Atmospheric Pollution Research 6 (2015) 389‑397

Atmospheric Pollution Research

www.atmospolres.com

Development of PLS–path model for understanding the role of precursors on ground level ozone concentration in Gulfport, Mississippi, USA

Amit Kumar Gorai [1], Francis Tuluri [2], Paul B. Tchounwou [3]

[1] *Department of Mining Engineering, National Institute of Technology, Rourkela, Odisha ‑769008, India*
[2] *Department of Technology, Jackson State University, Jackson, MS 39217, USA*
[3] *Department of Biology, Jackson State University, Jackson, MS 39217, USA*

ABSTRACT

Ground‑level ozone (GLO) is produced by a complex chain of atmospheric chemical reactions that depend on precursor emissions from natural and anthropogenic sources. GLO concentration in a particular location is also governed by local weather and climatic factors. In this work an attempt was made to explore a Partial Least Squares Path Modeling (PLS–PM) approach to quantify the inter–relationship between local conditions (weather parameters and primary air pollution) and GLO concentrations. PLS path modeling algorithm was introduced and applied to GLO concentration analyses at Gulfport, Mississippi, USA. In the present analysis, three latent variables were selected: PRC (photochemical reaction catalyst), MP (meteorological factor), and OPP (other primary air pollutants). The three latent variables included 14 indicators for the analysis; with PRC having two (extraterrestrial radiation on horizontal surface, and extraterrestrial radiation normal to the sun), MP having nine (temperature, dew point, relative humidity, pressure, visibility, maximum wind speed, average wind speed, precipitation, and wind direction) and OPP having three (NO_2, $PM_{2.5}$, and SO_2) parameters. The resulting model revealed that PRC had significant direct impact on GLO concentration but very small overall effect. This is because PRC had significant indirect negative impact on GLO via MP. Thus, when both direct and indirect effects were taken into account, PRC emerged as having the weakest effect on GLO. The third variable (OPP) also had a positive impact on GLO concentration.

Keywords: *Ground level ozone, precursor sources, Partial Least Square (PLS), Path Model*

Corresponding Author:
Amit Kumar Gorai
☎ : *+91‑0651‑2276587*
🖷 : *+91‑0651‑2275401*
✉ : *amit_gorai@yahoo.co.uk*

Article History:
Received: 08 July 2014
Revised: 07 November 2014
Accepted: 08 November 2014

doi: 10.5094/APR.2015.043

1. Introduction

Ground–level or tropospheric ozone (O_3) is listed as one of the criteria pollutants by many countries and organizations like World Health Organization (WHO) and United States Environmental Protection Agency (U.S. EPA). It is a major constituent of photochemical smog, an air pollution event that often occurs in mega cities. O_3 has been proved to have adverse effects on human health, especially to the respiratory system (Lippmann, 1993). It can also adversely affect crops and forest ecosystems (Bascom et al., 1996, Lippmann, 2009). The adverse effects of O_3 can be minimized by proper identification and reduction of its precursors.

Numerous studies on the evolution of tropospheric O_3 changes and the associated radiation forcing have been carried out using various chemical transport or climate models (Hauglustaine et al., 1994; Roelofs et al., 1997; Brasseur et al., 1998; Stevenson et al., 1998; Forster, 1999; Mickley et al., 1999; Berntsen et al., 2000; Grenfell et al., 2001; Hauglustaine and Brasseur, 2001). However, there is often a significant difference between the models in their predictions of ozone change (Houghton et al., 2001).

Variations in ozone concentration are controlled by a number of processes including photochemistry, physical and chemical process removal. Ozone is produced in the troposphere by photochemical oxidation of CO, methane and non–methane volatile organic compounds (NMVOCs) by the hydroxyl radical (OH) in the presence of reactive nitrogen oxides (NO_X=NO+NO_2). NMVOCs, CO and NO_X have large combustion sources.

Changes in climatic conditions also affect ozone concentration by perturbing ventilation rates (wind speed, mixing depth, convection and frontal passages), precipitation scavenging, dry deposition, chemical production and loss rates, natural emissions, and background concentrations. Temperature can serve as a proxy for other meteorological conditions, such as stagnation events, conducive for formation of elevated levels of ozone. Many model–based studies have revealed that temperature is the most important meteorological variable affecting ozone concentrations in polluted regions (Morris et al., 1989; Aw and Kleeman, 2003; Sanchez–Ccoyllo et al., 2006; Steiner et al., 2006; Dawson et al., 2007).

Atmospheric humidity, which is projected to increase overall in a warmer world, leads to increased ozone in high NO_X areas and decreased ozone in low NO_X areas. The correlation of ozone with relative humidity was studied by Camalier et al. (2007) who observed a good correlation. The observed correlation of ozone with solar radiation seen in some studies could reflect in part the association of clear sky with high temperatures (Ordonez et al., 2005). Wind can also be important in controlling ozone levels, as precursor species are dispersed and diluted, typically reducing the ozone‑forming reactions. Wind can also disperse ground‑level ozone that has already formed, reducing the amount of, and exposure time to, elevated ozone levels in local generation areas; it can also add to another downwind region that may or may not already have local ozone levels approaching critical health risk levels. Weaker wind speeds in polluted regions cause ozone to increase, as would be expected simply from a longer reaction time and increased aerodynamic resistance to dry deposition (Baertsch–Ritter et al., 2004; Sanchez–Ccoyllo et al., 2006; Dawson et al., 2007). Precipitation is generally not important for removing ground-level ozone as it is not water soluble.

It is imperative to accurately find the sources contributing to O_3 concentration in order to take corrective policy measures to reduce the ground level O_3 concentration. Since it is a secondary pollutant and is not directly emitted from any source, it is not possible to make source apportionment to develop cleaner technologies. It is formed as a result of complex photochemical reactions in the troposphere. Although VOCs and NO_X have been confirmed as the key precursors of O_3, the development of an effective strategy for reducing O_3 pollution in megacities is still problematic, due to the non–linear dependency of O_3 concentration on different factors. So it is of great importance to evaluate how different precursors/factors contribute to the high ozone concentration. Hence, it is of utmost importance to keep an account and control on the various precursors/factors leading to the formation of the ground level ozone in the city so that suitable measures can be taken to prevent the involvement and adverse effects of the different variables. In this work a path analysis was developed to describe the contributions of different variables to the formation of ground level ozone. Path analysis, which is a causal modeling approach for exploring the correlations within a defined network, was performed to compare and describe the cause–effect relationships between ground level ozone, endogenous variables and exogenous variables. The model development and performance test was carried out using "plspm" package of *r* software (http://cran.r–project.org/web/packages/plspm/index.html).

