European Geosciences Union
General Assembly 2010
Vienna, Austria, 02 – 07 May 2010

DIPARTIMENTO DI IDRAULICA
TRASPORTI E STRADE

SAPIENZA
UNIVERSITÀ DI ROMA

An entropy based rainfall disaggregation model

V. Montesarchio, F. Napolitano

valeria.montesarchio@uniroma1.it
D.I.T.S. Sapienza University of Rome

HS5.5/NP6.10

Stochastics in hydrometeorological processes: from point to global spatial scales and from minute to climatic time scales

Mon, 03 May 2010

Why a disaggregation model?

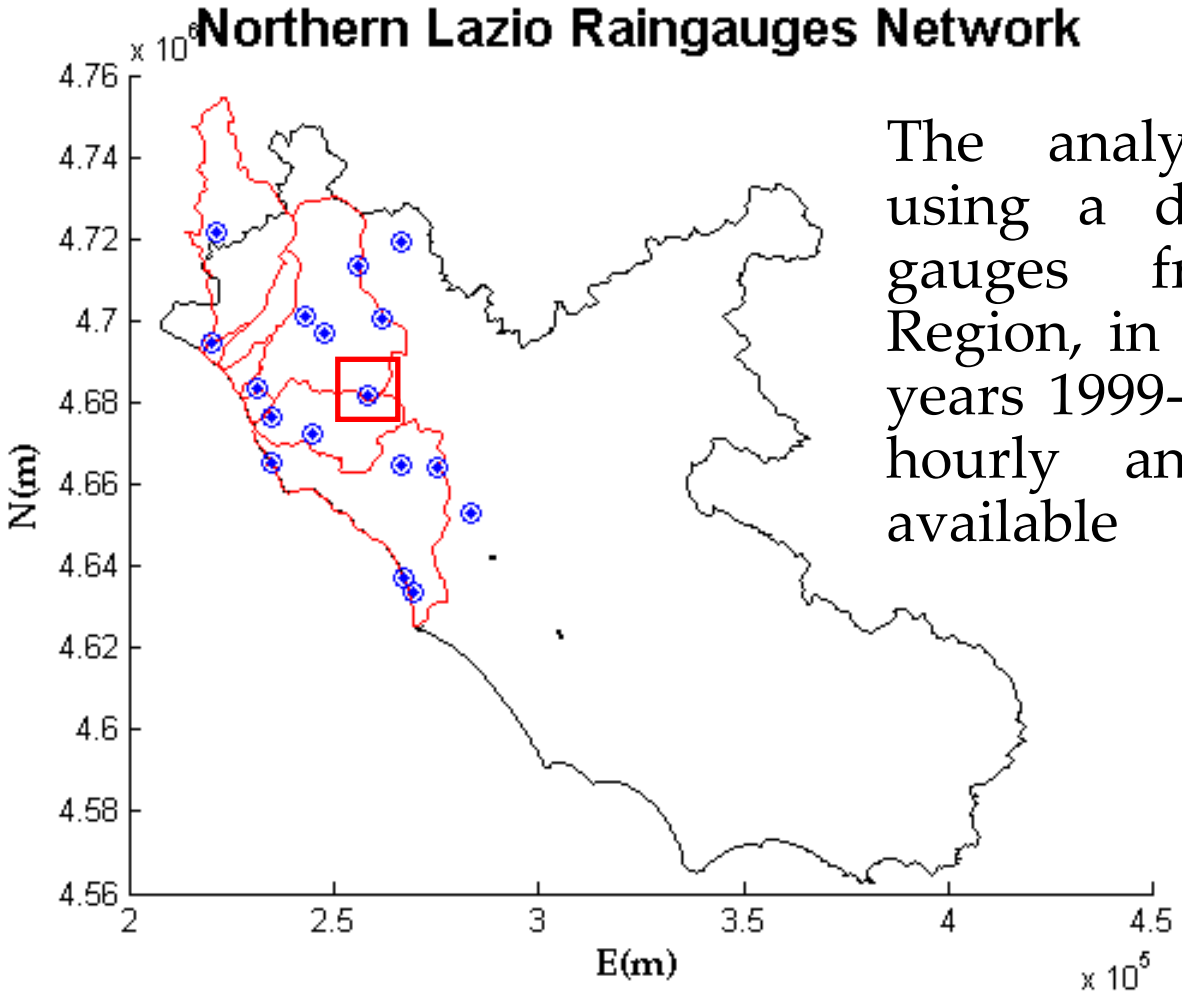
- Daily rainfall measurements are often the only available input for performing environmental, climate, hydrological, and water resources analysis
- Given that to simulate and forecast many hydrological phenomena the knowledge of space and time variability is need, rainfall disaggregation procedures can be used to overcome the problem of data limitations in terms of temporal resolution
- Entropy theory is suitable to explain and model hydrological phenomena, such as the observed dependence properties of the rainfall occurrence process, including the clustering behavior and persistence

entropy approach to disaggregate daily rainfall amounts to hourly level

Outline

- Study area
- Investigation of basic statistical properties performed for different time scales, ranging from 30 minutes to 24 hours
- Evaluation of properties based on entropy concept
- Disaggregation model
- Discussion of results.

Study Area



The analysis are performed using a data set of 20 rain gauges from North Lazio Region, in Central Italy, for the years 1999-2008, for which both hourly and daily data are available

