

1 **Detection of bow echoes in kilometer-scale forecasts using a**
2 **convolutional neural network (Supplemental Material)**

3 Arnaud MOUNIER *, Laure RAYNAUD, Lucie ROTTNER, Matthieu PLU

4 *CNRM, University of Toulouse, Météo-France, CNRS, Toulouse, France*

5 Philippe ARBOGAST

6 *Forecasting department, Météo-France, Toulouse, France*

7 Michaël KREITZ

8 *National meteorological school, Météo-France, Toulouse, France*

9 Léo MIGNAN and Benoît TOUZÉ

10 *Forecasting department, Météo-France, Rennes and Toulouse, France*

11 * *Corresponding author:* Arnaud MOUNIER, arnaud.mounier@meteo.fr

12 **A1. Procedure to collect forecasters' feedback**

13 As explained in section 6.c, a group of ten forecasters has examined several meteorological
14 cases with BEs from the validation AROME-EPS dataset:

- 15 • BE cases: 8 August 2017, 26 May 2018, 4 June 2019, 5 June 2019
- 16 • squall lines: 2 January 2018
- 17 • quasi-stationary convective systems: 5 August 2019

18 and cases from the deterministic AROME model:

- 19 • BE cases: 8 August 2017, 26 May 2018, 4 June 2019, 5 June 2019
- 20 • squall lines: 2 January 2018
- 21 • quasi-stationary convective systems: 5-6 August 2019
- 22 • supercells: 15 June 2019, 26 July 2019, 8-9 August 2019, 18 August 2019

23 A single case has been removed compared with the validation dataset described in section
24 2.b.2: the squall line case of 3 December 2018. On this case, the U-Net does not detect
25 any BE and consequently this case has no impact on the HR and FAR defined in the
26 section 2.h (i.e. this case is not relevant for the evaluation). Removing this case limits
27 also the number of pictures to examine. Some cases have been added (supercell cases) in
28 the deterministic AROME model dataset compared with the validation dataset in section 5
29 because the number of pictures was too limited after the split of the still small deterministic
30 AROME dataset into three smaller datasets. Consequently, the scores of section 5 are not
31 comparable with the ones of this section and are not mentioned in the section 6.c. The
32 group of ten forecasters is also divided into three subgroups (group 1: 4 forecasters, group

33 2: 3 forecasters, and group 3: 3 forecasters). Regarding AROME-EPS dataset, the 12¹ or
34 16 members of AROME-EPS are split into three groups to assign a third of AROME-EPS
35 runs to each group:

- 36 • group 1: member 1 to 4 (or 1-5 if 16 members)
- 37 • group 2: member 5 to 8 (or 6-10 if 16 members)
- 38 • group 3: member 9 to 12 (or 11-16 if 16 members)

39 This method for separating the AROME-EPS members is basic because the AROME-
40 EPS members are interchangeable and independent. Concerning the deterministic AROME
41 dataset, the previous five initializations before the event are available and divided between
42 the three expert groups (2 initializations, 2 initializations and 1 initialization). Only one
43 initialization is assigned to one of the three groups, but this group change according to the
44 date. For instance in the evaluation form below, only one AROME initialization is available
45 for the BE case in 2018 (26 May 2018 at 0h UTC), but two initializations in 2017 (7 August
46 2017 at 0h and 3h UTC).

47 Each forecaster is asked to fill in an evaluation form. The form assigned to group 1 is
48 added below for instance. Each row corresponds to one case, one runtime, one member (or
49 the deterministic AROME model), and multiple lead times. The forecaster should write the
50 number of pictures with correct detections of BEs, the number of pictures with false alarms,
51 and the number of pictures with misses based on the pseudo-reflectivity fields and the BE
52 detections of the U-Net. The forecaster can add in the last column 'Comments' the exact
53 lead times for the correct detections, false alarms, misses or any other relevant comment. At
54 the end of the evaluation, a HR, FAR and CSI are computed for each forecaster (Table.S1).

