Space weather monitoring and forecasting: a data-driven approach

Marco Cristoforetti

Bruno Kessler Foundation & Trento Institute for Fundamental Physics and Applications

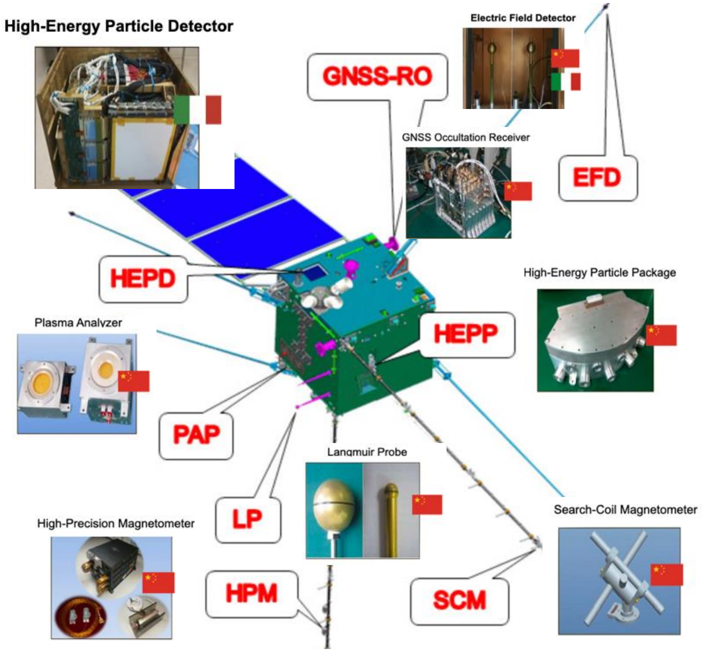
on behalf of the CSES-Limadou collaboration

CSES-01 satellite

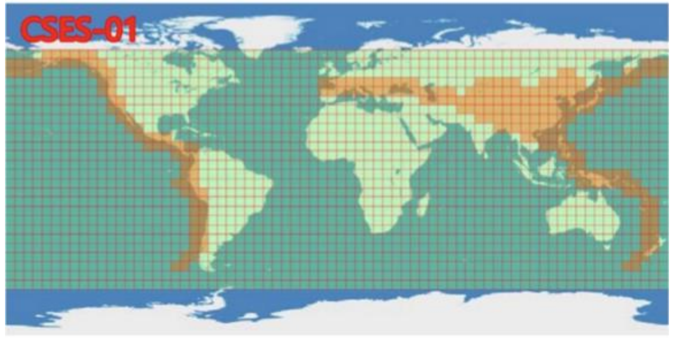
CSES-01 was launched on 02/02/2018

- Sun-Synchronous orbit at 500 km;
- Equipped with 9 instruments
- Payload operation range -65°/65° lat

Category	Payload Name	Observation Targets
Electro-Magnetic Field	Electric Field Detector	Electric Field: DC ～3.5MHz
	High Precision Magnetometer	Magnetic Field: DC ～15Hz
	Search Coil Magnetometer	Magnetic Field: 10Hz ～20kHz
In Situ Plasma	Plasma Analyzer Package	Composition：H^+、He^+、O^+ N_i：5×10^2~$1\times10^7 cm^{-3}$ T_i：500K~10000K
	Langmuir Probe	N_e：5×10^2~$1\times10^7 cm^{-3}$ T_e：500K~10000K
Plasma Construction	GNSS Occultation Receiver	TEC by GNSS Occultation Signal
	Tri-Band Beacon	TEC by transmit VH/U/L Signal
Energetic Particle	Italian HEPD(INFN Prod.) High Energy Particle Package	Proton：2MeV~200MeV Electron：100keV~100MeV



Multi-payload strategy



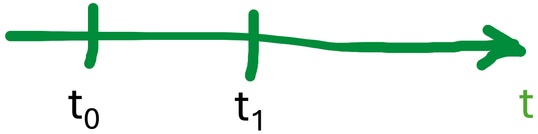
Operation area between lat [-65,65]

