Spatio-temporal analysis of night light data

BY-
PRANJAL DAVE
RACHIT AGRAWAL
KUMAR ANANT RAJ

Spatio-temporal analysis of Urban sprawl of North East India using DMSP-OLS and VIIRS-DNB night-light data from 1992 to 2017

PROBLEM STATEMENT & MOTIVATION

- Night-time lights are a good indicator of various socio-economic studies e.g.; Urban expansion, GDP estimation and so on.
- Sources are the images taken by the VIIRS and DMSP satellites.
- Data from these images are needed to be calibrated and scaled to study electrification dynamics of surface through time..

OBJECTIVES

- Generation of yearly composite VIIRS image.
- Extraction of Pseudo Invariant Features(PIFs) from DMSP and VIIRS data.
- Calibration between PIF(s) of VIIRS and DMSP/OLS images.
- Spatio-temporal quantification of electrification of North East India.

RELEVANCE

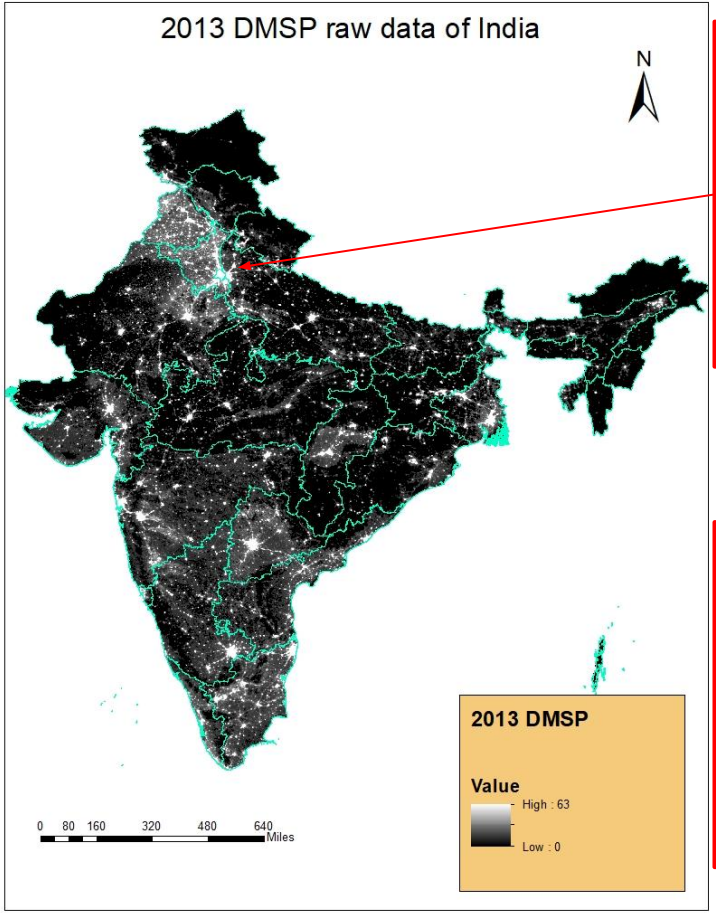
- Generated simulated DMSP image can be used with real DMSP images for time series analysis/change detection.

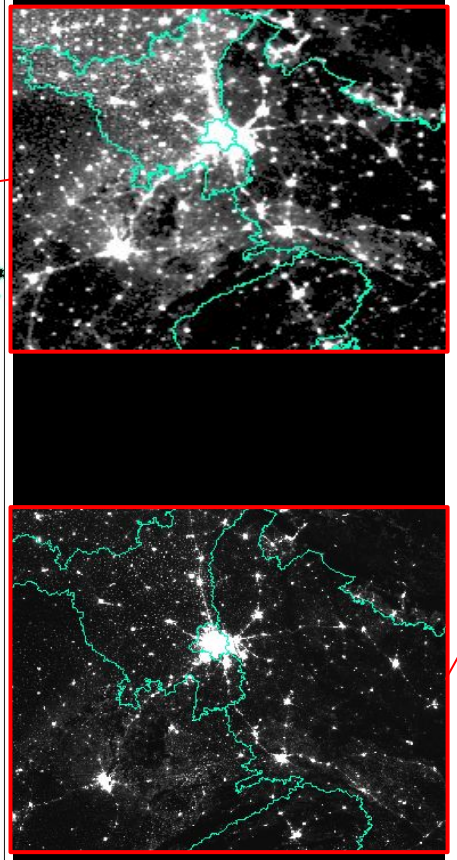
- Night-time images are used to find the degree of electrification over last 25 years in North East India.

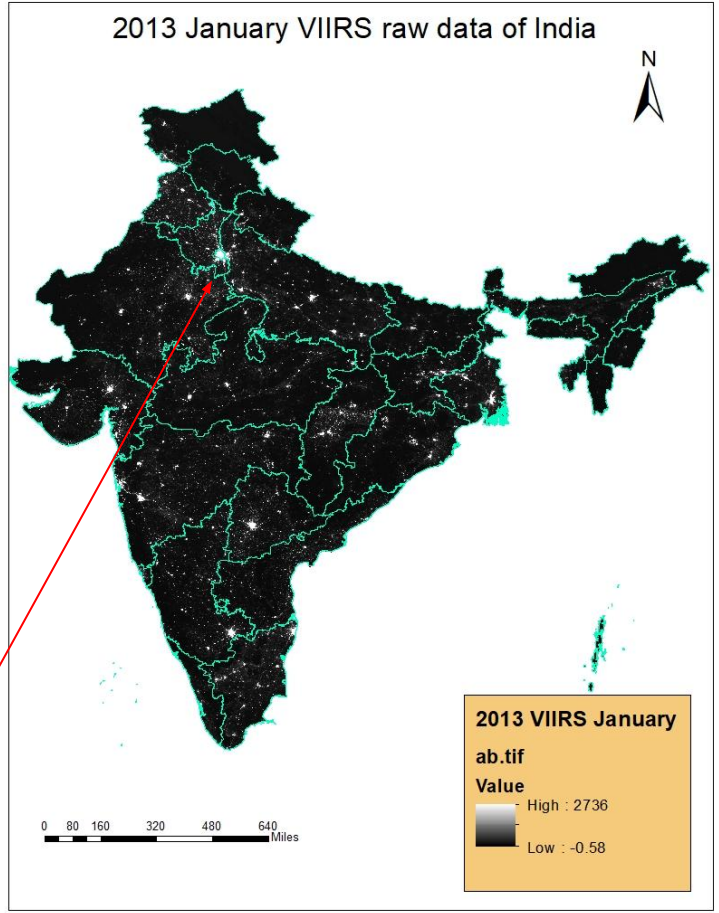
CHALLENGES

- The acquisition period of DMSP-OLS and VIIRS-DNB are non-identical.
- VIIRS-DNB data are provided in radiance values whereas DMSP-OLS data are provided in DNs.
- Spatial, Temporal and Radiometric resolution of DMSP-OLS and VIIRS-DNB are completely different from each other.
- Ground coverage of OLS sensors are provided with varying DMSP satellites(F10, F11, F12, F13, F14, F15,F16, F17, F18)through time;thus, Relative Radiometric Normalization is required.

STUDY AREA

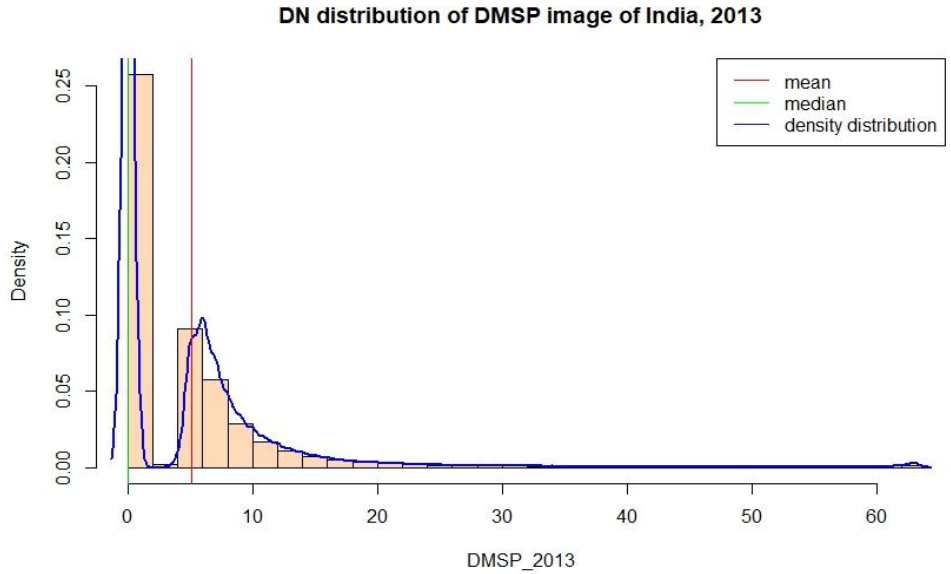

2013 DMSP raw data of India
N
0 80 160 320 480 640
Miles
2013 DMSP
Value
High : 63
Low : 0




2013 January VIIRS raw data of India
N
0 80 160 320 480 640
Miles
2013 VIIRS January
ab.tif
Value
High : 2736
Low : -0.58

