Mediation and Fuzzy Mediation Analysis for Multiple Covariates and Its Applications to Solar Power Data

Department of Mathematics and Statistics, Sejong University

Department of Digital Content, Sejong University

Yeon Kyung Lee[1], Jung Eun Lee[1], Sung Wook Baik[2] and Jin Hee Yoon [1*]

Nov,20,2021

KIIS 사단법인 한국지능시스템학회 Korean Institute of Intelligent Systems

조선대학교 CHOSUN UNIVERSITY

Contents

1 Introduction

- Background
- Need for our research

2 Preliminaries

- Simple Mediation Analysis
- Fuzzy Mediation Analysis

3 Our study

- Mediation Analysis for Multiple Covariates
- Fuzzy Mediation Analysis for Multiple Covariates

4 Application to data

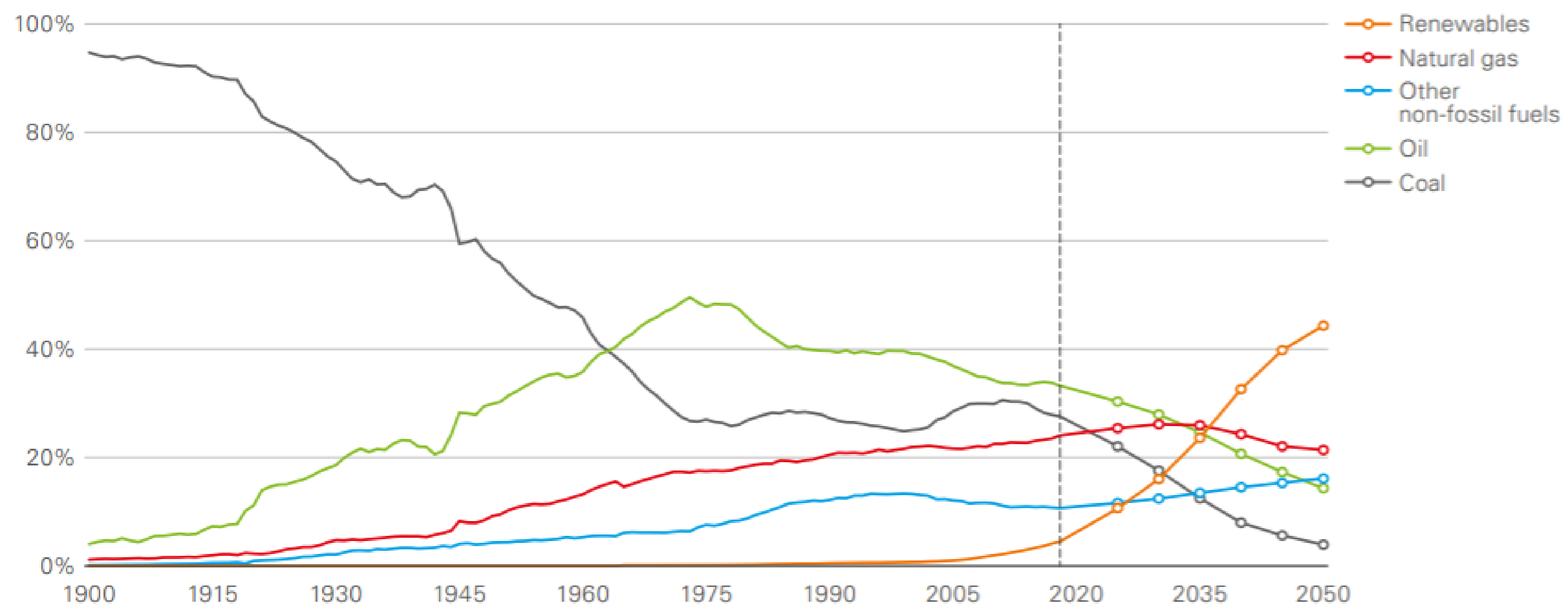
- Solar Power Data

1 Introduction

01 Introduction

[Background] The importance of accurately predicting solar energy

01 BP, Energy Outlook 2020 edition



The increasing amount of renewable energy, and solar energy which accounts for a large portion of it.

Solar power generation

- Advantages : free development location conditions, safe and long equipment life, infinite energy source
- Disadvantages : fluctuating output depending on the position of the sun, the season, and the weather

Many studies have been conducted on solar power prediction models

http://energytransitionkorea.org/sites/default/files/2020-09/bp-energy-outlook-2020.pdf

01 Introduction

Prior research and the need for our research

01 In Solar Prediction models,

lack of research <u>to understand the relationships with variables</u>

- Statistic model : ARMA, ARIMA
- Machine learning methods : SVM, ANN, LSTM

<u>Our study will prevent excessive interpretation of the results.</u>

02 In mediation models,

lack of model for <u>multiple independent variables or covariates</u> with mediators

various mediation models for multiple mediators, moderators or confounding variables have been proposed.
But, unfortunately, a **mediation model for multiple independent variable or covariates has not been dealt with.**

