

ISPRA - Roma 9-10 Luglio 2015
Idrologia Operativa - Workshop nazionale

L'Idrologia operativa e la comunità scientifica

Contributi e spunti di discussione sul tema:
**La misura delle grandezze idrologiche e la
statistica nell'idrologia operativa**

Roberto Deidda

Università di Cagliari

Dipartimento di Ingegneria Civile, Ambientale e Architettura

Contributi e spunti di discussione sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

- 1) Recupero di misure storiche:
reinterpretazione delle striscette cartacee contenenti i segnali pluviografici e idrografici
- 2) Caratterizzazione statistica degli eventi estremi e curve di possibili pluviometrica:
dagli approcci regionali a quelli geostatistici
- 3) Caratterizzazioni statistiche basate sulle osservazioni continue:
POT, invarianza di scala, etc.

Applicazioni in progetti prototipali: letture di serie continue ad alta risoluzione temporale 5-min (1-min) di serie storiche di precipitazione

Deidda, R., G. Mascaro, E. Piga, G. Querzoli (2007),
An automatic system for rainfall signal recognition from tipping bucket gage strip charts. *Journal of Hydrology*, **333**, 400-412

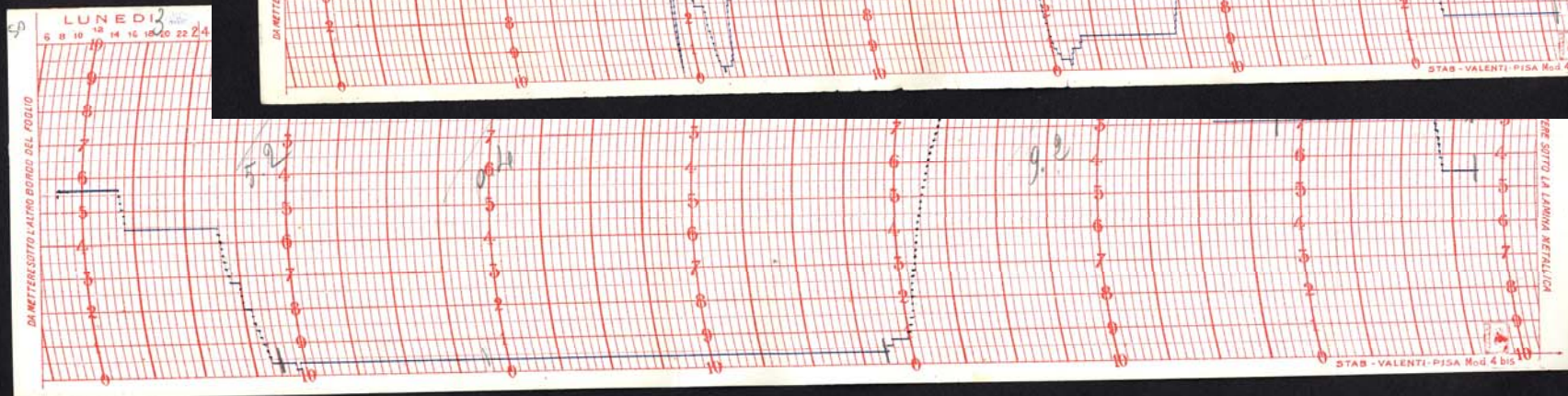
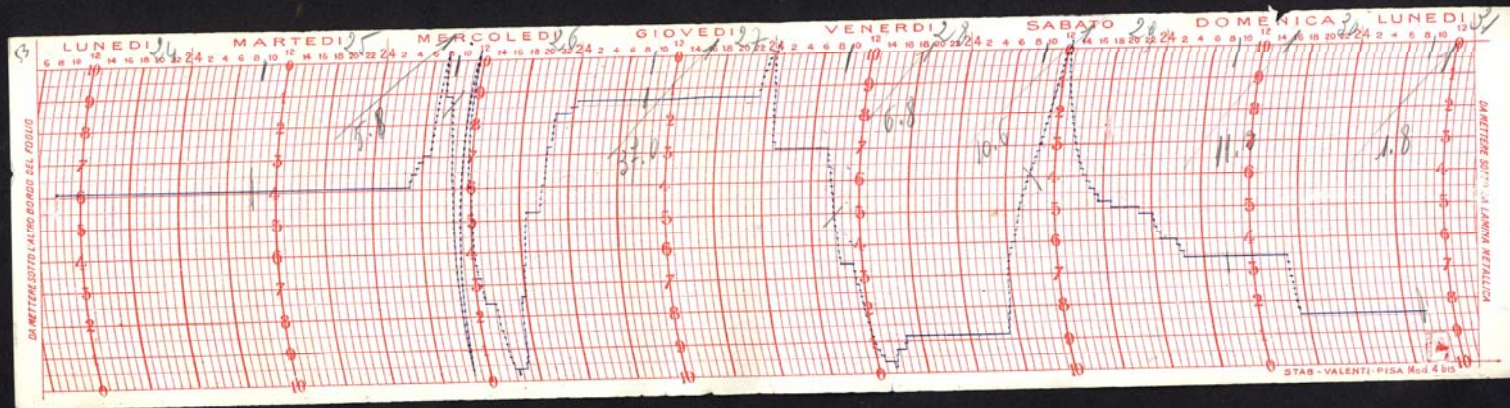
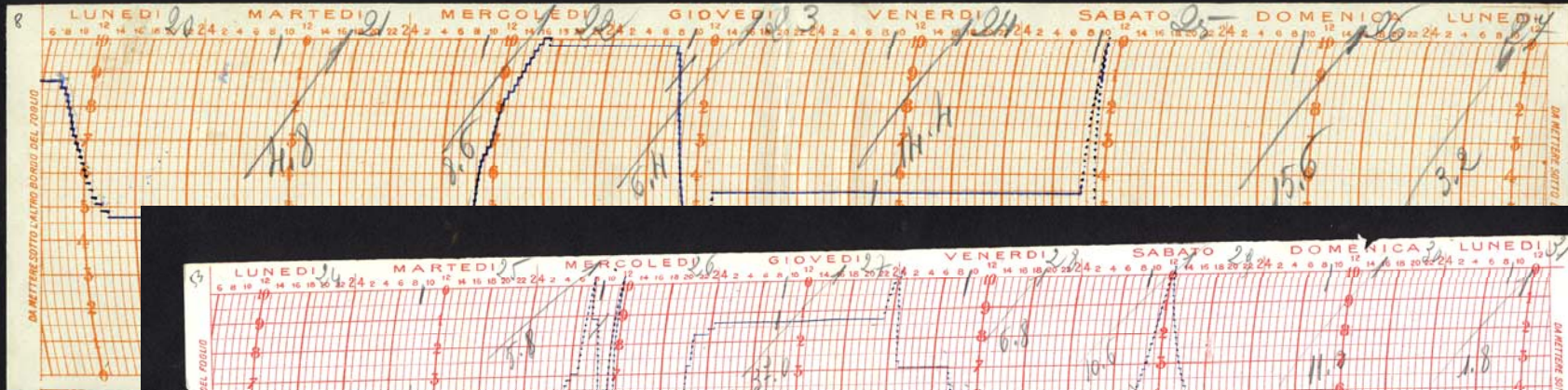
- 8' 000 strisce cartacee (circa 160 anni stazione)
- 4 stazioni nella Regione Veneto
- ISPRA, con finanziamento progetto FORALPS

- 16' 000 strisce cartacee (circa 320 anni stazione)
- 5 stazioni nella Regione Sardegna
- ARPA Sardegna, con finanziamento progetto RES-MAR

- 114' 000 strisce cartacee (circa 2200 anni stazione)
- Varie stazioni nella Regione Sardegna
- Agenzia del Distretto Idrografico della Regione Sardegna

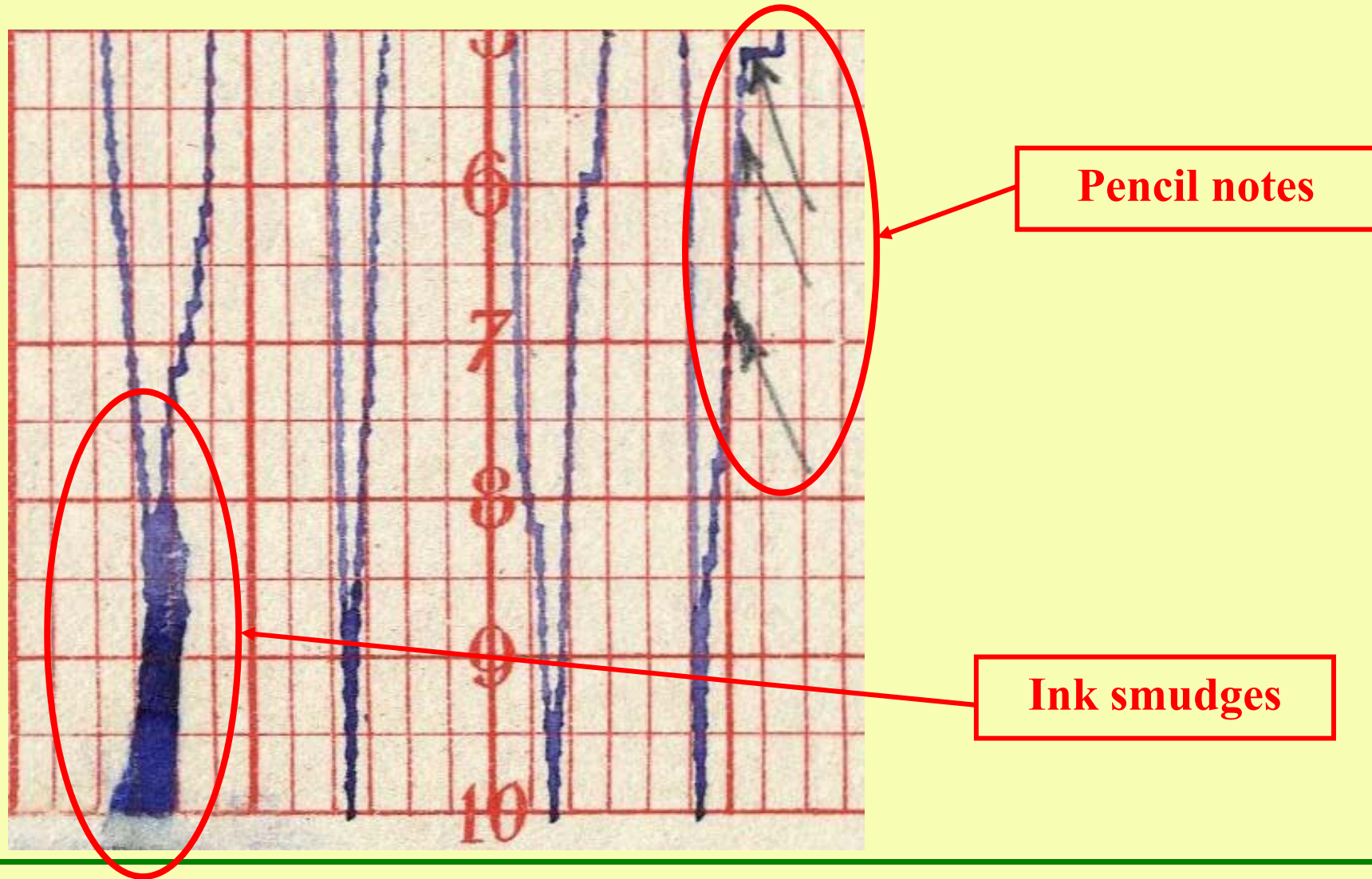
R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

Strip charts



R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

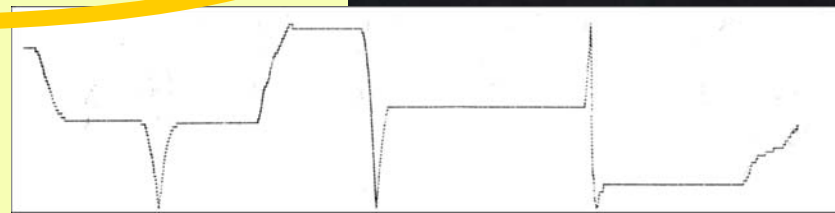
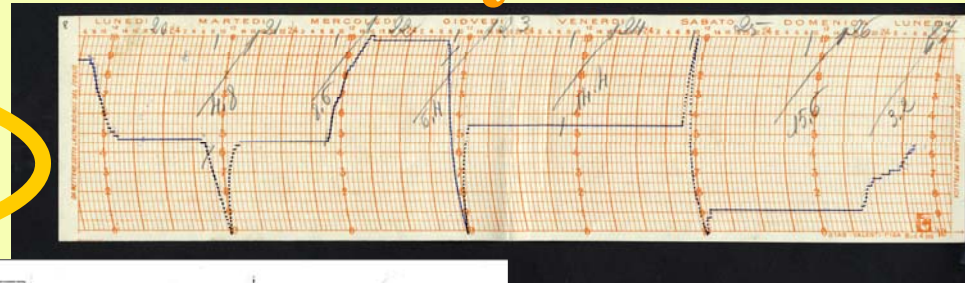
Strip charts - problems



