

G3-GNRAC

Quartiere Fieristico di Ferrara, 19 Settembre 2014

**CONFRONTO FRA LE
DISTRIBUZIONI DI WEIBULL E DI PARETO
NELL'ANALISI DELLE ALTEZZE D'ONDA ESTREME**

Motivazioni ed Obiettivi

- L'estensione temporale limitata delle fonti dati strumentali in ambito marino rende necessaria la verifica delle proprietà asintotiche della Teoria dei Valori Estremi
- La CDF è un'incognita dell'inferenza statistica, quindi il modello ottimale per la base dati a disposizione deve essere scelto rispetto ad un insieme di CDF candidate
- Analisi statistica degli eventi estremi del db della stazione di Alghero: il più energetico della RON (massima altezza d'onda significativa spettrale registrata, $H_{m0} \approx 10$ m)

Motivazioni ed Obiettivi

- Confronto distribuzione generalizzata di Pareto (GPA) e distribuzione di Weibull (WBL) all'interno di un modello POT – Poisson (Mazas & Hamm, 2010)
 - ✓ utilizzo caratteristiche GPA per la stima della soglia
 - ✓ valutazione performance per diversi metodi di stima e metodi di scelta del modello ottimale

- La GPA è meno cautelativa nella stima dei quantili delle altezze d'onda (Goda et al, 2010) (?)

Modello POT – POISSON

$$X_i \stackrel{N}{i=1} \text{ i.i.d.} \sim F(A, B, k) \Rightarrow \Pr[X \leq x]$$

$$M_N = \max_{i=1, \dots, N} X_i$$

$$\Pr[M_N \leq z] \approx G(A, B, k) \approx GEV(A, B, k)$$

Legge degli eccessi
sopra soglia (Pickands, 1975)

$$Y_i = X_i - u \quad u \text{ sufficientemente alta}$$

$$F(A, k) \Rightarrow \Pr[X < u + y | X > u] \approx \frac{F(u + y) - F(u)}{1 - F(u)} \approx GPA$$

Modello GPA - Poisson

$$\lambda = \frac{N}{K}; \text{POI} = \frac{\lambda^x}{x!} e^{-\lambda}$$

$$X_i = H_{m0} \text{ (m)}$$

A, parametro scala (m)

B, parametro posizione (m)

K, p. forma (-)

Y_i eccessi sopra soglia (m)
u soglia (m)

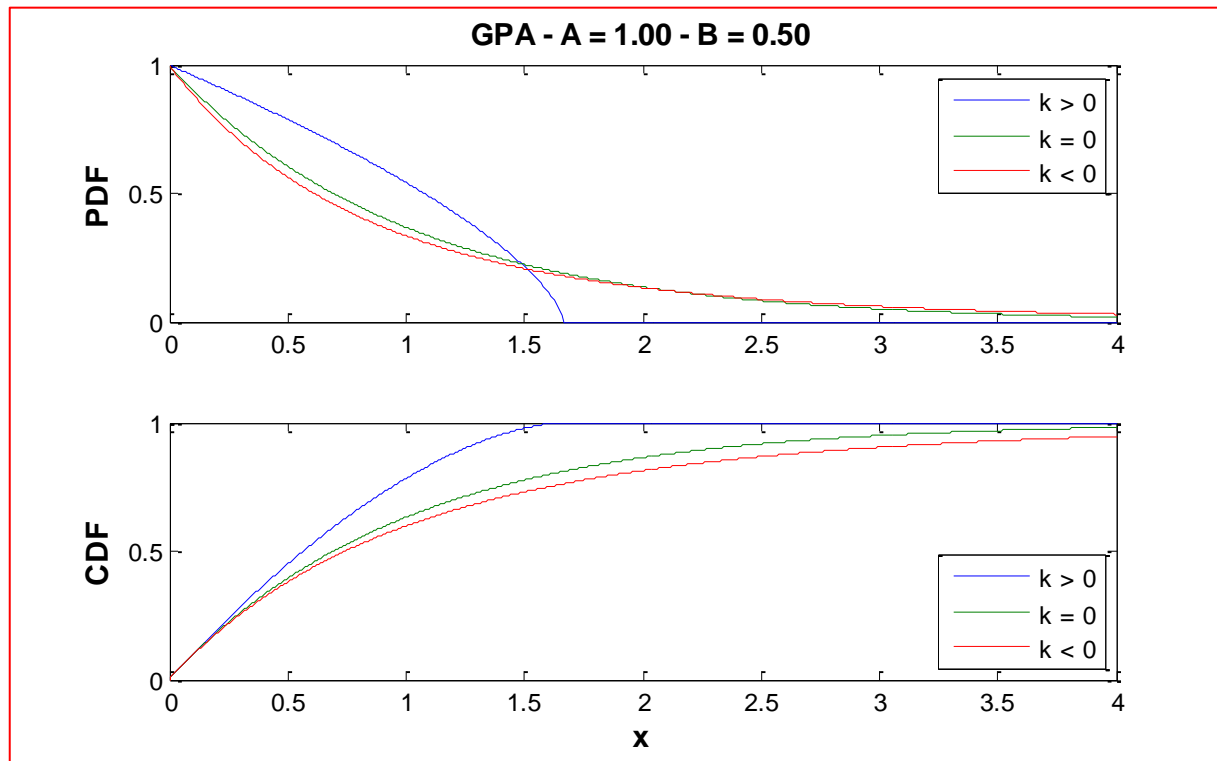
λ numero medio mareggiate
annue (eventi/anno)

K numero anni

CDF Pareto generalizzata (GPA)

$$F(x; A, k) = GPA = \begin{cases} 1 - \left(1 - k \frac{x-u}{A}\right)^{\frac{1}{k}} & k \neq 0 \\ 1 - \exp\left(-\frac{x-u}{A}\right) & k = 0 \end{cases} \quad A > 0, \quad u > 0, \quad 1 - k \frac{x-u}{A} > 0$$

