Final Term Project

2020.12.16(전산천문학)

202050135 홍지민

Contents

Research Purpose

- ‘Stansby’ used the magnetic field data of Helios E2 and E3 to create a data called ‘Corefit’ (E2’s data = E3’s data).

- However, E2 collected data 40 seconds on average, and E3 is 6 seconds.

- Stansby used azimuthal and elevation, but I considered the size of the three component (SSE coordinate).

- Also, I considered the approximation using the finite difference method (chapter 1) learned in class.

- I used Pearson correlation to determine how much correlation it had.

Applied Method

- I used the finite difference matrix I learned in class (chapter 1).

- However, although the book assumed a periodic function,
  the actual data are not a periodic function,
  so it is not possible to apply a ‘Central finite difference‘
  from the boundary value.

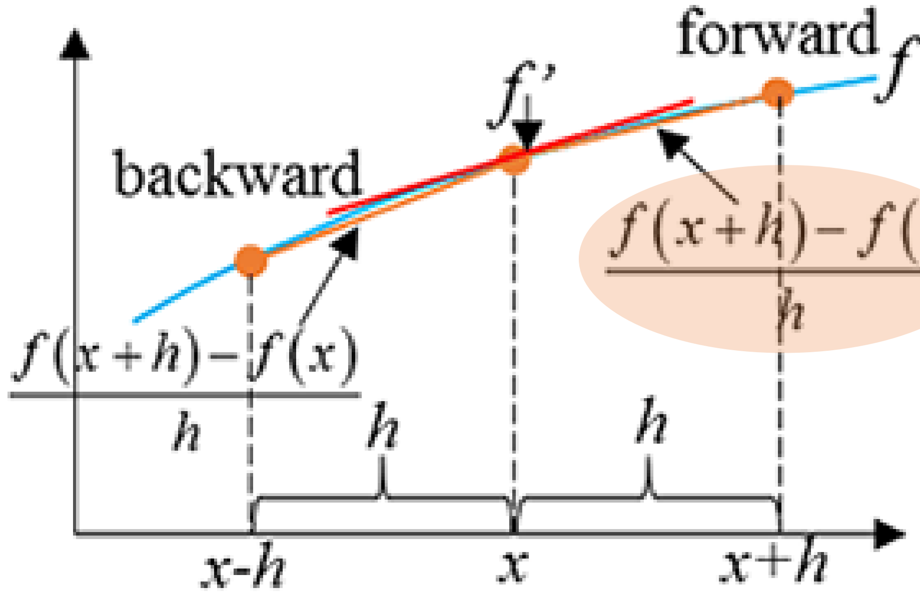
$$
h^{-1}\begin{pmatrix}
0 & \frac{1}{2} & & & -\frac{1}{2} \\
-\frac{1}{2} & 0 & \ddots & & \\
 & & \ddots & & \\
 & & \ddots & 0 & \frac{1}{2} \\
\frac{1}{2} & & & -\frac{1}{2} & 0
\end{pmatrix}
$$

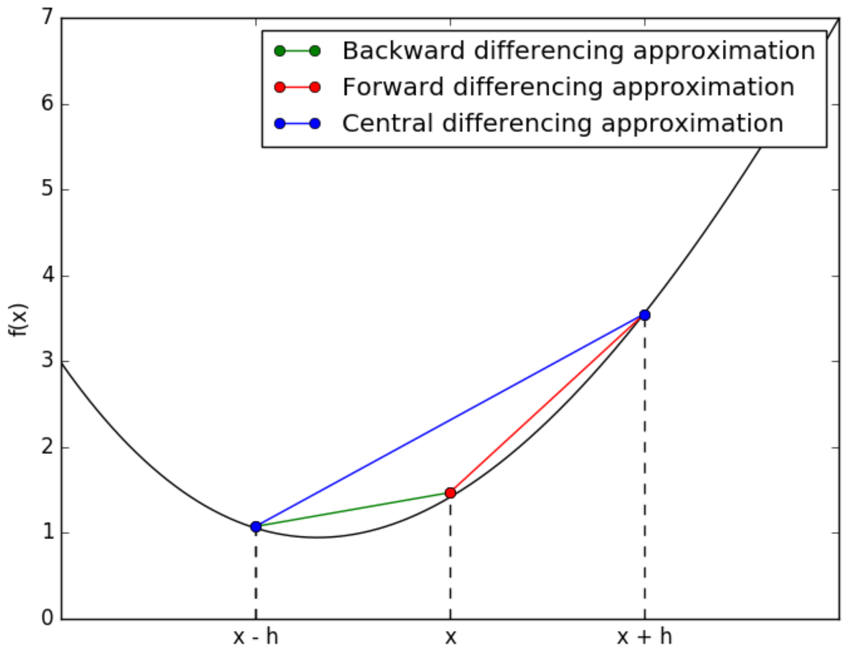
$$
\begin{pmatrix} w_1 \\ \\ \\ \vdots \\ \\ \\ w_N \end{pmatrix}
= h^{-1}\begin{pmatrix}
 & & \ddots & & & \frac{1}{12} & -\frac{2}{3} \\
 & & \ddots & -\frac{1}{12} & & & \frac{1}{12} \\
 & & \ddots & \frac{2}{3} & \ddots & & \\
 & & \ddots & 0 & \ddots & & \\
 & & \ddots & -\frac{2}{3} & \ddots & & \\
-\frac{1}{12} & & & \frac{1}{12} & \ddots & & \\
\frac{2}{3} & -\frac{1}{12} & & & \ddots & &
\end{pmatrix}
\begin{pmatrix} u_1 \\ \\ \\ \vdots \\ \\ \\ u_N \end{pmatrix}
$$

Applied Method

- So I found out through Googleing that there is a method of 'forward' and 'backward'.

