Journal of Research and Multidisciplinary
ISSN: 2622-9536 *Print*
ISSN: 2622-9544 *Online*
http://journal.alhikam.net/index.php/jrm
Volume 5, Issue 2, September 2022, Pages 584-601

Application of path modelling to identify causal relationships of *in vitro* production of sesquiterpenes in *Gyrinops walla* Gaetner

D.S.P. Munasinghe[1], S.Somaratne[2], S.R. Weerakoon[3], C. Ranasinghe[4]
[1]*Department of Botany, The Open University of Sri Lanka, Nawala, 10250, Sri Lanka.*
[2]*Department of Chemistry, The Open University of Sri Lanka, Nawala, 10250, Sri Lanka.*
[3]*Department of Chemistry, The Open University of Sri Lanka, Nawala, 10250, Sri Lanka.*
[4]*Department of Chemistry, The Open University of Sri Lanka, Nawala, 10250, Sri Lanka.*

Email : sachujd@gmail.com

Abstract

The recent recovery of *Gyrinops walla* as a potential producer of market-quality agarwood in mature damaged woods and branches, the intense illicit felling and exportation walla leading to the verge of extinction from Sri Lankan flora. The sustainable utilization of G. *walla* undoubtedly enhances the foreign exchange of the country by tissue culture-based techniques for the sustainable exploitation and conservation of the vulnerable G. *walla* species. Partial Least Square Structural Equation Model (PLS-SEM) was developed to elucidate the causal relationships between the latent variables; the cultured materials (CULTM), type of elicitation (ELI), concentration/intensities of the elicitors (CON), time of incubation (INCB) with the occurrence of sesquiterpenes (SES). The path coefficients of the PLS-SEM model showed direct effects of CULTM -> SES ($\beta = -0.509$, $t = 6.468$, $p < 0.05$) (hypothesis - H3), INCB -> SES ($\beta = 0.421$, $t = 5.037$, $p < 0.05$) (hypothesis - H1) and ELI + CON -> SES ($\beta = -0.282$, $t = 5.792$, $p < 0.05$) (hypothesis-H6) on induction of the production of sesquiterpenes. The PLS-SEM model developed the study proved the applicability in exploring the causal relationships between the deterministic factors that effect on artificial elicitation in the production of sesquiterpenes.

Keywords: Agarwood, Artificial elicitation, *Gyrinops walla*, Partial Least Square Structural Equation Model, Path modelling

1. Introduction

Gyrinops walla Gaetner of family Thymelaeacea, commonly known as 'Walla patta', is endemic to Sri Lanka and commonly occurs in the wet zone of the country (Townsend, 1981). It is mainly distributed in forest areas at lower elevations (below 1525m from sea level) in Sri Lanka and is restricted to the Western, Sabaragamuwa (southern margin, adjacent to the Sinharaja rain forest) and Southern Provinces of the country (Gunatilleke, 2005). "Walla patta" gained popularity in 2012, with the discovery of its similarity to the commercially valued agarwood, a fragrant resinous wood used in perfumery and medicine that obtained from the trunk and branches of *Aquilaria* and *Gonystulus*. Following this discovery, large scale illicit felling and exporting of *G. walla* chips has been reported. The prohibition was also brought into force and *G. walla* was categorised as a vulnerable species according to the IUCN Red List Categories (MOE, 2012). Conventionally, destructive harvesting is used to obtain agarwood chips and oil which naturally occur in trees or after inducing agarwood formation artificially. However, such destructive harvesting practices for obtaining agarwood from *G. walla* have to be halted because it is the only representative species of agarwood in Sri Lanka and its population is declining rapidly due to illicit harvesting and trading. As a result, it is necessary to protect this valuable and vulnerable species using biotechnological methods such as tissue culture to facilitate the production of agarwood constituents *ex situ* by elicitation of callus/cell cultures.

Callus and cell suspension culture systems are used nowadays for large-scale production of plant cells from which secondary metabolites are extracted. Cell cultures assist in two major ways in production of plant secondary metabolites by yielding defined standard phytochemicals in large volumes and eliminating the presence of interfering compounds that occur in the field-grown plants (Lila, 2005). The major advantage of the cell culture is that synthesis of bioactive secondary metabolites is possible running in controlled environment, independently from climate and soil conditions (Karuppusamy, 2009). Developments made in the callus and cell culture has made the possibility of producing wide range of medicinally important compounds such as alkaloids, terpenoids, steroids, saponins, phenolics, flavonoids and amino acids (Vijayasree et al. 2010).

Even though, optimization of factors required for the production of plant secondary metabolites such as the type of elicitors and their concentration, incubation time, etc., have been discussed, the casual relationship of the factors are not clearly demonstrated. For such approach, predictions should be suggested and evaluated. Therefore, a systematic study frame for better discovering causal relationship in *in vitro* production of plant secondary metabolites is still required.

Path analysis is a statistical procedure which was credited by Sewall Wright during his attempt in developing Shifting Balance Theory (SBT) of population genetics (Wright, 1960). This procedure allows researchers to investigate the cause-effect relationship within a system of composite variables. Even though, path analysis is compatible to multiple regression in which the influence of many predictors on a criterion variable can be predicted, in path analysis two or more criterion variables can be examined simultaneously.