03 About data,

Observations require fuzzification

Observations such as actual weather information, solar radiation, and solar energy generation require fuzzification **because they are observed with ambiguous values rather than crisp values**

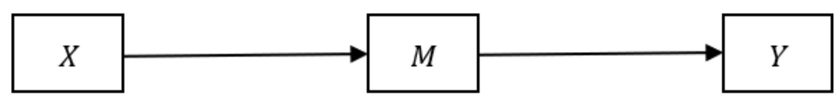
2

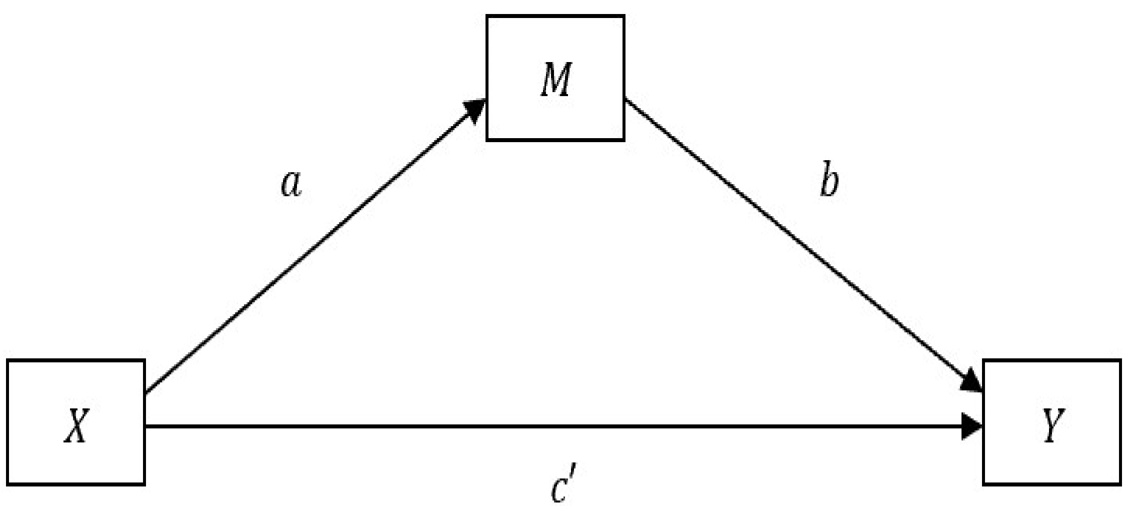
Preliminaries

02 Simple Mediation Analysis and Fuzzy Mediation Analysis

1. Simple Mediation Analysis

01 Simple Mediation Model by Baron and Kenny (1986)





$$Y = \beta_{10} + \beta_{11}X + \varepsilon_1$$

$$M = \beta_{20} + \beta_{21}X + \varepsilon_2$$

$$Y = \beta_{30} + \beta_{31}X + \beta_{32}M + \varepsilon_3$$

Total effect = indirect effect + direct effect

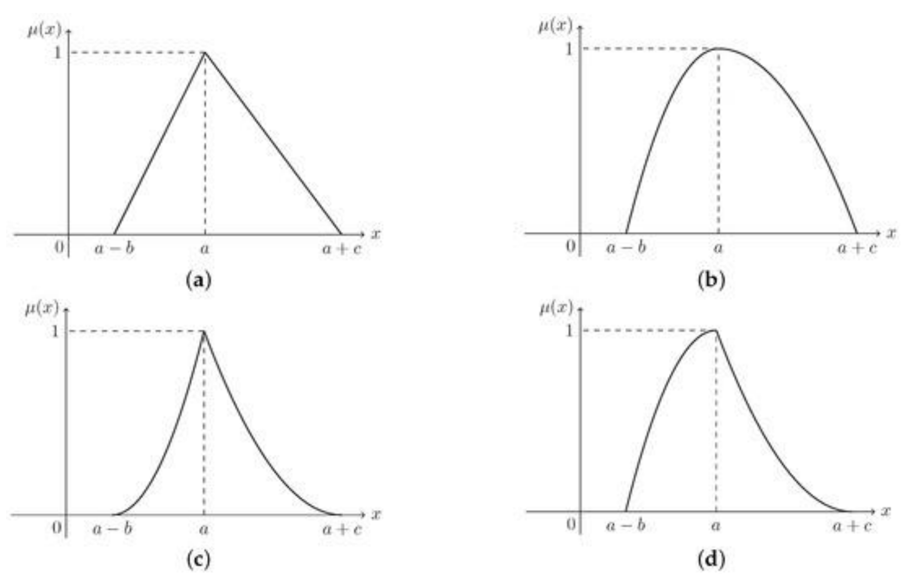
- Total effect : β_{11}
- Indirect effect : $\beta_{21}\beta_{32}$
- Direct effect : β_{31}

