

Final Term Project

To explore the feasibility of the magnetic field data in corefit of Stansby

202050135 홍지민

I. 서론

헬리오스 위성은 각각 1974년과 1976년에 발사하여 태양 및 태양풍을 관측, 측정 했습니다. 특히, E2와 E3에 해당하는 실험은 자력계를 이용하여 위성 근처의 자기장 변화를 측정 하였습니다. E2는 자기장 필드의 변동을 관측하기 위한 실험으로 평균 8초, 40.5초의 주기로 샘플링 됐습니다. E3의 실험은 평균적인 자기장 필드를 측정하였고 평균 6초의 주기로 샘플링 됐습니다. 이 두 자료를 사용하여 Stansby는 'Corefit'의 자기장 자료를 만들어 냈습니다. 여기서 저는 40.5초 샘플링 데이터(E2)와 6초의 샘플링 데이터(E3)를 사용하여 두 데이터의 연관성을 파악하고 'Corefit'의 자기장 데이터가 타당한 지 파악하였습니다.

II. Origin 데이터 비교

처음엔 1976년 4월 1일의 origin 데이터끼리 비교를 진행하였습니다. 각각 E2, E3 데이터를 산포도로 그린 후에 피어슨 상관계수를 이용하여 데이터 간의 연관성을 찾으려고 했습니다. 좌표계는 위성에서 사용한 SSE 좌표계로 x, y, z 의 세 성분과 자기장의 총 크기까지 총 네 성분으로 분석을 진행 했습니다. x 성분의 경우 16시에서 18시까지의 데이터가 서로 상반되는 것을 관측할 수가 있습니다. 이는 뒤에 서술할 차분으로 구한 근삿값에서도 확인할 수 있습니다. y, z 성분은 잘 맞는 것을 확인할 수가 있으며 크기 성분도 큰 차이 없이 잘 맞는 것을 알 수가 있습니다.

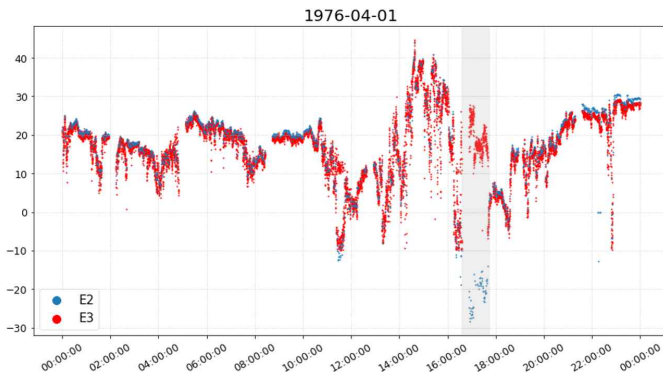
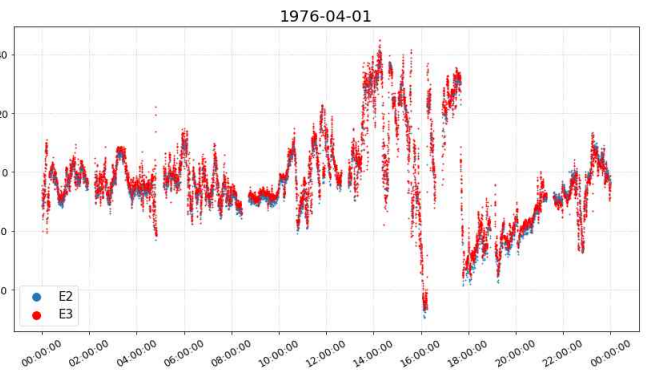
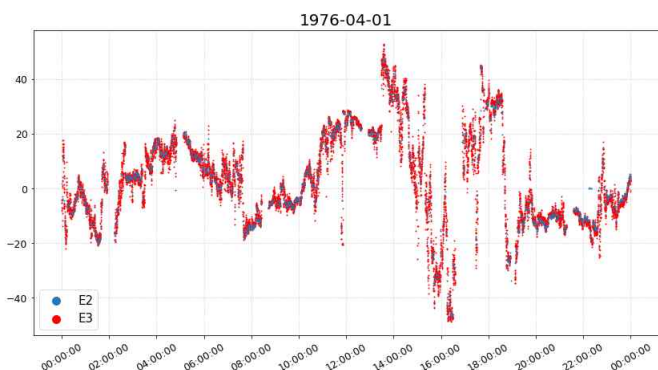
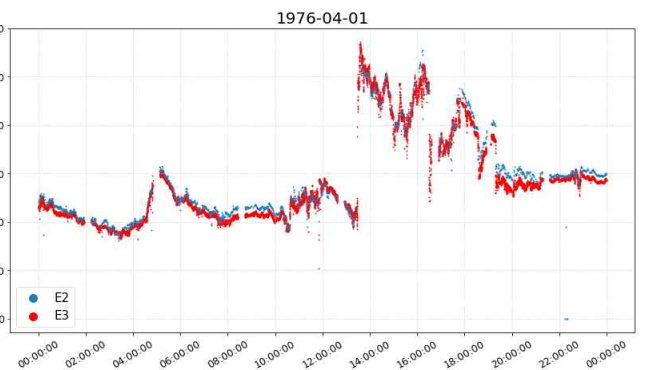
그림 1) x 성분(SSE 좌표계)그림 2) y 성분(SSE 좌표계)그림 3) z 성분(SSE 좌표계)