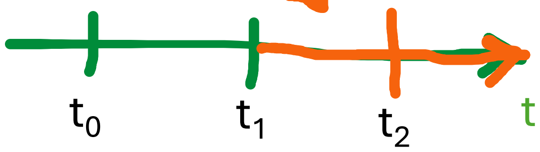
Use case

I have data

I train a model

To predict something in the future
or
extract info from new data

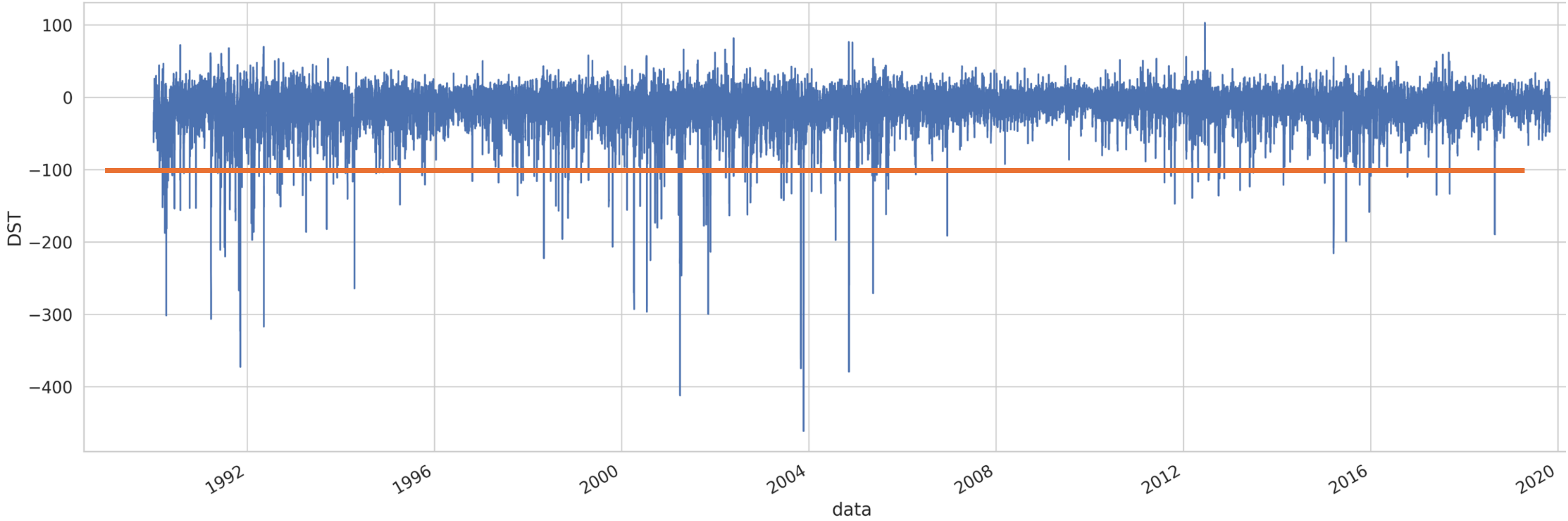




The Disturbance Storm Time Index

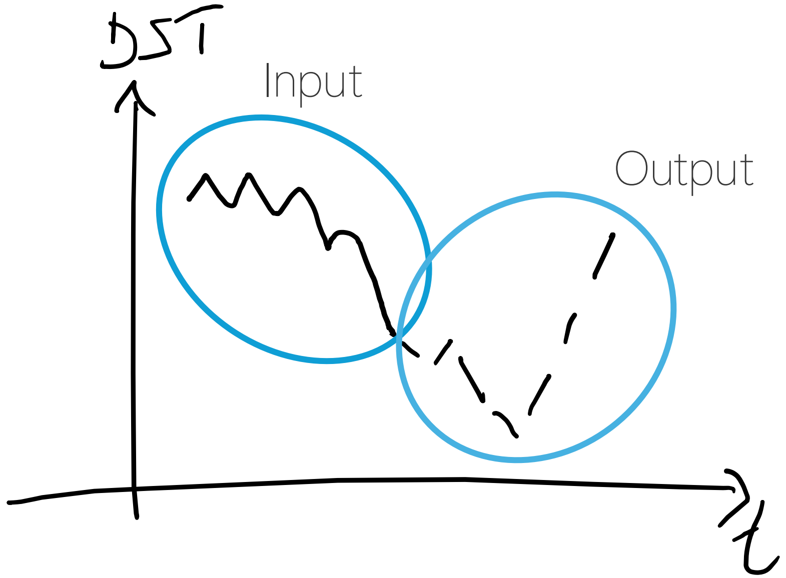
The Dst index is an index of magnetic activity derived from a network of near-equatorial geomagnetic observatories that measures the intensity of the globally symmetrical equatorial electrojet.
Dst is maintained at NGDC and is available via FTP from 1957 to the present.

DST > -20 nT	→ low
-20 nT ≥ DST > -50 nT	→ medium
-50 nT ≥ DST > -100 nT	→ high
DST ≤ -100 nT	→ intense



Forecast with Deep Learning

Machine Learning for DST forecast

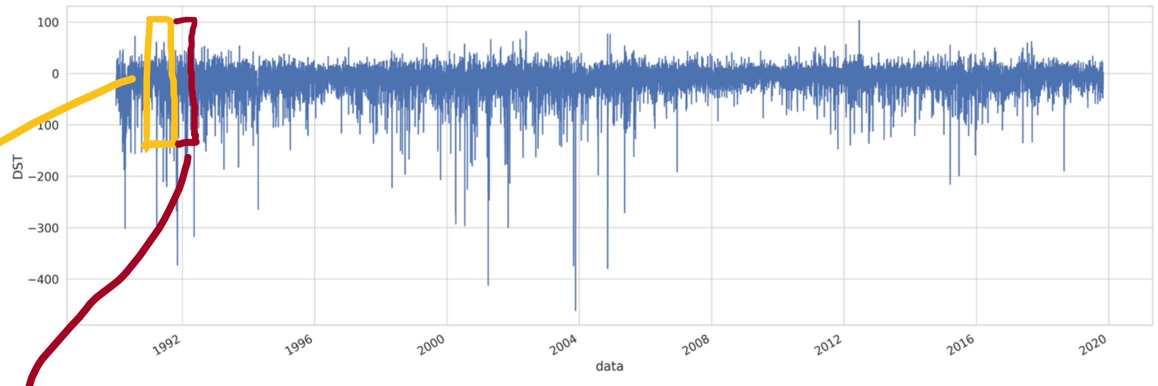


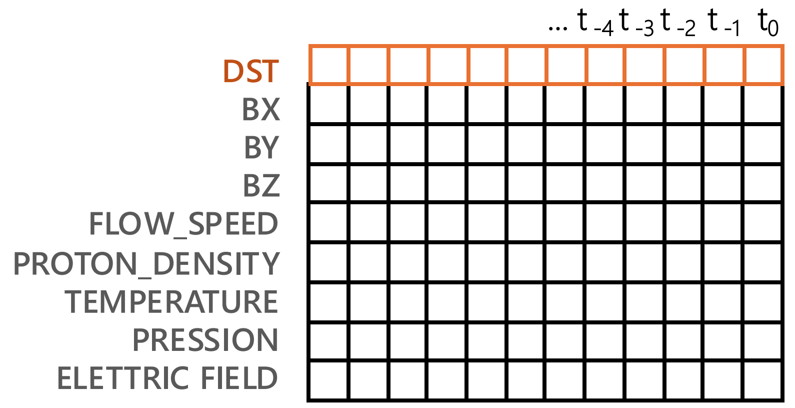
DST
BX, BY, BZ
FLOW_SPEED
PROTON_DENSITY
TEMPERATURE
PRESSURE
ELECTRIC FIELD

PRESSURE = α * PROTON_DENSITY*FLOW_SPEED^2
ELECTRIC = β * BZ * FLOW_SPEED

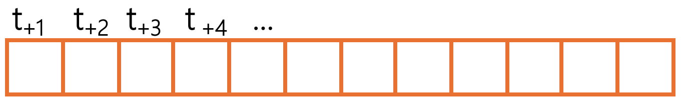
Input – output of the model

12 points (**hours**) to forecast the **next 12** points (**hours**)





