

석 사 스 터 디 그 룩 발 표

신재생에너지산업 추정 연구

발 표 자 김 동 현



CONTENTS

01

도입

- 신재생에너지개요
- 연구배경 및 필요성
- 연구 체계
- 추정 방법론

02

발전업 매출액 추정

- 가중단순선형 회귀분석
- 로버스트 회귀분석

03

발전업 종사자 수 추정

- 단순선형 회귀분석
- 로버스트 회귀분석

04

분석 결과 및 결론

- 추정 결과
- Q&A

CONTENTS

01 도입

- 신재생에너지 개요
- 연구 배경 및 필요성
- 연구 체계
- 추정 방법론

01

1. 도입

“ 신재생에너지 개요 ”

신재생에너지란?

- 석탄, 석유, 원자력 및 천연가스 등 화석연료가 아닌 태양에너지, 풍력, 소수력, 연료전지, 석탄의 액화, 가스화, 해양에너지, 폐기물에너지 및 기타로 구분되고 있고 이외에도 지열, 수소, 석탄에 의한 물질을 혼합한 유동성 연료를 의미함.
- 그러나 실질적인 신재생에너지란, 넓은 의미로는 석유를 대체하는 에너지원으로, 좁은 의미로는 신·재생에너지원을 나타냄.
- 국내에서 미래에 사용될 신재생에너지로 석유, 석탄, 원자력, 천연가스 등 화석연료가 아닌 에너지로 11개 분야를 지정(신재생에너지개발 및 이용·보급촉진법 제2조 근거)하였고, 세분하여 보면 아래와 같음.
- 신에너지 3개 분야 : 연료전지, 석탄액화·가스화, 수소에너지
- 재생에너지 8개 분야 : 태양열, 태양광발전, 바이오매스, 풍력, 소수력, 지열, 해양에너지, 폐기물에너지

01

1. 도입

“ 연구 배경 및 필요성 ”

배경

1. 전 세계적으로 신재생에너지산업 투자 및 일자리 증가한다는 점

- 전 세계 신재생에너지산업 투자는 최근 10년간 지속적인 성장세를 보이고 있으며, 개발도상국 및 신흥국은 약 2배, 중국의 경우 약 8배의 투자 증가
- BNEF(2018)에 따르면, 2050년까지 예상되는 전력부문의 투자 11.5조 달러 중 81%인 9.3조 달러가 재생에너지에 투자될 것으로 예상됨.
- IRENA(2018)에 따르면, 세계 재생에너지산업에 직간접적으로 종사하는 인원은 꾸준히 증가하여 2017년 기준 1천만 명을 넘어선 것으로 분석됨.

2. 국제적으로 탄소제로 요구가 거세지고 있는 점

- 국제적으로 탄소제로를 위해 노력하고 있으며, 국내에서도 신재생에너지의 확대가 불가피함.

01

1. 도입

“ 연구 배경 및 필요성 ”

