Tell me and I'll forget;
show me and I may remember;
involve me and I'll understand.
Chinese Proverb

Practice Session on Ensemble Verification

Reliability Diagram:

Task-A: Construct a Reliability Diagram for 3-day forecasts of positive 2m-temperature anomalies at London Heathrow using the data of the 28 cases given in Annex A.

1. Convert the "raw" data into a contingency table
   a) Determine ***n(k)***: the number of forecasts in each forecast probability category *k*, i.e. how many forecasts with probability p(*k*)=0.0, 0.1,…,1.0 with *k=1,11* exist in the sample?
   b) Determine ***nobs(k)***: the number of cases that the event occurred for the forecasts in each probability category *k*
   c) Calculate the observed relative frequency ***o(k)***: the number of observed events relative to the number of forecasts for each probability category *k:*

$$o(k) = \frac{nobs(k)}{n(k)}$$

2. Plot the observed relative frequency *o(k)* versus the probabilities *p(k)* of each category *k*.
3. Plot the sample climatology (total number of observed events / total number of forecasts):

$$\overline{o} = \frac{\sum_{k=1}^{K} nobs(k)}{\sum_{k=1}^{K} n(k)}$$

4. Plot separately the forecast frequency, i.e. number of forecasts in each forecast probability category *n(k)* versus the probability categories *p(k)*.

Question: Why do the results look so "weird"?

Task-B: Construct the Reliability Diagrams for:

- London Heathrow, 3-day forecast, MAM 01-06, pos. T2m anomalies
- 100 EU stations, 3-day forecast, MAM 01-06, pos. T2m anomalies
- 100 EU stations, 10-day forecast, MAM 01-06, pos. T2m anomalies

using the contingency tables with given *n(k)* and *nobs(k)* in Annex B1, B2, and B3.

Question: What is indicated by the differences in the diagrams?

Brier Score:

Task-C: Calculate the Brier Score (BS) and Brier Skill Score (BSS) for:

- London Heathrow, 3-day forecast, T2m anomalies > 0K
- London Heathrow, 3-day forecast, T2m anomalies > 3K

using the dataset of *N=28* cases given in Annex C.

1. For each of the *N* forecast-observation pairs compute the difference between the forecast probability p_i and observed occurrence o_i
2. Compute the Brier Score, i.e. the mean squared value of these differences:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

3. Compute the mean observed occurrence of the event (sample climatology, see also step 3 of Reliability Diagram):

$$\overline{o} = \frac{1}{N}\sum_{i=1}^{N} o_i$$

4. Compute the reference Brier Score, i.e. use the sample climatology as reference forecast:

$$BS_{ref} = \frac{1}{N}\sum_{i=1}^{N}(\overline{o} - o_i)^2$$

5. Compute the Brier Skill Score:

$$BSS = 1 - \frac{BS}{BS_{ref}}$$

Question: What is indicated by the differences in the results for BS and BSS for the two different events?

Task-D: Calculate the Brier Score (BS) for:

- 100 EU stations, 3-day forecast, MAM 01-06, pos. T2m anomalies
- 100 EU stations, 10-day forecast, MAM 01-06, pos. T2m anomalies

using the contingency tables given in Annex D1 and D2.

1. Compute the mean observed occurrence of the event (sample climatology, see also step 3 of Reliability Diagram):

   $$\bar{o} = \frac{1}{N}\sum_{i=1}^{N} o_i$$

2. Calculate the three components of the Brier Score:

   $$REL = \frac{1}{N}\sum_{k=1}^{K} n_k (p_k - o_k)^2$$

   $$RES = \frac{1}{N}\sum_{k=1}^{K} n_k (o_k - \bar{o})^2$$

   $$UNC = \bar{o}(1 - \bar{o})$$

3. Calculate the Brier Score:

   BS = REL - RES + UNC

4. Calculate the reference Brier Score (BS_{ref}).

   Tip: A climatological forecast has perfect reliability and no resolution…

5. Calculate the Brier Skill Score (BSS):

   $$BSS = 1 - \frac{BS}{BS_{ref}}$$

Questions:

- What is indicated by the differences in the values of the Brier Score and its components for the 3-day and 10-day forecast?
- How do you interpret the values for REL, RES and UNC?

