Comparison study of two independent precipitation networks in Piedmont, Italy

Authors:
F. Acquaotta[1], S. Fratianni[1] and V. Venema [2]

[1]Dipartimento di Scienze della Terra, Università di Torino, Italy
[2] Meteorological Institute, University of Bonn, Germany

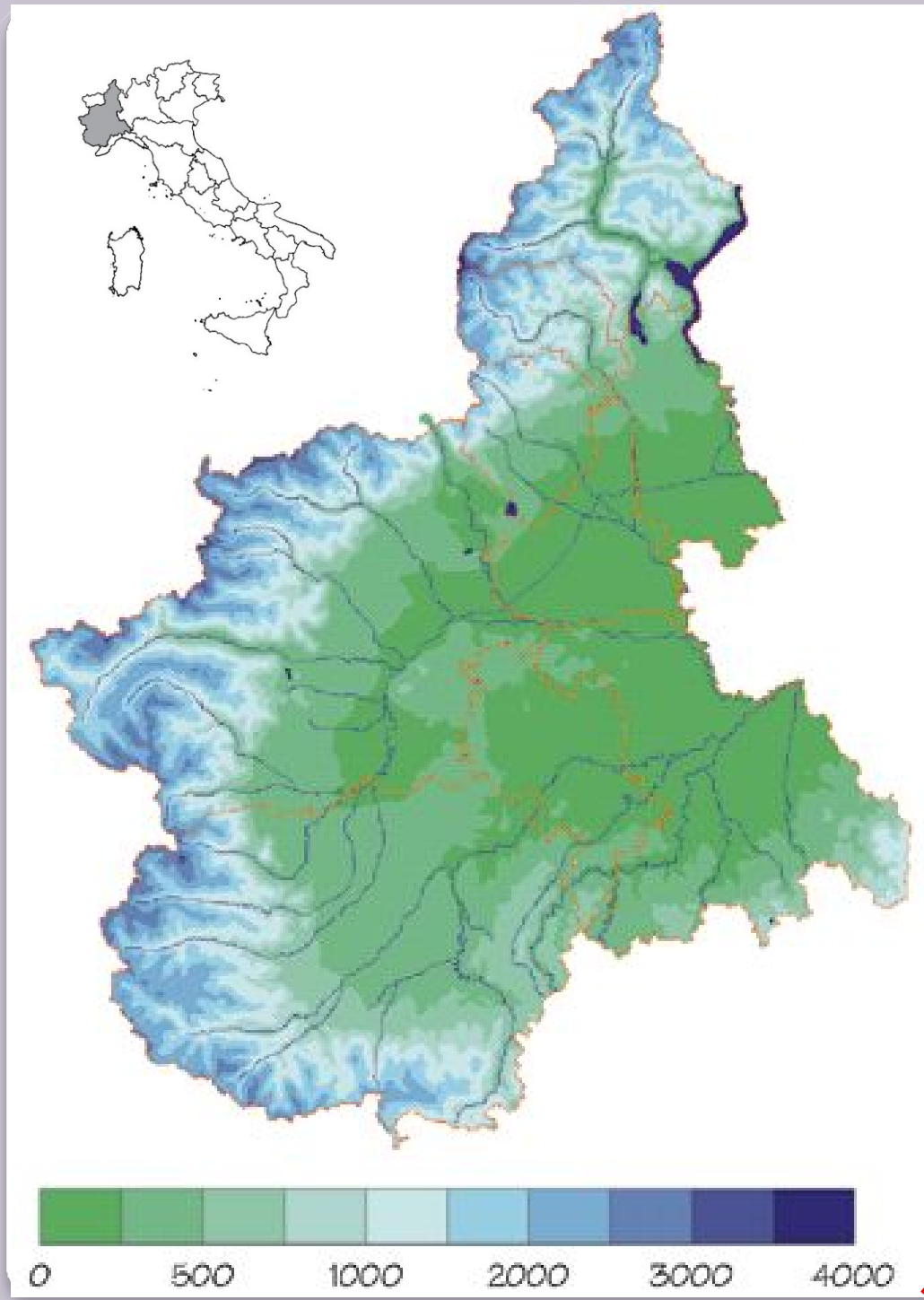
Outline

1 Area study

2 Dataset and instruments

3 Methodology

4 Results

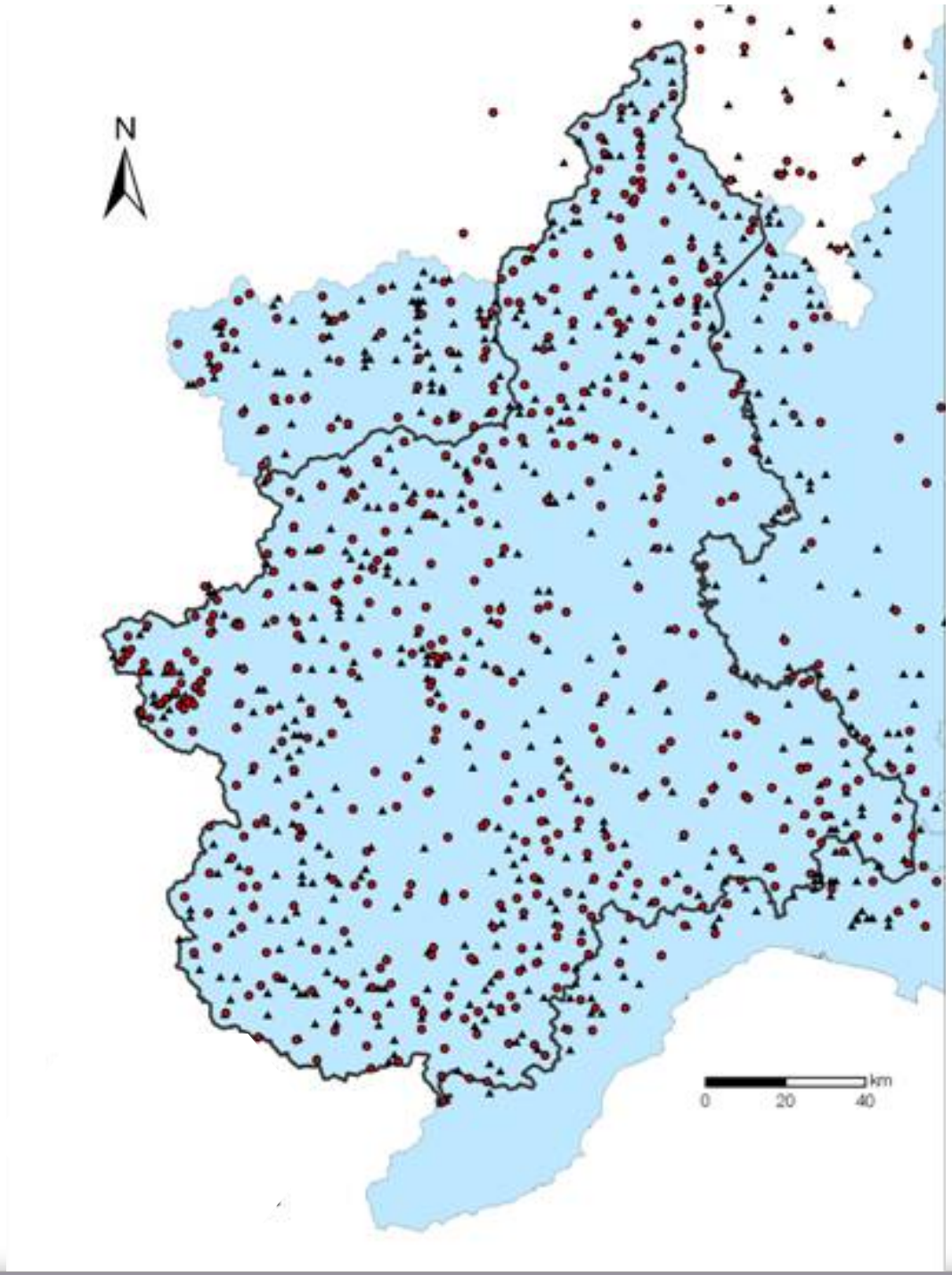


Piedmont area 25,399 km

Mountain 48.7%

Hill 25.9%

Plain 25.4%



From 1986 in Piedmont are present two independent climate network

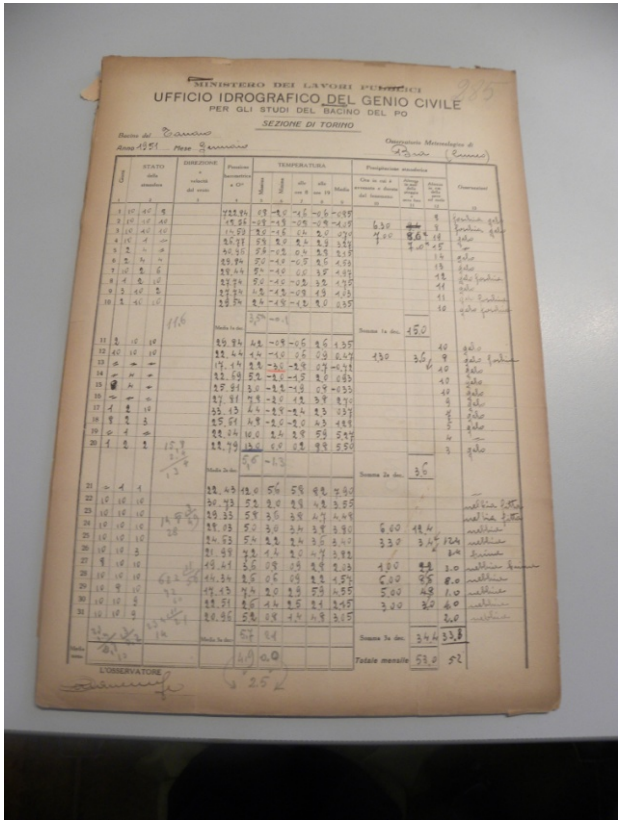
▲ **Hydrographic Mareographic Italian Service, SIMN from 1913**

● **Regional Agency for Environmental Protection Piedmont, ARPA from 1986**

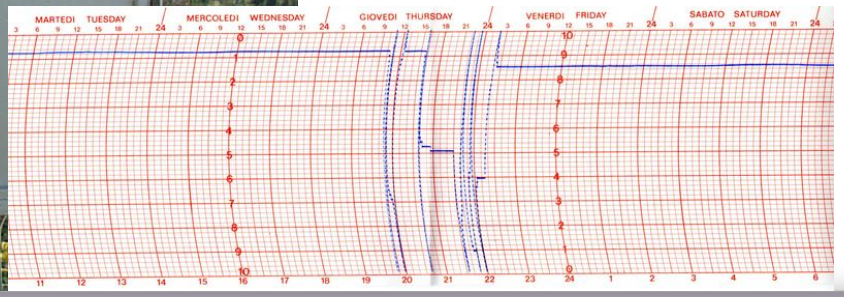
In 2002, a national law has forced the unification of the meteorological networks owned by the SIMN with those of the ARPA.
ARPA has decided to discontinue the SIMN stations.

Instruments

SIMN



The stations of SIMN require the presence of a operator for collecting the measurements



ARPA

The stations of ARPA don't require the presence of a operator for collecting the measurements.

The data are subjected to an immediate quality control that attaches a flag. The values used are indicated with the flag Z "correct data".

	A	B	C
1	DAY	RAIN	FLAG
2	01/01/2004	22.4	*
3	02/01/2004	2.4	Z
4	03/01/2004	24.6	Z
5	04/01/2004	0	Z
6	05/01/2004	0	Z
7	06/01/2004	0	Z
8	07/01/2004	0	Z
9	08/01/2004	0	Z
10	09/01/2004	0.2	Z
11	10/01/2004	0	Z
12	11/01/2004	0	Z
13	12/01/2004	0	Z
14	13/01/2004	0	Z
15	14/01/2004	0	Z
16	15/01/2004	0	Z
17	16/01/2004	0	Z
18	17/01/2004	0	Z
19	18/01/2004	0.4	Z
20	19/01/2004	19.8	Z
21	20/01/2004	2.2	Z
22	21/01/2004	0	Z
23	22/01/2004	0	Z
24	23/01/2004	0	Z
25	24/01/2004	0	Z
26	25/01/2004	0	Z

Methodology

Selection of stations pairs

✓A good overlapping period greater than 5 years (Vincent and Mekis 200

✓Difference in elevation ≤ 200m (Biancotti at al. 2005)

✓Difference in distance ≤ 20 Km (Isotta et al. 2013)

Historical Research and Quality Control

✓For the SIMN and ARPA stations a continuous and accurate historical research is available to detect potential breaks (metadata).