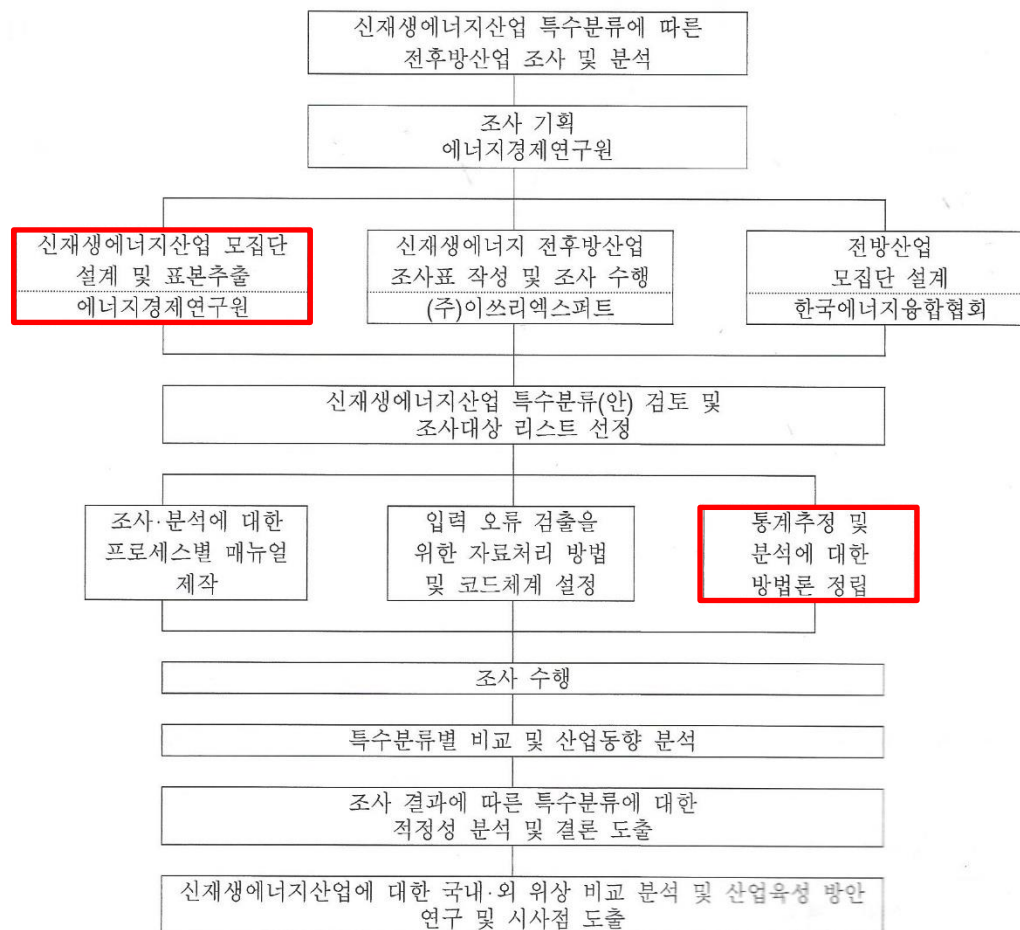
필요성

1. 신재생에너지산업의 융복합적 특성과 규모, 지원성과 및 파급효과 분석 등을 파악할 수 있는 산업분류체계 도입 필요
 - 후방산업보다 대규모 프로젝트 개발, 건설, 전력 생산 및 유지를 위한 전방산업에 대한 투자가 확대 중임.
 - 전방산업 : 건설/시공, 발전 등의 서비스 산업
 - 후방산업 : 제조기반 산업
 - 향후 신재생에너지 전후방산업의 매출액 및 고용 규모를 파악하고, 정부지원정책의 효과분석을 위한 산업분류체계 확립 및 조사 필요
2. 신재생에너지산업에 대한 분류체계 마련과 전후방 산업 조사 필요성 증가
 - 신재생에너지산업에 대한 공신력 있고 합리적인 분류체계의 마련을 통해 구조를 명확하게 파악하고, 객관적인 산업 통계 조사 필요
 - 산업 육성 및 파급 효과를 분석하고 지원 성과를 측정하여 성과 목표 및 발전 방향을 제시할 필요성 제기

01

1. 도입

“ 연구 체계 ”



01

1. 도입

“ 추정 방법론 ”

- 추정해야하는 대분류 : 제조업, 건설업, 공급업, 서비스업, **발전업**

발전업

- 발전소 용량에 따라 매출액과 종사자가 증가하는 특성을 반영, 가중단순선형 회귀분석, 로버스트 회귀분석, 단순선형 회귀분석 등 표본의 특성에 적합한 추정식을 적용함.
- 독립변수 X : 발전소 용량(발전량)
- 종속변수 Y : 매출액(단위: 백만원) or 종사자 수(단위: 명)

CONTENTS

02 발전업 매출액 추정

- 가중단순선형 회귀분석
- 로버스트 회귀분석

2. 발전업 매출액 추정

(1) 가중단순선형 회귀분석(Weighted Least Squares; WLS)

- 용량이 증가함에 따라, 매출액의 편차도 커지는 경향을 보임.
- 이를 보정하기 위하여 큰 용량에 대해서 가중치를 작게 주는 ‘가중단순선형 회귀분석’ 사용

개요

- 발전용량에 따른 매출액의 변화를 가중단순선형 회귀모형에 적용.

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \epsilon_i, & i &= 1, \dots, n \\ \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i, & i &= 1, \dots, n \end{aligned}$$

여기서 x 는 신재생에너지원별 발전용량(kW), y 는 매출액(백만원)임.

- 가중단순선형회귀란?

- * 단순선형회귀분석은 잔차의 제곱합(sum of squared residuals; $\sum_i (y_i - \beta_0 - \beta_1 x_i)^2$)

을 최소화하는데 반해, 가중단순 선형회귀는 이에 가중치를 곱해, 잔차의 제곱 가중합(sum of weighted squared residuals; $\sum_i w_i (y_i - \beta_0 - \beta_1 x_i)^2$)을 최소화함.

- * x_i 가 증가함에 따라 잔차들의 분산이 커지는 경우에 단순선형 회귀모형으로 분석하게 되면, 제대로 추정하지 못하는 문제가 발생하게 되는데, 이를 가중단순선형 회귀 분석에서 가중치(w_i)를 적용하여 보정함으로써 해결할 수 있음.

- * 본 분석에서 용량이 증가함에 따라 잔차들의 분산이 커지는 경향이 있음을 확인하였으며, 가중치를 적용하여 값이 증가하는 분산을 보정하였음. 여러 가중치 함수

(Weight function)를 적용해 본 결과, 가장 나은 가중치인 $w_i = \frac{1}{x_i}$ 를 사용하였음.

- * 추정된 β_1 이 양수이면, 발전용량과 매출액은 정(+)적인 관계를 가지며,
- * 추정된 β_1 이 음수이면, 발전용량과 매출액은 부(-)적인 관계를 가짐.
- * 아래 잔차도표와 절대잔차도표는 가중치로 보정된 잔차와 절대잔차를 이용하여 그렸음.

02

2. 발전업 매출액 추정

(1) 가중단순선형 회귀분석(Weighted Least Squares; WLS)

설명

- 가중회귀는 일반 선형회귀보다 조금 더 유연하다고 볼 수 있음.
- 기존 최소제곱법이 잔차의 제곱을 최소화 하는 것이었다면, 본 기법은 잔차에 가중치를 적용한 제곱 값을 최소화 하는 기법임.

[Example]

- 기존의 선형회귀분석은 ‘등분산성’을 가정하고 진행하게 됨.
- 오차항의 분산이 일정하다 -> 오차를 예측하는데 있어 좀 더 정확한 오차를 계산할 수 있음을 시사함.
- 그러나, 등분산성이 성립하지 않는 경우, 즉 어떠한 오차의 분산은 크고, 어떠한 분산은 작다면 여기에 가중치(weight)를 적용하여 등분산성을 맞춰주는 것임.