The SEM has originally used as a technique in marketing researches (Igbaria et al. 2015; Chin et al. 2003; Yi and Davis, 2003). However, it has been used widely in unraveling the research questions in a number of fields of studies. Recently, the SEM has gained acceptance and extensively used in biotechnology (Chen et al. 2013; Ghosh et al. 2020) and ecology (Fan et al. 2016).

The Partial Least Square method of SEM (PLS path modeling), originally developed by Wold (Wold, 1982 and 1958) and Lohmöller (Lohmöller, 1989), is an alternative to the Covariance Based SEM (CB-SEM), which is appropriate for circumstances where normality of the data is violated. The PLS path modelling has been referred to as “soft–modeling–technique” with minimum requirement on measurement scales, sample sizes and residual distributions (Monecke and Leisch, 2012). Even though other SEM tools available, the selection of Partial Least Square Structural Modeling (PLS-SEM) was made by considering the several factors. The PLS was reported to be developed to manage both formative and reflective indicators and in other SEM techniques this facility is not available. Therefore, this ability of the PLS-SEM allows the description of the nature of the relationship that the researcher believed between the manifest variables and the latent constructs. However, Wold, 1982 emphatically, is in the opinion that PLS is not be used for the confirmatory testing and he suggests that PLS-SEM can be used for predictive and the exploratory approaches for possible causality. Other techniques of SEM are basically depending on the accuracy of the parameters. Furthermore, PLS-SEM does not require assumption of multivariate normality as other SEM techniques such as LISREL (Goodhue et al. 2012) and AMOS (Lowry and Gaskin, 2014). PLS-SEM, being a nonparametric method, find no problem of multi-collinearity (Afthanorhan et al. 2016). More importantly, the sample size requirement of PLS-SEM is comparatively lower than that of other SEM techniques (Chin, 1998; Chin and Newsted, 1999). As a rule, PLS-SEM require a sample size of > = 10 fold of the number of indicators in complex formative construct or 10 times the largest number of independent constructs connecting to an endogenous construct (Chin et al. 2003).

In the present study the Structural Equation Modelling (SEM) was chosen for the data analysis. SEM is reported be capable in statistical testing of prior theoretical assumptions against empirical data and evaluates the characteristics of the scales used in measuring the theoretical constructs and estimates the hypothesized relationships among supposed constructs (Barclay et al. 1995; Chin et al. 2003). Thus, SEM can be employed in answering a set of interconnected research questions simultaneously *via* both measurement and structural model.

2. Materials and methods

2.1. Artificial elicitation of calli and cell suspension cultures of *Gyrinops walla*

Leaves and nodal explants were obtained from six-month-old G. *walla* saplings to raise calli and cell suspensions. Calli were raised according to Munasinghe et al. (2020) using varying concentrations of 1- Naphthaleneacetic acid (NAA) and 6-Benzylaminopurine (BAP) in Murashige and Skoog medium (MSM) and cultured under ambient culture conditions. The calli resulted were used to establish cell suspension. Healthy calli and cell suspensions were elicited with salicylic acid (SA) and methyl jasmonate (MJ), and sterilized fungal homogenate (carbohydrate equivalents) of *Fusariym oxysporum*, *Phaeocremonium parasitica*, *Aspergillus niger*, *Trichoderma viride*, *Penicillium commune* and *Lasidiplodia theobromae* fungal strains with varying concentrations. An electro-elicitation apparatus was designed and built to stimulate calli and cell suspensions under varying current intensities. The elicited and calli and cell suspensions were harvested at different time periods and chemical constituents of harvested calli and cell suspensions were extracted with 100% hexane to analyse with GC-MS. The data were served to developed Partial Least Square Structural Equation Model (PLS-SEM) to elucidate the causal relationships between the latent variables; the cultured materials (CULTM), type of elicitation (ELI), concentration/intensities of the elicitors (CON), time of incubation (INCB) with the occurrence of sesquiterpenes (SES).

2.2. Identification of causal relationship between latent variables

The data of artificial elicitation were subjected to Partial Least Squared Structural Equation Models (PLS-SEM) to explore causal relationship between factors using SmartPLS version 3.3.2 (SmartPLS GmbH, P.O. Box 1123. D-25474 Bönningstedt. Germany). Conceptual models of this study belonged to the Reflective-Formative type (Ringle et al. 2012), which has characteristics of reflective measurement models (outer models) and formative structural models (inner models), which all models were consist of latent variables and measurable variables (Figure 1).

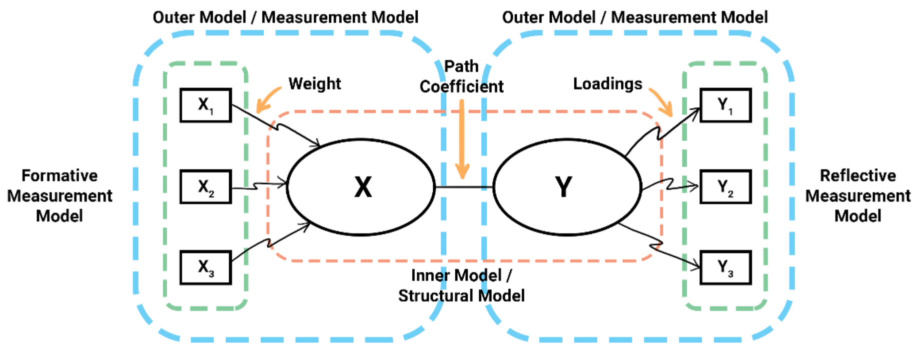

Figure 1: Typical PLS-SEM structure.