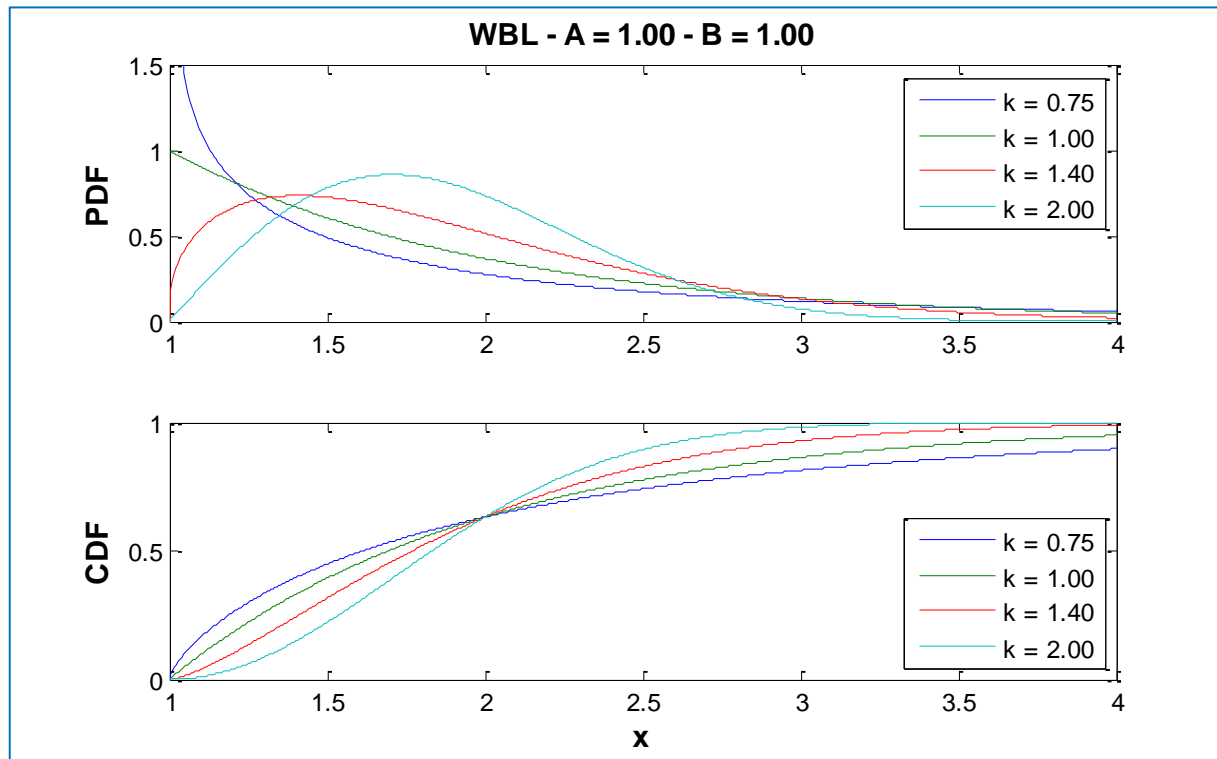
Hosking &
Wallis, 1987



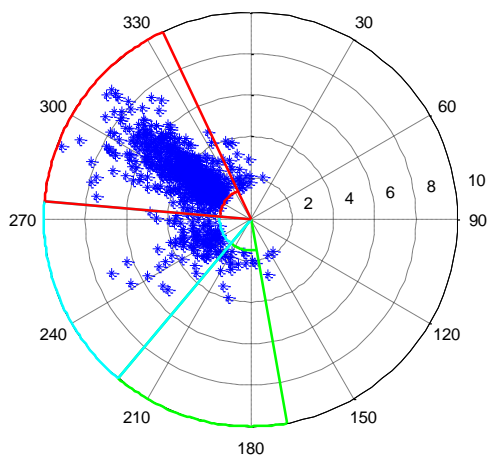
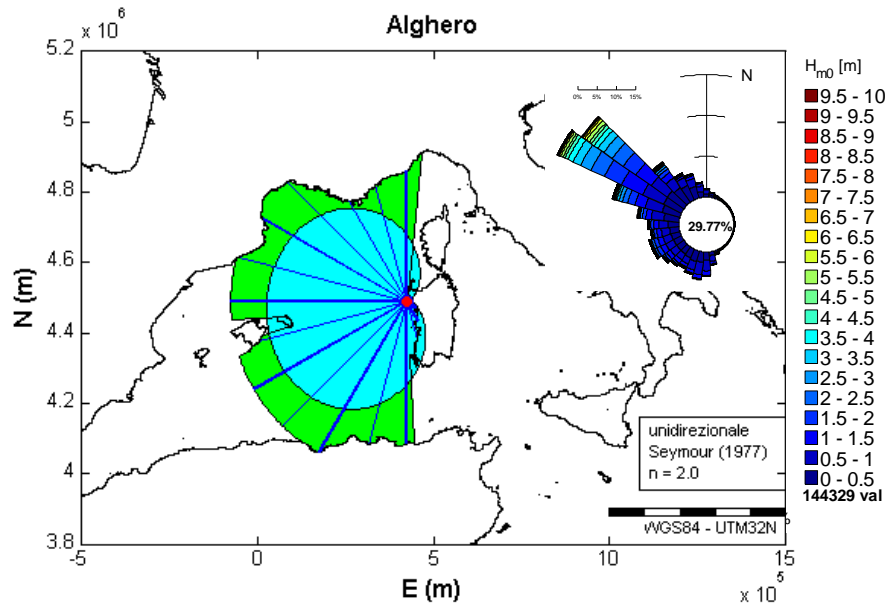
CDF Weibull (WBL)

$$WBL = F(x; A, B, k) = 1 - \exp\left[-\left(\frac{x-B}{A}\right)^k\right] \quad k > 0, A > 0, B \leq x < \infty$$

