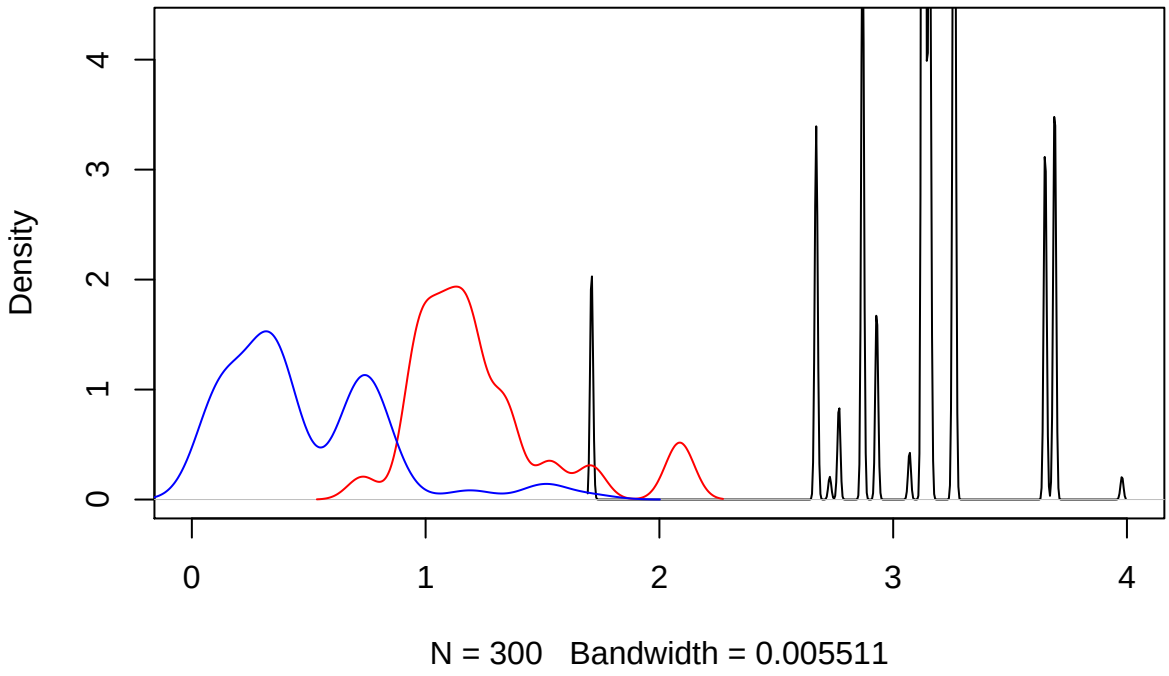
Poisson Change-point Application

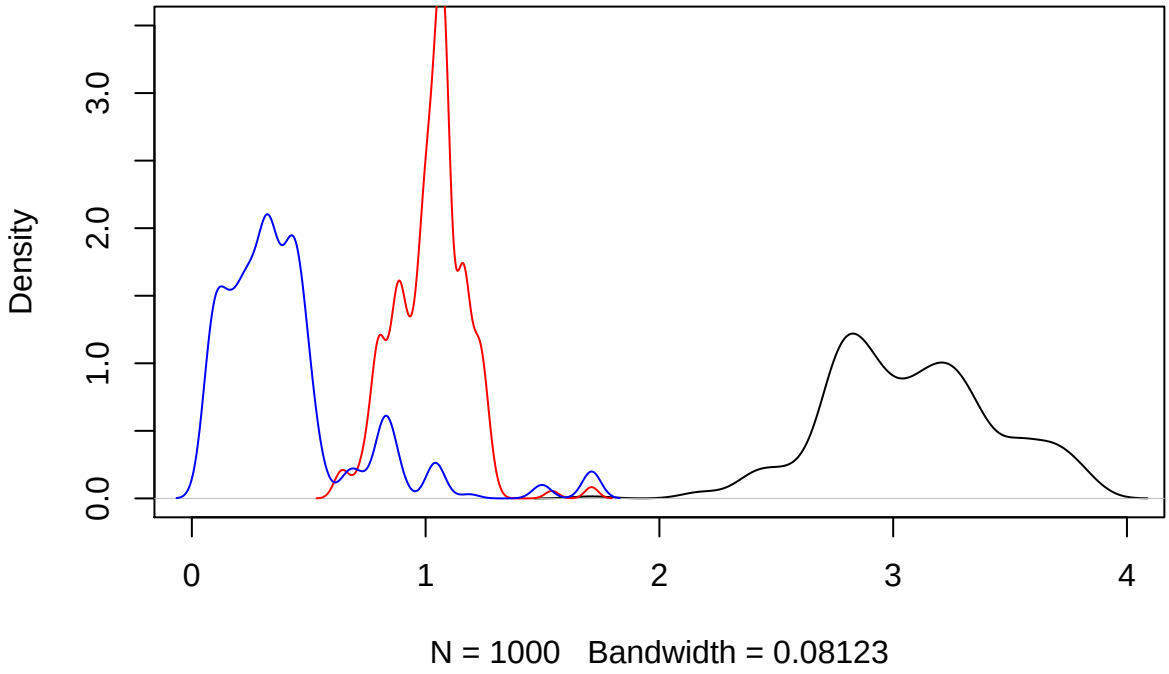
Aofei Liu

1) Fixed Change-point K=3

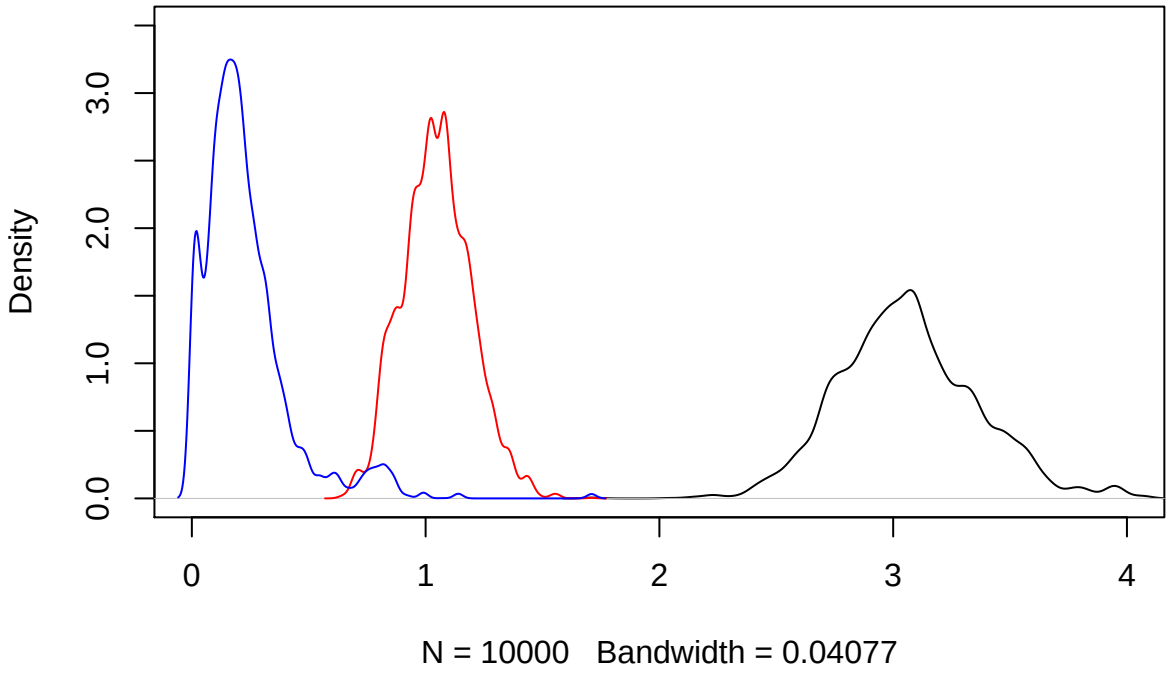
Density of Height with Running Length (N) = 300



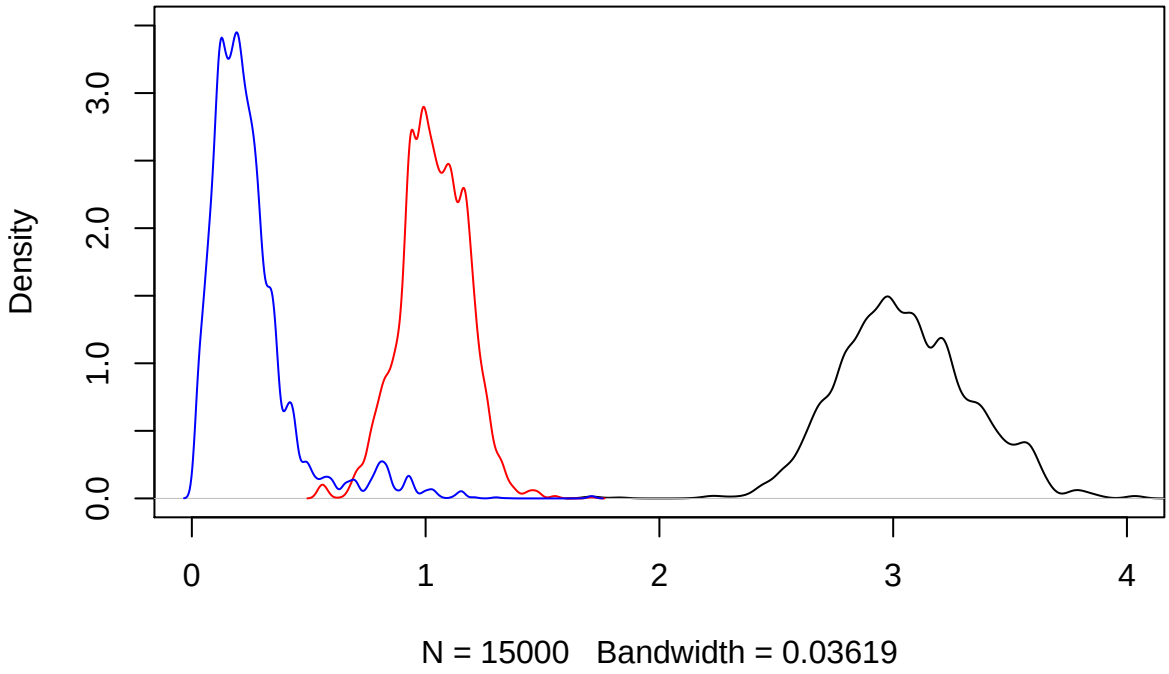
Density of Height with Running Length (N) = 1000



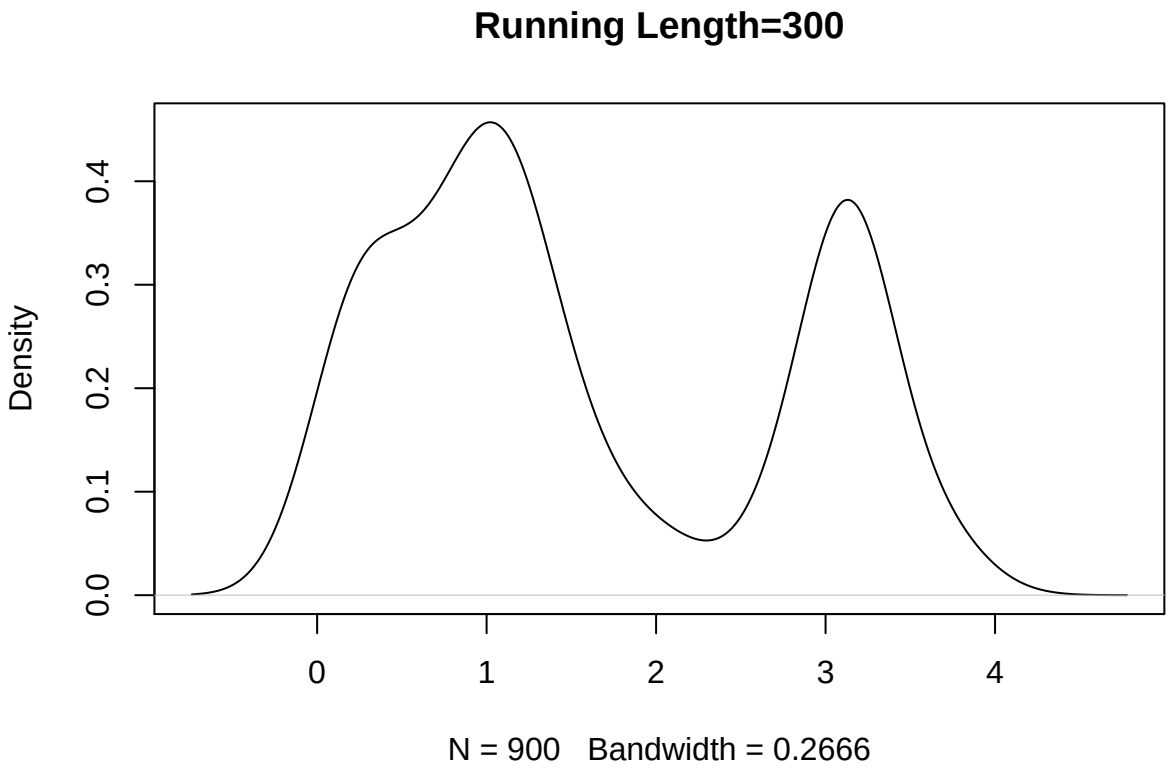
Density of Height with Running Length (N) = 10000