2. Structural Equation Model (SEM)

SEM model facilitates the estimation of causal relationships, defined according to a theoretical model, linking two or more latent complex concepts (i.e. the composite indicators), each measured through a number of observable indicators. The basic idea is that complexity inside a system can be studied taking into account a whole of causal relationships among latent concepts, called Latent Variables (LV), each measured by several observed indicators usually defined as Manifest Variables (MV). Moreover, path models are a logical extension of regression models as they involve the analysis of simultaneous multiple regression equations. More specifically, a path model is a relational model with direct and indirect effects among observed variables.

SEM has the ability to assess latent variables at the observation level (outer or measurement model) and to test relationships between latent variables on the theoretical level (inner or structural model) (Bollen, 1989). Most of the researchers generally analyzed two types of SEM methods: covariance–based techniques (CB–SEM) (Joreskog, 1978; Joreskog, 1993) and variance based partial least squares (PLS–PM) (Wold, 1982; Wold, 1985; Lohmoller, 1989). Although both methods share the same roots (Joreskog and Wold, 1982), the present research has focused primarily on a so–called component–based estimation method PLS–PM, because of the key role that is played by the estimation of the LVs in the model. In fact, the main aim of component–based methods is to provide an estimate of each LV in such a way that they are the most correlated with one another and the most representative of each corresponding block of manifest variables. This is of main importance in building system of composite indicators. As a matter of fact, according to PLS–PM approach, each composite indicator is obtained in order to be the most representative of each corresponding indicator and the most correlated with the others linked composite indicators.

PLS–PM maximizes the explained variance of the endogenous latent variables by estimating partial model relationships in an iterative sequence of ordinary least squares (OLS) regressions. An important characteristic of PLS–PM is that it estimates latent variable scores as exact linear combinations of their associated manifest variables (Fornell and Bookstein, 1982) and treats them as perfect substitutes for the manifest variables. The scores thus capture the variance that is useful for explaining the endogenous latent variable(s). Estimating models via a series of OLS regressions implies that PLS–PM relaxes the assumption of multivariate normality needed for maximum likelihood–based SEM estimations (Fornell and Bookstein, 1982; Wold, 1982; Lohmoller, 1989; Hwang et al., 2010). Furthermore, since PLS–PM is based on a series of OLS regressions, it has minimum demands regarding sample size and generally achieves high levels of statistical power (Reinartz et al., 2009). Furthermore, PLS–PM is not constrained by identification concerns, even if models become complex, a situation that typically restricts CB–SEM usage (Hair et al., 2011).

3. PLS Path Model

PLS path models are formally defined by two sets of linear equations: the inner model and the outer model. The inner model specifies the relationships between unobserved or latent variables, whereas the outer model specifies the relationships between a latent variable and its observed or manifest variables. A PLS path model is described by two models: (1) a measurement model relating the MVs to their own LV and (2) a structural model relating some endogenous LVs to other exogenous or endogenous LVs. The measurement model is also called the outer model and the structural model the inner model. Thus, the endogenous LVs can be seen not only as composite indicators, due to their relations with the corresponding indicators, but also as complex indicators, due to their causal relations with other composite indicators.

3.1. The measurement model

This represents the relationships between a latent variable and its block of manifest variables. There are two main measurement options for the outer model: reflective blocks and formative blocks. In reflective mode, the latent variables are considered as the cause of the manifest variables whereas in the formative mode, the manifest variables are considered to be the cause of the latent variables. In the present work, the model is constructed in formative mode. The outer model/measurement model relationship is also considered to be linear. In mathematical notation, the relationship can be represented as given in Equation (1):

$$LV_j = \lambda_{0j} + \lambda_{jk} X_{jk} + e_j \text{ (Formative mode)} \quad (1)$$

The coefficients λ_{jk} are called loadings, λ_{0j} is just the intercept term, and the e_j terms account for the residuals.

3.2. The structural model

The structural model represents to linear equations relating the LVs between them (the structural or inner model). Mathematically this can be represented as given in Equation (2):

$$LV_j = \beta_{0j} + \sum_{i \to j} \beta_{ji} LV_i + e_j \quad (2)$$

The subscript *i* of LV_i refer to all the latent variables that are supposed to predict LV_j. The coefficients β_{ji} are the path coefficients and they represent the "strength and direction" of the relations between the response LV_j and the predictors LV_i. β_{0j} is just the intercept term, and the e_j term accounts for the residuals errors.

4. Materials and Method

4.1. Study area

In the present work, the study area is Gulfport, Mississippi, USA. Gulfport is the second largest city in Mississippi after the state capital Jackson. It is the largest of the two principal cities of the Gulfport–Biloxi, Mississippi Metropolitan Statistical Area, which is included in the Gulfport–Biloxi–Pascagoula, Mississippi Combined Statistical Area. Gulfport is located at 30°24'6" N, 89°4'34" W. According to the 2010 census, the city of Gulfport had a total population of 69 220. The population density was 1 191.4 people per square mile (459.9 per square km). Gulfport has a humid subtropical climate, which is strongly moderated by the Gulf of Mexico. Winters are short and generally warm, cold spells do occur, but seldom last long. Summers are generally long, hot and humid, though the city's proximity to the Gulf prevents extreme summer highs, as seen farther inland. Gulfport is subject to extreme weather, most notably tropical storm activity through the Gulf of Mexico. The monitoring station is shown in Figure 1.

