ISSN 2581-3463

AJMS
Asian Journal of Mathematical Sciences

RESEARCH ARTICLE

Evaluation of the Weather-Influenza Pattern with a Regression Model Approximation to Causality

Lam Kai Shun Carson

Department of Mathematics, University of Hong Kong, Pokfulam, Hong Kong

Received: 28-07-2021 Revised: 13-08-2021 Accepted: 20-08-2021

ABSTRACT

Once upon a time, this author starts research in the relations between Australia's number of cases in influenza and weather. The outcome has been hypothesized with a structural equation model. In this article, the author tries to evaluate the model. It is true that one can apply the evaluation to both of the formative measurement model and structural model through certain suitable procedures. At the same time, this author approximates the model by the linear regression method. The result is one can apply the regression to the Hayes' Process model and find out the wanted model with mediation and moderation effects. In addition, one can also use the Granger Causality Test to examine all of the hypothesized causal relationships between those independent variables such as temperature, wettest_1, the concentration of carbon dioxide, strongest wind, and coolest and the number case of influenza infected. The final outcome is that Hayes' model 91 is the best mediated one with carbon dioxide as the moderated factor. This author will also explain in details why we have the above prescribed Hayes' model 91 as the proposed regression model approximation to causality from the SPSS data analysis.

Key words: Weather-influenza, regression, SPSS data analysis

INTRODUCTION

After this author's previous structural model in describing the relationship between the number of case in influenza and weather (Australia), one needs to evaluate it. Through the evaluation, one can modify the model and the most important thing is one can evaluate this author's HKLam Theory. In the following sections, this author will depict a brief review in how one should evaluate the formative measurement model, structural model together with the assessment of the causal relations from regression approximation. It is hoped that all of the above evaluation processes, one can verify the truthiness of my proposed HKLam Theory. While the main results focus on the models selection (Hayes' model analysis) and their corresponding equations. These events explain why this author chooses Hayes' model 91 as the wanted mediation model.

Address for correspondence:
Lam Kai Shun Carson
E-mail: h9361977@connect.hku.hk

THEORETICALLY BACKGROUND

Evaluation of the formative measurement model

First of all, one is required to know what convergent validity is. It is indeed a measure that correlates with other measures within the same construct under different indicators (such as formative vs. reflective) in the case of a formative measurement model. This is known as redundancy analysis (Chin, 1998). Or the formative measured construct acts as an exogenous latent variable that predicts an endogenous latent variable through some reflective indicators. This gives rise to the value 0.8 or higher (with 0.7 as minimum) for the path between Y (formative) and Y (reflective).

Next, when there are high correlation values occurred between formative indicators, this is known as collinearity. If there are more than two involved indicators, then this is referred as multicollinearity. In order to access the level of collinearity, researchers are required to compute the value of tolerance (TOL).[2] This can be done through two steps:

1. Regress the first formative indicator with all same block remaining indicators, calculate the proportion of variance of the first indicator that associated with others (R^2)
2. TOL can then be calculated from the formula using (1-R^2).

Indeed, the measure of collinearity is the variance inflation factor (VIF) which is just the reciprocal of the TOL. In the case of PLS- structural equation model, Hair *et al*. in 2011 told us that there may be a potential linear problem when the TOL value is lower than 0.2 and with a VIF value higher than 5. Finally, one should examine the indicators' outer weight and outer loading. When both of them are not significant, then the indicator is needed to be deleted.

Evaluation of the structural model

First, one needs to check the collinearity of those indicators [Figure 1]. The procedure is the same as those mentioned in the section "Evaluation of the Formative Measurement Model." Second, one is required to evaluate the structural model path coefficients.[3] Actually, when the path coefficients are more closely to +1, this represents a stronger positive relationship with statistically significant. While for the path values closer to zero, it means weaker relationship. For the bootstrap standard error, when an empirical t value is larger than the critical value, then the t-value is significant with a certain among of error probability. This may refer to most researchers' usage of *P*-values to assess significant levels. Practically, a p-value means the probability of getting a t-value when one observes conditionally from the supported null-hypothesis. Third, another value for us to evaluate the model is the coefficient of determination (R^2 value). It is used to measure the predictive power of the model and is just the squared correlation between the actual and predictive values in endogenous construct. Indeed, R^2 ranges from 0 to 1 where higher level of its value implies a more accurate prediction. In scholarly marketing research, 0.75, 0.5, and 0.25 represent substantial, moderate, or weak predictive power (Hair *et al*., 2011; Henseler *et al*., 2009). However, it is dangerous to select a model purely based on the R^2. With the multiple regression, one may apply the adjusted coefficient of determination (R^2_{adj}). It avoids those bias in complex model. Fourthly, one can test how the endogenous constructs may be impacted by the omitted constructs. This is referred to the f^2 — the effect size. Technically, one can calculate the change in R2 from the estimation of the Path Least Square model twice. While f^2 with values 0.02, 0.15, and 0.35 indicates small, medium, and large effect (Cohen, 1988). For those values smaller than 0.02, this means there is no effect of the exogenous latent variable. Fifthly, the Stone-Geisser's Q^2 value is a measure of the model's out-of-sample predictive power I.e., the predictive relevance. In other words, Q^2 tells us how well the path model can predict the originally observed values. Actually, Q^2 value uses the blinding procedure and performs sampling for the omission of every d-th data point. Hence, Q^2 computes those parameters for the remaining data points (Chin, 1998; Henseler *et al*., 2009; Tenenhaus *et al*., 2005). In fact, blindfolding is an iterative model re-estimation. Finally, one may compare the relative impact of predictive relevance through the measure to the q^2 effect size.

