The Multivariate Distributed Lag Non-Linear Model (MV_{DLNM}) that Accommodates Cumulative Outcomes

Yi-Hau Chen, Ph.D.*1, Xing-Yi Huang, MS*2, Pei-Cheng Kuo*3, MD, Chao-Yu Guo, Ph.D.2

NOTE: *the three authors contributes equally

1Institute of Statistical Science, Academia Sinica
2Institute of Public Health, National Yang-Ming University
3Department of Medicine, National Yang-Ming University

Corresponding authors:

Chao-Yu Guo, Ph. D.
Division of Biostatistics
Institute of Public Health
National Yang Ming University
Taipei, Taiwan, ROC

Email address:

yhchen@stat.sinica.edu.tw3
Loveu0129@gmail.com1
sam3236577@gmail.com2
cyguo@ym.edu.tw4
ABSTRACT

In recent years, the effects of meteorological factors on health outcomes have gained popularity due to the impact of climate change, which not only results in a general rise in temperature but also the abnormal climatic extremes. Instead of the conventional cross-sectional analyses that focus on the association between the main predictor and one single dependent variable, the distributed lag non-linear model (DLNM) has been widely adopted to examine the association between the multiple lagged environmental factors and the health outcome in epidemiology studies. In this research, we further investigate a more complex association structure between the lagged mortality and the lagged temperature. Five newly proposed strategies, which are derived by various statistical concepts, such as summation, autoregressive, principal component analysis, and adjustment or offset in the DLNM, are evaluated by the permutation study. The longitudinal climate and daily mortality data in Taipei Taiwan from 2012 to 2016 were permuted to simulate the null distribution. According to simulation results, only one strategy, named as $M_{V_{DLNM}}$, could yield valid Type-I errors. With an application to the real data, the $M_{V_{DLNM}}$ that incorporates both the current and lagged mortalities demonstrates a much more significant association comparing to the conventional DLNM that only relies on the current mortality.

Keyword: distributed lag non-linear model, multivariate analysis, temperature and mortality
Introduction

Extensive studies have indicated the association between temperature and human health, which arouses public health concerns as the climate, have changed drastically on a worldwide scale due to global warming in recent years. [1, 2]. After accounting for climate changes and other factors, how hot and cold weather, or their delayed effects, trigger human death were widely discussed in different areas, including the United States [3, 4], Europe [5], and Northeast Asia [6]. In addition to temperature, it has also been documented that exposure to air pollutants, which includes particulate matter (PM), ozone (O₃), nitrogen dioxide (NO₂), and sulfur dioxide (SO₂) according to The 2005 WHO Air Quality Guidelines, leads to adverse effects on human health, especially the respiratory and cardiovascular diseases. A number of researches have examined the relationship between PM₁₀, PM₂.₅ and daily mortality. Some showed that exposure to polluted air in a period would harm health conditions such as the development of lung or heart diseases, where the sources of pollution come from air, second hand smoke, ozone, or particle matters. [7, 8].

In 2010, Gasparrini [2] carried out the Distributed Lag Non-Linear Model (DLNM) to evaluate the lagged effect of predictors. The DLNM fits the non-linear association between the outcome variable and predictors. A cross-basis function depicts simultaneously the exposure-response relationship along the predictor space and lag-response relationship along the lag space. In 2018, a new approach assessed both the same day and one day lagged mortality in DLNM [9]. Therefore, associations in both lagged outcomes and exposures needs more attention in order to describe such complex structure.

In this research, we collected both weather and air pollution data as well as daily mortality in Taipei City from 2012 to 2016. Since the DLNM is widely adopted in
public health and environmental research [10], we aim to extend the DLNM with Poisson link function and natural cubic splines [11] to accommodate lagged outcomes. Validity and performance of the new methods would be evaluated by the permutation study. Finally, a real data application shows the improvement by the new method.

**Materials and Methods**

*Data collection*

All-cause mortality in Taipei city were obtained from the Cause of the Death Database published by Ministry of Health and Welfare. Daily mean temperature measures were obtained from Taipei Weather Station, available through the CWB Observation Data Inquiry System website [12]. Data on air pollution were obtained from Taipei Air Quality Monitoring System, available through Environmental Protection Administration Executive Yuan website [13], where we collected daily mean ozone concentration and daily mean PM$_{2.5}$ concentration (Table 1). Though some air pollutions were missing, we could just omitted these observations, since the missing rate are dismal with ignorable impact to the analyses.