그림 4) 총 크기(SSE좌표계)

그림 1)을 보면 16시에서 18시에 해당하는 시간에서 서로 다른 값을 나타냄을 확인할 수 있다.

Ⅲ. Origin 데이터의 상관계수 파악

또한 같은 시간대의 E2 데이터 값을 x 축으로 E3 데이터 값을 y 축으로 산포도를 그리고 피어슨 상관계수를 계산 했습니다. 산포도의 데이터를 좀 더 보기 쉽게 2D 히스토그램으로 나타내어 명확한 관계를 파악했습니다. 피어슨 상관계수가 0.7 이상으로 꽤 괜찮은 양의 상관계수를 가지는 것으로 나타났습니다.

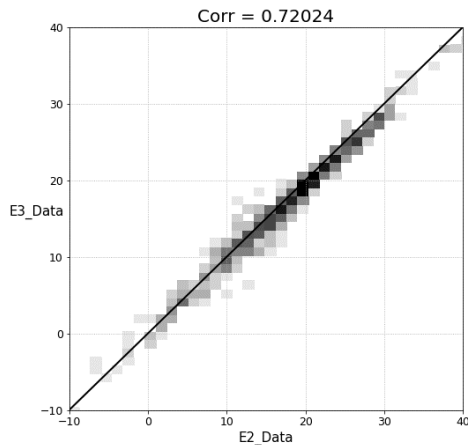


그림 5) x 성분(SSE 좌표계)

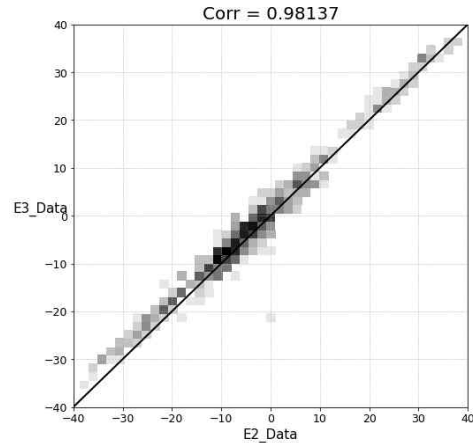


그림 6) y 성분(SSE 좌표계)

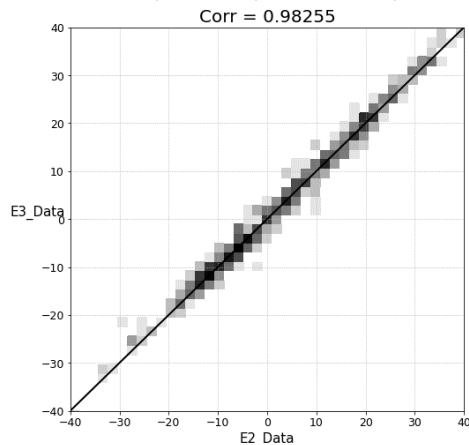


그림 7) z 성분(SSE 좌표계)

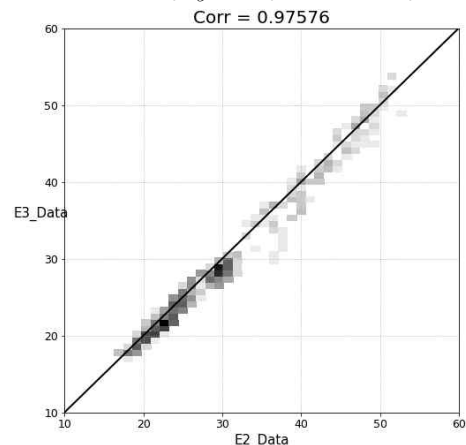


그림 8) 총 크기(SSE좌표계)

검은색 실선이 나타내고 있는 것은 $y = x$ 의 그래프입니다. 피어슨 상관계수는 Corr로 나타냈습니다.

