stat_decomposition

July 24, 2021

1 Statistical Decomposition for Satellite Images

Our purpose is to explore visualizations and the distributions of those images, and then finally, decompose them using PCA and Factorial Analysis.

1.1 Loading images

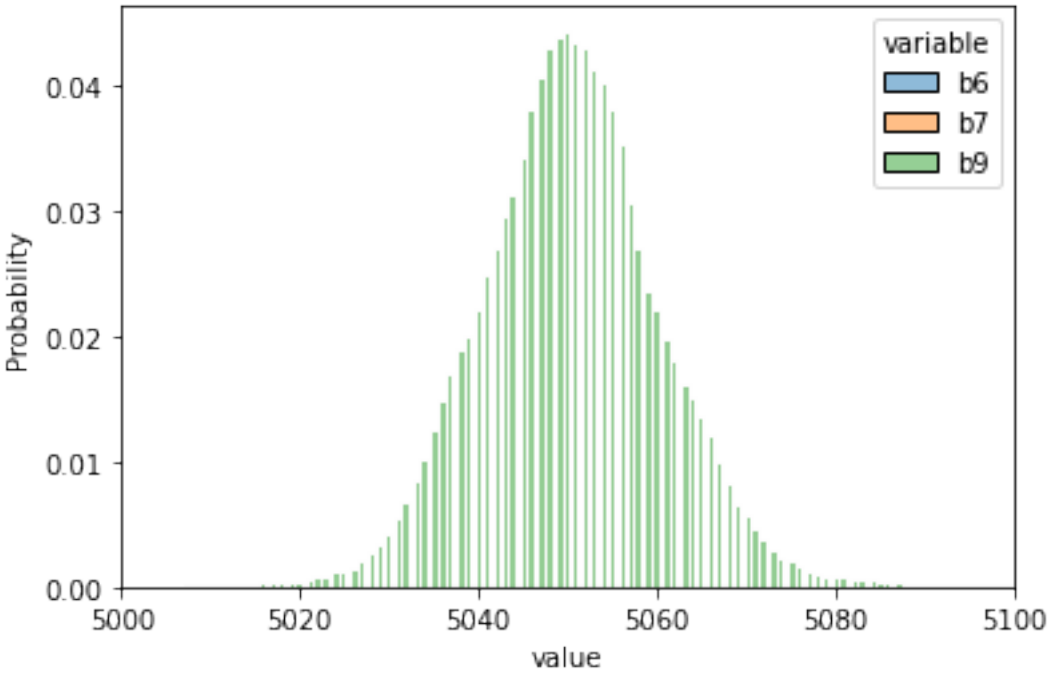
```
Lowered all files for data.

['data/lc08_l1tp_139045_20170304_20170316_01_t1_b10.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b9.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b2.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b5.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b1.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b6.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b4.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b11.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b3.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b7.tif',
 'data/lc08_l1tp_139045_20170304_20170316_01_t1_b8.tif']

(7771, 7611, 10)
uint16
Maximum value for the pixel: 65535
```