✓The hourly ARPA values were aggregated in daily data, from 9 am to 9 am, as the daily SIMN data.

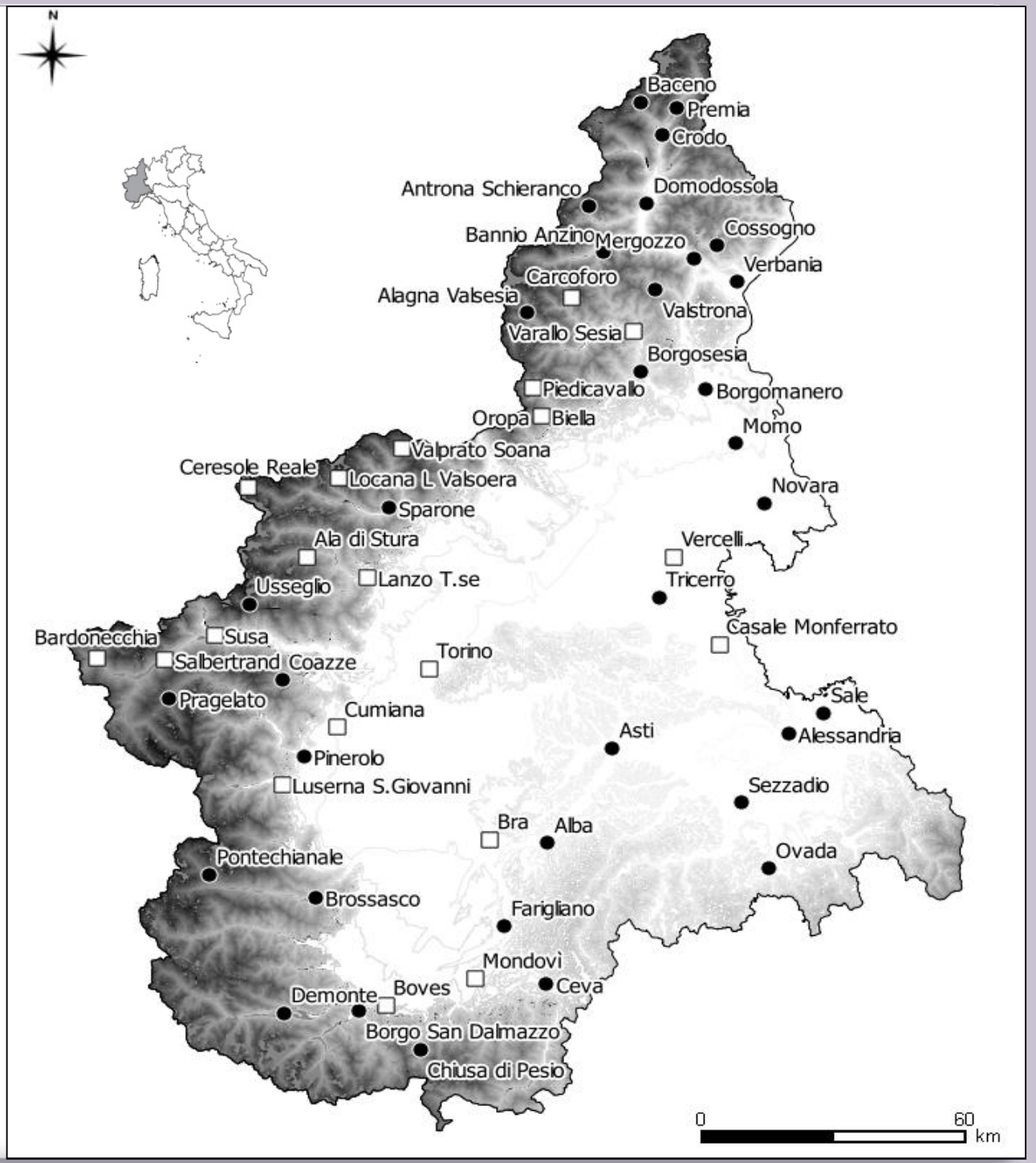
✓Manual quality control was carried out (RClimdex Zhang at al. 2004).

✓Identification of missing data (Gokturk et al 2008 and Tank et al. 2002). Missing values in one series were also set to be missing in its counterpart.

✓The data were not corrected for evaporation, wetting loss and splash.

For every location we have created a metadata file

From
55 locations
To
20 locations
(denote with square)



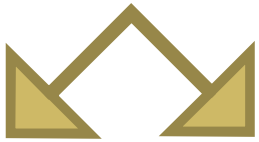
20 pairs of stations are available

Location	*SIMN elevation*	*ARPA elevation*	*Difference elevation*	*Distance*	*Period*
Ala di Stura	1006	1006	0	70	1993–2003
Bardonec					-2003
Boves					–2003
Bra					–2003
Carcofor					–2003
Casale M					–2000
Ceresole					–2003
Cumiana					–2003
Lanzo T.					–1999
Locana					-2003
Luserna					–2003
Mondovi					–2003
Oropa					–2002
Piedicava					–2003
Salbeltrand	1031	1010	21	1250	1991–2002
Susa	510	520	10	820	1991–2003
Torino	270	240	30	850	1990–2003
Valprato Soana	1550	1555	5	465	1993–1999
Varallo Sesia	453	470	17	2040	1989–2003
Vercelli	135	132	3	1360	1994–2003

	Mean
Overlapping period	12 (year)
Difference elevation	28 [m]
Distance	947 [m]

Comparison study on monthly data

For the precipitation daily series



R_ARPA and R_SIMN
the monthly precipitation sums

D_ARPA and D_SIMN
days with precipitation ≥ 1mm
were analyzed

- The Kolmogorov-Smirnov test (Sneyer 1990) is applied to the monthly precipitation series, and rainy days series.
- The Spearman correlation coefficient was calculated.
- On the pairs of series a 2-factorial ANOVA test was applied. One factor is the month and the second factor is the network.

The Shapiro_Wilk test is applied to test the normally distribution.

From the monthly precipitation amounts, the series of their ratios and relative percentage error (Kenneth et al. 2010) have been calculated

$$R = \frac{R_{ARPA\_monthly}}{R_{SIMN\_monthly}}$$

For the rain days series the difference $D = DR_{ARPA} - DR_{SIMN}$

Over the new series R and D a statistical analysis has been carried out

The extreme values of the R and D, values that fall in the distribution tails, have been checked by examining the daily values of SIMN and ARPA comparing them with the values of neighbouring stations.

The extreme values of the R and D monthly series allow to identify the time period when the instruments have not worked correctly.

Location	*Difference elevation*	*Distance*	SIMN	ARPA	R	Err R	ϱ	ANOVA
Ala di Stura	0	70	1413	1278	1.11	0.03	0.94	<0.001
Bardonecchia	**103**	**800**	**751**	**734**	**1.03**	**0.02**	**0.95**	**0.95**
Boves	15	1240	1338	1107	1.23	0.05	0.91	<0.001
Bra	5	15	731	616	1.20	0.02	0.95	<0.001
Carcoforo	140	2500	1683	1416	1.20	0.02	0.97	<0.001
Casale M.	5	20	673	570	1.30	0.02	0.95	<0.001
Ceresole R.	**44**	**920**	**988**	**903**	**1.11**	**0.11**	**0.93**	**0.23**
Cumiana	**38**	**2800**	**806**	**837**	**0.98**	**0.04**	**0.92**	**0.12**
Lanzo T.se	40	2200	1101	1428	0.78	0.03	0.95	<0.001
Locana –	45	250	1198	909	1.31	0.09	0.91	<0.001
Luserna S. G.	**3**	**760**	**996**	**1018**	**0.98**	**0.05**	**0.93**	**0.28**
Mondovì	18	390	839	760	1.10	0.03	0.94	0.04
Oropa	6	5	2240	1955	1.15	0.02	0.99	<0.001
Piedicavallo	**10**	**180**	**1798**	**1736**	**1.03**	**0.02**	**0.98**	**0.13**
Salbertrand	**21**	**1250**	**764**	**732**	**1.06**	**0.05**	**0.92**	**0.51**
Susa	**10**	**820**	**717**	**700**	**1.02**	**0.02**	**0.97**	**0.16**
Torino	30	850	823	851	0.97	0.02	0.98	0.007
Valprato S.	**5**	**465**	**1177**	**1142**	**1.05**	**0.06**	**0.86**	**0.76**
Varallo Sesia	**17**	**2040**	**1961**	**1792**	**1.10**	**0.02**	**0.95**	**0.06**
Vercelli	3	1360	827	763	1.10	0.03	0.95	0.02

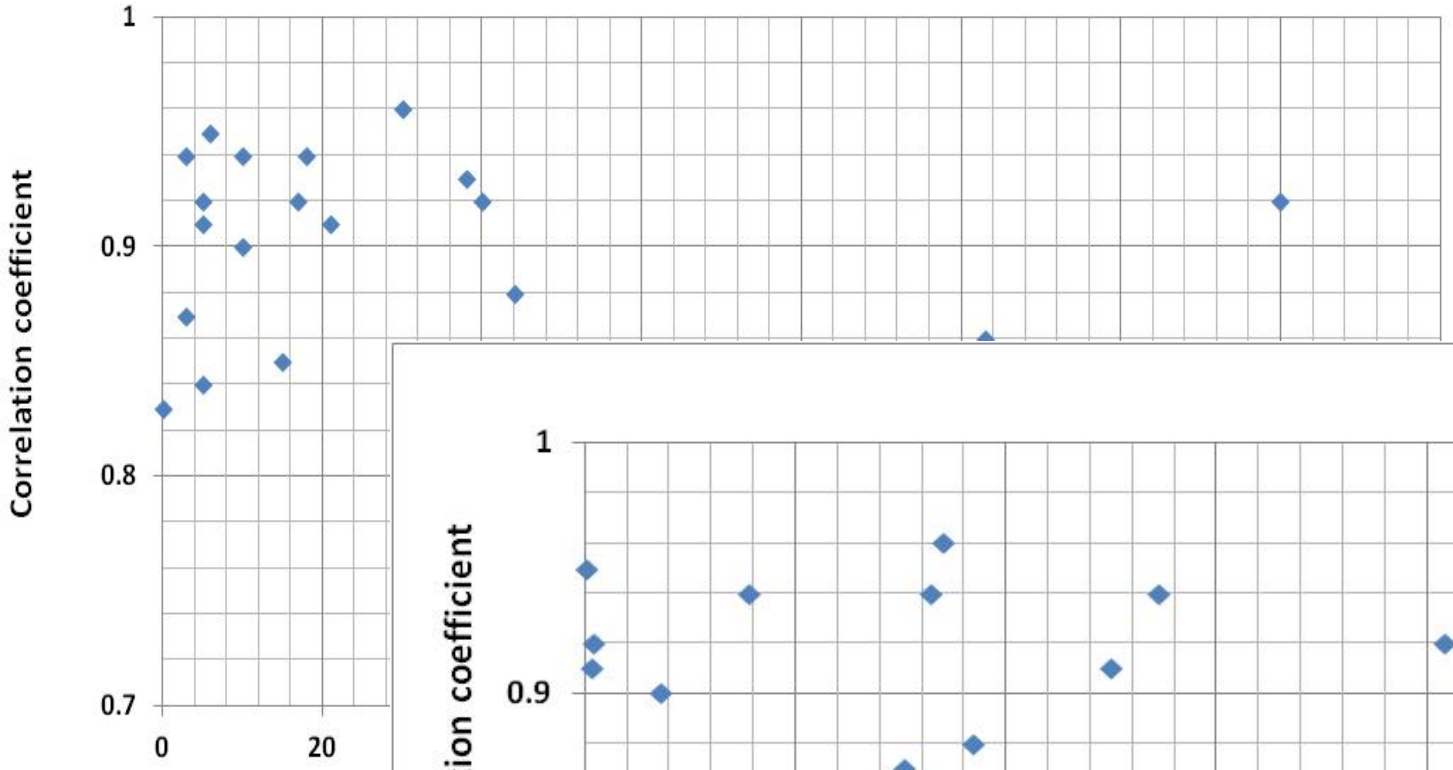
	Difference elevation	*Distance*	SIMN	ARPA	R
All location	28	947	1141.2	1062.4	1.07
Stable locations	**28**	**1115**	**1106.5**	**1066.0**	**1.04**
Variable locations	28	809	1169.6	1059.5	1.10

