and mining competitiveness were taken.



**Fig. 3. Scheme of indicators of the dimensions of sustainability and competitiveness of iron mining.**

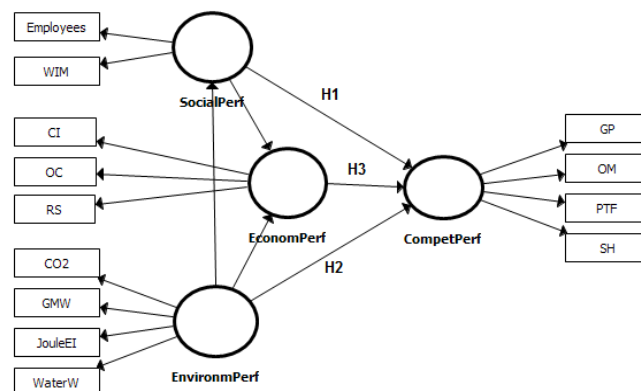
Table 1 shows the variables and their units. Figure 4 shows the diagram of the measurement and structural models. The hypotheses were as follows:

- H1: EconomPerf has a positive and significant effect on CompetPerf.

- H2: EnvironmPerf has a positive and significant effect on CompetPerf.
- H3: SocialPerf has a positive and significant effect on CompetPerf.

**Table 1. Description of the observable variables**

	<b>Description</b>	<b>Dimension</b>	<b>United</b>
JouleEI	Energy intensity	Environment	Gjoules/Tm produc
WIM	Women in Mina	Social	% of total Employees
PTF	Productivity	Competitiveness	Tm / Inputs
GP	Geological Potential	Competitiveness	% Fe, Reserves and Mine Weighted Useful Life
CO2	CO2 emission	Environment	Kg./Ton produc
OM	Operating margin	Competitiveness	EBIT/Revenues
GMW	Waste Generated	Environment	Tm Waste / Tm MENA
CI	Community Investment	Economic	% Profit before tax, depreciation and a mortization
OC	Operating costs	Economic	Costs Supplies and Serv.
RS	Sales revenue	Economic	Income / Tm sold
Employ	Jobs generated	Social	Employee/KTm
WaterW	Used water	Environment	M3 water / Tm produc
SH	Market Share	Competitiveness	Tm Sold / Global Tm



**Fig. 4. Diagram of the measurement and structural models**

#### *B. Data collected from the indicators.*

For data collection, annual data were taken from the reports of eight mining companies, national and international organizations. The mining companies were Assmang (South Africa), CAP (Chile), IOC (Canada), Kumba (South Africa), LKAB (Sweden), Rio Tinto (Australia), Vale (Brazil) and Ferrexpo (Ukraine). These companies represented approximately 45% of the world market for 2019. Table 2 presents the descriptive statistics of the data collected.

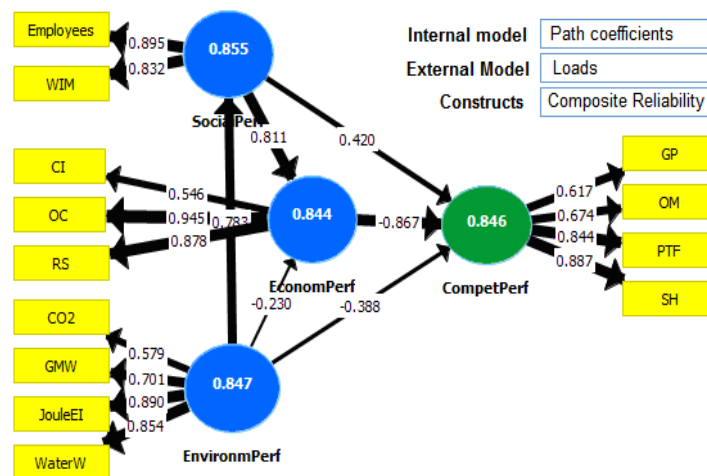


**Table 2. Descriptive statistics of the data of the observable variables**

	Mean	Min	Max	D. estánd	Asymmetr y	Kurtosis
JouleEI	0.64	0.20	1.77	0.45	1.458	1.013
WIM	17.11	7.30	29.00	6.77	0.008	-1.075
PTF	0.02	0.01	0.04	0.01	0.15	-0.562
GP	10.34	1.30	50.79	14.05	2.104	3.171
CO2	43.03	15.60	80.16	17.46	0.198	-1.094
OM	0.30	-0.07	0.60	0.16	-0.321	-0.646
GMW	3.55	0.78	9.98	2.30	0.664	-0.027
CI	0.04	0.00	0.35	0.07	3.214	11.395
OC	47.05	24.06	85.36	16.95	0.505	-0.83
RS	81.85	40.63	143.50	23.72	0.465	-0.466
Employees	0.24	0.03	0.88	0.25	1.904	2.371
WaterW	0.82	0.25	1.64	0.40	1.012	-0.406
SH	5.81	0.56	21.82	7.81	1.233	-0.374