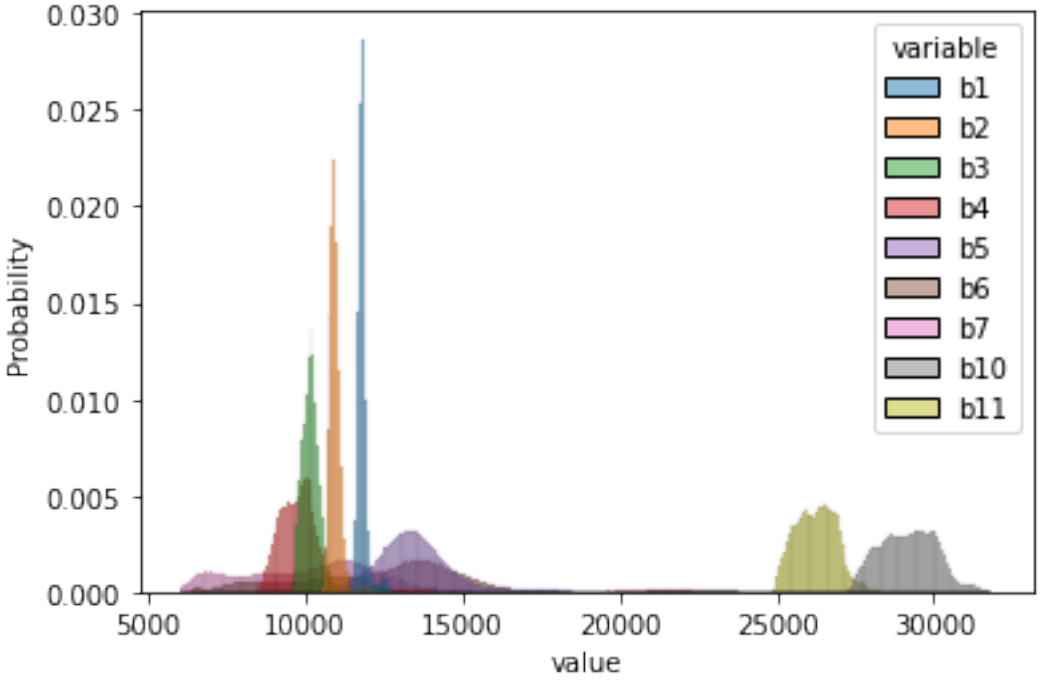
1.2 Cropping Figure

1.3 Exploratory Data Analysis

	b1	b2	b3	b4	b5	b6	b7	b9	\
0	11797.0	10866.0	9907.0	9229.0	13254.0	11461.0	8732.0	5031.0	
1	11810.0	10898.0	9933.0	9404.0	13029.0	12277.0	9517.0	5042.0	
2	11858.0	10977.0	10068.0	9704.0	13292.0	13226.0	10404.0	5059.0	
3	11842.0	10957.0	10053.0	9670.0	12832.0	12994.0	10243.0	5057.0	
4	11845.0	10959.0	10050.0	9671.0	13192.0	12669.0	10048.0	5044.0	
...	...	...	...	...	...	...	...	...	
249995	11603.0	10697.0	9896.0	9550.0	13257.0	12653.0	10196.0	5048.0	
249996	11525.0	10597.0	9796.0	9113.0	14050.0	11737.0	9198.0	5038.0	

```
249997  11556.0  10617.0   9790.0  9343.0  13285.0  12038.0   9563.0  5050.0
249998  11563.0  10664.0   9889.0  9346.0  13628.0  11801.0   9408.0  5044.0
249999  11532.0  10610.0   9811.0  9271.0  13477.0  11879.0   9349.0  5049.0

            b10      b11   index
0       28591.0  25865.0       0
1       28564.0  25830.0       1
2       28580.0  25857.0       2
3       28643.0  25934.0       3
4       28729.0  26017.0       4
...         ...      ...     ...
249995  28889.0  26031.0  249995
249996  28983.0  26095.0  249996
249997  29088.0  26157.0  249997
249998  29200.0  26216.0  249998
249999  29316.0  26275.0  249999

[250000 rows x 11 columns]

          index variable    value
0             0       b1  11797.0
1             1       b1  11810.0
2             2       b1  11858.0
3             3       b1  11842.0
4             4       b1  11845.0
...         ...      ...      ...
2499995  249995      b11  26031.0
2499996  249996      b11  26095.0
2499997  249997      b11  26157.0
2499998  249998      b11  26216.0
2499999  249999      b11  26275.0

[2500000 rows x 3 columns]
```


variable
b6
b7
b9
Probability
0.00
0.01
0.02
0.03
0.04
5000
5020
5040
5060
5080
5100
value


variable
b1
b2
b3
b4
b5
b6
b7
b10
b11
Probability
0.000
0.005
0.010
0.015
0.020
0.025
0.030
5000
10000
15000
20000
25000
30000
value

1.4 PCA Analysis

```
variable       b1      b10      b11       b2       b3      b4       b5  \
index
0         11797.0  28591.0  25865.0  10866.0   9907.0  9229.0  13254.0
1         11810.0  28564.0  25830.0  10898.0   9933.0  9404.0  13029.0
2         11858.0  28580.0  25857.0  10977.0  10068.0  9704.0  13292.0
3         11842.0  28643.0  25934.0  10957.0  10053.0  9670.0  12832.0
4         11845.0  28729.0  26017.0  10959.0  10050.0  9671.0  13192.0
...           ...      ...      ...      ...      ...     ...      ...
249995    11603.0  28889.0  26031.0  10697.0   9896.0  9550.0  13257.0
249996    11525.0  28983.0  26095.0  10597.0   9796.0  9113.0  14050.0
249997    11556.0  29088.0  26157.0  10617.0   9790.0  9343.0  13285.0
249998    11563.0  29200.0  26216.0  10664.0   9889.0  9346.0  13628.0
249999    11532.0  29316.0  26275.0  10610.0   9811.0  9271.0  13477.0

variable       b6       b7      b9
index
0         11461.0   8732.0  5031.0
1         12277.0   9517.0  5042.0
2         13226.0  10404.0  5059.0
3         12994.0  10243.0  5057.0
4         12669.0  10048.0  5044.0
...           ...      ...     ...
249995    12653.0  10196.0  5048.0
249996    11737.0   9198.0  5038.0
249997    12038.0   9563.0  5050.0
249998    11801.0   9408.0  5044.0
249999    11879.0   9349.0  5049.0

[250000 rows x 10 columns]
```