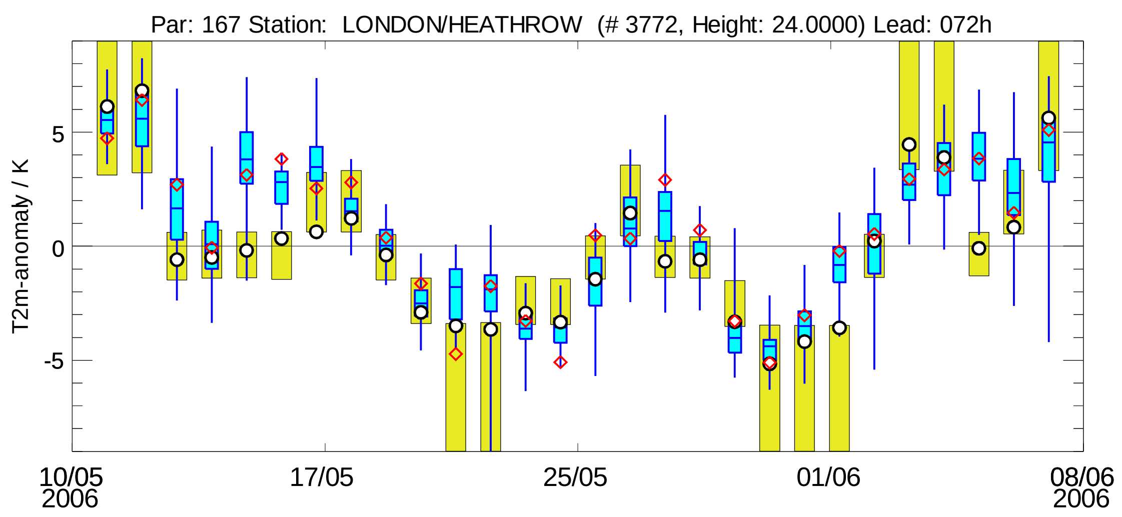
Task-E: Construct a ROC Diagram from the synthetic data in Annex E:

1. Compute the number of forecasts for each probability threshold, i.e. how many forecasts do we have in this sample with probabilities above 95%, 85%, 75%,…?
2. Do the same for the number of events observed. nobs_y(k) and not observed, nobs_n(k)
3. Calculate the Hit Rate (HR) by dividing every nobs_y(k) for each probability threshold k by the total number of nobs_y
4. Do the same for the False Alarm Rate (FR), i.e. nobs_n(k)
5. Plot HR over FR in the diagram

Annex A:

Data: 3-day forecasts for London Heathrow 2m temperatures

Date	observation	DET FC	EPS FC: p(T>0K)	EPS FC: p(T>3K)
11/05/2006	6.13	4.73	1.00	1.00
12/05/2006	6.82	6.41	1.00	0.88
13/05/2006	-0.59	2.70	0.73	0.25
14/05/2006	-0.49	-0.06	0.41	0.04
15/05/2006	-0.18	3.13	0.92	0.69
16/05/2006	0.34	3.83	1.00	0.25
17/05/2006	0.63	2.54	1.00	0.53
18/05/2006	1.22	2.80	0.98	0.04
19/05/2006	-0.38	0.38	0.39	0.00
20/05/2006	-2.90	-1.64	0.00	0.00
21/05/2006	-3.49	-4.72	0.00	0.00
22/05/2006	-3.64	-1.75	0.04	0.00
23/05/2006	-2.93	-3.27	0.00	0.00
24/05/2006	-3.33	-5.08	0.00	0.00
25/05/2006	-1.44	0.48	0.02	0.00
26/05/2006	1.45	0.34	0.69	0.08
27/05/2006	-0.66	2.91	0.73	0.08
28/05/2006	-0.59	0.71	0.25	0.00
29/05/2006	-3.32	-3.28	0.06	0.00
30/05/2006	-5.16	-5.10	0.00	0.00
31/05/2006	-4.18	-3.03	0.00	0.00
01/06/2006	-3.57	-0.21	0.20	0.00
02/06/2006	0.23	0.54	0.57	0.02
03/06/2006	4.46	2.95	0.98	0.31
04/06/2006	3.89	3.37	0.96	0.61
05/06/2006	-0.09	3.85	1.00	0.63
06/06/2006	0.83	1.46	0.84	0.31
07/06/2006	5.62	5.09	0.98	0.69