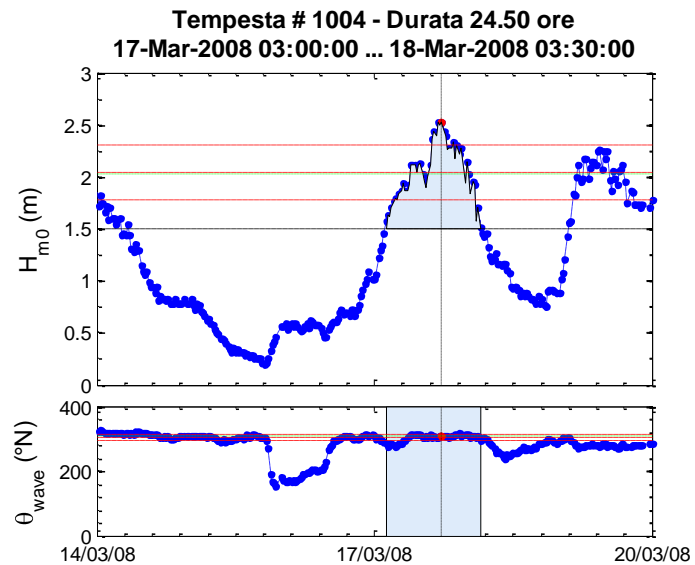
Weibull, 1959



Selezione eventi (iid)



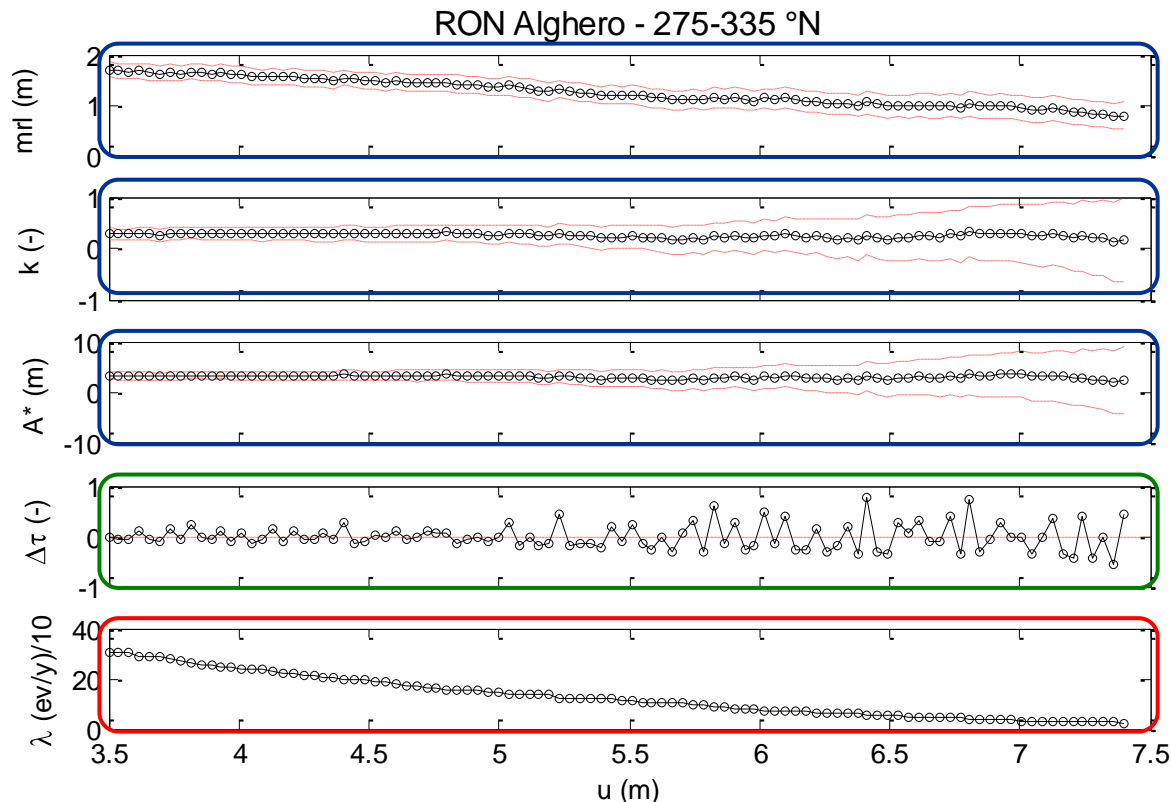
Ritardo 48 ore picchi consecutivi



Tre settori direzionali omogenei

Scelta della soglia (1)

Approcci mono-soglia sogettivi (Coles, 2001), oggettivi (Thompson, et,al, 2008), empirici (Mazas & Hamm, 2011)



$$mrl = E \left[\left\{ X - u \mid X > u \right\} \right] \frac{A(u)}{1+k}$$

$$k(u) \approx \text{cost}$$

$$A^* = A(u) \left\{ k(u) \approx \text{cost} \right.$$

$$u_1 < u_2 < \dots < u_n : \Delta u = \text{cost}$$

$$\tau(u_j) = \hat{A}(u_j) + \hat{k}(u_j)u_j \quad j=1, \dots, n$$

$$\left\{ \tau \right\} E \left[\tau \right] = 0 \text{ \& } \Delta \tau \sim N(0, \sigma^2)$$