1.4.1 Normalization

Mean of Columns:

```
variable
b1     11817.291212
b10    29196.351072
b11    26264.952468
b2     10948.185272
b3     10225.000512
b4      9855.851880
b5     13440.530704
b6     12655.372212
b7     10397.892800
b9      5050.385156
```

```
dtype: float64
```

```
Std. Dev. of Columns:
```

```
variable
b1      192.760962
b10     887.267230
b11     641.840973
b2      258.958874
b3      440.860891
b4      823.200569
b5     1403.974276
b6     2892.943615
b7     2736.677759
b9        9.692931
dtype: float64
```

```
Mean of Columns:
```

```
variable
b1     1.052518e-16
b10   -2.545218e-15
b11   -6.706966e-16
b2    -2.698668e-15
b3    -1.202562e-15
b4     1.307465e-16
b5    -8.780990e-17
b6    -9.088361e-16
b7     2.223435e-16
b9    -9.092658e-15
dtype: float64
```

```
Std. Dev. of Columns:
```

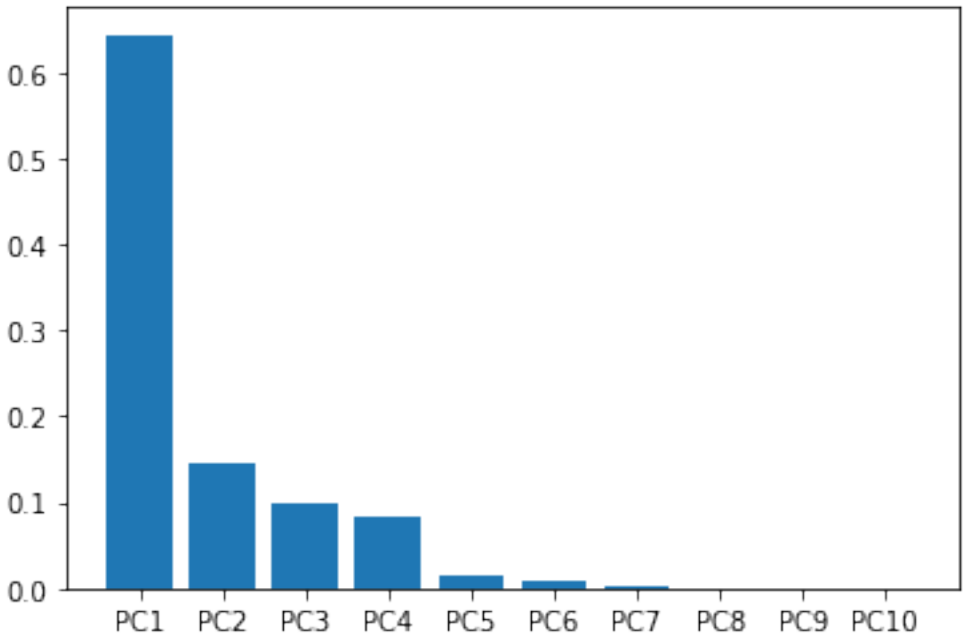
```
variable
b1     1.0
b10    1.0
b11    1.0
b2     1.0
b3     1.0
b4     1.0
b5     1.0
b6     1.0
b7     1.0
b9     1.0
dtype: float64
```

```
variable        b1       b10       b11        b2        b3        b4  \
index
0        -0.105266 -0.682265 -0.623133 -0.317368 -0.721317 -0.761481
1        -0.037825 -0.712695 -0.677664 -0.193796 -0.662342 -0.548896
2         0.211188 -0.694662 -0.635597  0.111271 -0.356123 -0.184465
3         0.128184 -0.623658 -0.515630  0.034039 -0.390147 -0.225767
4         0.143747 -0.526731 -0.386314  0.041762 -0.396952 -0.224553

variable        b5        b6        b7        b9
index
0        -0.132859 -0.412857 -0.608728 -1.999927
1        -0.293118 -0.130791 -0.321884 -0.865079
2        -0.105793  0.197248  0.002232  0.888776
3        -0.433434  0.117053 -0.056599  0.682440
4        -0.177019  0.004711 -0.127853 -0.658744
```

