Symulacje

Aleksandra Brodecka

October 31, 2016

```
set.seed(123456)
library(ggplot2)
library(gridExtra)
library(grid)
```

Implementacja funkcji, z których korzystam:

1. Funkcja wyliczajaca RMSE:

```
RMSE = function(fit, age){
  sqrt(mean((fit - age)^2))
}
```

2. Funkcja do wyznaczania współczynników β_0 i β_1 metoda momentów:

```
MOM_RC <- function(data){

  W = data$met
  Y = data$age
  m = data$m

  SW = sum(W - mean(W)^2)/(length(W)-1)
  SWY = sum((W - mean(W)) * (Y - mean(Y)))/(length(W)-1)

  sigma_u = sum((W*(1-W)/(m-1))^2)/length(W)

  b1 = SWY/(SW - sigma_u)
  b0 = mean(Y) - b1 * mean(W)

  bety = c(b0, b1)
  names(bety) <- c("b0", "b1")
  bety
}
```

3. Funkcja do generowania danych (metylacji) z rozkladu dwumianowego:

$$Y_i = \beta_0 + \beta_1 * met_i + \epsilon_i$$
$$met_i = Bin(m_i, p_i)/m_i$$

```
gen_bin <- function(b0, b1, n, sig, n1, n2){

  if (n1 == n2) {m = rep(n1, n)
  } else { m = sample(n1:n2, n, replace=T) }
```

```
  p = runif(n)
  Bin = rbinom(n, m, p)
  met = Bin/m
  e = rnorm(n, 0, sig)
  age = b0 + b1*met + e
  data <- data.frame(age, met, m)
  data[data$age > 0 & data$age < 100, ]
  }
```

4. Funkcja do generowania danych (metylacji) za pomoca rozkladu Beta:

$$age \sim runif(0, 80)$$
$$met \sim B(f(age), n)$$
$$n \sim Poiss(\lambda)$$

```
gen_beta <- function(n, lambda, n1, n2, fun){
  age = runif(n, 0, 80)
  if(n1==n2) { m = rep(n1, n)
  } else {m = sample(n1:n2, n, replace=T)}
  met = rbeta(n,  fun(age) , rpois(1, lambda))
  data <- data.frame(age, met, m)
  data[data$age > 0 & data$age < 100 , ]
  }
```

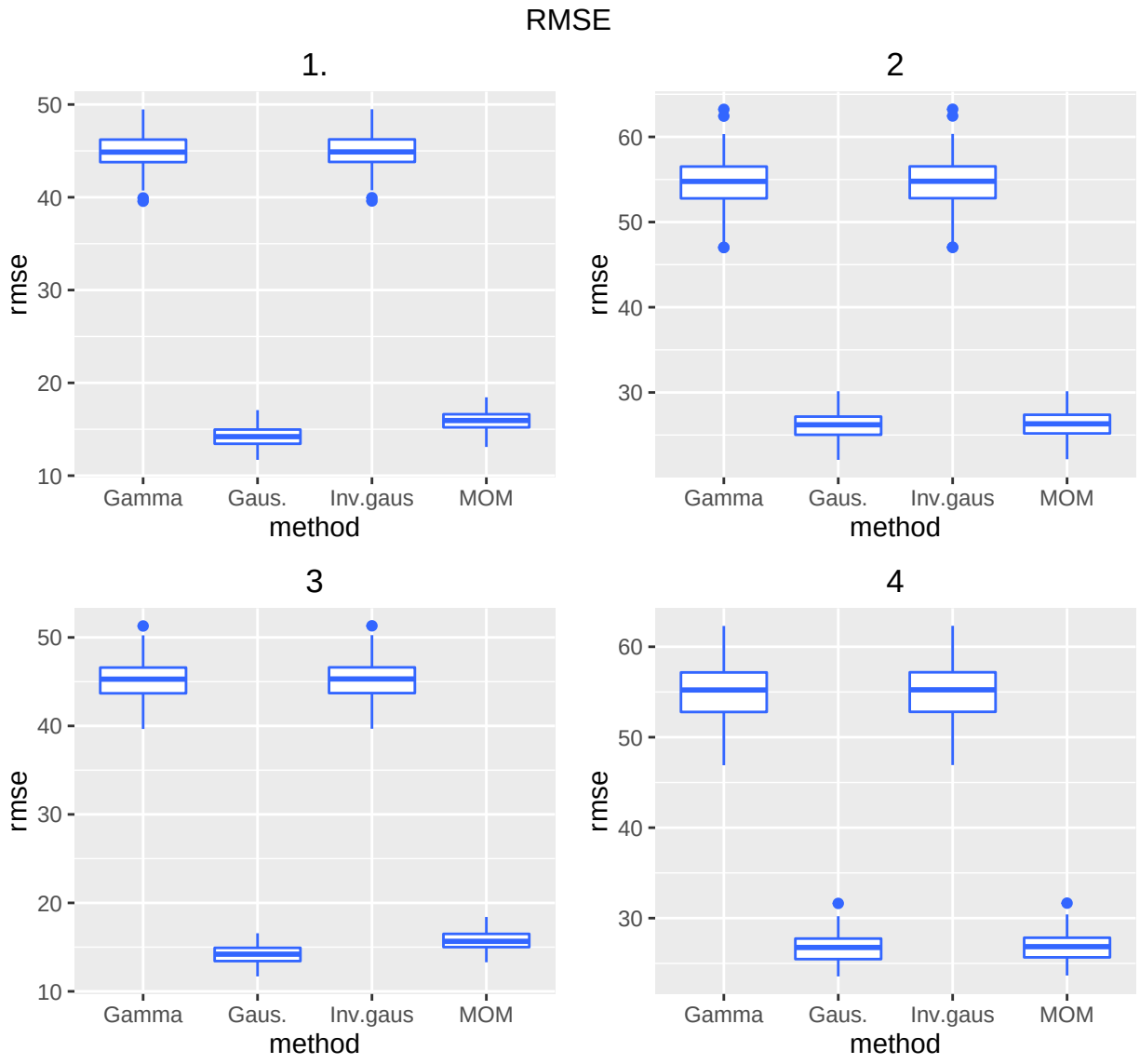
SYMULACJE DLA ROZKLADU DWUMIANOWEGO METYLACJI

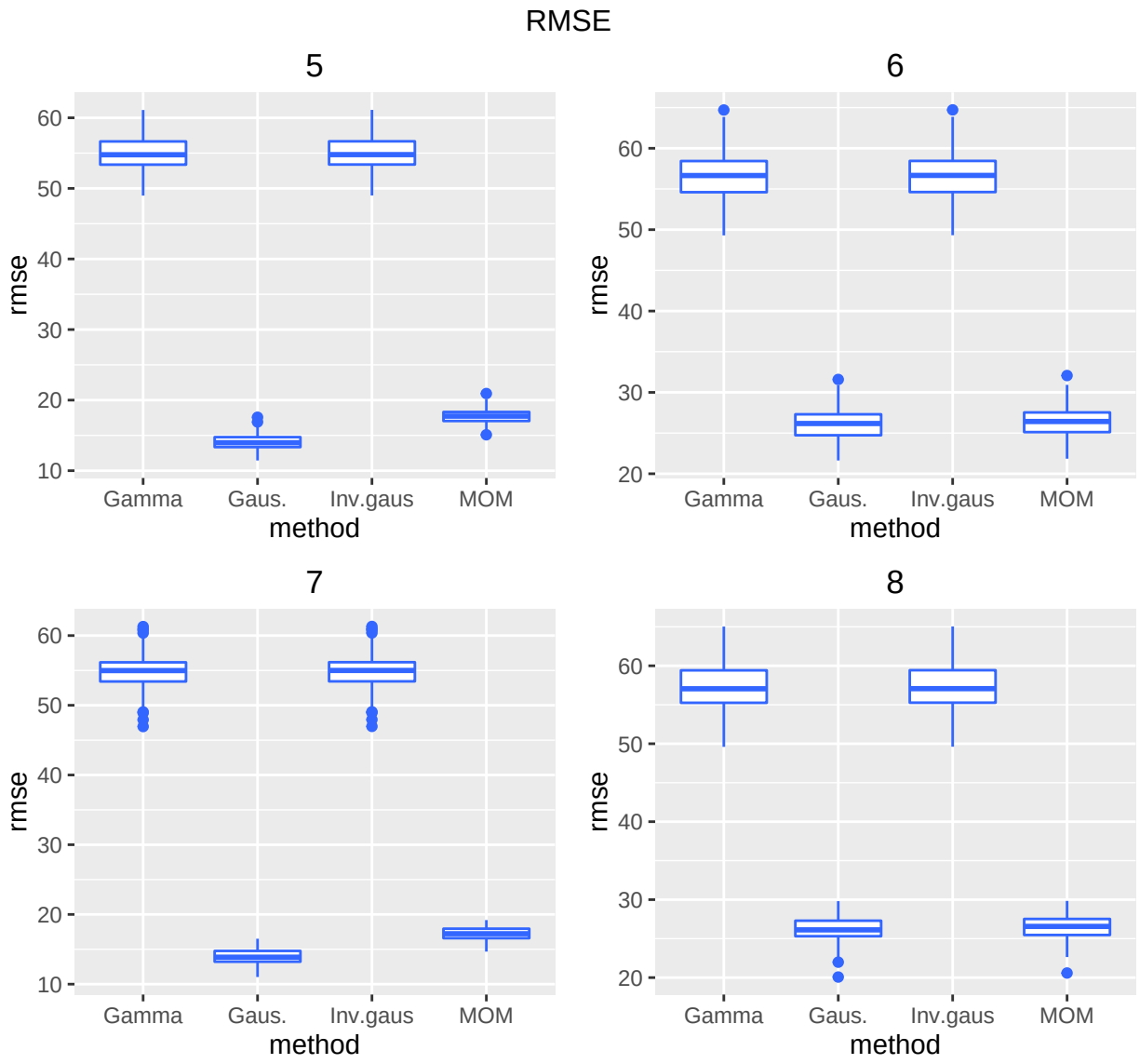
Tworze 200 próbek, kazda o licznosci 100:

