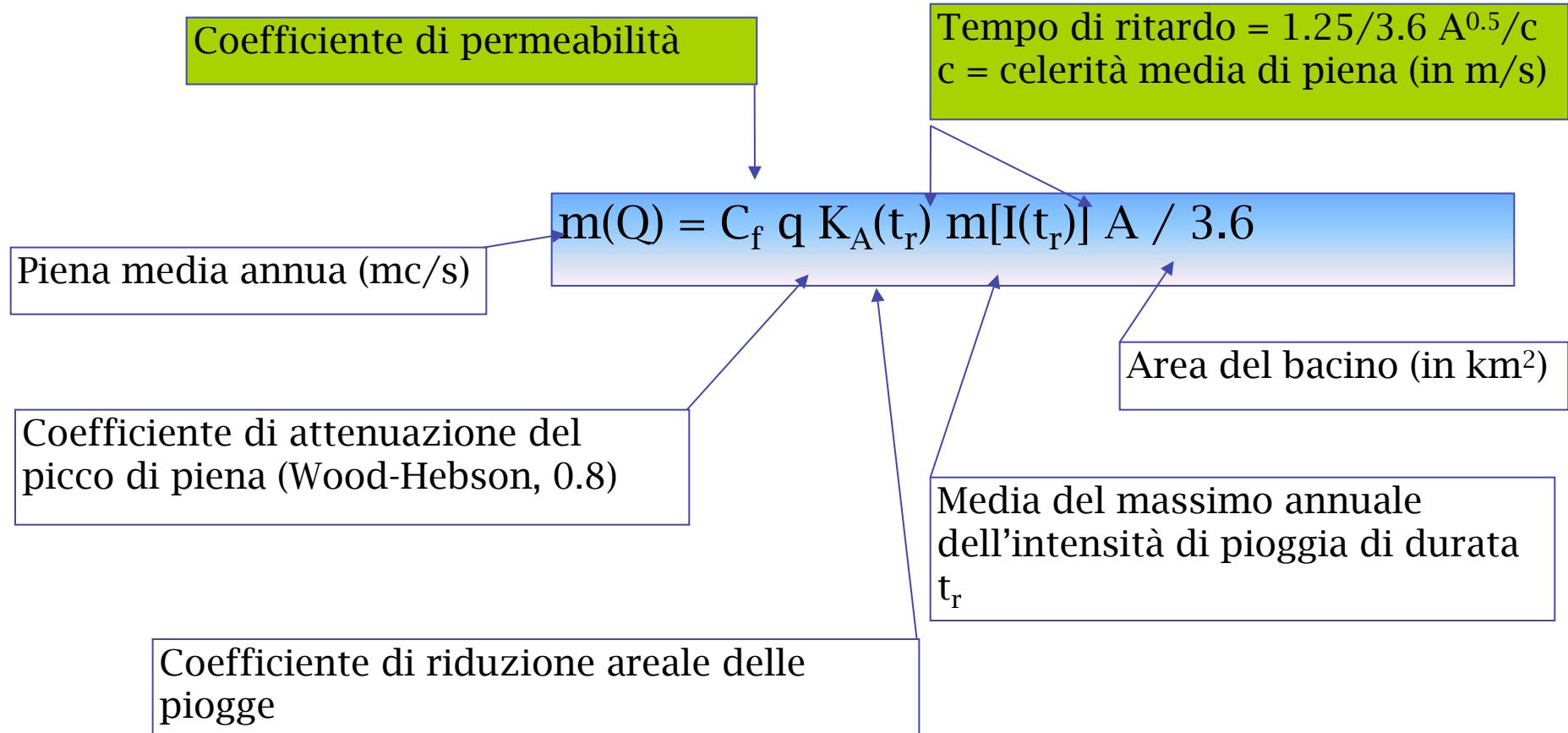


STIMA DELLA PIENA INDICE

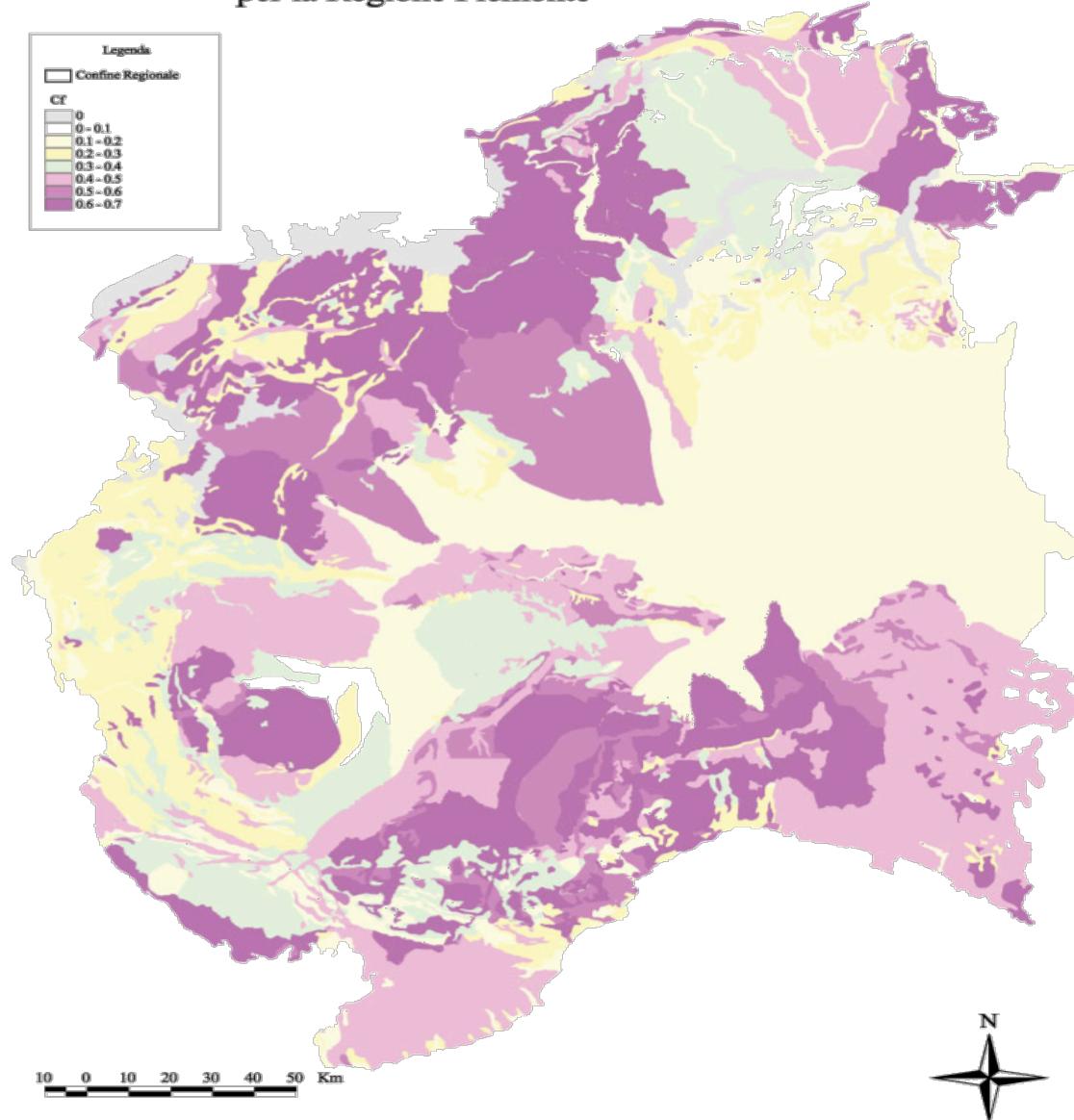
- 1) Metodi a base geomorfoclimatica (es. Formula razionale)
- 2) Metodi basati su regressioni rispetto a parametri morfo-climatici

