SEVIR: Storm EVent ImageRy Dataset

Mark Veillette

Nov 13th 2019

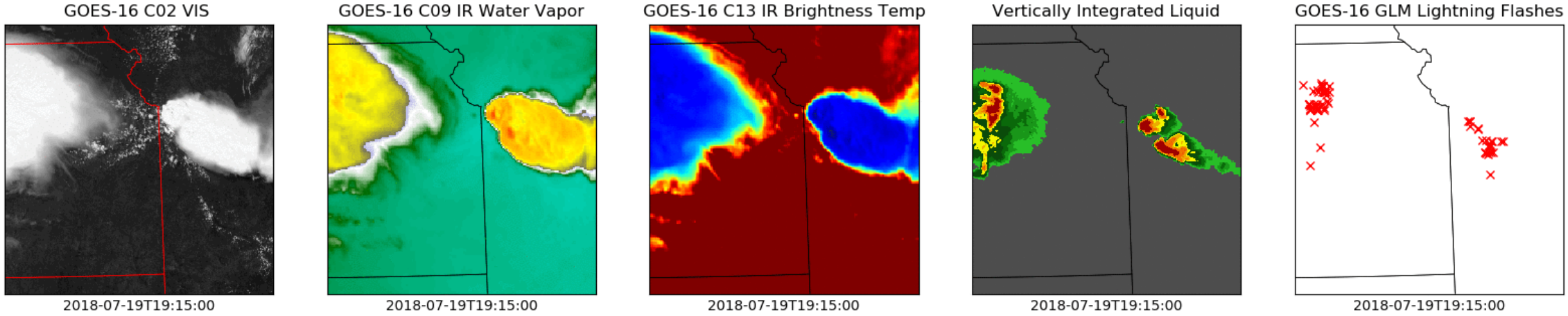
LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

This material is based upon work supported by the Federal Aviation Administration under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Federal Aviation Administration.

Storm EVent ImageRy (SEVIR) Dataset

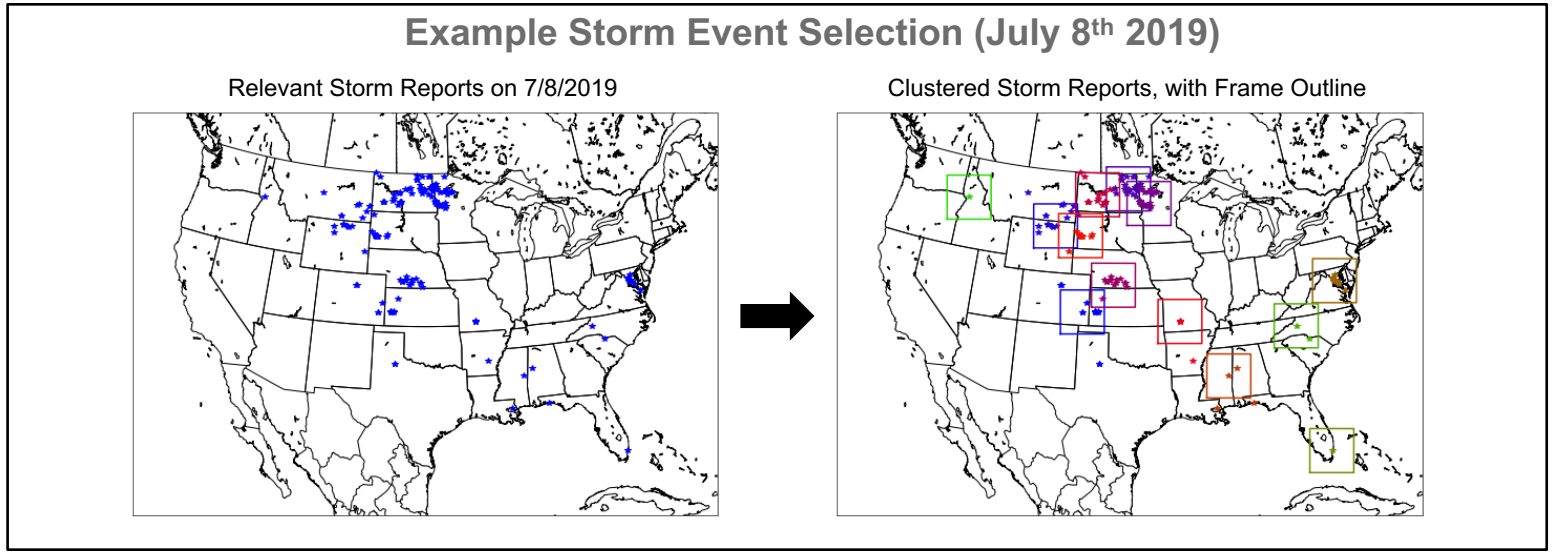
- **SEVIR is a dataset of approximately 15,000 spatially and temporally aligned image sequences from the GOES-16 satellite and NextGen radar mosaics. One such sequence is shown below.**
- **Each sequence, or "event" in SEVIR contains a 384 km x 384 km region spanning a 4 hours and containing one of more of the following:**
  - **Digital Vertically Integrated Liquid (VIL) extracted from the NextGen Weather Processor archives**
  - **GOES-16 0.64 μm Visible (C02), 6.9 μm Infrared (C09) and 10.7 μm Infrared (C13) imagery**
  - **GOES-16 Geostationary Lightning Mapper flashes**
- **The motivation for SEVIR is to advance Machine Learning (ML) in meteorology through interesting and labeled datasets that are "ML Ready" and easily accessible to the community**