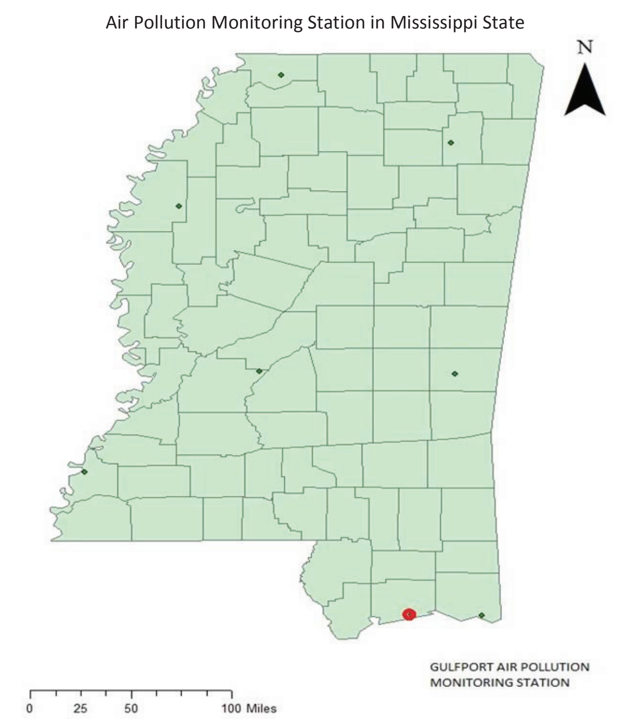

Figure 1. *Air pollution monitoring stations.*

4.2. Data collection and analysis

Table 1 provides a basic description of the model's variables, showing the latent exogenous variables (LEXV): PRC, MP, and OPP and latent endogenous variables (LENV): GLO from the PLS path model, and illustrates the potential determinants (MVs) used in the analysis. Factors that potentially influenced the ground level ozone concentration were selected from previous literature. Ozone is a secondary pollutant and there are many precursors like NO_X, VOCs etc. for the formation of this gaseous pollutant. Two indicators are used to measure the photochemical reaction catalyst (PRC). Indicator SR1 represents extraterrestrial radiation on horizontal surface in W/m^2, and SR2 represents extraterrestrial radiation normal to the sun in W/m^2. Nine indicators are used to measure meteorological factors (MP). These are TEMP (temperature in °F), DP (dew point temperature in °F), HUM (relative humidity in percentage), PRES (pressure in inch), VIS (visibility in distance mile), MWS (maximum wind speed in miles/hr), AWS (average wind speed in miles/hr), PREC (precipitation in inch), and WD (wind direction in degrees). Three indicators are used to measure other primary pollutant parameters (OPP). These are NO_2 (daily maximum 1–hour NO_2 concentration in ppb), $PM_{2.5}$ (24 hour average particulate matter less than and equal to 2.5 µm size in µg/m^3), and SO_2 (daily maximum 1–hour SO_2 concentration in ppb). There is only one indicator (GLOC: ground level ozone concentration in ppb) that was used to measure the GLO (ground level ozone) for understanding the degree of influence of various direct and indirect causal factors for the formation of ground level ozone. Though, VOC is considered as one of the major precursor for ozone formation, it is not considered in the model due to unavailability of data. Daily air quality data were collected from the website of U.S. EPA whereas the meteorological data were collected from website http://www.wunderground.com. In the present PLS–PM model analysis, an overall 672 day's data were compiled for the duration of 2008 to 2010.

4.3. Priori PLS–Path model formulation

The conceptual model behind the relations among latent and manifest variables is drawn as a path diagram (Figure 2) in which ellipses represent the manifest variables and rectangles or squares refer to the latent variables. Arrows show causations among the variables (either latent or manifest), and the direction of the array defines the direction of the relation, i.e. variables receiving the array are to be considered as endogenous variables in the specific relationship.

Table 1. *Description of the variables used in the model*

Latent Variables (LV)	Notation Used in the Equation	Manifest Variables (MV)	Notation Used in the Equation
Latent Exogenous Variables (LEXV)			
Photochemical Reaction Catalyst (PRC)	LV1	Extraterrestrial radiation on horizontal surface in W/m^2 (SR1)	X11
		Extraterrestrial radiation normal to the sun in W/m^2 (SR2)	X12
Meteorological Factors (MF)	LV2	Temperature in °F (TEMP)	X21
		Dew point temperature in °F (DP)	X22
		Relative humidity in percentage (HUM)	X23
		Pressure in inch (PRES)	X24
		Visibility in distance mile (VIS)	X25
		Maximum wind speed in miles/hr (MWS)	X26
		Average wind speed in miles/hr (AWS)	X27
		Precipitation in inch (PREC)	X28
		Wind direction in degrees (WD)	X29
Other Primary Pollutants (OPP)	LV3	Daily Maximum 1–hour NO_2 concentration in ppb (NO_X)	X31
		24 hour average particulate matter less than 2.5 µm size in µg/m^3 ($PM_{2.5}$)	X32
		Daily Maximum 1–hour SO_2 concentration in ppb (SO_2)	X33
Latent Endogenous Variables (LENV)			
Ground Level Ozone (GLO)	LV4	Ground Level Ozone Concentration in ppb (GLOC)	Y11

The measurement model indicates the relationship between the MVs as listed in Table 1 and the three LEXVs. The structural model reflects the relationship between the three LEXVs and one LENV as listed in Table 1. The main purpose of the PLS path model is to evaluate both the measurement and structural models. In particular, we focus on assessing the path coefficient (β) value in the structural model, which reflects the impact of the LEXVs on the LENV.

Measurement and structural model equations were formulated corresponding with the priori path model represented in Figure 2. The equations are the same as the multiple linear regression equations, and they are solved by the least square technique for determination of path coefficients. This facilitates direct comparison of values to reflect the relative importance of manifest variables in order to explain variation in the latent variables (Seker and Serin, 2004). PLS–PM's main objective is to get estimates of both the latent variables and the parameters (coefficients and loadings).

The formulations of the measurement model in the present case (formative mode) are represented in Equations (3) to (6) as follows:

$$LV_1 = \lambda_{01} + \lambda_{11}X_{11} + \lambda_{12}X_{12} + e_1 \tag{3}$$

$$\begin{aligned} LV_2 = \lambda_{02} + \lambda_{21}X_{21} + \lambda_{22}X_{22} + \lambda_{23}X_{23} + \lambda_{24}X_{24} \\ +\lambda_{25}X_{25} + \lambda_{26}X_{26} + \lambda_{27}X_{27} + \lambda_{28}X_{28} + e_2 \end{aligned} \tag{4}$$

$$LV_3 = \lambda_{03} + \lambda_{31}X_{31} + \lambda_{32}X_{32} + \lambda_{33}X_{33} + e_3 \tag{5}$$

$$LV_4 = \lambda_{04} + \lambda_{41}Y_{11} + e_4 \tag{6}$$

where, λ_{jk} is a coefficient linking each manifest variable to the corresponding latent variable and the error term e_j represents the fraction of the corresponding latent variable not accounted for by the block of manifest variables.