The developed model in the present study included three constructs i.e. nature of the cultured material (CULTM), type of elicitation + concentrations/intensities of the elicitors (ELI + CON) and duration of incubation (INCB) and four dependent variables (γ-selinene, β-caryophyllene, α-cadinol and α-guaiene) encompassed only one construct, sesquiterpenes (SES).

In PLS-SEM, analysis and interpretation carry out in two phases successively. The PLS-SEM, initially assess and refine the adequacy of the measurement model and followed by the assessment and evaluation of the structural model. This process ensures the reliability and validity of the measures prior to any attempts in making conclusion on the structural model. The assessment of measurement model is essential step in the PLS-SEM modelling since it provides systematic assessment for the reliability and validity of the scales used in measuring the latent constructs and their manifest variables (Loehlin, 1998). A number of steps were used in the assessment of the measurement model such as the assessment of convergent and discriminant validity, and evaluation of the measure's reliability. In order to test the validity and reliability of the constructs (Rossiter, 2002) procedure for scale development was followed. First, convergent, and discriminate validity were determined and finally, reliability of the scale items was evaluated.

(a) Assessment of measurement/outer model

(i) Convergent Validity

According to Hair et al. (2006), the convergent validity specifies the items that are indicators of a construct required to share a high proportion of variance. Therefore, convergent validity of the items in the scale was assessed based on three criteria: a). Factor loadings should be greater than 0.50 as proposed by Hair et al. (2017), composite reliability for each construct should exceed 0.70, c) Average Variance Extracted (AVE) for each construct should be above the recommended cut-off 0.50 (Fornell and Larker, 1981).

(ii) Discriminant validity

Discriminant validity examine the extent of the uniqueness of the measures (Peter and Churchill, 1986) and each measurement in a construct required to be unique and differ from the other measures. The discriminant validity determined by using several methods. Examination of Average Variance Extracted (AVE) often uses as a common method of assesse discriminant validity (Gerbing and Anderson, 1988). However, recently, Hair et al. (2017) and Henseler et al. (2015) proposed a new criterion for the determination of discriminant validity. According to Henseler et al. (2015), for a satisfactory cut-off value was > 0.90. Therefore, discriminant validity was established by examining the cross loadings of each item in the constructs, the square root of AVE calculated for each construct and Heterotrait-monotrait (HTMT) ratio of correlation. Based on prior research and their study results, Henseler et al. (2015) suggest a threshold value of 0.90 if the path model includes constructs that are conceptually very similar. In other words, an HTMT value above 0.90 suggests a lack of discriminant validity. When the constructs in the path model are conceptually more distinct, a lower and thus more conservative threshold value of 0.85 seems warranted (Henseler et al. 2015). All the items should have higher loading on their corresponding construct than the cross loadings on the other constructs in the model. The square root of AVE for all factors should be greater than all the correlations between that construct and other constructs.

(iii) Reliability of Measures

The final phase of the examination of the construct validity is the determination of reliability of the construct items. Reliability is a measure of a set of indicators' internal consistency and possibility of obtaining the same results on repeated trials. Though reliability is essential, however, insufficient for determination of validity of a measure, even the measures with high reliability would not be valid in measuring the importance of the constructs (Hair et al. 2006). Reliable indicators should measure the same construct is determined by Cornbach α or composite reliability i.e., composite α value. These indices were used to assess the reliability of the constructs used in the model building. According to Hair et al. (1998), construct reliability coefficients should surpass the 0.70 lower limits. In order to assess the collinearity issues, according Hair et al. (2017) the VIF score for each latent variable should be less than 5. Subsequently, the PLS-SEM model was bootstrapped to obtain path coefficients.

(b) Assessment of the outer Structural Model

The structural models are used to capture the mutual linear regression relationships of the endogenous constructs (Hair et al. 1998) and has the capability in specifying pattern of the relationships among the constructs (Loehlin, 1998). Therefore, structural modeling is a growing area and popular among researches because of its ability to achieve direct testing of the theory of interest (Cheng, 2001).

The model developed in the study was assessed allying the criteria of: 1) path coefficients (β); 2) significant of path coefficient (probability value); and 3) variance explain (R^2). The validation of the structural model was achieved using SmartPLS 3.3.2. The designing of the model in PLS according to the procedures provide in the SmartPLS Guide (Ringle et al. 2005). Bootstrap re-sampling method was used to test the statistical significance of each path coefficient (Chin, 1998). Five thousand (5000) iterations using randomly selected sub-samples were used to estimate the theoretical model and hypothesized relationships. The criterion put forth by Rossiter (2002) states that for the structural model all paths should result in a T-statistic value greater than 2 and latent variable R^2 greater than 50%.

(c) Evaluation of the Inner Structural Model

This included observing the model's predictive relevancy and the relationships between the constructs. The coefficient of determination (R^2), Path coefficient (β value) and T-statistic value, Effect size (f^2), the Predictive relevance of the model (Q^2) , q^2 and RMSR indexes are the key standards for evaluating the inner structural model (Hair et al. 2017).

(i) Value of R^2 and Q^2

The coefficient of determination (R^2) indicates the total effect size and variance explained in the endogenous construct of the structural model. Therefore, R^2 could be considered as a measure of the predictive accuracy of the model. According to Henseler et al. (2015) and Hair et al. (2017), an R^2 value of 0.75 is considered substantial, an R^2 value of 0.50 is regarded as moderate, and an R^2 value of 0.25 is considered as weak.

