Work Package I
Observation quality control strategy
UERRA General Assembly 2

Linden Ashcroft
Manola Brunet, Peter Domonkos, Enric Aguilar, Javier Sigró, Joan Ramon Coll and Mercè Castellà

Centre for Climate Change, Universitat Rovira i Virgili

27 January 2015

UERRA Uncertainties in Ensembles of Regional ReAnalyses

Centre for Climate Change

Universitat Rovira i Virgili
Campus Terres de l'Ebre

It’s nice to meet you

- I am a new post-doc here at URV
- Completed a PhD at The University of Melbourne (Australia) in 2014 on recovery and homogenisation of historical Australian climate data.

Today's presentation

1. Method of data digitisation
2. Visual cross-checking (VCC) procedure
3. Automated quality control (QC) methods
4. What next?

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Digitisation process

- Eleven science students digitising 15 hours a week
- Tutorials given in person and online about digitising method
- Digitisation process is different for each data source…

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Morocco April 1968

STATIONS	ALTITUDE (m) (dé la cuvette du baromètre)	OBSERVATIONS DE 1800 TU la veille: Caractère du temps (voir légende)	Nébulosité (en octas)	VENT Direction	VENT Vitesse en nœuds	Pression barométrique à 0 degré C et au niveau de la mer en millibars et dixièmes de millibars	Température de l'air en degré C	Température du point de rosée en degré C	Hauteur de pluie tombée dans les 12 h. (de 6 à 18 h.) en millimètres	OBSERVATIONS DE 0600 TU le jour: Caractère du temps (voir légende)	Nébulosité (en octas)	VENT Direction	VENT Vitesse en nœuds	Pression barométrique à 0 degré C et au niveau de la mer en millibars et dixièmes de millibars	Température de l'air en degré C	Température du point de rosée en degré C	Hauteur de pluie tombée dans les 12 h. (de 18 à 6 h.) en millimètres	Hauteur de pluie tombée dans les 24 h. (de 6 à 6 h.) en millimètres	Minimum relevé à 6 h. du matin en degrés C	Maximum relevé à 18 h. la veille en degré C
Tanger-Aérodrome	15		3	WSW	06	1017.5	17	13			5	SSW	10	1018.1	12	12			12	18
Tétouan	10		3	WSW	12	1017.9	18	06												21
Al Hoceima	12		2	C		1015.8	17	13			0	C		1016.3	11	09			11	21
Larache	30		5	W	07	1017.7	17	14			8	SSW	06	1018.5	13	12			11	18
Oujda	478		5	NE	11	1016.6	13	08			8	WSW	10	1017.0	12	09	01	01	09	16
Kénitra	15		5	WNW	07	1017.7	17	13		II	0	C		1018.5	10	09			10	19
Taza	510		6	W	10	1017.3	13	11			6	W	02	1018.9	11	09			10	15
Rabat-Salé	76		4	NW	04	1018.0	16	14		III	4	S	02	1018.7	10	10			10	18
Fès-Saïs	579		7	SW	09	1018.0	13	11			0	S	09	1020.8	06	05	04	04	04	16
Meknès	549		7	C		1018.0	13	11	06		0	C		1020.2	07	06		06	07	15
Casablanca	58		1	NNW	06	1017.8	16	14			0	S	04	1018.6	10	09			09	20
Ifrane-Aérodrome (1)	1664		7	SW	20		06	04	03		1	C			03	01		03	00	10
Safi	45		3	NNW	10	1017.4	17	14			6	E	06	1018.3	12	12			10	19
Kasba-Tadla	507		7	N	10	1016.4	17	10	00,3		1	E	06	1020.0	09	07		00,3	09	19
Midelt (1)	1515		7	NW	03		10	04			2	WSW	06		04	03			04	13
Essaouira	8		2	N	11	1017.5	16	14			2	N	07	1018.2	14	12			14	20
Marrakech	466		6	ESE	03	1016.7	18	09			1	SSE	03	1019.4	09	09			09	19
Agadir	19		3	W	10	1015.3	19	14			5	C		1017.7	09	09			09	20
Ouarzazate (1)	1136		7	NNW	16		14	06	00,3		1	C			06	04	TR	00,3	06	20
Tarfaya	5																			
Oran	113		7	WSW	04	1016.4	14	09	00,5		6	C		1016.0	09	08		00,5	08	16
Colomb-Béchar (1)	806		4	N	06		18	03			0	NNE	04		10	05			10	18
Tindouf	431		1	N	06	1009.2	26	08												26
Fort-Trinquet	360																			
Madrid	601		2	WNW	08	1013.4	17	05			3	C		1014.1	05	02			04	21
Séville	7		4	C		1017.3	22	08			[illegible]	C		1017.6	12	11			11	22
Lajes (Açores)	54		6	SW	12	1022.0	17	11			7	SSE	07	1017.3	15	12			11	18
Funchal (Madère)	56		[illegible]	NE	08	1022.5	15	11			3	NNE	11	1020.3	14	09			XX	XX
Porto	79		2	NW	10	1017.9	15	13		II	8	NNW	06	1016.3	11	09			08	18
Lisbonne	106		0	N	12	1017.7	19	09			7	NNW	10	1018.3	11	07			10	23