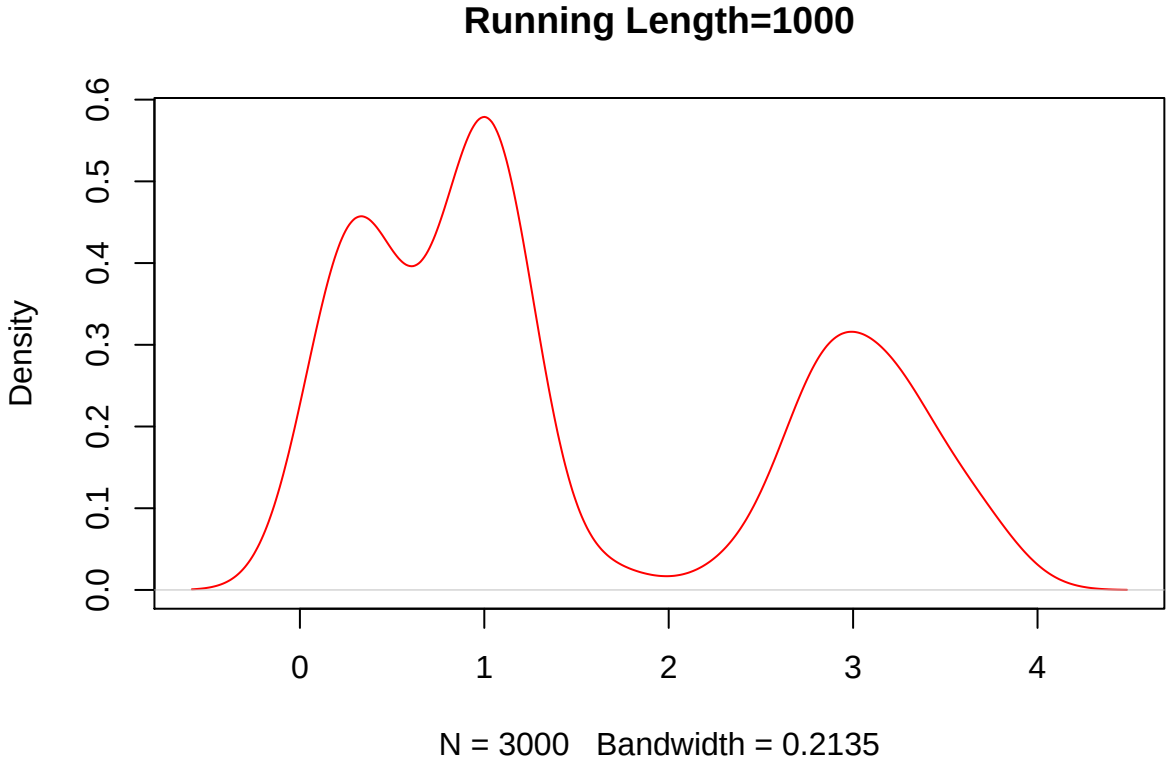


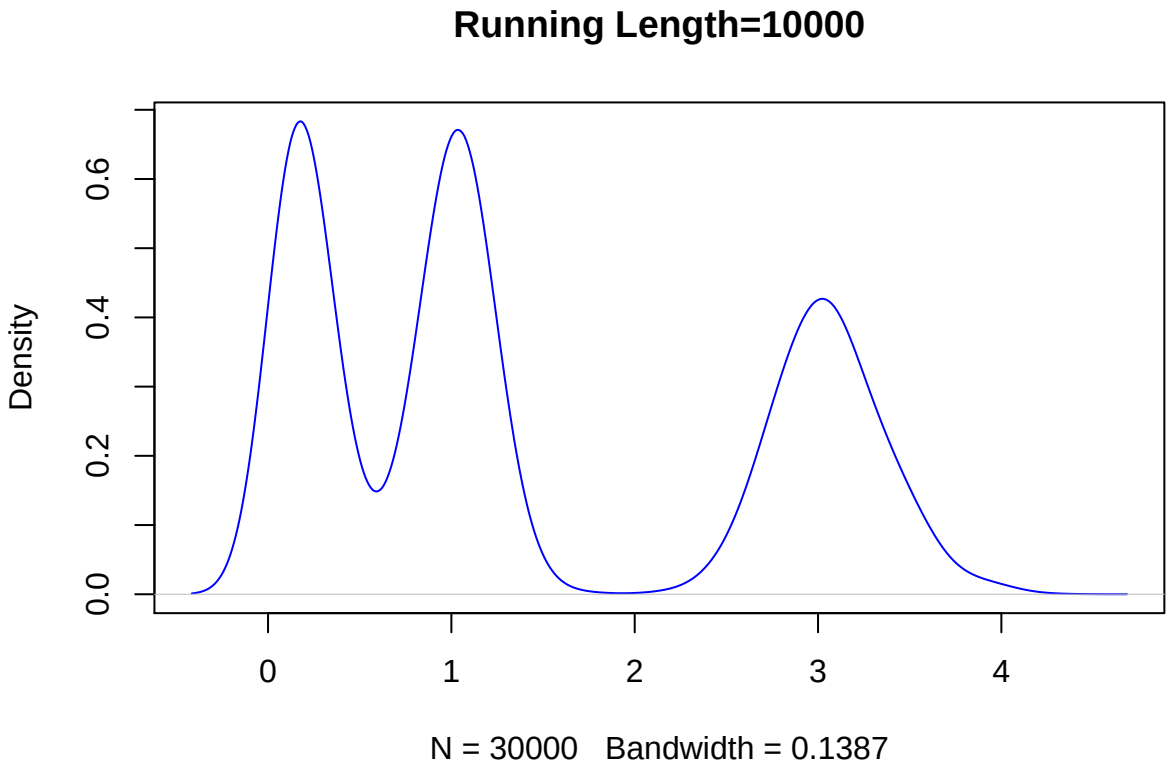
Density of Height with Running Length (N) = 15000