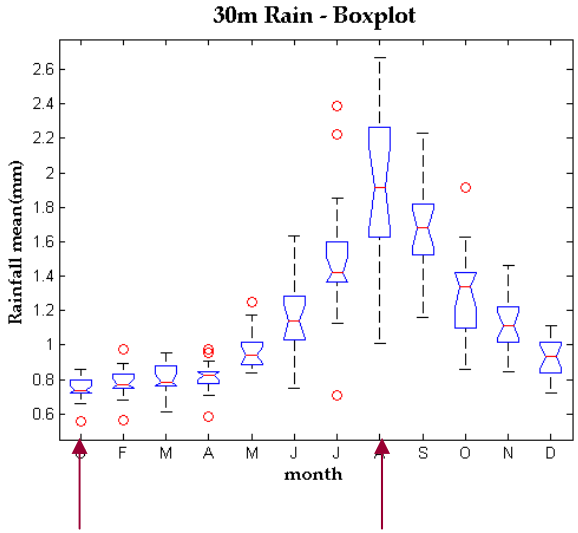
R341-Barbarano

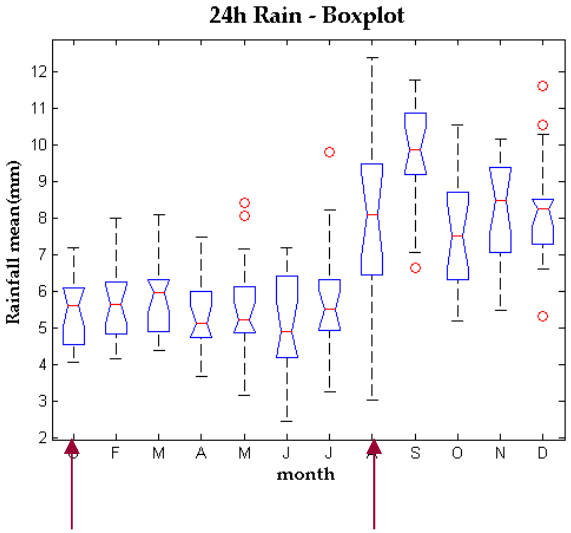
Preliminary analysis

- Rainfall model is required to preserve rainfall amounts and some process characteristics:
  - mean values, variances, skewness;
  - probability and length of dry intervals;
  - dependence structure of rainfall.