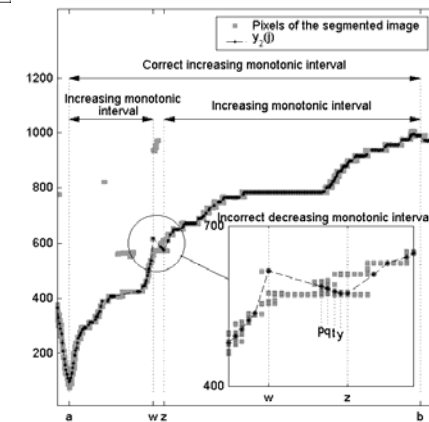
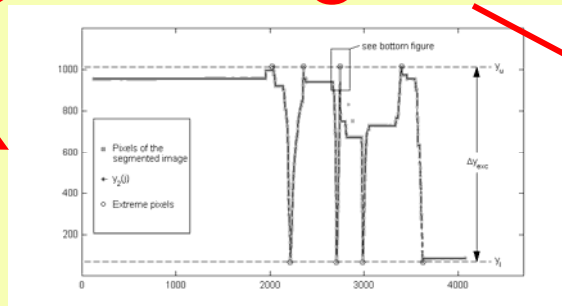
R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

Main modules of the system

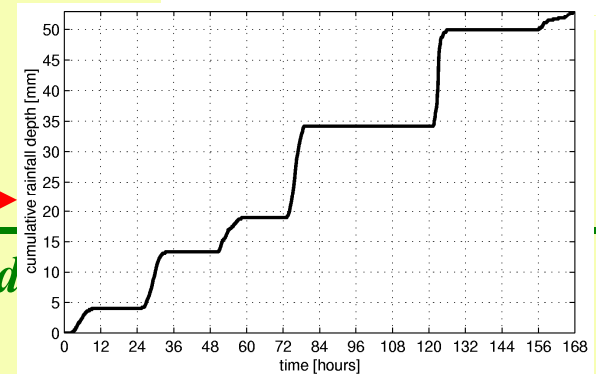
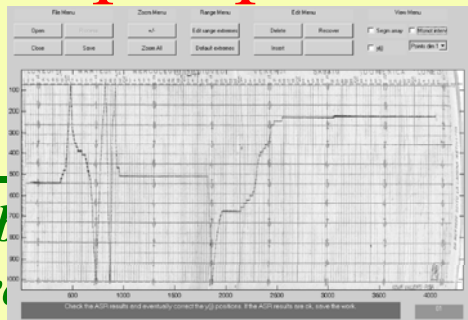
A. Image Segmentation



B. Automatic Signal Recognition



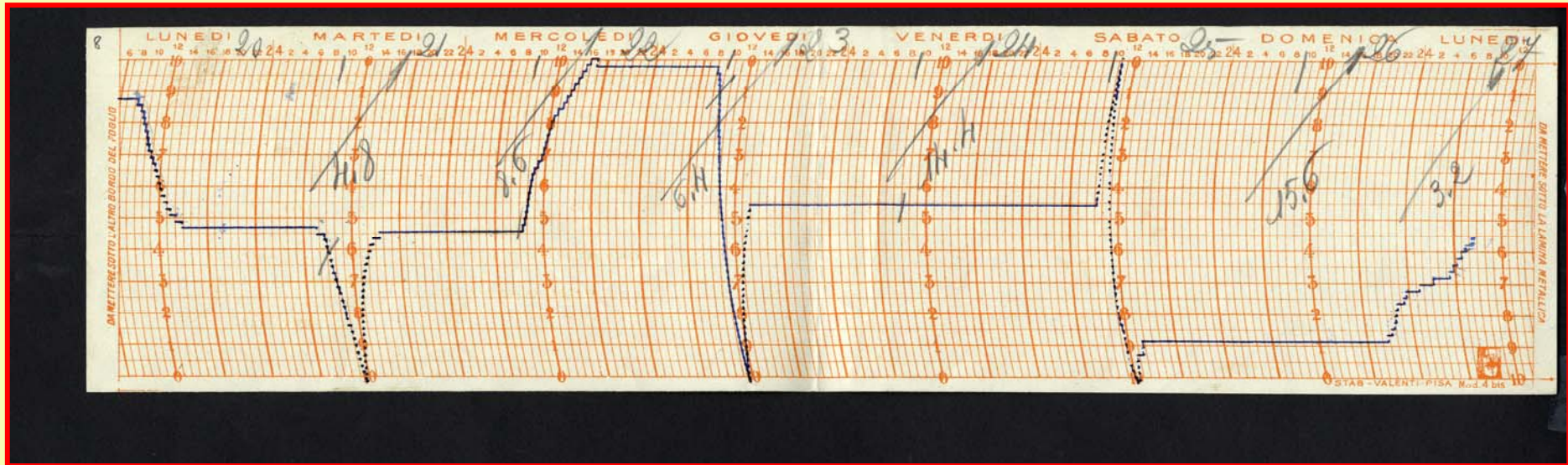
C. Interactive postprocessing



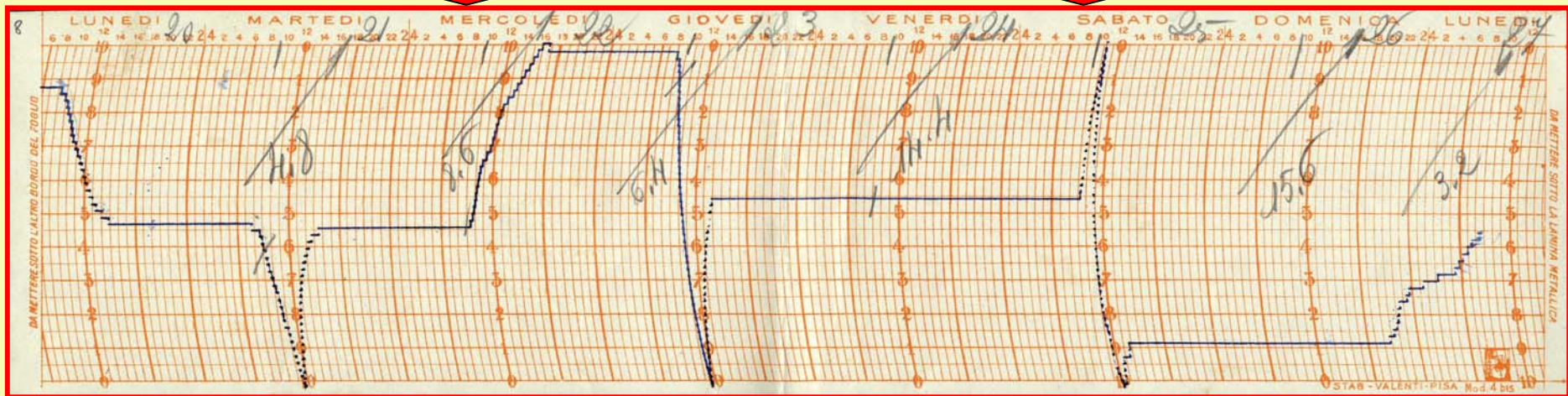
*R. Deidda – Contrib
la statistica nell'idr*

ira delle grand

A.1) Rotation Procedure



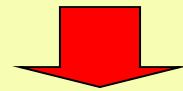
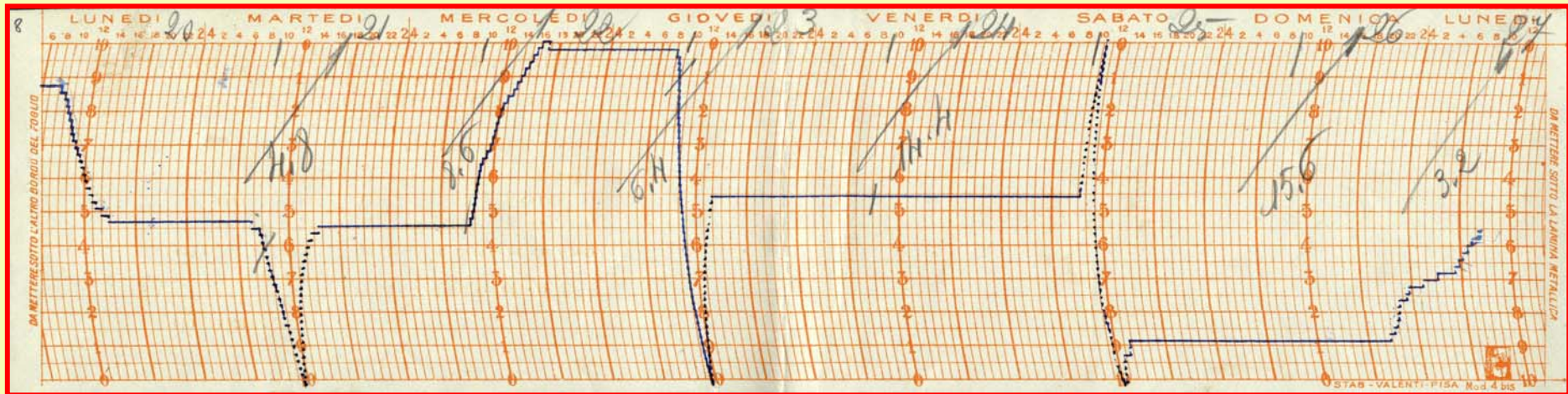
Rotation



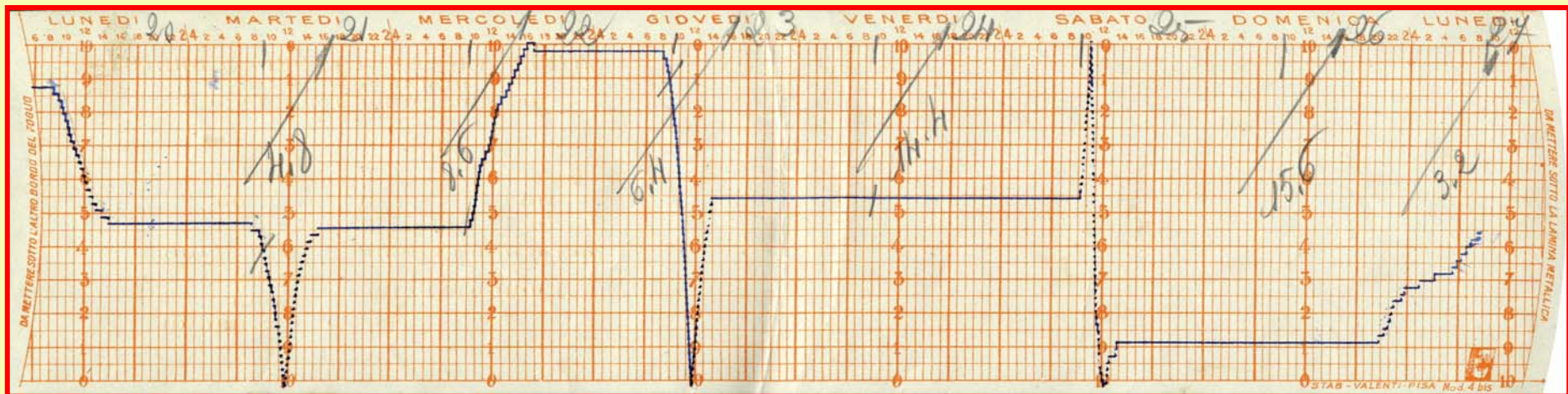
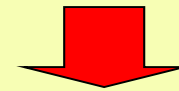
R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