*Statistical analysis*

The DLNM model is defined as the following:

$$\log(\mu_t) = \alpha + s(x_t, l, \beta) + \beta_{O_3} O_{3t} + \beta_{PM_{2.5}} PM_{2.5t} + \sum_{i=1}^{p} f(z_t^i ; \theta)$$

The independent variable ($x_t$) is daily mean temperature and other pollutant variables ($O_{3t}$ and $PM_{2.5t}$) are treated as potential confounders. The outcome variable ($\mu_t$) was all-cause mortality. The DLNM model was fitted through a cross basis functions $s(x_t, l, \beta)$ simultaneously describing the effect of the daily mean
temperature $x_t$ and its lag structure with maximum lag $l$ on the expected mortality. Daily mean ozone concentration $O_{3t}$ and daily mean concentration $PM2.5_t$ are fixed effects. A natural cubic spline $f(z_t^i : \theta)$ with 8 degrees of freedom for each year is used to adjust for the seasonal effect. We selected 10, 20, and 30 days for the maximum exposure lag $l$. The cross basis consists of a quadratic B-spline for temperature with the knots placed at 10, 75, and 90 percentile and a natural cubic spline for the lag with 5 degrees of freedom, which indicates 3 internal knots equally spaced in the log scale.

In order to extend the DLNM to accommodate the lagged outcomes, we propose five different approaches to transform the lagged outcomes ($n \times l$) into a 1-dimensional dependent variable ($n \times 1$) to be integrated into the DLNM.

For illustration purpose, assume that the Y matrix consists of four days of mortality with two lagged days. Hence, the dimension of Y is $(4 \times 3)$. The second column of Y is the one-day lagged mortality. The third column of Y represents the two-day lagged mortality.

Let $Y = \begin{bmatrix} 60 & 51 & 62 \\ 73 & 60 & 51 \\ 61 & 73 & 60 \\ 55 & 61 & 73 \end{bmatrix}$, with $eigenvector = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & e \end{bmatrix}$, $eigenvalue = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$ in the principal component analysis (PCA).

Method 1:

$MV_{sum}$: The simplest idea is to obtain the total mortalities from today to previous lagged days. The new Y matrix $(4 \times 1)$ contains the summation of mortalities from the current day to the maximal lag day:

$$MV_{sum} = \begin{bmatrix} 60 + 51 + 62 \\ 73 + 60 + 51 \\ 61 + 73 + 60 \\ 55 + 61 + 73 \end{bmatrix}_{4+1} = \begin{bmatrix} 173 \\ 184 \\ 194 \\ 189 \end{bmatrix}_{4+1}$$
Method 2:

MVAR: A commonly used longitudinal structure is Autoregressive (AR). The lagged mortality could be integrated into the current mortality by this weighted summation. The earlier a day lags, the less impact of mortality would contribute. We assumed a geometric progression with different ratios (0.8, 0.9, and 0.98):

i. The n-day lagged mortality is multiplied by coefficients \(0.8^n\)

\[
M_{VAR1} = \begin{bmatrix} 60 & 51 & 62 \\ 73 & 60 & 51 \\ 61 & 73 & 60 \\ 55 & 61 & 73 \end{bmatrix}_{4 \times 3} \cdot \begin{bmatrix} 0.8^0 \\ 0.8^1 \\ 0.8^2 \end{bmatrix}_{3 \times 1} = \begin{bmatrix} 140.48 \\ 153.64 \\ 157.8 \\ 150.52 \end{bmatrix}_{4 \times 1}
\]

ii. The n-day lagged mortality is multiplied by coefficients \(0.9^n\)

\[
M_{VAR2} = \begin{bmatrix} 60 & 51 & 62 \\ 73 & 60 & 51 \\ 61 & 73 & 60 \\ 55 & 61 & 73 \end{bmatrix}_{4 \times 3} \cdot \begin{bmatrix} 0.9^0 \\ 0.9^1 \\ 0.9^2 \end{bmatrix}_{3 \times 1} = \begin{bmatrix} 156.12 \\ 168.31 \\ 175.3 \\ 169.03 \end{bmatrix}_{4 \times 1}
\]