¹The AROME-EPS is a 16-member ensemble only since July 2019

55 The HR and FAR vary greatly from one forecaster to another (FAR of '1-C' forecaster is
56 particularly high). However, the CSI is around 0.5-0.6 according to the majority of forecast-
57 ers because the HR and FAR are correlated (i.e. high HR are most of the time associated
58 with high FAR and low HR with low FAR). Some explanations concerning this behavior
59 are proposed in the section 6.c (second item). However, the 'general comments' reported at
60 the end of the evaluation form are similar with the same conclusions. These conclusions are
61 reported in the section 6.c.

Evaluation grid : Group 1 (Mb 1 to 4) - AROME-EPS

CD : Correct detections of BE, FA : False alarms, M : Misses

62

Date	Run	MB	Nb_images with CD	Nb_images with FA	Nb_images with M	Comments
07/08/2017	09h	1				
07/08/2017	09h	2				
07/08/2017	09h	3				
07/08/2017	09h	4				

07/08/2017	21h	1				
07/08/2017	21h	2				
07/08/2017	21h	3				
07/08/2017	21h	4				

25/05/2018	15h	1				
25/05/2018	15h	2				
25/05/2018	15h	3				
25/05/2018	15h	4				

25/05/2018	21h	1				
25/05/2018	21h	2				
25/05/2018	21h	3				
25/05/2018	21h	4				

26/05/2018	03h	1				
26/05/2018	03h	2				
26/05/2018	03h	3				
26/05/2018	03h	4				

26/05/2018	09h	1				
26/05/2018	09h	2				
26/05/2018	09h	3				
26/05/2018	09h	4				

03/06/2019	15h	1				
03/06/2019	15h	2				
03/06/2019	15h	3				
03/06/2019	15h	4				

03/06/2019	21h	1				
03/06/2019	21h	2				
03/06/2019	21h	3				
03/06/2019	21h	4				

04/06/2019	03h	1				
04/06/2019	03h	2				
04/06/2019	03h	3				
04/06/2019	03h	4				

Date	Run	MB	Nb_images with CD	Nb_images with FA	Nb_images with M	Comments
04/06/2019	09h	1				
04/06/2019	09h	2				
04/06/2019	09h	3				
04/06/2019	09h	4				

04/06/2019	21h	1				
04/06/2019	21h	2				
04/06/2019	21h	3				
04/06/2019	21h	4				

05/06/2019	03h	1				
05/06/2019	03h	2				
05/06/2019	03h	3				
05/06/2019	03h	4				

05/06/2019	09h	1				
05/06/2019	09h	2				
05/06/2019	09h	3				
05/06/2019	09h	4				

05/06/2019	15h	1				
05/06/2019	15h	2				
05/06/2019	15h	3				
05/06/2019	15h	4				

02/01/2018	09h	1				
02/01/2018	09h	2				
02/01/2018	09h	3				
02/01/2018	09h	4				

05/08/2019	15h	1				
05/08/2019	15h	2				
05/08/2019	15h	3				
05/08/2019	15h	4				
05/08/2019	15h	5				

Total AROME-EPS			0	0	0	
-----------------	--	--	---	---	---	--

HR	#DIV/0!
----	---------

FAR	#DIV/0!
-----	---------

CSI	#DIV/0!
-----	---------

Evaluation grid : Group 1 (Mb 1 to 4) - deterministic AROME

CD : Correct detections of BE, FA : False alarms, M : Misses

64

Date	Run	MB	Nb_images with CD	Nb_images with FA	Nb_images with M	Comments
07/08/2017	0h	ARO				
07/08/2017	3h	ARO				
26/05/2018	0h	ARO				
04/06/2019	3h	ARO				
04/06/2019	6h	ARO				
04/06/2019	18h	ARO				
05/06/2019	0h	ARO				
02/01/2018	12h	ARO				
02/01/2018	18h	ARO				
06/08/2019	6h	ARO				
06/08/2019	12h	ARO				
15/06/2019	0h	ARO				
26/07/2019	3h	ARO				
26/07/2019	6h	ARO				
08/08/2019	18h	ARO				
09/08/2019	0h	ARO				
18/08/2019	0h	ARO				
Total Arome			0	0	0	