Assessment of the causal relationships

Theoretically, this author's suggested HKLam Net-Seizing Theory can be expressed mathematically in the following ways:

1. The Bayesian Probability part: For all conditional probabilities to events, they can be expressed in terms of the corresponding Bayesian trees. While these trees can be expressed in terms of matrices
2. The Linear Mapping part: One can evaluate the linear mapping through the selection of a suitable linear transformation. Through the transformation, one can map it to the proposed casual relationship. This author notes that in order to verify a linear transformation, the transformation should have the following properties: $T: U \rightarrow V$

I. $T(U + V) = T(U) + T(V)$and

II. $T(cU) = cT(U)$

Abstractly, for a collection of all linear maps, $T: V \rightarrow V$, denoted by End(V) is a (non-commutative) ring, where addition is a point-wise addition:

$$(T_1 + T_2)(V) \mapsto T_1(V) + T_2(V)$$

And the respective multiplication is the composition of:

$$(T_1 . T_2) \mapsto T_1(T_2(V))$$

Now choose some particular $T \in END(V)$. By the universal property of the polynomial ring k[x], one can define a ring homomorphism $k[x] \rightarrow END(V)$ by simply declaring that

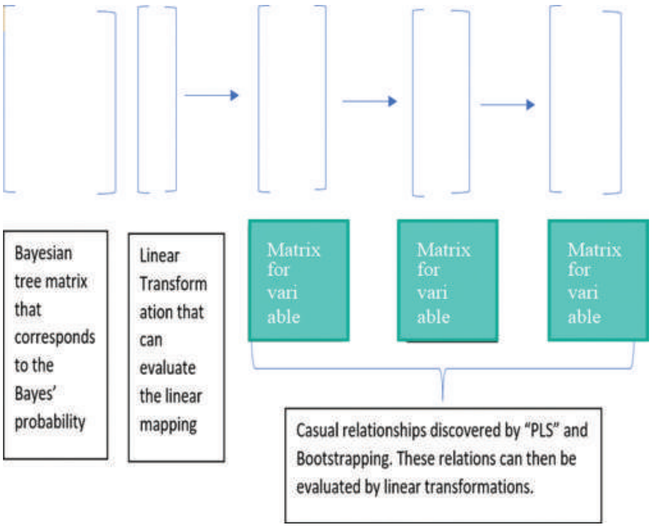
x should go to T. The result is the evaluation homomorphism

$$e_v T: k[x] \rightarrow END(V),\ e_v T(p) = p(T)$$

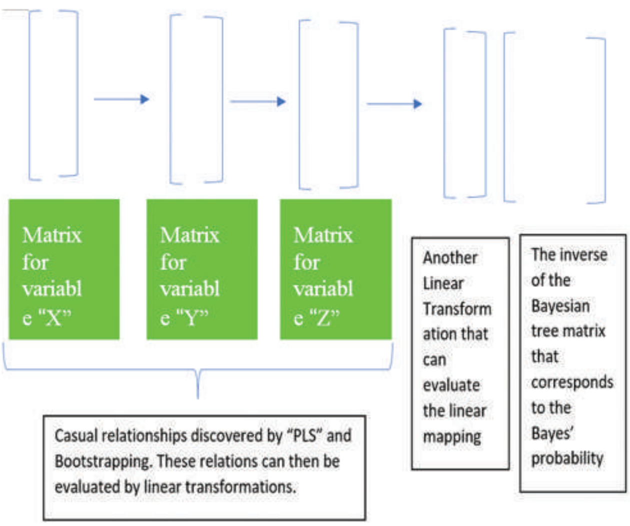
To be precise, one may have the following:
If V is a k-vector space, and $T \in End(V)$ is some linear map $T: V \rightarrow V$,
then the polynomial:

$$p = a_0 + a_1 x + \ldots + a_n x^n \in k[x]$$

Evaluated at T is just: $a_0 + a_1 T + \ldots + a_n T^n \in End(V)$ where T^k is T composed with itself k times, and (aT) is the element defined by $(aT)(v) = a*(TV)$



Actually, partial least square is only used to find the relations between two matrices (X and Y). That is PLS is a latent variable approach which one can model the covariance structures of two spaces. In practice, a collection of data for different variables (like this author's data file — Australia_weather_influenza.xlsx) can be expressed in form of a matrix. Then, one may perform Bootstrapping in statistics with the use of software SmartPLS. The purpose is to find out those causal relationship together with the expression of these relations in form of the required matrices.



This author notes that for the domino effects (or the later part of my proposed philosophy), it suffices to find out those causal relations. Theoretically, suppose there are variable matrices X, Y, and Z, they can be expressed as follows:

$$X = TP^T + E$$
$$Y = UQ^T + F|$$
$$Z = VR^T + G$$

Where X is an n X m matrix of predictors, Y is an n X p matrix of the corresponding responses (to X). T and U are n X l matrices that are, respectively, projections of X and projection of Y. P and Q are respectively, m X l and p X l orthogonal loading matrices; and matrices E and F are the error terms. Similarly, one may apply the same decomposition method to Y and Z. The aim is to maximize the covariance between T and U together with V.
Next, one may try to estimate the factor and loading matrices T, U, V and P, Q, R. One may then construct the linear regression between X and Y, Y and Z as

$$Y = XB + B_0 \text{ and } Z = YD + D_0.$$

The above way is known as partial least squares method for column vector Y and Z or matrices Y and Z. Actually, for a series of domino effect, one will have:

$$Y = XB + B0$$
$$Z = YD + D_0 \text{ or } Z = \{XB + B_0\}\ D + D_0$$

This means one can always express the series of domino effect in a sequence of recursive approximated manner or a partial least square regression.
To be precise, the Bayesian Matrix, say [M], can be expressed by the regression as:

$$[M]\ [LT] = X + (XB + B_0) + \{(XB + B_0)D + D_0\} \qquad \text{———— Eqt} \quad (1)$$

where [LT] is the associated linear transformation; while the converse is also true:

$$X + (XB + B_0) + \{(XB + B_0)D + D_0\} = [M]\ [LT]$$

Hence, from the above mathematical expression, the causal relationships that found from my proposed Net-Seizing Theory can be assessed by Baron and Kenny regression method (1986) — Testing for Mediation. The steps are listed as below:
Step I: Conduct a simple regression analysis with X predicting Y in order to test for path "c" alone. Or, one may have: "path "c"

$$Y = B_0 + B_1 X + e_0 \qquad X \rightarrow Y$$

Step II: Conduct a simple regression analysis with X predicting M to test for path "a". Or one may have:

$$M = B_2 + B_3X + e_1 \qquad X \xrightarrow{\text{path``a''}} M$$

Step III: Conduct a simple regression analysis with M Predicting Y to test the significance of path “b” alone. Or one may have:

$$Y = B_4 + B_5M + e_2 \qquad M \xrightarrow{\text{path``b''}} Y$$

Step IV: Conduct a multiple regression analysis with X and M Predicting Y. Or one may have:

$$Y = B_6 + B_7X + B_8M + e_3$$

$$\frac{X \xrightarrow{\text{path ``c'}}}{M \xrightarrow{\text{Path ``b''}} Y}$$

In step I to III, when one or more of these relationships are non-significant, then one may usually conclude that mediation is not possible or likely with exceptions from MacKinnon *et al*., 2007. If one assumes there are significant relationships from step I to step III, one can proceed to step IV. There is some form of mediations when path b remains significant after controlling for X. If M is controlled and X is no longer significant, the finding gives full mediation. When both X and M significantly predict Y, the finding provides partial mediation.

