

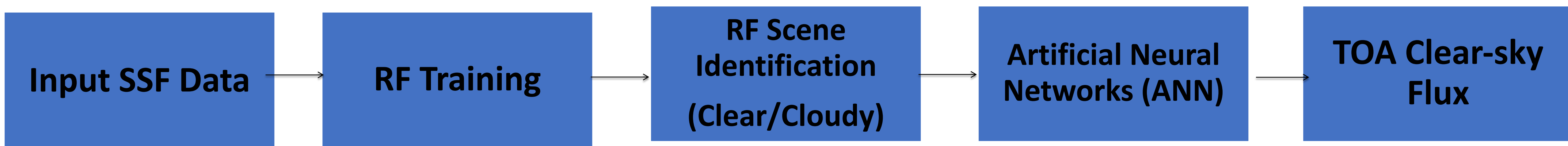
Objectives:

The motivation for this study is to develop a Machine learning methodology for the determination of CERES clear-sky scenes and the estimation of clear-sky TOA fluxes from CERES TOA radiance measurements without using any imager information.

Background:

Clouds and the Earth’s Radiant Energy System (CERES) provides top of the atmosphere (TOA) radiative flux estimates on a global scale. This is achieved by combining broadband footprint radiance measurements from CERES instrument with radiances from a high-resolution, multispectral imager on the same spacecraft. To estimate TOA radiative fluxes, CERES radiances are converted to fluxes using empirical angular distribution models (ADMs) and they depend on variables like viewing geometry, surface type and atmospheric conditions, etc. However, in the absence of imager information over CERES footprints, accurate scene identification and subsequent estimation of TOA fluxes are difficult. The unavailability of imager data leads to a situation similar to that encountered in Earth Radiation Budget Experiment (ERBE) TOA flux measurements. The empirical ADMs developed for the estimation of ERBE TOA flux depends on parameters like radiometer geometry, broadband radiance measurements, etc., and have relatively large uncertainties compared to CERES ADMs (Loeb et al., 2009). In this study, our objective is to develop a statistical methodology based on the most accurate modern technique to solve the CERES “ERBE like” problem.

Methodology:



Machine learning Algorithms

Random Forests (RF)

- Developed by Breiman and Cutler(2000)
- is an ensemble learning method for classification and regression.
- Random forests operate by constructing a multitude of decision trees and outputting the class that gets maximum number of votes from the forest.
- Main advantages of RF method are:
 - they have faster runtimes
 - can deal with unbalanced and missing data

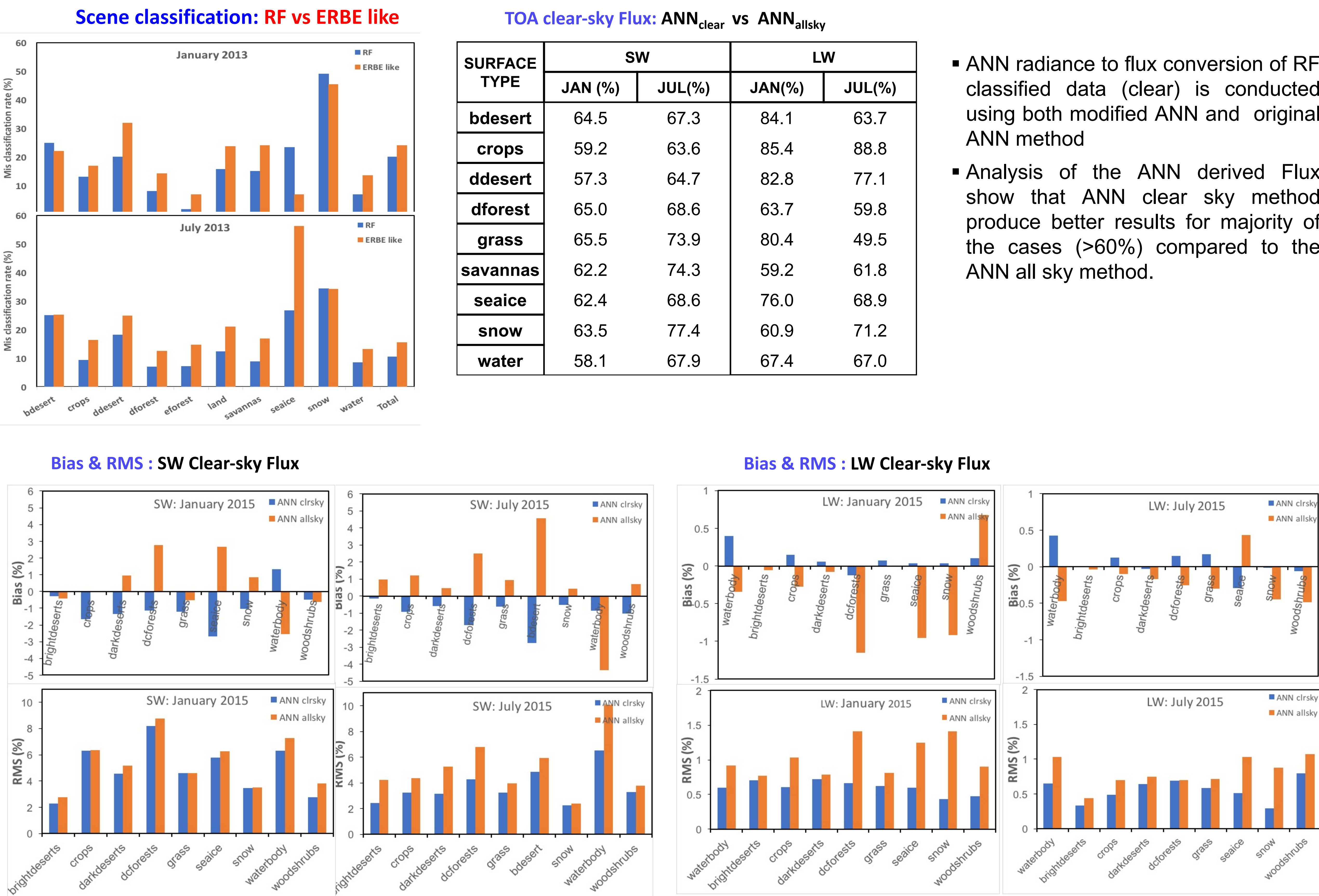
Artificial Neural networks (ANN)

