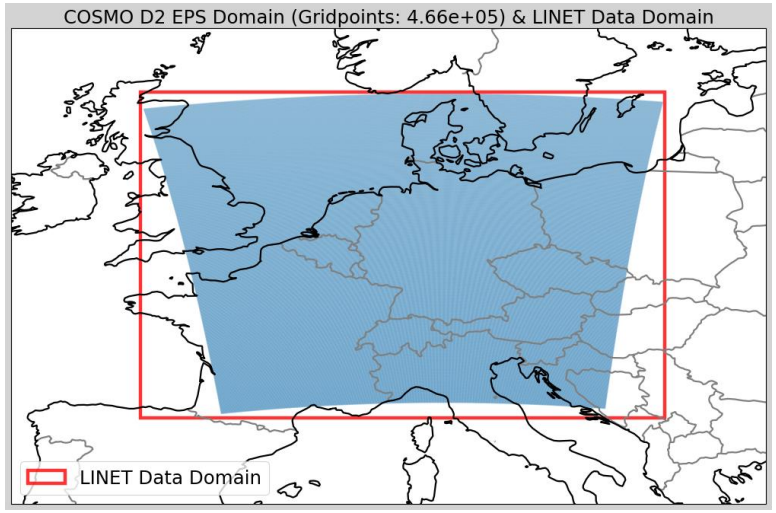
Predictability analysis and verification of the Lightning Potential Index (LPI) in the COSMO-D2 high resolution EPS

Advisors: Dr. Chiara Marsigli – DWD Ass.-Prof. Mag. Dr. Manfred Dorninger – IMGW

SUMMER (JJA) 2019

Observation Dataset

- Data from LINET/DWD lightning detection Network
- hourly aggregated and gridded to match COSMO's fields



Forecast Dataset

- COSMO D2 EPS hourly LPI fields [J/kg]
- 00Z Run, 24 hours forecast

- Verification of the ENS Mean with the Symmetric Extremal Dependence Index (SEDI)

$$SEDI = \frac{lnF - lnH - ln(1-F) + ln(1-H)}{lnF + lnH + ln(1-F) + ln(1-H)}$$

H = Hit Rate
F = False Alarm Rate
Ferro, Stephenson 2011

- Predictability and spread-error analysis through the dispersion Fractions Skill Score (dFSS)

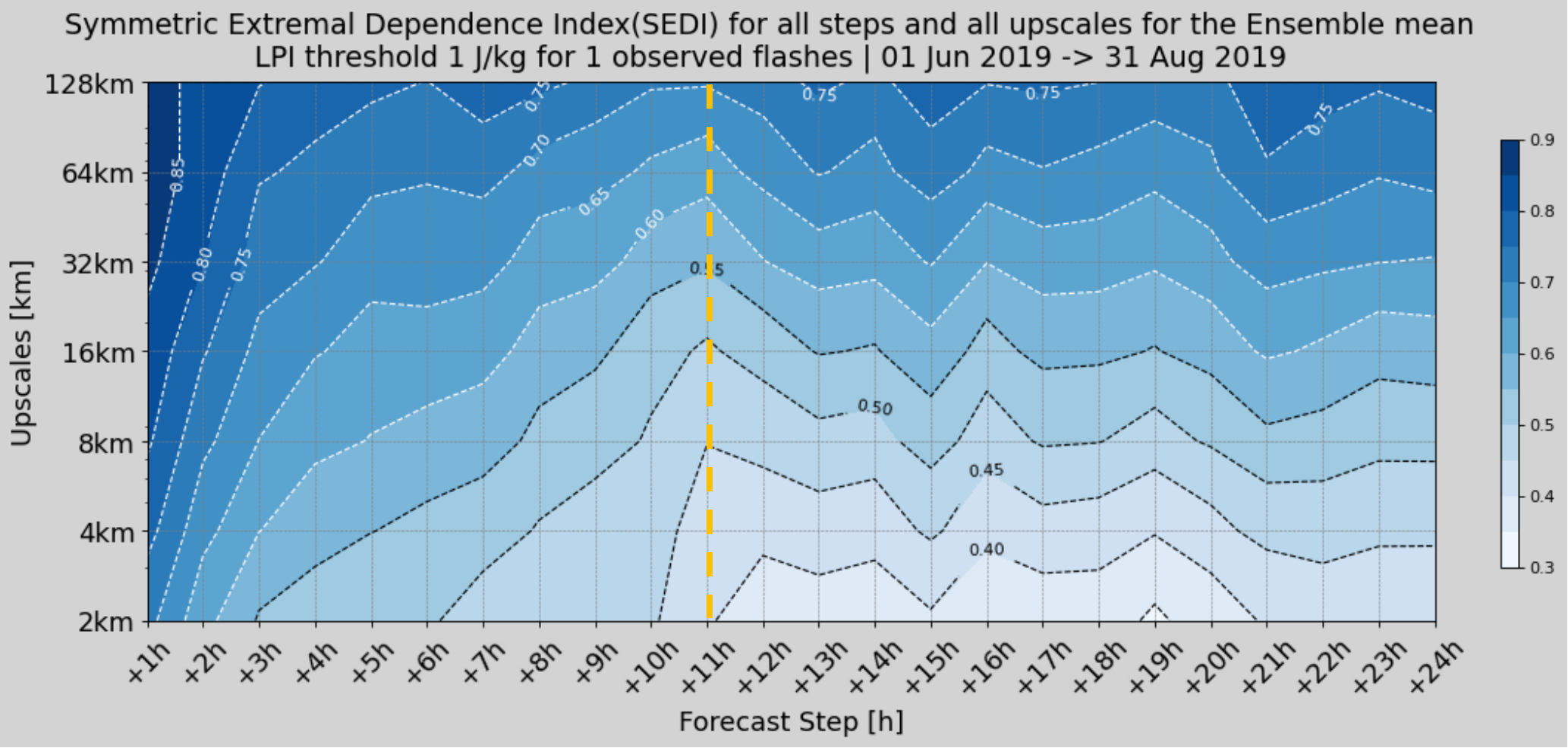
dFSS = FSS calculated between members of the ensemble, then averaged. Dey et al. 2014 – **Analysis filtered for days with strong lightning activity**

- Further analysis of the EPS by adapting the Structure Amplitude Location measure to the ensemble (eSAL)

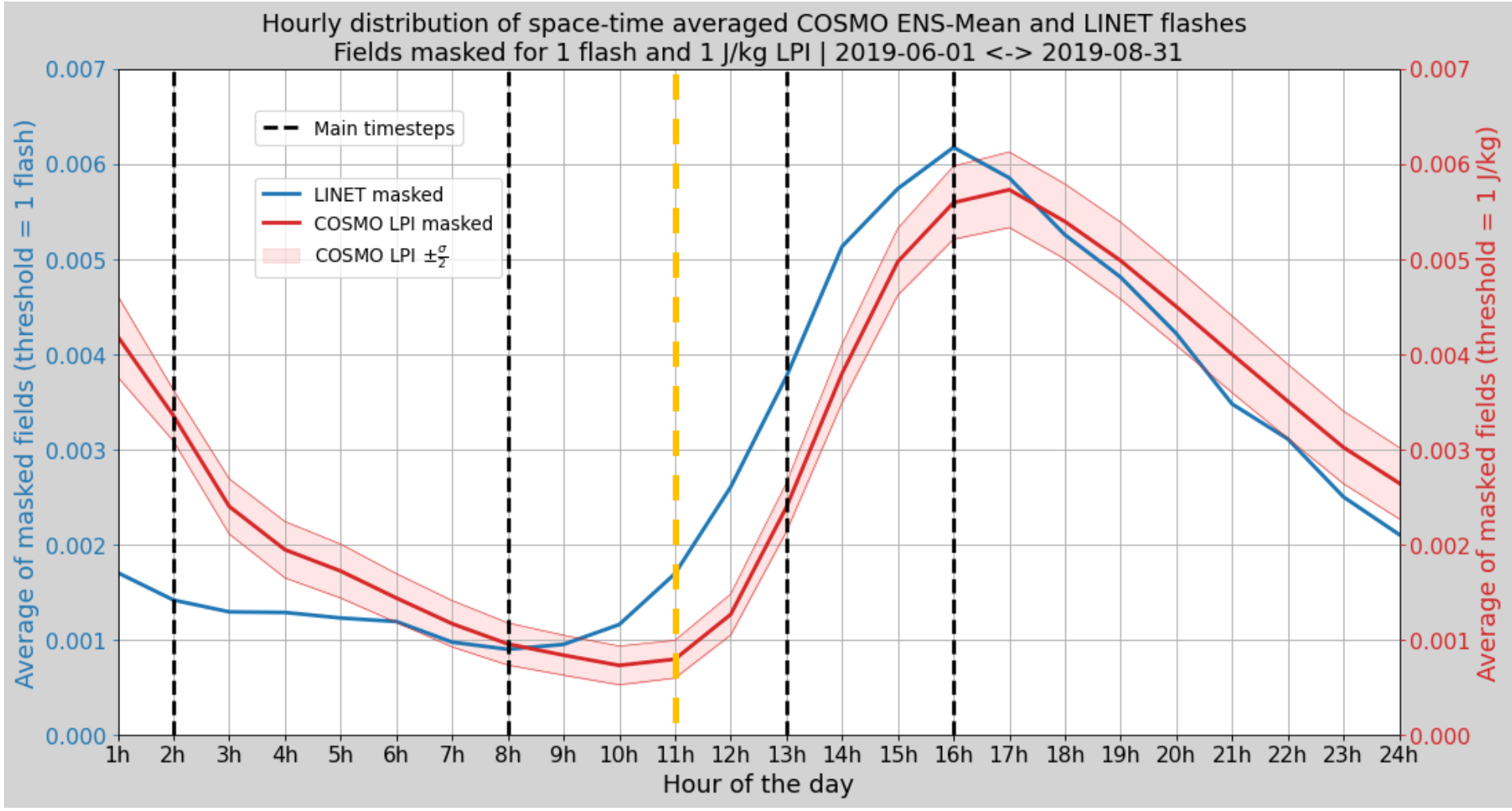
eSAL = SAL calculated between members of the ensemble, then averaged. Adapted from Radanovics et al. 2018 - **Analysis filtered for days and regions with strong lightning activity**