No. 9 Stable locations
No. 11 Variable locations

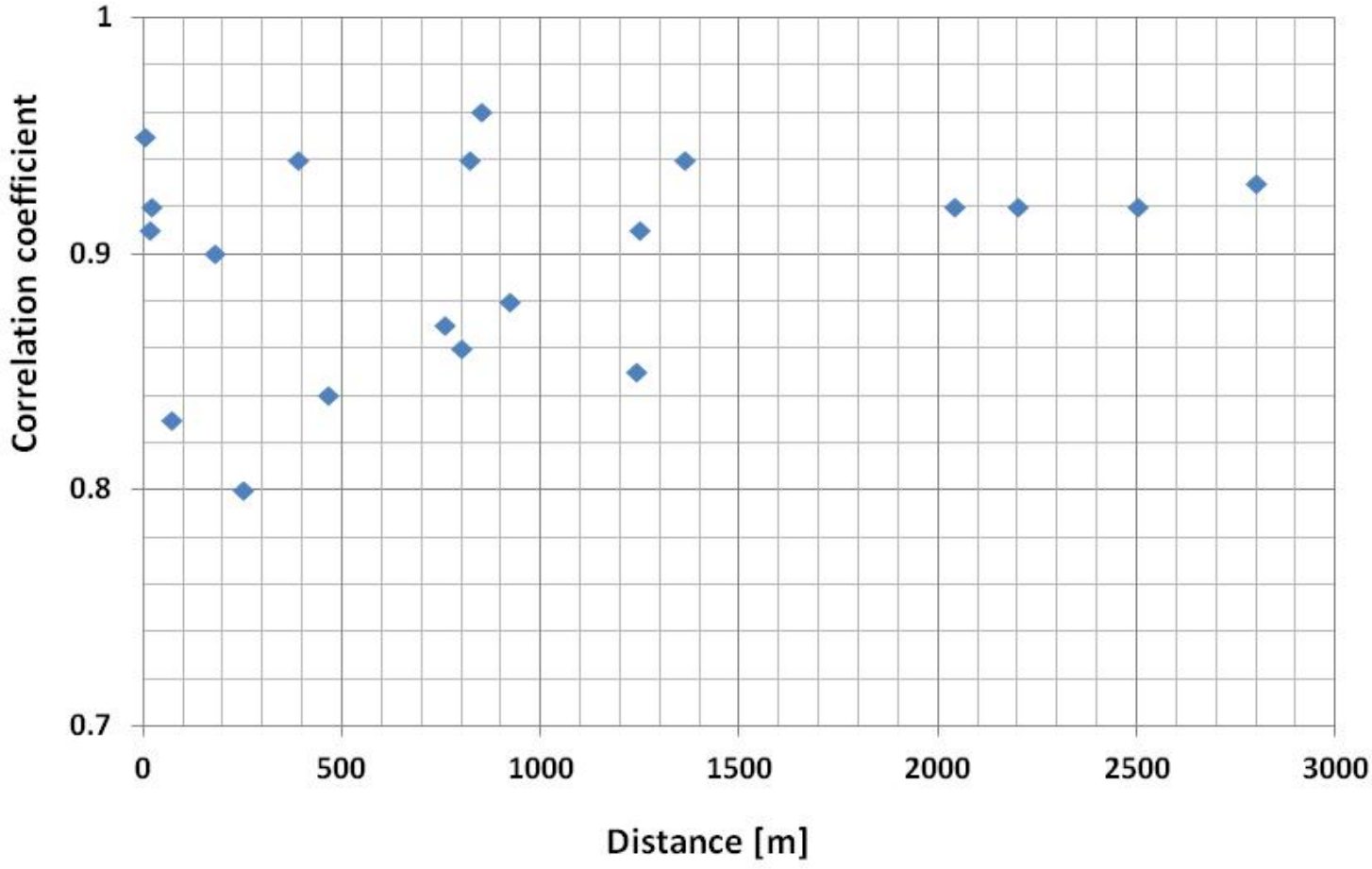
Location	*Difference elevation*	*Distance*	N-SIMN	N-ARPA	D	Err D	ϱ	ANOVA
Ala di Stura	**0**	**70**	**91**	**92**	**-1**	**4**	**0.83**	**0.37**
Bardonecchia	**103**	**800**	**89**	**91**	**-2**	**3**	**0.86**	**0.15**
Boves	15	1240	52	82	-30	3	0.85	<0.001
Bra	5	15	68	61	7	1	0.91	0.002
Carcoforo	140	2500	66	96	-30	5	0.92	<0.001
Casale M.	**5**	**20**	**63**	**61**	**2**	**1**	**0.92**	**0.23**
Ceresole Reale	44	920	99	101	2	1	0.88	0.04
Cumiana	**38**	**2800**	**69**	**71**	**-2**	**1**	**0.93**	**0.16**
Lanzo T.se	40	2200	73	93	-20	4	0.92	<0.001
Locana	**45**	**250**	**99**	**94**	**5**	**2**	**0.80**	**0.23**
Luserna S. G.	3	760	71	78	-7	3	0.87	0.03
Mondovì	**18**	**390**	**65**	**65**	**-0.01**	**0.12**	**0.94**	**0.95**
Oropa	6	5	105	100	5	2	0.95	0.01
Piedicavallo	10	180	93	106	-14	2	0.90	<0.001
Salbertrand	**21**	**1250**	**84**	**87**	**-3**	**2**	**0.91**	**0.30**
Susa	**10**	**820**	**75**	**76**	**-1**	**1**	**94**	**0.36**
Torino	30	850	69	73	-3	1	0.96	0.01
Valprato Soana	**5**	**465**	**102**	**111**	**-9**	**3**	**0.84**	**0.25**
Varallo Sesia	**17**	**2040**	**95**	**94**	**-0.2**	**1.4**	**0.92**	**0.50**
Vercelli	**3**	**1360**	**69**	**67**	**1**	**2**	**0.94**	**0.50**

	Difference elevation	*Distance*	N-SIMN	N-ARPA	D
All location	28	947	**80**	**85**	-5
Stable locations	**24**	**933**	82	83	-1
Variable locations	33	963	77	88	-10

No. 11 Stable locations
No. 9 Variable locations



left scatter plot between correlations coefficient and difference elevation.



right scatter plot between correlations coefficient and distance.

For the monthly comparison

✓6 locations, Good_Locations
show stable results for both the variables

✓6 locations, Bad_Locations
show unstable results for both the variables

✓3 locations, Yes_monthly _rain
show stable results for the monthly sum rain

✓5 locations, Yes_rainy_days
show stable results for the number of rainy day

Comparison study on precipitation class

For every location two “new” precipitation series, n_simn and n_arpa, were created neglecting the values ≤ 0.4 mm, error associated to the instrument*

On the two new series, n_simn and n_arpa, a statistical analysis has been carried out (the length, the mean, the median, 1st quantile, 3rd quantile and the maximum)

Test	Plot
Kolmogorov Smirnov test	Histogram
Wilcow rank sum test	QQ-plot

(*Evaluation of measurement data - Guide to the expression of uncertainty in measurement, JCGM 2008)

For every locations the percentile were calculated on the historical series from 1961 to 1990 (30 years)

5 class of precipitation were selected

name	range
weak rain (w_r)	R < 50th
mean rain (m_r)	50th ≤ R < 80th
heavy rain (h_r)	80th ≤ R ≤ 95th
very heavy rain (R95p)	R95p = Rclimdex; R>95p
extremely rain (R99p)	R99p =Rclimdex; R>99p

For each class were calculated

- ✓the numbers of events including in every precipitation class
- ✓the sum of precipitation.
- ✓the difference between the number of events of n_arpa and n_simn
- ✓the ratio between the sum of precipitation

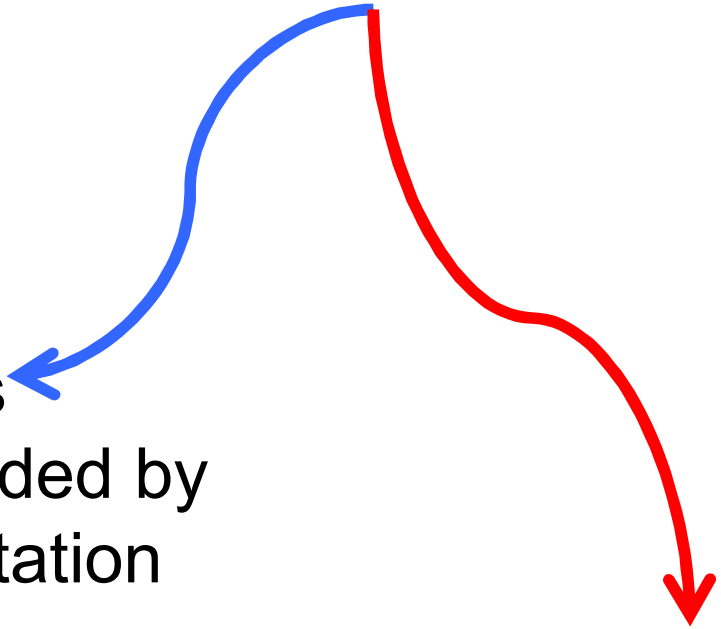
Plot	Test
Box_plot	Kolmogorov Smirnov test
QQ_plot	the Wilcow test
	the Friedman test

For R95p and R99p belonging SIMN and ARPA

- ✓the date of the event.
- ✓From every year the number of events
  the cumulate heavy precipitation
  and their percentage on annual precipitation.

RClimdex to evaluate the variations of precipitation for the location.

The program has been applied on two series:



SIMN_HOM

Homogeneous series

the variable has been recorded by a unique meteorological station

ARPA_not_HOM

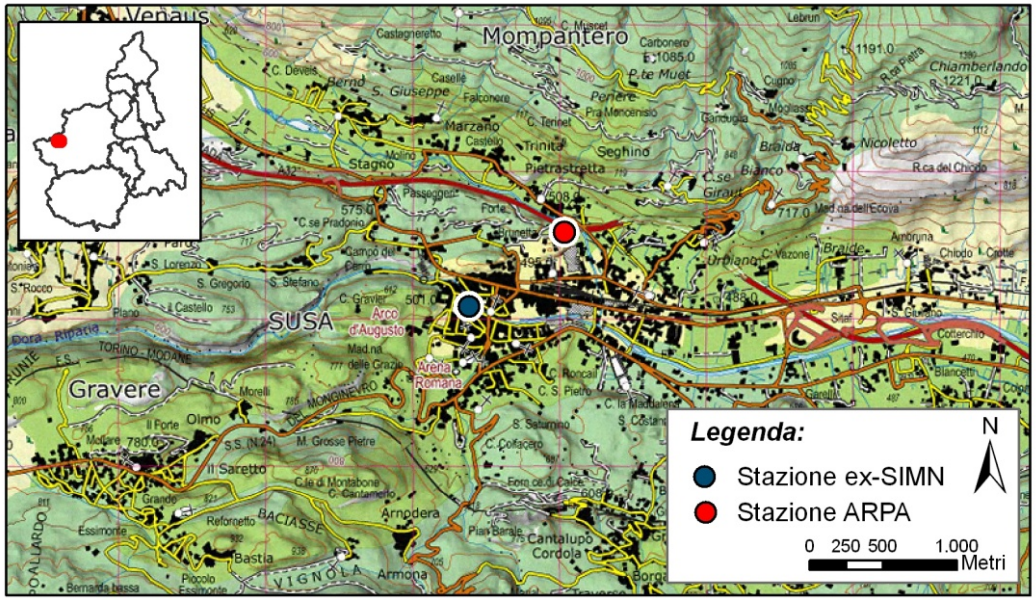
inhomogeneous series

with a break, change of instrument and position, in the year of union

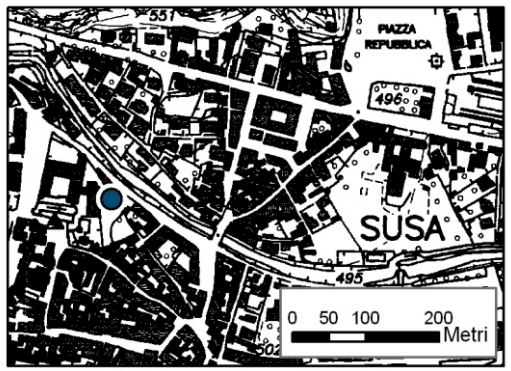
Susa
Good Location

SUSA	*SIMN station*	*ARPA station*
Name or technical code	Susa	Susa -cod. 146-
Municipalities	Susa (TO)	Susa (TO)
	entrale ENEL	Pietrastretta
	Dora Riparia	Dora Riparia
	510	520
	45° 08’’’	45° 08’ 34’’
	° 24’ W M M	7° 03’ 18’’
	346442	347088
	5000250	5000758
	1943 ulized rain gauge	05/12/1990 rain gauge PMB2
	31/12/2003	active
		05/12/1990 thermograph
		active
	820	
	10	
	1991 – 2003	