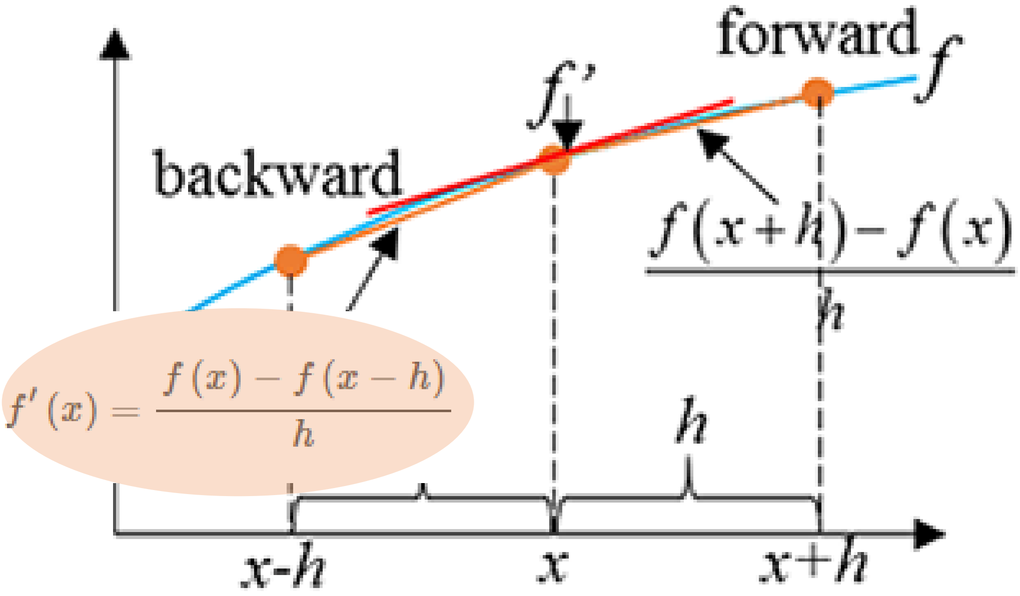
- The forward finite difference method is a approximating values ahead of the reference point.

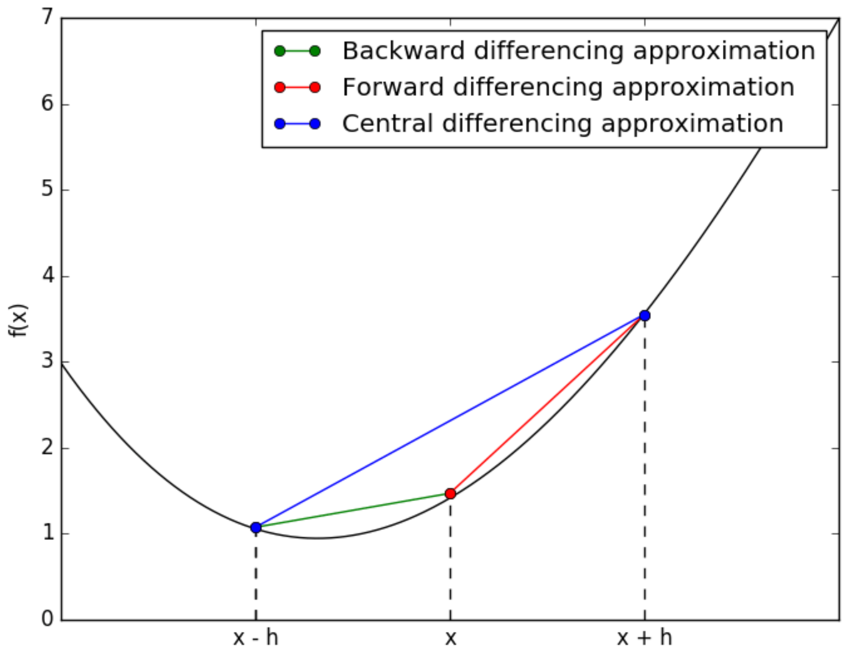




Applied Method

- Conversely, the backward difference method is an approximation of the points located behind the reference point.
- Also, I followed the accuracy order by referring to Wikipedia.





Applied Method

Central

Derivative	Accuracy	−5	−4	−3	−2	−1	0	1	2	3	4	5
1	2					−1/2	0	1/2				
	4				1/12	−2/3	0	2/3	−1/12			
	6			−1/60	3/20	−3/4	0	3/4	−3/20	1/60		

Forward

Derivative	Accuracy	0	1	2	3	4	5	6	7	8
1	1	−1	1							
	2	−3/2	2	−1/2						
	3	−11/6	3	−3/2	1/3					
	4	−25/12	4	−3	4/3	−1/4				
	5	−137/60	5	−5	10/3	−5/4	1/5			

Backward

Derivative	Accuracy	−8	−7	−6	−5	−4	−3	−2	−1	0
1	1								−1	1
	2							1/2	−2	3/2
	3						−1/3	3/2	−3	11/6

Code

```python
def centered_finite_diff(N,h): #centered
  line = np.zeros(N)
  line[1] = 2/3 ; line[N-1] = -2/3 ; line[2] = -1/12 ; line[N-2] = 1/12

  matrix = 1/h * toep(-line,line)

  return matrix

def forward_finite_diff(N,h): #forward
  line = np.zeros(N)
  line[0] = -25/12 ; line[1] = 4 ; line[2] = -3 ; line[3] = 4/3 ; line[4] = -1/4

  line2 = np.zeros(N)
  line2[0] = -25/12 ; line2[N-1] = 4 ; line2[N-2] = -3 ; line2[N-3] = 4/3 ; line2[N-4] = -1/4

  matrix = 1/h * toep(line2,line)

  return matrix

def backward_finite_diff(N,h): #backward
  line = np.zeros(N)
  line[0] = 25/12 ; line[1] = -4 ; line[2] = 3 ; line[3] = -4/3 ; line[4] = 1/4

  line2 = np.zeros(N)
  line2[0] = 25/12 ; line2[N-1] = -4 ; line2[N-2] = 3 ; line2[N-3] = -4/3 ; line2[N-

  matrix = 1/h * toep(line,line2)

  return matrix
```

$$
h^{-1}\begin{pmatrix}
 & & \ddots & & & \frac{1}{12} & -\frac{2}{3} \\
 & & \ddots & -\frac{1}{12} & & & \frac{1}{12} \\
 & & \ddots & \frac{2}{3} & \ddots & & \\
 & & \ddots & 0 & \ddots & & \\
 & & \ddots & -\frac{2}{3} & \ddots & & \\
-\frac{1}{12} & & & \frac{1}{12} & \ddots & & \\
\frac{2}{3} & -\frac{1}{12} & & & \ddots & &
\end{pmatrix}
$$

Accuracy	0	1	2	3	4	5
1	−1	1				
2	−3/2	2	−1/2			
3	−11/6	3	−3/2	1/3		
4	−25/12	4	−3	4/3	−1/4	

Code

```
N = len(E2_x)
h = 1/90

A = centered_finite_diff(N,h)
B = forward_finite_diff(N,h)
C = backward_finite_diff(N,h)

L = np.zeros_like(A)
L[0] = B[0] ; L[1] = B[1]
L[N-1] = C[N-1] ; L[N-2] = C[N-2]
L[2:N-2] = A[2:N-2]
```