Loznica (Serbia) January 2000

88

ЛОЗНИЦА ширина 44°33', дужина 19°14', висина 121м ЈАНУАР 2000

Дан	Ваздушни притисак у мб 07	14	21	ср.	Температура ваздуха у °C Екстреми мах	мин	амп	мин 5см	Терминске 07	14	21	ср.	Релативна влажност Термини 07	14	21	ср.
1	1012,2	1013,1	1013,9	1013,1	4,0	-3,2	7,2	-2,1	-1,4	2,3	-3,2	-1,4	90	76	85	84
2	1013,6	1014,7	1017,5	1015,3	1,8	-6,8	8,6	-8,5	-5,0	1,4	-0,8	-1,3	95	72	86	84
3	1018,1	1016,1	1014,0	1016,1	3,3	-2,7	6,0	-2,4	-1,5	1,7	-2,7	-1,3	89	75	93	86
4	1011,5	1010,3	1010,5	1010,8	4,6	-5,2	9,8	-9,1	-5,1	3,3	-3,1	-2,0	93	64	89	82
5	1010,6	1010,4	1011,5	1010,8	1,9	-5,1	7,0	-9,4	-4,7	1,0	-0,6	-1,2	98	83	92	91
6	1011,5	1011,6	1012,8	1012,0	3,8	-1,1	4,9	-1,9	0,4	2,9	-0,9	0,4	83	75	97	85
7	1015,0	1015,6	1016,0	1015,5	-0,9	-3,1	2,2	-2,8	-2,8	-2,2	-2,8	-2,7	95	92	93	93
8	1015,6	1014,7	1014,9	1015,1	-2,7	-4,2	1,5	-3,9	-4,0	-3,0	-3,5	-3,5	94	91	93	93
9	1014,5	1014,5	1015,1	1014,7	-2,8	-3,9	1,1	-4,0	-3,8	-3,1	-3,4	-3,4	94	89	93	92
10	1017,8	1019,2	1021,8	1019,6	-1,9	-3,9	2,0	-3,5	-3,9	-2,1	-1,9	-2,5	96	88	90	91
11	1023,1	1023,5	1024,2	1023,6	0,0	-1,9	1,9	-2,5	-1,7	-0,1	-0,5	-0,7	91	86	96	91
12	1023,0	1021,7	1020,5	1021,7	0,1	-1,0	1,1	-0,8	-0,9	-0,1	-0,7	-0,6	95	91	85	90
13	1017,5	1016,0	1014,7	1016,1	1,5	-1,7	3,2	-2,7	-1,7	1,1	-1,0	-0,7	86	88	91	88
14	1012,5	1012,2	1013,2	1012,6	1,6	-3,2	4,8	-6,0	-3,2	0,6	-2,7	-2,0	100	88	96	95
15	1014,1	1015,2	1015,4	1014,9	-0,4	-3,3	2,9	-6,4	-1,8	-0,8	-0,4	-0,9	93	81	83	86
16	1015,5	1015,8	1015,8	1015,7	0,6	-1,4	2,0	-1,9	-0,5	-0,6	-0,9	-0,7	83	87	79	83