02 Simple Mediation Analysis and Fuzzy Mediation Analysis

2. Simple Fuzzy Mediation Analysis

01 Fuzzy numbers by Zadeh (1965)

<LR-fuzzy number>



$$\mu_A(x) = \begin{cases} L\left(\frac{m-x}{l}\right) & if\ x \leq m \\ R\left(\frac{x-m}{l}\right) & if\ x > m \end{cases}$$

Generalization of a real number by representing a continuous set of possible values between 0 and 1, not one value

L : left shape function
R : right shape function

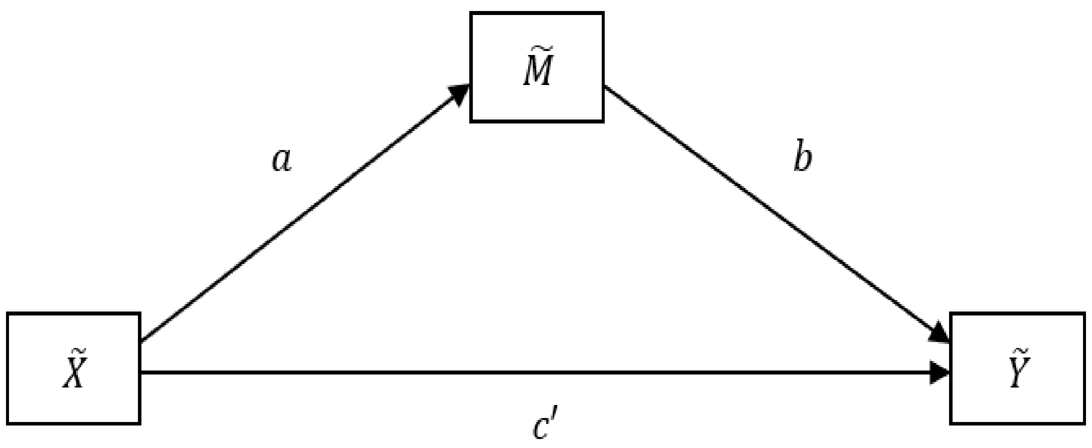
m : mode of the fuzzy number
l : width of the left
r : width of the right

Zhou, J.; Huang, C.; Zhao, M.; Li, H. Entropy and Semi-Entropies of LR Fuzzy Numbers' Linear Function with Applications to Fuzzy Programming. *Entropy* 2019, *21*, 697.

02 Simple Mediation Analysis and Fuzzy Mediation Analysis

2. Simple Fuzzy Mediation Analysis

02 Simple Fuzzy Mediation Model by Yoon (2020)



$$\widetilde{Y} = \beta_{10} \oplus \beta_{11}\widetilde{X} \oplus \widetilde{E}_1,$$

$$\widetilde{M} = \beta_{20} \oplus \beta_{21}\widetilde{X} \oplus \widetilde{E}_2,$$

$$\widetilde{Y} = \beta_{30} \oplus \beta_{31}\widetilde{X} \oplus \beta_{32}\widetilde{M} \oplus \widetilde{E}_3$$

Total effect = indirect effect + direct effect

- Total effect : β_{11}
- Indirect effect : $\beta_{21}\beta_{32}$
- Direct effect : β_{31}

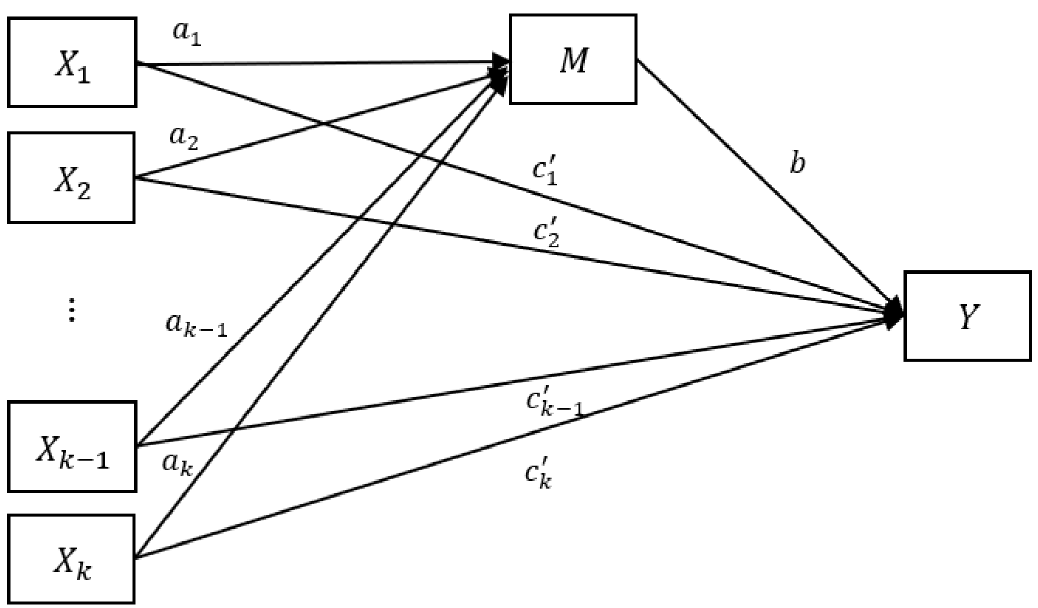
3

Our study

03 Mediation Analysis for Multiple Covariates

1. Mediation Analysis for Multiple Covariates with one mediator

01 One mediator



$$Y = \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \cdots + \beta_{1p}X_p + \varepsilon_1$$