All Accuracy = 4

1~2 rows: Forward finite difference Matrix

$$L = h^{-1} \begin{bmatrix} \ddots & \frac{2}{3} & \ddots \\ \ddots & 0 & \ddots \\ \ddots & -\frac{2}{3} & \ddots \end{bmatrix}$$

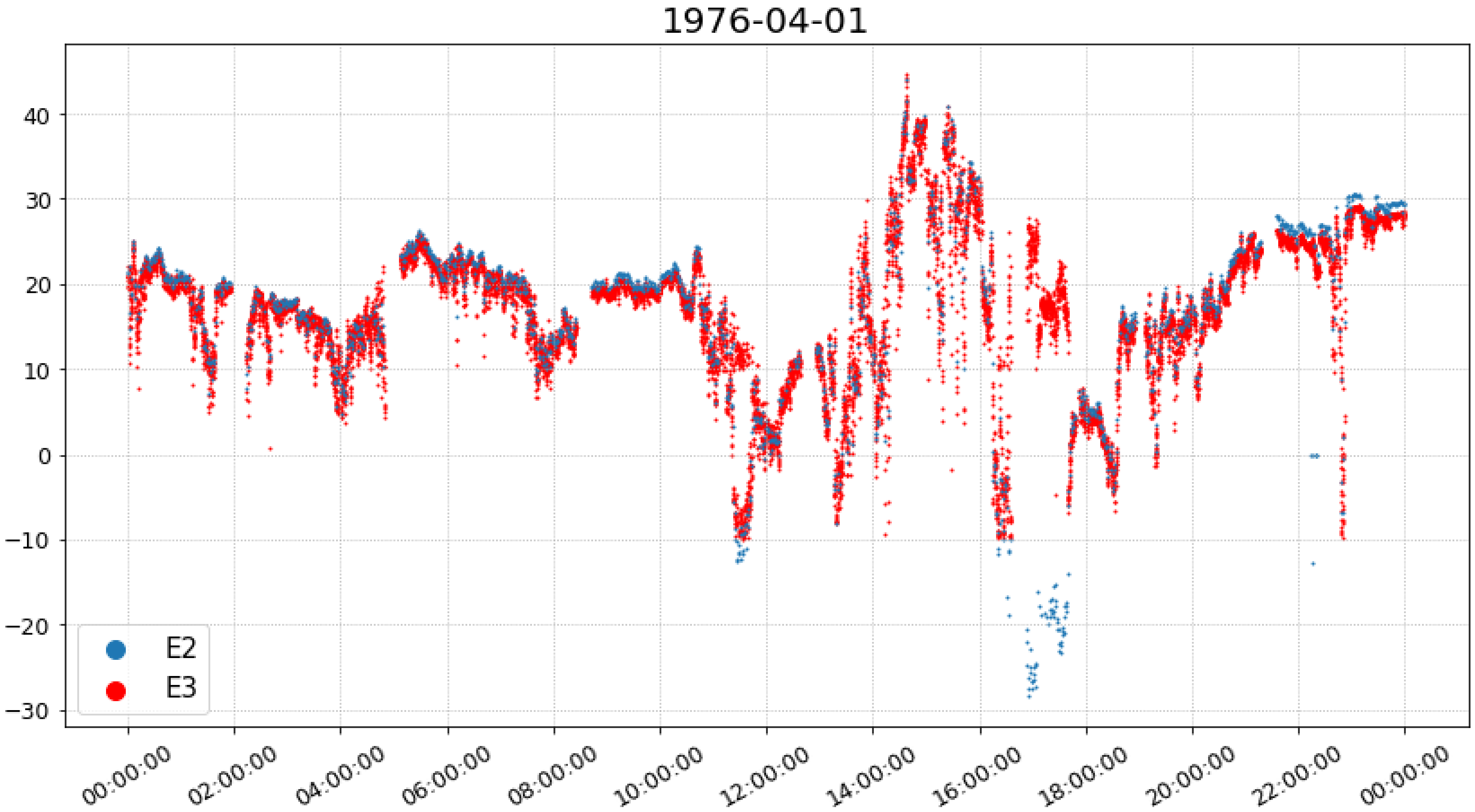
(N-1)~N rows: Backward finite difference Matrix

4				1/12	−2/3	0	2/3	−1/12

The approximation was calculated using this L matrix.

Result

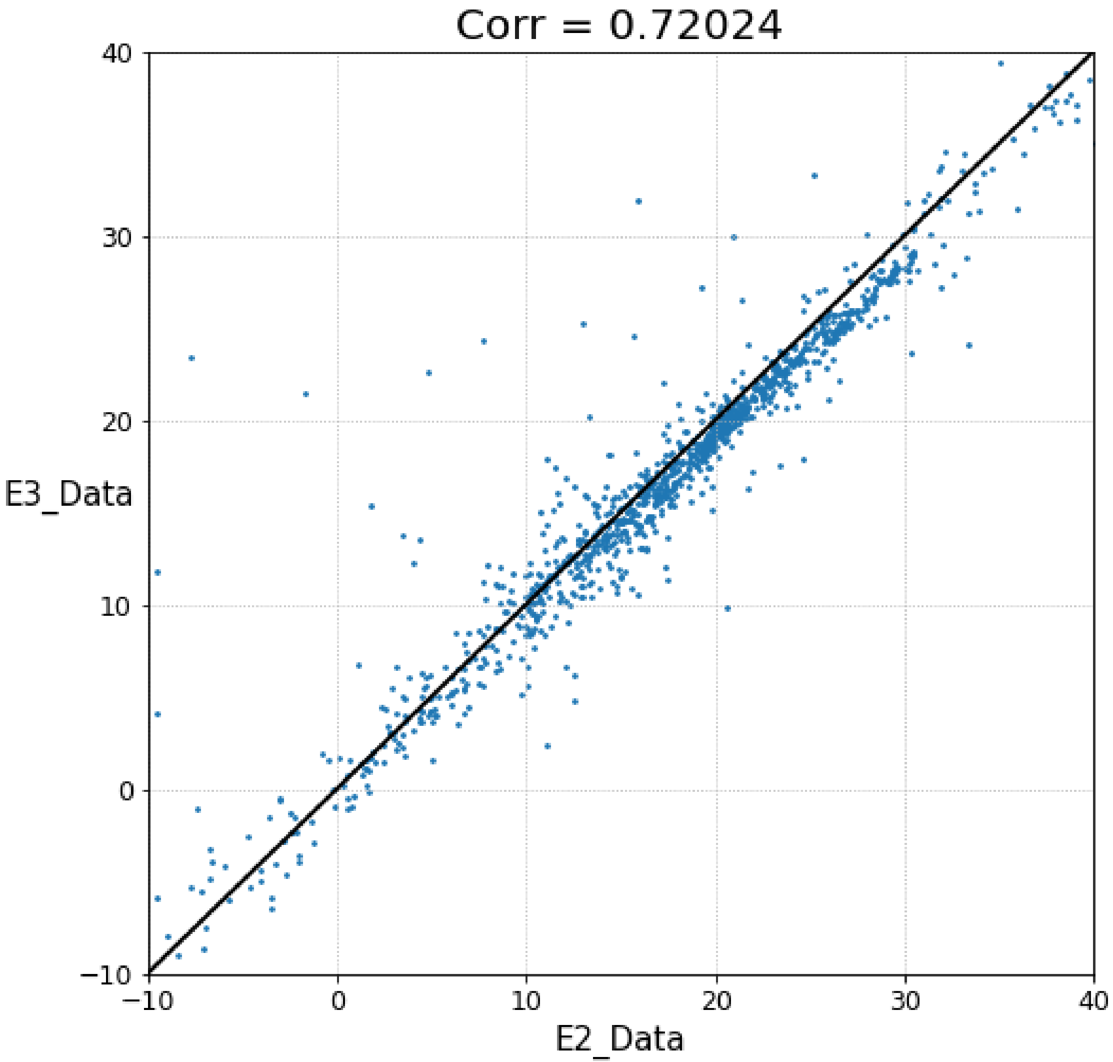
- First, I compared E2 and E3 data by drawing them in a scatter plot.