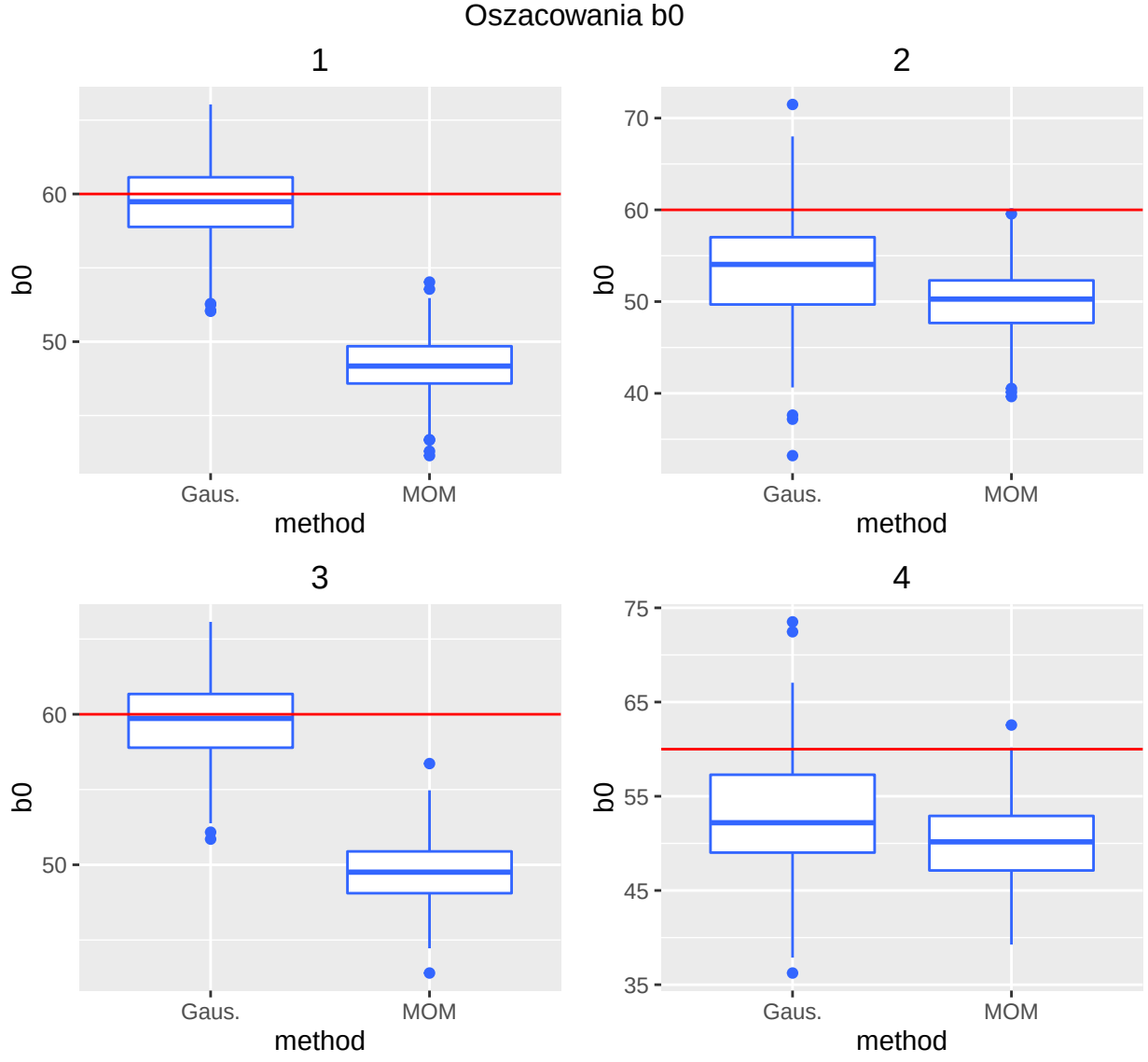
$$Y_i = \beta_0 + \beta_1 * met_i + \epsilon_i$$
$$met_i = Bin(m_i, p_i)/m_i$$

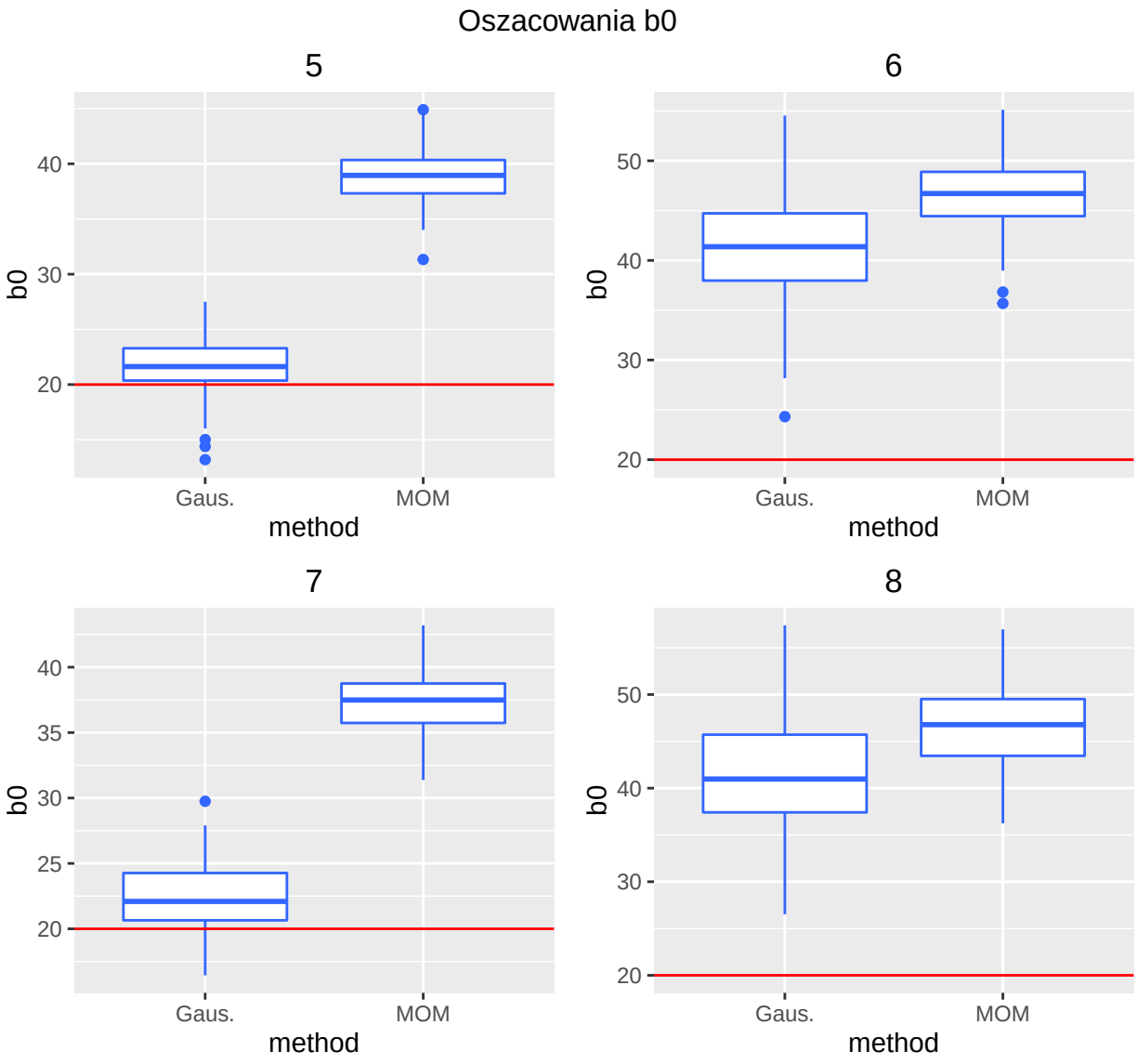
1. $\epsilon \sim N(0, 15), m_i\ sample(5:15), \beta_0 = 60, \beta_1 = -40$
2. $\epsilon \sim N(0, 50), m_i\ sample(5:15), \beta_0 = 60, \beta_1 = -40$
3. $\epsilon \sim N(0, 15), m_i == 5, \beta_0 = 60, \beta_1 = -40$
4. $\epsilon \sim N(0, 50), m_i == 5, \beta_0 = 60, \beta_1 = -40$
5. $\epsilon \sim N(0, 15), m_i\ sample(5:15), \beta_0 = 20, \beta_1 = 60$
6. $\epsilon \sim N(0, 50), m_i\ sample(5:15), \beta_0 = 20, \beta_1 = 60$
7. $\epsilon \sim N(0, 15), m_i == 5, \beta_0 = 20, \beta_1 = 60$
8. $\epsilon \sim N(0, 50), m_i == 5, \beta_0 = 20, \beta_1 = 60$