INPUT

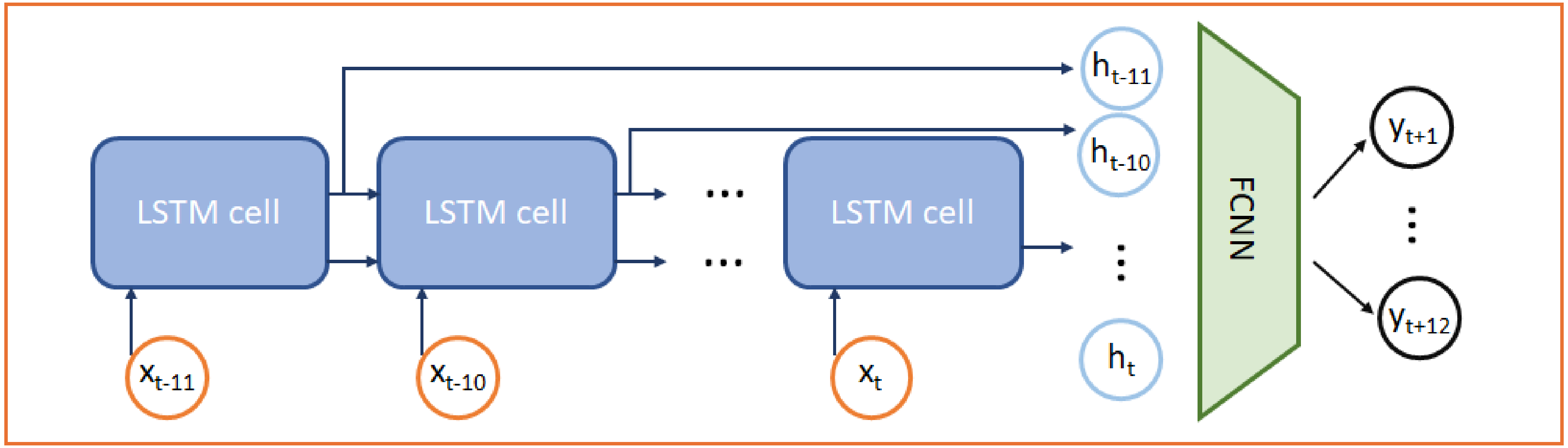


OUTPUT

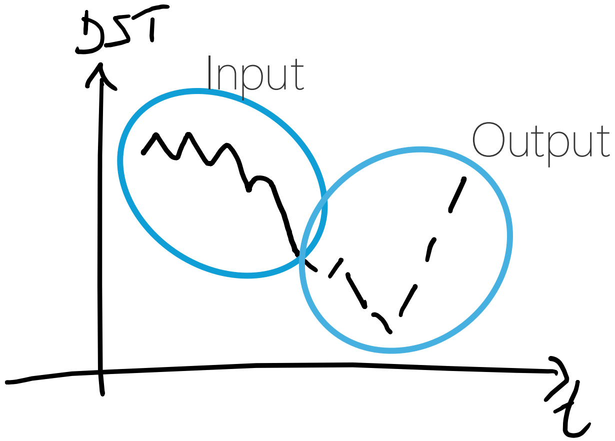
DATASET

186 534 samples

Machine Learning for DST forecast



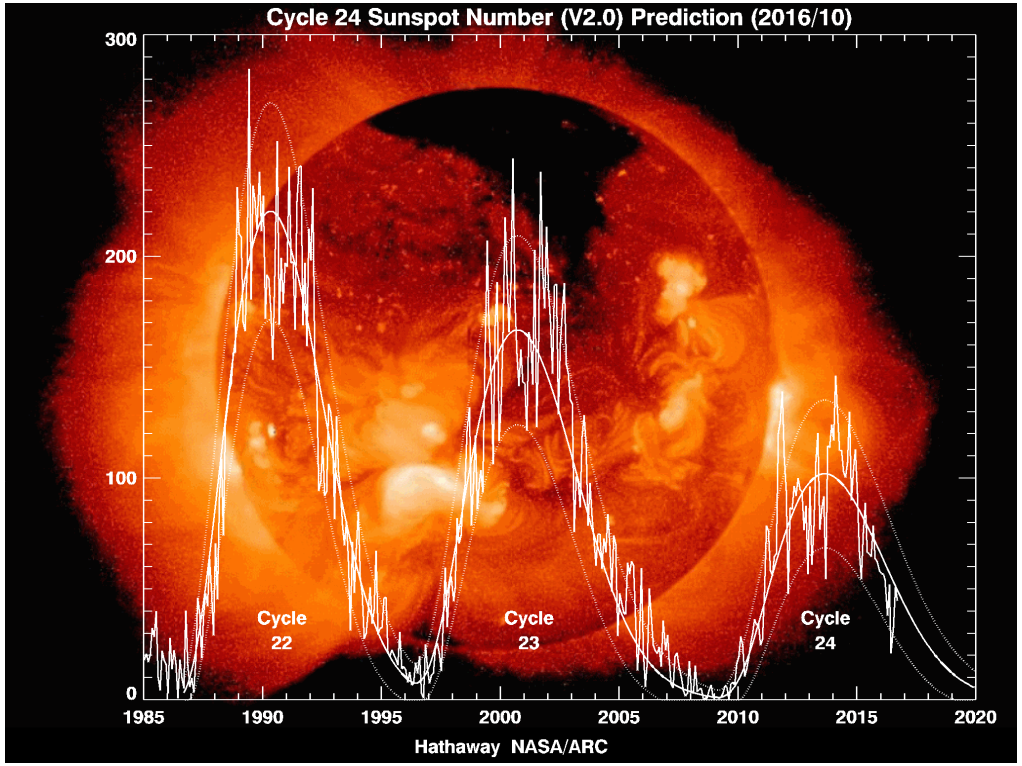
DST
BX, BY, BZ
FLOW_SPEED
PROTON_DENSITY
TEMPERATURE
PRESSURE
ELECTRIC FIELD



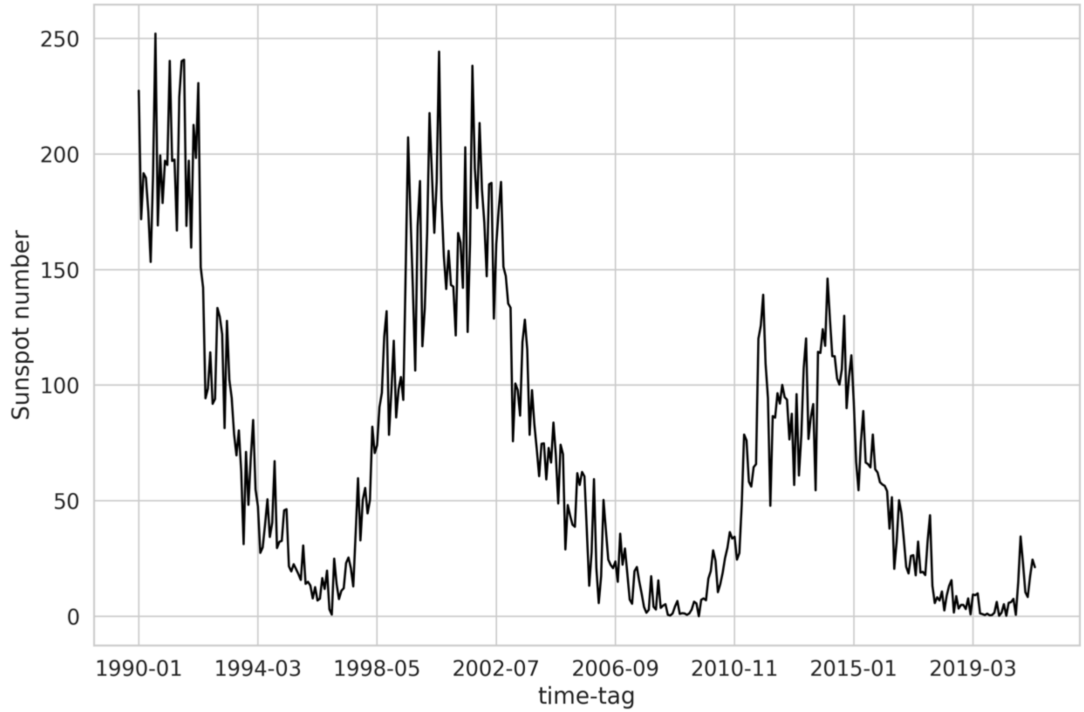
Dealing with time series and rare events