17	1010,4	1001,1	993,7	1001,7	5,1	-4,6	9,7	-8,8	-3,9	1,5	5,1	2,0	86	51	54	64
18	991,5	986,5	995,0	991,0	9,0	2,3	6,7	2,6	4,2	3,1	2,3	3,0	74	91	68	78
19	998,6	1002,7	1007,3	1002,9	2,8	0,5	2,3	0,0	2,2	2,7	1,6	2,0	62	78	62	67
20	1011,8	1013,1	1009,4	1011,4	3,0	-0,9	3,9	-3,0	-0,9	1,6	1,3	0,8	74	53	54	60
21	999,1	994,3	1002,3	998,6	9,0	-1,0	10,0	-1,4	0,7	7,6	0,7	2,4	61	53	81	65
22	1008,9	1004,5	999,6	1004,3	1,2	-3,0	4,2	-3,5	-2,6	1,1	-1,3	-1,0	92	32	57	60
23	990,9	990,6	995,5	992,3	-1,3	-5,0	3,7	-3,0	-2,4	-3,7	-4,7	-3,9	88	94	96	93
24	998,3	1001,9	1006,8	1002,3	-4,0	-7,6	3,6	-5,5	-5,1	-5,7	-7,6	-6,5	91	77	85	84
25	1011,7	1014,0	1017,8	1014,5	-1,0	-14,0	13,0	-19,8	-13,7	-3,8	-14,0	-11,4	88	74	87	83
26	1018,5	1016,6	1017,3	1017,5	-0,3	-19,0	18,7	-25,0	-16,0	-1,3	-6,7	-7,7	86	36	63	62
27	1016,5	1015,3	1013,8	1015,2	9,7	-7,0	16,7	-7,1	-4,5	6,6	-4,5	-1,7	58	39	76	58
28	1011,3	1009,0	1010,6	1010,3	5,1	-9,7	14,8	-14,0	-9,7	3,3	-1,5	-2,4	97	52	90	80
29	1009,7	1004,8	1000,9	1005,1	9,0	-8,0	17,0	-13,8	-7,1	6,0	8,8	4,1	96	50	46	64
30	996,9	998,3	1007,4	1000,9	15,1	5,5	9,6	3,5	10,3	14,3	5,5	8,9	50	39	87	59
31	1007,7	1004,6	1011,6	1008,0	14,6	3,3	11,3	1,1	3,7	14,1	4,7	6,8	94	54	84	77
дек1	1014,0	1014,0	1014,8	1014,3	1,1	-3,9	5,0	-4,8	-3,2	0,2	-2,3	-1,9	93	81	91	88
дек2	1011,8	1010,8	1010,9	1011,2	2,3	-1,5	3,9	-2,9	-0,8	0,9	0,4	0,2	84	79	77	80
дек3	1006,3	1004,9	1007,6	1006,3	5,2	-6,0	11,1	-8,0	-4,2	3,5	-1,9	-1,1	82	55	77	71
мес	1010,6	1009,7	1011,0	1010,4	3,0	-3,9	6,8	-5,3	-2,8	1,6	-1,3	-0,9	86	71	82	80