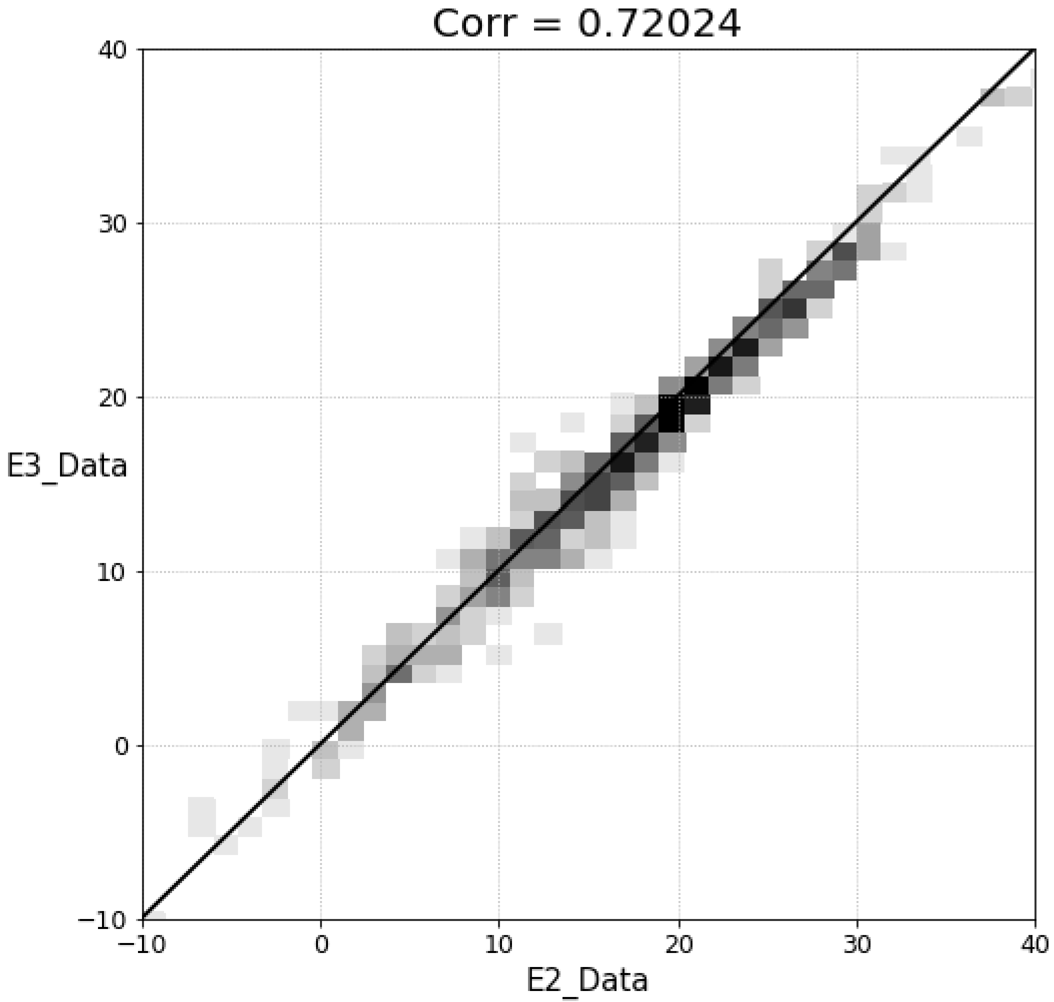


X-axis
in SSE coordinate

There is a
similar tendency.