1.4.2 PC estimation and exploration

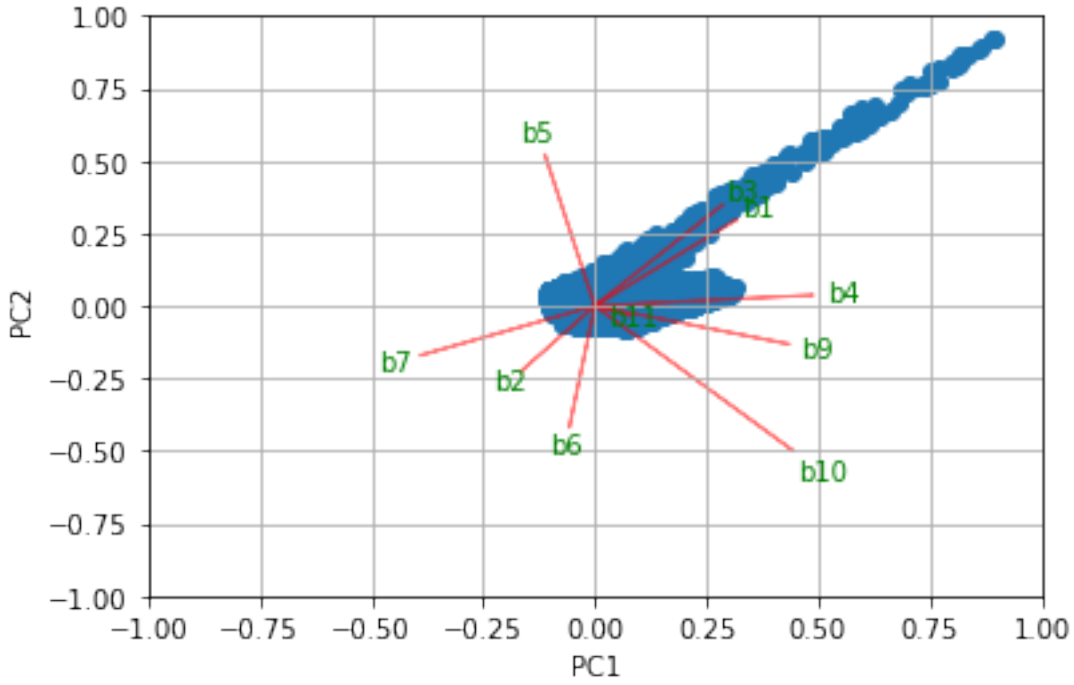
```
PCA(n_components=10)

               PC1       PC2       PC3       PC4       PC5       PC6  \
variable
b1        0.317347  0.297197  0.298769  0.341724  0.362918  0.380150
b10       0.443109 -0.498231 -0.497093  0.379027  0.281173  0.109792
b11       0.076152 -0.033782 -0.029953  0.064878  0.040767  0.029629
b2       -0.165716 -0.226449 -0.221468 -0.164924 -0.083926 -0.143950
b3        0.284582  0.346801  0.322844  0.154369  0.017089 -0.249758
b4        0.485591  0.036777 -0.048136  0.215465 -0.543098 -0.482071
b5       -0.111410  0.519450 -0.544640  0.085337 -0.125379  0.313306
b6       -0.056834 -0.416726  0.404034  0.187690 -0.477314  0.510792
b7       -0.387899 -0.169535  0.178091  0.375913  0.407338 -0.408333
b9        0.432193 -0.131017  0.124708 -0.673854  0.275331  0.035421

               PC7       PC8       PC9      PC10
variable
b1        0.233071  0.362462  0.376506  0.035281
b10       0.039248 -0.188480 -0.142577 -0.127089
b11      -0.144191 -0.086673 -0.060648  0.976438
b2        0.848814  0.242522  0.095253  0.169914
b3        0.404393 -0.429912 -0.509376 -0.013766
b4       -0.129078  0.396982  0.111528  0.008015
b5       -0.005533  0.309829 -0.451595 -0.001354
b6        0.059017  0.136576 -0.337242 -0.005741
b7       -0.114306  0.428629 -0.343972  0.000059
b9       -0.078018  0.354935 -0.342453 -0.003496

PC1: 0.6439
```

```
PC2: 0.1467
PC3: 0.099
PC4: 0.0832
PC5: 0.0159
PC6: 0.0087
PC7: 0.001
PC8: 0.0007
PC9: 0.0005
PC10: 0.0004
```