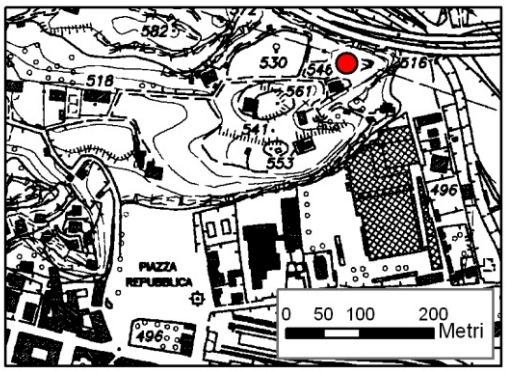
STAZIONI METEOROLOGICHE DI SUSA (TO)



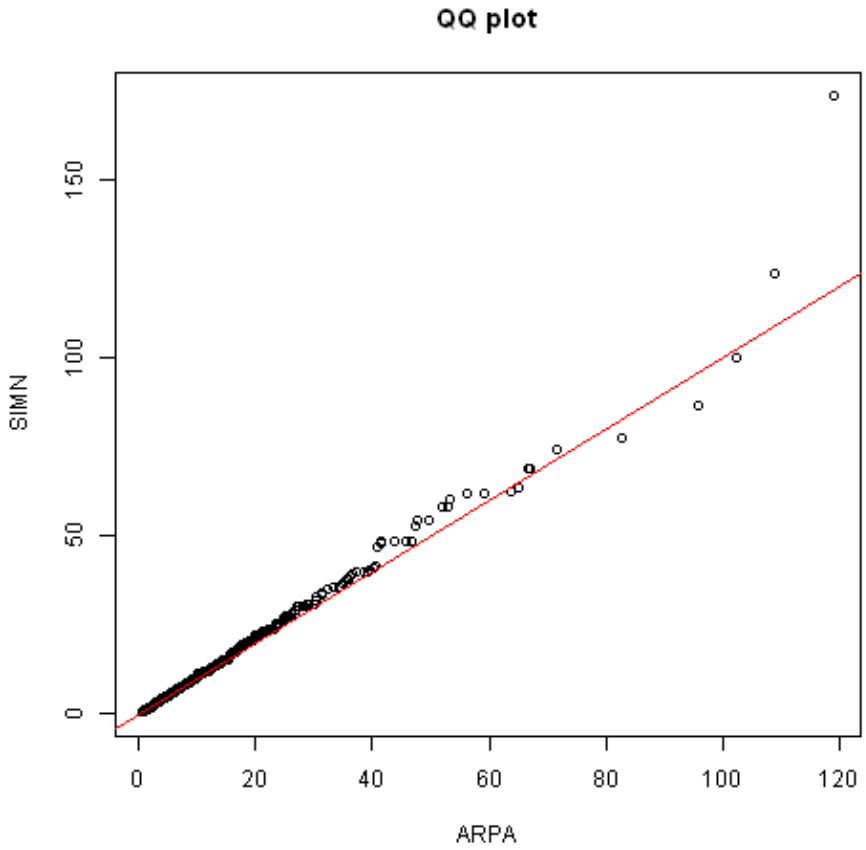
Dettaglio della stazione ex-SIMN

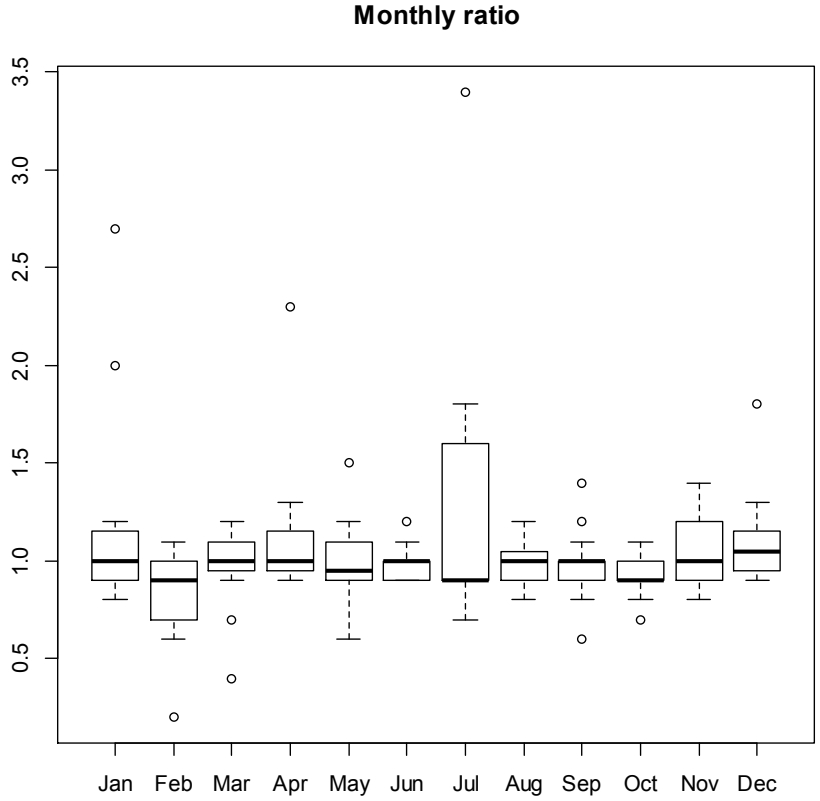


Dettaglio della stazione ARPA



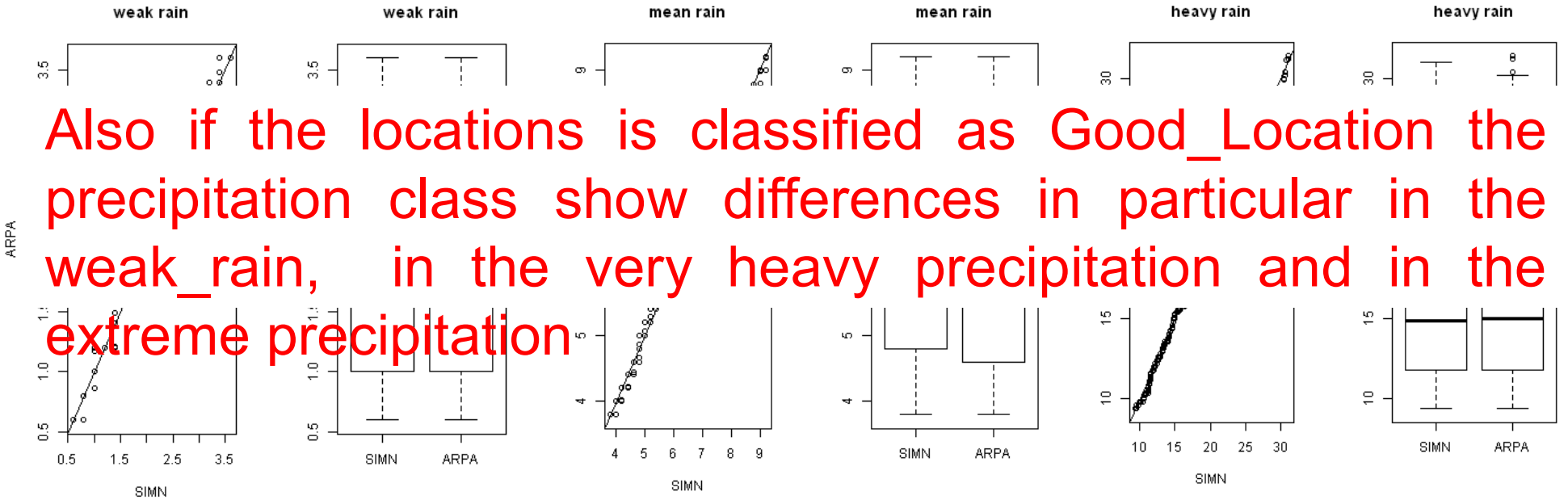
Rain	Arpa	SIMN
Mean annual rain (mm)	763.6	790.1
Max annual rain (mm)	977.4 (1994)	1054 (1994)
Min annual rain (mm)	475 (2003)	461.8 (2003)
Daily maximum rain (mm)	119(October 15, 2000) 109 (November 6, 1994)	173.6 (October 15, 2000) 123.8 (November 6,
Number of rainy days	**ARPA**	**SIMN**
Mean annual rainy days	81	81
Max annual rainy days	104 (1996)	102 (1996)
Min annual rainy days	68 (1991)	59 (1991)