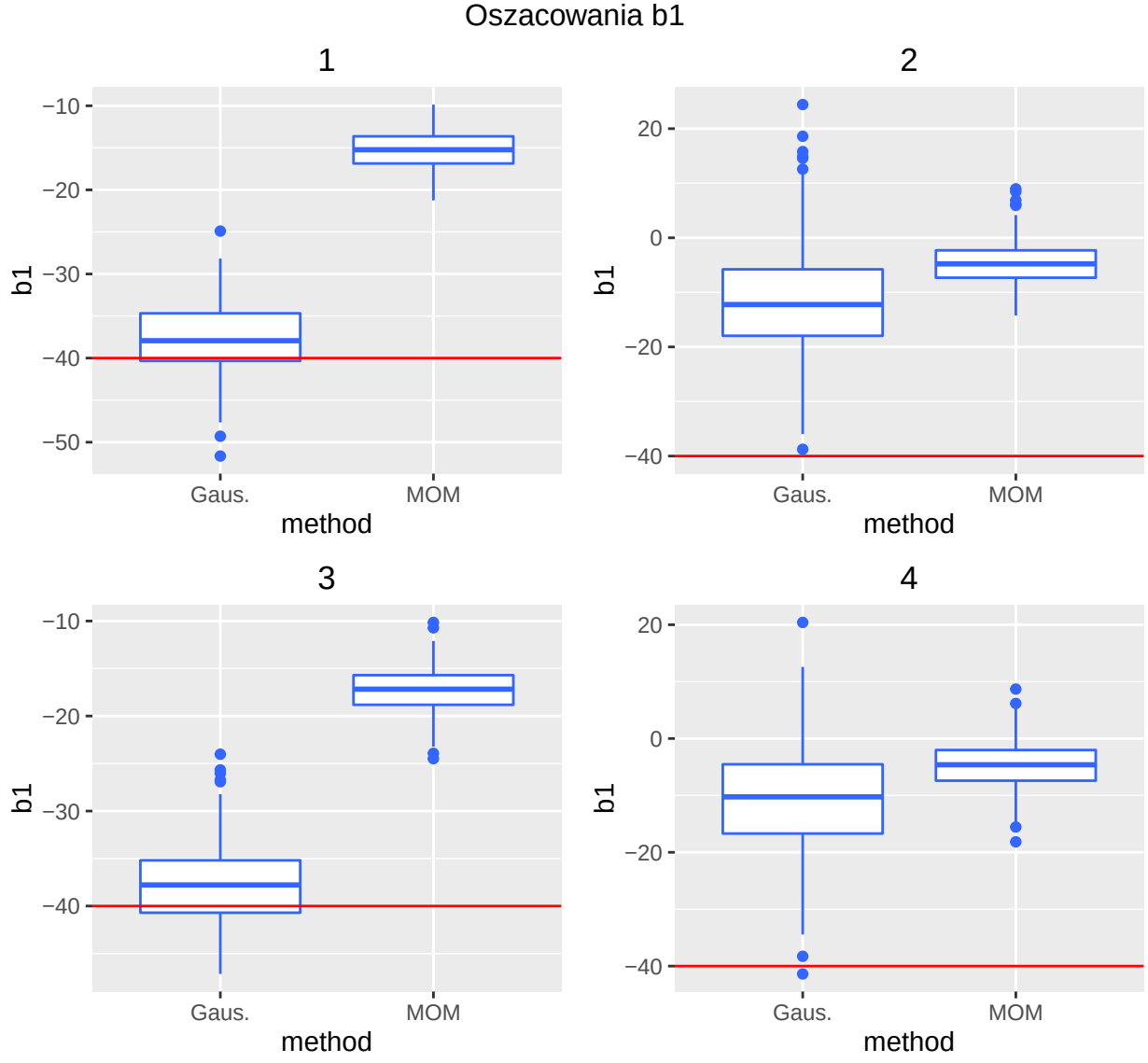
BOX-PLOTY:

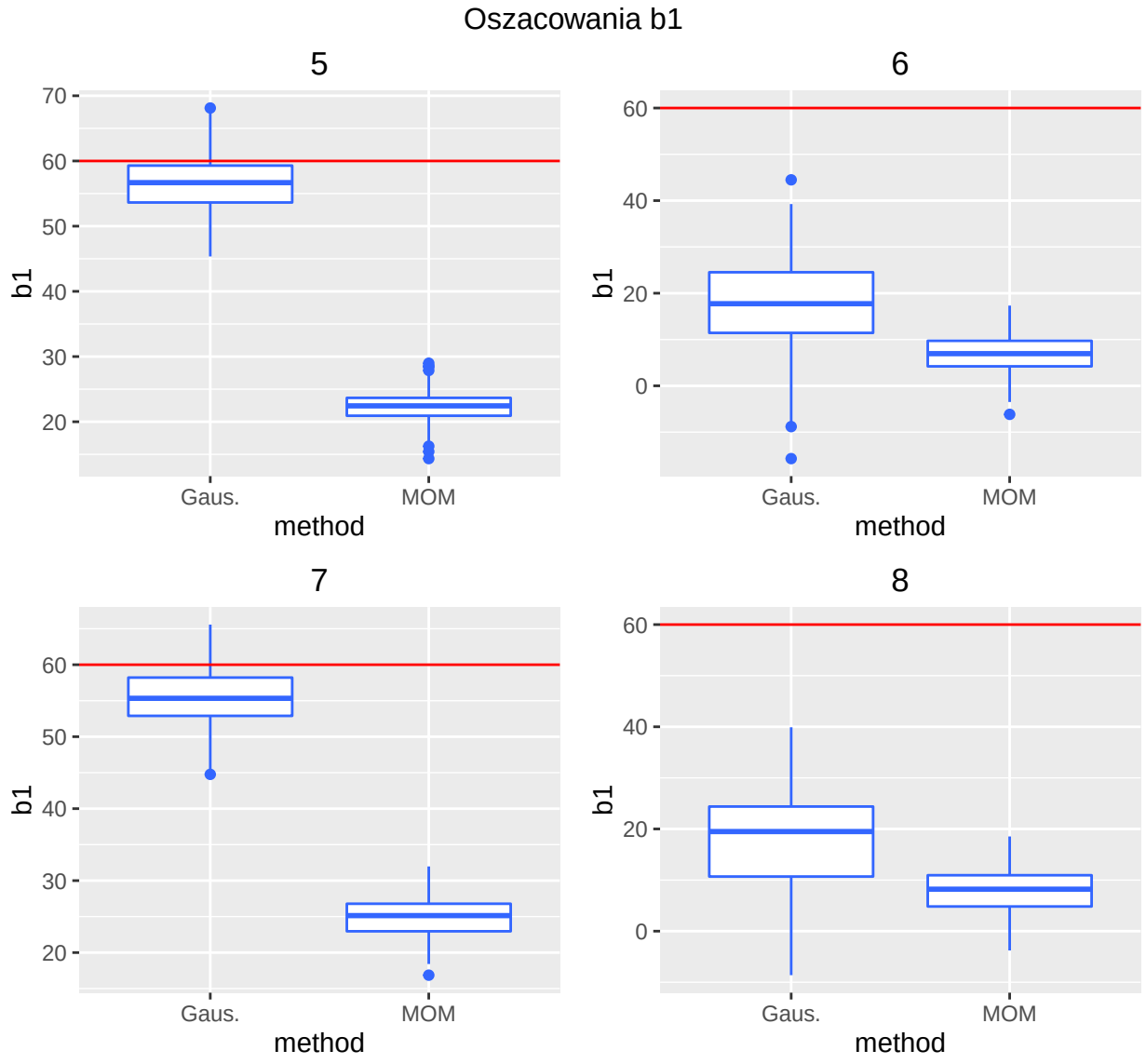



RMSE
5
6
7
8
rmse
method
Gamma
Gaus.
Inv.gaus
MOM
10
20
30
40
50
60

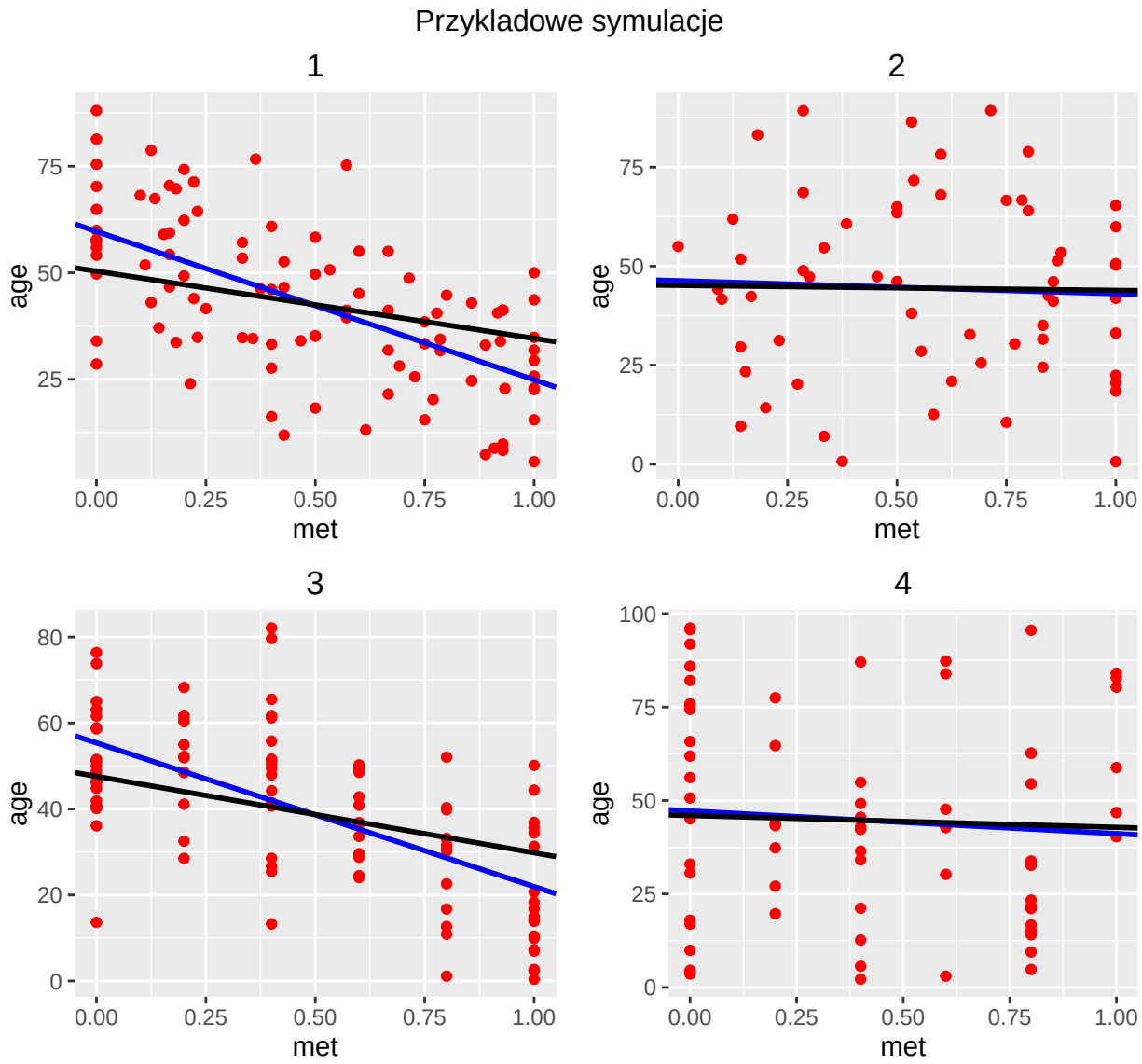

Oszacowania b0
1
2
3
4
b0
method
Gaus.
MOM
60
50
70
40
75
65
55
45
35


Oszacowania b0
5
6
7
8
b0
method
Gaus.
MOM
20
30
40
50
25
35

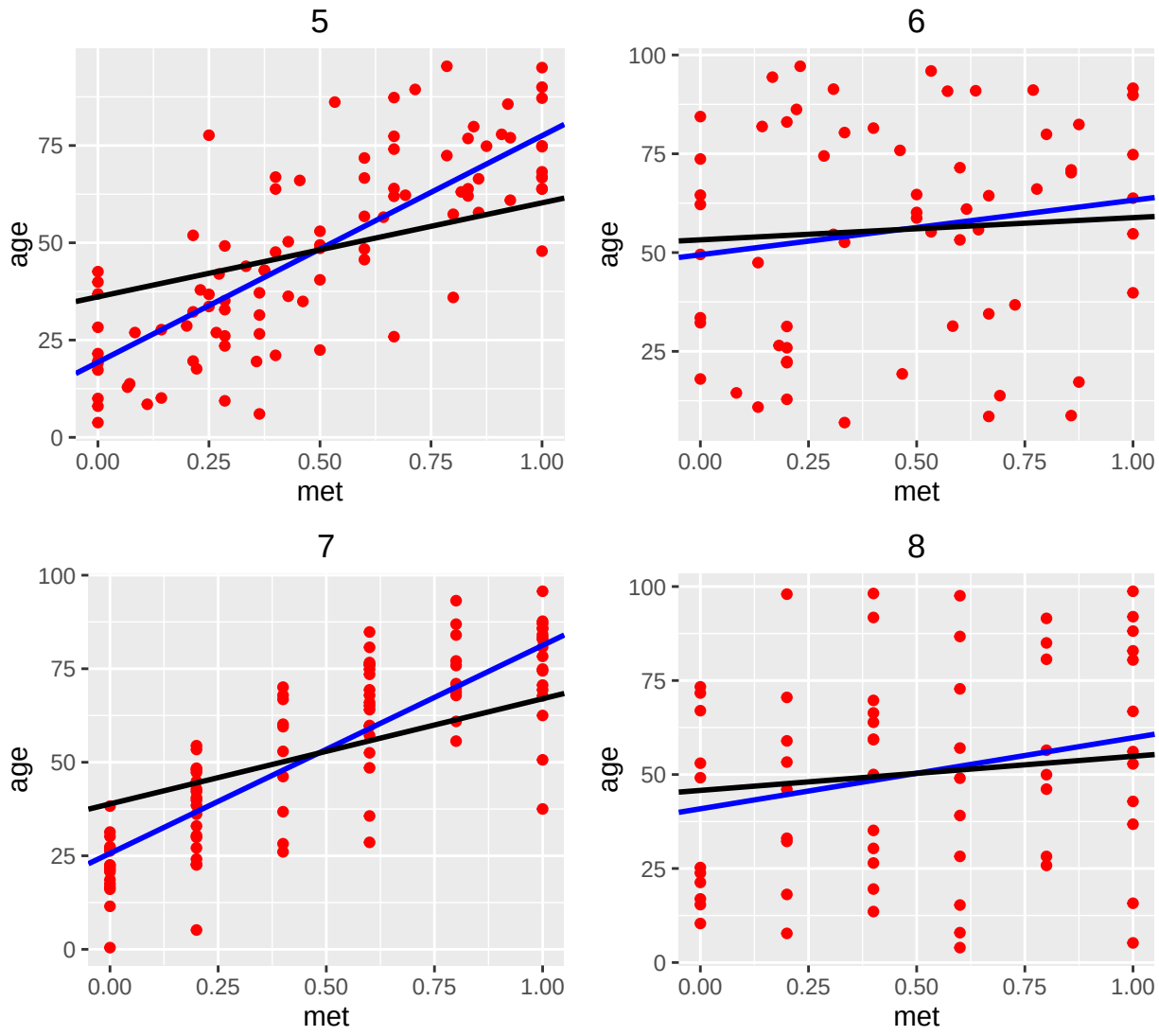

Oszacowania b1
1
2
3
4
b1
method
Gaus.
MOM
−10
−20
−30
−40
−50
20
0


Oszacowania b1
5
6
7
8
b1
method
Gaus.
MOM
70
60
50
40
30
20
0

Postac danych wraz z krzywymi dopasowania:



Przykladowe symulacje

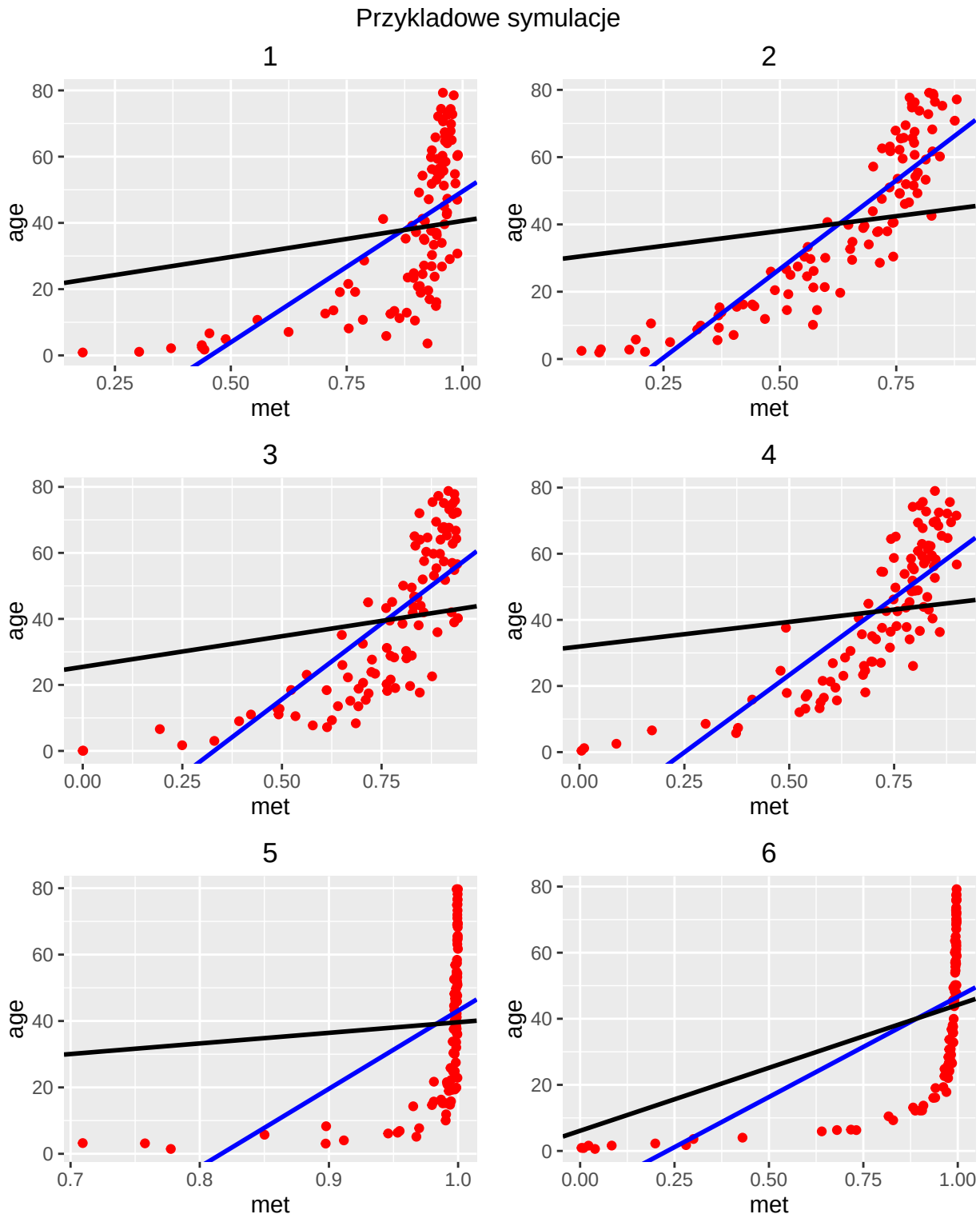


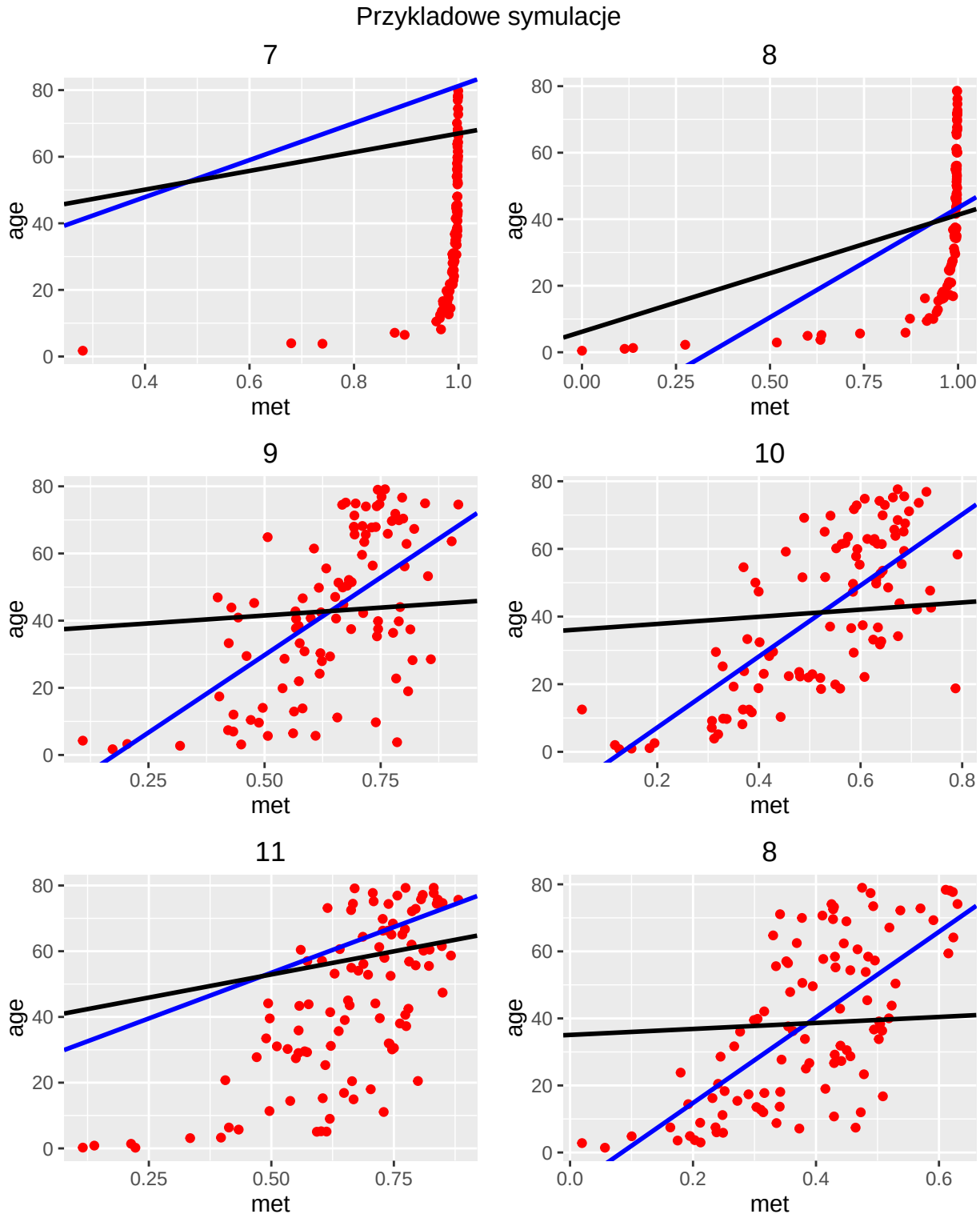
SYMULACJE DLA ROZKLADU BETA

$$age \sim runif(0, 80)$$
$$met \sim B(f(age), n)$$
$$n \sim Poiss(\lambda)$$

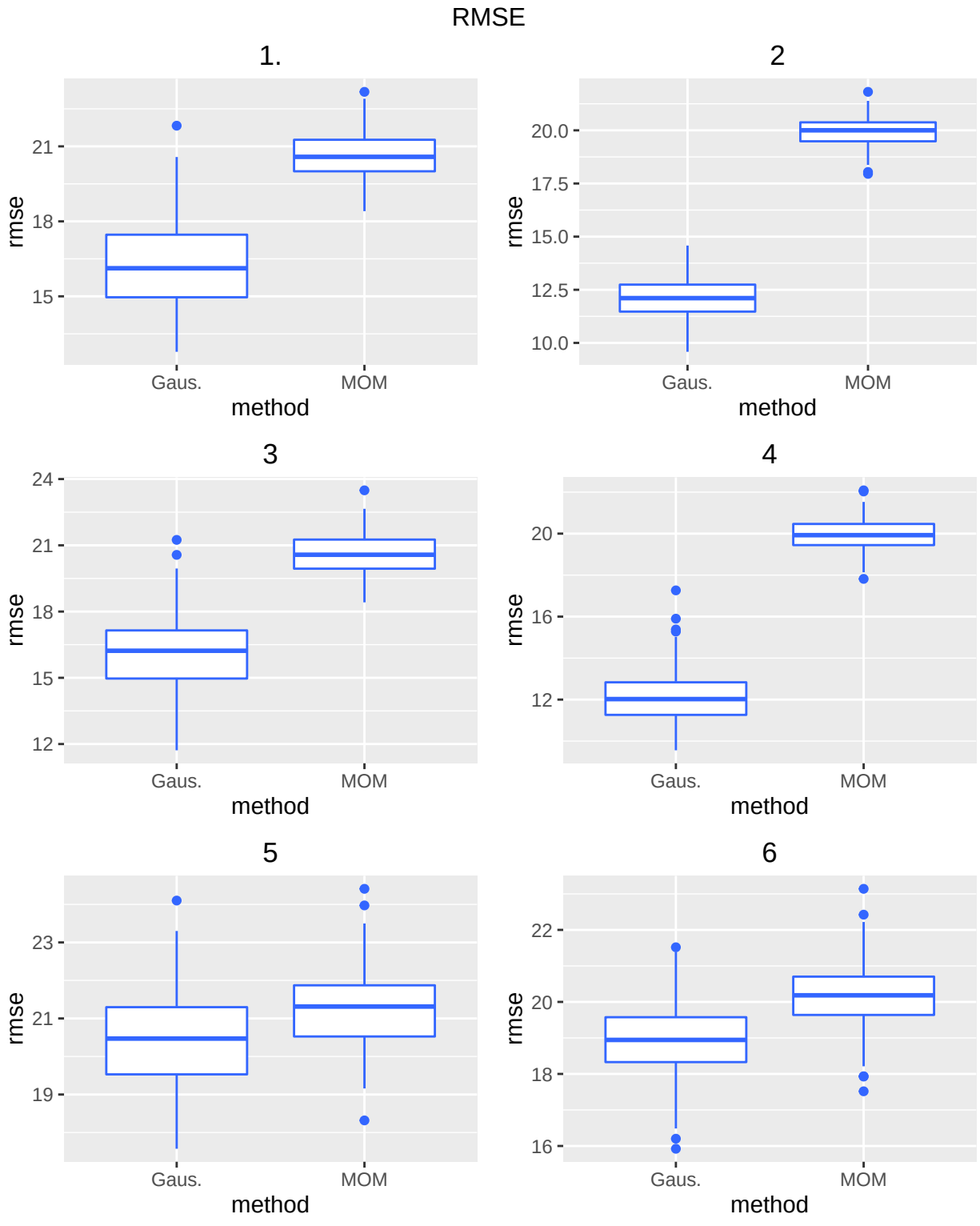
1. $f(age) = age, n \sim Poiss(5), m = sample(5:15)$
2. $f(age) = age, n \sim Poiss(15), m = sample(5:15)$
3. $f(age) = age, n \sim Poiss(5), m == 5$
4. $f(age) = age, n \sim Poiss(15), m == 5$
5. $f(age) = age^2, n \sim Poiss(5), m = sample(5:15)$