When we go a further step, compare the above Eqt (1) with the the Eqt in step IV, we get:

$$[M][LT] = (X + XB + XBD) + (B_0 + B_0D) + D_0$$
$$= (I + B + BD)\,X + B_0D + B_0D_0 \text{ ——————(2)}$$

Obviously, from the regression equation in Step IV, the matrix (or vector) D in the term “B_0D”of equation (2) can be viewed as a wanted mediator. In other words, rather than the regression equation for prediction, we may thus construct a hypotheses Hayes model with a feasible moderator.

Indeed, there are defects of Baron and Kenny’s mediation test. First, the test never assesses the significance of the indirect path or how X affects Y through the paths “a” and “b.” Second, the test suffers much from the Type II error of some true mediation effects (Mackinnon *et al*., 2007). Alternatively, there are two other approaches to test the mediation:

1. Judd and Kenny Difference of CoefficientsApproach;
2. Sobel Product of CoefficientApproach.

Although there are drawbacks, this author will still employ Baron and Kenny’s regression method as a way to evaluate the finding causal relationships in chapter fifteen.

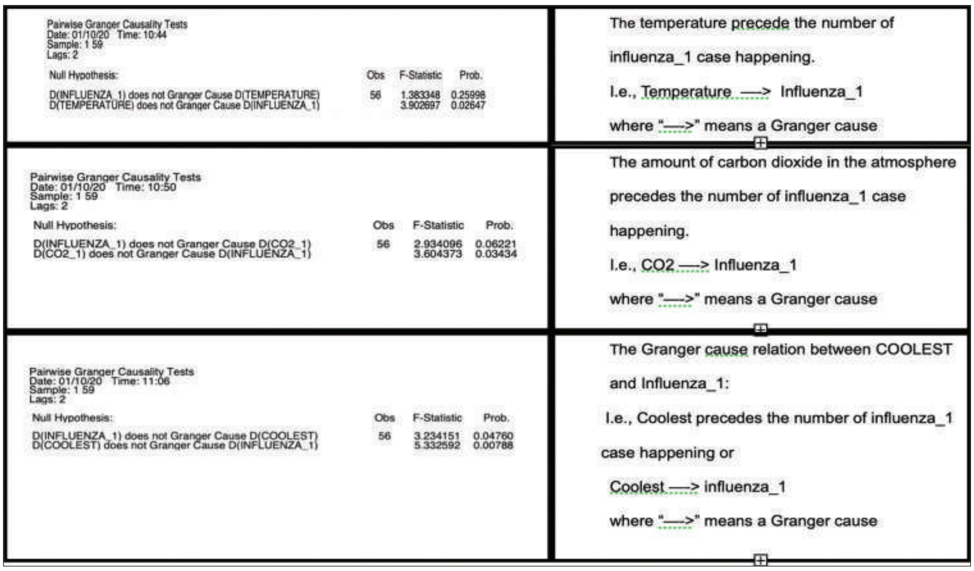
In brief, one can evaluate the Bayesian tree matrix by linking it to the casual relationship matrices through the linear transformation in forms of a polynomial matrices.

Major results — a regression model approximation to causality

Granger test for causality

In order to test the proposed causal relationships, one may need to perform the Granger Causality Test. This author employs the software E Views for such Granger examination. One can find out the corresponding related diagram and hence the linked causal relation diagrams as shown in the next page. There are several possible Ganger causal paths, three of them are [Table 1]:

1. Coolest —-> Temperature —-> Strongest Wind —-> Influenza_1
2. CO_2 <—-> Temperature —-> Wettest_1 —-> Influenza_1
3. CO_2 <—-> Temperature —-> Strongest Wind —-> Influenza_1.



Pairwise Granger Causality Tests
Date: 01/10/20 Time: 10:44
Sample: 1 59
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D(INFLUENZA_1) does not Granger Cause D(TEMPERATURE)	56	1.383348	0.25998
D(TEMPERATURE) does not Granger Cause D(INFLUENZA_1)		3.902697	0.02647

The temperature precede the number of influenza_1 case happening.
I.e., Temperature —> Influenza_1
where “—>” means a Granger cause

Pairwise Granger Causality Tests
Date: 01/10/20 Time: 10:50
Sample: 1 59
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D(INFLUENZA_1) does not Granger Cause D(CO2_1)	56	2.934096	0.06221
D(CO2_1) does not Granger Cause D(INFLUENZA_1)		3.604373	0.03434

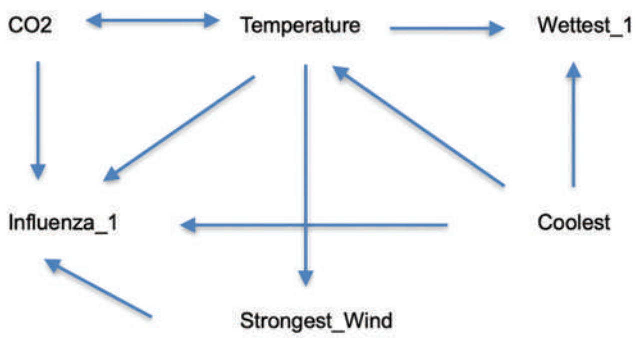
The amount of carbon dioxide in the atmosphere precedes the number of influenza_1 case happening.
I.e., CO2 —> Influenza_1
where “—>” means a Granger cause

Pairwise Granger Causality Tests
Date: 01/10/20 Time: 11:06
Sample: 1 59
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
D(INFLUENZA_1) does not Granger Cause D(COOLEST)	56	3.234151	0.04760
D(COOLEST) does not Granger Cause D(INFLUENZA_1)		5.332592	0.00788

The Granger cause relation between COOLEST and Influenza_1:
I.e., Coolest precedes the number of influenza_1 case happening or
Coolest —> influenza_1
where “—>” means a Granger cause