$$M = \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \cdots + \beta_{2p}X_p + \varepsilon_2$$

$$Y = \beta_{30} + \sum_{j=1}^{p} \beta_{31}^{j} X_j + \beta_{32}M + \varepsilon_3, \ \text{where } j = 1, \cdots, p.$$

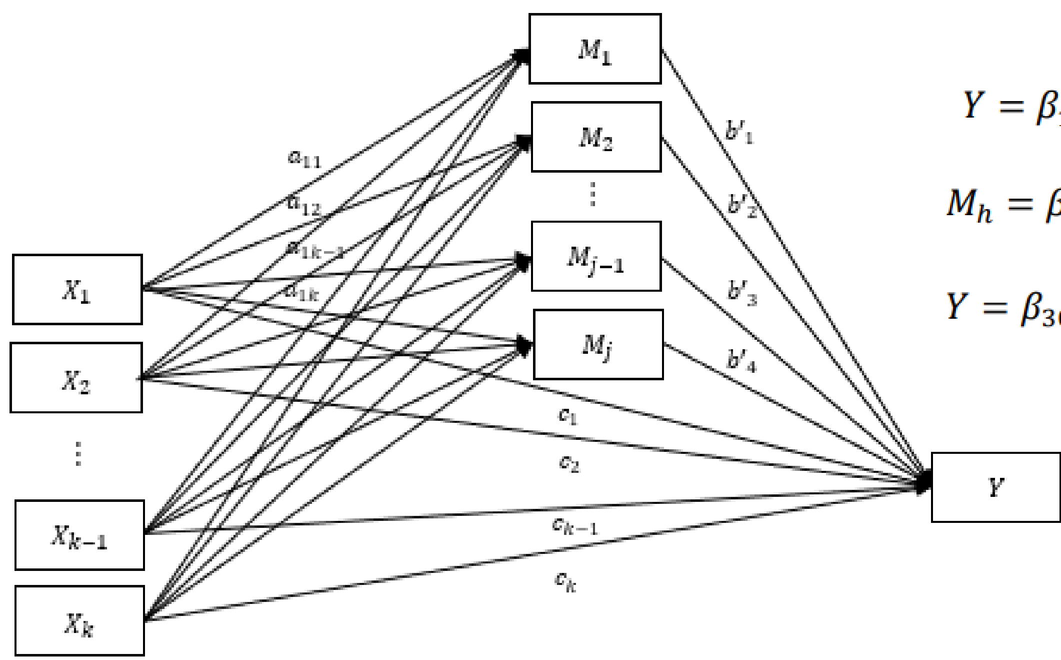
Total effect = indirect effect + direct effect

- Total effect : β_{1j}
- Indirect effect : $\beta_{2j}\beta_{32}$
- Direct effect : β_{31}^{j}

03 Mediation Analysis for Multiple Covariates

2. Mediation Analysis for Multiple Covariates with multiple mediators

02 Multiple mediators



$$Y = \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \cdots + \beta_{1p}X_p + \varepsilon_1$$

$$M_h = \beta_{20}^h + \beta_{21}^h X_1 + \beta_{22}^h X_2 + \cdots + \beta_{2p}^h X_p + \varepsilon_2^h$$

$$Y = \beta_{30} + \sum_{j=1}^{p} \beta_{31}^{j} X_j + \sum_{h=1}^{k} \beta_{32}^{h} M_h + \varepsilon_3 \quad \text{where } j = 1, \cdots, p \text{ and } h = 1, \cdots, k.$$

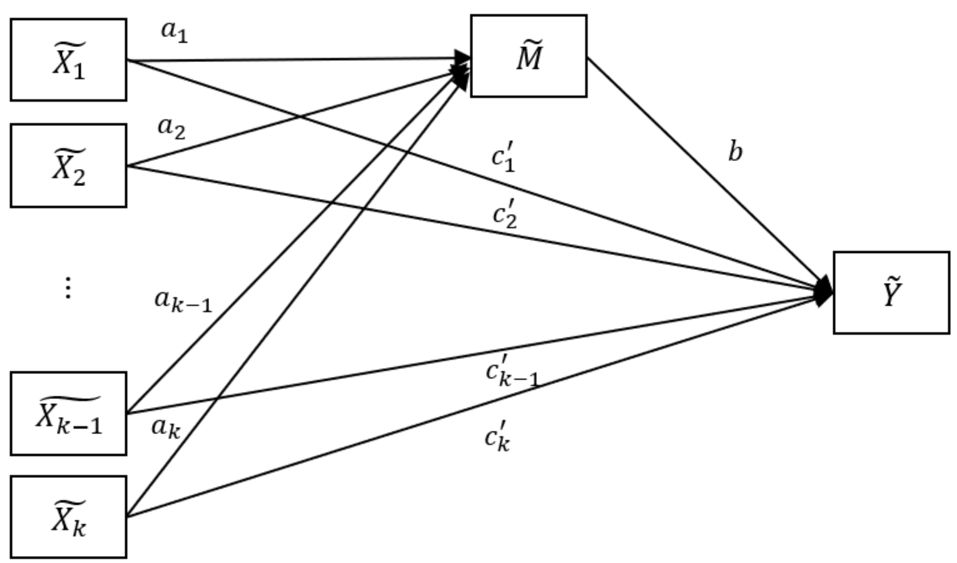
Total effect = indirect effect + direct effect