The Q^2 statistics are used to measure the quality of the PLS path model, which is calculated using blindfolding procedures, and cross-validated redundancy was performed. The Q^2 criterion recommends that the conceptual model can predict the endogenous latent constructs. In the SEM, the Q^2 values measured must be greater than zero for a particular endogenous latent construct. A Q^2 value larger than zero for a certain endogenous latent variable indicates that the PLS path model has predictive relevance for this construct (Hair et al. 2017) and they are in the opinion that Q^2 values higher than 0, 0.25, and 0.5 depict small, medium, and large predictive relevance of the PLS-path model while q^2 values measure the effect size of the Q^2 values. Chin (1998) is in the opinion that the effect size (in %) of PLS constructs, falls within the range of small (0.02), medium (0.15), or large (0.35). For the sake of completeness of the analysis, Q^2 and q^2 values also calculated and reported.

(ii) Estimation of Path Coefficients (β) and T-statistics

The constructed hypotheses were tested through the standardized values of path coefficients and their significance obtained from the bootstrapping procedure of the SmartPLS. The path coefficient β denoted relationship between the unit variations in the dependent construct and the unit variation in the independent construct. The greater the path coefficient value, the more the substantial effect on the endogenous latent construct. The path coefficient coefficients for all path in the hypothesized model was computed. However, the β value should be verified for its significance level through the T-statistics to test the hypotheses. The bootstrapping procedure was used to evaluate the statistical significance of the hypothesis. The levels of significance of path coefficient and T-statistics values calculated through a bootstrapping procedure using 5000 subsamples (without sign changes).

(iii) Measuring the Effect Size (f^2)

The f^2 is the size of the impact of each exogenous latent construct on the endogenous latent construct. When an independent construct is deleted from the path model, it changes the value of the coefficient of determination (R^2) and defines whether the removed latent exogenous construct has a significant influence on the value of the latent endogenous construct.

(d) Testing of Model Fit

Model fitness is usually accompanied by diagnostic three model fitting parameters: a) Standardized Root Mean Square Residual (SRMR), b) Normed Fit Index (NFI) and c) exact model fit based on the bootstrapped statistical inference. The SRMR is the difference between the observed correlation and the model implied correlation matrix and values less than 0.08 are considered a good fit (Hu et al. 1998). In addition, Henseler et al. (2015) have introduced the SRMR as a goodness of fit measure for PLS-SEM in order to avoid the model misspecification. The NFI fit index with an incremental fit measure computes the Chi-square value of the proposed model and compares value against a benchmark (Bentler and Bonett, 1980) and NFI > 0.9 often considered as an acceptable fit. The NFI value obtained for the model was 0.89. However, NFI is not applicable formative constructs. The exact model fit can be used to examine the difference between the empirical covariance matrix and the covariance matrix inferred from the composite factor model. In this regard, Dijkstra and Henseler (2015a and 2015b) proposed the d_ULS (the squared Euclidean distance) and d_G (the geodesic distance) as the two different methods to compute discrepancy. The values of d_ULS and d_G be non-significant ($p > 0.05$) for an acceptable model fit. Further, Henseler et al. (2017) suggested that d_ULS and d_G < 95% bootstrapped quartile in model fitness (HI 95% of d_ULS = and HI 95% of d_G).A number of models were developed and the most parsimonious model often selected by using various information criteria. In the present study, Akaike Information Criterion (AIC), Corrected AIC (AICc) and Bayesian Information Criterion (BIC) were supported to estimate the relative expected Kullback–Leibler distance (Tumminello et al. 2007) to the unknown true model, correct AIC's tendency to over-fit (select a complicated model) under small samples and derived using Bayesian argument; adjusts AIC for model complexity by using a stronger penalty for over-fitting, respectively.

$$AIC_{PLS} = n\,[\log(SS_{errork} / n) + 2pk / n] \quad (1)$$

$$AICc_{PLS} = n\,[\log(SS_{errork} / n) + n + pk / n - pk - 2] \quad (2)$$

$$BIC_{PLS} = n\,[\log(SS_{errork} / n) + pk\log(n) / n] \quad (3)$$

where, SS =sums of squared error, n= number of samplesize, p = probability, k= number of parameters.

(e) Testing of hypothesis

Based on the theoretical framework, a model (Figure 2) was developed with a formative construct and tested with SmartPLS software. An initial model was developed with all constructs and necessary optimizations were made on the model according to the Manual of the software (Hair et al. 2017) model included the following hypotheses and was tested.

- H1: Time of incubation has direct effect on the occurrence of sesquiterpenes
- H2: Type of cultured material has direct effect on time of incubation
- H3: Type of cultured material has direct effect on the occurrence of sesquiterpenes
- H4: Type of cultured material has direct effect on concentration/intensities of elicitors
- H5: Concentration/intensities of elicitors has direct effect on time of incubation
- H6: Concentration/intensities of elicitors has direct effect on occurrence of sesquiterpenes