6. $f(age) = age^2, n \sim Poiss(15), m = sample(5:15)$
7. $f(age) = age^2, n \sim Poiss(5), m == 5$
8. $f(age) = age^2, n \sim Poiss(15), m == 5$
9. $f(age) = 2sqrt(age), n \sim Poiss(5), m = sample(5:15)$
10. $f(age) = 2sqrt(age), n \sim Poiss(15), m = sample(5:15)$
11. $f(age) = 2sqrt(age), n \sim Poiss(5), m == 5$
12. $f(age) = 2sqrt(age), n \sim Poiss(15), m == 5$

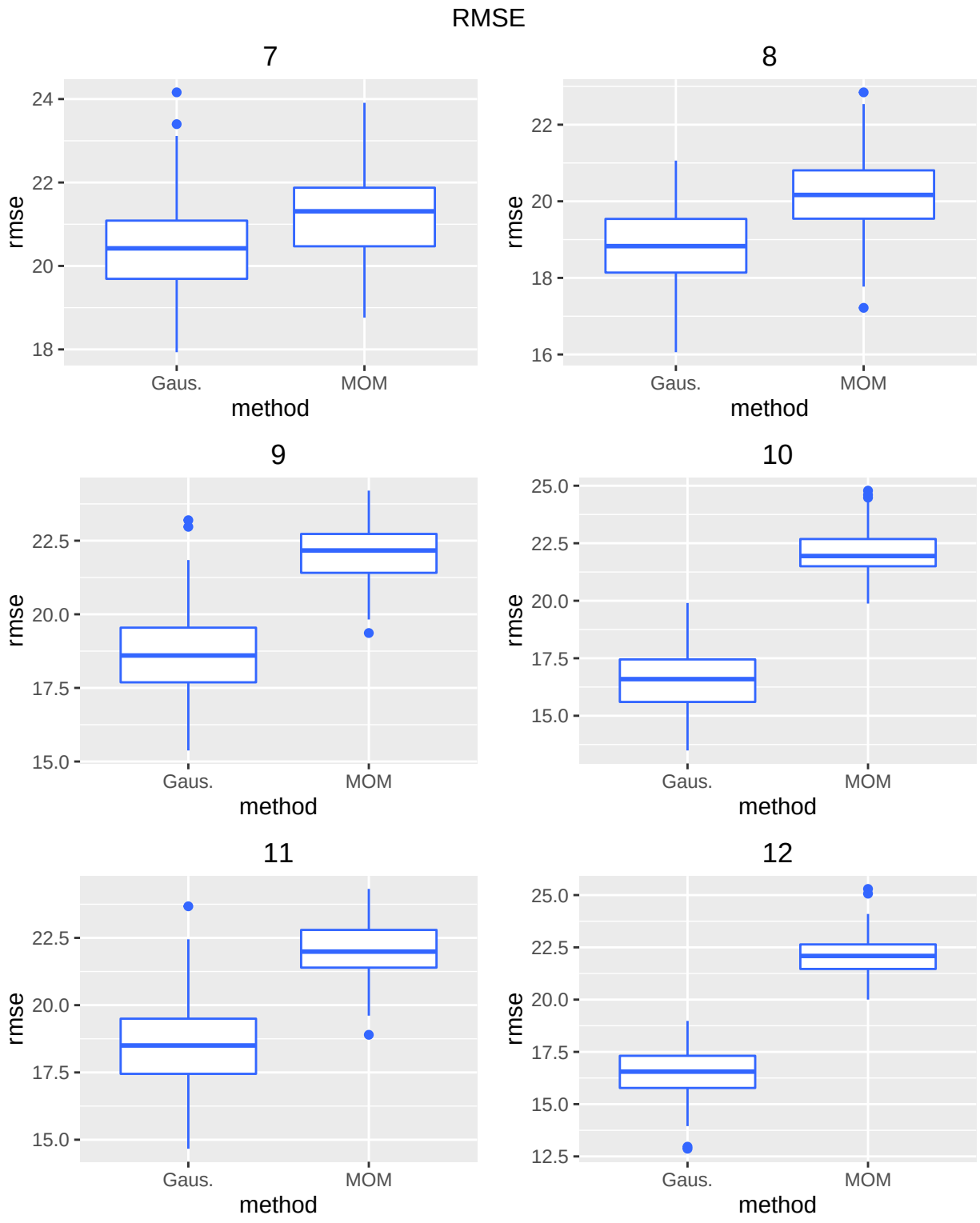
Postac danych wraz z krzywymi dopasowania:

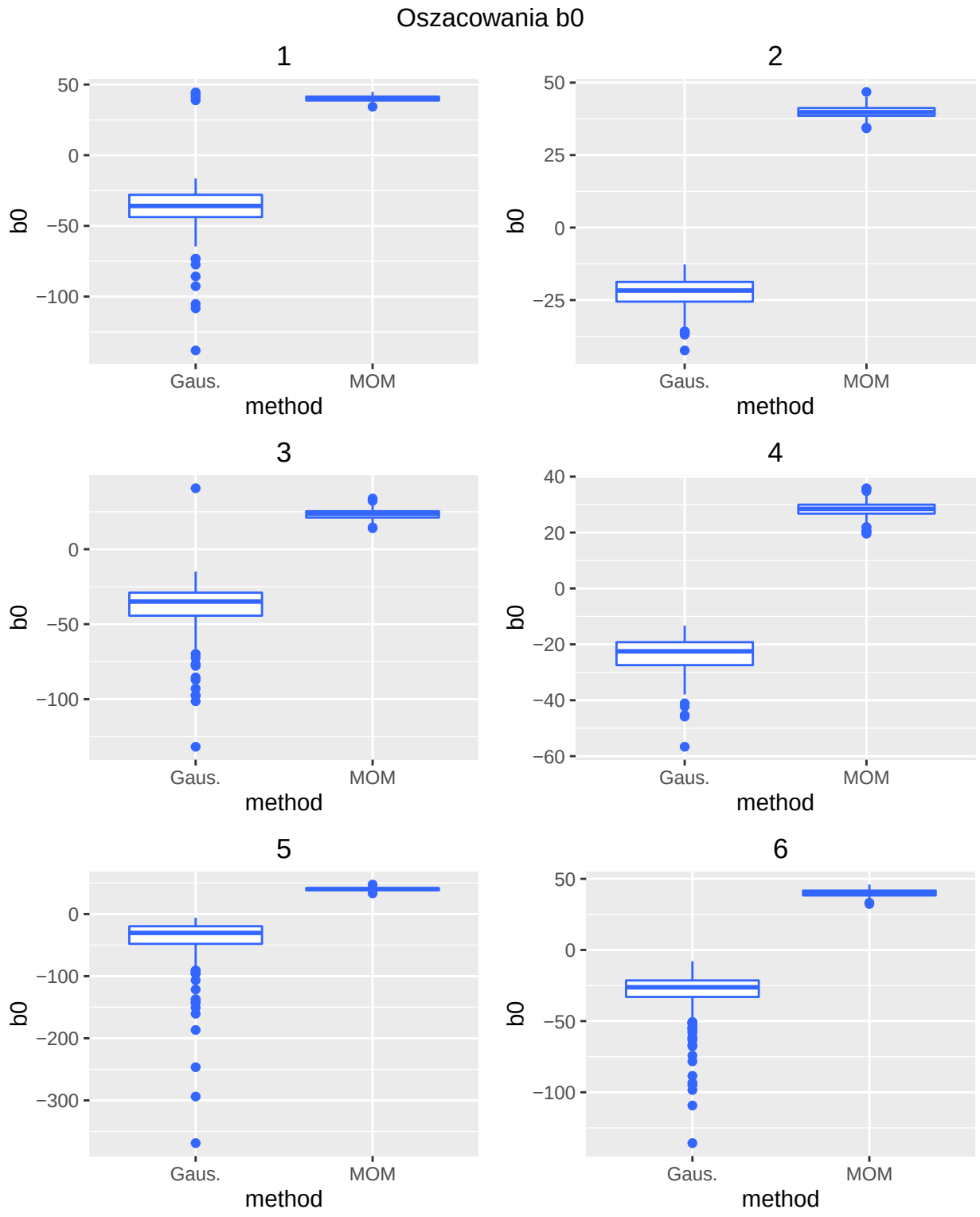


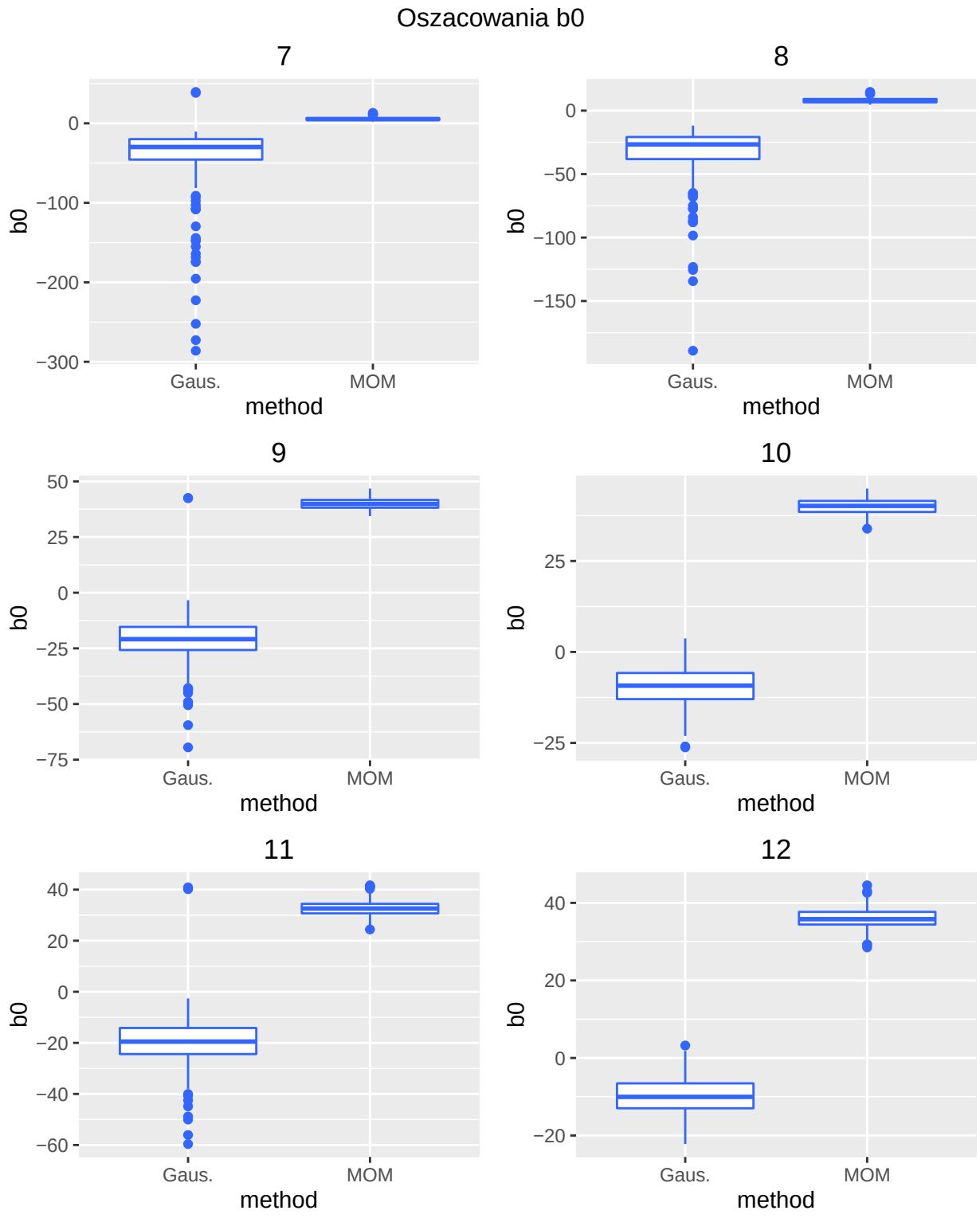

Przykladowe symulacje
7
8
9
10
11
8
age
met

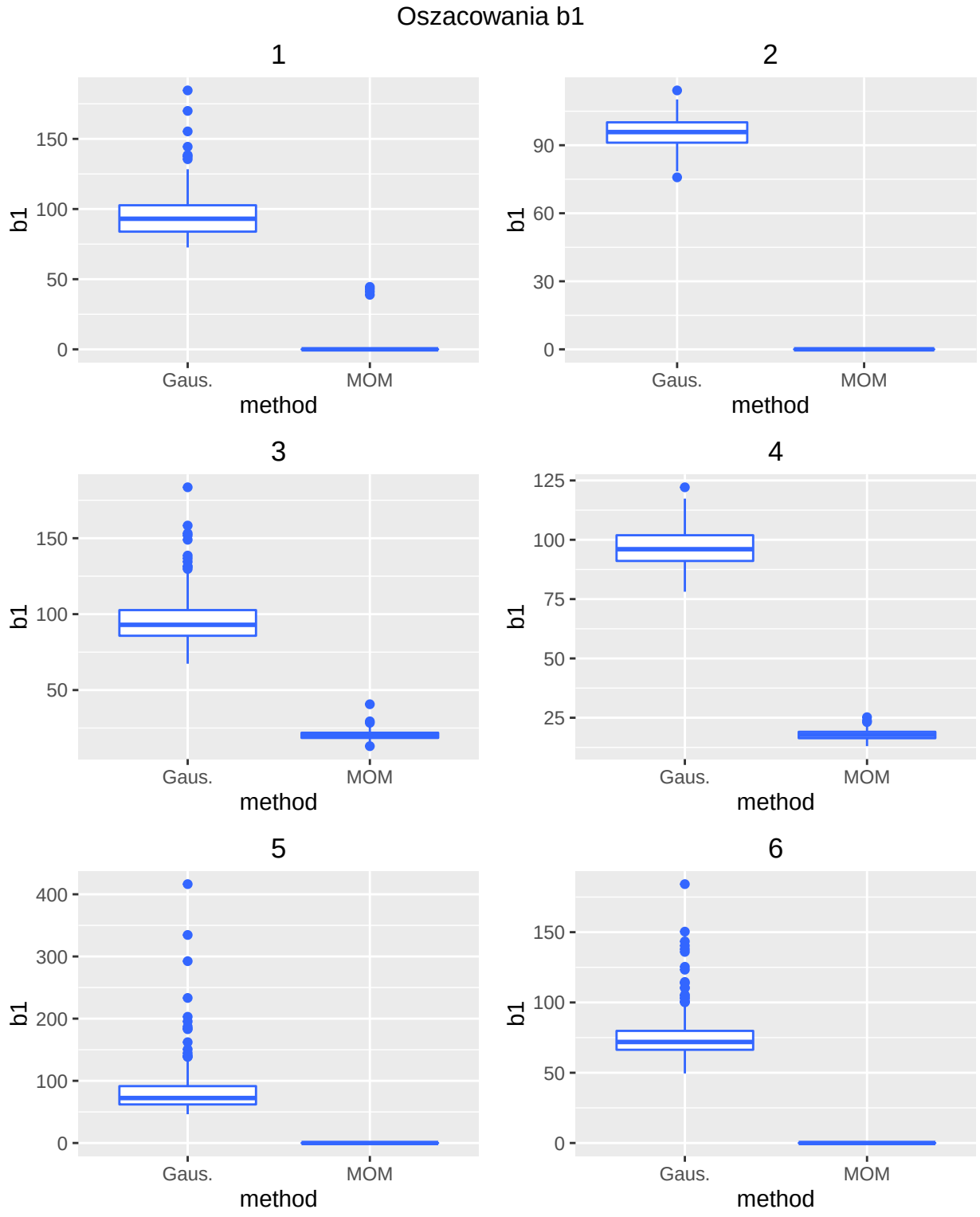
SYMYLACJE:

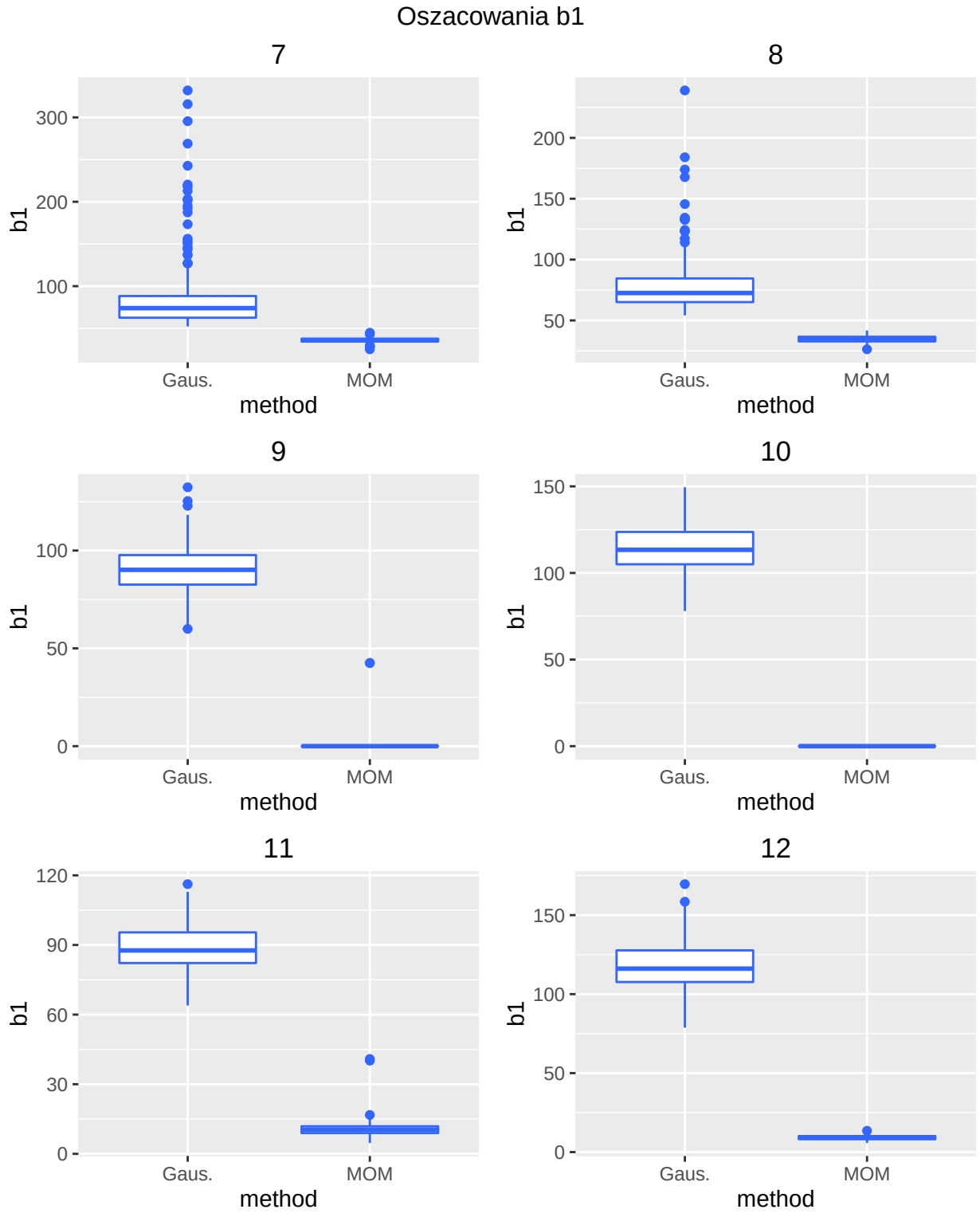



RMSE
7
8
9
10
11
12
rmse
method
Gaus.
MOM
24
22
20
18
16
22.5
20.0
17.5
15.0
25.0
12.5


Oszacowania b0
1
2
3
4
5
6
b0
method
Gaus.
MOM
50
0
−50
−100
25
−25
40
20
−20
−40
−60
−200
−300


Oszacowania b0
7
8
9
10
11
12
b0
method
Gaus.
MOM
0
−100
−200
−300
−50
−150
50
25
−25
−75
20
−20
−40
−60
40


Oszacowania b1
1
2
3
4
5
6
b1
method
Gaus.
MOM
0
50
100
150
30
60
90
25
75
125
200
300
400


Oszacowania b1
7
8
9
10
11
12
b1
method
Gaus.
MOM
300
200
100
150
50
120
90
60
30
0