Verification of the ensemble mean for all forecast steps and «simple» upscalings of the LPI fields

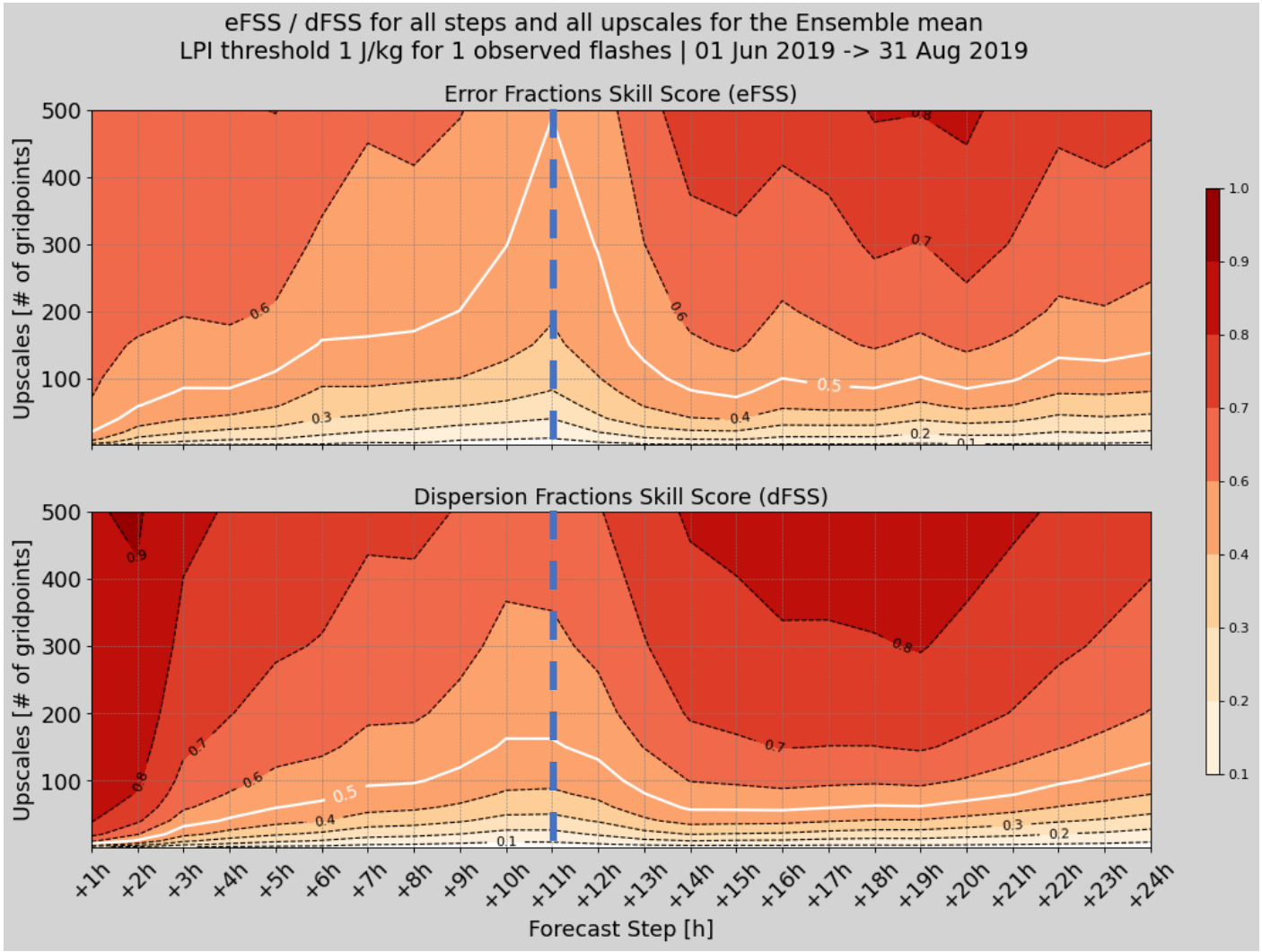
SEDI [-1,+1] is a base-rate independent measure specifically developed for rare events | For random forecast SEDI = 0



Space-time average of masked (0,1) observation/forecast fields for all the forecast steps (= hour of the day in our case)