➤ 잔차제곱합 및 가중변수 식 : $S = \sum_{i=1}^n W_{ii} r_i^2, \quad W_{ii} = \frac{1}{\sigma_i^2}$

➤ 추정된 회귀계수 : $\hat{\beta} = (X^T W X)^{-1} X^T W y$

2. 발전업 매출액 추정

(2) 로버스트 회귀분석(Robust Regression Analysis)

- 이상치(outliers)가 있는 경우에 가중단순선형 회귀모형으로 분석하게 되면, 제대로 추정하지 못하는 문제가 발생하게 됨.
- 이를 LTS(Least Trimmed Squares) 방법을 적용하여 해결할 수 있음.

개요

- 본 분석에서는 로버스트(robust) 회귀분석 중 하나인 Least trimmed squares (LTS) 방법을 적용하여 분석하였음.
- 발전용량에 따른 매출액의 변화를 가중단순선형 회귀모형에 적용.

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \epsilon_i, & i &= 1, \dots, n \\ \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i, & i &= 1, \dots, n \end{aligned}$$

여기서 x 는 신재생에너지원별 발전용량(kW), y 는 매출액(백만원)임.

- Least trimmed squares (LTS) 란?

- * 단순선형회귀분석은 잔차의 제곱합 (sum of squared residuals;

$$\sum_i (y_i - \beta_0 - \beta_1 x_i)^2 \text{을 최소화하는데 반해, LTS는 전체 잔차 중 그 값이 작은 일부}$$

분만을 이용해, 잔차 일부의 제곱 합(trimmed sum of squared residuals;

$$\sum_{i \in T} (y_i - \beta_0 - \beta_1 x_i)^2 \text{을 최소화함.}$$

- * 이상치(Outliers)가 있는 경우에 단순선형 회귀모형으로 분석하게 되면, 제대로 추정하지 못하는 문제가 발생하게 되는데, 이를 LTS 방법을 적용하여 해결할 수 있음.
- * 본 분석에서 의무대상자의 종사자 수 예상값 도출 시, 풍력과 연료전지 분야에서 이상치가 예상값에 크게 영향을 주고 있음을 확인하였으며, LTS 방법을 적용하여, 이러한 이상치의 영향을 효과적으로 줄일 수 있음.
- * 추정된 β_1 이 양수이면, 발전용량과 매출액은 정(+)적인 관계를 가지며,
- * 추정된 β_1 이 음수이면, 발전용량과 매출액은 부(-)적인 관계를 가짐.
- * 아래 잔차도표는 $y_i - \beta_0 - \beta_1 x_i$ 를 x_i 에 대해 도식화 한 것이며,
절대잔차도표는 $|y_i - \beta_0 - \beta_1 x_i|$ 를 x_i 에 대해 도식화 한 것임.

02

2. 발전업매출액 추정

(2) 로버스트 회귀분석(Robust Regression Analysis)

설명

[일반 회귀분석의 가정에서의 '이상치·특이점 문제']

- 일반 회귀분석은 예측식을 구성하는 독립변수들의 예측계수(회귀계수)를 구할 때, 잔차의 제곱합이 최소가 되게 하는 최소제곱법(OLS) 방법을 사용함.
- OLS로 회귀선을 추정하는 과정에서 특이점이 존재하는 경우, 회귀직선의 결정에 큰 영향을 미치게 됨.

[로버스트 회귀분석]

- OLS를 적용하는 모든 예측상황에서 사용 가능하며, 가중치 메커니즘을 사용해 영향력 있는 관측치를 낮춰줄 수 있음.
- 특히, 데이터에서 이상치를 제외시킬 강력한 이유가 없는 경우에 유용함.
- 잔차의 제곱을 이용하는 최소제곱법 대신에 절댓값의 합이 최소가 되도록 계수를 추정하는 방식임.

CONTENTS

03 발전업 종사자 수 추정

- 단순선형 회귀분석
- 로버스트 회귀분석

03

3. 발전업 종사자 수 추정

(1) 단순선형 회귀분석(Ordinary Least Squares; OLS)

- 선형성이라는 기본 가정이 충족된 상태에서 독립변수와 종속변수의 관계를 설명하거나 예측하는 통계방법

개요

- 발전용량에 따른 매출액의 변화를 단순선형 회귀모형에 적용.

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \epsilon_i, & i &= 1, \dots, n \\ \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i, & i &= 1, \dots, n \end{aligned}$$

여기서 x 는 신재생에너지원별 발전용량(kW), y 는 종사자 수임.

- 단순선형회귀란?

- * 잔차의 제곱합(sum of squared residuals; $\sum_i (y_i - \beta_0 - \beta_1 x_i)^2$)을 최소화하는 β_0 과 β_1 을 찾는 분석 방법임.
- * 추정된 β_1 이 양수이면, 발전용량과 매출액은 정(+)적인 관계를 가지며,
- * 추정된 β_1 이 음수이면, 발전용량과 매출액은 부(-)적인 관계를 가짐.
- * 아래 잔차도표는 $y_i - \beta_0 - \beta_1 x_i$ 를 x_i 에 대해 도식화 한 것이며, 절대잔차도표는 $|y_i - \beta_0 - \beta_1 x_i|$ 를 x_i 에 대해 도식화 한 것임.

3. 발전업 종사자 수 추정

(2) 로버스트 회귀분석(Robust Regression Analysis)

- 이상치(outliers)가 있는 경우에 가중단순선형 회귀모형으로 분석하게 되면, 제대로 추정하지 못하는 문제가 발생하게 됨.
- 이를 LTS(Least Trimmed Squares) 방법을 적용하여 해결할 수 있음.