A.2) Warping Procedure

Image Segmentation Module

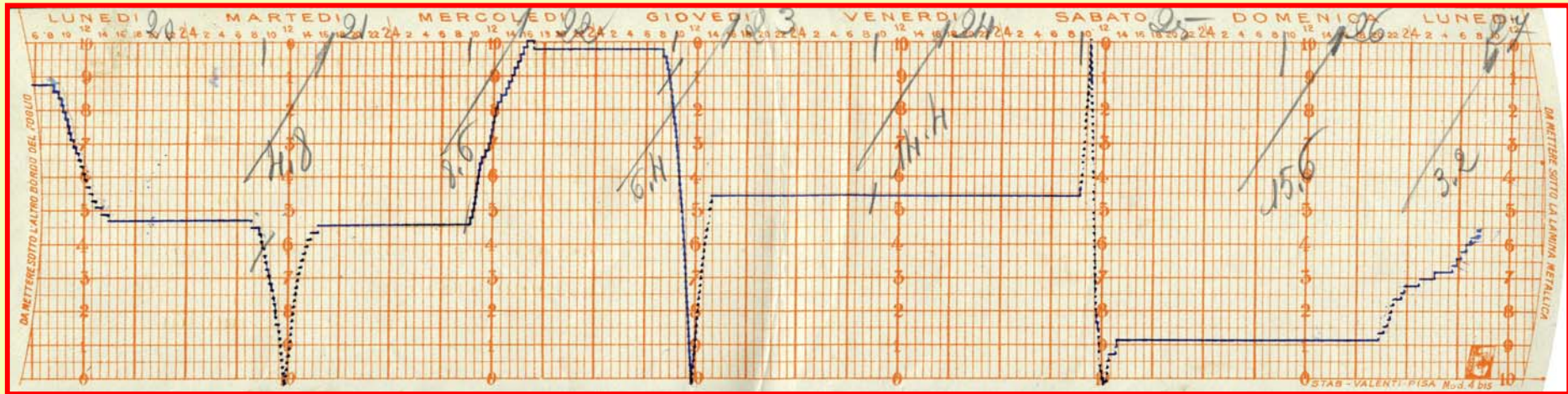


Warping

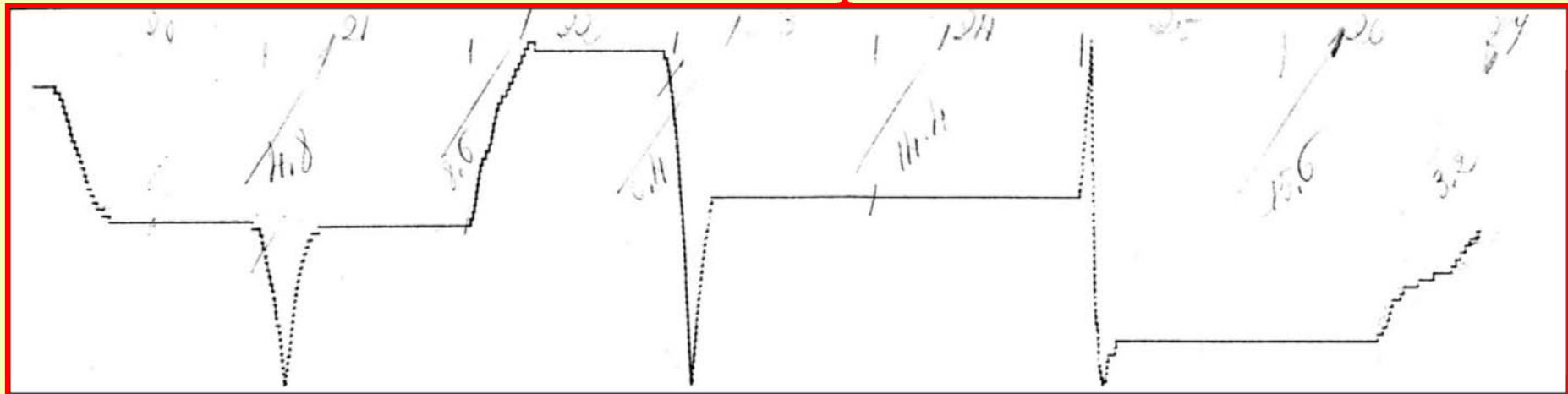


R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

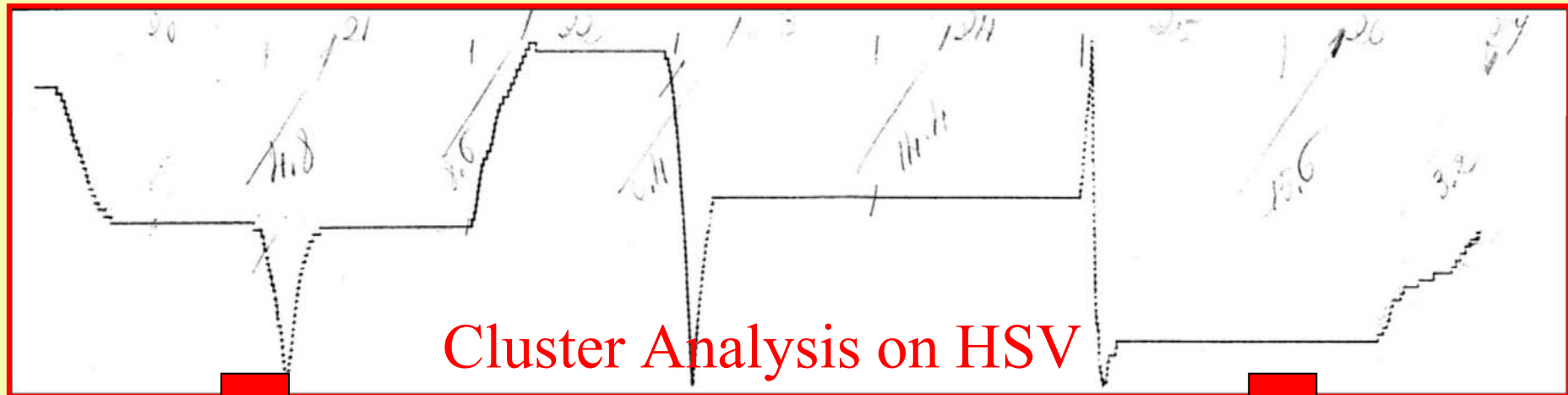
A.3) Thresholding Procedure



Thresholding on
RGB components



A.4) Cluster analysis on HSV space



(Hue, Saturation, Value) space

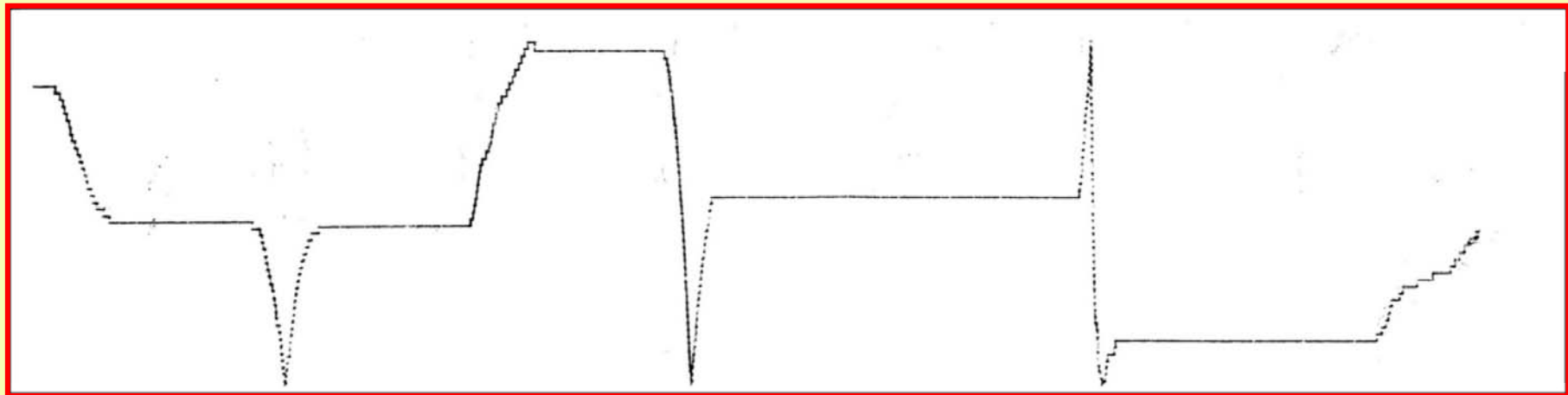
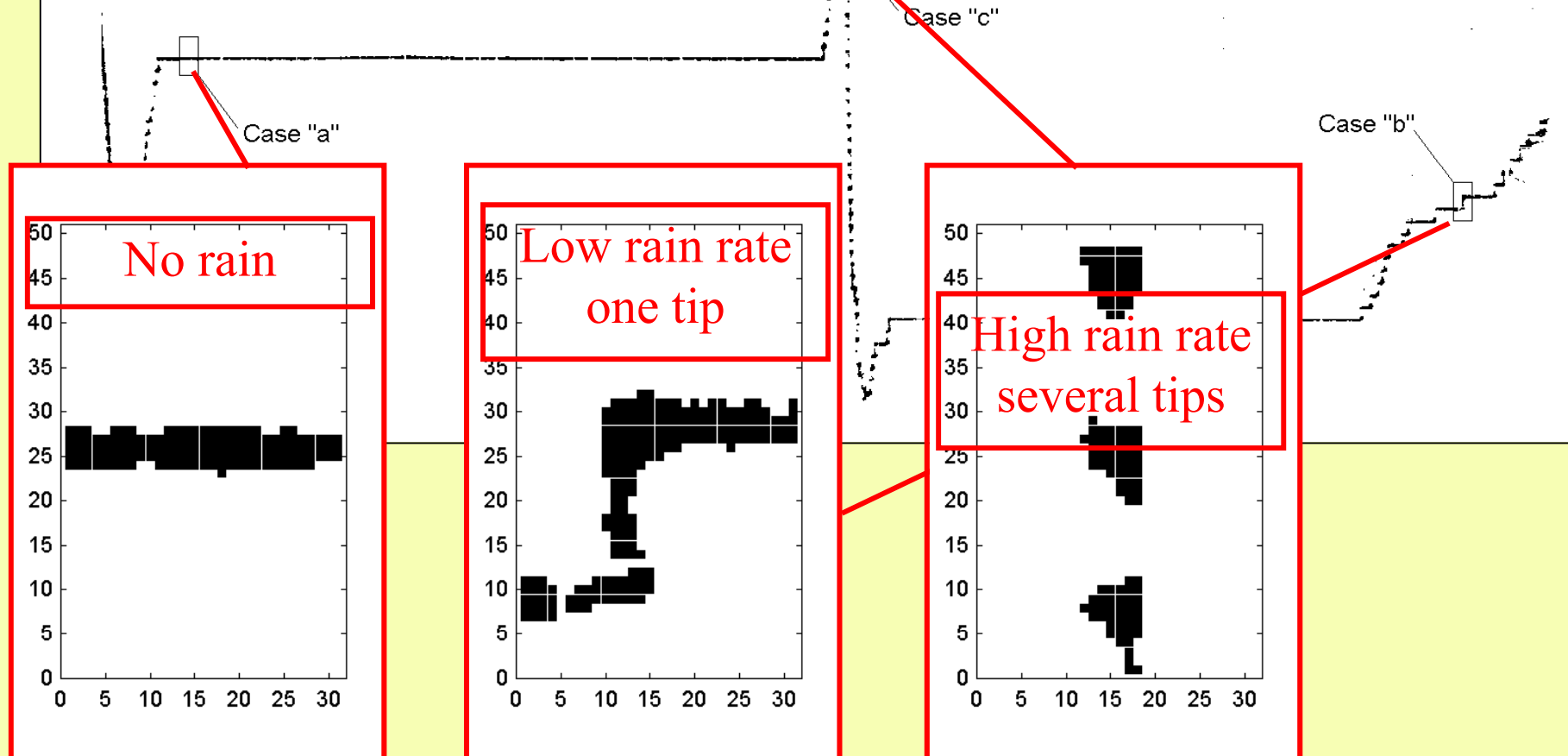