### C. Estimated PLS parameters of the models.

For PLS calculations, SmartPLS version 3.2.7 software was used. In Figure 5, the model is presented with the values of the loads of the measurement model, the path coefficients of the structural model and the composite reliability values of the latent variables.



**Fig. 5. Models with the estimated parameters: Loads, Coefficients and composite reliability indices.**

### D. Evaluation of the results

*Validity of the Measurement Model.* The individual reliability of each of the items is assessed by examining the loads ( $\lambda$ ), as can be seen in Figure 5, all the loads are greater than 0.5, which satisfies the criterion of the minimum required load value  $\lambda > 0.5$ . Table 3 shows the statistical significance of all loads ( $\lambda$ ), it is observed that they are significant; therefore all items are accepted as valid.

*Internal Consistency and Convergent Validity.* Table 4 shows the results of the composite reliability index and the mean extracted variance (AVE) for each latent variable. The measurement model is considered to have internal consistency and convergent validity, since the composite reliabilities are greater than .80 and the AVEs are greater than 0.5.

**Table 3. Statistical significance of the loads of the observable variables**

<b>Loads (Standardized <math>\lambda</math>)</b>	<b>Statistics t</b>	<b>P Values</b>
CI <- EconomPerf	5.056	0.000
CO2 <- EnvironmPerf	5.748	0.000
Employees <- SocialPerf	19.316	0.000
GMW <- EnvironmPerf	4.161	0.000
GP <- CompetPerf	4.611	0.000
JouleEI <- EnvironmPerf	26.370	0.000
OC <- EconomPerf	60.485	0.000
OM <- CompetPerf	7.137	0.000
PTF <- CompetPerf	30.127	0.000
RS <- EconomPerf	19.277	0.000
SH <- CompetPerf	19.428	0.000
WIM <- SocialPerf	7.834	0.000
WaterW <- EnvironmPerf	8.941	0.000

**Table 4. Internal consistency and convergent validity**

	<b>Composite reliability</b>	<b>AVE</b>
CompetPerf	0.846	0.583
EconomPerf	0.844	0.654
EnvironmPerf	0.847	0.587
SocialPerf	0.855	0.747

*Discriminant validity.* There are different criteria for determining the discriminant validity, among which are the analysis of the extracted variance (AVE) and the cross loads. Table 5 shows the correlation matrix between constructs, where the diagonal shows that the square root of the extracted variance is greater than the shared variance between constructs, therefore, according to the Fornell-Larcker criterion, can affirm that there is discriminant validity. In Appendix A1, the cross-load matrix is presented, which also confirms the discriminant validity of the measurement model.

**Table 5. Discriminant validity (Fornell-Larcker criterion).**

<b>Constructs</b>	<b>CompetPerf f</b>	<b>EconomPerf f</b>	<b>EnvironmPerf f</b>	<b>SocialPerf f</b>
CompetPerf	0.764			
EconomPerf	-0.759	0.809		
EnvironmPerf	-0.410	0.405	0.766	
SocialPerf	-0.430	0.631	0.763	0.864

**Validity of the Structural Model.** For the evaluation of the structural model, the collinearity coefficients, the magnitude and statistical significance of the path coefficients, the effect sizes  $f^2$  and the predictive relevance  $Q^2$  were verified.

The  $R^2$  values  $<0.8$  and tolerance  $> 0.2$ , shown in Table 6, indicate the absence of multicollinearity, this is also corroborated with the VIF values shown in Appendix A2, by satisfying the criterion of VIF  $<4.0$ .

**Table 6. Reliability and construct validity**

	$R^2$	Tolerance
CompetPerf	0.636	0.3869
EconomPerf	0.419	0.581
EnvironmPerf		
SocialPerf	0.613	0.395

Table 7 shows the standardized path coefficients, the t statistics and the corresponding statistical significance, it is observed that the coefficients are significant.