Pairwise Granger Causality Tests Date: 01/10/20 Time: 12:57 Sample: 1 59 Lags: 5 Null Hypothesis: / Obs / F-Statistic / Prob. D(INFLUENZA_1) does not Granger Cause D(STRONGEST_WIND) / 53 / 1.074871 / 0.38803 D(STRONGEST_WIND) does not Granger Cause D(INFLUENZA_1) / / 4.900275 / 0.00127	the relation between Strongest_wind and Influenza_1 I.e., Strongest_wind precedes the Influenza_1 or Strongest_wind —> Influenza_1 where "—>" means a Ganger cause
Pairwise Granger Causality Tests Date: 01/10/20 Time: 11:20 Sample: 1 59 Lags: 2 Null Hypothesis: / Obs / F-Statistic / Prob. D(CO2_1) does not Granger Cause D(TEMPERATURE) / 56 / 16.6114 / 3.E-06 D(TEMPERATURE) does not Granger Cause D(CO2_1) / / 9.44135 / 0.0003	The symmetric Granger cause relations between carbon dioxide and temperature: I.e., CO2 precedes the temperature happening Temperature precedes the CO2 happening CO2 —> Temperature Temperature —> CO2 where "—>" means a Granger cause
Pairwise Granger Causality Tests Date: 01/10/20 Time: 11:18 Sample: 1 59 Lags: 2 Null Hypothesis: / Obs / F-Statistic / Prob. D(WETTEST_1) does not Granger Cause D(TEMPERATURE) / 56 / 0.358226 / 0.70065 D(TEMPERATURE) does not Granger Cause D(WETTEST_1) / / 4.303410 / 0.01874	The relation between temperature and wettest_1, I.e., Temperature precedes the wettest_1 case happening or Temperature —> wettest_1 where "—>" means a Granger cause
Pairwise Granger Causality Tests Date: 01/10/20 Time: 11:43 Sample: 1 59 Lags: 2 Null Hypothesis: / Obs / F-Statistic / Prob. D(WETTEST_1) does not Granger Cause D(COOLEST) / 56 / 0.503310 / 0.60749 D(COOLEST) does not Granger Cause D(WETTEST_1) / / 4.093170 / 0.02245	The relation between coolest and wettest_1 I.e., Coolest precedes the wettest_1 case happening or Coolest —> wettest_1 where "—>" means a Granger cause



Indeed, one can first convert the data set into time series by statistical programming software R first. Then, one can apply the maximum likelihood estimation for the fitting of the above Granger Causal relation to build the corresponding models for the prediction. As the process is similar to the following mediation analysis by software SPSS add-on Hayes PROCESS, this author decides to omit the approximation.

Mediation analysis for the causal relationships found by smartPLS (Lam, 2019)

Formally speaking, in order to test a causal relationship, one should perform causal analysis. Indeed, causal analysis = regression analysis (or any test) + theory (and hypothesis) or logical analysis. Usually, in statistics and economics, causality is often tested by regression (i.e. the present influenza relationship research case). Actually, one should perform exploratory causal analysis which is known as data causality or causal discovery. It is the use of statistical algorithms to infer associations in observed data set. In addition, these data sets are potentially causal under strict assumptions.

Therefore, according to the results of chapter fifteen, there are potentially several cases of causal relationships. According to the aforementioned Baron and Kenny's Testing of mediation method, one obtains the following partial least square regression results (by R programming) [Table 2]:

To sum up, all of the four steps of Baron and Kenny's method are fulfilled. In other words, we have:

- Step I: Wettest → Influenza cases
- Step II: Wettest → Temperature
- Step III: Temperature → Influenza cases
- Step IV: Wettest + Temperature → Influenza cases

Hence, the proposed causal relationship — wettest and influenza is actually a fully mediated one with temperature as the immediate mediator. This event indirectly implies that wettest and influenza is actually a causal relationship.

The second one is the evaluation of the causal relationship "wind and influenza": Wind →

Table 1: Detecting the Granger causality between different weather variables and the influenza in-flected.

X NOT Granger Cause Y	Probability	F-Statistics	Result (T/F)
Influenza_1, temperature	0.25998	1.383348	1
Temperature, influenza_1	0.06221	2.934096	FALSE Temperature granger causes Influenza_1
influenza_1, coolest	0.04790	3.234151	1
Coolest, influenza_1	0.00788	5.332592	FALSE Coolest granger causes influenza_1
Influenza_1, Strongest_Wind	0.38803	1.074871	1
Strongest_Wind, Influenza_1	0.00127	4.900275	FALSE Strongest Wind granger causes influenza_1
CO2_1, Temperature	3E-06	16.6114	FALSE CO2_1 granger causes Temperature
Temperature, CO2_1	0.0003	9.44135	FALSE Temperature granger causes CO2_1
Wettest_1, Temperature	0.70065	0.358226	1
Temperature, Wettest_1	0.011874	4.303410	FALSE Temperature granger causes Wettest_1
Wettest_1, Coolest	0.60749	0.503310	1
Coolest, Wettest_1	0.02245	4.093170	FALSE Coolest granger causes Wettest_1

1. CO2_1 and Temperature constitutes a symmetric granger cause relationship.
2. The above table's data values are obtained from running E-View software where the raw data is downloaded from Australian government web-site.)

Table 2: Using Baron & Kenny Steps to identify the mediation relationships demonstrated by using the JASP & R coding.

Baron & Kenny Step	JASP&R-Code	Coefficients	P-Value	Error Term	Model Equations	Reject the Null Hypothesis
I	Influenza~wettest, data=australia_ weather_influenza	Intercept=3.7699516 Wettest = -0.0001773	0.0884	0.4926	influenza=3.7699516-0.0001773* Wettest	Influenza infected is unaffected by Wettest.
II	Temperature~wettest, data=australia_ weather_influenza	Intercept=1.631e+0 Wettest=1.705e-03	0.01884	3.398	Temperature=1.631e+01+1.705e-03 *Wettest	Temperature is unaffected by Wettest.
III	Influenza~temperature, data=australia_ weather_influenza	intercept=5.21750 temperature = -0.08998	6.532E-08	0.3904	Influenza=5.21750-0.08998* temperature	Influenza inflected is unaffected by temperature.
IV	Influenza ~ wettest + temperature, data =	Intercept = 5.214 Wettest = -2.627e-05 Temperature = -8.855e-02	4.988E-07	0.3935	Influenza = 5.214-2.627e-05 * Wettest + (-8.855e-02) * temperature	Influenza is unaffected by the combined effects of Wettest and Temperature

Wettest →Temperature →Influenza.
The following list are the partial least square result (Hayes Process Model Macro) that obtained through software SPSS [Tables 3-5]:
Wettest and Wind

Temperature and (Wind Together with Wettest)