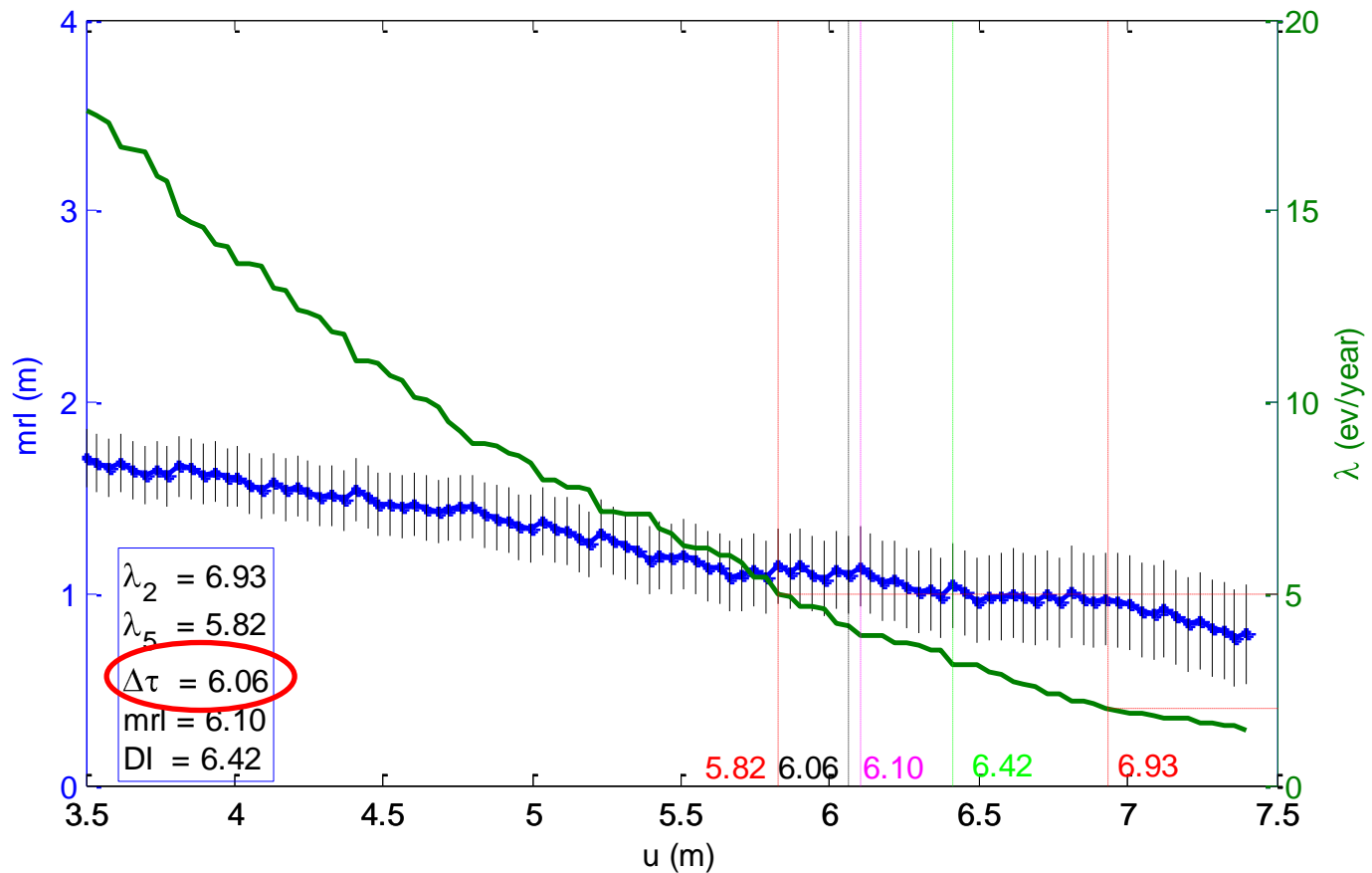
$$\lambda \in \left[\frac{1}{5} \right]$$

$$DI = \sigma^2 / \mu$$

$$\text{POI, } DI = 1$$

Scelta della soglia (2)

Soglia ottimale

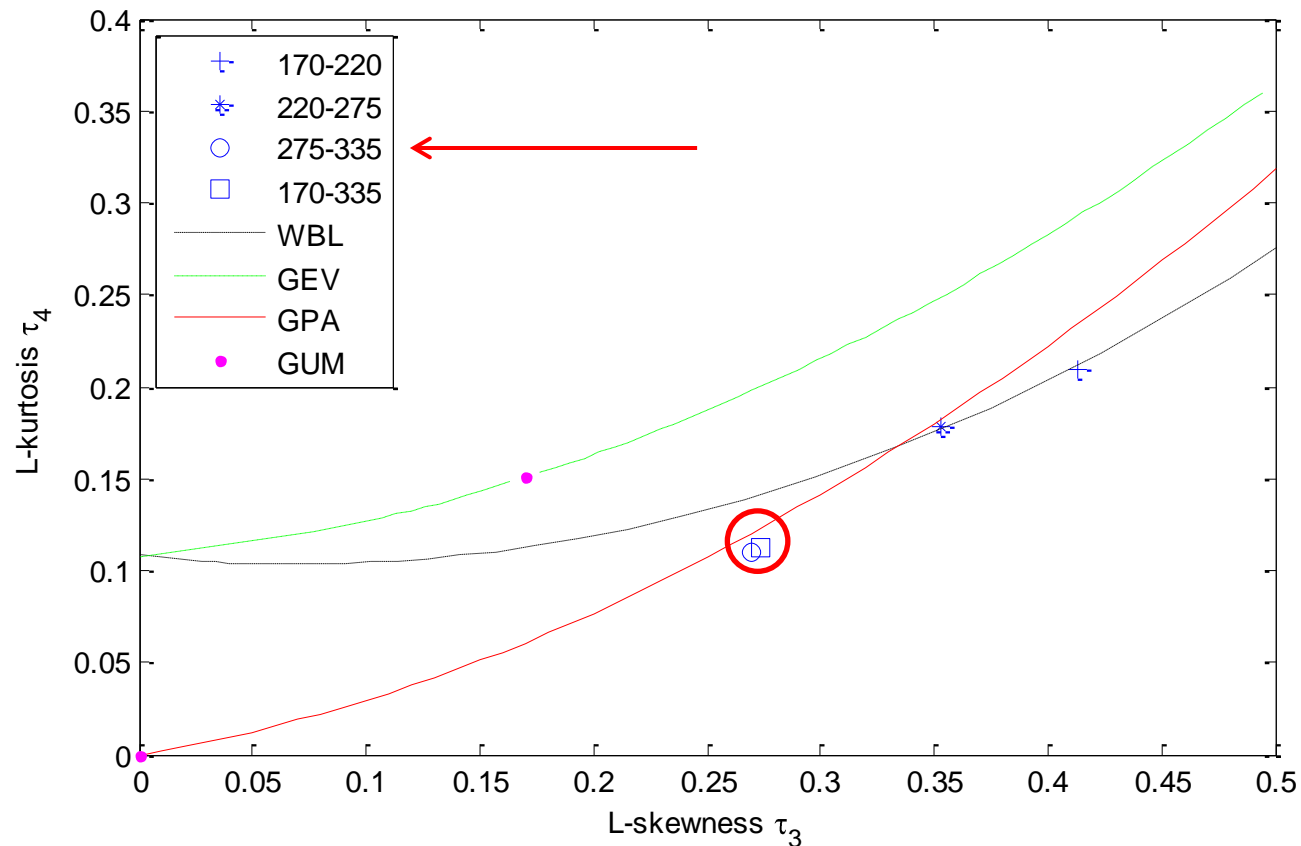


Scelta del modello (1)