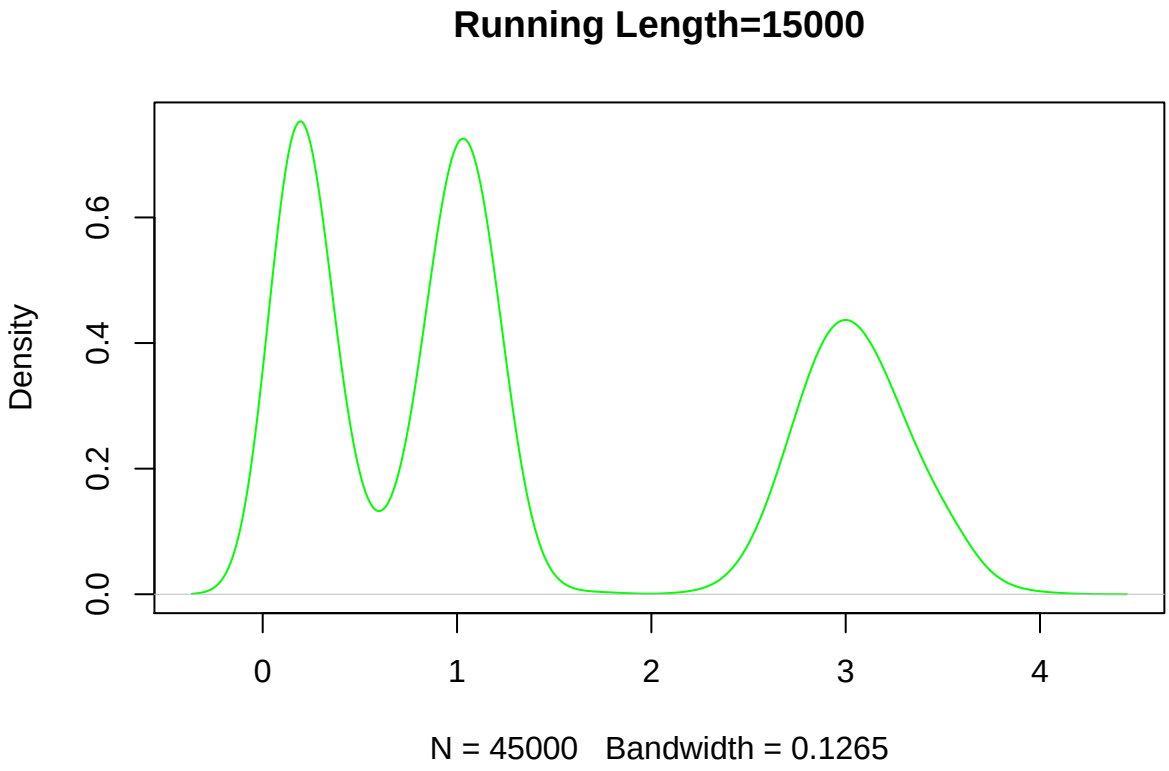


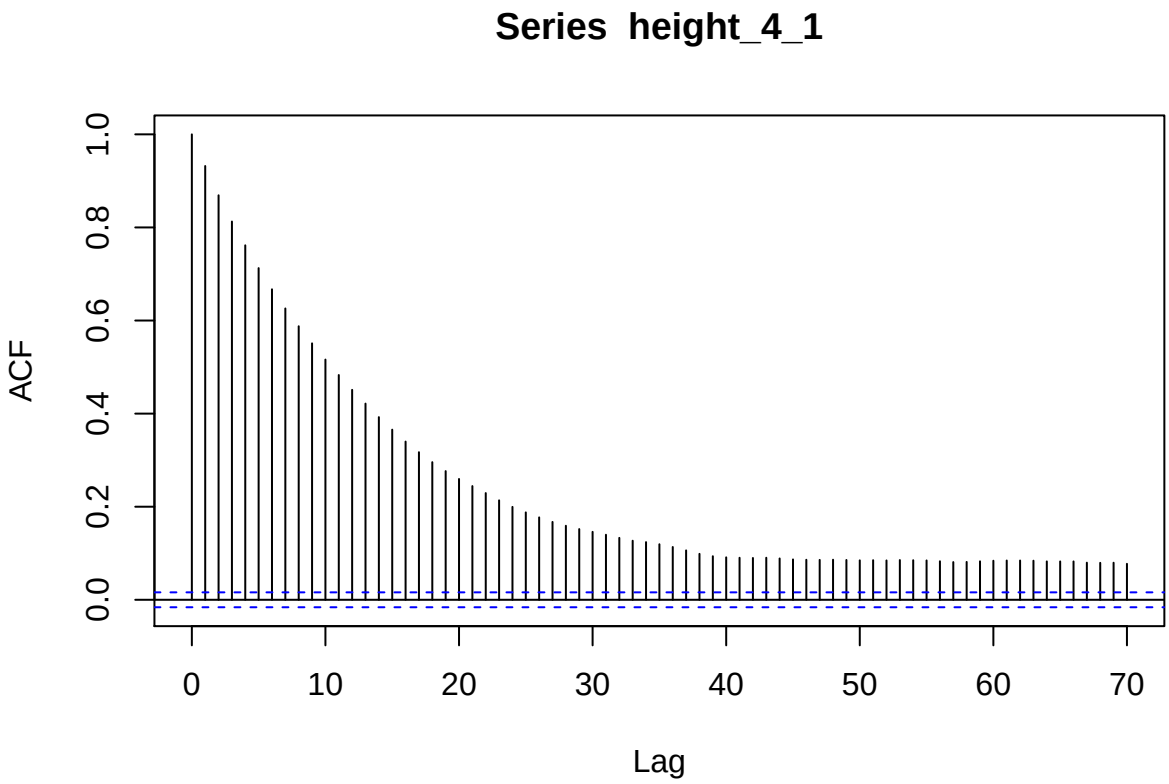
Comparsion of density of height

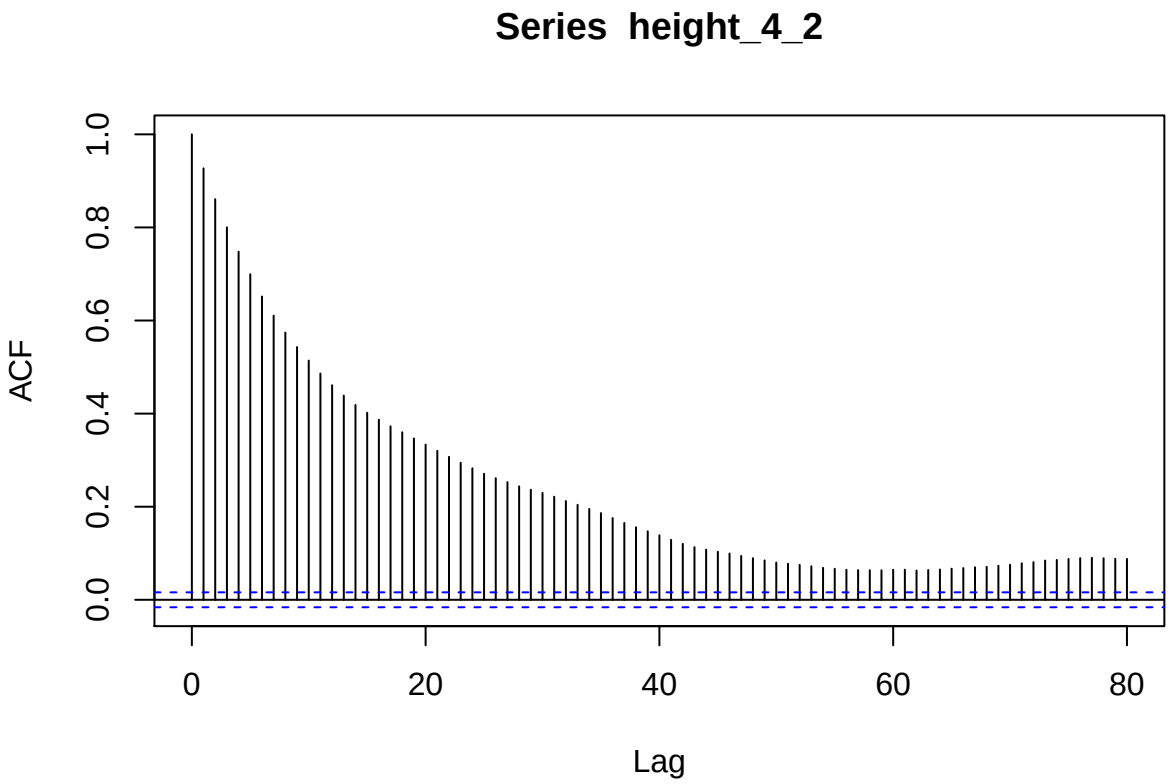


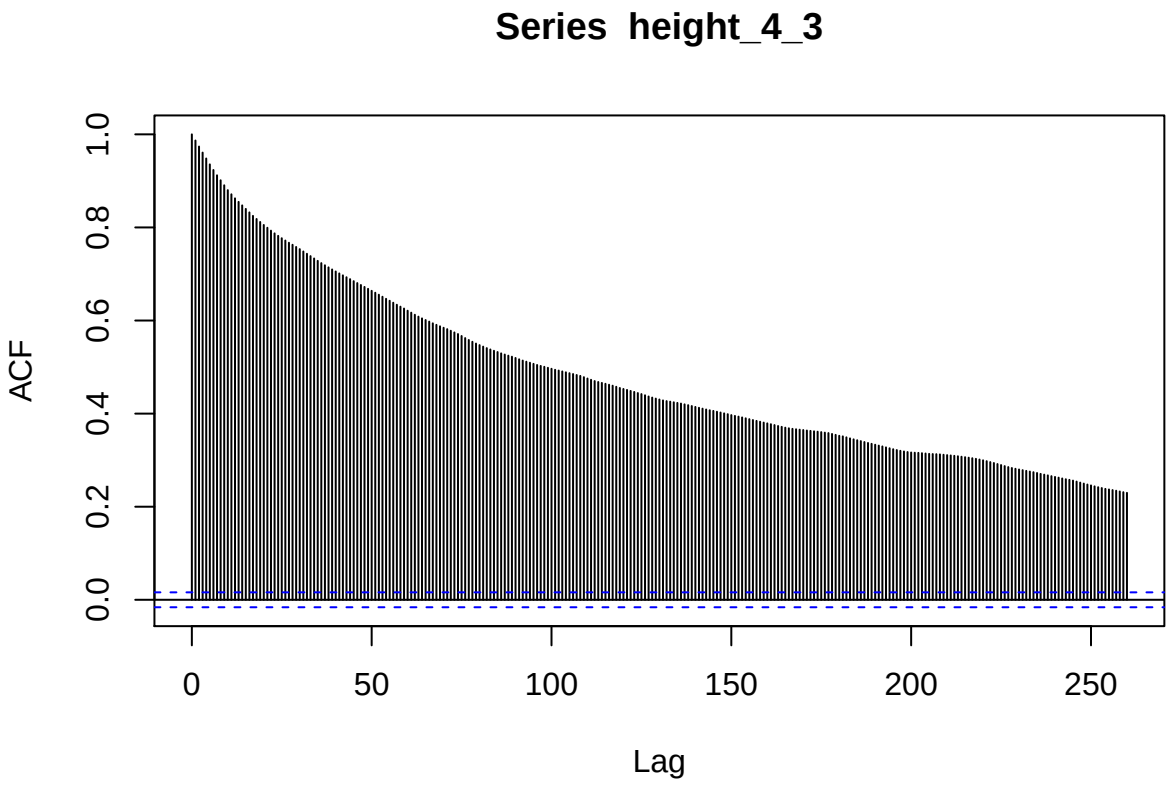



Running Length=10000
Density
0.0
0.2
0.4
0.6
0
1
2
3
4
N = 30000 Bandwidth = 0.1387


Running Length=15000
Density
0.0
0.2
0.4
0.6
0
1
2
3
4
N = 45000 Bandwidth = 0.1265


Series height_4_1
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
70
Lag

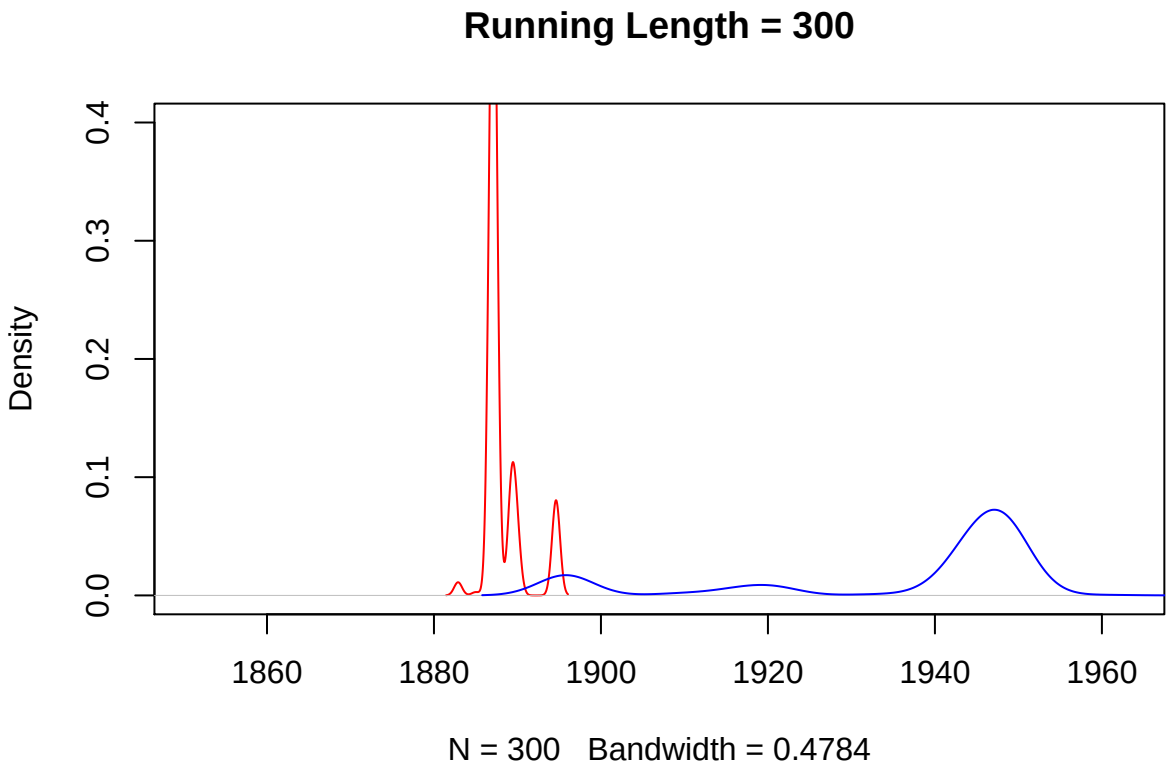

Series height_4_2
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
Lag



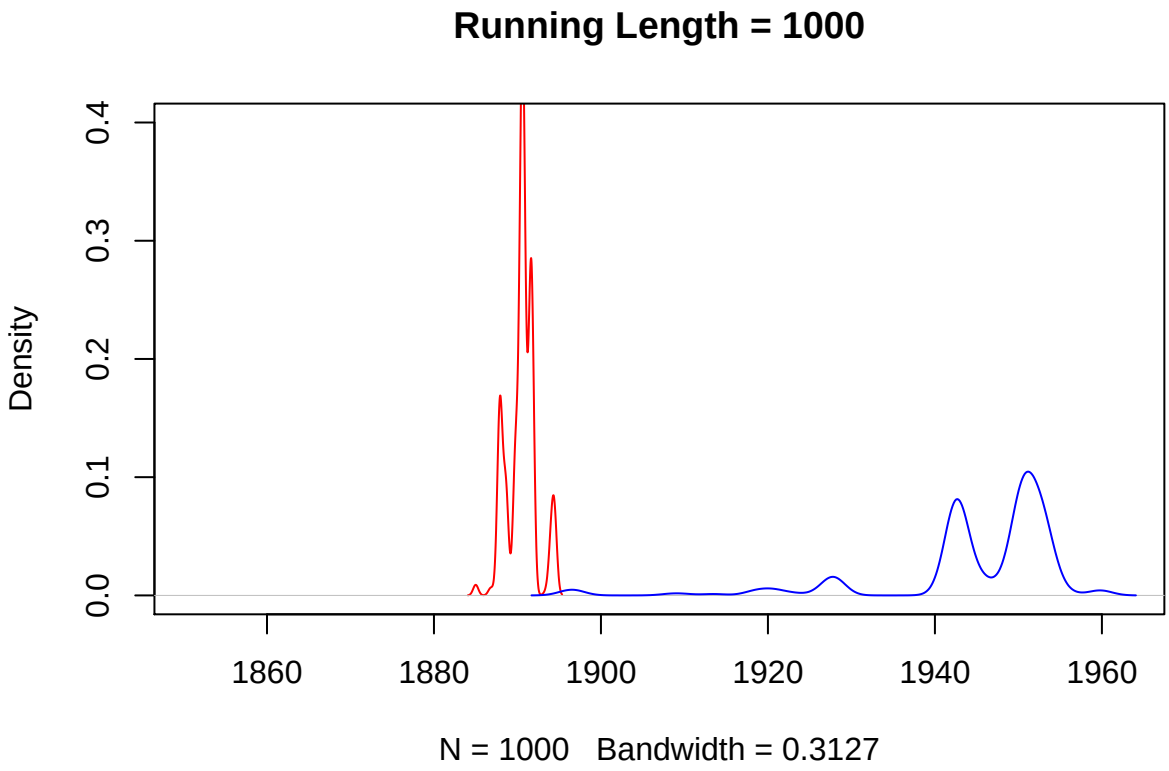
Result:

For $N < 10000$, the modes are not very clear. Thus, running length is too short if it is less than 10000. Besides, the modes tends to be clearer as running length reachs and beyonds 10000.

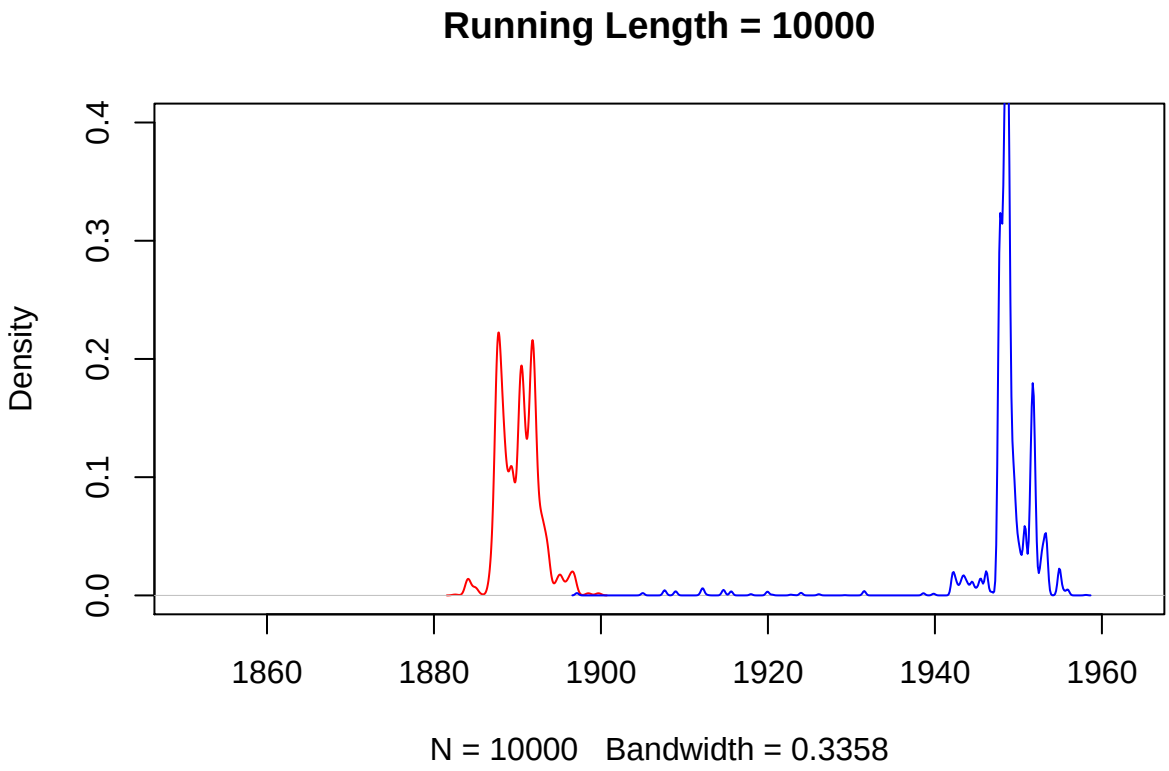
Density of Position with Running Length(N) = 300