For the structural relationships we constructed four equations for four latent variables. The first equation for LV1 which is not influenced by any other latent variables is given by Equation (7) as follows:

$$LV_1 = \beta_{01} + E_1 \tag{7}$$

Similarly the second the relationship in which LV2 depends on LV1 the structural model is represented in Equation (8) as follows:

$$LV_2 = \beta_{02} + \beta_{21}LV_1 + E_2 \tag{8}$$

The third inner or structural relationship in which LV_3 depends on LV_1 and LV_2 is represented in Equation (9) as follows:

$$LV_3 = \beta_{03} + \beta_{31}LV_1 + \beta_{32}\,LV_2 + E_3 \tag{9}$$

The fourth inner relationship in which LV_4 depends on LV_1, LV_2 and LV_3 is represented in Equation (10) as follows:

$$LV_4 = \beta_{04} + \beta_{41}LV_1 + \beta_{42}\,LV_2 + \beta_{43}\,LV_3 + E_4 \tag{10}$$

4.4. Path analysis algorithm

The algorithm for the path analysis is represented in Figure 3. PLS Path modeling follows a sequential procedure that can be divided in three major stages. In first stage, the main objective is to determine the scores of the latent variables in the model that serve as a numerical representation of the latent construct. This is an iterative process between inner and outer method to determine the weights until convergence. The purpose of determining the weights is that these values further used to get the scores of the latent variables. This process begins with an outside approximation assuming initial weight of 1 so that each LV factor score is initially a simple sum of its item scores. From the starting point, the algorithm iterates between the inside and outside approximation methods in the calculation of LV scores. The outside method provides an estimate of the LV score via an aggregation of its indicators, whereas the inside method yields an estimate based on the adjacent (neighboring) LV scores.

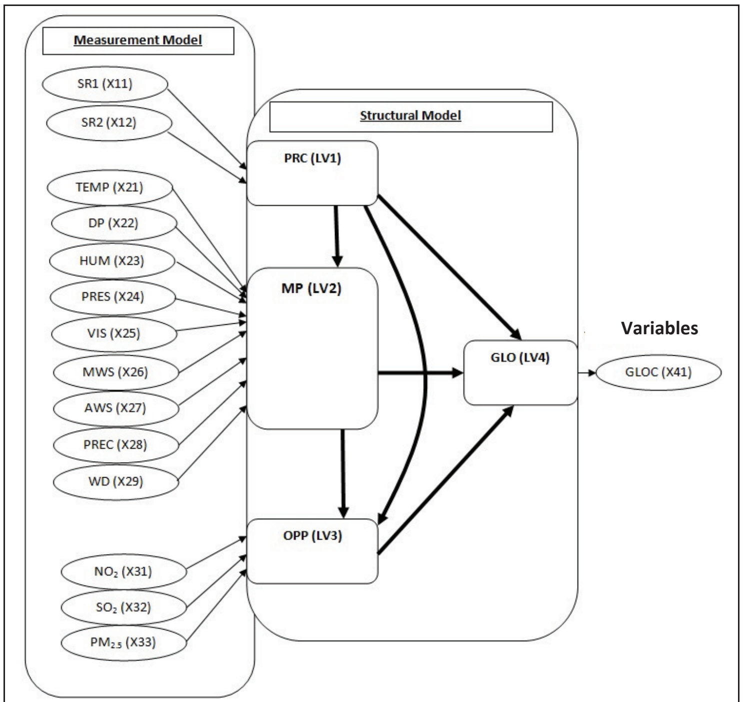

Figure 2. *A proposed path model showing the causal relations among the variables.*

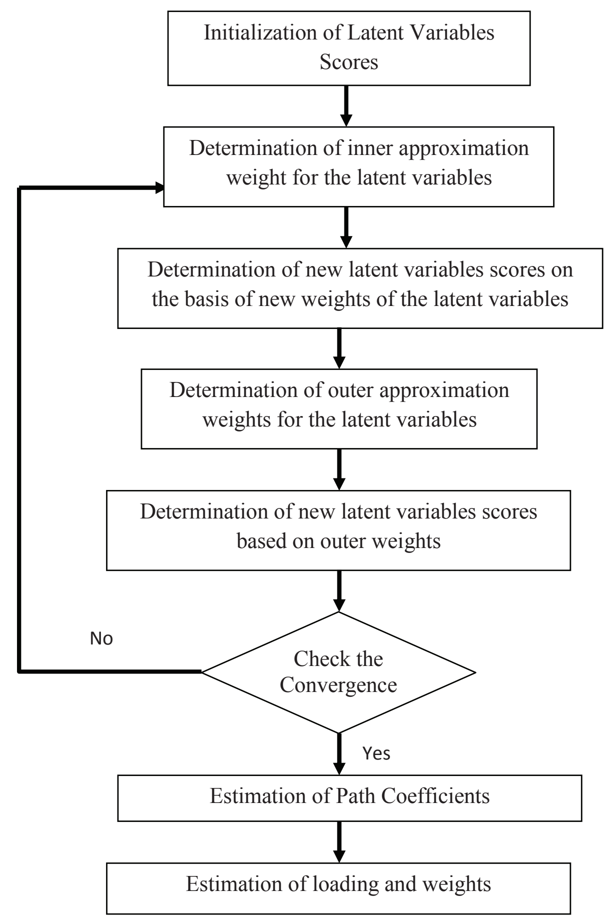

Figure 3. *Flow diagram of PLS–Path analysis.*

Once the iterative process completes, the second stage has to do with estimating the path coefficients of the inner model using the LV scores obtained in the previous step. The final estimates of the loadings or weights (measurement model) are also determined using ordinary least squares (OLS) regression. And the third stage involves the computation of the loadings or weights (outer model/ measurement model). The loadings are determined by computing simple correlations. The model was run for the analysis using plspm package of R-Software.

5. Results and Analysis of Model

The idea is to calculate estimates of latent variables as linear combinations of their associated indicators using a special linear combination. We look for a linear combination in such a way that the obtained latent variables take into account the relationships of the structural and the measurement models in order to maximize the explained variance of the dependent variables (both latent and observed variables). The model results are analyzed in two stages. In first stage outer model or measurement model was analyzed. Once the validation of the outer model was done, the inner model was tested for determining the path coefficients and factor loadings.