Image segmentation outcome: typical features

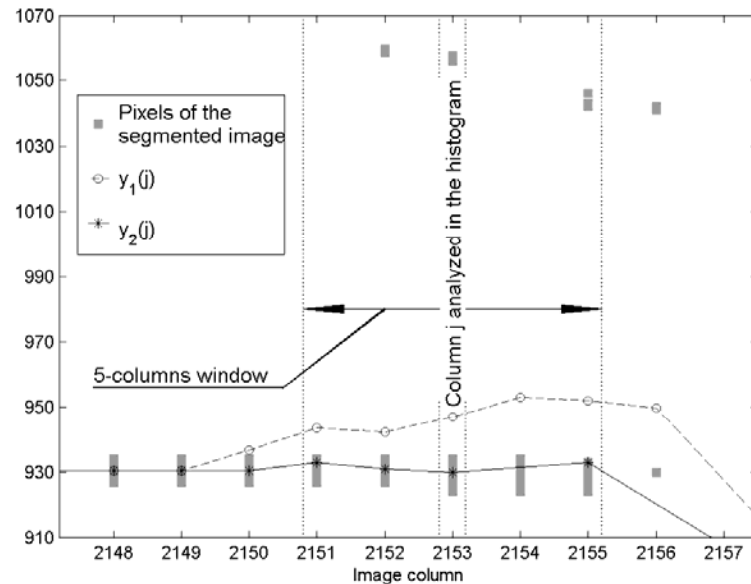


B.1) Robust Line Detection procedure

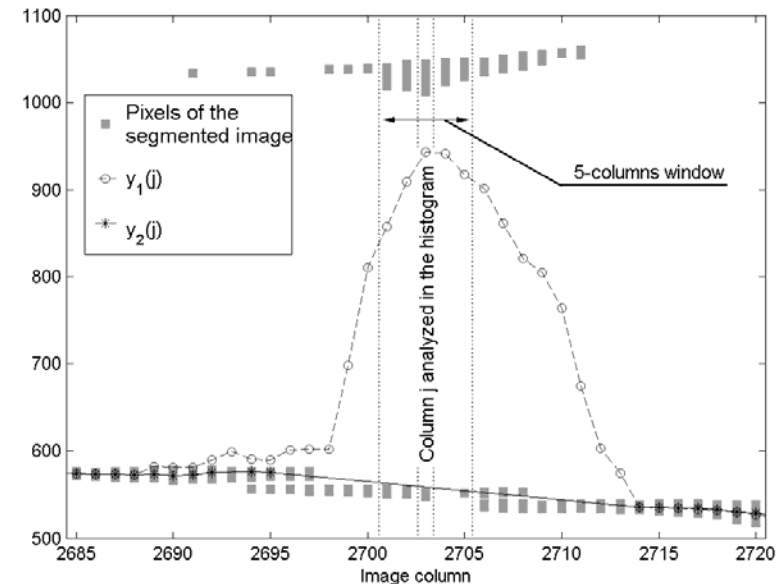
The **y-centroid** is a robust estimate of the plotline.

A double step computation reject most outliers:

1st unbounded + 2nd bounded (vertical) y-centroid computation



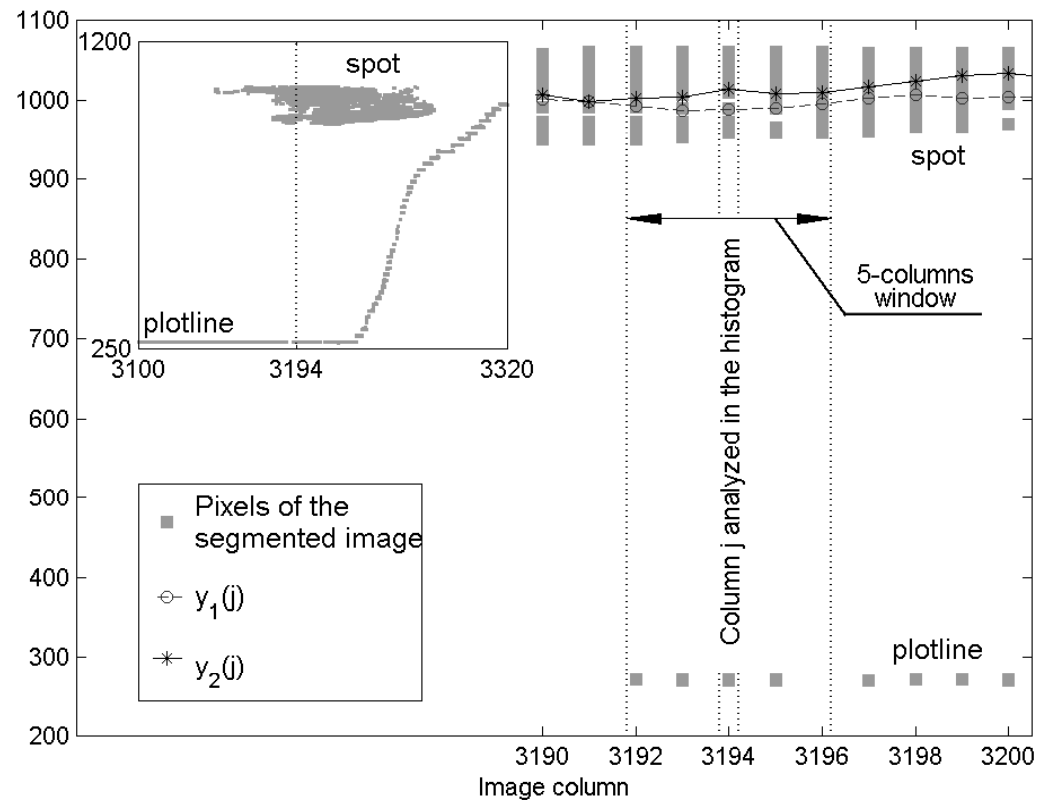
Small outliers:
signal correctly detected



Large outliers:
missing value assigned

B.1) Robust Line Detection procedure

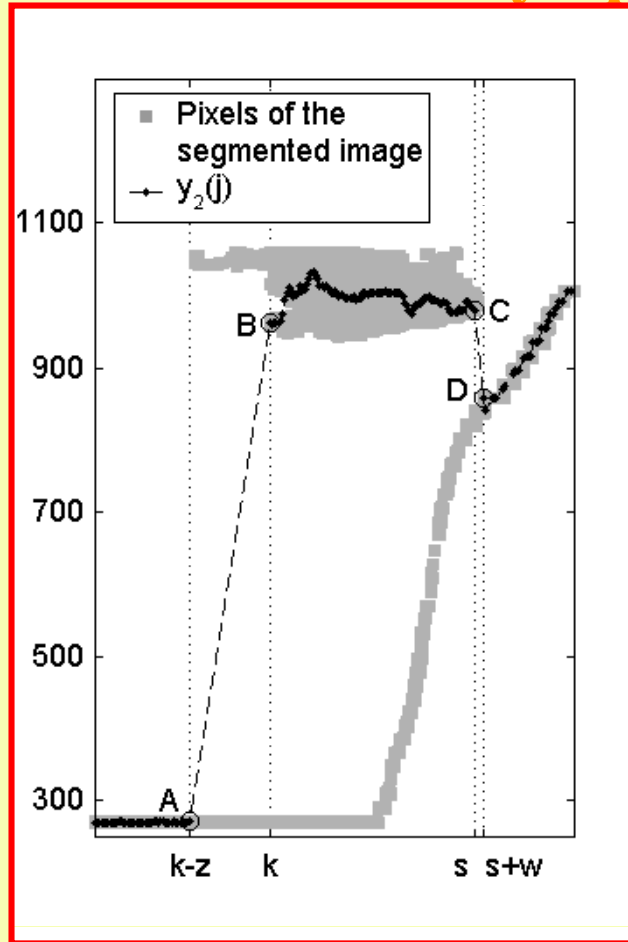
... may fail with very large smudges



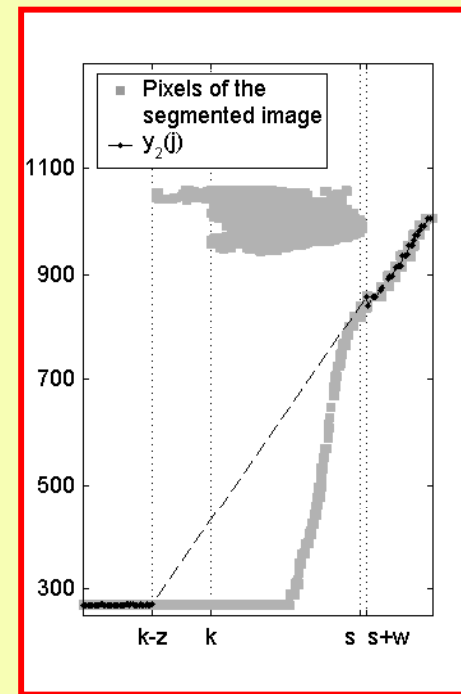
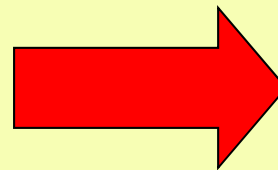
y-centroid results within the ink smudge