개요

- 본 분석에서는 로버스트(robust) 회귀분석 중 하나인 Least trimmed squares (LTS) 방법을 적용하여 분석하였음.
- 발전용량에 따른 매출액의 변화를 가중단순선형 회귀모형에 적용.

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \epsilon_i, & i &= 1, \dots, n \\ \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i, & i &= 1, \dots, n \end{aligned}$$

여기서 x 는 신재생에너지원별 발전용량(kW), y 는 매출액(백만원)임.

- Least trimmed squares (LTS) 란?

- * 단순선형회귀분석은 잔차의 제곱합 (sum of squared residuals;

$\sum_i (y_i - \beta_0 - \beta_1 x_i)^2$)을 최소화하는데 반해, LTS는 전체 잔차 중 그 값이 작은 일부

분만을 이용해, 잔차 일부의 제곱 합(trimmed sum of squared residuals;

$\sum_{i \in T} (y_i - \beta_0 - \beta_1 x_i)^2$)을 최소화함.

- * 이상치(Outliers)가 있는 경우에 단순선형 회귀모형으로 분석하게 되면, 제대로 추정하지 못하는 문제가 발생하게 되는데, 이를 LTS 방법을 적용하여 해결할 수 있음.
- * 본 분석에서 의무대상자의 종사자 수 예상값 도출 시, 풍력과 연료전지 분야에서 이상치가 예상값에 크게 영향을 주고 있음을 확인하였으며, LTS 방법을 적용하여, 이러한 이상치의 영향을 효과적으로 줄일 수 있음.
- * 추정된 β_1 이 양수이면, 발전용량과 매출액은 정(+)적인 관계를 가지며,
- * 추정된 β_1 이 음수이면, 발전용량과 매출액은 부(-)적인 관계를 가짐.
- * 아래 잔차도표는 $y_i - \beta_0 - \beta_1 x_i$ 를 x_i 에 대해 도식화 한 것이며,
절대잔차도표는 $|y_i - \beta_0 - \beta_1 x_i|$ 를 x_i 에 대해 도식화 한 것임.

CONTENTS

04 분석 결과 및 결론

- 추정 결과
- Q&A

04

4. 분석 결과 및 결론

부연 설명: 추정 알고리즘

1. 에너지원별로 데이터를 나눔.
2. 데이터 전처리를 진행함.
 - 의무대상자의 데이터는 사업체별로 용량 및 매출액, 종사자 수를 합하여 하나의 사업체로 고려함.
 - 미의무대상자의 데이터는 그대로 두되, 수력 표본 중 ‘한국농어촌공사’만 용량 및 매출액, 종사자 수를 합하여 하나의 사업체로 고려함.
3. 나눈 데이터를 이용하여 최종 예상 매출액 또는 종사자 수의 합을 도출함.
 - 조사된 값이 있는 경우, 그 값을 그대로 이용하여 최종 예상 매출액 또는 종사자 수에 합함.
 - 조사된 값이 없는 경우, 위에서 설명된 방법을 이용하여 회귀선을 도출, 적용하여 예상 값을 도출한 후 최종 예상 매출액 또는 종사자 수에 합함. 이때, 예상 값이 0보다 작은 경우, 0으로 합함.

4. 분석 결과 및 결론

(1) 추정 결과

발전업 매출액(2018)

(단위 : 백만원)

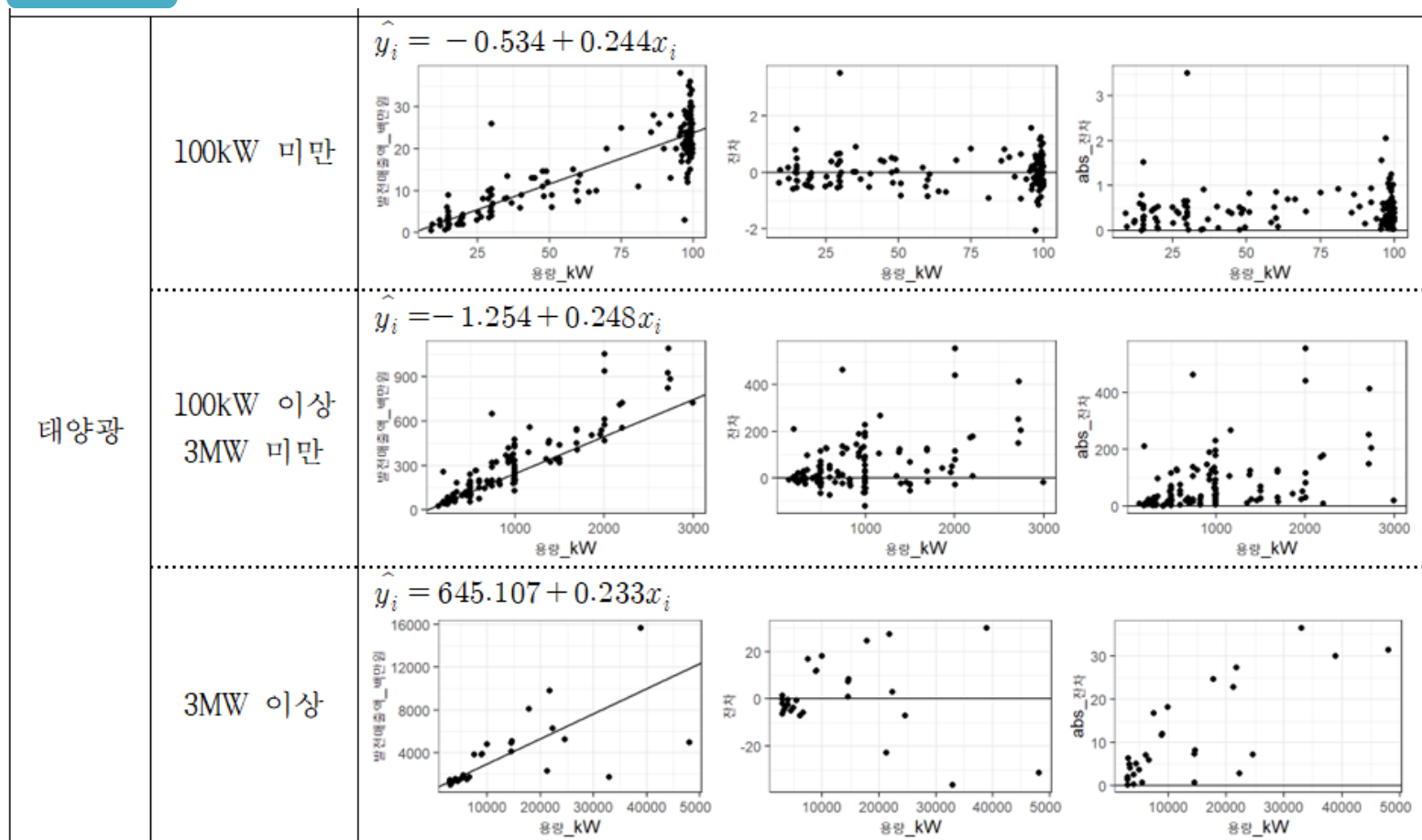
에너지원		조사 O	조사 X (추정)	합계
태양광	100kW 미만	3,523	791,763	795,286
	100kW 이상 3MW 미만	33,295	1,411,294	1,444,589
	3MW 이상	67,625	159,190	226,815
	계	104,443	2,362,247	2,466,690
풍력		194,952	406,398	601,350
수력	1MW 미만	475	1,043	1,518
	1MW 이상	3,482	11,698	15,180
	계	3,957	12,741	16,698
바이오에너지		690,443	333,956	1,024,399
폐기물		306,591	155,144	461,735
연료전지		152,947	164,743	317,690
계		1,453,333	3,435,229	4,888,562