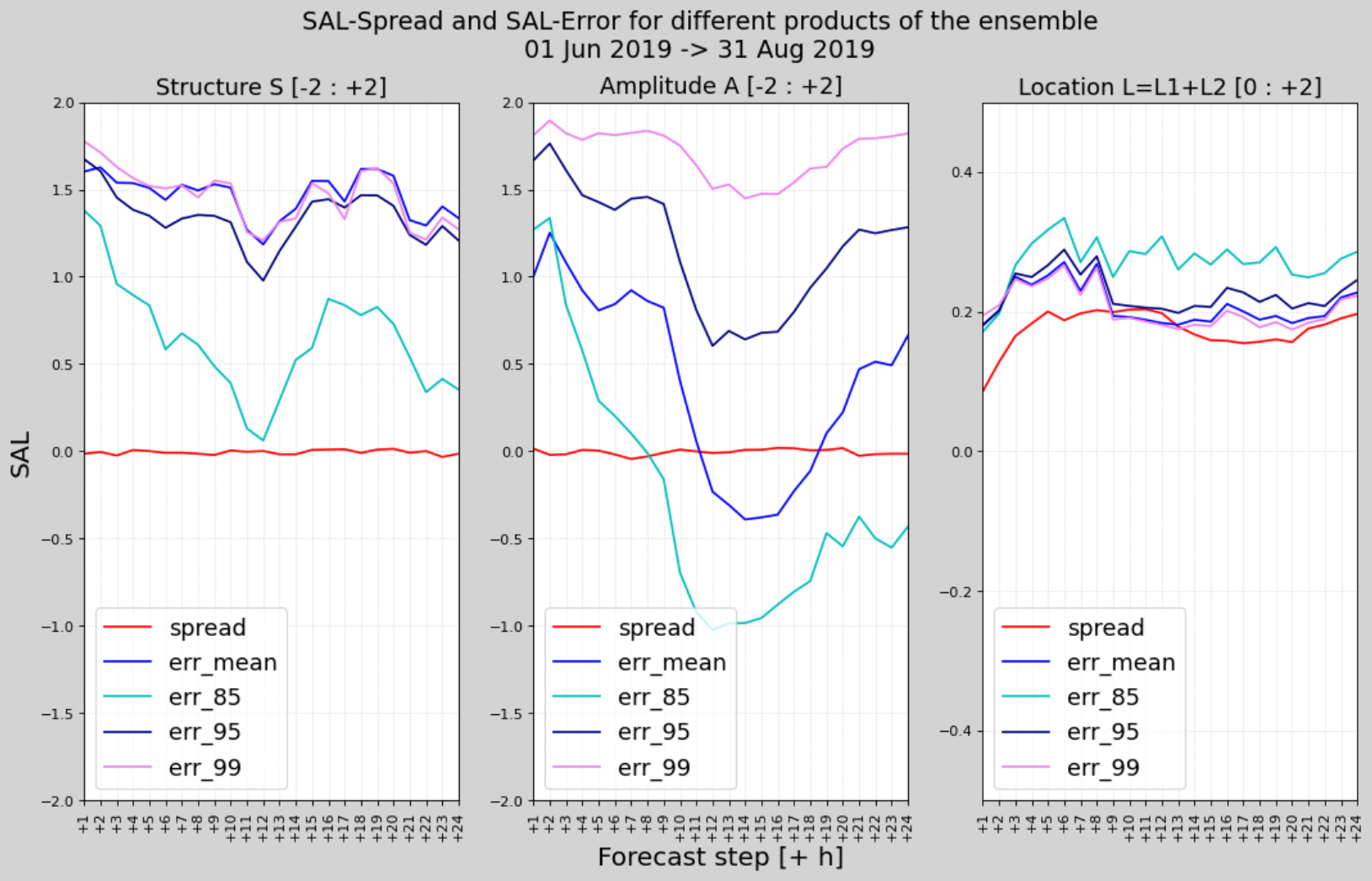


Assessment of the spread-error relationship based on the Fractions Skill Score



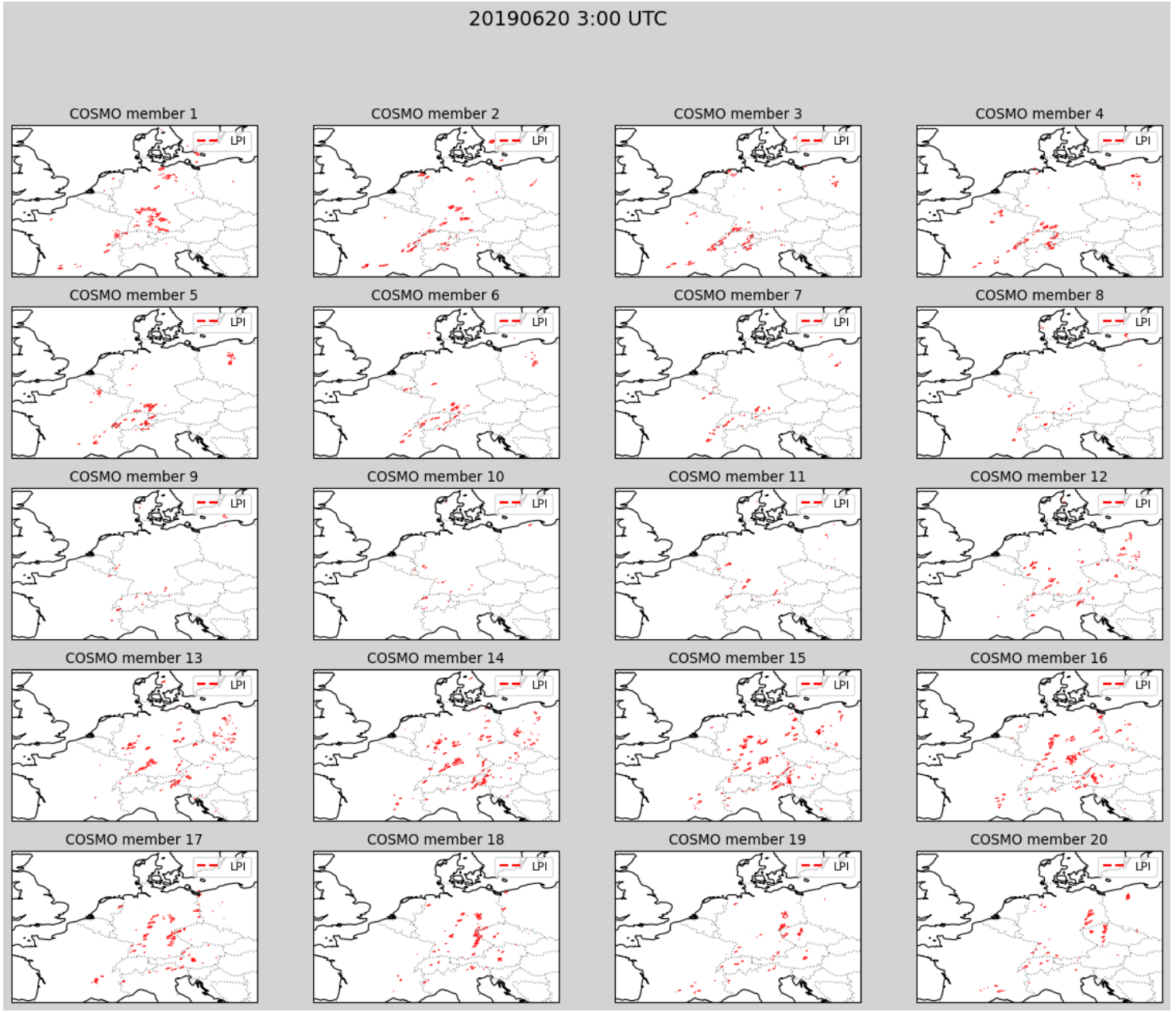
- Maximum bias at the start of the convective cycle (eFSS, above). The ensemble correctly predict a decrease in the predictability (dFSS below, equivalent to ENS-spread).
- The ensemble is overall slightly underdispersive.
- “Useful Forecast” (FSS > 0.5 + fo/2, white line) during the main convective window at a scale of around 100 gridpoints (~200 km). The ensemble would rate the forecast as “useful” already at 60 gridpoints (~120 km).

Assessment of the spread-error relationship based on the SAL scores



- S and A components investigate model biases that will not be addressed by the EPS (red line, dSAL).
- S: areas with lightning activity are too wide and/or intense compared to observations.
- A: nighttime positive bias remains evident.
- L component confirms a slight underdispersion in the EPS, with better spread-error ratio at the beginning of the convective cycle (+9/+12).

Thank you for your attention!



Main references

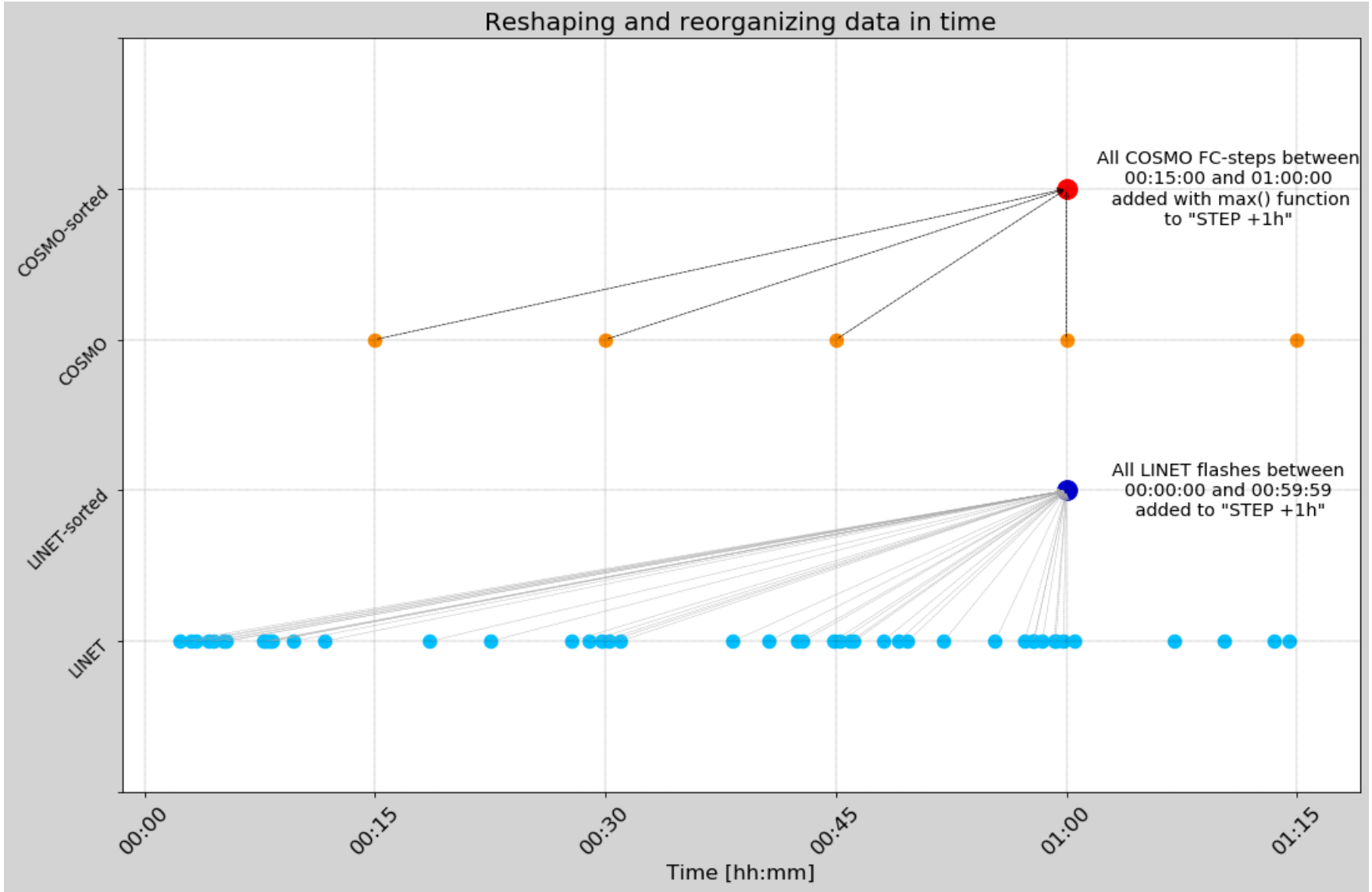
- **SEDI** – Ferro, C. A. T., and Stephenson D. B., 2011: **Extremal Dependence Indices: Improved Verification Measures for Deterministic Forecasts of Rare Binary Events.** *Wea. Forecasting*, 26, 699–713.

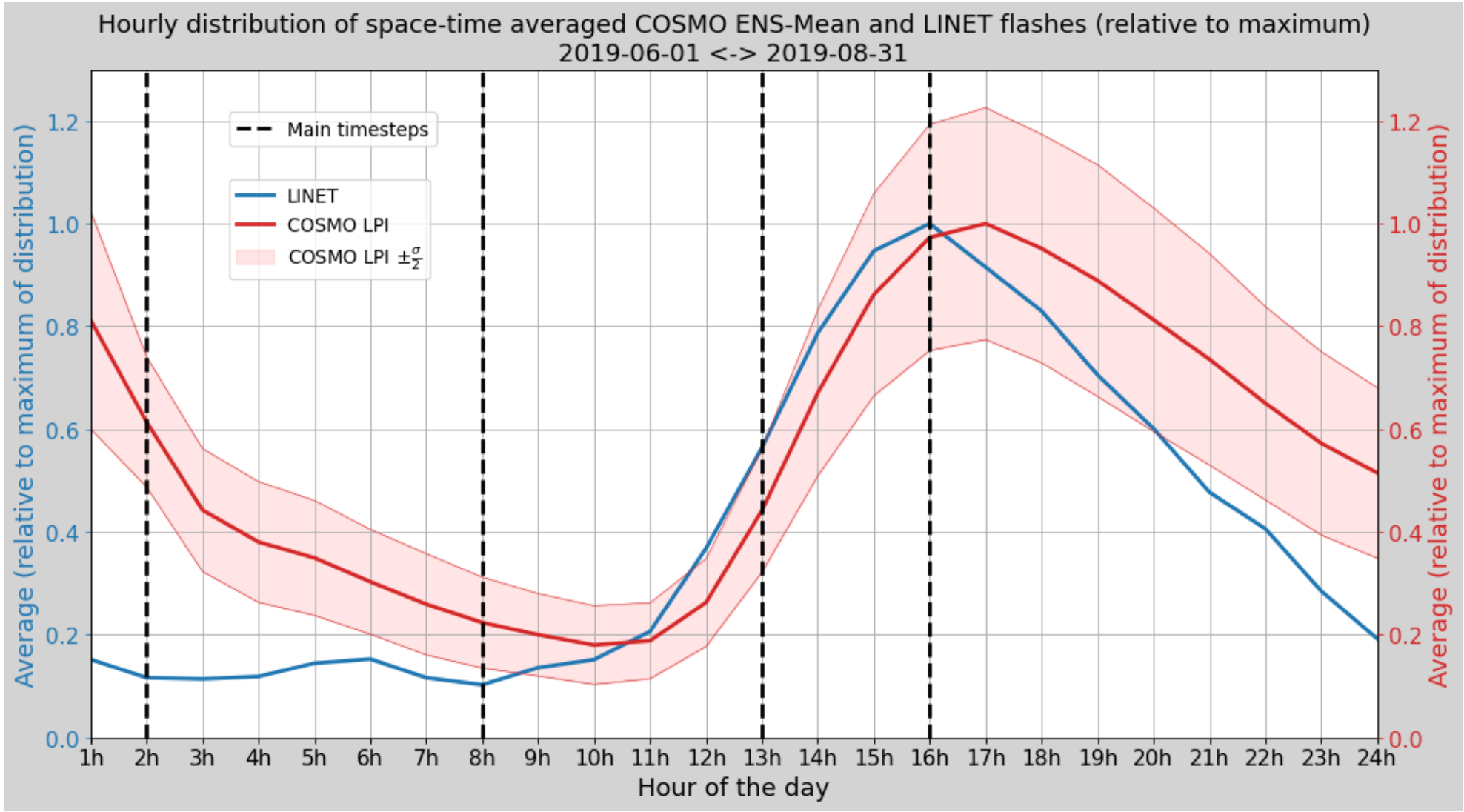
- **dFSS** – Dey, S. R. A., Leoncini G., Roberts N. M., Plant R. S., Migliorini S., 2014: **A Spatial View of Ensemble Spread in Convection Permitting Ensembles.** *Mon. Wea. Rev.*, 142, 4091–4107.