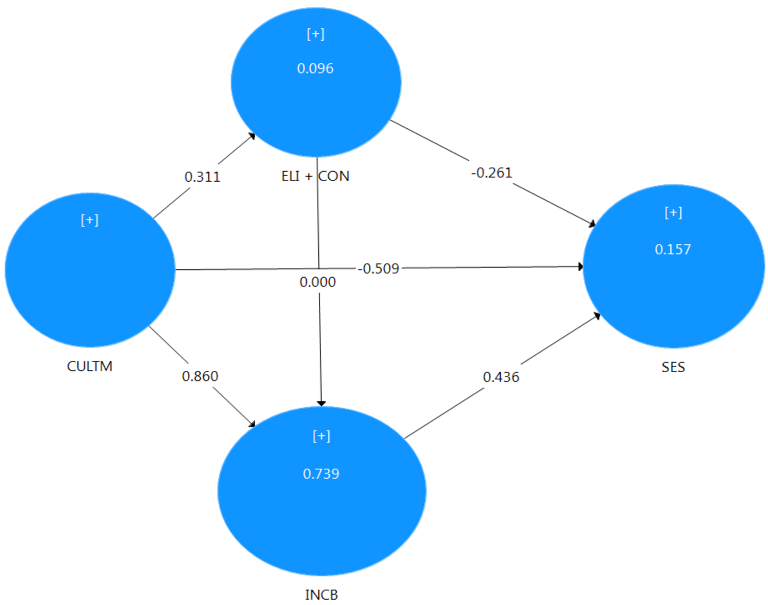

Figure 2: Path model of causal relationships of artificial elicitation

3. Results

3.1. Assessment of measurement model

The loadings lesser than 0.70 were removed from the initial model and therefore, α-guaiene was removed from the construct SES. The path coefficients which are > 0.70 of the finalised path model are summarized in Table 1.

Table 1: Outer loadings

	INCB	CULTM	SES	ELI + CON
Concentrations /intensities				0.997
Type of elicitation				0.877
Plant material		1.000		
Incubation time	1.000			
α-cadinol			0.904	
β-caryophyllene			0.941	
γ-selinene			0.913	

The loadings obtained for the items chosen from the artificial elicitation supported the convergent validity of the four constructs. The higher loadings lead to conclude that the measures satisfied the convergent validity. The composite reliability, i.e., Cronbach α values ranged from the 0.913 to 0.948 (Table 2). All values exceed the cut-off specified suggesting that convergent validity is preserved.

Table 2: The Cronbach α values and composite reliability values

	Cronbach's Alpha	Composite Reliability
INCB	Single item construct	Single item construct
CULTM	Single item construct	Single item construct
SES	0.918	0.948
ELI + CON	0.913	0.941

The between construct correlations and the measures of the validity between the variables and construct are shown in Table 3 and shows the AVE and cross factor loading extracted for all latent variables. All the items show higher loadings on their corresponding construct compared to the cross loadings on the other constructs in the model. The AVE for each latent factor exceeded the respective squared correlation between factors and HTMT ratios were below 0.9 (Table 3 and 4.), thus providing evidence of discriminant validity (Fornell and Larker, 1981).

Table 3: Fornell-Lacker Criterion. Correlation and AVE values (squared correlation values) in diagonal bold

	INCB	CULTM	SES	ELI + CON
INCB	1.000			
CULTM	0.860	1.000		
SES	-0.088	-0.231	0.858	
ELI + CON	0.255	0.296	-0.325	0.889

Table 4: Heterotrait-monotrait (HTMT) ratio of correlation

	INCB	CULTM	SES	ELI + CON
INCB				
CULTM	0.860			
SES	0.093	0.242		
ELI + CON	0.195	0.227	0.332	

According to values in Table 5, it is evident that VIF of all variables included in the study were < 5. Thus, all values were within the acceptable range and led to conclude satisfactory reliability.

Table 5: Outer VIF values

	VIF
Concentrations/intensities	3.396
Type of elicitation	3.396
Nature of the cultured material	1.000
Time of incubation	1.000
α-cadinol	3.610
β-caryophyllene	3.316
γ-selinene	3.021

3.2. Evaluation of outer structure model

According to the result obtained from estimates of path coeffcients (β) and T-statistics of the paths INCB ->SES, CULTM ->INCB, CULTM ->SES, CULTM ->ELI +

CON and ELI + CON ->SES were significant ($p < 0.05$) (Table 6). The β values of the path ELI + CON ->INCB were statistically insignificant ($p > 0.05$).

Table 6: Path coefficients with corresponding T-statistics obtained from the bootstrap procedure

Hypothesis		Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values
H1	INCB -> SES	0.436	0.424	0.084	5.037	S
H2	CULTM ->INCB	0.860	0.860	0.011	76.861	S
H3	CULTM -> SES	-0.509	-0.512	0.079	6.468	S
H4	CULTM ->ELI + CON	0.311	0.302	0.043	6.862	S
H5	ELI + CON ->INCB	0.000	0.000	0.025	0.000	NS
H6	ELI + CON -> SES	-0.261	-0.281	0.049	5.792	S

S - Significant at $p < 0.05$; NS = Not significant at $p> 0.05$.

3.3. Evaluation of the Inner Structural Model

(a) R^2 and Q^2

The value of R^2 of the present model was 0.74 for the INCB.The PLS-SEM algorithm results indicated R^2 value of 0.74 for duration of incubation time which explains 74% variance. R^2 of SES was 0.17 and which denotes 17% variance is explained by SES. The R^2 value of ELI + CON was 0.08 which is the lowest in the PLS-SEM model. Although, researchers believe that r-squared value is always contextual and is based upon research settings (Hair et al. 2017), Falk and Miller (1992) suggest that if a square value, in the social science context, reaches up to 10 percent, it may be considered accepted. However, in the field of biological sciences, this criterion cannot be applied. From Table 7, it is clear that the Q^2 values obtained for proposed model was greater than 0 and was higher than the threshold limit, and supports that the path model's predictive relevance was adequate for the endogenous construct. The effect size (f^2) of all paths in the model fall within the 'small' category.