Valutazione della piena media

Formula razionale in senso Geomorfoclimatico



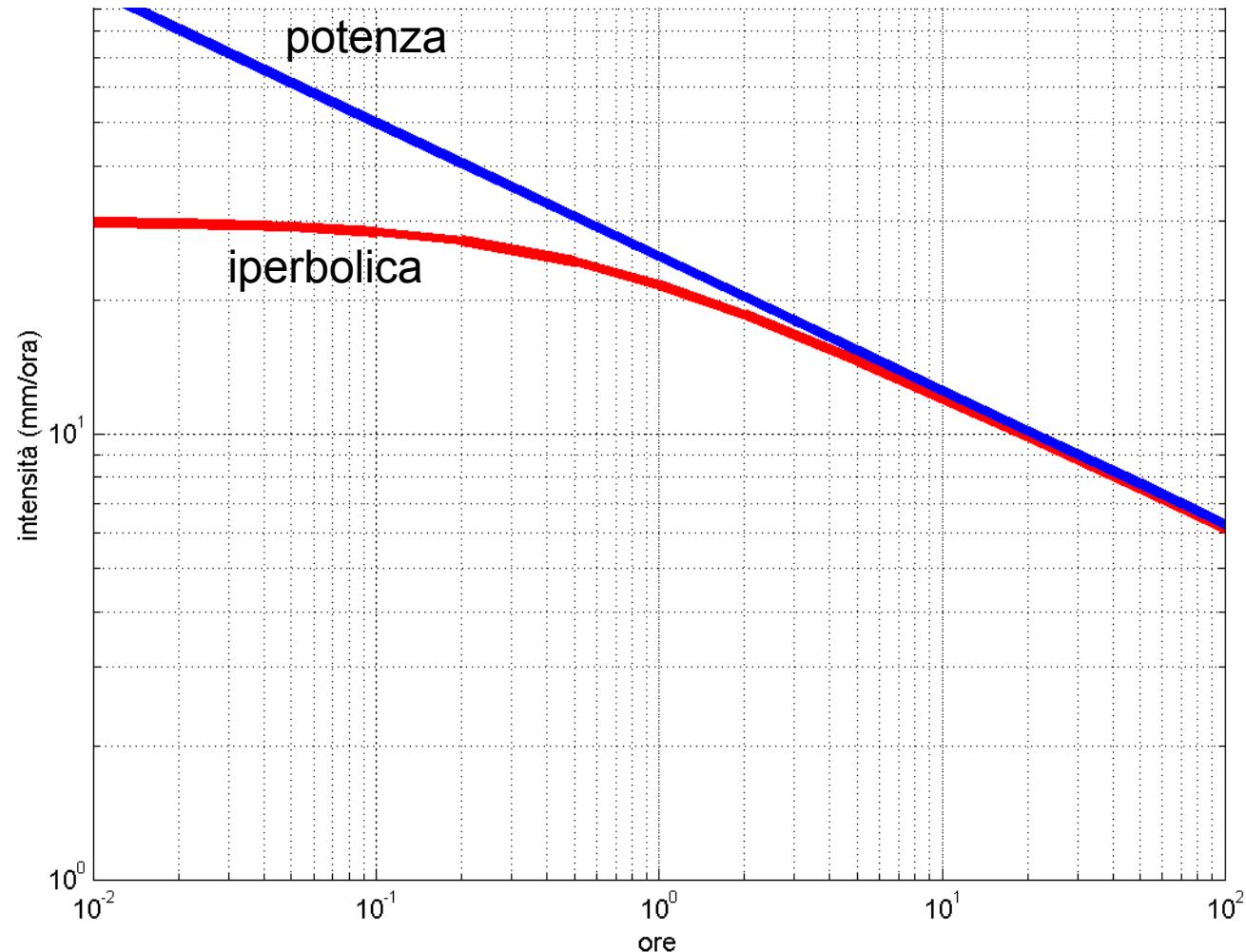
Carta del coefficiente di afflusso di piena, cf
per la Regione Piemonte