Result

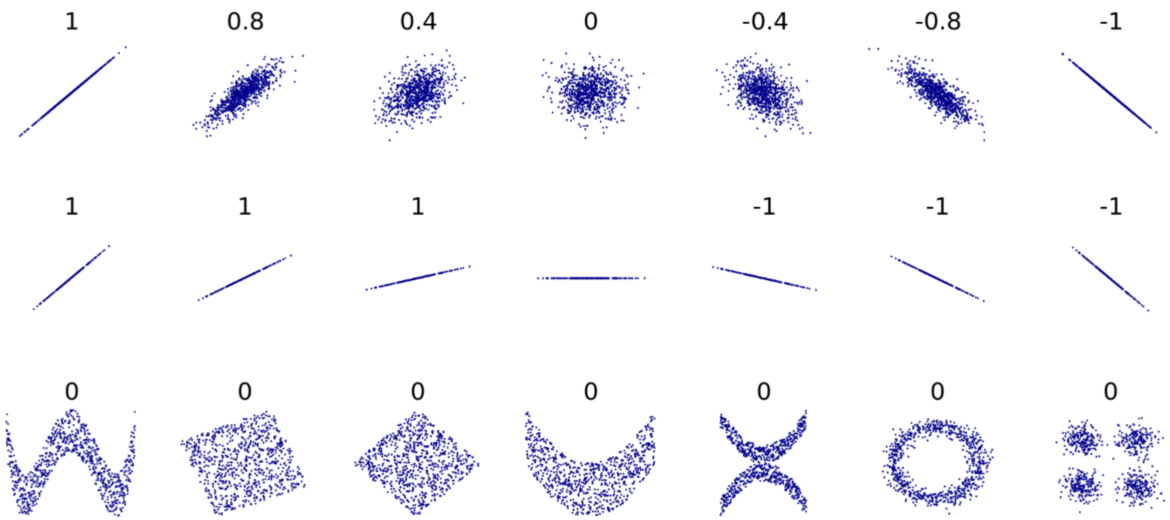


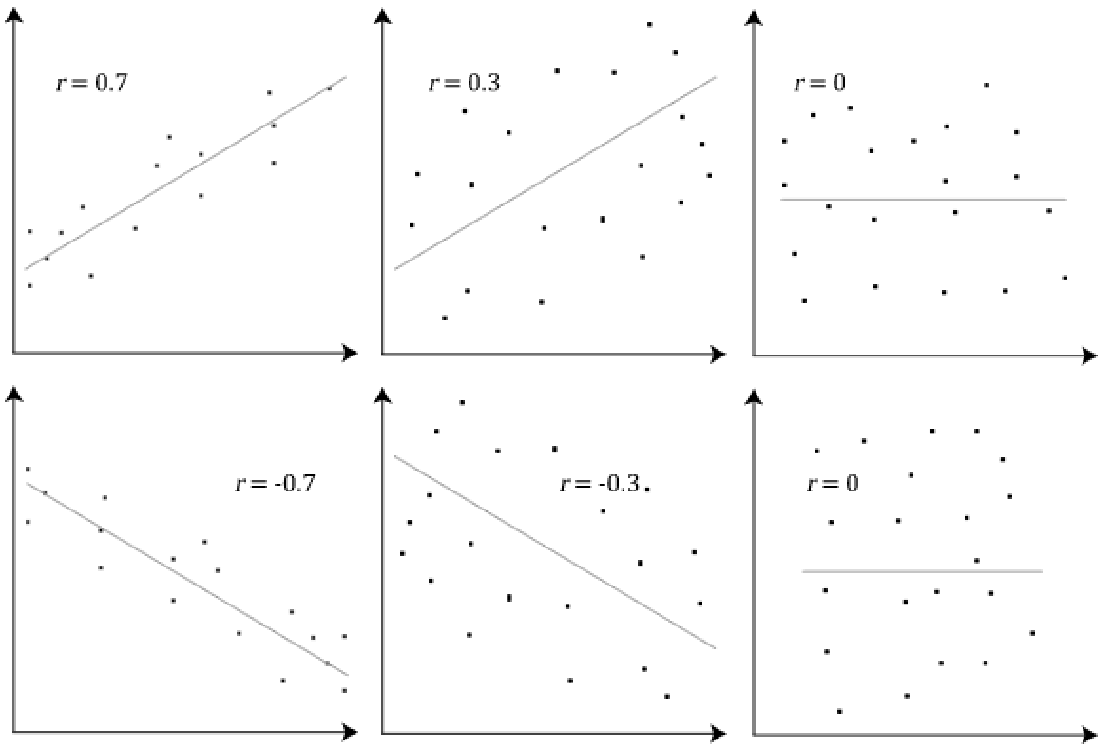


The black line represents y=x.

Pearson C.C

- I understood that x increases as positive correlation y increases.
- Conversely, the negative correlation was understood to decrease y as x increased.
- I haven't studied much about this part because it's statistical. I'm sorry.



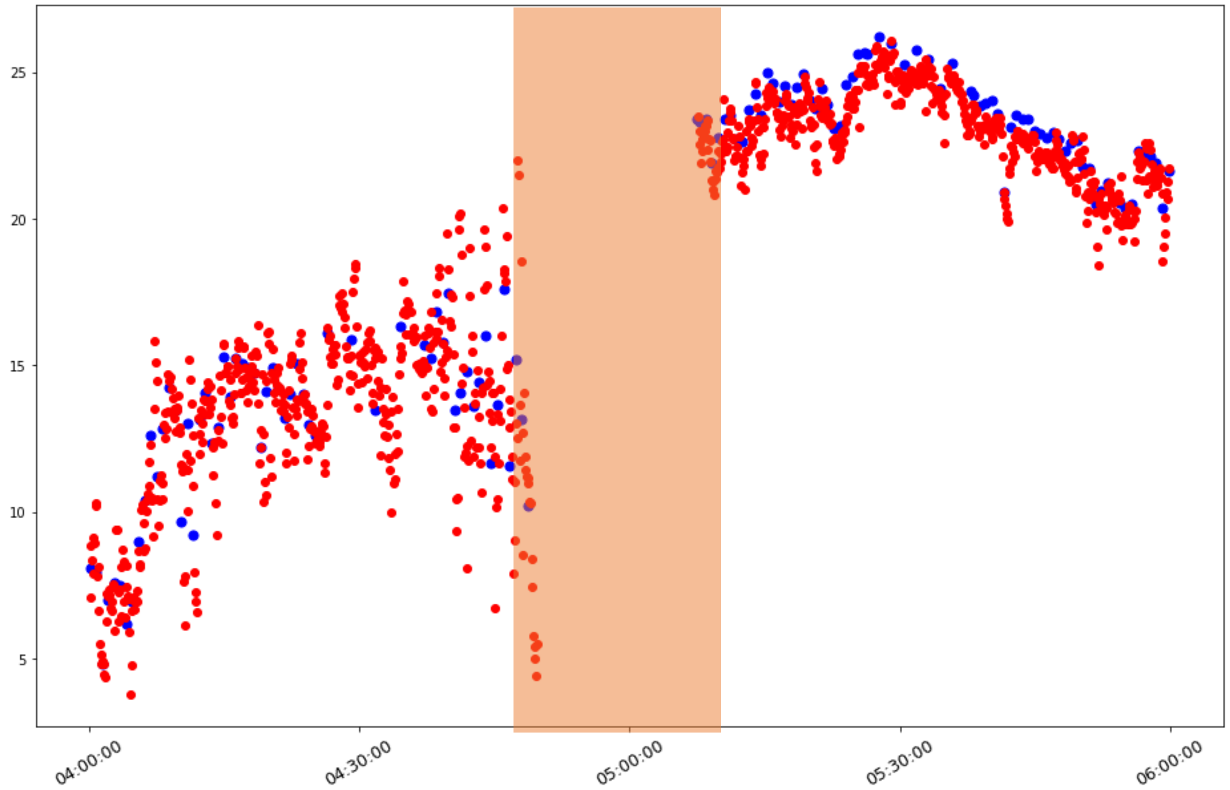


Result

Second, I tried to interpolation and regression

Because there are empty gaps
and not uniform in the data,
interpolation was implemented as a whole.