5.1. Analysis of measurement model

Two cases were critically examined for deciding on whether a particular indicator should enter into the index or not. An indicator can be irrelevant for the construction of the formative index because it either does not have a significant impact on the formative index, or because it exhibits high multicollinearity, which could mean that the indicator's information is redundant. In order to check for the first case, the significance of the estimated indicator weights were determined by means of bootstrapping (Chin, 1998; Davison et al., 2003; Tenenhaus and Vinzi, 2005). The bootstrap results indicated that the weights of all the indicators are significant at 5% level. Thus on the basis of this test no manifest variables were eliminated from the priori path model.

In order to assess the degree of multicollinearity among the formative indicators, correlation coefficients (for two variables) and variance inflation factor (VIF) (for more than two variables in a particular block) values should be computed (Cassel et al., 2000; Diamantopoulos and Winklhofer, 2001; Grewal et al., 2004). A rule of thumb states that VIFs greater than 10 reveal a critical level of multicollinearity. However, any VIF substantially greater than 3.3 (Diamantopoulos and Siguaw, 2006; Petter et al., 2007) indicates multicollinearity and should be remediated. The correlation coefficient between two variables greater than 0.8 indicates the problem of collinearity (Kennedy, 2003).

In the present model, there are three blocks of manifest variables. In first block, there are only two variables or indicators. The second and third block is represented by nine and three indicators respectively. The correlation between the variables or indicators in first block (SR1, SR2) was found to be 0.993 which exceeds the recommended value of 0.8 indicating the problem of multicollinearity. Thus, the variable SR2 removed as remediation from the model during subsequent path analysis. The estimated magnitudes of multicollinearity for (Block 2: TEMP, DP, HUM, PRES, VIS, MWS, AWS, PREC, WD and Block 3: NO_2, $PM_{2.5}$, SO_2) with more than 2 indicators were tested using variation inflation factor (VIF). The estimated magnitudes of multicollinearity for block 2 (MP) manifest variables indicated high value of VIF with dew point temperature (DP). After removing the variables (DP), the VIF values found to be within 2 for all the combinations and this indicates that there is no further problem of multicollinearity in second block (MP). The VIF values for third block (OPP) were found to be within 2 and this clearly indicated no problem of multicollinearity. After remediating action, the subsequent results indicated no further problem of multicollinearity and the measurement model is valid.

The standardized factor loading reflects the explanatory power of MVs to their corresponding LVs. These are calculated on the basis of MVs datasets (excluding SR2 and DP) as represented in Figure 2. The loadings of each manifest variable with its associated latent variable and its cross loading on other latent variables are shown in Table 2. The manifest variable, SR1 (X11) is positively correlated to the latent variable, PRC. The LV, MP is positively affected by TEMP (X21), HUM (X23), VIS (X25), MWS (X26), AWS (X27), PREC (X28) and negatively affected by the PRES (X24). All the MVs [NO_2 (X31), SO_2 (X32), and $PM_{2.5}$ (X33)] in block 3 have positive impact on the LV (OPP). The results shown in Table 2 also clearly indicate that TEMP (X21) has maximum impact (0.792) on LV MP whereas VIS (X25) has least impact. In block 3, NO_2 has the maximum impact (0.997) on the LV OPP. The manifest variables $PM_{2.5}$ (X32) and SO_2 (X33) have very less impact in comparison to the variable NO_2 (X31). Generally, a MV's loading on its associated latent variable is greater than its cross–loading on other latent variables in the model. The loading and cross–loading results represented in Table 2 clearly indicates that the loading on its associated latent variables are greater than its cross–loadings on other latent variables except VIS (X25) in block 2 (MP) and SO_2 (X32) in block 3. This may indicate that these variables are more associated with the latent variable GLO.

Table 2. *Latent variables loadings and cross loadings*

	PRC	MP	OPP	GLO
		PRC		
X11	1	0.434	–0.316	0.056
		MP		
X21	0.552	0.792	–0.362	–0.157
X23	0.093	0.439	–0.147	–0.354
X24	–0.223	–0.396	0.325	–0.013
X25	–0.003	0.012	–0.049	0.029
X26	–0.039	0.267	–0.197	–0.203
X27	–0.130	0.318	–0.292	–0.267
X28	0.023	0.255	–0.095	–0.226
		OPP		
X31	–0.319	–0.512	0.997	0.299
X32	0.082	–0.024	0.139	0.215
X33	0.023	–0.097	0.211	0.166
		GLO		
Y11	0.056	–0.403	0.304	1

5.2. Analysis of structural model

Reliable and valid outer model estimations permit an evaluation of the inner path model estimates. The essential criterion for this assessment is the coefficient of determination (R^2) of endogenous latent variables. Falk and Miller (1992) recommended the R^2 for variable's variance explained by the independent variables. They also recommended the R^2 for endogenous variables be greater and equal to 0.10. An R^2 greater and equal to 0.10 ensures that the variance explained by the endogenous variables has practical, as well as statistical significance. The observed R^2 value for three endogenous latent variables, MP, OPP and GLO were found to be 0.189, 0.275, and 0.25 respectively. Thus, the observed value is sufficiently higher than the recommended value (0.10). The path coefficients in the model decomposed into direct and indirect effects, corresponding, to direct and indirect paths represented in the arrows in the model. This is based on the rule that in a linear system, the total causal effect of LV_i on LV_j is the sum of the values of all the paths from i to j. The PLS PM results for path coefficients and its statistical significance are represented in Table 3. In order to determine the confidence intervals of the path coefficients and statistical inference, bootstrap method was used (Tenenhaus and Vinzi, 2005). The path coefficient results for each indicated paths along with the confidence interval are represented in Table 3. The bootstrap analysis was carried out using 200 sample data sets. The path coefficients results represented in Table 3 clearly indicate that the original path coefficients values are closely matches with the path coefficients values obtained from bootstrap results. This indicates the path drawn in the model having consistent relationship. Also, the confidence intervals shown in Table 3 clearly indicate that the path coefficients values fall in this range. Thus, all the paths defined in the diagram are significant. From Table 3, it can inferred that the LEXV, MP has the highest negative impact (–0.448) on the LENV, GLO. According to the structural model results, the ground level ozone concentration will decrease with the meteorological factors. Similarly, the photochemical reaction catalyst (PRC) and other primary pollutant (OPP) also have significant positive impact on the ground level ozone concentration.