- 가중단순선형 회귀분석
- 에너지원 : 태양광 100kW 미만, 태양광 3MW 이상, 수력, 바이오, 풍력, 폐기물, 연료전지
- 로버스트 회귀분석
- 에너지원 : 태양광 100kW 이상 3MW 미만

4. 분석 결과 및 결론

(1) 추정 결과 – 매출액

태양광

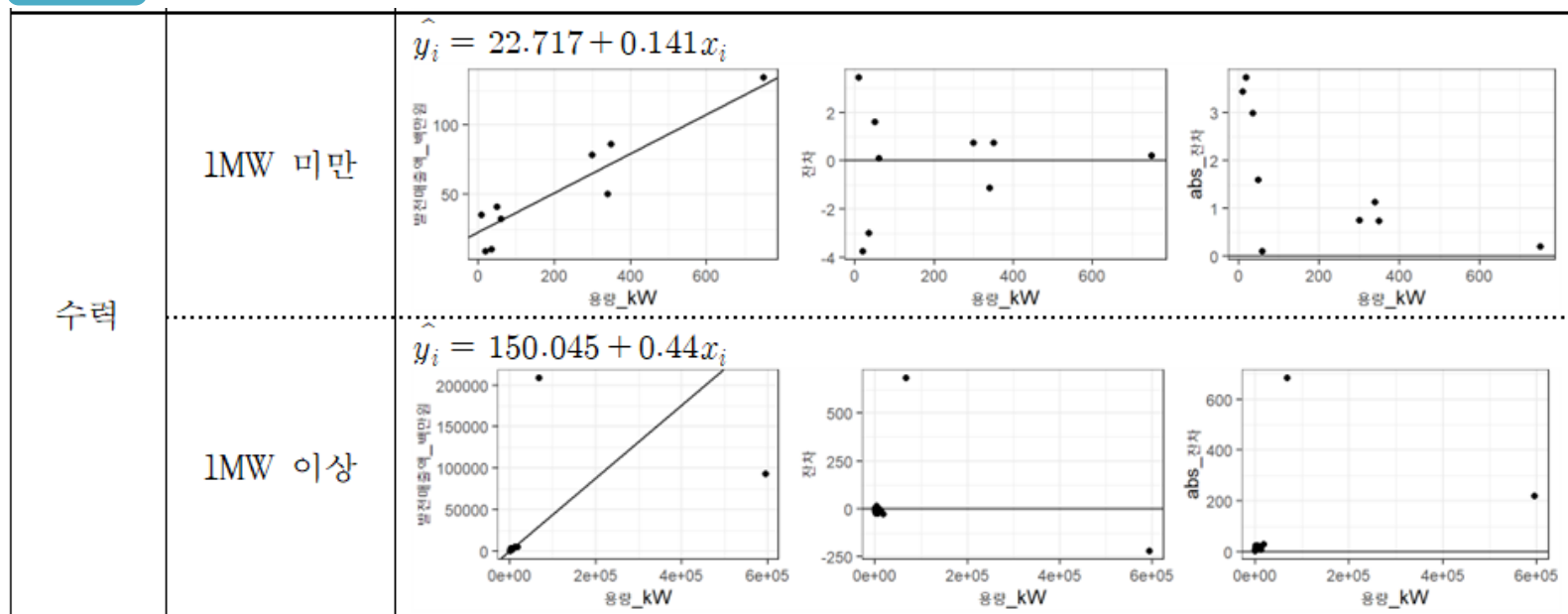


04

4. 분석 결과 및 결론

(1) 추정 결과 – 매출액

수력



04

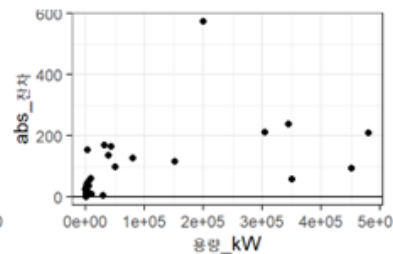
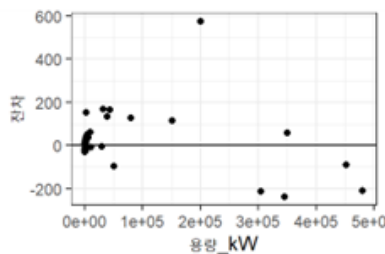
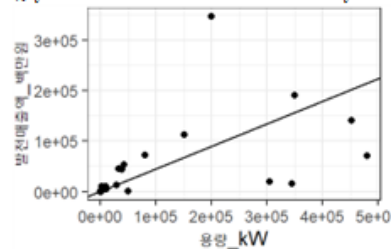
4. 분석 결과 및 결론

(1) 추정 결과 – 매출액

바이오에너지

바이오에너지

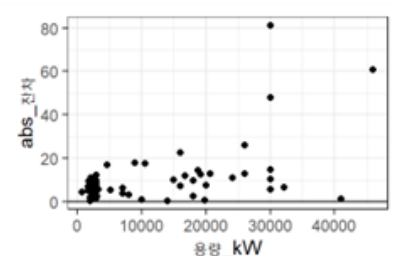
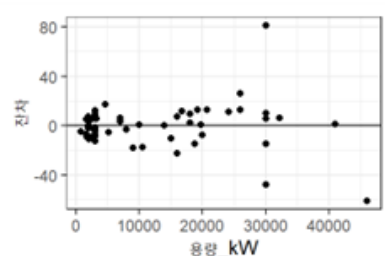
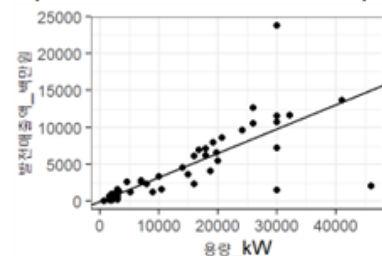
$$\hat{y}_i = 461.357 + 0.446x_i$$



풍력

풍력

$$\hat{y}_i = -76.176 + 0.328x_i$$



04

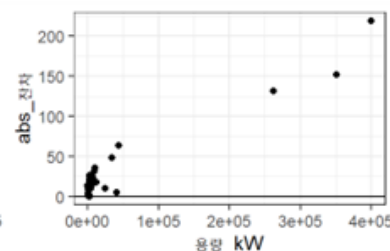