DATASETS USED

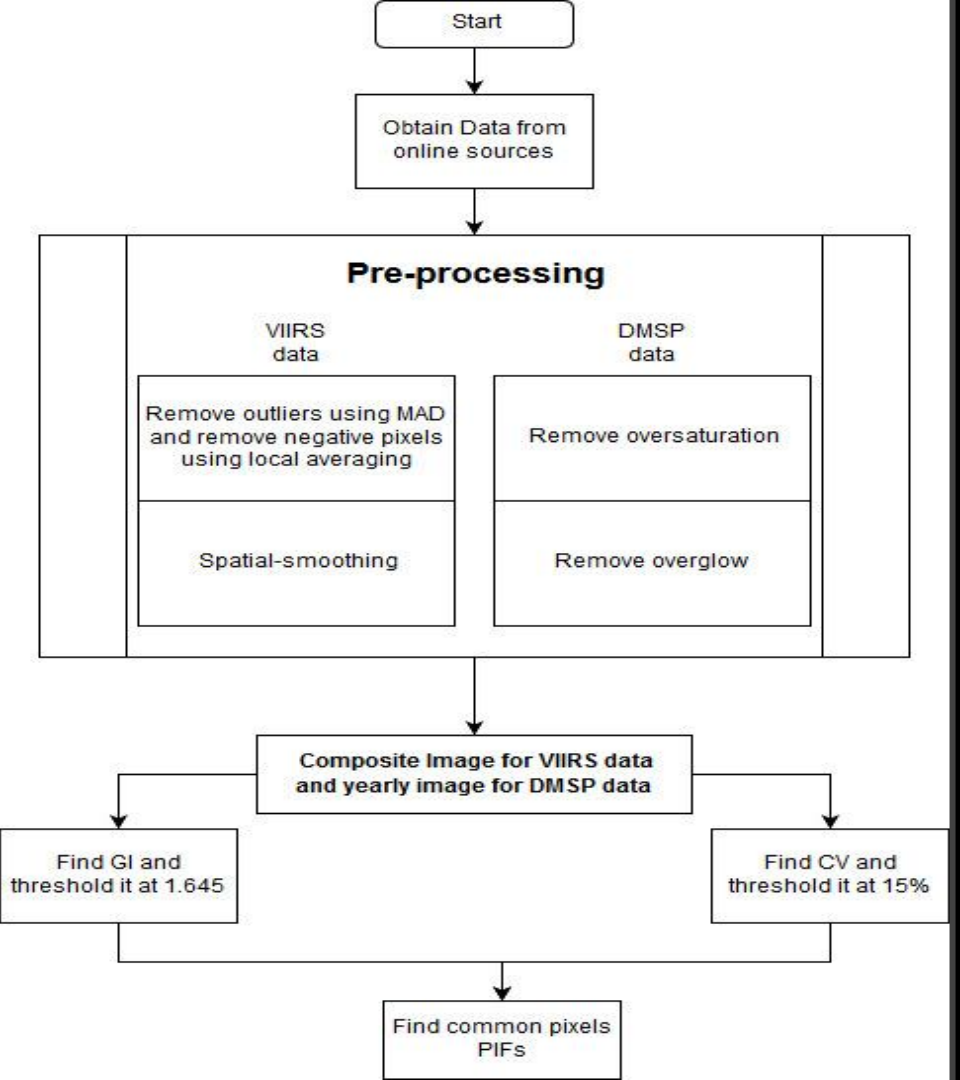
Specification	DMSP-OLS	VIIRS-DNB
Spectral Band	0.5-0.9 μm	0.5-0.9 μm
Number of bits	6 bits	14 bits
Saturation	Urban areas	No saturation problem
Dynamic range(DN)	0 to 63	0 to 16384
Additional spectral bands	Thermal infrared	21 bands spanning 0.4 to 12 um
Calibration	No on-board calibration	Solar Diffuser
Spatial Resolution	1 km	500 m
Swath	3000 km	3040 km
Local night-time overpass	~19:30	~01:30

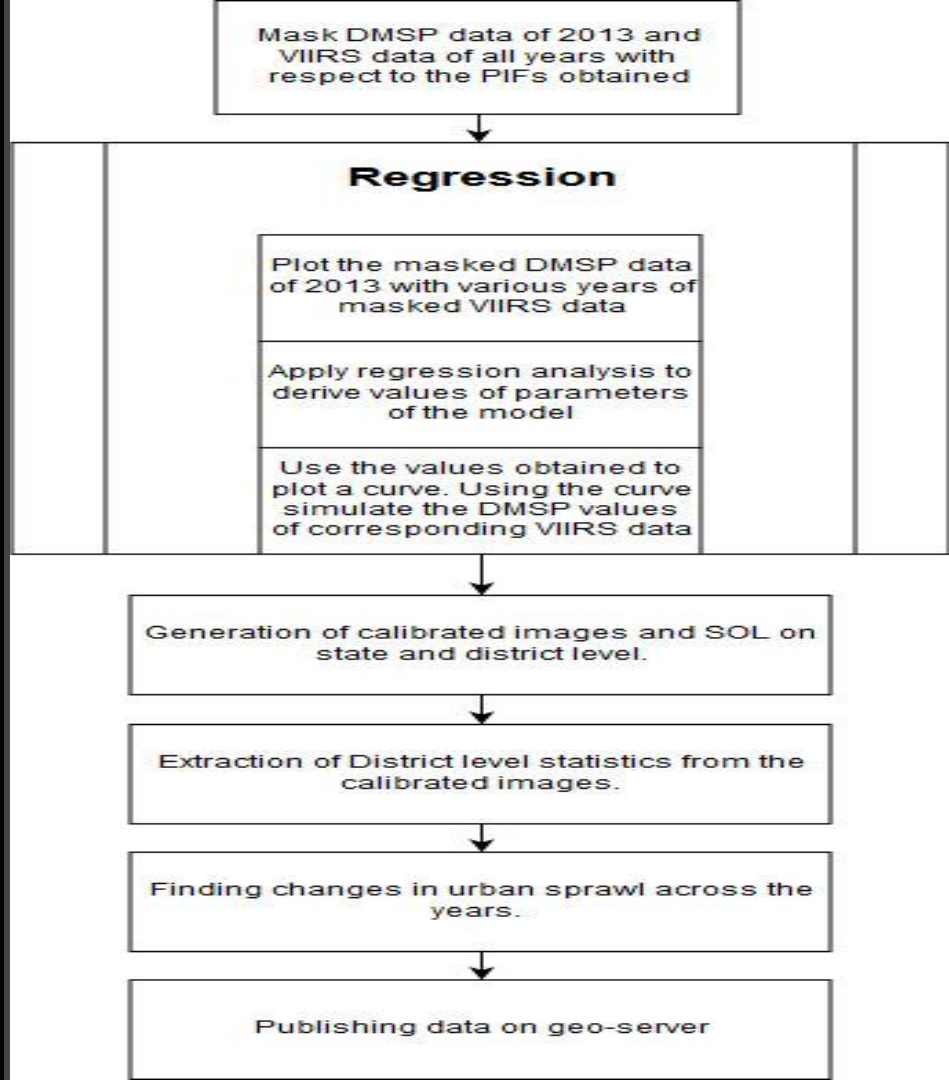


DATA	MIN.	MEDIAN	MEAN	MAX
DMSP 2013	0	0	5.097	63
VIIRS Jan 2013	-0.58	0.2167	0.4106	2736

Dissimilarities of VIIRS-DNB & DMSP-OLS

Type of resolution	VIIRS-DNB	DMSP-OLS	Normalization method
Spatial	500 m	1 km	Spatial-smoothing
Temporal	Available on monthly basis	Available on yearly basis	Annual composite for VIIRS data by averaging
Radiometric	14bit	6 bit	Regression


Start
Obtain Data from online sources
Pre-processing
VIIRS data
DMSP data
Remove outliers using MAD and remove negative pixels using local averaging
Remove oversaturation
Spatial-smoothing
Remove overglow
Composite Image for VIIRS data and yearly image for DMSP data
Find GI and threshold it at 1.645
Find CV and threshold it at 15%
Find common pixels PIFs


Mask DMSP data of 2013 and VIIRS data of all years with respect to the PIFs obtained
Regression
Plot the masked DMSP data of 2013 with various years of masked VIIRS data
Apply regression analysis to derive values of parameters of the model
Use the values obtained to plot a curve. Using the curve simulate the DMSP values of corresponding VIIRS data
Generation of calibrated images and SOL on state and district level.
Extraction of District level statistics from the calibrated images.
Finding changes in urban sprawl across the years.
Publishing data on geo-server

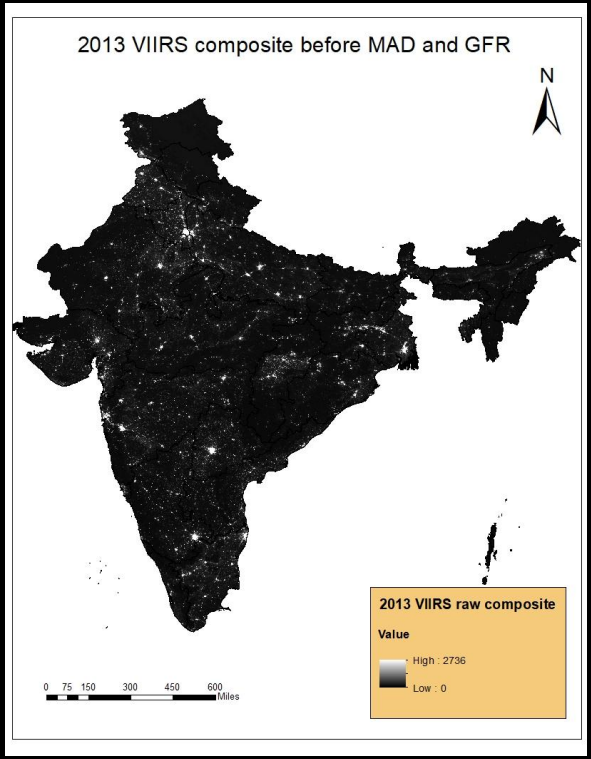
PRE-PROCESSING OF DATA

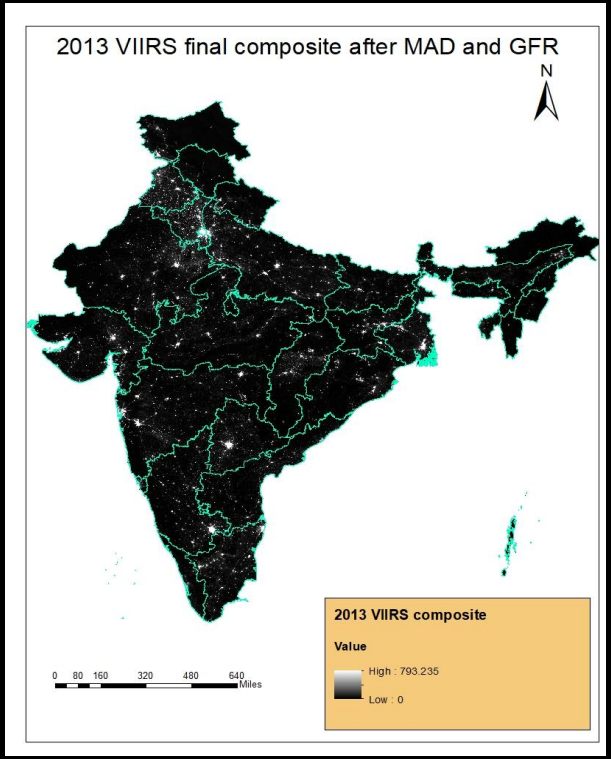
FOR DMSP-OLS

- All pixels with Digital Number(DN) value 63 were removed because of oversaturation of DMSP data for DN values greater than 63.
- All pixels with DN value < 5 were removed to remove overflow around city region.