iii. The n-day lagged mortality is multiplied by coefficients \(0.98^n\)

\[
M_{VAR3} = \begin{bmatrix} 60 & 51 & 62 \\ 73 & 60 & 51 \\ 61 & 73 & 60 \\ 55 & 61 & 73 \end{bmatrix}_{4 \times 3} \cdot \begin{bmatrix} 0.98^0 \\ 0.98^1 \\ 0.98^2 \end{bmatrix}_{3 \times 1} = \begin{bmatrix} 169.5249 \\ 180.7804 \\ 190.164 \\ 184.8892 \end{bmatrix}_{4 \times 1}
\]

Method 3:

MVPCA: The principal component analysis (PCA) [14] is an unsupervised methodology to reduce the dimensionality of numerous variables. The first component represents the maximum variability explained. Therefore, we use only the first component in the first attempt. The second employs all eigenvector such that all variabilities are maintained.

i. Only multiply the first eigenvector (to obtain the first principal component):

\[
M_{PCA1} = \begin{bmatrix} 60 & 51 & 62 \\ 73 & 60 & 51 \\ 61 & 73 & 60 \\ 55 & 61 & 73 \end{bmatrix}_{4 \times 3} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix}_{3 \times 1} = \begin{bmatrix} 60a + 51b + 62c \\ 73a + 60b + 51c \\ 61a + 73b + 60c \\ 55a + 61b + 73c \end{bmatrix}_{4 \times 1}
\]

ii. Multiply all eigenvector (to obtain all principal components) and the corresponding percentage:
Method 4:

\[ MV_{\text{adjust}} : \text{Separate the current mortality from the lag mortalities. Create a reduced} \]
\[ \text{matrix that sums over } L \text{ lagged mortalities but not the current mortality, } \sum(x_{stL}) = \]
\[
\begin{bmatrix}
51 + 62 \\
60 + 51 \\
73 + 60 \\
61 + 73
\end{bmatrix}_{4 \times 1}
= \begin{bmatrix}
113 \\
111 \\
133 \\
134
\end{bmatrix}_{4 \times 1}, \text{and adjust } \sum(x_{stL}) \text{ as a covariate in the DLNM.}
\]

\[
\begin{bmatrix}
60 \\
73 \\
61 \\
55
\end{bmatrix}_{4 \times 1} \text{ is the outcome & } \sum(x_{stL}) = \begin{bmatrix}
113 \\
111 \\
133 \\
134
\end{bmatrix}_{4 \times 1} \text{ is adjusted in the DLNM}
\]

Method 5:

\[ MV_{\text{DLNM}} : \text{Similar to method 4, but instead of treating the sum of previous lagged} \]
\[ \text{mortalities as a covariate, } \sum(x_{stL}) = \begin{bmatrix}
113 \\
111 \\
133 \\
134
\end{bmatrix}_{4 \times 1} \text{ is considered as the offset of the} \]
\[ \text{current mortality in the DLNM.}
\]

\[
\begin{bmatrix}
60 \\
73 \\
61 \\
55
\end{bmatrix}_{4 \times 1} \text{ is the outcome & } \sum(x_{stL}) = \begin{bmatrix}
113 \\
111 \\
133 \\
134
\end{bmatrix}_{4 \times 1} \text{ is the offset in the DLNM}
\]

The simulation study

To validate the performance of above approaches, we conducted simulation study under the null hypothesis for 1000 times. The null distribution was generated by permutations of mortality such that the outcome mortality and temperature measures were not correlated. The validity of each model was assessed. If the proportion of rejecting the null hypothesis does not exceed the significance level of 5%, the proposed strategy is a valid test.
Software

All the statistical analyses and simulations were conducted by the software R, equipped with the package “dlm” by Gasparrini [2].

Results

According to the permuted samples, the observed type-I error for $MV_{sum}$ is presented in Table 2. Method 1 and 2 are based on summation of previous outcomes, but with different weights. Therefore, results of $MV_{AR1}$, $MV_{AR2}$, and $MV_{AR3}$ are similar and not shown.