Min and max time resolution

Winter and Summer

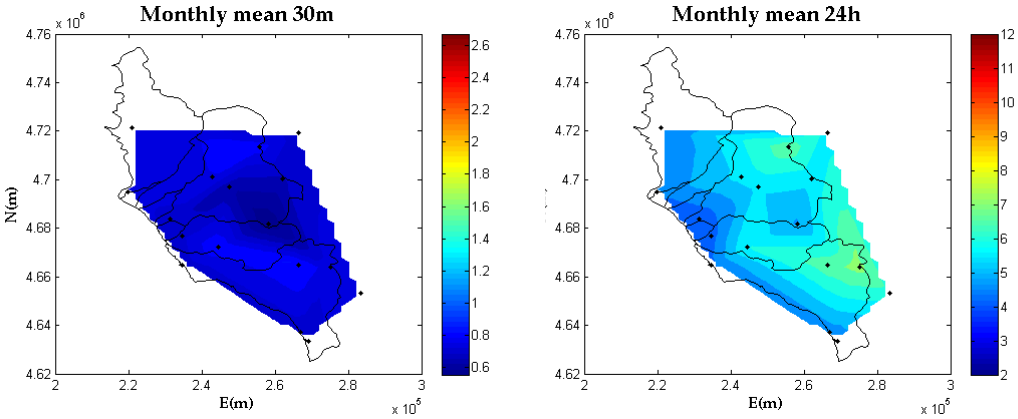




Statistical properties

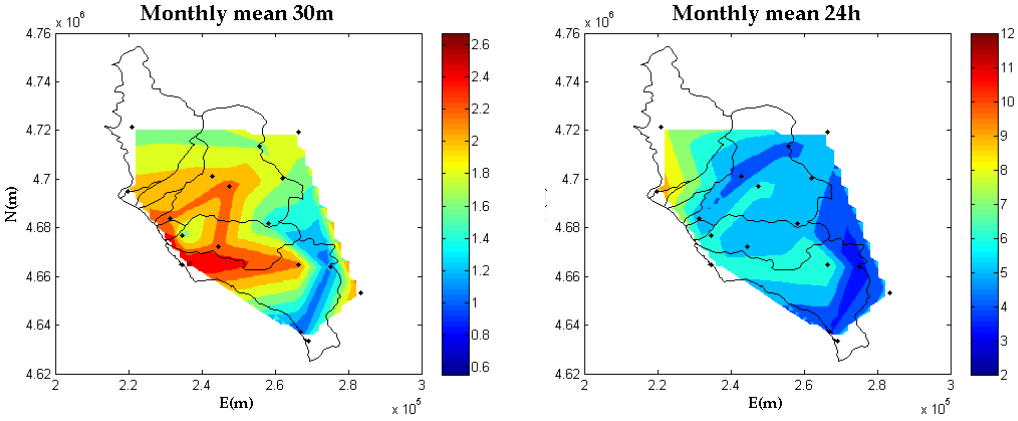
January

	30 min	24 h
Mean	0.80	4.20
Var	1.78	49.91
Skew	3.19	3.83



August

	30 min	24 h
Mean	2.60	9.80
Var	41.01	574.35
Skew	3.19	8.37



Dry intervals

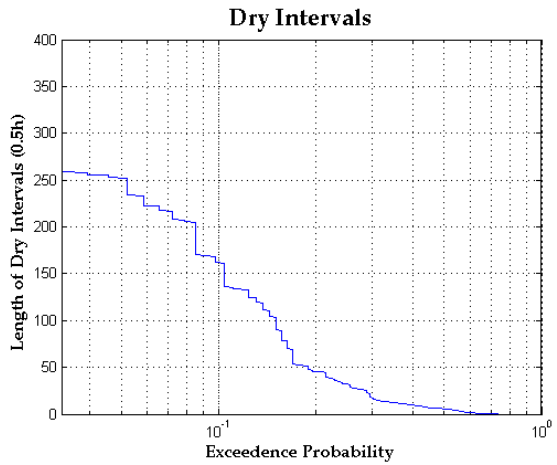
R341-Barbarano

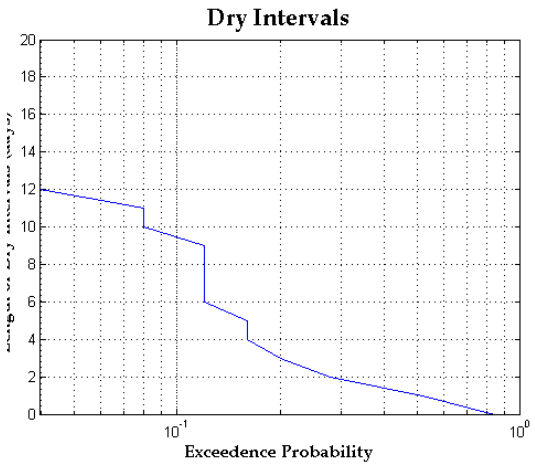
- January

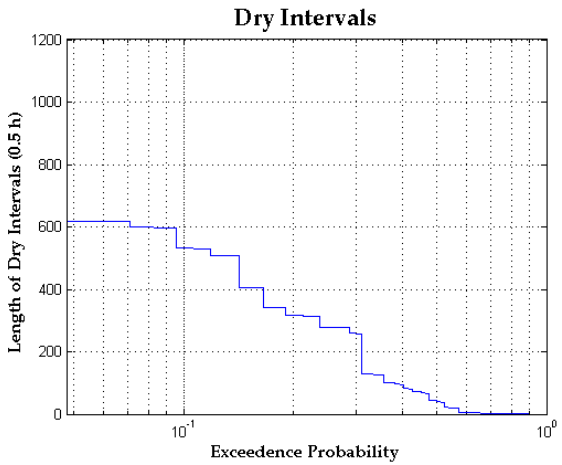
	30 min	24 h
P_dry	0.95	0.71

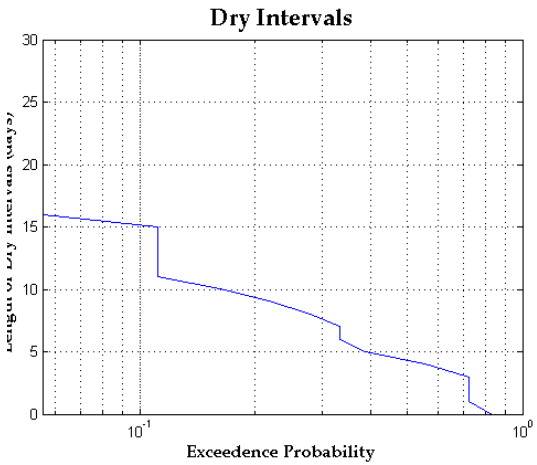
- August

	30 min	24 h
P_dry	0.99	0.88





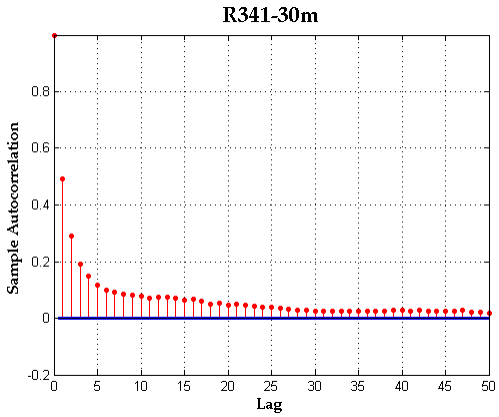


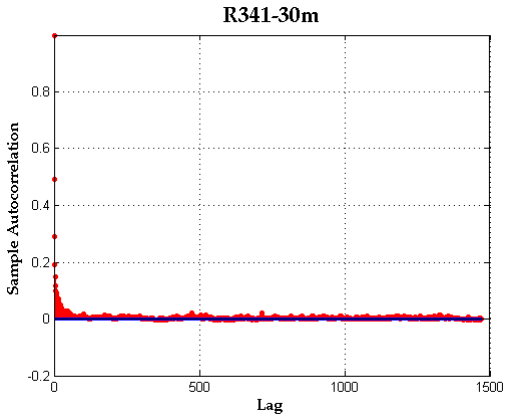


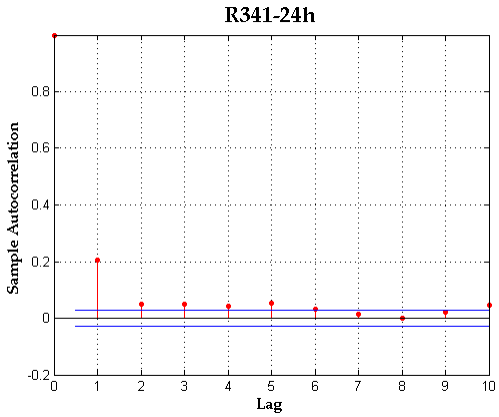
Dependence structure

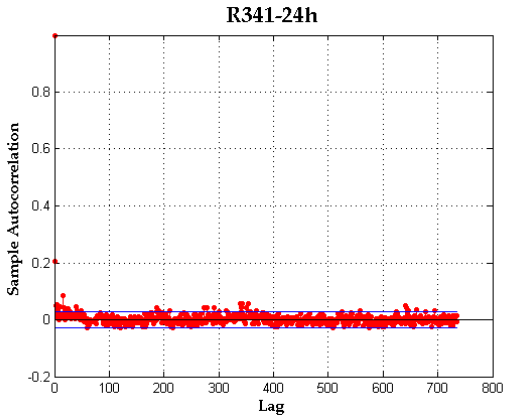
- 30 minutes
  - 1 day
  - 1 month
- 24 hours
  - 10 days
  - 2 years

R341-Barbarano









Information Entropy

- Shannon (1948) generalize the mathematical formulation of entropy in the field of information and uncertainty

$$H(X) = -\sum_i p_i \ln p_i \qquad \textit{Discrete R.V.}$$

$$H(X) = -\int_{-\infty}^{\infty} f(x) \ln f(x) dx \qquad \textit{Continuous R.V.}$$

- The entropy is a measure of the disorder associated to all probable states of a systems

- The entropy theory can be applied to the observed rainfall data to quantify the disorder in rainfall, in intensity and occurrence (Maruyama and Kawachi,1998, Kawachi et al.,2001, Maruyama et al., 2005) and to explain dry/wet sequences (Koutsoyiannis,2006)

Entropy-based Investigation of Rainfall Characteristics

- The recorded rainfall series of r_1, r_2,…r_k can be regarded as accumulated occurrence frequencies of unit rains for 1,2, ..rth intervals.
- The relative frequency can be calculated as:

$$p_i = \frac{r_{hi}}{R_d}$$

r_{hi} hourly (or sub-daily) rainfall

$$R_d = \sum_{i=1}^{24} r_{hi}$$ daily rainfall

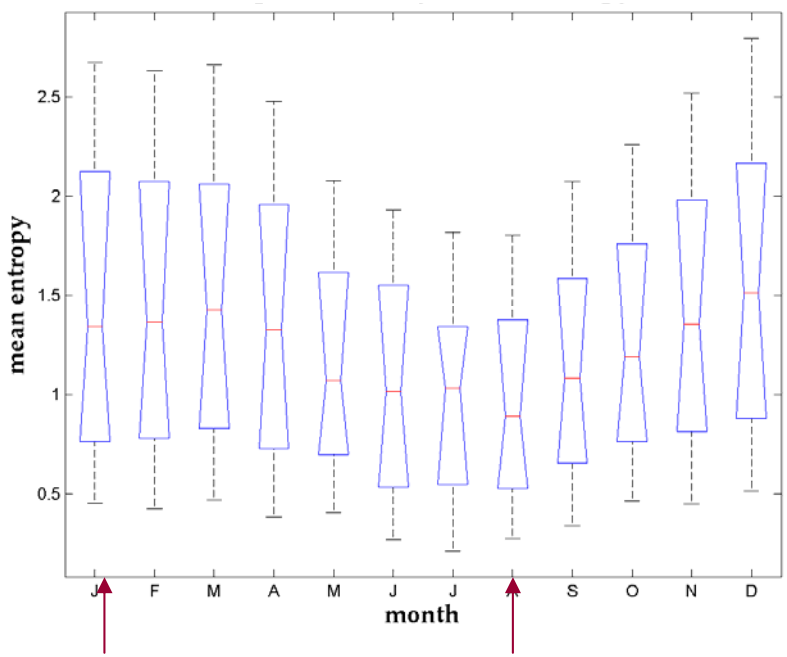
- The relative frequency p_i is an occurrence probability for the rainfall amount on the ith interval, and its distribution represents the uncertainty of rainfall occurrence.
- The daily Apportionment Entropy (AE) can be calculated as:

$$AE = -\sum_{i=1}^{24}\left(\frac{r_{hi}}{R_d}\right)\log_2\left(\frac{r_{hi}}{R_d}\right) = -\sum_{i=1}^{24} p_i \log_2 p_i = H$$