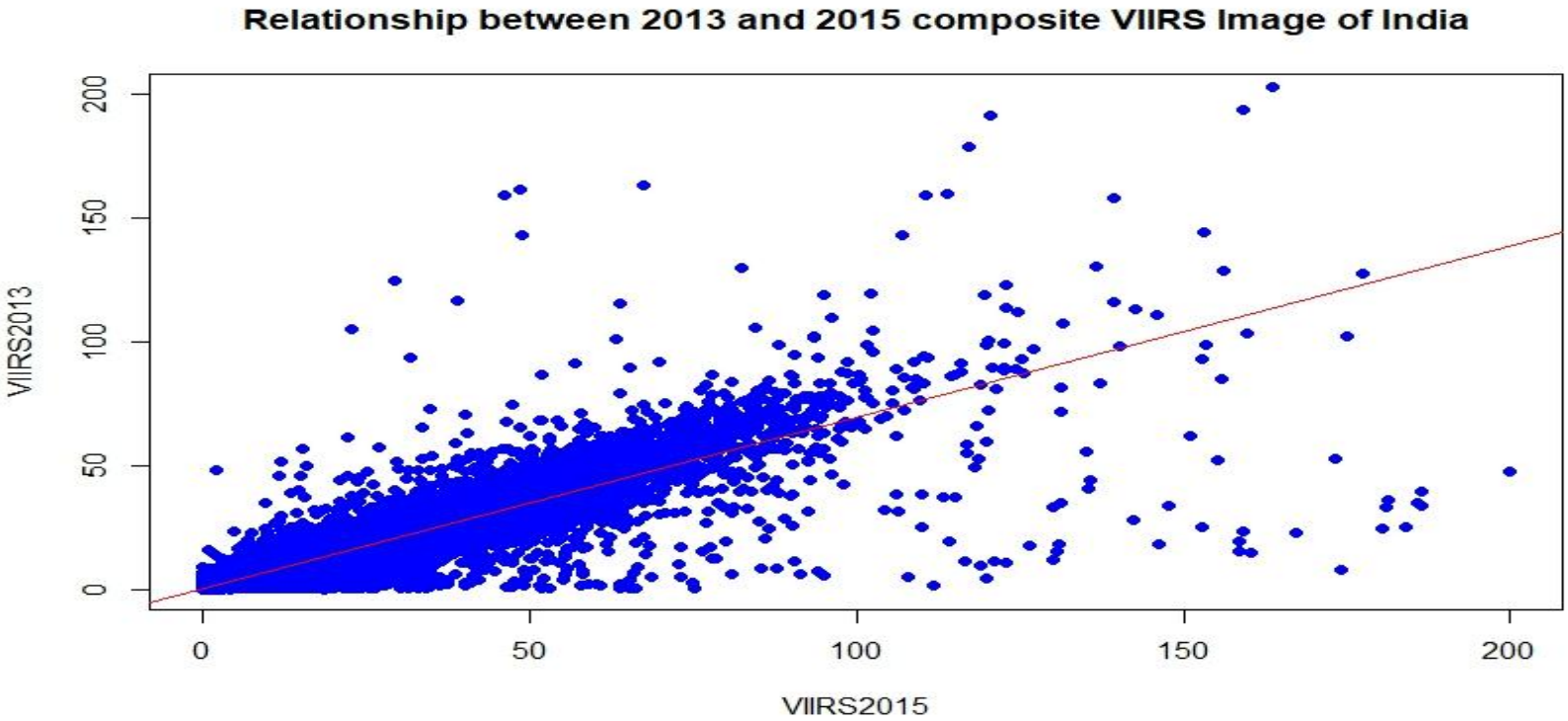
FOR VIIRS

- Outliers were removed by using median absolute deviation(MAD)
- Faint and vague light features (including pixels of negative radiance values) were removed by using a process of local averaging.
- The data was also brought to the same spatial resolution(spatial smoothing to the resolution of DMSP-OLS data)
- Annual composites are generated as the average of monthly data.





Data	Mean	Standard Deviation
Before MAD and GFR	0.524	3.597
After MAD and GFR	0.529	2.266


Relationship between 2013 and 2015 composite VIIRS Image of India
VIIRS2013
VIIRS2015
0
50
100
150
200

$y = a*x + b$

$R^2=0.85$

RESULT

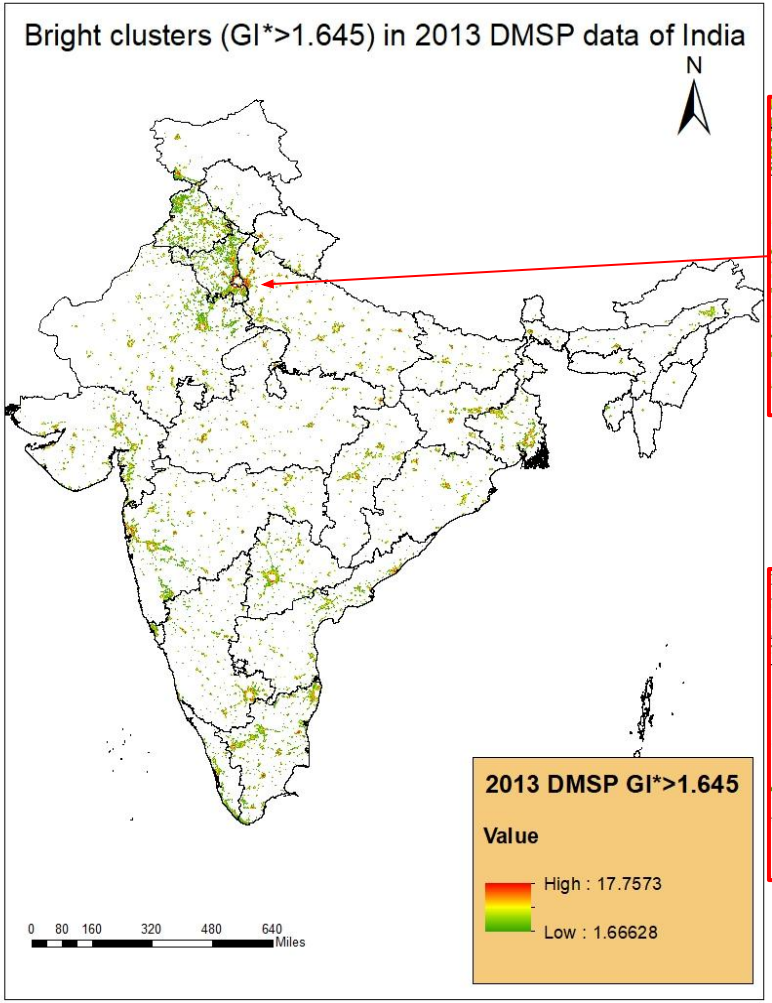
GENERATING PIFs

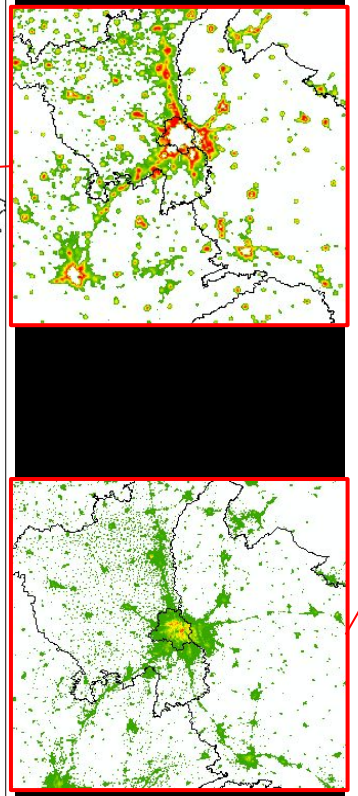
- PIFs are identified as regions with low local spatial variability (CV) and spatial homogeneity (Gi*).
- Common Pixel with low CV (<15%) and high GI*(>1.645) are identified as the PIFs.