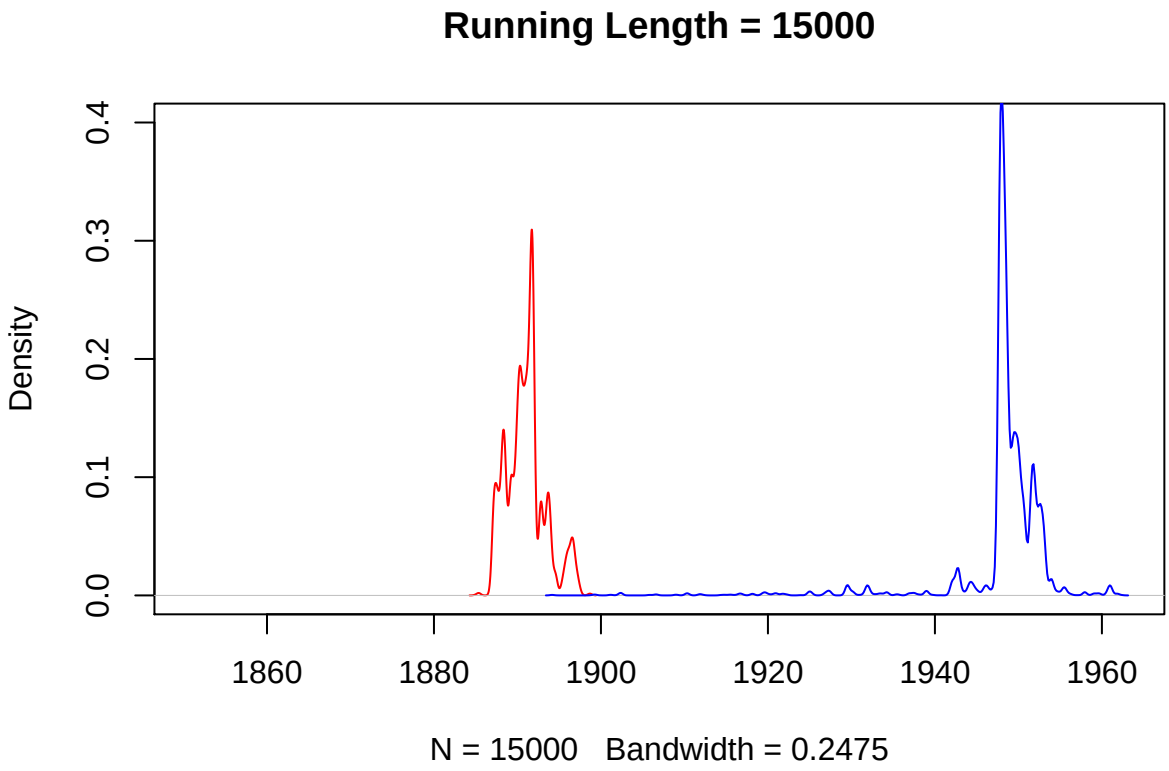
Density of Height with Running Length (N) = 1000

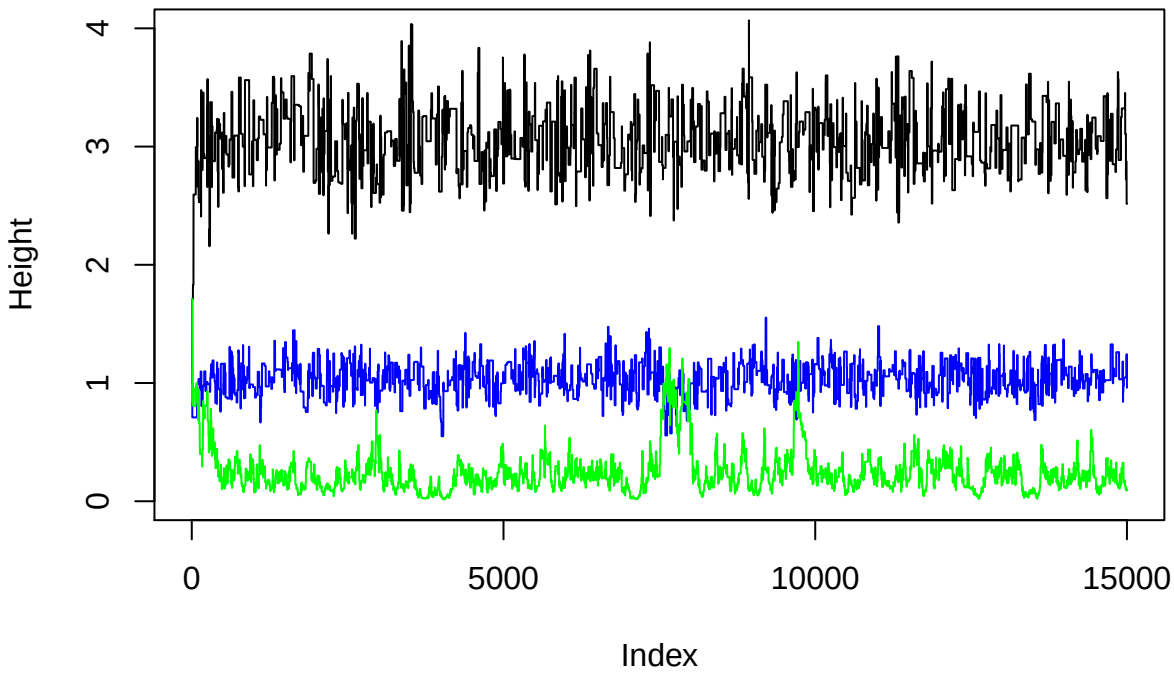


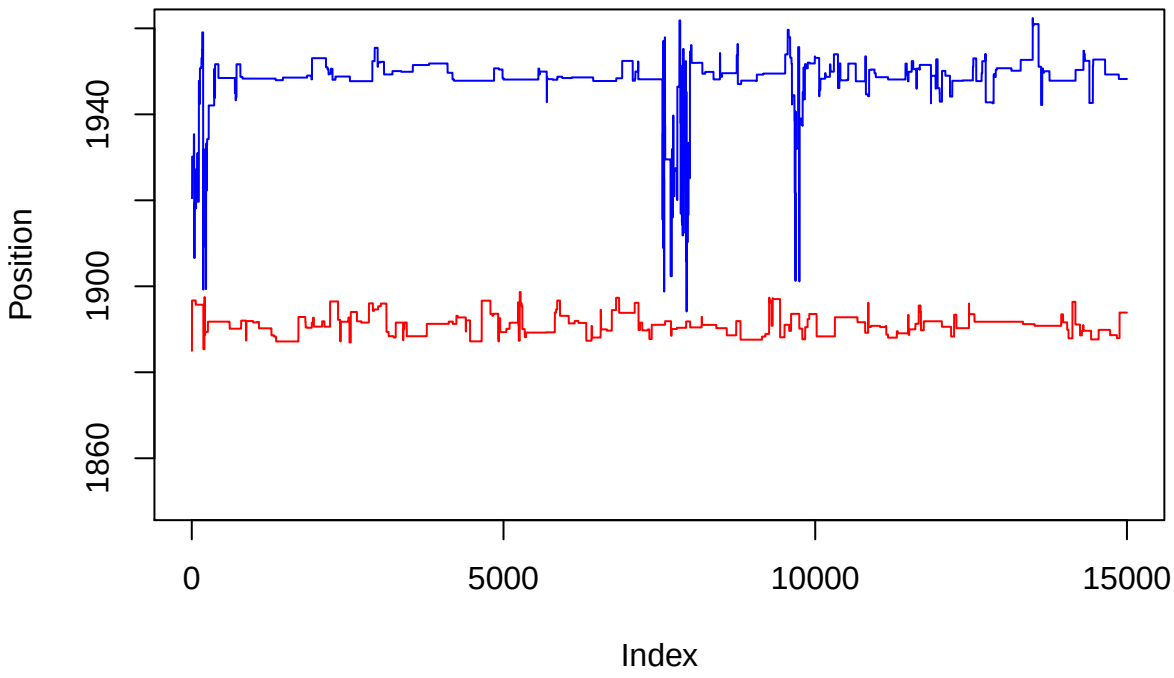
Density of Height with Running Length (N) = 10000

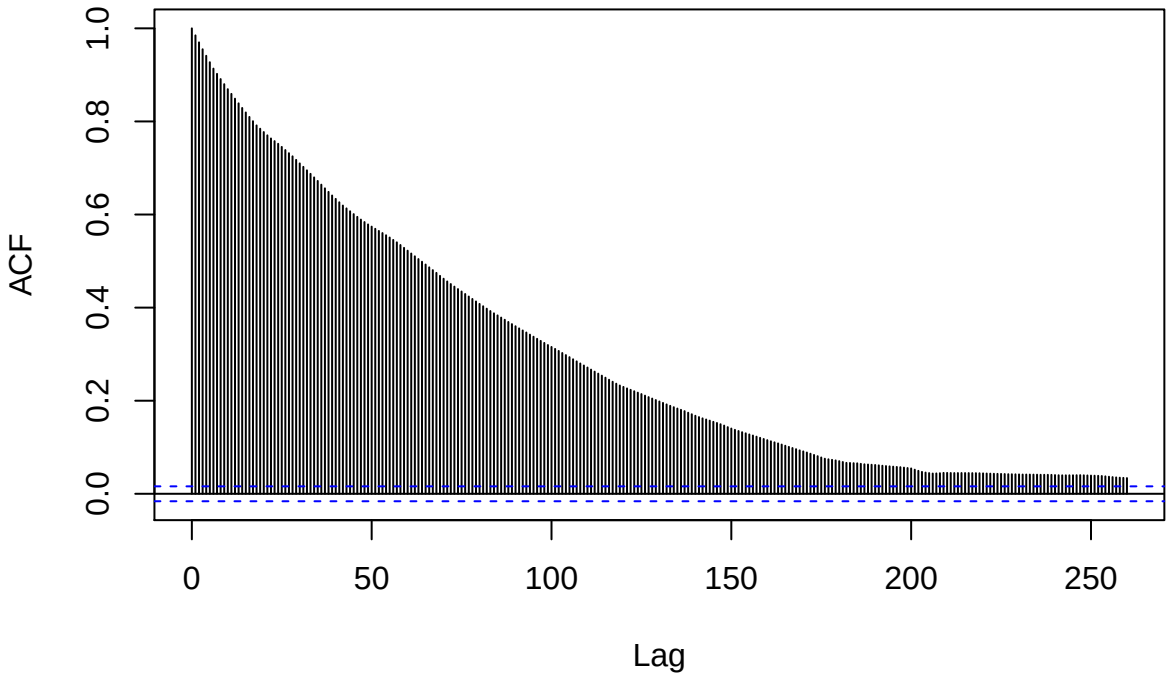


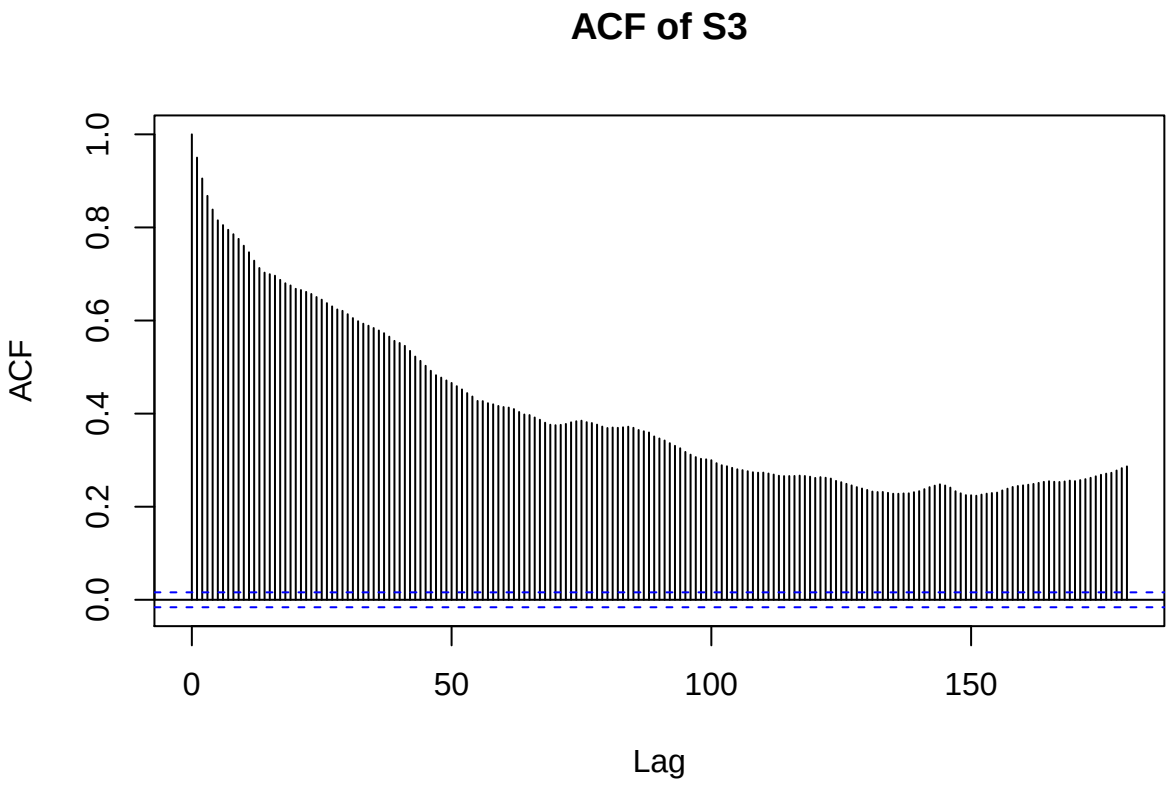
Density of Height with Running Length (N) = 15000