- Total effect : β_{1j}
- Indirect effect : $\beta_{2j}^h \beta_{32}^h$
- Direct effect : β_{31}^j

04 Fuzzy Mediation Analysis for Multiple Covariates

1. Fuzzy Mediation Analysis for Multiple Covariates with one mediator

01 One mediator



$$\tilde{Y} = \beta_{10} \oplus \beta_{11}\tilde{X}_1 \oplus \cdots \oplus \beta_{1p}\tilde{X}_p \oplus \tilde{E}_1,$$

$$\tilde{M} = \beta_{20} \oplus \beta_{21}\tilde{X}_1 \oplus \cdots \oplus \beta_{2p}\tilde{X}_p \oplus \tilde{E}_2,$$

$$\tilde{Y} = \beta_{30} \oplus \textstyle\sum_{j=1}^{p} \beta_{31}^{j}\tilde{X}_j \oplus \beta_{32}\tilde{M} \oplus \tilde{E}_3, \text{ where } j = 1, \cdots, p.$$

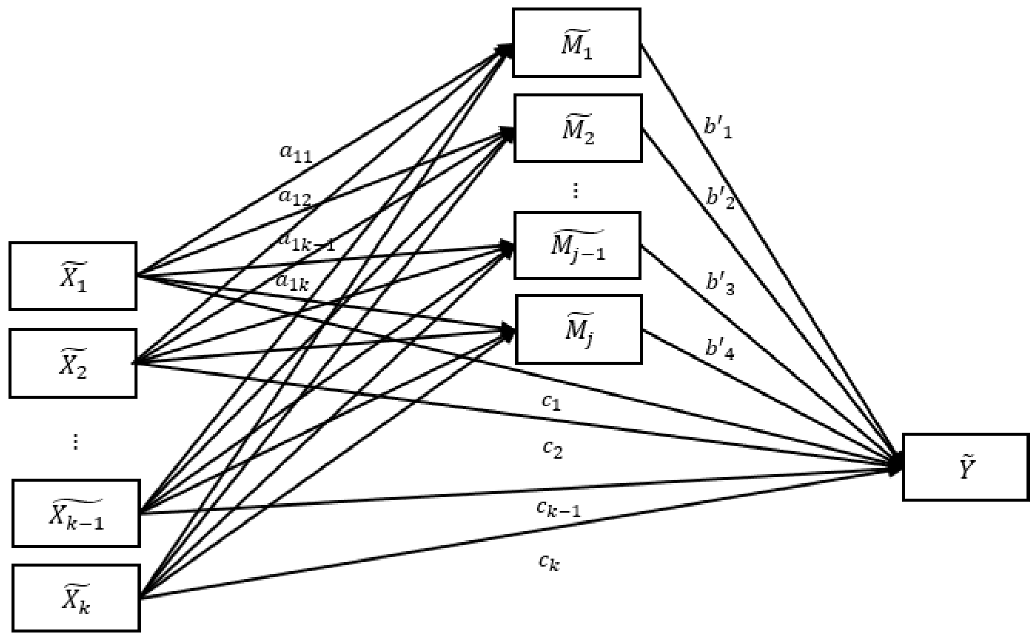
Total effect = indirect effect + direct effect

- Total effect : β_{1j}
- Indirect effect : $\beta_{2j}\beta_{32}$
- Direct effect : β_{31}^{j}

04 Fuzzy Mediation Analysis for Multiple Covariates

2. Fuzzy Mediation Analysis for Multiple Covariates with multiple mediators

02 Multiple mediators



$$\tilde{Y} = \beta_{10} \oplus \beta_{11}\tilde{X}_1 \oplus \cdots \oplus \beta_{1p}\tilde{X}_p \oplus \tilde{E}_1,$$

$$\widetilde{M}_h = \beta_{20} \oplus \beta_{21}^h\tilde{X}_1 \oplus \cdots \oplus \beta_{2p}^h\tilde{X}_p \oplus \tilde{E}_2,$$

$$\tilde{Y} = \beta_{30} \oplus \sum_{j=1}^{p}\beta_{31}^j\widetilde{X}_i \oplus \sum_{h=1}^{k}\beta_{32}^h\widetilde{M}_h \oplus \tilde{E}_3, \text{ where } j = 1,\cdots,p \text{ and } h = 1,\cdots,k.$$