Дан	Напон водене паре (мб) 07	14	21	ср.	Правац и брзина ветра (м/с) 07	14	21	ср.	Инсолација (h)	Облачност у десетинама 07	14	21	ср.	Падавине (мм)	Снег (см) У	Н	Појаве
1																	
2																	
3																	
4																	
5																	
6																	
7																	
8																	
9																	
10																	
11																	
12																	
13																	
14																	
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	
31																	
дек1																	
дек2																	
дек3																	
мес																	

Cuxhaven-Altenbruch (Germany) snow June 1978

Original Deutscher Wetterdienst

Stations-Nr. .....

Monat Juni 1978

Niederschlagsstation: Cuxhaven-Altenbruch Kreis: Cuxhaven Reg.-Bez. (Land): Stade/Niedersachsen

Tag	1a Schneedecke Höhe insgesamt cm	1b Schneedecke Neuschnee cm	1c E*)	2 Fallender Niederschlag ●, •, ✳, ↔, ✳↑, ●✳, △, ✳, ▲, △, ●▽, ✳▽ usw. Stärke (0—2) Zeit (Anfang und Ende genau angeben)	3 Sonstige Wettererscheinungen ≡, ⌓, ⌴, V, ⊠, ⊠, ⌶, ⌶, (⌶), (⌶), T, ⌶ usw. Stärke (0—2) Zeit (Anfang und Ende genau angeben)	4a Niederschlagshöhe Teilmessungen Zeit	4b Teilmessungen Höhe mm	4c Messungs-Tag	4c in 24 Stunden mm
1			0					2	•
2			0	▽⁰ p	(⌶) 14⁴⁵–15¹⁵			3	0.1
3			0		≡⁰			4	•
4			0		≡⁰ na, fr			5	•
5			0	▽¹ abd, ▽⁰ np	(⌶) 14¹⁰–15	22³⁰	5,5	6	6,0
6			1	▽⁰ m ▽⁰⁻¹ abd	(⌶) 20⁵⁵–21³⁰			7	4,5
7			1	▽¹ na ▽⁰ m, p	≡⁰			8	4,4
8			1	▽⁰ na				9	2,2
9			1	▽⁰⁻¹ na ▽⁰ np				10	0,5
10			1	▽⁰ a, p	⌶	17⁰⁰	0.7	11	1,2
Sum.	×		×	×	×	×	×	Summe	189
11			1	▽⁰ na ▽⁰ m	⌶			12	2,2
12			1	▽⁰ na ▽⁰ a, m	⌶			13	0,5
13			0					14	0,3
14			0					15	•

Egypt January 1939

CAIRO
Time of Sunrise 06h 52m
" " Sunset 17h 09m

MINISTRY OF PUBLI
DAILY

Month JANUARY 1939

At 14 h. on 4th

STATION	PRESENT WEATHER ww	VISIBILITY V	WIND Dir. DD	WIND Force F	PAST WEATHER W	CLOUD Amount N	CLOUD Form C	BAR. PRESS. M.S.L. 1000+ mbs. PPP	TEMP TTT	% HUMIDITY UU
BALOM	02	8	10	3	1	4	Sc	21.8	16.5	36
[illegible]UH	02	7	06	2	2	5	Cu	19.6	16.6	62
[illegible]EILA	01	8	02	4	1	4	Cu	20.4	15.9	55
[illegible]XANDRIA	02	9	04	4	2	4	Ac	19.9	19.2	51
DAMIETTA										
PORT SAÎD	01	7	02	3	1	3	Ac	19.1	18.0	57
EL 'ARISH	00	7	02	4	0	0	—	20.3	16.2	51
DAMANHUR										
MANŞURA										
TANTA	01	7	00	0	1	3	Cu	19.9	17.4	57
ZAGAZIG	02	7	04	2	1	4	Cc	19.2	17.6	60
CAIRO ALMAZA	02	8	24	5	2	5	Cu	18.2	18.8	48
CAIRO EZBEKIYA	01	8	04	1	1	3	Cu	18.6	18.0	46
CAIRO GIZA	01	8	04	2	1	3	Cu	20.0	17.2	52
HELWAN	01	8	02	5	0	2	Cu	18.2	17.1	37
SUEZ	00	9	02	3	0	0	—	17.9	17.2	45
FAIYUM	00	7	02	4	0	0	—	17.9	18.6	45
MINYA	01	8	02	2	1	1	Cu	17.2	17.2	42
ASYUT	00	8	26	3	0	0	—	16.6	19.0	38
QENA	00	8	24	3	0	0	—	16.0	21.0	47
LUXOR	00	7	30	1	0	0	—	15.2	22.5	41
ASWAN	01	7	32	2	1	3	Cu	13.1	23.7	28
SIWA	00	7	26	1	0	0	—	—	17.8	62
BAHARIYA	00	9	00	0	0	0	—	17.5	18.0	36
KHARGA	00	8	04	3	0	0	—	16.3	20.7	34
DAKHLA	00	7	02	1	1	2	Ci	17.2	20.0	55
TOR	01	6	28	3	0	2	Cc	15.5	19.5	69
HURGHADA										
QUSEIR	00	5	04	3	0	0	—	17.4	20.3	66