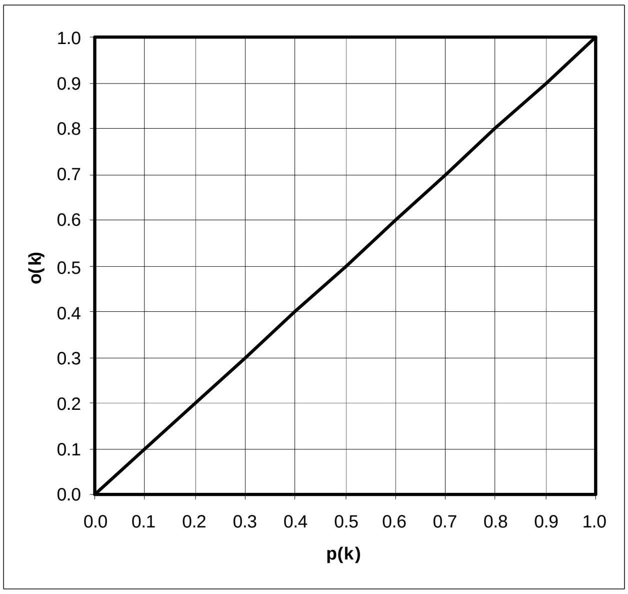


Results Task-A:

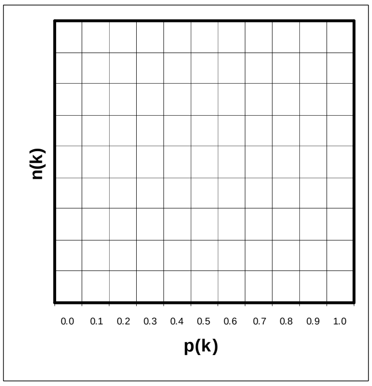
1. Contingency Table:

p(k)	probability category	n(k)	nfc_fr(k)	nobs(k)	o(k)
0.0	$0.00 \le p < 0.05$				
0.1	$0.05 \le p < 0.15$				
0.2	$0.15 \le p < 0.25$				
0.3	$0.25 \le p < 0.35$				
0.4	$0.35 \le p < 0.45$				
0.5	$0.45 \le p < 0.55$				
0.6	$0.55 \le p < 0.65$				
0.7	$0.65 \le p < 0.75$				
0.8	$0.75 \le p < 0.85$				
0.9	$0.85 \le p < 0.95$				
1.0	$0.95 \le p \le 1.00$				

2. Reliability Diagram:



4. Frequency Diagram:



3. Sample Climatology:

$\bar{o}$ =

Annex B:

Data: Contingency Tables from data over the period MMA 2001-2006 and the event of positive 2m temperature anomalies.

B1: London Heathrow, 3-day forecast:

p(k)	probability category	n(k)	nfc_fr	nobs(k)	o(k)
0.0	0.00 ≤ p < 0.05	170		15	
0.1	0.05 ≤ p < 0.15	42		11	
0.2	0.15 ≤ p < 0.25	26		6	
0.3	0.25 ≤ p < 0.35	24		7	
0.4	0.35 ≤ p < 0.45	24		15	
0.5	0.45 ≤ p < 0.55	23		15	
0.6	0.55 ≤ p < 0.65	32		25	
0.7	0.65 ≤ p < 0.75	37		26	
0.8	0.75 ≤ p < 0.85	35		32	
0.9	0.85 ≤ p < 0.95	63		53	
1.0	0.95 ≤ p ≤ 1.00	76		75	
		Σn		Σnobs	ō