HR	#DIV/0!
----	---------

FAR	#DIV/0!
-----	---------

CSI	#DIV/0!
-----	---------

General comments

65 **LIST OF TABLES**

66 **Table S1.** Subjective scores of ten forecasters for the AROME-EPS dataset
67 (EPS) and the deterministic AROME dataset (ARO). The first col-
68 umn is an ID for each forecaster (group number and letter from 'A'
69 to 'C' or 'D'). The HR, FAR and CSI for every dataset are reported. . . . 9

70 TABLE S1. Subjective scores of ten forecasters for the AROME-EPS dataset (EPS) and the
71 deterministic AROME dataset (ARO). The first column is an ID for each forecaster (group number
72 and letter from 'A' to 'C' or 'D'). The HR, FAR and CSI for every dataset are reported.

Forecaster	HR (EPS)	FAR (EPS)	CSI (EPS)	HR (ARO)	FAR (ARO)	CSI (ARO)
1-A	0.56	0.03	0.55	0.64	0.06	0.61
1-B	0.83	0.37	0.56	0.80	0.06	0.76
1-C	0.61	0.40	0.43	0.77	0.53	0.41
1-D	0.87	0.27	0.66	0.85	0.27	0.65
2-A	0.61	0.13	0.56	0.51	0.15	0.47
2-B	0.68	0.11	0.63	0.73	0.16	0.64
2-C	0.44	0.05	0.43	0.54	0.16	0.49
3-A	0.82	0.21	0.67	0.60	0.17	0.53
3-B	0.56	0.08	0.53	0.62	0.14	0.56
3-C	0.72	0.21	0.60	0.45	0.14	0.42

73 **LIST OF FIGURES**

74 **Fig. S1.**The entire domain of AROME and AROME-EPS 11

75 **Fig. S2.**All hand-labeled BEs corresponding to 21/09/2019. 12

76 **Fig. S3.**Example of correct detections in the validation dataset. BE contours and pseudo-
77 reflectivity fields are overlaid. Hand-labeled objects are magenta contours. U-Net
78 objects are blue contours. 13

79 **Fig. S4.**Example of false alarms (blue contours) in the validation dataset. BE contours
80 and pseudo-reflectivity fields are overlaid. 14

81 **Fig. S5.**Example of misses (magenta contours) in the validation dataset. BE contours
82 and pseudo-reflectivity fields are overlaid. 15