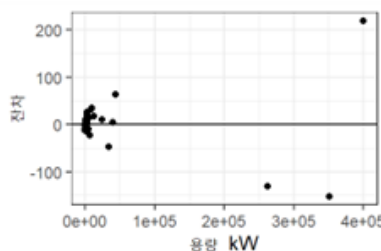
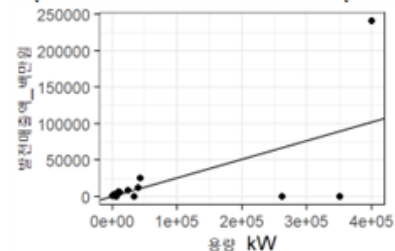
4. 분석 결과 및 결론

(1) 추정 결과 – 매출액

폐기물

폐기물

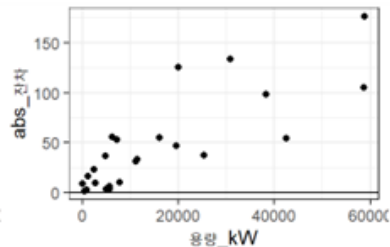
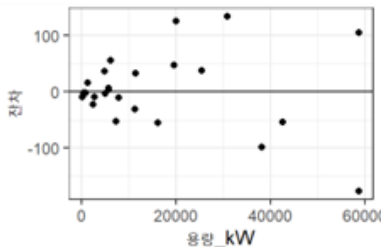
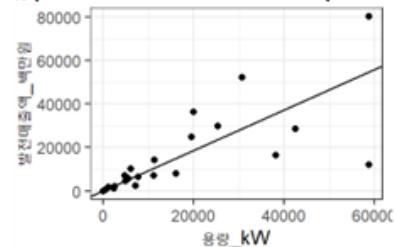
$$\hat{y}_i = 177.515 + 0.256x_i$$



연료전지

연료전지

$$\hat{y}_i = 25.106 + 0.934x_i$$



04

4. 분석 결과 및 결론

(1) 추정 결과

발전업 종사자 수(2018)

(단위 : 명)

에너지원		조사 O	조사 X (추정)	합계
태양광	1MW 미만	314	56,141	56,455
	1MW 이상	134	2,128	2,262
	계	448	58,269	58,717
풍력		157	330	487
수력	1MW 미만	23	24	47
	1MW 이상	44	55	99
	계	67	79	146
바이오에너지		613	771	1,384
폐기물		244	319	563
연료전지		67	105	172
계		1,596	59,873	61,469