Ⅳ. 문제점

데이터를 파악하던 중 일정 간격이 빈 데이터거나 샘플링 간격이 일정하지 않는 곳을 발견 했습니다. 이것을 해결하기 위해서 데이터 전체를 각각 40초와 6초로 보간법을 시행하였습니다. 그리고 origin 자료에서 너무 많은 변동이 관측 되므로 머신러닝을 통해서 얻은 회귀선의 데이터로 차분 근사값을 구했습니다.

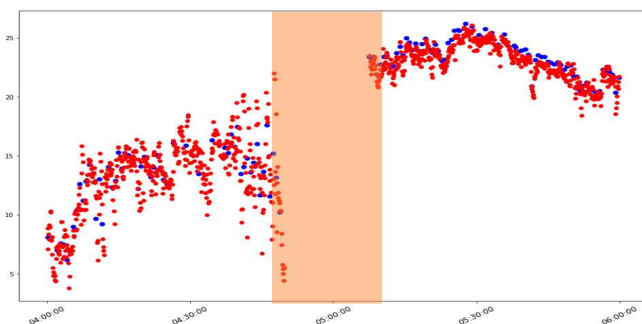


그림 9) 데이터 중 비어있는 부분

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

그림 10) 일정하지 않은 샘플링의 개수

5시에 큰 공백이 존재하며 E2의 데이터는 약 8.6%, 샘플링 간격이 일정하지 않은 구간이 있습니다.

V. 차분법 적용 방법

수업에서 배운 중앙 차분법(4계 정확도)은 책에서 주기 함수로 가정한 후에 주기성을 이용하여 각각 끝 값에도 중앙 차분법을 적용할 수 있었습니다. 하지만 여기서 사용한 데이터는 주기성을 가지지 않으므로 끝 값에 해당하는 부분은 중앙 차분법을 적용할 수 없습니다. 그리하여 검색을 통해 전방(후방) 차분법의 방법을 적용해서 근삿값을 구했습니다. 첫 번째와 두 번째 행에 대해서는 전방 차분법에 대한 행렬을 N-1번째와 N번째 행에 대해서는 후방 차분법에 대한 행렬을 적용 하였습니다. 나머지의 값들은 중앙 차분법을 적용할 수 있으므로 해당하는 부분을 중앙 차분법의 행렬로 대체 했습니다.

$$L = h^{-1} \begin{bmatrix} \text{Forward finite difference Matrix} \\ \text{Backward finite difference Matrix} \end{bmatrix}$$

1~2 rows

(N-1)~N rows

Forward finite difference Matrix

Backward finite difference Matrix

Matrix elements:

- Row 1: $\frac{2}{3}$
- Row 2: 0
- Row 3: $-\frac{2}{3}$

그림 11) 적용된 차분 행렬의 모식도

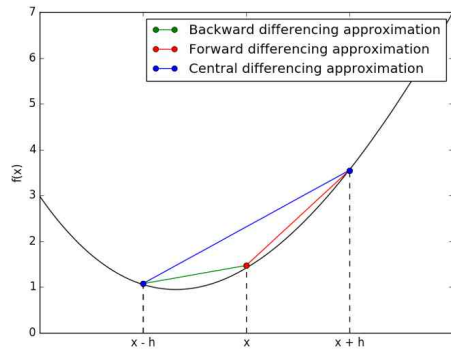


그림 12) 전방(후방) 차분의 모식도

VI. 근삿값 계산

이렇게 구한 근삿값들을 시간에 따라서 그래프를 그렸습니다. 또한 상관성을 파악하기 위해서 앞에서 시행했던 방법처럼 산포도를 바탕으로 2D 히스토그램을 그렸습니다. 16시에서 18시에 해당하는 x 성분에서 오차가 많이 나는 것을 찾을 수 있었습니다.