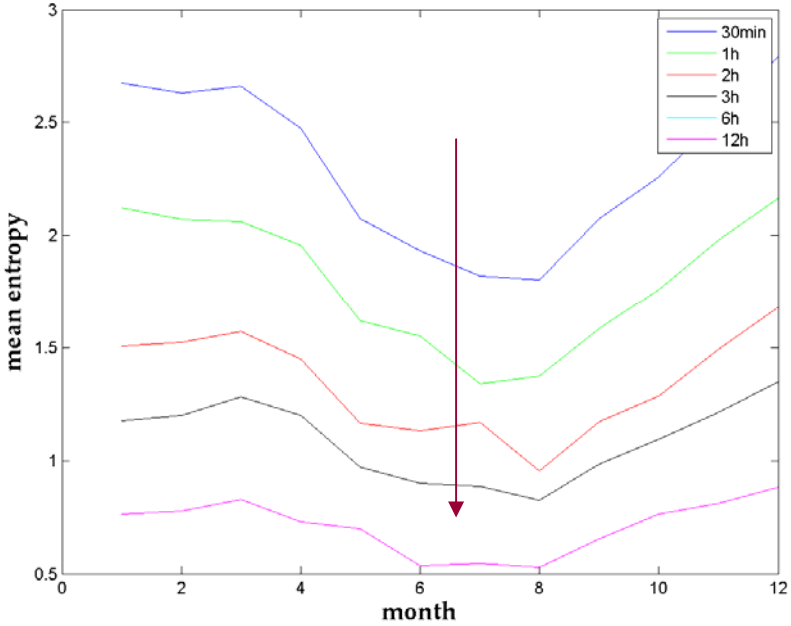
Entropy of the studied area

- Daily Apportionment entropy

Monthly variability

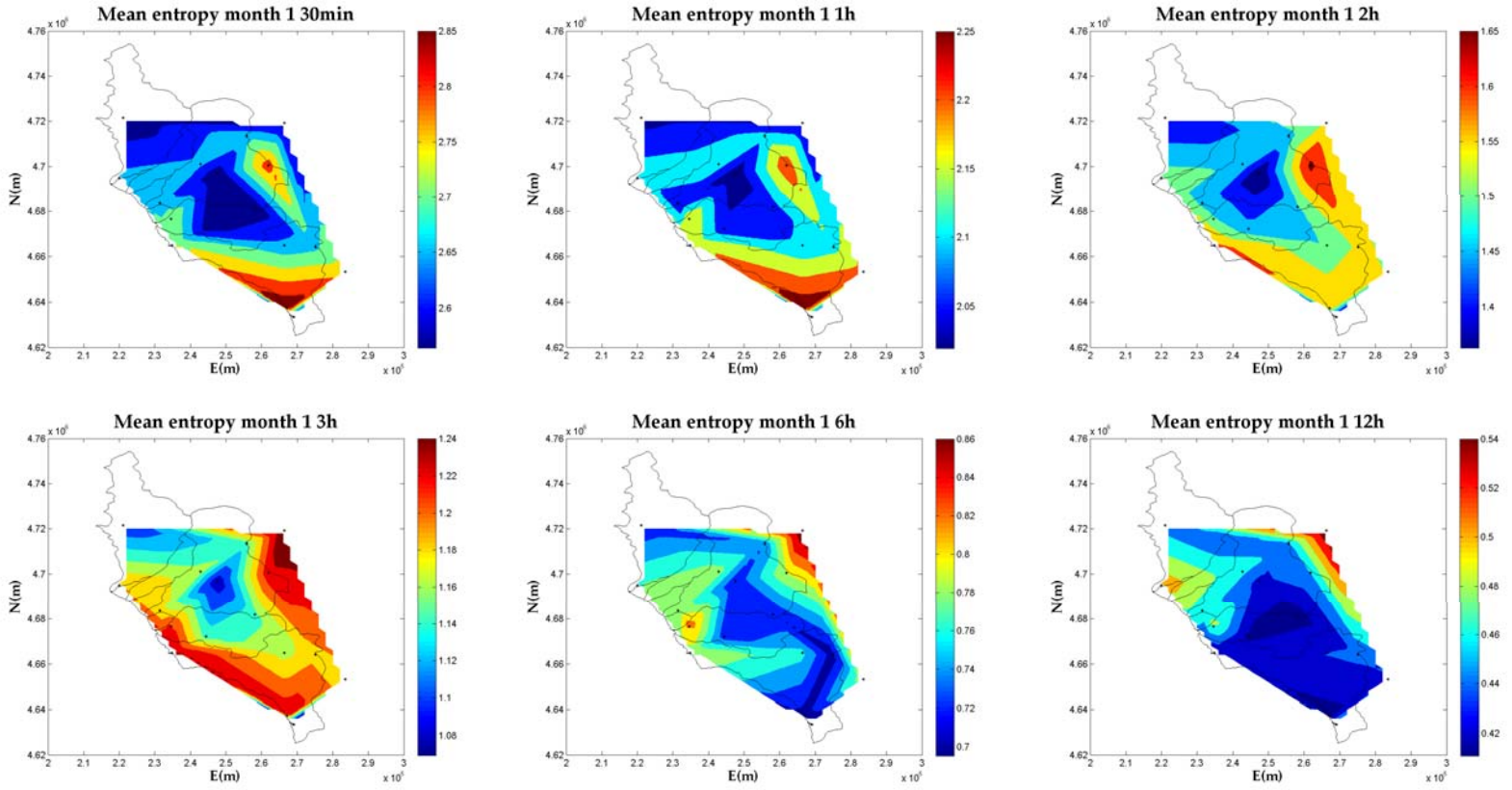


Time scale aggregation variability



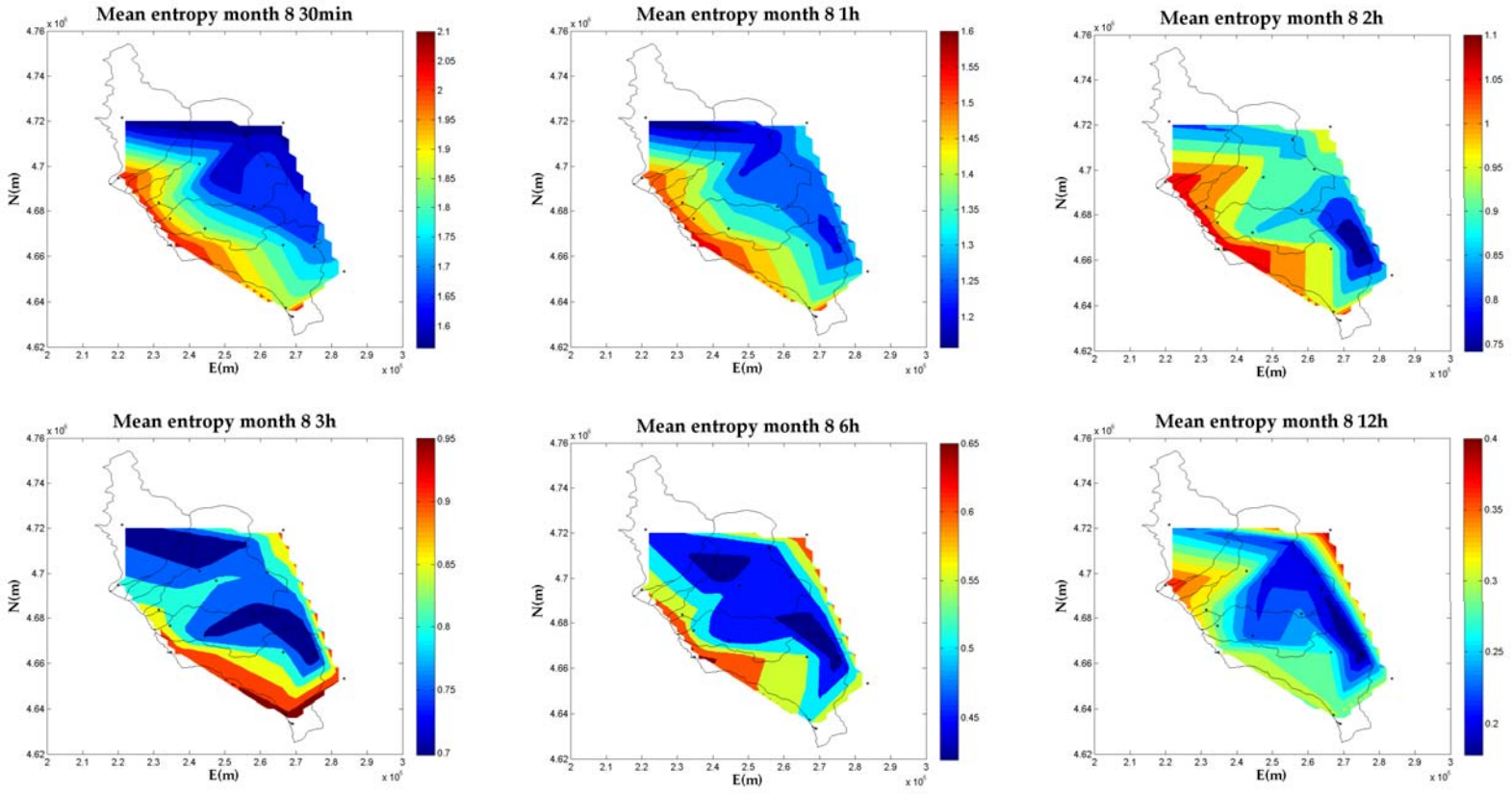
Time scale variability

- January



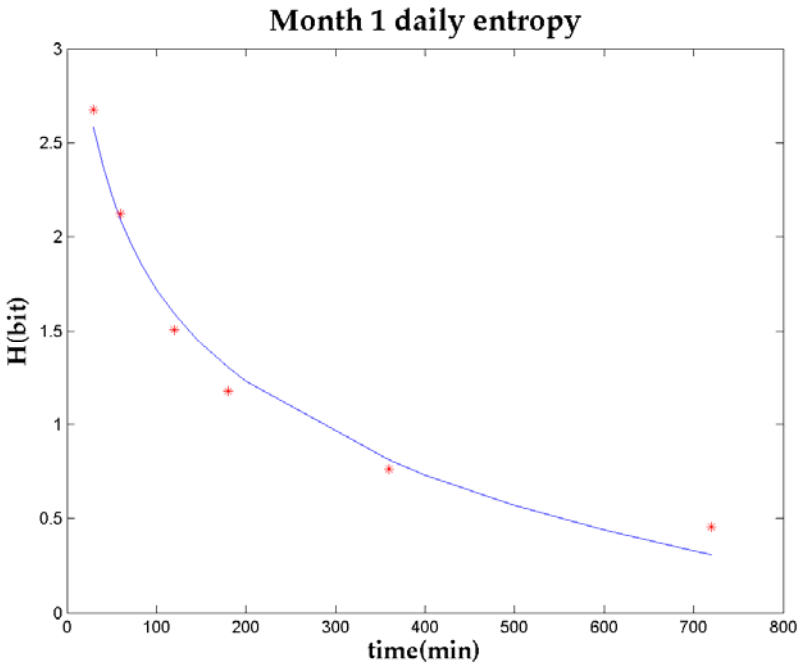
Time scale variability

- August



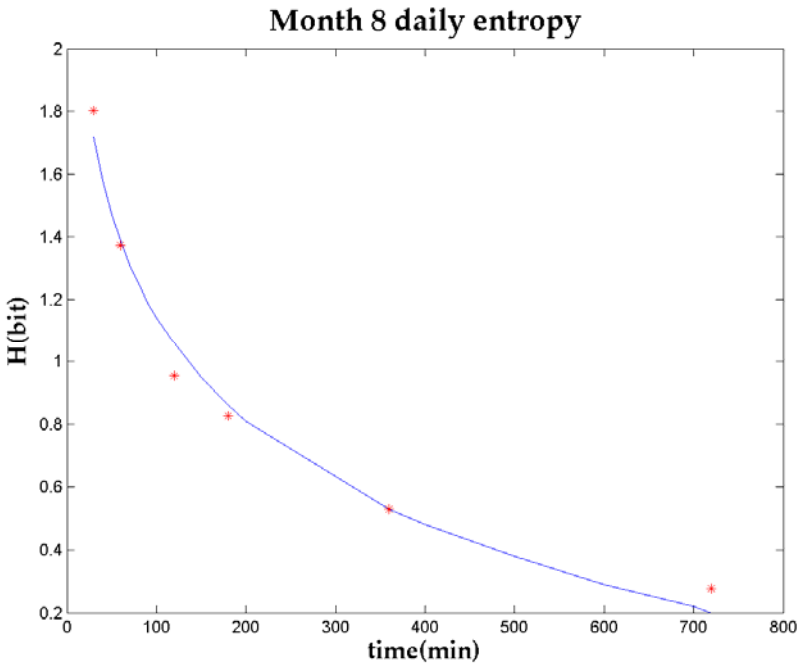
Regression

- January



$$H_{daily} = -0.714 \log_2(time) + 5.0109$$

- August



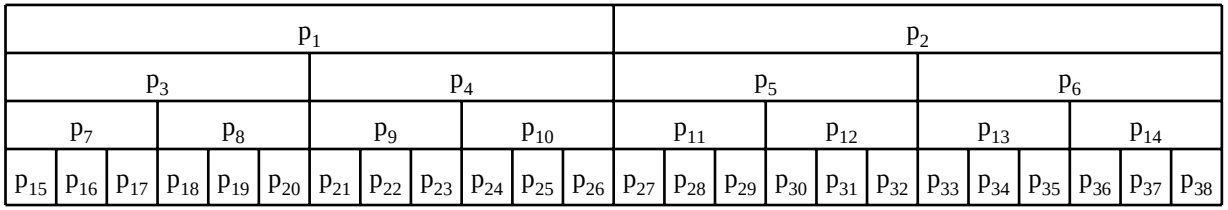
$$H_{daily} = -0.476 \log_2(time) + 3.3359$$

Disaggregation scheme

- A function $H_n(p_1 \ldots p_n)$ can be broken down in many different ways, relating it to the lower order functions by a large number of equations. The H_n are additive (Jaynes, 2003):