- Rapporti L-Moment (Hosking & Wallis, 1997);
- BIC (Bayesian Information Criteria), AIC (Akaike Information Criteria) (Mazas & Hamm, 2010);
- MIR – PPC (Minimum Residual – Probability Plot Correlation coefficient), MIR – QQC (Minimum Residual – Quantile Quantile Plot) (Goda, 2011);
- Norma di errore (Li, et. al., 2008) modificata

Scelta del modello (2)

Diagramma rapporti L-Moment



Scelta del modello (3)

Criteri parametrici: BIC, AIC; MIR – PPC e MIR-QQC

$$BIC = -2 \ln L + M_p \ln N \quad AIC = -2 \ln L + 2M_p$$

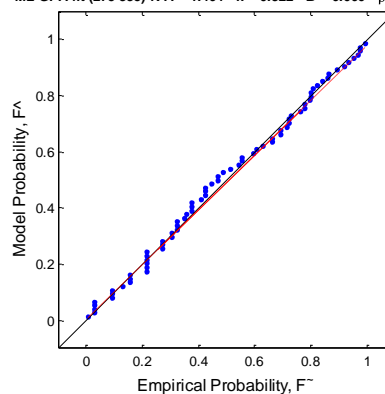
$$M.I.R = 1 - \rho = 1 - \frac{Cov(x, y)}{\sqrt{Var(x)Var(y)}}$$

L funzione verosimiglianza

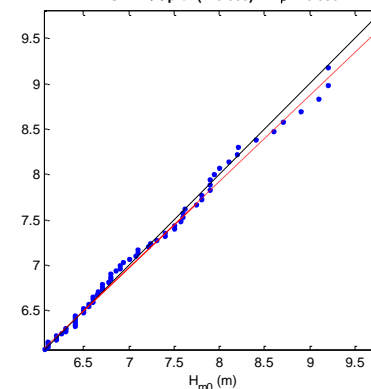
N numerosità

M_p # parametri

ML-GPA fit (275-335)*N A = 1.464 - k = 0.322 - B = 6.060 - $\rho = 0.999$



ML-GPA QQ-plot (275-335)*N - $\rho = 0.998$



Procedura Monte Carlo, criteri non parametrici: Error Norm, EN; Bias e RMSE

$$EN = E \left[\max_{i=1}^n \left| \tilde{F}^h - \hat{F}(r_i)^h \right| \right]$$

$$Bias(\hat{\theta}) = E \left[\hat{\theta} \right] - \theta$$

$$RMSE(\hat{\theta}) = \left\{ Var \left[\hat{\theta} \right] + Bias(\hat{\theta})^2 \right\}^{1/2}$$