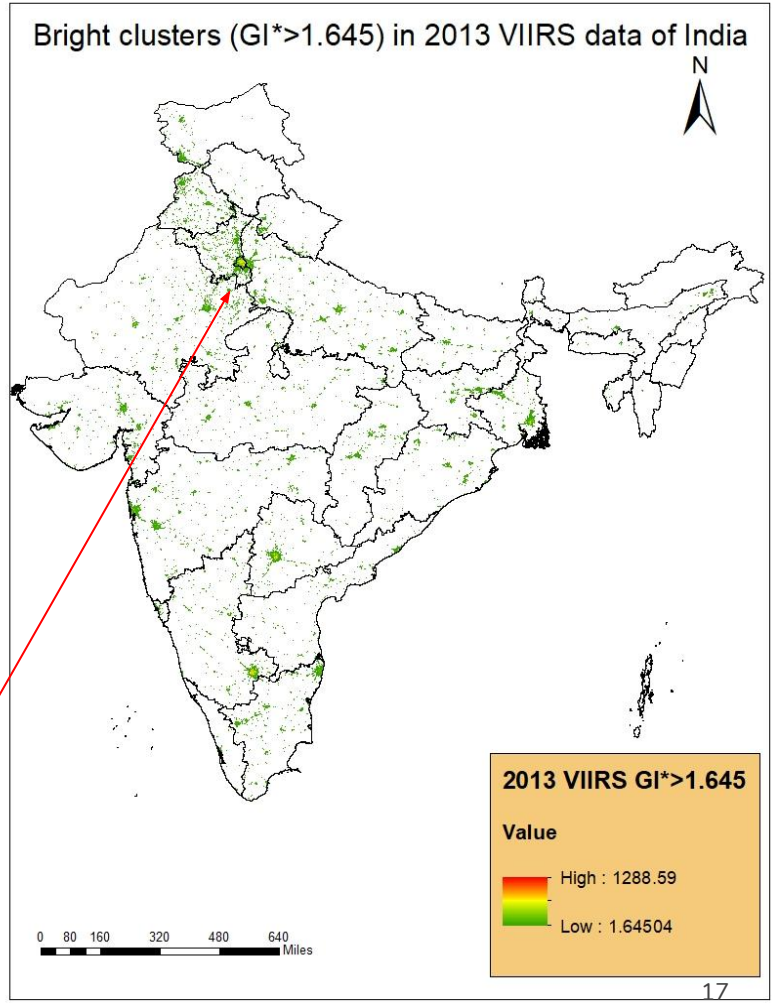
DMSP PIFs = Gi* AND CV

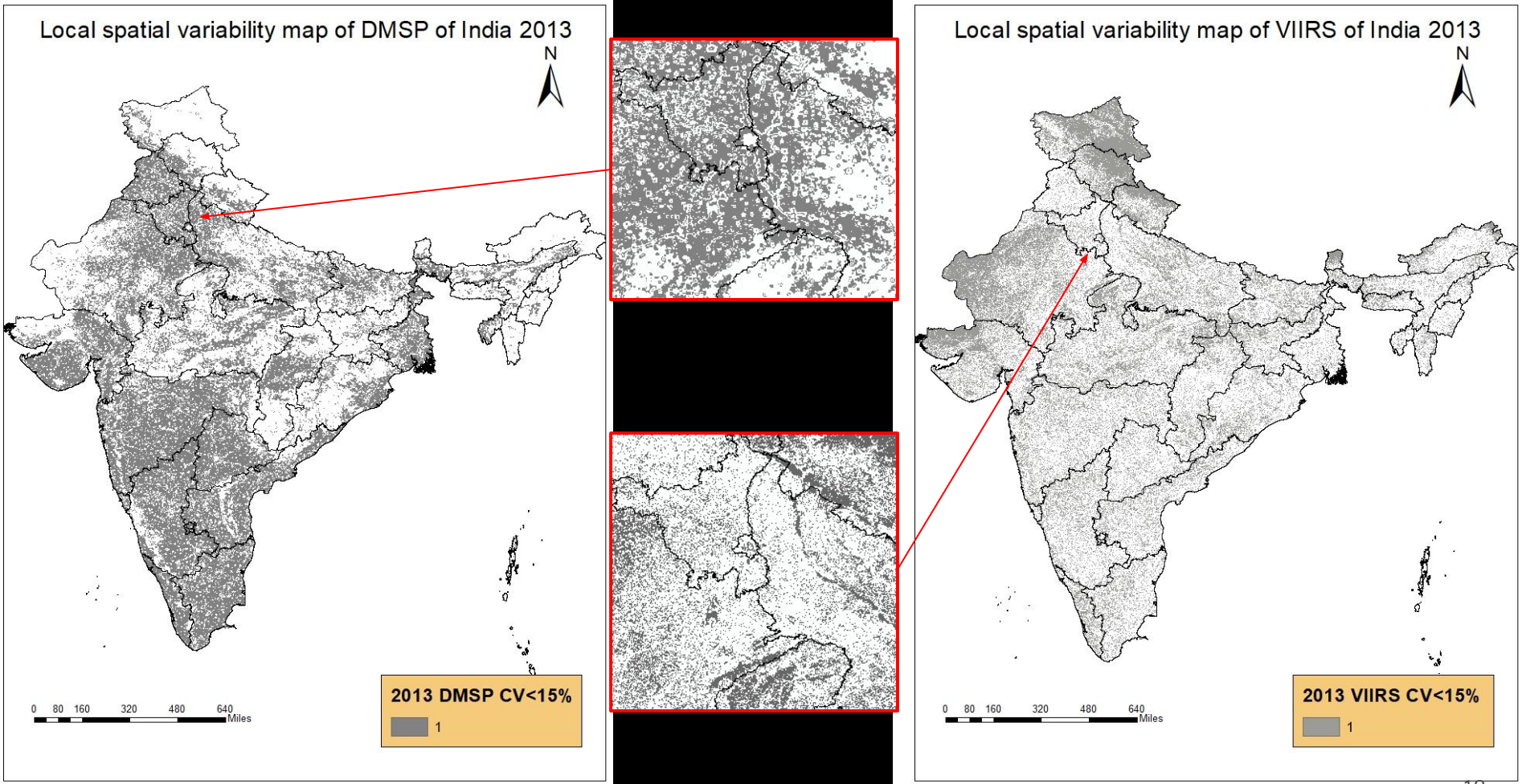
VIIRS PIFs = Gi* AND CV

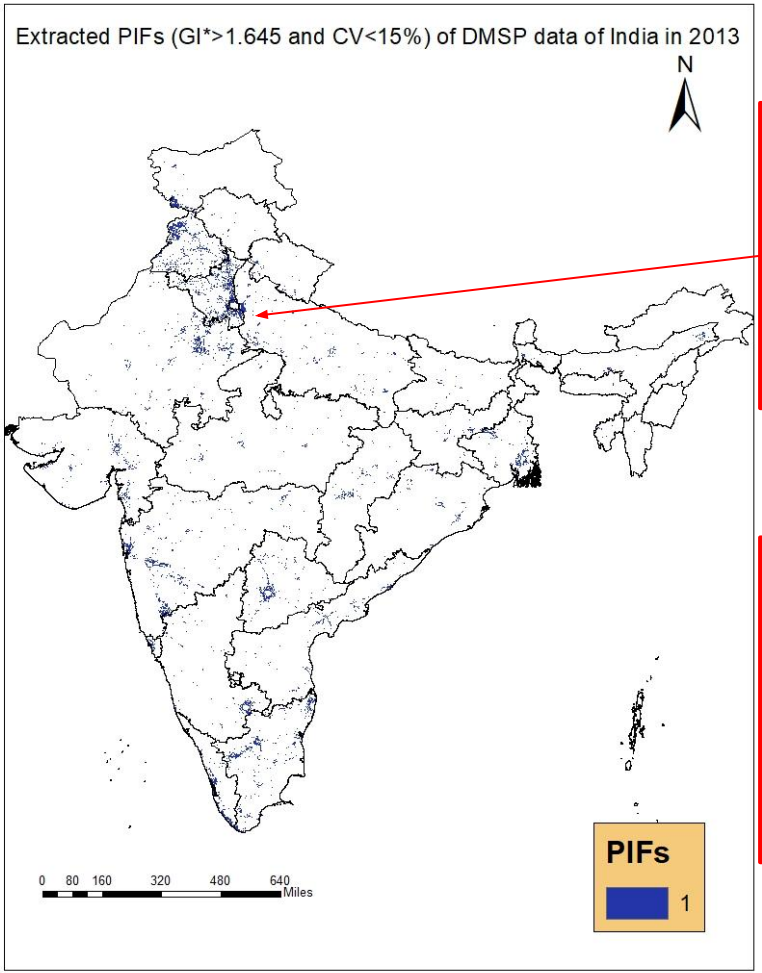
Total PIFs = DMSP PIFs AND VIIRS PIFs

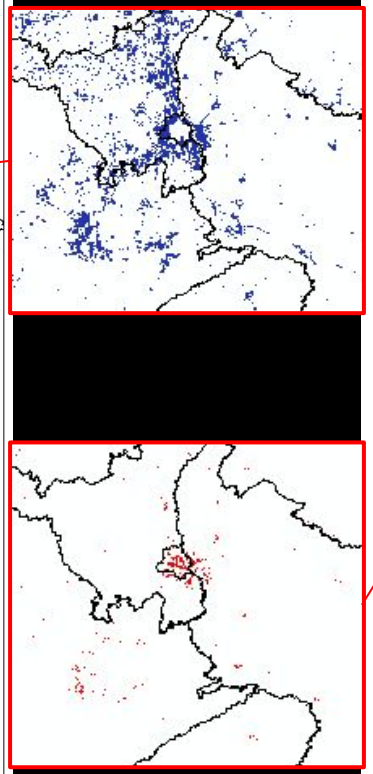

Bright clusters (GI*>1.645) in 2013 DMSP data of India
N
0 80 160 320 480 640
Miles
2013 DMSP GI*>1.645
Value
High : 17.7573
Low : 1.66628