$$H(p_1,\ldots,p_n) = H(w_1,\ldots w_k) + w_1 H\left(\frac{p_1}{w_1},\ldots\frac{p_k}{w_1}\right) + w_2 H\left(\frac{p_{k+1}}{w_2},\ldots\frac{p_{k+m}}{w_2}\right) + \ldots$$

1 day



- From 24 to 12 hours:

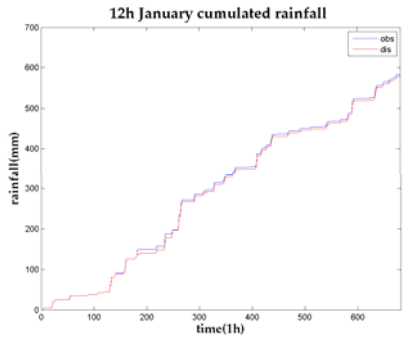
$$\begin{cases} H(p_1, p_2) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 \\ p_1 + p_2 = 1 \end{cases}$$

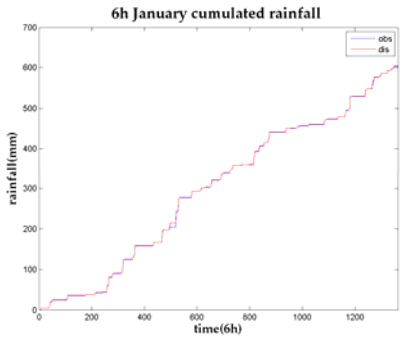
- From 12 to 6 hours:

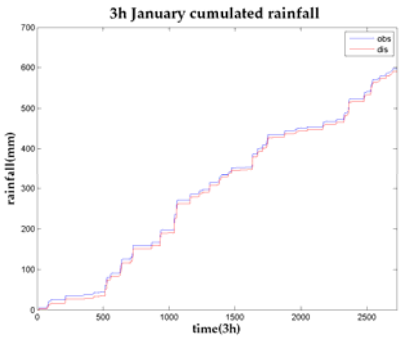
$$\begin{cases} H(p_1, p_5, p_6) = H(p_1, p_2) + p_2 H\left(\frac{p_5}{p_2}, \frac{p_6}{p_2}\right) \\ H(p_2, p_3, p_4) = H(p_1, p_2) + p_1 H\left(\frac{p_3}{p_1}, \frac{p_4}{p_1}\right) \\ p_5 + p_6 = p_2 \\ p_3 + p_4 = p_1 \end{cases}$$