Total effect = indirect effect + direct effect

- Total effect : β_{1j}
- Indirect effect : $\beta_{2j}^h\beta_{32}^h$
- Direct effect : β_{31}^j

04 Fuzzy Mediation Analysis for Multiple Covariates

3. Estimation for Fuzzy Mediation Analysis for Multiple Covariates with mediators

03

$$\tilde{Y}_i = \beta_0 \oplus \beta_1 \tilde{X}_{1i} \oplus \beta_2 \tilde{X}_{2i} \oplus \cdots \oplus \beta_p \tilde{X}_{pi} \oplus \tilde{E}_i.$$

The variables are represented by

$X_{ij} = \left(l_{x_{ij}}, x_{ij}, r_{x_{ij}}\right)$ and $Y_i = (l_{y_i}, y_i, r_{y_i})$ for $i = 1, \ldots, n$, $j = 1, \ldots, p$.

It is assumed that $\tilde{E}_i$ are the fuzzy random errors for expressing fuzziness.

The objective function can be obtained based on the LSE, and here the LSE can be expressed as follows

$$d^2(\tilde{Y}_i, \textstyle\sum_{j=0}^{p} \beta_{kj} \tilde{X}_{ij}) = (l_{y_i} - \sum_{j=0}^{p} \beta_{kj} l_{x_{ij}})^2 + (y_i - \sum_{j=0}^{p} \beta_{kj} x_{ij})^2 + (r_{y_i} - \sum_{j=0}^{p} \beta_{kj} r_{x_{ij}})^2$$

J. H. Yoon, Fuzzy mediation analysis, Int. J. Fuzzy Syst. vol. 22, no. 1, pp.338-349 (2020)
J. H. Yoon, Fuzzy Moderation and Moderated-Mediation Analysis. Int. J. Fuzzy Syst. vol 22, no. 1, pp. 338-349 (2020)

04 Fuzzy Mediation Analysis for Multiple Covariates

3. Estimation for Fuzzy Mediation Analysis for Multiple Covariates with mediators

03

To find the solution vector, we define a ***triangular fuzzy matrix (t.f.m.)***

which is expressed by

$$\tilde{X} = \begin{bmatrix} (1,1,1) & (l_{x_{11}}, x_{11}, r_{11}) & \cdots & (l_{x_{1p}}, x_{1p}, r_{1p}) \\ & \vdots & \ddots & \vdots \\ (1,1,1) & (l_{x_{n1}}, x_{n1}, r_{n1}) & \cdots & (l_{x_{np}}, x_{np}, r_{np}) \end{bmatrix}$$

$$\tilde{\boldsymbol{y}} = \left[(l_{y_1}, y_1, r_{y_1}), \cdots, (l_{y_n}, y_n, r_{y_n})\right]^t$$

$$\tilde{X} \diamond \tilde{Y} = l_x l_y + xy + r_x r_y$$

$$\widehat{\boldsymbol{\beta}_k} = \left(\tilde{\boldsymbol{X}}^t \diamond \tilde{\boldsymbol{X}}\right)^{-1} \tilde{\boldsymbol{X}}^t \diamond \tilde{\boldsymbol{y}}$$

where

$$\tilde{X}^t \diamond \tilde{X} = \left[\sum_{i=1}^n (l_{x_{il}} l_{x_{ij}} + x_{il} x_{ij} + r_{x_{il}} r_{x_{ij}})\right]_{(p+1)\times(p+1)}$$

$\tilde{X}^t \diamond \tilde{y} = \left[\sum_{i=1}^n (l_{x_{il}} l_{y_i} + x_{il} y_i + r_{x_{il}} r_{y_i})\right]_{(p+1)\times 1}$, for $l = 0,1,\ldots,p$.

4 Application to data

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Data

- This data was collected every hour from 1 a.m. on January 1, 2015 to 11 p.m. on December 31, 2017 at a solar energy facility in Dangjin, Korea.

1. Processing missing values => Linear interpolation

$$f(x_k) = f(x_{k-1}) + \frac{f(x_{k+1}) - f(x_{k-1})}{x_{k+1} - x_{k-1}}(x_k - x_{k-1})$$

2. Normalization

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Data

- This data was collected every hour from 1 a.m. on January 1, 2015 to 11 p.m. on December 31, 2017 at a solar energy facility in Dangjin, Korea.