FIG. S1. The entire domain of AROME and AROME-EPS

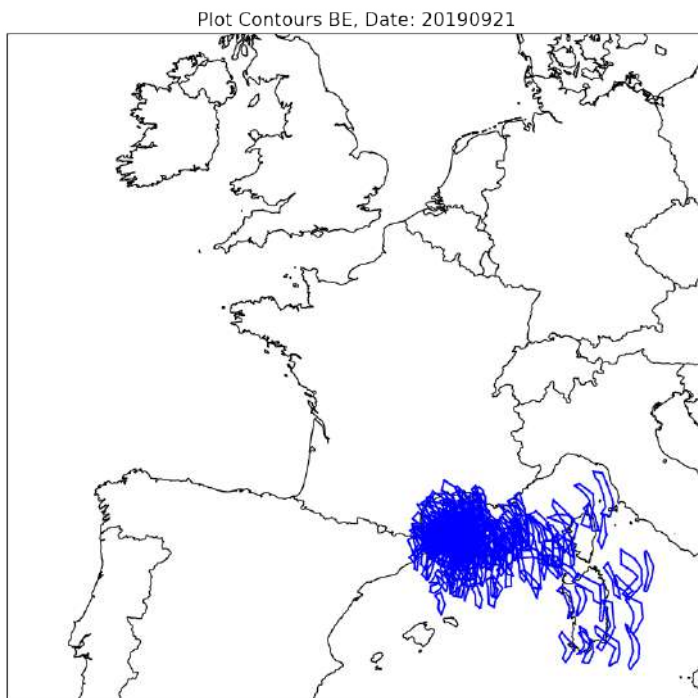