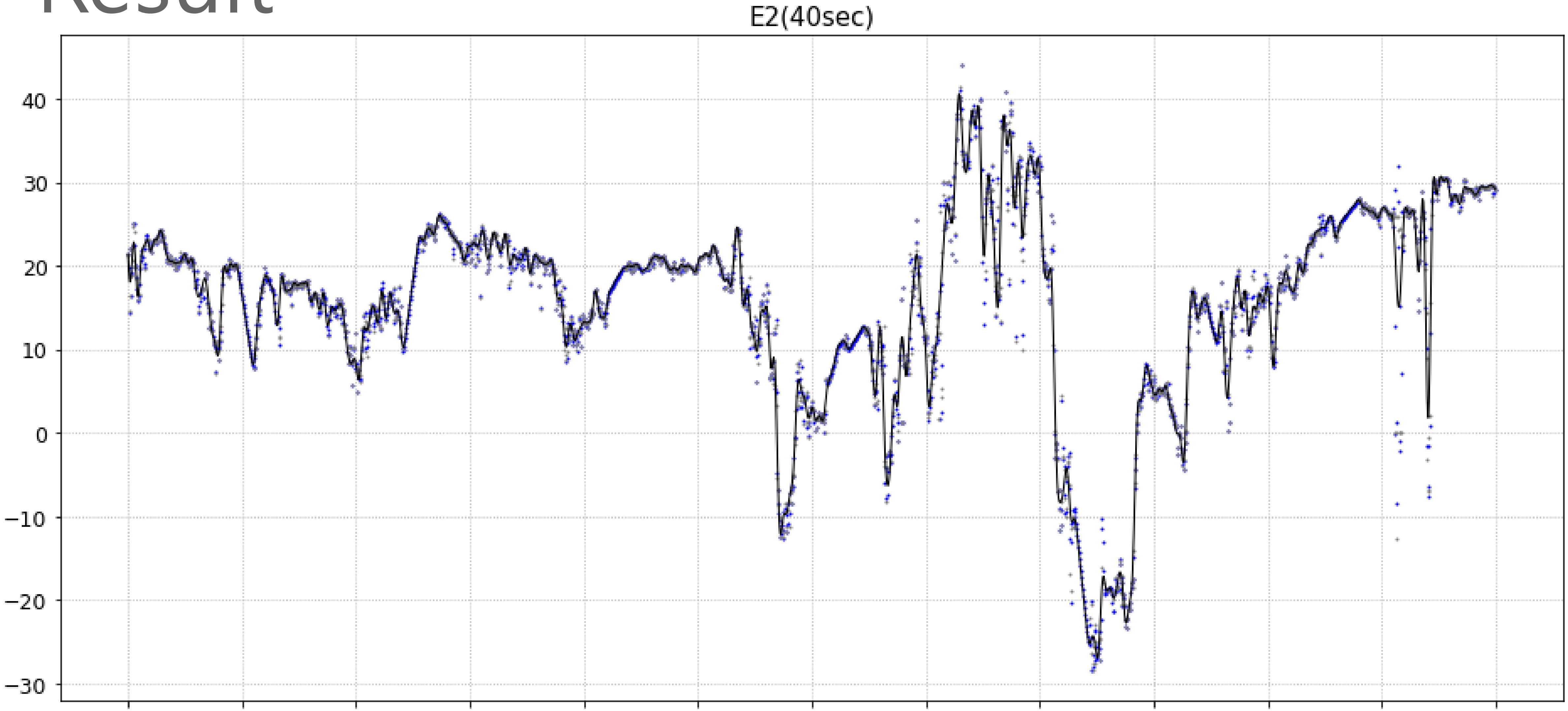
Later, an approximation was obtained
using a regression equation.



	E2	E3
Grid Point	Ave 40 sec	Ave 6 sec
No. Strange Data (1976.4.1)	152 / 1772	45 / 13004