**Table 7. Path coefficients (standardized regression coefficients)**

	Path coefficient (Standardized $\beta$ )	Statistics t	P Values
EconomPerf->CompetPerf	-0.867	10.144	0.000
EnvironmPerf->CompetPerf	-0.388	2.547	0.011
SocialPerf->CompetPerf	0.420	2.240	0.026

In Table 8, the values of  $f^2$  are presented, which measure the change in  $R^2$  when a certain exogenous construct is omitted from the model. As can be seen, the EconomPerf has a large effect with the CompetPerf; however, the EnvironmPerf has a medium effect with the CompetPerf, and the SocialPerf has a small effect with the CompetPerf.

**Table 8. F square**

Constructs	Compet Perf	Econom Perf	Environm Perf	Social Perf
CompetPerf				
EconomPerf	1.202			
EnvironmPerf	0.154	0.035		1.585
SocialPerf	0.131	0.438		

In Table 9, the Q squared values are presented. According to the criterion of Cohen (1988), affirming that the model has a high degree of predictive relevance with respect to the endogenous CompetPerf and SocialPerf factors. In the case of EconomPerf, it presents a medium degree of predictive relevance.

**Table 9. Q squared**

	SSO	SSE	$Q^2 (=1-SSE/SSO)$
CompetPerf	192.0	132.8	0.308
EconomPerf	144.0	110.6	0.232
SocialPerf	96.0	55.3	0.424

*E. Interpret results and draw conclusions.*

The results of the path coefficients in Table 7 indicate that EconomPerf → CompetPerf and EnvironmPerf → CompetPerf, even when they have statistical significance, have signs contrary to those postulated in hypotheses H1 and H2, which indicates that these hypotheses are rejected. On the other hand, the path SocialPerf → CompetPerf, corresponding to hypothesis H3, has statistical significance and a positive sign, therefore this hypothesis is accepted (see Table 10).

**Table 10. Summary of results**

<b>Hypotheses</b>	<b>Independent variable</b>	<b>Results</b>
<b>H1</b>	EconomPerf has a positive and significant effect on the CompetPerf	Rejected
<b>H2</b>	EnvironmPerf has a positive and significant effect on the CompetPerf.	Rejected
<b>H3</b>	SocialPerf has a positive and significant effect on the CompetPerf	Accepted

The results obtained reveal that the economic sustainability and environmental sustainability dimensions of iron mining have a negative influence on competitiveness. This result, far from being a conflict of interest between latent variables, represents the effect of the observable variables. In the case of the environmental dimension, the emission of CO<sub>2</sub>, the waste generated, the use of water and the use of energy results in a negative effect on competitiveness. In the case of the economic dimension, the operating costs involved in the acquisition of goods and services in the localities, generating indirect jobs for the mining activity, and the community investment, if it increases, benefits the community, but in turn affects the profitability by being expenditures.

It is then a reality, in which the mining companies must give responsible treatment to the socio-environmental systems affected by the operations, and where necessarily part of the income and benefits generated will necessarily have to be allocated to their remediation.

## **V.CONCLUSIONS**

The following conclusions emerge from the research carried out:

1. It has been shown that the PLS SEM is a technique that facilitates the development of research models from theoretical concepts and latent variables, with a limited number of observations.

2. With the present case of a multidisciplinary nature of mining, industrial, environmental engineering and statistical science, applied to the mining industry, where indicators prepared from objective data of the observed reality were used, we can affirm that the PLS SEM technique, constitutes an excellent support tool for research in the field of engineering sciences.

3. The ability to model the relationships between latent variables in a flexible way and not subject to rigorous parametric assumptions of the PLS SEM, allows us to forecast for this recent technique many applications in the field of engineering sciences.

4. Finally, as future research work, there are possible applications of the PLS SEM to the study of important aspects of the industry such as: productivity, efficiency, innovation, quality, corporate social responsibility, operation of industrial plants, organizational climate, ergonomics, industrial safety, among others.

## APPENDIX

## A1. Cross Factor Loads.

	<b>Compet Perf</b>	<b>Econom Perf</b>	<b>Environm Perf</b>	<b>Social Perf</b>
SH	0.887	-0.565	-0.389	-0.394
PTF	0.844	-0.803	-0.401	-0.407
OM	0.674	-0.410	-0.071	-0.044
GP	0.617	-0.387	-0.345	-0.465
OC	-0.798	0.945	0.399	0.596
RS	-0.597	0.878	0.483	0.631
CI	-0.368	0.546	-0.054	0.199
JouleEI	-0.401	0.562	0.890	0.821
WaterW	-0.294	0.260	0.854	0.559
GMW	-0.217	0.105	0.701	0.601
CO2	-0.358	0.106	0.579	0.202
Employees	-0.394	0.486	0.843	0.895
WIM	-0.347	0.623	0.475	0.832