Due to a negative value in principal components, $MV_{PCA1}$ and $MV_{PCA2}$ failed to satisfy the assumption of Poisson distribution model and did not generate any results from the DLNM package. Hence, the type-I errors were not obtained. Note that type-I error rates for $MV_{sum}$, $MV_{AR1}$, $MV_{AR2}$, and $MV_{AR3}$ were much larger than the nominal level of 0.05. The inflation is increasing with respect to the number of lagged outcomes. Therefore, these methods are not valid, although the idea is simple and could be easily implemented.

Table 3 showed that the type-I error rate of $MV_{adj}$ was between 0.056-0.065 when the lagged exposure was up to 10 days. The type-I error becomes 0.083-0.183 when the lagged exposures was up to 20 days. Finally, Type-I errors are 0.076-0.295 for 30 lagged temperature measures. Although the type-I error rate was consistently larger than 0.05, $MV_{adj}$ yielded much smaller type-I errors comparing to summation based methods.

Finally, Table 4 showed that the type-I error rate of $MV_{DLNM}$ was smaller than 0.05 when the lagged exposure was 10 days. The type-I error ranges from 0 to 0.078 if the lagged exposure was 20 days. When the lagged exposure was 30 days, the type-I
error ranges from 0 to 0.102., Therefore, the results indicated that MV\textsubscript{DLNM} is the only valid test. For the 10, 20, and 30 lagged exposures, the cumulative outcome mortality could be implemented up to 10, 10, and 13 days, respectively.

Applications to the real data

Our previous work using 6 six major cities in Taiwan [15] reported a significant temperature impact on mortality. In this research, only Taipei city is available for the recent years. Hence the association based on the conventional DLNM is not statistically significant. However, the MV\textsubscript{DLNM} could provide more significant overall p-values (Table 5). Regarding the temperature measure up to 10 lagged days, significant p-values were observed for 4 or more lagged mortalities incorporated in the model. For temperature in 20 and 30 lagged days, if 5 or more lagged mortalities are used in the MV\textsubscript{DLNM}, the result would suggest significant associations. Hence, the cumulative outcomes could contribute to the association with lagged exposures.

Discussions

Through simulation studies, we examined several novel approaches to characterize the association between the lagged mortality and the lagged temperature measures. Results suggested that most methods are invalid, although these statistical concepts are intuitive and could be implemented effortlessly. The only valid model is MV\textsubscript{DLNM}, where the log transformed summation of lagged outcomes is treated as the offset in the DLNM model. The MV\textsubscript{DLNM} model is:

$$\log(\mu_t) = \alpha + s(x_t, t, \beta) + \beta_0 \theta + \beta_{PM2.5} PM2.5_t + \sum_{i=1}^{p} f(z_{t_i} : \theta) + \text{offset} \left( \log(\text{sum}(x_{sl})) \right)$$

The illustration of real data analysis of Taipei City from 2012-2016 confirmed that the cumulative mortality is significantly associated with lagged temperature measures.
In contrast, the conventional DLNM method failed to provide significant results as we previously reported[15]. Nevertheless, this new strategy is a very useful tool and could be adopted by various research fields when the cumulative effect of the outcome is desired.

Limitations:

The data used in this study is limited to Taipei, the capital of Taiwan, while the relationship between temperature and mortality may consist of various profiles in other regions. For example, the accessibility and quality of medical care may be different in smaller towns. Additionally, we considered the all-cause mortality, since we could not further classify the causes of death into more categories, such as sudden cardiac death or myocardial infraction, which are more likely to be related to temperature and air pollution. As for the temperature, only daily mean temperature was considered in this study. We didn’t explore the highest, lowest temperature, and intraday temperature variation in the contribution to human death. Finally, some researchers proposed that a threshold is needed to differentiate impact of hot and cold temperature on mortality. In contrast, we use the continuous temperature measures to employ spline functions and polynomials.
• Ethics approval and consent to participate
  o The data satisfy ethics approval and consent to participate
• Consent for publication
  o Each author approves consent for publication
• Competing interests
  o None
• Funding
  o None
• Authors’ contributions
  o Yi-Hau Chen
    • Research concepts and manuscript edits
  o Xing-Yi Huang
    • Carried out analysis, simulations, generates results
  o Pei-Cheng Kuo
    • Draft the manuscript
  o Chao-Yu Guo
    • Supervise study, draft and edits manuscript
• Acknowledgements
  o NA
References

1. Basu R: High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environmental health* 2009, 8(1):40.

2. Gasparrini A, Armstrong B, Kenward MG: Distributed lag non-linear models. *Statistics in medicine* 2010, 29(21):2224-2234.