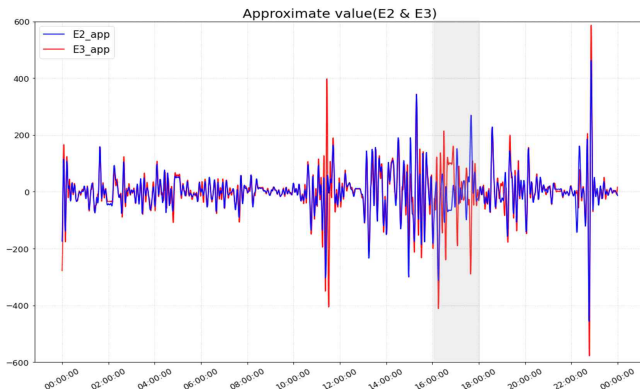


그림 13) x 성분(SSE 좌표계)

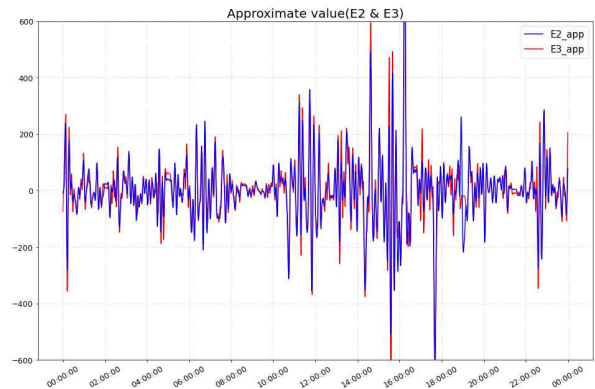


그림 14) y 성분(SSE 좌표계)

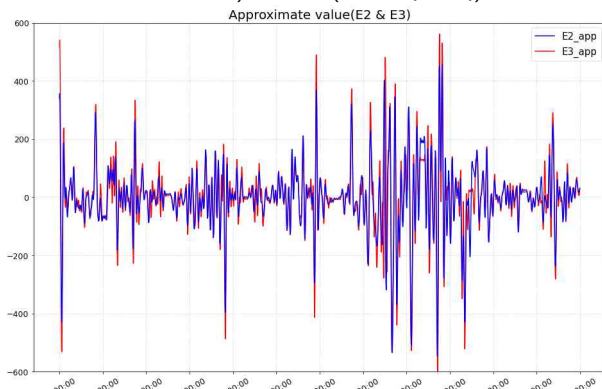


그림 15) z 성분(SSE 좌표계)

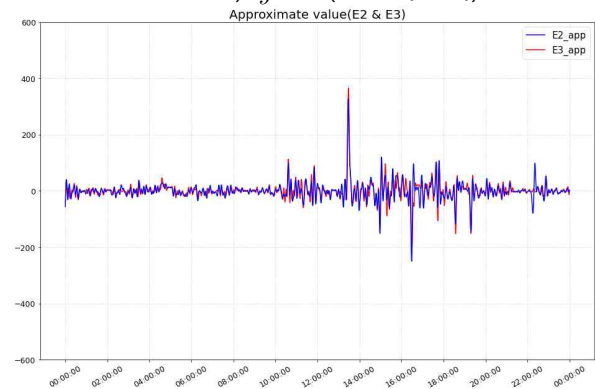


그림 16) 총 크기(SSE좌표계)

그림 13)의 16시~18시 잘 맞지 않는 경향을 보인다.

VII. 근삿값의 상관계수 파악

이는 E2와 E3에서 상반되는 데이터가 관측되기 때문에 발생한 것으로 생각됩니다. y, z 성분은 꽤 높은 상관 계수를 가지고 있으므로 E2와 E3를 이용한 'Corefit' 데이터가 충분히 타당하다고 생각합니다.