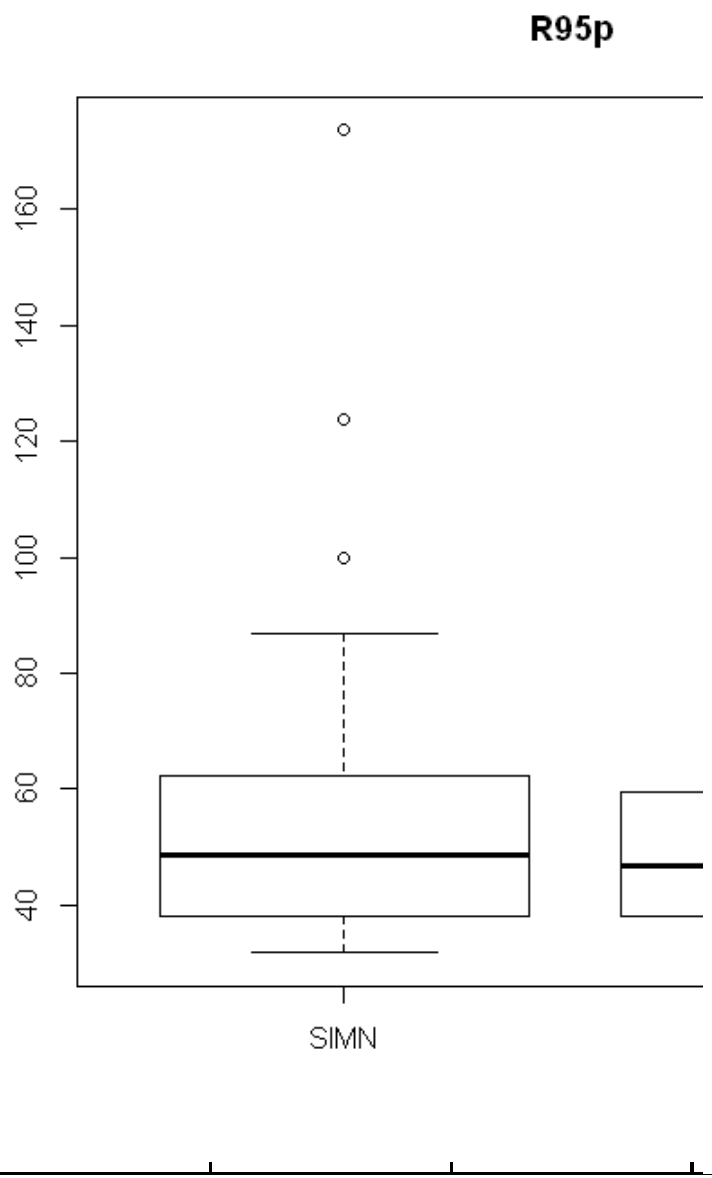


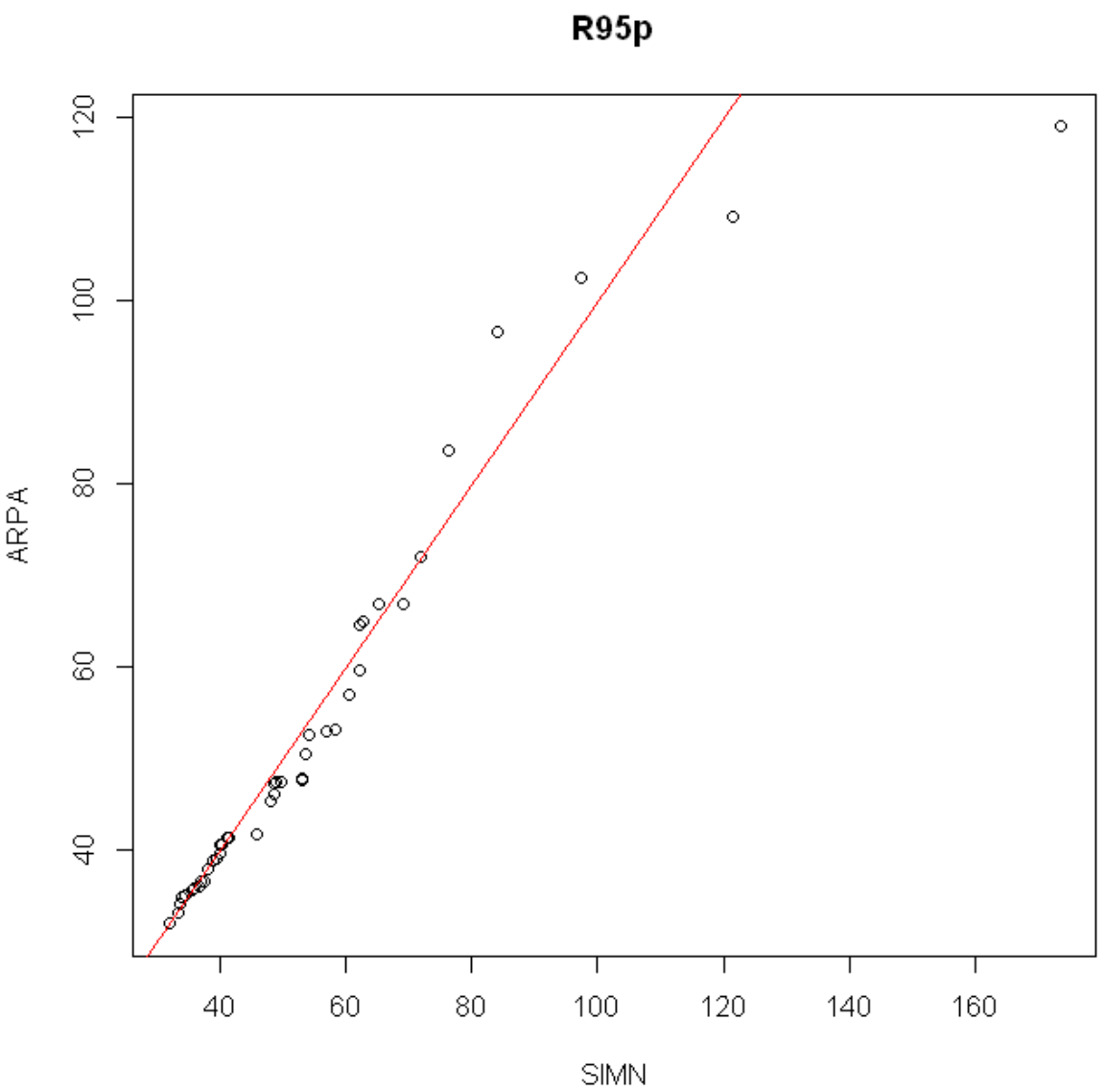
	N_SIMN	N_ARPA
length	1121	1143
var	157.9	131.2
mean	8.4	8.0
RMSE	3.1	

name	Range (mm)	Events_SIMN	Sum_SIMN (mm)	Events_ARPA	SUM_ARPA (mm)	Diffrence	Ratio
w_r	0.4-3.6	507	868	529	915.6	22	1.05
m_r	3.6-9.2	316	1929.6	317	1922.4	1	1.00
h_r	9.2-31.6	252	4127	255	4070	3	0.99
R95p	>31.65	46	2501	42	2213	-4	0.88
R99p	>52.9	19	1415	14	1069	-5	0.76

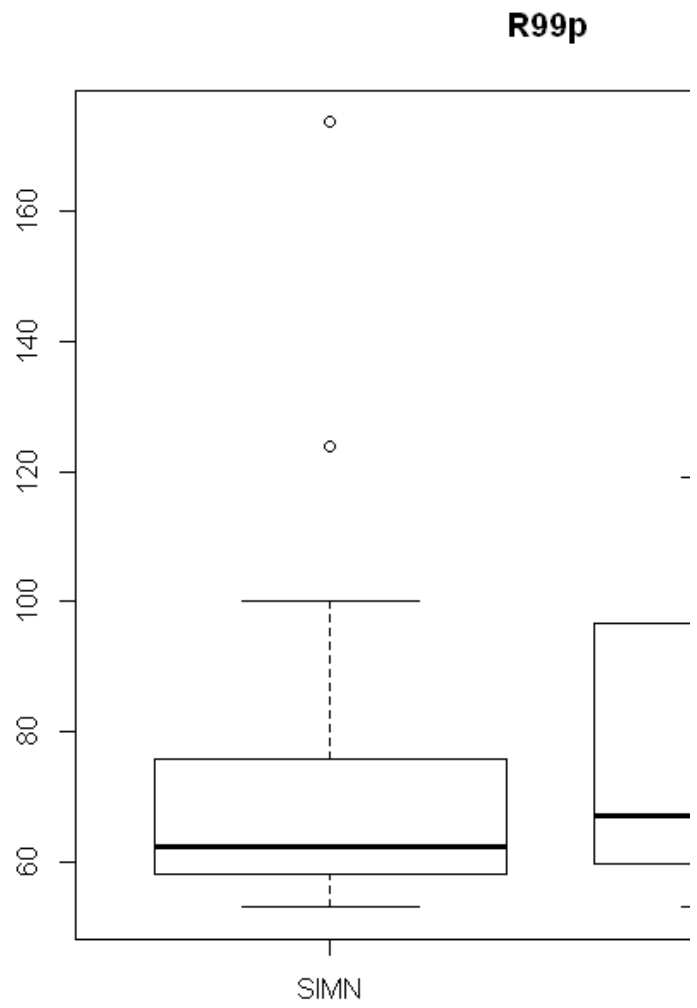


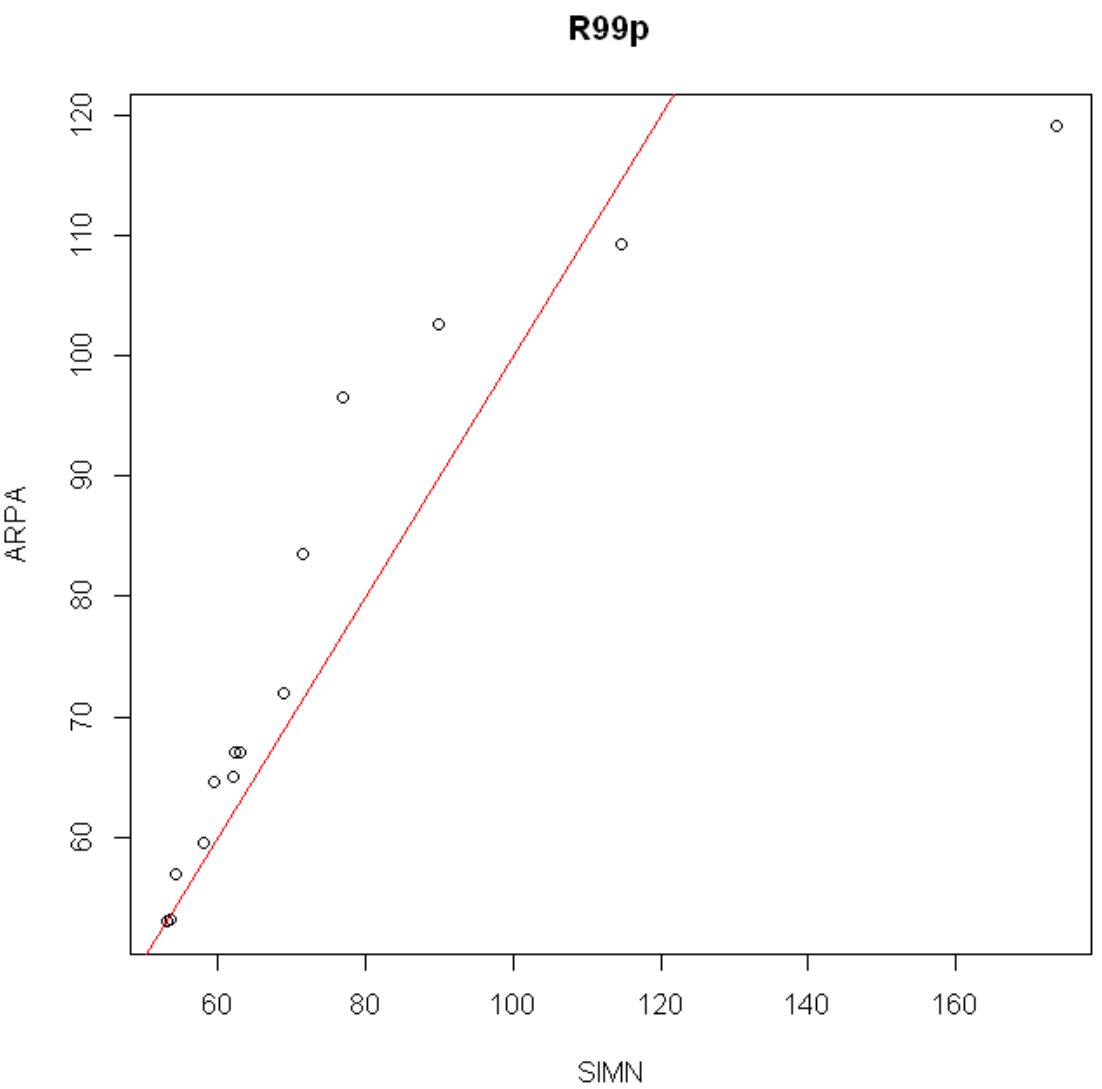
Also if the locations is classified as Good_Location the precipitation class show differences in particular in the weak_rain, in the very heavy precipitation and in the extreme precipitation





R95p	K_s test	W test	Frid. Test
p_value	0.61	0.75	0.03





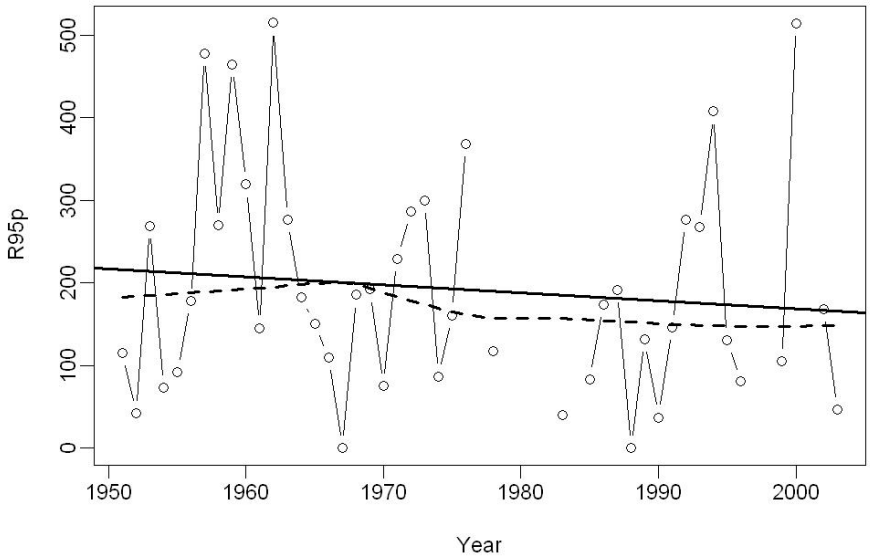
2003	0	0				

R99p	K_s test	W test	Frid. Test
p_value	0.49	0.5	0.05

RClimdex

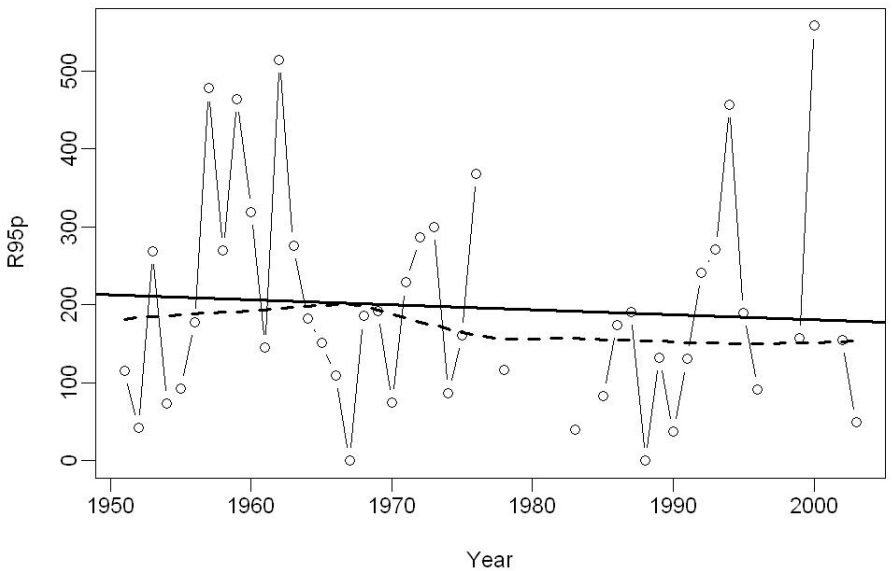
SUSA				ARPA_no HOM			SIMN_HOM		
Indices	Unit	SYear	EYear	Slope	STD of Slope	P_Value	Slope	STD of Slope	P_Value
rx1day	Mm	1951	2003	-0.23	0.24	0.34	-0.12	0.25	0.65
rx5day	Mm	1951	2003	-0.11	0.62	0.86	-0.05	0.63	0.94
sdii	Mm	1951	2003	-0.02	0.02	0.24	-0.01	0.02	0.51
r10mm	Days	1951	2003	-0.07	0.07	0.32	-0.04	0.07	0.58
r20mm	Days	1951	2003	-0.07	0.04	0.09	-0.04	0.04	0.39
R25mm	Days	1951	2003	-0.05	0.03	0.11	-0.04	0.03	0.25
cdd	Days	1951	2003	-0.05	0.11	0.65	0.09	0.13	0.48
cwd	Days	1951	2003	0.01	0.02	0.76	0.01	0.02	0.61
r95p	Mm	1951	2003	-0.98	1.32	0.46	-0.64	1.35	0.64
r99p	Mm	1951	2003	-0.33	0.98	0.74	0.29	1.03	0.78
prcptot	Mm	1951	2003	-1.36	2.04	0.51	-0.71	2.12	0.74