3. Curriero FC, Heiner KS, Samet JM, Zeger SL, Strug L, Patz JA: Temperature and mortality in 11 cities of the eastern United States. *American journal of epidemiology* 2002, 155(1):80-87.

4. Mills D, Schwartz J, Lee M, Sarofim M, Jones R, Lawson M, Duckworth M, Deck L: Climate change impacts on extreme temperature mortality in select metropolitan areas in the United States. *Climatic Change* 2015, 131(1):83-95.

5. Baccini M., Biggeri A., Accetta G., Kosatsky T., Katsouyanni K., Analitis A., Anderson H.R., Bisanti L., D'Ippoliti D., Danova J. *et al.* Heat effects on mortality in 15 European cities. *Epidemiology* 2008, 19(5):711-719.

6. Chung Y, Lim Y-H, Honda Y, Guo Y-LL, Hashizume M, Bell ML, Chen B-Y, Kim H: Mortality related to extreme temperature for 15 cities in northeast Asia. *Epidemiology* 2015, 26(2):255-262.

7. Janssen N, Fischer P, Marra M, Ameling C, Cassee F: Short-term effects of PM2.5, PM10 and PM2.5 – 10 on daily mortality in the Netherlands. *Science of the Total Environment* 2013, 463:20-26.

8. Dai H, Song W, Gao X, Chen L: Study on relationship between ambient PM10, PM2.5 pollution and daily mortality in a district in Shanghai. *Wei sheng yan jiu= Journal of hygiene research* 2004, 33(3):293-297.

9. Chen C-C, Lin B-C, Yap L, Chiang P-H, Chan T-C: The Association between Ambient Temperature and Acute Diarrhea Incidence in Hong Kong, Taiwan, and Japan. *Sustainability* 2018, 10(5):1417.

10. Vicedo-Cabrera AM, Forsberg B, Tobias A, Zanobetti A, Schwartz J, Armstrong B, Gasparrini A: Associations of Inter- and Intraday Temperature Change With Mortality. *American journal of epidemiology* 2016, 183(4):286-293.

11. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B: Time series regression studies in environmental epidemiology. *International journal of epidemiology* 2013, 42(4):1187-1195.

12. [https://eservice.cwb.gov.tw/HistoryDataQuery/index.jsp]

13. [https://erdb.epa.gov.tw/DataRepository/EnvMonitor/AirQualityMonitorDayData.aspx?topic1=%E5%A4%A7%E6%B0%A3&topic2=%E7%92%B0%E5%A2%83%E5%8F%8A%E7%94%9F%E6%85%8B%E7%9B%A3%E6%B8%AC&subject=%E7%A9%BA%E6%B0%A3%E5%93%81%E8%B3%AA]
14. Jolliffe IT, Cadima J: Principal component analysis: a review and recent developments. Philosophical transactions Series A, Mathematical, physical, and engineering sciences 2016, 374(2065):20150202.

15. Guo CY, Pan WC, Chen MJ, Tsai CW, Chen NT, Su HJ: When are we most vulnerable to temperature variations in a day? PLoS One 2014, 9(12):e113195.
Table 1: Daily all-cause deaths and three primary parameters, mean temperature, ozone concentration, PM$_{2.5}$ concentration, in Taipei from 2012 to 2016.