Another important evaluation relates to the indirect effects of the LEXVs on other LEXVs or LENVs. This relationships evaluates the effect of predecessor of a certain endogenous latent variable involves in mediating (Helm et al., 2009) or moderating (Henseler and Fassott, 2009). The indirect effects can be calculated from the above results using the Equation (11) as:

$$\text{Indirect Effect} = \text{Total effects} - \text{Direct effects} \tag{11}$$

Mathematically, the indirect effects can be calculated as the multiplication of path coefficients of indirect paths. Considering “GLO” as the dependent in the model above, and considering “MP” as the independent, the indirect effects were calculated by multiplying the path coefficients for each path from MP to GLO. Thus the indirect effects of the MP to GLO is calculated as:

$$\beta_{MP\rightarrow GLO} = \beta_{MP\rightarrow OPP} \times \beta_{OPP\rightarrow GLO} = -0.464 \times 0.170 = -0.0791$$

Similarly, indirect effects of PRC to GLO is:

$$\beta_{PRC\rightarrow GLO} = \beta_{PRC\rightarrow MP} \times \beta_{MP\rightarrow GLO} + \beta_{PRC\rightarrow MP} \times \beta_{MP\rightarrow OPP} \times \beta_{OPP\rightarrow GLO} = -0.2486$$

Table 3. *Path coefficients results*

Paths	Path Coefficients (using 672 data sets)	Path Coefficients (using 200 data sets in bootstrap analysis)	Standard Error	95 LCI	95 UCI
		Direct Effects			
$\beta_{PRC\rightarrow MP}$	0.434	0.434	0.036	0.364	0.506
$\beta_{PRC\rightarrow OPP}$	–0.115	–0.115	0.037	–0.192	–0.046
$\beta_{PRC\rightarrow GLO}$	0.305	0.308	0.037	0.237	0.383
$\beta_{MP\rightarrow OPP}$	–0.464	–0.467	0.038	–0.542	–0.399
$\beta_{MP\rightarrow GLO}$	–0.448	–0.454	0.032	–0.527	–0.397
$\beta_{OPP\rightarrow GLO}$	0.170	0.171	0.036	0.105	0.241
		Total Effects			
$\beta_{PRC\rightarrow MP}$	0.434	0.434	0.036	0.364	0.506
$\beta_{PRC\rightarrow OPP}$	–0.316	–0.318	0.042	–0.392	–0.222
$\beta_{PRC\rightarrow GLO}$	0.056	0.056	0.042	–0.025	0.132
$\beta_{MP\rightarrow OPP}$	–0.464	–0.467	0.038	–0.542	–0.399
$\beta_{MP\rightarrow GLO}$	–0.527	–0.534	0.027	–0.588	–0.483
$\beta_{OPP\rightarrow GLO}$	0.170	0.171	0.036	0.105	0.241

The sum of the direct and the indirect effects gives total effect of each variable on GLO. The summary of direct effects, indirect effects and total effects of various paths of the structural model is represented in Table 4. The same is also graphically represented in Figure 4. The characteristics of GLO responses obtained by path analysis in the Gulfport area showed that GLO concentration is most strongly related to the meteorological factors (MP) in both the ways (direct effect and indirect effect) but in negative way. That is, meteorological factors (MP) and other primary pollutant parameter (OPP) have reduced the GLO concentrations. The model results also indicate that PRC has significant direct impact on ground level ozone concentration but very small overall effects. This is because PRC has significant indirect negative impact on GLO via MP. Thus, the combined effect of direct and indirect compensates each other leads to weakest effect of PRC on GLO concentration. Thus, when both direct and indirect effects are taken into account, PRC emerges as having the weakest effect on GLO. The third block, OPP also has a positive impact on GLO concentration. The structural model is demonstrated by the path coefficients ($\beta_{i \to j}$) and the loading of the MVs to corresponding LVs represented in Figure 5.