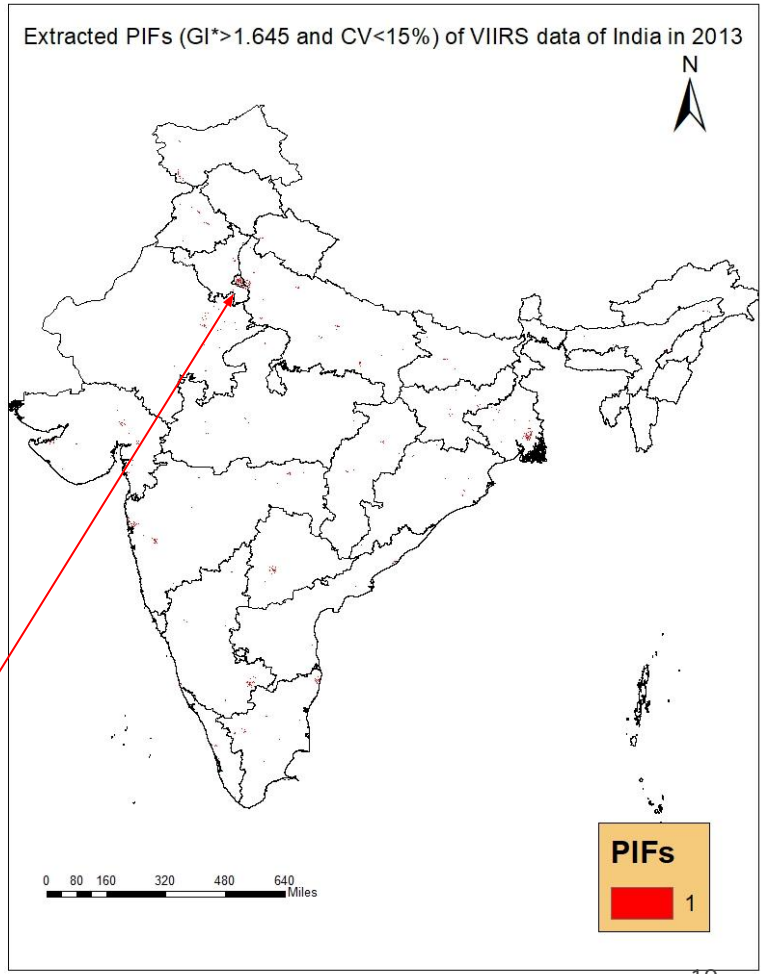


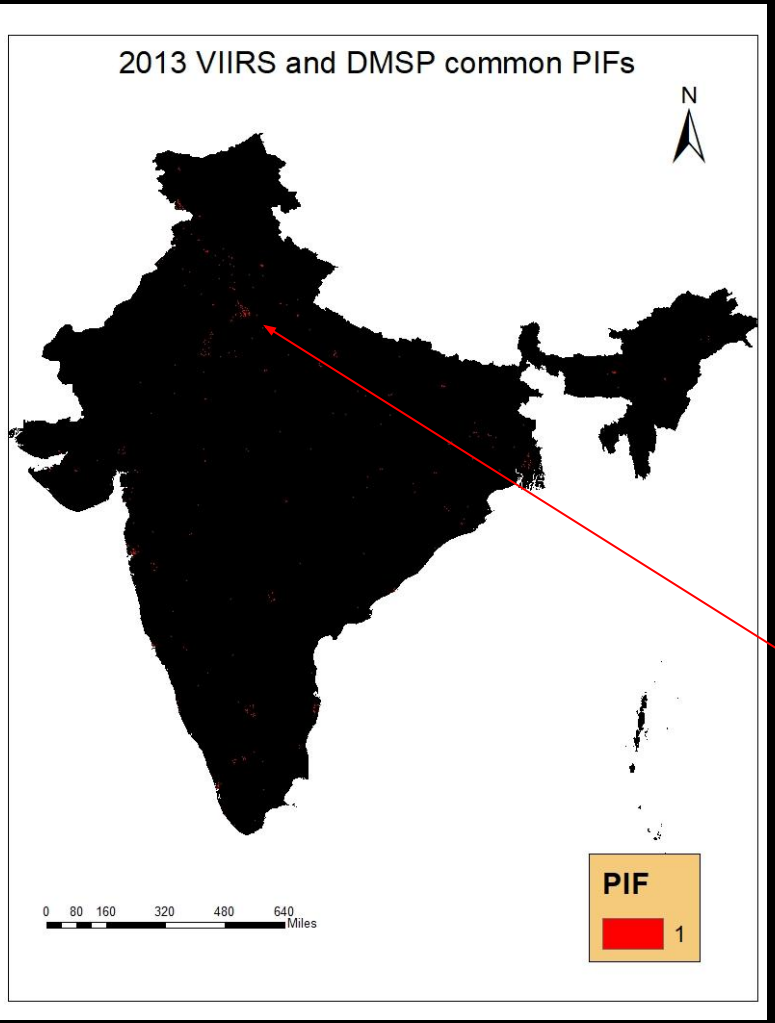

Bright clusters (GI*>1.645) in 2013 VIIRS data of India
N
0 80 160 320 480 640
Miles
2013 VIIRS GI*>1.645
Value
High : 1288.59
Low : 1.64504

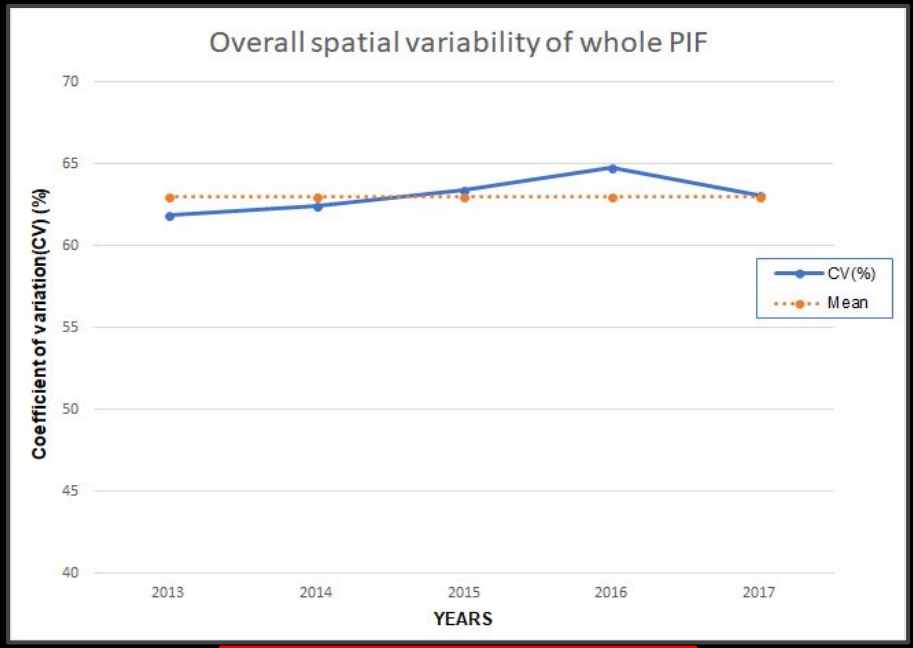

Local spatial variability map of DMSP of India 2013
N
0 80 160 320 480 640
Miles
2013 DMSP CV<15%
1
Local spatial variability map of VIIRS of India 2013
N
0 80 160 320 480 640
Miles
2013 VIIRS CV<15%
1

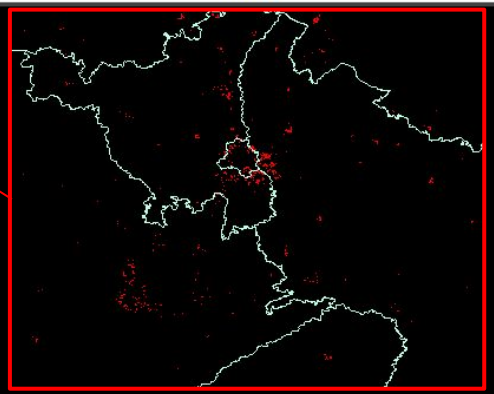

Extracted PIFs (GI*>1.645 and CV<15%) of DMSP data of India in 2013
N
0 80 160 320 480 640
Miles
PIFs
1



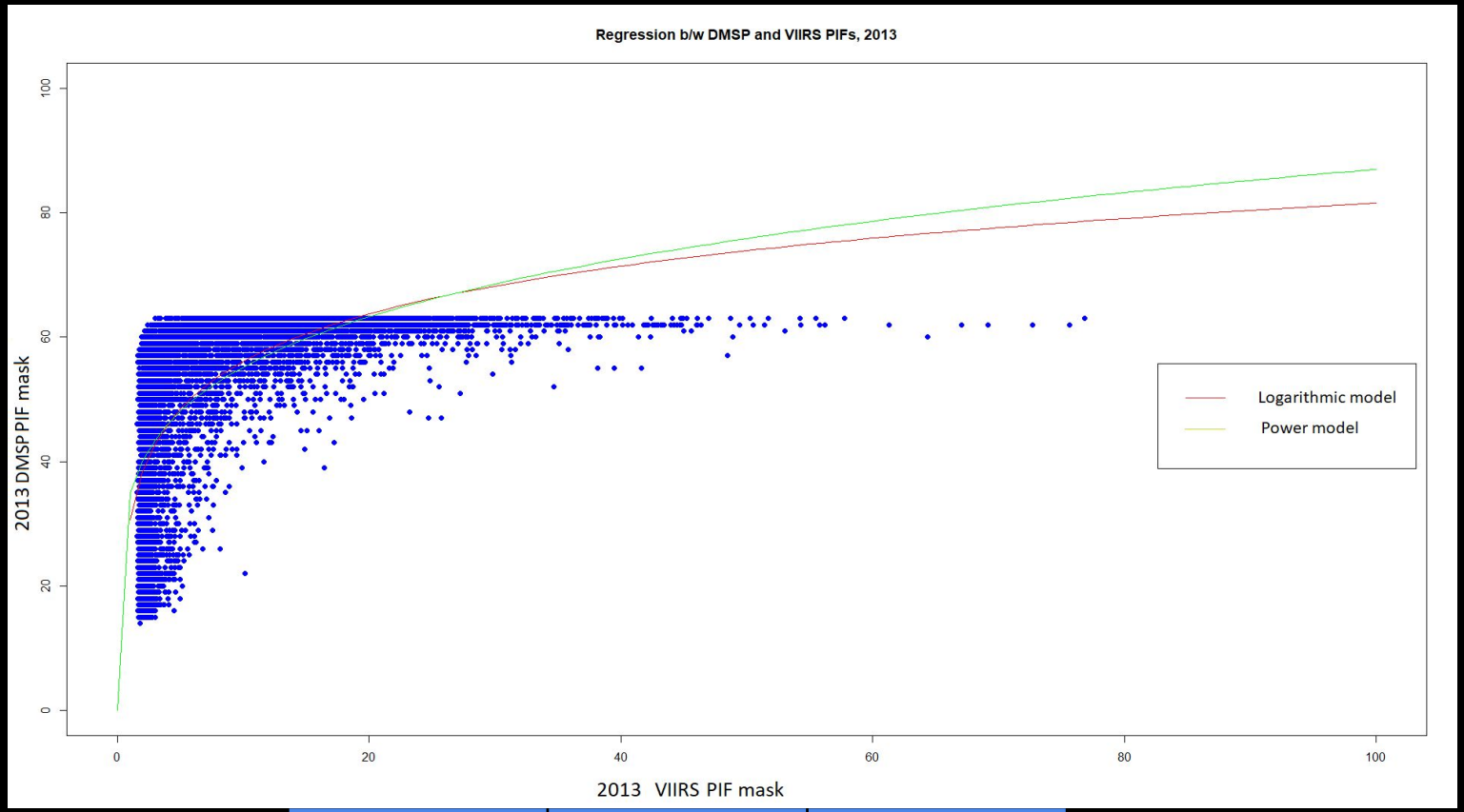

Extracted PIFs (GI*>1.645 and CV<15%) of VIIRS data of India in 2013
N
0 80 160 320 480 640
Miles
PIFs
1