Table 7: The R^2 and Q^2, and f^2 of the PLS-SEM model

	R^2	Q^2	f^2			
			INCB	**CULTM**	**SES**	**ELI + CON**
INCB	0.739	0.731			0.056*	
CULTM			2.586		0.080*	0.096
SES	0.172	0.143				
ELI + CON	0.088	0.029	0.000		0.087*	

Effect size (f^2) of 0.02, 0.15, and 0.35 indicates small, medium, and large effect, respectively [28] and are shown with * mark. Values of effect size are: * = higher than 0.02 – Small; ** = 0.15 – Medium; *** = 0.35 - Large.

The effect size of the predictive relevance (q^2) according to Latan and Ramli (2013) range from 0.02, 0.15, and 0.35 indicates small, medium, and large effect, respectively. In the present model, SES -> Concentration/intensities and type of elicitation -> SES were -0.0045 and 0.0022 respectively.

(b) **Measuring the Effect Size (*f^2*)**

The f^2 values obtained after boostrapping were shown in the Table 8. According to the table, except Type of elicitation/concentration -> Incubation time, the rest of the paths were significant at $p < 0.05$.

Table 8: Path coefficient and effect size (f^2).

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values
Incubation time -> Sesquiterpenes	0.056	0.059	0.023	2.429	S
Plant material -> Incubation time	2.586	2.587	0.210	12.316	S
Plant material -> Sesquiterpenes	0.080	0.083	0.026	3.012	S
Plant material -> Type of elicitation/concentration	0.096	0.103	0.035	2.761	S
Type of elicitation/concentration -> Incubation time	0.000	0.002	0.003	0.000	NS
Type of elicitation/concentration -> Sesquiterpenes	0.087	0.091	0.034	2.563	S

S - Values in effect size are significant $p < 0.05$; NS – Not significant $p > 0.05$.

3.4. Testing model fit

Since proposed model was a saturated model without free paths, the saturated model (measurement) fit values and the estimated model (structural model) fit values were the same. The SRMR value was 0.094 (< 0.06) and the d_ULS was 0.249 (< bootstrapped HI 95% of d_ULS = 0.130 and d_G = 0.131 < bootstrapped HI 95% of d_G = 0.166) indicating the data fits the model well for the data (Table 9). Further, HTMT ratio of the model falls within the range of acceptable values. The HTMT values in the paths of the models do not indicate 1 under 2.5 % and 97.5% bounds (Table 10).

Table 9: Model fit summery of bootstrapping

	Saturated Model	Estimated Model
SRMR	0.094	0.094
d_ULS	0.249	0.249
d_G	0.238	0.238

Table 10: HTMT ratio of the final model obtained after boostrapping.

	Original Sample (O)	Sample Mean (M)	2.5%	97.5%
CULTM ->INCB	0.860	0.860	0.843	0.875
SES ->INCB	0.093	0.094	0.019	0.177
SES ->INCB	0.242	0.242	0.160	0.323
ELI + CON ->INCB	0.195	0.199	0.173	0.232
ELI + CON ->CULTM	0.227	0.229	0.204	0.260
ELI + CON ->SES	0.332	0.331	0.233	0.423

In general, the lower the values of AIC, AICc and BIC, the better the model fit and reflect the parsimony of the model (Akaike, 1985; Schwarz and Gideon, 1978) (Table 11).

Table 11: Model selection criteria.

	AIC (Akaike's Information Criterion)	AICc (Corrected Akaikes Information Criterion)	BIC (Bayesian Information Criteria)
INCB	-615.973	-151.885	-603.566
SES	-80.131	384.000	-63.589
ELI + CON	-39.422	424.631	-31.150

3.5. Testing of Hypotheses

Summary of the results (β) obtained from the bootstrapping is reported in Table 6. According to the table, H1 - INCB ->SES (β = 0.436, t = 5.037, p < 0.05), H2- CULTM ->INCB (β = 0.860, t = 76.861, p < 0.05), H3 -CULTM ->SES (β = -0.509, t = 6.468, p < 0.05), H4 -CULTM ->ELI + CON (β = 0.311, t = 6.862, p < 0.05) and H6–ELI + CON ->SES (β = -0.261, t = 5.792, p < 0.05) were statistically significant except H5. Therefore, the five hypotheses, H1, H2, H3, H4 and H6 were accepted and H5 was rejected.