| Daily Data    | Descriptive Statistics |
|---------------|------------------------|
|               | N  | Mean  | SD   | Min | Max  |
| Mortality     | 1818 | 61.397 | 9.584 | 32  | 101  |
| Mean temperature | 5455 | 20.55 | 9.22  | 1.53 | 73.21 |
| Mean $O_3$   | 1816 | 28.030 | 9.628 | 6.317 | 81.967 |
| Mean $PM_{2.5}$ | 1801 | 20.459 | 10.413 | 3.714 | 87.143 |
Table 2: Type-I errors of $MV_{sum}$ ($MV_{AR1}$, $MV_{AR2}$, and $MV_{AR3}$ are similar)

| Lag | (10)  | (20)  | (30)  |
|-----|-------|-------|-------|
| ~Lag1 | 0.557 | 0.712 | 0.74  |
| ~Lag2 | 0.797 | 0.937 | 0.951 |
| ~Lag3 | 0.856 | 0.98  | 0.989 |
| ~Lag4 | 0.879 | 0.997 | 0.996 |
| ~Lag5 | 0.872 | 0.996 | 0.998 |
| ~Lag6 | 0.869 | 0.996 | 0.999 |
| ~Lag7 | 0.864 | 0.997 | 1     |
| ~Lag8 | 0.872 | 0.999 | 1     |
| ~Lag9 | 0.871 | 0.998 | 1     |
| ~Lag10 | 0.874 | 0.997 | 1     |
| ~Lag11 | 0.996 | 1     |
| ~Lag12 | 0.996 | 1     |
| ~Lag13 | 0.995 | 1     |
| ~Lag14 | 0.994 | 1     |
| ~Lag15 | 0.993 | 1     |
| ~Lag16 | 0.994 | 1     |
| ~Lag17 | 0.995 | 1     |
| ~Lag18 | 0.993 | 1     |
| ~Lag19 | 0.995 | 1     |
| ~Lag20 | 0.995 | 1     |
| ~Lag21 | 1     |
| ~Lag22 | 1     |
| ~Lag23 | 1     |
| ~Lag24 | 1     |
| ~Lag25 | 1     |
| ~Lag26 | 1     |
| ~Lag27 | 1     |
| ~Lag28 | 1     |
| ~Lag29 | 1     |
| ~Lag30 | 1     |
Table 3: Type-I errors of MV\textsubscript{adjust}

| ~Lag   | lag(10) | lag(20) | lag(30) |
|--------|---------|---------|---------|
| ~Lag1  | 0.058   | 0.083   | 0.076   |
| ~Lag2  | 0.06    | 0.097   | 0.105   |
| ~Lag3  | 0.065   | 0.11    | 0.129   |
| ~Lag4  | 0.066   | 0.12    | 0.148   |
| ~Lag5  | 0.062   | 0.129   | 0.163   |
| ~Lag6  | 0.062   | 0.133   | 0.18    |
| ~Lag7  | 0.067   | 0.141   | 0.192   |
| ~Lag8  | 0.065   | 0.142   | 0.213   |
| ~Lag9  | 0.064   | 0.136   | 0.225   |
| ~Lag10 | 0.065   | 0.135   | 0.231   |
| ~Lag11 |        | 0.144   | 0.232   |
| ~Lag12 |        | 0.148   | 0.239   |
| ~Lag13 |        | 0.153   | 0.241   |
| ~Lag14 |        | 0.165   | 0.244   |
| ~Lag15 |        | 0.162   | 0.245   |
| ~Lag16 |        | 0.165   | 0.256   |
| ~Lag17 |        | 0.162   | 0.253   |
| ~Lag18 |        | 0.163   | 0.256   |
| ~Lag19 |        | 0.174   | 0.264   |
| ~Lag20 |        | 0.183   | 0.268   |
| ~Lag21 |        |         | 0.284   |
| ~Lag22 |        |         | 0.295   |
| ~Lag23 |        |         | 0.297   |
| ~Lag24 |        |         | 0.308   |
| ~Lag25 |        |         | 0.302   |
| ~Lag26 |        |         | 0.299   |
| ~Lag27 |        |         | 0.294   |
| ~Lag28 |        |         | 0.301   |
| ~Lag29 |        |         | 0.299   |
| ~Lag30 |        |         | 0.295   |
Table 4: Type-I errors of MVDLN