2013 VIIRS and DMSP common PIFs
N
0 80 160 320 480 640
Miles
PIF
1


Overall spatial variability of whole PIF
Coefficient of variation(CV) (%)
70
65
60
55
50
45
40
2013
2014
2015
2016
2017
YEARS
CV(%)
Mean



REGRESSION

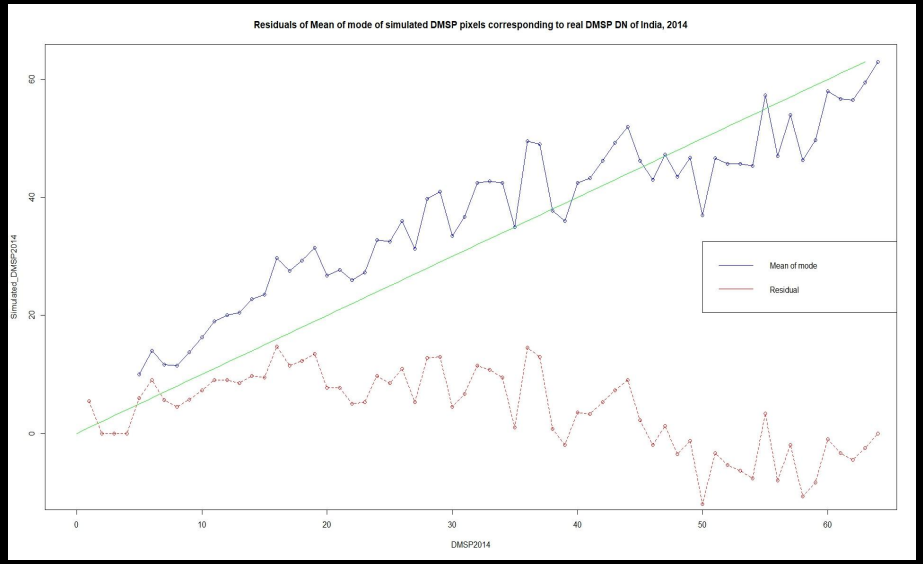


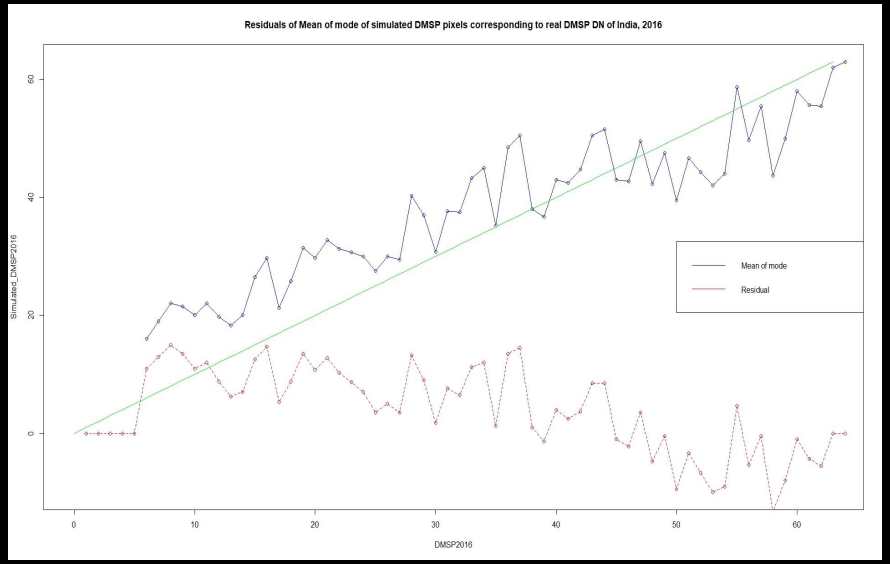
Model	RMSE	R^2
Logarithmic	9.86	0.81
Power	10.17	0.76

COEFFICIENTS OF LOGARITHMIC MODEL					
Year	2013	2014	2015	2016	2017
a	11.0556	11.4623	12.2764	11.9637	12.7412
b	30.6120	27.8722	23.7286	25.8012	23.4470
RMSE	9.86	9.913	9.787	9.832	9.991
R^2	0.81	0.69	0.72	0.73	0.62

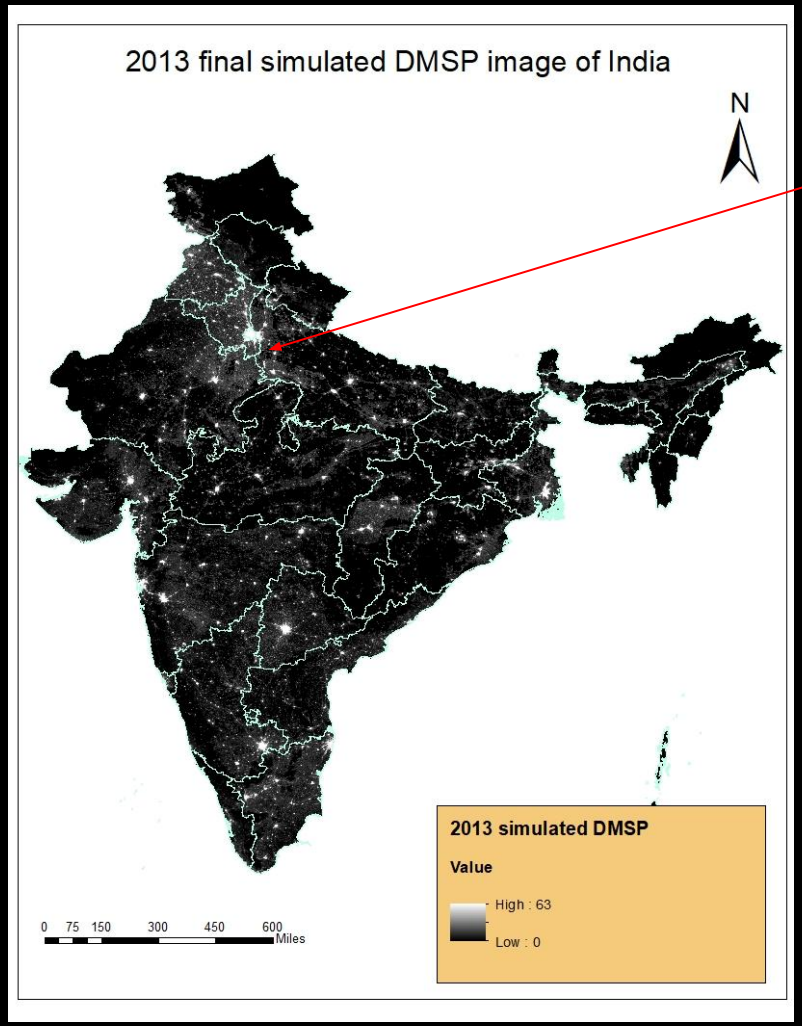
Equation is:

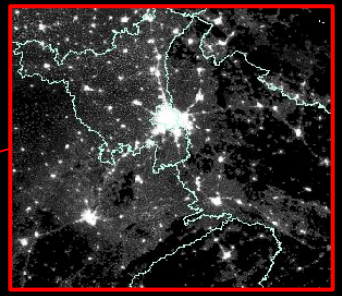
$$y(DMSP) = a * \log(x(VIIRS)) + b$$

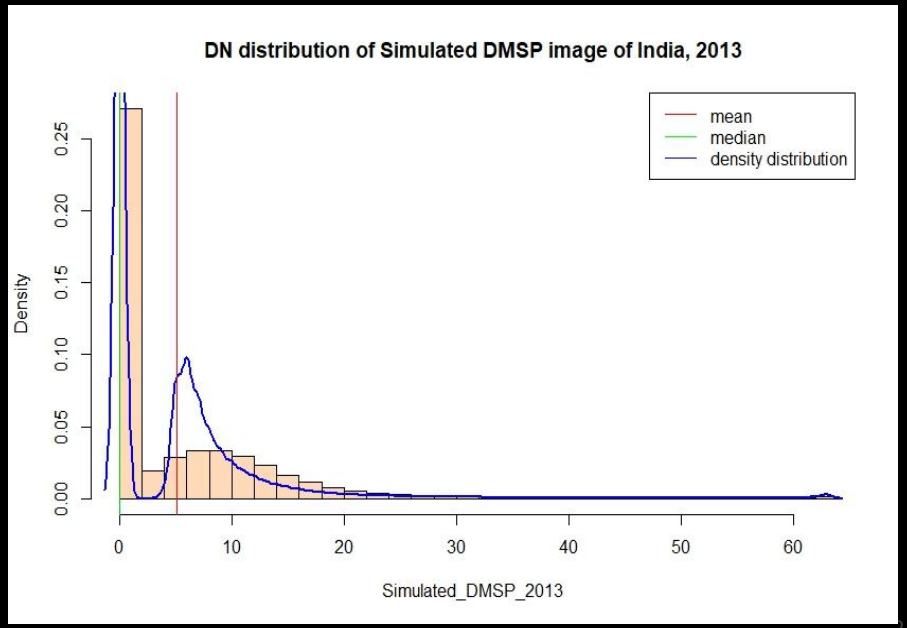


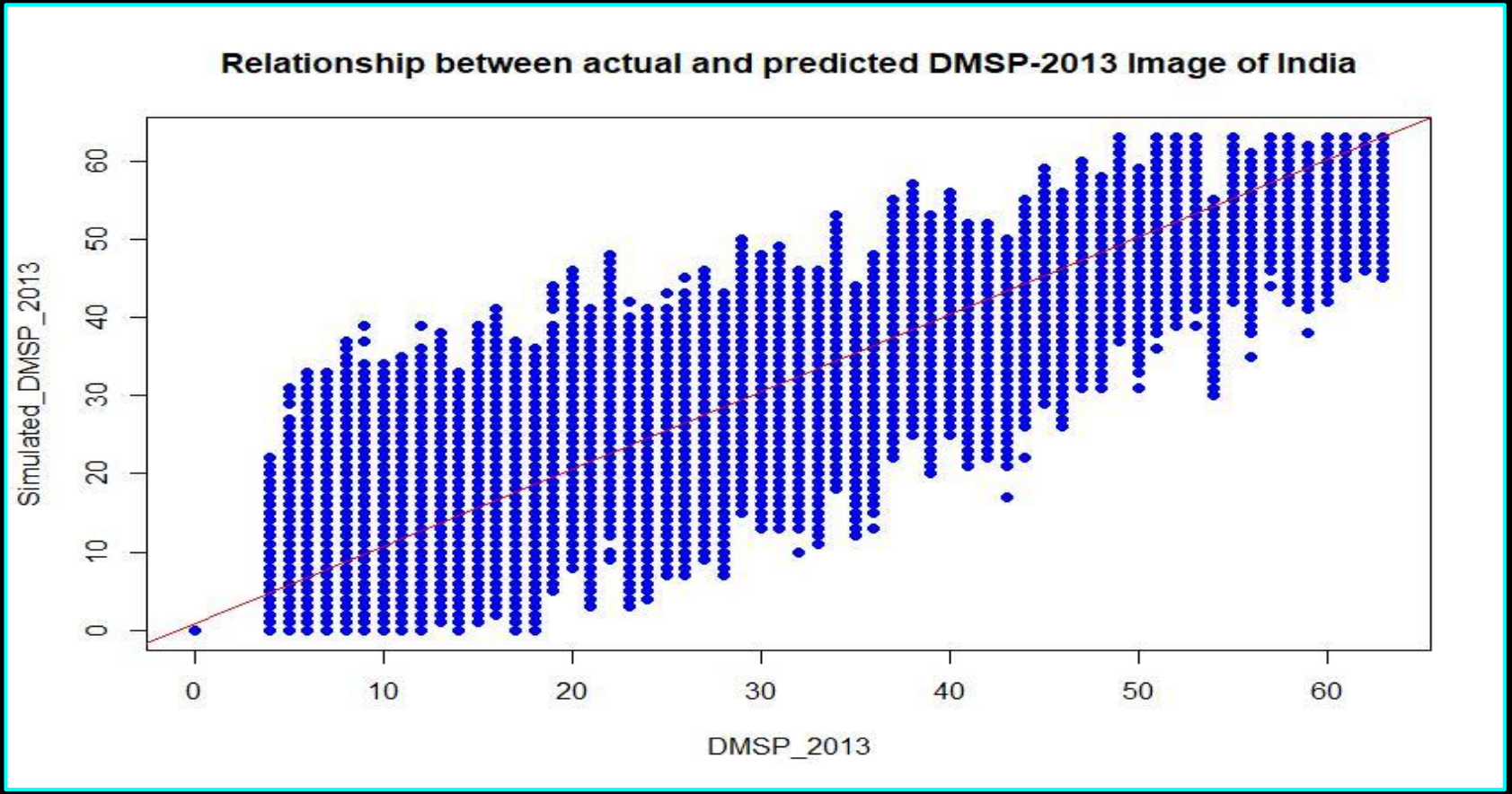


$$y(DMSP_i) = a * \log(x(VIIRS)) + b - (\text{Residual of mean of mode}_i)$$


2013 final simulated DMSP image of India
N
2013 simulated DMSP
Value
High : 63
Low : 0
0 75 150 300 450 600
Miles

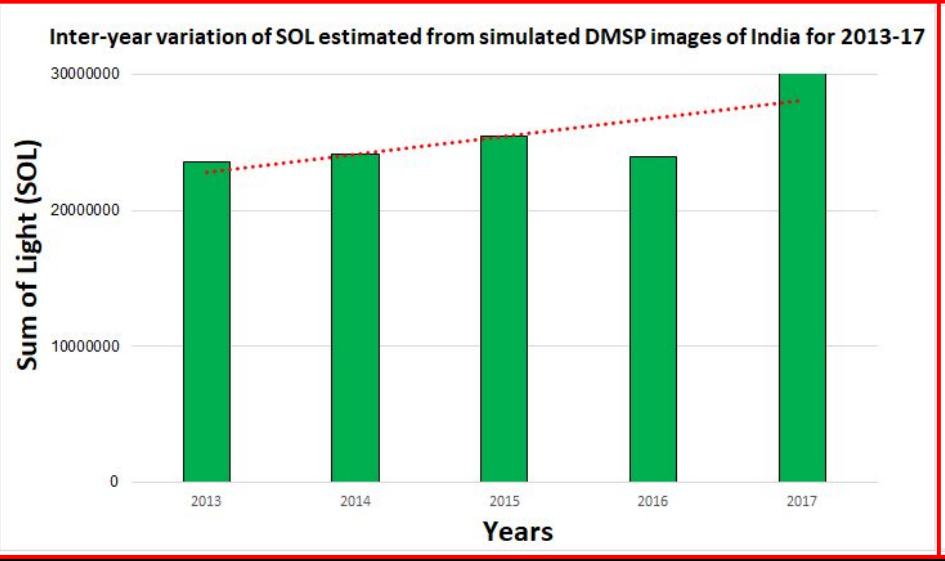


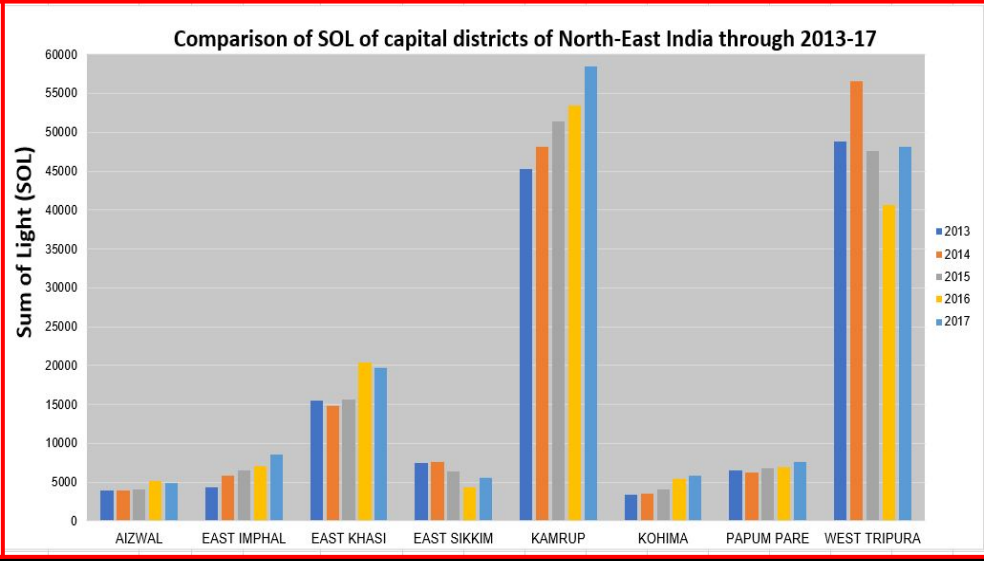

DN distribution of Simulated DMSP image of India, 2013
mean
median
density distribution
Density
0.00
0.05
0.10
0.15
0.20
0.25
0
10
20
30
40
50
60
Simulated_DMSP_2013


Relationship between actual and predicted DMSP-2013 Image of India
Simulated_DMSP_2013
0
10
20
30
40
50
60
DMSP_2013

$y = a*x + b$

$R^2 = 0.86$


Inter-year variation of SOL estimated from simulated DMSP images of India for 2013-17
Sum of Light (SOL)
30000000
20000000
10000000
0
2013
2014
2015
2016
2017
Years