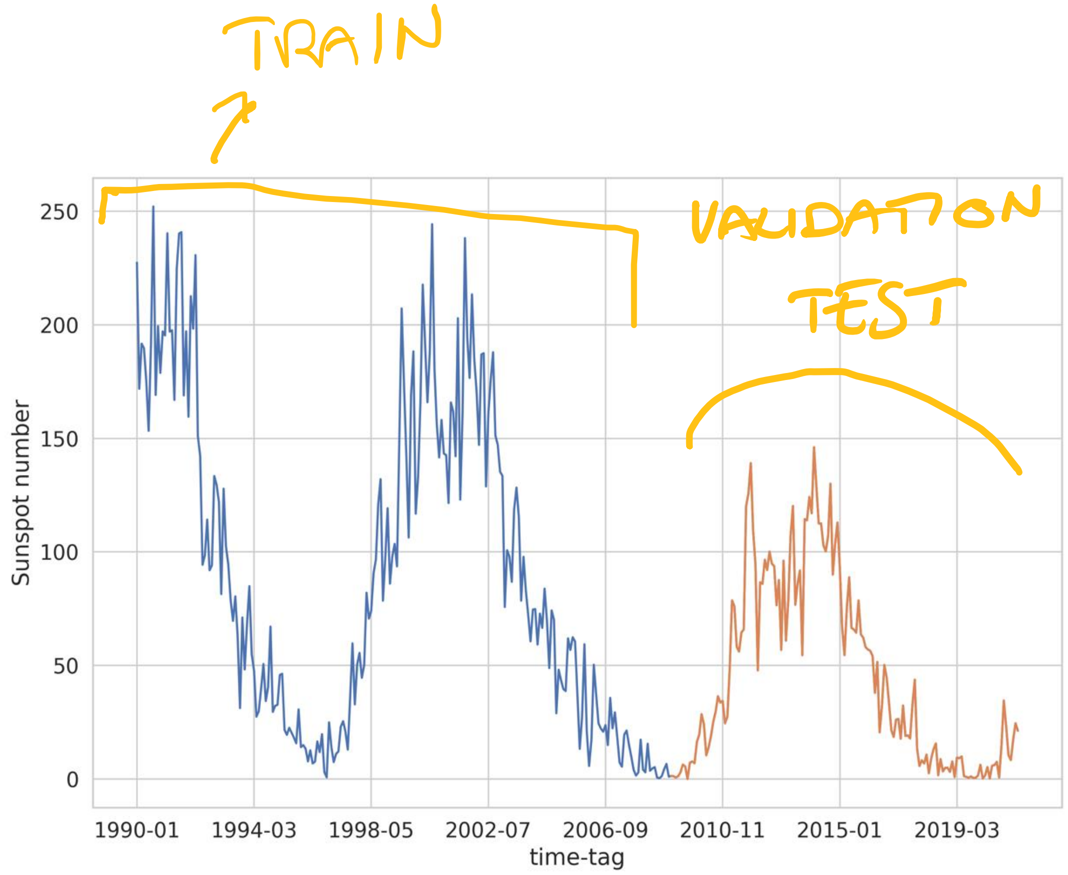
Periodicity and arrow of time. If data are periodic, it is safe to train the model considering at least one complete period and test it on different periods. being the arrow of time fixed and the future unknown, the training operation that make use of points that follow the data used in the test can introduce bias