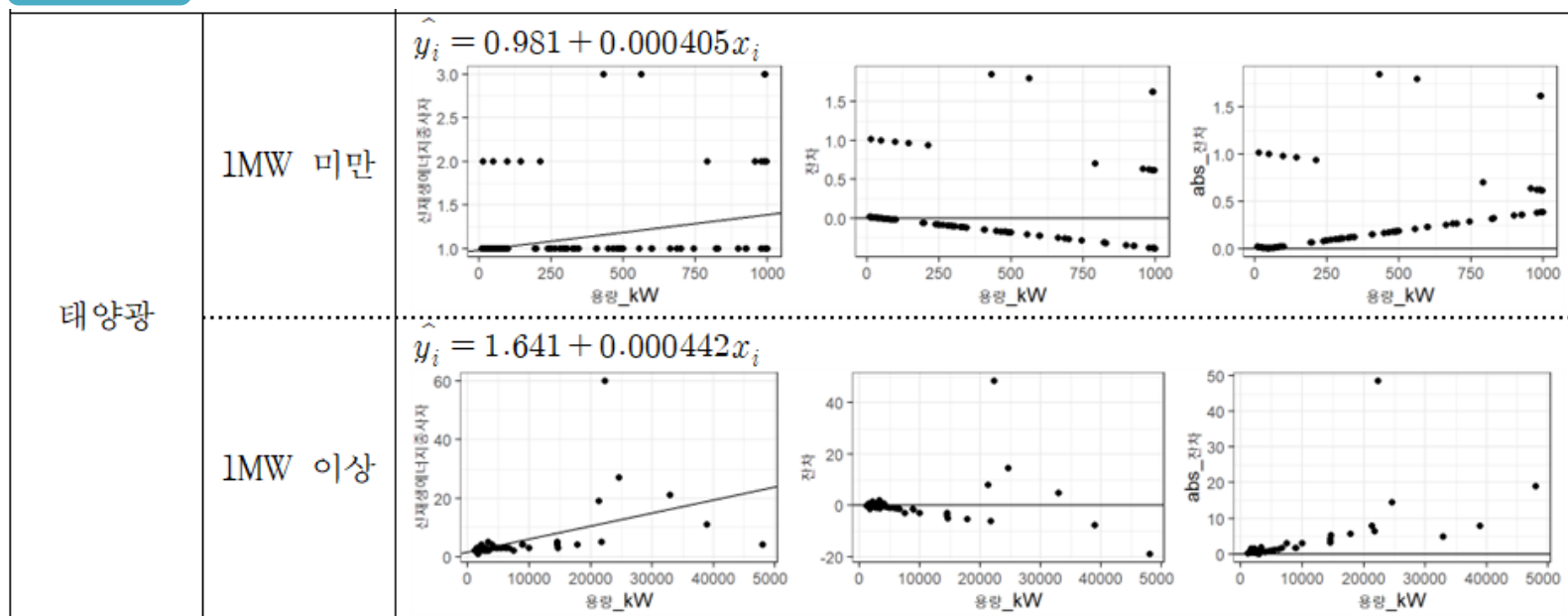
- 단순선형 회귀분석
- 에너지원 : 태양광, 수력 1MW 이상, 바이오, 풍력, 폐기물, 연료전지
- 로버스트 회귀분석
- 에너지원 : 수력 1MW 미만