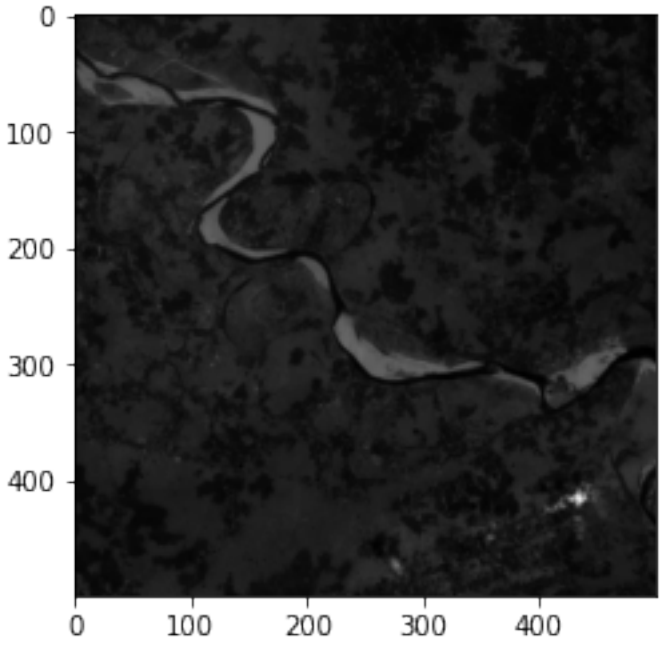

```
              PC1       PC2       PC3       PC4       PC5       PC6       PC7  \
0       -1.562415  0.609884 -1.899809 -0.078258  0.122494  0.413571 -0.028087
1       -1.208980  0.524218 -0.785919  0.026129 -0.267289  0.513689  0.005252
2       -0.411637  0.522821  0.911681  0.410430 -0.469799  0.513241  0.017641
3       -0.570336  0.384522  0.748034  0.064857 -0.497465  0.493759 -0.013019
4       -0.552427  0.491838 -0.590167 -0.033458 -0.210152  0.407038 -0.035174
...           ...       ...       ...       ...       ...       ...       ...
249995 -1.375408 -0.720279 -0.383280  0.459485 -0.635417 -0.153267  0.070941
249996 -2.015455 -1.005404 -1.509943  0.832216 -0.048046 -0.302312  0.010768
249997 -1.768111 -1.202350 -0.221547  0.491351 -0.338202 -0.198739  0.070891
249998 -1.562621 -1.050738 -0.833876  0.464015 -0.052082 -0.340831  0.070507
249999 -1.722653 -1.443406 -0.358569  0.503899 -0.086494 -0.297146  0.100101

              PC8       PC9      PC10
0       0.086845 -0.009167  0.033750
1       0.103105 -0.015590  0.011003
```