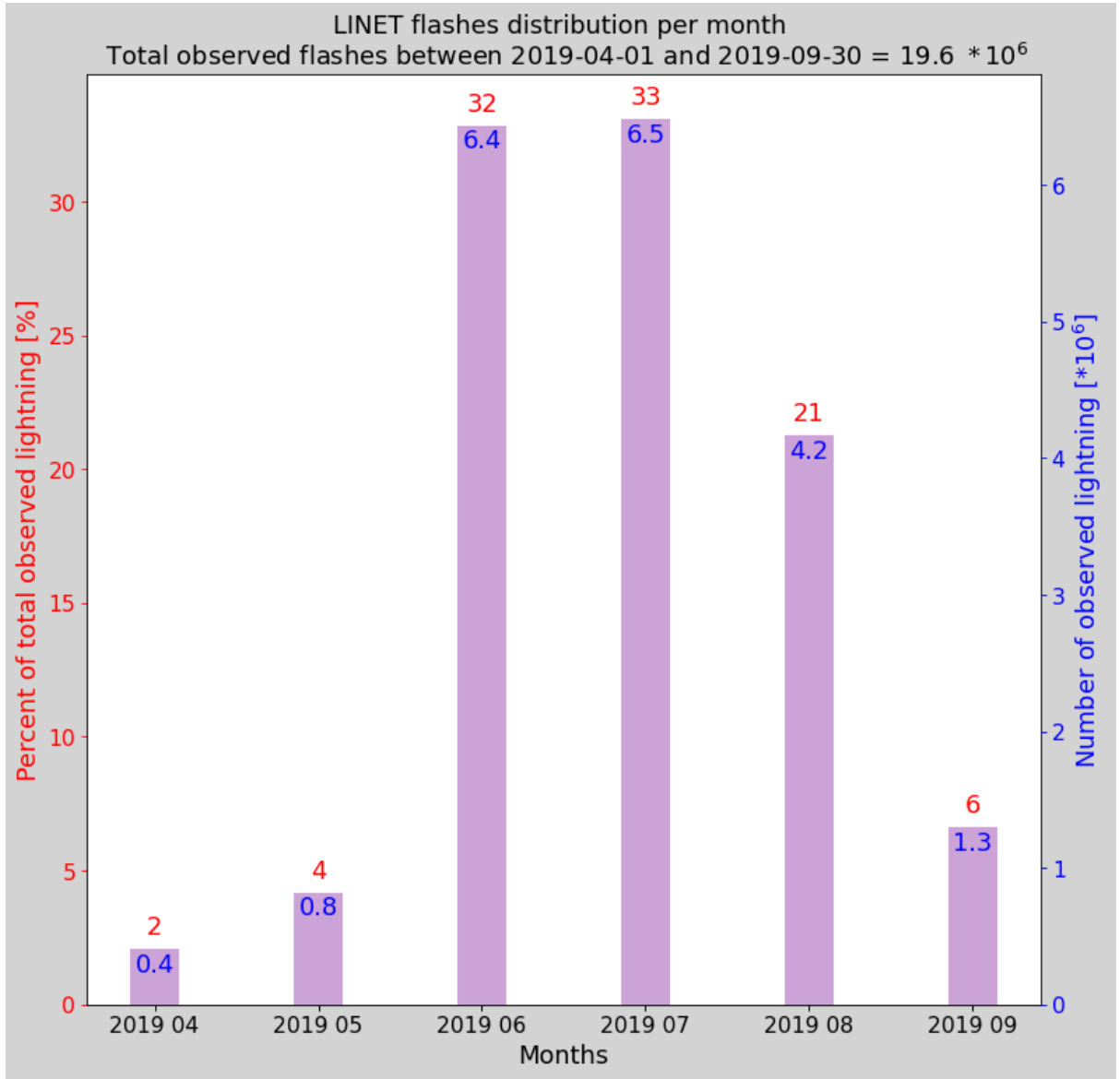
- **eSAL** – Radanovics, S., Vidal J., and Sauquet E., 2018: **Spatial Verification of Ensemble Precipitation: An Ensemble Version of SAL.** *Wea. Forecasting*, 33, 1001–1020.