Height
0
1
2
3
4
0
5000
10000
15000
Index


Position
1860
1900
1940
0
5000
10000
15000
Index


ACF of S2
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
250
Lag

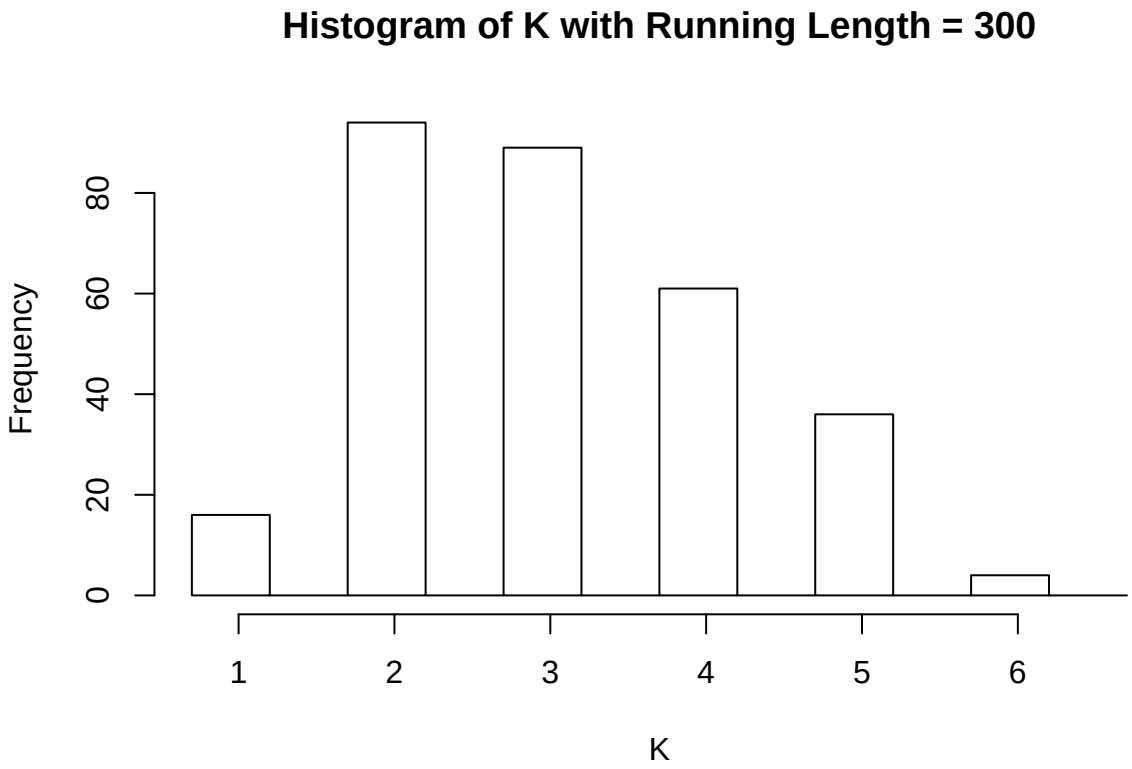


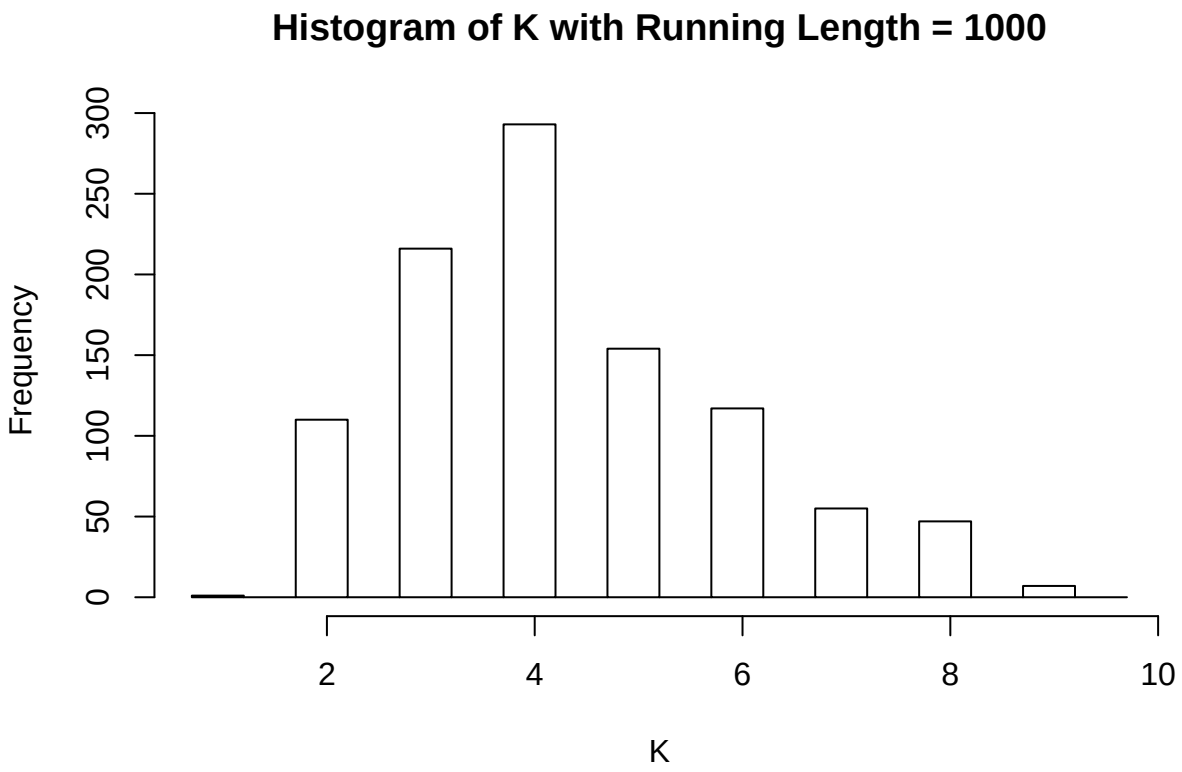
Result:

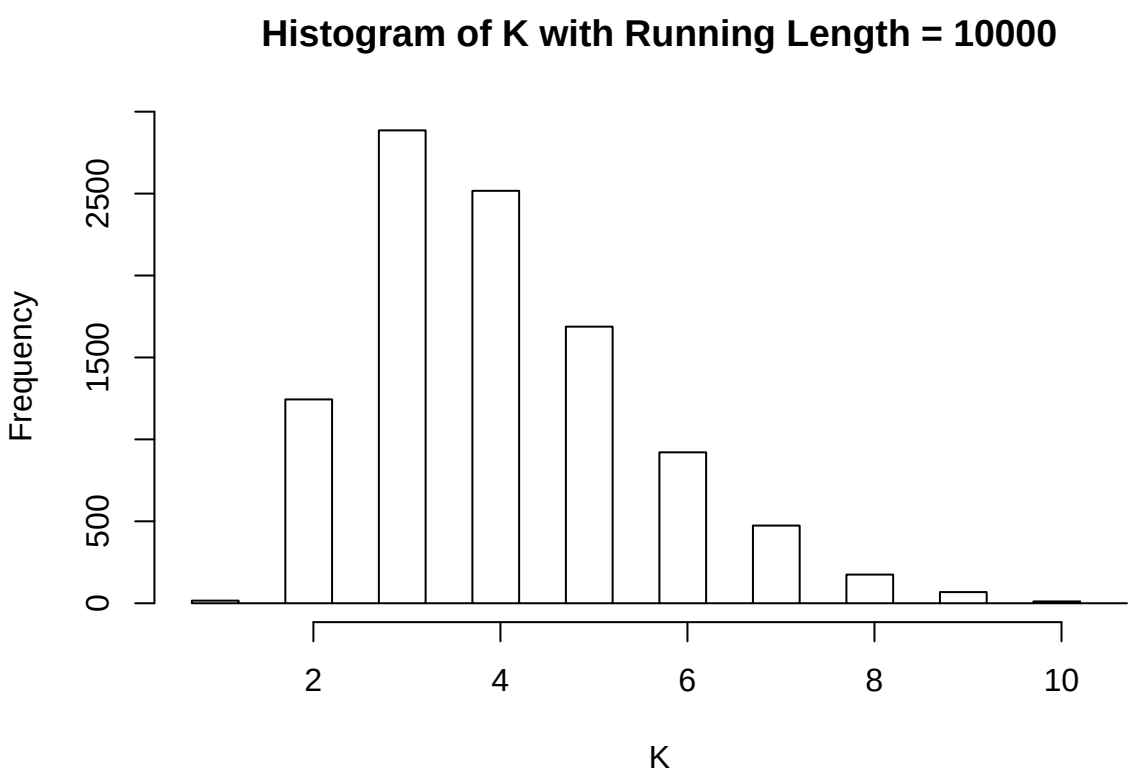
1) The density of s_2 and s_3 overlapp around 1900, which makes it be an important change-point. The plot showes that h_2 and h_3 also jumps to h_1 at same running length. Thus, the jump of position may caused by the change of height.

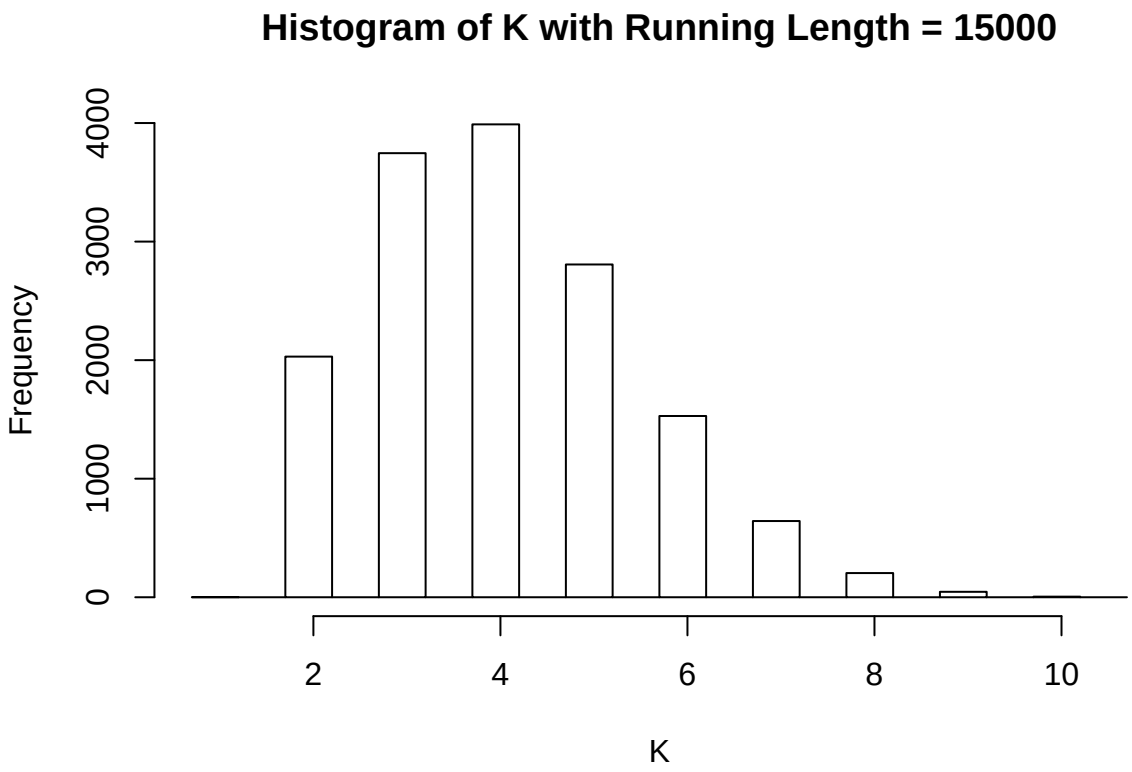
2) The distance between the mode of s_1 and s_2 dataset gets farther as running time increases. Besides, one mode finally "wins" with the highest probability for the choice of both s_1 and s_2 . Thus, the choice of s_1 and s_2 converges to the "true" value with the increase of running length.