R95p susa_arpa



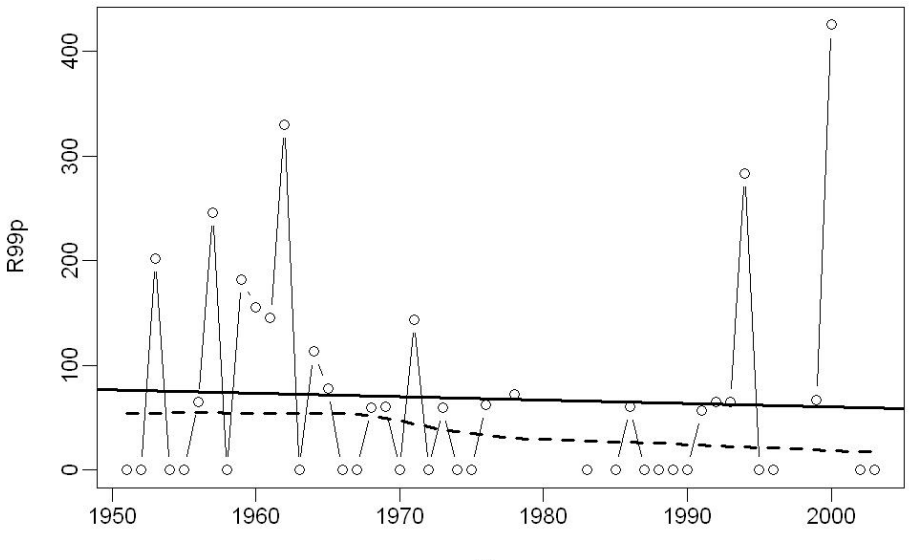
R2= 1.3 p-value= 0.463 Slope estimate= -0.975 Slope error= 1.318

R95p susa_simn



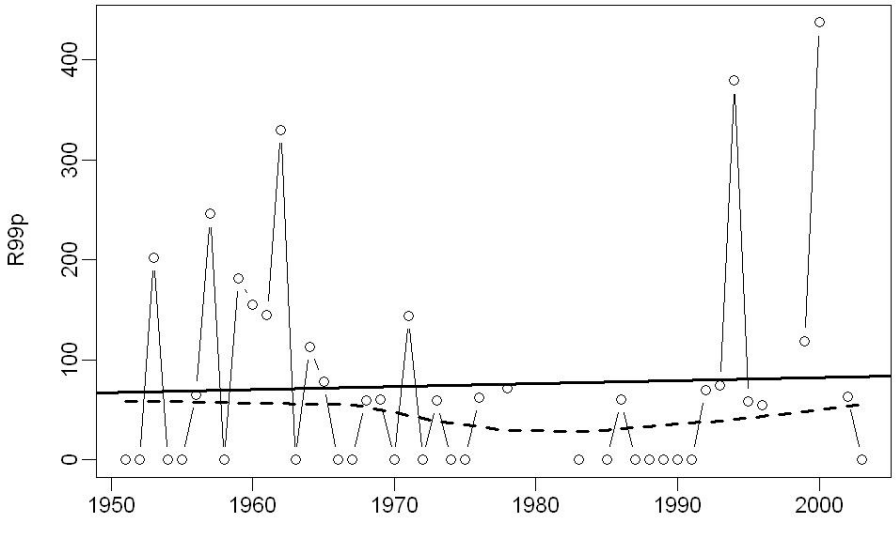
R2= 0.5 p-value= 0.64 Slope estimate= -0.638 Slope error= 1.354

R99p susa_arpa



R2= 0.3 p-value= 0.737 Slope estimate= -0.33 Slope error= 0.976

R99p susa_simn

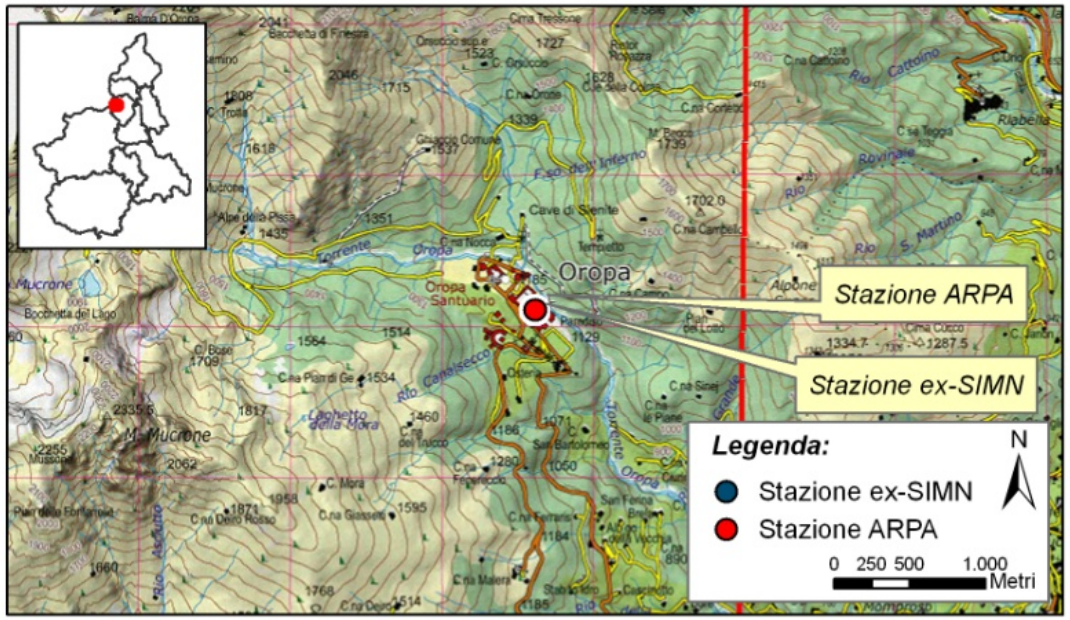


R2= 0.2 p-value= 0.777 Slope estimate= 0.294 Slope error= 1.032

Oropa
Bad Location

Oropa	*SI*
Name or technical code	

STAZIONI METEOROLOGICHE DI OROPA (BI)

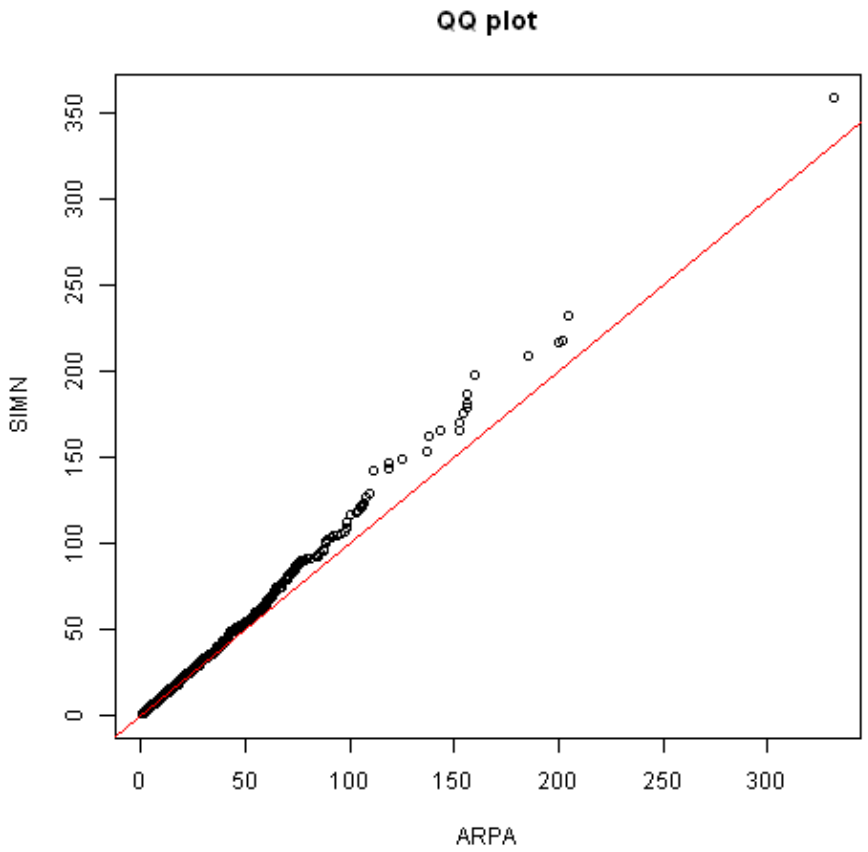


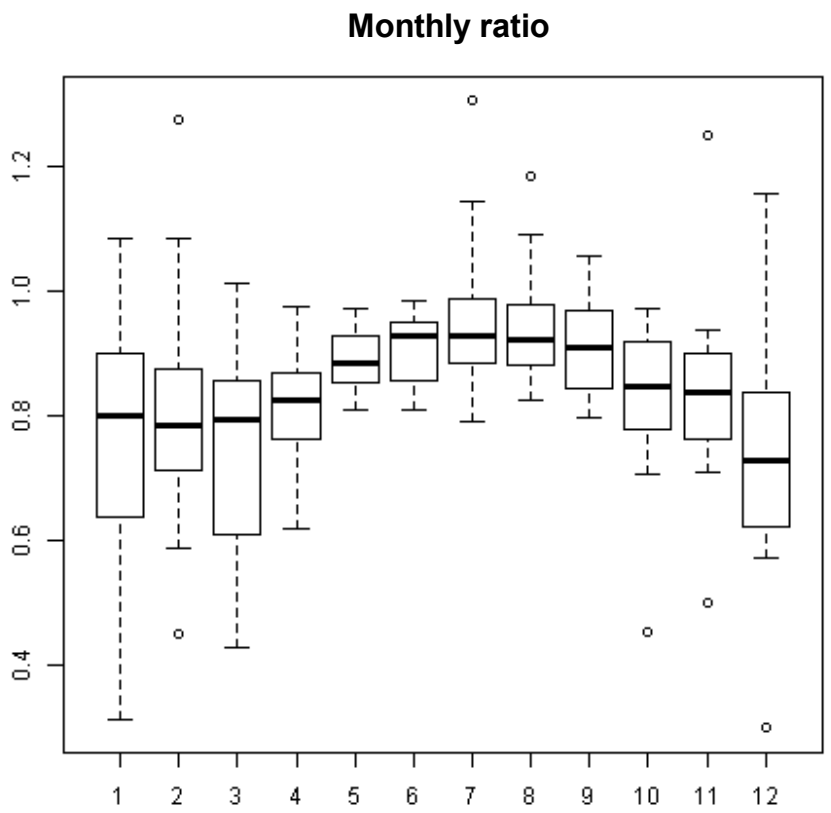
Carta di dettaglio

W M M	58' 57
20669	420668
053279	5053282
1941 n gauge	19/05/1988 rain gauge PMB2
12/2001	active
1867	19/05/1988 thermograph
12/2001	active
2	
6	
1991 – 2002	