We observe that the model equation for Temperature from (Wind and Wettest) is:
Temperature = 14.7740 + 0.0107*Wind + 0.0016*Wettest
Influenza and (Wind and Wettest and Temperature)
We observe that the model equation for Influenza from (Wind, Wettest and Temperature) is:
Influenza = 43685.53 + 26.71*Wind - 2.185*Wettest- 2054.05 * Temperature
Once we observe that the model equation for Wind and Wettest is: Wettest = -322.2481 + 7.5870*Wind
Temperature = 14.7740 + 0.0107*Wind + 0.0016*Wettest
Influenza = 43685.53 + 26.71*Wind– 2.185*Wettest- 2054.05 * Temperature
All of the above data shows that there should be a Hayes PROCESS model 6 established as like the following:[4]

Table 3: Evaluation of the causal relationship between Wettest and Wind by the software SPSS. The model equation is: Wettest = (-322.2481) + 7.5870*Wind

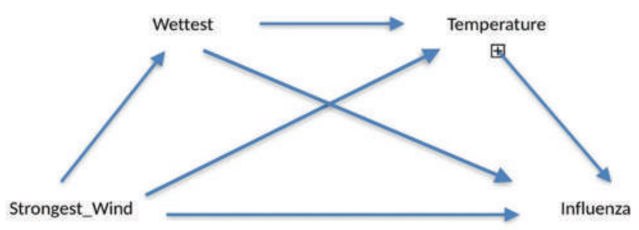
A. Wettest and Wind			
Outcome Variable: Wettest			
Model	**Coeff**	**P-Value**	**Standardised Coefficients**
Constant	-322.2481	0.5471	Nil
Strongest Wind	7.5870	0.0320	0.2795
Covariance matrix of regression parameter estimates:			
	Constant	**Strongest Wind**	
Constant	283057.659	-1815.9068	
Strongest Wind	-1815.9068	11.9175	

Table 4: Evaluation of the causal relationship between temperature and (Wind & Wettest) by the software SPSS. The model equation is: Temperature=14.7740+0.0107*Wind+0.0016*Wettest

B. Temperature and (Wind together with Wettest)			
Outcome Variable: Wettest			
Model	**Coeff**	**P-Value**	**Standardised Coefficients**
Constant	14.7740	0.0000	Nil
Strongest Wind	0.0107	0.0201	0.0703
Wettest	0.0016	0.0352	0.2853
Covariance matrix of regression parameter estimates:			
	Constant	**Strongest Wind**	**Wettest**
Constant	8.8727	-0.0579	0.0002
Strongest Wind	-0.0579	0.0004	0.000
Wettest	0.0002	0.000	0.000

Table 5: Evaluation of the causal relationship between cases of influenza infected and (Wind & Wettest & Temperature) by the software SPSS.
The model equation is: Influenza=43685.53+26.71*Wind - 2.185*Wettest - 2054.05*Temperature.

C. Influenza infected and (Wind and Wettest and Temperature)				
Outcome Variable: Wettest				
Model	**Coeff**		**P-Value**	**Standardised Coefficients**
Constant	43685.5285		0.0107	Nil
Strongest Wind	26.7064		0.7751	0.0361
Wettest	-2.1845		0.5417	-0.0803
Temperature	-2054.0522		0.0016	-0.4219
Covariance matrix of regression parameter estimates:				
	Constant	**Strongest Wind**	**Wettest**	**Temperature**
Constant	273111110	-1177928.9	12766.3258	-5642138.7
Outcome Variable: Wettest				
Strongest Wind	-1177928.9	8654.2225	-82.1602	-4074.2063
Wettest	12766.3258	-82.1602	12.6576	-609.2217
Temperature	-5642138.7	-4074.2063	-609.2217	381897.287



This author notes that although the bootstrap confidence interval straddles zero which means that the mediation is not significant (or actually a border case), it does not imply that thing we are estimating is zero. Thus, this author finally concludes wettest and temperature are the mediators for wettest and influenza.

Furthermore, if we add the concentration of carbon dioxide as the moderator that lays between temperature and influenza, one may obtain the results below:

The aforementioned outcome tells us that the index of moderated mediation is referring to the

Step I

```
> model = lm(influenza ~ Wettest, data = australia_weather_influenza)
> summary(model)

Call:
lm(formula = influenza ~ Wettest, data = australia_weather_influenza)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.6999 -0.3504 -0.1150  0.2660  1.3127 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.7699516  0.1066826  35.338   <2e-16 ***
Wettest     -0.0001773  0.0001022  -1.734   0.0884 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4926 on 57 degrees of freedom
Multiple R-squared:  0.05008,	Adjusted R-squared:  0.03342 
F-statistic: 3.005 on 1 and 57 DF,  p-value: 0.0884
```

From the statistical data, one discovers that the p-value of the relationship is 0.0884 with the model equation equals to:

influenza = 3.7699516 - 0.0001773*Wettest

Hence one will reject the null hypothesis: number of influenza infected is unaffected by Wettest.

Step II

```
> model = lm(temperature ~ Wettest, data = australia_weather_influenza)
> summary(model)

Call:
lm(formula = temperature ~ Wettest, data = australia_weather_influenza)

Residuals:
    Min      1Q  Median      3Q     Max 
-7.5544 -2.7871  0.1083  3.2289  5.0736 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.631e+01  7.359e-01  22.162   <2e-16 ***
Wettest     1.705e-03  7.053e-04   2.418   0.0188 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.398 on 57 degrees of freedom
Multiple R-squared:  0.09301,	Adjusted R-squared:  0.0771 
F-statistic: 5.845 on 1 and 57 DF,  p-value: 0.01884
```

One observes that the p-value of the relationship is 0.01884 with the model equation equals to:

Temperature = 1.631e+01 + 1.705e-03*Wettest

Hence one will reject the null hypothesis:

Temperature is unaffected by Wettest.

Step III

```
> model = lm(influenza ~ temperature, data = australia_weather_influenza)
> summary(model)

Call:
lm(formula = influenza ~ temperature, data = australia_weather_influenza)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.73352 -0.24386 -0.04267  0.24618  1.14992 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.21750    0.26193  19.919  < 2e-16 ***
temperature -0.08998    0.01449  -6.209 6.53e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3904 on 57 degrees of freedom
Multiple R-squared:  0.4034,	Adjusted R-squared:  0.393 
F-statistic: 38.55 on 1 and 57 DF,  p-value: 6.532e-08
```

One finds that the p-value of the above relationship is 6.532e-08 with the model equation equals to:

Influenza = 5.21750 - 0.08998*Temperature

Hence one will reject the null hypothesis:

Influenza is unaffected by Temperature.