3. Significant variables (after regression analysis)

	temp	rain	wind speed	humidity	sun hour	solar radiation	snow	cloud
p-value	0.000	0.019	0.000	0.000	0.000	0.000	0.000	0.000

4. fuzzification

For the last data, we use diff before that

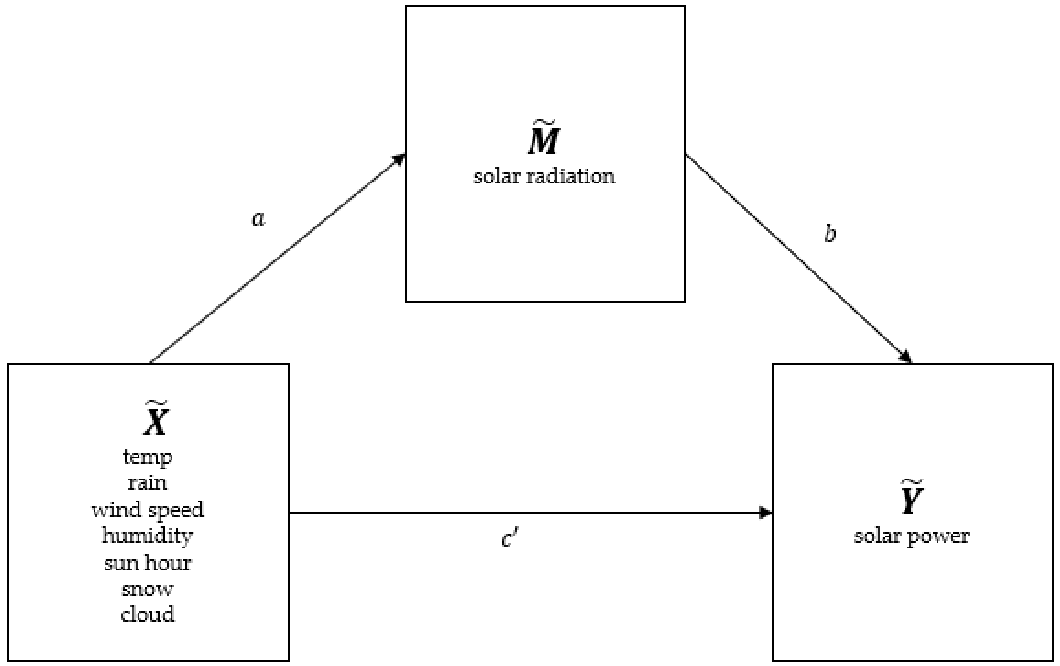
Date	Before fuzzification	m, l, r	After fuzzification
2015-01-01-1:00AM	-4.4	M : -4.4, l, r : |-4.6-(-4.4)|/2 = 0.1	(-4.5,-4.4,-4.3)
2015-01-01-2:00AM	-4.6	M : -4.6, l, r : |-4.7-(-4.6)|/2 = 0.05	(-4.65, -4.6, -4.55)
2015-01-01-3:00AM	-4.7	M : -4.7, l, r : |-5-(-4.7)|/2 = 0.15	(-4.85, -4.7, -4.55)

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Research hypothesis

01 Solar radiation mediates weather conditions with solar power

independent variables : weather data
mediator : radiation
dependent variable : solar power



02 Fuzzy Mediation Analysis for Multiple Covariates will have more significant result than Mediation.

- H. Lee. Use of the Moving Average of the Current Weather Data for the Solar Power Generation Amount Prediction. KMMS. vol. 19, no. 8, pp. 1530-1537 (2016)
-. J. Jeong, Y. Tae. Comparison Analysis for Characteristic of Short-Term Power generation Forecasting models for Building Integrated photovoltaics. AIK. vol. 36, no.2, pp. 665-666 (2016)
- K. Lee, W. Kim. Forecasting of 24_hours Ahead Photovoltaic Power Output Using Support Vector Regression. JKIIT. vol. 14, no. 3, pp. 175-183 (2016)
- J. Jeong, Y. Chae. Improvement for Forecasting of Photovoltaic Power Output using Real Time Weather Data based on Machine Learning. KSLES. vol. 25, no. 1, pp. 119-125 (2018)

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Model

<Fuzzy Mediation Analysis for Multiple Covariates with one mediator>

$$\tilde{Y} = \beta_{10} \oplus \beta_{11}\widetilde{X_1} \oplus \beta_{12}\widetilde{X_2} \oplus \beta_{13}\widetilde{X_3} \oplus \beta_{14}\widetilde{X_4} \oplus \beta_{15}\widetilde{X_5} \oplus \beta_{16}\widetilde{X_6} \oplus \beta_{17}\widetilde{X_7} + \varepsilon_1$$

$$\tilde{M} = \beta_{20} \oplus \beta_{21}\widetilde{X_1} \oplus \beta_{22}\widetilde{X_2} \oplus \beta_{23}\widetilde{X_3} \oplus \beta_{24}\widetilde{X_4} \oplus \beta_{25}\widetilde{X_5} \oplus \beta_{26}\widetilde{X_6} \oplus \beta_{27}\widetilde{X_7} + \varepsilon_2$$

$$\tilde{Y} = \beta_{30} \oplus \beta_{31}\widetilde{X_1} \oplus \beta_{32}\widetilde{X_2} \oplus \beta_{33}\widetilde{X_3} \oplus \beta_{34}\widetilde{X_4} \oplus \beta_{35}\widetilde{X_5} \oplus \beta_{36}\widetilde{M_6} \oplus \beta_{37}\widetilde{X_7} \oplus \beta_{38}\widetilde{X_8} + \varepsilon_3$$