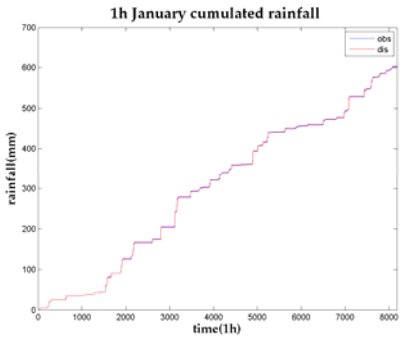
Disaggregation results

- Cumulated rainfall
  - January

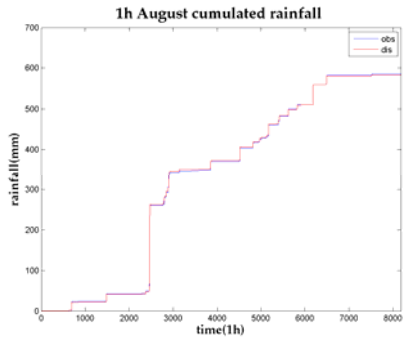


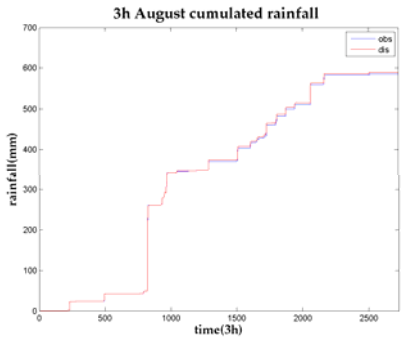


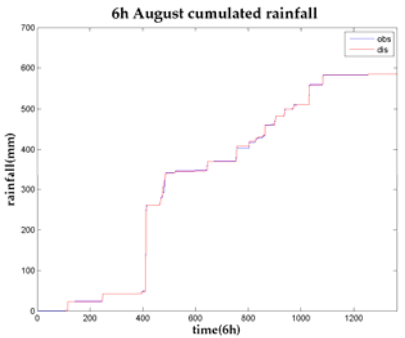


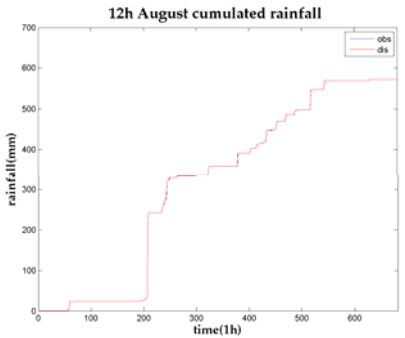


  - August



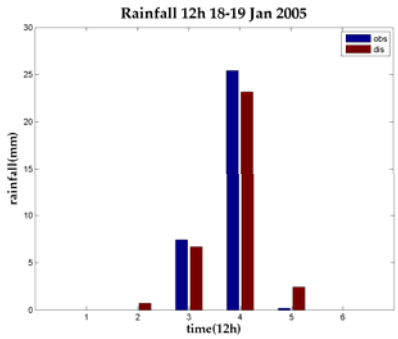


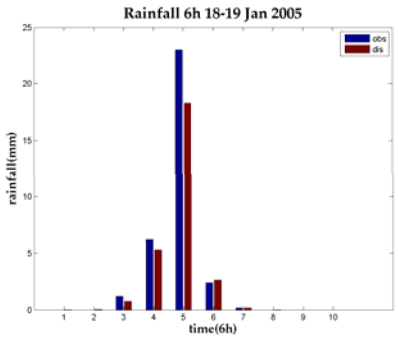


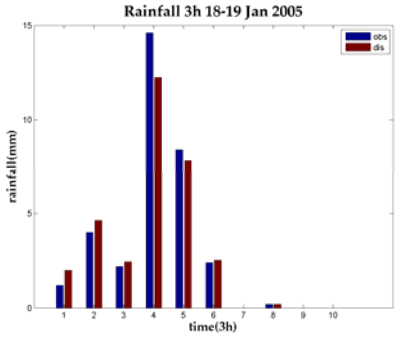


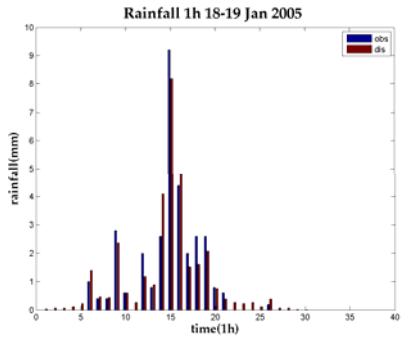
Disaggregation Results

- Rainfall event
  - January

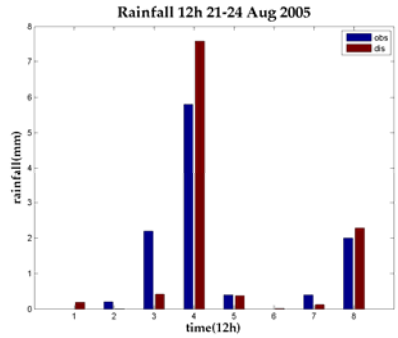


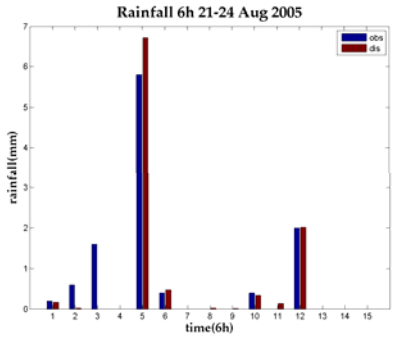


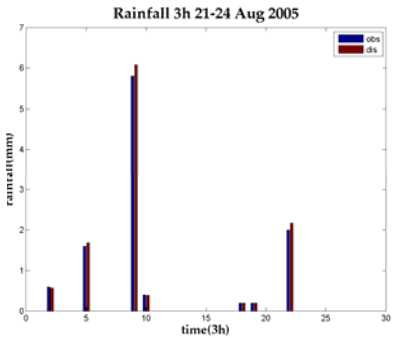


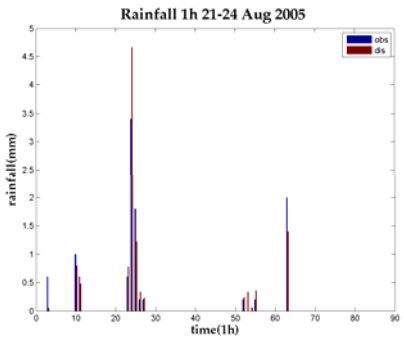


  - August









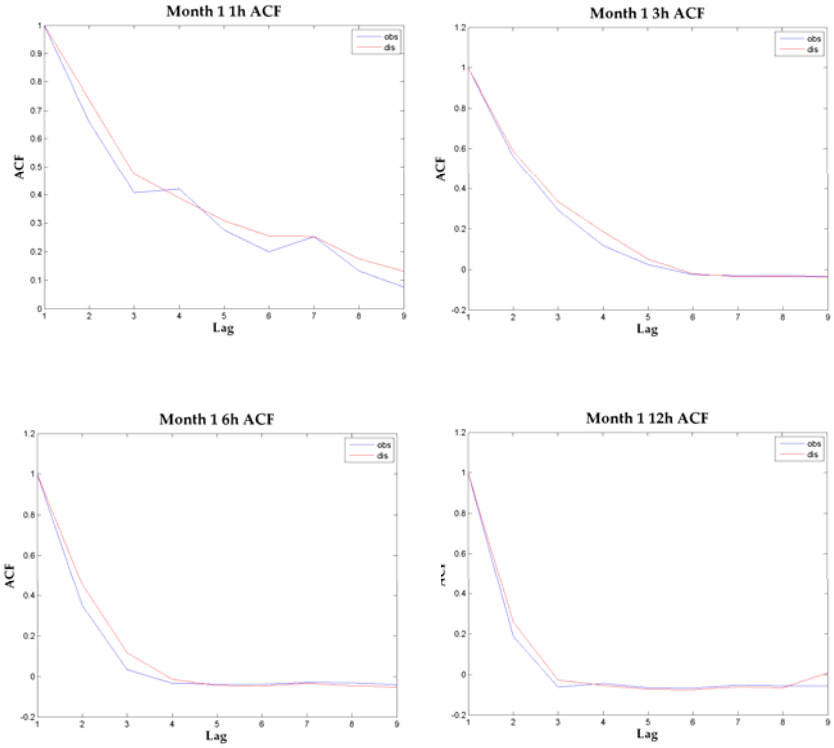
Disaggregation Results

❑ Statistical properties

	1h		3h		6h		12h	
	obs	**dis**	**obs**	**dis**	**obs**	**dis**	**obs**	**dis**
max	63.1	72.24	80.5	97.6	157.8	161.9	183.2	182.69
mean	3.59	2.42	5.57	4.64	6.8	3.78	8.29	5.02
var	56.6	43.96	158.2	138.81	337.85	215.91	533.91	342.38
stdec	7.52	6.63	12.58	11.78	18.38	14.69	23.11	18.5
CV	2.1	2.74	2.26	2.54	2.7	3.89	2.79	3.69
skew	4.3	6.88	4.27	5.45	6.71	9.12	6.5	8.06
kurt	28.37	63.41	23.8	38.33	54.53	96.85	49.14	76.27
iq	3.1	1.53	4.05	2.79	8.4	1.33	9.47	2.07
p_dry	0.98	0.97	0.96	0.95	0.94	0.9	0.9	0.83

January

❑ Correlation structure

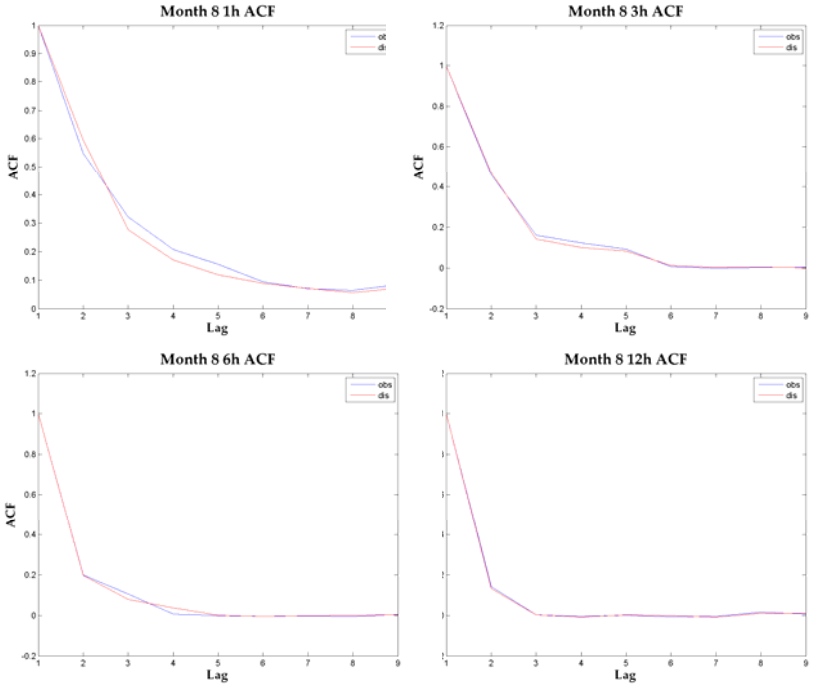


Disaggregation Results

❑ Statistical properties

	1h		3h		6h		12h	
	oss	**dis**	**oss**	**dis**	**oss**	**dis**	**oss**	**dis**
max	14	13.47	21	20.14	31.6	24.2	33.8	30.6
mean	0.93	0.63	1.7	1.53	2.52	1.74	3.47	2.43
var	2.34	1.52	9.1	7.31	19.84	12.71	32.09	23.83