Backup slides

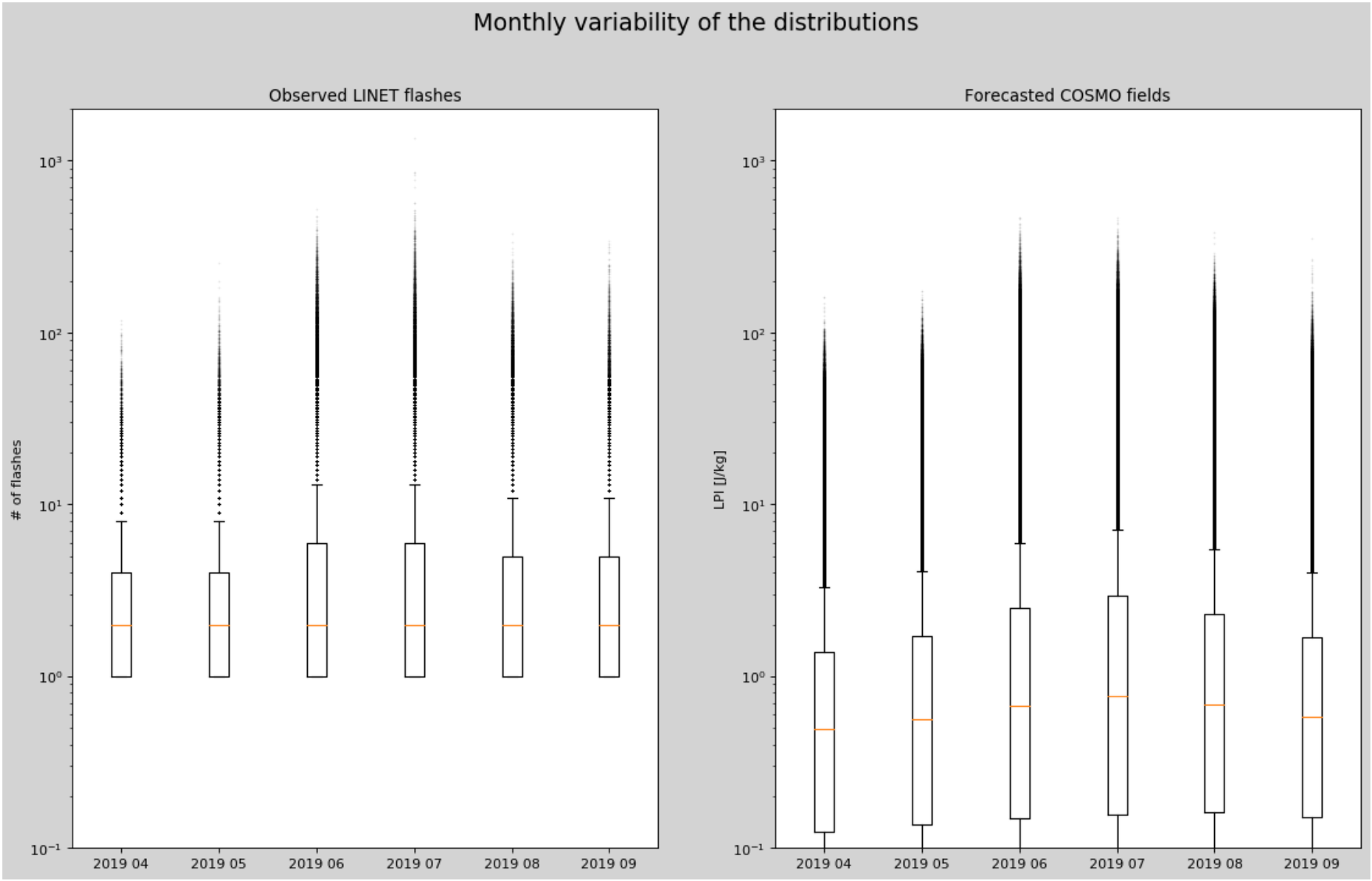
Further data description

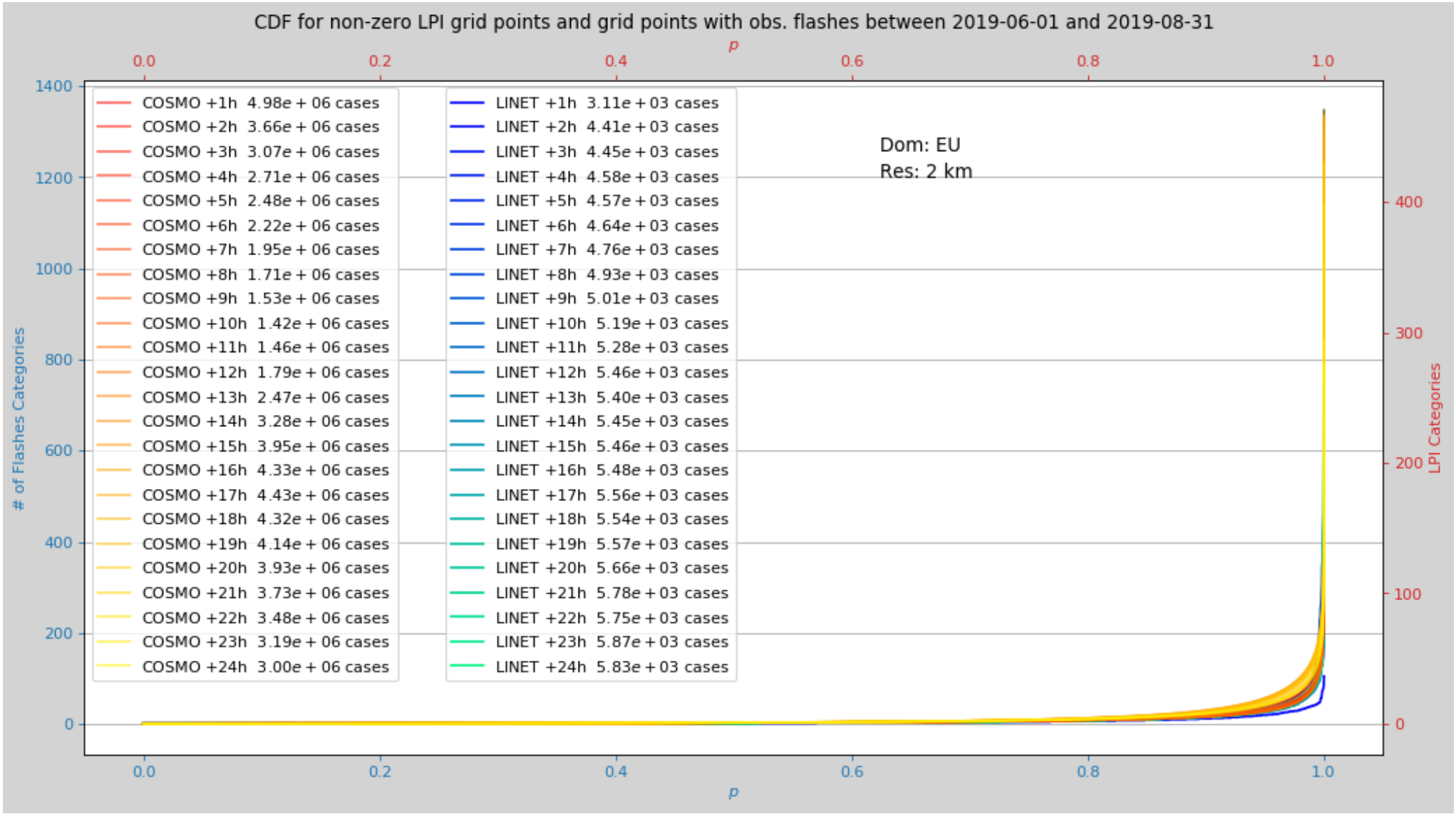

Reshaping and reorganizing data in time
COSMO-sorted
COSMO
LINET-sorted
LINET
All COSMO FC-steps between
00:15:00 and 01:00:00
added with max() function
to "STEP +1h"
All LINET flashes between
00:00:00 and 00:59:59
added to "STEP +1h"
00:00
00:15
00:30
00:45
01:00
01:15
Time [hh:mm]


Hourly distribution of space-time averaged COSMO ENS-Mean and LINET flashes (relative to maximum)
2019-06-01 <-> 2019-08-31
Main timesteps
LINET
COSMO LPI
COSMO LPI $\pm\frac{\sigma}{2}$
Average (relative to maximum of distribution)
Average (relative to maximum of distribution)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1h
2h
3h
4h
5h
6h
7h
8h
9h
10h
11h
12h
13h
14h
15h
16h
17h
18h
19h
20h
21h
22h
23h
24h
Hour of the day

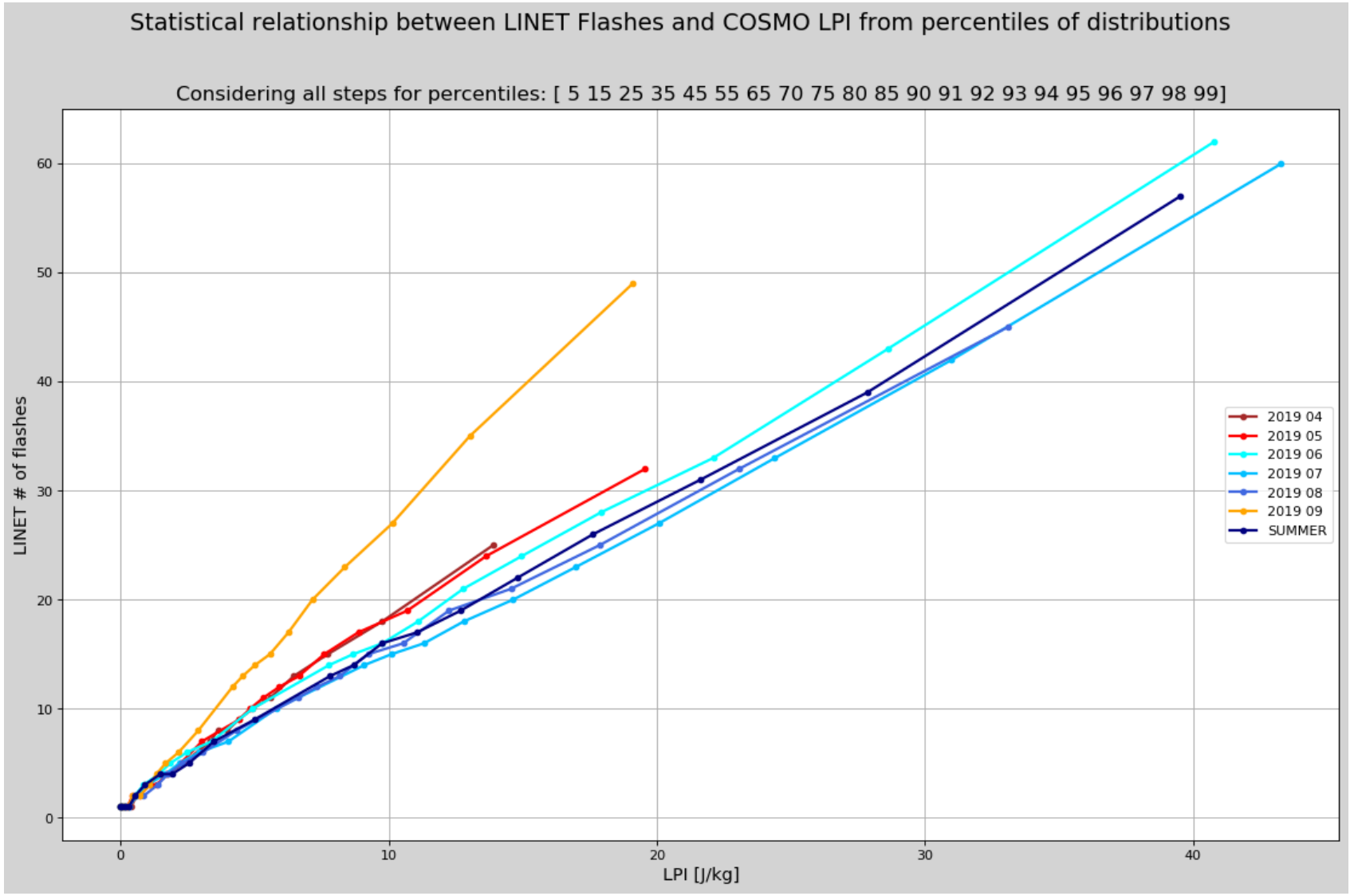

LINET flashes distribution per month
Total observed flashes between 2019-04-01 and 2019-09-30 = 19.6 *10^6
Percent of total observed lightning [%]
Number of observed lightning [*10^6]
Months
2019 04
2019 05
2019 06
2019 07
2019 08
2019 09
2
0.4
4
0.8
32
6.4
33
6.5
21
4.2
6
1.3

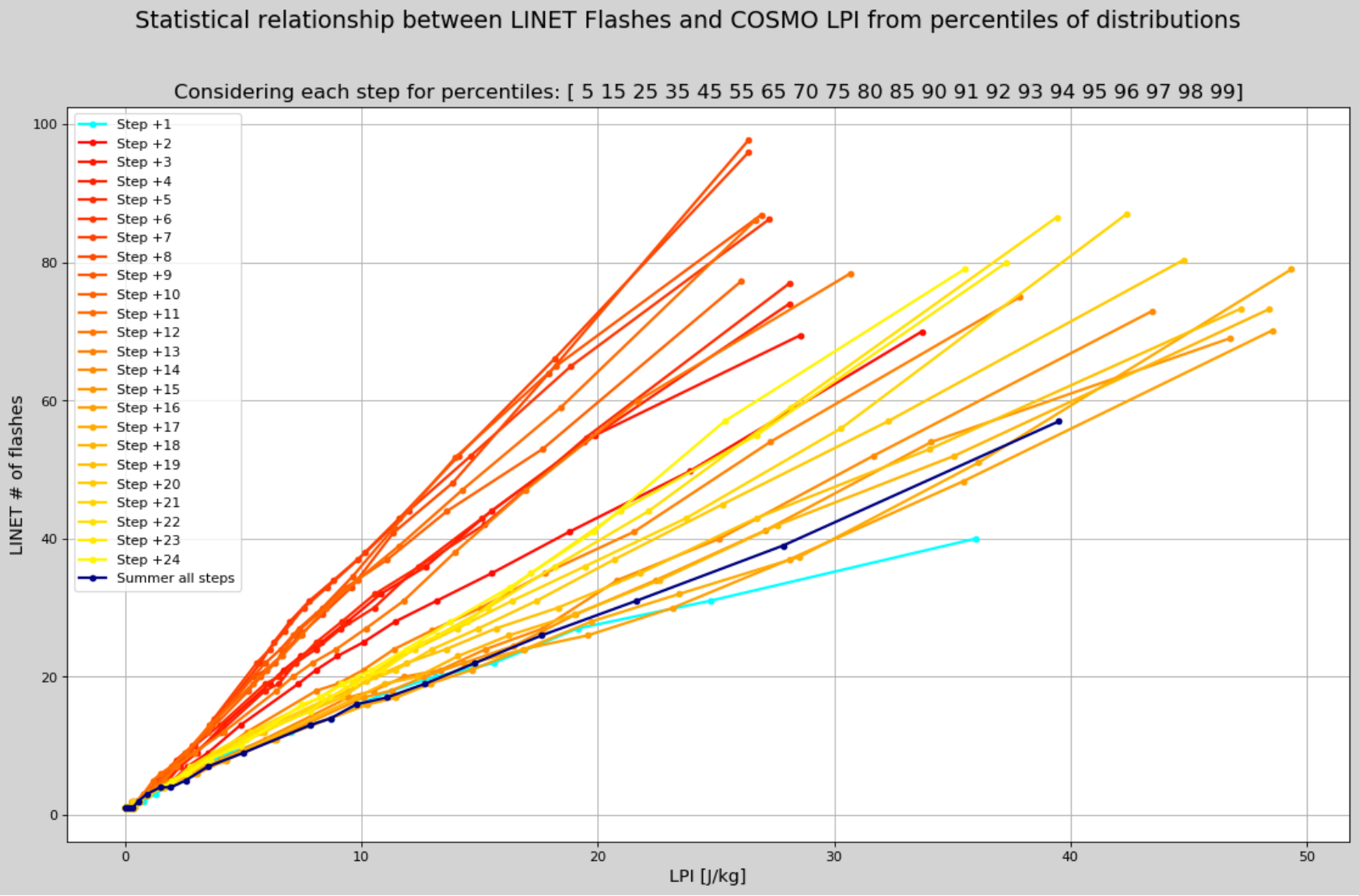

Percentiles of distributions from 2019-06-01 to 2019-08-31
LINET [# of flashes]
COSMO LPI [J/kg]
50th Percentile
80th Percentile
90th Percentile
95th Percentile
98th Percentile
99th Percentile
57.0
39.0
22.0
13.0
7.0
2.0
39.5
27.8
14.8
7.8
3.5
0.7