Forecast of rare events (storms). Training supervised DL model requires a **balanced sampling** of data referring to **quiet** and **storm** periods and **proper metrics** to measure the performances.



Dataset preparation: solar activity



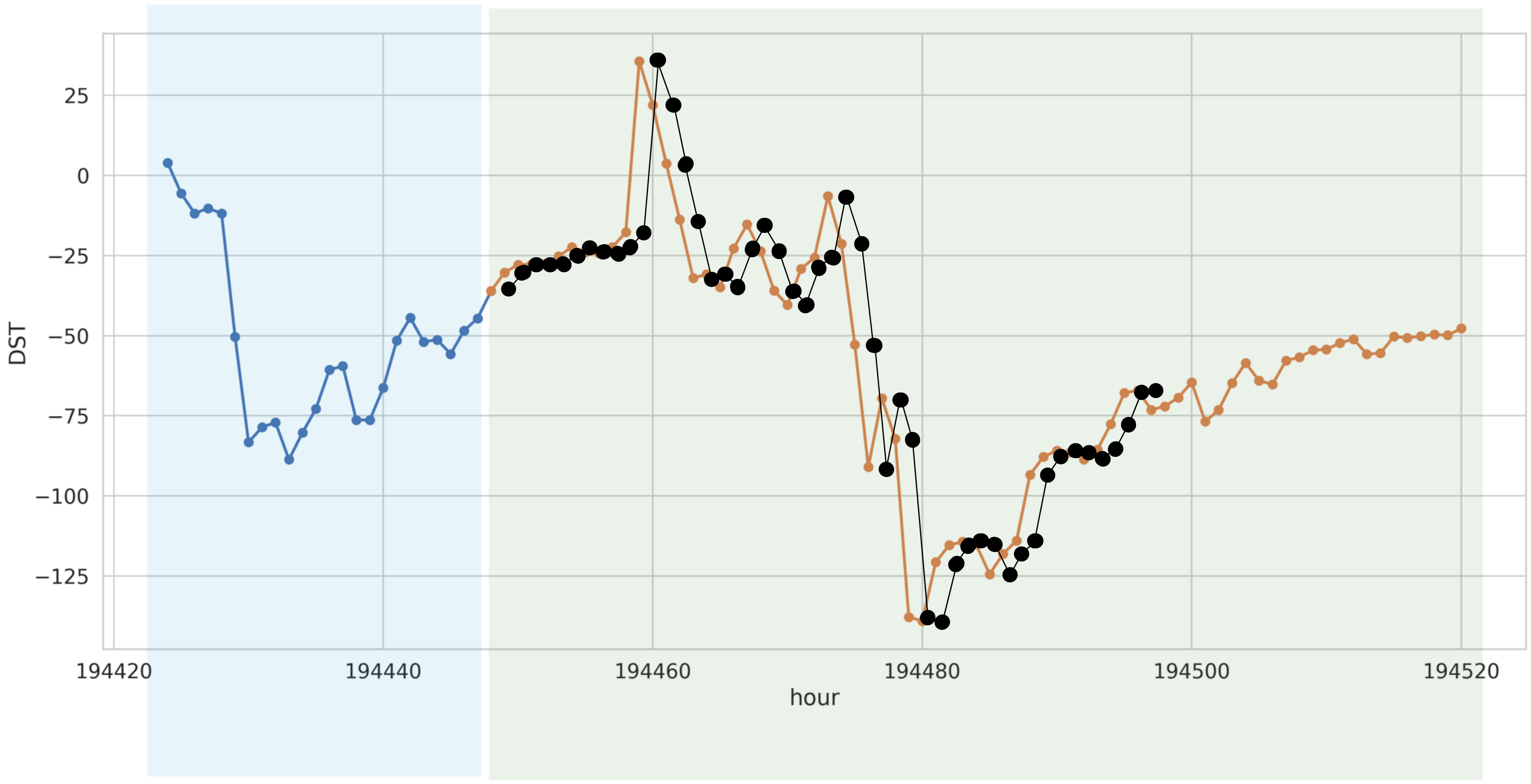


The loss function

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \left(y_{pred_i} - y_{true}\right)^2}{N}}$$

Persistence model

Looks good,
reproduce perfectly the shape,
but the algorithm
does not really learn