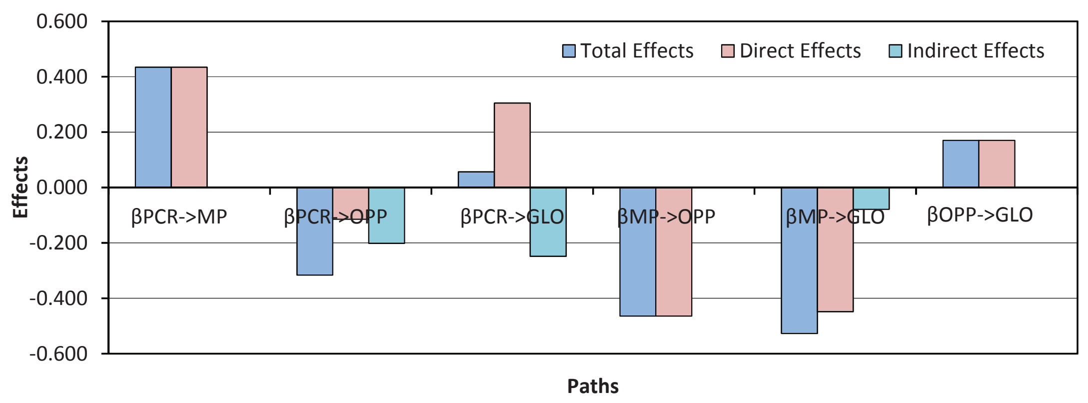

Figure 4. *Direct, indirect and total effects of LEXVs on LEXVs/LEDVs.*

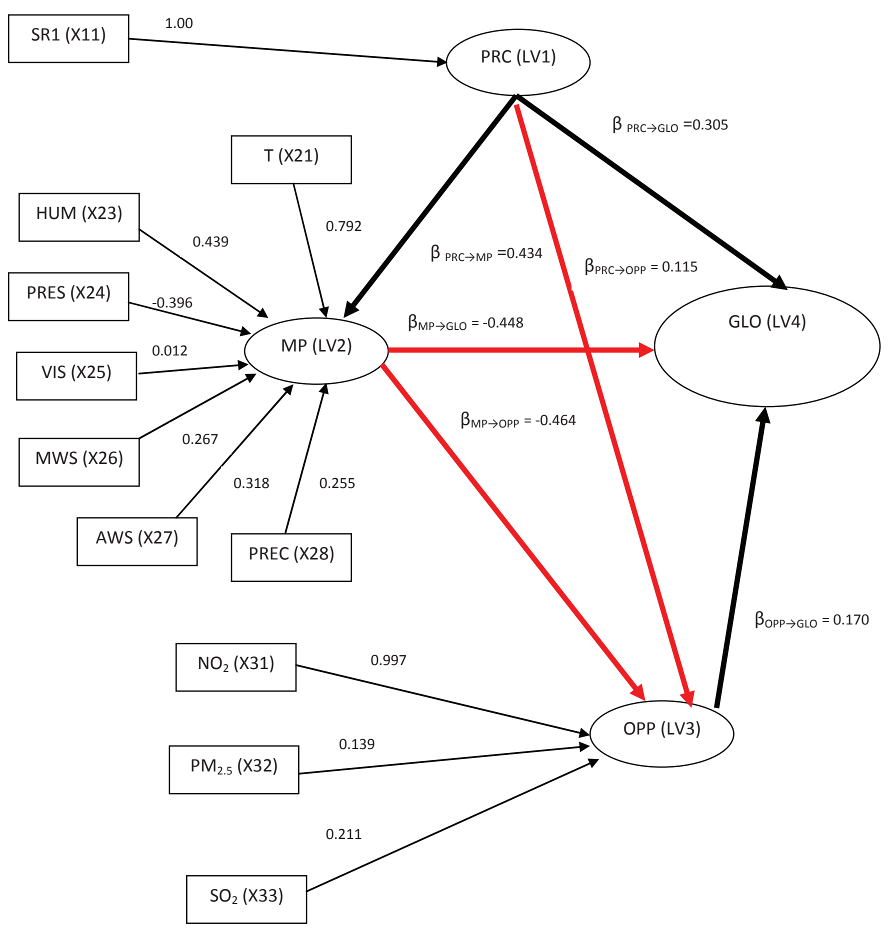

Figure 5. *Estimated loadings and path coefficients.*

Table 4. *Direct effects, indirect effects and total effects*

Paths	Direct Effects	Total Effects	Indirect Effects=Total Effects–Direct Effects
$\beta_{PRC \rightarrow MP}$	0.434	0.434	0.000
$\beta_{PRC \rightarrow OPP}$	–0.115	–0.316	–0.202
$\beta_{PRC \rightarrow GLO}$	0.305	0.056	–0.249
$\beta_{MP \rightarrow OPP}$	–0.464	–0.464	0.000
$\beta_{MP \rightarrow GLO}$	–0.448	–0.527	–0.079
$\beta_{OPP \rightarrow GLO}$	0.170	0.170	0.000

6. Conclusions

The results of this exploratory study suggest that the PLS path modeling approach for constructing ground level ozone concentration index is promising. This model not only permits construction of the index, but also provides insights into how different parameters differentially affect ground level ozone concentration in a particular geographical area. The results provide useful information for controlling the concentration of ground level ozone concentration. The PLS path modeling approach can be extended by incorporating an expanded set of variables related to the different dimensions. However, the results of this study illustrate the potential advantages of these approaches generally to better understanding conceptualizations and measures of ground level ozone concentration, which will ultimately aid for its management and impact reduction. The present work is limited by data availability for all the variables involved in ground level ozone formation and dispersion.

Acknowledgments

This research work was carried out in part of the corresponding author's Raman Postdoctoral Fellowship awarded by UGC, New Delhi, India. Authors are also thankful to U.S. EPA for making air pollution data available on the website for public use. The support of NIH/NIMHD Grant No. G12MD007581 through the RCMI–Center for Environmental Health at Jackson State University is also acknowledged.

References

Aw, J., Kleeman, M.J., 2003. Evaluating the first–order effect of intraannual temperature variability on urban air pollution. *Journal of Geophysical Research–Atmospheres* 108, art. no. 4365.

Baertsch–Ritter, N., Keller, J., Dommen, J., Prevot, A.S.H., 2004. Effects of various meteorological conditions and spatial emission resolutions on the ozone concentration and ROG/NO_X limitation in the Milan area (I). *Atmospheric Chemistry and Physics* 4, 423–438.

Bascom, R., Bromberg, P.A., Costa, D.A., Devlin, R., Dockery, D.W., Frampton, M.W., Lambert, W., Samet, J.M., Speizer, F.E., Utell, M., 1996. Health effects of outdoor air pollution. *American Journal of Respiratory and Critical Care Medicine* 153, 3–50.

Berntsen, T.K., Myhre, G., Stordal, F., Isaksen, I.S.A., 2000. Time evolution of tropospheric ozone and its radiative forcing. *Journal of Geophysical Research–Atmospheres* 105, 8915–8930.

Bollen, K.A., 1989. *Structural Equations with Latent Variables*, John Wiley & Sons, New York.

Brasseur, G.P., Kiehl, J.T., Muller, J.F., Schneider, T., Granier, C., Tie, X.X., Hauglustaine, D., 1998. Past and future changes in global tropospheric ozone: Impact on radiative forcing. *Geophysical Research Letters* 25, 3807–3810.

Camalier, L., Cox, W., Dolwick, P., 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. *Atmospheric Environment* 41, 7127–7137.

Cassel, C.M., Hackl, P., Westlund, A.H., 2000. On measurement of intangible assets: A study of robustness of partial least squares. *Total Quality Management* 11, S897–S907.

Chin, W.W., 1998. The partial least squares approach for structural equation modeling, in *Modern Methods for Business Research*, edited by Marcoulides, G.A., Lawrence Erlbaum Associates, London, pp. 295–236.

Davison, A.C., Hinkley, D.V., Young, G.A., 2003. Recent developments in bootstrap methodology. *Statistical Science* 18, 141–157.

Dawson, J.P., Adams, P.J., Pandis, S.N., 2007. Sensitivity of ozone to summertime climate in the Eastern USA: A modeling case study. *Atmospheric Environment* 41, 1494–1511.

Diamantopoulos, A., Siguaw, J.A., 2006. Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management* 17, 263–282.

Diamantopoulos, A., Winklhofer, H.M., 2001. Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research* 38, 269–277.

Falk, R.F., Miller, N.B., 1992. A Primer for Soft Modeling, University of Akron Press, Akron, Ohio, pp. 80.

Fornell, C., Bookstein, F.L., 1982. Two structural equation models: LISREL and PLS applied to consumer exit–voice theory. *Journal of Marketing Research* 19, 440–452.

Forster, P.T.D., 1999. Radiative forcing due to stratospheric ozone changes 1979–1997, using updated trend estimates. *Journal of Geophysical Research–Atmospheres* 104, 24395–24399.