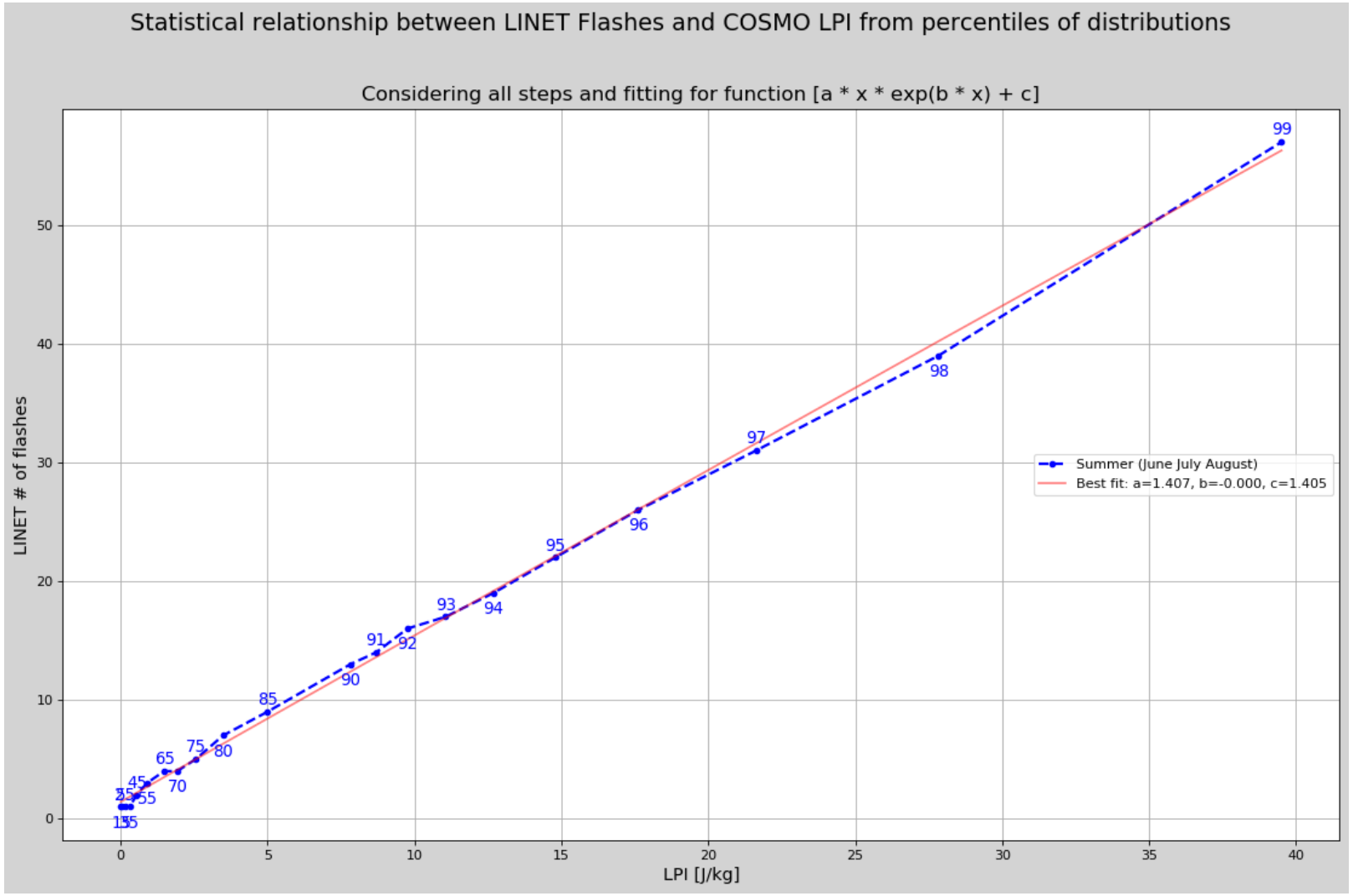

Monthly variability of the distributions
Observed LINET flashes
Forecasted COSMO fields
# of flashes
LPI [J/kg]
10^3
10^2
10^1
10^0
10^{-1}
2019 04
2019 05
2019 06
2019 07
2019 08
2019 09

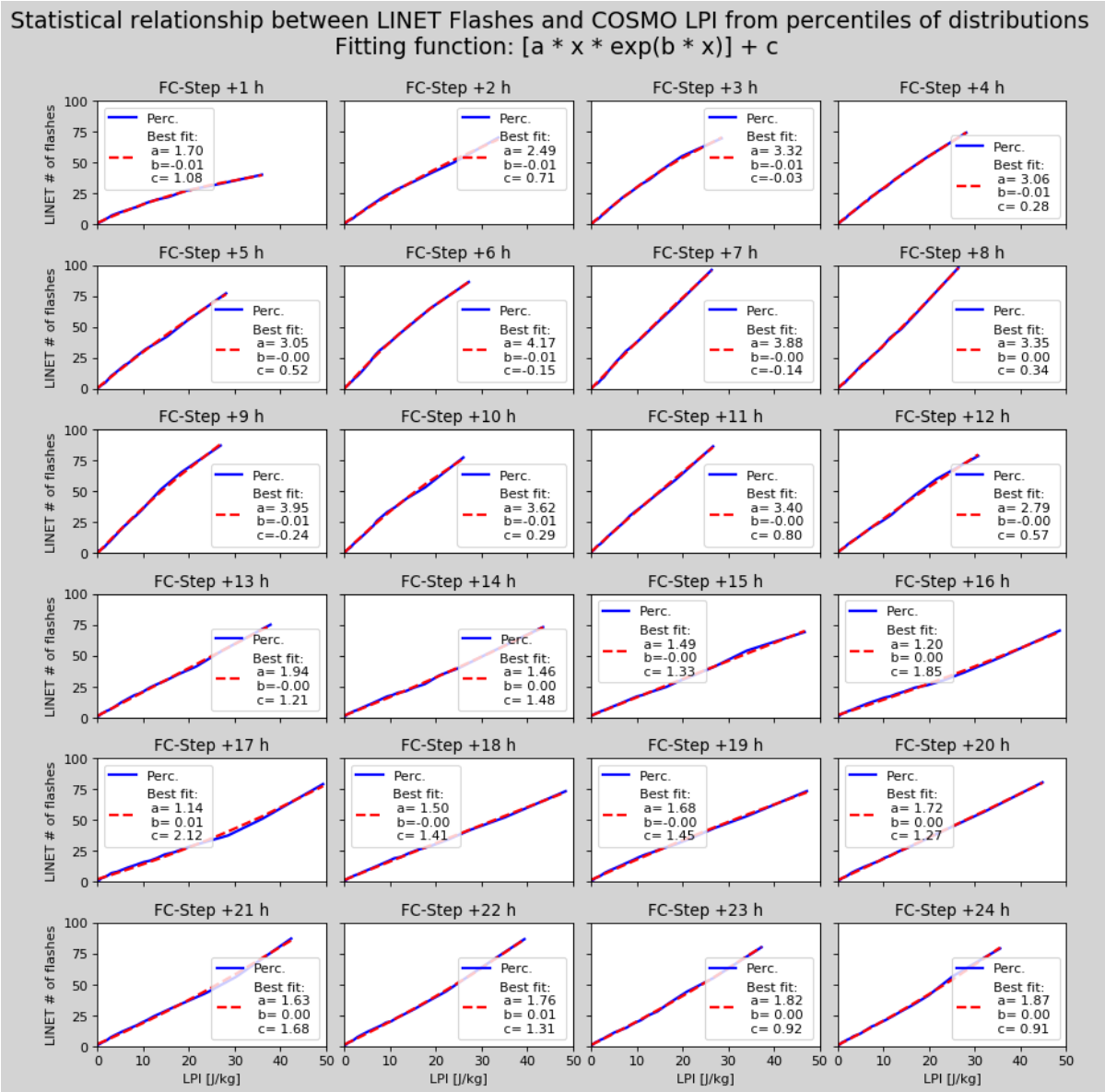

CDF for non-zero LPI grid points and grid points with obs. flashes between 2019-06-01 and 2019-08-31
p
# of Flashes Categories
LPI Categories
Dom: EU
Res: 2 km
COSMO +1h 4.98e+06 cases
COSMO +2h 3.66e+06 cases
COSMO +3h 3.07e+06 cases
COSMO +4h 2.71e+06 cases
COSMO +5h 2.48e+06 cases
COSMO +6h 2.22e+06 cases
COSMO +7h 1.95e+06 cases
COSMO +8h 1.71e+06 cases
COSMO +9h 1.53e+06 cases
COSMO +10h 1.42e+06 cases
COSMO +11h 1.46e+06 cases
COSMO +12h 1.79e+06 cases
COSMO +13h 2.47e+06 cases
COSMO +14h 3.28e+06 cases
COSMO +15h 3.95e+06 cases
COSMO +16h 4.33e+06 cases
COSMO +17h 4.43e+06 cases
COSMO +18h 4.32e+06 cases
COSMO +19h 4.14e+06 cases
COSMO +20h 3.93e+06 cases
COSMO +21h 3.73e+06 cases
COSMO +22h 3.48e+06 cases
COSMO +23h 3.19e+06 cases
COSMO +24h 3.00e+06 cases
LINET +1h 3.11e+03 cases
LINET +2h 4.41e+03 cases
LINET +3h 4.45e+03 cases
LINET +4h 4.58e+03 cases
LINET +5h 4.57e+03 cases
LINET +6h 4.64e+03 cases
LINET +7h 4.76e+03 cases
LINET +8h 4.93e+03 cases
LINET +9h 5.01e+03 cases
LINET +10h 5.19e+03 cases
LINET +11h 5.28e+03 cases
LINET +12h 5.46e+03 cases
LINET +13h 5.40e+03 cases
LINET +14h 5.45e+03 cases
LINET +15h 5.46e+03 cases
LINET +16h 5.48e+03 cases
LINET +17h 5.56e+03 cases
LINET +18h 5.54e+03 cases
LINET +19h 5.57e+03 cases
LINET +20h 5.66e+03 cases
LINET +21h 5.78e+03 cases
LINET +22h 5.75e+03 cases
LINET +23h 5.87e+03 cases
LINET +24h 5.83e+03 cases