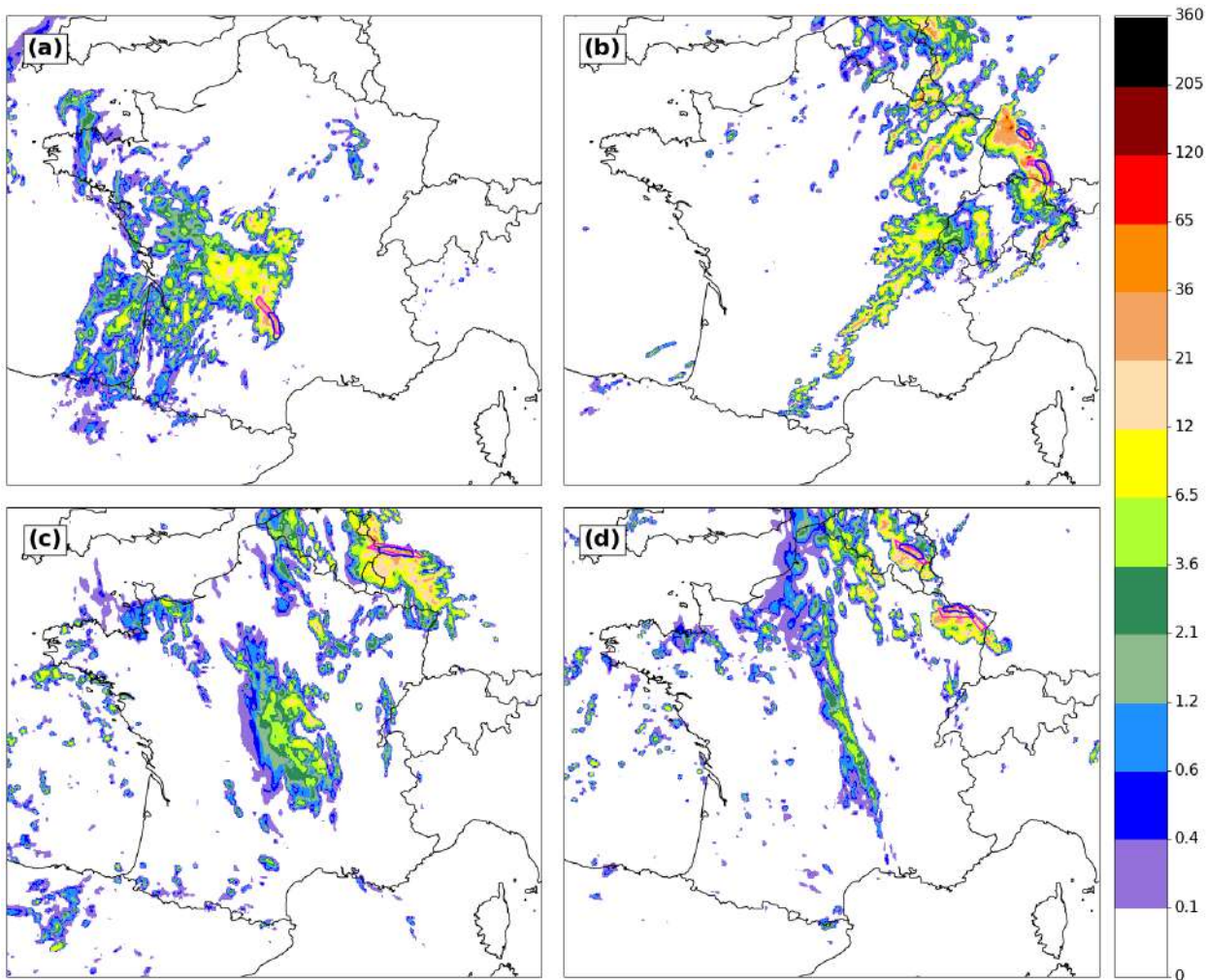
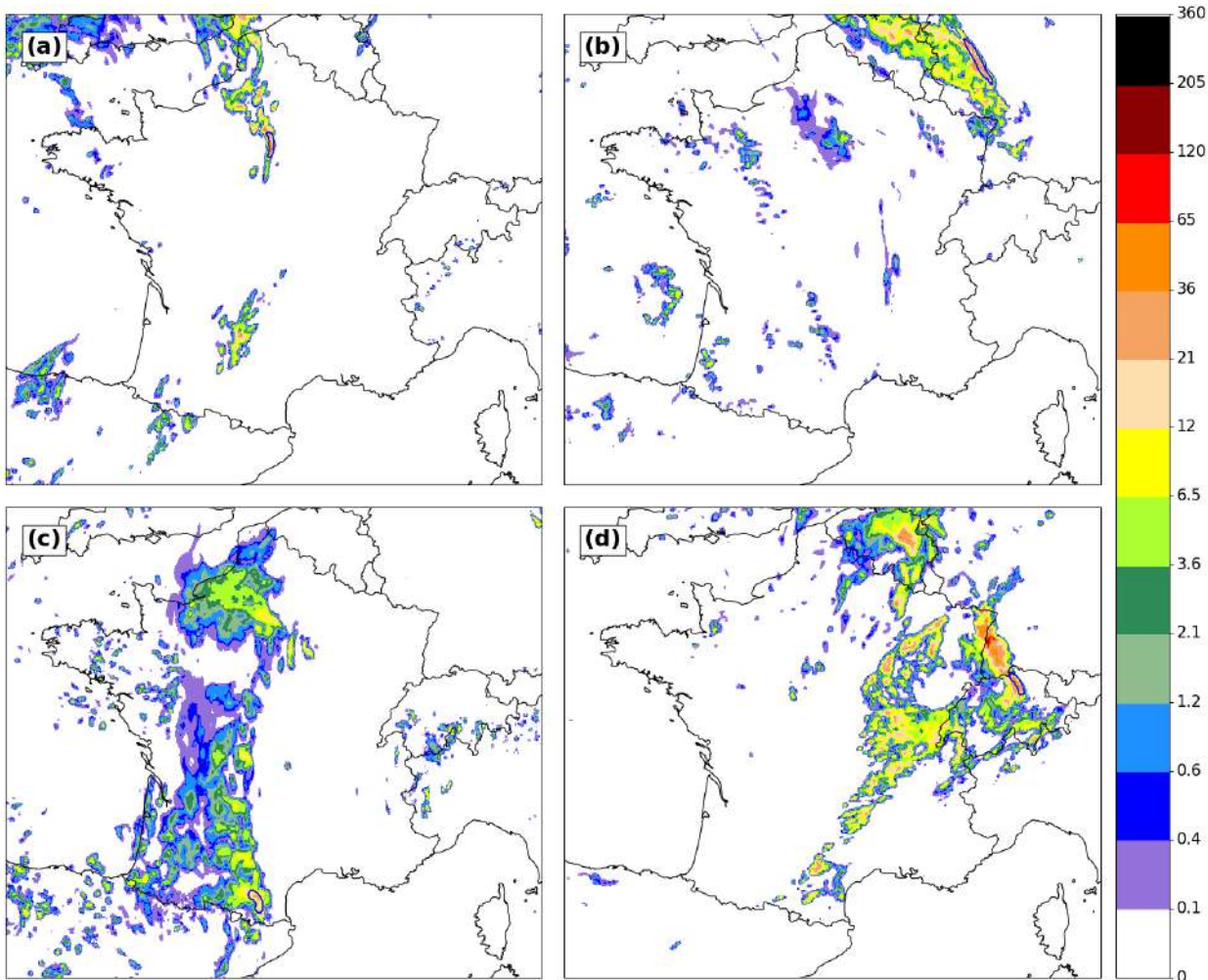