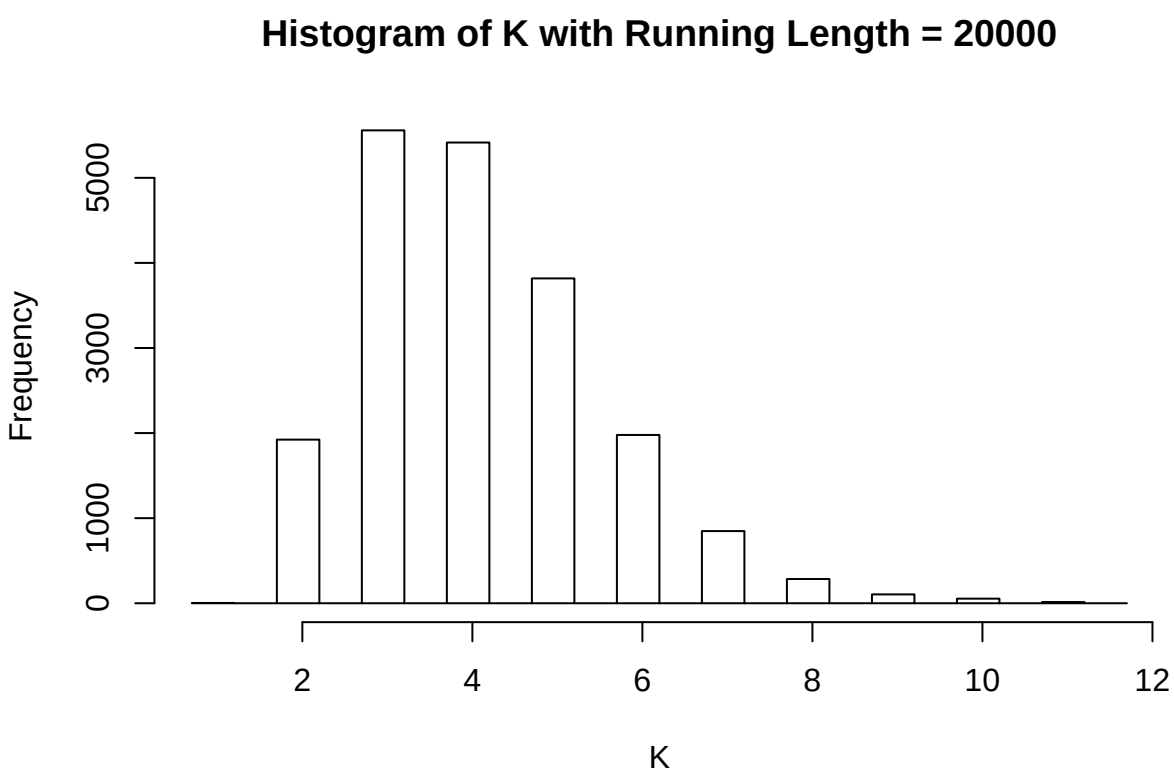
2) Varied K


Histogram of K with Running Length = 300
Frequency
0
20
40
60
80
1
2
3
4
5
6
K


Histogram of K with Running Length = 1000
Frequency
0
50
100
150
200
250
300
2
4
6
8
10
K


Histogram of K with Running Length = 10000
Frequency
0
500
1500
2500
2
4
6
8
10
K

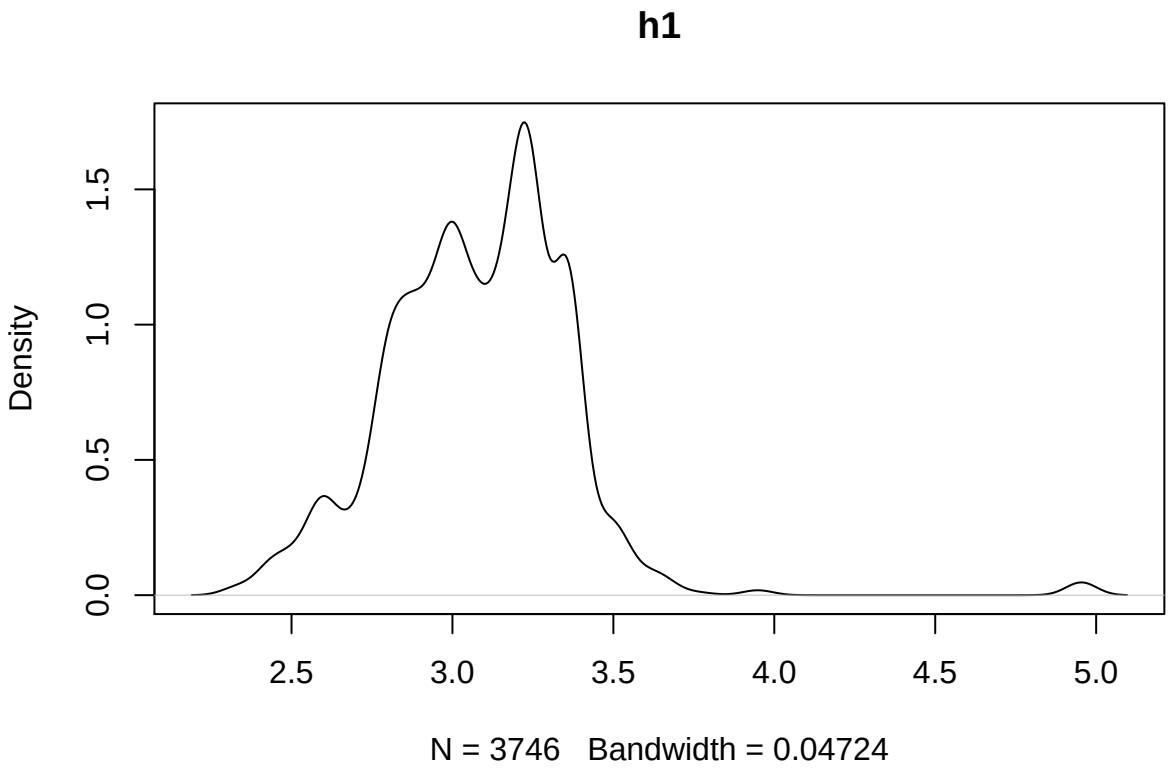

Histogram of K with Running Length = 15000
Frequency
0
1000
2000
3000
4000
2
4
6
8
10
K

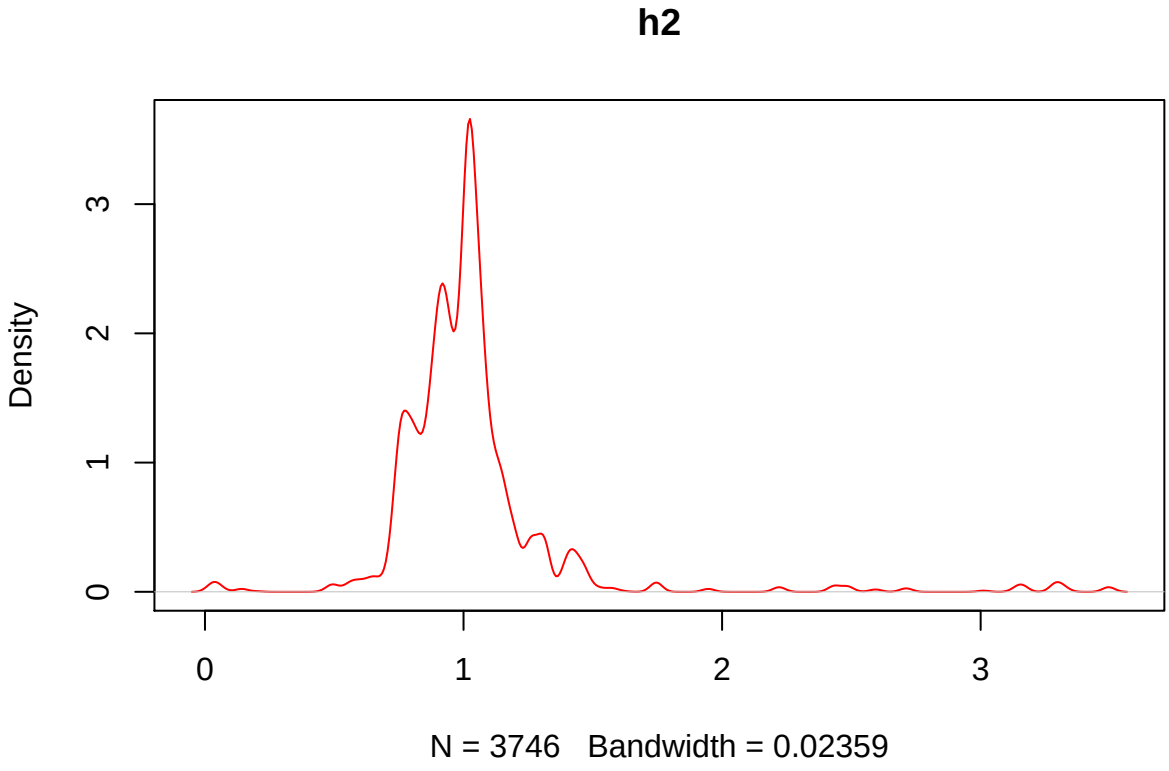


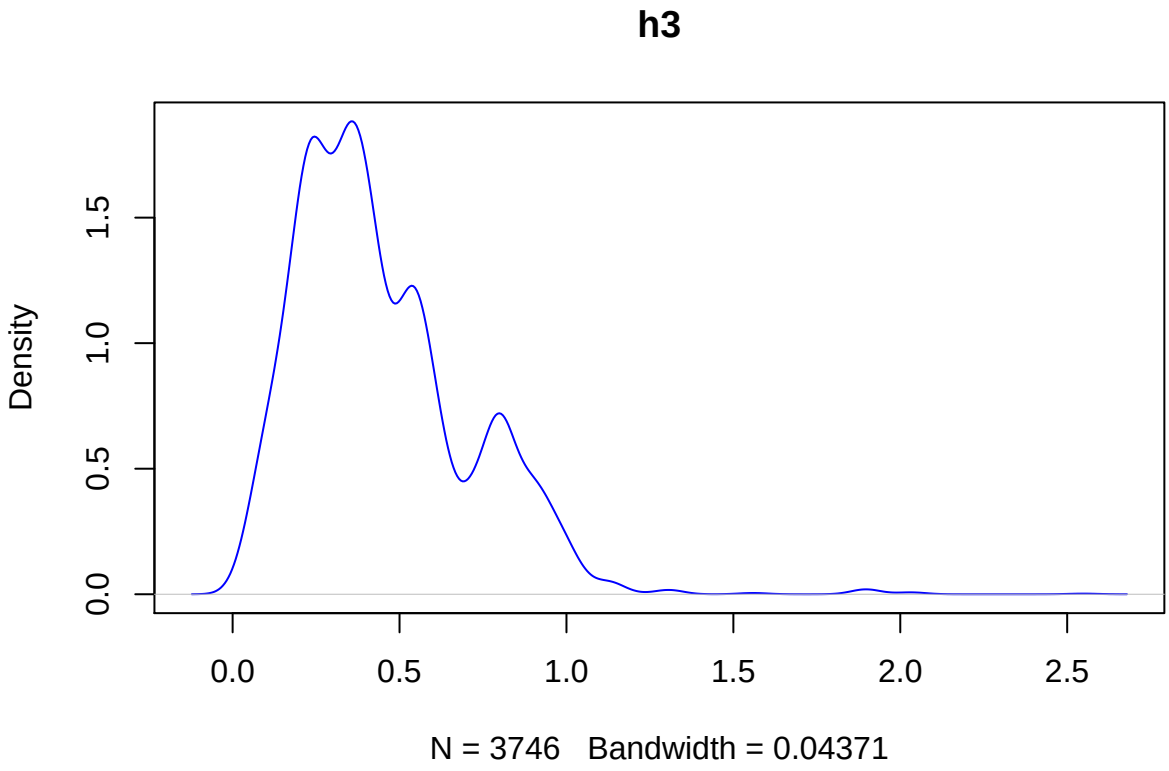
Result:

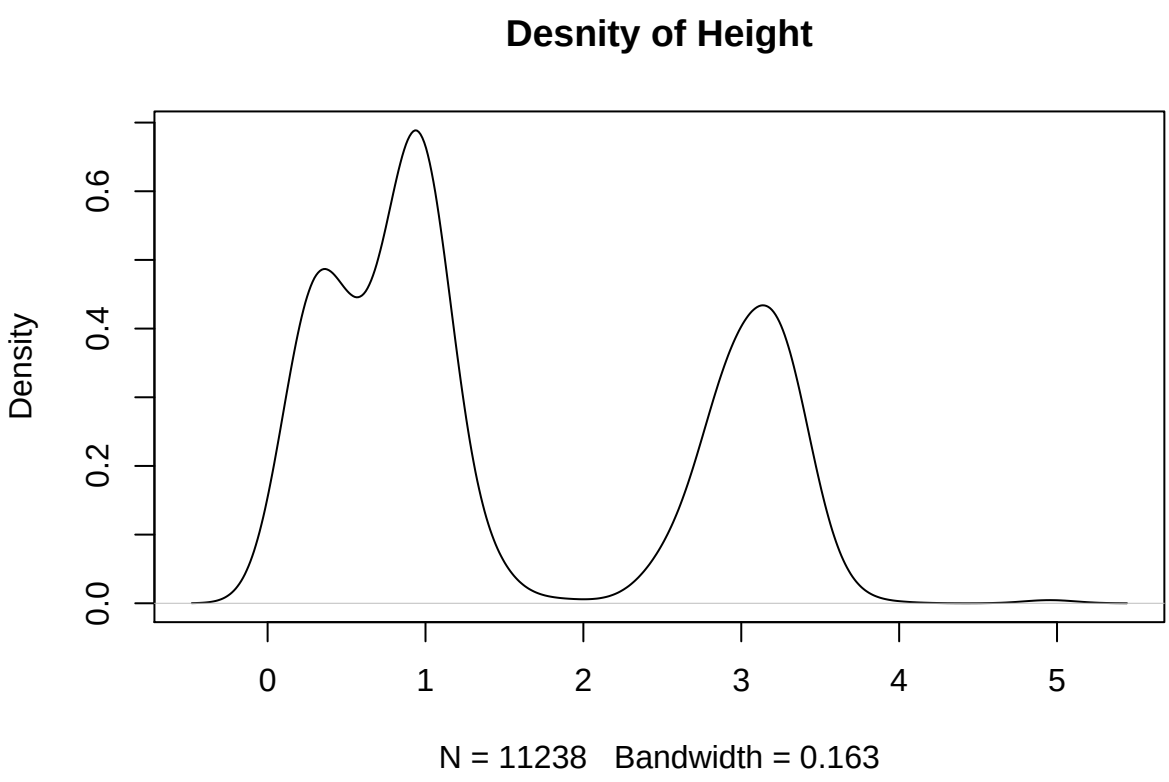
The range of k values tends to be wider as running length increases Besides, the propotion of extreme values tends to be smaller as running length increases and $k = 3$, $k = 4$ are always the most frequent choice of k value.

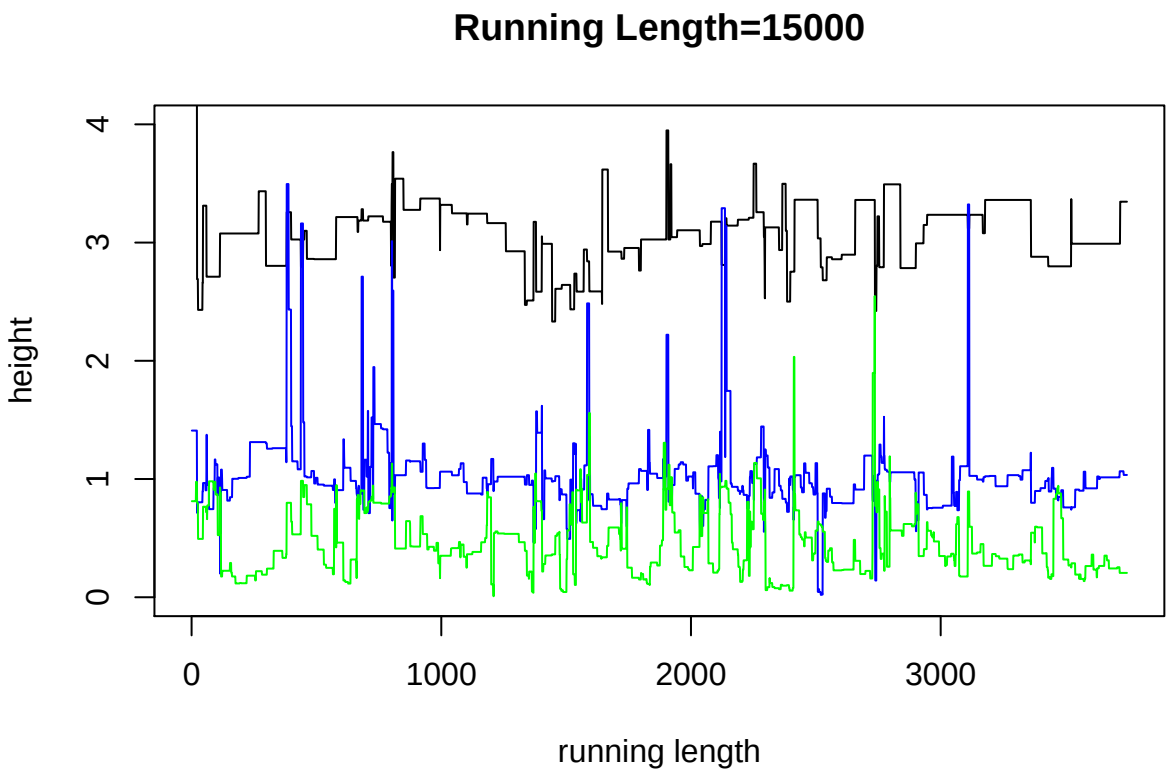
Density of Height with K=3 and Running length=15000

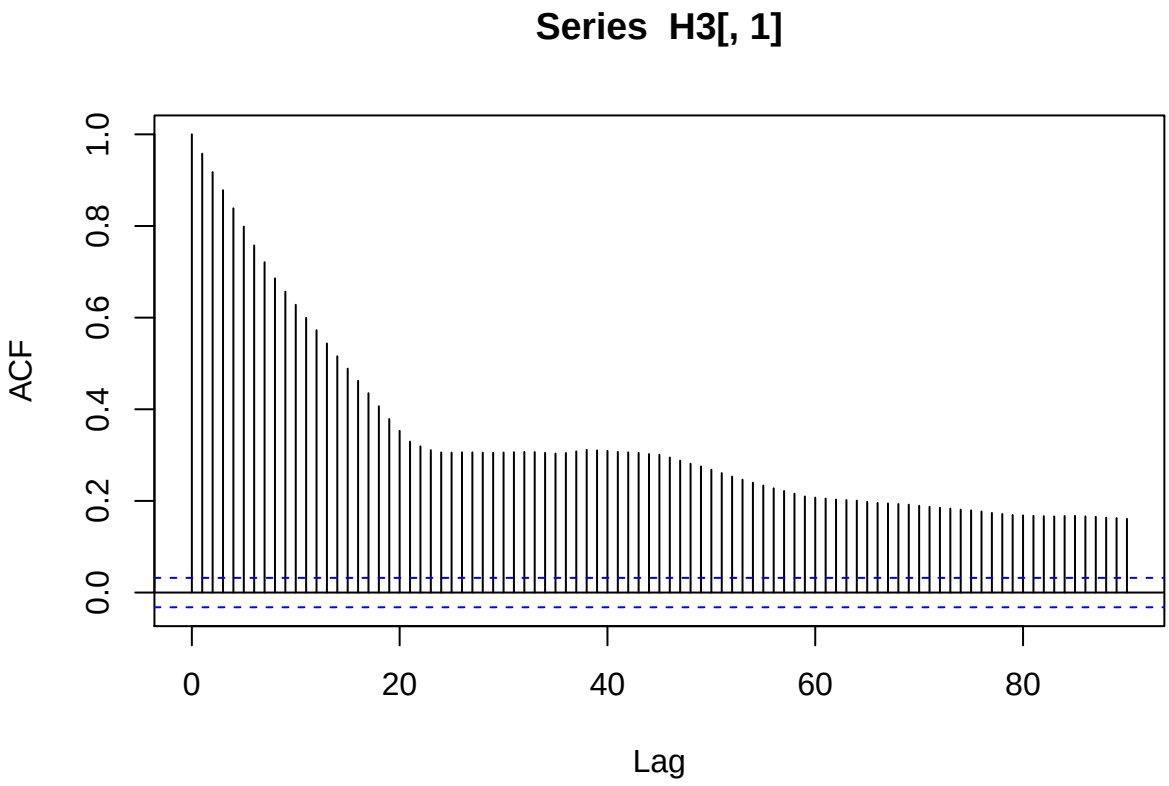


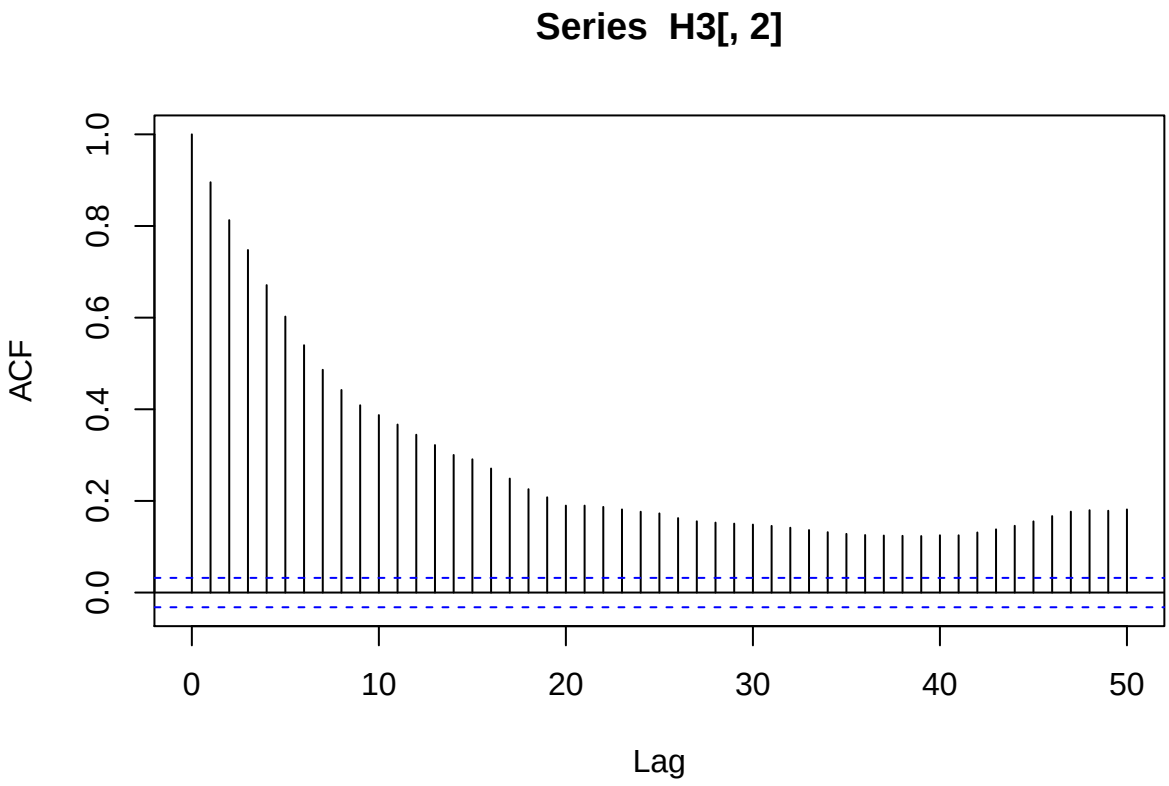


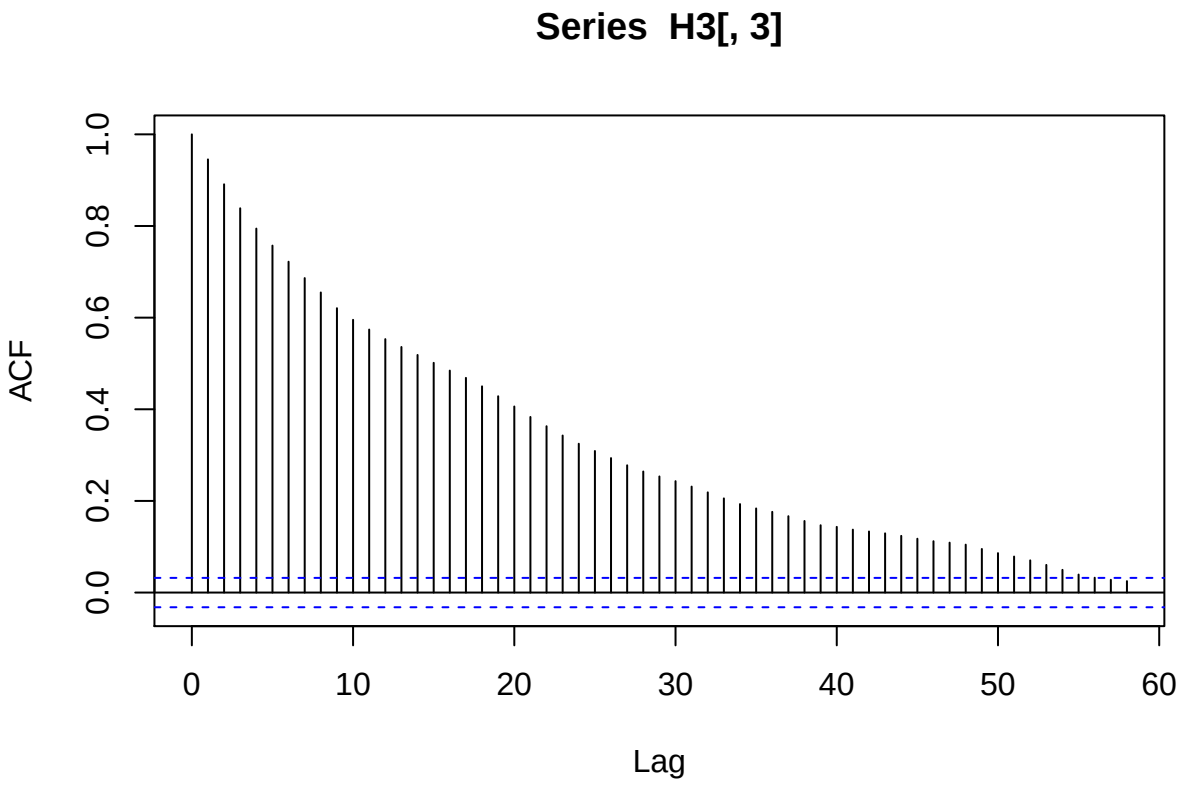

h3
Density
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
2.0
2.5
N = 3746 Bandwidth = 0.04371