Sample SEVIR "Event"



SEVIR Event Selection

- **SEVIR Events were chosen in one of two ways:**
  - **Based on events in NOAA's National Centers for Environmental Information (NCEI) Storm Event Database (https://www.ncdc.noaa.gov/stormevents/)**
  - **Or, randomly chosen in areas of NEXRAD coverage over CONUS**
- **Events listed in NCEI's Storm event database from 2017*, 2018 & 2019 with categories Flash Flood, Flood, Hail, Heavy Rain, Lightning, Thunderstorm Wind, or Tornado were clustered based on geographical and temporal proximity**
- **A 384 km x 384 km frame was drawn around the center of each cluster, and a 4 hour sequence of images were extracted for each case**



*GOES-16 data only available in 2018 & 2019

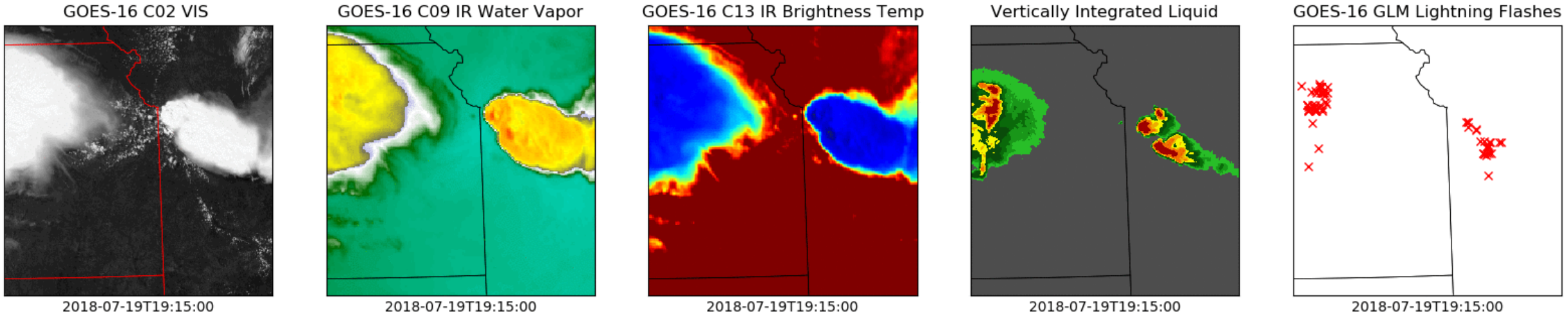
SEVIR File Details

- The SEVIR dataset consists of
  - 107 files in HDF5 format containing imagery
    - Files separated by image type and date range for easier access
    - File sizes ranging between 1GB and 20 GB
  - A catalog containing meta data of each image, including
    - Unique ID assigned to each event
    - Times of each image
    - Lat / Lon bounds of slice
    - Map projection and image extent in projection coordinates for exact georeferencing
    - NCEI Storm Event ID and Episode ID (for non-random cases)