Total effect = $\sum_{i=1}^{7} \beta_{1i}$

Total effect = indirect effect + direct effect

Direct effect = $\sum_{i=1, i\neq 6}^{8} \beta_{3i}$

Indirect effect = $\sum_{i=1}^{7} \beta_{2i} \cdot \beta_{36} = \beta_{36} \cdot (\sum_{i=1}^{7} \beta_{2i})$

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Parameters estimates

CMA: **classic mediation analysis**
FMA: **fuzzy mediation analysis**

(a) Parameters estimates between weather conditions and solar power

Method	Parameter estimates							
	β_{10} const	β_{11} temp	β_{12} rain	β_{13} windspeed	β_{14} humidity	β_{15} sun hour	β_{16} snow	β_{17} cloud
CMA	0.114	0.240	- 0.718	0.153	- 0.291	0.327	0.038	0.064
FMA	0.112	0.248	- 0.585	0.170	- 0.291	0.316	0.043	0.056

(b) Parameters estimates between weather conditions and solar radiation

Method	Parameter estimates							
	β_{20} const	β_{21} temp	β_{22} rain	β_{23} windspeed	β_{24} humidity	β_{25} sun hour	β_{26} snow	β_{27} cloud
CMA	0.107	0.201	-0.601	0.113	-0.255	0.360	0.063	0.054
FMA	0.104	0.208	- 0.487	0.130	- 0.254	0.350	0.067	0.047

(c) Parameters estimates between weather conditions + solar radiation and solar power

Method	Parameter estimates								
	β_{30} const	β_{31} temp	β_{32} rain	β_{33} windspeed	β_{34} humidity	β_{35} sun hour	β_{36} solar radiation	β_{37} snow	β_{38} cloud
CMA	0.005	0.033	- 0.101	0.037	- 0.029	- 0.042	1.027	- 0.027	0.009
FMA	0.006	0.035	- 0.088	0.038	- 0.031	- 0.041	1.027	- 0.026	0.008

05 Fuzzy mediation Analysis with Multiple Covariates for Solar Power Data

Effects of the solar radiation

Method	Effect		
	Total effect	Direct effect	Indirect effect
CMA	-0.187	-0.120	-0.067[a]
FMA	-0.043	-0.105	0.062[b]

[a] $(-0.065)(1.027)= -\ 0.067$

[b] $(0.061)(1.027)=\ 0.062$

FMA considers the spread of data rather than CMA using summarized data

FMA effect will be more reliable than CMA.

06 Conclusions

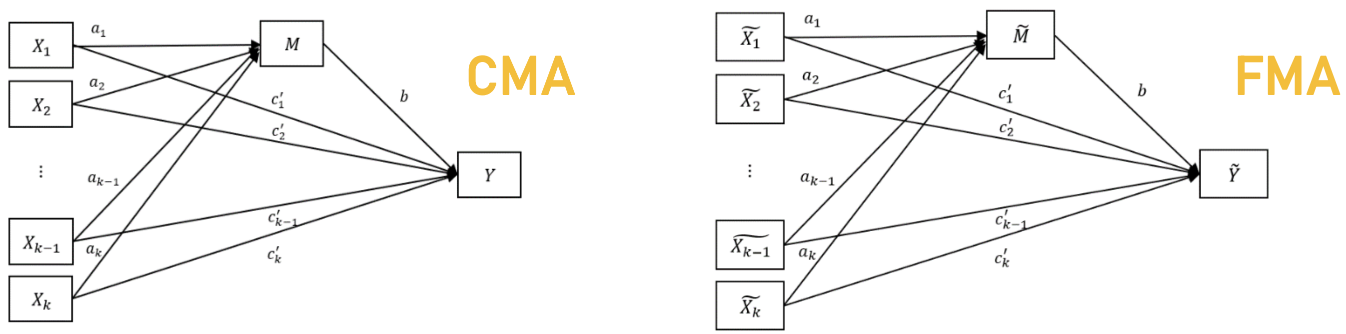
01 Need for our research

In Solar Prediction models,
lack of research to understand the relationships with variables

02 Research hypothesis

1.Solar radiation will mediate weather conditions and solar power.
2.Fuzzing the data will have more meaningful results than using crisp data

03 Model



04 Result

Method	Effect		
	Total effect	Direct effect	Indirect effect
CMA	-0.187	-0.120	-0.067^a
FMA	-0.043	-0.105	0.062^b

FMA effect will be more reliable than CMA.

05 Follow-up study

- Statistical Inferences about direct effect and indirect effect
- This theory is to be extended to other various mediation, moderation, mediation-moderation analysis

Thank you