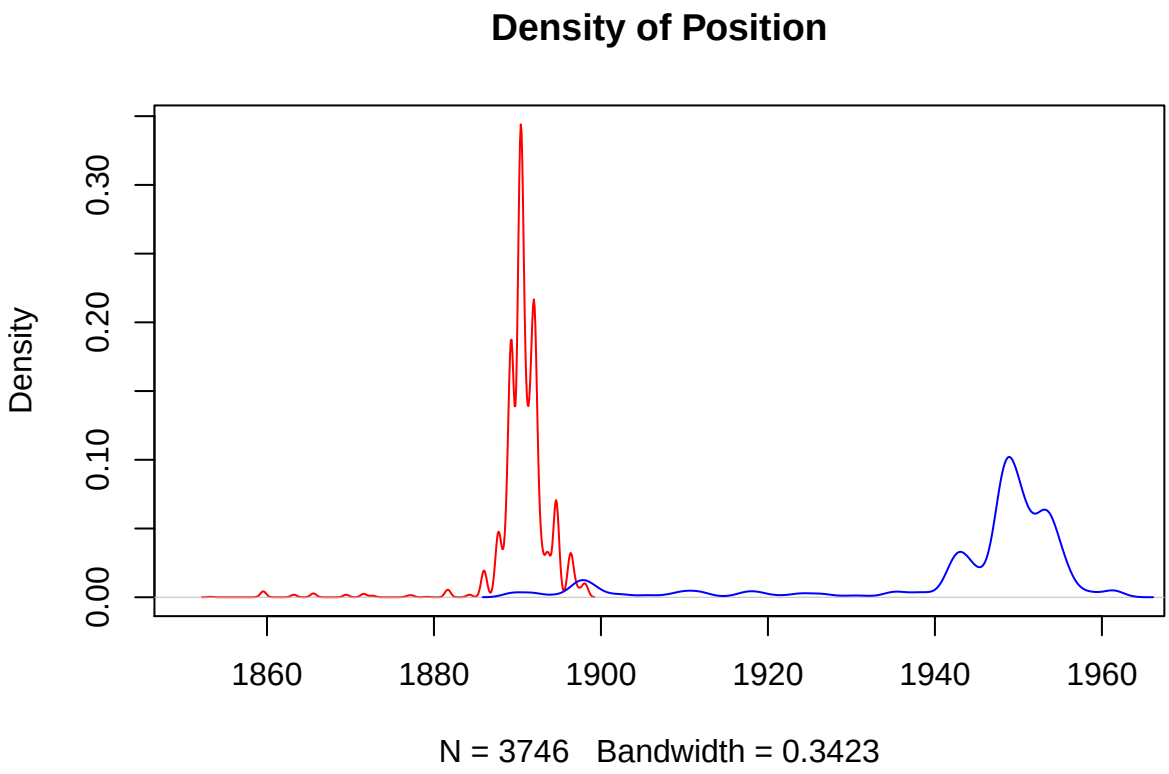

Desnity of Height
Density
0.0
0.2
0.4
0.6
0
1
2
3
4
5
N = 11238 Bandwidth = 0.163

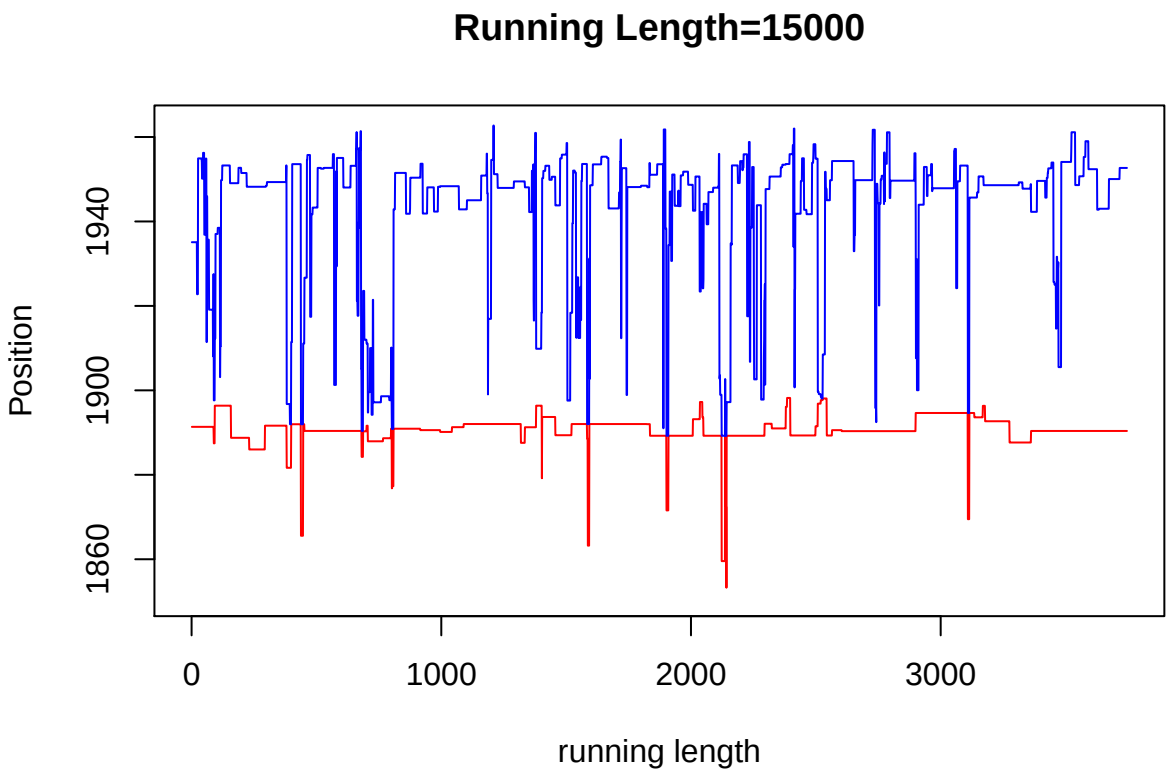

Running Length=15000
height
0
1
2
3
4
0
1000
2000
3000
running length

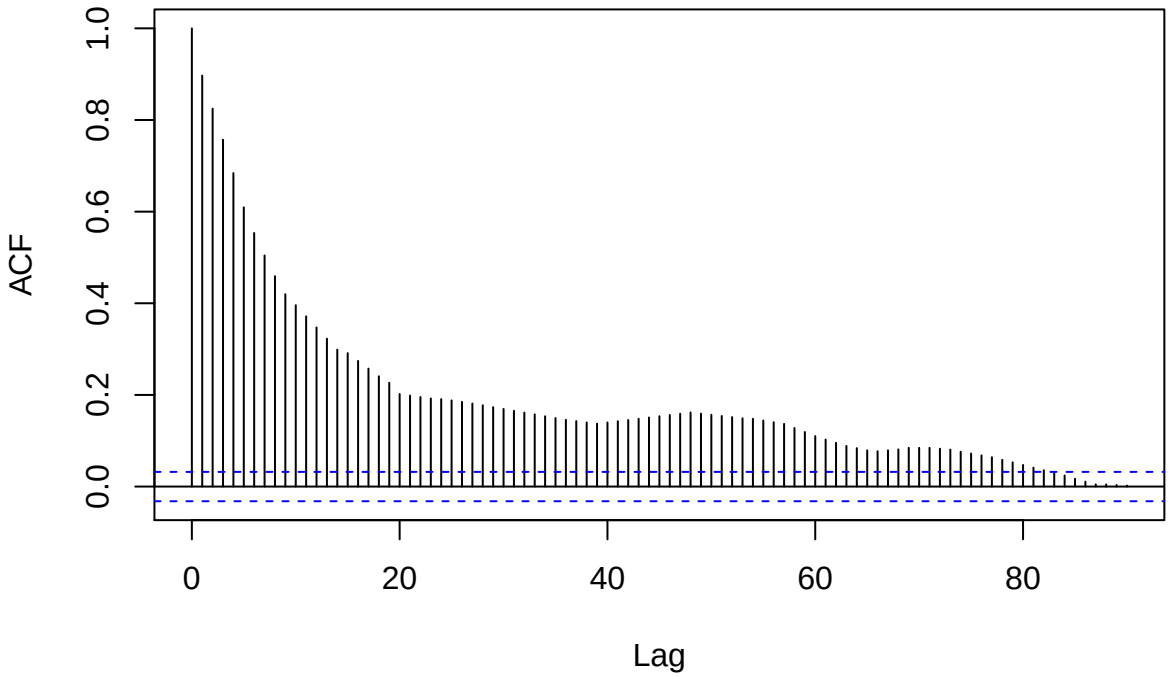

Series H3[, 1]
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
Lag

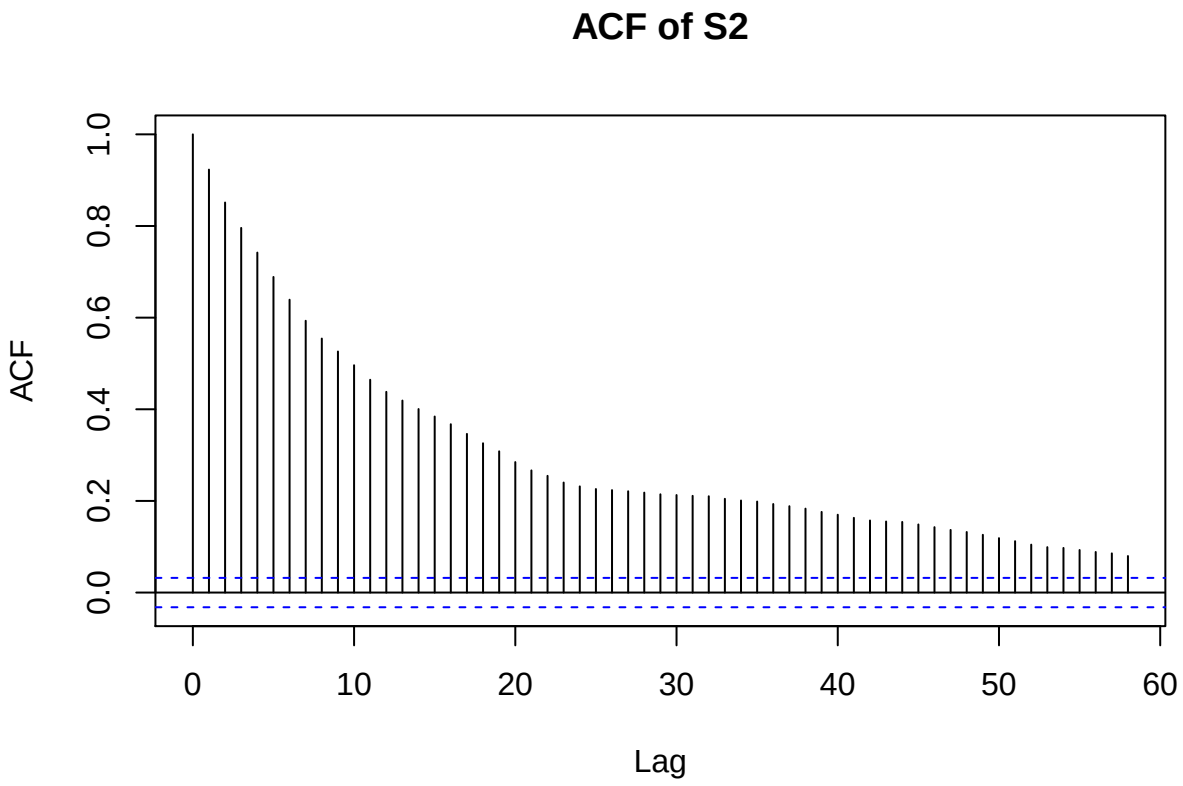

Series H3[, 2]
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
Lag


Series H3[, 3]
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
Lag


Density of Position
Density
0.00
0.10
0.20
0.30
1860
1880
1900
1920
1940
1960
N = 3746 Bandwidth = 0.3423


Running Length=15000
Position
1860
1900
1940
0
1000
2000
3000
running length


ACF of S2
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
Lag



Result:

Compared with fixed $k = 3$ case, s_2 jumps close to s_1 more frequently as well as h_3, h_2 and h_1, although the density of height are similar.