At 8 h on THURSDAY 5th

STATION	PRESENT WEATHER ww	VISIBILITY V	WIND Dir DD	WIND Force F	PAST WEATHER W	CLOUD Amount N	CLOUD Form C	BAR PRESSURE M.S.L. 1000+ mbs. PPP	Diff. from normal mbs.	TEMPERATURE Dry Bulb °C TTT	MAXIMUM Yesterday °C MM	MAXIMUM Diff. from norm. °C	MINIMUM Today °C mm	MINIMUM Diff. from norm. °C	% HUMIDITY UU	PRECIPITATION Past 24 hrs.	PRECIPITATION Total from
BALOM	01	6	26	4	0	2	Cu	21.4	+2.9	9.4	17	—	7	—	54	0	35
[illegible]UH	02	7	20	3	2	5	Cu	19.2	+0.6	11.2	18	0	11	+3	81	0	64
[illegible]EILA	01	8	00	0	1	4	Cu	19.6	—	14.0	17	—	7	—	63	0	45
[illegible]XANDRIA	02	8	00	0	2	5	Ac	19.0	+0.1	16.3	20	+1	12	+1	64	0	111
DAMIETTA	04	6	26	1	0	2	Cu	18.2	—	11.4	17	—	8	—	83	0	28
PORT SAÎD	01	7	02	2	1	3	Ac	17.6	+0.2	14.7	19	0	10	-1	76	0	23
EL 'ARISH	00	7	00	0	0	0	—	19.4	—	11.5	16	—	9	—	—	0	69
DAMANHUR	01	5	32	1	1	2	Cu	19.4	+0.7	8.2	19	-2	6	-2	83	0	10
MANŞURA	01	8	12	1	1	3	Sc	18.4	—	10.9	18	-3	8	0	71	0	24
TANTA	00	8	00	0	0	0	—	19.1	-0.4	8.0	20	0	3	-4	82	0	15
ZAGAZIG	00	7	30	1	1	0	—	19.0	—	9.4	18	-3	2	-4	87	0	10
CAIRO ALMAZA	05	7	08	1	1	2	Sc	18.8	—	8.2	19	—	4	—	81	0	2
CAIRO EZBEKIYA	01	6	02	1	4	2	Cu	18.7	-0.5	7.6	20	+1	6	-2	89	0	3
CAIRO GIZA	04	4	00	0	4	3	As	18.3	-0.9	6.6	18	-1	4	-2	91	0	3
HELWAN	08	3	32	2	0	1	Ac	18.4	-1.2	8.6	17	-1	6	-2	79	0	9
SUEZ	00	7	32	2	0	0	—	18.2	-0.5	9.0	18	-2	6	-4	87	0	13
FAIYUM	00	6	30	2	0	0	—	18.8	-0.6	7.5	20	-1	4	-2	65	0	dr
MINYA	01	5	32	1	0	2	Cu	18.4	-1.9	6.1	18	-2	4	-3	87	0	2
ASYUT	00	6	22	3	0	0	—	18.4	-1.8	7.1	21	+1	5	-1	84	0	2
QENA	00	8	30	1	0	0	—	17.5	-1.7	11.0	22	-1	8	+1	75	0	dr
LUXOR	00	7	06	2	0	0	—	17.1	—	8.1	24	—	4	—	98	0	dr
ASWAN	00	7	02	1	0	0	—	16.0	-3.2	11.6	24	0	9	-1	46	0	1
SIWA	00	7	00	0	0	0	—	—	—	8.0	18	-1	1	-3	85	0	dr
BAHARIYA	00	9	00	0	0	0	—	20.4	—	7.0	18	—	—	—	71	0	0
KHARGA	00	8	32	1	0	0	—	17.0	-2.6	9.2	21	-2	4	-2	58	0	dr
DAKHLA	00	5	24	1	0	0	—	18.8	-1.5	7.0	21	-1	5	0	84	0	dr
TOR	01	6	32	3	0	3	Cu	16.3	-1.9	9.5	21	0	6	-3	74	0	19
HURGHADA	01	9	30	3	1	3	Ci	16.3	-2.1	15.3	20	-2	9	-2	47	0	1
QUSEIR	00	5	24	2	0	0	—	16.7	-1.2	16.2	21	-2	14	0	49	0	tr

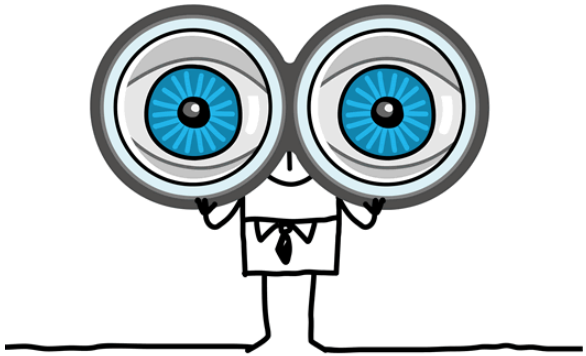
Visual cross checks

Step 1: Digitiser self-assessment

- 10th, 20th and 30th of each month checked
- Monthly means calculated and compared with source if possible
- All missing data and illegible observations recorded in metadata