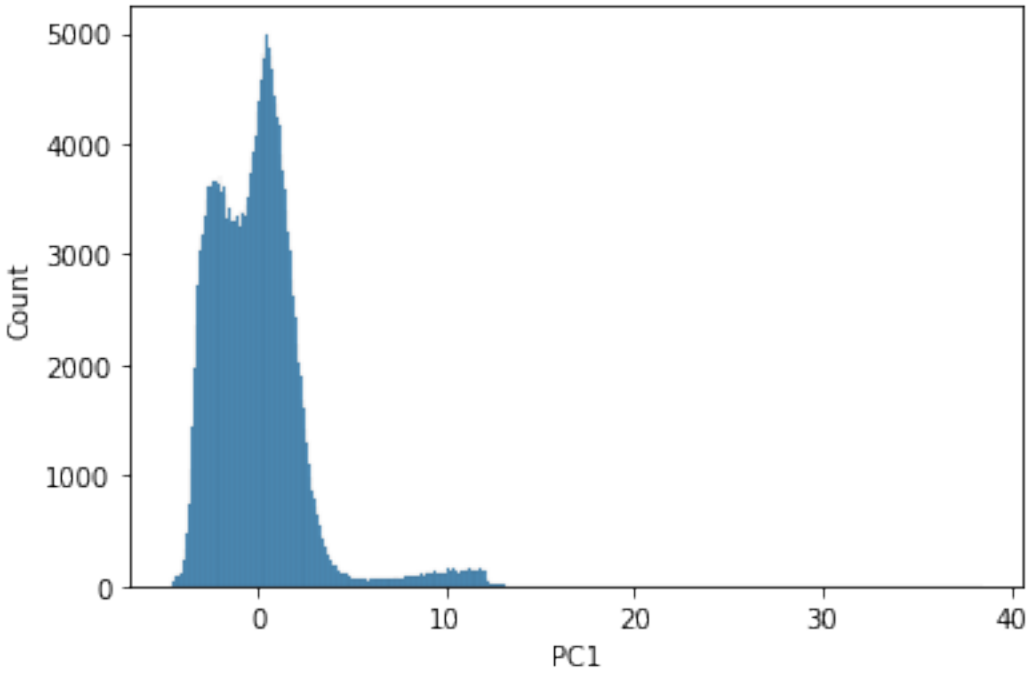
```
2       0.132162 -0.009331 -0.002152
3       0.127144  0.029468  0.026811
4       0.174953 -0.023369 -0.000859
...          ...       ...       ...
249995  0.063099 -0.051910 -0.009512
249996 -0.034135  0.009152 -0.034841
249997  0.040109 -0.085611  0.014750
249998 -0.022793  0.009888 -0.061552
249999 -0.030578 -0.015481 -0.024606

[250000 rows x 10 columns]
```