Step IV:

```
> model = lm(influenza ~ Wettest + temperature, data = australia_weather_influenza)
> summary(model)

Call:
lm(formula = influenza ~ Wettest + temperature, data = australia_weather_influenza)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.7376 -0.2486 -0.0272  0.2711  1.1425 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.214e+00  2.643e-01  19.729  < 2e-16 ***
Wettest     -2.627e-05  8.577e-05  -0.306    0.761    
temperature -8.855e-02  1.534e-02  -5.772 3.56e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3935 on 56 degrees of freedom
Multiple R-squared:  0.4044,	Adjusted R-squared:  0.3832 
F-statistic: 19.01 on 2 and 56 DF,  p-value: 4.988e-07
```

One observes that the p-value is 4.988e-07 with the model equation equals to:

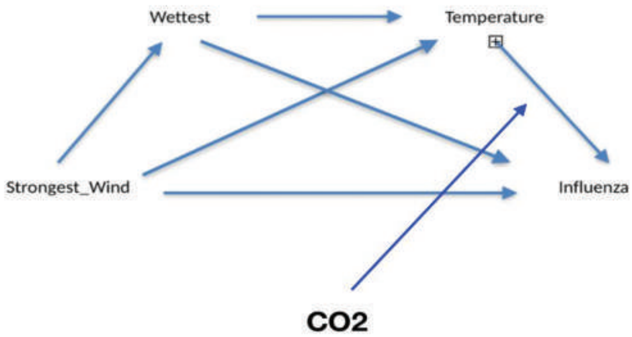
influenza = 5.214e+00 - 2.627e-05*Wettest - 8.855e-02*Temperature

Hence, one will reject the null hypothesis: influenza is unaffected by the combined effects of Wettest and Temperature.

"weight for the moderator in a linear function relating to the size of indirect effect of X on Y to the moderator" (Hayes, 2018, p.491).

When the index is not zero, this will mean the indirect effect relates linearly with the moderator. Hence, one can claim there is a moderated mediation (Hayes, 2018). On the contrary, if the bootstrap confidence interval does not includes zero, this event implies the indirect effect will not relate linearly with the moderator (Hayes, 2018) [Table 6]. Therefore, we conclude that carbon dioxide is the moderator to temperature and influenza. This event is because the above data shows that both the index is greater than zero and the confidence interval contains a zero.

Then, one may obtain the Hayes PROCESS model 87 as shown below:



In addition, if one changes the independent variable from wettest_1 to temperature (while wettest_1

```
OUTCOME VARIABLE:
 Wettest

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .2795      .0781 375399.450     4.8301     1.0000    57.0000      .0320

Model
              coeff         se          t          p       LLCI       ULCI
constant  -322.2481   532.0316     -.6057      .5471 -1387.6297   743.1335
Stronges     7.5870     3.4522     2.1977      .0320      .6741    14.4999

Standardized coefficients
              coeff
Stronges      .2795

Covariance matrix of regression parameter estimates:
           constant   Stronges
constant 283057.659 -1815.9068
Stronges -1815.9068    11.9175

*************************************************************************
```

We observe that the model equation for Wind and Wettest is:

Wettest = -322.2481 + 7.5870*Wind

```
OUTCOME VARIABLE:
 temperat

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .3124      .0976    11.6920     3.0272     2.0000    56.0000      .0564

Model
              coeff         se          t          p       LLCI       ULCI
constant    14.7740     2.9787     4.9599      .0000     8.8069    20.7411
Stronges      .0107      .0201      .5317      .5971     -.0295      .0509
Wettest       .0016      .0007     2.1581      .0352      .0001      .0031

Standardized coefficients
              coeff
Stronges      .0703
Wettest       .2853

Covariance matrix of regression parameter estimates:
           constant   Stronges    Wettest
constant     8.8727     -.0579      .0002
Stronges     -.0579      .0004      .0000
Wettest       .0002      .0000      .0000

*************************************************************************
```

We observe that the model equation for Temperature from (Wind and Wettest) is:

Temperature = 14.7740 + 0.0107*Wind + 0.0016*Wettest

```
OUTCOME VARIABLE:
 influenz

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4475      .2002  250048440     4.5897     3.0000    55.0000      .0061

Model
              coeff         se          t          p       LLCI       ULCI
constant 43685.5285 16526.0736     2.6434      .0107 10566.3151 76804.7418
Stronges    26.7064    93.0281      .2871      .7751  -159.7273   213.1401
Wettest     -2.1845     3.5578     -.6140      .5417    -9.3144     4.9455
temperat -2054.0522   617.9784    -3.3238      .0016 -3292.5168  -815.5875

Standardized coefficients
              coeff
Stronges      .0361
Wettest      -.0803
temperat     -.4219

Covariance matrix of regression parameter estimates:
           constant   Stronges    Wettest   temperat
constant  273111110 -1177928.9 12766.3258 -5642138.7
Stronges -1177928.9  8654.2225   -82.1602 -4074.2063
Wettest  12766.3258   -82.1602    12.6576  -609.2217
temperat -5642138.7 -4074.2063  -609.2217 381897.287

************************** TOTAL EFFECT MODEL ****************************
```

We observe that the model equation for Influenza from (Wind, Wettest and Temperature) is:

Influenza = 43685.53 + 26.71*Wind - 2.185*Wettest- 2054.05 * Temperature

and strongest_wind become the mediators), we will get the following results [Table 7]:

```
****************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
 -1898.6041   646.8798    -2.9350      .0049 -3194.9887  -602.2194

Conditional and unconditional indirect effects of X on Y:

INDIRECT EFFECT:
 temperat    ->    Wettest_    ->    influenz

     Effect     BootSE   BootLLCI   BootULCI
  -290.2719   220.5588  -776.4535    87.7065

INDIRECT EFFECT:
 temperat    ->    Stronges    ->    influenz

     Effect     BootSE   BootLLCI   BootULCI
     -.0614    63.2151  -111.2673   155.1546

INDIRECT EFFECT:
 temperat    ->    Wettest_    ->    Stronges    ->    influenz

      CO2_1     Effect     BootSE   BootLLCI   BootULCI
   398.9220    14.1275    44.2673   -61.9476   116.5677
   403.6400    36.6441    62.5905   -42.0169   194.8251
   408.9040    61.7665   109.6117   -58.7893   370.4109

       Index of moderated mediation:
           Index     BootSE   BootLLCI   BootULCI
CO2_1     4.7725    10.6077    -6.2494    35.7283
---
```

It is obviously observed from the above data that the index of moderated mediation has increased greatly from 2 to nearly 5 if we made the amendment of the change to temperature as the independent variables. Then the concentration of carbon dioxide becomes the moderator between wettest_1 and strongest_wind. This means that the indirect effect depends heavily on the moderator carbon dioxide. According to Hayes in 2018, the index of moderated mediation is slope of the equation formed by indirect effect. If it is equal to zero (i.e. flat) then the indirect effect is not related to the moderator. However, if the index is strongly greater than zero, the indirect effect depends heavily on the moderator. Thus, carbon dioxide acts as the wanted moderator. Therefore,

Table 6: Carbon Dioxide acts as the moderator for temperature and influenza (Direct & Indirect Effects Data Analysis) by the software SPSS.