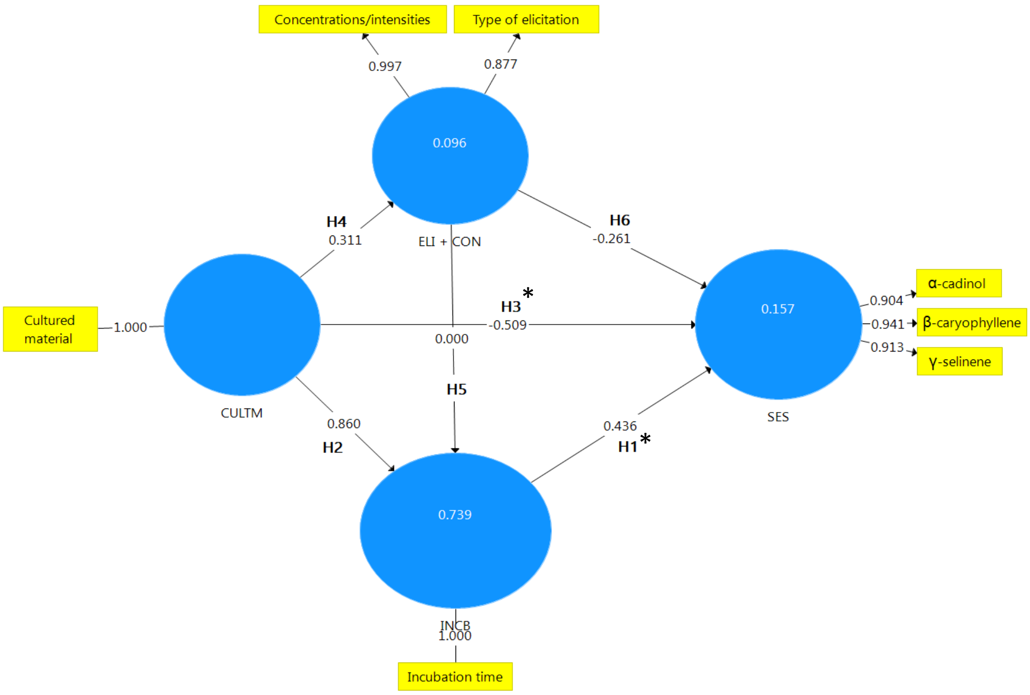

Figure 3: The final model indicating the significant hypotheses.

*indicates statistical significant at p < 0.05.

4. Discussion

Path coefficients with corresponding T-statistics obtained from the bootstrap procedure imply higher the magnitude of the path coefficient, greater the effectiveness of the path. Path coefficients of the hypotheses of the final model had direct effect of Incubation time (INCB) -> Induction of sesquiterpenes (SES) (β = 0.436, t = 5.037, p < 0.05), culture material (CULTM) -> Induction of sesquiterpenes (SES) (β = -0.509, t = 6.468, p < 0.05) and Elicitors and concentration /intensities (ELI + CON -> Induction of sesquiterpenes (SES) (β = -0.261, t = 5.792, p< 0.05) (Figure 3, Table 6). These results led to conclude that there is a significant direct effect of the nature of the cultured material (H3) than that of with the time of incubation (H1) and type of elicitation + concentration/intensities of elicitors (H6) on sesquiterpene production. As far as the indirect effects of the path model are concerned, two effects (H4 x H6 and H2 x H1) were identified (Figure 3). Among them H2 x H1 = β = 0.375, indicating a significant indirect effect for CULTM -> INCB -> SES path.

5. Conclusion

The present study reveals the first time application of path modelling to explore the causal effect of deterministic factors affecting sesquiterpene synthesis in calli and cell suspensions of G. *walla* using artificial elicitation. Comparatively, incubation time has greater effect on the production of sesquiterpens.

Acknowledgment

Financial assistance provided by OUSL Competitivee Research Grant 2016 (OU201601) is highly appreciated.

References

Afthanorhan, A., Awang, Z., & Mamat, M., 2016. A comparative study between GSCA-SEM and PLS-SEM. MJ Journal on Statistics and Probability, 1(1), 63-72.

Akaike, H., 1985. Prediction and entropy. In *Selected Papers of Hirotugu Akaike* (pp. 387-410). Springer, New York, NY.

Barclay, B., Higgins, C., &Thompson, R., 1995. The Partial Least Squares (PLS) Approach to Causal Modelling: Personal Computer Adoption and Use as an Illustration. Technology Studies, 2 (2),285 – 309.

Bentler, P.M. & Bonett, D.G., 1980. Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88 (3), 588-606.

Chen, W., Li, J., Zhu, H., Xu, P., Chen, J., & Yao, Q., 2017. Arbuscular mycorrhizal fungus enhances lateral root formation in *Poncirus trifoliata* (L.) as revealed by RNA-Seq analysis. Frontiers in Plant Science, 8, 2039.

Cheng, E. W., 2001. SEM being more effective than multiple regression in parsimonious model testing for management development research. Journal of Management Development, 20(7), 650-667.

Chin, W. W., Marcolin, B. L., & Newsted, P. R., 2003. A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Information systems research, 14(2), 189-217.

Chin, W.W., 1998. The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), Modern methods for business research (pp. 295-336). Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W., & Newsted, P., 1999. Structural equation modeling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), Statistical Strategies for Small Sample Research (pp. 307–342). Thousand Oaks: SAGE.

Dijkstra, T. K., & Henseler, J., 2015a. Consistent partial least squares path modeling. MIS quarterly, 39(2).

Dijkstra, T. K., & Henseler, J., 2015b. Consistent and asymptotically normal PLS estimators for linear structural equations. Computational statistics & data analysis, 81, 10-23.

Falk, R. F. & Miller, N. B., 1992. A Primer for Soft Modeling. Akron, OH: University of Akron Press.

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C., 2016. Applications of structural equation modeling (SEM) in ecological studies: an updated review. Ecological Processes, 5(1), 19.

Fornell, C., & Larker, D., 1981. Structural equation modeling and regression: guidelines for research practice. Journal of Marketing Research, 18(1), 39-50.

Gerbing, D. W., & Anderson, J. C., 1988. An updated paradigm for scale development incorporating uni-dimensionality and its assessment. Journal of marketing research, 25(2), 186-192.