stdec	1.53	1.23	3.02	2.7	4.45	3.56	5.67	4.88
CV	1.65	1.95	1.77	1.76	1.76	2.05	1.63	2.01
skew	3.27	4.5	3.18	3.33	3.34	3.36	2.83	3.19
kurt	17.23	30.84	15.03	16.27	16.52	15.98	11.97	14.17
iqrange	0.9	0.51	1.6	1.42	2.6	1.93	3.9	1.88
p_dry	0.92	0.88	0.87	0.86	0.82	0.74	0.75	0.65

❑ Correlation structure



August

Conclusions

- The disaggregation model is based on Apportionment Entropy, giving the definition of pi as occurrence probability for the rainfall amount on the ith interval
  - It preserves the rainfall amounts
  - Good event representation, both in winter and summer
- For the chosen additive properties there are some trace values, that influenced both skewness and proportion of dry intervals
  - Need for trace values correction
- More analysis are to come, including using disaggregated data set in rainfall runoff modelling

References

- Jaynes, E.T. Probability Theory. The Logic of Science. *Cambridge University Press*. Cambridge, 2003.
- Kawachi, T., Maruyama, T., and Singh, V.P. Rainfall entropy for delineation of water resources zones in Japan. *Journ. of Hydrol.*, 246, 36-44, 2001.
- Koutsoyiannis, D., An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, *Water Resources Research*, 42 (1), W01401, 2006.
- Maruyama, T., and Kawachi, T. Evaluation of rainfall characteristics using entropy. *Journal of Rainwater Catchment Systems*. 4(1), 7-10, 1998.
- Maruyama, T., Kawachi, T., Singh, V.P. Entropy-based assessment and clustering of potential water resources availability. *Journ. of Hydrol.*, 309, 104-113, 2005.
- Shannon, C.E. A Mathematical Theory of Communication. *Bell System Technical Journal*. 1948.