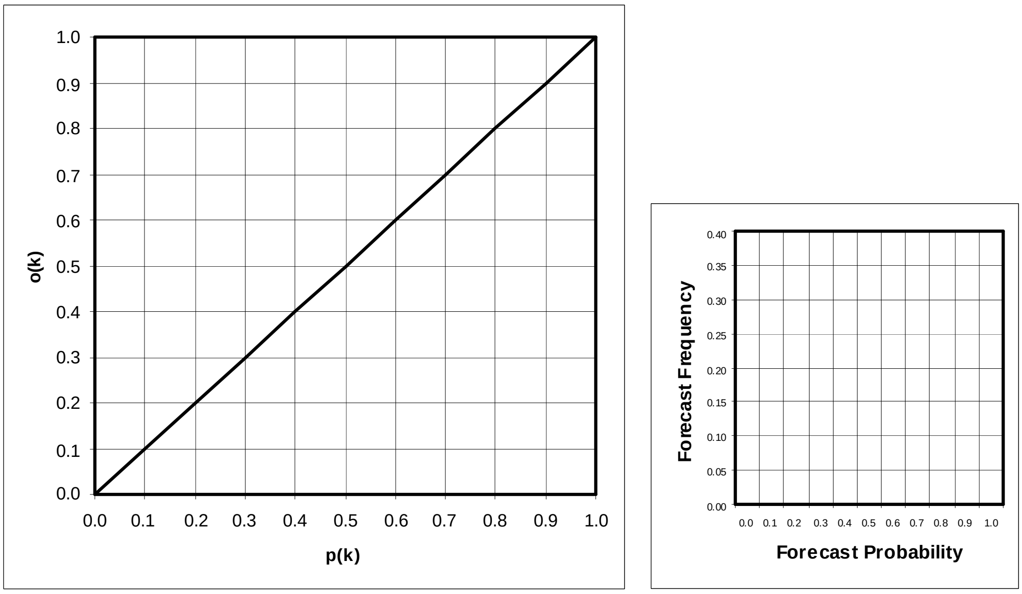
B2: 100 EU stations, 3-day forecast:

p(k)	probability category	n(k)	nfc_fr	nobs(k)	o(k)
0.0	0.00 ≤ p < 0.05	15541		1471	
0.1	0.05 ≤ p < 0.15	3041		662	
0.2	0.15 ≤ p < 0.25	2443		707	
0.3	0.25 ≤ p < 0.35	2233		818	
0.4	0.35 ≤ p < 0.45	2245		1026	
0.5	0.45 ≤ p < 0.55	2340		1268	
0.6	0.55 ≤ p < 0.65	2621		1590	
0.7	0.65 ≤ p < 0.75	2784		1890	
0.8	0.75 ≤ p < 0.85	3577		2650	
0.9	0.85 ≤ p < 0.95	7169		5939	
1.0	0.95 ≤ p ≤ 1.00	11206		10213	
		Σn		Σnobs	ō

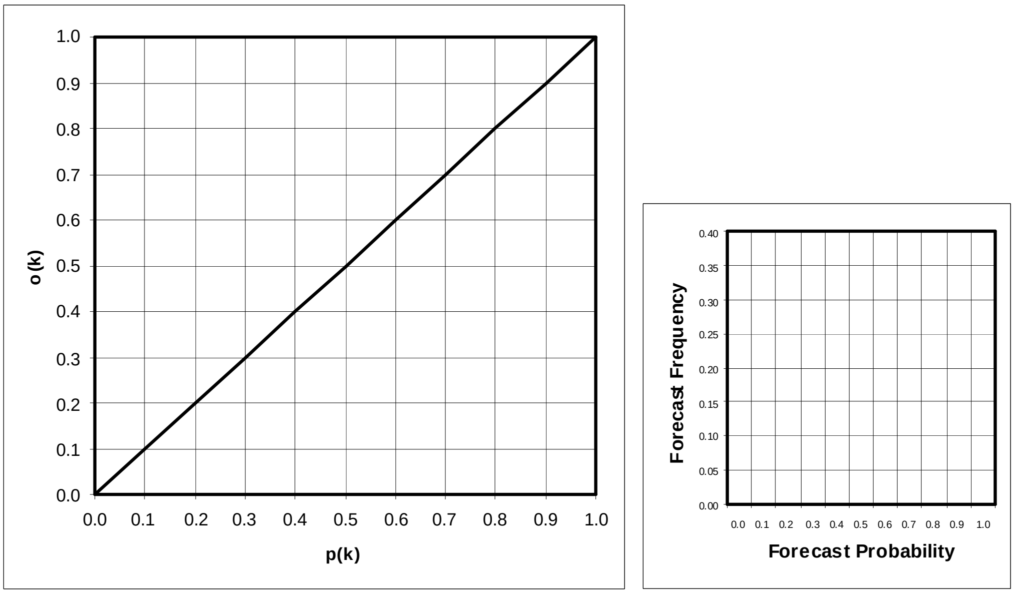
B3: 100 EU stations, 10-day forecast:

p(k)	probability category	n(k)	nfc_fr	nobs(k)	o(k)
0.0	0.00 ≤ p < 0.05	2764		701	
0.1	0.05 ≤ p < 0.15	3170		872	
0.2	0.15 ≤ p < 0.25	5043		1658	
0.3	0.25 ≤ p < 0.35	6488		2592	
0.4	0.35 ≤ p < 0.45	7306		3376	
0.5	0.45 ≤ p < 0.55	7832		4094	
0.6	0.55 ≤ p < 0.65	7506		4427	
0.7	0.65 ≤ p < 0.75	6498		4220	
0.8	0.75 ≤ p < 0.85	5022		3535	
0.9	0.85 ≤ p < 0.95	3201		2459	
1.0	0.95 ≤ p ≤ 1.00	370		300	
		Σn		Σnobs	ō

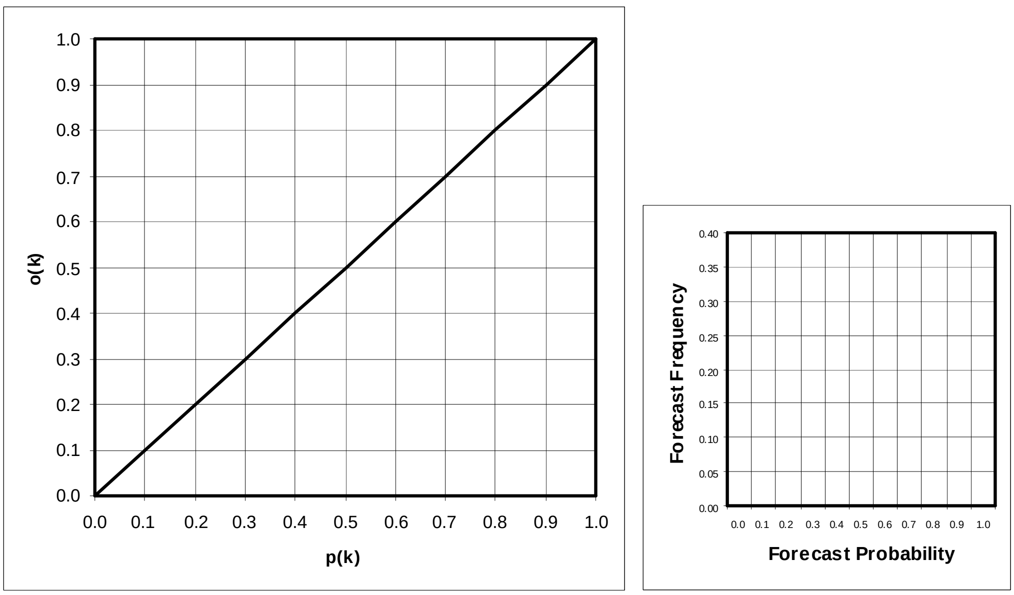
Results Task-B:



B1:



B2:



B3:

Annex C:

Data: 3-day forecasts for London Heathrow 2m temperatures

Date	obs	OBS: p(T>0K)	EPS: p(T>0)	BS	BS_REF	OBS: p(T>3K)	EPS: p(T>3K)	BS	BS_REF
11/05/2006	6.13		1.00				1.00		
12/05/2006	6.82		1.00				0.88		
13/05/2006	-0.59		0.73				0.25		
14/05/2006	-0.49		0.41				0.04		
15/05/2006	-0.18		0.92				0.69		
16/05/2006	0.34		1.00				0.25		
17/05/2006	0.63		1.00				0.53		
18/05/2006	1.22		0.98				0.04		
19/05/2006	-0.38		0.39				0.00		
20/05/2006	-2.90		0.00				0.00		
21/05/2006	-3.49		0.00				0.00		
22/05/2006	-3.64		0.04				0.00		
23/05/2006	-2.93		0.00				0.00		
24/05/2006	-3.33		0.00				0.00		
25/05/2006	-1.44		0.02				0.00		
26/05/2006	1.45		0.69				0.08		
27/05/2006	-0.66		0.73				0.08		
28/05/2006	-0.59		0.25				0.00		
29/05/2006	-3.32		0.06				0.00		
30/05/2006	-5.16		0.00				0.00		
31/05/2006	-4.18		0.00				0.00		
01/06/2006	-3.57		0.20				0.00		
02/06/2006	0.23		0.57				0.02		
03/06/2006	4.46		0.98				0.31		
04/06/2006	3.89		0.96				0.61		
05/06/2006	-0.09		1.00				0.63		
06/06/2006	0.83		0.84				0.31		
07/06/2006	5.62		0.98				0.69		
	Σo =		AVG:			Σo =		AVG:	
	ō =					ō =			

BSS=		BSS=	

Annex D:

Data: Contingency Tables from data over the period MMA 2001-2006 and the event of positive 2m temperature anomalies.

D1: 100 EU stations, 3-day forecast:

p(k)	n(k)	nobs(k)	o(k)	REL	RES
0.0	15541	1471			
0.1	3041	662			
0.2	2443	707			
0.3	2233	818			
0.4	2245	1026			
0.5	2340	1268			
0.6	2621	1590			
0.7	2784	1890			
0.8	3577	2650			
0.9	7169	5939			
1.0	11206	10213			

Σn	Σnobs	ō

REL =

RES =

UNC =

BS =

BSS =

D2: 100 EU stations, 10-day forecast:

p(k)	n(k)	nobs(k)	o(k)	REL	RES
0.0	2764	701			
0.1	3170	872			
0.2	5043	1658			
0.3	6488	2592			
0.4	7306	3376			
0.5	7832	4094			
0.6	7506	4427			
0.7	6498	4220			
0.8	5022	3535			
0.9	3201	2459			
1.0	370	300			

Σn	Σnobs	ō

REL =

RES =

UNC =

BS =

BSS =

Annex E:

Data: Synthetic Contingency Table to construct ROC Diagram.

p(k)	probability category	n_fc	rf_fc	nobs_y	fobs_y	nobs_n	fobs_n
0.0	0.00 ≤ p < 0.05	50	0.082	0	0	50	1
0.1	0.05 ≤ p < 0.15	100	0.164	10	0.1	90	0.9
0.2	0.15 ≤ p < 0.25	100	0.164	20	0.2	80	0.8
0.3	0.25 ≤ p < 0.35	100	0.164	30	0.3	70	0.7
0.4	0.35 ≤ p < 0.45	100	0.164	40	0.4	60	0.6
0.5	0.45 ≤ p < 0.55	100	0.164	50	0.5	50	0.5
0.6	0.55 ≤ p < 0.65	100	0.164	60	0.6	40	0.4
0.7	0.65 ≤ p < 0.75	100	0.164	70	0.7	30	0.3
0.8	0.75 ≤ p < 0.85	100	0.164	80	0.8	20	0.2
0.9	0.85 ≤ p < 0.95	100	0.164	90	0.9	10	0.1
1.0	0.95 ≤ p ≤ 1.00	50	0.082	50	1	0	0
		1000		500		500	

p-thresh	acc_n_fc	acc_nobs_y	HR		acc_nobs_n	FR
p>0						
p>5						
p>15						
p>25						
p>35						
p>45						
p>55						
p>65						
p>75						
p>85						
p>95