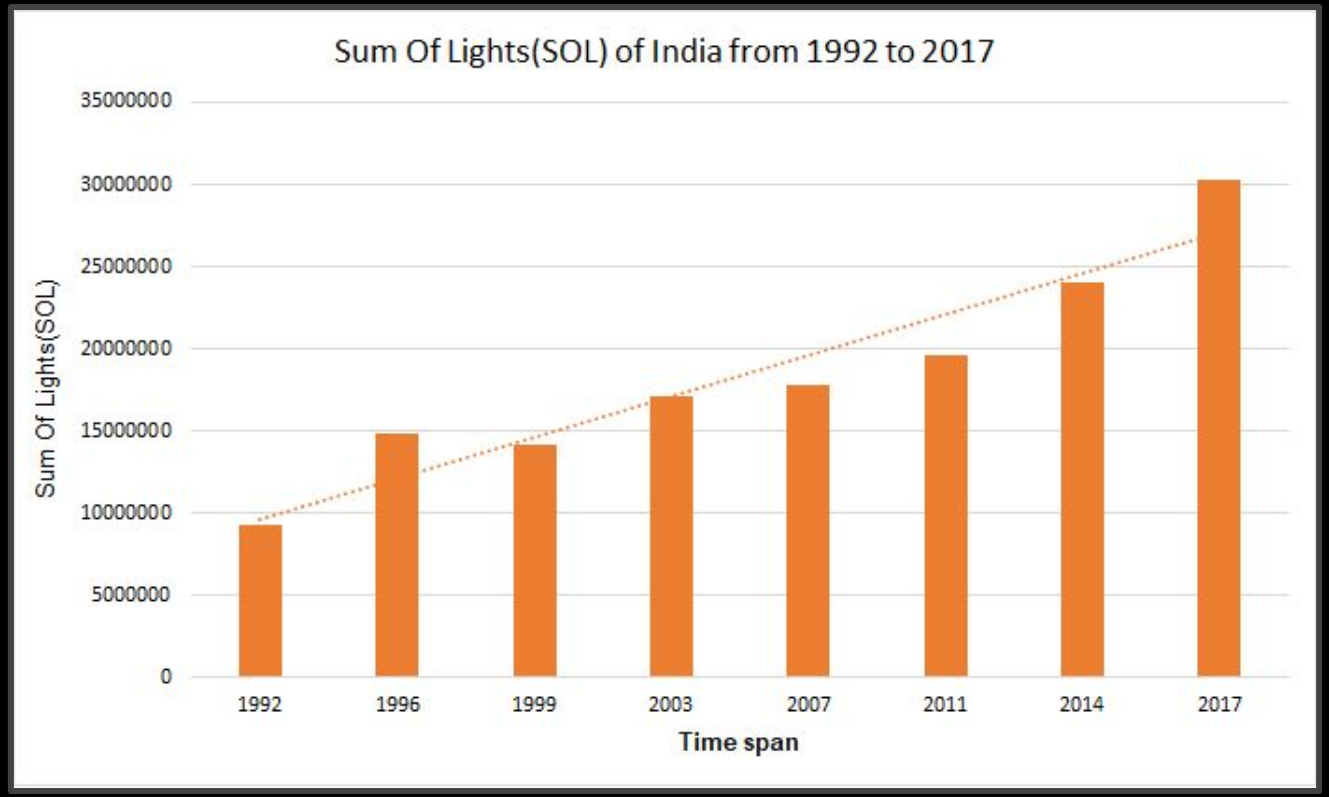

Comparison of SOL of capital districts of North-East India through 2013-17
Sum of Light (SOL)
60000
55000
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
AIZWAL
EAST IMPHAL
EAST KHASI
EAST SIKKIM
KAMRUP
KOHIMA
PAPUM PARE
WEST TRIPURA
2013
2014
2015
2016
2017

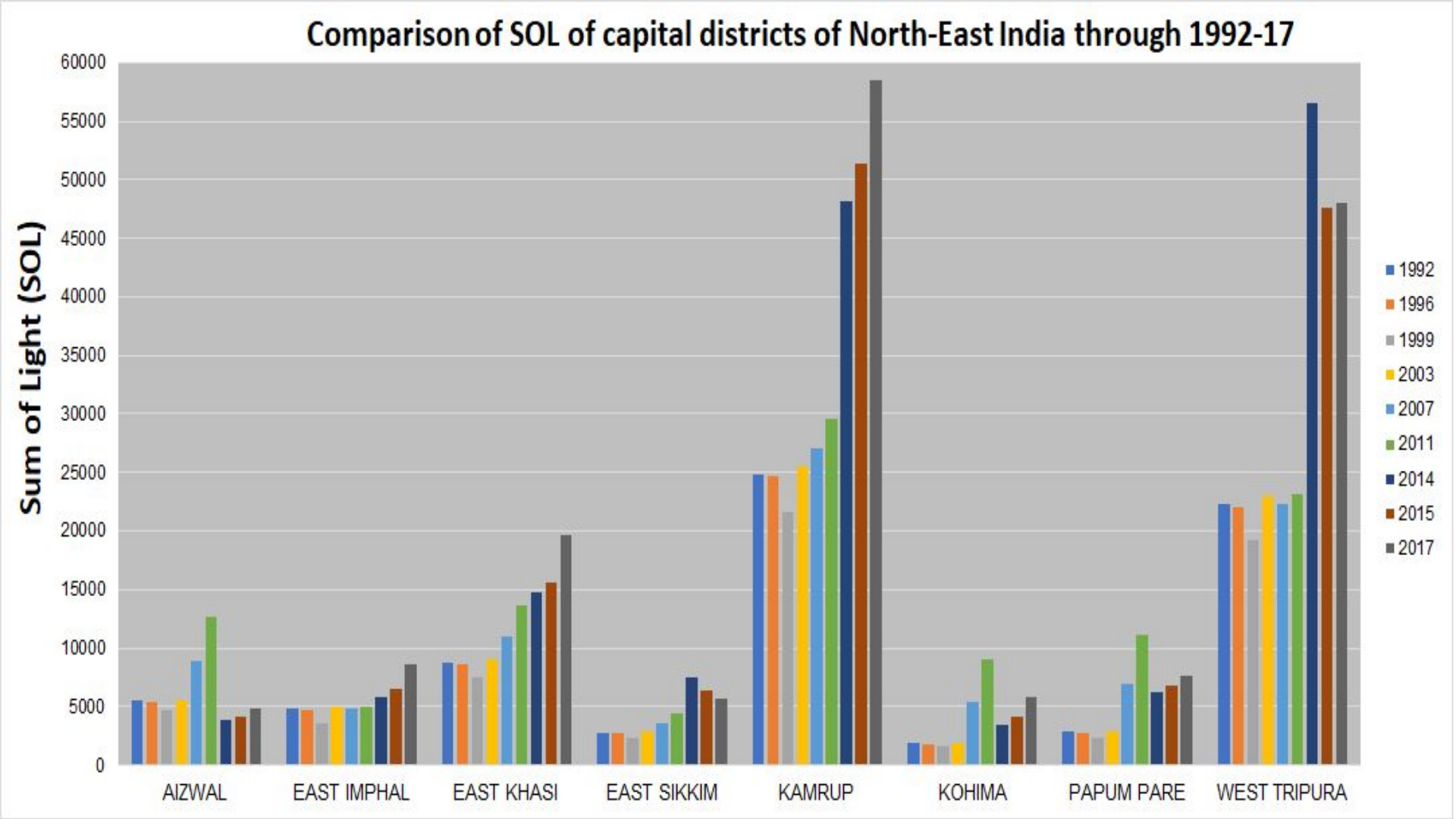
INTERCALIBRATION OF DMSP-OLS DATA

Satellite	Year	C_o	C_1	C_2	R^2
F10	1992	−2.0570	1.5903	−0.0090	0.9075
F12	1996	−0.0959	1.2727	−0.0040	0.9319
F12	1999(Ref.)	0.0000	1.0000	0.0000	1.0000
F14	2003	0.7390	1.2416	−0.0040	0.9432
F16	2007	0.6394	0.9114	0.0014	0.9511
F18	2011	1.8956	0.7345	0.0030	0.9095

Equation of intercalibration is:

$$\mathbf{Y = C_o + C_1X + C_2X^2}$$


Sum Of Lights(SOL) of India from 1992 to 2017
35000000
30000000
25000000
20000000
15000000
10000000
5000000
0
Sum Of Lights(SOL)
1992
1996
1999
2003
2007
2011
2014
2017
Time span


Comparison of SOL of capital districts of North-East India through 1992-17
Sum of Light (SOL)
60000
55000
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
AIZWAL
EAST IMPHAL
EAST KHASI
EAST SIKKIM
KAMRUP
KOHIMA
PAPUM PARE
WEST TRIPURA
1992
1996
1999
2003
2007
2011
2014
2015
2017

THANK YOU

MEDIAN ABSOLUTE DEVIATION (MAD)

It involves finding median of the absolute deviation around the median of the data

$$MAD = b*M_i (|x_i - M_j(x_j)|)$$

where M_i denotes median of x_i

- The multiplication by b is crucial, as otherwise the formula for the MAD would only estimate the scale up to a multiplicative constant . The b = 1.4826 assuming normality of data.
- This method is used because it is not affected by presence of outliers in data. It has a breakdown point of 0.5 whereas the breakdown point of the classical interquartile range is 0.25
- Data points outside the range of (M - 2.5* MAD , M + 2.5*MAD) were removed as outliers.

GAS FLARES REMOVAL

- Gas flaring is widely used to dispose of dissolved natural gas present in petroleum in production and processing facilities where there is no infrastructure to make use of the gas
- Most gas flaring occurs at remote petroleum production or processing facilities, many of which are offshore
- Flaring sites are indicated for grid cells with average temperature exceeding 1200 K and per cent detection frequencies of 1% or more.
- Currently no precise algorithm is there. A gas flaring vector set is available, with more than 12,000 flaring sites identified in 2015
- The site present in India were identified and were successfully removed from the yearly composites.

SPATIAL SMOOTHING OF VIIRS DATA

- The spatial resolution of VIIRS composites is 0.004167°, while that of the DMSP/OLS is 0.008333°. Therefore, the VIIRS composites were aggregated to a spatial resolution of 0.008333° .

- This was done using the concept of local averaging where a 2X2 sub-matrix of VIIRS data was aggregated and the result was put into a single pixel of matrix of dimension equal to that of DMSP/OLS.

- This process was necessary to bring the area corresponding to same geographical location to the same indexed pixel in both DMSP/OLS and VIIRS data.

Getis-ord GI* statistic

$$G_i^*(d) = \frac{\sum w_{ij}(d)x_j - W_i^*\bar{x}}{s[W_i^*(n - W_i^*)/(n-1)]^{1/2}}$$

Where,

$$W_i^* = \sum_j w_{ij}(d)$$

$$\bar{x} = \frac{\sum_j x_j}{n}$$

- The getis statistic gives a measure of clustering and it is used to identify clusters of comparatively higher values (GI*>0; hotspot) and cluster of comparatively lower values (GI*<0; cold spot) than the global mean of the image attribute.
- We applied GI* statistics to each annual composite and then thresholded them to values greater than 1.645.
- It does ensure the maximum probability of random cluster formation is limited by only 10% of all the clusters have formed and a 90% probability of the formed cluster to have high spatial autocorrelation.

Coefficient of Variability(CV)

$$CV = \frac{S}{\bar{x}}$$

Where, S is the measure of standard deviation of luminosity measurements, is the mean of luminosity measurements in a pre-defined window.

- Coefficient of variability(CV) is used to select local spatially stable regions in the study area. The areas with a low CV over different years are considered as regions with stable and low spatial variability.
- We did CV analysis on all the annual composites and thresholded them for values less than 5%, 10% and 15%.
- Due to less no. of points available after thresholding for 5% and 10% , we pushed the percentage to 15 for better regression.

COMPOSITE GENERATION

- The VIIRS data are available for each month.
- The pre -processing steps are applied to each months' data.
- Then the average of the monthly composites is taken to find the VIIRS data for the whole year.
- The averaging process takes care of the 'No data' values to ensure they don't interfere in the process of finding the composite.