| Mortality     | lag(10) | lag(20) | lag(30) |
|---------------|---------|---------|---------|
| ~Lag1         | 0       | 0       | 0       |
| ~Lag2         | 0.001   | 0       | 0       |
| ~Lag3         | 0.008   | 0       | 0       |
| ~Lag4         | 0.026   | 0.001   | 0       |
| ~Lag5         | 0.03    | 0.004   | 0       |
| ~Lag6         | 0.034   | 0.009   | 0.002   |
| ~Lag7         | 0.035   | 0.019   | 0.002   |
| ~Lag8         | 0.036   | 0.028   | 0.005   |
| ~Lag9         | 0.043   | 0.031   | 0.009   |
| ~Lag10        | 0.049   | 0.04    | 0.014   |
| ~Lag11        |         | 0.052*  | 0.02     |
| ~Lag12        |         | 0.055*  | 0.029    |
| ~Lag13        |         | 0.052*  | 0.038    |
| ~Lag14        |         | 0.056*  | 0.056*   |
| ~Lag15        |         | 0.058*  | 0.065*   |
| ~Lag16        |         | 0.064*  | 0.072*   |
| ~Lag17        |         | 0.068*  | 0.079*   |
| ~Lag18        |         | 0.073*  | 0.087*   |
| ~Lag19        |         | 0.074*  | 0.089*   |
| ~Lag20        |         | 0.078*  | 0.087*   |
| ~Lag21        |         |         | 0.085*   |
| ~Lag22        |         |         | 0.086*   |
| ~Lag23        |         |         | 0.091*   |
| ~Lag24        |         |         | 0.092*   |
| ~Lag25        |         |         | 0.091*   |
| ~Lag26        |         |         | 0.098*   |
| ~Lag27        |         |         | 0.099*   |
| ~Lag28        |         |         | 0.099*   |
| ~Lag29        |         |         | 0.098*   |
| ~Lag30        |         |         | 0.102*   |

*These p-values indicate the inflated type-I errors
Table 5: Real data application of MV_DLMN: Overall P-value

| ~Lag1 | lag(10)  | lag(20)  | lag(30)  |
|-------|----------|----------|----------|
| ~Lag2 | 0.807516 | 0.996482 | 0.999951 |
| ~Lag3 | 0.100355 | 0.489865 | 0.945508 |
| ~Lag4 | 0.009709 | 0.045561 | 0.338491 |
| ~Lag5 | 0.001388# | 0.009588 | 0.060874 |
| ~Lag6 | 0.000231# | 0.001842# | 0.012078# |
| ~Lag7 | 0.000107# | 0.000605# | 0.004511# |
| ~Lag8 | 0.000120# | 0.000350# | 0.001656# |
| ~Lag9 | 0.000086# | 0.000088# | 0.000184# |
| ~Lag10 | 0.000031# | 0.000019# | 0.000045# |
| ~Lag11 | 0.000009* | 0.000017# | 0.000157# |
| ~Lag12 | 0.000009* | 0.000010# | 0.000157# |
| ~Lag13 | 0.000014* | 0.000011# | 0.000157# |
| ~Lag14 | 0.000020* | 0.000015* | 0.000157# |
| ~Lag15 | 0.000019* | 0.000015* | 0.000157# |
| ~Lag16 | 0.000009* | 0.000010* | 0.000157# |
| ~Lag17 | 0.000003* | 0.000005* | 0.000157# |
| ~Lag18 | 0.000002* | 0.000003* | 0.000157# |
| ~Lag19 | 0.000001* | 0.000001* | 0.000157# |
| ~Lag20 | 0.000001* | 0.000001* | 0.000157# |
| ~Lag21 | 0.000001* | 0.000001* | 0.000157# |
| ~Lag22 | 0.000001* | 0.000001* | 0.000157# |
| ~Lag23 | 0.000002* | 0.000002* | 0.000157# |
| ~Lag24 | 0.000003* | 0.000003* | 0.000157# |
| ~Lag25 | 0.000004* | 0.000004* | 0.000157# |
| ~Lag26 | 0.000006* | 0.000006* | 0.000157# |
| ~Lag27 | 0.000007* | 0.000007* | 0.000157# |
| ~Lag28 | 0.000009* | 0.000009* | 0.000157# |
| ~Lag29 | 0.000014* | 0.000014* | 0.000157# |
| ~Lag30 | 0.000021* | 0.000021* | 0.000157# |

# denotes significant p-values
*denotes invalid type-I error rates indicated by the permutation study in Table 4