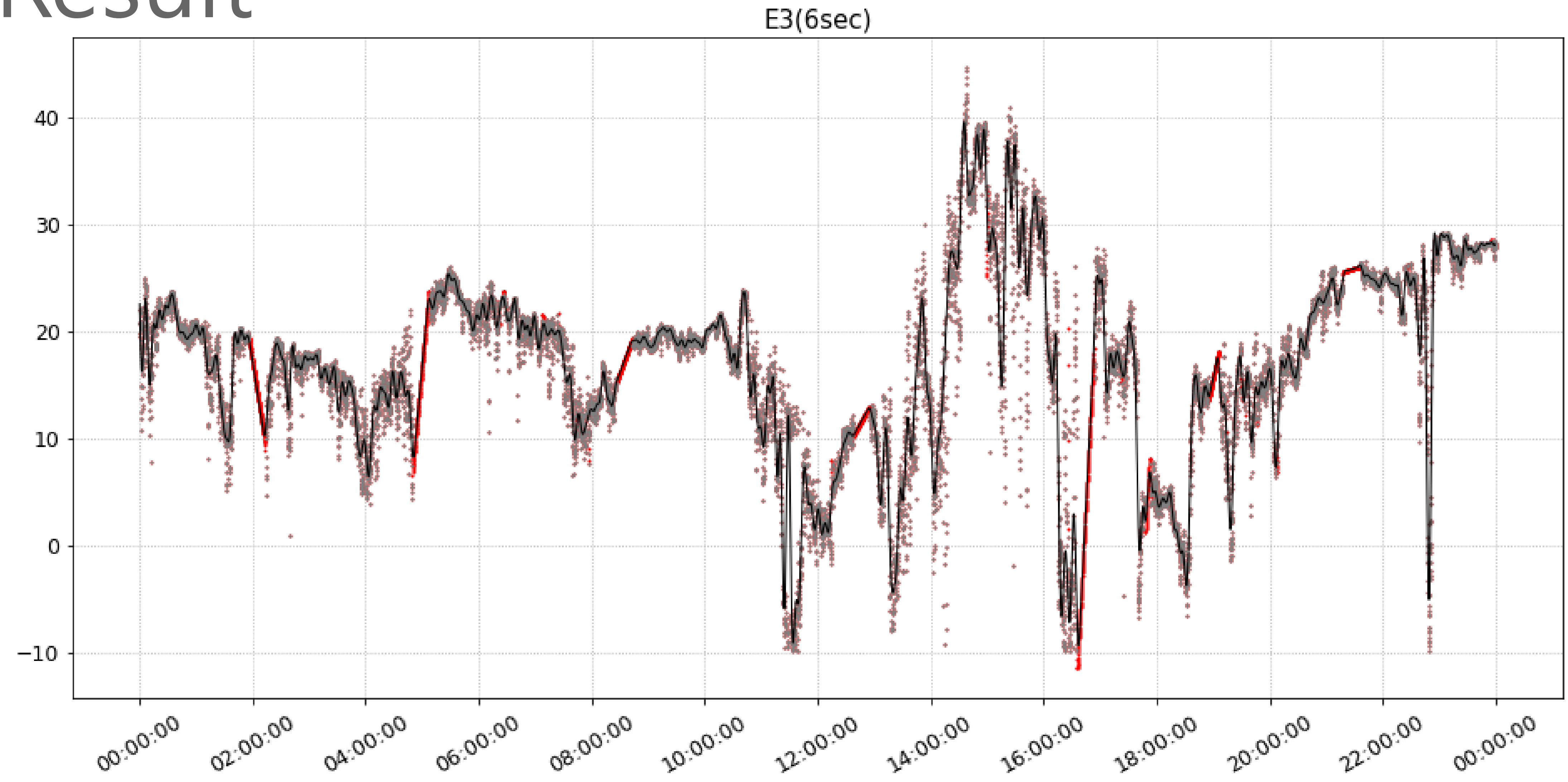
Not uniform

Result



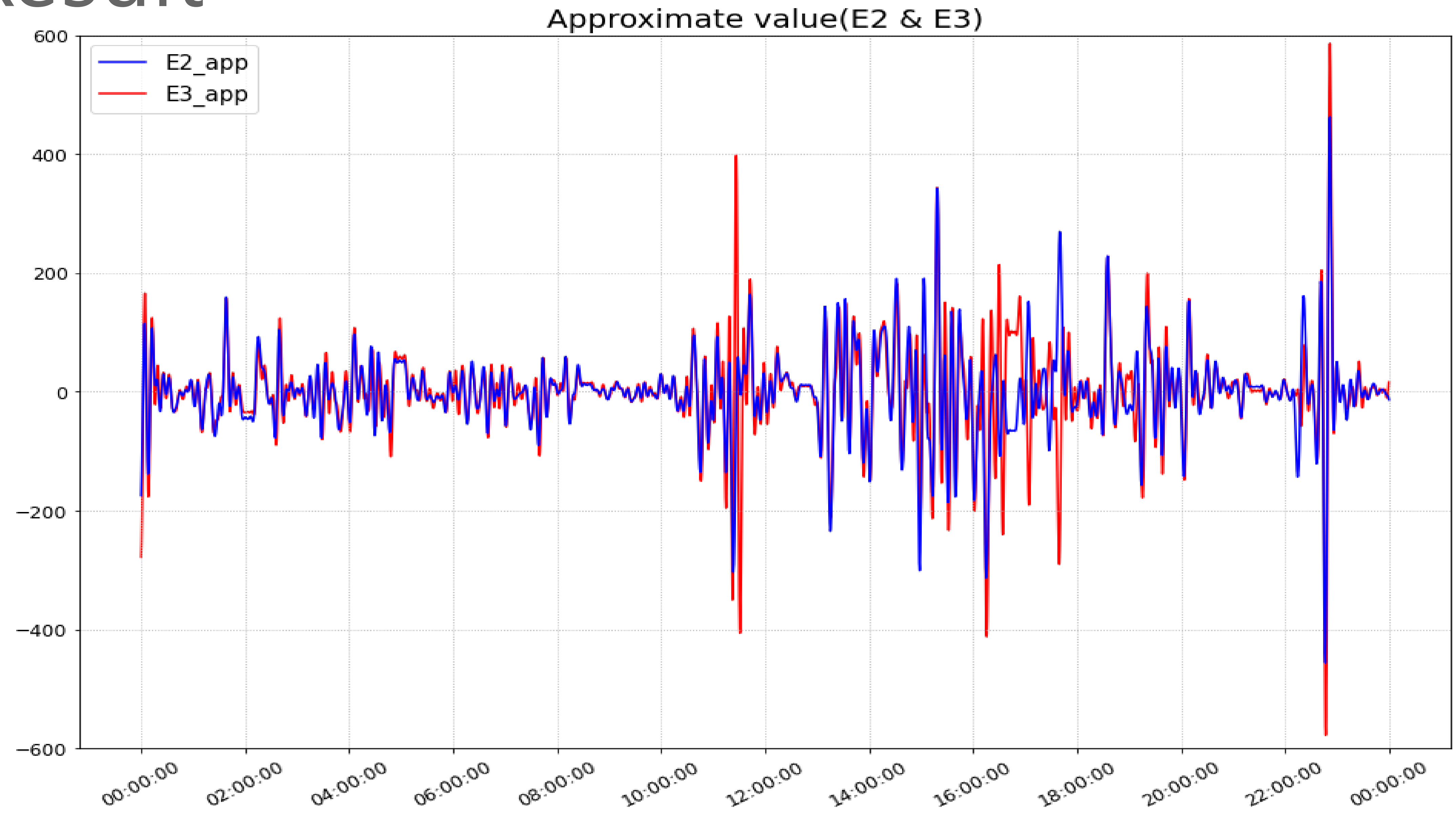
Black line : Reg(SVR), Grey point : Origin point, Blue point : Inter point