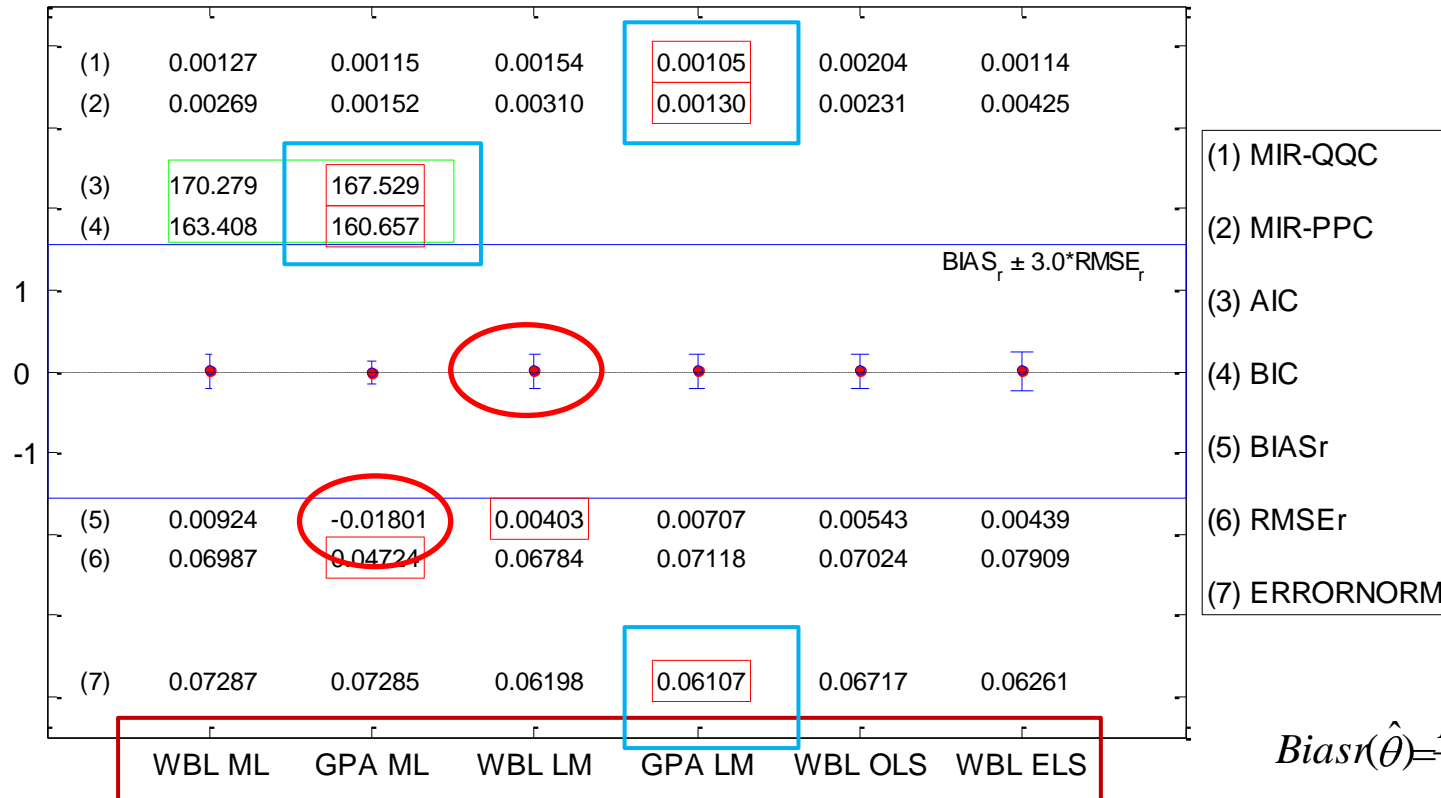
F^{\wedge} CDF ottimale

F^{\sim} CDF empirica (plotting position non distorta)

h parametro di peso (>1 maggiore enfasi sulla coda destra, serie di valori pari a 0.25, 0.5, 0.75, 1, 1.25, 1.5 nelle simulazioni)

Risultati (1)