Total Size of SEVIR Dataset: ~900 GB

SEVIR Sample

NCEI Event ID: 772996

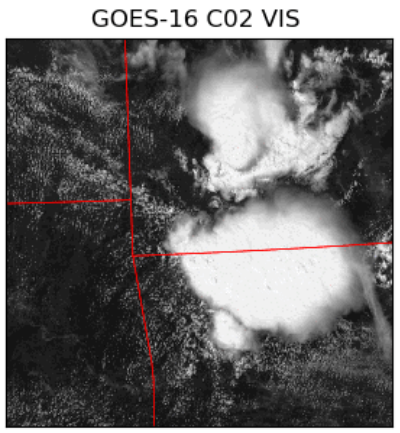


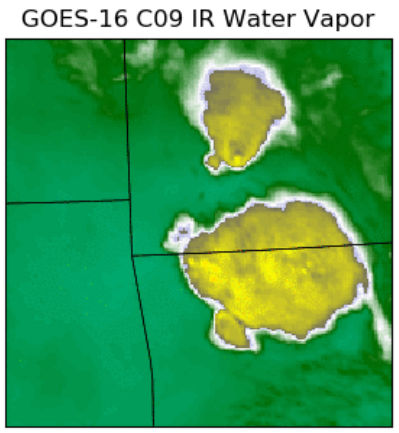
Associated NCEI Database Entry

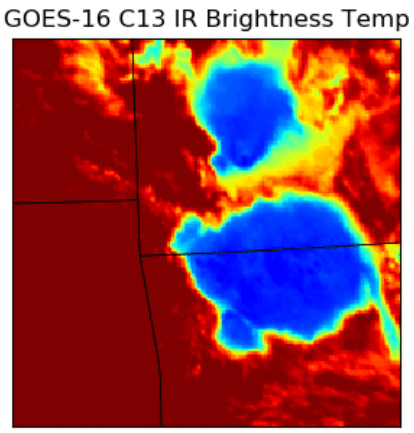
BEGIN_YEARMONTH	BEGIN_DAY	BEGIN_TIME	END_YEARMONTH	END_DAY	END_TIME	EPISODE_ID	EVENT_ID	STATE	STATE_FIPS	
201807	19	1515	201807	19	1518	128915	772996	KANSAS	20	
YEAR	MONTH_NAME	EVENT_TYPE	CZ_TYPE	CZ_FIPS	CZ_NAME	WFO	BEGIN_DATE_TIME	CZ_TIMEZONE	END_DATE_TIME	
2018	July	Thunderstorm Wind	C	121	MIAMI	EAX	7/19/2018 15:15	CST-6	7/19/2018 15:18	
INJURIES_DIRECT	INJURIES_INDIRECT	DEATHS_DIRECT	DEATHS_INDIRECT	DAMAGE_PROPERTY	DAMAGE_CROPS	SOURCE	MAGNITUDE	MAGNITUDE_TYPE	FLOOD_CAUSE	CATEGORY
0	0	0	0	1.00K		Law Enforcement	56	EG		
TOR_F_SCALE	TOR_LENGTH	TOR_WIDTH	TOR_OTHER_WFO	TOR_OTHER_CZ_STATE	TOR_OTHER_CZ_FIPS	TOR_OTHER_CZ_NAME	BEGIN_RANGE	BEGIN_AZIMUTH	BEGIN_LOCATION	END_RANGE
							1	W	OSAWATOMIE	1
END_AZIMUTH	END_LOCATION	BEGIN_LAT	BEGIN_LON	END_LAT	END_LON	EPISODE_NARRATIVE	EVENT_NARRATIVE	DATA_SOURCE		
W	OSAWATOMIE	38.5	-94.95	38.5	-94.95	On July 19, a large complex of storms moved through NW and west central Missouri. The western edge of this storm complex produced some wind damage in far eastern Kansas.	Power lines down on 6th Street due to strong winds.	CSV		