Rain	Arpa	SIMN
Mean annual rain (mm)	1949.4	2224.5
Max annual rain (mm)	2692.2 (2000)	2993.4 (1993)
Min annual rain (mm)	1214.8 (2001)	1328 (2001)
Daily maximum rain (mm)	331.4 (November 6, 1994) 204.6 (May 26, 1998)	359 (November 6, 1994) 233.4 (May 26, 1998)

Number of rainy days	ARPA	SIMN
Mean annual rainy days	106	110
Max annual rainy days	118 (1996)	137 (1996)
Min annual rainy days	87 (1997)	93 (1998)



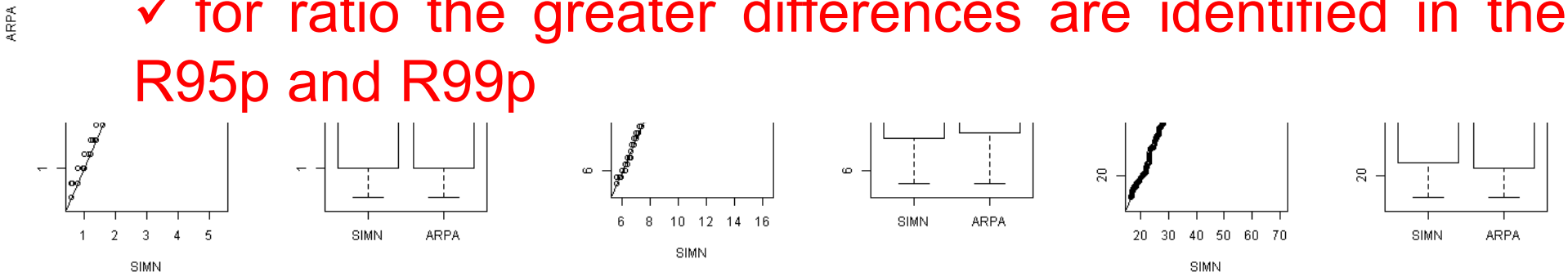


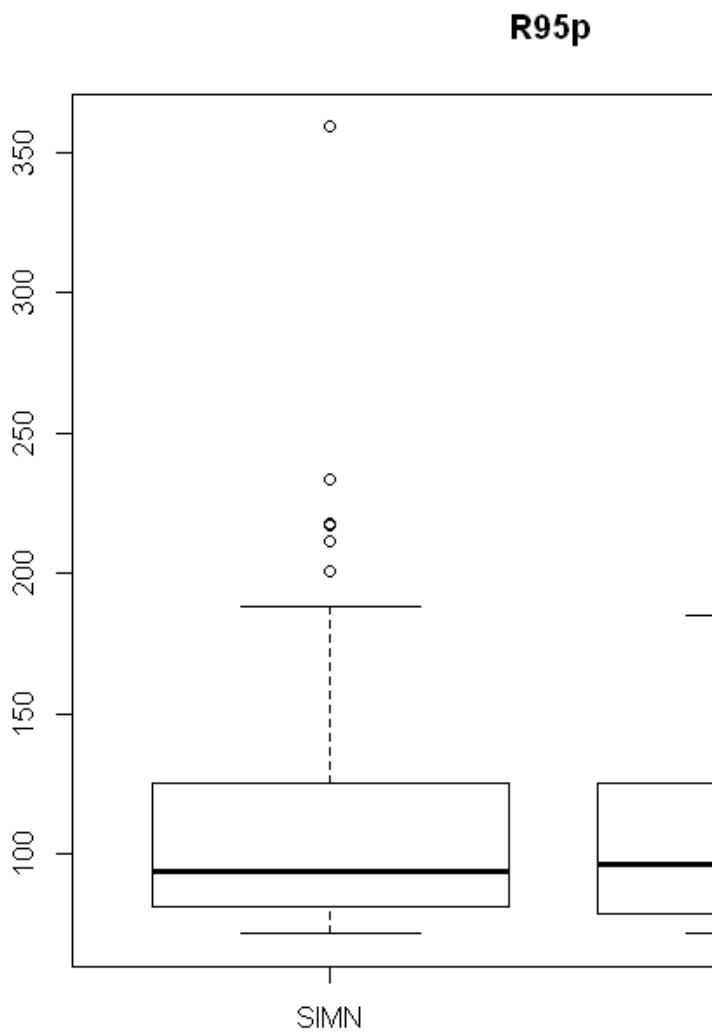
	N_SIMN	N_ARPA
length	1393	1327
var	956.9	754.2
mean	19.3	17.6
RMSE	5.3	

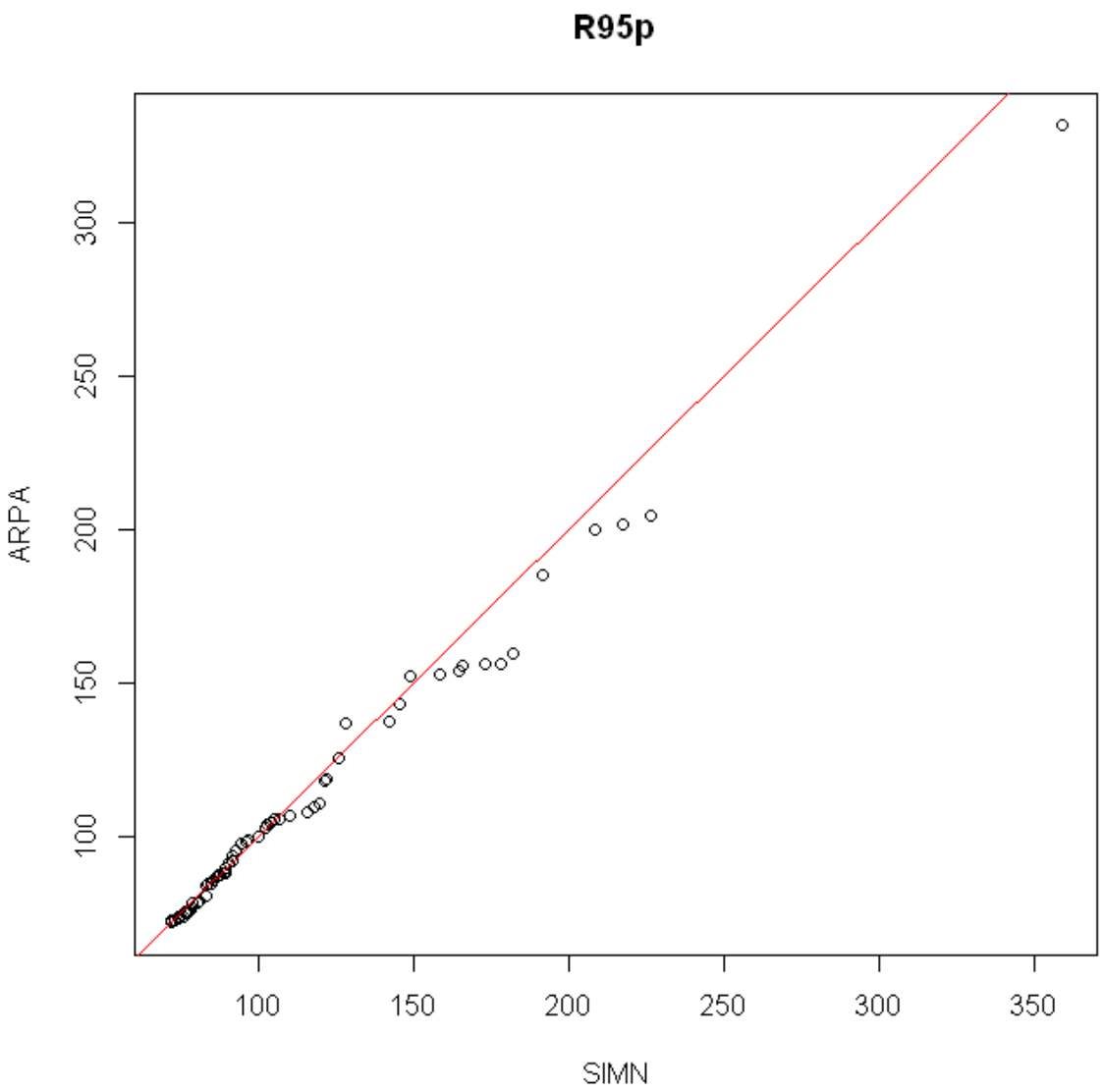
name	Range (mm)	Events_SIMN	Sum_SIMN (mm)	Events_ARPA	SUM_ARPA (mm)	Diffrence	Ratio
w_r	0.4-5.4	585	1342.2	575	1347.8	-10	1.00
m_r	5.4-16.4	371	3670	350	3433.6	-21	0.94
h_r	16.4-71.2	349	11855	340	11753	-9	0.99
R95p	>71.2	88	9955	62	6861	-26	0.69
R99p	>147.7	17	3296	12	2210	-5	0.67

The locations is classified as Bad_Location

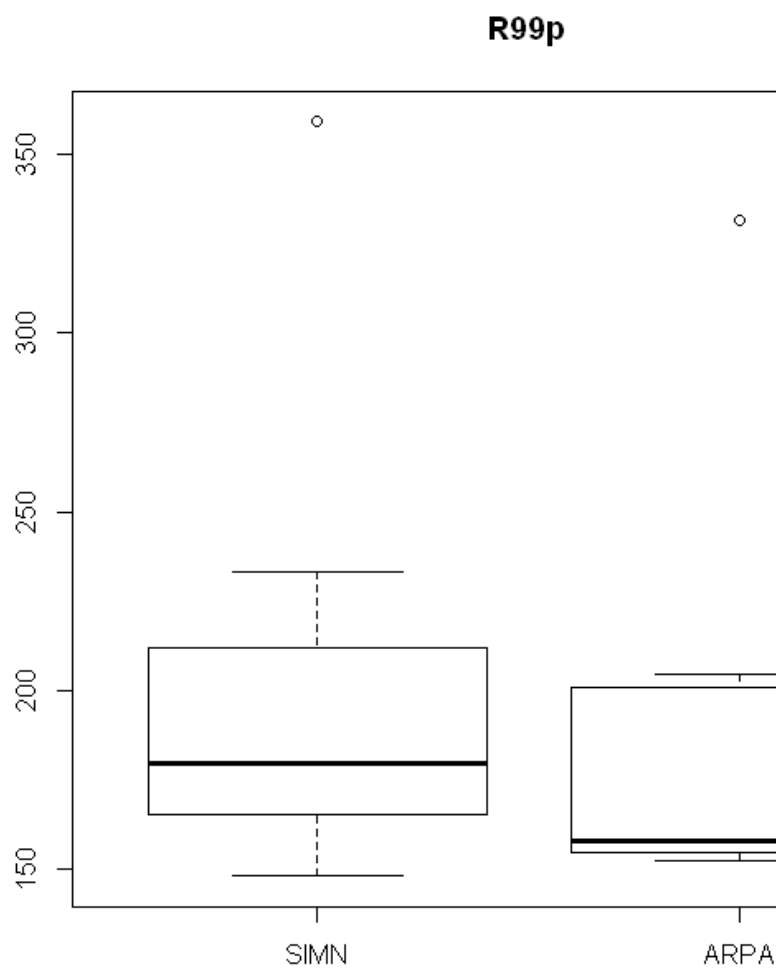
- ✓for the number of events the greater differences are identified in the m_r and R95p
- ✓ for ratio the greater differences are identified in the R95p and R99p