H_{100} (m) - (275-335) °N - 1000 SIM

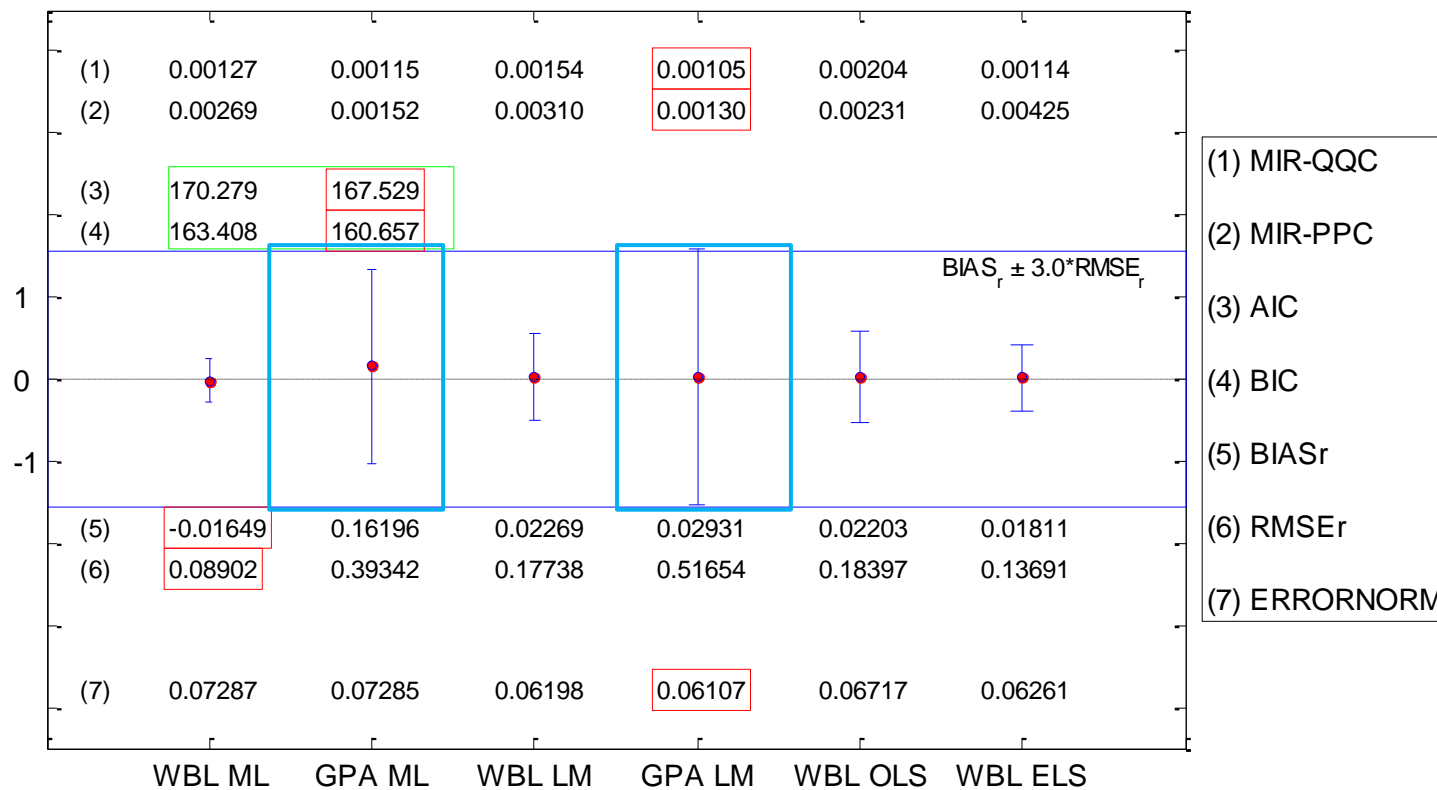


$$Biasr(\hat{\theta}) = \frac{Bias(\hat{\theta})}{\theta}$$

$$RMSEr(\hat{\theta}) = \frac{RMSE(\hat{\theta})}{\theta}$$

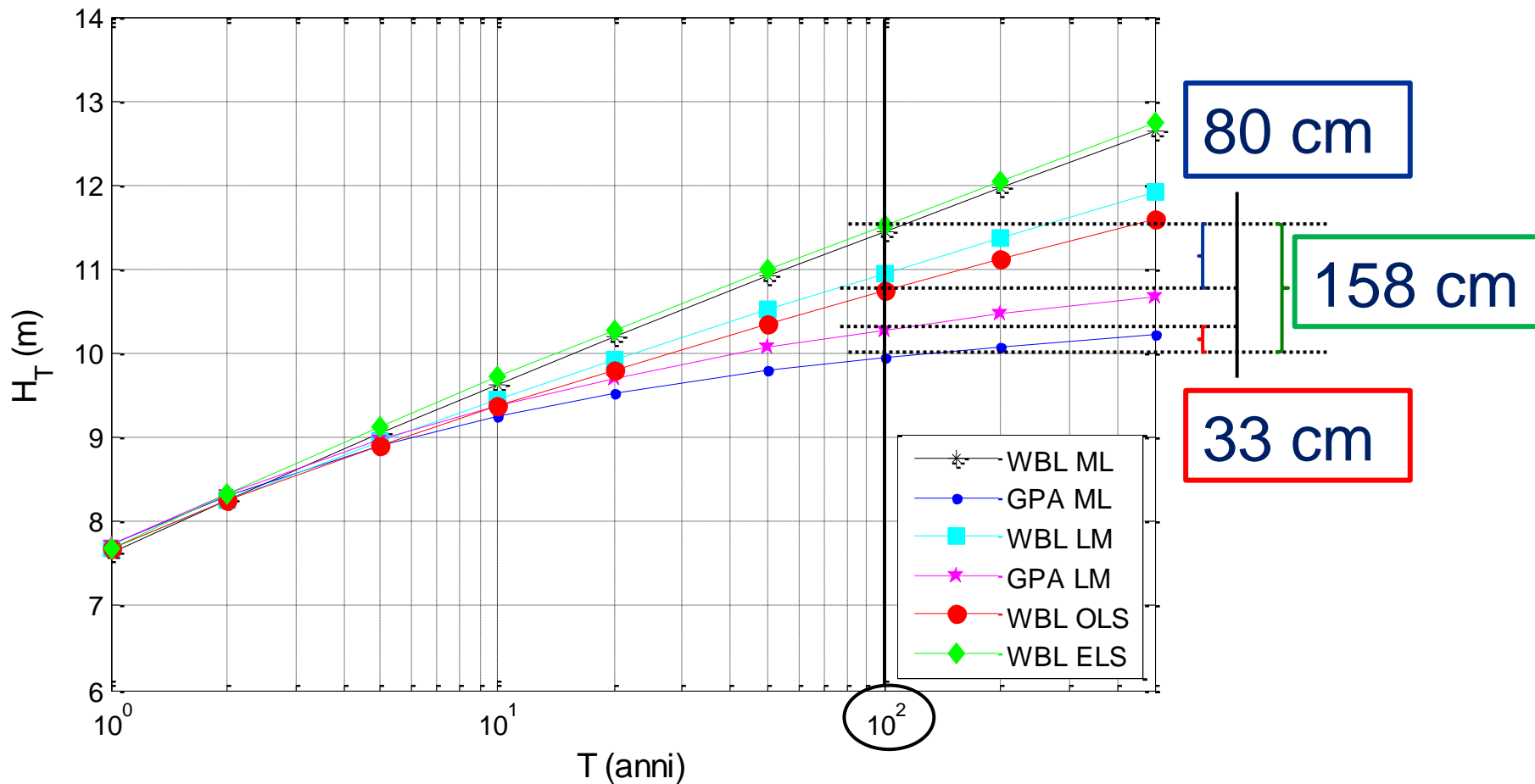
Risultati (2)

PARAMETRO DI FORMA k (-) - (275-335) °N - 1000 SIM



Risultati (3)

Confronto Quantili CDF - 275-335 °N



Conclusioni (1)

- Importanza della validazione base dati iniziale
 - ✓ controllo climatologico di consistenza con altre fonti dati
 - ✓ classificazione e verifica della selezione delle mareggiate

- Approccio multi-soglia (necessità di ulteriori studi)
 - ✓ Adattato ai massimi di H_m0 che presentano un comportamento diverso rispetto ai massimi di precipitazione giornaliera (indipendenza ?, processo poissoniano ?)

Conclusioni (2)

- La GPA è flessibile nell'identificazione della soglia ottimale
- La GPA è la più indicata per il criterio LM, per i criteri di scelta parametrici e per alcuni non parametrici
- La GPA è la meno efficiente nella stima dei parametri
- La GPA determina quantili relativamente bassi, quindi poco cautelativi, in accordo con altri autori (Goda, et al, 2010)
- La combinazione GPA-LM (3 parametri) determina stime dei quantili maggiori rispetto alla combinazione GPA-ML (2 parametri), in accordo con altri autori (Mazas & Hamm, 2012)

Conclusioni (3)

- Conferma validità WBL a scopi applicativi
 - ✓ generale buon adattamento per tutti i metodi di stima
 - ✓ quantili cautelativi rispetto alla GPA

- Criterio di sicurezza per la pratica progettuale:
 - ✓ utilizzare una CDF robusta nella stima dei parametri ma soprattutto nella stima dei quantili

Sviluppi futuri

- Possibile estensione approccio multi-soglia ai db marini
- Possibile integrazione modelli GPA – GEV
 - ✓ irrobustimento dei parametri della GEV seguendo le relazioni esistenti tra i rispettivi parametri
 - ✓ utilizzo della GEV per la stima dei quantili
- Estensione dell'analisi statistica ai livelli marini