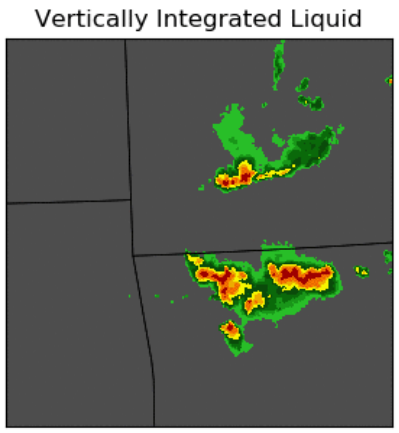
SEVIR Sample

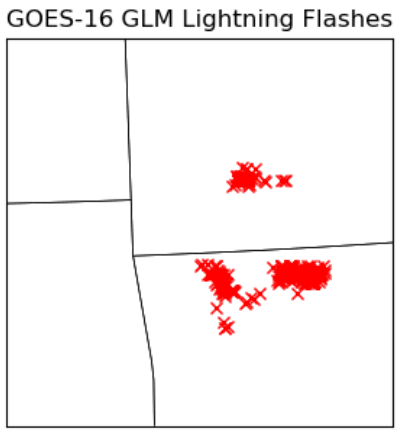
NCEI Event ID: 835047











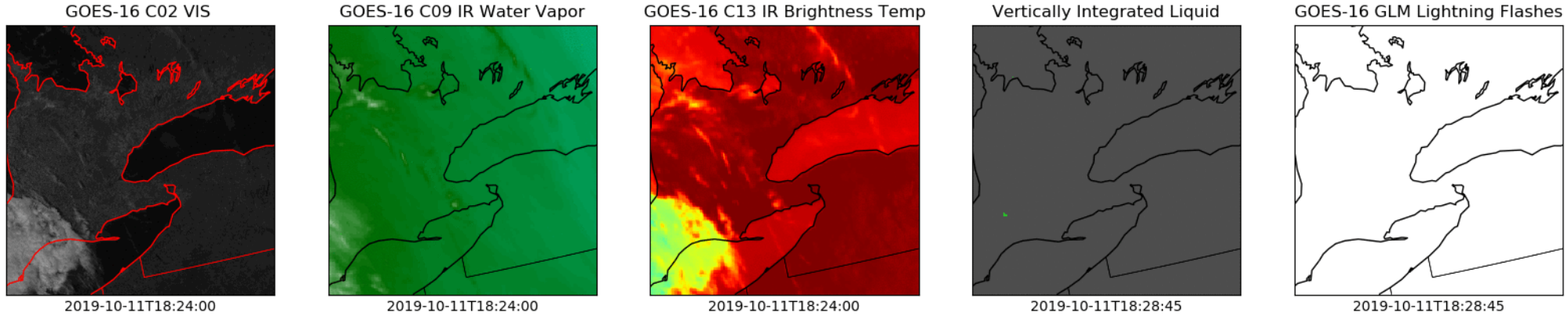
Associated NCEI Database Entry

BEGIN_YEARMONTH	BEGIN_DAY	BEGIN_TIME	END_YEARMONTH	END_DAY	END_TIME	EPISODE_ID	EVENT_ID	STATE	STATE_FIPS	
201906	26	1532	201906	26	1532	138836	835047	MISSOURI	29	
YEAR	MONTH_NAME	EVENT_TYPE	CZ_TYPE	CZ_FIPS	CZ_NAME	WFO	BEGIN_DATE_TIME	CZ_TIMEZONE	END_DATE_TIME	
2019	June	Hail	C	9	BARRY	SGF	6/26/2019 15:32	CST-6	6/26/2019 15:32	
INJURIES_DIRECT	INJURIES_INDIRECT	DEATHS_DIRECT	DEATHS_INDIRECT	DAMAGE_PROPERTY	DAMAGE_CROPS	SOURCE	MAGNITUDE	MAGNITUDE_TYPE	FLOOD_CAUSE	CATEGORY
0	0	0	0	0.00K	0.00K	Law Enforcement	1			
TOR_F_SCALE	TOR_LENGTH	TOR_WIDTH	TOR_OTHER_WFO	TOR_OTHER_CZ_STATE	TOR_OTHER_CZ_FIPS	TOR_OTHER_CZ_NAME	BEGIN_RANGE	BEGIN_AZIMUTH	BEGIN_LOCATION	END_RANGE
							1	N	CASSVILLE	1
END_AZIMUTH	END_LOCATION	BEGIN_LAT	BEGIN_LON	END_LAT	END_LON	EPISODE_NARRATIVE	EVENT_NARRATIVE	DATA_SOURCE		