Direct And Indirect Effects of X and Y

Effect	Se	t	p	LLCI	ULCI
8.6924	94.3905	0.0921	0.927	-180.6324	198.0172
Conditional	And	Unconditional	Indirect Effects	Of X on Y	

Indirect Effect: Strongest Wind —>Wettest —>Influenza

Effect	BootSE	BootLLCI	BootULCI
-9.8542	19.5547	-62.4032	13.8536

Indirect Effects: Strongest Wind —>Temperature —>Influenza

	CO2_1	Effect	BootSE	BootLLCI	BootULCI
	398.9220	-29.5673	64.3348	-127.4806	136.0163
	403.6400	-21.9630	49.1415	-100.2869	102.1029
	408.9040	-13.4787	37.3614	-83.1024	73.7515

Index of moderated Mediation

	Index	BootSE	BootLLCI	BootULCI
CO2_1	1.6118	4.1492	-9.1840	8.621

Indirect Effect: Strongest Wind —> Wettest —> temperature —> influenza

	CO2_1	Effect	BootSE	BootLLCI	BootULCI
	398.9220	-33.5439	31.4431	-124.8538	-3.0546
	403.6400	-24.9169	24.9493	-97.3552	-2.1555
	408.9040	-15.2915	23.3582	-83.9560	3.3906

Index of moderated Mediation

	Index	BootSE	BootLLCI	BootULCI
CO2_1	1.8285	2.5126	-1.3135	8.5707

Table 7: Carbon dioxide acts as the moderator between wettest_1 and strongest_wind (Direct & Indirect Effects Data Analysis) by the software SPSS.

Direct And Indirect Effects of X and Y

Effect	Se	t	p	LLCI	ULCI
-1898.6041	646.8798	-2.9350	0.0049	-3194.9887	-602.2194
Conditional	And	Unconditional	Indirect Effects	Of X on Y	

Indirect Effect: Temperature —> Wettest_1 —> Influenza

Effect	BootSE	BootLLCI	BootULCI
-290.2719	220.5588	-776.4535	87.7065

Indirect Effect: Temperature —>Strongest_Wind —>Influenza

Effect	BootSE	BootLLCI	BootULCI
-0.614	63.2151	-111.2673	155.1546

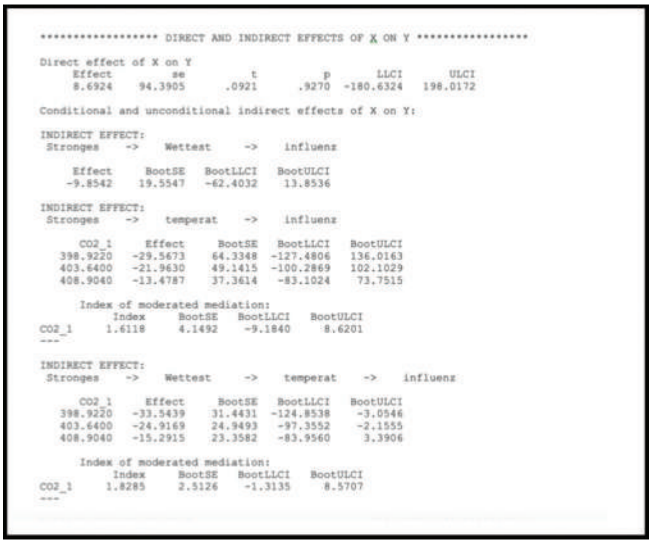
Indirect Effects: Temperature —> Wettest_1 —> Strongest_Wind —> influenza

	CO2_1	Effect	BootSE	BootLLCI	BootULCI
	398.9220	14.1275	44.2673	-61.9476	116.5677
	403.6400	36.6441	62.5905	-42.0169	194.8251
	408.9040	61.7665	109.6117	-58.7893	370.4109

Index of moderated Mediation

	Index	BootSE	BootLLCI	BootULCI
CO2_1	4.7725	10.6077	-6.2494	35.7283

we conclude that the Hayes PROCESS model 91 is the most suitable one for describing the relationship between different weather variables and the number of case in influenza in Australia when compared to the previous one. The final and conclusive model is shown in below:



```
****************** DIRECT AND INDIRECT EFFECTS OF X ON Y *****************

Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
     8.6924    94.3905      .0921      .9270  -180.6324   198.0172

Conditional and unconditional indirect effects of X on Y:

INDIRECT EFFECT:
 Stronges    ->    Wettest     ->    influenz

     Effect     BootSE   BootLLCI   BootULCI
    -9.8542    19.5547   -62.4032    13.8536

INDIRECT EFFECT:
 Stronges    ->    temperat    ->    influenz

      CO2_1     Effect     BootSE   BootLLCI   BootULCI
   398.9220   -29.5673    64.3348  -127.4806   136.0163
   403.6400   -21.9630    49.1415  -100.2869   102.1029
   408.9040   -13.4787    37.3614   -83.1024    73.7515

      Index of moderated mediation:
           Index     BootSE   BootLLCI   BootULCI
CO2_1     1.6118     4.1492    -9.1840     8.6201
---

INDIRECT EFFECT:
 Stronges    ->    Wettest     ->    temperat    ->    influenz

      CO2_1     Effect     BootSE   BootLLCI   BootULCI
   398.9220   -33.5439    31.4431  -124.8538    -3.0546
   403.6400   -24.9169    24.9493   -97.3552    -2.1555
   408.9040   -15.2915    23.3582   -83.9560     3.3906

      Index of moderated mediation:
           Index     BootSE   BootLLCI   BootULCI
CO2_1     1.8285     2.5126    -1.3135     8.5707
---
```