Curve Intensità-Durata delle precipitazioni

- Relazione di potenza: $m[I(d)] = a \ d^{n-1}$
 - Porta ad intensità infinite per $d = 0$
- Relazione iperbolica: $m[I(d)] = Io / (1 + d/d_c)^\beta$
 - Fissa un limite alle intensità pari ad Io per $d = 0$
 - Particolarmente utile in ambito urbano
 - Richiede dati per durate inferiori all'ora

Curve Intensità-Durata delle precipitazioni



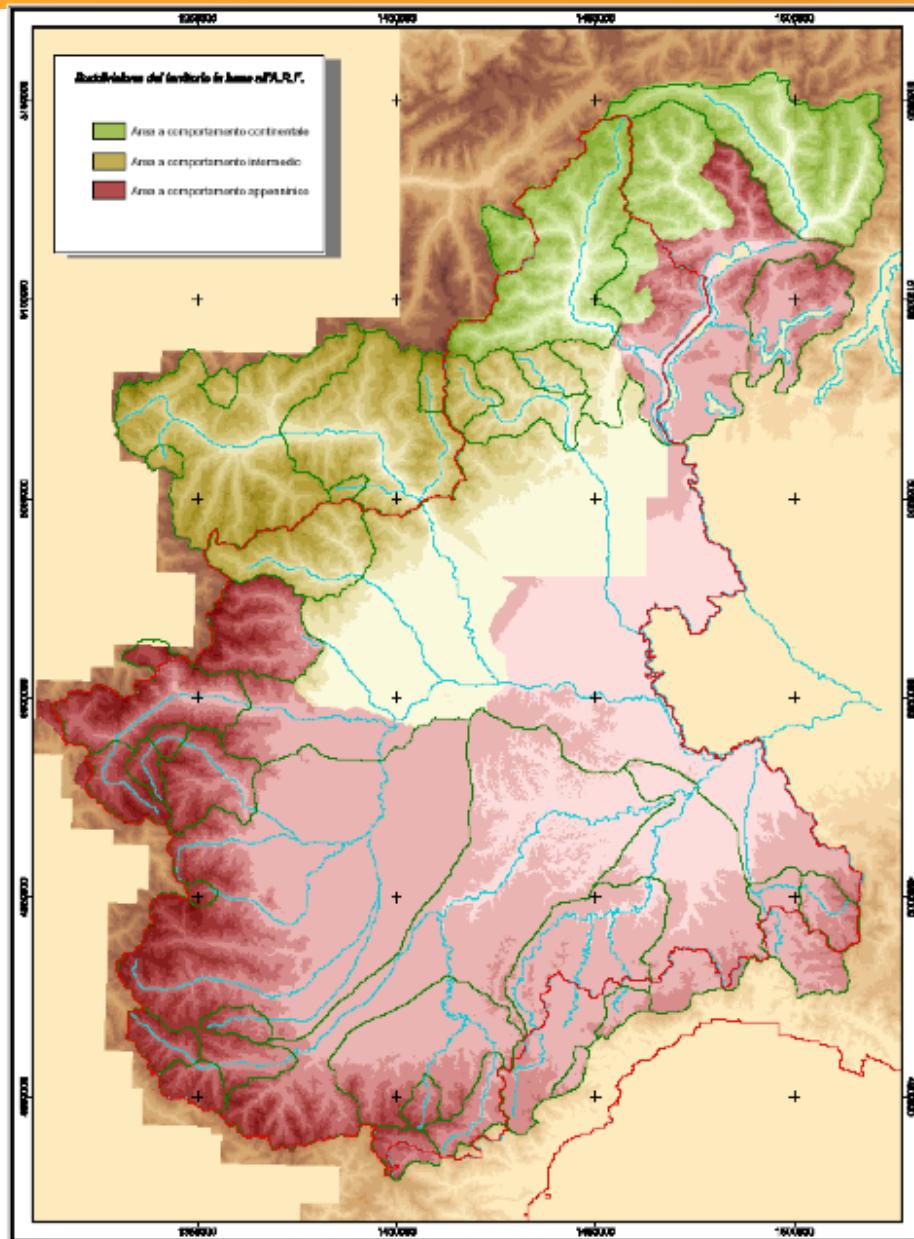
Riduzione areale delle precipitazioni

Caso Piemonte

$$KA(d) = 1 - (1 - \exp(-c_1 A)) \exp(-c_2 d^{c_3})$$

| | c_1 | c_2 | c_3 |
|-----------------------------------|--------|-------|-------|
| Area a comportamento continentale | 0.0043 | 0.45 | 0.25 |
| Area a comportamento intermedio | 0.0012 | 0.56 | 0.25 |
| Area a comportamento Appenninico | 0.0021 | 0.53 | 0.25 |

Tab. 4: parametri della legge di variabilità dell'ARF con durata (in ore) ed area (in kmq), nelle diverse regioni climatiche omogenee (valido per $A < 4000$ kmq).



Suddivisione in aree climatiche omogenee nei confronti del coefficiente di riduzione areale delle precipitazioni.

Metodi basati su regressioni rispetto a parametri morfo-climatici

$$y = Xb + e$$

Dove:

$$y = \begin{cases} \bullet Q_{\text{indice}} \\ \bullet Q_{\text{indice}} / A \\ \bullet \ln(Q_{\text{indice}}) \\ \bullet \ln(Q_{\text{indice}} / A) \end{cases} \quad \text{con} \quad Q_{\text{indice}} = \begin{cases} \bullet Q_{\text{media}} \\ \bullet Q_{\text{mediana}} \end{cases}$$

$$X = \begin{cases} \bullet \text{parametri morfologici} \\ \bullet \text{parametri climatici} \\ \bullet \text{parametri di uso del suolo} \end{cases}$$

b = coefficienti di regressione

e = residui della regressione

Regressione lineare semplice: definizione del modello

Variabile aleatoria Y funzione in una variabile non-random x

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

β_0 e β_1 sono parametri incogniti, coefficienti di regressione;
 β_0 è l'intercetta e β_1 è il coefficiente angolare (relazione lineare tra Y e x);

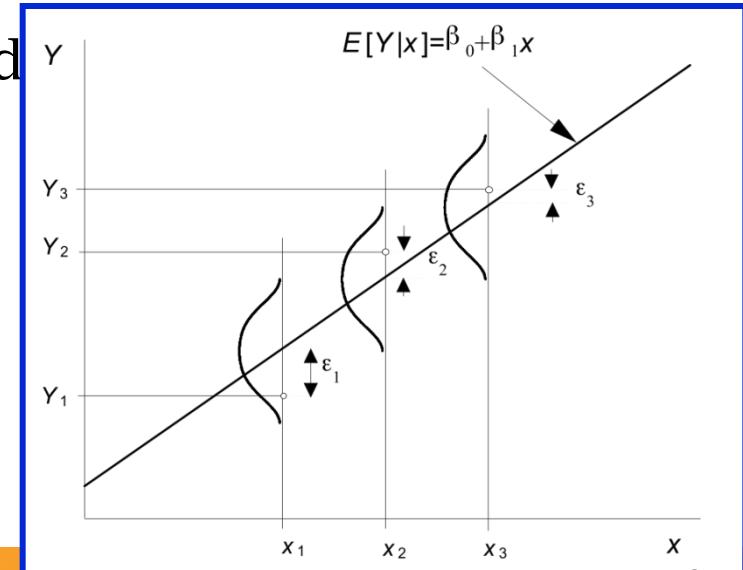
Y variabile dipendente, x variabile indipendente.

La v. a. ε rappresenta la differenza tra la variabile aleatoria Y e la componente deterministica, $\beta_0 + \beta_1 x$; termine di errore.

$E[\varepsilon] = 0$ e $Var[\varepsilon] = \sigma^2$, costante e indip. d.

↓ si ricava la seguente *relazione*:

$$E[Y | x] = \beta_0 + \beta_1 x$$

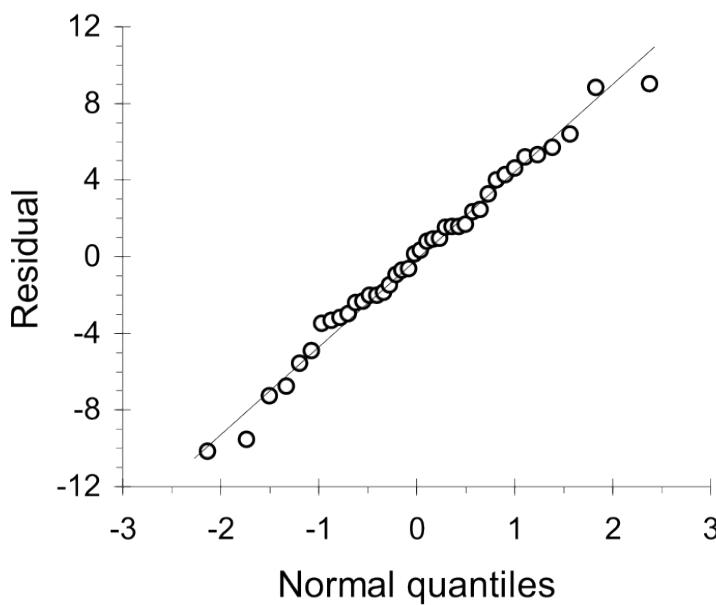


Regressione Lineare Semplice: Controllo delle ipotesi

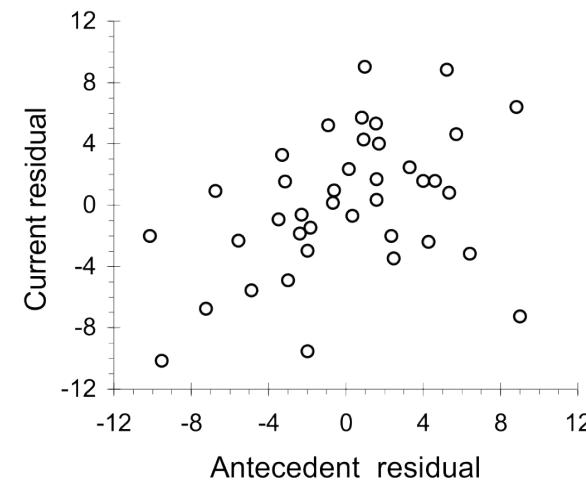
1. Controllo che $\varepsilon \sim N(0, \sigma^2)$
2. Controllo di indipendenza dei residui

Questa analisi dovrebbe mostrare la presenza di outliers o osservazioni “viziate” da errori di misura

1.Carta probabilistica
Normale



2.Controllo
indipendenza dei residui



Regressione Lineare Multipla: Definizione del modello

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{p-1} x_{p-1} + \varepsilon$$

Y è la variabile da spiegare, x_1, x_2, \dots, x_{p-1} sono $p-1$ variabili esplicative, $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$ sono p parametri (coefficienti di regressione). I parametri $\beta_1, \beta_2, \dots, \beta_{p-1}$ sono chiamati coefficienti di regressione parziale.

ε rappresenta il termine di errore, $\varepsilon \sim N(0, \sigma^2)$

Regressione Lineare Multipla: stima dei parametri attraverso il metodo dei minimi quadrati

I p parametri incogniti $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$ sono rappresentati dal vettore colonna β di dimensione $(p \times 1)$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}$$

Regressione Lineare Multipla: stima dei parametri

Le osservazioni x_{ij} sono contenute nella matrice di dimensione $(n \times p)$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p-1} \\ 1 & x_{21} & x_{22} & \dots & x_{2p-1} \\ \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot \\ 1 & x_{n1} & x_{n2} & \dots & x_{np-1} \end{bmatrix}$$

Gli n termini di errore ε_i e le n osservazioni della variabile Y sono contenuti nei vettori ε and Y di dimensione $(n \times 1)$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

Il modello di regressione multipla può essere scritto in forma matriciale come

$$y = X\beta + \varepsilon$$

Regressione Lineare Multipla: Stima dei parametri

Il valore medio $E[Y]$ di Y è

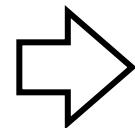
$$E[Y] = X\beta$$

$$S^2 = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_{p-1} x_{ip-1})^2$$

Minimizzazione
somma quadratica
degli scarti

$$\left\{ \begin{array}{l} \frac{\partial S^2}{\partial \beta_0} = 0 \quad n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{i2} + \dots + \hat{\beta}_{p-1} \sum_{i=1}^n x_{ip-1} = \sum_{i=1}^n y_i \\ \frac{\partial S^2}{\partial \beta_1} = 0 \quad \hat{\beta}_0 \sum_{i=1}^n x_{i1} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{i1}x_{i2} + \dots + \hat{\beta}_{p-1} \sum_{i=1}^n x_{i1}x_{ip-1} = \sum_{i=1}^n y_i x_{i1} \\ \dots \\ \dots \\ \frac{\partial S^2}{\partial \beta_{p-1}} = 0 \hat{\beta}_0 \sum_{i=1}^n x_{ip-1} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}x_{ip-1} + \hat{\beta}_2 \sum_{i=1}^n x_{i2}x_{ip-1} + \dots + \hat{\beta}_{p-1} \sum_{i=1}^n x_{ip-1}x_{ip-1} = \sum_{i=1}^n y_i x_{ip-1} \end{array} \right.$$

$$X^T X \hat{\beta} = X^T y$$



$$\hat{\beta} = (X^T X)^{-1} X^T y$$

Adattamento dei dati al modello (1)

Coefficiente di determinazione, R^2 (coefficiente di correlazione multipla)

$$R^2 = 1 - \frac{SS_E}{SS_y}$$

dove $SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ $SS_y = \sum_{i=1}^n (y_i - \bar{y})^2$

Fornisce la quota di variabilità di Y spiegata attraverso le variabili esplicative x_i .