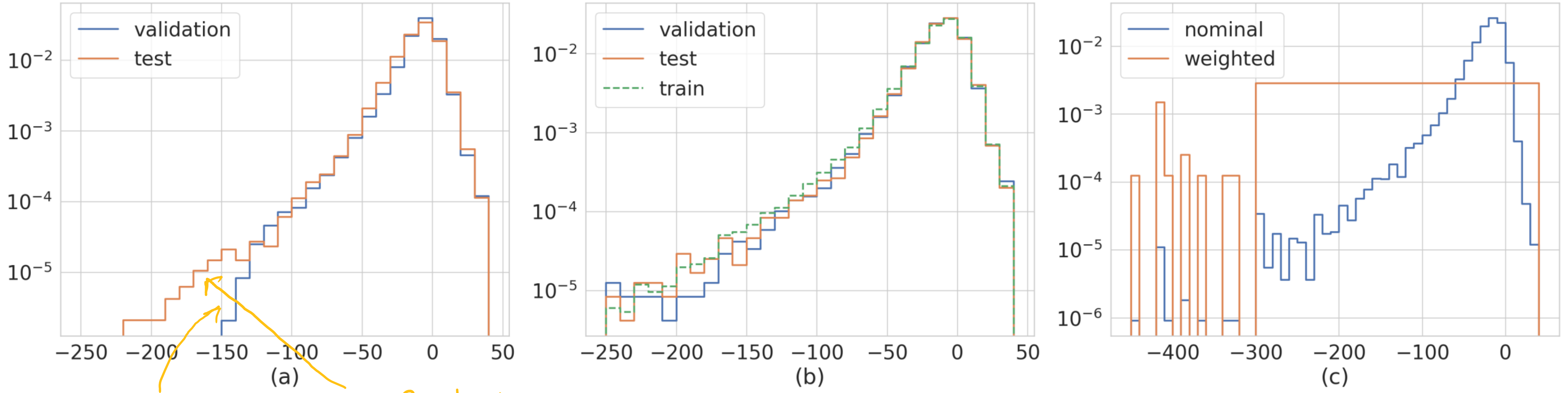


Storm 1

Storm 2

Dataset preparation

Random sampling and reweighting

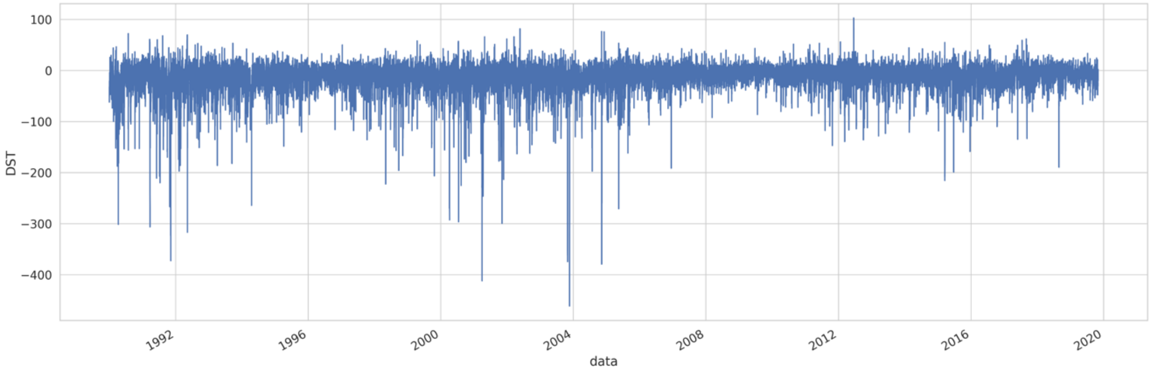


½ cycle val ½ cycle test

RANDOM

TRAIN

Results



Adopted metric: $$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \left(y_{pred_i} - y_{true}\right)^2}{N}}$$

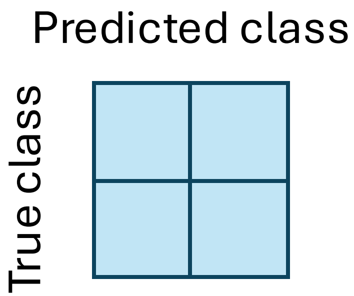
Performance with default sampling looks better

	Train			Valid			Test		
	Pers	Nom	Weight	Pers	Nom	Weight	Pers	Nom	Weight
t + 1h	5.2	4.7	6.7	4.1	3.8	6.2	4.0	**3.8**	6.0
t + 2h	8.3	6.8	7.8	6.4	5.5	7.1	6.4	**5.4**	6.8
t + 3h	10.2	8.3	9.0	8.0	6.8	8.2	8.0	**6.7**	7.9
t + 4h	11.7	9.3	10.0	9.1	7.6	9.2	9.1	**7.5**	8.9
t + 5h	12.9	10.0	10.8	10.0	8.2	10.2	10.0	**8.1**	9.9
t + 6h	13.9	10.7	11.9	10.7	8.6	11.4	10.8	**8.7**	11.0
t + 7h	14.9	11.4	13.2	11.5	9.1	12.6	11.5	**9.2**	12.3
t + 8h	15.7	12.0	14.7	12.1	9.5	14.1	12.2	**9.7**	13.9
t + 9h	16.4	12.6	16.1	12.7	9.9	15.5	12.8	**10.1**	15.5
t + 10h	17.0	13.1	17.3	13.2	10.3	16.8	13.2	**10.4**	16.7
t + 11h	17.6	13.6	18.1	13.6	10.7	17.4	13.6	**10.8**	17.4
t + 12h	18.1	14.1	18.4	14.0	10.9	17.4	14.0	**11.0**	17.6