1.4.3 Compressed Image through PCA

```
(250000,)
(500, 500)
```


0
100
200
300
400
0
100
200
300
400


5000
4000
3000
2000
1000
0
Count
0
10
20
30
40
PC1

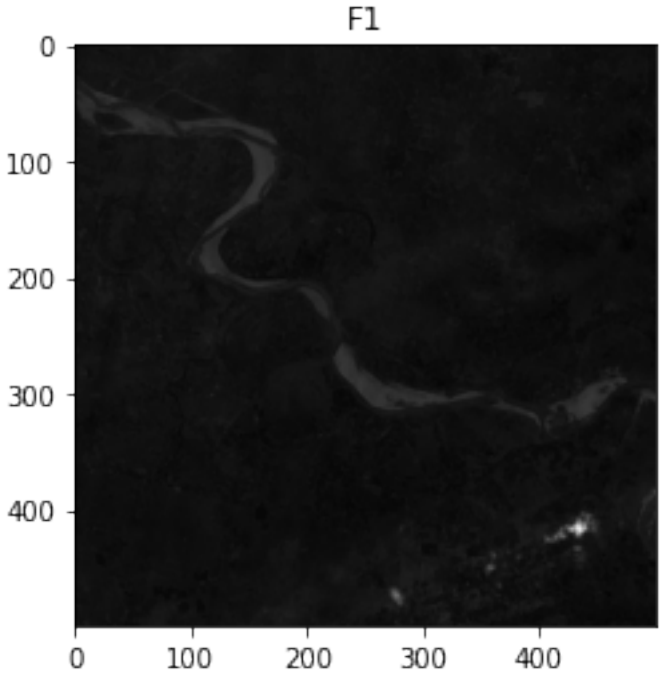
1.5 Factorial Analysis

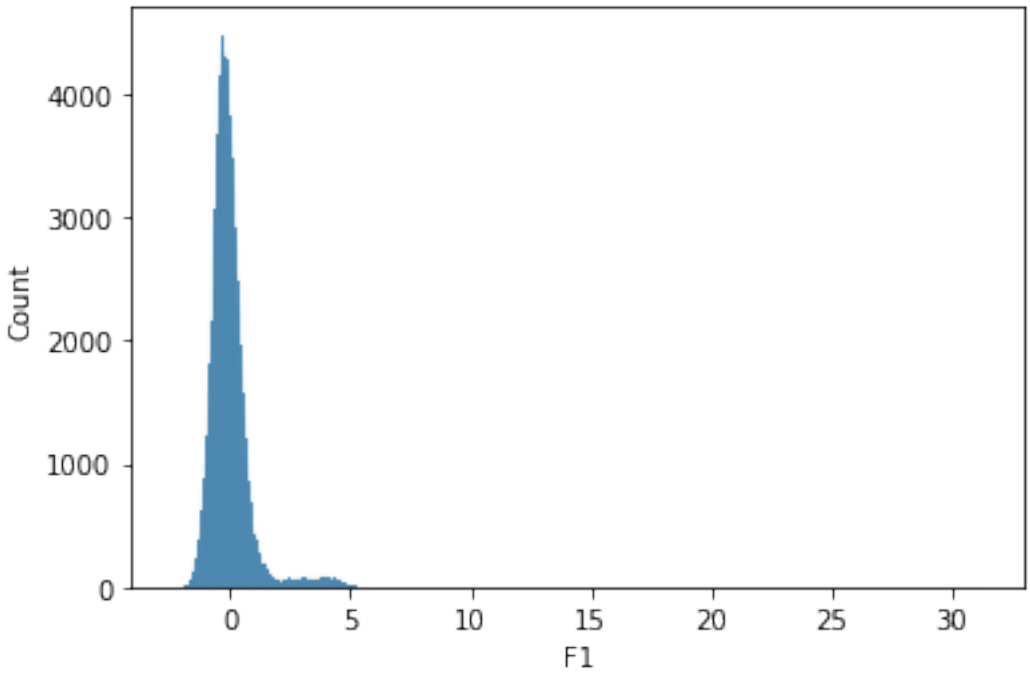
```
variable        b1       b10       b11        b2        b3        b4  \
index
0        -0.105266 -0.682265 -0.623133 -0.317368 -0.721317 -0.761481
1        -0.037825 -0.712695 -0.677664 -0.193796 -0.662342 -0.548896
2         0.211188 -0.694662 -0.635597  0.111271 -0.356123 -0.184465
3         0.128184 -0.623658 -0.515630  0.034039 -0.390147 -0.225767
4         0.143747 -0.526731 -0.386314  0.041762 -0.396952 -0.224553
...            ...       ...       ...       ...       ...       ...
249995   -1.111694 -0.346402 -0.364502 -0.969981 -0.746268 -0.371540
249996   -1.516340 -0.240459 -0.264789 -1.356143 -0.973097 -0.902395
249997   -1.355519 -0.122118 -0.168192 -1.278911 -0.986707 -0.622997
249998   -1.319205  0.004113 -0.076269 -1.097415 -0.762146 -0.619353
249999   -1.480026  0.134851  0.015654 -1.305942 -0.939073 -0.710461

variable        b5        b6        b7        b9
index
0        -0.132859 -0.412857 -0.608728 -1.999927
1        -0.293118 -0.130791 -0.321884 -0.865079
2        -0.105793  0.197248  0.002232  0.888776
3        -0.433434  0.117053 -0.056599  0.682440
4        -0.177019  0.004711 -0.127853 -0.658744
...            ...       ...       ...       ...
249995   -0.130722 -0.000820 -0.073773 -0.246072
249996    0.434103 -0.317453 -0.438449 -1.277751
249997   -0.110779 -0.213406 -0.305075 -0.039736
249998    0.133528 -0.295330 -0.361713 -0.658744
249999    0.025976 -0.268368 -0.383272 -0.142904

[250000 rows x 10 columns]

array([[0.94681213, 0.12387987],
       [0.19370344, 0.92915095],
       [0.19574387, 0.93492442],
       [0.96323093, 0.21229673],
       [0.92414434, 0.33998778],
       [0.81899463, 0.52723959],
       [0.4152459 , 0.32890942],
       [0.54382131, 0.76099074],
       [0.60936102, 0.75527474],
       [0.02494798, 0.08875818]])

                F1        F2
variable
b1        0.946812  0.123880
b10       0.193703  0.929151
```

```
b11       0.195744  0.934924
b2        0.963231  0.212297
b3        0.924144  0.339988
b4        0.818995  0.527240
b5        0.415246  0.328909
b6        0.543821  0.760991
b7        0.609361  0.755275
b9        0.024948  0.088758

                F1        F2
variable
b2        0.963231  0.000000
b1        0.946812  0.000000
b3        0.924144  0.000000
b4        0.818995  0.527240
b7        0.609361  0.755275
b6        0.543821  0.760991
b11       0.000000  0.934924
b10       0.000000  0.929151
b5        0.000000  0.000000
b9        0.000000  0.000000

                     F1        F2
SS Loadings     4.265013  3.457001
Proportion Var  0.426501  0.345700
Cumulative Var  0.426501  0.772201
```