$R^2=1$ perfetto adattamento del modello ai dati

Coefficiente di determinazione R^2_{adj} (corretto in base al numero p dei parametri)

n è il numero delle osservazioni

$$R^2_{adj} = 1 - \frac{SS_E}{SS_y} \frac{n-1}{n-p}$$

Esempio: stima della piena indice per i bacini piemontesi

$$q_{indice} = c_0 \times A^{c1}$$

$$q_{indice} = c_0 \times A^{c1} \times a_1^{c2}$$

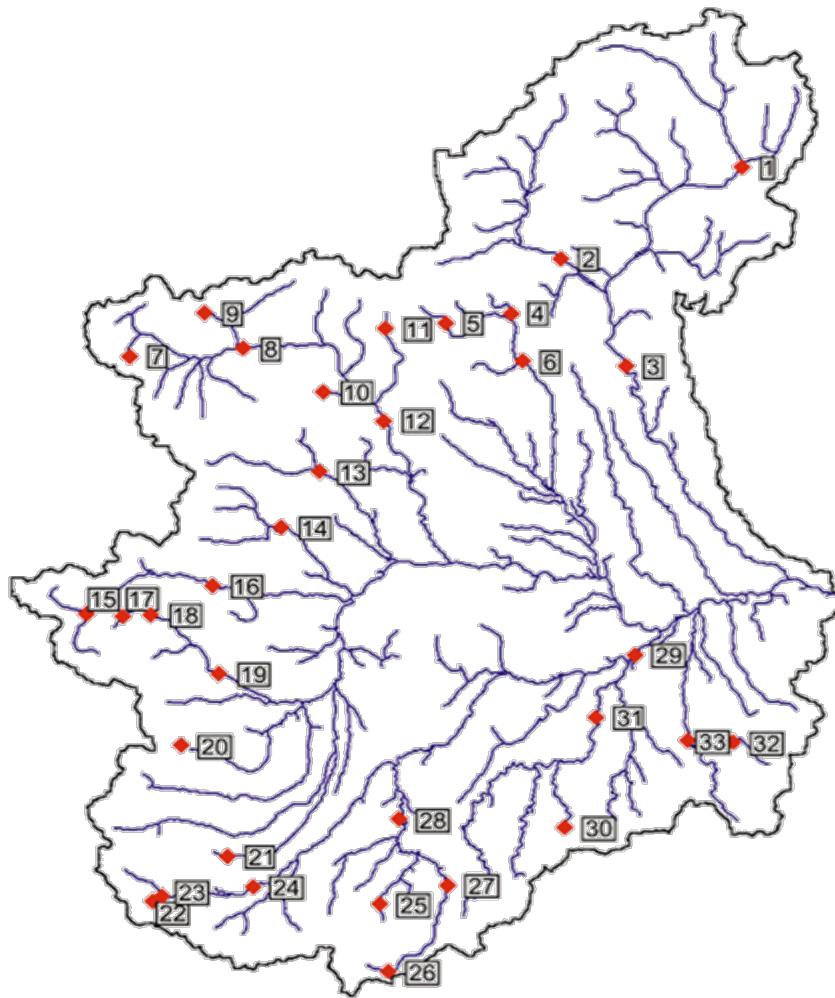
$$q_{indice} = c_0 \times A^{c1} \times a_1^{c2} \times n^{c3}$$

$$q_{indice} = c_0 \times A^{c1} \times a_1^{c2} \times n^{c3} \times DH^{c4}$$

$$q_{indice} = c_0 \times A^{c1} \times a_1^{c2} \times n^{c3} \times DH^{c4} \times S^{c5}$$

esame di una serie di modelli di regressione con numero variabile di parametri

Esempio: stima della piena indice per i bacini piemontesi

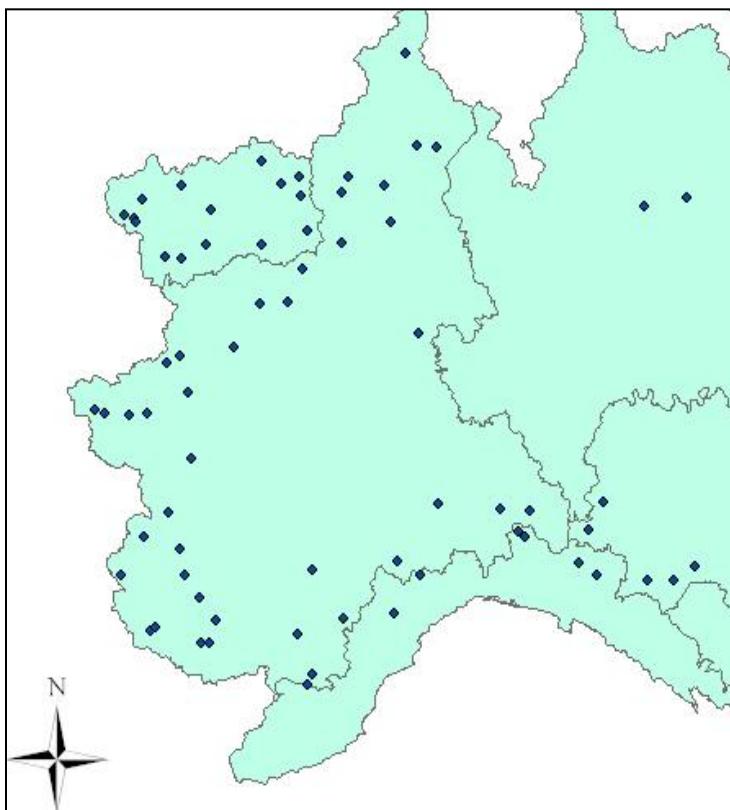


Dati utilizzati in Piemonte

Massimi annui delle portate al colmo di piena e delle portate medie giornaliere in 81 stazioni, di cui:

- 58 storiche (S.I.M.N)
- 13 Enel
- 26 sbarramenti (sui 46 di interesse) di cui 10 con misure al colmo

Parametri geomorfoclimatici considerati (Caso Piemonte)



- 44 parametri geomorfoclimatici:
- morfologia (30): area, lunghezza asta principale, rapporti di Horton, pendenze, coord. punti caratteristici del bacino, etc.;
 - suolo (7): indici di uso del suolo (Corine), CN, permeabilità apparente;
 - piovosità (7): a, n, afflusso medio annuo, regimi pluviometrici.



Criteri seguiti (Caso Piemonte)

Si sono considerati modelli di regressione:

- *parsimoniosi*: le variabili esplicative sono al massimo 4
- caratterizzati da una buona *capacità descrittiva* (misurabile in termini di R^2)
- per cui tutte le variabili esplicative coinvolte siano *significative* (test di Student con livello di significatività = 1 %)
- *stabili* al variare dei dati di taratura (*assenza di multicollinearità*, cioè dipendenti da grandezze non correlate tra loro)

ANALISI DEI RISULTATI MEDIANTE:

- verifica delle caratteristiche dei residui
- calcolo delle *statistiche d'errore* (*RMSE, MAE e MAPE*) riferite ad una y convenzionale (Q/A)

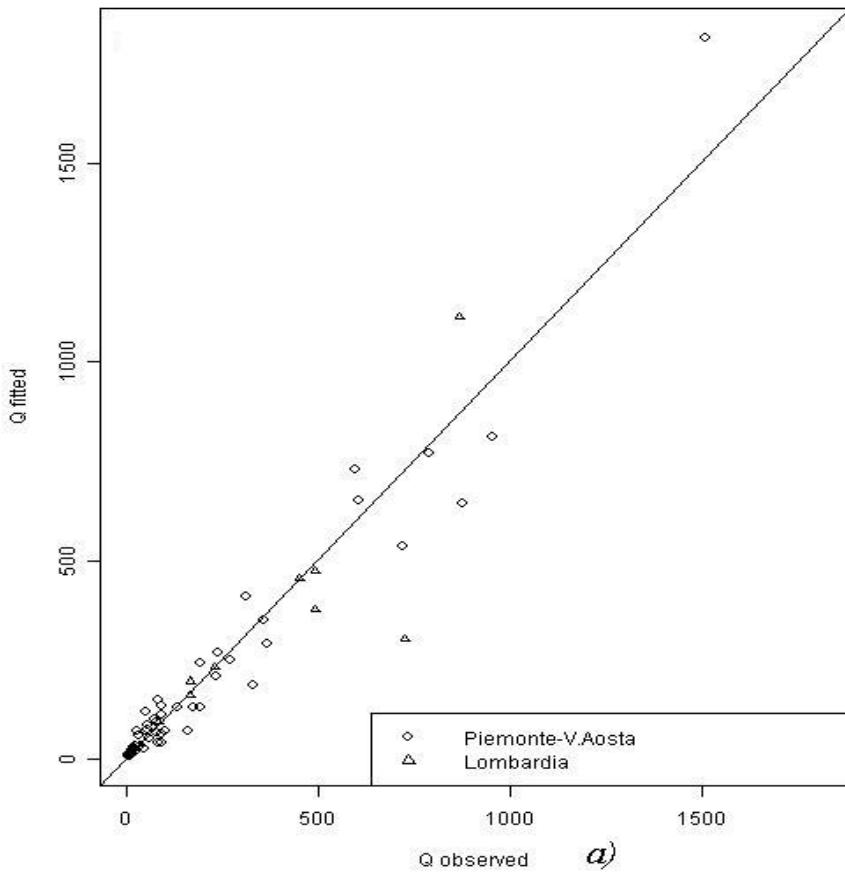
Esempio: stima della piena indice per i bacini piemontesi