Fitting LPI values to observed flashes (for the SAL calculation)

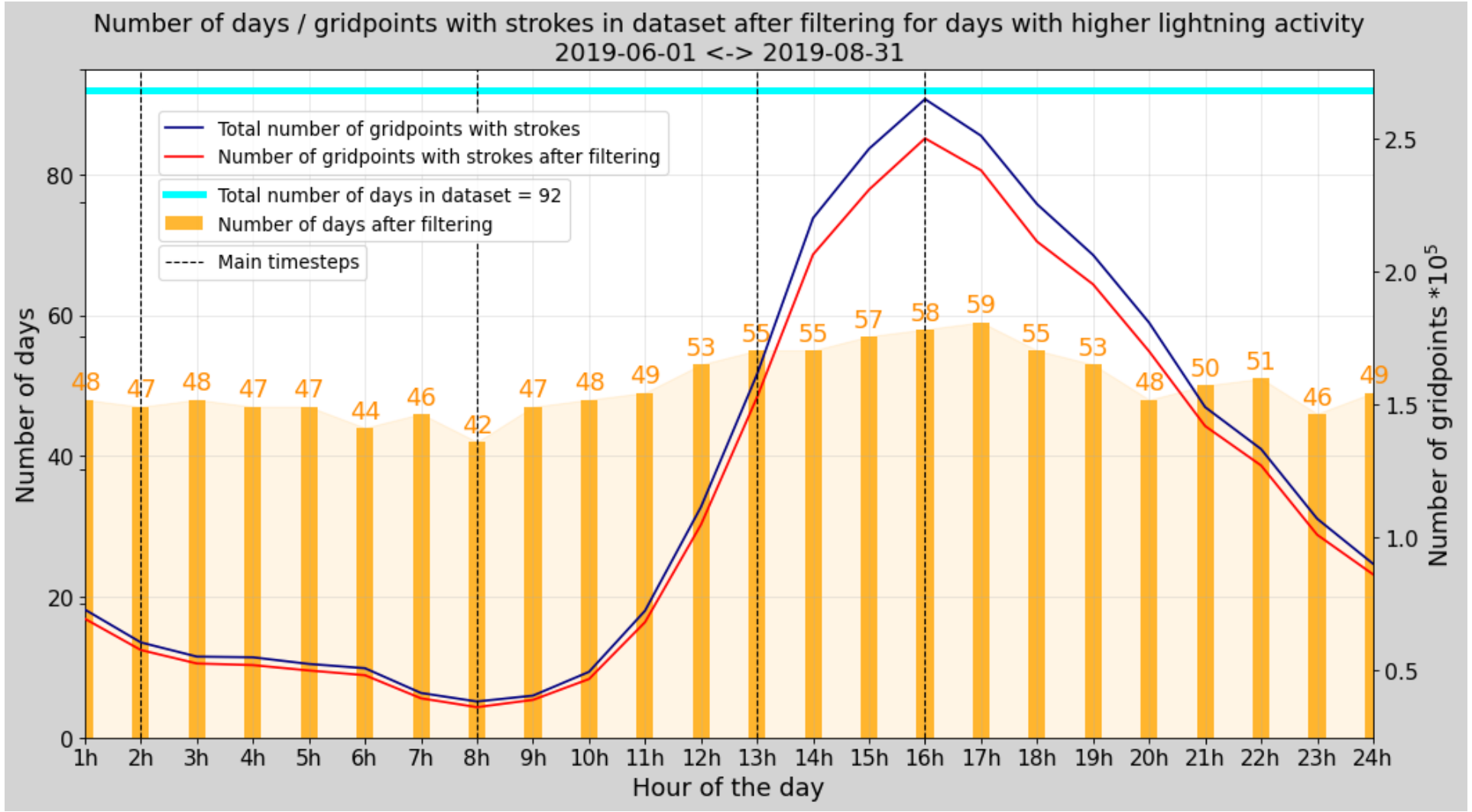

Statistical relationship between LINET Flashes and COSMO LPI from percentiles of distributions
Considering all steps for percentiles: [5 15 25 35 45 55 65 70 75 80 85 90 91 92 93 94 95 96 97 98 99]
LINET # of flashes
LPI [J/kg]
0
10
20
30
40
50
60
2019 04
2019 05
2019 06
2019 07
2019 08
2019 09
SUMMER


Statistical relationship between LINET Flashes and COSMO LPI from percentiles of distributions
Considering each step for percentiles: [5 15 25 35 45 55 65 70 75 80 85 90 91 92 93 94 95 96 97 98 99]
Step +1
Step +2
Step +3
Step +4
Step +5
Step +6
Step +7
Step +8
Step +9
Step +10
Step +11
Step +12
Step +13
Step +14
Step +15
Step +16
Step +17
Step +18
Step +19
Step +20
Step +21
Step +22
Step +23
Step +24
Summer all steps
LINET # of flashes
LPI [J/kg]
0
20
40
60
80
100
0
10
20
30
40
50

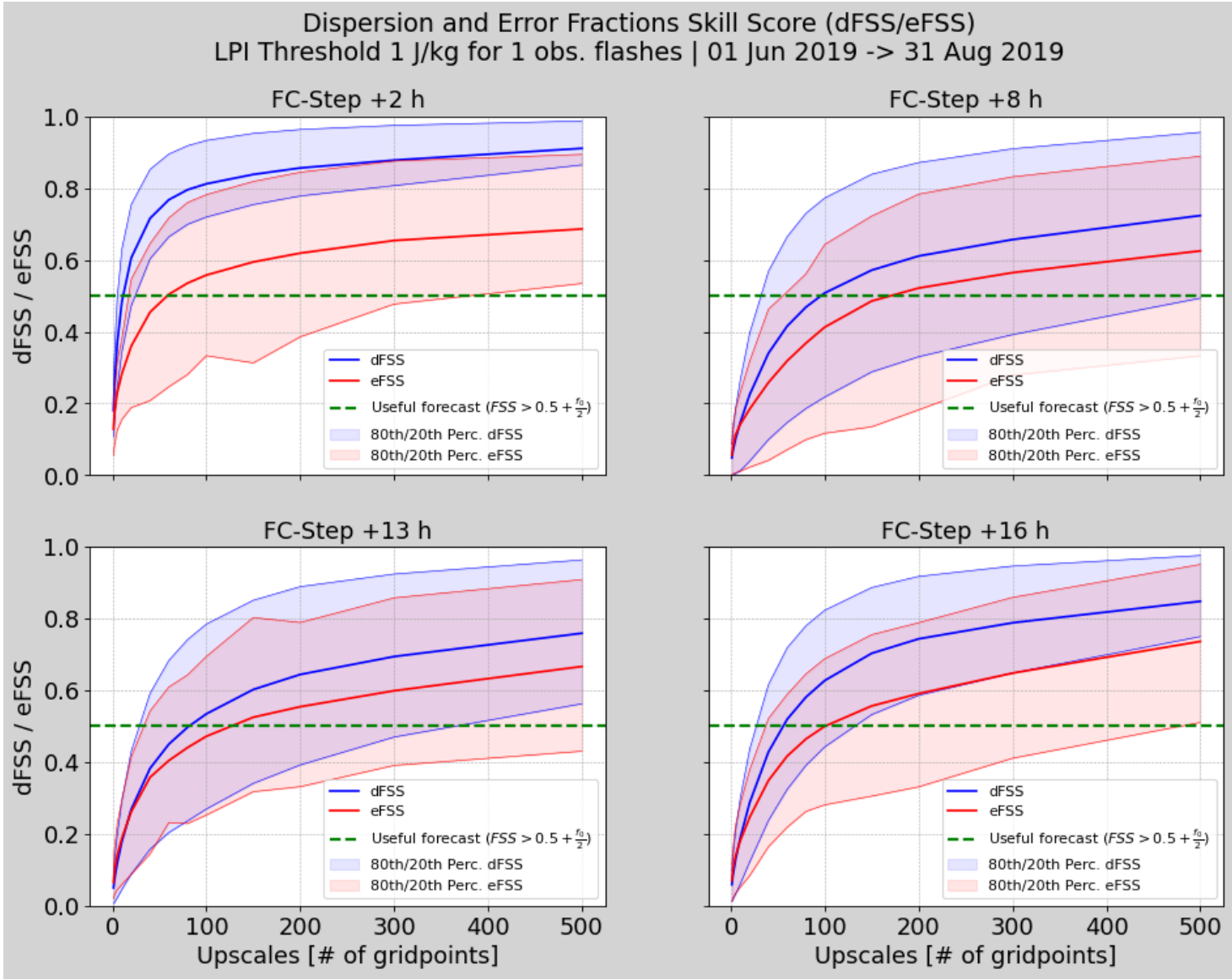

Statistical relationship between LINET Flashes and COSMO LPI from percentiles of distributions
Considering all steps and fitting for function [a * x * exp(b * x) + c]
LINET # of flashes
LPI [J/kg]
Summer (June July August)
Best fit: a=1.407, b=-0.000, c=1.405
0
5
10
15
20
25
30
35
40
50
55
65
70
75
80
85
90
91
92
93
94
95
96
97
98
99