B.2) Spot Rejection procedure

Spots are detected by opposite & large Δy in close columns

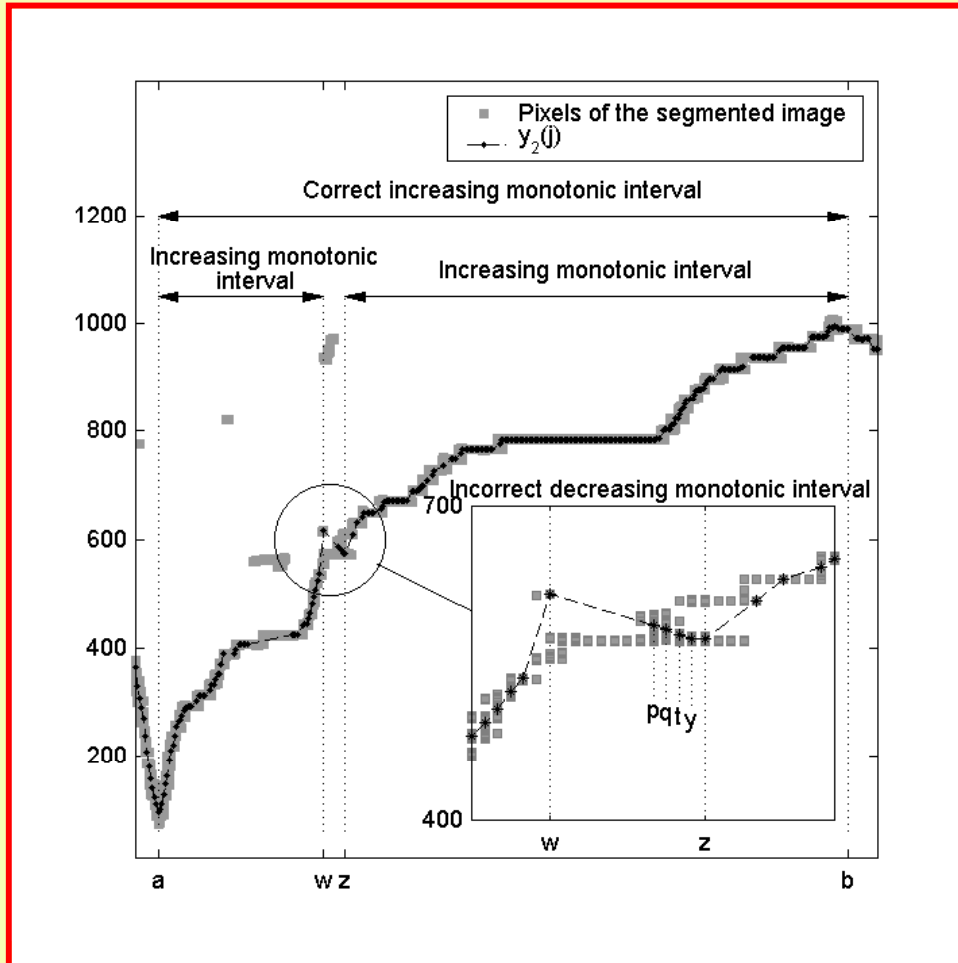


Missing values assigned to y-centroids from A to D



B.3) Monotonic Constraint procedure

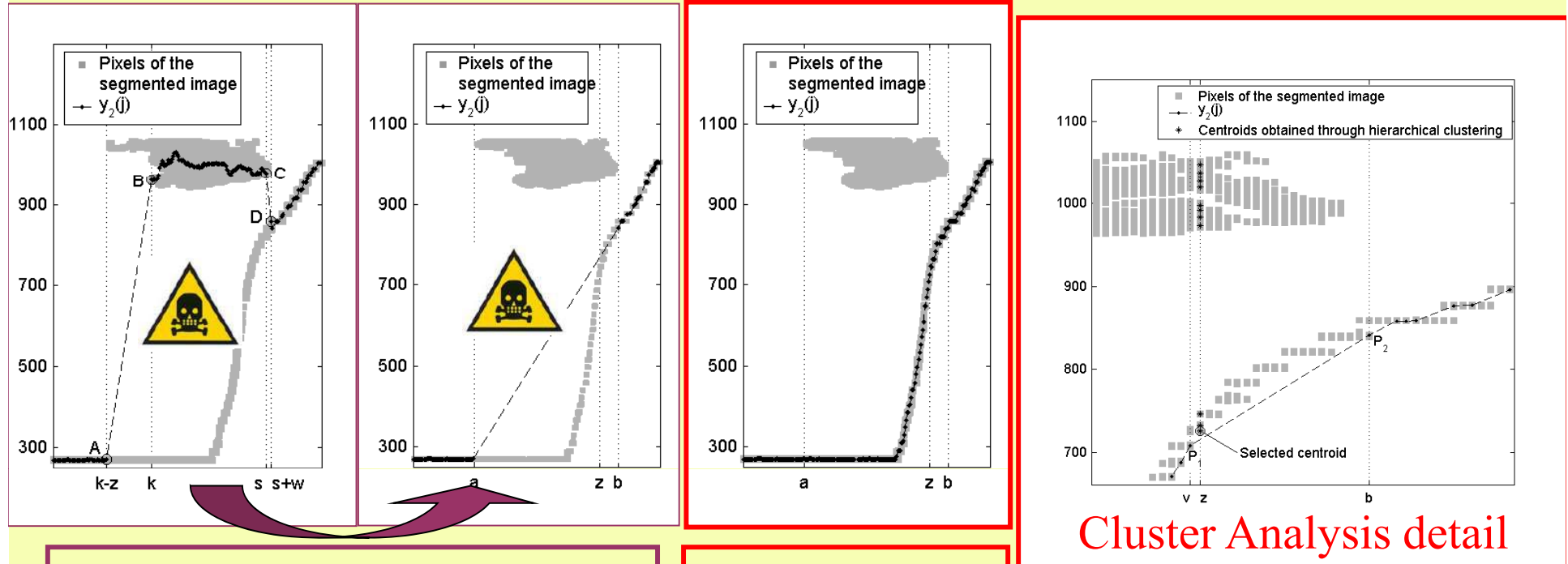
Spots close to the plotline are detected



Missing values assigned to y-centroids from W to Z

B.4) Signal Recovering procedure

Cluster analysis on the columns allows recovering



Missing values assigned by Spot Rejection procedure

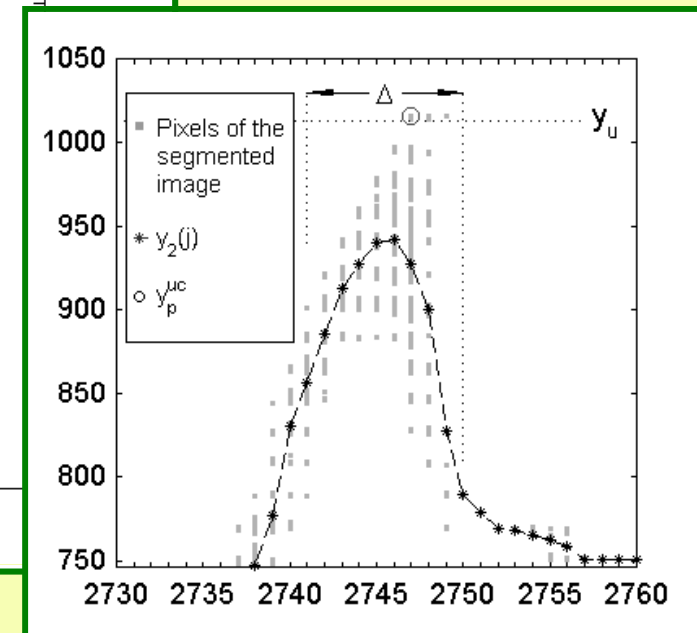
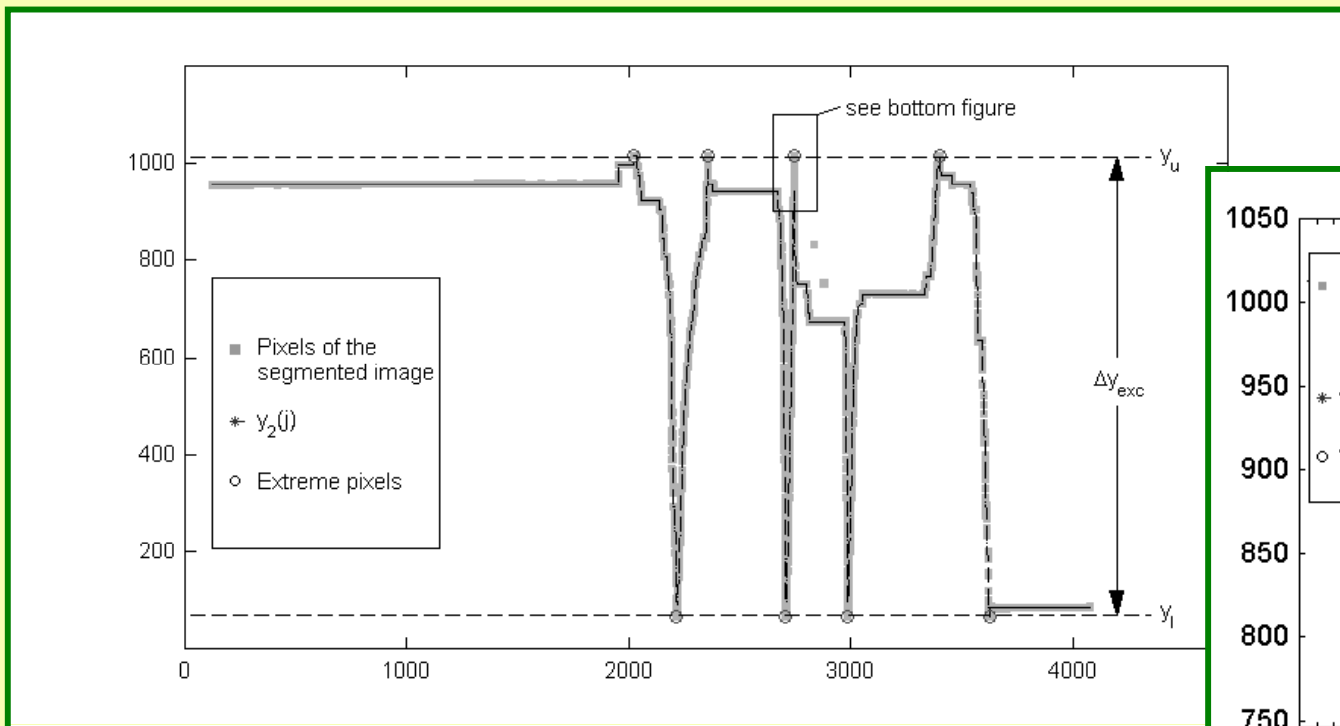
Recovered signal

Remaining y-centroid missing values, if any, are filled by interpolation

R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

B.5) Inversion Point Search procedure

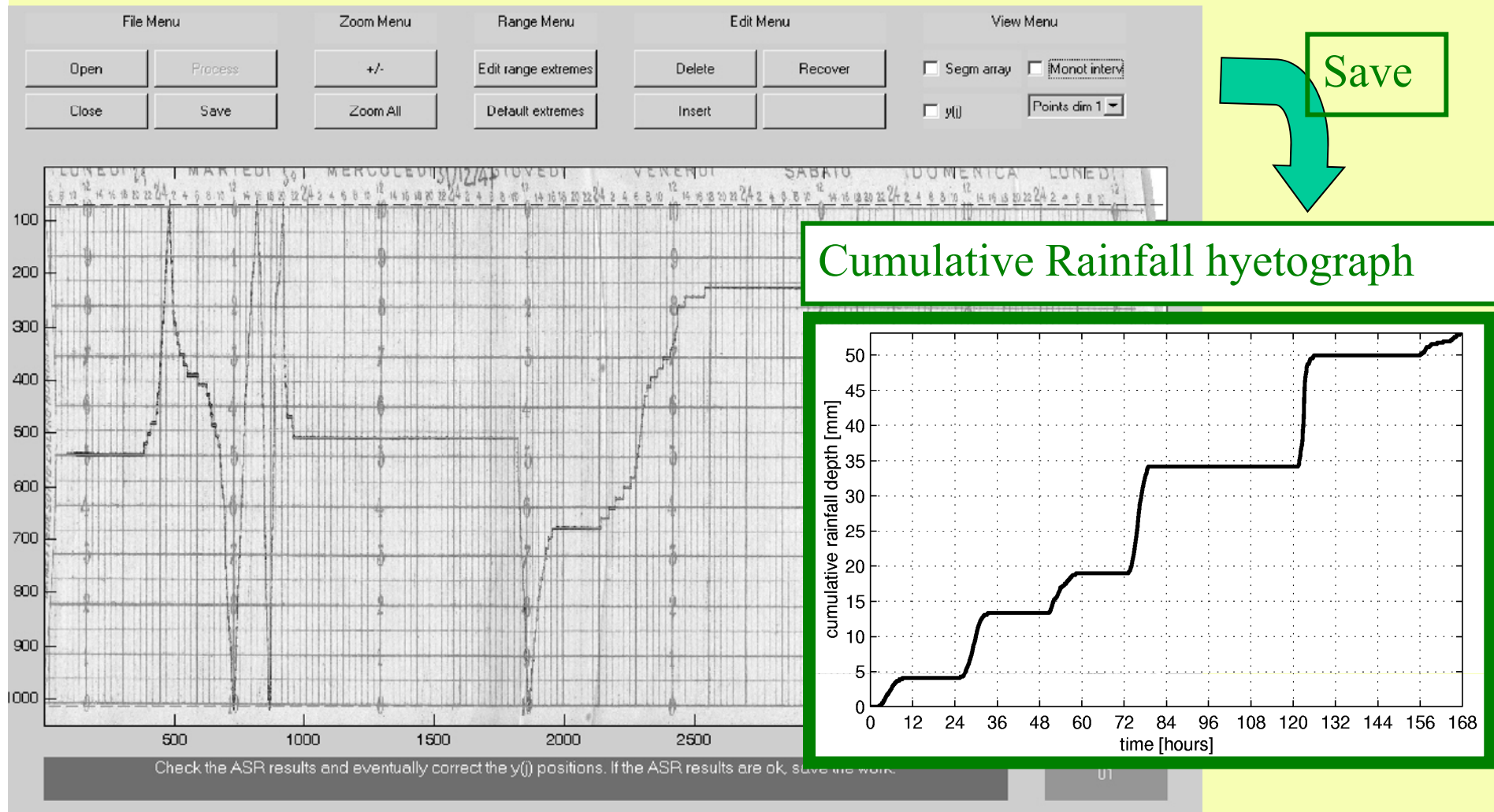
... cusps result often smoothed by y-centroids



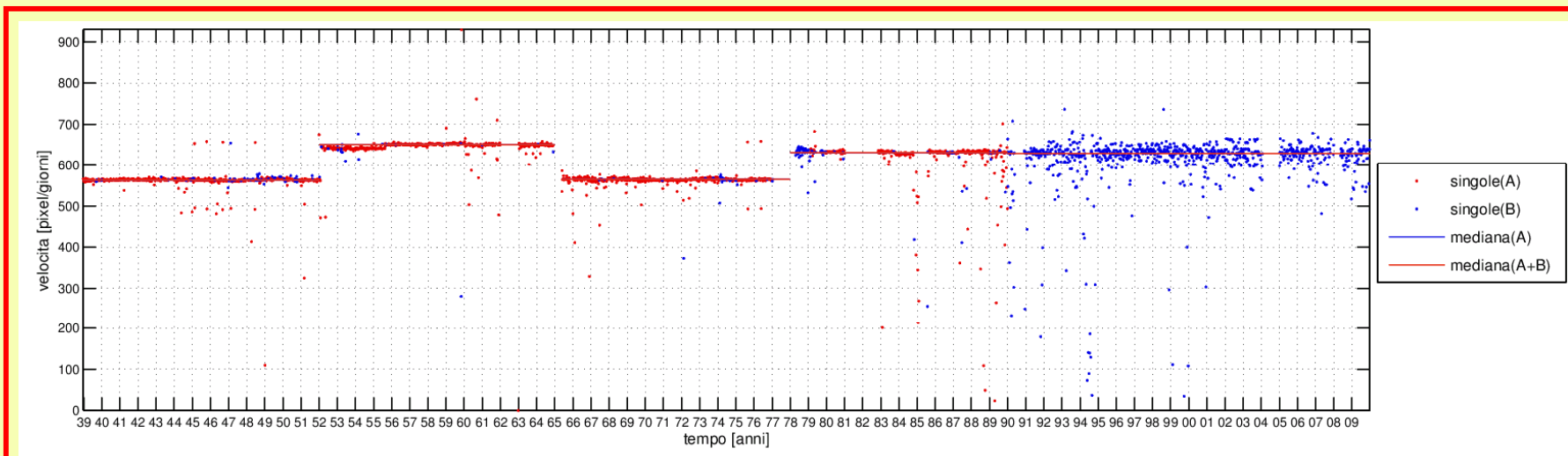
Procedure identifies all cusps
and adjusts excursion