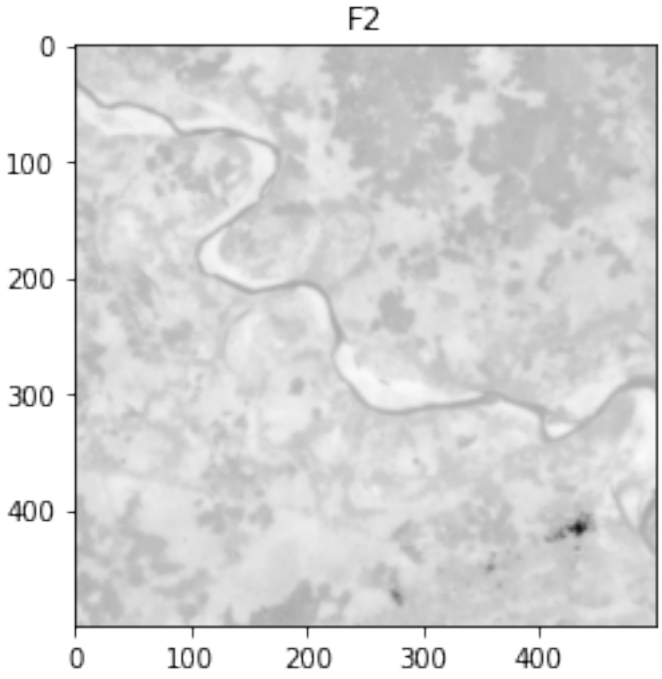
1.6 Compressed image through FA

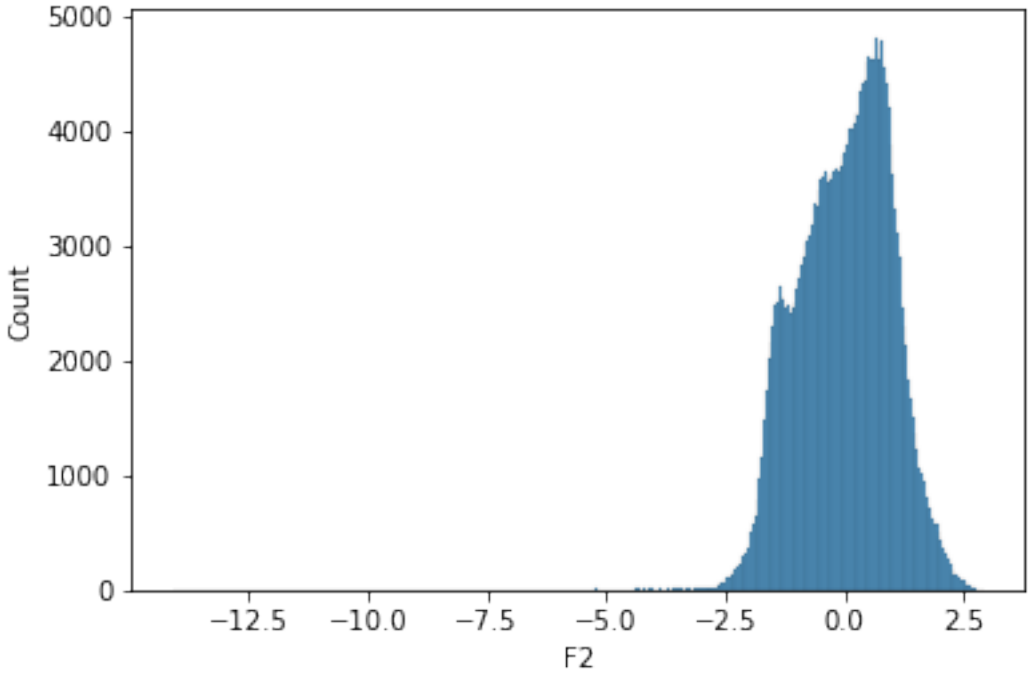
```
               F1        F2
0       -0.423327 -0.574774
1       -0.245363 -0.526909
2        0.122958 -0.489821
3        0.090803 -0.431901
4       -0.116046 -0.345796
...           ...       ...
249995 -0.936696  0.164351
249996 -1.161653  0.151110
249997 -1.452235  0.392258
249998 -0.933572  0.183353
249999 -1.274365  0.394117

[250000 rows x 2 columns]

(250000,)
(500, 500)
```






```
(250000,)
(500, 500)
```


F2
0
100
200
300
400
0
100
200
300
400


5000
4000
3000
2000
1000
0
Count
−12.5
−10.0
−7.5
−5.0
−2.5
0.0
2.5
F2