- ANN methodology outlined in Lukashin and Loeb(2003) is used
- ANN is based on a large number of neural units loosely modelling the way a biological brain solves problem.
- They are exceptionally good at performing pattern recognition and other tasks that are very difficult to program using conventional techniques.
- Programs that employ neural nets are also capable of learning on their own and adapting to changing conditions

Input Variables	IGBP surface types
Solar & viewing zenith- angles	Water bodies
Relative azimuth angle	Bright Desert
CERES Shortwave (SW) and Longwave (LW) broadband radiances	Dark Desert
LW surface emissivity	Grasslands
Broadband surface- albedo	Croplands and cities
Surface skin temperature	Evergreen Forests
Precipitable water	Deciduous Forests
Wind speed	Woody Savannas and Shrub lands
	Permanent and Fresh- snow
	Sea Ice

CERES Aqua SSF data is used in the RF and ANN analysis

Training data : 2003-2014
Test data : 2015



- ANN radiance to flux conversion of RF classified data (clear) is conducted using both modified ANN and original ANN method
- Analysis of the ANN derived Flux show that ANN clear sky method produce better results for majority of the cases (>60%) compared to the ANN all sky method.

Comparison of TOA Clear-sky Fluxes and CRF (SSF Vs ANN Vs ERBE-like)

Mean(arithmetic) clear-sky TOA flux							Global mean TOA flux					Global day-time mean TOA CRF																	
Surface Type	Month	SW			LW			Month	All-sky			Clear-sky			Month	SSF	ANN-SSF	ERBE-SSF											
		SSF	ANN-SSF	ERBE-SSF	SSF	ANN-SSF	ERBE-SSF		SSF	ANN-SSF	ERBE-SSF	SSF	ANN-SSF	ERBE-SSF															
Desert	JAN	287.5	0.6	-47.5	302.9	-2.0	5.5	SW	JAN	257.0	0.5	-10.7	114.6	-1.8	4.0	SWCRF	JAN	-133.1	-2.5	5.3									
	APR	354.0	6.0	-70.1	338.2	2.3	-8.8		APR	246.2	0.4	-10.0	119.7	-2.1	3.1		APR	-117.7	2.8	4.8									
	JUL	340.6	3.3	-73.5	348.6	5.1	-11.4		JUL	233.7	-0.4	-12.6	103.7	-1.8	6.8		JUL	-120.9	-0.8	9.1									
Land	JAN	190.3	2.5	-23.8	320.2	-3.6	-15.8	LW	JAN	243.0	-0.6	2.0	274.6	0.9	-1.8	NETCRF	JAN	31.0	-2.4	-3.8									
	APR	204.5	3.6	-23.5	318.4	1.3	-11.4		APR	243.1	-0.6	1.8	276.0	1.0	-1.3		APR	32.6	-3.0	-3.5									
	JUL	170.7	3.1	-19.3	316.1	-4.1	-16.6		JUL	252.4	-0.8	0.8	285.5	1.0	-4.0		JUL	31.8	-1.6	-4.1									
Ocean	JAN	86.1	6.0	8.1	287.6	0.6	0.7	SWCRF	APR	JAN	243.1	-0.6	1.8	276.0	1.0	-1.3	NETCRF	JAN	-102.1	-4.9	1.5								
	APR	84.4	8.5	4.7	284.7	-3.5	2.1															APR	-117.7	2.8	4.8	APR	-85.1	-0.1	1.3
	JUL	79.4	8.2	7.6	284.4	-2.4	2.4															JUL	-120.9	-0.8	9.1	JUL	-89.1	-2.4	5.0
Snow	JAN	293.7	-33.9	31.8	196.0	-1.3	-1.0	NETCRF	JAN	243.1	-0.6	1.8	276.0	1.0	-1.3	NETCRF	JAN	-102.1	-4.9	1.5									
	APR	259.7	-7.3	-7.5	201.3	-2.5	2.3														APR	-117.7	2.8	4.8	APR	-85.1	-0.1	1.3	
	JUL	329.9	-30.5	-87.5	233.9	-1.3	6.8														JUL	-120.9	-0.8	9.1	JUL	-89.1	-2.4	5.0	

SUMMARY:

- ANN clear-sky method produce more accurate TOA flux values most of the time (>60% of data) compared to all-sky ANN method
- Mean clear-sky TOA flux estimated using ANN clear-sky method show good comparison with SSF derived TOA Fluxes
- Global mean weighted CRF estimated using ANN method compares well with that with of SSF for different months.
- Global mean CRF estimated using the ANN method generally performs better compared to ERBE-like