Bibliografia

- Arns, A. et al., 2013. Estimating extreme water level probabilities: a comparison of the direct methods and recommendations for best practise. *Coastal Engineering*, Volume 81, pp. 51 - 66.
- Coles, S., 2001. *An introduction to statistical modeling of extreme values*. London: Springer-Verlag.
- Corsini, S., Inghilesi, R., Franco, L. & Pisocopia, R., 2004. *Atlante delle onde nei mari italiani. Italian wave atlas*, Roma: ISPRA, Università degli Studi Roma Tre.
- Deidda, R., 2010. A multiple threshold method for fitting the generalized Pareto distribution and a simple representation of the rainfall process. *Hydrology and Earth System Sciences*, Volume 14, pp. 2559-2575.
- Goda, Y., 2010. *Random seas and design of maritime structures. Advanced Series on Ocean Engineering*. Singapore: World Scientific.
- Goda, Y., 2011. Plotting position estimator for the L-Moment method and quantile confidence interval for the GEV, GPA, and Weibull distributions applied for extreme wave analysis. *Coastal Engineering Journal*, 53(2), pp. 111-149.
- Goda, Y., Kudaka, M. & Kawai, H., 2010. *Incorporation of Weibull distribution in L-Moments method for regional frequency analysis of Peaks-Over-Threshold wave heights*. s.l., Proceedings of 32nd International Conference on Coastal Engineering.
- Hawkes, P. J., Gonzalez-Marco, D., Sanchez-Arcilla, A. & Prinos, P., 2008. Best practice for the estimation of extremes: a review. *Journal of Hydraulic Research*, 46(Extra Issue 2), pp. 324-332.
- Hosking, J. R. M., 1990. L-Moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1), pp. 105-124.
- Hosking, J. R. M. & Wallis, J. R., 1987. Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics*, 29(3), pp. 339-349.
- Li, Y., Simmonds, D. & Reeve, D., 2008. Quantifying uncertainty in extreme values of design parameters with resampling techniques. *Ocean Engineering*, pp. 1029-1038.
- Mathiesen, J. et al., 1994. Recommended practice for extreme wave analysis. *Journal of Hydraulic Research*, 32(6), pp. 803-814.
- Mazas, F. & Hamm, L., 2011. A multi-distribution approach to POT methods for determining extreme wave heights. *Coastal Engineering*, Volume 58, pp. 385-394.
- Mazas, F., Hamm, L. & Garat, P., 2012. *POT methods: a new insight into the estimation of extreme value distributions - applications to extreme wave heights*. Santander, Spain, ICCE.
- Pickands, J., 1975. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1), pp. 119-131.
- Thompson, P., Cai, Y., Reeve, D. & Stander, J., 2009. Automated threshold selection methods for extreme wave analysis. *Coastal Engineering*, Volume 56, pp. 1013-1021.
- Weibull, W., 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, Volume 18, pp. 293-297.

Ringraziamenti



UNIONE EUROPEA
Fondo sociale europeo



REGIONE AUTONOMA DELLA SARDEGNA



Dot.IT

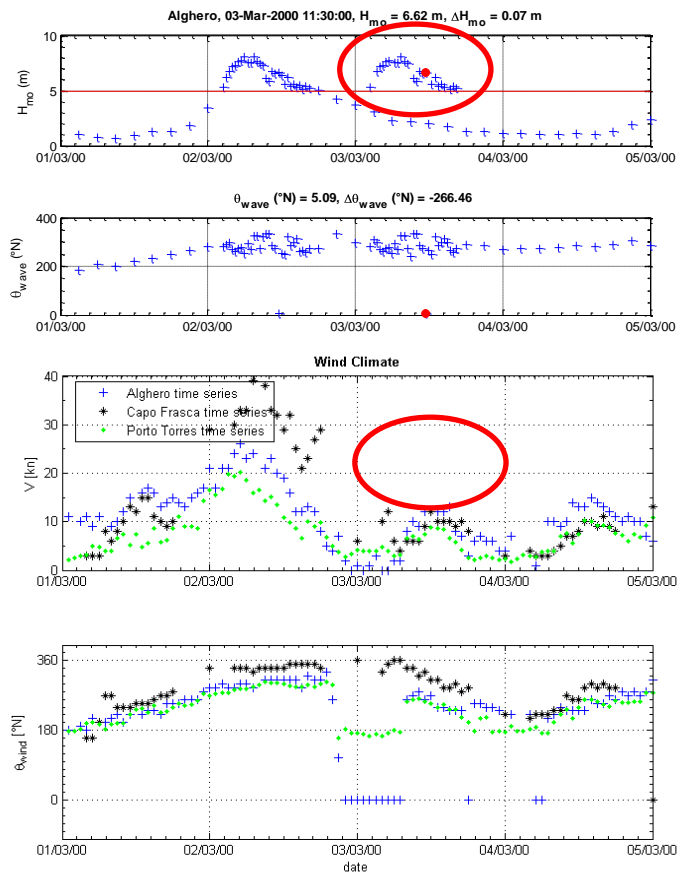


ISPRA

Grazie per l'attenzione

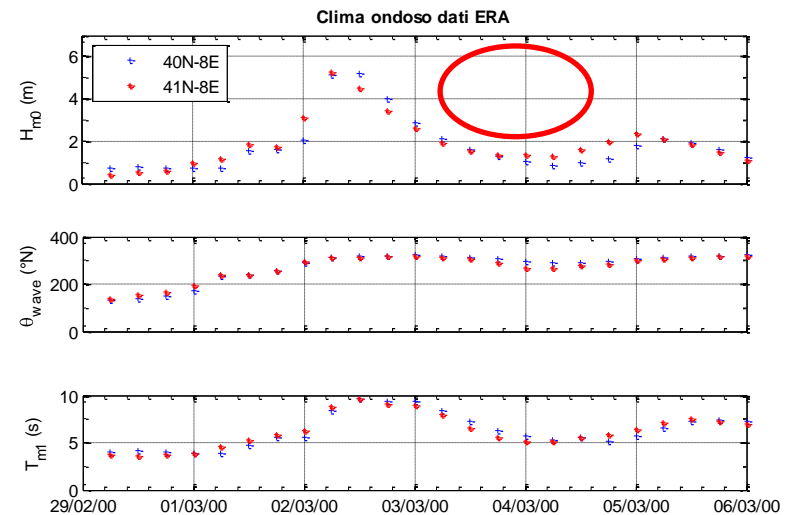
Validazione Serie Storica

Clima ondoso RON



Confronto climatologico di
consistenza fra DB diversi

Clima ondoso dati ERA



Clima anemometrico

Soglia ottimale

Approccio multi soglia (Deidda, 2010)