Statistical relationship between LINET Flashes and COSMO LPI from percentiles of distributions
Fitting function: [a * x * exp(b * x)] + c
FC-Step +1 h
Perc.
Best fit:
a= 1.70
b=-0.01
c= 1.08
FC-Step +2 h
Perc.
Best fit:
a= 2.49
b=-0.01
c= 0.71
FC-Step +3 h
Perc.
Best fit:
a= 3.32
b=-0.01
c=-0.03
FC-Step +4 h
Perc.
Best fit:
a= 3.06
b=-0.01
c= 0.28
FC-Step +5 h
Perc.
Best fit:
a= 3.05
b=-0.00
c= 0.52
FC-Step +6 h
Perc.
Best fit:
a= 4.17
b=-0.01
c=-0.15
FC-Step +7 h
Perc.
Best fit:
a= 3.88
b=-0.00
c=-0.14
FC-Step +8 h
Perc.
Best fit:
a= 3.35
b= 0.00
c= 0.34
FC-Step +9 h
Perc.
Best fit:
a= 3.95
b=-0.01
c=-0.24
FC-Step +10 h
Perc.
Best fit:
a= 3.62
b=-0.01
c= 0.29
FC-Step +11 h
Perc.
Best fit:
a= 3.40
b=-0.00
c= 0.80
FC-Step +12 h
Perc.
Best fit:
a= 2.79
b=-0.00
c= 0.57
FC-Step +13 h
Perc.
Best fit:
a= 1.94
b=-0.00
c= 1.21
FC-Step +14 h
Perc.
Best fit:
a= 1.46
b= 0.00
c= 1.48
FC-Step +15 h
Perc.
Best fit:
a= 1.49
b=-0.00
c= 1.33
FC-Step +16 h
Perc.
Best fit:
a= 1.20
b= 0.00
c= 1.85
FC-Step +17 h
Perc.
Best fit:
a= 1.14
b= 0.01
c= 2.12
FC-Step +18 h
Perc.
Best fit:
a= 1.50
b=-0.00
c= 1.41
FC-Step +19 h
Perc.
Best fit:
a= 1.68
b=-0.00
c= 1.45
FC-Step +20 h
Perc.
Best fit:
a= 1.72
b= 0.00
c= 1.27
FC-Step +21 h
Perc.
Best fit:
a= 1.63
b= 0.00
c= 1.68
FC-Step +22 h
Perc.
Best fit:
a= 1.76
b= 0.01
c= 1.31
FC-Step +23 h
Perc.
Best fit:
a= 1.82
b= 0.00
c= 0.92
FC-Step +24 h
Perc.
Best fit:
a= 1.87
b= 0.00
c= 0.91
LINET # of flashes
LPI [J/kg]

Datasets after filtering out days with low convective activity


Number of days / gridpoints with strokes in dataset after filtering for days with higher lightning activity
2019-06-01 <-> 2019-08-31
Total number of gridpoints with strokes
Number of gridpoints with strokes after filtering
Total number of days in dataset = 92
Number of days after filtering
Main timesteps
Number of days
Number of gridpoints *10^5
Hour of the day
1h 2h 3h 4h 5h 6h 7h 8h 9h 10h 11h 12h 13h 14h 15h 16h 17h 18h 19h 20h 21h 22h 23h 24h
48 47 48 47 47 44 46 42 47 48 49 53 55 55 57 58 59 55 53 48 50 51 46 49

FSS specific time steps

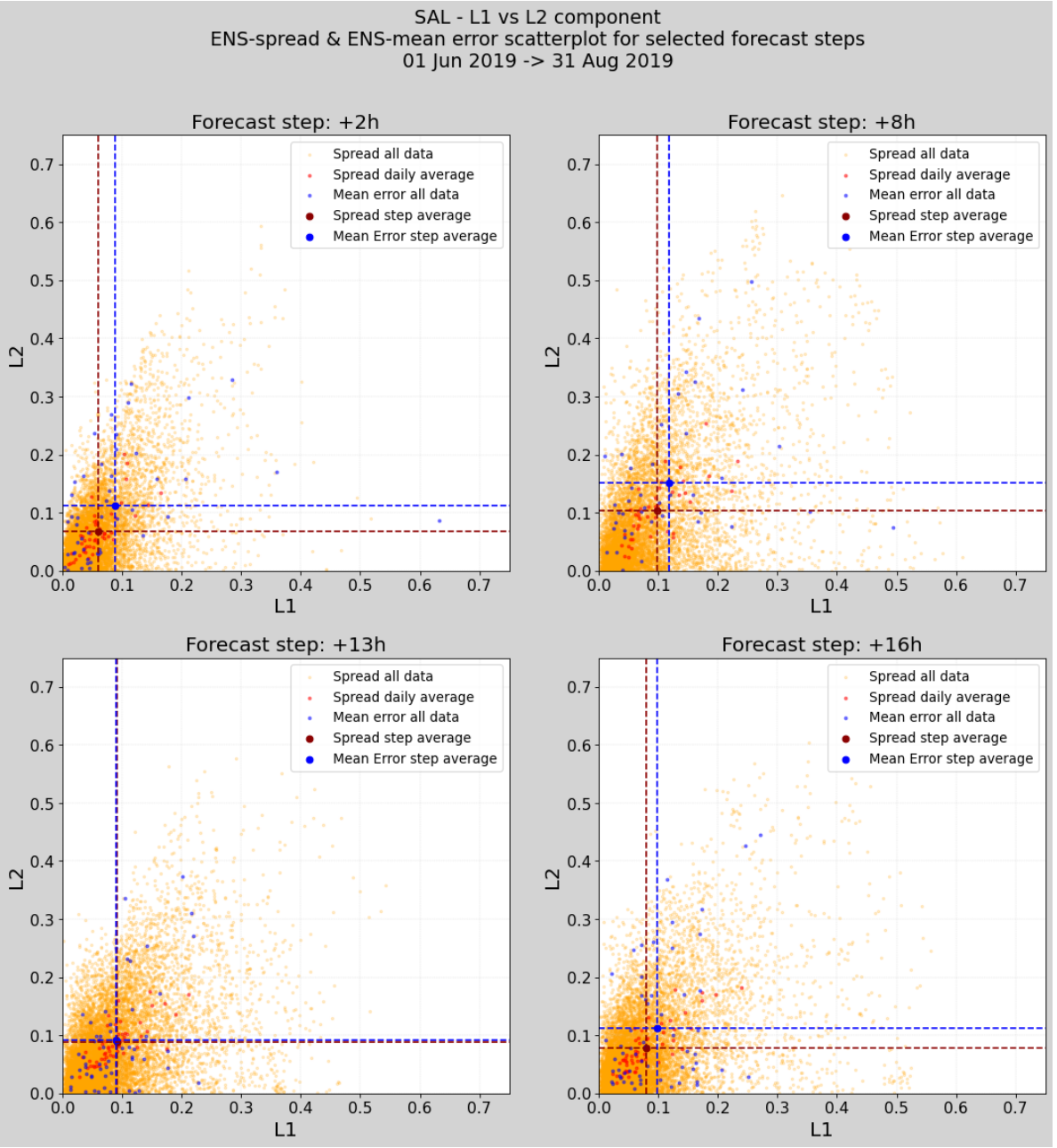

Dispersion and Error Fractions Skill Score (dFSS/eFSS)
LPI Threshold 1 J/kg for 1 obs. flashes | 01 Jun 2019 -> 31 Aug 2019
FC-Step +2 h
FC-Step +8 h
FC-Step +13 h
FC-Step +16 h
dFSS / eFSS
Upscales [# of gridpoints]
dFSS
eFSS
Useful forecast ($FSS > 0.5 + \frac{f_0}{2}$)
80th/20th Perc. dFSS
80th/20th Perc. eFSS

SAL specific time steps

- By plotting the whole dataset for both the dSAL (190 couples of members per day, orange and small red dots) and the eSAL (one value per day, small blue dots) we can better understand if the EPS is able to handle outliers as well.

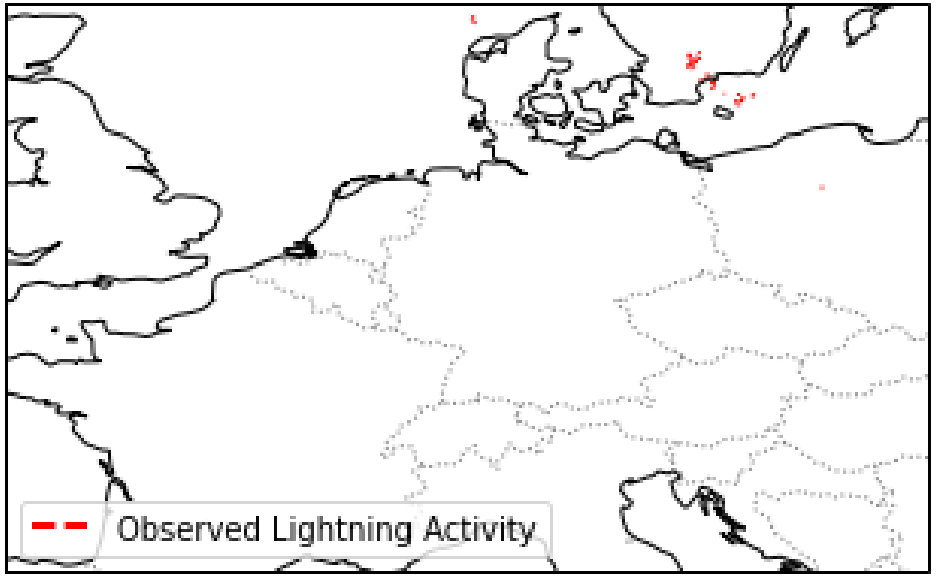
- Despite a general overestimation of the predictability, most of the error-outliers are covered by the EPS.

- For a more robust analysis, case-specific data (or on the other end a much wider pool of cases) would be ideal.



SAL “zoom on region” method

Original Field



Zoomed Region