04

4. 분석 결과 및 결론

(1) 추정 결과 – 종사자 수

태양광

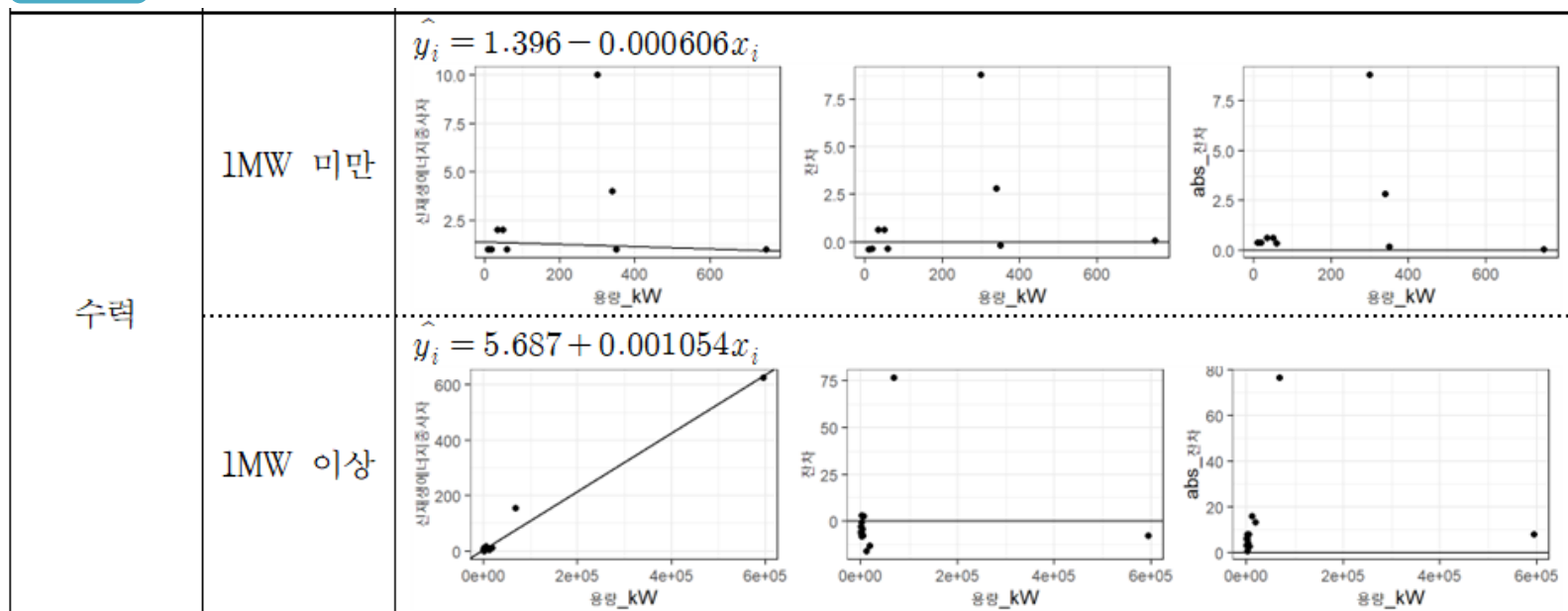


04

4. 분석 결과 및 결론

(1) 추정 결과 – 종사자 수

수력



04

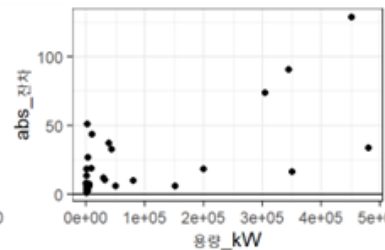
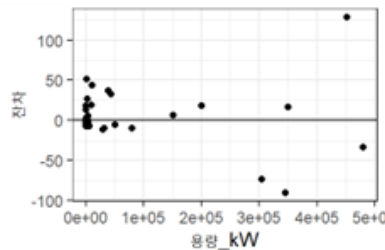
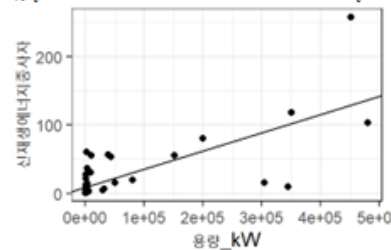
4. 분석 결과 및 결론

(1) 추정 결과 – 종사자 수

바이오에너지

바이오에너지

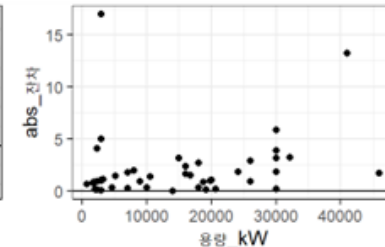
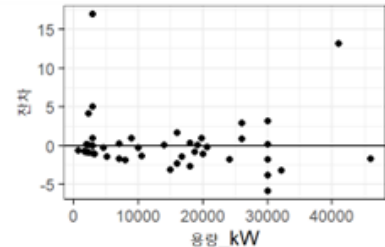
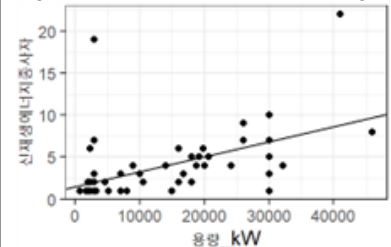
$$\hat{y}_i = 8.676 + 0.000267x_i$$



풍력

풍력

$$\hat{y}_i = 1.496 + 0.000178x_i$$



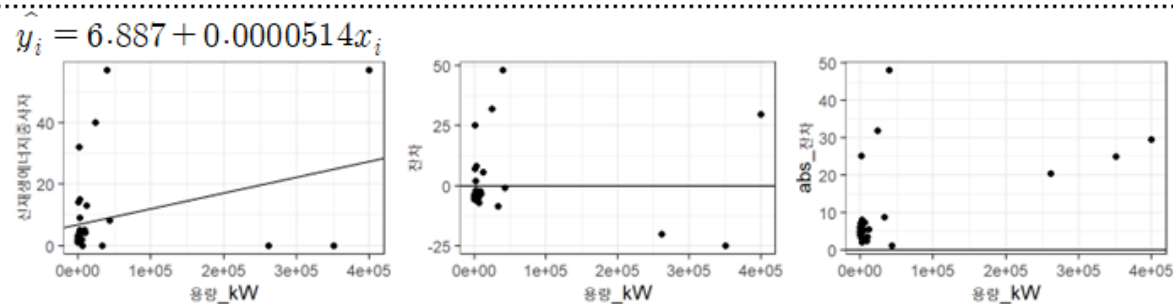
04

4. 분석 결과 및 결론

(1) 추정 결과 – 종사자 수

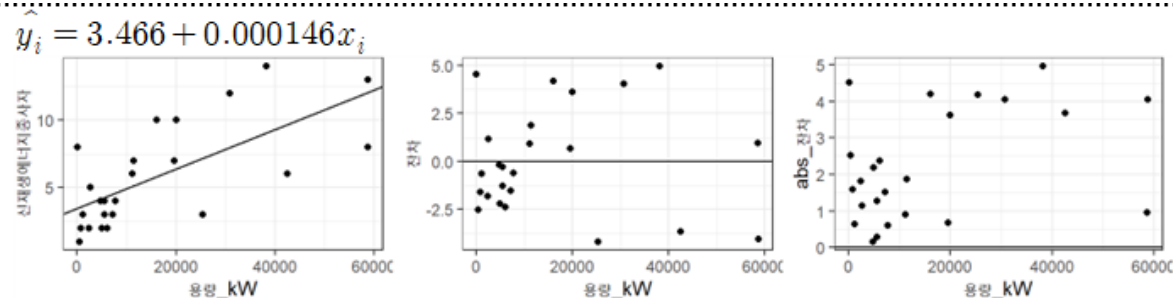
폐기물

폐기물



연료전지

연료전지



04

4. 분석 결과 및 결론

(2) Q & A

“ Q & A ”

THANK
YOU

발 표 자 김 동 현