Using the weighted dataset the performances on the storms improve

	Dst> −20 nT			−20 nT > *Dst* > −50 nT			−50 nT > *Dst* > −100 nT			Dst< −100 nT		
	Pers	Nom	Weight	Pers	Nom	Weight	Pers	Nom	Weight	Pers	Full	Weight
t + 1h	**3.4**	**3.4**	5.8	5.6	**4.1**	6.2	10.0	**9.9**	11.8	**17.0**	18.2	23.6
t + 2h	5.2	**4.7**	6.6	8.8	**6.7**	7.8	17.5	14.5	**13.8**	31.4	28.6	**27.9**
t + 3h	6.3	**5.6**	7.5	11.0	**8.5**	9.1	22.9	18.6	**16.8**	43.2	40.0	**35.0**
t + 4h	7.0	**6.1**	8.5	12.5	**9.7**	10.4	27.4	21.5	**18.4**	54.1	49.8	**43.2**
t + 5h	7.5	**6.4**	9.2	13.7	**10.5**	11.7	31.0	24.4	**20.9**	63.8	56.8	**46.6**
t + 6h	7.9	**6.6**	10.2	14.9	**11.1**	13.0	34.3	26.9	**22.6**	73.2	63.5	**50.4**
t + 7h	8.2	**6.7**	11.2	15.9	**11.7**	14.7	37.6	29.7	**24.1**	81.4	68.3	**53.8**
t + 8h	8.5	**6.9**	12.7	16.8	**12.4**	16.3	40.5	32.1	**25.5**	87.9	73.0	**56.8**
t + 9h	8.8	**7.0**	14.0	17.6	**12.9**	17.7	42.9	34.3	**27.4**	93.4	77.3	**60.1**
t + 10h	9.0	**7.1**	15.3	18.2	**13.3**	18.7	44.9	36.1	**29.0**	97.4	82.4	**64.6**
t + 11h	9.2	**7.2**	15.9	18.9	**13.7**	19.4	46.7	37.8	**30.6**	100.7	86.7	**69.9**
t + 12h	9.4	**7.3**	15.9	19.5	**14.0**	19.9	48.1	39.3	**31.8**	103.4	90.4	**73.8**

Confusion matrix
Full dataset



persistence t + 1

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.97 (61842)	0.03 (1620)	0.00 (0)	0.00 (0)
-20>dst>-50	0.14 (1633)	0.84 (10063)	0.02 (251)	0.00 (0)
-50>dst>-100	0.00 (3)	0.18 (244)	0.80 (1095)	0.02 (34)
dst<-100	0.00 (0)	0.01 (1)	0.22 (33)	0.77 (115)

nominal t + 1

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.98 (62054)	0.02 (1407)	0.00 (1)	0.00 (0)
-20>dst>-50	0.11 (1361)	0.88 (10516)	0.01 (70)	0.00 (0)
-50>dst>-100	0.00 (1)	0.31 (426)	0.68 (940)	0.01 (9)
dst<-100	0.00 (0)	0.00 (0)	0.42 (63)	0.58 (86)

weight t + 1

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.95 (60131)	0.05 (3328)	0.00 (3)	0.00 (0)
-20>dst>-50	0.15 (1817)	0.82 (9810)	0.03 (319)	0.00 (1)
-50>dst>-100	0.00 (0)	0.24 (326)	0.73 (1008)	0.03 (42)
dst<-100	0.00 (0)	0.00 (0)	0.36 (53)	0.64 (96)

persistence t + 6

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.94 (59398)	0.06 (4113)	0.00 (7)	0.00 (0)
-20>dst>-50	0.32 (3847)	0.62 (7340)	0.06 (701)	0.00 (4)
-50>dst>-100	0.15 (209)	0.32 (447)	0.45 (621)	0.07 (103)
dst<-100	0.17 (24)	0.19 (28)	0.35 (50)	0.29 (42)

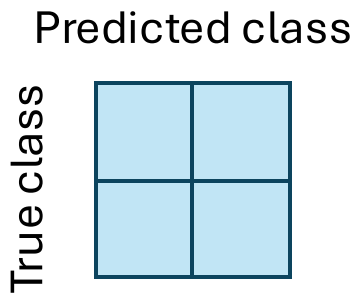
nominal t + 6

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.97 (61337)	0.03 (2168)	0.00 (13)	0.00 (0)
-20>dst>-50	0.37 (4343)	0.62 (7431)	0.01 (118)	0.00 (0)
-50>dst>-100	0.09 (121)	0.56 (776)	0.34 (469)	0.01 (14)
dst<-100	0.09 (13)	0.19 (27)	0.58 (84)	0.14 (20)

weight t + 6

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.87 (55532)	0.12 (7699)	0.00 (272)	0.00 (15)
-20>dst>-50	0.20 (2322)	0.72 (8620)	0.08 (911)	0.00 (39)
-50>dst>-100	0.04 (57)	0.32 (439)	0.59 (811)	0.05 (73)
dst<-100	0.03 (4)	0.15 (21)	0.49 (71)	0.33 (48)

Confusion matrix Data with input DST > -20nT



	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	1.00 (51499)	0.00 (0)	0.00 (0)	0.00 (0)
-20>dst>-50	1.00 (1380)	0.00 (0)	0.00 (0)	0.00 (0)
-50>dst>-100	1.00 (50)	0.00 (0)	0.00 (0)	0.00 (0)
dst<-100	1.00 (5)	0.00 (0)	0.00 (0)	0.00 (0)

persistence t + 3

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	1.00 (51357)	0.00 (140)	0.00 (2)	0.00 (0)
-20>dst>-50	0.85 (1175)	0.15 (204)	0.00 (1)	0.00 (0)
-50>dst>-100	0.38 (19)	0.50 (25)	0.12 (6)	0.00 (0)
dst<-100	0.40 (2)	0.60 (3)	0.00 (0)	0.00 (0)