Step 2: Systematic checks

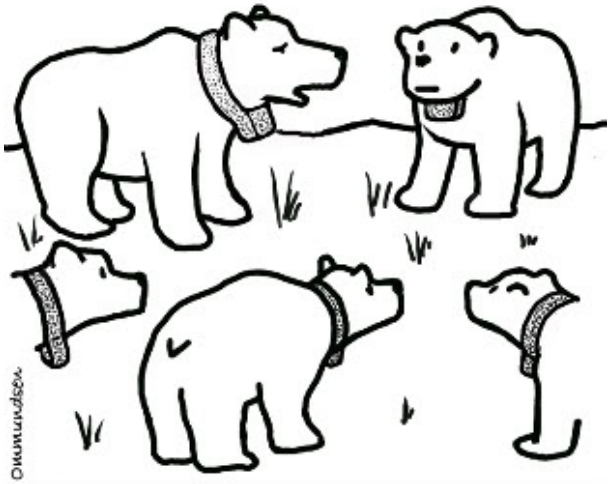
- Typical errors identified and reported back to digitisers



UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Automated quality control process

- Three main components so far:
  - Temperatures (T), relative humidity (RH), wind speed and direction (W) (P. Domonkos)
  - Atmospheric pressure (E. Aguilar)
  - Snowfall and snow depth (J. Sigró)



"The purpose of meeting here is to create a type two error."

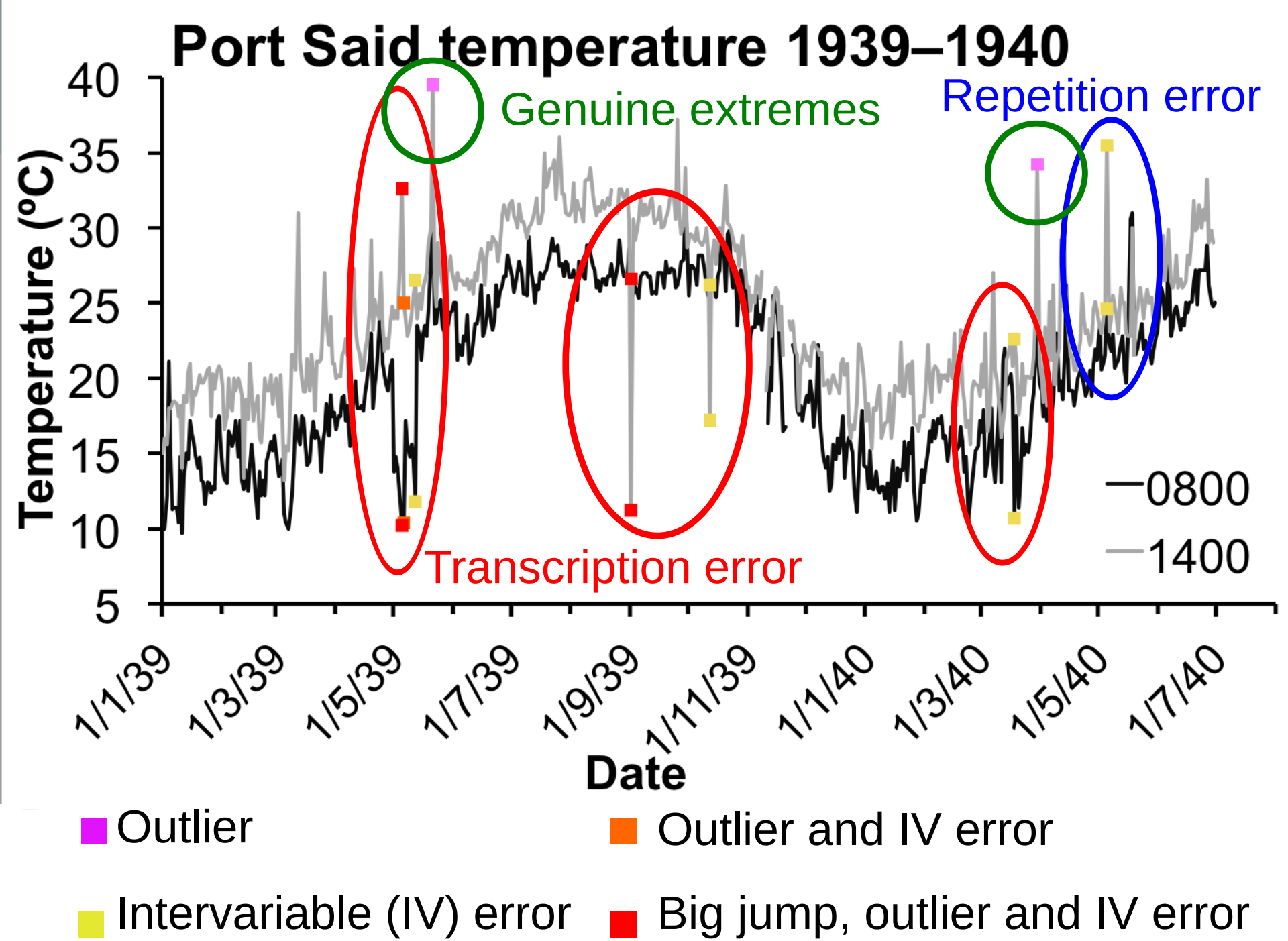
Temperature, relative humidity and wind quality control:13 tests