Ghosh, A., Pandey, B., Agrawal, M., & Agrawal, S. B., 2020. Interactive effects and competitive shift between *Triticum aestivum* L.(wheat) and *Chenopodium album* L.(fat-hen) under ambient and elevated ozone. Environmental Pollution, 114764.

Goodhue, D., Lewis, W., & Thompson, R., 2012. Does PLS Have Advantages for Small Sample Size or Non-Normal Data? MIS Quarterly, 36(3), 981-1001.

Gunatilleke, I.A.U.N., Gunatilleke, C.V.S., Dilhan, M.A.A.B., 2005. Plant biogeography and conservation of the south-western hill forests of Sri Lanka. Raf. Bul. Zoo. ; 12(1): 9-22.

Hair. J. F., Black., W. C., Babin., B. J., Anderson., R. E., &Tatham., R., 2006. Multivariant Data Analysis. New Jersey: Pearson International Edition. 9.

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M., 2017. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)., 2nd Ed., Thousand Oakes, CA: Sage.

Hair, J. F., Anderson, R. E., Tatham, R. L. and Black, W. C., 1998. Multivariate data analysis. (5^{th} ed.). Prentice-Hall, New Jersey.

Henseler, Jr., Ringle, C. M., and Sarstedt, M., 2015. A New Criterion for assessing discriminant Validity in Variance-based Structural Equation Modeling., Journal of the Academy of Marketing Science, 43(1): 115-135.

Hu, Li-tze, Bentler, Peter M., 1998. Fit indices in covariance structure modeling: sensitivity to under parameterized model misspecification. Psychological Methods 3 (4), 424–453.

Igbaria, M., Guimaraes, T., & Davis, G. B., 1995. Testing the determinants of microcomputer usage via a structural equation model. Journal of management information systems, 11(4), 87-114.

Karuppusamy, S., 2009. A review on trends in production of secondary metabolites from higher plants by *in vitro* tissue, organ and cell cultures. Journal of Medicinal Plant Research, 3(13), 1222-1239.

Lila, K. M., 2005. Valuable secondary products from *in vitro* culture, Secondary Products *In Vitro*. CRC Press LLC.

Loehlin, J.C. (1998). Latent variable models: An introduction to factor, path, and structural analysis. Hillsdale, NJ: Erlbaum.

Lohmöller, J.B., 1989. Latent Variables Path Modeling with Partial Least Squares. Physica-Verlag, Heildelberg.

Lowry, P. B., & Gaskin, J., 2014. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. IEEE transactions on professional communication, 57(2), 123-146.

MOE, 2012. The Notational Red List 2012 of Sri Lanka. Conservation status of the Fauna and Flora. Ministry of Environment, Colombo, Sri Lanka. Viii :476.

Monecke, A. & Leisch, F., 2012. SEM-PLS: structural equation modeling using partial least squares. Journal of Statistical Software, 48 (3), 1-32.

Munasinghe, S. P., Somaratne, S., Weerakoon, S. R., & Ranasinghe, C., 2020. Prediction of chemical composition for callus production in *Gyrinops walla* Gaetner through machine learning. *Information Processing in Agriculture*, 7(4), 511-522.

Peter, J. P., & Churchill, G. A., 1986. Relationships among research design choices and psychometric properties of rating scales: a meta-analysis. Journal of Marketing Research, 23(1), 1–10.

Ringle, C. M., Sarstedt, M., & Straub, D. W., 2012. Editor's Comments: A Critical Look at the Use of PLS-SEM in" MIS Quarterly". MIS quarterly, iii-xiv.

Ringle, C. M., Wende, S., & Will, A., 2005. SmartPLS 2.0 (M3) beta software. Hamburg, Germanay.

Rossiter, J. R., 2002. The C-OAR-SE procedure for scale development in marketing. International Journal of Research in Marketing, 19(4): 305-335.

Schwarz, Gideon E., 1978. "Estimating the dimension of a model", Annals of Statistics, **6** (2): 461–464.

Townsend, C. C., 1981. Thymeliaceae. In A revised Handbook to the Flora of Ceylon, Vol. II, M. D. Dassanyake and F. R. Fosberg (edts.), Amerind Publishing Co. Pvt. Ltd., New Delhi.

Tumminello, M., Lillo, F., & Mantegna, R. N., 2007. Kullback-Leibler distance as a measure of the information filtered from multivariate data. Physical Review E, 76(3), 031123.

Yi, M. Y., & Davis, F. D., 2003. Developing and validating an observational learning model of computer software training and skill acquisition. Information systems research, 14(2), 146-169.

Vijayasree, N., Udayasri, P., Aswani, K.Y., Ravi, B. B., Phani, K .Y, and Vijay, V. M., 2010. Advancements in the Production of Secondary Metabolites. Journal of Natural Products,3, 112-123.

Wright, S., 1960. Path coefficients and path regressions: alternative or complementary concepts?. Biometrics, 16(2), 189-202.

Wold, H., (1982). Soft modeling: the basic design and some extensions. In: J Loreskog, K.G., Wold, H. (Eds.), Systems under Indirect Observation, Part 2. North-Holland, Amsterdam, 1–54.

Wold, H., (1985). Partial least squares. In: Kotz, S., Johnson, N.L. (Eds.), Encyclopedia of Statistical Sciences, 6. Wiley, New York, 581–591.