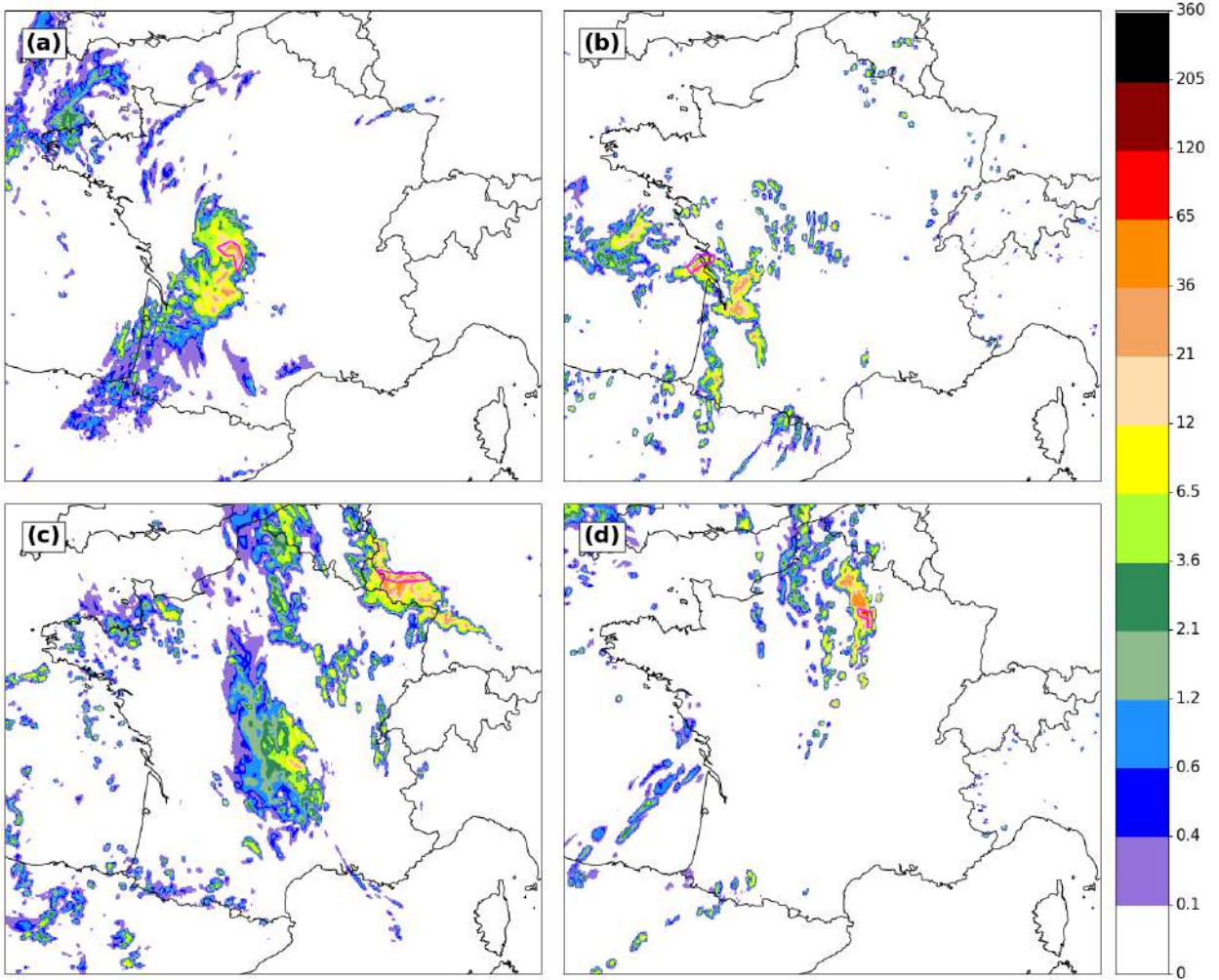
FIG. S2. All hand-labeled BEs corresponding to 21/09/2019.



83 FIG. S3. Example of correct detections in the validation dataset. BE contours and pseudo-
 84 reflectivity fields are overlaid. Hand-labeled objects are magenta contours. U-Net objects are blue
 85 contours.



86 FIG. S4. Example of false alarms (blue contours) in the validation dataset. BE contours and
 87 pseudo-reflectivity fields are overlaid.



88 FIG. S5. Example of misses (magenta contours) in the validation dataset. BE contours and
 89 pseudo-reflectivity fields are overlaid.