Gross errors

1. Date order check
2. Repetition
3. Physical impossibilities
4. Flat lines
5. Frequently occurring values
6. Monthly increments
7. Distribution errors

Relative errors

8. Outliers
9. Logical checks
10. Big jumps
11. Sudden spikes (T)
12. Intervariable test
13. Dew point calculation


Port Said temperature 1939–1940
Genuine extremes
Repetition error
Transcription error
Temperature (°C)
40
35
30
25
20
15
10
5
0800
1400
1/1/39
1/3/39
1/5/39
1/7/39
1/9/39
1/11/39
1/1/40
1/3/40
1/5/40
1/7/40
Date
Outlier
Outlier and IV error
Intervariable (IV) error
Big jump, outlier and IV error

Atmospheric pressure quality control

- Checking for similar things
  1. Duplicate days
  2. Gross outliers (incorrect units)
  3. Statistical outliers using interquartile range
  4. Bivariate normal distribution tests comparing multiple time steps
- To be added
  1. Combination of data from similar hours to improve distribution statistics
  2. Corrections for clear errors (e.g. mmHg to hPa)
  3. Multivariate checks using wind data

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Snowfall and snow depth quality control

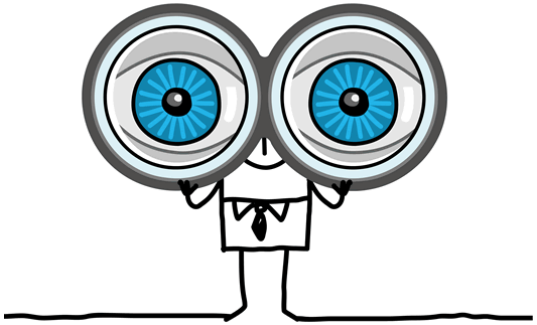
1. Changes in snow depth compared to snowfall for days before and after t=0
2. Fresh snow to precipitation ratio
3. Snow depth and snowfall data recorded during May–October

- To do
  1. Incorporate additional neighbouring station and multivariable checks, testing snowfall probability against temperature data.

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Next steps: data and digitisation

- More data to be obtained, both for digitisation and auxiliary data from MARS and ECA&D to improve QC methods
- Use of Optical Character Recognition when appropriate
- Providing more detailed templates and feedback to the digitisers



UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Next steps: automated QC

- Incorporate additional multivariate comparisons and neighbouring checks
- Improve graphical outputs to improve manual assessment
- Objective numerical assessment of source quality for future use
- Comparison on automated QC procedures with other international packages (e.g. HadISD)
- Watch this space…

UERRA
Uncertainties in Ensembles of Regional ReAnalyses

Questions and ideas?

Email: lindenclaire.ashcroft@urv.cat

Twitter: @lindenashcroft

UERRA Uncertainties in Ensembles of Regional ReAnalyses

Centre for Climate Change

Universitat Rovira i Virgili

Campus Terres de l'Ebre