nominal t + 3

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.98 (50597)	0.02 (894)	0.00 (7)	0.00 (1)
-20>dst>-50	0.69 (954)	0.29 (404)	0.02 (22)	0.00 (0)
-50>dst>-100	0.32 (16)	0.50 (25)	0.16 (8)	0.02 (1)
dst<-100	0.20 (1)	0.80 (4)	0.00 (0)	0.00 (0)

weight t + 3

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	1.00 (50808)	0.00 (0)	0.00 (0)	0.00 (0)
-20>dst>-50	1.00 (1941)	0.00 (0)	0.00 (0)	0.00 (0)
-50>dst>-100	1.00 (165)	0.00 (0)	0.00 (0)	0.00 (0)
dst<-100	1.00 (20)	0.00 (0)	0.00 (0)	0.00 (0)

persistence t + 6

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	1.00 (50646)	0.00 (159)	0.00 (3)	0.00 (0)
-20>dst>-50	0.90 (1744)	0.10 (194)	0.00 (3)	0.00 (0)
-50>dst>-100	0.58 (95)	0.35 (57)	0.08 (13)	0.00 (0)
dst<-100	0.50 (10)	0.40 (8)	0.10 (2)	0.00 (0)

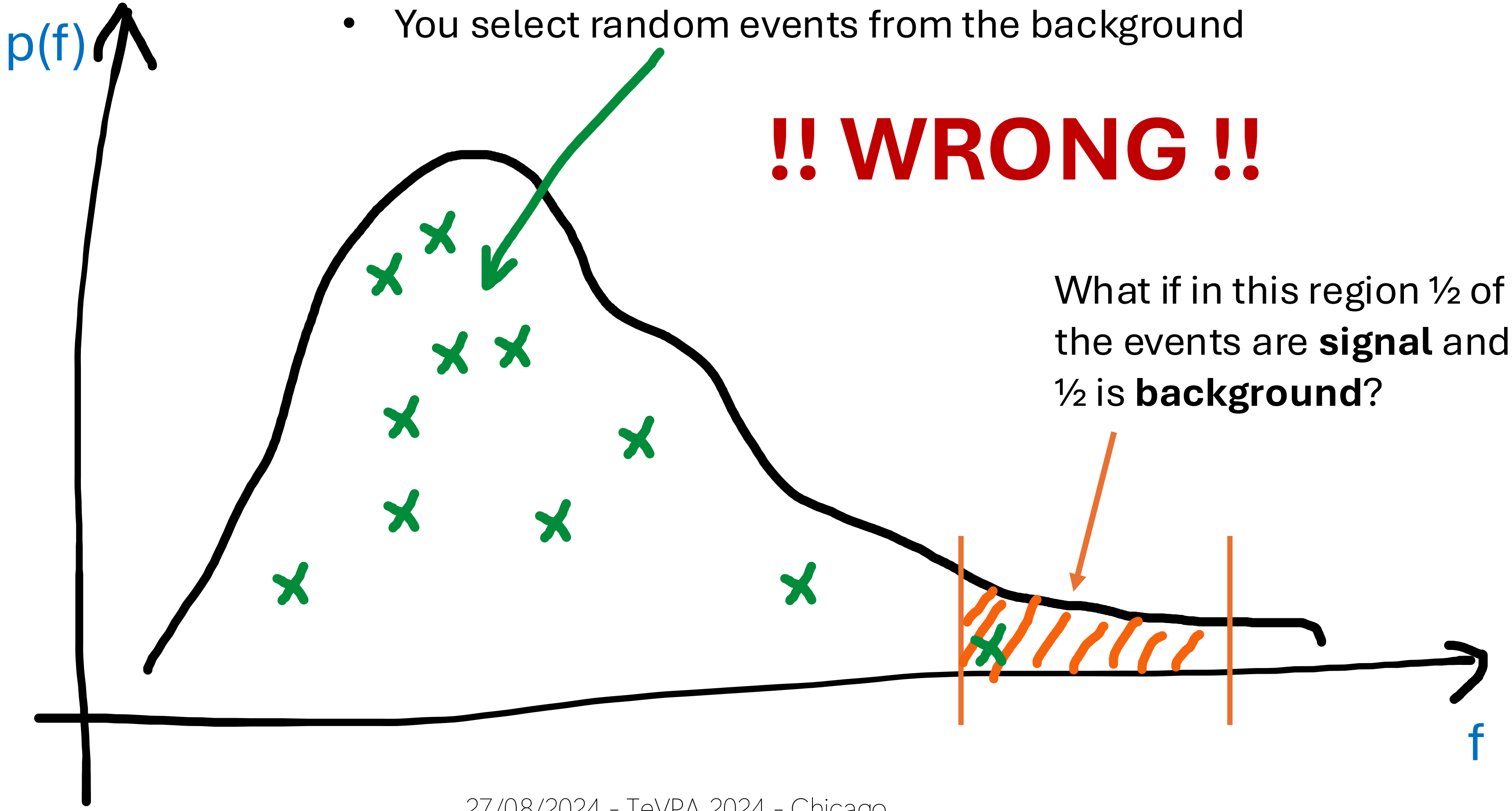
nominal t + 6

	dst>-20	-20>dst>-50	-50>dst>-100	dst<-100
dst>-20	0.96 (48867)	0.04 (1824)	0.00 (110)	0.00 (7)
-20>dst>-50	0.66 (1276)	0.30 (583)	0.03 (65)	0.01 (17)
-50>dst>-100	0.33 (54)	0.39 (65)	0.24 (40)	0.04 (6)
dst<-100	0.10 (2)	0.65 (13)	0.15 (3)	0.10 (2)

weight t + 6

Searching for rare events

- You want a classifier that given f signal vs background
- You have the signal (all data in the orange region)
- You select random events from the background



!! WRONG !!

What if in this region ½ of the events are **signal** and ½ is **background**?