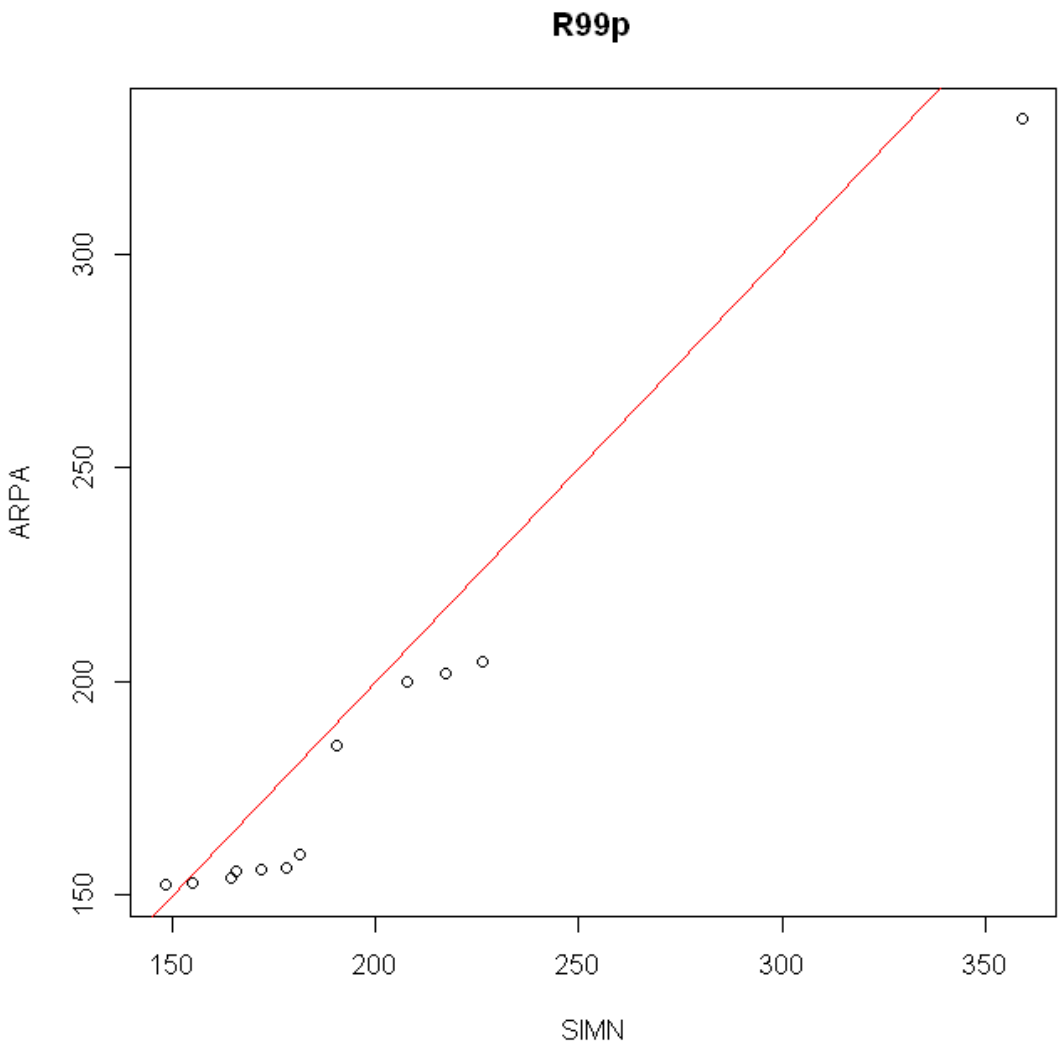






R95p	K_s test	W test	Frid. Test
p_value	0.93	0.82	0.03

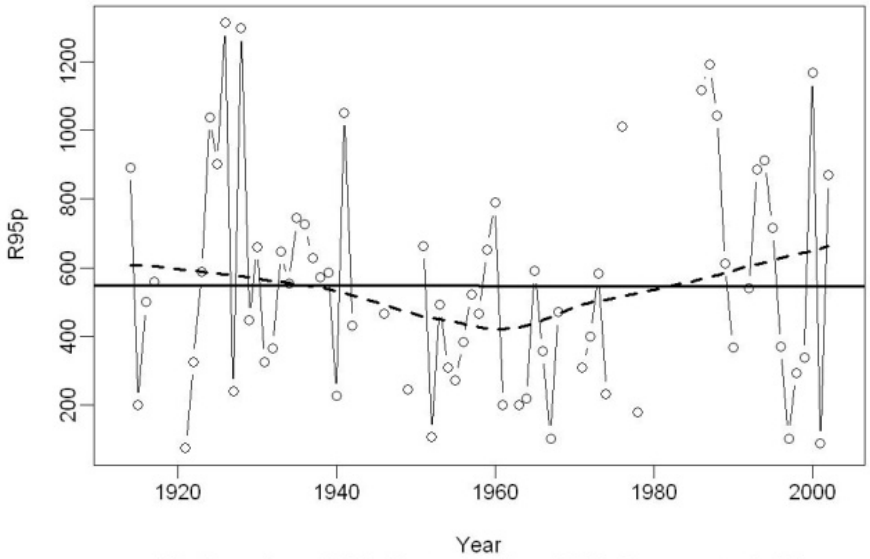


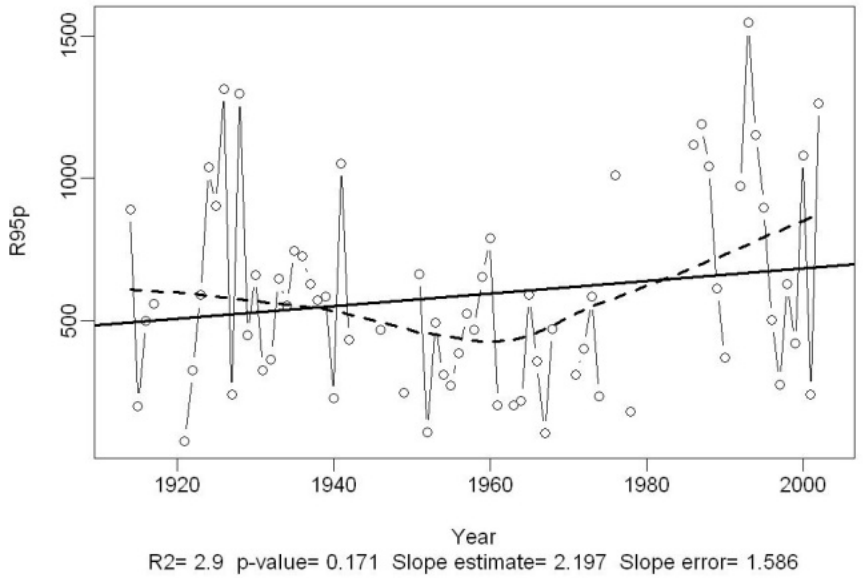


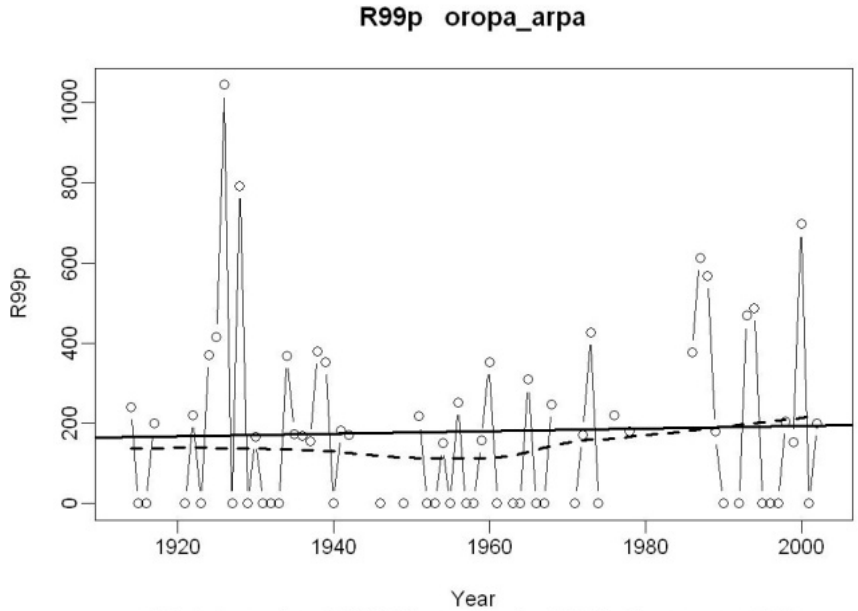
R99p	K_s test	W test	Frid. Test
p_value	0.06	0.25	0.04

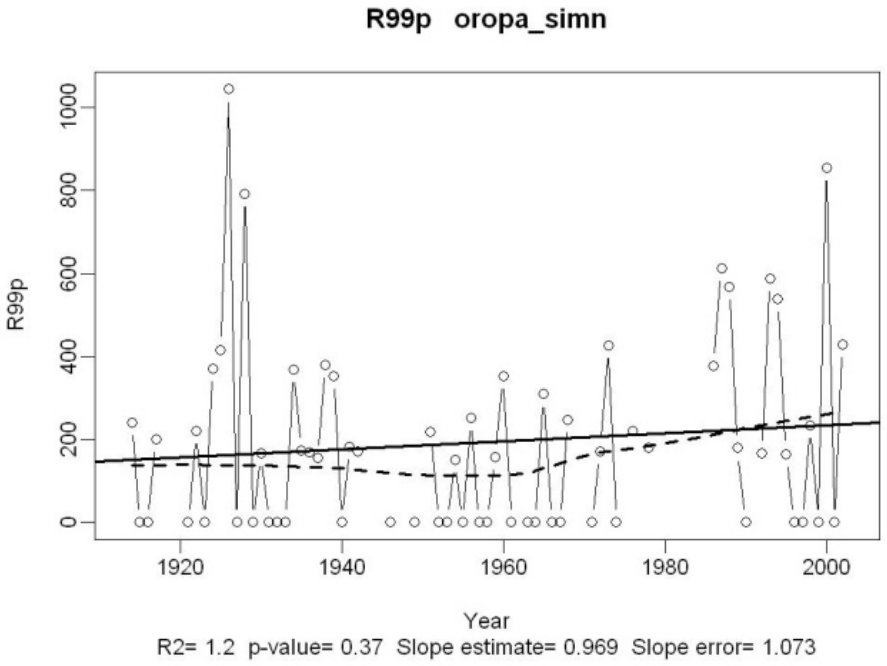
RClimdex

OROPA				ARPA_no_HOM			SIMN_HOM		
Indices	Unit	SYear	EYear	Slope	STD o _Slope	P Value	Slope	STD of Slope	P Value
rx1day	Mm	1913	2003	0.081	0.293	0.784	0.249	0.3	0.41
rx5day	Mm	1913	2003	0.575	0.576	0.322	0.905	0.597	0.135
sdii	Mm	1913	2003	-0.021	0.019	0.268	-0.005	0.019	0.776
r10mm	Days	1913	2003	-0.118	0.045	0.01	-0.073	0.046	0.116
r20mm	Days	1913	2003	-0.057	0.034	0.098	-0.029	0.035	0.404
R25mm	Days	1913	2003	-0.028	0.028	0.315	-0.007	0.029	0.807
cdd	Days	1913	2003	0.014	0.053	0.799	-0.017	0.052	0.743
cwd	Days	1913	2003	-0.006	0.015	0.711	-0.003	0.015	0.607
r95p	Mm	1913	2003	-0.033	1.479	0.982	2.197	1.586	0.171
r99p	Mm	1913	2003	0.327	1.028	0.751	0.969	1.073	0.37
prcptot	Mm	1913	2003	-1.373	2.182	0.531	1.122	2.275	0.624


R95p oropa_arpa
R95p
Year
R2= 0 p-value= 0.982 Slope estimate= -0.033 Slope error= 1.479


R95p oropa_simn
R95p
Year
R2= 2.9 p-value= 0.171 Slope estimate= 2.197 Slope error= 1.586


R99p oropa_arpa
R99p
Year
R2= 0.2 p-value= 0.751 Slope estimate= 0.327 Slope error= 1.028


R99p oropa_simn
R99p
Year
R2= 1.2 p-value= 0.37 Slope estimate= 0.969 Slope error= 1.073

Conclusions

Comparison study on monthly data

✓In some locations, stable locations, the analysis of the precipitation series have showed a variations less or equal to 10% in the years

✓The cause of the discrepancies among the two network could be explained by breakdown of the instrumentation
To explain these variations we have taken into consideration the precipitation data of the neighboring stations

✓The identification of the inhomogeneities and in particular of the causes of the bias allows us a possible correction by homogenization.

✓In the other locations, variable locations, the analysis of the precipitation series have shown significant deviation between the meteorological stations. The sources of the difference between the two networks are unknown. Perhaps the geographical conditions have contributed to increase the differences or a malfunctions not detected in the instruments have increased the discrepancy in the series.

Work in progress

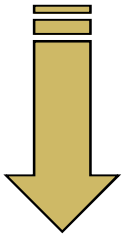
Improve the quality control on daily data

Apply a homogenisation test on the monthly series before the comparison to identify the unknown discontinued period and to evaluate their influence between the two series.

Conclusions

Comparison study on precipitation class

✓For all the stations classificated as Good_location, Yes_monthly _rain, Yes_rainy_days and Bad_location, the comparisons between the class have highlighted important differences in particular in the first class, weak_rain, in the very heavy rain and extreme precipitation..



These differences are important for the indices

Work in progress

The comparison study on precipitation class will be done on seasonal and monthly scale

Thank you