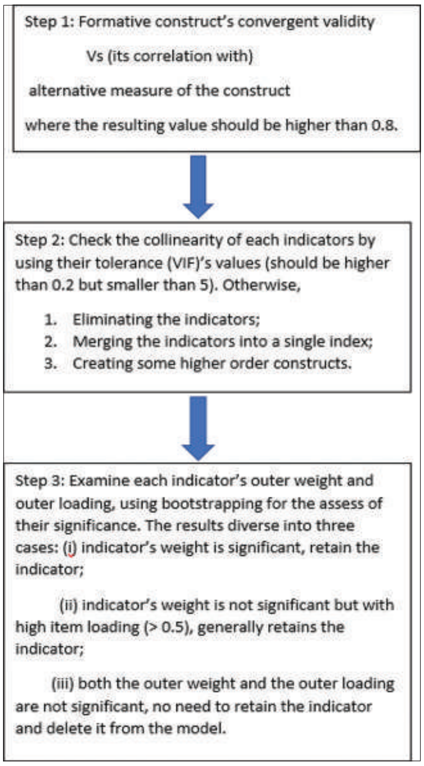
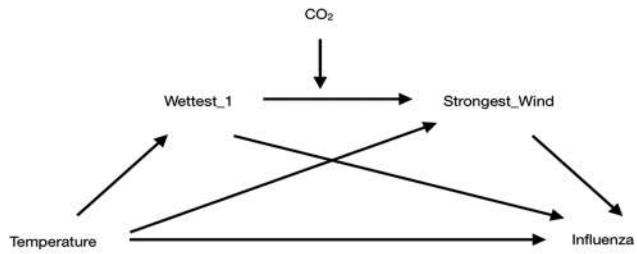


Figure 1: A summary of procedures to evaluate the formative measurement indicators



CONCLUSION

To conclude, the above test and model tell us that wettest_1 and the temperature are the most feasible causal relationship to the number case of influenza. The results are obtained from the mediation analysis (Baron and Kenny's Testing and the Haye's PROCESS modelling for SPSS). Similarly, one can show that the sequenced domino effects is indeed a list of causal relations. Practically, one should find out all of the possible regression models (in this case, one should employ the carbon dioxide and strongest_wind as the independent variables respectively for the dependent variables number of case in influenza, other variables such as temperature and Wettest_1 will be used as mediator together with suitable moderators). Then, one should compare the r-square, r-square(adjusted), and r-square(predicted) for each of these calculated models. The aim is to find out the best goodness fit in these models from these r values and hence selects the best fitted model for the wanted causal relations. While at the same time, the concentration of carbon dioxide acts as the moderator that lays between Wettest_1 and strongest_wind provide that temperature as the independent variable (model 91). On the other hand, the concentration of carbon dioxide also acts as the moderator

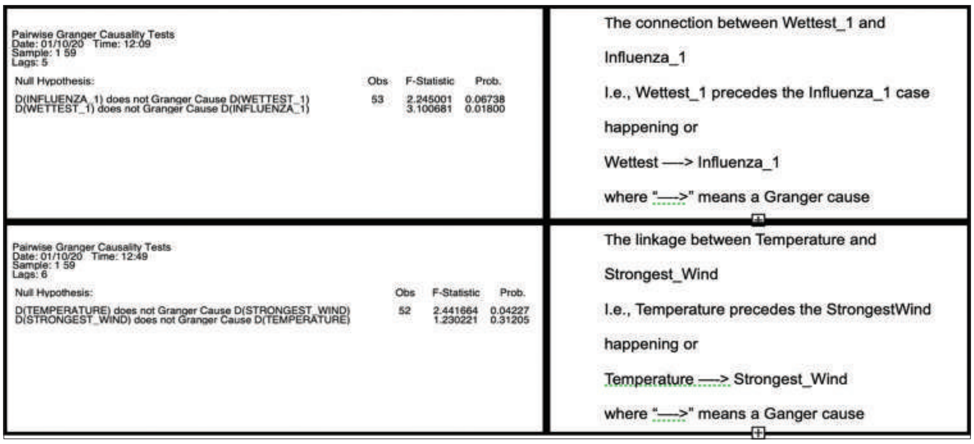


Figure 2: Granger Causal relationships obtained from the above nine corresponding linkings

between the mediator temperature and the dependent variable number of case in influenza (model 87). The most significant discovery in this paper is that the role of carbon dioxide is indeed a moderator. It gives a moderated effect to other mediators. This author remarks that a moderator is difference from a mediator in that moderator only affects the strength of the concerning variables. While the mediator can explain the relationship between two variables. The existence of CO2 as a moderator implies that the gas has a conditional effect to the number of cases of influenza infected. Hence, there may be a thermal degradation from CO2 to CO at around 20^{o}C (Asperen *et al*., 2015).[1] However, the symptoms of carbon monoxide poisoning are similar to the common flu infected. In addition, the dissolved carbon dioxide can have an influence during the production of recombinant hemagglutinin component that induced from an influenza vaccine by insect cells (Meghrous *et al*., 2015).[5] These are the reasons for the conditional effects of CO2 in the number of case of influenza infected. Therefore, I suggest there should be a reduction in the emission of pollutants such as CO2 before and during the peak months of common flu infection.

All in all, the above Hayes Process model 91 can help us verify the truthiness of this author's HKLam Net-Seizing Theory. In other words, all of the causal relations can be find by a mediation analysis. In addition, all of the causal relationships can be examined by Granger Causality Test followed by Maximum Likelihood estimation for fitting into the models. Hence the second part of the philosophy — domino effects can be expressed as a causal relationship is proposed to be correct.

REFERENCES

1. Asperen HV, Warneke T, Sabbatini S, Nicolini G, Papale D, Notholt J. The role of photo and thermal degradation for CO_2 and CO fluxes in an arid ecosystem. Biogeosciences 2015;12:4161-74.
2. Joseph HF, Sarstedt M, Ringle CM, Gudergan SP. Advanced Issues in Partial Least Squares Structural Equation Modeling. Thousand Oaks, California: SAGE Publication; 2018.
3. Joseph HF Jr., Hult GT, Ringle CM, Sarstedt M. A primer on partial lease squares structural equation modeling. Thousand Oaks, California: SAGE Publication; 2017.
4. Hayes FA. Introduction to Mediation, Moderation, and Conditional Process Analysis a Regression-Based Approach. New York: Guilford Press; 2018.
5. Meghrous J, Khramtsov N, Buckland BC, Cox MM, Palomares LA, Srivastava IK. Dissolved Carbon Dioxide Determines the Productivity of a Recombinant Hemagglutinin Component of an Influenza Vaccine Produced by Insect Cells. Hoboken, New Jersey: Wiley Publication; 2015.