Interactive Graphical Interface

- Visual inspection of Automatic Signal Recognition outcomes
- Adjust some pieces of signal with aided procedures



Allineamento temporale

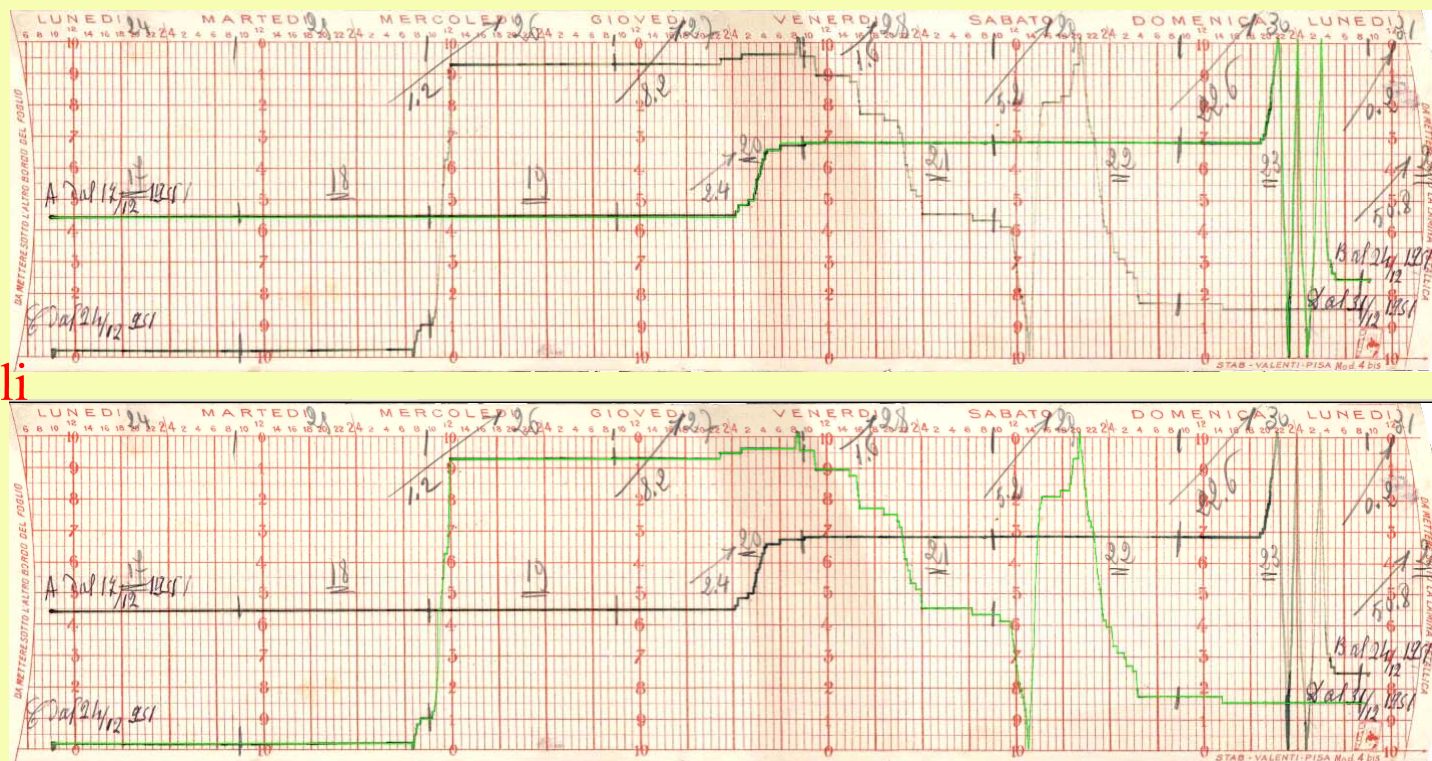


Velocità di rotazione del tamburo per le strisce pluviografiche della stazione di Arzana con sovrapposta la velocità adottata

“A”: strisce con l’ora di inizio e fine dichiarata dall’operatore;
“B”: strisce con l’ora non dichiarata dall’operatore e posta convenzionalmente uguale a 9:00.

Segnali multipli

Trattamento
segnali multipli
nella stessa
striscia



R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

COSTI DEI RECUPERI

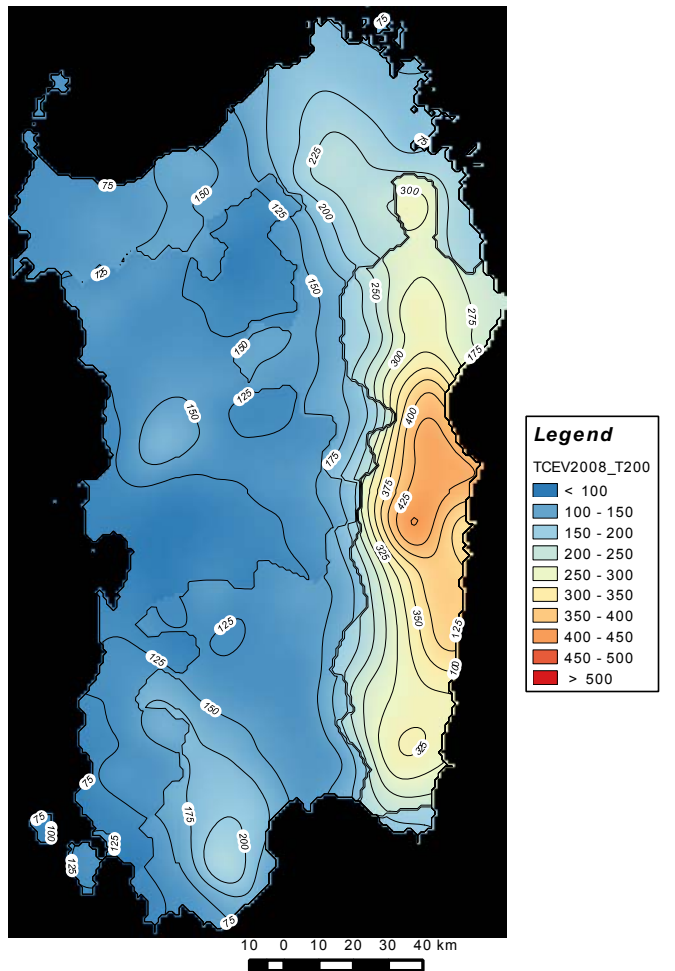
Circa 200-250 € per anno-stazione

Contributi e spunti di discussione sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

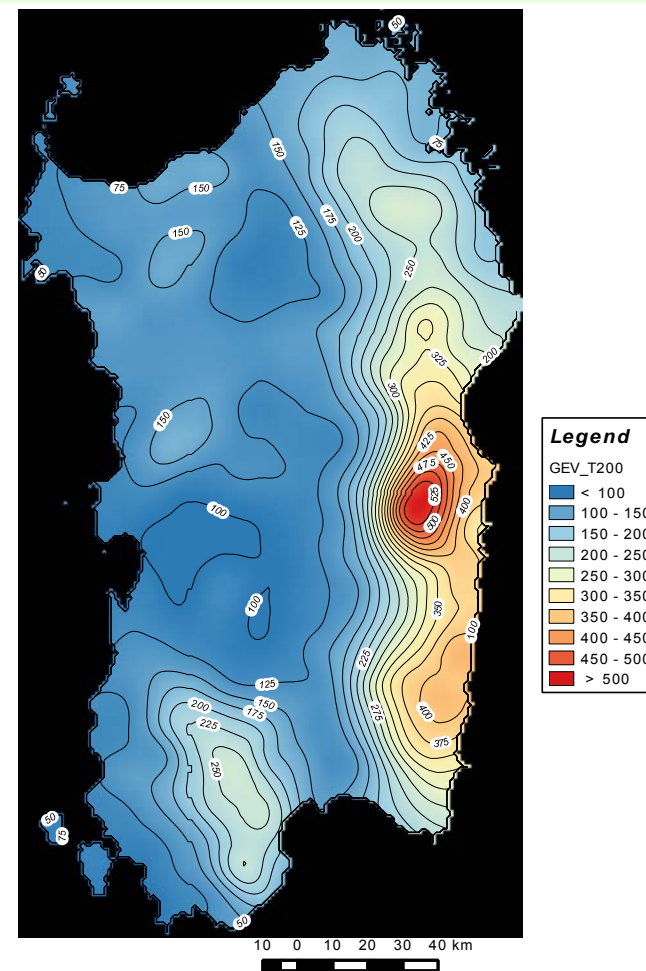
- 1) Recupero di misure storiche:
*reinterpretazione delle striscette cartacee contenenti i
segnali pluviografici e idrografici*
- 2) Caratterizzazione statistica degli eventi estremi e curve di
possibili pluviometrica:
dagli approcci regionali a quelli geostatistici
- 3) Caratterizzazioni statistiche basate sulle osservazioni
continue:
POT, invarianza di scala, etc.

Quantili precipitazione giornaliera T = 200 anni

Regioni omogenee



Approccio geostatistico



Contributi e spunti di discussione sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

- 1) Recupero di misure storiche:
reinterpretazione delle striscette cartacee contenenti i segnali pluviografici e idrografici
- 2) Caratterizzazione statistica degli eventi estremi e curve di possibili pluviometrica:
dagli approcci regionali a quelli geostatistici
- 3) Caratterizzazioni statistiche basate sulle osservazioni continue:
POT, invarianza di scala, etc.

The Generalized Pareto Distribution (GPD)

$$\Pr\{X \leq x \mid X > u\} = F(x; u, \alpha, \xi) = \begin{cases} 1 - \left(1 + \xi \frac{x-u}{\alpha}\right)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp\left(-\frac{x-u}{\alpha}\right) & \xi = 0 \end{cases}$$

- The shape parameter ξ :

- $\xi > 0$ “heavy tailed” distribution
- $\xi = 0$ exponential distribution
- $\xi < 0$ “bounded” distribution