Result



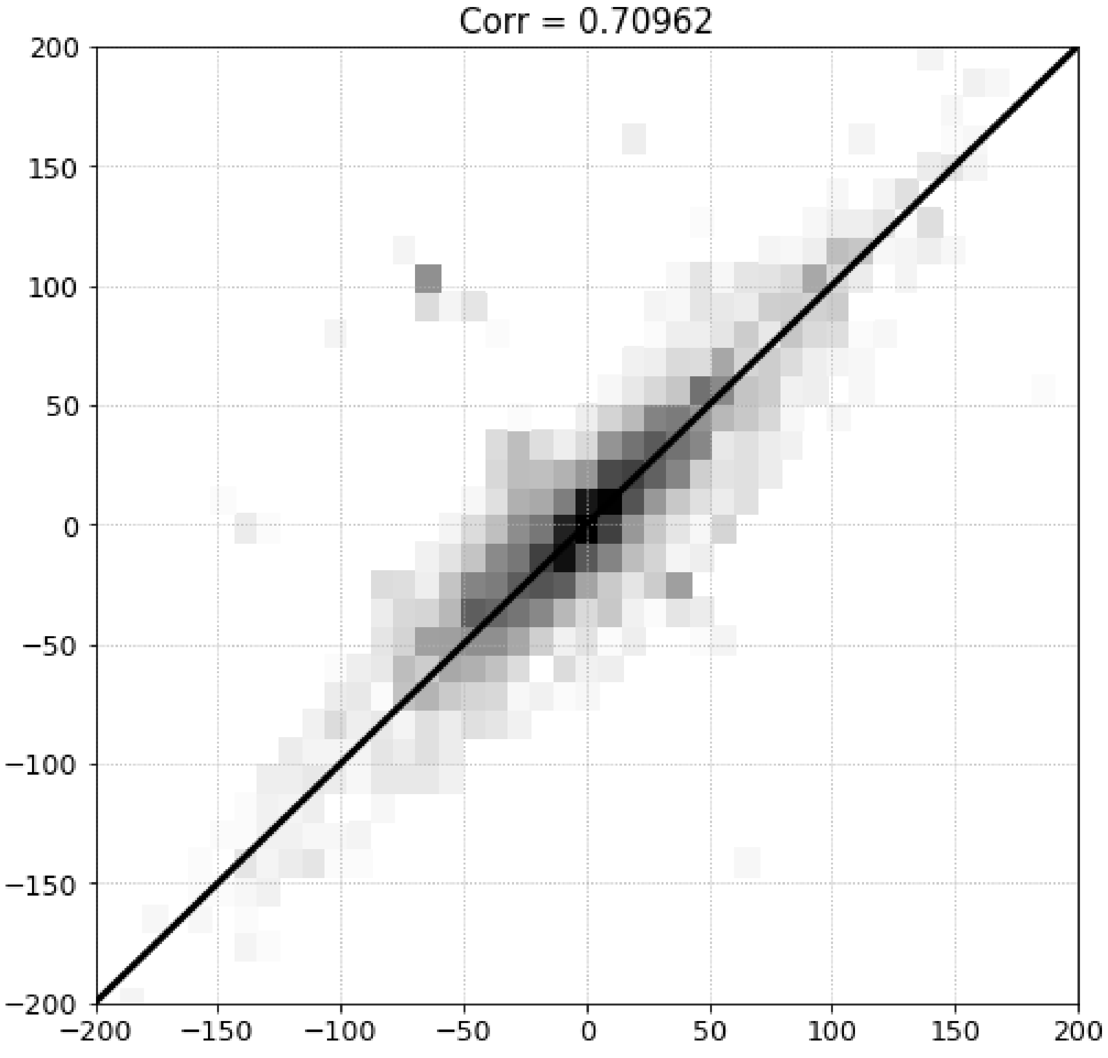
Black line : Reg(SVR), Grey point : Origin point, Red point : Inter point

Result

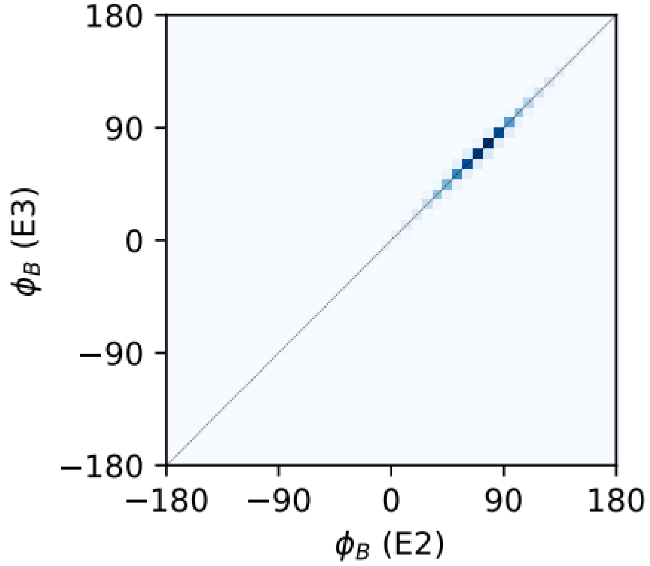


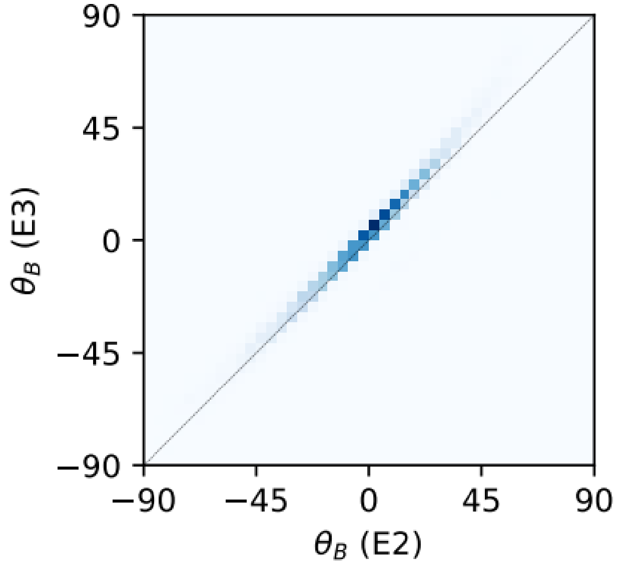
Approximate value(E2 & E3)
E2_app
E3_app
600
400
200
0
−200
−400
−600
00:00:00
02:00:00
04:00:00
06:00:00
08:00:00
10:00:00
12:00:00
14:00:00
16:00:00
18:00:00
20:00:00
22:00:00
00:00:00

Result



If the count in bin is less than five, it is not drawn.





It can be seen that it is more widely spread than the values considered for azimuthal and elevation.

Discussion

- The time from 16H to 18H tends to not fit well. This is thought to be the difference between grid period or by measurement method.

- The Pearson correlation coefficient is 0.7 to 0.9 and indicates good correlation just by the magnitude of each vector component other than the azimuthal and elevation.

- It seems that data from empty part can be interpolated through machine learning (Requires further study).

- 'Corefit' using E2 and E3 data by Stansby is thought to be reasonable enough data.

- Regardless of the grid period, the data in E2 and E3 have similar values and the approximate differences have similar values, so the two data can be used interchangeably.