$$y = \ln(Q_{indice} / A)$$

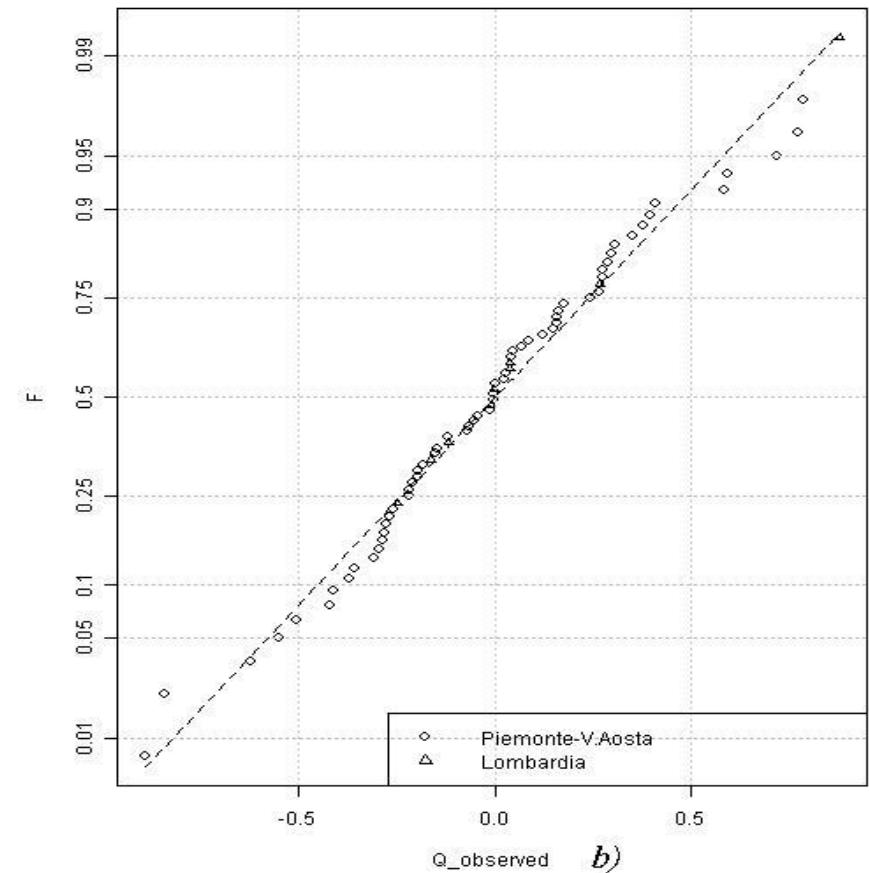
| X₁ | X₂ | X₃ | X₄ | Intercett a | Coeff. X₁ | Coeff. X₂ | Coeff. X₃ | Coeff. X₄ | R² | RMSE postCV | MAPE postCV |
|----------------------|----------------------|----------------------|----------------------|------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|----------------------|------------------------|------------------------|
| a | | | | -6.004 | 1.805 | | | | 0.72 | 0.37 | 0.31 |
| Slope_medio1 | afflusso | | | -16.301 | -0.963 | 2.68 | | | 0.78 | 0.32 | 0.34 |
| Y_sc | Lungh_vett_or | a | | -282.74 | 17.92 | -0.36 | 2.21 | | 0.83 | 0.32 | 0.39 |
| Lungh_vett_or | a | n | A _c | -7.52 | -0.23 | 2.82 | 2.45 | -0.63 | 0.85 | 0.34 | 0.31 |

Verifica di qualità di adattamento e delle caratteristiche dei residui

Y magnitudine + a + n + A2



RESIDUI magnitudine + a + n + A2





Qualità di adattamento

$$RMSE = \sqrt{\frac{1}{s} \sum_{j=1}^s (y_j - \hat{y}_j)^2} \quad MAE = \frac{1}{s} \sum_{j=1}^s |y_j - \hat{y}_j| \quad MAPE = \frac{1}{s} \sum_{j=1}^s \left| \frac{y_j - \hat{y}_j}{y_j} \right|$$

- ✓ RMSE= Root Mean Square Error (conferisce un peso maggiore agli errori più elevati)
- ✓ MAE= Mean Absolute Error (conferisce un peso lineare agli errori più elevati)
- ✓ MAPE= Mean Absolute Percent Error (conferisce ad ogni errore lo stesso peso)
- ✓ R^2_{adj} (corretto in base al numero dei parametri)

$$R^2_{adj} = 1 - \frac{y^T y - \beta^T X^T y}{y^T y - \frac{1}{s} \left(\sum_{j=1}^s y_j \right)^2} \cdot \frac{(s-1)}{(s-p)}$$

- ✓ Importante valutare il comportamento in siti non strumentati:

CALCOLO DELLE STATISTICHE DOPO LA
PROCEDURA DI *CROSS - VALIDAZIONE*



Relazione finale di stima

Considerati:

- ✓ gli errori;
- ✓ l'allineamento dei punti (y_{fitted} , $y_{observed}$) sulla bisettrice;
- ✓ la facilità di determinare i descrittori

la migliore relazione stimata è:

$$\overline{Q} = 1.5243 \cdot 10^{-4} \cdot C_{H_{med}} \cdot \text{aff} \cdot a \cdot A^{0.5(1+n)}$$

\overline{Q} = piena indice in m^3/s

$$C_{H_{med}} = \ln \left(e - (e - 1) \frac{H_{med}}{3500} \right)$$

A = area del bacino in Km^2

H_{med} = quota media del bacino

a [mm/h] ed n [] = coefficienti della CPP

aff = afflusso medio annuo

Relazione da modello regressivo ARPIEM 2012

$$Q_{\text{ind}} = 0.01324 \cdot \text{area}^{0.7995} \cdot \text{IDF}_a^{2.82089} \cdot \text{IDF}_n^{2.06805} \cdot \text{LCV}_{1h}^{1.33232}$$