Grenfell, J.L., Shindell, D.T., Koch, D., Rind, D., 2001. Chemistry–climate interactions in the Goddard Institute for Space Studies general circulation model 2. New insights into modeling the preindustrial atmosphere. *Journal of Geophysical Research–Atmospheres* 106, 33435–33451.

Grewal, R., Cote, J.A., Baumgartner, H., 2004. Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science* 23, 519–529.

Hair, J.F., Ringle, C.M., Sarstedt, M., 2011. PLS–SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice* 19, 139–151.

Hauglustaine, D.A., Brasseur, G.P., 2001. Evolution of tropospheric ozone under anthropogenic activities and associated radiative forcing of climate. *Journal of Geophysical Research–Atmospheres* 106, 32337–32360.

Hauglustaine, D.A., Granier, C., Brasseur, G.P., Megie, G., 1994. Impact of present aircraft emissions of nitrogen–oxides on tropospheric ozone and climate forcing. *Geophysical Research Letters* 21, 2031–2034.

Helm, S., Eggert, A., Garnefeld, I., 2009. Modelling the impact of corporate reputation on customer satisfaction and loyalty using PLS, in *Handbook of Partial Least Squares: Concepts, Methods, and Applications,* edited by Esposito Vinzi, V., Chin, W.W., Henseler, J., Wang, H., Springer, Berlin.

Henseler, J., Fassott, G., 2009. Testing moderating effects in PLS path models: An illustration of available procedures, in *Handbook of Partial Least Squares: Concepts, Methods, and Applications*, edited by Esposito Vinzi, V., Chin, W.W., Henseler, J., Wang, H., Springer, Berlin.

Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A., 2001. *Climate Change 2001: The Scientific Basis*, Cambridge University Press, Cambridge, UK.

Hwang, H., Malhotra, N.K., Kim, Y., Tomiuk, M.A., Hong, S.J., 2010. A comparative study on parameter recovery of three approaches to structural equation modeling. *Journal of Marketing Research* 47, 699–712.

Joreskog, K.G., 1993. Testing structural equation models, in *Testing Structural Equation Models*, edited by Bollen, K.A., Long, J.S., Sage Publication, Newbury Park.

Joreskog, K., 1978. Structural analysis of covariance and correlation matrices. *Psychometrika* 43, 443–477.

Joreskog, K.G., Wold, H., 1982. The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects, in *Systems under Indirect Observation: Part I*, edited by Joreskog, K.G., Wold, H., Amsterdam, North–Holland, pp. 263–270.

Kennedy, P., 2003. *A Guide to Econometrics*, Fifth Edition, MIT Press, pp. 209.

Lippmann, M., 2009. *Environmental Toxicants: Human Exposures and Their Health Effects*, John Wiley and Sons, Hoboken, New Jersey, USA.

Lippmann, M., 1993. Health effects of tropospheric ozone: A review of recent research findings and their implications to ambient air quality standards. *Journal of Exposure Analysis and Environmental Epidemiology* 3, 103–129.

Lohmoller, J.B., 1989. *Latent Variable Path Modeling with Partial Least Squares*, Physica–Verlag HD.

Mickley, L.J., Murti, P.P., Jacob, D.J., Logan, J.A., Koch, D.M., Rind, D., 1999. Radiative forcing from tropospheric ozone calculated with a unified chemistry–climate model. *Journal of Geophysical Research–Atmospheres* 104, 30153–30172.

Morris, R.E., Gery, M.S., Liu, M.K., Moore, G.E., Daly, C., Greenfield, S.M., 1989. Sensitivity of a regional oxidant model to variation in climate parameters, in *The Potential Effects of Global Climate Change on the United States*, edited by Smith, J.B., Tirpak, D.A., US Environmental Protection Agency, Office of Policy, Planning and Evaluation, Washington DC.

Ordonez, C., Mathis, H., Furger, M., Henne, S., Huglin, C., Staehelin, J., Prevot, A.S.H., 2005. Changes of daily surface ozone maxima in Switzerland in all seasons from 1992 to 2002 and discussion of summer 2003. *Atmospheric Chemistry and Physics* 5, 1187–1203.

Petter, S., Straub, D., Rai, A., 2007. Specifying formative constructs in information systems research. *MIS Quarterly* 31, 623–656.

Reinartz, W., Haenlein, M., Henseler, J., 2009. An empirical comparison of the efficacy of covariance–based and variance–based SEM. *International Journal of Research in Marketing* 26, 332–344.

Roelofs, G.J., Lelieveld, J., vanDorland, R., 1997. A three–dimensional chemistry general circulation model simulation of anthropogenically derived ozone in the troposphere and its radiative climate forcing. *Journal of Geophysical Research–Atmospheres* 102, 23389–23401.

Sanchez–Ccoyllo, O.R., Ynoue, R.Y., Martins, L.D., Andrade, M.D., 2006. Impacts of ozone precursor limitation and meteorological variables on ozone concentration in Sao Paulo, Brazil. *Atmospheric Environment* 40, S552–S562.

Seker, H., Serin, Y., 2004. Explanation of the relationships between seed yield and some morphological traits in smooth bromegrass (Bromus inermis Leyss.) by path analysis. *European Journal of Agronomy* 21, 1–6.

Steiner, A.L., Tonse, S., Cohen, R.C., Goldstein, A.H., Harley, R.A., 2006. Influence of future climate and emissions on regional air quality in California. *Journal of Geophysical Research–Atmospheres* 111, art. no. D18303.

Stevenson, D.S., Johnson, C.E., Collins, W.J., Derwent, R.G., Shine, K.P., Edwards, J.M., 1998. Evolution of tropospheric ozone radiative forcing. *Geophysical Research Letters* 25, 3819–3822.

Tenenhaus, M., Vinzi, V.E., 2005. PLS regression, PLS path modeling and generalized Procrustean analysis: A combined approach for multiblock analysis. *Journal of Chemometrics* 19, 145–153.

Wold, H., 1985. Partial least squares, in *Encyclopedia of Statistical Sciences*, edited by Kotz, S., Johnson, N.L., Wiley, New York.

Wold, H., 1982. Soft modeling: The basic design and some extensions, in *Systems under Indirect Observations: Part II*, edited by Joreskog, K.G., Wold H., Amsterdam, North Holland.