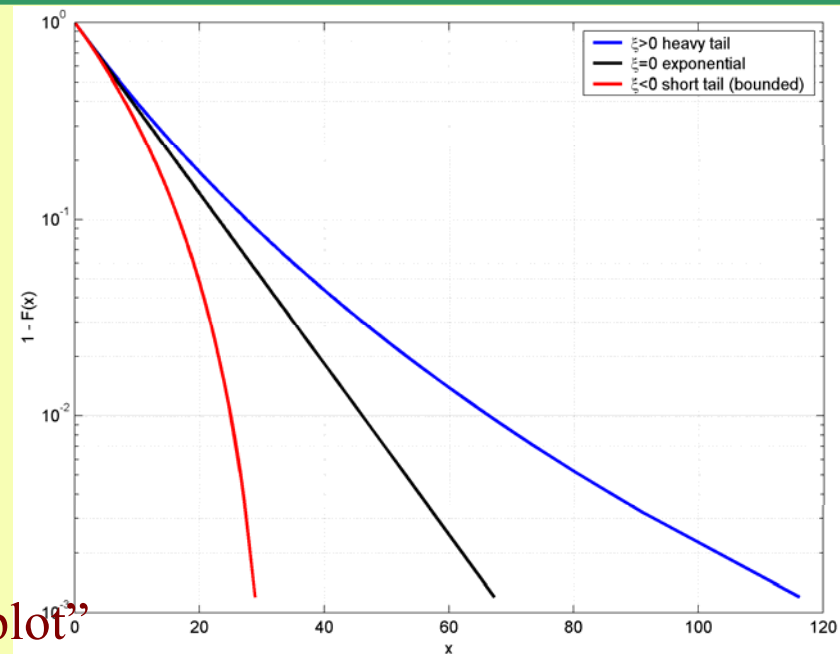
- The scale parameter α

> ML, PWM ... estimate α and ξ

- The “position” or “location” parameter u

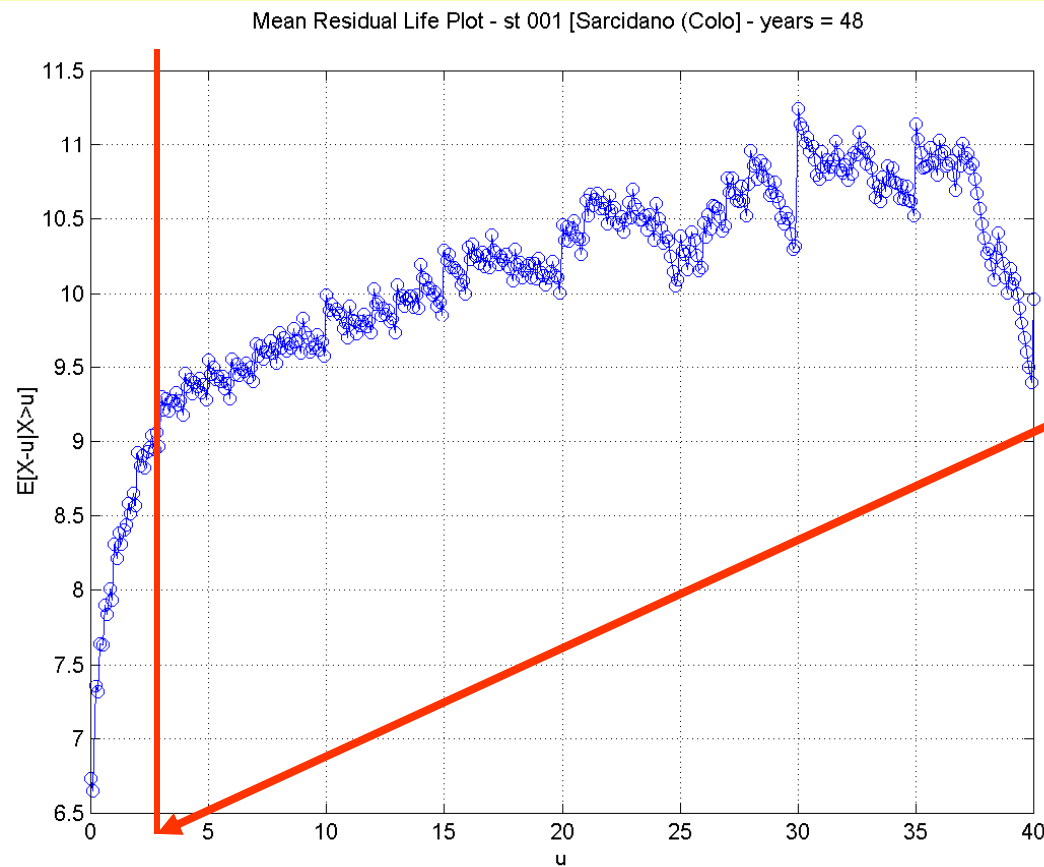
> A) graphical methods: “mean residual life plot”

> B) statistical methods based on GoF tests: “failure to reject”

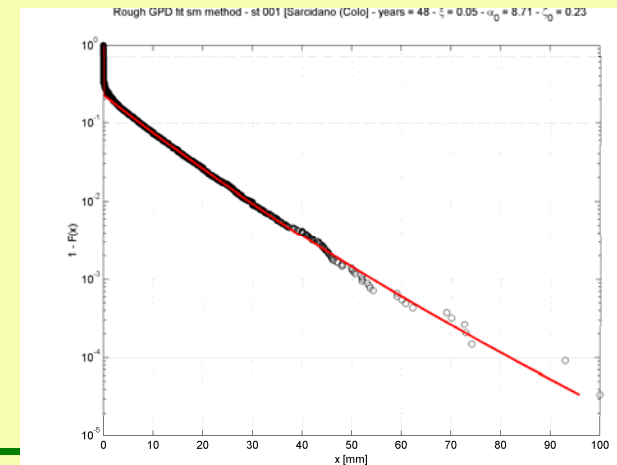


A) Estimate of the threshold u by “mean residual life plot”

Plots of $E[X - u | X > u]$ versus different thresholds u

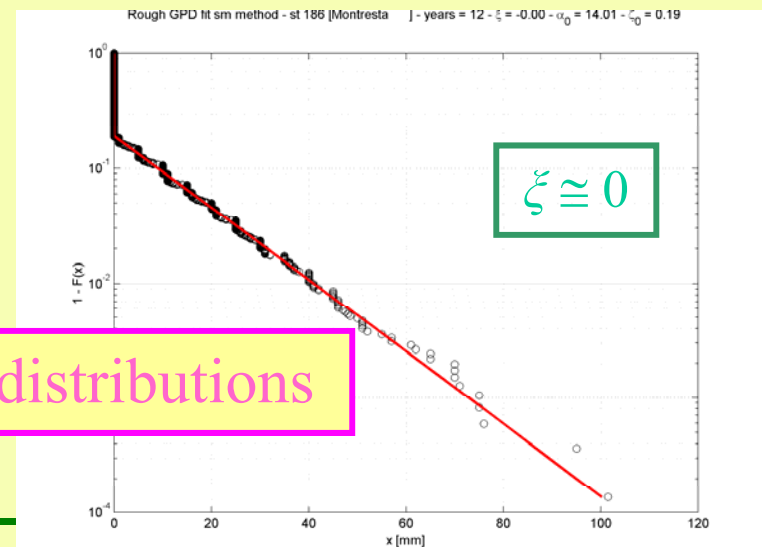
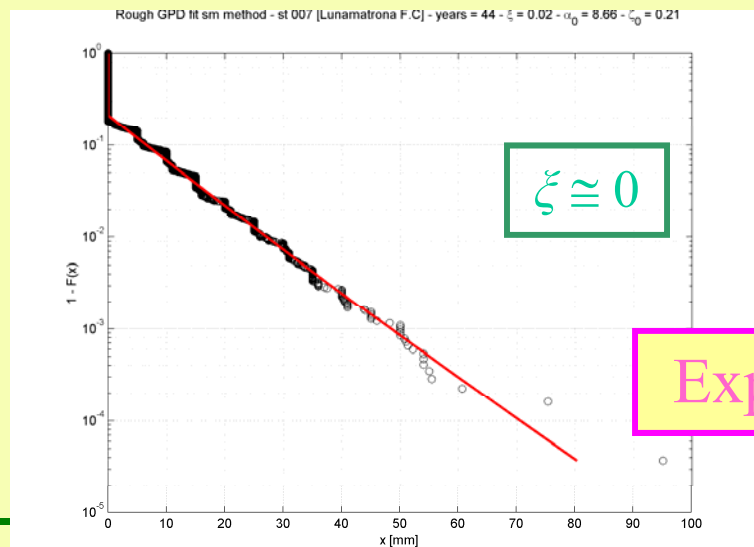
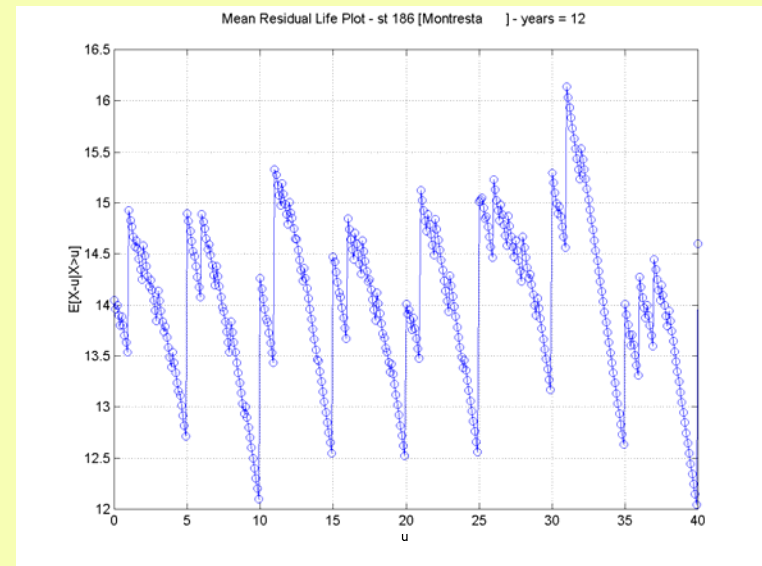
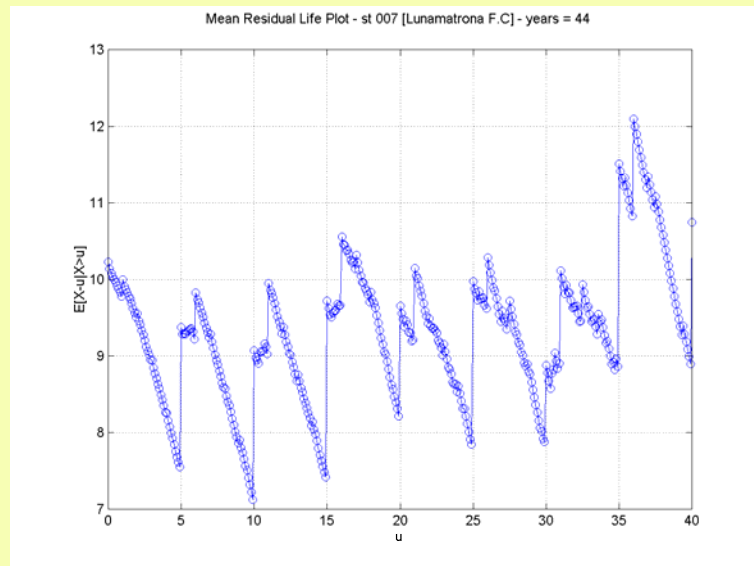


Optimum threshold u
is the minimum value
beyond which
the plot shows a
straight line behaviour



R. Deidda – Contributi sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

“mean residual life plot” on rounded off records



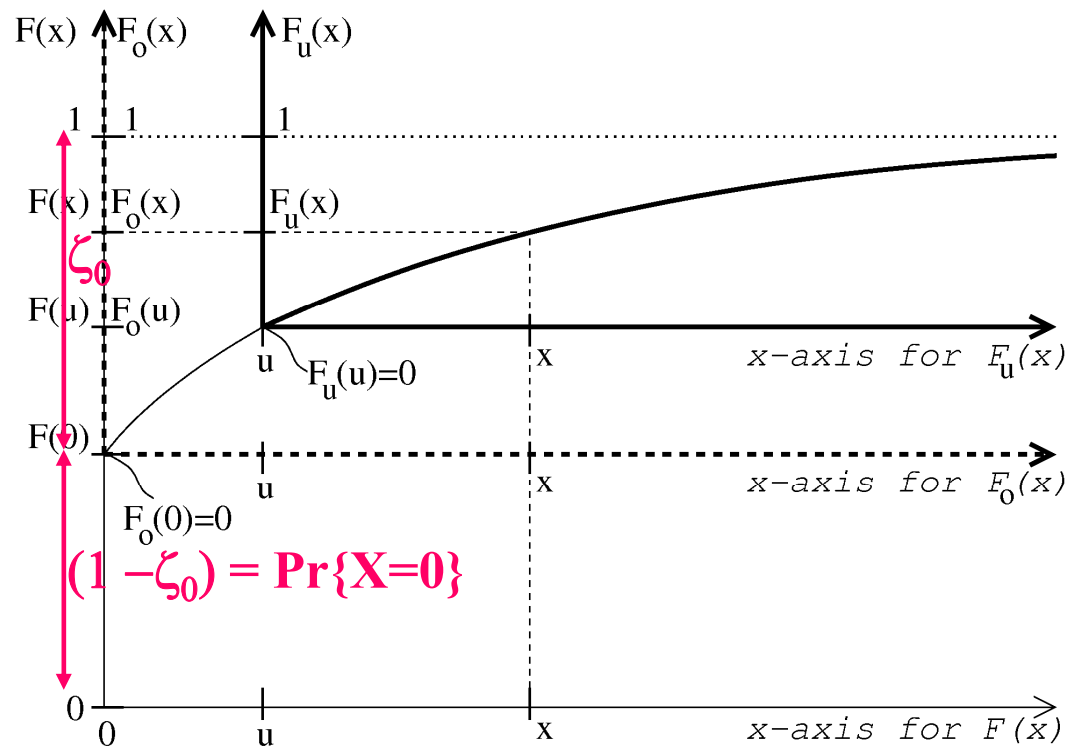
Exponential distributions

Threshold-invariant GPD reparameterization - I

We look for a general equation $F(x)$ for zero and non-zero rainfall

$$F(x) = (1 - \zeta_0) + \zeta_0 F_0(x) \quad (1)$$

where $\zeta_0 = \Pr\{X > 0\}$; $F_0(x) = \Pr\{X \leq x | X > 0\}$



Fitting a GPD $F_u(x) = \Pr\{X \leq x | X > u\}$
for any threshold u

larger than the optimum one:
 ξ (shape) = threshold-invariant
 α_u (scale) = linear function of u
 $\zeta_u = \Pr\{X > u\} = \text{non-linear } f(u)$