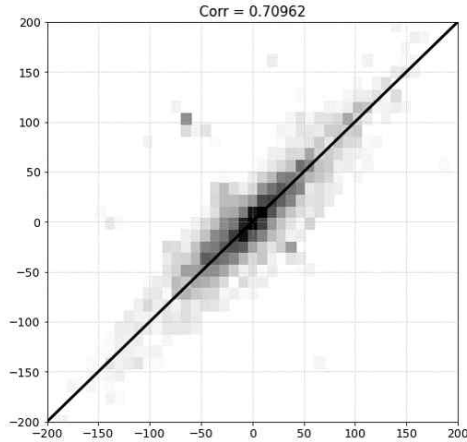


그림 17) x 성분(SSE 좌표계)

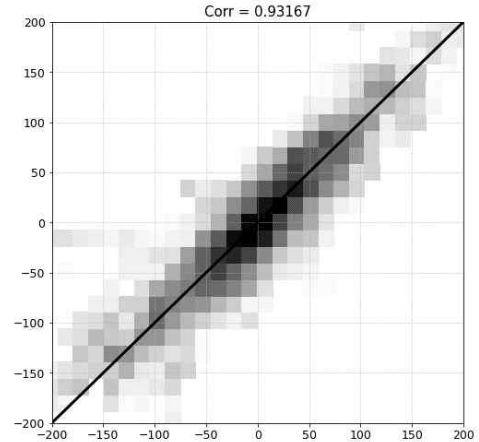


그림 18) y 성분(SSE 좌표계)

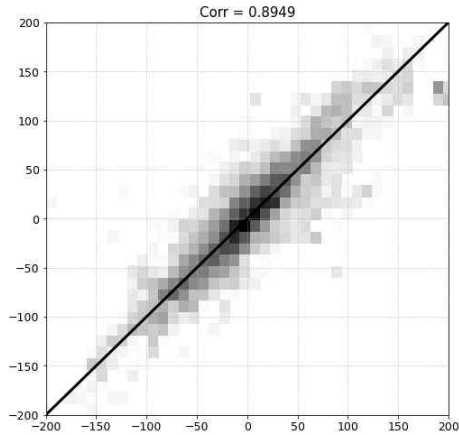


그림 19) z 성분(SSE 좌표계)

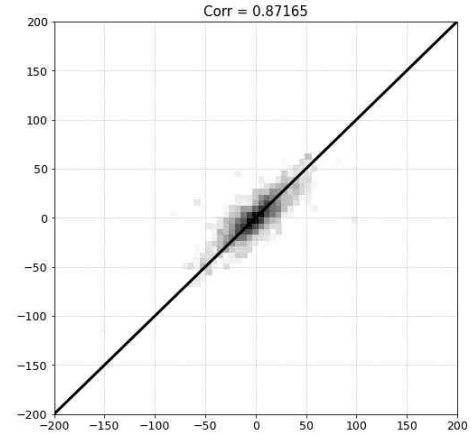


그림 20) 총 크기(SSE좌표계)

검은색 실선이 나타내고 있는 것은 $y = x$ 의 그래프입니다. 피어슨 상관계수는 Corr로 나타냈습니다.

VIII. Stansby와 비교 및 결론

Stansby는 E2와 E3의 각각 자기장 데이터를 Azimuthal과 Elevation으로 상관계수를 파악 했습니다. Stansby는 1일부터 10일까지의 10일간의 데이터로 파악 했습니다. 현재 저는 1일의 데이터만으로 파악을 하였습니다. 추후에 기간을 늘려서 다른 날짜에도 'Corefit' 데이터가 타당한 지 검증할 필요가 있다고 생각합니다. 또한 머신러닝을 통한 회귀가 좋은 성능을 가지고 있음을 알 수 있었습니다. 빈 데이터 구간을 머신러닝의 방법으로 학습을 통해 채울 수 있다면 더욱 좋은 성능을 낼 것으로 예상됩니다.

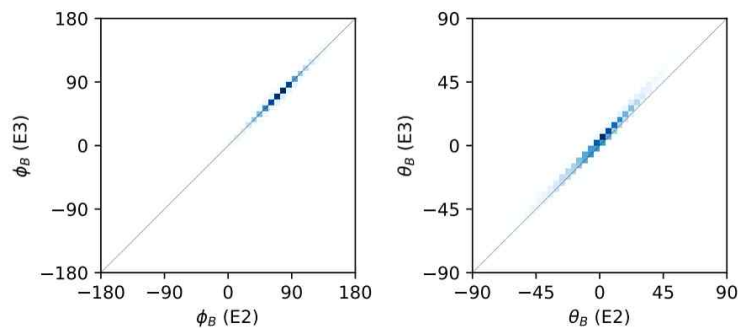


그림 21) Stansby가 파악한 E2와 E3의 상관관계

회색 실선이 나타내고 있는 것은 $y = x$ 의 그래프입니다.