## A2. VIF values of the structural model

<b>Constructs</b>	<b>CompetPerf</b>	<b>EconomPerf</b>	<b>EnvironmPerf</b>	<b>SocialPerf</b>
CompetPerf				
EconomPerf	1.720			
EnvironmPerf	2.677	2.585		1.000
SocialPerf	3.718	2.585		

## REFERENCES

- [1]J. Hair, G. Hult, C. Ringle and M. Sarstedt. A Primer on Partial Least Square Structural Equation Modeling (PLS-SEM). California: United States. Sage, 2017.
- [2]H. Wold. Model Construction and Evaluation when Theoretical Knowledge Is Scarce: An Example of the Use of Partial Least Squares. Genève. Faculté des Sciences Économiques et Sociales, Université de Genève. 1979.
- [3]J. Henseler, G. Hubona & P. Ray. "Using PLS path modeling new technology research: updated guidelines". *Industrial Management & Data Systems*, 116(1), 2-20. 2016.
- [4]G. Cepeda and Roldán J. "Aplicando en la Práctica la Técnica PLS en la Administración de Empresas". Congreso de la ACEDE, Murcia, España, 2004.
- [5]D. Garson. Partial Least Squares. Regresión and Structural Equation Models. USA. Statistical Associates Publishing: 2016.
- [6]D. Barclay, C. Higgins & R. Thompson. "The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration". *Technology Studies. Special Issue on Research Methodology*. (2:2), pp. 285-309. 1995
- [7]J. Medina, N. Pedraza & M. Guerrero. "Modelado de Ecuaciones Estructurales. Un Enfoque de Partial Least Square Aplicado en las Ciencias Sociales y Administrativas". XIV Congreso Internacional de la Academia de Ciencias Administrativas A.C. (ACACIA). EGADE – ITESM. Monterrey, México, 2010.
- [8]J. Medina & J. Chaparro. "The Impact of the Human Element in the Information Systems Quality for Decision Making and User Satisfaction". *Journal of Computer Information Systems*. (48:2), pp. 44-52. 2008.
- [9]D. Leidner, S. Carlsson, J. Elam & M. Corrales. "Mexican and Swedish Managers' Perceptions of the Impact

- of EIS on Organizational Intelligence, Decision Making, and Structure”. *Decision Science*. (30:3), pp. 633-658. 1999.
- [10]W. Chin. “The partial least squares approach for structural equation modeling”. Chapter Ten, pp. 295-336 in *Modern methods for business research*. Edited by Macoulides, G. A., New Jersey: Lawrence Erlbaum Associates, 1998.
- [11]M. Höck & C. Ringle M. “Strategic networks in the software industry: An empirical analysis of the value continuum”. IFSAM VIIIth World Congress, Berlin 2006.
- [12]J. Henseler, Ch. Ringle & M. Sarstedt. *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields*. Berlin: Springer, 2012.
- [13]S. Daskalakis & J. Mantas. “Evaluating the impact of a service-oriented framework for healthcare interoperability”. *Studies in Health Technology and Informatics*. pp. 285-290. 2008.
- [14]C. Fornell & D. Larcker: “Evaluating Structural Equation Models with Unobservable Variables and Measurement Error”, *Journal of Marketing Research*, vol. 18, pp. 39-50. February 1981.
- [15]C. Fornell. *A Second Generation of Multivariate Analysis: An Overview*. Vol. 1. New York, U.S.A. Praeger Publishers: 1982.
- [16]R. Falk and N. Miller. *A Primer for Soft Modeling*. Ohio: The University of Akron. 1992.
- [17]M. Martínez. Aplicación de la técnica PLS-SEM en la gestión del conocimiento: un enfoque técnico práctico. *Revista Iberoamericana para Investigación y el Desarrollo Educativo*. Vol. 8, Núm. 16. 2018.
- [18]S. Geisser. “A predictive approach to the random effects model”. *Biometrika*, Vol. 61(1), pp. 101-107. 1974.
- [19]J. Cohen. *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum, 1988.
- [20]GRI (2013). *G4 Sustainability Reporting Guidelines*. Global Reporting Initiative. Available: [www.globalreporting.org](http://www.globalreporting.org)

#### CURRICULUM SUMMARY



**Juan E. Villalva.** Dr. in Engineering Sciences, MSc in Electronic Engineering, Esp in Operations and Production, Esp in Automation and Electrical Engineer. Researcher, Teacher and field experience in the mining and metal processing industries.