Letting $F_0(x)$ be also a GPD we
reparameterize eq.(1):

$$\alpha_0 = f(\alpha_u, \xi, u)$$

$$\zeta_0 = f(\zeta_u, \alpha_0, \xi, u)$$

Perfect overlapping for $x > u$

α_0 and ζ_0 are threshold-invariant reparameterization of eq.(1):

A) Basis for the Multiple Threshold Method

B) Important property for spatial mapping of parameters

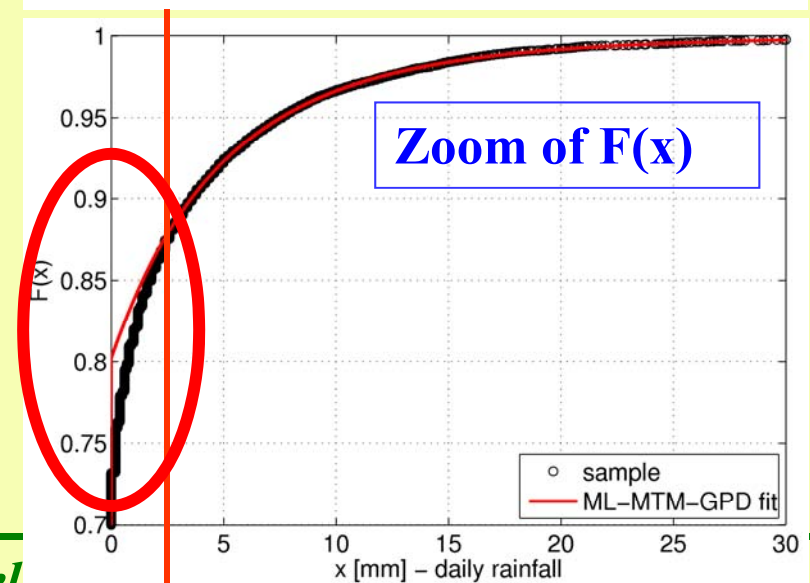
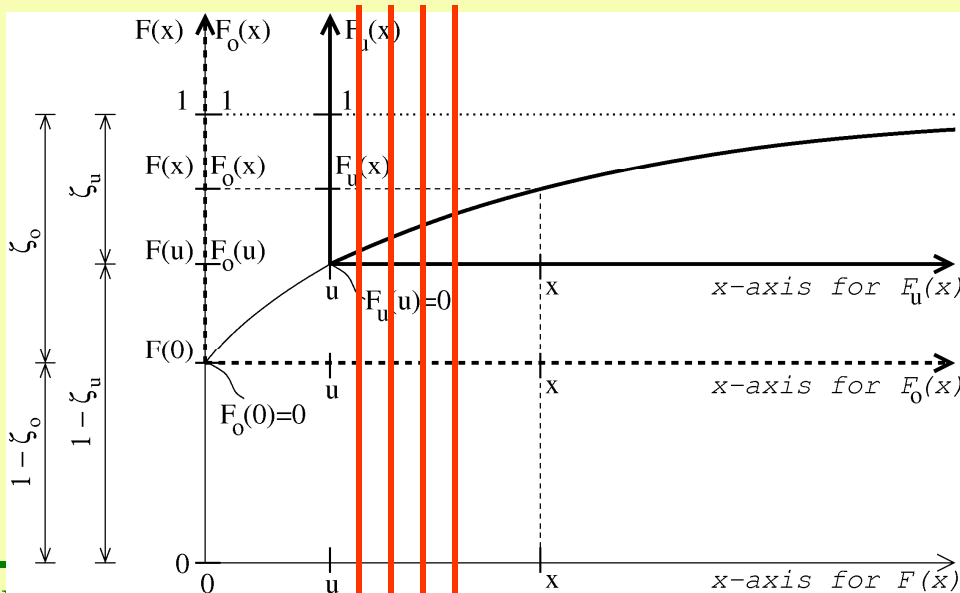
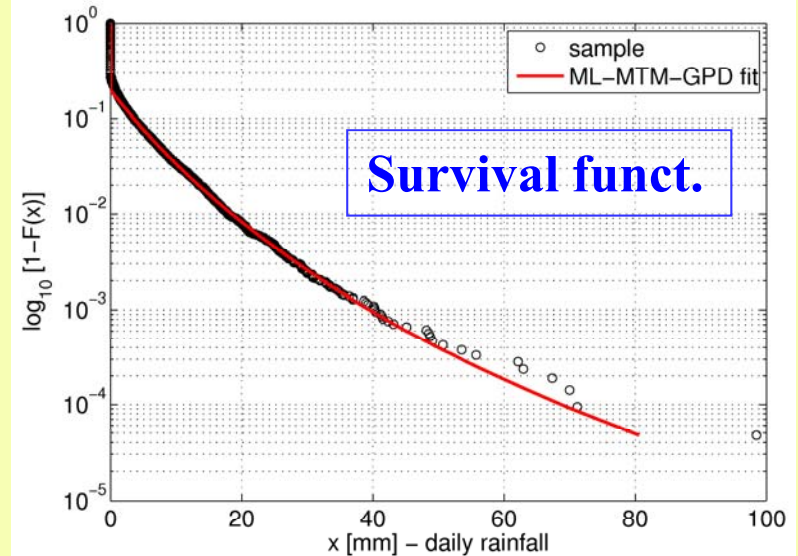
Multiple Threshold Method for GPD fitting - II

A good time series
(no rounding)

Multiple Threshold Method
(MTM)

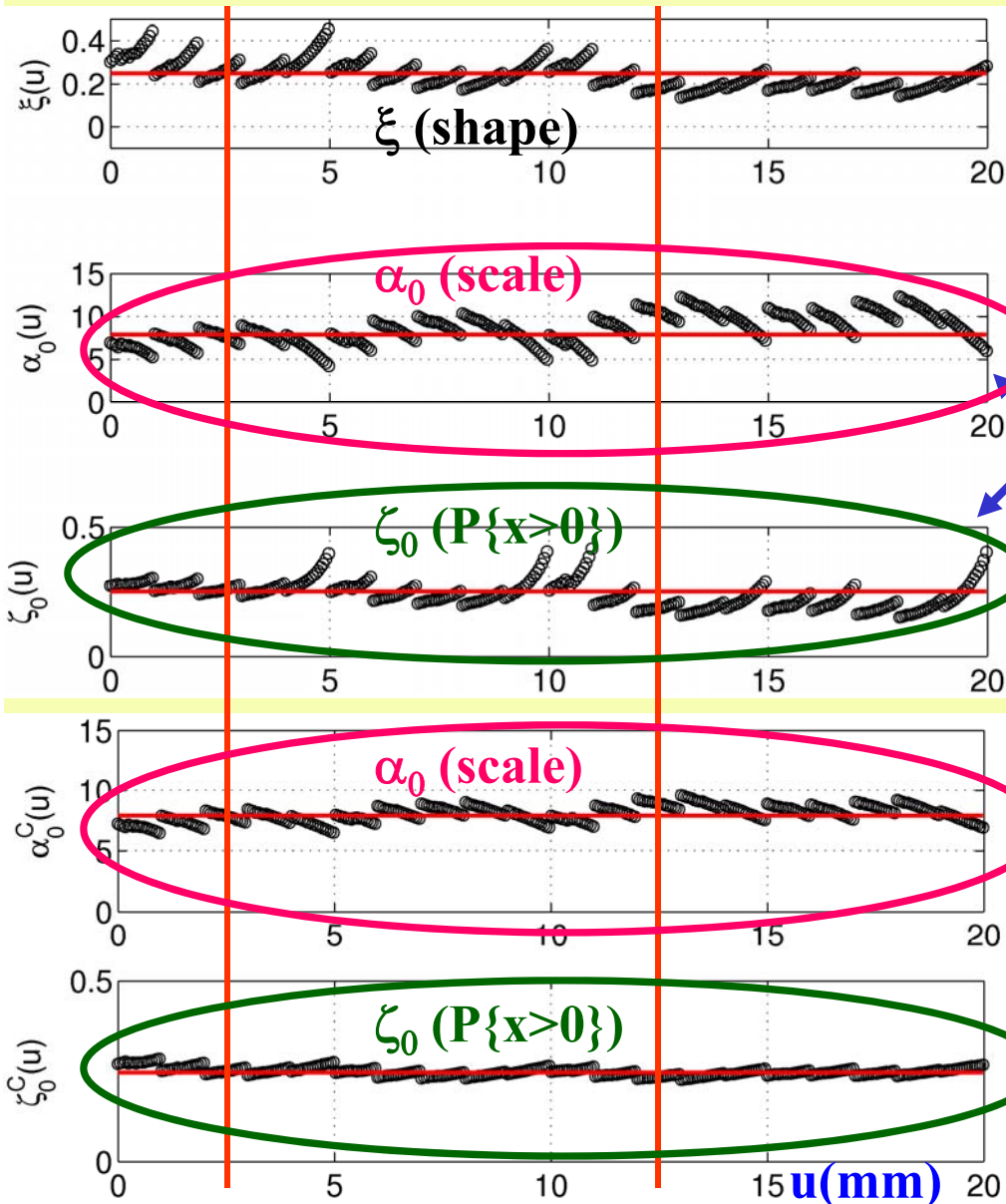
RESULTS

ML-MTM-GPD fit st074-58y $\xi^M = 0.15 - \alpha_0^M = 4.95 - \zeta_0^M = 0.20$



R. Deidda – Contributi sul tema: La misura delle grandine e l'idrologia e la statistica nell'idrologia operativa

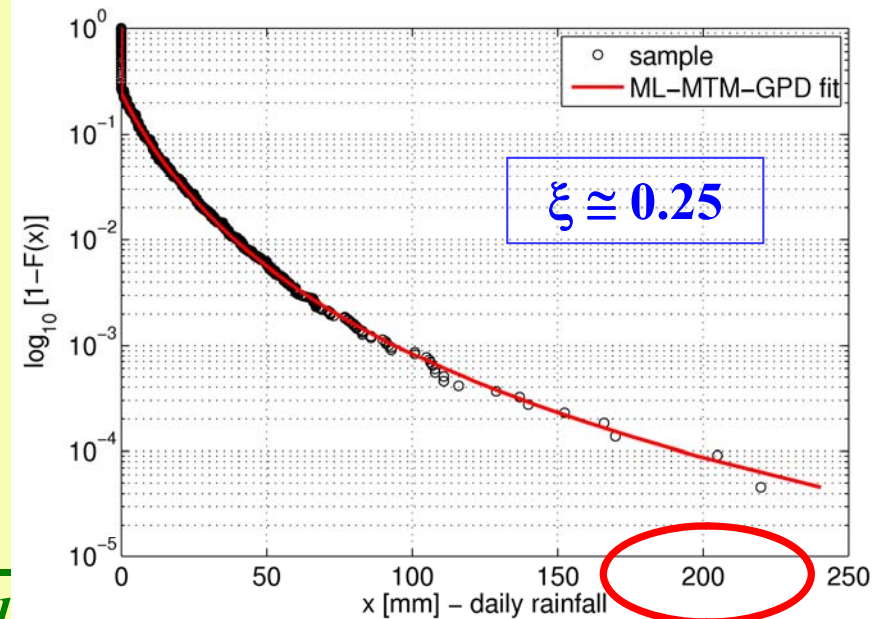
Multiple Threshold Method for GPD fitting - IV



Other MTM examples on rounded time series: a lot of 1 mm roundings

Simple reparameterization

ML-MTM-GPD fit st285-60y $\xi^M = 0.25 - \alpha_0^M = 7.92 - \zeta_0^M = 0.25$



Contributi e spunti di discussione sul tema: La misura delle grandezze idrologiche e la statistica nell'idrologia operativa

- 1) Recupero di misure storiche:
reinterpretazione delle striscette cartacee contenenti i segnali pluviografici e idrografici
- 2) Caratterizzazione statistica degli eventi estremi e curve di possibili pluviometrica:
dagli approcci regionali a quelli geostatistici
- 3) Caratterizzazioni statistiche basate sulle osservazioni continue:
POT, invarianza di scala, etc.