N	CASSVILLE	36.69	-93.87	36.69	-93.87	Another MCS that developed over northern Missouri moved southeast into the Ozarks. This system brought rain to the western half of the area during the early morning and scattered storms to locations along and south of Interstate 44 during the afternoon and early evening. Most locations received between one half inch and one inch of rain,	Hail up to quarter size was reported.	CSV		

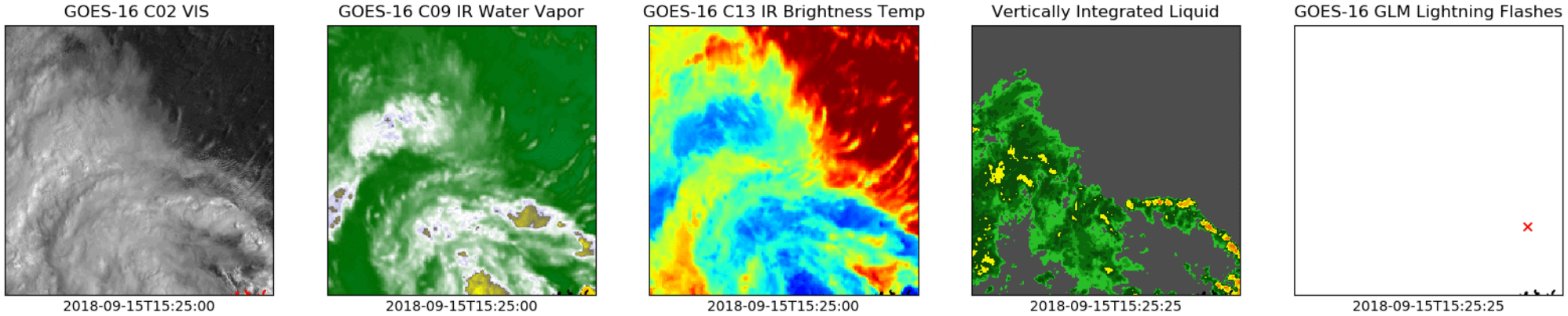
SEVIR Samples

Random Events

Random Event R19101120288358



Random Event R18091517257459



Random Events have no associated NCEI Event ID

SEVIR Applications

- **SEVIR provides an open and reasonably sized yet statistically diverse dataset that AI and Machine learning researchers can use to improve algorithms in several research areas:**
  - **Storm object characterization**
  - **Storm tracking**
  - **Event analysis**
  - **Short-term forecasting (nowcasting)**
  - **Synthetic storm simulation**
  - **Forecast verification**
  - **Many others…**
- **SEVIR can also be combined with other datasets, such as numerical prediction model output, to enable further capabilities**